MEASURE AND INTEGRATION

Lecture notes by Mark Veraar, complemented by Emiel Lorist.

Contents

Date: October 27, 2023.

Introduction

These notes have been created for the "Measure and integration theory" part of a course on real analysis at the TU Delft. Sections 1-6 and 8-9 were mainly written by Mark and Section 7 by Emiel. Together with the first part of the course on metric spaces, these notes form the mathematical basis for several bachelor and master courses in applied mathematics at TU Delft.

We would like to thank Zoe Nieraeth, Esmée Theewis, Gerrit Vos and Carel Wagenaar for their support in the preparation of these notes. We would also like to thank the students of the course on real analysis for pointing out the typos in the manuscript.

In Section 1 and 2 we introduce σ-algebras and measures. The Lebesgue measure is constructed in Section 3 and is based on Appendix B on Carathéodory's theorem. Uniqueness questions are addressed in Appendix A on Dynkin's monotone class theorem. The amount of books on measure theory is almost not measurable. The lecture notes are based on [1], [8], [15] and [16]. A very complete treatment of measure theory is given in the impressive works [5].

In Sections 5, 6, 7 and 8 we introduce the integration theory and the Lebesgue spaces L^p. This theory is fundamental in modern (applied) mathematics. There are many excellent books which give more detailed treatments on the subject. See for instance [1], [4], [6], [14] for detailed treatments.

In Section 9 we give a brief introduction to the theory of Fourier series. More thorough treatments can be found in for example [9], [10], [12], [13] and [18]. A full course on Fourier Analysis is offered as a third year elective course based on the lecture notes [10]. The theory of Fourier series will be used in the second year bachelor course on Partial Differential Equations [7], but also in several other parts of Mathematical Physics and Numerical Analysis.

We end this brief introduction with a quote from a historical note of Zygmund [17]:

"The Lebesgue integral did not arise via the theory of Fourier series but was created through the necessities of measuring geometric figures. But once it was introduced, it had an enormous impact on analysis through Fourier series"

1. σ-ALGEBRAS

We want to introduce a way to assign a size to subsets of a general set S, reflecting our intuition about the size of subsets of $\mathbb{R}^d$. Before we can do so, we will first need to decide to which subsets of S we would like to assign a size. In the upcoming sections, we will see that it is not always possible or desirable to assign a size to *all* subsets of S. Therefore, in this first section, we introduce collections of subsets of S with some structure. We will assign a size to sets in these collections in Section 2. We let $\mathcal{P}(S)$ denote the power set of S, i.e. the collection of all subsets of S.

Definition 1.1. *Let S be a set. A collection $\mathcal{R} \subseteq \mathcal{P}(S)$ is called a **ring** if*

(i) $\varnothing \in \mathcal{R}$;
(ii) $A, B \in \mathcal{R} \implies B \setminus A \in \mathcal{R}$;
(iii) $A, B \in \mathcal{R} \implies A \cup B \in \mathcal{R}$.

Remark 1.2. By induction, it follows from (iii) that for $A_1, A_2, \ldots, A_n \in \mathcal{R}$ we have $\bigcup_{k=1}^n A_k \in \mathcal{R}$.

If $\mathcal{R} \subseteq \mathcal{P}(S)$ is a ring, then for all $A, B \in \mathcal{R}$ one has $A \cap B \in \mathcal{R}$, which follows from the identity $A \cap B = A \setminus (A \setminus B)$. Moreover, defining the **symmetric difference** of A and B as

$$A \Delta B := (A \setminus B) \cup (B \setminus A),$$

we also have $A \Delta B \in \mathcal{R}$.[1]

Definition 1.3. *Let S be a set. A collection $\mathcal{A} \subseteq \mathcal{P}(S)$ is called a σ-**algebra**[2] if*

(i) $\varnothing, S \in \mathcal{A}$;
(ii) $A \in \mathcal{A} \implies A^c \in \mathcal{A}$;
(iii) $A_1, A_2, \ldots \in \mathcal{A} \implies \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$.

*The pair $(S, \mathcal{A})$ is called a **measurable space**. The sets $A \in \mathcal{A}$ are called **measurable sets**.*

If $\mathcal{A} \subseteq \mathcal{P}(S)$ is a σ-algebra, then for $A_1, A_2, \ldots \in \mathcal{A}$ one has $\bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$, which follows from De Morgan's law

$$\bigcap_{k=1}^{\infty} A_k = \Big(\bigcup_{k=1}^{\infty} A_k^c \Big)^c.$$

Every σ-algebra is a ring. Indeed, (iii) in the definition of a ring follows by taking $A_k = \varnothing$ for all $k \geq 3$ in (iii) in the definition of a σ-algebra and (ii) in the definition of a ring follows from the identity $B \setminus A = (B^c \cup A)^c$. Conversely, not every ring is a σ-algebra. For example, $\mathcal{R} := \{\varnothing\}$ is a ring, but not a σ-algebra.

Next, let us turn to a some examples of rings and σ-algebras.

Example 1.4. Let S be a set.

(a) $\mathcal{A} := \{S, \varnothing\}$ is the smallest possible σ-algebra on S.
(b) $\mathcal{A} := \mathcal{P}(S)$ is the largest possible σ-algebra on S.

Example 1.5. Let $S := \{1, 2, 3\}$.

(a) Let $\mathcal{A} := \{\varnothing, S, \{1\}, \{2, 3\}\}$. Then $\mathcal{A}$ is a σ-algebra.
(b) Let $\mathcal{A} := \{\varnothing, S, \{1\}, \{2\}, \{1, 2\}\}$. Then $\mathcal{A}$ is a ring, but not a σ-algebra.

Example 1.6. Let S be a set.

(a) The collection $\mathcal{R} := \{A \subseteq S : A \text{ is finite}\}$ is a ring.
(b) The collection $\mathcal{A} := \{A \subseteq S : A \text{ is countable or } A^c \text{ is countable}\}$ is a σ-algebra (see Exercise 1.4). It is called the countable-cocountable σ-algebra and can be used to construct counterexamples.

---

[1] Using Δ as addition with additive identity $\varnothing$ and $\cap$ as multiplication with multiplicative identity S, one can check that $\mathcal{R}$ is a ring in the algebraic sense.

[2] In part of the literature a σ-algebra is also called a σ-field.

The following example will play a crucial role in the construction of the Lebesgue measure in Section 3.

Example 1.7.

(a) Let $S := \mathbb{R}$. For $a, b \in \mathbb{R}$ with $a \leq b$ we call $(a, b]$ a **half-open interval**. Let $\mathcal{I}^1$ be the collection of all half-open intervals. Then $\mathcal{I}^1$ is not a ring, since, for instance, $(0, 3] \setminus (1, 2] = (0, 1] \cup (2, 3]$ is not in $\mathcal{I}^1$.

(b) Let $S := \mathbb{R}$. Let $\mathcal{F}^1$ be the collection of sets which can be written as a finite union of half-open intervals. Then $\mathcal{F}^1$ is not a σ-algebra, since, for instance $\bigcup_{n=1}^{\infty}(0, 1 - \frac{1}{n}] = (0, 1)$ is not an element of $\mathcal{F}^1$. We will check that $\mathcal{F}^1$ is a ring. (i) follows from $\varnothing = (1, 1] \in \mathcal{F}^1$. (iii) is clear since a finite union of a finite union of intervals of the form $(a, b]$ is again a finite union. It remains to check (ii). For this we first note that it is simple to check that for $A, B \in \mathcal{F}^1$ one has $A \cap B \in \mathcal{F}^1$ and by induction this extends to the intersections of finitely many sets. For two intervals $(a, b]$ and $(c, d]$ using $B \setminus A = B \cap A^c$ and $\mathbb{R} \setminus (a, b] = (-\infty, a] \cup (b, \infty)$ we find

$$\begin{aligned}(c, d] \setminus (a, b] &= (c, d] \cap (\mathbb{R} \setminus (a, b]) \\ &= \big((c, d] \cap (-\infty, a]\big) \cup \big((c, d] \cap (b, \infty)\big) \\ &= (c, a \wedge d] \cup (b \vee c, d].\end{aligned}$$

This is in $\mathcal{F}^1$ again. Now if $A = \bigcup_{j=1}^{m}(a_j, b_j]$ and $B = \bigcup_{k=1}^{n}(c_k, d_k]$ are in $\mathcal{F}^1$, then

$$B \setminus A = \bigcup_{k=1}^{n}(c_k, d_k] \setminus A = \bigcup_{k=1}^{n}\bigcap_{j=1}^{m}(c_k, d_k] \setminus (a_j, b_j].$$

and, by the previous observations, this is in $\mathcal{F}^1$ again.

(c) Let $S := \mathbb{R}^d$. For $a, b \in \mathbb{R}^d$ with $a = (\alpha_1, \ldots, \alpha_d)$ and $b = (\beta_1, \ldots, \beta_d)$ with $\alpha_j \leq \beta_j$ for $j \in \{1, \ldots, d\}$ the **half-open rectangle** $(a, b]$ is given by

$$(a, b] = (\alpha_1, \beta_1] \times \cdots \times (\alpha_d, \beta_d].$$

Let $\mathcal{I}^d$ be the collection of all half-open rectangles and let $\mathcal{F}^d$ be the collection of sets which can be written as a finite union of half-open rectangles. Then $\mathcal{F}^d$ is a ring (see Exercise 1.9).

On a set S there can be various σ-algebras. One may therefore wonder what happens when we take the union or intersection of two or more σ-algebras. It turns out that the union of two σ-algebras may not be a σ-algebra. For the intersection of σ-algebras things are much better. Indeed, the intersection of arbitrarily many σ-algebras is again a σ-algebra, which we formalize in the next proposition. We refer to Exercise 1.5 for a proof of both facts.

Proposition 1.8 (Intersection of σ-algebras)**.** *Let I be an index set and suppose that $\mathcal{A}_i$ is a σ-algebra on S for every $i \in I$. Then $\bigcap_{i \in I} \mathcal{A}_i$ is a σ-algebra.*

Using Proposition 1.8, we can construct the smallest σ-algebra containing any collection $\mathcal{F} \subseteq S$.

Definition 1.9. *Let S be a set and let $\mathcal{F} \subseteq \mathcal{P}(S)$. We write $\sigma(\mathcal{F})$ for the smallest σ-algebra containing $\mathcal{F}$, i.e.*

$$\sigma(\mathcal{F}) := \bigcap\{\mathcal{A} \subseteq \mathcal{P}(S) : \mathcal{A} \text{ is a } \sigma\text{-algebra on } S \text{ and } \mathcal{F} \subseteq \mathcal{A}\}.$$

Then $\sigma(\mathcal{F})$ is called the σ-algebra ***generated*** *by $\mathcal{F}$.*

Note that Proposition 1.8 implies that $\sigma(\mathcal{F})$ is a σ-algebra. Moreover, the intersection ensures we obtain the smallest possible one. Since $\sigma(\mathcal{F})$ is a σ-algebra, one way to show that $A \subseteq S$ is an element of $\sigma(\mathcal{F})$ is to write it as the countable union or intersection of sets in $\mathcal{F}$ or their complements. However, in certain cases, not all elements of $\sigma(\mathcal{F})$ can be written in this way.

Let us give a few elementary examples.

Example 1.10. Let $S := \{1, 2, 3\}$.

(a) For $\mathcal{F} := \{\{1, 2\}\}$, we have $\sigma(\mathcal{F}) = \{\varnothing, S, \{1, 2\}, \{3\}\}$.
(b) For $\mathcal{F} := \{\{2, 3\}, \{1, 2\}\}$ we have $\sigma(\mathcal{F}) = \mathcal{P}(S)$. Indeed, $\{2, 3\}^c = \{1\}$, $\{1, 2\}^c = \{3\}$ and $\{2, 3\} \cap \{1, 2\} = \{2\}$. Thus the singletons $\{1\}$, $\{2\}$ and $\{3\}$ are in $\sigma(\mathcal{F})$. Therefore, the required result follows since we can form every subset of S by taking suitable finite unions.

Example 1.11. Let $S := \mathbb{N}$ and $\mathcal{F} := \{\{n\} : n \in \mathbb{N}\}$. Then $\sigma(\mathcal{F}) = \mathcal{P}(\mathbb{N})$.

The following definition introduces some of the most important σ-algebras.

Definition 1.12. *Let (M, d) be a metric space. Let $\mathcal{B}(M)$ be the σ-algebra generated by the open sets in M. Thus*

$$\mathcal{B}(M) := \sigma\{O \subseteq M : O \text{ is open}\}.$$

The σ-algebra $\mathcal{B}(M)$ is called[3] *the* ***Borel σ-algebra*** *of M. The elements of $\mathcal{B}(M)$ are called* ***Borel sets****.*

We will mainly use Definition 1.12 with $M = \mathbb{R}$ or $M = \mathbb{R}^d$ equipped with the Euclidean norm. Taking $\mathcal{I}^d$ and $\mathcal{F}^d$ be as in Example 1.7, we will see in Exercise 1.10 that

$$\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{I}^d) = \sigma(\mathcal{F}^d).$$

Moreover, in Exercise 3.7 we will show that $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$.

We end this section with a useful lemma for some of the exercises on the Borel σ-algebra of $\mathbb{R}$ and $\mathbb{R}^d$.

Lemma 1.13 (Lindelöf)**.** *Let $A \subseteq \mathbb{R}^d$. Assume that for each $i \in I$, $O_i \subseteq \mathbb{R}^d$ is open. If $A \subseteq \bigcup_{i\in I} O_i$, then there exists a countable $J \subseteq I$ such that $A \subseteq \bigcup_{i\in J} O_i$*

Proof. Choose for each $x \in A$, $i_x \in I$ and $r_x > 0$ such that $B(x, r_x) \subseteq O_{i_x}$. For each $x \in A$ choose $a_x \in \mathbb{Q}^d$ and $s_x \in \mathbb{Q} \cap (0, \infty)$ such that $x \in B(a_x, s_x) \subseteq B(x, r_x) \subseteq O_{i_x}$. Let

$$\mathcal{F} := \{B(a_x, s_x) : x \in A\}.$$

Then clearly $A \subseteq \bigcup_{x\in A} B(a_x, s_x)$. Moreover, $\mathcal{F}$ contains at most countably many sets, which follows from the fact that it is a subset of the countable collection

$$\{B(q, r) : q \in \mathbb{Q}^d, r \in \mathbb{Q} \cap (0, \infty)\}.$$

Therefore, we can write $\mathcal{F} = \{B(a_{x_n}, s_{x_n}) : n \in \mathbb{N}\}$ with $x_n \in A$ for each $n \in \mathbb{N}$.

Now let $J := \{i_{x_n} \in I : n \in \mathbb{N}\}$. Then $A \subseteq \bigcup_{i\in J} O_i$. Indeed, if $x \in A$, then $x \in B(a_x, s_x)$ and choosing $n \in \mathbb{N}$ such that $a_{x_n} = a_x$ and $s_{x_n} = s_x$, we find $x \in O_{i_{x_n}} \subseteq \bigcup_{i\in J} O_i$. ☺

Exercises

Exercise 1.1. Let $S = \mathbb{R}$ and $\mathcal{F} = \{A \subseteq \mathbb{R} : A \subseteq [0, 1] \text{ or } A^c \subseteq [0, 1]\}$. Is $\mathcal{F}$ a ring?

Exercise* 1.2. Let S be a set. Suppose that $\mathcal{R} \subseteq \mathcal{P}(S)$ is nonempty and

(i) $A, B \in \mathcal{R} \Longrightarrow A \cap B \in \mathcal{R}$.
(ii) $A, B \in \mathcal{R} \Longrightarrow A \triangle B \in \mathcal{R}$.

Show that $\mathcal{R}$ is a ring.

Exercise 1.3. Let S be a set. Suppose that $\mathcal{A} \subseteq \mathcal{P}(S)$ satisfies

(i) $\varnothing, S \in \mathcal{A}$;
(ii) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$;
(iii) $A_1, A_2, \ldots \in \mathcal{A}$ disjoint[4] $\Longrightarrow \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$.

Show that $\mathcal{A}$ is a σ-algebra.

[3] Named after the French mathematician Félix Borel 1871-1956

[4] Here we mean $A_j \cap A_k = \varnothing$

Exercise 1.4. Prove that the collection in Example 1.6(b) is a σ-algebra.
Hint: Use the following facts: The subset of a countable set is again countable and the countable union of countable sets is again countable.

Exercise 1.5.
(a) Prove Proposition 1.8.
(b) Give an example of two σ-algebras $\mathcal{A}$ and $\mathcal{B}$ on $S := \{1, 2, 3\}$ such that $\mathcal{A} \cup \mathcal{B}$ is not a σ-algebra.

Exercise* 1.6. Let S be a set and let $\mathcal{F} := \{\{s\} : s \in S\}$. Show that $\sigma(\mathcal{F})$ coincides with the countable-cocountable σ-algebra of Example 1.6.

Exercise 1.7. Show that $\mathbb{N}, \mathbb{Q}, \mathbb{R} \setminus \mathbb{Q} \in \mathcal{B}(\mathbb{R})$. That is $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are Borel subsets of $\mathbb{R}$.

Exercise* 1.8. Consider the following collection $\mathcal{B}_0 := \{(-\infty, x) : x \in \mathbb{R}\}$ of subsets of $\mathbb{R}$.
(a) Show that $\sigma(\mathcal{B}_0)$ contains all open intervals.
(b) Show that every open set in $\mathbb{R}$ can be written as the union of countably many open intervals.
Hint: Use Lindelöf's Lemma 1.13.
(c) Conclude that $\sigma(\mathcal{B}_0) = \mathcal{B}(\mathbb{R})$.

Exercise* 1.9. Let $\mathcal{I}^d$ and $\mathcal{F}^d$ as in Example 1.7. Prove the following assertions.
(a) If $I, J \in \mathcal{I}^d$ then $I \cap J \in \mathcal{I}^d$.
(b) If $I, J \in \mathcal{I}^d$, then $I \setminus J$ is the union of finitely many *disjoint* sets from $\mathcal{I}^d$, and thus $I \setminus J \in \mathcal{F}^d$.
Hint: Use induction on the dimension d. Use Example 1.7(b) for $d = 1$.
(c) Each $A \in \mathcal{F}^d$ can be written as union of finitely many *disjoint* sets in $\mathcal{I}^d$.
Hint: Use induction on n to prove this for all sets of the form $A = \bigcup_{k=1}^n I_k$ with $I_1, \ldots, I_n \in \mathcal{I}^d$.
(d) $\mathcal{F}^d$ is a ring.

Exercise* 1.10. Let $\mathcal{I}^d$ and $\mathcal{F}^d$ as in Example 1.7. Show that $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{I}^d) = \sigma(\mathcal{F}^d)$.
Hint: Use Lindelöf's Lemma 1.13.

Exercise** 1.11. Prove that a σ-algebra is either finite or uncountable.[5]
Hint: Recall that $\mathcal{P}(\mathbb{N})$ is uncountable.

[5]This shows that σ-algebras are either easy finite collections of sets or quite complicated.

2. Measures

Having introduced the notion of a measurable space, i.e. a set S with a σ-algebra $\mathcal{A} \subseteq \mathcal{P}(S)$, we will now assign a size to each set in $\mathcal{A}$. That means that we will need a map from $\mathcal{A}$ to $[0,\infty]$, since sizes should be positive and could be infinite (think of the size of $\mathbb{R}$). Such a map will be called a measure if it satisfies a few properties that model our intuition about the size of sets in $\mathbb{R}^d$. For instance, the empty set should have size 0. Moreover, the size of a (possibly infinite) union of disjoint sets should be equal to the sum of the sizes of the individual sets, which we will call (σ-)-additivity:

Definition 2.1. *Let S be a set and let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring. Let $\mu : \mathcal{R} \to [0,\infty]$ be a function with $\mu(\varnothing) = 0$.*

*(i) μ is called **additive** if for $A, B \in \mathcal{R}$ with $A \cap B = \varnothing$ one has*

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

*(ii) μ is called σ-**additive** if for all disjoint[6] $A_1, A_2, \ldots \mathcal{R}$ which satisfy $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$, one has*

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

Note that any σ-additive map is additive, as follows by taking $A_m = \varnothing$ for $m \geq 3$. Moreover, if $\mathcal{R}$ a σ-algebra, the assumption $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ in the definition of σ-additivity is always satisfied.

Remark 2.2. By Remark 1.2 and an induction argument, for an additive map $\mathcal{R} \to [0,\infty]$ and disjoint $A_1, \ldots, A_n \in \mathcal{R}$ one has

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mu(A_k),$$

see Exercise 2.1.

Let us give a few basic properties of (σ-)additive maps. Further properties can be found in Exercise 2.2.

Proposition 2.3. *Let S be a set and let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring. Let $\mu : \mathcal{R} \to [0,\infty]$ be additive with $\mu(\varnothing) = 0$. The following assertions hold:*

(i) (Monotonicity) If $A, B \in \mathcal{R}$ with $A \subseteq B$, then $\mu(A) \leq \mu(B)$.

(ii) (Subadditivity) If $A, B \in \mathcal{R}$, then $\mu(A \cup B) \leq \mu(A) + \mu(B)$.

(iii) (σ-subadditivity) If $A_1, A_2, \ldots \in \mathcal{R}$ with $\bigcup_{k=1}^{\infty} A_k \in \mathcal{R}$ and μ is σ-additive, then

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Proof. (i): Using $A \subseteq B$, we can write $B = A \cup (B \setminus A)$. Then

$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

(ii): Using (i) we have

$$\mu(A \cup B) = \mu(A) + \mu(B \setminus A) \leq \mu(A) + \mu(B).$$

(iii): For $n \geq 1$ define

$$B_n := A_n \setminus \Big(\bigcup_{k=1}^{n-1} A_k\Big). \tag{2.1}$$

Then $(B_n)_{n\geq 1}$ is a disjoint sequence in $\mathcal{R}$ and $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Therefore, by the σ-additivity of μ and (i), we have

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \sum_{n=1}^{\infty} \mu(B_n) \overset{\text{(i)}}{\leq} \sum_{n=1}^{\infty} \mu(A_n).$$ ☺

[6] Here we mean $A_j \cap A_k = \varnothing$ if $j \neq k$.

We are now ready to define the size of the sets in a σ-algebra, which is a map that we will call a measure.

Definition 2.4. *Let $(S, \mathcal{A})$ be a measurable space. A function $\mu : \mathcal{A} \to [0, \infty]$ is called a* ***measure*** *if*

(i) $\mu(\varnothing) = 0$.
(ii) μ *is σ-additive.*

In this case the triple $(S, \mathcal{A}, \mu)$ is called a ***measure space****. If additionally $\mu(S) = 1$, then μ is called a* ***probability measure*** *and $(S, \mathcal{A}, \mu)$ is called a* ***probability space****.*[7]

Let us give some examples of measures.

Example 2.5 (Counting measure). Let S be a set and $\mathcal{A} = \mathcal{P}(\mathbb{N})$. We write $\#A$ for the number of elements of a finite set A, and we set $\#A = \infty$ if A is infinite. Let $\mu : \mathcal{A} \to [0, \infty]$ be given by $\mu(A) = \#A$. Then μ is a measure. Often μ is denoted by τ and is called the **counting measure**.

Example 2.6 (Dirac measure). Let $S = \mathbb{R}$, $\mathcal{A} = \mathcal{P}(\mathbb{R})$. Let $x \in \mathbb{R}$. Let $\delta_x : \mathcal{A} \to [0, \infty]$ be given by $\delta_x(A) = 1$ if $x \in A$ and $\delta_x(A) = 0$ if $x \notin A$. Then δ_x is a measure. It is usually called the **Dirac measure**[8] at x.

Example 2.7 (Lebesgue measure). For a half-open interval $(a, b]$ with $a \leq b$, we could define $\lambda((a, b]) = b - a$ (length or size of the interval). This definition can be extended to unions of disjoint half-open intervals, i.e. to all sets in $\mathcal{F}^1$ as in Example 1.7. In Section 3 we will see that $\lambda\colon \mathcal{F}^1 \to [0, \infty]$ is σ-additive and has an extension to a measure on $\sigma(\mathcal{F}^1) = \mathcal{B}(\mathbb{R})$. This measure will be called the **Lebesgue measure**[9] on $\mathbb{R}$.

Let $(a_n)_{n\geq 1}$ be a sequence of real numbers and $a \in \mathbb{R}$. We write $a_n \uparrow a$ if $(a_n)_{n\geq 1}$ is an increasing sequence which converges to a. Similarly, we write $a_n \downarrow a$ if it decreases and converges to a. This notation can be extended to sets as follows.

Definition 2.8. *Let S be a set.*

(i) A sequence $(A_n)_{n\geq 1}$ of subsets of S will be called ***increasing*** *if $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$. In this case we write $A_n \uparrow A$, where $A = \bigcup_{n=1}^\infty A_n$.*
(ii) A sequence $(A_n)_{n\geq 1}$ of subsets of S will be called ***decreasing*** *if $A_n \supseteq A_{n+1}$ for all $n \in \mathbb{N}$. In this case we write $A_n \downarrow A$, where $A = \bigcap_{n=1}^\infty A_n$.*

Note that we have *not* given a meaning to $A_n \to A$ for a sequence $(A_n)_{n\geq 1}$ of subsets of S.

Theorem 2.9. *Let $(S, \mathcal{A}, \mu)$ be a measure space and let $(A_n)_{n\geq 1}$ be a sequence in $\mathcal{A}$.*

(i) If $A_n \uparrow A$, then $\mu(A_n) \uparrow \mu(A)$.
(ii) If $A_n \downarrow A$ and $\mu(A_1) < \infty$, then $\mu(A_n) \downarrow \mu(A)$.

Proof. (i): Define $(B_n)_{n\geq 1}$ as in (2.1). Then $(B_n)_{n\geq 1}$ is a disjoint sequence and the following identities hold: $\bigcup_{k=1}^\infty B_k = A$ and $\bigcup_{k=1}^n B_k = A_n$. Therefore, the σ-additivity of μ gives

$$\mu(A) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n).$$

(ii): See Exercise 2.4. ☺

The following result will help us to check σ-additivity on rings. It will be used in the construction of the Lebesgue measure in Lemma 3.8.

Lemma 2.10 (Sufficient condition for σ-additivity)**.** *Let S be a set and let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring. Let $\mu : \mathcal{R} \to [0, \infty]$ be additive with $\mu(\varnothing) = 0$. Suppose that for each sequence $(A_n)_{n\geq 1}$ in $\mathcal{R}$ with $A_n \downarrow \varnothing$ one has $\mu(A_n) \to 0$. Then μ is σ-additive on $\mathcal{R}$.*

[7] Measure theory is at the very heart of probability theory.

[8] Named after the English theoretical physicist Paul Dirac 1902-1984. Physicists often call it the Dirac delta function, although it is actually not a function.

[9] In the literature also sometimes called the Borel–Lebesgue measure

Proof. Let $(B_k)_{k\geq 1}$ be a disjoint sequence in $\mathcal{R}$ with $B := \bigcup_{k=1}^{\infty} B_k \in \mathcal{R}$. We need to show that

$$\mu(B) = \sum_{k=1}^{\infty} \mu(B_k). \tag{2.2}$$

Let $A_n = \bigcup_{k=n}^{\infty} B_k = B \setminus (B_1 \cup \ldots \cup B_{n-1})$. Then $A_n \in \mathcal{R}$ and $A_n \downarrow \varnothing$. Now the assumption yields $\mu(A_n) \to 0$. Since μ is additive, we have

$$\mu(B) = \mu(A_n \cup B_1 \cup B_2 \cup \ldots \cup B_{n-1}) = \mu(A_n) + \sum_{k=1}^{n-1} \mu(B_k).$$

Taking the limit $n \to \infty$, (2.2) follows. ☺

Exercises

Exercise 2.1. Let $\mathcal{R}$ be a ring on a set S. Let $\mu : \mathcal{R} \to [0, \infty]$ be additive with $\mu(\varnothing) = 0$. Show that for disjoint $A_1, \ldots, A_n \in \mathcal{R}$ we have

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mu(A_k).$$

Exercise 2.2. Let $\mathcal{R}$ be a ring on a set S. Let $\mu : \mathcal{R} \to [0, \infty]$ be additive with $\mu(\varnothing) = 0$.

(a) Prove that for $A, B \in \mathcal{R}$ with $\mu(A) < \infty$ and $A \subseteq B$ one has
$$\mu(B \setminus A) = \mu(B) - \mu(A).$$
(b) Prove that for $A, B \in \mathcal{R}$ with $\mu(A) < \infty$ one has
$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$
(c) Proof that for disjoint $A_1, A_2, \ldots \in \mathcal{R}$ with $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$ one has
$$\sum_{n=1}^{\infty} \mu(A_n) \leq \mu\Big(\bigcup_{n=1}^{\infty} A_n\Big).$$

Exercise 2.3. Let $(a_n)_{n\geq 1}$ be numbers in $[0, \infty)$. Define $\mu : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ by $\mu(A) = \sum_{n\in A} a_n$. Prove that μ is a measure.

Exercise* 2.4.

(a) Prove Theorem 2.9(ii).
(b) Give an example of a measure space $(S, \mathcal{A}, \mu)$ and sets $A_n \in \mathcal{A}$ such that $A_n \downarrow \varnothing$ and $\mu(A_n) = \infty$ for all $n \in \mathbb{N}$.
Hint: Use the counting measure.

Exercise* 2.5. Let $(S, \mathcal{A}, \mu)$ be a measure space. For $A_1, A_2, \ldots \subseteq S$ define

$$\limsup_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n.$$

(a) Show that $s \in \limsup_{n\to\infty} A_n$ if and only if there are infinitely many $n \in \mathbb{N}$ such that $s \in A_n$.
(b) Assume $A_1, A_2, \ldots \in \mathcal{A}$. Show that $\limsup_{n\to\infty} A_n \in \mathcal{A}$.
(c) Assume $A_1, A_2, \ldots \in \mathcal{A}$ satisfy $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Show that $\mu(\limsup_{n\to\infty} A_n) = 0$.[10]

Exercise* 2.6. Let $\mathcal{A}$ be the σ-algebra from Example 1.6(b) with $S = \mathbb{R}$. Define $\mu : \mathcal{A} \to [0, \infty]$ by $\mu(A) = 0$ if A is countable and $\mu(A) = 1$ if A^c is countable. Show that μ is a measure.

[10] This is called the Borel-Cantelli lemma in probability theory.

3. Construction of measures

It is not a simple task to construct a measure. In this section we will construct the Lebesgue measure on $\mathbb{R}^d$, which we already briefly encountered in Example 2.7. To extend it from $\mathcal{F}^d$ to the Borel σ-algebra, we will use a deep result of Carathéodory.[11] His result basically says that it is enough to check that a measure is σ-additive on a ring generating the desired σ-algebra. A detailed proof can be found in Theorem B.4 in the appendix, but it will do no harm if one takes the result for granted.

