MATH 303 — Measure Theory
Lecture Notes, Fall 2025

Ethan Ackelsberg

Contents

Part 1

Motivation and Basics of Abstract Measure Theory

CHAPTER 1

Motivating Problems of Measure Theory

Learning Objectives

At the end of this chapter, you will be able to:

- Compare and contrast different approaches to the "problem of measurement" in Euclidean space and identify the advantages and disadvantages of different methods
- Describe applications of measure theory to other areas of mathematics

1. The Problem of Measurement

A basic (and very old) problem in mathematics is to compute the size (length, area, volume) of geometric objects. In this chapter, we will trace the history (in a highly abbreviated form) of the mathematical developments related to the measurement of the size of geometric objects from ancient times up to the 20th century. The guide throughout will be the following open-ended questions:

- What are the "geometric objects" to which we want to (and are able to) assign a notion of size?
- What properties should size (length, area, volume) satisfy?
- How do we compute sizes of geometric objects?

This loosely-defined problem is what we will call the "problem of measurement" in Euclidean space.

2. Ancient Mathematics - Polygons, Polygons, Polygons!

In the Greek school of mathematics of antiquity[1], the computations of areas and volumes of regions was carried out by reducing general regions for which the area or volume was unknown to polygon or polyhedral regions for which the area or volume was easily computed. This consisted of two primary methods: *quadrature* (or *squaring*) and the *method of exhaustion*.

2.1. Quadrature. Quadrature (or squaring) is the process of constructing, from a given two-dimensional region, a square of equal area. This is easily carried out for simple regions such as rectangles (see Example 1.1), parallelograms, and triangles, but quickly becomes much more difficult for curved regions. The problem of "squaring the circle," i.e. carrying out this procedure for a circle in a finite number of steps using only straightedge and compass, stumped ancient mathematicians and for good reason: the fact that π is a transcendental number (proved by Lindemann in 1882) makes a solution to the problem impossible.

Example 1.1: Quadrature of a rectangle

Given a rectangle with sides a, b, we can square the rectangle as follows. Place segments of length a and b end to end and form a (semi)circle with diameter given by the two segments

[1] A very enlightening discussion of the history of "Greek mathematics" recently took place in the pages of the *Notices of the American Mathematical Society*; see [**Kim25**, **Net25**].

(of total length $a + b$), and draw a segment perpendicular to the diameter at the meeting point of the two segments (see Figure 1.1).

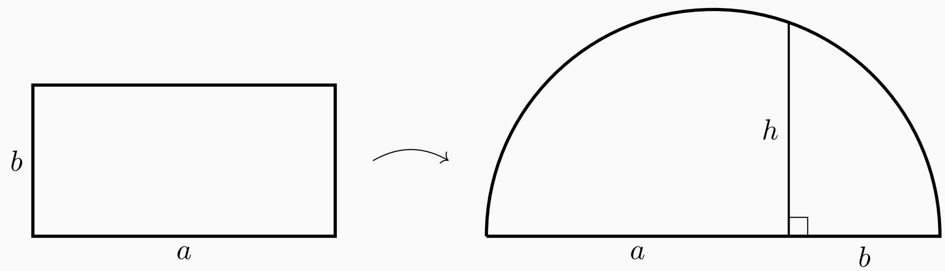

FIGURE 1.1.

To compute the height h, form two right triangles over bases a and b (see Figure 1.2). The two hypotenuses meet at a right angle by Thales's theorem. Then by three applications of the Pythagorean theorem,

$$\underbrace{(a+b)^2}_{a^2+b^2+2ab} = \underbrace{(a^2+h^2)+(b^2+h^2)}_{a^2+b^2+2h^2},$$

whence $h^2 = ab$. The square with base h shown in red thus has the same area as the original rectangle, so we have successfully squared the rectangle.

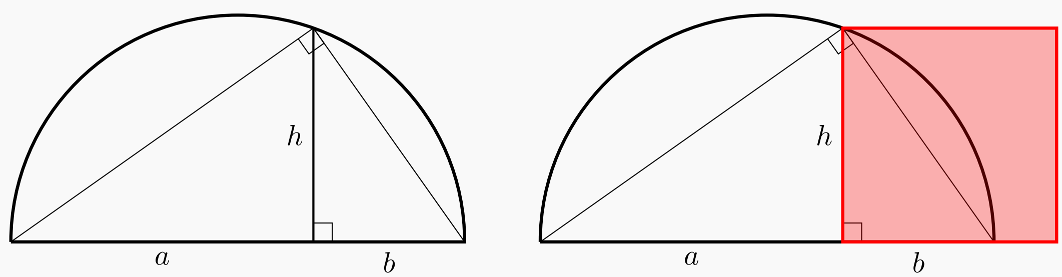

FIGURE 1.2.

