Measure theory in a nutshell

Ulrik Skre Fjordholm

May 4, 2026

About this document

This text is meant as a supplement to Tom Lindstrøm's *Spaces*. Only the theory is presented here, so for explanations, motivation and exercises, please consult *Spaces*. For most proofs, only sketches and ideas are given. To get the most out of this material, try to fill in all the details in each proof on your own.

Chapter 1 corresponds to Chapter 7 in *Spaces*; each subsection therein corresponds roughly to those in *Spaces.* Chapter 2 consists of a sketch of the construction of the Lebesgue measure (see Chapter 8 in *Spaces*, or Chapter 9 of R. G. Bartle, *The Elements of Integration and Lebesgue Measure*), as well as a comparison of the Riemann and Lebesgue integrals.

Chapter 1

Measure and integration

1.1 Measure spaces

Summary. For a set X, we wish to measure the "size" of X and its subsets. A *σ-algebra* is the collection of subsets which we will allow ourselves to measure (the *measurable sets*), and a *measure* is a function which computes the size of a set. We show that measures satisfy several very natural properties. Finally, we show that any collection of subsets of X generates a unique smallest σ-algebra. In particular, the *Borel σ-algebra* on a metric space is the σ-algebra generated by the collection of open sets.

Definition 1.1.1. Let X be a set. A *σ-algebra on X* is a collection $\mathcal{A}$ of subsets of X satisfying:

(i) $\emptyset \in \mathcal{A}$

(ii) if $E \in \mathcal{A}$ then also $E^c \in \mathcal{A}$ (where $E^c := X \setminus E$)

(iii) if $E_1, E_2, \ldots \in \mathcal{A}$ then also $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$.

The pair $(X, \mathcal{A})$ is a *measurable space*. Each set $E \in \mathcal{A}$ is *$\mathcal{A}$-measurable*, or simply *measurable*.

Proposition 1.1.2. *Let $(X, \mathcal{A})$ be a measurable set. Then*

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

(ii) if $E_1, E_2, \ldots \in \mathcal{A}$, then also $\bigcap_{n=1}^{\infty} E_n \in \mathcal{A}$

(iii) if $E_1, \ldots, E_N \in \mathcal{A}$ then also $\bigcup_{n=1}^{N} E_n \in \mathcal{A}$ and $\bigcap_{n=1}^{N} E_n \in \mathcal{A}$

(iv) if $E, F \in \mathcal{A}$ then also $E \setminus F \in \mathcal{A}$ (where $E \setminus F := E \cap F^c$).

Idea of proof. Use De Morgan's laws

$$\left(\bigcup_{n\in I} E_n\right)^c = \bigcap_{n\in I} E_n^c, \qquad \left(\bigcap_{n\in I} E_n\right)^c = \bigcup_{n\in I} E_n^c$$

for any family of sets $(E_n)_{n\in I}$. For *(iii)*, set $E_{N+1}, E_{N+2}, \ldots = \emptyset$. ∎

We will be working with *extended real numbers*, partially treating ∞ and $-\infty$ as numbers. We let $\overline{\mathbb{R}} := [-\infty, \infty] = \mathbb{R} \cup \{-\infty, \infty\}$ and $\overline{\mathbb{R}}_+ := [0, \infty] = [0, \infty) \cup \{\infty\}$. We can extend the usual rules of addition, subtraction and multiplication from $\mathbb{R}$ to $\overline{\mathbb{R}}$, as long as we are careful to avoid the expression "$\infty - \infty$". Thus, for every $x \in \overline{\mathbb{R}}$ we let

$$\begin{cases} x + \infty = \infty & \text{for } x \neq -\infty, \\ x - \infty = -\infty & \text{for } x \neq \infty. \end{cases}$$

Definition 1.1.3. Let $(X, \mathcal{A})$ be a measurable space. A *measure* on $(X, \mathcal{A})$ is a function $\mu \colon \mathcal{A} \to \overline{\mathbb{R}}_+$ satisfying

(i) $\mu(\emptyset) = 0$

(ii) if $E_1, E_2, \ldots \in \mathcal{A}$ are disjoint, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n). \tag{1.1}$$

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

Note that both sides of (1.1) could be ∞.

Proposition 1.1.4. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

(i) if $E_1, \ldots, E_N \in \mathcal{A}$ then $\mu\left(\bigcup_{n=1}^{N} E_n\right) = \sum_{n=1}^{N} \mu(E_n)$

(ii) if $E, F \in \mathcal{A}$ and $F \subseteq E$, then $\mu(F) \leqslant \mu(E)$

(iii) if $E, F \in \mathcal{A}$, $F \subseteq E$ and $\mu(F) < \infty$, then $\mu(E \setminus F) = \mu(E) - \mu(F)$

(iv) if $E_1, E_2, \ldots \in \mathcal{A}$, then $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) \leqslant \sum_{n=1}^{\infty} \mu(E_n)$.

Idea of proof. For *(ii)* and *(iii)*, write E as the disjoint union $E = (E \setminus F) \cup F$. For *(iv)*, replace the sets by $F_1 := E_1$ and $F_n := E_n \setminus F_{n-1}$ (for $n \geqslant 2$). ■

Proposition 1.1.5 (Continuity of measures)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $E_1, E_2, \ldots \in \mathcal{A}$.*

(i) If $E_1 \subseteq E_2 \subseteq \cdots$ then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n).$$

(ii) If $E_1 \supseteq E_2 \supseteq \cdots$ and $\mu(E_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{n \to \infty} \mu(E_n).$$

Idea of proof. For *(i)*, write $\bigcup_{n=1}^{\infty} E_n$ as the disjoint union $\bigcup_{n=1}^{\infty} F_n$, where $F_1 := E_1$ and $F_n := E_n \setminus E_{n-1}$ (for $n \geqslant 2$). ■

Proposition 1.1.6. *Let X be a set and let $\mathcal{B}$ be a collection of subsets of X. Then there exists a σ-algebra, which we denote as $\sigma(\mathcal{B})$, satisfying*

(i) $\mathcal{B} \subseteq \sigma(\mathcal{B})$

(ii) if $\mathcal{A}$ is any σ-algebra such that $\mathcal{B} \subseteq \mathcal{A}$, then $\sigma(\mathcal{B}) \subseteq \mathcal{A}$.

We call $\sigma(\mathcal{B})$ the σ-algebra generated by $\mathcal{B}$.

Idea of proof. Let

$$\sigma(\mathcal{B}) := \bigcap_{\substack{\mathcal{A} \text{ is a } \sigma\text{-algebra} \\ \mathcal{A} \supseteq \mathcal{B}}} \mathcal{A}. \qquad \blacksquare$$

Thus, for any collection of sets $\mathcal{B}$ there is a "smallest" σ-algebra containing $\mathcal{B}$.

Definition 1.1.7. For a metric space (X, d), the *Borel σ-algebra* is the σ-algebra

$$\sigma\big(\{\text{all open subsets of } X\}\big).$$

The sets in this σ-algebra are the *Borel measurable sets*.

Remark 1.1.8. If we equip $\mathbb{R}$ with the standard metric $d(x, y) := |x - y|$, then every interval $(a, b) \subseteq \mathbb{R}$ is open, and therefore Borel measurable. Hence, also closed intervals $[a, b] \subseteq \mathbb{R}$ are Borel measurable (exercise!). The *Lebesgue measure* is a measure μ on the Borel σ-algebra on $\mathbb{R}$ satisfying

$$\mu([a, b]) = b - a \qquad \forall\, a, b \in \mathbb{R},\ a \leqslant b$$

(so the measure of an interval is the length of the interval). It is not at all obvious that such a measure μ exists.

