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).

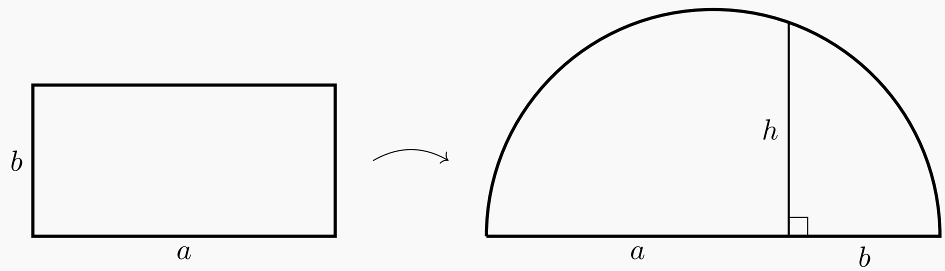

FIGURE 1.1.

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.

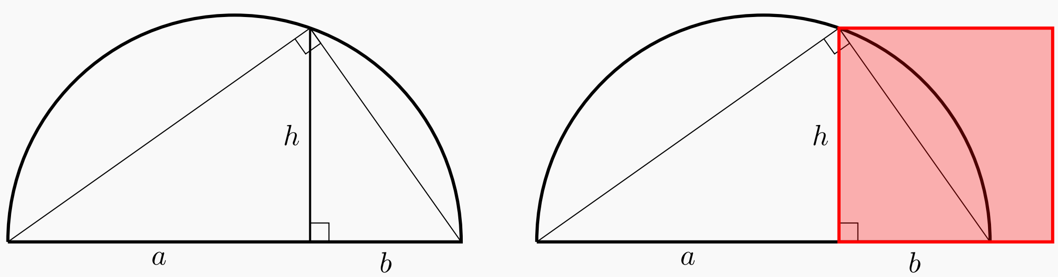

FIGURE 1.2.

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.

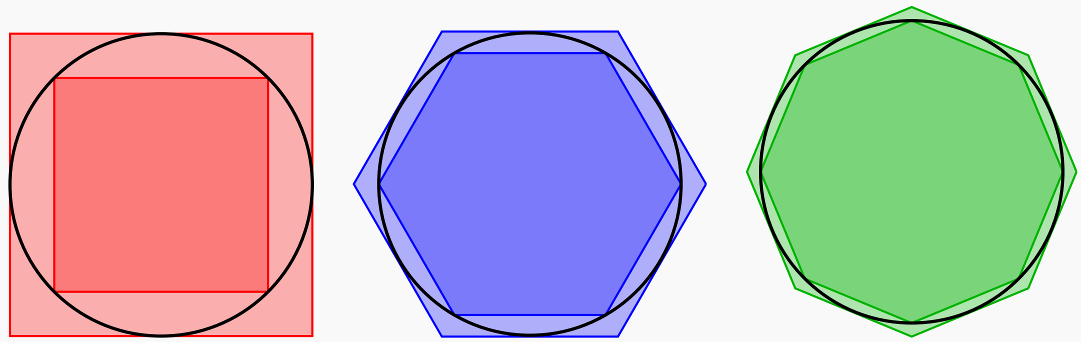

FIGURE 1.3. The method of exhaustion applied to a circle. Archimedes used this method (with regular 96-gons) to show $3 + \frac{10}{71} < \pi < 3 + \frac{1}{7}$.

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).

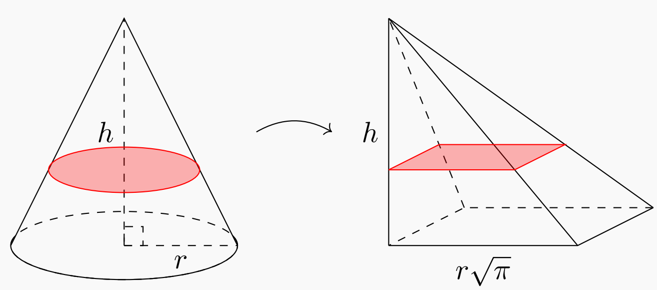

FIGURE 1.4. A demonstration of Cavalieri's principle. The two regions are built out of cross-section of equal area so have equal volume.