2.2. Method of Exhaustion. Another method utilized in antiquity and containing the seeds of later developments in analysis was the method of exhaustion. Credited to Eudoxus for establishing the method rigorously, the method of exhaustion consists of inscribing and circumscribing sequences of polygons that converge to a given shape (see Figure 1.3). When used in conjunction with the method of squaring, which can be used to compute polygonal areas, the method of exhaustion was a powerful method for measurement in Greek mathematics. Polygonal approximations (rediscovered and improved in various locations and times) continued to be the best-known method for computing π until the end of the 17th century.

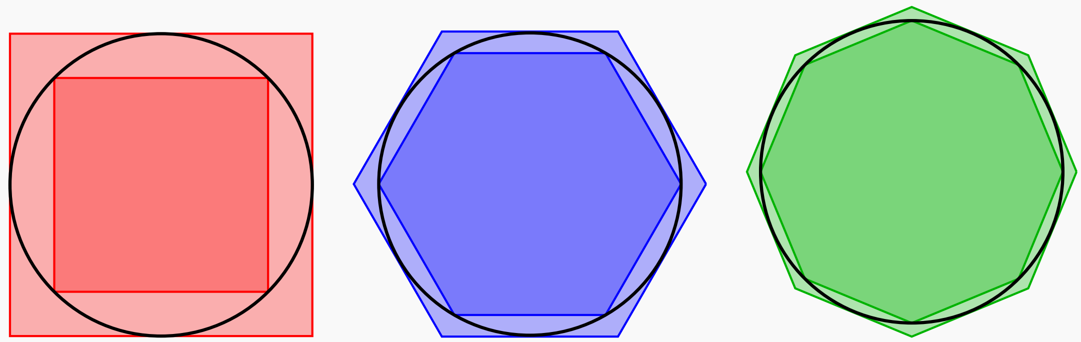

FIGURE 1.3. The method of exhaustion applied to a circle. Archimedes used this method (with regular 96-gons) to show $3 + \frac{10}{71} < \pi < 3 + \frac{1}{7}$.

3. Indivisibles and Infinitesimals

A significant breakthrough in the computation of areas and volumes was formalized in the 17th century by Bonaventura Cavalieri. (Similar methods had been used in antiquity by Archimedes and in the 5th century CE by Chinese mathematicians Zu Chongzhi and Zu Gengzhi, but the method was not in common usage in early modern Europe.) Cavalieri's principle can be expressed as follows:

> *Suppose two regions in the plane are bounded between two parallel lines. If the two regions have cross-sections of equal length, then they have equal area.*

A corresponding statement also holds in three dimensions: if two regions have planar cross-sections of equal area, then they have equal volume.

EXAMPLE 1.2

Cavalieri's principle can be used to compute the volume of a cone. First, by slicing parallel to the base, Cavalieri's principle shows that the volume of a pyramidal region depends only on the area of the base and the height. In particular, the computation of the volume of a cone (or any other starting pyramid) can be reduced to computing the volume of a square pyramid (Figure 1.4).

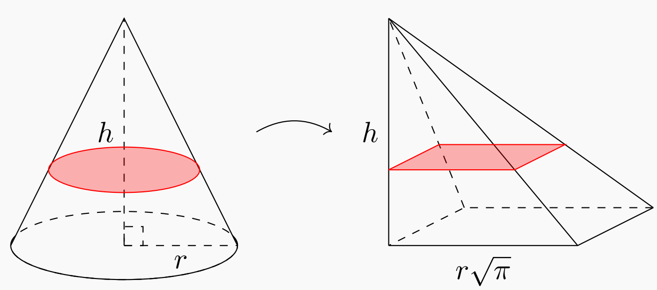

FIGURE 1.4. A demonstration of Cavalieri's principle. The two regions are built out of cross-section of equal area so have equal volume.

The volume of the resulting pyramid can be computed by observing that three pyramids for which the height is equal to the side length of the base can be combined into a cube (Figure 1.5).

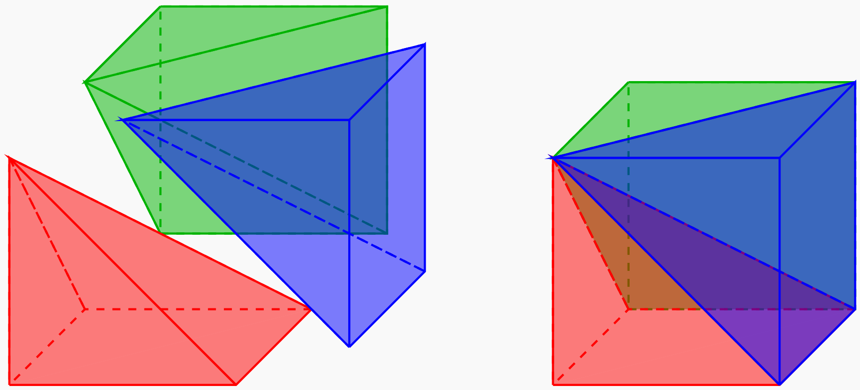

FIGURE 1.5. Three square pyramids combine to form a cube, so the volume of each pyramid is one third the volume of the cube.