1.2 Complete measure spaces

Summary. A measurable set N can be thought of as being "small" if $\mu(N) = 0$. For instance, $\emptyset$ is small in this measure theoretic sense. It would be reasonable to expect that if N is small and $M \subseteq N$, then M is also small. However, M might not be measurable! Luckily, there is a rather simple procedure to add these missing sets, leading to a *complete measure space* (not to be confused with a complete metric space).

The theory in this section is not used until the next chapter (Chapter 8 in *Spaces*), and so may safely be skipped for now.

Definition 1.2.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. A set $M \subseteq X$ is a *null set* if there is some $N \in \mathcal{A}$ such that $M \subseteq N$ and $\mu(N) = 0$. We denote

$$\mathcal{N} := \{\text{all null sets in } X\}.$$

The measure space $(X, \mathcal{A}, \mu)$ is *complete* if $\mathcal{A}$ contains all null sets.

Teorem 1.2.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let*

$$\overline{\mathcal{A}} := \{E \cup N : E \in \mathcal{A},\ N \in \mathcal{N}\}.$$

Define $\overline{\mu}\colon \overline{\mathcal{A}} \to \overline{\mathbb{R}}_+$ by

$$\overline{\mu}(E \cup N) := \mu(E) \qquad \text{for } E \cup N \in \overline{\mathcal{A}}.$$

Then $(X, \overline{\mathcal{A}}, \overline{\mu})$ is a complete measure space, and $\overline{\mu}(E) = \mu(E)$ for all $E \in \mathcal{A}$.

Idea of proof. First, show that $\overline{\mathcal{A}}$ is a σ-algebra on X. It is a good idea to start by proving that $\bigcup_{n\in\mathbb{N}} N_n \in \mathcal{N}$ for all $N_1, N_2, \ldots \in \mathcal{N}$. Next, show that $\overline{\mu}$ is well-defined — that is, if a set $E \cup N \in \overline{\mathcal{A}}$ can also be written as $E \cup N = F \cup M$ (for $E, F \in \mathcal{A}$ and $N, M \in \mathcal{N}$), then $\mu(E) = \mu(F)$. Checking that $\overline{\mu}$ is a measure is rather straightforward. Finally, one checks that $(X, \overline{\mathcal{A}}, \overline{\mu})$ is complete. ■

1.3 Measurable functions

Summary. On a measurable space X we introduce *measurable functions*, which are those functions $f : X \to \mathbb{R}$ which we eventually will be able to integrate. The first example of a measurable function is the *indicator function* over a measurable set E, that is, the function $\mathbb{1}_E : X \to \mathbb{R}$,

$$\mathbb{1}_E(x) := \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{else.} \end{cases}$$

Linear combinations of indicator functions gives rise to *simple functions*, which are also measurable.

Although the concept of measurability of functions is difficult to grasp at first, we deduce many different equivalent definitions of measurability, of which the final one is the most enlightening: The measurable functions are precisely those that can be approximated by simple functions.

The purpose of the rest of this chapter is to develop a theory of integration of functions $f : X \to \mathbb{R}$, where X is equipped with a σ-algebra. It will be convenient to develop this theory for functions which might take the values ∞ or $-\infty$, and we will therefore consider *extended real-valued functions* $f : X \to \overline{\mathbb{R}}$.

Remark 1.3.1. For a function $f : X \to Y$, we define the *preimage* of a subset $B \subseteq Y$ as

$$f^{-1}(B) := \{x \in X : f(x) \in B\}.$$

It is a good exercise to check the following properties: For any $B, C \subseteq Y$,

$$f^{-1}(B \setminus C) = f^{-1}(B) \setminus f^{-1}(C), \qquad f^{-1}(B^c) = f^{-1}(B)^c,$$

and if $(B_n)_{n\in I}$ is a collection of subsets of Y, then

$$f^{-1}\left(\bigcup_{n\in I} B_n\right) = \bigcup_{n\in I} f^{-1}(B_n), \qquad f^{-1}\left(\bigcap_{n\in I} B_n\right) = \bigcap_{n\in I} f^{-1}(B_n).$$

Definition 1.3.2. Let $(X, \mathcal{A})$ be a measurable space. A function $f : X \to \overline{\mathbb{R}}$ is $\mathcal{A}$-*measurable* (or simply *measurable*) if

$$f^{-1}\big([-\infty, r)\big) \in \mathcal{A} \qquad \forall\, r \in \mathbb{R}, \tag{1.2}$$

that is, the set $\{x \in X : f(x) < r\}$ is $\mathcal{A}$-measurable for any $r \in \mathbb{R}$. We denote the set of measurable functions by

$$M(X, \mathcal{A}) := \big\{\text{all measurable } f : X \to \overline{\mathbb{R}}\big\}$$

and the set of non-negative measurable functions by

$$M_+(X, \mathcal{A}) := \big\{\text{all measurable } f : X \to \overline{\mathbb{R}}_+\big\}.$$

Example 1.3.3. The indicator function $\mathbb{1}_E$ of a set $E \subseteq X$ is measurable *if and only if* E is measurable.

Lemma 1.3.4. *Let $(X, \mathcal{A})$ be a measurable space, and let $f : X \to \overline{R}$. Then the following are equivalent:*

(i) f *is measurable*
(ii) $f^{-1}\big([-\infty, r]\big) \in \mathcal{A}$ *for all* $r \in \mathbb{R}$
(iii) $f^{-1}\big([r, \infty]\big) \in \mathcal{A}$ *for all* $r \in \mathbb{R}$
(iv) $f^{-1}\big((r, \infty]\big) \in \mathcal{A}$ *for all* $r \in \mathbb{R}$.

In addition, if f does not take the values ∞ and $-\infty$, then the following are also equivalent to the above:

(v) $f^{-1}\big([a, b]\big) \in \mathcal{A}$ *for all* $a, b \in \mathbb{R}$
(vi) $f^{-1}\big((a, b]\big) \in \mathcal{A}$ *for all* $a, b \in \mathbb{R}$
(vii) $f^{-1}\big([a, b)\big) \in \mathcal{A}$ *for all* $a, b \in \mathbb{R}$
(viii) $f^{-1}\big((a, b)\big) \in \mathcal{A}$ *for all* $a, b \in \mathbb{R}$
(ix) $f^{-1}(E) \in \mathcal{A}$ *for all open* $E \subseteq \mathbb{R}$
(x) $f^{-1}(F) \in \mathcal{A}$ *for all closed* $F \subseteq \mathbb{R}$.

If f does *take the values ∞ or $-\infty$, then the equivalence still holds if we add the condition $f^{-1}(\{\infty\}) \in \mathcal{A}$ to (v)–(x).*

Idea of proof. For *(ii)*, write $[-\infty, r] = \bigcap_{n\in\mathbb{N}} [-\infty, r + 1/n)$. For *(iii)*, write $[r, \infty] = [-\infty, r)^c$. Use the elementary properties of the preimage from Remark 1.3.1. For *(ix)*, write an open set E as the countable union $E = \bigcup_{\substack{p,q\in\mathbb{Q} \\ (p,q)\subseteq E}} (p, q)$. ■

Proposition 1.3.5. *Let (X, d) be a metric space, $f : X \to \mathbb{R}$ a continuous function, and $\mathcal{A}$ a σ-algebra on X containing all open sets (such as the Borel σ-algebra). Then f is $\mathcal{A}$-measurable.*

Idea of proof. Use the fact that f is continuous if and only if $f^{-1}(E)$ is open for every open $E \subseteq \mathbb{R}$ (see Prop. 3.3.10 in *Spaces*). ■