Theorem 3.1 (Carathéodory's extension theorem)**.** *Let S be a set and let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring. Suppose that $\mu(\varnothing) = 0$ and $\mu\colon \mathcal{R} \to [0,\infty]$ is σ-additive. Then there exists a measure $\overline{\mu}\colon \sigma(\mathcal{R}) \to [0,\infty]$ such that $\overline{\mu}(A) = \mu(A)$ for all $A \in \mathcal{R}$.*

Remark 3.2. The measure $\overline{\mu}$ is often unique.[12] When there is no danger of confusion we will write μ again for the extension to $\sigma(\mathcal{R})$. However, in general one has to be careful about uniqueness. For instance if we define μ on the ring $\mathcal{F}^1$ by $\mu(\varnothing) = 0$ and $\mu(A) = \infty$ if $A \in \mathcal{F}^1$ is nonempty, then μ has at least two extensions: the counting measure on $\mathcal{B}(\mathbb{R})$ is an extension of μ, but also the measure $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ given by $\nu(A) = \infty$ if $A \neq \varnothing$ is an extension of μ.

We continue with a uniqueness result. For the statement we need collections of subsets of S that are closed under taking intersections.

Definition 3.3 (π-system)**.** *A collection $\mathcal{E} \subseteq \mathcal{P}(S)$ is called a **π-system** if for all $A, B \in \mathcal{E}$ one has $A \cap B \in \mathcal{E}$.*

Example 3.4.

(a) Every ring is a π-system.
(b) The half-open rectangles $\mathcal{I}^d$ are a π-system on $\mathbb{R}^d$.

The following result will be proved in Proposition A.7 in the appendix, but it will once again do no harm if one takes the result for granted. It is based on a so-called monotone class argument.

Proposition 3.5 (Uniqueness)**.** *Let μ_1 and μ_2 both be measures on a measurable space $(S, \mathcal{A})$. Assume the following conditions:*

(i) $\mathcal{E} \subseteq \mathcal{A}$ is a π-system with $\sigma(\mathcal{E}) = \mathcal{A}$;
(ii) $\mu_1(S) = \mu_2(S) < \infty$ and $\mu_1(E) = \mu_2(E)$ for all $E \in \mathcal{E}$.

Then $\mu_1 = \mu_2$ on $\mathcal{A}$.

3.1. The Lebesgue measure.

With Theorem 3.1 and Proposition 3.5 at our disposal, we are now ready to construct the **Lebesgue measure** λ.[13] In Example 1.7(a) we introduced the half-open intervals $\mathcal{I}^1$. For a half-open interval $(a, b] \in \mathcal{I}^1$ with $a \leq b$ we define

$$\lambda((a, b]) := b - a,$$

which is the length of $(a, b]$. Note that $\lambda(\varnothing) = 0$.

Next we extend λ to the collection of all finite unions of half-open intervals $\mathcal{F}^1$. We have seen that $\mathcal{F}^1$ is a ring and that every set in $\mathcal{F}^1$ can be written as a finite union of *disjoint* half-open intervals (see Example 1.7(b)).

Definition 3.6 (Lebesgue measure on $\mathcal{F}^1$)**.** *For $A \in \mathcal{F}^1$ of the form $A = I_1 \cup \cdots \cup I_n$ with disjoint $(I_k)_{k=1}^n$ in $\mathcal{I}^1$, define $\lambda : \mathcal{F}^1 \to [0, \infty)$ as*

$$\lambda(A) := \sum_{k=1}^{n} \lambda(I_k).$$

---

[11] Constantin Carathéodory 1873-1950 was a Greek mathematician working in Analysis, but also on Thermodynamics.

[12] For instance when μ is a finite measure. See Proposition A.7.

[13] Henri Lebesgue 1875–1941 was a French mathematician well-known for his integration theory. See Section 5.

The map $\lambda\colon \mathcal{F}^1 \to [0,\infty)$ is well-defined. Indeed, λ clearly takes valued in $[0,\infty)$, so it remains to prove that the definition of $\lambda(A)$ does not depend on the choice of $I_1,\dots,I_n$. We first prove this for $A=(a,b]\in\mathcal{I}^1$. Suppose that $(a,b]=(a_1,b_1]\cup\ldots\cup(a_n,b_n]$ is a disjoint union with $a_k\le b_k$ for all $k\in\{1,\dots,n\}$. Without loss of generality we can assume $b_k\le a_{k+1}$ for each $k\in\{1,\dots,n-1\}$. Since $(b_k,a_{k+1}]\cap(a,b]$ is nonempty if $b_k<a_{k+1}$, we find that $b_k=a_{k+1}$ for each $k\in\{1,\dots,n-1\}$, $a_1=a$ and $b_n=b$. Therefore

$$\lambda((a,b]) = b-a = \sum_{k=1}^n b_k - a_k = \sum_{k=1}^n \lambda((a_k,b_k]), \tag{3.1}$$

so λ is well defined on $\mathcal{I}^1$.

Next, let $A\in\mathcal{F}^1$ and write

$$A = I_1\cup\cdots\cup I_n = I_1'\cup\cdots\cup I_m'$$

with $I_1,\dots,I_n\in\mathcal{I}^1$ disjoint and $I_1',\dots,I_m'\in\mathcal{I}^1$ disjoint. Let $I_{jk}=I_j'\cap I_k$. Then each I_{jk} is a half open interval, $\bigcup_{j=1}^m I_{jk}=I_k$ and $\bigcup_{k=1}^n I_{jk}=I_j'$. From (3.1) we therefore obtain

$$\begin{aligned}\sum_{k=1}^n \lambda\big(I_k\big) = \sum_{k=1}^n \lambda\Big(\bigcup_{j=1}^m I_{jk}\Big) &\overset{(3.1)}{=} \sum_{k=1}^n\sum_{j=1}^m \lambda(I_{jk})\\ &= \sum_{j=1}^m\sum_{k=1}^n \lambda(I_{jk}) \overset{(3.1)}{=} \sum_{j=1}^m \lambda\Big(\bigcup_{k=1}^n I_{jk}\Big) = \sum_{j=1}^m \lambda(I_j'),\end{aligned}$$

which proves the well-definedness.[14] Moreover, it is clear from the definition that λ is additive.

Next we do the same construction in higher dimensions, i.e. for $d\ge 2$. In Example 1.7(c) we introduced the half-open rectangles $(a,b]\in\mathcal{I}^d$. Also recall that $\mathcal{F}^d$ denotes the collection of all finite unions of half-open rectangles. By Exercise 1.9(d), $\mathcal{F}^d$ is a ring. Moreover, in Exercise 1.9(c) it was shown that every set in $\mathcal{F}^d$ can be written as a finite union of disjoint half-open rectangles.

Definition 3.7 (Lebesgue measure on $\mathcal{F}^d$)**.** *For a half-open rectangle $I=(a,b]\in\mathcal{I}^d$ with $a=(\alpha_1,\dots,\alpha_d)$ and $b=(\beta_1,\dots,\beta_d)$ and $\alpha_j\le\beta_j$ for $j\in\{1,\dots,d\}$ we define its **volume** by*

$$|I| := \prod_{j=1}^d(\beta_j-\alpha_j).$$

For $A\in\mathcal{F}^d$ of the form $A=I_1\cup\ldots\cup I_n$ with disjoint $I_1,\dots,I_n\in\mathcal{I}^d$ define $\lambda:\mathcal{F}^d\to[0,\infty]$ by

$$\lambda(A) := \sum_{k=1}^n |I_k|.$$

The map $\lambda\colon\mathcal{F}^d\to[0,\infty]$ is well-defined and additive, which can be proven analogously to the case $d=1$ above. We want to extend λ to $\sigma(\mathcal{F}^d)=\mathcal{B}(\mathbb{R}^d)$. To be able to apply Theorem 3.1, we first need to check the σ-additivity of λ on $\mathcal{F}^d$. This will be done via Lemma 2.10.

Lemma 3.8. *The map $\lambda\colon\mathcal{F}^d\to[0,\infty]$ is σ-additive on $\mathcal{F}^d$.*

Proof. By Lemma 2.10 it suffices to take a sequence $(A_n)_{n\ge1}$ in $\mathcal{F}^d$ with $A_n\downarrow\varnothing$ and prove $\lambda(A_n)\to0$. Fix $\varepsilon>0$. We have to find $N\in\mathbb{N}$ such that $\lambda(A_n)<\varepsilon$ for all $n\ge N$.

For each $n\in\mathbb{N}$ choose a $B_n\in\mathcal{F}^d$ such that $\overline{B_n}\subseteq A_n$ and $\lambda(A_n\setminus B_n)\le 2^{-n}\varepsilon$.[15] Then

$$\bigcap_{n=1}^\infty \overline{B_n} = \bigcap_{n=1}^\infty A_n = \varnothing.$$

Note that $\overline{A_1}$ is compact by the Heine–Borel theorem and thus has the finite intersection property. Since $\overline{B_n}\subseteq\overline{A_1}$ for all $n\ge1$, we can find an $N\ge1$ such that $\bigcap_{n=1}^N\overline{B_n}=\varnothing$.

[14] Alternatively, one can observe that $\lambda(A)$ coincides with the Riemann integral of $\mathbf{1}_A$ and use the linearity of the Riemann integral.

[15] So we choose a set B_n which is slightly smaller than A_n.

Fix $n \geq N$. Since $\bigcap_{k=1}^n B_k \subseteq \bigcap_{k=1}^N B_k = \varnothing$, we have

$$A_n = A_n \setminus \bigcap_{k=1}^n B_k = \bigcup_{k=1}^n A_n \setminus B_k \subseteq \bigcup_{k=1}^n A_k \setminus B_k.$$

Therefore, using Proposition 2.3(i) and Exercise 2.2(c), we find

$$\lambda(A_n) \leq \lambda\Big(\bigcup_{k=1}^n (A_k \setminus B_k)\Big) \leq \sum_{k=1}^n \lambda(A_k \setminus B_k) \leq \sum_{k=1}^n 2^{-k}\varepsilon = \varepsilon,$$

which finishes the proof. ☺

We can now deduce the main result of this section, which is the extension of the Lebesgue measure to the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$.

Theorem 3.9 (Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$)**.** *There exists a unique measure λ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that $\lambda(I) = |I|$ for all half-open rectangles $I \in \mathcal{I}^d$. Moreover, for all $h \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$, $\lambda(A + h) = \lambda(A)$.*[16]

In the above $A + h := \{x + h : x \in A\}$ for $A \subseteq \mathbb{R}^d$ and $h \in \mathbb{R}^d$.

Proof. Step 1: Existence. In Lemma 3.8 we have shown that λ is σ-additive on the ring $\mathcal{F}^d$. Therefore, by Theorem 3.1, λ extends to a measure on $\sigma(\mathcal{F}^d)$. Since $\sigma(\mathcal{F}^d) = \mathcal{B}(\mathbb{R}^d)$ by Exercise 1.10, the existence follows.

Step 2: Uniqueness. Let μ be another measure such that $\mu(I) = |I|$ for half-open rectangles $I \in \mathcal{I}^d$. Fix $n \in \mathbb{N}$ and let $S_n = (-n, n]^d$. Define $\lambda^{(n)}$ and $\mu^{(n)}$ on $\mathcal{B}(\mathbb{R}^d)$ by

$$\lambda^{(n)}(A) := \lambda(A \cap S_n) \quad \text{and} \quad \mu^{(n)}(A) := \mu(A \cap S_n).$$

Then $\lambda^{(n)}$ and $\mu^{(n)}$ are measures and $\lambda^{(n)}(\mathbb{R}^d) = \lambda(S_n) = |S_n|$ and similarly $\mu^{(n)}(\mathbb{R}^d) = |S_n|$. Since $\lambda^{(n)}$ and $\mu^{(n)}$ coincide on $\mathcal{I}^d$, it follows from Example 3.4(b) and Proposition 3.5 that $\lambda^{(n)} = \mu^{(n)}$ on $\mathcal{B}(\mathbb{R}^d)$. Therefore, for any $A \in \mathcal{B}(\mathbb{R}^d)$, since $A \cap S_n \uparrow A$, Theorem 2.9 yields

$$\lambda^{(n)}(A) = \lambda(A \cap S_n) \to \lambda(A) \quad \text{and} \quad \mu^{(n)}(A) = \mu(A \cap S_n) \to \mu(A).$$

Thus $\lambda(A) = \mu(A)$.

Step 3: Translation invariance: Let $h \in \mathbb{R}^d$. We claim that for every $A \in \mathcal{B}(\mathbb{R}^d)$ one has $A + h \in \mathcal{B}(\mathbb{R}^d)$. For this let $\mathcal{A}_h := \{A \in \mathcal{B}(\mathbb{R}^d) : A + h \in \mathcal{B}(\mathbb{R}^d)\}$. By definition $\mathcal{A}_h \subseteq \mathcal{B}(\mathbb{R}^d)$. One can easily check that $\mathcal{A}_h$ is a σ-algebra. Moreover, for each open set A one has $A + h$ is open and hence $A + h \in \mathcal{B}(\mathbb{R}^d)$. Thus, $\mathcal{B}(\mathbb{R}^d) = \sigma\{O \subseteq \mathbb{R}^d : O \text{ is open}\} \subseteq \mathcal{A}_h$, from which the claim follows.

Define μ_h on $\mathcal{B}(\mathbb{R}^d)$ by $\mu_h(A) := \lambda(A + h)$. Then μ_h is a measure and for any half-open rectangle I, $\mu_h(I) = |I + h| = |I| = \lambda(I)$. By the uniqueness of step 2, we find $\mu_h(A) = \lambda(A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$ and this proves the result. ☺

Remark 3.10. From Theorem B.4 one can actually see that for any $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\lambda(A) = \inf\Big\{\sum_{j=1}^\infty |I_j| : A \subseteq \bigcup_{j=1}^\infty I_j \text{ for } I_1, \ldots, I_n \in \mathcal{I}^d\Big\},$$

but we will not use this formula.

Exercises on the Lebesgue measure

Exercise 3.1.

(a) Show that any countable subset $A \subseteq \mathbb{R}^d$ is in $\mathcal{B}(\mathbb{R}^d)$.
Hint: First show that $\{x\} \in \mathcal{B}(\mathbb{R}^d)$ for every $x \in \mathbb{R}^d$.

(b) Show that any countable subset $A \subseteq \mathbb{R}^d$ satisfies $\lambda(A) = 0$. In particular, $\lambda(\mathbb{Q}^d) = 0$.
Hint: First show that $\lambda(\{x\}) = 0$.

[16] This is called **translation invariance**. Up to a scaling factor λ is the only measure on $\mathcal{B}(\mathbb{R}^d)$ which satisfies this property. See Exercise 3.5.

Exercise* 3.2. For $a, b \in \mathbb{R}^d$ with $a = (\alpha_1, \ldots, \alpha_d)$ and $b = (\beta_1, \ldots, \beta_d)$ with $\alpha_j \le \beta_j$ for $j \in \{1, \ldots, d\}$ let

$$(a, b) := (\alpha_1, \beta_1) \times \ldots \times (\alpha_d, \beta_d) \text{ and } [a, b] := [\alpha_1, \beta_1] \times \ldots \times [\alpha_d, \beta_d]$$

be the open and closed rectangle, respectively. Prove that

$$\lambda\big((a, b)\big) = \lambda\big([a, b]\big) = \lambda\big((a, b]\big)$$

and thus all coincide with the volume of the rectangle.
Hint: Use Theorem 2.9.

Exercise* 3.3 (Uncountable sets can have measure zero). Read the construction of the Cantor set in [4, p25-26]. Show that the Cantor set Δ is in $\mathcal{B}(\mathbb{R})$ and satisfies $\lambda(\Delta) = 0$.

Exercise* 3.4. For $A \subseteq \mathbb{R}$ and $t \ge 0$ let $tA := \{tx : x \in A\}$. Show that for each $A \in \mathcal{B}(\mathbb{R})$, $\lambda(tA) = t\lambda(A)$.
Hint: Use the same method as in the proof of Theorem 3.9.

Exercise* 3.5. Let $\mu\colon \mathcal{B}(\mathbb{R}) \to [0, \infty]$ be a measure such that for all $h \in \mathbb{R}$ and $A \in \mathcal{B}(\mathbb{R})$, $\mu(A + h) = \mu(A)$. Let $c := \mu((0, 1])$ and assume $c \in (0, \infty)$. Prove the following assertions:[17]

(a) For each $x \ge 0$ and $q \in \mathbb{N}$, $\mu\big((0, qx]\big) = q\mu\big((0, x]\big)$.
(b) For each $p, q \in \mathbb{N}$, $\mu\big((0, \frac{p}{q}]\big) = c\frac{p}{q}$.
(c) For each $x \ge 0$, $\mu\big((0, x]\big) = cx$.
(d) For each $a \le b$, $\mu\big((a, b]\big) = c(b - a)$.
(e) For each $A \in \mathcal{B}(\mathbb{R})$, $\mu(A) = c\lambda(A)$.

Exercise** 3.6 (Lebesgue-Stieltjes measure[18]). Let $a < b$ and let $F : \mathbb{R} \to \mathbb{R}$ be right-continuous and increasing. Show that $\mu((a, b]) = F(b) - F(a)$ for $a \le b$ extends to a measure on $\mathcal{B}(\mathbb{R})$.

Exercise** 3.7 (A non-Borel set). For $x, y \in [0, 1]$ define $x \sim y$ if $x - y \in \mathbb{Q}$, which is an equivalence relation on $[0, 1]$. Let $A \subseteq [0, 1]$ contain exactly one representative of each equivalence class in $[0, 1]/\sim$.

(a) Show that
$$[0, 1] \subseteq \bigcup_{x \in [-1,1] \cap \mathbb{Q}} (A + x) \subseteq [-1, 2].$$
(b) Show that $A \notin \mathcal{B}(\mathbb{R})$.

Exercises on general measures

Exercise* 3.8. Let $\mathcal{A}$ be a σ-algebra on S and let $T \subseteq S$. Define the restricted σ-algebra $\mathcal{A}_T$ on T by $\mathcal{A}_T := \{A \cap T : A \in \mathcal{A}\}$.

(a) Show that $\mathcal{A}_T$ is a σ-algebra.
(b) For $T \in \mathcal{A}$, show that $\mathcal{A}_T = \{A \subseteq T : A \in \mathcal{A}\}$.
(c) Let μ be a measure on $(S, \mathcal{A})$. For $T \in \mathcal{A}$ show that the restriction of μ to $\mathcal{A}_T$ is a measure.

Exercise 3.9 (Non-uniqueness of extensions I). Let $S := \{1, 2, 3, 4\}$ and let $\mathcal{F} := \{\{1, 2\}, \{1, 3\}\}$. Define $\mu : \mathcal{F} \to [0, \infty]$ by $\mu(\{1, 2\}) = \mu(\{1, 3\}) = \frac{1}{2}$. Find two different extensions of μ to $\sigma(\mathcal{F}) = \mathcal{P}(S)$. Why does this not contradict Proposition 3.5?

Exercise* 3.10 (Non-uniqueness of extensions II). Let $S := \mathbb{N}$ and let $\mathcal{F} := \big\{\{n, n + 1, \ldots\} : n \in \mathbb{N}\big\}$.

(a) Show that $\mathcal{F}$ is a π-system.
(b) Show that $\sigma(\mathcal{F}) = \mathcal{P}(\mathbb{N})$.
(c) Let τ be the counting measure and let $\mu : \mathcal{P}(\mathbb{N}) \to [0, \infty]$ be defined by $\mu(A) = 2\tau(A)$. Show that $\tau = \mu$ on $\mathcal{F}$. Why does this not contradict Proposition 3.5?

---

[17] From this exercise we see that up to a scaling factor, λ is the only translation invariant measure on $\mathcal{B}(\mathbb{R})$. The case for dimensions $d \ge 2$ holds as well and can be proved in a similar way.

[18] Thomas Stieltjes 1856-1894 was a Dutch mathematician working in Analysis. He has even worked in Delft.

4. Measurable functions

Let $(S, \mathcal{A}, \mu)$ be a measure space. One of our main goals is to integrate functions $f : S \to \mathbb{R}$. To accomplish this, we will use discretization in the range of f.[19] Therefore, we are interested in determining the measure of sets such as

$$A_{x,y} = \{s \in S : f(s) \in [x, y]\} = f^{-1}([x, y]), \qquad x, y \in \mathbb{R}.$$

For this to make sense, we will need that these sets are elements of $\mathcal{A}$. This motivates the following definition.

Definition 4.1. *Let $(S, \mathcal{A})$ and $(T, \mathcal{B})$ be two measurable spaces. A function $f : S \to T$ is called* ***measurable*** *if for each $B \in \mathcal{B}$, one has $f^{-1}(B) \in \mathcal{A}$.*

Recall that $f^{-1}(B) = \{s \in S : f(s) \in B\}$ is called the inverse image of B by f. One sometimes writes $\{f \in B\}$ for the same set.

The following functions will play an important role in our integration theory.

Example 4.2. Let $(S, \mathcal{A})$ be a measurable space and for $A \subseteq S$ define $\mathbf{1}_A : S \to \mathbb{R}$ as

$$\mathbf{1}_A(s) := \begin{cases} 1, & s \in A, \\ 0, & s \notin A. \end{cases}$$

Then $\mathbf{1}_A$ is measurable if and only if $A \in \mathcal{A}$, see Exercise 4.4(a).

Let us examine some properties of measurable functions. To begin, we can note that the composition of two measurable functions is also measurable. The proof of this fact is Exercise 4.1.

Lemma 4.3. *Let $(S_j, \mathcal{A}_j)$ be measurable spaces for $j = 1, 2, 3$. If $f : S_1 \to S_2$ and $g : S_2 \to S_3$ are both measurable, then their composition $g \circ f : S_1 \to S_3$ is measurable.*

It can be challenging to verify that $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$. Fortunately, it is sufficient to confirm this for B in a generating collection $\mathcal{F} \subseteq \mathcal{B}$. When $(T, \mathcal{B}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, common choices of $\mathcal{F}$ include $\mathcal{F} = \mathcal{F}^d$ and $\mathcal{F} = \{O \subseteq \mathbb{R}^d : O \text{ is open}\}$.

Lemma 4.4. *Let $(S, \mathcal{A})$ and $(T, \mathcal{B})$ be two measurable spaces and let $f : S \to T$. Suppose $\mathcal{F} \subseteq \mathcal{B}$ is such that $\sigma(\mathcal{F}) = \mathcal{B}$. If $f^{-1}(F) \in \mathcal{A}$ for all $F \in \mathcal{F}$, then f is measurable.*

Proof. Define $\widetilde{\mathcal{B}} := \{B \in \mathcal{B} : f^{-1}(B) \in \mathcal{A}\}$. Our aim is to show that $\widetilde{\mathcal{B}} = \mathcal{B}$. We claim that $\widetilde{\mathcal{B}}$ is a σ-algebra. Indeed, since $f^{-1}(\varnothing) = \varnothing \in \mathcal{A}$ also $\varnothing \in \widetilde{\mathcal{B}}$. Similarly, since $f^{-1}(T) = S \in \mathcal{A}$, we find $T \in \widetilde{\mathcal{B}}$. If $B \in \widetilde{\mathcal{B}}$, then $f^{-1}(T \setminus B) = S \setminus f^{-1}(B) \in \mathcal{A}$, which implies that $B^c \in \widetilde{\mathcal{B}}$. If $B_1, B_2, \ldots \in \widetilde{\mathcal{B}}$, then

$$f^{-1}\Big(\bigcup_{n=1}^{\infty} B_n\Big) = \bigcup_{n=1}^{\infty} f^{-1}(B_n) \in \mathcal{A}.$$

Therefore, the claim follows. Since $\mathcal{F} \subseteq \widetilde{\mathcal{B}}$, the claim yields $\mathcal{B} = \sigma(\mathcal{F}) \subseteq \widetilde{\mathcal{B}} \subseteq \mathcal{B}$, which implies $\widetilde{\mathcal{B}} = \mathcal{B}$. ☺

In the sequel, a metric space (M, d) will always be equipped with its Borel σ-algebra $\mathcal{B}(M)$, unless otherwise stated.

Proposition 4.5 (Continuous mappings are measurable)**.** *Let (M, d) and (N, ρ) be metric spaces. If $f : M \to N$ is continuous, then f is measurable.*[20]

Proof. By the continuity of f, we find that for all open $O \subseteq Y$ the inverse image is $f^{-1}(O)$ open in (M, d) and hence $f^{-1}(O) \in \mathcal{B}(M)$. Since the open sets of N generate the Borel σ-algebra, the result follows from Lemma 4.4 with $\mathcal{F} = \{O \subseteq N : O \text{ is open}\}$. ☺

[19] In Riemann integration of functions $f : \mathbb{R}^d \to \mathbb{R}$ the discretization is always done in the domain of the function. This is one of the major differences with Lebesgue integration.

[20] In the setting of Borel-σ-algebras, measurable functions are often called Borel measurable.

Example 4.6. Let $a < b$ and suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous. Then $g\colon \mathbb{R} \to \mathbb{R}$ given by

$$g(x) := \begin{cases} f(x), & x \in [a,b], \\ 0, & x \notin [a,b], \end{cases}$$

is measurable, see Exercise 4.4(b).

4.1. **Real-valued functions.** The most frequent case we will encounter is when $f\colon S \to \mathbb{R}$, where $(S, \mathcal{A})$ is a measurable space and $\mathbb{R}$ is equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. The following characterization of measurability will prove useful in this context.

Proposition 4.7 (Real-valued functions)**.** *Let $(S, \mathcal{A})$ be a measurable space. For $f\colon S \to \mathbb{R}$ the following are equivalent:*

(i) f is measurable.
(ii) For all $a, b \in \mathbb{R}$, one has $f^{-1}((a,b)) \in \mathcal{A}$.
(iii) For all $r \in \mathbb{R}$, one has $f^{-1}((-\infty, r)) \in \mathcal{A}$.

Proof. (i) $\Rightarrow$ (ii) is trivial. For (ii) $\Rightarrow$ (iii) we note that

$$f^{-1}((-\infty, r)) = \bigcup_{n=1}^{\infty} f^{-1}((-r-n, r)).$$

For (iii) $\Rightarrow$ (i), let $\mathcal{F} := \{(-\infty, r) : r \in \mathbb{R}\}$. In Exercise 1.8 we have seen that $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R})$. Therefore, f is measurable by Lemma 4.4. ☺

Measurability is preserved under operations like addition, (scalar) multiplication and division, as we will demonstrate in the following theorem. For $x, y \in \mathbb{R}$ we define

$$x \vee y := \max\{x, y\},$$
$$x \wedge y := \min\{x, y\}.$$

Theorem 4.8. *Let $(S, \mathcal{A})$ be a measurable space. Let $f, g : S \to \mathbb{R}$ be measurable functions and let $\alpha \in \mathbb{R}$. Then the following functions are measurable as well:*

$$f+g,\ f-g,\ f\cdot g,\ f\vee g,\ f\wedge g,\ f^+ := f\vee 0,\ f^- := (-f)\vee 0,\ |f|,\ \alpha\cdot f,\ \frac{1}{f}\ (\textit{if } f\neq 0 \textit{ on } S).$$

Proof. We first claim that $h : S \to \mathbb{R}^2$ given by $h(s) := (f(s), g(s))$ is measurable. To prove this, observe that for all half-open rectangles $I = I_1 \times I_2 \subseteq \mathbb{R}^2$, one has

$$h^{-1}(I) = \{s \in S : f(s) \in I_1 \text{ and } g(s) \in I_2\} = f^{-1}(I_1) \cap g^{-1}(I_2) \in \mathcal{A}.$$

By Exercise 1.10 we have $\sigma(\mathcal{I}^2) = \mathcal{B}(\mathbb{R}^2)$, so we can use Lemma 4.4 to deduce that h is measurable.

To prove the statements, we use that continuous functions are measurable by Proposition 4.5 and that the composition of measurable functions is measurable by Lemma 4.3. For instance let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be given by $\varphi(x,y) := x + y$. Then φ is continuous and $f + g = \varphi \circ h$, from which the claimed measurability of $f + g$ follows.

The proofs for the difference, product, maximum and minimum are similar. The measurability of $f^{\pm}$, $|f|$ and $\alpha \cdot f$ follow in the same way by rewriting them as $\varphi \circ f$ for a suitable continuous function $\varphi\colon \mathbb{R} \to \mathbb{R}$.

The statement for $\frac{1}{f}$ requires some comment. Let $\varphi : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be given by $\varphi(x) := \frac{1}{x}$. Note that $\mathbb{R} \setminus \{0\}$ is a metric space on which φ is continuous. Now, for all open sets $B \in \mathcal{B}(\mathbb{R})$, $C := \varphi^{-1}(B)$ is open in $\mathbb{R} \setminus \{0\}$ and hence in $\mathcal{B}(\mathbb{R})$. Thus $(\varphi \circ f)^{-1}(B) = f^{-1}(C) \in \mathcal{A}$. Therefore, the measurability of $\varphi \circ f$ follows from Lemma 4.4. ☺

In applications it will be useful to work with functions $f : S \to \overline{\mathbb{R}}$, where

$$\overline{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}.$$

For this, we introduce an analogue of the Borel σ-algebra on $\overline{\mathbb{R}}$.

Definition 4.9. *Let $\mathcal{B}(\overline{\mathbb{R}})$ be the σ-algebra generated by the sets $\{\infty\}, \{-\infty\}$ and $\mathcal{B}(\mathbb{R})$. Similarly, let $\mathcal{B}([0,\infty])$ be the σ-algebra generated by the set $\{\infty\}$ and $\mathcal{B}([0,\infty))$.*

We will always equip $\overline{\mathbb{R}}$ and $[0,\infty]$ with their Borel σ-algebras $\mathcal{B}(\overline{\mathbb{R}})$ and $\mathcal{B}([0,\infty])$ respectively. From Lemma 4.4 we see that a function $f : S \to \overline{\mathbb{R}}$ is measurable if and only if

$$\{s \in S : f(s) = \infty\}, \{s \in S : f(s) = -\infty\} \in \mathcal{A}$$

and $f^{-1}(B) \in \mathcal{A}$ for each $B \in \mathcal{B}(\mathbb{R})$.

Remark 4.10. One could view $\overline{\mathbb{R}}$ as a metric space with metric

$$d(x,y) := |\arctan(x) - \arctan y|, \qquad x, y \in \overline{\mathbb{R}},$$

where $\arctan(\pm\infty) = \pm\frac{\pi}{2}$. Then $\mathcal{B}(\overline{\mathbb{R}})$ defined in Definition 4.9 coincides with the Borel σ-algebra of the metric space $(\overline{\mathbb{R}}, d)$.

We extend addition, multiplication, and division to $\overline{\mathbb{R}}$ in the following way:

$$\begin{aligned}
\infty + a = a + \infty &= \infty, && \text{for all } a \in (-\infty,\infty] \\
-\infty + a = a - \infty &= -\infty, && \text{for all } a \in [-\infty,\infty) \\
\infty \cdot a = a \cdot \infty &= \infty, && \text{for all } a \in (0,\infty] \\
\infty \cdot a = a \cdot \infty &= -\infty, && \text{for all } a \in [-\infty,0) \\
\infty \cdot 0 = 0 \cdot \infty = \frac{a}{\infty} = \frac{a}{-\infty} &= 0 && \text{for all } a \in (-\infty,\infty)
\end{aligned}$$

Moreover, for a sequence $(x_n)_{n\to\infty}$ in $\overline{\mathbb{R}}$ we say that $\lim_{n\to\infty} x_n = \pm\infty$ if it either diverges to $\pm\infty$ or is eventually equal to $\pm\infty$. In this setting Theorem 4.8 remains true[21] for functions $f, g : S \to \overline{\mathbb{R}}$.