Since the volume of the pyramid scales proportionally to its height, we conclude that the pyramid (and hence the cone) has volume $\frac{1}{3}\pi r^2 h$.

While the volume of a cone was known in antiquity using the method of exhaustion (it appeared, for example, in the 12th book of Euclid's *Elements*), Cavalieri's principle provides a much simpler proof. As a result of the negative resolution of Hilbert's third problem, it is also now known that the computation of the volume of a cone or even certain pyramids requires some form of infinitary argument (using Cavalieri's indivisibles, integral calculus, or a limiting process as in the method of exhaustion), since polyhedra of equal volume cannot always be transformed into each other via finitely many polyhedral cuts and rearrangements.

4. Integral Calculus

Notwithstanding earlier developments from Menaechmus and Apollonius in ancient Greek mathematics and Omar Khayyam in 11th century Persian mathematics, the introduction of coordinate systems by Descartes in the 17th century set forth the discipline of *analytic geometry* and revolutionized geometric calculations by uniting geometry with algebra. The variety of "geometric objects" was no longer limited to polygons, polyhedra, conic sections, and other classical objects; mathematicians had been unleashed to describe an endless assortment of new shapes by means of algebraic formulae. But how was one to compute their sizes?

Following earlier contributions by Cavalieri (who computed the area under $y = x^n$ for $n \leq 9$), Wallis (who extended Cavalieri's work to general $n \in \mathbb{Z}$), and many others, Newton and Leibniz discovered an astonishing link between the computation of areas (integration) and differentiation, namely the *fundamental theorem of calculus*.

5. Introducing ε and δ

Early work in calculus was based on infinitesimals and does not meet our present-day standards for mathematical rigor. Though a rigorous foundation for the theory of infinitesimals was eventually established by Abraham Robinson in the 1960s (dubbed "nonstandard analysis" as a result of its later historical development), calculus was first put on firm foundations by the "ε-δ" formalism established in the 19th century by Cauchy, Weierstrass, and others. Using the newly rigorous notions of limits, the ancient method of exhaustion could finally reach its full potential with the integration theory developed by Riemann and Darboux. For purposes of exposition, we will focus on Darboux's approach to integration, which is very similar to Riemann's but with some simplifications.

Definition 1.3: Darboux Integration

Let $B = \prod_{i=1}^d [a_i, b_i]$ be a closed 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 $i \in \{1, \ldots, d\}$.

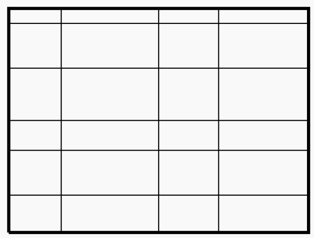

FIGURE 1.6. A Darboux partition in dimension $d = 2$ with $n_1 = 4$ and $n_2 = 6$.

- 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 $B_{\mathbf{j}}$.

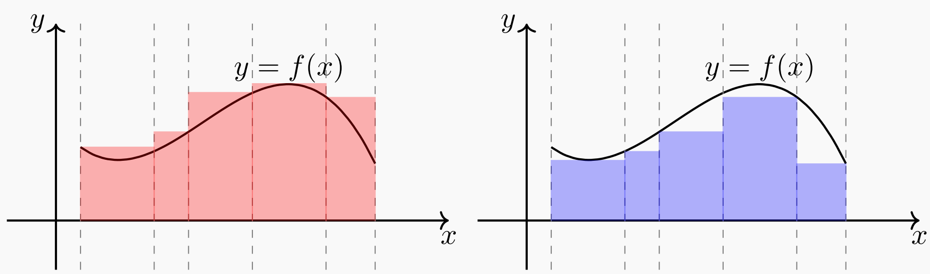

FIGURE 1.7. Upper (red) and lower (blue) Darboux sums of a function f over an interval ($d = 1$).

- 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 $\int_B f(\boldsymbol{x})\, d\boldsymbol{x}$.

Since Riemann integration is more commonly taught, we mention that the Darboux integral and the Rieman integral define the same quantity.

Proposition 1.4

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.

Because of its flexibility in terms of the dimension of the ambient Euclidean space, the Riemann–Darboux integral comes with an attendant notion of size or "hyper-volume" for regions in Euclidean space: the *Jordan content*.

Definition 1.5

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 closed box containing E.

Jordan measurable sets include basic geometric objects such as polyhedra, conic sections, regions bounded by finitely many smooth curves/surfaces, etc. The basic building blocks for Jordan measurable sets are what are called *simple sets* (or *elementary sets*).