Proposition 1.3.6. *Let $(X, \mathcal{A})$ be a measurable space, let $f, g \in M(X, \mathcal{A})$, and assume that g only takes real values. Then $f+g$, $f-g$, fg, $|f|$, αf and $\varphi \circ f$ are measurable for any $\alpha \in \mathbb{R}$ and continuous function $\varphi \colon \mathbb{R} \to \mathbb{R}$.*

Idea of proof. For, say, $\alpha > 0$ write $(\alpha f)^{-1}([-\infty, r)) = f^{-1}([-\infty, r/\alpha))$. For $f+g$, write

$$(f+g)^{-1}\big([-\infty, r)\big) = \bigcup_{s \in \mathbb{Q}} f^{-1}\big([-\infty, s)\big) \cap g^{-1}\big([-\infty, r-s)\big).$$

■

Example 1.3.7. If $E_1, \ldots, E_N \in \mathcal{A}$ and $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$, then the function

$$\varphi := \sum_{n=1}^{N} \alpha_n \mathbb{1}_{E_n}$$

is measurable. Functions of the above form are called *simple functions.*

Definition 1.3.8. Let $(X, \mathcal{A}, \mu)$ be a measure space. We say that some property $P(x)$ holds *μ-almost everywhere* (or *for μ-almost every x*, or simply *a.e.*) if there is a measurable set $N \subseteq X$ with measure $\mu(N) = 0$ such that $P(x)$ is true for all $x \notin N$. Some common uses are:

(i) a function $f : X \to \overline{\mathbb{R}}$ is *zero almost everywhere* if

$$\mu\big(\{x \in X : f(x) \neq 0\}\big) = 0$$

(ii) a function $f : X \to \overline{\mathbb{R}}$ is *finite almost everywhere* if

$$\mu\big(\{x \in X : f(x) \notin \mathbb{R}\}\big) = 0$$

(iii) two functions $f, g \colon X \to \overline{\mathbb{R}}$ are *equal almost everywhere* if

$$\mu\big(\{x \in X : f(x) \neq g(x)\}\big) = 0$$

(iv) a sequence of functions $f_n \colon X \to \overline{\mathbb{R}}$ *converges to* $f : X \to \overline{\mathbb{R}}$ *almost everywhere* if

$$\mu\big(\{x \in X : f_n(x) \not\to f(x) \text{ as } n \to \infty\}\big) = 0.$$

Remark 1.3.9. For a sequence of real numbers $\alpha_n \in \mathbb{R}$ (for $n \in \mathbb{N}$), the limit $\lim_{n\to\infty} \alpha_n$ might or might not exist. The *limit inferior* of the sequence is

$$\liminf_{n\to\infty} \alpha_n := \lim_{n\to\infty} \inf_{k \geqslant n} \alpha_k.$$

Since the sequence $\big(\inf_{k \geqslant n} \alpha_n\big)_{n \in \mathbb{N}}$ is increasing, then its limit always exists (at least as an extended real number — it might converge to ∞). Similarly, the *limit superior* is the extended real number

$$\limsup_{n\to\infty} \alpha_n := \lim_{n\to\infty} \sup_{k \geqslant n} \alpha_k.$$

If $(\alpha_n)_{n\in\mathbb{N}}$ does converge, then

$$\liminf_{n\to\infty} \alpha_n = \limsup_{n\to\infty} \alpha_n = \lim_{n\to\infty} \alpha_n,$$

while if the sequence does *not* converge, then

$$\liminf_{n\to\infty} \alpha_n < \limsup_{n\to\infty} \alpha_n.$$

For a sequence of functions, we define "inf", "sup", "lim inf" and "lim sup" pointwise, so that e.g.

$$f = \liminf_{n\to\infty} f_n \qquad \text{means} \qquad f(x) = \liminf_{n\to\infty} f_n(x) \quad \forall\, x \in X.$$

Lemma 1.3.10. *Let $(X, \mathcal{A})$ be a measurable space and let $f_n \in M(X, \mathcal{A}, \overline{\mathbb{R}})$ for $n \in \mathbb{N}$. Then the functions*

$$\inf_{n\in\mathbb{N}} f_n, \qquad \sup_{n\in\mathbb{N}} f_n, \qquad \liminf_{n\to\infty} f_n, \qquad \limsup_{n\to\infty} f_n$$

are all measurable. If $(f_n)_{n\in\mathbb{N}}$ converges pointwise to a function f, then f is measurable.

Idea of proof. Write $\big(\inf_{n\in\mathbb{N}} f_n\big)^{-1}\big([r,\infty]\big) = \bigcap_{n\in\mathbb{N}} f_n^{-1}\big([r,\infty]\big)$, and similarly for the supremum. For the limit inferior, note that we can write $\liminf_{n\to\infty} f_n(x) = \sup_{n\in\mathbb{N}} \inf_{k\geqslant n} f_k(x)$. Being the supremum of an infimum of a sequence of measurable functions, the limit inferior is measurable. ■

Definition 1.3.11. A function $\varphi\colon X \to \mathbb{R}$ is a *simple function* if it takes only finitely many values.

Letting $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$ be the (distinct) values attained by φ, and defining $E_n := \varphi^{-1}(\{\alpha_n\}) = \{x \in X : \varphi(x) = \alpha_n\}$, we see that the simple functions are those that can be written as

$$\varphi = \sum_{n=1}^{N} \alpha_n \mathbb{1}_{E_n} \tag{1.3}$$

for *distinct* $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and *disjoint* $E_1, \ldots, E_N$. (See also Example 1.3.7.) The *standard form* of a simple function φ is the expression (1.3) for distinct scalars α_n and disjoint, measurable sets E_n.

Teorem 1.3.12. *Let $(X, \mathcal{A})$ be a measurable space. Then a function $f : \mathcal{A} \to \overline{\mathbb{R}}_+$ is measurable if and only if there exist simple functions $\varphi_n\colon X \to \mathbb{R}_+$ (for $n \in \mathbb{N}$) such that*

(i) $0 \leqslant \varphi_n \leqslant \varphi_{n+1} \leqslant f$ *for all* $n \in \mathbb{N}$

(ii) $(\varphi_n)_{n\in\mathbb{N}}$ *converges pointwise to* f.

Idea of proof. Necessity follows from the final part of Lemma 1.3.10. For sufficiency, define

$$\varphi_n := \sum_{k=0}^{4^n} 2^{-n} k \, \mathbb{1}_{E_k},$$

where $E_k := f^{-1}\big([2^{-n}k, 2^{-n}(k+1))\big)$ (for $k = 0, \ldots, 4^n - 1$) and $E_{4^n} := f^{-1}\big([2^n, \infty)\big)$. ∎

Thus, the measurable, nonnegative functions are precisely those that can be approximated (in the sense of pointwise convergence) by increasing sequences of nonnegative, simple functions.

1.4 Integration of simple functions

Summary. Starting with the intuition that the integral of a constant over a set should equal the constant times the size of the set, we find that the integral of an indicator function $\mathbb{1}_E$ should be $\mu(E)$. Demanding that the integral should be linear, this yields the integral of simple functions, which are linear combinations of indicator functions.

To avoid ill-defined expressions of the form "$\infty - \infty$", we will only consider *nonnegative* functions, and later extend to arbitrary functions.

Definition 1.4.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $\varphi \in M_+(X, \mathcal{A})$ be a simple function with standard form $\varphi = \sum_{n=1}^{N} \alpha_n \mathbb{1}_{E_n}$. The *integral of φ with respect to μ* is the nonnegative number

$$\int \varphi \, d\mu := \sum_{n=1}^{N} \alpha_n \mu(E_n).$$

(Note that the integral might be ∞.)