The next result will show that measurability is preserved under taking countable suprema, countable infima and limits of sequences. For a sequence of numbers $(x_n)_{n\geq 1}$ in $\overline{\mathbb{R}}$, let

$$\limsup_{n\to\infty} x_n := \lim_{k\to\infty} \sup_{n\geq k} x_n \qquad \text{and} \qquad \liminf_{n\to\infty} x_n := \lim_{k\to\infty} \inf_{n\geq k} x_n.$$

The limit of $y_k := \sup_{n\geq k} x_n$ exists since $(y_k)_{k\geq 1}$ is decreasing. Moreover,

$$\limsup_{n\to\infty} x_n = \lim_{k\to\infty} y_k = \inf_{k\geq 1} y_k = \inf_{k\geq 1} \sup_{n\geq k} x_n. \tag{4.1}$$

Similar formulas hold for the $\liminf_{n\to\infty} x_n$. Therefore $\limsup_{n\to\infty} x_n$ and $\liminf_{n\to\infty} x_n$ always exist and

$$\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n.$$

Moreover they both coincide with $\lim_{n\to\infty} x_n$ if and only if $(x_n)_{n\geq 1}$ converges in $\overline{\mathbb{R}}$.

Theorem 4.11. *Let $(S, \mathcal{A})$ be a measurable space. For each $n \in \mathbb{N}$ let $f_n : S \to \overline{\mathbb{R}}$ be a measurable function. Then each of the following functions is measurable as well:*[22]

$$\sup_{n\geq 1} f_n, \quad \inf_{n\geq 1} f_n, \quad \limsup_{n\to\infty} f_n, \quad \liminf_{n\to\infty} f_n.$$

Moreover, if $f_n \to f$ pointwise[23]*, then f is measurable again.*

Proof. Let $g = \sup_{n\geq 1} f_n$. Then for each $r \in \mathbb{R}$,

$$g^{-1}([-\infty, r]) = \{s \in S : g(s) \leq r\} = \{s \in S : f_n(s) \leq r \text{ for all } n \in \mathbb{N}\} = \bigcap_{n=1}^{\infty} f_n^{-1}([-\infty, r]) \in \mathcal{A}.$$

Using a similar argument as in Exercise 1.8, one can show that $\sigma(\{[-\infty, r] : r \in \mathbb{R}\}) = \mathcal{B}(\overline{\mathbb{R}})$. Therefore, the measurability of g follows from Lemma 4.4. The case of infima follows from $\inf_{n\geq 1} f_n = -\sup_{n\geq 1}(-f_n)$.

[21] We do not define $\infty - \infty$, so some cases need to be excluded. For the proof one additionally needs to check in each of the cases that inverse images of $\{\infty\}$ and $\{-\infty\}$ are measurable. We leave this to the reader.

[22] Here it is important what we work with *countable* suprema and infima.

[23] Here we allow divergence to $\pm\infty$.

By (4.1), we can write $\limsup_{n\to\infty} f_n = \inf_{k\geq 1}\sup_{n\geq k} f_k$. Therefore, the measurability follows from the previous cases. The remaining cases follow from $\liminf_{n\to\infty} f_n = -\limsup_{n\to\infty}(-f_n)$ and $\lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$. ☺

4.2. **Simple functions.** In Section 5, the Lebesgue integral will first be defined for linear combinations of indicator functions, which were defined in Example 4.2. Then, we will approximate any measurable function by using linear combinations of indicator functions. For ease of reference, we will give such functions a descriptive name.

Definition 4.12. *A function $f : S \to \mathbb{R}$ is called a* ***simple function***[24] *if f is measurable and takes only finitely many values.*

Letting $x_1, \dots, x_n \in \mathbb{R}$ denote the *distinct* values of f and $A_k = \{s \in S : f(s) = x_k\}$, we can always represent a simple function as

$$f = \sum_{k=1}^{n} x_k \cdot \mathbf{1}_{A_j},$$

i.e. as a linear combination of indicator functions. Of course, if $x_k = 0$ for some $k \in \{1, \dots, n\}$, we can leave it out from the sum.

Example 4.13. Let $S = \mathbb{R}$ and $\mathcal{A} = \mathcal{B}(\mathbb{R})$. The following functions are simple functions:

(a) $f = \pi \cdot \mathbf{1}_{(0,1)} - 4 \cdot \mathbf{1}_{(13,14]} + 5 \cdot \mathbf{1}_{\mathbb{Z}\cap(-\infty,0)}$.
(b) $f = \mathbf{1}_{\mathbb{Q}}$.

As a preparation for the Lebesgue integral in Section 5, we show that measurable functions can be written as pointwise limits of simple functions. For this we discretize in the range space in a suitable way.

In the sequel, we write $f_n \uparrow f$ if $(f_n(s))_{n\geq 1}$ is increasing and $\lim_{n\to\infty} f_n(s) = f(s)$ for all $s \in S$.

Theorem 4.14. *Let $(S, \mathcal{A})$ be a measurable space.*

(i) Let $f : S \to [0, \infty]$ be measurable. Then there exists a sequence of simple functions $(f_n)_{n\geq 1}$ such that $0 \leq f_n \uparrow f$.
(ii) Let $f : S \to \overline{\mathbb{R}}$ be measurable. Then there exists a sequence of simple functions $(f_n)_{n\geq 1}$ such that $f_n \to f$ pointwise.

Proof. (i):[25] For each $n \in \mathbb{N}$ and $j \in \{0, 1, \dots, 4^n - 1\}$, let

$$\begin{aligned} A_{n,m} &:= \{s \in S : m2^{-n} < f(s) \leq (m+1)2^{-n}\} = f^{-1}\big((\tfrac{m}{2^n}, \tfrac{m+1}{2^n}]\big), \\ A_n &:= \{s \in S : f(s) > 2^n\} = f^{-1}\big((\tfrac{1}{2^n}, \infty]\big). \end{aligned}$$

Then for each n, m one has $A_{n,m}, A_n \in \mathcal{A}$. Define[26]

$$f_n := 2^n \mathbf{1}_{A_n} + \sum_{m=0}^{4^n-1} \frac{m}{2^n} \mathbf{1}_{A_{n,m}}.$$

It is clear that each f_n takes finitely many values. Moreover, by Example 4.2 and Theorem 4.8, each f_n is measurable. Thus each f_n is a simple function.

Now fix $s \in S$. We first prove $0 \leq f_n(s) \leq f_{n+1}(s)$ for each $n \in \mathbb{N}$. First assume $f(s) < 2^n$. Then selecting the unique $m \in \{0, \dots, 4^n - 1\}$ such that $s \in A_{n,m}$, we find that $f_n(s) = m2^{-n}$. Similarly, we can select $k \in \{0, \dots, 4^{n+1} - 1\}$ such that $s \in A_{n+1,k}$ and thus $f_{n+1}(s) = k2^{-(n+1)}$. Since $s \in A_{n,m}$, we know that $f(s) > m2^{-n} = 2m2^{-(n+1)}$. Since $s \in A_{n+1,k}$, we also have $f(s) \leq (k+1)2^{-(n+1)}$. This implies that $2m < k+1$, which is equivalent to $2m \leq k$. We conclude that

$$f_n(s) = m2^{-n} \leq k2^{-(n+1)} = f_{n+1}(s).$$

[24] In part of the literature this is called a step function, but we will use this name for a different class of functions.

[25] To visualize the proof, you could make a picture where put the set S on the x-axis and partition the y-axis into intervals of length 2^{-n}.

[26] The idea is that for each $n \in \mathbb{N}$ we approximate f up to 2^{-n} on the set $\{f < 2^n\}$.

The case $f(s) > 2^n$ can be treated similarly and is left to the reader.

To prove that $f_n(s) \to f(s)$, first assume $f(s) < \infty$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ so large that $f(s) \leq 2^N$ and $2^{-N} < \varepsilon$. Let $n \geq N$. Selecting $m \in \{0, \dots, 4^n - 1\}$ such that $s \in A_{n,m}$, we find that

$$|f(s) - f_n(s)| \leq 2^{-n} \leq 2^{-N} < \varepsilon.$$

Therefore, $f_n(s) \to f(s)$ in this case. If $f(s) = \infty$, then $f_n(s) = 2^n$ for every $n \in \mathbb{N}$ and thus $f_n(s) \to \infty = f(s)$.

(ii): Write $f = f^+ - f^-$. By (i) we can find simple functions $f_{n,+}, f_{n,-} : S \to \mathbb{R}$ such that $f_{n,+} \to f^+$ and $f_{n,-} \to f^-$. Let $f_n = f_{n,+} - f_{n,-}$ for $n \in \mathbb{N}$. For $s \in S$ we have

$$f(s) = f^+(s) - f^-(s) = \lim_{n\to\infty} f_{n,+}(s) - \lim_{n\to\infty} f_{n,-}(s) = \lim_{n\to\infty} f_n(s),$$

proving the statement. ☺

Exercises

Exercise 4.1. Prove Lemma 4.3.

Exercise 4.2. Let $(S, \mathcal{A})$ be a measurable space. Let $f, g : S \to \mathbb{R}$ be measurable functions. Show that $\{s \in S : f(s) = g(s)\} \in \mathcal{A}$ and $\{s \in S : f(s) < g(s)\} \in \mathcal{A}$.

Exercise 4.3. Let $(S, \mathcal{A})$ be a measurable space. Let $f : S \to \mathbb{R}$ be a measurable function and $p \in (0, \infty)$. Show that the function $|f|^p$ is measurable.

Exercise 4.4.

(a) Prove Example 4.2.
(b) Prove Example 4.6.
Hint: $\mathcal{B}([a, b]) \subseteq \mathcal{B}(\mathbb{R})$.

Exercise 4.5. Let $(S, \mathcal{A})$ be a measurable space. Let $f_1, f_2, \dots : S \to \mathbb{R}$ be measurable functions and $A_1, A_2, \dots \in \mathcal{A}$ be disjoint. Show that the function $f = \sum_{n=1}^{\infty} \mathbf{1}_{A_n} f_n$ is measurable.

Exercise* 4.6 (Vector-valued functions). Let $(S, \mathcal{A})$ be a measurable space. For each $j \in \{1, \dots, d\}$ let $f_j : S \to \mathbb{R}$ be a function and let $f : S \to \mathbb{R}^d$ be given by $f := (f_1, \dots, f_d)$. Prove that f is measurable if and only if f_j is measurable for each $j \in \{1, \dots, d\}$.
Hint: For the "if part" one can use the same technique as in Theorem 4.8.

Exercise* 4.7 (Monotone functions). Assume that $f : \mathbb{R} \to \mathbb{R}$ is increasing. Show that f is measurable, where as usually on $\mathbb{R}$ we consider the Borel σ-algebra.
Hint: Use Proposition 4.7.

Exercise* 4.8 (Set of convergence). Let $(S, \mathcal{A})$ be a measurable space. Let $f_1, f_2, \dots : S \to \mathbb{R}$ be measurable functions. Define

$$A := \{s \in S : (f_n(s))_{n\geq 1} \text{ is convergent}\}.$$

(a) Explain why $A = \{s \in S : (f_n(s))_{n\geq 1} \text{ is a Cauchy sequence}\}$.
(b) Show that for each $k, n, m \in \mathbb{N}$ the set $A(k, n, m) := \{s \in S : |f_n(s) - f_m(s)| < \frac{1}{k}\}$ is in $\mathcal{A}$.
(c) Show that $A = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{m=N}^{\infty} \bigcap_{n=N}^{\infty} A(k, n, m)$.
Hint: $s \in A$ if and only if $\forall k \in \mathbb{N}, \exists N \in \mathbb{N} : \forall n \in \mathbb{N}, \forall m \in \mathbb{N} : |f_n(s) - f_m(s)| < \frac{1}{k}$ and connect the universal quantifier with an intersection and the existential quantifier with a union.
(d) Conclude that $A \in \mathcal{A}$.

Exercise** 4.9 (Pointwise convergence to uniform convergence). Let $(S, \mathcal{A}, \mu)$ be a measure space and suppose $\mu(S) < \infty$. Let $f_n, f : S \to \mathbb{R}$ be measurable functions for $n \geq 1$ such that $f_n \to f$ pointwise. For $n, m \geq 1$ define

$$A_{n,m} := \big\{s \in S : |f_k(s) - f(s)| < \tfrac{1}{m} \text{ for all } k \geq n\big\}.$$

(a) Show that $A_{n,m} \in \mathcal{A}$ for all $n, m \geq 1$.
(b) Prove that for any $m \geq 1$ we have
$$\lim_{n\to\infty} \mu(A_{n,m}) = \mu(S).$$

Now fix $\varepsilon > 0$.

(c) Show that for any $m \geq 1$ there exists an $n_m \geq 1$ such that
$$\mu(S \setminus A_{n_m,m}) < \frac{\varepsilon}{2^m}.$$
(d) Let $A = \bigcap_{m=1}^{\infty} A_{n_m,m}$, where n_m is as in (c). Show that $\mu(S \setminus A) < \varepsilon$.
(e) Prove that
$$\lim_{n\to\infty} \sup_{s\in A} |f_n(s) - f(s)| = 0,$$
i.e. $f_n \to f$ uniformly on A.

5. Construction of the Lebesgue integral

Throughout this section, $(S, \mathcal{A}, \mu)$ is a measure space. Our goal is to construct an integral, called the Lebesgue integral, for measurable functions $f : S \to \overline{\mathbb{R}}$. We will denote this integral by

$$\int_S f \, \mathrm{d}\mu \quad \text{or} \quad \int_S f(s) \, \mathrm{d}\mu(s).$$

The Lebesgue integral will be built in three steps:

(1) For simple functions $f : S \to [0, \infty)$;
(2) For measurable functions $f : S \to [0, \infty]$;
(3) For (certain) measurable functions $f : S \to \overline{\mathbb{R}}$.

The advantage of this setting is that it works for any measure space $(S, \mathcal{A}, \mu)$. Moreover, in the special case of the Lebesgue measure, we will show that it extends the Riemann integral.

Before we begin our construction, we introduce the following convenient terminology.

Definition 5.1. *For measurable functions $f, g : S \to \overline{\mathbb{R}}$ we say that $f = g$* ***almost everywhere*** *if $\mu(\{s \in S : f(s) \neq g(s)\}) = 0$, which is denoted as $f = g$ a.e.*[27] *Similarly, one can define $f \leq g$ a.e. and $f < g$ a.e.*

5.1. Lebesgue integral for simple functions.

Definition 5.2 (Lebesgue integral for simple functions)**.** *Let $f : S \to [0, \infty)$ be a simple function. Let $x_1, \ldots, x_n \in [0, \infty)$ and let $(A_k)_{k=1}^n$ be disjoint sets in $\mathcal{A}$ such that*

$$f = \sum_{k=1}^{n} x_k \cdot \mathbf{1}_{A_k}. \tag{5.1}$$

For $E \in \mathcal{A}$ we define

$$\int_E f \, \mathrm{d}\mu := \sum_{k=1}^{n} x_k \cdot \mu(E \cap A_k),$$

which is called the ***Lebesgue integral*** *of f over the set E.*

If f as in (5.1) is represented in a different way as $f = \sum_{j=1}^m y_j \cdot \mathbf{1}_{B_j}$ with $(y_1, \ldots, y_m) \in [0, \infty)$ and $(B_j)_{j=1}^m$ disjoint, one can check that

$$\sum_{k=1}^{n} x_k \cdot \mu(A_k \cap E) = \sum_{j=1}^{m} y_j \cdot \mu(B_j \cap E),$$

see the proof of Lemma 5.3(ii) below. Thus, the Lebesgue integral of f is well-defined.

Clearly $\int_E f \, \mathrm{d}\mu \in [0, \infty]$ for every $E \in \mathcal{A}$. We continue with some other basic properties of the Lebesgue integral for simple functions

Lemma 5.3. *Let $f, g \colon S \to [0, \infty)$ be simple functions. Then the following hold:*

(i) For all $E \in \mathcal{A}$, $\int_E f \, \mathrm{d}\mu = \int_S \mathbf{1}_E f \, \mathrm{d}\mu$.
(ii) (Monotonicity I) If $E \in \mathcal{A}$ and $f \leq g$ on E, then $\int_E f \, \mathrm{d}\mu \leq \int_E g \, \mathrm{d}\mu$.
(iii) (Monotonicity II) If $E, F \in \mathcal{A}$ satisfy $E \subseteq F$, then $\int_E f \, \mathrm{d}\mu \leq \int_F f \, \mathrm{d}\mu$.
(iv) (Linearity) For all $E \in \mathcal{A}$ and $\alpha, \beta \in [0, \infty)$, $\int_E \alpha f + \beta g \, \mathrm{d}\mu = \alpha \int_E f \, \mathrm{d}\mu + \beta \int_E g \, \mathrm{d}\mu$.
(v) (Additivity) For all disjoint sets $E, F \in \mathcal{A}$, $\int_{E \cup F} f \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu + \int_F f \, \mathrm{d}\mu$.
(vi) $\int_S f \, \mathrm{d}\mu = 0$ if and only if $f = 0$ almost everywhere.

Proof. Write

$$f = \sum_{j=1}^{m} x_j \cdot \mathbf{1}_{A_j}, \qquad x_1, \ldots, x_m \in [0, \infty), \qquad A_1, \ldots, A_m \in \mathcal{A} \text{ disjoint},$$

$$g = \sum_{k=1}^{n} y_k \cdot \mathbf{1}_{B_k} \qquad y_1, \ldots, y_n \in [0, \infty), \qquad B_1, \ldots, B_n \in \mathcal{A} \text{ disjoint}.$$

[27]In probability theory this is usually called almost surely and this is abbreviated as a.s.

Assume without loss of generality that $\bigcup_{j=1}^m A_j = S$ and $\bigcup_{k=1}^n B_k = S$. Let $C_{j,k} = A_j \cap B_k$ for $k = 1, \dots, n$ and $j = 1, \dots, m$. Then $(C_{j,k})_{j,k=1}^{m,n}$ are disjoint sets in $\mathcal{A}$ with $\bigcup_{k=1}^n C_{j,k} = A_j$ and $\bigcup_{j=1}^m C_{j,k} = B_k$.

(i): Note that $\mathbf{1}_{E\cap A} = \mathbf{1}_E \cdot \mathbf{1}_A$ and thus $\mathbf{1}_E f = \sum_{j=1}^m x_j \cdot \mathbf{1}_{E\cap A_j}$. Therefore

$$\int_E f \,\mathrm{d}\mu = \sum_{j=1}^m x_j \cdot \mu(E \cap A_k) = \int_S \mathbf{1}_E f \,\mathrm{d}\mu.$$

(ii): Note that by the additivity of μ

$$\int_E f \,\mathrm{d}\mu = \sum_{j=1}^m x_j \cdot \mu(E \cap A_j) = \sum_{j=1}^m \sum_{k=1}^n x_j \cdot \mu(E \cap C_{j,k}) \tag{5.2}$$

$$\int_E g \,\mathrm{d}\mu = \sum_{k=1}^n y_k \cdot \mu(E \cap B_k) = \sum_{k=1}^n \sum_{j=1}^m y_k \cdot \mu(E \cap C_{j,k}). \tag{5.3}$$

Therefore, we need to show $x_j \le y_k$ for all j, k such that $E \cap C_{j,k} \neq \varnothing$. But this is immediate. Indeed, for $s \in E \cap C_{j,k}$ we have

$$x_j = f(s) \le g(s) = y_k.$$

(iii): This follows from $\mu(E \cap A_j) \le \mu(F \cap A_j)$ for $j = 1, \dots, m$, which is immediate from the monotonicity of μ.

(iv): For disjoint sets $A, B \subseteq S$ one has $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$, so we can write

$$\alpha f + \beta g = \sum_{j=1}^m \sum_{k=1}^n (\alpha x_j + \beta y_k) \cdot \mathbf{1}_{C_{j,k}}.$$

Therefore,

$$\begin{aligned}\int_E \alpha f + \beta g \,\mathrm{d}\mu &= \sum_{j=1}^m \sum_{k=1}^n (\alpha x_j + \beta y_k)\mu(E \cap C_{j,k}) \\ &= \alpha \sum_{j=1}^m \sum_{k=1}^n x_j \cdot \mu(E \cap C_{j,k}) + \beta \sum_{k=1}^n \sum_{j=1}^m y_k \cdot \mu(E \cap C_{j,k}) = \alpha \int_E f \,\mathrm{d}\mu + \beta \int_E g \,\mathrm{d}\mu,\end{aligned}$$

where we used (5.2) and (5.3).

(v): Since $\mathbf{1}_{E\cup F} f = \mathbf{1}_E f + \mathbf{1}_F f$, we have

$$\int_{E_1\cup E_2} f \,\mathrm{d}\mu \overset{(i)}{=} \int_S \mathbf{1}_{E_1\cup E_2} f \,\mathrm{d}\mu \overset{(iv)}{=} \int_S \mathbf{1}_{E_1} f \,\mathrm{d}\mu + \int_S \mathbf{1}_{E_2} f \,\mathrm{d}\mu \overset{(i)}{=} \int_{E_1} f \,\mathrm{d}\mu + \int_{E_2} f \,\mathrm{d}\mu.$$

(vi): By leaving out some of those i for which $x_i = 0$, we can assume $x_j > 0$ for $j = 1, \dots, m$. Let $A = \{s \in S : f(s) > 0\}$ and observe that $A = \bigcup_{j=1}^m A_j$. Now $\int_S f \,\mathrm{d}\mu = \sum_{j=1}^m x_j \mu(A_j)$. Noting that $x_j >$ and $\mu(A_j) \ge 0$ for $j = 1, \dots, m$, we deduce that $\int_S f \,\mathrm{d}\mu = 0$ if and only if $\mu(A) = \sum_{j=1}^n \mu(A_j) = 0$. ☺

Example 5.4. Let λ be the Lebesgue measure on $\mathbb{R}$. Since $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$ by Exercise 1.7, it follows from Example 4.2 that $\mathbf{1}_{\mathbb{Q}}$ is measurable and for any $E \in \mathcal{B}(\mathbb{R})$ we have

$$\int_E \mathbf{1}_{\mathbb{Q}} \,\mathrm{d}\lambda = \lambda(E \cap \mathbb{Q}) = 0$$

by Exercise 3.1. Note that $\mathbf{1}_{\mathbb{Q}}$ is not Riemann integrable on any interval $[a, b]$ with $a < b$. More on the Riemann integral can be found in Theorem 5.17.

5.2. **Lebesgue integral for positive measurable functions.** In order to extend the definition of the Lebesgue integral to arbitrary measurable functions $f\colon S \to [0,\infty]$, we need the following lemma.

Lemma 5.5 (Consistency)**.** *Let $f : S \to [0,\infty]$ be a measurable function. Suppose that $(f_n)_{n\geq 1}$ is sequence of simple functions such that $0 \leq f_n \uparrow f$. Suppose $g : S \to [0,\infty)$ is a simple function such that $0 \leq g \leq f$. Then $\displaystyle\int_E g\,\mathrm{d}\mu \leq \lim_{n\to\infty} \int_E f_n\,\mathrm{d}\mu$ for every $E \in \mathcal{A}$.*

Observe that $\lim_{n\to\infty} \int_E f_n\,\mathrm{d}\mu$ exists in $[0,\infty]$, as it is an increasing sequence of real numbers.

Proof. Let $\varepsilon \in (0,1)$ and set $E_n = \{s \in E : (1-\varepsilon)g(s) \leq f_n(s)\}$ for $n \in \mathbb{N}$. Then using the indicated parts of Lemma 5.3 in the estimates below, we find

$$(1-\varepsilon)\int_{E_n} g\,\mathrm{d}\mu \overset{\text{(iv)}}{=} \int_{E_n} (1-\varepsilon)g\,\mathrm{d}\mu \overset{\text{(ii)}}{\leq} \int_{E_n} f_n\,\mathrm{d}\mu \overset{\text{(iii)}}{\leq} \int_E f_n\,\mathrm{d}\mu \overset{\text{(ii)}}{\leq} \lim_{n\to\infty}\int_E f_n\,\mathrm{d}\mu. \tag{5.4}$$

We still need to get rid of n and ε on the left-hand side.

We claim that $E_n \uparrow E$. Indeed, $E_n \subseteq E_{n+1}$ for all $n \geq 1$ since $f_n \leq f_{n+1}$. Moreover, by definition, we have $\bigcup_{n=1}^\infty E_n \subseteq E$. For the converse inclusion, take $s \in E$. We consider two cases

- If $f(s) < \infty$, since $f_n(s) \uparrow f(s)$ and $\varepsilon > 0$, we can find a $k \geq 1$ such that
$$f_k(s) \geq (1-\varepsilon)f(s) \geq (1-\varepsilon)g(s).$$
- If $f(s) = \infty$, since $f_n(s) \to \infty$ and $g(s) < \infty$, we can find a $k \geq 1$ such that
$$f_k(s) \geq g(s) \geq (1-\varepsilon)g(s).$$

Thus, in both cases, $s \in E_k \subseteq \bigcup_{n=1}^\infty E_n$. This proves the claim.

Now write $g = \sum_{j=1}^m x_j \cdot \mathbf{1}_{A_j}$ with $x_1,\dots,x_m \in [0,\infty)$ and $(A_j)_{j=1}^m$ in $\mathcal{A}$ disjoint. For $j = 1,\dots,m$ we have $E_n \cap A_j \uparrow E \cap A_j$ as $n \to \infty$, so Theorem 2.9 yields

$$\int_{E_n} g\,\mathrm{d}\mu = \sum_{j=1}^m x_j \cdot \mu(E_n \cap A_j) \to \sum_{i=1}^m x_j \cdot \mu(E \cap A_j) = \int_E g\,\mathrm{d}\mu. \tag{5.5}$$

Combined with (5.4), this shows

$$(1-\varepsilon)\int_E g\,\mathrm{d}\mu \leq \lim_{n\to\infty}\int_E f_n\,\mathrm{d}\mu.$$

Since $\varepsilon > 0$ was arbitrary, this finishes the proof. ☺

Definition 5.6 (Lebesgue integral for positive functions)**.** *Let $f : S \to [0,\infty]$ be a measurable function. Let $(f_n)_{n\geq 1}$ be a sequence of simple functions with $0 \leq f_n \uparrow f$. We define the **Lebesgue integral** of f over $E \in \mathcal{A}$ as*

$$\int_E f\,\mathrm{d}\mu := \lim_{n\to\infty}\int_E f_n\,\mathrm{d}\mu \quad \text{in } [0,\infty].$$

The above limit exists in $[0,\infty]$ since the numbers $a_n := \int_E f_n\,\mathrm{d}\mu$ form an increasing sequence $(a_n)_{n\geq 1}$ in $[0,\infty]$. Note that, by Theorem 4.14, we can always find a sequence of simple functions $f_n : S \to [0,\infty)$ such that $0 \leq f_n \uparrow f$. However, we need to check that the above definition does not depend on the choice of the sequence $(f_n)_{n\geq 1}$. Let $(g_m)_{m\geq 1}$ be another sequence of simple functions such that $0 \leq g_m \uparrow f$ and let $b_m = \int_E g_m\,\mathrm{d}\mu$. It suffices to show that $\lim_{n\to\infty} a_n = \lim_{m\to\infty} b_m$. By Lemma 5.5, for each $m \geq 1$, we have

$$b_m = \int_E g_m\,\mathrm{d}\mu \leq \lim_{n\to\infty}\int_E f_n\,\mathrm{d}\mu = \lim_{n\to\infty} a_n.$$

From this we obtain $\lim_{m\to\infty} b_m \leq \lim_{n\to\infty} a_n$. Reversing the roles of g_m and f_n, one sees that the converse holds as well.

We can extend the properties of the Lebesgue integral in Lemma 5.3 to positive measurable functions.

Proposition 5.7. *Let $f, g : S \to [0,\infty]$ be measurable functions. Then the following hold:*

(i) For all $E \in \mathcal{A}$, $\int_E f \,\mathrm{d}\mu = \int_S \mathbf{1}_E f \,\mathrm{d}\mu$.
(ii) (Monotonicity I) If $E \in \mathcal{A}$ and $f \le g$ on E, then $\int_E f \,\mathrm{d}\mu \le \int_E g \,\mathrm{d}\mu$.
(iii) (Monotonicity II) If $E, F \in \mathcal{A}$ satisfy $E \subseteq F$, then $\int_E f \,\mathrm{d}\mu \le \int_F f \,\mathrm{d}\mu$.
(iv) (Linearity) For all $E \in \mathcal{A}$ and $\alpha, \beta \in [0,\infty)$, $\int_E \alpha f + \beta g \,\mathrm{d}\mu = \alpha \int_E f \,\mathrm{d}\mu + \beta \int_E g \,\mathrm{d}\mu$.
(v) (Additivity) For all disjoint sets $E_1, E_2 \in \mathcal{A}$, $\int_{E_1 \cup E_2} f \,\mathrm{d}\mu = \int_{E_1} f \,\mathrm{d}\mu + \int_{E_2} f \,\mathrm{d}\mu$.
(vi) $\int_S f \,\mathrm{d}\mu = 0$ if and only if $f = 0$ almost everywhere.

Proof. (i): Let $(f_n)_{n\ge1}$ be simple functions such that $0 \le f_n \uparrow f$. Then by Lemma 5.3(i),

$$\int_E f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_E f_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int_S \mathbf{1}_E f_n \,\mathrm{d}\mu = \int_S \mathbf{1}_E f \,\mathrm{d}\mu.$$

(ii)-(v): See Exercise 5.2.

(vi): Assume $f = 0$ a.e. Choose a sequence of simple functions $(f_n)_{n\ge1}$ such that $0 \le f_n \uparrow f$. Then $f_n = 0$ a.e. for each $n \in \mathbb{N}$ and thus by Lemma 5.3 (vi),

$$\int_S f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_S f_n \,\mathrm{d}\mu = 0.$$

For the converse we use contraposition. Assume one does not have $f = 0$ a.e., which means that $E = \{s \in S : f(s) > 0\}$ satisfies $\mu(E) > 0$. Letting $E_n = \{s \in S : f(s) \ge \frac{1}{n}\}$ we find $E_n \uparrow E$, so by Theorem 2.9, $\mu(E_n) \to \mu(E)$. Therefore, there exists an $n \in \mathbb{N}$ such that $\mu(E_n) > 0$ and thus

$$\int_S f \,\mathrm{d}\mu \overset{\text{(iii)}}{\ge} \int_{E_n} f \,\mathrm{d}\mu \overset{\text{(ii)}}{\ge} \int_{E_n} \tfrac{1}{n} \,\mathrm{d}\mu = \tfrac{1}{n} \cdot \mu(E_n) > 0.$$ ☺

Example 5.8 (Series are integrals). Let $S = \mathbb{N}$ and $\mathcal{A} = \mathcal{P}(\mathbb{N})$. Let $\tau : \mathcal{P}(\mathbb{N}) \to [0,\infty]$ denote the counting measure. Let $f : \mathbb{N} \to [0,\infty)$ be arbitrary. Then f is clearly measurable. Now define for each $n \ge 1$, $f_n = \sum_{k=1}^n f(k)\mathbf{1}_{\{k\}}$. Since each f_n is a simple function and $0 \le f_n \uparrow f$ we find

$$\int_{\mathbb{N}} f \,\mathrm{d}\tau = \lim_{n\to\infty} \int_{\mathbb{N}} f_n \,\mathrm{d}\tau = \lim_{n\to\infty} \sum_{k=1}^n f(k) = \sum_{k=1}^\infty f(k).$$

5.3. **Lebesgue integral for measurable functions.** The final step is to define the Lebesgue integral for measurable functions $f : S \to \overline{\mathbb{R}}$ by using the splitting $f = f^+ - f^-$. Recall that $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$.