Definition 1.6

An *interval* in $\mathbb{R}$ is a set of the form (a, b), $[a, b)$, $(a, b]$, or $[a, b]$ for some real numbers $a \leq b$. A *box* in $\mathbb{R}^d$ is a set of the form $B = \prod_{i=1}^d I_i$, where $I_1, \ldots, I_d$ are intervals. A set $S \subseteq \mathbb{R}^d$ is a *simple set* (or *elementary set*) if it is a finite union of boxes $S = \bigcup_{j=1}^k B_j$.

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

$$\text{Vol}(S) = \sum_{j=1}^k \text{Vol}(B_j) - \sum_{1 \leq j_1 < j_2 \leq k} \text{Vol}(B_{j_1} \cap B_{j_2}) + \sum_{1 \leq j_1 < j_2 < j_3 \leq k} \text{Vol}(B_{j_1} \cap B_{j_2} \cap B_{j_3}) - \ldots$$

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 1.7

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

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

and

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

The inner Jordan content can be viewed as a generalization of the method of approximation by inscribed polygons and the outer Jordan content as a generalization of the method of approximation by circumscribed polygons. In order to make sense of the size of an object using the Jordan content, the inscribed and circumscribed regions must approach the same size. In other words, Jordan measurable sets are those for which this extended method of exhaustion successfully converges. This is made precise by the following theorem.

Theorem 1.8

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) $J^*(\partial E) = 0$.

We do not include a proof of Theorem 1.8 but indicate its core content in Figure 1.8.

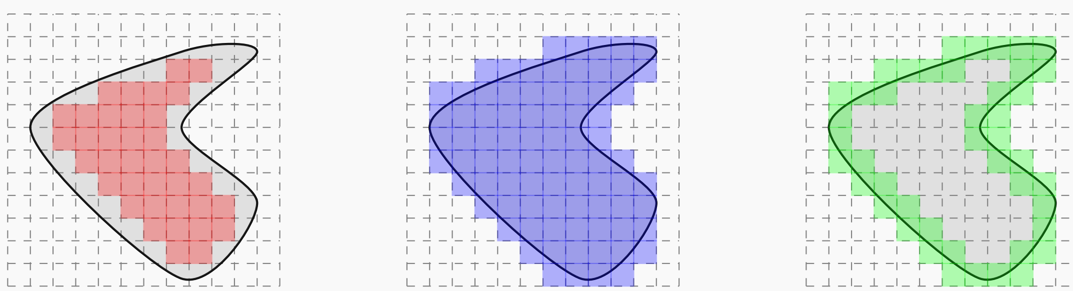

Figure 1.8. Simple sets approximating the inner (red) and outer Jordan content (blue) of a region in dimension $d = 2$. With the red boxes removed from the blue, we get a simple set covering the boundary (in green).

While the family of Jordan measurable sets is quite vast, a consequence of Theorem 1.8 is that not all sets are Jordan measurable.

Example 1.9

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

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 1.10

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.

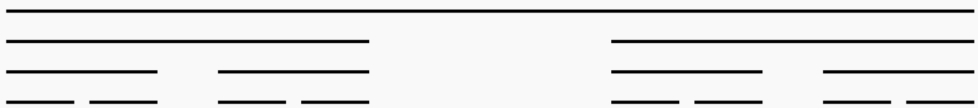

Figure 1.9. Iterative construction of the fat Cantor set.

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

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 $J^*(U) = J^*([0,1]) = 1$.

6. Set Theory, Choice, and the Impossibility of Measuring Everything

As mathematicians continued the quest for formalization, the notion of "geometric object" continued to expand in possible meaning. With set theory taking its place at the foundations of mathematics, one could now dream of perhaps assigning a size to arbitrary subsets of Euclidean space. Giuseppe Vitali dashed such hopes with a clever construction in 1905.

Thus far, our discussion of the notion of size has been largely based on geometric intuition. In order to say that there exists a set incapable of being assigned a sensible notion of size, we now take an axiomatic approach and reformulate (a version of) the problem of measurement as a concrete mathematical statement.

Problem 1.11: The Problem of Measurement (Strong Form)

Let $d \in \mathbb{N}$. Does there exists a notion of d-dimensional volume Vol, defined for all subsets of $\mathbb{R}^d$, such that

- normalized: $\operatorname{Vol}([0,1)^d) = 1$;
- isometry-invariant: if A and B are isometric, then $\operatorname{Vol}(A) = \operatorname{Vol}(B)$;
- countably additive: if $E_1, E_2, \dots$ are pairwise disjoint, then $\operatorname{Vol}\left(\bigsqcup_{n\in\mathbb{N}} E_n\right) = \sum_{n=1}^{\infty} \operatorname{Vol}(E_n)$.

Vitali showed that Problem 1.13 has a negative answer even for $d = 1$.

Theorem 1.12

There is no normalized, translation-invariant, countably additive function defined for all subsets 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 $x \in E_t$.

Suppose for contradiction that L is a normalized, translation-invariant, countably additive function defined for all subsets of $\mathbb{R}$.