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).

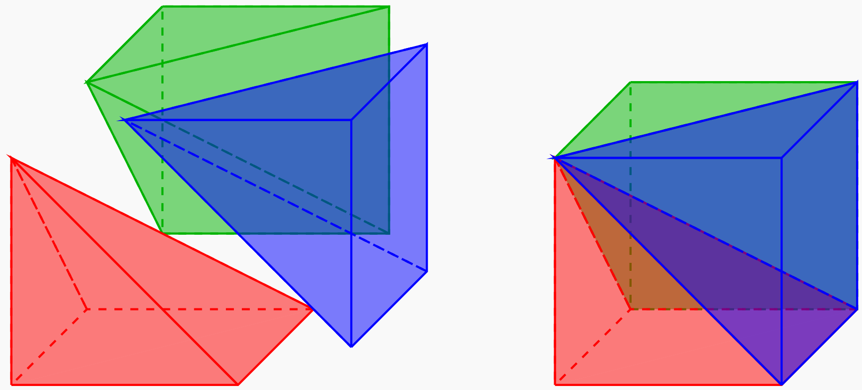

FIGURE 1.5. Three square pyramids combine to form a cube, so the volume of each pyramid is one third the volume of the cube.

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 $i \in \{1, \ldots, d\}$.

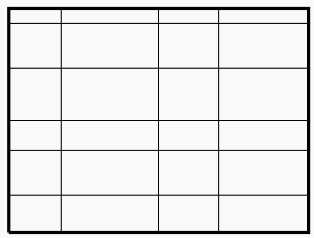

FIGURE 1.6. A Darboux partition in dimension $d = 2$ with $n_1 = 4$ and $n_2 = 6$.

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

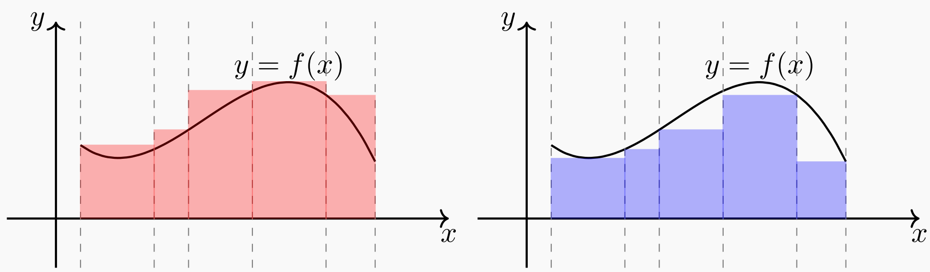

FIGURE 1.7. Upper (red) and lower (blue) Darboux sums of a function f over an interval ($d = 1$).

- 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 $a \le b$. A *box* in $\mathbb{R}^d$ is a set of the form $B = \prod_{i=1}^d I_i$, where $I_1, \dots, I_d$ 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 $B_1, \dots, B_k$ 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 1.7

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

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

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.

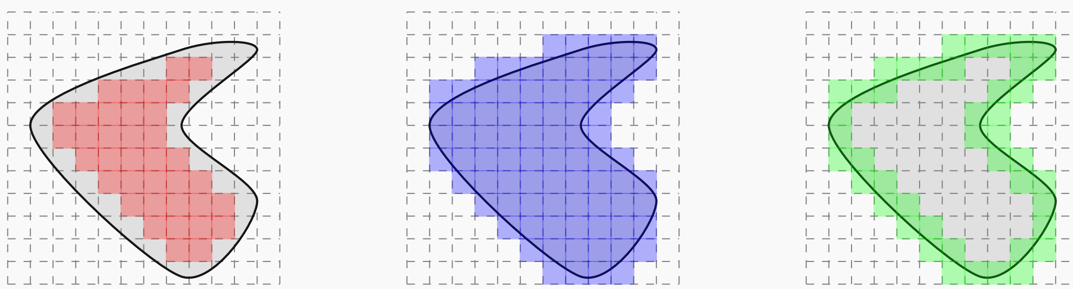

Figure 1.8. Simple sets approximating the inner (red) and outer Jordan content (blue) of a region in dimension $d = 2$. With the red boxes removed from the blue, we get a simple set covering the boundary (in green).