Definition 5.9. *A measurable function $f : S \to \overline{\mathbb{R}}$ is called **integrable** when both $\int_S f^+ \,\mathrm{d}\mu$ and $\int_S f^- \,\mathrm{d}\mu$ are finite. In this case, the **Lebesgue integral** of f over $E \in \mathcal{A}$ is defined as*

$$\int_E f \,\mathrm{d}\mu := \int_E f^+ \,\mathrm{d}\mu - \int_E f^- \,\mathrm{d}\mu.$$

Remark 5.10.

(1) If $f : S \to \overline{\mathbb{R}}$ is integrable, then by monotonicity $\int_E f^+ \,\mathrm{d}\mu$ and $\int_E f^- \,\mathrm{d}\mu$ are finite for every $E \in \mathcal{A}$. Therefore, $\int_E f \,\mathrm{d}\mu$ is a number in $\mathbb{R}$.
(2) A measurable function $f : S \to [0,\infty]$ is integrable if and only if $\int_S f \,\mathrm{d}\mu < \infty$.
(3) If $f : S \to \overline{\mathbb{R}}$ is integrable, then f is finite a.e. (see Exercise 5.1).

To check integrability, the following characterization is very useful.

Proposition 5.11. *For a measurable function $f : S \to \overline{\mathbb{R}}$ the following are equivalent:*

(i) f is integrable.
(ii) $|f|$ is integrable.

Moreover, in this case for each $E \in \mathcal{A}$,

$$(\textit{triangle inequality}) \qquad \Big|\int_E f \,\mathrm{d}\mu\Big| \le \int_E |f| \,\mathrm{d}\mu.$$

Proof. First observe that $|f| = f^+ + f^-$. Therefore, both assertions are equivalent to $I^+, I^- < \infty$, where $I^\pm = \int_S f^\pm \,\mathrm{d}\mu$ and hence the result follows. To prove the required estimate note that by the triangle inequality and linearity of the Lebesgue integral for positive functions:

$$\Big|\int_E f\,\mathrm{d}\mu\Big| = |I^+ - I^-| \le I^+ + I^- = \int_E f^+ + f^- \,\mathrm{d}\mu = \int_E |f|\,\mathrm{d}\mu. \qquad ☺$$

By considering the positive and negative part separately one can check Proposition 5.7(i), (ii) and (v) hold for all integrable functions $f, g : S \to \overline{\mathbb{R}}$ again (see Exercise 5.3). The following example shows that (iii) and (vi) do not extend to this setting.

Example 5.12. Let $S = \mathbb{R}$ and λ the Lebesgue measure. Let $f = \mathbf{1}_{(0,1]} - \mathbf{1}_{(1,2]}$. Then $f^+ = \mathbf{1}_{(0,1]}$ and $f^- = \mathbf{1}_{(1,2]}$ and

$$\int_{(0,1]} f\,\mathrm{d}\lambda = \int_{(0,1]} f^+\,\mathrm{d}\lambda = \lambda((0,1]) = 1,$$

$$\int_{(0,2]} f\,\mathrm{d}\lambda = \int_{(0,2]} f^+\,\mathrm{d}\lambda - \int_{(0,2]} f^-\,\mathrm{d}\lambda = \lambda((0,1]) - \lambda((1,2]) = 1 - 1 = 0.$$

The extension of the linearity in Proposition 5.7(iv) is more difficult and proved below.

Proposition 5.13. *Let $f, g : S \to \overline{\mathbb{R}}$ be integrable functions and $\alpha, \beta \in \mathbb{R}$. Then*[28] *$\alpha f + \beta g$ is integrable, and for all $E \in \mathcal{A}$*

$$\int_E \alpha f + \beta g\,\mathrm{d}\mu = \alpha\int_E f\,\mathrm{d}\mu + \beta\int_E g\,\mathrm{d}\mu.$$

Proof. Note that, by Proposition 5.11, each of the following functions is integrable $f^\pm$, $g^\pm$, $|f|, |g|$. Therefore, $|\alpha|\,|f| + |\beta|\,|g|$ is integrable by Proposition 5.7(iv). Since $|\alpha f + \beta g| \le |\alpha|\,|f| + |\beta|\,|g|$, the function $\alpha f + \beta g$ is integrable as well (see Exercise 5.6(a)).

Next, we prove the identity for each $E \in \mathcal{A}$. To do this, we first consider the case $\alpha = \beta = 1$. Write $h := f + g = \phi - \psi$, where $\phi := f^+ + g^+$ and $\psi := (f^- + g^-)$. Then, by Proposition 5.7(iv), ϕ and ψ are integrable. Moreover, since $h^+ + \psi = (f+g)^- + \phi$, we have

$$\begin{aligned}\int_E h^+\,\mathrm{d}\mu &= \int_E h^+ + \psi\,\mathrm{d}\mu - \int_E \psi\,\mathrm{d}\mu\\ &= \int_E h^- + \phi\,\mathrm{d}\mu - \int_E \psi\,\mathrm{d}\mu\\ &= \int_E h^-\,\mathrm{d}\mu + \int_E f^+\,\mathrm{d}\mu + \int_E g^+\,\mathrm{d}\mu - \int_E f^-\,\mathrm{d}\mu - \int_E g^-\,\mathrm{d}\mu.\end{aligned}$$

Rearranging the terms yields

$$\int_E f + g\,\mathrm{d}\mu = \int_E h\,\mathrm{d}\mu = \int_E f\,\mathrm{d}\mu + \int_E g\,\mathrm{d}\mu.$$

It remains to show that for all $\alpha \in \mathbb{R}$,

$$\int_E \alpha f\,\mathrm{d}\mu = \alpha\int_E f\,\mathrm{d}\mu. \tag{5.6}$$

Note that $\alpha f = (\alpha f)^+ - (\alpha f)^- = \alpha f^+ - \alpha f^-$. Thus, Proposition 5.7(iv) yields

$$\int_E \alpha f\,\mathrm{d}\mu = \int_E \alpha f^+\,\mathrm{d}\mu - \int_E \alpha f^-\,\mathrm{d}\mu = \alpha\int_E f^+\,\mathrm{d}\mu - \alpha\int_E f^-\,\mathrm{d}\mu,$$

which proves (5.6). ☺

We give a few examples of integrable functions.

[28] Of course, we only consider those functions for which $\alpha f + \beta g$ is well-defined.

Example 5.14. Every simple function $f : S \to \mathbb{R}$ given by $f = \sum_{k=1}^{n} x_k \mathbf{1}_{A_k}$ with $x_1, \ldots, x_n \in \mathbb{R}$, $A_1, \ldots, A_n \in \mathcal{A}$ and $\mu(A_k) < \infty$ for $k = 1, \ldots, n$, is integrable and by Proposition 5.13,

$$\int_E f \, \mathrm{d}\mu = \sum_{k=1}^{n} x_k \int_E \mathbf{1}_{A_k} \, \mathrm{d}\mu = \sum_{k=1}^{n} x_k \cdot \mu(E \cap A_k).$$

Example 5.15. Let $S = \mathbb{R}$, $\mathcal{A} = \mathcal{P}(\mathbb{R})$. Let $x \in \mathbb{R}$ and let δ_x be the Dirac measure from Example 2.6. Then any $f : \mathbb{R} \to \mathbb{R}$ is integrable and

$$\int_{\mathbb{R}} f \, \mathrm{d}\delta_x = f(x),$$

see Exercise 5.7.

Example 5.16. Let $S = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ and let $\tau\colon \mathcal{P}(\mathbb{N}) \to [0, \infty]$ denote the counting measure. A function $f : \mathbb{N} \to \mathbb{R}$ is integrable if and only if $\sum_{n=1}^{\infty} |f(n)| < \infty$. In that case

$$\int_{\mathbb{N}} f \, \mathrm{d}\tau = \sum_{n=1}^{\infty} f(n),$$

see Exercise 5.8.

In the next theorem, we will show that, for continuous functions, the Lebesgue integral with respect to the Lebesgue measure coincides with the Riemann integral. It is even known that every Riemann integrable function f is integrable with respect to the Lebesgue measure and the integrals coincide (see [1, Theorem 23.6]). The proof uses a similar method. However, one should be aware that this breaks down for *improper* Riemann integrals. See Theorem 6.11 and Exercise 6.10 for more details on improper Riemann integrals.

Theorem 5.17 (Lebegue and Riemann integrals coincide)**.** *Let $a, b \in \mathbb{R}$ with $a < b$. Let λ be the restriction of the Lebesgue measure to $\mathcal{B}([a, b])$ and let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then*

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

Proof. Note that f is measurable by Proposition 4.5 and, since there is a constant $M \geq 0$ such that $|f| \leq M$, f is integrable by Exercise 5.6. Moreover, since f is continuous, f is Riemann integrable.

To prove the identity, let $\varepsilon > 0$. Since f is uniformly continuous we can find a $\delta > 0$ such that for all $x, y \in [a, b]$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$. Let $n \in \mathbb{N}$ be such that $\frac{b-a}{n} < \delta$ and let $x_k = a + k\frac{b-a}{n}$ for $k = 0, \ldots, n$. Set

$$m_k := \min\{f(x) : x \in [x_{k-1}, x_k]\},$$
$$M_k := \max\{f(x) : x \in [x_{k-1}, x_k]\},$$

and let

$$g(x) = \mathbf{1}_{\{a\}} f(a) + \sum_{k=1}^{n} \mathbf{1}_{(x_{k-1}, x_k]} m_k, \qquad x \in [a, b],$$
$$G(x) = \mathbf{1}_{\{a\}} f(a) + \sum_{k=1}^{n} \mathbf{1}_{(x_{k-1}, x_k]} M_k, \qquad x \in [a, b].$$

Then we have

$$\int_{[a,b]} g \, \mathrm{d}\lambda = \int_a^b g(x) \, \mathrm{d}x = \sum_{k=1}^{n} (x_k - x_{k-1}) m_k =: \alpha,$$
$$\int_{[a,b]} G \, \mathrm{d}\lambda = \int_a^b G(x) \, \mathrm{d}x = \sum_{k=1}^{n} (x_k - x_{k-1}) M_k =: \beta.$$

Since $g \leq f \leq G$, by monotonicity of both the Lebesgue and Riemann integral we find that

$$\Big| \int_{[a,b]} f \,\mathrm{d}\lambda - \int_a^b f(x)\,\mathrm{d}x \Big| \leq \beta - \alpha = \sum_{k=1}^n (x_k - x_{k-1})(M_k - m_k) \leq \sum_{k=1}^n \frac{b-a}{n}\frac{\varepsilon}{b-a} = \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the claimed identity follows. ☺

To conclude this section, we briefly indicate how one can extend the Lebesgue integral to complex functions $f\colon S \to \mathbb{C}$. On the complex numbers $\mathbb{C}$ we consider its Borel σ-algebra. If $f : S \to \mathbb{C}$ is measurable, we write $f = u + iv$ with $u, v : S \to \mathbb{R}$. Then u and v are measurable (see Exercise 5.11).

Definition 5.18. *A measurable function $f : S \to \mathbb{C}$ given by $f = u + iv$ with $u, v : S \to \mathbb{R}$ is called **integrable** if both u and v are integrable. In this case, the **Lebesgue integral** of f over $E \in \mathcal{A}$ is defined as*

$$\int_E f \,\mathrm{d}\mu = \int_E u \,\mathrm{d}\mu + i \int_E v \,\mathrm{d}\mu.$$

Exercise 5.11 yields that Propositions 5.11 and 5.13 extend to the complex setting.

Exercises

Exercise 5.1. Let $f : S \to \overline{\mathbb{R}}$ be an integrable function. Show that $\mu(\{s \in S : f(s) = \pm\infty\}) = 0$.

Exercise 5.2.

(a) Prove Proposition 5.7(ii).
Hint: Approximate by simple functions as in the definition of the Lebesgue integral. Make sure that the approximating sequence for f is less than or equal to the approximating sequence for g.
(b) Prove Proposition 5.7(iii).
(c) Prove Proposition 5.7(iv).
(d) Prove Proposition 5.7(v).

Exercise 5.3. Extend the following results to integrable functions $f, g : S \to \overline{\mathbb{R}}$:

(a) Proposition 5.7(i).
Hint: First show that $\mathbf{1}_E f$ is integrable for $E \in \mathcal{A}$.
(b) Proposition 5.7(ii).
(c) Proposition 5.7(v).

Exercise* 5.4. Let $f, g : S \to \overline{\mathbb{R}}$ both be measurable functions. Assume f is integrable and $f = g$ a.e. Show that g is integrable and

$$\int_E f \,\mathrm{d}\mu = \int_E g \,\mathrm{d}\mu, \quad E \in \mathcal{A}.$$

Hint: Define $A^{\pm} = \{s \in E : f^{+}(s) = g^{+}(s)\}$, split $\mathbf{1}_E = \mathbf{1}_{A^{\pm}} + \mathbf{1}_{E\setminus A^{\pm}}$ and use Proposition 5.7(vi).

Exercise 5.5. Let λ denote the Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Let $f : \mathbb{R} \to \mathbb{R}$ be an integrable function. Show that for all $-\infty < a < b < \infty$

$$\int_{[a,b]} f \,\mathrm{d}\lambda = \int_{(a,b)} f \,\mathrm{d}\lambda.$$

Does this also hold if we replace λ by an arbitrary measure μ on $\mathcal{B}(\mathbb{R})$?

Exercise* 5.6 (Domination). Let $f, g : S \to \mathbb{R}$ be measurable functions.

(a) Assume $|f| \leq g$ and g is integrable. Prove that f is integrable.

(b) Show that f is integrable if and only if

$$\sum_{n=-\infty}^{\infty} 2^n \mu(\{s \in S : 2^n < |f(s)| \le 2^{n+1}\}) < \infty.$$

.

Hint: Use (a) for a suitable function $g : S \to [0, \infty]$.

In the next exercises, you are asked to give the details of Examples 5.15 and 5.16.

Exercise 5.7. Let $S = \mathbb{R}$, $\mathcal{A} = \mathcal{P}(\mathbb{R})$. Let $x \in \mathbb{R}$ and let δ_x be the Dirac measure from Example 2.6.

(a) For a simple function $f \colon \mathbb{R} \to [0, \infty)$, show that $\int_{\mathbb{R}} f \, \mathrm{d}\delta_x = f(x)$.
(b) For a function $f \colon \mathbb{R} \to [0, \infty]$, show that f is measurable and $\int_{\mathbb{R}} f \, \mathrm{d}\delta_x = f(x)$.
(c) For a function $f \colon \mathbb{R} \to \mathbb{R}$, show that f is integrable and $\int_{\mathbb{R}} f \, \mathrm{d}\delta_x = f(x)$.

Exercise* 5.8. Let $S = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ and let $\tau \colon \mathcal{P}(\mathbb{N}) \to [0, \infty]$ denote the counting measure.

(a) Show that a function $f : \mathbb{N} \to \mathbb{R}$ is integrable if and only if $\sum_{n=1}^{\infty} |f(n)| < \infty$.
(b) Assume $\sum_{n=1}^{\infty} |f(n)| < \infty$. Show that

$$\int_{\mathbb{N}} f \, \mathrm{d}\tau = \sum_{n=1}^{\infty} f(n).$$

Exercise* 5.9 (Chebyshev's inequality). Let $f : S \to [0, \infty]$ be a measurable function. Prove that for every $t > 0$,

$$\mu(\{s \in S : f(s) \ge t\}) \le \frac{1}{t} \int_S f \, \mathrm{d}\mu.$$

Hint: Let $A_t = \{s \in S : f(s) \ge t\}$ and write $\mu(A_t) = \int_S \mathbf{1}_{A_t} \, \mathrm{d}\mu$.

Exercise* 5.10. Let λ be the Lebesgue measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. For $h \in \mathbb{R}^d$ define the translation $f_h : \mathbb{R}^d \to \mathbb{R}$ by $f_h(x) = f(x - h)$.

(a) Show that f_h is measurable.
(b) Assume that f is integrable. Show that f_h is integrable and

$$\int_{\mathbb{R}^d} f_h \, \mathrm{d}\lambda = \int_{\mathbb{R}^d} f \, \mathrm{d}\lambda.$$

Exercise** 5.11 (Complex-valued functions).

(a) Let $f : S \to \mathbb{C}$ be a measurable function and write $f = u + iv$ with $u, v : S \to \mathbb{R}$. Show that u and v are measurable.
(b) Prove Proposition 5.13 for integrable functions $f, g : S \to \mathbb{C}$ and $\alpha, \beta \in \mathbb{C}$.
(c) Prove Proposition 5.11 for integrable functions $f : S \to \mathbb{C}$.
(d) Which parts of Proposition 5.7 remain true for integrable functions $f, g : S \to \mathbb{C}$?

6. Convergence theorems and applications

One of the problems with the Riemann integral is that the cases where limit and integral can be interchanged are rather limited. Indeed, we have seen that we can interchange limit and the Riemann integral if we have uniform continuity, which is a quite strong assumption. In modern mathematics it is crucial to have better "tools" for this. We use the integration theory developed so far in order to obtain these tools. In Section 6.1 we prove three famous convergence results. In Section 6.2 we give several consequences and applications. Throughout this section, $(S, \mathcal{A}, \mu)$ is a measure space.

We start with a simple example, which illustrates some difficulties.

Example 6.1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by one of the following $n\mathbf{1}_{(0,\frac{1}{n}]}$, $\frac{1}{n}\mathbf{1}_{(n,2n)}$ or $\mathbf{1}_{(n,\infty)}$ for $n \geq 1$. Then $f_n \to 0$, but in each case $\int_{\mathbb{R}} f_n d\lambda \not\to 0$.

Example 6.1 shows that, even with the Lebesgue integral in our arsenal, we should be careful when interchanging limits and integrals.

6.1. **The three main convergence results.**

In this section we prove the three most famous convergence results of integration theory.

- Monotone Convergence Theorem (MCT).[29]
- Fatou's lemma.[30]
- Dominated Convergence Theorem (DCT).[31]

We start with the monotone convergence theorem, which is a statement about a an increasing sequence of positive-valued measurable functions.

Theorem 6.2 (Monotone Convergence Theorem (MCT))**.** *Let $f_n\colon S \to [0,\infty]$ be measurable functions for $n \geq 1$ such that $0 \leq f_n \uparrow f$. Then f is measurable and*

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

If $(f_n)_{n\geq 1}$ would be a sequence of *simple* functions, the monotone convergence theorem is just the definition of the Lebesgue integral of f. To prove the statement for a sequence of measurable functions, we will of course approximate the measurable functions by simple functions.

Proof of Theorem 6.2. By Theorem 4.11, $f\colon S \to [0,\infty]$ is measurable. For each fixed $n \geq 1$ choose a sequence of simple functions $0 \leq f_{n,m} \uparrow f_n$. For $n, m \geq 1$ set

$$g_{n,m} := \max\{f_{1,m}, \dots, f_{n,m}\}$$

and note that

$$g_{n,n} \leq \max\{f_1, \dots, f_n\} = f_n.$$

We claim that $g_{n,n}$ is a simple function for all $n \geq 1$ and $0 \leq g_{n,n} \uparrow f$. Assuming this claim for the moment, it follows from Proposition 5.7(ii) that

$$\int_S f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_S g_{n,n} \,\mathrm{d}\mu \leq \lim_{n\to\infty} \int_S f_n \,\mathrm{d}\mu \leq \int_S f \,\mathrm{d}\mu,$$

which proves the theorem.

To prove the claim, note that the maximum of a finite number of simple functions is simple. Moreover, since $f_{n,m} \leq f_{n,m+1}$ for all $n, m \geq 1$, we have

$$g_{n,n} = \max\{f_{1,n}, \dots, f_{n,n}\} \leq \max\{f_{1,n+1}, \dots, f_{n,n+1}, f_{n+1,n+1}\} = g_{n+1,n+1},$$

so $(g_{n,n})_{n\geq 1}$ is increasing. Furthermore, we have

$$f_n = \lim_{m\to\infty} g_{n,m} = \lim_{m\to\infty} g_{n,m+n} \leq \lim_{m\to\infty} g_{m+n,m+n} = \lim_{m\to\infty} g_{m,m} \leq f,$$

so since $f_n \uparrow f$, we obtain that $g_{n,n} \to f$. This finishes the proof of the claim. ☺

[29]This result is due to Beppo Levi 1875–1961, who was an Italian mathematician.

[30]This result is named after the French mathematician Pierre Fatou 1878–1929.

[31]This is due to Lebesgue (see footnote 13).

We continue with a statement for a, not necessarily increasing, sequence of positive-valued measurable functions. In this case, we can only prove an inequality.

Lemma 6.3 (Fatou's Lemma)**.** *Let $f_n\colon S \to [0,\infty]$ be measurable functions for $n \geq 1$. Then*

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

Proof. Note that $f = \liminf_{n\to\infty} f_n$ is a measurable function by Theorem 4.11. Fix $n \in \mathbb{N}$ and let $g_n = \inf_{k\geq n} f_k$. For all $m \geq n$, we have $g_n \leq f_m$ and by the monotonicity of the Lebesgue integral this gives $\int_S g_n \,\mathrm{d}\mu \leq \int_S f_m \,\mathrm{d}\mu$. Therefore, taking the infimum over all $m \geq n$, we find

$$\int_S g_n \,\mathrm{d}\mu \leq \inf_{m\geq n} \int_S f_m \,\mathrm{d}\mu. \tag{6.1}$$

Note that $0 \leq g_n \uparrow f$ and thus

$$\int_S f \,\mathrm{d}\mu = \int_S \lim_{n\to\infty} g_n \,\mathrm{d}\mu \overset{\text{(MCT)}}{=} \lim_{n\to\infty} \int_S g_n \,\mathrm{d}\mu \overset{(6.1)}{\leq} \lim_{n\to\infty} \inf_{m\geq n} \int_S f_m \,\mathrm{d}\mu = \liminf_{n\to\infty} \int_S f_n \,\mathrm{d}\mu.$$

☺

Our final convergence theorem is for real-valued functions.

Theorem 6.4 (Dominated Convergence Theorem (DCT))**.** *Let $f_n\colon S \to \mathbb{R}$ be measurable functions for $n \geq 1$ such that $f_n \to f$ pointwise. If there exists an integrable function $g : S \to [0,\infty)$ such that $|f_n| \leq g$ for all $n \geq 1$, then f_n and f are integrable and*

$$\lim_{n\to\infty} \int_S f_n \, d\mu = \int_S f \, d\mu. \tag{6.2}$$

Proof. Since $|f_n| \leq g$, we see that f_n is integrable for each $n \geq 1$. Moreover, $f\colon S \to \mathbb{R}$ is measurable by Theorem 4.11 and integrable since $|f| \leq g$. Let $x_n := \int_S f_n \,\mathrm{d}\mu$ and $x := \int_S f \,\mathrm{d}\mu$. It suffices to show that $\limsup_{n\to\infty} x_n \leq x \leq \liminf_{n\to\infty} x_n$. It follows that

$$\begin{aligned}
\int_S g \,\mathrm{d}\mu \pm x = \int_S g \pm f \,\mathrm{d}\mu &= \int_S \lim_{n\to\infty} (g \pm f_n) \,\mathrm{d}\mu && \text{(linearity)} \\
&\leq \liminf_{n\to\infty} \int_S g \pm f_n \,\mathrm{d}\mu && \text{(Fatou's lemma with } g \pm f_n \geq 0) \\
&= \liminf_{n\to\infty} \Big(\int_S g \,\mathrm{d}\mu \pm x_n \Big) && \text{(linearity).} \\
&= \int_S g \,\mathrm{d}\mu + \liminf_{n\to\infty} (\pm x_n).
\end{aligned}$$

Since $\int_S g \,\mathrm{d}\mu < \infty$, we find that $\pm x \leq \liminf_{n\to\infty}(\pm x_n)$. This implies the estimates

$$x \leq \liminf_{n\to\infty} x_n \quad \text{and} \quad \limsup_{n\to\infty} x_n \leq x,$$

where for the second part we used $\liminf_{n\to\infty}(-x_n) = -\limsup_{n\to\infty} x_n$. ☺

For functions $f, f_1, f_2 : S \to \overline{\mathbb{R}}$ we say that $f_n \to f$ a.e. if there exists a set $A \in \mathcal{A}$ such that $\mu(A) = 0$ and $f_n(s) \to f(s)$ for all $s \in S \setminus A$. For the details of the following remark we refer to Exercise 6.5.

Remark 6.5. Let $f, g, f_1, f_2, \ldots : S \to \overline{\mathbb{R}}$ be measurable functions.

(a) Theorem 6.2 also holds under the weaker assumption that $0 \leq f_n \uparrow f$ a.e.
(b) Theorem 6.4 also holds under the weaker assumptions that $f_n \to f$ a.e. and $|f_n| \leq g$ a.e. for all $n \geq 1$.

Example 6.6. Let $f : \mathbb{R} \to \mathbb{R}$ be integrable with respect to the Lebesgue measure and let $F : \mathbb{R} \to \mathbb{R}$ be given by $F(x) = \int_{(-\infty,x]} f \,\mathrm{d}\lambda$. Then F is continuous on $\mathbb{R}$. Indeed, if $x_n \to x$, then $\mathbf{1}_{(-\infty,x_n)}(y) \to \mathbf{1}_{(-\infty,x]}(y)$ for all $y \in \mathbb{R} \setminus \{x\}$ and thus by Remark 6.5

$$F(x_n) = \int_{\mathbb{R}} \mathbf{1}_{(-\infty,x_n]} f \,\mathrm{d}\lambda \overset{\text{(DCT)}}{\longrightarrow} \int_{\mathbb{R}} \mathbf{1}_{(-\infty,x]} f \,\mathrm{d}\lambda = F(x).$$

One can show that F is actually differentiable a.e and $F' = f$ a.e., which is a Lebesgue integral version of the Fundamental theorem of calculus. The proof requires some knowledge of Harmonic Analysis, a master elective course.

6.2. **Consequences and applications.** We continue with several consequences and applications of the three main convergence theorems. We start with a result on integration of series of positive functions. The proof is Exercise 6.4.

Corollary 6.7 (Series and integrals)**.** *Let $f_1, f_2, \ldots : S \to [0,\infty]$ be measurable functions. Then*

$$\int_S \sum_{n=1}^{\infty} f_n \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_S f_n \,\mathrm{d}\mu. \tag{6.3}$$

Next, we note that from a measure μ one can build many other measures in the following way:

Theorem 6.8 (Density)**.** *Let $f : S \to [0,\infty]$ be a measurable function. Define $\nu : \mathcal{A} \to [0,\infty]$ by*

$$\nu(A) = \int_A f \,\mathrm{d}\mu.$$

Then ν is a measure.[32] *Moreover, $g : S \to \overline{\mathbb{R}}$ is integrable with respect to ν if and only if fg is integrable with respect to μ. In this case*

$$\int_S g \,\mathrm{d}\nu = \int_S fg \,\mathrm{d}\mu. \tag{6.4}$$

Proof. *Step 1:* Clearly, $\nu(\varnothing) = 0$. For $(A_n)_{n\geq 1}$ a disjoint sequence in $\mathcal{A}$, and $A := \bigcup_{n=1}^{\infty} A_n$,

$$\nu(A) = \int_S \mathbf{1}_A f \,\mathrm{d}\mu = \int_S \sum_{n=1}^{\infty} \mathbf{1}_{A_n} f \,\mathrm{d}\mu \overset{(6.3)}{=} \sum_{n=1}^{\infty} \int_S \mathbf{1}_{A_n} f \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \nu(A_n).$$

Step 2: First we show that (6.4) holds for all measurable $g : S \to [0,\infty]$. For $g = \mathbf{1}_A$ this is immediate from $\int_S \mathbf{1}_A \,\mathrm{d}\nu = \nu(A) = \int_S \mathbf{1}_A f \,\mathrm{d}\mu$. For simple functions $g : S \to [0,\infty)$ this follows by linearity. For a measurable function $g : S \to [0,\infty]$, by Theorem 4.14 we can find a sequence of simple functions $(g_n)_{n\geq 1}$ such that $0 \leq g_n \uparrow g$. By the previous case, we obtain

$$\int_S g \,\mathrm{d}\nu = \lim_{n\to\infty} \int_S g_n \,\mathrm{d}\nu = \lim_{n\to\infty} \int_S f g_n \,\mathrm{d}\mu \overset{\text{(MCT)}}{=} \int_S fg \,\mathrm{d}\mu.$$

Step 3: To prove the "if and only if" assertion and (6.4), let $g : S \to \overline{\mathbb{R}}$ be a measurable function. Since step 2 yields that

$$\int_S g^{\pm} \,\mathrm{d}\nu = \int_S f g^{\pm} \,\mathrm{d}\mu,$$

both the equivalence and (6.4) follows by writing $g = g^+ - g^-$. ☺

Example 6.9. Let $f : \mathbb{R} \to [0,\infty]$ be given by $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, which is continuous and therefore measurable by Proposition 4.5. Define $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty)$ by

$$\nu(A) = \int_A f \,\mathrm{d}\lambda, \quad A \in \mathcal{B}(\mathbb{R}),$$

which is the standard Gaussian measure on $\mathbb{R}$.

[32] The function f is usually called the density of ν and plays an important role in probability theory.

To show that $\mathbb{R}$ with the Gaussian measure from Example 6.9 is a probability space, we need to prove that $\nu(\mathbb{R}) = 1$, i.e. we need to calculate $\int_{\mathbb{R}} f \,\mathrm{d}\lambda$. To do so, we would like to relate this Lebesgue integral to the improper Riemann integral

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \,\mathrm{d}x = 1.$$

However, Theorem 5.17 only relates Riemann and Lebesgue integrals on closed and bounded intervals $[a, b]$, for $a < b$. Using the monotone convergence theorem, we will extend this to theorem to improper Riemann integrals. We start with the definition of an improper Riemann integral.

Definition 6.10. *Let $-\infty \leq a < b \leq \infty$ and let $f\colon (a, b) \to \mathbb{R}$ be a continuous function. We say that $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$ exists as an **improper** Riemann integral if the limits $\lim_{t\uparrow a} \int_c^t f(x)\,\mathrm{d}x$ and $\lim_{t\downarrow b} \int_t^c f(x)\,\mathrm{d}x$ exist in $\mathbb{R}$ for some (equivalently all) $c \in (a, b)$. In that case define*

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{t\uparrow a} \int_c^t f(x) + \lim_{t\downarrow b} \int_t^c f(x)\,\mathrm{d}x.$$

Note that Definition 6.10 includes both the improper Riemann integral of continuous functions $f\colon \mathbb{R} \to \mathbb{R}$ and the improper Riemann integral of continuous functions $f\colon (a, b) \to \mathbb{R}$ with $a < b$ for which $\lim_{t\downarrow a} f(t)$ or $\lim_{t\uparrow b} f(t)$ diverges to $\pm\infty$.