Claim 3. For every $t \in \mathbb{Q} \cap [0,1)$, $L(E_t) = L(E)$.

We can write

$$E_t = ((E + t) \cap [0,1)) \sqcup ((E + t) \cap [1,2) - 1).$$

Therefore, by translation-invariance,

$$L(E_t) = L(E + t) = L(E).$$

Combining Claims 1–3 and using countable additivity of L,

$$1 = L([0,1)) = \sum_{t \in \mathbb{Q} \cap [0,1)} L(E_t) = \sum_{t \in \mathbb{Q} \cap [0,1)} L(E) = \infty \cdot L(E).$$

But there is no value of $L(E)$ that can satisfy this equation, so we have reached a contradiction. □

Confronted with Vitali's example, one must make some compromise. In order to comport with an intuitive meaning of "size," normalization and isometry-invariance appear absolutely essential. This leaves two options: (1) restrict the domain of the volume function to only assign size to a certain subclass of "nice" sets and hope to avoid the pathologies of the Vitali sets, or (2) sacrifice countable additivity for the weaker notion of finite additivity. We address the two possibilities in turn, starting with the latter. Relaxing our additivity assumption to *finite additivity*, we arrive at a new form of the problem of measurement.

Problem 1.13: The Problem of Measurement (Weak Form)

Let $d \in \mathbb{N}$. Does there exists a notion of d-dimensional volume Vol, defined for all subsets of $\mathbb{R}^d$, such that

- NORMALIZED: $\text{Vol}([0,1)^d) = 1$;
- ISOMETRY-INVARIANT: if A and B are isometric, then $\text{Vol}(A) = \text{Vol}(B)$;
- FINITELY ADDITIVE: if A and B are disjoint, then $\text{Vol}(A \sqcup B) = \text{Vol}(A) + \text{Vol}(B)$.

Surprisingly, the solvability of this weak form of the problem of measurement depends on the dimension d. The problem was solved by Banach in dimensions $d = 1$ and $d = 2$ using a version of the Hahn–Banach theorem from functional analysis. However, in dimensions 3 and higher, a paradoxical situation emerges.

Theorem 1.14: Banach–Tarski Theorem

Let $d \geq 3$. Given any two bounded regions $A, B \subseteq \mathbb{R}^d$, both with nonempty interior, there exist partitions $A = A_1 \sqcup \ldots \sqcup A_k$ and $B = B_1 \sqcup \ldots B_k$ for some $k \in \mathbb{N}$ such that A_i and B_i are congruent for each $i \in \{1, \ldots, k\}$. In particular, the unit ball can be decomposed into finitely many pieces and reassembled into two congruent copies of the unit ball.

Thus, at least in high dimensional situations, even with weaken the notion of "size" to only be finitely additive, there is no consistent way to measure every subset of $\mathbb{R}^d$. There is also good reason to insist on the property of countable additivity. For example, finitely-additive notions of measure produce integrals that do not interact with limits in the way that one might hope. For ease of exposition, we give an example with the Riemann integral, but similar examples can be constructed for any notion of integration that fails to be countably additive (including Banach's notion of integration in dimensions 1 and 2).

Example 1.15

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

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

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.

Gathering all of our observations thus far, the best we can hope for in addressing the problem of measurement is to exhibit a rich class of "measurable sets" (including, for example, polygons, circles, open sets[2], and general Jordan-measurable sets) for which we can define a normalized, translation-invariant, and countably additive notion of measure.

7. The Solution of Lebesgue

The Jordan non-measurable set in Example 1.10 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 $\mathbb{R}$.

[2]Example 1.10 shows that there are open sets that are not Jordan-measurable, so we need a more general construction to handle arbitrary open sets.

Proposition 1.16

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

By Proposition 1.16, 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 1.16 needs to be modified, but one can still reasonably talk about the d-dimensional volume of open sets in $\mathbb{R}^d$. The key is to replace open intervals with half-open boxes $\prod_{i=1}^d [a_i, b_i)$.

Definition 1.17

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 1.18

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 1.19

(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 A and B are congruent, then $\lambda(A) = \lambda(B)$.
(4) If $E_1 \subseteq E_2 \subseteq \dots \subseteq \mathbb{R}^d$ are Lebesgue measurable sets, then $\lambda\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \lambda(E_n)$.
(5) If $E_1 \supseteq E_2 \supseteq \dots$ are Lebesgue measurable subsets of $\mathbb{R}^d$ and $\lambda(E_1) < \infty$, then $\lambda\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \lambda(E_n)$.

The "continuity" properties expressed in items (4) and (5) result in a corresponding notion of integration that is able to interact desirably with pointwise limits, overcoming the shortcomings of the Riemann integral illustrated in Example 1.15.

8. 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 1.10) 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.