Simple functions with standard form $\varphi = \sum_{n=1}^{N} \alpha_n \mathbb{1}_{E_n}$ can also be written in many other forms, say, $\sum_{m=1}^{M} \beta_m \mathbb{1}_{F_m}$ for scalars β_m and measurable sets F_m, where either some of the scalars are not distinct, or some of the sets F_m are not disjoint. It is not too hard to show that the formula for the integral is still valid:

$$\int \varphi \, d\mu := \sum_{n=1}^{N} \alpha_n \mu(E_n) = \sum_{m=1}^{M} \beta_m \mu(F_m).$$

Proposition 1.4.2. *Let $\varphi, \psi \in M_+(X, \mathcal{A})$ be simple functions. Then*

(i)

$$\int \varphi + \psi \, d\mu = \int \varphi \, d\mu + \int \psi \, d\mu$$

(ii) for any $\alpha \in \mathbb{R}_+$

$$\int \alpha\varphi \, d\mu = \alpha \int \varphi \, d\mu$$

(iii) if $\varphi \leqslant \psi$ then

$$\int \varphi \, d\mu \leqslant \int \psi \, d\mu.$$

1.5 Integration of nonnegative functions

Summary. We now have a consistent theory of integrals of simple functions. Since the measurable functions are limits of simple functions, we now take limits of integrals of simple functions, arriving at the integral of an arbitrary (nonnegative) measurable function. From our construction, this new concept of integrals satisfies several nice, intuitive properties: The integral is linear, and is compatible with taking (certain) pointwise limits.

Definition 1.5.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f \in M_+(X, \mathcal{A})$. The *integral of f with respect to μ* (or just the *integral of f*) is the nonnegative number

$$\int f \, d\mu \coloneqq \sup\left\{ \int \varphi \, d\mu : \varphi \text{ is simple and } 0 \leqslant \varphi \leqslant f \right\}.$$

For a set $E \in \mathcal{A}$, the *integral of f over E* is

$$\int_E f \, d\mu \coloneqq \int f \mathbb{1}_E \, d\mu.$$

Lemma 1.5.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Then*

(i) if $f, g \in M_+(X, \mathcal{A})$ and $f \leqslant g$ then $\int f \, d\mu \leqslant \int g \, d\mu$
(ii) if $f \in M_+(X, \mathcal{A})$, $E, F \in \mathcal{A}$ and $E \subseteq F$, then $\int_E f \, d\mu \leqslant \int_F f \, d\mu$.

Idea of proof. For *(i)*, if $\varphi \leqslant f$ is a simple function, then also $\varphi \leqslant g$. For *(ii)*, use the inequality $f \mathbb{1}_E \leqslant f \mathbb{1}_F$ in *(i)*. ■

Lemma 1.5.3. *Let $(X, \mathcal{A}, \mu)$ be a measure space, let $\varphi \in M_+(X, \mathcal{A})$ be simple and define*

$$\lambda(E) \coloneqq \int_E \varphi \, d\mu.$$

Then λ is a measure on $(X, \mathcal{A})$.

Teorem 1.5.4 (The Monotone Convergence Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $(f_n)_{n \in \mathbb{N}} \in M_+(X, \mathcal{A})$ be an increasing sequence, that is, each f_n is measurable, and $0 \leqslant f_n \leqslant f_{n+1}$ for all $n \in \mathbb{N}$. Define $f(x) \coloneqq \lim_{n \to \infty} f_n(x)$ for all $x \in X$. Then f is well-defined and measurable, and*

$$\int f \, d\mu = \lim_{n \to \infty} \int f_n \, d\mu. \tag{1.4}$$

Sketch of proof. Lemma 1.5.2 yields "$\geqslant$" in (1.4). For "$\leqslant$", let $\alpha \in (0, 1)$ and let $\varphi \leqslant f$ be simple. Define $A_n \coloneqq \{x \in X : \alpha\varphi(x) \leqslant f_n(x)\}$ (for $n \in \mathbb{N}$). Then $A_n \subseteq A_{n+1}$, and $\bigcup_{n \in \mathbb{N}} A_n = X$. We get

$$\int_{A_n} \alpha\varphi \, d\mu \leqslant \int_{A_n} f_n \, d\mu \leqslant \int f_n \, d\mu. \tag{$*$}$$

Since $\lambda(E) \coloneqq \int_E \varphi \, d\mu$ is a measure (Lemma 1.5.3), Proposition 1.1.5 gives

$$\int \varphi \, d\mu = \lambda\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \lambda(A_n) = \lim_{n\to\infty} \int_{A_n} \varphi \, d\mu.$$

Thus, passing $n \to \infty$ in $(*)$ gives

$$\alpha \int \varphi \, d\mu \leqslant \lim_{n\to\infty} \int f_n \, d\mu.$$

Since α and φ were arbitrary, we can conclude (1.4). ■

Corollary 1.5.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, let $f \in M_+(X, \mathcal{A})$, and let $\varphi_1, \varphi_2, \ldots$ be the simple functions from Theorem 1.3.12. Then*

$$\int f \, d\mu = \lim_{n\to\infty} \int \varphi_n \, d\mu.$$

Corollary 1.5.6. *Let $f, g \in M_+(X, \mathcal{A})$. Then*

(i)

$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu$$

(ii) for any $\alpha \in \mathbb{R}_+$

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu.$$

Sketch of proof. For *(i)*, let $(\varphi_n)_{n\in\mathbb{N}}$ and $(\psi_n)_{n\in\mathbb{N}}$ be increasing sequences of simple functions converging to f and g, respectively. Then $(\varphi_n + \psi_n)_{n\in\mathbb{N}}$ is an increasing sequence of simple functions converging to $f + g$. The result now follows from Corollary 1.5.5. ■

Teorem 1.5.7 (Fatou's lemma)**.** *Let $f_1, f_2, \ldots \in M_+(X, \mathcal{A})$. Then*

$$\int \liminf_{n\to\infty} f_n \, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu.$$

Sketch of proof. Let $g_n \coloneqq \inf_{k\geqslant n} f_k$. By the Monotone Convergence Theorem,

$$\int \liminf_{n\to\infty} f_n \, d\mu = \int \lim_{n\to\infty} g_n \, d\mu = \lim_{n\to\infty} \int g_n \, d\mu.$$

Since $g_n \leqslant f_k$ for all $k \geqslant n$, we get $\int g_n \, d\mu \leqslant \inf_{k\geqslant n} \int f_k \, d\mu$. Pass $n \to \infty$ to conclude. ■

Corollary 1.5.8. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

(i) Let $f \in M_+(X, \mathcal{A})$. Then $f = 0$ μ-a.e. if and only if $\int f \, d\mu = 0$.

(ii) Let $g \in M(X, \mathcal{A})$. Then $g = 0$ μ-a.e. if and only if $\int |g| \, d\mu = 0$.

(iii) Let $h_1, h_2 \in M(X, \mathcal{A})$. Then $h_1 = h_2$ μ-a.e. if and only if $\int |h_1 - h_2| \, d\mu = 0$.

Proof. For *(i)*, let $E := \{x : f(x) > 0\}$. If $f = 0$ μ-a.e. then $\mu(E) = 0$. Since $f \leqslant \liminf_{n\to\infty} n\mathbb{1}_E$, we get

$$\begin{aligned} 0 \leqslant \int f\,d\mu &\leqslant \int \liminf_{n\to\infty} n\mathbb{1}_E\,d\mu \\ &\leqslant \liminf_{n\to\infty} \int n\mathbb{1}_E\,d\mu = \liminf_{n\to\infty} n\mu(E) = 0. \end{aligned}$$

Conversely, if $\int f\,d\mu = 0$, let $E_n := \{x : f(x) > 1/n\}$. Then $f \geqslant \mathbb{1}_E/n$, so

$$0 = \int f\,d\mu \geqslant \int \mathbb{1}_E/n\,d\mu = \mu(E_n)/n,$$

whence $\mu(E_n) = 0$. Thus, $\mu(E) = \mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) \leqslant \sum_{n=1}^{\infty} \mu(E_n) = 0$.