Theorem 6.11 (Improper Lebesgue and Riemann integrals coincide)**.** *Let $-\infty \leq a < b \leq \infty$ and let $f\colon (a, b) \to \mathbb{R}$ be a continuous function. The following are equivalent:*

(i) f is integrable;

(ii) $\int_a^b |f(x)|\,\mathrm{d}x$ exists as an improper Riemann integral.

Moreover, in this case $\int_a^b f(x)\,\mathrm{d}x$ exists as an improper Riemann integral and

$$\int_{(a,b)} f \,\mathrm{d}\lambda = \int_a^b f(x)\,\mathrm{d}x. \tag{6.5}$$

It may happen that $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$ exists as an improper Riemann integral without f being integrable, see Exercise 6.10.

Proof. (i)⇒(ii): First we make a general observation. Let $g\colon (a, b) \to \mathbb{R}$ be a continuous and integrable function and $c \in (a, b)$. In Theorem 5.17, we have seen that $\int_c^t g(x)\,\mathrm{d}x = \int_{[c,t]} g \,\mathrm{d}\lambda$ for $t \in (c, b)$. We find

$$\Big| \int_{[c,b)} g \,\mathrm{d}\lambda - \int_c^t g(x)\,\mathrm{d}x \Big| = \Big| \int_{(a,b)} \mathbf{1}_{[c,b)} g \,\mathrm{d}\lambda - \int_{(a,b)} \mathbf{1}_{[c,t]} g \,\mathrm{d}\lambda \Big| \leq \int_{(a,b)} \mathbf{1}_{(t,b)} |g| \,\mathrm{d}\lambda.$$

Now the DCT yields that $\int_{(a,b)} \mathbf{1}_{(t,b)} |g| \,\mathrm{d}\lambda \to 0$ as $t \to b$, and we may conclude that $\int_c^b g(x)\,\mathrm{d}x$ exists. The part on $(a, c]$ goes similarly. Now (i)⇒(ii) follows by letting $g = |f|$. Moreover, (6.5) follows by taking $g = f$.

(ii)⇒(i): Let $(a_n)_{n\geq 1}$ and $(b_n)_{n\geq 1}$ be sequences in (a, b) with $a_n \downarrow a$ and $b_n \uparrow b$. By Theorem 5.17 we have for $n \geq 1$

$$\int_{(a,b)} \mathbf{1}_{[a_n,b_n]} |f| \,\mathrm{d}\lambda = \int_{a_n}^{b_n} |f(x)|\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x =: M < \infty.$$

Therefore, by the MCT, $\int_{(a,b)} |f| \,\mathrm{d}\lambda = \lim_{n\to\infty} \int_{(a,b)} \mathbf{1}_{[a_n,b_n]} |f| \,\mathrm{d}\lambda \leq M$, so f is integrable. ☺

Example 6.12. Let $f_n : [0, \infty) \to \mathbb{R}$ be defined by $f_n(x) = \left(1 + \frac{x}{n}\right)^n e^{-2x}$. Below we show that $\lim_{n\to\infty} \int_0^n f_n(x)\,\mathrm{d}x = 1$. Recall the standard limit $0 \leq \left(1 + \frac{x}{n}\right)^n \uparrow e^x$ for every $x \in [0, \infty)$ and thus $0 \leq \mathbf{1}_{[0,n]} f_n \uparrow f$, where $f(x) = e^{-x}$. Therefore, by Theorem 5.17 and Theorem 6.11

$$\int_0^n f_n(x)\,\mathrm{d}x = \int_{[0,\infty)} \mathbf{1}_{[0,n]} f_n \,\mathrm{d}\lambda \overset{\text{(MCT)}}{\longrightarrow} \int_{[0,\infty)} f \,\mathrm{d}\lambda \overset{(6.5)}{=} \int_0^\infty f(x)\,\mathrm{d}x = \lim_{t\to\infty} (1 - e^{-t}) = 1.$$

We end this section with a standard application of the DCT to calculus.

Theorem 6.13 (Differentiating under the integral sign)**.** *Suppose $f : \mathbb{R} \times S \to \mathbb{R}$ is such that the following hold:*

(i) f is continuous and differentiable with respect to its first coordinate and for each $y_0 \in \mathbb{R}$ there exists a $\delta > 0$ and integrable function $g : S \to [0, \infty)$ such that

$$|\tfrac{\partial f}{\partial y}(y, s)| \leq g(s), \qquad y \in (y_0 - \delta, y_0 + \delta),\ s \in S.$$

(ii) $s \mapsto f(y, s)$ is integrable with respect to μ.

Then for all $y \in \mathbb{R}$,

$$\frac{d}{dy} \int_S f(y, s)\, \mathrm{d}\mu(s) = \int_S \frac{\partial f}{\partial y}(y, s)\, \mathrm{d}\mu(s).$$

Proof. Fix $y_0 \in \mathbb{R}$ and let δ and g be as in (i). Let $h_n \in (0, \delta)$ for $n \geq 1$ be such that $h_n \to 0$. Let $\phi_n \colon (y_0 - \delta, y_0 + \delta) \times S \to \mathbb{R}$ and $F \colon (y_0 - \delta, y_0 + \delta) \to \mathbb{R}$ be given by

$$\phi_n(y, s) = \frac{f(y + h_n, s) - f(y, s)}{h_n} \qquad \text{and} \qquad F(y) = \int_S f(y, s)\, \mathrm{d}\mu(s).$$

Then $\phi_n(y, s) \to \frac{\partial f}{\partial y}(y, s)$ for each $y \in \mathbb{R}$ and $s \in S$. Therefore, Theorem 4.14 yields that $s \mapsto \phi_n(y, s)$ is measurable. From the mean value theorem[33] we obtain $\phi_n(y_0, s) = \frac{\partial f}{\partial y}(y_n, s)$ for some $y_n \in (y_0, y_0 + \delta)$, and hence $|\phi_n(y_0, s)| \leq g(s)$ for all $s \in S$. It follows that

$$F'(y_0) = \lim_{n\to\infty} \frac{F(y_0 + h_n) - F(y_0)}{h_n} = \lim_{n\to\infty} \int_S \phi_n(y_0, s)\, \mathrm{d}\mu(s) \overset{\text{(DCT)}}{=} \int_S \frac{\partial f}{\partial y}(y_0, s)\, \mathrm{d}\mu(s). \qquad ☺$$

Exercises

Exercise 6.1. Let $f_n \colon \mathbb{R} \to \mathbb{R}$ be functions for $n \geq 1$. Calculate $\lim_{n\to\infty} \int_{\mathbb{R}} f_n \, \mathrm{d}\lambda$ in each of the following cases:

(a) $f_n(x) = \mathbf{1}_{[2,45]}(x)\big(1 + \frac{\log(2x)}{n}\big)^n$ (use MCT; answer is 2021).

(b) $f_n(x) = \mathbf{1}_{(1,\infty)}(x)\dfrac{\sin(nx)}{nx^2}$ (use DCT; answer is 0).

(c) $f_n(x) = \mathbf{1}_{[0,1]}(x)\dfrac{nx^n \sin(nx) - n}{\sqrt{x + 2n^2}}$ (use DCT; answer is $-\frac{1}{2}\sqrt{2}$).

Exercise 6.2. Compute $\lim_{n\to\infty} \sum_{m=1}^{\infty} \frac{1}{m^2+n^2}$.
Hint: Use the counting measure and the DCT.

Exercise 6.3. Assume $f, f_1, f_2, \ldots : S \to \overline{\mathbb{R}}$ are measurable functions such that

(i) There is a constant $M \geq 0$ such that $\int_S |f_n| \, \mathrm{d}\mu \leq M$ for all $n \in \mathbb{N}$.

(ii) $f_n \to f$ pointwise.

Show that $\int_S |f| \, \mathrm{d}\mu \leq M$.

Exercise 6.4. Deduce Corollary 6.7.
Hint: Apply the MCT to $g_n = \sum_{k=1}^{n} f_k$ for $n \geq 1$.

Exercise* 6.5.

(a) Prove Remark 6.5(a).
Hint: Use Exercise 5.4.

(b) Prove Remark 6.5(b).
Hint: Use Exercise 5.1.

Exercise* 6.6. Let $f : S \to \mathbb{R}$ be measurable. Define $\nu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ by $\nu(B) = \mu(f^{-1}(B))$. Prove the following:

(a) ν is a measure.

[33] If $g : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists a point $c \in (a, b)$ such that $g'(c) = \frac{g(b)-g(a)}{b-a}$.

(b) For every measurable $g : \mathbb{R} \to [0, \infty]$ one has

$$\int_{\mathbb{R}} g(t)\, \mathrm{d}\nu(t) = \int_S g(f(s))\, \mathrm{d}\mu(s).$$

(c) A function $g : \mathbb{R} \to \overline{\mathbb{R}}$ is integrable with respect to ν if and only if $g \circ f$ is integrable with respect to μ. Moreover, in this case

$$\int_{\mathbb{R}} g(t)\, \mathrm{d}\nu(t) = \int_S g(f(s))\, \mathrm{d}\mu(s).$$

Exercise* 6.7. Assume $f : \mathbb{R} \to \mathbb{R}$ is integrable with respect to λ.

(a) Show that for each $y \in \mathbb{R}$ the function $x \mapsto \sin(xy) f(x)$ is integrable with respect to λ.

Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(y) = \int_{\mathbb{R}} \sin(xy) f(x)\, \mathrm{d}\lambda(x).$$

(b) Show that g is continuous.
Hint: Use the sequential characterization of continuity.

Exercise* 6.8 (DCT for complex-valued functions). Let $f_n : S \to \mathbb{C}$ be measurable functions for $n \geq 1$ such that $f_n \to f$ pointwise. Assume there exists an integrable function $g : S \to [0, \infty)$ such that $|f_n| \leq g$.

(a) Show that $\int_S |f_n - f|\, \mathrm{d}\mu \to 0$ as $n \to \infty$.
(b) Prove that $\int_S f_n\, \mathrm{d}\mu \to \int_S f\, \mathrm{d}\mu$.
Hint: Use Exercise 5.11(c).

Exercise** 6.9. Use induction and Theorem 6.13 to show that for each integer $n \geq 0$,

$$\int_0^\infty x^n e^{-yx}\, \mathrm{d}x = \frac{n!}{y^{n+1}}, \quad y > 0.$$

In particular, setting $y = 1$, one obtains $\int_0^\infty x^n e^{-x}\, \mathrm{d}x = n!$.

Exercise** 6.10. Let $f : [0, \infty) \to \mathbb{R}$ be given by $f(x) = \frac{\sin(x)}{x}$ if $x \neq 0$ and $f(0) = 1$.

(a) Show that f is not integrable with respect to λ.
Hint: Use that $|\sin(x)| \geq \frac{1}{2}$ on $[\pi n + \frac{1}{3}\pi, \pi n + \frac{2}{3}\pi]$ for all integers $n \geq 0$.
(b) Show that $\int_0^\infty f(x)\, \mathrm{d}x$ exists as an improper Riemann integral.[34]

[34] Using some smart tricks one can actually show that the integral equals $\frac{\pi}{2}$.

7. Product measures and iterated integrals

In a multivariate calculus course, you may have learned that for a rectangle $R := [a, b] \times [c, d] \subseteq \mathbb{R}^2$ and a continuous function $f : R \to \mathbb{R}$, we can compute the Riemann integral of f over R in different ways as an iterated integral

$$\int_R f(x, y)\, \mathrm{d}(x, y) = \int_c^d \int_a^b f(x, y)\, \mathrm{d}x\, \mathrm{d}y = \int_a^b \int_c^d f(x, y)\, \mathrm{d}y\, \mathrm{d}x. \tag{7.1}$$

In this section, we will prove an analogous result for measurable functions and the Lebesgue integral. However, we should be careful. For example, the continuous function $f : [0, 1]^2 \backslash \{(0, 0)\} \to \mathbb{R}$ given by

$$f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2} = -\frac{\partial^2}{\partial x \partial y} \arctan(y/x)$$

satisfies

$$\int_0^1 \int_0^1 f(x, y)\, \mathrm{d}x\, \mathrm{d}y = -\frac{\pi}{4},$$
$$\int_0^1 \int_0^1 f(x, y)\, \mathrm{d}y\, \mathrm{d}x = \frac{\pi}{4},$$

see Exercise 7.1.

7.1. **The product measure.** Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be measure spaces. In order to make sense of measurable and integrable functions $f : S \times T \to \mathbb{R}$, we need to define a σ-algebra and a measure on the product $S \times T$. In the specific case

$$(S, \mathcal{A}, \mu) = (T, \mathcal{B}, \nu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda),$$

our construction should yield $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$.

Our first hurdle is the fact that

$$\{A \times B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$$

is not a σ-algebra in general (see Exercise 7.4). Luckily, this is easily solved by defining the product σ-algebra

$$\mathcal{A} \times \mathcal{B} := \sigma\big(\{A \times B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}\big).$$

In the rest of this section, we will always equip $S \times T$ with this σ-algebra, yielding the measurable space $(S \times T, \mathcal{A} \times \mathcal{B})$.

To familiarize ourselves with the typical proof method that we will use throughout this section, we will prove that measurability of a function $f : S \times T \to \mathbb{R}$ directly implies that the "coordinate functions", obtained by fixing one of the variables, are measurable.

Proposition 7.1. *Let $(S, \mathcal{A})$ and $(T, \mathcal{B})$ be measurable spaces and let $f : S \times T \to \mathbb{R}$ be measurable. Then*

(i) For $t \in T$ the function $s \mapsto f(s, t)$ is measurable;
(ii) For $s \in S$ the function $t \mapsto f(s, t)$ is measurable.

Proof. The collection

$$\mathcal{C} := \big\{C \in \mathcal{A} \times \mathcal{B} : \text{(i) and (ii) hold for } f = \mathbf{1}_C\big\} \subseteq \mathcal{A} \times \mathcal{B}$$

is a σ-algebra, see Exercise 7.3. Moreover, $A \times B \in \mathcal{C}$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Since $\mathcal{A} \times \mathcal{B}$ is the smallest σ-algebra containing $A \times B$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$, it follows that $\mathcal{A} \times \mathcal{B} = \mathcal{C}$. Therefore, the proposition holds for all indicator functions $\mathbf{1}_C$ with $C \in \mathcal{A} \times \mathcal{B}$. By linearity, this extends to all simple functions $f : S \times T \to \mathbb{R}$.

For a measurable $f : S \times T \to \mathbb{R}$, we can find a sequence of simple functions $(f_n)_{n \geq 1}$ such that $f_n \to f$ pointwise by Theorem 4.14. For a fixed $t \in T$, note that $s \mapsto f_n(s, t)$ is a simple function as well, converging pointwise to $s \mapsto f(s, t)$. Thus (i) follows from Theorem 4.11 and (ii) follows similarly. ☺

Next, we want to define a product measure $\mu\times\nu$ on the measurable space $(S\times T,\mathcal{A}\times\mathcal{B})$, which should satisfy

$$(\mu\times\nu)(A\times B)=\mu(A)\cdot\nu(B),\qquad A\in\mathcal{A},\ B\in\mathcal{B}. \tag{7.2}$$

We will show that such a measure exists and is unique. Our strategy will be to apply Carathéodory's extension theorem (Theorem 3.1) on the ring $\mathcal{R}$ consisting of all finite unions of sets of the form $A\times B$ for $A\in\mathcal{A}$ and $B\in\mathcal{B}$ and $\mu\colon\mathcal{R}\to[0,\infty]$ satisfying (7.2). Moreover, to obtain uniqueness, we will use Proposition 3.5. Recall that in Proposition 3.5 we need to work with finite measures, for which we will require the following definition.

Definition 7.2. *A measure space $(S,\mathcal{A},\mu)$ is called **σ-finite** if there exists a sequence $(A_n)_{n\geq 1}$ in $\mathcal{A}$ such that $A_n\uparrow S$ and $\mu(A_n)<\infty$ for all $n\geq 1$.*

Let us give a few examples:

Example 7.3. The measure space $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\lambda)$ is σ-finite. Indeed, take $A_n=(-n,n)^d$ for $n\geq 1$.

Example 7.4. The measure space $(S,\mathcal{P}(S),\tau)$, where τ denotes the counting measure, is σ-finite if and only if S is countable, see Exercise 7.5.

We are now ready for the main result of this subsection, which asserts the existence of the (unique) product measure.

Proposition 7.5. *Let $(S,\mathcal{A},\mu)$ and $(T,\mathcal{B},\nu)$ be measure spaces. There exists a measure $\mu\times\nu\colon\mathcal{A}\times\mathcal{B}\to[0,\infty]$ such that*

$$(\mu\times\nu)(A\times B)=\mu(A)\cdot\nu(B),\qquad A\in\mathcal{A},\ B\in\mathcal{B}.$$

Moreover, if $(S,\mathcal{A},\mu)$ and $(T,\mathcal{B},\nu)$ are σ-finite, then $\mu\times\nu$ is unique and the measure space $(A\times B,\mathcal{A}\times\mathcal{B},\mu\times\nu)$ is σ-finite.

Proof. Let $\mathcal{R}\subseteq\mathcal{A}\times\mathcal{B}$ consist of all finite unions of sets of the form $A\times B$ for $A\in\mathcal{A}$ and $B\in\mathcal{B}$. Then we have $\sigma(\mathcal{R})=\mathcal{A}\times\mathcal{B}$ by definition. Moreover, by Exercise 7.4, $\mathcal{R}$ is a ring and we can write any $R\in\mathcal{R}$ as a *disjoint* union of sets of the form $A\times B$ for $A\in\mathcal{A}$ and $B\in\mathcal{B}$.

Take $R\in\mathcal{R}$ and write $R\in\mathcal{R}$ as

$$R=\bigcup_{k=1}^n(A_k\times B_k),\qquad A_1,\dots,A_n\in\mathcal{A},\ B_1,\dots,B_n\in\mathcal{B},$$

such that $(A_j\times B_j)\cap(A_k\times B_k)=\varnothing$ for $1\leq j\neq k\leq n$. Since

$$\mathbf{1}_R(s,t)=\sum_{k=1}^n\mathbf{1}_{A_k}(s)\mathbf{1}_{B_k}(t),\qquad s\in S,\ t\in T,$$

we note that $s\mapsto\mathbf{1}_R(s,t)$ is a simple function and we have for all $t\in T$

$$\int_S\mathbf{1}_R(s,t)\,\mathrm{d}\mu(s)=\sum_{k=1}^n\mu(A_k)\mathbf{1}_{B_k}(t),$$

so $t\mapsto\int_S\mathbf{1}_R(s,t)\,\mathrm{d}\mu(s)$ is a simple function as well. Therefore, we can define

$$(\mu\times\nu)(R):=\int_T\int_S\mathbf{1}_R(s,t)\,\mathrm{d}\mu(s)\,\mathrm{d}\nu(t),$$

which clearly satisfies

$$(\mu\times\nu)(A\times B)=\mu(A)\cdot\nu(B),\qquad A\in\mathcal{A},\ B\in\mathcal{B}.$$

To check σ-additivity of $\mu \times \nu$, take disjoint $R_1, R_2, \ldots \in \mathcal{R}$ with $\bigcup_{k=1}^{\infty} R_k \in \mathcal{R}$. Then we have

$$\begin{aligned}(\mu \times \nu)\Big(\bigcup_{k=1}^{\infty} R_k\Big) &= \int_T \int_S \sum_{k=1}^{\infty} \mathbf{1}_{R_k}(s,t)\, \mathrm{d}\mu(s)\, \mathrm{d}\nu(t) \\ &= \lim_{n\to\infty} \int_T \int_S \sum_{k=1}^{n} \mathbf{1}_{R_k}(s,t)\, \mathrm{d}\mu(s)\, \mathrm{d}\nu(t) \\ &= \sum_{k=1}^{\infty} \int_T \int_S \mathbf{1}_{R_k}(s,t)\, \mathrm{d}\mu(s)\, \mathrm{d}\nu(t) = \sum_{k=1}^{\infty} (\mu \times \nu)(R_k),\end{aligned}$$

so $\mu \times \nu$ is indeed σ-additive. Thus, $\mu \times \nu$ satisfies the assumptions of Theorem 3.1, so we can extend $\mu \times \nu$ to a measure $\overline{\mu \times \nu}\colon \mathcal{A} \times \mathcal{B} \to [0, \infty]$, which we once again write as $\mu \times \nu$.

Next, assume that $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ are σ-finite and let $(A_n)_{n\geq 1}$ in $\mathcal{A}$ and $(B_n)_{n\geq 1}$ in $\mathcal{B}$ such that $A_n \uparrow S$, $B_n \uparrow T$ and $\mu(A_n) < \infty$ and $\nu(B_n) < \infty$ for $n \geq 1$. Setting $C_n := A_n \times B_n$ for $n \geq 1$, we have $(\mu \times \nu)(C_n) < \infty$. Moreover $C_n \uparrow S \times T$, proving the σ-finiteness claim.

For uniqueness, let τ be another measure on $\mathcal{A} \times \mathcal{B}$ such that $\tau(A \times B) = \mu(A) \cdot \nu(B)$ for $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Fix $n \in \mathbb{N}$ and define $(\mu \times \nu)^{(n)}$ and $\tau^{(n)}$ on $\mathcal{A} \times \mathcal{B}$ by

$$(\mu \times \nu)^{(n)}(R) := (\mu \times \nu)(R \cap C_n) \qquad \text{and} \qquad \tau^{(n)}(R) := \tau(R \cap C_n).$$

Then $(\mu \times \nu)^{(n)}$ and $\tau^{(n)}$ are measures and

$$(\mu \times \nu)^{(n)}(S \times T) = (\mu \times \nu)(C_n) < \infty$$

and similarly $\tau^{(n)}(S \times T) < \infty$. Since $(\mu \times \nu)^{(n)}$ and $\tau^{(n)}$ coincide on $\mathcal{R}$, it follows from Example 3.4(a) and Proposition 3.5 that $(\mu \times \nu)^{(n)} = \tau^{(n)}$ on $\sigma(\mathcal{R}) = \mathcal{A} \times \mathcal{B}$. Therefore, for $R \in \mathcal{A} \times \mathcal{B}$, Theorem 2.9 yields

$$(\mu \times \nu)(R) = \lim_{n\to\infty} (\mu \times \nu)(R \cap C_n) = \lim_{n\to\infty} (\mu \times \nu)^{(n)}(R) = \lim_{n\to\infty} \tau^{(n)}(R) \lim_{n\to\infty} \tau(R \cap C_n) = \tau(R),$$

finishing the proof ☺

In the remainder of this section, when we write $\mu \times \nu$ we will always mean the product measure provided by Proposition 7.5.

Example 7.6. We have $(\mathbb{R} \times \mathbb{R}, \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}), \lambda \times \lambda) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda)$.

Proof. We have to show that $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$ and, writing λ_1 for the 1-dimensional Lebesgue measure and λ_2 for the 2-dimensional Lebesgue measure, we have to show that $\lambda_1 \times \lambda_1 = \lambda_2$.

To show $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$, recall that $\mathcal{I}^d$ denotes the collection of all half-open rectangles in $\mathbb{R}^d$ from Example 1.7. Since

$$\mathcal{I}^2 = \{A \times B : A, B \in \mathcal{I}^1\} \subseteq \{A \times B : A, B \in \mathcal{B}(\mathbb{R})\}$$

we obtain from Exercise 1.10 that

$$\mathcal{B}(\mathbb{R}^2) = \sigma(\mathcal{I}^2) \subseteq \sigma(\{A \times B : A, B \in \mathcal{B}(\mathbb{R})\}) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}).$$

Conversely, we note that $\{A \subseteq \mathbb{R} : A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)\}$ is a σ-algebra on S which contains $\sigma(\mathcal{I}^1) = \mathcal{B}(\mathbb{R})$, since it contains $\mathcal{I}^1$. It follows that $A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$ for all $A \in \mathcal{B}(\mathbb{R})$. Similarly $\mathbb{R} \times B \in \mathcal{B}(\mathbb{R}^2)$ for all $B \in \mathcal{B}(\mathbb{R})$. Taking intersections, we conclude that

$$\{A \times B : A, B \in \mathcal{B}(\mathbb{R})\} \subseteq \mathcal{B}(\mathbb{R}^2),$$

and thus $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) = \sigma(\{A \times B : A, B \in \mathcal{B}(\mathbb{R})\}) \subseteq \mathcal{B}(\mathbb{R}^2)$.

Next, writing λ_1 for the 1-dimensional Lebesgue measure and λ_2 for the 2-dimensional Lebesgue measure, we have for $I = (a_1, b_1] \times (a_2, b_2] \in \mathcal{I}^2$

$$(\lambda_1 \times \lambda_1)(I) = \lambda_1((a_1, b_1]) \cdot \lambda_1((a_2, b_2]) = (b_1 - a_1)(b_2 - a_2) = \lambda_2(I).$$

Thus by the uniqueness claim in Theorem 3.9, it follows that $\lambda_1 \times \lambda_1 = \lambda_2$. ☺

7.2. **The Fubini–Tonelli theorem.** After the preparations in the previous subsection, we are now ready to discuss the conditions under which we are allowed to change the order of integration. We start with a version for positive-valued functions, which was first proven by Tonelli.[35]

Theorem 7.7 (Tonelli's theorem). *Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be σ-finite measure spaces. If $f\colon S \times T \to [0, \infty]$ is measurable, then*

(i) The function $t \mapsto \int_S f(s,t)\,\mathrm{d}\mu(s)$ is measurable;
(ii) The function $s \mapsto \int_T f(s,t)\,\mathrm{d}\nu(t)$ is measurable;
(iii) We have

$$\int_{S\times T} f \,\mathrm{d}(\mu\times\nu) = \int_T \int_S f \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_S \int_T f\,\mathrm{d}\nu\,\mathrm{d}\mu.$$

Proof. First assume that $\mu(S) = \nu(T) = 1$. We claim that the collection

$$\mathcal{C} := \big\{C \in \mathcal{A}\times\mathcal{B} : \text{(i)-(iii) hold for } f = \mathbf{1}_C\big\} \subseteq \mathcal{A}\times\mathcal{B}$$

is a σ-algebra. It is clear that $\varnothing, S\times T \in \mathcal{C}$. If $C \in \mathcal{C}$, then $\mathbf{1}_{C^c} = 1 - \mathbf{1}_C$, which implies that (i) and (ii) hold for $\mathbf{1}_{C^c}$ and furthermore

$$\int_{S\times T} \mathbf{1}_{C^c}\,\mathrm{d}(\mu\times\nu) = 1 - \int_{S\times T} \mathbf{1}_C\,\mathrm{d}(\mu\times\nu) = 1 - \int_T\int_S \mathbf{1}_C\,\mathrm{d}\mu\,\mathrm{d}\nu = \int_T\int_S \mathbf{1}_{C^c}\,\mathrm{d}\mu\,\mathrm{d}\nu.$$

A similar calculation for the other iterated integral yields that (iii) holds for $\mathbf{1}_{C^c}$ and thus $C^c \in \mathcal{C}$. Finally, for $C_1, C_2, \ldots \in \mathcal{C}$ set $C := \bigcup_{k=1}^\infty C_k$. Then $\mathbf{1}_{\bigcup_{k=1}^n C_k} \uparrow \mathbf{1}_C$ as $n\to\infty$, so (i)-(iii) for $\mathbf{1}_C$ follow from the definition of the integral and therefore $C \in \mathcal{C}$.

We have shown that $\mathcal{C}$ is a σ-algebra. Furthermore, since $(\mu\times\nu)(A\times B) = \mu(A)\cdot\nu(B)$, it is clear that $A \times B \in \mathcal{C}$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Since $\mathcal{A}\times\mathcal{B}$ is the smallest σ-algebra containing $A\times B$ for all $A\in\mathcal{A}$ and $B \in \mathcal{B}$, it follows that $\mathcal{A}\times\mathcal{B} = \mathcal{C}$. This proves (i)-(iii) for $f = \mathbf{1}_C$ with $C \in \mathcal{A}\times\mathcal{B}$, which extends to simple functions by linearity.

For a measurable $f\colon S\times T\to[0,\infty]$, we can find a sequence of simple functions $(f_n)_{n\geq 1}$ such that $0 \leq f_n \uparrow f$ pointwise by Theorem 4.14. Then, for $t \in T$ we have that $(f_n(\cdot, t))_{n\geq1}$ is a sequence of simple functions such that $0 \leq f_n(\cdot,t)\uparrow f(\cdot,t)$ pointwise and thus

$$0 \leq \int_S f_n(s,\cdot)\,\mathrm{d}\mu(s) \uparrow \int_S f(s,\cdot)\,\mathrm{d}\mu(s).$$

A similar statement holds for the integral over T. Therefore, (i)-(ii) for f follow from Theorem 4.11 and (iii) for f follows from a similar approximation argument, proving the theorem in the case $\mu(S) = \nu(T) = 1$.

For the general case, by σ-finiteness, we can find a sequences $(A_n)_{n\geq1}$ in $\mathcal{A}$ and $(B_n)_{n\geq1}$ in $\mathcal{B}$ such that $A_n\uparrow S$, $B_n \uparrow T$ and $\mu(A_n) < \infty$ and $\nu(B_n) < \infty$ for all $n \geq 1$. Now note that

$$\mu^{(n)}(A) := \frac{\mu(A\cap A_n)}{\mu(A_n)} = \int_A \frac{\mathbf{1}_{A_n}}{\mu(A_n)}\,\mathrm{d}\mu, \qquad A \in \mathcal{A},$$
$$\nu^{(n)}(B) := \frac{\nu(B\cap B_n)}{\nu(B_n)} = \int_B \frac{\mathbf{1}_{B_n}}{\nu(B_n)}\,\mathrm{d}\nu, \qquad B \in \mathcal{B},$$

are measures satisfying $\mu^{(n)}(S) = \nu^{(n)}(T) = 1$. To prove (i), observe that for $t \in T$ we have by Theorem 6.8

$$\mu(A_n)\int_S f(s,t)\,\mathrm{d}\mu^{(n)}(s) = \int_S \mathbf{1}_{A_n}(s) f(s,t)\,\mathrm{d}\mu(s) \uparrow \int_S f(s,t)\,\mathrm{d}\mu(s).$$

Therefore, as we already know that $t \mapsto \int_S f(s,t)\,\mathrm{d}\mu^{(n)}(s)$ is measurable, the measurability of $t \mapsto \int_S f(s,t)\,\mathrm{d}\mu(s)$ follows from Theorem 4.11. The proof of (ii) is similar. Lastly, for (iii) observe that for $C = A\times B$ with $A \in \mathcal{A}$ and $B\in\mathcal{B}$ we have

$$\mu^{(n)}\times\nu^{(n)}(C) = \int_C \frac{\mathbf{1}_{A_n\times B_n}}{\mu(A_n)\cdot\nu(B_n)}\,\mathrm{d}(\mu\times\nu). \tag{7.3}$$

[35] Leonida Tonelli (1885-1946) was an Italian mathematician. He is one of the founders of the modern theory of functions of real variables.