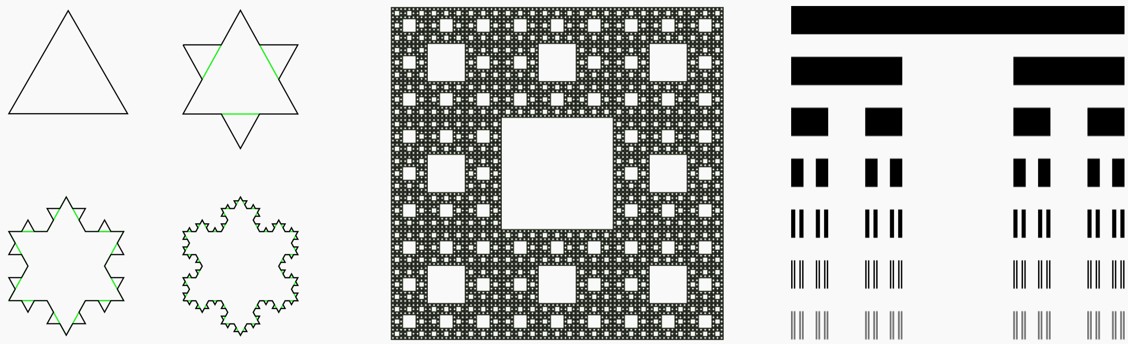

FIGURE 1.10. Fractal shapes: the Koch snowflake (left) of Hausdorff dimension $\frac{\log 4}{\log 3} \approx 1.26$, Sierpiński carpet (middle) of dimension $\frac{\log 8}{\log 3} \approx 1.89$, and middle-thirds Cantor set (right) of dimension $\frac{\log 2}{\log 3} \approx 0.63$.

CHAPTER 2

Measure Spaces

Learning Objectives

At the end of this chapter, you will be able to:

- Define the fundamental objects in measure theory (measurable sets, measurable functions, and measures)
- Identify and utilize tools for proving measurability of functions
- Prove basic properties of measures

1. σ-Algebras

Before defining measures, we must determine which subsets of a given set X we would like to be able to measure. Of course, in the best case scenario, we may hope to measure every subset of X. However, as demonstrated by Theorem 1.12, attempting to measure every set is often incompatible with other desirable properties for a measure. Thus, instead of insisting that a measure be defined for arbitrary subsets, we will be satisfied with having a sufficiently rich class of "measurable" subsets. What properties should we impose on this class? Certainly, we want the full space X to be measurable, and we should allow ourselves to perform the basic set-theoretic operations (complements, unions, and intersections). Allowing *finite* unions and intersections leads to the concept of an *algebra* of sets. Algebras are a very useful notion, but (as we saw with the Jordan content 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 countable intersection of sets:

$$\bigcap_{n\in\mathbb{N}} B_n = X \setminus \left(\bigcup_{n\in\mathbb{N}} (X \setminus B_n) \right).$$

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 $\mathcal{B} = \left\{\bigcup_{n\in S}[n, n+1) : S \subseteq \mathbb{Z}\right\}$.

The basic object of study in abstract measure theory is a *measurable space*, which is a set for which we have designated a σ-algebra of measurable sets. More formally, we have the following definition.

Definition 2.3

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*.

In order to produce a wider variety of examples of σ-algebras than what appears in Example 2.2, it is helpful to have a general construction for producing a σ-algebra from a given family of sets. For example, in a topological space, it is natural to insist that all open sets be measurable. But then we need a method for producing a σ-algebra that contains all of the open sets (and is not the power set $\mathscr{P}(X)$, as this may contain pathological examples like the Vitali sets that make it impossible to define interesting measures). The following property of σ-algebras enables the desired general constructing of σ-algebras.

Proposition 2.4

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, $\bigcup_{n\in\mathbb{N}} B_n \in \mathcal{B}$. □

Definition 2.5

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 2.4:

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

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

Definition 2.6

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

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.

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

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.

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 $\mathcal{B} = \mathrm{Borel}(X)$ and $\mathcal{C} = \mathrm{Borel}(Y)$ 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. □

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 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. □

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

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

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

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

Chapter Notes

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* [**Rud87**, Chapter 1]. Alternative presentations can be found in [**Fol99**, Sections 1.2, 1.3, and 2.1], and [**Tao11**, 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 [**Kec95**] and/or the lecture notes of Tserunyan [**Tse22**], which draw quite heavily on [**Kec95**].