For *(ii)*, let $f := |g|$, and observe that $g = 0$ μ-a.e. if and only if $|g| = 0$ μ-a.e.

For *(iii)*, let $g := h_1 - h_2$. ∎

1.6 Integrable functions

Summary. Knowing how to integrate nonnegative functions, we now write an arbitrary function as the difference between nonnegative functions, leading to a natural definition of their integral. Along with several natural linearity and monotonicity properties, we prove the most important theorem about convergence of integrals and integrands, the Dominated Convergence Theorem.

Definition 1.6.1. Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f \in M_+(X, \mathcal{A})$ is *μ-integrable* (or just *integrable*) if $\int f\,d\mu < \infty$.

For $f \in M(X, \mathcal{A})$ we define its *positive part* and *negative part,*

$$f_+(x) := \begin{cases} f(x) & \text{if } f(x) > 0 \\ 0 & \text{else,} \end{cases} \qquad f_-(x) := \begin{cases} -f(x) & \text{if } f(x) < 0 \\ 0 & \text{else.} \end{cases}$$

Note that $f = f_+ - f_-$, that $|f| = f_+ + f_-$, and that $f_- = (-f)_+$.

Definition 1.6.2. A function $f \in M(X, \mathcal{A})$ is *μ-integrable* (or just *integrable*) if f_+ and f_- are integrable; if so, we define its integral as

$$\int f\,d\mu := \int f_+\,d\mu - \int f_-\,d\mu.$$

We denote $\mathcal{L}^1(X, \mathcal{A}, \mu) := \big\{\text{all integrable } f \in M(X, \mathcal{A})\big\}$.

Teorem 1.6.3. *Let $(X, \mathcal{A}, \mu)$ be a measure space.*

(i) If $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ and $\alpha \in \mathbb{R}$, then $\alpha f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, with

$$\int \alpha f\,d\mu = \alpha \int f\,d\mu.$$

(ii) If $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, then $f + g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, with

$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

(iii) If $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ then $\left|\int f \, d\mu\right| \leqslant \int |f| \, d\mu$.

(iv) If $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ are such that $f \leqslant g$, then $\int f \, d\mu \leqslant \int g \, d\mu$.

(v) If $f \in M(X, \mathcal{A})$ satisfies $|f| \leqslant g$ for some nonnegative $g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, then $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$.

Sketch of proof. For *(ii)*, note first that $|f + g| \leqslant |f| + |g|$, and since the latter is integrable, $f + g$ is integrable. We have

$$f + g = (f_+ + g_+) - (f_- - g_-) = (f + g)_+ - (f + g)_-,$$

or after reorganizing,

$$f_+ + g_+ + (f + g)_- = f_- + g_- + (f + g)_+,$$

and since all terms are nonnegative, Corollary 1.5.6 yields

$$\int f_+ \, d\mu + \int g_+ \, d\mu + \int (f + g)_- \, d\mu = \int f_- \, d\mu + \int g_- \, d\mu + \int (f + g)_+ \, d\mu.$$

Reorganizing now yields *(ii)*. ■

Teorem 1.6.4 (Lebesgue's Dominated Convergence Theorem)**.** *Let $f, f_n \in M(X, \mathcal{A})$ (for $n \in \mathbb{N}$) be such that $f_n \to f$ as $n \to \infty$ almost everywhere, and assume that $|f_n| \leqslant g$ for all $n \in \mathbb{N}$ for some nonnegative $g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$. Then $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$, and*

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

Sketch of proof. Redefining f_n, f on a set of measure zero, we may assume that $f_n \to f$ pointwise everywhere. Then also $|f| \leqslant g$, whence $f \in \mathcal{L}^1$. Since $g + f_n \geqslant 0$, Fatou's lemma yields

$$\int g \, d\mu + \int f \, d\mu = \int \lim_{n\to\infty} (g + f_n) \, d\mu \leqslant \liminf_{n\to\infty} \int g + f_n \, d\mu.$$

Reorganizing yields $\int f \, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu$. A similar computation with $g - f \geqslant 0$ yields $\int f \, d\mu \geqslant \limsup_{n\to\infty} \int f_n \, d\mu$. This concludes the proof. ■

1.7 Spaces of integrable functions

Summary. The L^p spaces are the vector spaces of functions f for which $|f|^p$ is integrable. We prove that the L^p norm is indeed a norm, and that the resulting normed vector space is complete. Along the way we meet a technical obstacle: The integral can only tell the difference between functions *almost everywhere* (that is, outside of sets of measure zero), so we need to treat functions that are equal almost everywhere as one and the same function.

We say that $p, q \in [1, \infty]$ are *conjugate exponents* if

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

(where we interpret $1/\infty$ as 0). For $p \in [1, \infty)$, the *p-norm* of a vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is

$$\|x\|_p \coloneqq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p}.$$

We also define the ∞-norm on $\mathbb{R}^n$ as

$$\|x\|_\infty \coloneqq \max_{k=1,\ldots,n} |x_k|.$$

We recall that $\|\cdot\|_p$ is a norm on $\mathbb{R}^n$ for any $p \in [1, \infty]$, and moreover, we have *Hölder's inequality*: If $p, q \in [1, \infty]$ are conjugate exponents, then

$$\|xy\|_1 \leqslant \|x\|_p \|y\|_q \qquad \forall\ x, y \in \mathbb{R}^n \tag{1.5}$$

where $xy = (x_1 y_1, \ldots, x_n y_n)$.

Definition 1.7.1. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p \in [1, \infty)$. We define the *L^p-norm* as

$$\|f\|_{L^p} \coloneqq \left(\int |f|^p \, d\mu \right)^{1/p} \qquad \text{for } f \in M(X, \mathcal{A}). \tag{1.6}$$

We write $\mathcal{L}^p(X, \mathcal{A}, \mu) \coloneqq \big\{ f \in M(X, \mathcal{A}) : \|f\|_{L^p} < \infty \big\}$.

The case $p = \infty$ is somewhat more intricate.

Definition 1.7.2. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $f \in M(X, \mathcal{A})$. The *essential supremum* of f is

$$\operatorname{ess\,sup} f \coloneqq \inf\Big\{ \sup\big\{ f(x) : x \in X \setminus N \big\} : N \in \mathcal{A},\ \mu(N) = 0 \Big\}. \tag{1.7}$$

The *L^∞-norm* of f is

$$\|f\|_{L^\infty} \coloneqq \operatorname{ess\,sup} |f|,$$

and we let $\mathcal{L}^\infty(X, \mathcal{A}, \mu) \coloneqq \big\{ f \in M(X, \mathcal{A}) : \|f\|_{L^\infty} < \infty \big\}$.

It is not too hard to show that for every measurable function f, there exists a set N of measure zero such that $\operatorname{ess\,sup} f = \sup_{x \in X \setminus N} f(x)$, and that there exists no other sets of measure zero for which the right-hand side is larger. (Hint: Take a minimizing sequence $N_1, N_2, \ldots$ in (1.7) and consider its union.)