Since the right-hand side of (7.3) defines a measure, (7.3) must hold for any $C \in \mathcal{A} \times \mathcal{B}$ by the uniqueness claim in Proposition 7.5. Therefore, since we already know that (iii) holds for $\mu^{(n)}$ and $\nu^{(n)}$, we obtain by Theorem 6.8

$$\int_{S\times T} \mathbf{1}_{A_n\times B_n} f \,\mathrm{d}(\mu\times\nu) = \int_T\int_S \mathbf{1}_{A_n\times B_n} f \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_S\int_T \mathbf{1}_{A_n\times B_n} f \,\mathrm{d}\nu\,\mathrm{d}\mu.$$

Taking the limit $n \to \infty$, (iii) now follows from the MCT. ☺

Remark 7.8. Tonelli's theorem shows that for any $C \in \mathcal{A} \times \mathcal{B}$ we have

$$(\mu\times\nu)(C) = \int_T\int_S \mathbf{1}_C \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_S\int_T \mathbf{1}_C \,\mathrm{d}\nu\,\mathrm{d}\mu, \tag{7.4}$$

which we already knew for $C = A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Instead of using Carathéodory's extension theorem in Proposition 7.5 and then using the product measure to prove Tonelli's theorem, we could also first have proven Tonelli's theorem (without the left-hand side in (iii)) and then use either of the iterated integrals in (7.4) as the definition of the product measure.

As we have already seen in the introduction of this section, one has to be careful when interchanging integrals of real-valued functions. It turns out that, as long as $f\colon S\times T \to \mathbb{R}$ is integrable, one is allowed to interchange the order of integration. This result was first proven by Fubini[36] and predates Tonelli's theorem.

Theorem 7.9 (Fubini's theorem)**.** *Let $(S, \mathcal{A}, \mu)$ and $(T, \mathcal{B}, \nu)$ be σ-finite measure spaces. If $f\colon S\times T \to \mathbb{R}$ is integrable, then*

$$\int_{S\times T} f \,\mathrm{d}(\mu\times\nu) = \int_T\int_S f \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_S\int_T f \,\mathrm{d}\nu\,\mathrm{d}\mu.$$

Proof. Note that for the iterated integrals to be well-defined, we need that

(i) For a.e. $t \in T$ the function $s \mapsto f(s,t)$ is integrable;
(ii) For a.e. $s \in S$ the function $t \mapsto f(s,t)$ is integrable;
(iii) The function $t \mapsto \int_S f(s,t)\,\mathrm{d}\mu(s)$ is well-defined a.e. and integrable;
(iv) The function $s \mapsto \int_T f(s,t)\,\mathrm{d}\nu(t)$ is well-defined a.e. and integrable.

Write $f = f^+ - f^-$ and note that, by Theorem 7.7 and the integrability assumption, $t \mapsto \int_S f^+\,\mathrm{d}\mu$ is measurable and

$$\int_T\int_S f^+ \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_{S\times T} f^+ \,\mathrm{d}(\mu\times\nu) < \infty.$$

This means that $t \mapsto \int_S f^+(s,t)\,\mathrm{d}\mu(s)$ is integrable and thus finite for a.e. $t \in T$ by Remark 5.10(3). Consequently, the function $s \mapsto f^+(s,t)$ is integrable for a.e. $t \in T$. Applying the same argument to f^- yields (i) and (iii). Furthermore, (ii) and (iv) follow similarly. Knowing that all (iterated) integrals under consideration are well-defined and equal for f^+ and f^-, the theorem follows. ☺

Remark 7.10. Fubini's theorem and Tonelli's theorem are often used in tandem. Indeed, let $f\colon S\times T \to \mathbb{R}$ be a measurable function. Suppose one wishes to show that

$$\int_T\int_S f \,\mathrm{d}\mu\,\mathrm{d}\nu = \int_S\int_T f \,\mathrm{d}\nu\,\mathrm{d}\mu. \tag{7.5}$$

Then one first checks that e.g. $\int_T\int_S |f|\,\mathrm{d}\mu\,\mathrm{d}\nu < \infty$, which by Tonelli's theorem and Proposition 5.11 implies that f is integrable. Afterwards Fubini's theorem yields that (7.5) holds. This combination is called the Fubini–Tonelli theorem.

[36]Guido Fubini (1879-1943) was an Italian mathematician, mainly known for his theorem on the iterated Lebesgue integrals.

As a corollary of Tonelli's theorem, we can deduce a generalization of Cavalieri's[37] principle for the Lebesgue integral, also known as the layer cake representation. Heuristically speaking, Cavalieri's principle states that if two objects have the same height and the same cross-sectional area at every point along that height, they have the same volume (see also Figure 7.2).

FIGURE 7.1. Cavalieri's principle visualized with poker chips. The two stacks, although differently shaped, have the same volume.

Corollary 7.11. *Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space and let $f\colon S \to [0,\infty)$ be measurable. Then*

$$\int_S f \,\mathrm{d}\mu = \int_{[0,\infty)} \mu\big(\{s \in S : f(s) \geq t\}\big) \,\mathrm{d}\lambda(t).$$

Now suppose that we have a set $V \subseteq \mathbb{R}^3$ parametrized by

$$V = \big\{(x,y,z) \in \mathbb{R}^3 : f_1(x,y) \leq z \leq f_2(x,y)\big\}$$

for some measurable $f_1, f_2\colon \mathbb{R}^2 \to \mathbb{R}$. Applying Corollary 7.11, we obtain

$$\begin{aligned} \lambda(V) &= \int_{\mathbb{R}^2} f_2(x,y) - f_1(x,y) \,\mathrm{d}\lambda(x,y) \\ &= \int_{[0,\infty)} \lambda\big(\{(x,y) \in \mathbb{R}^2 : f_1(x,y) < t \leq f_2(x,y)\}\big) \,\mathrm{d}\lambda(t). \end{aligned}$$

Noting that $\lambda(V)$ is the volume of V and

$$\lambda(\{(x,y) \in \mathbb{R}^2 : f_1(x,y) < t \leq f_2(x,y)\})$$

is the area of cross-section of V at height $z = t$, we recover Cavalieri's principle.

Proof of Corollary 7.11. Since f is measurable, $F\colon S \times [0,\infty) \to \mathbb{R}$ given by $F(s,t) = f(s) - t$ is measurable as well. It follows that

$$A := \big\{(s,t) \in S \times [0,\infty) : f(s) \geq t\big\} = F^{-1}([0,\infty)) \in \mathcal{A} \times \mathcal{B}([0,\infty)),$$

and therefore the function

$$\mathbf{1}_A(s,t) = \mathbf{1}_{\{f \geq t\}}(s) = \mathbf{1}_{[0,f(s)]}(t), \qquad (s,t) \in S \times [0,\infty)$$

is measurable. By Tonelli's theorem, we obtain

$$\begin{aligned} \int_S f \,\mathrm{d}\mu &= \int_S \int_{[0,\infty)} \mathbf{1}_A(s,t) \,\mathrm{d}\lambda(t) \,\mathrm{d}\mu(s) \\ &= \int_{[0,\infty)} \int_S \mathbf{1}_A(s,t) \,\mathrm{d}\mu(s) \,\mathrm{d}\lambda(t) = \int_{[0,\infty)} \mu\big(\{s \in S : f(s) \geq t\}\big) \,\mathrm{d}\lambda(t), \end{aligned}$$

finishing the proof. ☺

[37]Bonaventura Cavalieri (1598-1647) was another Italian mathematician, known for his work on precursors of infinitesimal calculus.

Exercises

Exercise 7.1. Let $f\colon (0,1]^2 \to \mathbb{R}$ be given by

$$f(x,y) = \frac{x^2 - y^2}{(x^2+y^2)^2} = -\frac{\partial^2}{\partial x \partial y} \arctan(y/x).$$

Show the following identities:

$$\int_0^1 \int_0^1 |f(x,y)| \,\mathrm{d}x\,\mathrm{d}y = \infty,$$
$$\int_0^1 \int_0^1 f(x,y) \,\mathrm{d}x\,\mathrm{d}y = -\frac{\pi}{4},$$
$$\int_0^1 \int_0^1 f(x,y) \,\mathrm{d}y\,\mathrm{d}x = \frac{\pi}{4}.$$

Exercise 7.2. Define $f\colon \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ by

$$f(m,n) := \begin{cases} 1, & \text{if } m = n, \\ -1, & \text{if } m = n+1, \\ 0, & \text{otherwise.} \end{cases}$$

Show that

$$\sum_{m,n\geq 1}^{\infty} |f(m,n)| \neq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} f(m,n) \neq \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} f(m,n).$$

Exercise 7.3. Show that the collection $\mathcal{C}$ in the proof of Proposition 7.1 is a σ-algebra.

Exercise 7.4. Let $(S,\mathcal{A})$ and $(T,\mathcal{B})$ be measurable spaces and define

$$\mathcal{R} := \Big\{ \bigcup_{k=1}^{n} (A_k \times B_k) : A_1,\dots,A_n \in \mathcal{A} \text{ and } B_1,\dots,B_n \in \mathcal{B} \Big\}.$$

(a) Show that $\mathcal{R}$ is a ring.
(b) Show that every $R \in \mathcal{R}$ can be written as

$$R = \bigcup_{k=1}^{n} (A_k \times B_k), \qquad A_1,\dots,A_n \in \mathcal{A},\ B_1,\dots,B_n \in \mathcal{B},$$

such that $(A_j \times B_j) \cap (A_k \times B_k) = \varnothing$ for $1 \leq j \neq k \leq n$.
(c) Give an example of $(S,\mathcal{A})$ and $(T,\mathcal{B})$ such that $\mathcal{R}$ is not a σ-algebra.

Exercise 7.5. Show that the measure space $(S,\mathcal{P}(S),\tau)$, where τ denotes the counting measure, is σ-finite if and only if S is countable.

In the next exercise, we will show that the σ-finiteness assumption in Tonelli's theorem is necessary.

Exercise* 7.6. Let $(S,\mathcal{A},\mu) = ([0,1],\mathcal{B}([0,1]),\lambda)$ and $(T,\mathcal{B},\nu) = ([0,1],\mathcal{P}([0,1]),\tau)$, where τ denotes the counting measure. Define the diagonal

$$D := \big\{(x,x) : x \in [0,1]\big\} \subseteq [0,1] \times [0,1].$$

Show that

$$\int_{[0,1]} \int_{[0,1]} \mathbf{1}_D \,\mathrm{d}\lambda\,\mathrm{d}\tau = 0,$$
$$\int_{[0,1]} \int_{[0,1]} \mathbf{1}_D \,\mathrm{d}\tau\,\mathrm{d}\lambda = 1,$$

and, as an additional challenge, show that

$$\int_{[0,1]\times[0,1]} \mathbf{1}_D \,\mathrm{d}(\lambda \times \tau) = \infty.$$

Hint: Use the definition of $\lambda \times \tau$, given by (B.1).

Exercise* 7.7. Let $(S, \mathcal{A}, \mu)$ be a σ-finite measure space and let $f\colon S \to [0,\infty)$ be measurable. Show that for $p \in [1,\infty)$

$$\int_S f^p \, \mathrm{d}\mu = p \int_{[0,\infty)} t^{p-1} \mu\big(\{s \in S : f(s) \geq t\}\big) \, \mathrm{d}\lambda(t).$$

Hint: Adapt the proof of Corollary 7.11.

8. L^p-spaces

In this section, we will introduce spaces of integrable functions and equip these spaces with a norm. In this way, we will combine metric space theory with measure theory. The most important example will be the space of square integrable functions $L^2(S)$, which plays a fundamental role in e.g. the analysis of partial differential equations and quantum physics. Throughout this section, $(S, \mathcal{A}, \mu)$ is a measure space.

Definition 8.1. *For $p \in [1, \infty)$ let*

$$L^p(S) := \Big\{ f\colon S \to \mathbb{R} : f \text{ is measurable and} \int_S |f|^p \, d\mu < \infty \Big\}.$$

For $f \in L^p(S)$ let

$$\|f\|_p := \Big(\int_S |f|^p \, d\mu \Big)^{\frac{1}{p}}.$$

Note that $L^1(S)$ coincides with the set of integrable functions $f : S \to \mathbb{R}$. Moreover, it is clear that for scalars $\alpha \in \mathbb{R}$ we have

$$\|\alpha f\|_p = |\alpha| \|f\|_p.$$

If $\|f\|_p = 0$, then from Proposition 5.7(vi) we can conclude $f = 0$ a.e. However, we would like $\|\cdot\|_p$ to be a norm on $L^p(S)$ and thus we want to have $f = 0$. Therefore, we will identify two functions f and g whenever $f = g$ a.e.[38] Therefore, one has to be rather careful if one talks about $f(s)$ for a certain fixed $s \in S$. In integration theory this usually does not lead to any problems since $f = g$ a.e. implies that $\int_E f \, \mathrm{d}\mu = \int_E g \, \mathrm{d}\mu$ for any $E \in \mathcal{A}$.

Example 8.2. Let $S = \mathbb{R}$ with the Lebesgue measure λ. Then $\mathbf{1}_{\{0\}} = \mathbf{1}_{\mathbb{Q}} = 0$ and $\mathbf{1}_{[0,1]\setminus\mathbb{Q}} = \mathbf{1}_{[0,1]}$ in $L^p(\mathbb{R})$.

8.1. Minkowski and Hölder's inequalities.

The final property to check in order to show that $\|\cdot\|_p$ is a norm on $L^p(S)$ is the triangle inequality, which is the content of the following proposition.

Proposition 8.3 (Minkowski's inequality[39])**.** *For all $f, g \in L^p(S)$ we have $f + g \in L^p(S)$ and*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

Proof. By Exercise 8.1, we have for all $a, b \in [0, \infty)$

$$(a + b)^p = \inf_{\theta \in (0,1)} \theta^{1-p} a^p + (1 - \theta)^{1-p} b^p. \tag{8.1}$$

It follows from (8.1) that for all $\theta \in (0, 1)$ and all $s \in S$,

$$|f(s) + g(s)|^p \le (|f(s)| + |g(s)|)^p \le \theta^{1-p} |f(s)|^p + (1 - \theta)^{1-p} |g(s)|^p.$$

Therefore, by monotonicity and linearity of the Lebesgue integral, we obtain that for all $\theta \in (0, 1)$,

$$\int_S |f + g|^p \, \mathrm{d}\mu \le \theta^{1-p} \int_S |f|^p \, \mathrm{d}\mu + (1 - \theta)^{1-p} \int_S |g|^p \, \mathrm{d}\mu.$$

Stated differently, this says that for all $\theta \in (0, 1)$,

$$\|f + g\|_p^p \le \theta^{1-p} \|f\|_p^p + (1 - \theta)^{1-p} \|g\|_p^p.$$

Now the result follows by taking the infimum over all $\theta \in (0, 1)$ and applying (8.1). ☺

Another famous inequality, which can be proved with a similar method is the following. In case $p = q = 2$, this is called the Cauchy-Schwarz inequality.

---

[38] More precisely, one can build an equivalent relation $f \sim g$ if $f = g$ almost everywhere and then consider a quotient space. We will use the above imprecise but more intuitive definition.

[39] Hermann Minkowski (1864-1909) was a German mathematician who worked in geometry. He also was Albert Einstein's teacher and provided the 4-dimensional mathematical framework for part of Einstein's relativity theory.

Proposition 8.4 (Hölder's inequality[40])**.** *Let $p, q \in (1, \infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$.[41] If $f \in L^p(S)$ and $g \in L^q(S)$, then $fg \in L^1(S)$ and*

$$\|fg\|_1 \le \|f\|_p \|g\|_q.$$

Proof. See Exercise 8.2. ☺

8.2. **Completeness of L^p.** Having established that $\|\cdot\|_p$ is a norm on $L^p(S)$, the next order of business is to show that $L^p(S)$ with this norm is complete and thus a Banach space.[42] So we will show that a sequence of functions $(f_n)_{n\ge1}$ in $L^p(S)$, which is Cauchy with respect to the norm $\|\cdot\|_p$, also converges with respect to the norm $\|\cdot\|_p$

Theorem 8.5 (Riesz–Fischer[43])**.** *Let $p \in [1, \infty)$. Then $(L^p(S), \|\cdot\|_p)$ is a Banach space.*[44]

Proof. Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in $L^p(S)$. It suffices to show that $(f_n)_{n=1}^\infty$ has a subsequence which is convergent in $L^p(S)$, since any Cauchy sequence with a convergent subsequence is convergent. Recursively, we can find a subsequence $(f_{n_k})_{k\ge1}$ such that

$$\|f_{n_{k+1}} - f_{n_k}\|_p \le \frac{1}{2^k}, \quad k = 1, 2, \dots$$

We will show that there exists an $f \in L^p(S)$ such that $\|f_{n_k} - f\|_p \to 0$.

Define $g, g_1, g_2, \dots : S \to [0, \infty]$ by

$$g := |f_{n_1}| + \sum_{k=1}^{\infty} |f_{n_{k+1}} - f_{n_k}|,$$

$$g_m := |f_{n_1}| + \sum_{k=1}^{m-1} |f_{n_{k+1}} - f_{n_k}|.$$

Then $0 \le g_m \uparrow g$. Therefore, by using Minkowski's inequality, we obtain

$$\begin{aligned}
\|g\|_p &= \Big(\int_S |g|^p \,\mathrm{d}\mu\Big)^{\frac{1}{p}} \\
&\overset{\text{(MCT)}}{=} \lim_{m\to\infty} \Big(\int_S |g_m|^p \,\mathrm{d}\mu\Big)^{\frac{1}{p}} = \lim_{m\to\infty} \|g_m\|_p \\
&\le \lim_{m\to\infty} \big(\|f_{n_1}\|_p + \sum_{k=0}^{m-1} \|f_{n_{k+1}} - f_{n_k}\|_p\big) \\
&\le \|f_{n_1}\|_p + \sum_{k\ge1} 2^{-k} = \|\phi_1\|_p + 1 < \infty.
\end{aligned}$$

Letting $A = \{s \in S : g(s) = \infty\}$, Exercise 5.1 yields $\mu(A^c) = 0$. Therefore, we can define $f : S \to \mathbb{R}$ by

$$f(s) := \begin{cases} f_{n_1}(s) + \sum_{k=1}^{\infty} (f_{n_{k+1}}(s) - f_{n_k}(s)), & s \in S \setminus A, \\ 0, & s \in A, \end{cases}$$

where the series is absolutely convergent and f is measurable by Theorem 4.11. By a telescoping argument it follows that for $s \in S \setminus A$

$$f(s) = \lim_{m\to\infty} \big(f_{n_1}(s) + \sum_{k=1}^{m-1} (f_{n_{k+1}}(s) - f_{n_k}(s))\big) = \lim_{m\to\infty} f_{n_m}(s),$$

[40]Otto Hölder (1859-1937) is most famous for this result and for the notion of Hölder continuity of a function.

[41]These exponents are called conjugate exponents.

[42]$L^p(\mathbb{R})$ could also be defined using Riemann integrability. However, the resulting space would not be suitable for analysis, as it is not complete.

[43]Frigyes Riesz (1880–1956) was a Hungarian mathematician who worked in functional analysis. Ernst Fischer (1875–1954) was a Austrian mathematician who worked in analysis.

[44]Recall that a Banach space is a complete normed vector space. Stefan Banach (1892–1945) was a Polish mathematician who is one of the world's most important 20th-century mathematicians. He is most famous for his book on functional analysis [2].

so $f_{n_m} \to f$ pointwise a.e. Clearly,

$$|f_{n_m}| = |f_{n_1}| + \sum_{k=1}^{m-1} |f_{n_{k+1}} - f_{n_k}| \leq |g|$$

and, by letting $m \to \infty$, we see that also $|f| \leq |g|$ and in particular $f \in L^p(S)$. It follows that

$$|f - f_{n_m}|^p \leq (|f| + |f_{n_m}|)^p \leq (2|g|)^p = 2^p|g|^p.$$

Since $|f - f_{n_m}|^p \to 0$ a.e. and $2^p|g|^p$ is integrable, it follows from the DCT (see also Remark 6.5) that

$$\lim_{m\to\infty} \|f - f_{n_m}\|_p^p = \lim_{m\to\infty} \int_S |f - f_{n_m}|^p \, d\mu = 0,$$

which finishes the proof. ☺

For $f, g \in L^2(S)$ define

$$\langle f, g\rangle := \int_S f(s)g(s) \, d\mu(s).$$

Then $\langle \cdot, \cdot \rangle$ is an inner product that induces the L^2-norm, i.e. $\langle f, f\rangle = \|f\|_2^2$. Thus, Theorem 8.5 states in particular that $L^2(S)$ is a Hilbert space.[45]

From the proof of Theorem 8.5 we deduce the following result.

Corollary 8.6. *Let $p \in [1, \infty)$. Suppose $f, f_1, f_2, \ldots \in L^p(S)$. If $f_n \to f$ in $L^p(S)$, then there exists a subsequence $(f_{n_k})_{k\geq 1}$ such that $f_{n_k} \to f$ a.e.*

In general one does not have $f_n \to f$ a.e. (see Exercise 8.9).

Proof. Indeed, since $(f_n)_{n\geq 1}$ is convergent it is a Cauchy sequence in $L^p(S)$. By the proof of Theorem 8.5 there exists a subsequence $(f_{n_k})_{k\geq 1}$ and a function $\tilde{f}$ such that $f_{k_n} \to \tilde{f}$ in $L^p(S)$ and $f_{n_k} \to \tilde{f}$ a.e. Now Minkowski's inequality yields:

$$\|f - \tilde{f}\|_p \leq \|f - f_{n_k}\|_p + \|f_{n_k} - \tilde{f}\|_p \to 0.$$

Thus $f = \tilde{f}$ a.e. and therefore $f_{n_k} \to f$ a.e. ☺

Example 8.7. Let $S = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$ and let $\mu = \tau$ denote the counting measure on $\mathbb{N}$. For $p \in [1, \infty)$ we define $\ell^p := L^p(\mathbb{N})$, which is the space of sequences such that

$$\|(a_n)_{n\geq 1}\|_p = \Big(\sum_{n\geq 1} |a_n|^p\Big)^{\frac{1}{p}} < \infty.$$

If $p \leq q < \infty$, then $\ell^p \subseteq \ell^q$ (see Exercise 8.10).

Example 8.8. Assume $\mu(S) < \infty$. If $1 \leq p \leq q < \infty$, then $L^q(S) \subseteq L^p(S)$ (see Exercise 8.3).

We end this subsection with a simple density result. The following result shows that any L^p-function can be approximated by simple functions.

Proposition 8.9 (Density of simple functions)**.** *Let $p \in [1, \infty)$. The set of simple functions is dense in $L^p(S)$.*

Proof. By Theorem 4.14 we can find simple functions g_n and h_n such that $0 \leq g_n \uparrow f^+$ and $0 \leq h_n \uparrow f^-$. Then $f_n := g_n - h_n$ is a simple function and $f_n \to f$ pointwise. Now since

$$|f_n - f|^p \leq \big(|g_n| + |h_n| + |f|\big)^p \leq 3^p|f|^p$$

and the latter is integrable, it follows from the DCT that $\|f_n - f\|_p \to 0$ as $n \to \infty$. ☺

[45] Recall that a Hilbert space is a complete normed vector space with a norm induced by an inner product. David Hilbert (1862–1943) is sometimes said to be the last universal mathematician (which means he knew "all" mathematics of his time). He was one of the most influential mathematicians of the 19th and early 20th century.

8.3. **L^p-spaces on intervals.** In the remaining part of this section, we discuss $L^p(I)$, where I is an interval equipped with the Borel σ-algebra $\mathcal{B}(I)$ and the Lebesgue measure λ. For an interval $I \subseteq \mathbb{R}$, let the set of step functions $\mathrm{Step}(I)$ be defined by

$$\mathrm{Step}(I) := \mathrm{span}\{\mathbf{1}_J : J \subseteq I \text{ is an interval with finite length}\}.$$

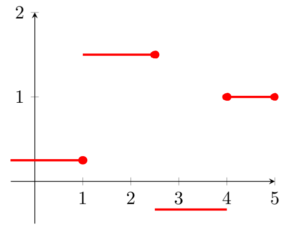

FIGURE 8.1. The step function $\phi = \frac{1}{4}\mathbf{1}_{(-\frac{1}{2},1]} + \frac{3}{2}\mathbf{1}_{(1,\frac{5}{2}]} - \frac{1}{3}\mathbf{1}_{(\frac{5}{2},4)} + \mathbf{1}_{[4,5]}$.

Since intervals are Borel sets, each $\phi \in \mathrm{Step}(I)$ is a simple function. The converse does not hold. For instance $\mathbf{1}_{\mathbb{Q}\cap(0,1)}$ is not a step function. Step function have a lot of structure and often questions about functions in $L^p(I)$ can be reduced to this special class by the following density result.

Theorem 8.10. *For $-\infty \leq a < b \leq \infty$ set $I = (a,b)$ and let $p \in [1,\infty)$. Then* $\mathrm{Step}(I)$ *is dense in $L^p(I)$.*

Proof. Let $f \in L^p(I)$ and let $\varepsilon > 0$. We will construct a step function ϕ such that $\|f-\phi\|_p < 3\varepsilon$.

Step 1: Reduction to bounded I. Let $f_n = \mathbf{1}_{(-n,n]}f$ for $n \in \mathbb{N}$. Then $f_n \to f$ pointwise. Moreover, $|f_n - f|^p \leq |f|^p$. Therefore, the DCT yields $\|f_n - f\|_p \to 0$. So, for $n \in \mathbb{N}$ large enough, we have $\|f-g\|_p < \varepsilon$ for $g := f_n \in L^p(I)$.

Step 2: Set $J = (-n,n]$ and note that $g \in L^p(J)$. By Proposition 8.9, there exists a simple function $h : J \to \mathbb{R}$ such that $\|g-h\|_p < \varepsilon$. We can write $h = \sum_{k=1}^n x_k \cdot \mathbf{1}_{A_k}$ with $(A_k)_{k=1}^n$ in $\mathcal{B}(J)$ disjoint.

Step 3: By Exercise 8.6 there exist $F_k \in \mathcal{F}_J$ such that $\lambda(A_k \triangle F_k) < \big(\frac{\varepsilon}{n|x_k|}\big)^p$ for $k = 1,\ldots,n$. Now let $\phi := \sum_{k=1}^n x_k \cdot \mathbf{1}_{F_k}$. Observe that $|\mathbf{1}_{A_j} - \mathbf{1}_{F_j}| = \mathbf{1}_{A_j \triangle F_j}$. Therefore, Minkowski's inequality yields that

$$\|h-\phi\|_p \leq \sum_{k=1}^n |x_k| \|\mathbf{1}_{A_k \triangle F_k}\|_p = \sum_{k=1}^n |x_k| \lambda(A_k \triangle F_k)^{\frac{1}{p}} < \sum_{k=1}^n |x_k| \frac{\varepsilon}{n|x_k|} < \varepsilon.$$

Clearly, ϕ is a step function. Moreover, by Minkowski's inequality, we find

$$\|f-\phi\|_p \leq \|f-g\|_p + \|g-h\|_p + \|h-\phi\|_p \leq 3\varepsilon,$$

finishing the proof. ☺

We finish this section with the density of continuous functions in $L^p(I)$ if I is a bounded interval.

Corollary 8.11. *For $-\infty < a < b <$ set $I = [a,b]$ and let $p \in [1,\infty)$. Then $C([a,b])$ is dense in $L^p(I)$.*

Proof. Let $f \in L^p(I)$ and let $\varepsilon > 0$. By Theorem 8.10, there exists a step function ϕ such that $\|f-\phi\|_p < \varepsilon$. It remains find $\psi \in C([a,b])$ such that $\|\phi - \psi\|_p < \varepsilon$. For this, it suffices to approximate $\mathbf{1}_J$ for an arbitrary interval $J \subseteq I$. Using Figure 8.2 and the DCT, the reader can easily convince themselves that this can indeed be done. ☺

Exercises

Exercise 8.1. Prove the identity (8.1).
Hint: Minimize the right-hand side using calculus techniques.

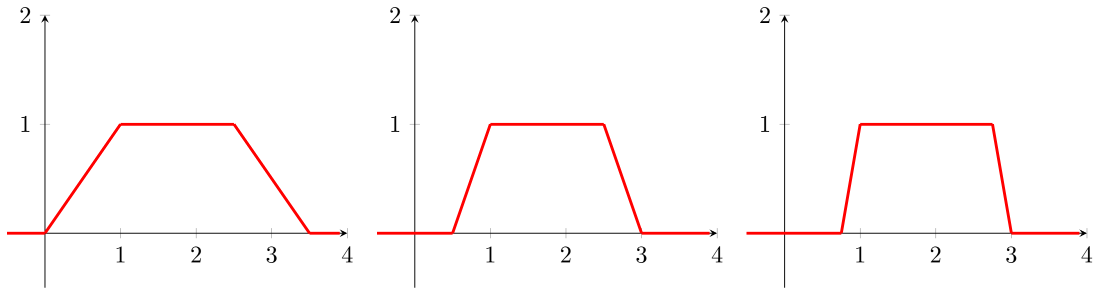

FIGURE 8.2. Approximation of $\mathbf{1}_{(1,\frac{5}{2}]}$ in L^p by continuous functions

Exercise* 8.2. Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

(a) Show that for all $a, b \geq 0$,
$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$
(b) Use the above identity to prove Proposition 8.4.
Hint: First argue why it suffices to consider $\|f\|_p = \|g\|_q = 1$.

Exercise 8.3. Assume $\mu(S) < \infty$ and let $1 \leq q \leq p < \infty$. Show that $L^p(S) \subseteq L^q(S)$ and $\|f\|_q \leq \mu(S)^{\frac{1}{q}-\frac{1}{p}} \|f\|_p$ for all $f \in L^p(S)$.
Hint: Apply Hölder's inequality to $|f|^q \cdot \mathbf{1}_S$.

Exercise 8.4. Let $p \in [1, \infty)$. Determine for which $\alpha \in \mathbb{R}$ one has $f \in L^p(\mathbb{R})$ in each of the following cases:

(a) $f(x) := \mathbf{1}_{(0,1)}(x) x^\alpha$.
(b) $f(x) := \mathbf{1}_{(1,\infty)}(x) x^\alpha$.

Explain why $L^p(\mathbb{R}) \not\subseteq L^q(\mathbb{R})$ for all $p, q \in [1, \infty)$ with $p \neq q$.

Exercise* 8.5. Let $1 \leq p < q < r < \infty$.