CHAPTER 3

Integration Against a Measure

Learning Objectives

At the end of this chapter, you will be able to:

- Interpret measurability of a function in terms of approximation by simple functions
- Define the integral of a measurable function against a measure
- Prove properties of integration (linearity and fundamental limit theorems)
- Compare and contrast the measure-theoretic approach to integration with the Riemann–Darboux approach
- Apply the definitions and limit theorems to other problems in combinatorics, probability, and analysis

Our next task is to develop an integration theory for integrating measurable functions on abstract measure spaces. In the Riemann–Darboux approach to integration, the strategy is to approximate a general (integrable) function $f : [a, b] \to [0, \infty)$ by step functions, for which the integral is easily defined. This approach has two serious drawbacks that we wish to overcome. One, which we have seen in Example 1.15, is that the Riemann integral does not interact favorably with pointwise limits. The other more severe limitation is that step functions are built as piecewise constant functions on *intervals* (or boxes in higher dimensions), which uses the underlying geometry of Euclidean space. In the setting of abstract measure spaces, there is no geometry on which to rely. In this chapter, we will see that the more general notion of a *simple function* overcomes both of the aforementioned issues with the Riemann integral.

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 $a_1, \dots, a_m$ be the finite collection of values taken by s, and let $A_k = \{s = a_k\}$ for $k = 1, \dots, m$. Then the sets $A_1, \dots, A_k$ 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 $J \subseteq \{1, \dots, n\}$. 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 (see Figure 3.1).

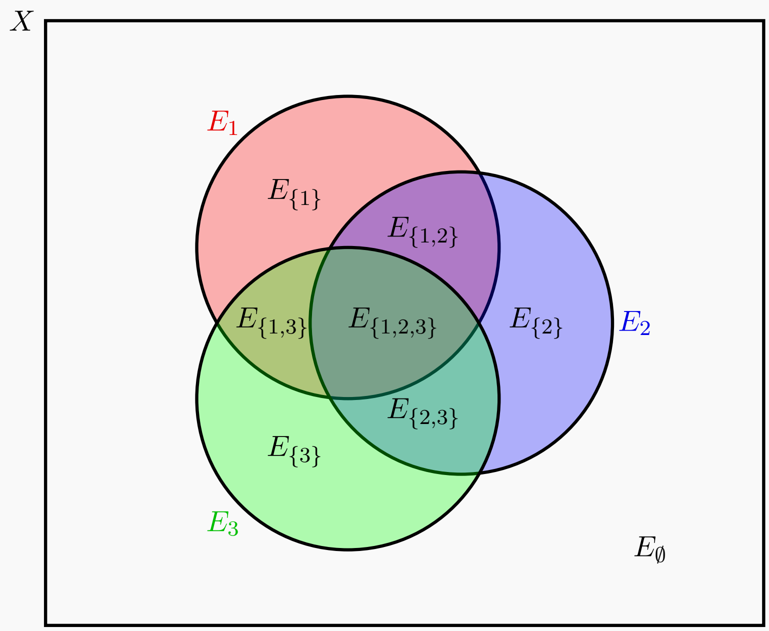

Figure 3.1. Construction of a partition $(E'_J)_{J \subseteq \{1,\dots,n\}}$ from sets $(E_j)_{j \in \{1,\dots,n\}}$.

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

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

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) and the values $c_1, \ldots, c_n$ are distinct.

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 $0 \leq s_1 \leq s_2 \leq \cdots \leq f$, and $s_n \to f$ pointwise.

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

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

The functions s_n are nondecreasing and satisfy $s_n \leq f$ by construction. If $f(x) < \infty$, then $f(x) - s_n(x) < 2^{-n}$ for all $n > f(x)$, so $s_n(x) \to f(x)$ as $n \to \infty$. On the other hand, if $f(x) = \infty$, then $s_n(x) = n$ for every $n \in \mathbb{N}$, so $s_n(x) \to \infty = f(x)$ as $n \to \infty$.