Recalling the requirements of being a norm, it is clear that $\|f\|_{L^p} \geqslant 0$ for all $f \in \mathcal{L}^p$, and that $\|\alpha f\|_{L^p} = |\alpha| \|f\|_{L^p}$ for all $f \in \mathcal{L}^p$ and $\alpha \in \mathbb{R}$. Proving the triangle inequality will, as usual, require more work. We return to the requirement that $\|f\|_{L^p} = 0$ implies $f = 0$ later.

Teorem 1.7.3 (Hölder's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p, q \in [1, \infty]$ be conjugate exponents. Then*

$$\|fg\|_{L^1} \leqslant \|f\|_{L^p} \|g\|_{L^q} \qquad \forall\ f \in \mathcal{L}^p,\ g \in \mathcal{L}^q. \tag{1.8}$$

In particular, $fg \in \mathcal{L}^1(X, \mathcal{A}, \mu)$.

Sketch of proof. The case $p = 1$, $q = \infty$ (or vice versa) is handed separately, and is simpler, so assume $p, q \in (1, \infty)$. If f, g are simple, write $f = \sum_{k=1}^K \alpha_k \mathbb{1}_{G_k}$, $g = \sum_{k=1}^K \delta_k \mathbb{1}_{G_k}$ for $\gamma_k, \delta_k \in \mathbb{R}$ and disjoint sets $G_k \in \mathcal{A}$. Then

$$\|fg\|_{L^1} = \sum_{k=1}^K |\gamma_k \delta_k| \mu(G_k) \overset{(\text{as } \frac{1}{p}+\frac{1}{q}=1)}{=} \sum_{k=1}^K \Big(|\gamma_k| \mu(G_k)^{1/p}\Big)\Big(|\delta_k| \mu(G_k)^{1/q}\Big)$$
$$\overset{(\text{Hölder})}{\leqslant} \left(\sum_{k=1}^K \Big(|\gamma_k| \mu(G_k)^{1/p}\Big)^p\right)^{1/p} \left(\sum_{k=1}^K \Big(|\delta_k| \mu(G_k)^{1/q}\Big)^q\right)^{1/q} = \|f\|_{L^p} \|g\|_{L^q}.$$

For general f, g, approximate them by increasing sequences of simple functions. ■

Teorem 1.7.4 (Minkowski's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p \in [1, \infty]$. If $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ then also $f + g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$, and*

$$\|f + g\|_{L^p} \leqslant \|f\|_{L^p} + \|g\|_{L^p}.$$

Idea of proof. Same idea as in Hölder's inequality: Prove it first for simple functions using the $\|\cdot\|_p$ norm on $\mathbb{R}^n$, and then approximate general functions by simple functions. ■

We can now conclude that $\|\cdot\|_{L^p}$ is a *seminorm* on $\mathcal{L}^p(X, \mathcal{A}, \mu)$: It satisfies the conditions of begin a norm *except* the implication $\|f\|_{L^p} = 0 \ \Rightarrow\ f = 0$. Indeed, if $\|f\|_{L^p} = 0$ then, by Corollary 1.5.8 *(ii)*, we can only conclude that $f = 0$ μ-almost everywhere.

Henceforth, we will treat two functions f, g as the same function if they are equal μ-almost everywhere. The set of all functions that equal $f \in M(X, \mathcal{A})$ μ-almost everywhere is denoted

$$[f] := \big\{g \in M(X, \mathcal{A}) : f = g\ \mu\text{-a.e.}\big\}$$

(this is the *equivalence class of* f), and we define the L^p-spaces (for $p \in [1, \infty]$) as

$$L^p(X, \mathcal{A}, \mu) := \big\{[f] : f \in \mathcal{L}^p(X, \mathcal{A}, \mu)\big\}$$

(or often just $L^p(X)$, if the σ-algebra and the measure are clear from the context). Since integrals cannot tell the difference between functions that are almost everywhere (that is, $\int f\, d\mu = \int g\, d\mu$ for any $g \in [f]$), the definition

$$\big\|[f]\big\|_{L^p} := \|g\|_{L^p} \qquad \text{for any } g \in [f]$$

is well-defined. For the most part, it is safe to think of the elements of $L^p(X, \mathcal{A}, \mu)$ as functions in the usual sense.

Corollary 1.7.5. *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p \in [1, \infty]$. Then $L^p(X, \mathcal{A}, \mu)$ is a vector space with norm $\| \cdot \|_{L^p}$.*

Idea of proof. One needs to check that addition and scalar multiplication are well-defined, that is, $[f] + [g] = [f + g]$ and $[\alpha f] = \alpha[f]$ for all $f, g \in \mathcal{L}^p$ and $\alpha \in \mathbb{R}$, and that the three conditions for being a norm are satisfied. The hard work has already been done, in Minkowski's inequality. ■

Teorem 1.7.6 (The Riesz–Fischer theorem)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space, and let $p \in [1, \infty]$. Then $\big(L^p(X, \mathcal{A}, \mu), \| \cdot \|_{L^p}\big)$ is a complete normed vector space.*

Sketch of proof. We look only at the (more difficult) case $p, q \in (1, \infty)$. Let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence, pick a subsequence $(f_{n_k})_{k \in \mathbb{N}}$ such that $\| f_{n_{k+1}} - f_{n_k} \|_{L^p} < 2^{-k}$, and denote $g_k := f_{n_k}$. Define $g := |g_1| + \sum_{k \geqslant 1} |g_{k+1} - g_k|$ and use Fatou's lemma to show that $\|g\|_{L^p} < \infty$. In particular, g is finite a.e., so the function

$$f(x) := g_1(x) + \sum_{k=1}^{\infty} g_{k+1}(x) - g_k(x) = \lim_{k \to \infty} g_k(x)$$

is well-defined and finite almost everywhere. The Dominated Convergence Theorem ensures that $g_k \to f$ in $\| \cdot \|_{L^p}$. Finally, since the original sequence $(f_n)_n$ is Cauchy, and it has a subsequence converging to f in $\| \cdot \|_{L^p}$, the whole sequence must converge to f in $\| \cdot \|_{L^p}$. ■

1.8 Modes of convergence

We have now several different modes of convergence for sequences of functions: Pointwise convergence, almost everywhere convergence, uniform convergence, and convergence in any of the L^p norms (for $p \in [1, \infty]$). It is natural to ask how these relate to one another. We will only look at a few examples of connections between these modes of convergence in the case of *finite measure spaces*, that is, when $\mu(X) < \infty$.

The first result relates uniform and L^p convergence:

Proposition 1.8.1. *Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$, and let $f_1, f_2, \ldots \in M(X, \mathcal{A})$ be a sequence of functions converging uniformly to some bounded function $f : X \to \mathbb{R}$. Then f is integrable, and $f_n \to f$ in $L^p(X, \mathcal{A}, \mu)$ for every $p \in [1, \infty]$.*

Proof. Since f is the pointwise limit of measurable functions, it is measurable, and since f is bounded, there is some $C > 0$ such that $|f(x)| \leqslant C$ for all $x \in X$, so

$$\int |f| \, d\mu \leqslant \int C \, d\mu = C\mu(X) < \infty,$$

so f is integrable. For any $p \in [1, \infty)$,

$$\| f_n - f \|_{L^p}^p = \int |f_n - f|^p \, d\mu \leqslant \mu(X) \sup_{x \in X} |f_n(x) - f(x)| \to 0$$

as $n \to \infty$, whereas for $p = \infty$,

$$\|f_n - f\|_{L^\infty} = \operatorname{ess\,sup}_{x \in X} |f_n(x) - f(x)| \leqslant \sup_{x \in X} |f_n(x) - f(x)| \to 0$$

as $n \to \infty$. ■

The second result relates L^p convergence for different values of p with one another. The proof is a rather typical application of Hölder's inequality.