(a) Suppose that $f \in L^p(S) \cap L^r(S)$. Prove that $f \in L^q(S)$.
Hint: Set $A := \{s \in S : f(s) \geq 1\}$ and split $f = f\mathbf{1}_A + f\mathbf{1}_{A^c}$.
(b) Suppose that $f \in L^q(S)$. Prove that $f \in L^p(S) + L^r(S)$, where
$$L^p(S) + L^r(S) := \{g + h : g \in L^p(S) \text{ and } h \in L^r(S)\}.$$

Recall that for sets $A, B \subseteq \mathbb{R}$ we defined the symmetric difference as
$$A \triangle B := (A \setminus B) \cup (B \setminus A).$$

Exercise* 8.6. Let $\mathcal{F}_{(a,b]}$ be the finite unions of half-open intervals in $(a, b]$.[46] Define
$$\mathcal{A} := \{A \in \mathcal{B}((a, b]) : \forall \varepsilon > 0, \exists F \in \mathcal{F}_{(a,b]} \text{ such that } \lambda(A \triangle F) < \varepsilon\}.$$

(a) For $A, B \in \mathcal{B}((a, b])$ show that $A \triangle B = A^c \triangle B^c$.
(b) Let I be an index set and $A_i, B_i \subseteq \mathcal{B}((a, b])$ for all $i \in I$. Let $A = \bigcup_{i \in I} A_i$ and $B = \bigcup_{i \in I} B_i$. Show that $A \triangle B \subseteq \bigcup_{i \in I} A_i \triangle B_i$.
(c) Deduce from (a) that $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$.
(d) Assume $(A_n)_{n \geq 1}$ is a disjoint sequence in $\mathcal{A}$. Deduce from (b) that $\bigcup_{n=1}^\infty A_n \in \mathcal{A}$.
Hint: Use that $\lim_{n \to \infty} \lambda(\bigcup_{k=n}^\infty A_k) = 0$ in order to reduce to finitely many sets.
(e) Show that $\mathcal{A} = \mathcal{B}((a, b])$.
Hint: Use Exercises 1.10 and 3.8.

Exercise* 8.7. Let $-\infty < a < b < \infty$, $f \in L^1((a, b])$ and assume $\int_{(a,t]} f \, \mathrm{d}\lambda = 0$ for all $t \in (a, b]$. We will derive that $f = 0$ in $L^1((a, b])$.

(a) Show that for all $A \in \mathcal{F}^1$ with $A \subseteq (a, b]$, $\int_A f \, \mathrm{d}\lambda = 0$.

[46] By Exercise 1.7, we can take disjoint unions.

(b) Let $A \in \mathcal{B}(\mathbb{R})$ be such that $A \subseteq (a,b]$. Construct sets $A_1, A_2, \ldots \in \mathcal{F}^1$ such that $A_n \subseteq (a,b]$ and $\mathbf{1}_{A_n} \to \mathbf{1}_A$ a.e.
Hint: Use Exercise 8.6 and Corollary 8.6.

(c) Show that for all $A \in \mathcal{B}(\mathbb{R})$ with $A \subseteq (a,b]$, we have $\int_A f \, \mathrm{d}\lambda = 0$.
Hint: Use (a) and (b) and DCT.

(d) Derive that $f = 0$ in $L^1((a,b])$.
Hint: Consider $A := \{x \in (a,b] : f(x) \geq 0\}$ and $B := \{x \in (a,b] : f(x) \leq 0\}$ separately.

Exercise* 8.8. Let $f, g \in L^1(\mathbb{R})$ and define $h\colon \mathbb{R}^2 \to \mathbb{R}$ by $h(x,y) = f(x-y)g(y)$.

(a) Show that $h \in L^1(\mathbb{R}^2)$.
Hint: Use Tonelli's theorem.

(b) Show that there exists a set $E \subseteq \mathbb{R}$ with $\lambda(E) = 0$ such that $h(x,\cdot) \in L^1(\mathbb{R})$ for all $x \in \mathbb{R} \setminus E$.

Now define

$$f * g(x) := \int_{\mathbb{R}} f(x-y)g(y) \, \mathrm{d}\lambda(y) = \int_{\mathbb{R}} h(x,y) \, \mathrm{d}\lambda(y), \qquad x \in \mathbb{R} \setminus E,$$

and set $f * g(x) = 0$ for $x \in E$.

(c) Show that $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$.

Exercise** 8.9. Note that each $n \in \mathbb{N}$ can be uniquely written as $n = 2^k + j$ with $k \in \mathbb{N}_0$ and $j \in \{0, \ldots, 2^k - 1\}$. Now, for such n, define $f_n := \mathbf{1}_{(j2^{-k},(j+1)2^{-k}]}$. Show that $f_n \to 0$ in $L^1(\mathbb{R})$, but for all $x \in (0,1]$, there exists infinitely many $n \in \mathbb{N}$ such that $f_n(x) = 1$.

Exercise** 8.10. Let $1 \leq p \leq q < \infty$).

(a) Prove $\ell^p \subseteq \ell^q$ and that for all $(a_n)_{n\geq 1} \in \ell^p$ one has $\|(a_n)_{n\geq 1}\|_q \leq \|(a_n)_{n\geq 1}\|_p$.
Hint: By homogeneity, one can assume $\|(a_n)_{n\geq 1}\|_p = 1$, and therefore $|a_n| \leq 1$ for all $n \in \mathbb{N}$.

(b) Let $a_n = n^\alpha$ for $n \in \mathbb{N}$. For which $\alpha \in \mathbb{R}$ does one have $(a_n)_{n\geq 1} \in \ell^p$?

There is a natural limiting space of $L^p(S)$ for $p \to \infty$:

Exercise** 8.11. A measurable function $f : S \to \mathbb{R}$ is said to be in $L^\infty(S)$ if there exists an $M \geq 0$ such that $\mu(\{|f| > M\}) = 0$.[47] For $f \in L^\infty(S)$ define

$$\|f\|_\infty = \inf\{M \geq 0 : \mu(\{|f| > M\}) = 0\}.$$

As usual, we identify functions f and g in $L^\infty(S)$ if $f = g$ a.e.

(a) Show that $(L^\infty(S), \|\cdot\|_\infty)$ is a Banach space.

(b) Show that for $f \in L^\infty(S)$ and $g \in L^1(S)$, prove that $fg \in L^1(S)$ and

$$\|fg\|_1 \leq \|f\|_\infty \|g\|_1.$$

(c) Assume $\mu(S) < \infty$ and $p \in [1,\infty)$. Show that $L^\infty(S) \subseteq L^p(S)$ and $\|f\|_p \leq \mu(S)^{\frac{1}{p}} \|f\|_\infty$ for $f \in L^\infty(S)$.

(d) Assume $S = \mathbb{N}$ with the counting measure and $p \in [1,\infty)$. Define $\ell^\infty := L^\infty(\mathbb{N})$. Show that $\ell^p \subseteq \ell^\infty$ and $\|(a_n)_{n\geq 1}\|_\infty \leq \|(a_n)_{n\geq 1}\|_p$ for $(a_n)_{n\geq 1} \in \ell^p$.

[47] In other words $|f| \leq M$ a.e.

9. Applications to Fourier series

In this section we want to allow the scalar field to be complex as well and we use the notation $\mathbb{K}$ for this. So $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. Recall that we defined in the integral for a complex-valued function in Definition 5.18. Moreover, the theory developed in the subsequent sections extends to complex-valued functions by considering real and imaginary parts separately, see e.g. Exercise 5.11.

Fourier[48] analysis plays a role in a large part of mathematics. In particular, it is one of the central mathematical tools in Physics and Electrical Engineering. A Fourier series is of the form[49]

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad \text{or equivalently} \quad \frac{a_0}{2} + \sum_{n\geq 1} a_n \cos(nx) + b_n \sin(nx),$$

where $x \in [0, 2\pi]$. Of course one can also use $x \in \mathbb{R}$ here. Clearly, the above functions will be periodic whenever they are well-defined.

One of the reasons that Fourier series naturally arise in mathematics is that each of the functions $e^{\pm inx}, \cos(nx), \sin(nx)$ as an eigenfunction of $\frac{d^2}{dx^2}$ with eigenvalue $-n^2$. Indeed, for instance $\cos(nx)'' = -n^2 \cos(nx)$.

We have seen that Taylor series can be used to represent functions which are smooth enough.[50] Fourier series provides another tool to represent functions. The class of functions which can be represented as a Fourier series will turn out to be enormous.

In this section we will prove a couple of central results in the theory of Fourier series. The interested reader can read more on the subject in [9], [10], [12], [13] and [18]. In particular, very interesting but mostly elementary applications to geometry, ergodicity, number theory and PDEs can be found in [12].

9.1. **Fourier coefficients.** In this section $S = [0, 2\pi]$ and λ is the Lebesgue measure on $[0, 2\pi]$. For notational convenience we will write

$$\int_a^b f(x)\,\mathrm{d}x := \int_{[a,b]} f\,\mathrm{d}\lambda.$$

for $f \in L^1(0, 2\pi)$ and $[a, b] \subseteq [0, 2\pi]$.

Definition 9.1. *Let $e_k : [0, 2\pi] \to \mathbb{C}$ be given by $e_k(x) = e^{ikx}$ for $k \in \mathbb{Z}$.*[51] *Let $f \in L^1(0, 2\pi)$.*

*(i) For $k \in \mathbb{Z}$ the **k-th order Fourier coefficient** is defined by*[52]

$$\widehat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(x)\overline{e_k(x)}\,\mathrm{d}x.$$

*(ii) for $n \in \mathbb{N}_0$, the **n-th partial sum of the Fourier series** $s_n(f) : [0, 2\pi] \to \mathbb{C}$ is defined by*

$$s_n(f) = \sum_{|k|\leq n} \widehat{f}(k) e_k. \tag{9.1}$$

*(iii) A function of the form $\sum_{|k\leq n} c_k e_k$ with $n \in \mathbb{N}_0$ and $(c_k)_{|k|\leq n}$ in $\mathbb{C}$ is called a **trigonometric polynomial**.*

Remark 9.2.

(1) The reason to use the notion "trigonometric polynomial" is that $e_k(x) = (e^{ix})^k$. Also note that $e_k(x) = \cos(kx) + i \sin(kx)$ by Euler's formula.

---

[48]Fourier analysis is named after the French mathematician Jean-Baptiste Fourier (1768–1830), and were introduced in order to solve differential equations such as the heat equation.

[49]Recall Euler's formula: $e^{ix} = \cos(x) + i\sin(x)$ for $x \in \mathbb{R}$.

[50]In Complex Function Theory these functions will be characterized as the so-called analytic functions

[51]Note that $e_0 = \mathbf{1}_{[0,2\pi]}$.

[52]To calculate Fourier transforms numerically one can use the so-called fast Fourier transform FFT (see [12, Section 7.1.3])

(2) Observe that by (the complex version of) Proposition 5.11 for each $k \in \mathbb{Z}$,

$$|\widehat{f}(k)| \leq \frac{1}{2\pi}\int_0^{2\pi} |f(x)e_k(x)|\,\mathrm{d}x = \frac{1}{2\pi}\|f\|_1. \tag{9.2}$$

In Exercise 9.4 it will be shown that one even has $\lim_{k\to\infty} \widehat{f}(k) = 0$.

(3) If $f, g \in L^1(0, 2\pi)$, the following linearity property holds: $\widehat{(f+g)}(k) = \widehat{f}(k) + \widehat{g}(k)$ for all $k \in \mathbb{Z}$. This follows directly from the linearity of the Lebesgue integral.

Example 9.3. Given $f \in L^1(0, 2\pi)$, the function $s_n(f)$ is a trigonometric polynomial for each $n \in \mathbb{N}$. Each of the functions

$$\cos(nx) = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin(nx) = \frac{e^{inx} - e^{-inx}}{2i}$$

is a trigonometric polynomial as well.

Finally note that if P is a trigonometric polynomial, then for any $j \in \mathbb{N}_0$, P^j is a trigonometric polynomial as well.

The main questions in this section is whether we can reconstruct f from its Fourier coefficients. More precisely:

- (Representation) Which functions $f : [0, 2\pi] \to \mathbb{C}$ can we write as $f = \sum_{k\in\mathbb{Z}} c_k e_k$ for certain coefficients $(c_k)_{k\in\mathbb{Z}}$?
- (Convergence) In what sense does the above series converge?
- (Uniqueness) Does $\widehat{f}(k) = \widehat{g}(k)$ for all $k \in \mathbb{Z}$ imply $f = g$?

When considering convergence of Fourier series we will always consider the convergence of

$$\sum_{|k|\leq n} c_k e_k = \sum_{k=-n}^{n} c_k e_k \quad \text{as } n \to \infty.$$

9.2. **Weierstrass' approximation result and uniqueness.** Before we consider convergence of Fourier series, we first we prove a fundamental result about the approximation by trigonometric polynomials. It will be an essential ingredient in the uniqueness result in Theorem 9.5.

Theorem 9.4 (Weierstrass' approximation theorem[53] for periodic functions)**.** *The trigonometric polynomials are dense in* $\{f \in C([0, 2\pi]) : f(0) = f(2\pi)\}$.

Proof. Let $f \in C([0, 2\pi])$ be such that $f(0) = f(2\pi)$ and let $\varepsilon > 0$ be arbitrary. It suffices to show that there exists a trigonometric polynomial P such that $\|f - P\|_\infty < \varepsilon$. We extend f periodically to a function $f : \mathbb{R} \to \mathbb{C}$. Since f is also uniformly continuous we can choose $\delta \in (0, \pi)$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon/2$.

Define[54] $F_n = \frac{1}{n}\sum_{k=1}^{n}\sum_{|j|\leq k-1} e_j$ which is a periodic function as well. Define $P_n : [0, 2\pi] \to \mathbb{C}$ by $P_n(x) = \frac{1}{2\pi}\int_0^{2\pi} F_n(x - y)f(y)\,\mathrm{d}y$. Then since $e_j(x - y) = e^{2\pi ix}e^{-2\pi iy}$ the following identity holds

$$\begin{aligned} P_n(x) &= \frac{1}{2\pi}\int_0^{2\pi} F_n(x - y)f(y)\,\mathrm{d}y \\ &= \frac{1}{2\pi n}\sum_{k=1}^{n}\sum_{|j|\leq k-1} e_j(x)\int_0^{2\pi} e_j(-y)f(y)\,\mathrm{d}y \\ &= \frac{1}{n}\sum_{k=1}^{n}\sum_{|j|\leq k-1} e_j(x)\widehat{f}(j). \end{aligned}$$

This shows that P_n is a trigonometric polynomial.

---

[53] Originally Weierstrass proved that the polynomials are dense in $C([a, b])$. This can be derived from our version of the theorem as indicated in Exercise 9.11

[54] This is called Féjer's kernel

By Exercise 9.3 the following identity holds

$$F_n(z) = \frac{\sin^2(nz/2)}{n\sin^2(z/2)}, \quad z \in (0, 2\pi). \tag{9.3}$$

Therefore, $F_n(z) \geq 0$, and thus we can write

$$\|F_n\|_1 = \int_0^{2\pi} F_n(z)\,\mathrm{d}z = \frac{1}{n}\sum_{k=1}^n \sum_{|j|\leq k-1} \int_0^{2\pi} e_j(z)\,\mathrm{d}z = \frac{1}{n}\sum_{k=1}^n 2\pi = 2\pi.$$

Fix $x \in [0, 2\pi]$. It follows that

$$\begin{aligned} 2\pi|f(x) - P_n(x)| &= \Big|f(x)\int_0^{2\pi} F_n(x-y)\,\mathrm{d}y - \int_0^{2\pi} f(y)F_n(x-y)\,\mathrm{d}y\Big| \\ &= \Big|\int_0^{2\pi} F_n(x-y)(f(x) - f(y))\,\mathrm{d}y\Big| \\ &\leq \int_0^{2\pi} |f(x) - f(y)|F_n(x-y)\,\mathrm{d}y\Big| \leq T_1 + T_2, \end{aligned}$$

where T_1 and T_2 are the integrals over $I := [x-\delta, x+\delta]$ and $J := [0, x-\delta]\cup[x+\delta, 2\pi]$, respectively. On the interval I we can use the uniform continuity of f to estimate

$$T_1 \leq \frac{\varepsilon}{2}\int_I F_n(x-y)\,\mathrm{d}y \leq \pi\varepsilon.$$

On $J := [0, x-\delta]\cup[x+\delta, 2\pi]$ we can estimate

$$T_2 \leq 2\|f\|_\infty \int_J F_n(x-y)\,\mathrm{d}y = 2\|f\|_\infty \int_{B_\delta} F_n(z)\,\mathrm{d}z,$$

where we substituted $z := x - y$ and where $B_\delta = [\delta, 2\pi - \delta]$. Therefore, using (9.3) again and the fact that $|\sin(z/2)| \geq \sin(\delta/2)$ for $z \in B_\delta$ (recall that $\delta \leq \pi$) we obtain

$$\int_{B_\delta} F_n(z)\,\mathrm{d}z \leq \frac{|B_{\delta|}}{n\sin^2(\delta/2)} \leq \frac{2\pi}{n\sin^2(\delta/2)}.$$

So choosing $n \geq 1$ so large that $\frac{2\|f\|_\infty}{n\sin^2(\delta/2)} < \frac{\varepsilon}{2}$ we obtain $T_2 < \pi\varepsilon$.

Therefore, combining the the estimates can conclude that $|f(x) - P_n(x)| \leq \frac{T_1+T_2}{2\pi} < \varepsilon$. Since $x \in [0, 2\pi]$ was arbitrary it follows that $\|f - P_n\|_\infty < \varepsilon$ as required. ☺

Now we can deal with the uniqueness question for Fourier series. This is the most technical part of this section and could be skipped it at first reading.

Theorem 9.5 (Uniqueness)**.** *If $f \in L^1(0, 2\pi)$ satisfies $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$, then $f = 0$ in $L^1(0, 2\pi)$.*

Proof. [55] *Step 1:* First assume $f \in C([0, 2\pi])$ and $f(0) = f(2\pi)$. By linearity it follows that for each trigonometric polynomial P we have

$$\int_0^{2\pi} f(x)\overline{P(x)}\,\mathrm{d}x = 0.$$

By Theorem 9.4 we can find trigonometric polynomials (P_n) such that $P_n \to f$ uniformly. Therefore, it follows that

$$\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x = \lim_{n\to\infty}\int_0^{2\pi} f(x)\overline{P_n(x)}\,\mathrm{d}x = 0.$$

This implies $f = 0$.

[55] There is a much better proof in the literature using the Féjer kernel as an approximate identity. Since approximate identities are not part of these lecture notes, we proceed differently.

Step 2: Next let $f \in L^1(0,2\pi)$ and assume $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$. Let $F : [0,2\pi] \to \mathbb{R}$ be defined by

$$F(t) = \int_0^t f(x)\,\mathrm{d}x.$$

Then $F \in C([0,2\pi])$ (see Example 6.6), and $F(0) = F(2\pi) = 0$.[56] By Exercise 9.6(b) $\widehat{F}(k) = \frac{\widehat{f}(k)}{ik} = 0$ for all $k \neq 0$. Now let $g = F - C$, where $C = \widehat{F}(0)$. Then $g \in C([0,2\pi])$, $g(0) = g(2\pi)$ and $\widehat{g}(k) = 0$ for all $k \in \mathbb{Z}$. Therefore, $g = 0$ by step 1 and hence $F = C$. Since $F(0) = 0$, this yields $F(x) = 0$ for all $x \in [0,2\pi]$. By Exercise 8.7 we find $f = 0$. ☺

9.3. **Fourier series in $L^2(0,2\pi)$.** In this section we will consider Fourier series in the Hilbert space $L^2(0,2\pi)$. Here the inner product is given by

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\,\mathrm{d}x.$$

Note that $L^2(0,2\pi) \subseteq L^1(0,2\pi)$ (see Exercise 8.3). Therefore, if $f \in L^2(0,2\pi)$ the Fourier coefficients are well-defined and $\widehat{f}(k) = (2\pi)^{-1}\langle f, e_k\rangle$.

Let us recall as special case of Proposition 8.3 for $f, g \in L^2(0,2\pi)$,

$$|\langle f, g\rangle| \leq \|f\|_2\|g\|_2 \qquad \text{(Cauchy–Schwarz inequality)}. \tag{9.4}$$

We will say that $f, g \in L^2(0,2\pi)$ are orthogonal if $\langle f, g\rangle = 0$. Note that in this case the following form of Pythagoras theorem holds[57]

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2. \tag{9.5}$$

Lemma 9.6 (Orthogonality)**.** *For $j, k \in \mathbb{Z}$,*

$$\langle e_j, e_k\rangle = \begin{cases} 2\pi, & \text{if } j = k; \\ 0, & \text{if } j \neq k. \end{cases}$$

Consequently, if finitely many $(c_j)_{j\in\mathbb{Z}}$ in $\mathbb{C}$ are nonzero, then

$$\Big\|\sum_{j\in\mathbb{Z}} c_j e_j\Big\|_2 = (2\pi)^{\frac{1}{2}}\Big(\sum_{j\in\mathbb{Z}}|c_j|^2\Big)^{\frac{1}{2}}. \tag{9.6}$$

Proof. Indeed, if $j \neq k$, then using $\overline{e_k(x)} = e^{-ikx}$ we find that

$$\langle e_j, e_k\rangle = \int_0^{2\pi} e^{i(j-k)x}\,\mathrm{d}x = \int_0^{2\pi}\cos((j-k)x)\,\mathrm{d}x + i\int_0^{2\pi}\sin((j-k)x)\,\mathrm{d}x = 0$$

by periodicity of cos and sin. Similarly, one sees $\langle e_j, e_j\rangle = 2\pi$.

The final statement follows from

$$\Big\|\sum_{j\in\mathbb{Z}} c_j e_j\Big\|_2^2 = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} c_j\overline{c_k}\langle e_j, e_k\rangle = 2\pi\sum_{j\in\mathbb{Z}}|c_j|^2. \qquad ☺$$

We extend this result to series using the completeness of $L^2(0,2\pi)$.

Theorem 9.7 (Riesz–Fischer, Convergence of Fourier series in L^2)**.**

(i) If $(c_n)_{n\in\mathbb{Z}} \in \ell^2$, then $g := \sum_{n\in\mathbb{Z}} c_n e_n$ converges in $L^2(0,2\pi)$, and $\widehat{g}(n) = c_n$ for all $n \in \mathbb{Z}$, and

$$\|g\|_2 = (2\pi)^{\frac{1}{2}}\|(c_n)_{n\in\mathbb{Z}}\|_{\ell^2} \qquad \textit{(Parseval's identity)} \tag{9.7}$$

(ii) If $f \in L^2(0,2\pi)$, then $(\widehat{f}(n))_{n\in\mathbb{Z}} \in \ell^2$ and $f = \sum_{n\in\mathbb{Z}}\widehat{f}(n)e_n$ in $L^2(0,2\pi)$ and (9.7) holds with $g = f$ and $c_n = \widehat{f}(n)$ for $n \in \mathbb{Z}$.

[56] Note that $F(2\pi) = (2\pi)\widehat{f}(0) = 0$.

[57] This follows by writing out $\|f+g\|_2^2 = \langle f+g, f+g\rangle$.

Part (ii) shows that every L^2-function can be represented as a Fourier series. A similar result holds for series of sine and cosine functions and can be derived as a consequence of the above result (see Exercise 9.8).

Proof. (i): Let $g_n = \sum_{|k|\leq n} c_k e_k$ for $n \in \mathbb{N}$. We show that $(g_n)_{n\geq 1}$ is a Cauchy sequence. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ such that

$$\Big(\sum_{|k|\geq N} |c_k|^2\Big)^{\frac{1}{2}} < \frac{\varepsilon}{(2\pi)^{\frac{1}{2}}}.$$

Then for all integers $n > m \geq N$ by Lemma 9.6,

$$\|g_n - g_m\|_2 = \Big\|\sum_{m<|k|\leq n} c_k e_k\Big\|_2 = (2\pi)^{\frac{1}{2}}\Big(\sum_{m<|k|\leq n} |c_k|^2\Big)^{\frac{1}{2}} \leq (2\pi)^{\frac{1}{2}}\Big(\sum_{|k|\geq N} |c_k|^2\Big)^{\frac{1}{2}} < \varepsilon.$$

This proves that $(g_n)_{n\geq 1}$ is a Cauchy sequence. By completeness (see Theorem 8.5), we know that $g := \lim_{n\to\infty} g_n$ exists in $L^2(0, 2\pi)$.

To check (9.7) note that by the continuity of $\|\cdot\|_2$ and Lemma 9.6,

$$\|g\|_2 = \lim_{n\to\infty} \|g_n\|_2 = \lim_{n\to\infty} (2\pi)^{\frac{1}{2}}\Big(\sum_{|k|\leq n} |c_k|^2\Big)^{\frac{1}{2}} = (2\pi)^{\frac{1}{2}}\|(c_n)\|_{\ell^2}.$$

Finally, note that $\widehat{g}(k) = (2\pi)^{-1}\langle g, e_k\rangle = \lim_{n\to\infty}(2\pi)^{-1}\langle g_n, e_k\rangle = c_k$.[58]

(ii): Fix $n \in \mathbb{N}$. Since $\langle f - s_n(f), e_k\rangle = 0$ for each $|k| \leq n$, also $\langle f - s_n(f), s_n(f)\rangle = 0$ and hence (9.5) yields

$$\|f\|_2^2 = \|f - s_n(f) + s_n(f)\|_2^2 = \|f - s_n(f)\|_2^2 + \|s_n(f)\|_2^2 \geq \|s_n(f)\|_2^2. \tag{9.8}$$

Let $c_k = \widehat{f}(k)$ for $k \in \mathbb{Z}$. Then (9.8) yields:

$$\|f\|_2^2 \geq \|s_n(f)\|_2 \overset{(9.6)}{=} \sum_{|k|\leq n} 2\pi |c_k|^2.$$

Letting $n \to \infty$, we find $\|(c_k)_{k\in\mathbb{Z}}\|_{\ell^2} \leq (2\pi)^{-1}\|f\|_2 < \infty$.

By (i) we can define $g = \sum_{n\in\mathbb{Z}} c_n e_n$ where the series converges in $L^2(0, 2\pi)$. We claim that $f = g$ in $L^2(0, 2\pi)$. To see this note that by (i), $\widehat{g}(n) = c_n = \widehat{f}(n)$. Therefore, the claim follows from the uniqueness Theorem 9.5 applied to $f - g$. ☺

For $f \in L^2(0, 2\pi)$ Theorem 9.7 yields that $f - s_n(f) = \sum_{|k|>n} \widehat{f}(k)e_k$ and by (9.7)

$$\|f - s_n(f)\|_2^2 = 2\pi \sum_{|k|>n} |\widehat{f}(k)|^2 \qquad (L^2\text{-error estimate}). \tag{9.9}$$

Moreover, since $s_n(f)$ is a trigonometric polynomial, we also find the following:[59]

Corollary 9.8. *The trigonometric polynomials are dense in $L^2(0, 2\pi)$.*

Example 9.9 (Sawtooth function). Let $f : [0, 2\pi) \to \mathbb{R}$ be defined by $f(x) = x - \pi$ and extended periodically on $\mathbb{R}$. For $k \in \mathbb{Z} \setminus \{0\}$, by the Fundamental Theorem of Calculus and integration by parts,

$$\widehat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} (x - \pi)e^{-ikx}\,\mathrm{d}x = \frac{1}{2\pi}\Big[\frac{(x-\pi)e^{-ikx}}{-ik}\Big]_0^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{-ikx}}{-ik}\,\mathrm{d}x = -\frac{1}{ik}. \tag{9.10}$$

Clearly, $\widehat{f}(0) = 0$. Therefore, Theorem 9.7 yield that $f = -\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{e_k}{ik}$ with convergence in $L^2(0, 2\pi)$. Moreover, using that $2\sin(kx) = \frac{e_k - e_{-k}}{i}$ we also find that $f = -2\sum_{k=1}^{\infty} \frac{\sin(k\cdot)}{k}$ with convergence in $L^2(0, 2\pi)$ (see Figure 9.1) for plots of the partial sums).

[58] Here we used $|\langle g - g_n, e_k\rangle| \leq \|g - g_n\|_2\|e_k\|_2$ as follows from (9.5).

[59] With more advanced techniques one can show that the trigonometric polynomials are dense in any $L^p(0, 2\pi)$ with $p \in [1, \infty)$.

The L^2-error can be estimated using (9.9):

$$\|f - s_n(f)\|_2^2 = 2\pi \sum_{|k|>n} |\widehat{f}(k)|^2 \le 4\pi \sum_{k\ge n+1} \frac{1}{k^2} \le 4\pi \int_n^\infty \frac{1}{x^2}\,\mathrm{d}x = \frac{4\pi}{n}.$$

One can show that $s_n(f)$ will not converge to f uniformly (also see Figure 9.1). This is in particular clear for $x = 0$ and $x = 2\pi$, because $f(0) = \pi$ and $f(2\pi) = -\pi$, but $s_n(f)(0) = s_n(f)(2\pi) = 0$.

By applying (9.7) one can obtain a remarkable identity:

$$\|f\|_2^2 = 2\pi \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k^2}.$$

On the other hand, if we calculate $\|f\|_2^2$ with the fundamental theorem of calculus, we obtain

$$\|f\|_2^2 = \int_0^{2\pi} (x-\pi)^2\,\mathrm{d}x = \Big[\tfrac{1}{3}(x-\pi)^3\Big]_0^{2\pi} = \frac{2}{3}\pi^3.$$

Combining both identities gives $\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k^2} = \frac{1}{3}\pi^2$, and so we find $\sum_{k=1}^\infty \frac{1}{k^2} = \frac{1}{6}\pi^2$.

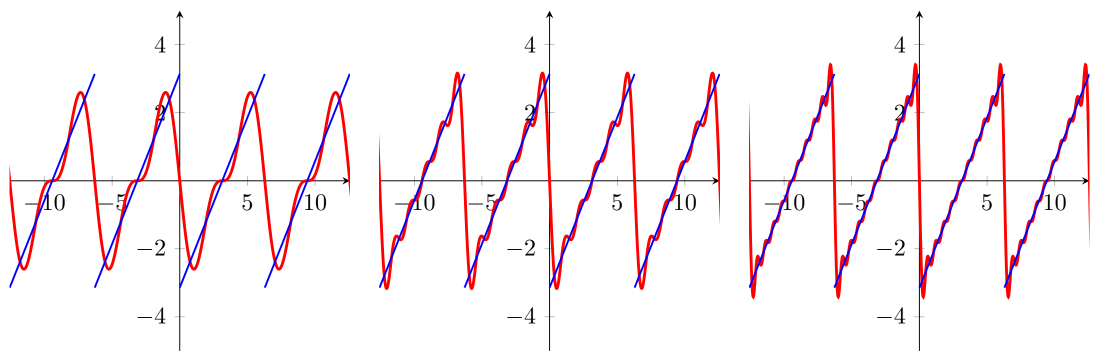

FIGURE 9.1. The Fourier series of the sawtooth function with $n = 2$, $n = 5$ and $n = 10$.

9.4. **Fourier series in $C([0, 2\pi])$.** In this section we will give some sufficient condition on f which imply the Fourier series is uniformly convergent (or equivalently convergent in $C([0, 2\pi])$ with the supremum norm $\|\cdot\|_\infty$). Note that $f_n \to f$ uniformly implies that $f_n \to f$ in $L^2(0, 2\pi)$. Indeed, this follows from

$$\|f_n - f\|_2^2 \le \int_0^{2\pi} |f_n(x) - f(x)|^2\,\mathrm{d}x \le \|f_n - f\|_\infty^2 \int_0^{2\pi} 1\,\mathrm{d}x = 2\pi\|f_n - f\|_\infty^2. \tag{9.11}$$

From the above we see that convergence of Fourier series in $C([0, 2\pi])$ is stronger than convergence in $L^2(0, 2\pi)$. However, there are example of functions $f \in C([0, 2\pi])$ with $f(0) = f(2\pi)$ for which the uniform convergence (and even the pointwise convergence) fails (see [1, Example 35.11] and [10, Example 2.5.1]). So apparently more restrictive conditions are needed.