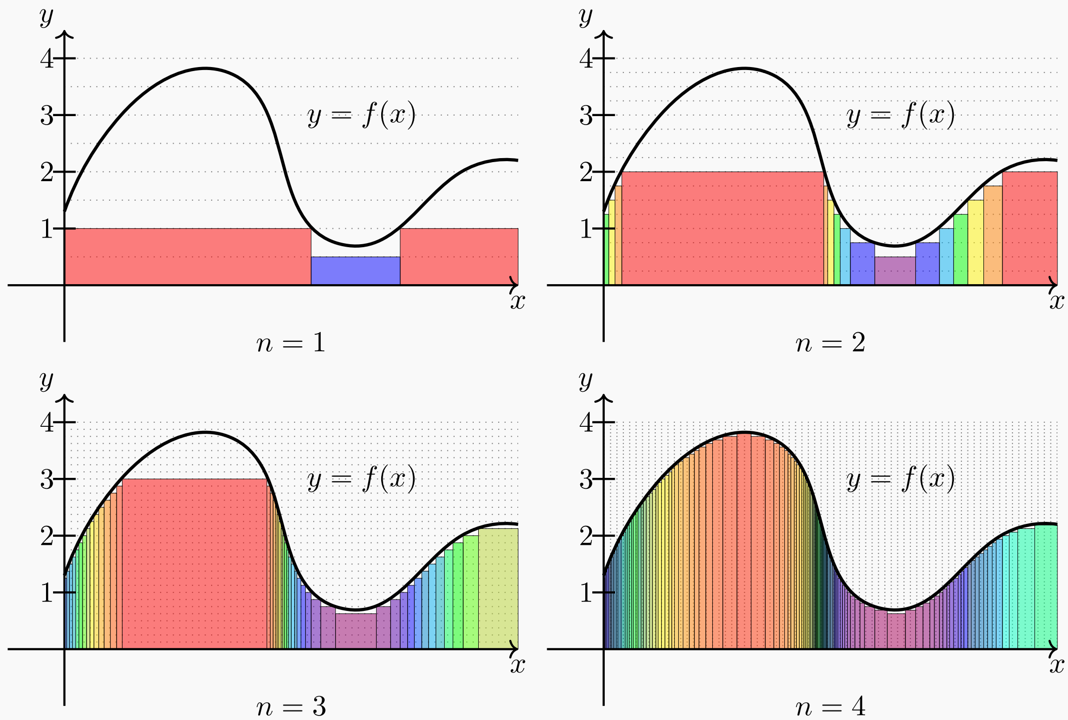

FIGURE 3.2. Successive approximations of a function f by simple functions.

□

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

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

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 $0 \leq s_1 \leq s_2 \leq \dots$, 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 $f \leq g$, then

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

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

Theorem 3.10: Monotone Convergence Theorem

Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions $0 \leq f_1 \leq f_2 \leq \dots$, 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 $0 \leq s_1 \leq s_2 \leq \dots \to f$ 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,

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

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

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

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, $0 \leq s \leq f$. For $n \in \mathbb{N}$, let $E_n = \{f_n \geq cs\}$. 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

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

Taking a supremum over all such simple functions, we conclude

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

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 $0 \le s_1 \le s_n \le \cdots \le f$ with $s_n \to f$ and $0 \le t_1 \le t_2 \le \cdots \le g$ with $t_n \to g$. Then $0 \le s_1 + t_1 \le s_2 + t_2 \le \cdots \le f + g$ 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

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

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: 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

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

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. $c \in \mathbb{R}$

Combine Case 1 and Case 2.

Case 4. $c = i$

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 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 3.16 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 3.17

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 in Section 5 below.

As another basic property of integration, we establish a version of the triangle inequality:

Proposition 3.18: 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 by Proposition 3.16,

$$\left| \int_X f \, d\mu \right| = \int_X \lambda f \, d\mu.$$

Since this is a real number, we then have $\int_X \lambda f \, d\mu = \int_X \mathrm{Re}\,(\lambda f) \, d\mu$, so by monotonicity of the integral,

$$\left| \int_X f \, d\mu \right| = \int_X \mathrm{Re}\,(\lambda f) \, d\mu \le \int_X |\lambda f| \, d\mu = \int_X |f| \, d\mu.$$

□

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),

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

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 of the measure μ. 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,

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

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

□

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 **??**),

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

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

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

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. □

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

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

Chapter Notes

For other presentations of integration on abstract measures spaces, see [**Fol99**, Section 2.1–2.3], [**Rud87**, Chapter 1], [**SS05**, Sections 2.1 and 6.2], and/or [**Tao11**, Section 1.3 and Subsection 1.4.4]. The development of integration in the books of Folland [**Fol99**] and Rudin [**Rud87**] is very similar to the presentation in these notes. By contrast, Stein and Shakarchi [**SS05**] and Tao [**Tao11**] first develop integration in the special case of the Lebesgue measure before moving to abstract spaces. The book of Stein and Shakarchi [**SS05**] 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 [**Oxt80**] 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.*

Bibliography

[Fol99] Gerald B. Folland. *Real Analysis*, 2nd edition (John Wiley & Sons, Inc., 1999).

[Kec95] Alexander S. Kechris. *Classical Descriptive Set Theory*, Grad. Texts in Math., **156** (Springer, 1995).

[Kim25] Minhyong Kim. History, identity, and ownership in mathematics. *Notices Amer. Math. Soc.* **72** (2025) 846–853.

[Net25] Reviel Netz. History, identity, and ownership in mathematics: a response to Minhyong Kim. *Notices Amer. Math. Soc.* **72** (2025) 854–857.

[Oxt80] John C. Oxtoby. *Measure and Category*, 2nd edition, Grad. Texts in Math., **2** (Springer, 1980).

[Rud87] Walter Rudin. *Real and Complex Analysis*, 3rd edition (McGraw-Hill Book Co., 1987).

[SS05] Elias M. Stein and Rami Shakarchi. *Real Analysis* (Princeton University Press, 2005).

[Tao11] 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

[Tse22] Anush Tserunyan. *Introduction to Descriptive Set Theory.* Online lecture notes (2022). Available at: https://www.math.mcgill.ca/atserunyan/Teaching_notes/dst_lectures.pdf