Proposition 1.8.2. *Let $(X, \mathcal{A}, \mu)$ be a measure space with $\mu(X) < \infty$, let $p_1 \in [1, \infty]$, and let $f_1, f_2, \ldots \in M(X, \mathcal{A})$ be a sequence of functions converging in L^{p_1} to some function f. Then $(f_n)_{n \in \mathbb{N}}$ also converges in L^{p_2} for all $p_2 \leqslant p_1$.*

Proof. If $p_1 = \infty$ then $|f_n(x) - f(x)| \leqslant \|f_n - f\|_{L^\infty}$ for almost every x, so if $p_2 < \infty$ then

$$\|f_n - f\|_{L^{p_2}}^{p_2} = \int |f_n(x) - f(x)|^{p_2} \, d\mu \leqslant \int \|f_n - f\|_{L^\infty}^{p_2} \, d\mu = \|f_n - f\|_{L^\infty}^{p_2} \mu(X) \to 0$$

as $n \to \infty$. If $p_1 < \infty$ and $p_2 < p_1$ then we use Hölder's inequality with exponents $p \coloneqq p_1/p_2$ and $q = p_1/(p_1 - p_2)$ to get

$$\begin{aligned}
\|f_n - f\|_{L^{p_2}}^{p_2} &= \int |f_n - f|^{p_2} \, d\mu = \int |f_n - f|^{p_2} \cdot 1 \, d\mu \\
&\leqslant \left(\int \left| |f_n - f|^{p_2} \right|^{p_1/p_2} d\mu \right)^{p_2/p_1} \left(\int 1^{p_1/(p_1 - p_2)} \, d\mu \right)^{(p_1 - p_2)/p_1} \\
&\leqslant \left(\int |f_n - f|^{p_1} \, d\mu \right)^{p_2/p_1} \mu(X)^{(p_1 - p_2)/p_1} \\
&= \|f_n - f\|_{L^{p_1}}^{p_2} \mu(X)^{(p_1 - p_2)/p_1} \to 0
\end{aligned}$$

as $n \to \infty$. ■

Chapter 2

The Lebesgue measure

2.1 Constructing the Lebesgue measure

Summary. For $a, b \in \mathbb{R}$, $a < b$, the *length* of an interval with endpoints a and b (such as (a, b), $[a, b]$, etc.) is $b - a$. The *Lebesgue measure* is a measure on $\mathbb{R}$ which assigns to any interval its length. The goal of this section is to present a general procedure to construct measure spaces with certain prescribed properties, and in particular the Lebesgue measure.

Step 1: Measures on algebras

For $a, b \in \mathbb{R}$, $a < b$, let $\mathcal{R}$ be the collection of all finite unions of intervals of the form

$$(a, b], \qquad (-\infty, b], \qquad (a, \infty), \qquad (-\infty, \infty).$$

(It will be slightly easier to work with intervals $(a, b]$ rather than, say, (a, b).) We can easily assign a measure μ to any element of $\mathcal{R}$ corresponding to the length of an interval. If $R \in \mathcal{R}$ is the union of at least one infinitely long interval, then $\mu(R) = \infty$, while if not — say, $R = \bigcup_{n=1}^{N} (a_n, b_n] \in \mathcal{R}$ for disjoint intervals $(a_n, b_n]$ — then

$$\mu\left(\bigcup_{n=1}^{N} (a_n, b_n]\right) = \sum_{n=1}^{N} b_n - a_n.$$

The collection $\mathcal{R}$ is an example of an *algebra*, and μ is a *measure* on $\mathcal{R}$, in the following sense (compare these definitions carefully with Definitions 1.1.1 and 1.1.3!).

Definition 2.1.1. Let X be a set. An *algebra on X* is a collection $\mathcal{A}$ of subsets of X satisfying:

(i) $\emptyset \in \mathcal{A}$

(ii) if $E \in \mathcal{A}$ then also $E^c \in \mathcal{A}$

(iii) if $E_1, \ldots, E_N \in \mathcal{A}$ then also $\bigcup_{n=1}^{N} E_n \in \mathcal{A}$.

Definition 2.1.2. Let $\mathcal{A}$ be an algebra on X. A *measure* on $(X, \mathcal{A})$ is a function $\mu : \mathcal{A} \to \overline{\mathbb{R}}_+$ satisfying

(i) $\mu(\emptyset) = 0$

(ii) if $E_1, E_2, \ldots \in \mathcal{A}$ are disjoint and $\bigcup_{n=1}^{\infty} E_n \in \mathcal{A}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n). \tag{2.1}$$

Lemma 2.1.3. *Let $\mathcal{R}$ and μ be as described above. Then $\mathcal{R}$ is an algebra on $\mathbb{R}$, and μ is a measure on $\mathcal{R}$.*

Step 2: Extending measures

Given an algebra $\mathcal{A}$ on X with a measure μ, we now extend μ to a function μ^* defined on *all* subsets of X. The function μ^* will generally not be a measure, so we must restrict it to a smaller collection of sets $\mathcal{A}^*$. This will yield the measure we are after; in particular, $\mu^*(E) = \mu(E)$ for every $E \in \mathcal{A}$ (so μ^* is an *extension* of μ).

Definition 2.1.4. For an algebra $\mathcal{A}$ on a set X with a measure μ, the *outer measure* corresponding to μ is

$$\mu^*(F) := \inf\left\{\sum_{n=1}^{\infty} \mu(E_n) : E_1, E_2, \ldots \in \mathcal{A},\ F \subseteq \bigcup_{n=1}^{\infty} E_n\right\} \qquad \forall F \subseteq X.$$

Thus, the outer measure of a set F is the infimum of the combined measure of all coverings of F using sets in $\mathcal{A}$. The outer measure has several natural properties:

Lemma 2.1.5. *The outer measure μ^* satisfies*

(i) $\mu^*(\emptyset) = 0$

(ii) $\mu^*(F) \in \overline{\mathbb{R}}_+$ *for all* $F \subseteq X$

(iii) *if* $F \subseteq G$, *then* $\mu^*(F) \leqslant \mu^*(G)$

(iv) *if* $F \in \mathcal{A}$, *then* $\mu^*(F) = \mu(F)$

(v) *if* $F_1, F_2, \ldots \subseteq X$, *then*

$$\mu^*\left(\bigcup_{n=1}^{\infty} F_n\right) \leqslant \sum_{n=1}^{\infty} \mu^*(F_n). \tag{2.2}$$

Thus, μ^* is "almost" a measure, but in general there might be a strict inequality in (2.2), even when the sets $F_1, F_2, \ldots$ are disjoint. The next task is to find a collection of sets $\mathcal{A}^*$ which yield equality in (2.2) when the sets are disjoint.

Definition 2.1.6. Let X, $\mathcal{A}$, μ and μ^* be as above. A set $E \subseteq X$ is *μ^*-measurable* if

$$\mu^*(A) = \mu^*\big(A \cap E\big) + \mu^*\big(A \setminus E\big) \qquad \forall\ A \subseteq X.$$

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

One might think of μ^*-measurability of a set E as the property that E and its complement E^c are sufficiently nicely separated so that they split an arbitrary set A in an "additive" manner.