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.

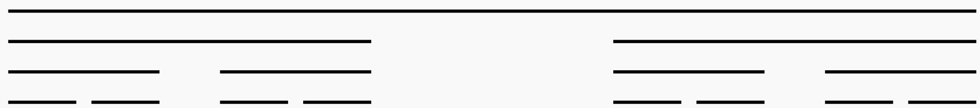

FIGURE 1.9. Iterative construction of the fat Cantor set.

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

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

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: $\text{Vol}([0,1)^d) = 1$;
- ISOMETRY-INVARIANT: if A and B are isometric, then $\text{Vol}(A) = \text{Vol}(B)$;
- COUNTABLY ADDITIVE: if $E_1, E_2, \dots$ are pairwise disjoint, then $\text{Vol}\left(\bigsqcup_{n \in \mathbb{N}} E_n\right) = \sum_{n=1}^{\infty} \text{Vol}(E_n)$.

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.

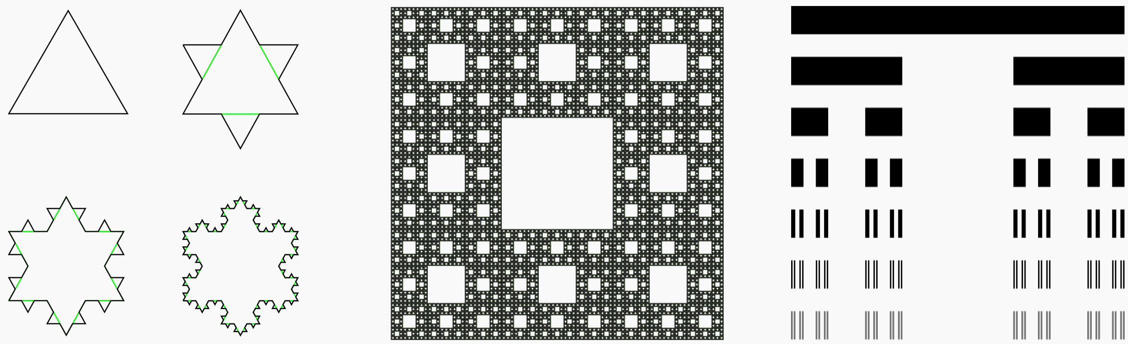

FIGURE 1.10. Fractal shapes: the Koch snowflake (left) of Hausdorff dimension $\frac{\log 4}{\log 3} \approx 1.26$, Sierpiński carpet (middle) of dimension $\frac{\log 8}{\log 3} \approx 1.89$, and middle-thirds Cantor set (right) of dimension $\frac{\log 2}{\log 3} \approx 0.63$.

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.

$$\begin{array}{ccccccccccccc}
\Sigma_1^0 = G & \xrightarrow{\bigcap} & \Pi_2^0 = G_\delta & \xrightarrow{\bigcup} & \Sigma_3^0 = G_{\delta\sigma} & \xrightarrow{\bigcap} & \Pi_4^0 = G_{\delta\sigma\delta} & \xrightarrow{\bigcup} & \dots & \xrightarrow{\bigcup} & \Sigma_\omega^0 & \xrightarrow{\bigcap} & \dots \\
\updownarrow c & & \updownarrow c & & \updownarrow c & & \updownarrow c & & & & \updownarrow c & & \\
\Pi_1^0 = F & \xrightarrow{\bigcup} & \Sigma_2^0 = F_\sigma & \xrightarrow{\bigcap} & \Pi_3^0 = F_{\sigma\delta} & \xrightarrow{\bigcup} & \Sigma_4^0 = F_{\sigma\delta\sigma} & \xrightarrow{\bigcap} & \dots & \xrightarrow{\bigcap} & \Pi_\omega^0 & \xrightarrow{\bigcup} & \dots
\end{array}$$

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 $\mathcal{B} = \text{Borel}(X)$ 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. □

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**].

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.

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

[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