Teorem 2.1.7 (Carathéodory's extension theorem)**.** *Let X, $\mathcal{A}$, $\mathcal{A}^*$, μ and μ^* be as above. Then $(X, \mathcal{A}^*, \mu^*)$ is a complete measure space. Moreover, $(X, \mathcal{A}^*, \mu^*)$ is an extension of $(X, \mathcal{A}, \mu)$, in the sense that $\mathcal{A}^*$ contains $\mathcal{A}$, and $\mu^*(E) = \mu(E)$ for all $E \in \mathcal{A}$.*

It is also possible to obtain uniqueness of the extension μ^*, with one extra condition:

Teorem 2.1.8 (Hahn's extension theorem)**.** *Let X, $\mathcal{A}$, $\mathcal{A}^*$, μ and μ^* be as above. Assume in addition that μ is σ-finite, that is, there exist $E_1, E_2, \ldots \in \mathcal{A}$ such that $\mu(E_n) < \infty$ for all n, and $\bigcup_{n=1}^{\infty} E_n = X$. Then μ^* is the only extension of μ to $\mathcal{A}^*$.*

Step 3: Construction of the Lebesgue measure

We can now apply this theory to the construction of the Lebesgue measure on $\mathbb{R}$.

Teorem 2.1.9. *There exists a complete measure space $(\mathbb{R}, \mathcal{R}^*, \mu^*)$ such that $\mathcal{R}^*$ contains all intervals in $\mathbb{R}$ (an in particular, all open or closed sets), and the measure of every interval equals its length.*

Proof. Let $\mathcal{R}$ be the algebra of intervals constructed in Step 1, and let μ be the measure on $\mathcal{R}$ corresponding to length. According to Lemma 2.1.3, $\mathcal{R}$ is indeed an algebra, and μ is a measure on $\mathcal{R}$. Let $\mathcal{R}^*$, μ^* be as given in Carathéodory's extension theorem.

The observation that $\mathcal{R}^*$ contains all intervals, open sets and closed sets, follows from the same type of computations as in Lemma 1.3.4. For instance, $(a, b) = \bigcup_{n\in\mathbb{N}}(a, b - 1/n]$. ■

The σ-algebra $\mathcal{R}^*$ is the *Lebesgue σ-algebra on* $\mathbb{R}$, and the measure μ^* is the *Lebesgue measure on* $\mathbb{R}$. Note that the Lebesgue σ-algebra contains the Borel σ-algebra on $\mathbb{R}$, since the latter is the smallest σ-algebra containing all open sets.

We can restrict the Lebesgue measure (or any other measure) to a subset of $\mathbb{R}$, say, an interval $[a, b]$. Letting

$$\mathcal{R}^*\big|_{[a,b]} := \big\{E \in \mathcal{R}^* : E \subseteq [a, b]\big\}$$

and $\mu^*\big|_{[a,b]}(E) := \mu^*(E)$ for $E \in \mathcal{R}^*\big|_{[a,b]}$, it is straightforward to check that $\big([a, b],\ \mathcal{R}^*\big|_{[a,b]},\ \mu^*\big|_{[a,b]}\big)$ is a complete measure space. We almost always simplify this notation and write e.g. $\int_{[a,b]} f\, d\mu^*$ for the integral of a function $f : [a, b] \to \overline{\mathbb{R}}$.

2.2 Riemann integrals vs. Lebesgue integrals

Consider a function $f : [a, b] \to \mathbb{R}$. A *partition* of $[a, b]$ is a list $\mathcal{P} = (x_0, \ldots, x_N)$ such that

$$a = x_0 < x_1 < \cdots < x_N = b.$$

The *size* (or *resolution*) of a partition is denoted $|\mathcal{P}| := \max_{k=1,\dots,N} |x_k - x_{k-1}|$. A *Riemann sum* of f is a sum of the form

$$\sum_{k=1}^{N} f(y_k)(x_k - x_{k-1})$$

for $y_k \in [x_{k-1}, x_k]$ (for $k = 1, \dots, N$), and where $\mathcal{P} = (x_0, \dots, x_N)$ is a partition of $[a, b]$. Riemann sums have natural lower and upper bounds,

$$\mathcal{L}_{\mathcal{P}}(f) := \sum_{k=1}^{N} m_k(x_k - x_{k-1}), \qquad \mathcal{U}_{\mathcal{P}}(f) := \sum_{k=1}^{N} M_k(x_k - x_{k-1}),$$

where $m_k := \inf_{x \in [x_{k-1}, x_k]} f(x)$ and $M_k := \sup_{x \in [x_{k-1}, x_k]} f(x)$ (for $k = 1, \dots, N$). We say that f is *Riemann integrable* if the limits of $\mathcal{L}_{\mathcal{P}}(f)$ and $\mathcal{U}_{\mathcal{P}}(f)$ exist and are equal in the limit $|\mathcal{P}| \to 0$. We then define the *Riemann integral of* f as

$$\int_a^b f(x)\,dx := \lim_{|\mathcal{P}| \to 0} \mathcal{L}_{\mathcal{P}}(f) = \lim_{|\mathcal{P}| \to 0} \mathcal{U}_{\mathcal{P}}(f).$$

Teorem 2.2.1. *Let $f : [a, b] \to \mathbb{R}$ be Riemann integrable, and let μ be the Lebesgue measure on $[a, b]$. Then f is Lebesgue integrable, and the Riemann and Lebesgue integrals coincide:*

$$\int_a^b f(x)\,dx = \int_{[a,b]} f\,d\mu. \tag{2.3}$$

Proof. For a partition $\mathcal{P}$, let $\varphi_{\mathcal{P}}, \Phi_{\mathcal{P}} : [a, b] \to \mathbb{R}$ be the Lebesgue measurable simple functions

$$\varphi_{\mathcal{P}} := \sum_{k=1}^{N} m_k \mathbb{1}_{(x_{k-1}, x_k]}, \qquad \Phi_{\mathcal{P}} := \sum_{k=1}^{N} M_k \mathbb{1}_{(x_{k-1}, x_k]},$$

and observe that

$$\varphi_{\mathcal{P}} \leqslant f \leqslant \Phi_{\mathcal{P}} \tag{$*$}$$

and that

$$\mathcal{L}_{\mathcal{P}}(f) = \int_{[a,b]} \varphi_{\mathcal{P}}\,d\mu, \qquad \mathcal{U}_{\mathcal{P}}(f) = \int_{[a,b]} \Phi_{\mathcal{P}}\,d\mu.$$

By $(*)$ and Fatou's lemma, we have

$$\begin{aligned} 0 &\leqslant \int_{[a,b]} \liminf_{|\mathcal{P}| \to 0} \big(\Phi_{\mathcal{P}} - \varphi_{\mathcal{P}}\big)\,d\mu \leqslant \liminf_{|\mathcal{P}| \to 0} \int_{[a,b]} \Phi_{\mathcal{P}} - \varphi_{\mathcal{P}}\,d\mu \\ &= \liminf_{|\mathcal{P}| \to 0} \big(\mathcal{U}_{\mathcal{P}}(f) - \mathcal{L}_{\mathcal{P}}(f)\big) = 0. \end{aligned}$$

Hence, by Corollary 1.5.8, we have $\liminf_{|\mathcal{P}| \to 0} \big(\Phi_{\mathcal{P}} - \varphi_{\mathcal{P}}\big) = 0$ μ-almost everywhere. From $(*)$ we then conclude that f is equal almost everywhere to the Lebesgue

measurable function $\liminf_{|\mathcal{P}|\to 0} \Phi_{\mathcal{P}} = \limsup_{|\mathcal{P}|\to 0} \varphi_{\mathcal{P}}$, and is therefore Lebesgue measurable. Finally, again using $(*)$ we get

$$\mathcal{L}_{\mathcal{P}}(f) = \int_{[a,b]} \varphi_{\mathcal{P}}\, d\mu \leqslant \int_{[a,b]} f\, d\mu \leqslant \int_{[a,b]} \Phi_{\mathcal{P}}\, d\mu = \mathcal{U}_{\mathcal{P}}(f),$$

and taking limits $|\mathcal{P}| \to 0$ results in (2.3). ■