All the different types of convergence can be confusing. Let us summarize some convergence results for a sequence $(f_n)_{n\ge1}$ in $L^2(0, 2\pi)$.[60]

L^2-convergent $\Longrightarrow$ L^1-convergent $\Longrightarrow$ a.e.-convergent subsequence.

uniform convergent $\Nearrow$ / $\Searrow$

pointwise convergent $\Longrightarrow$ a.e.-convergent

Conversely, L^1 and L^2-convergence are only implied by a.e. convergence under additional assumptions on the $(f_n)_{n\ge1}$, as given for instance in the DCT.

The following result provides sufficient conditions for uniform convergence.

[60] For completeness we note that a.e.-conv. $\Longrightarrow$ conv. in measure $\Longrightarrow$ conv. in distribution.

Theorem 9.10 (Absolute and uniform convergence of Fourier series)**.** *Let $f \in C([0,2\pi])$. If $(\widehat{f}(k))_{k\in\mathbb{Z}} \in \ell^1$, then $f(x) = \sum_{k\in\mathbb{Z}} \widehat{f}(k)e_k(x)$, $x \in [0,2\pi]$ where the series is absolutely and uniformly convergent.*

As a consequence we see that $f(0) = f(2\pi)$ holds in this situation, because $e_k(0) = e_k(2\pi)$.

Proof. For all $x \in [0,2\pi]$,

$$\sum_{k\in\mathbb{Z}} |\widehat{f}(k)e_k(x)| = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)| < \infty.$$

Therefore, we can let $g = \sum_{k\in\mathbb{Z}} \widehat{f}(k)e_k$, where the series is absolutely convergent. Moreover,

$$\|g - s_n(f)\|_\infty \le \sum_{|k|>n} |\widehat{f}(k)| \to 0,$$

and hence $g \in C([0,2\pi])$ and $s_n(f) \to g$ uniformly. By (9.11) the convergence holds in $L^2(0,2\pi)$ as well, and hence

$$2\pi\widehat{g}(k) = \langle g, e_k\rangle = \lim_{n\to\infty} \langle s_n(f), e_k\rangle = 2\pi\widehat{f}(k)$$

and therefore, $g = f$ a.e. by Theorem 9.5. Let $A = \{s \in [0,2\pi] : f(s) = g(s)\}$. Then A is closed and $\lambda(A) = 2\pi$. We claim that A is dense. Indeed, if not then there exists an nonempty open interval $I \subseteq [0,2\pi] \setminus A$. It follows that $0 < \lambda(I) \le \lambda([0,2\pi] \setminus A) = \lambda([0,2\pi]) - \lambda(A) = 0$. This is a contradiction and thus the claim follows. Since A is also closed in $[0,2\pi]$, the claim implies that $A = [0,2\pi]$. ☺

From the proof we see that the following error estimate holds:

$$\|f - s_n(f)\|_\infty \le \sum_{|k|>n} |\widehat{f}(k)| \qquad \text{(uniform error estimate).} \tag{9.12}$$

The condition of Theorem 9.10 holds in the following situation:

Corollary 9.11. *Assume $f \in L^2(0,2\pi)$ satisfies $\widehat{f}(0) = 0$. Suppose $c_0 \in \mathbb{C}$ and $F : [0,2\pi] \to \mathbb{K}$ is given by $F(t) = c_0 + \int_0^t f(x)\,\mathrm{d}x$. Then $(\widehat{F}(k))_{k\in\mathbb{Z}} \in \ell^1$ and $F = \sum_{k\in\mathbb{Z}} \widehat{F}(k)e_k$ where the series is absolutely and uniformly convergent.*

Proof. Since $f \in L^2(0,2\pi) \subseteq L^1(0,2\pi)$, it follows from Example 6.6 that F is continuous on $[0,2\pi]$. Moreover, for every $t \in [0,2\pi]$,

$$|F(t)| \le |c_0| + \int_0^{2\pi} |f(x)|\,\mathrm{d}x \le |c_0| + \|f\|_1.$$

In particular, by monotonicity we see that $\|F\|_2 \le (2\pi)^{\frac{1}{2}}(|c_0| + \|f\|_1)$.

By Exercise 9.6, $\widehat{F}(k) = \frac{\widehat{f}(k)}{ik}$ for $k \ne 0$. By Theorem 9.7, $\|\widehat{f}(k))_{k\in\mathbb{Z}}\|_{\ell^2} = \|f\|_2$. Therefore, by the Cauchy–Schwarz inequality (9.4)

$$\sum_{k\in\mathbb{Z}} |\widehat{F}(k)| = |\widehat{F}(0)| + \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{|\widehat{f}(k)|}{|k|} \le |\widehat{F}(0)| + C\|(\widehat{f}(k))_{k\in\mathbb{Z}}\|_{\ell^2} \le |\widehat{F}(0)| + \frac{C}{(2\pi)^{\frac{1}{2}}}\|f\|_2,$$

where used (9.7) in the last step. Therefore, the absolute and uniform convergence follows from Theorem 9.10. ☺

Example 9.12. If $g \in C([0,2\pi])$ satisfies $g(0) = g(2\pi)$, g is piecewise continuously differentiable on $(0,2\pi)$ and $g' \in L^2(0,2\pi)$, then $g = \sum_{k\in\mathbb{Z}} \widehat{g}(k)e_k$ where the series is absolutely and uniformly convergent. Indeed, let $F = g - g(0)$. Then $F(t) = -g(0) + \int_0^t g'(x)\,\mathrm{d}x$, where $f := g'$ satisfies the assumptions of Corollary 9.11.

Exercises

Exercise 9.1. Let $f : [0,2\pi) \to \mathbb{R}$ be given by $f = \mathbf{1}_{[0,\pi]}$.

(a) Show that $\widehat{f}(k) = 0$ for even $k \ne 0$, $\widehat{f}(k) = \frac{1}{\pi i k}$ for odd k, and $\widehat{f}(0) = \frac{1}{2}$.

(b) Write f as a series of sines as in Example 9.9 and give an estimate of L^2-error given by (9.9).
(c) Evaluate $\sum_{n\in\mathbb{Z}} \frac{1}{(2n+1)^2}$.
Hint: Argue as in Example 9.9.

Exercise 9.2. Let $f : [0, 2\pi) \to \mathbb{R}$ be given by $f(x) = |x - \pi|$.

(a) Deduce from Example 9.12 that the Fourier series of f is absolutely and uniformly convergent.
(b) Calculate $\widehat{f}(k)$ for $k \in \mathbb{Z}$.
(c) Show that the L^2-error (9.9) is of order $O(\frac{1}{n^3})$ and uniform error (9.12) is of order $O(\frac{1}{n})$.
(d) Use Parseval's identity to show that

$$\sum_{n\in\mathbb{Z}} \frac{1}{(2n+1)^4} = \frac{\pi^4}{48}.$$

(e) Evaluate f and its Fourier series in $x = 0$ to obtain another way to calculate $\sum_{n\in\mathbb{Z}} \frac{1}{(2n+1)^2}$.

Exercise 9.3 (Special Fourier series and kernels)*.* The following kernel's play a central role in more advanced theory of Fourier series. Prove the identities below for $x \in (0, 2\pi)$. For each exercise you should use the geometric sum $\sum_{k=0}^{n} a^k = \frac{1-a^{n+1}}{1-a}$ for $a \in \mathbb{C} \setminus \{1\}$.

(a) (Dirichlet kernel): Show that

$$D_n(x) := \sum_{|k|\leq n-1} e_k(x) = \frac{\sin((n-\frac{1}{2})x)}{\sin(\frac{1}{2}x)}, \qquad n \geq 1.$$

(b) (Féjer kernel): Show that

$$F_n(x) := \frac{1}{n}\sum_{j=1}^{n} D_j(x) = \frac{1}{n}\frac{\sin^2(n\frac{x}{2})}{\sin^2(\frac{1}{2}x)}, \qquad n \geq 1$$

.

Exercise* 9.4 (Riemann-Lebesgue lemma)*.*

(a) Show that for any step function $f : [0, 2\pi] \to \mathbb{C}$ (see Section 8.3) one has $\lim_{|k|\to\infty} \widehat{f}(k) = 0$.
Hint: By linearity it suffices to consider $f = \mathbf{1}_{(a,b)}$, where $(a, b) \subseteq [0, 2\pi]$.
(b) Show that for any $f \in L^1(0, 2\pi)$ one has $\lim_{|k|\to\infty} \widehat{f}(k) = 0$.
Hint: Use Theorem 8.10 and ((a)).

Exercise* 9.5.

(a) Let $(\langle H, \langle\cdot,\cdot\rangle)$ be a Hilbert space (over the complex scalars). Prove that for all $u, v \in H$,

$$\text{(polarization)}\quad 4\langle u, v\rangle = \|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2.$$

(b) Use (a) and (9.7) to prove that for all $f, g \in L^2(0, 2\pi)$:

$$\int_0^{2\pi} f(x)\overline{g(x)}\,\mathrm{d}x = 2\pi \sum_{k\in\mathbb{Z}} \widehat{f}(k)\overline{\widehat{g}(k)}.$$

Exercise* 9.6. Let $g \in C^1([0, 2\pi])$[61] and $f \in L^1(0, 2\pi)$. Define $F : [0, 2\pi] \to \mathbb{C}$ by $F(t) = \int_0^t f(x)\,\mathrm{d}x$. By Example 6.6, F is continuous.

(a) Prove the following integration by parts formula:

$$\int_0^{2\pi} f(x)g(x)\,\mathrm{d}x = F(2\pi)g(2\pi) - F(0)g(0) - \int_0^{2\pi} F(x)g'(x)\,\mathrm{d}x.$$

Hint: For continuous f this is just the standard integration by parts formula. Use Corollary 8.11 and approximation to deduce the general case.
(b) Show that $\widehat{f}(k) = \widehat{f}(0) + ik\widehat{F}(k)$ for all $k \in \mathbb{Z} \setminus \{0\}$.
Hint: Apply (a) with $g = e_{-k}$.

[61]That means g is differentiable and its derivative is continuous on $[0, 2\pi]$

Exercise* 9.7. Assume $F : [0, 2\pi] \to \mathbb{C}$ is continuously differentiable and satisfies $F(0) = F(2\pi)$ and $\widehat{F}(0) = 0$. Let $f = F'$.

(a) Show that $\widehat{f}(k) = ik\widehat{F}(k)$ for all $k \in \mathbb{Z}$.
Hint: Apply Exercise 9.6(b).

(b) Show that $\|F\|_2 \leq \|f\|_2$ and that equality holds if and only if $F = c_1 e_1 + c_{-1} e_{-1}$ for $c_1, c_{-1} \in \mathbb{C}$.
Hint: Apply (9.7).

Exercise* 9.8. Consider $\Gamma = \{\frac{1}{2}\mathbf{1}_{[0,2\pi]}\} \cup \{\cos(n\cdot) : n \in \mathbb{N}\} \cup \{\sin(n\cdot) : n \in \mathbb{N}\} \subseteq L^2(0, 2\pi)$.

(a) Show that $\phi, \psi \in \Gamma$ with $\phi \neq \psi$ are orthogonal and $\|\phi\|_2 = \pi$.

(b) Let $f \in L^1(0, 2\pi)$ be such that $\langle f, \phi\rangle = 0$ for all $\phi \in \Gamma$. Show that $f = 0$.
Hint: Use Theorem 9.5

(c) Show that for every $(a_n)_{n\geq 0}, (b_n)_{n\geq 1} \in \ell^2$ the following series converges in $L^2(0, 2\pi)$.
$$g := \frac{a_0}{2} + \sum_{n\geq 1} a_n \cos(n\cdot) + \sum_{n\geq 1} b_n \sin(n\cdot)$$
Hint: Argue as in Theorem 9.7.

(d) Show that $\|g\|^2_{L^2(0,2\pi)} = \pi \sum_{n\geq 0} |a_n|^2 + \pi \sum_{n\geq 1} |b_n|^2$.
Hint: Argue as in Theorem 9.7.

(e) Show that $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 1}$ satisfy
$$a_n = \frac{1}{\pi}\int_0^{2\pi} \cos(nx)g(x)\,\mathrm{d}x, \;\; b_n = \frac{1}{\pi}\int_0^{2\pi} \sin(nx)g(x)\,\mathrm{d}x.$$
Hint: Argue as in Theorem 9.7.

(f) Show that every $f \in L^2(0, 2\pi)$ can be written as
$$f = \frac{a_0}{2} + \sum_{n\geq 1} a_n \cos(n\cdot) + b_n \sin(n\cdot)$$
with converges in $L^2(0, 2\pi)$.
Hint: Apply Theorem 9.7 or argue as in Theorem 9.7.

It follows from the previous exercise and (9.2) that for C^1-functions F one has $|\widehat{F}(k)| \leq \frac{\|f\|_1}{2\pi|k|}$ for all $k \in \mathbb{Z} \setminus \{0\}$. Moreover, in Exercise 9.4 we have seen that for general $F \in L^1(0, 2\pi)$ one has $\widehat{F}(k) \to 0$ as $|k| \to \infty$. In the next exercise we show that the convergence can be arbitrary slow even for periodic functions $F \in C([0, 2\pi])$.

Exercise** 9.9. Show that for any sequence $(c_k)_{k\geq 1}$ with $c_k \neq 0$ and $c_k \to 0$ there exists a function $F \in C([0, 2\pi])$ with $F(0) = F(2\pi)$ such that $|\widehat{F}(k)| \geq |c_k|$ for infinitely many $k \in \mathbb{N}$.
Hint: Choose a subsequence such that $\sum_{n=1}^{\infty} |c_{k_n}| < \infty$.

Finally we deduce Weierstrass' classical approximation result. We first need an elementary result about even trigonometric polynomials.

Exercise** 9.10. Let $P = \sum_{k=-n}^{n} c_k e_k$ be a trigonometric polynomial.

(a) Show that there exists a polynomial q_n of degree n such that $\cos(nx) = q_n(\cos(x))$.
Hint: Use induction and the recursion formula $\cos(kx)+\cos((k-2)x) = 2\cos((k-1)x)\cos(x)$.

(b) Show that there exists a polynomial q_n of degree n such that $\frac{\sin(nx)}{\sin(x)} = r_n(\cos(x))$.
Hint: Use induction and the recursion formula $\sin((k+1)x)+\sin((k-1)x) = 2\cos(kx)\sin(x)$.

(c) From (a) and (b) derive that there there exit polynomials $q, r : [-\pi, \pi] \to \mathbb{R}$ of degree n such that $P(x) = q(\cos(x)) + r(\cos(x))\sin(x)$.

(d) If additionally P is even (i.e. $P(-x) = P(x)$ for $x \in [-\pi, \pi]$). Show that there exists a polynomial q of degree n such that $P(x) = q(\cos(x))$.
Hint: Write $2P(x) = P(x) + P(-x)$ and use (c).

Exercise** 9.11 (Weierstrass' approximation theorem for continuous functions). Let $f : [-1, 1] \to \mathbb{R}$ be continuous and let $\varepsilon > 0$. Let $g : [-\pi, \pi] \to \mathbb{R}$ given by $g(x) = f(\cos(x))$.

(a) Show that there is an *even* trigonometric polynomial P such that $\|g - P\|_\infty < \varepsilon$.
Hint: First apply Theorem 9.4 on $[-\pi, \pi]$ to obtain a trigonometric polynomial P such that $\|g - P\|_\infty < \varepsilon$. Now consider $\widetilde{P}(x) = \frac{P(-x)+P(x)}{2}$.
(b) Use Exercise 9.10(d) to find a a polynomial q such that $\|f - q\|_\infty < \varepsilon$.
(c) The above shows that the polynomials are dense in $C([-1, 1])$. Use a scaling argument to show that the polynomials are dense in $C([a, b])$.

Final remarks

- One can show that $\|f - F_n * f\|_p \to 0$ for all $f \in L^p(0, 2\pi)$ with $p \in [1, \infty)$. In Theorem 9.4 the convergence holds uniformly if $f \in C([0, 2\pi])$ satisfies $f(0) = f(2\pi)$. Here $F_n * f$ is the so-called convolution product of F_n and f and is defined by $F_n * f(t) = \int_0^{2\pi} F_n(t-x) f(x)\,\mathrm{d}x$. Using the definition of F_n one can check that $F_n * f$ is a trigonometric polynomial.
- Similar results for D_n are true as well as long as $p \in (1, \infty)$, but this a much deeper result. Since $D_n * f = s_n(f)$, this implies that $\|f - s_n(f)\|_p \to 0$ for all $f \in L^p(0, 2\pi)$ with $p \in (1, \infty)$.
- Finally, we note that $s_n(f) \to f$ a.e. for any $f \in L^p(0, 2\pi)$ with $p > 1$. This is one of the deepest result in the theory of Fourier series and was proved by Carleson for $p = 2$ and extended to $p > 1$ by Hunt in 1968. It was proved a long time before that the result fails for $p = 1$ by Kolmogorov in 1923.

`For details we refer to the elective Bachelor course on Fourier analysis!`

Appendix A. Dynkin's lemma

The results of this section are not part of the exam material. We will proof the uniqueness result of Proposition 3.5. In this section S denotes a set.

Definition A.1 (π-system)**.** *A collection $\mathcal{E} \subseteq \mathcal{P}(S)$ is called a* ***π-system*** *if for all $A, B \in \mathcal{E}$ one has $A \cap B \in \mathcal{E}$.*

Example A.2. Every ring is a π-system.

Definition A.3. *A collection $\mathcal{D} \subseteq \mathcal{P}(S)$ is called a* ***Dynkin-system***[62] *if the following conditions hold:*
(D1) $S \in \mathcal{D}$;
(D2) $A, B \in \mathcal{D}$ and $A \subseteq B$ implies $B \setminus A \in \mathcal{D}$;
(D3) If $(A_n)_{n\geq 1}$ in $\mathcal{D}$ and $A_n \uparrow A$[63], then $A \in \mathcal{D}$.

Example A.4. Let $S = \{1, 2, 3, 4\}$ and $\mathcal{D} = \{\varnothing, S, \{1, 2\}, \{3, 4\}, \{1, 3\}, \{2, 4\}\}$. Then $\mathcal{D}$ is a Dynkin system, but it is not a π-system.

Proposition A.5. *For a collection $\mathcal{F} \subseteq \mathcal{P}(S)$ the following are equivalent:*
(i) $\mathcal{F}$ is a σ-algebra;
(ii) $\mathcal{F}$ is a Dynkin system and a π-system.

Proof. (i)⇒(ii): This is Exercise A.1.

(ii)⇒(i): $\varnothing, S \in \mathcal{F}$ follows from (D1) and (D2) of the definition of a Dynkin system. Let $(A_n)_{n\geq 1}$ be a sequence in $\mathcal{F}$. Let $A = \bigcup_{j=1}^{\infty} A_j$ and $B_n = \bigcup_{j=1}^{n} A_j$ for $n \geq 1$. Since $\mathcal{F}$ is a π-system it is closed under finite intersections. Therefore, using (D2) we obtain $B_n = \left(\bigcap_{j=1}^{n} A_j^c\right)^c \in \mathcal{F}$. Since $B_n \uparrow A$, it follows from (D3) that $A \in \mathcal{F}$. ☺

Lemma A.6 (Dynkin)**.** *Let $\mathcal{E} \subseteq \mathcal{P}(S)$ be a π-system and $\mathcal{D} \subseteq \mathcal{P}(S)$ be a Dynkin system. If $\mathcal{E} \subseteq \mathcal{D}$, then $\sigma(\mathcal{E}) \subseteq \mathcal{D}$.*

Proof. Let $\mathcal{D}_0$ denote the intersection of all Dynkin system which contain $\mathcal{E}$. Observation: $\mathcal{E} \subseteq \mathcal{D}_0 \subseteq \mathcal{D}$ and D_0 is a Dynkin system (see Exercise A.2). We claim that $\mathcal{D}_0$ is a π-system as well. As soon as we have proved this claim, Proposition A.5 yields that $\mathcal{D}_0$ is a σ-algebra. Therefore, from the observation it follows that $\sigma(\mathcal{E}) \subseteq \mathcal{D}_0 \subseteq \mathcal{D}$. To prove the claim we need two steps.

Step 1: Define a new collection by

$$\mathcal{D}_1 = \{D \in \mathcal{D}_0 : D \cap E \in \mathcal{D}_0 \text{ for each } E \in \mathcal{E}\}.$$

Since $\mathcal{E}$ is a π-system also $\mathcal{E} \subseteq \mathcal{D}_1$. The collection $\mathcal{D}_1$ is a Dynkin-system again. Indeed, $S \in \mathcal{D}_1$ is clear. If $A, B \in \mathcal{D}_1$ and $A \subseteq B$, then for each $E \in \mathcal{E}$ we find $(B \setminus A) \cap E = (B \cap E) \setminus (A \cap E) \in \mathcal{D}_0$, because $A \cap E, B \cap E \in \mathcal{D}_0$ and $\mathcal{D}_0$ is a Dynkin system, and thus $B \setminus A \in \mathcal{D}_1$. Next let $(A_n)_{n\geq 1}$ in $\mathcal{D}_1$ with $A_n \uparrow A$. Then for each $E \in \mathcal{E}$, $A \cap E = \bigcup_{n=1}^{\infty} (A_n \cap E) \in \mathcal{D}_0$ since $A_n \cap E \in \mathcal{D}_0$. Since $\mathcal{D}_0 \subseteq \mathcal{D}_1$ and $\mathcal{D}_1$ is a Dynkin system which contains $\mathcal{E}$, we find $\mathcal{D}_1 = \mathcal{D}_0$.

Step 2: Define a new collection by

$$\mathcal{D}_2 = \{D \in \mathcal{D}_0 : D \cap C \in \mathcal{D}_0 \text{ for each } C \in \mathcal{D}_0\}.$$

Since $\mathcal{D}_1 = \mathcal{D}_0$, we find that $\mathcal{E} \subseteq \mathcal{D}_2$. As before one checks that $\mathcal{D}_2$ is a Dynkin system. Moreover, as before this yields $\mathcal{D}_2 = \mathcal{D}_0$. This proves the claim. ☻

Proposition A.7 (Uniqueness)**.** *Let μ_1 and μ_2 both be measures on measurable space $(S, \mathcal{A})$. Assume the following conditions:*
(i) $\mathcal{E} \subseteq \mathcal{A}$ is a π-system with $\sigma(\mathcal{E}) = \mathcal{A}$;

---

[62] Eugene Dynkin 1924–2014 was a Russian mathematician who worked on Algebra and Probability theory.
[63] See Definition 2.8 for the meaning of $A_n \uparrow A$

(ii) $\mu_1(S) = \mu_2(S) < \infty$ *and* $\mu_1(E) = \mu_2(E)$ *for all* $E \in \mathcal{E}$.

Then $\mu_1 = \mu_2$ *on* $\mathcal{A}$.

Proof. Let $\mathcal{D} = \{A \in \mathcal{A} : \mu_1(A) = \mu_2(A)\}$. Then $\mathcal{E} \subseteq \mathcal{D}$. We claim that $\mathcal{D}$ is a Dynkin system. From the claim and Lemma A.6 it follows that $\mathcal{A} = \sigma(\mathcal{E}) \subseteq \mathcal{D} \subseteq \mathcal{A}$. This implies $\mathcal{D} = \mathcal{A}$ and the required result follows from the Definition of $\mathcal{D}$.

To prove the claim note that $S \in \mathcal{D}$ by assumption. If $A, B \in \mathcal{D}$ with $A \subseteq B$, then $\mu_1(B \setminus A) = \mu_1(B) - \mu_1(A) = \mu_2(B) - \mu_2(A) = \mu_2(B \setminus A)$ and hence $B \setminus A \in \mathcal{D}$. Finally, if $(A_n)_{n \geq 1}$ in $\mathcal{D}$ with $A_n \uparrow A$, then by Theorem 2.9, $\mu_j(A_n) \uparrow \mu_j(A)$ for $j = 0, 1$. Since $\mu_1(A_n) = \mu_2(A_n)$, this yields $\mu_1(A) = \mu_2(A)$ and thus $A \in \mathcal{D}$. ☺

Exercises

Exercise A.1. Prove Proposition A.5 (i) ⇒ (ii).

Exercise* A.2. Prove that the intersection of Dynkin systems is again a Dynkin system.

Exercise* A.3. Find a version of Proposition A.7 which for measures with $\mu_1(S) = \mu_2(S) = \infty$.
Hint: See the proof of Theorem 3.9.

Appendix B. Carathéodory's extension theorem

The results of this section are not part of the exam material, although the statement of the Carathéodory's extension in Theorem 3.1 is part of the exam.

Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring and $\mu : \mathcal{R} \to [0,\infty]$ be σ-additive on $\mathcal{R}$ and $\mu(\varnothing) = 0$, for example $\mathcal{R} = \mathcal{F}^d$ and $\mu = \lambda$. Our goal is to extend μ to a measure $\overline{\mu}$ defined on (at least) $\sigma(\mathcal{R})$. Since $\overline{\mu}$ should be a measure, it should be σ-subadditive, i.e. for $A \in \sigma(\mathcal{R})$ and any choice of $B_1, B_2 \ldots \in \mathcal{R}$ with $A \subseteq \bigcup_{n=1}^\infty B_n$ we should have

$$\overline{\mu}(A) \le \sum_{n=1}^\infty \overline{\mu}(B_j) = \sum_{n=1}^\infty \mu(B_j).$$

Note that the right-hand side of this inequality is defined purely in terms of μ. Therefore, as an initial guess, we define $\overline{\mu}\colon \mathcal{P}(S) \to [0,\infty]$ by

$$\overline{\mu}(A) := \inf\Big\{ \sum_{j=1}^\infty \mu(B_j) : A \subseteq \bigcup_{j=1}^\infty B_j \text{ with } \ B_n \in \mathcal{R} \text{ for } n \ge 1 \Big\}, \tag{B.1}$$

where we set $\overline{\mu}(A) = \infty$ if the set on the right-hand side is empty.

However, this is a little too optimistic, as $\overline{\mu}$ is not necessarily additive (see Exercise B.1) and therefore not a measure. We do have σ-subadditivity:

Lemma B.1. *Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring, $\mu : \mathcal{R} \to [0,\infty]$ be σ-additive on $\mathcal{R}$ and $\mu(\varnothing) = 0$. Then $\overline{\mu}$ is σ-subadditive on $\mathcal{P}(S)$, i.e. for $A_1, A_2, \ldots \in \mathcal{P}(S)$ we have*

$$\overline{\mu}\Big(\bigcup_{n=1}^\infty A_n\Big) \le \sum_{n=1}^\infty \overline{\mu}(A_n).$$

Proof. If $\overline{\mu}(A_n) = \infty$ for some $n \ge 1$, then the statement is trivial. Assume $\overline{\mu}(A_n) < \infty$ for all $n \ge 1$ and let $\varepsilon > 0$. Then, by definition of $\overline{\mu}$, for each fixed $n \ge 1$ we can find $B_{n,j} \in \mathcal{R}$ such that

$$A_n \subseteq \bigcup_{j=1}^\infty B_{n,j} \qquad \text{and} \qquad \overline{\mu}(A_n) + 2^{-n}\varepsilon \ge \sum_{j=1}^\infty \mu(B_{n,j}).$$

Then $\bigcup\limits_{n=1}^\infty A_n \subseteq \bigcup\limits_{n,j=1}^\infty B_{n,j}$ and by the definition of $\overline{\mu}$ and the σ-additivity of μ on $\mathcal{R}$ we find

$$\overline{\mu}\Big(\bigcup_{n=1}^\infty A_n\Big) \le \sum_{n,j=1}^\infty \mu(B_{n,j}) \le \sum_{n=1}^\infty \overline{\mu}(A_n) + 2^{-n}\varepsilon = \varepsilon + \sum_{n=1}^\infty \overline{\mu}(A_n).$$

Since $\varepsilon > 0$ was arbitrary, this finishes the proof. ☺

Next to σ-subadditivity, we trivially have that $\overline{\mu}(\varnothing) = 0$ and, if $A \subseteq B$, then $\overline{\mu}(A) \le \overline{\mu}(B)$. Mappings $\alpha\colon \mathcal{P}(S) \to [0,\infty]$ with these three properties are called outer measures.

Definition B.2. *Let S be a set. A function $\alpha : \mathcal{P}(S) \to [0,\infty]$ is called an **outer measure** if*

(i) $\alpha(\varnothing) = 0$*;*
(ii) (monotonicity) $A \subseteq B \implies \alpha(A) \le \alpha(B)$*;*
(iii) (σ-subadditivity) For $A_1, A_2, \ldots, \mathcal{P}(S)$ one has $\alpha\Big(\bigcup\limits_{n=1}^\infty A_n\Big) \le \sum\limits_{n=1}^\infty \alpha(A_n)$.

We would like to make any outer measure α (and therefore $\overline{\mu}$) σ-additive. To do so, we will restrict ourselves to a subset $\mathcal{M}_\alpha$ of $\mathcal{P}(S)$ where α is σ-additive and hope that we have $\sigma(R) \subseteq \mathcal{M}_\alpha$ for $\alpha = \overline{\mu}$. Define

$$\mathcal{M}_\alpha := \big\{A \in \mathcal{P}(S) : \alpha(Q) = \alpha(Q \cap A) + \alpha(Q \cap A^c) \text{ for all } Q \in \mathcal{P}(S)\big\}$$

and note that for $A, B \in \mathcal{M}_\alpha$ disjoint, we have

$$\alpha(A \cup B) = \alpha((A \cup B) \cap A) + \alpha((A \cup B) \cap A^c) = \alpha(A) + \alpha(B),$$

i.e. α is additive on $\mathcal{M}_\alpha$. It turns out that $\mathcal{M}_\alpha$ is a σ-algebra and we also have σ-additivity of α on $\mathcal{M}_\alpha$.

Lemma B.3. *Let S be a set and let $\alpha : \mathcal{P}(S) \to [0,\infty]$ be an outer measure. Then $\mathcal{M}_\alpha$ is a σ-algebra and α is a measure on $(S, \mathcal{M}_\alpha)$.*

Proof. We note that by definition $S, \varnothing \in \mathcal{M}_\alpha$ and if $A \in \mathcal{M}_\alpha$, then $A^c \in \mathcal{M}_\alpha$. Thus, to prove the proposition, we only need to show the following properties:

- Property (iii) of a σ-algebra, i.e. for $A_1, A_2, \ldots \in \mathcal{M}_\alpha$ we have $\bigcup_{n=1}^\infty A_n \in \mathcal{M}_\alpha$. As we have seen in Exercise 1.3, it suffices to check this for disjoint sets.
- σ-additivity of α on $\mathcal{M}_\alpha$, i.e. for disjoint $A_1, A_2, \ldots \in \mathcal{M}_\alpha$ we have

$$\alpha\Big(\bigcup_{n=1}^\infty A_n\Big) = \sum_{n=1}^\infty \alpha(A_n).$$

We proceed in several steps.

Step 1: Let $A, B \in \mathcal{M}_\alpha$. We will show that $A \cap B \in \mathcal{M}_\alpha$, from which $A \cup B \in \mathcal{M}_\alpha$ follows by taking complements. Set $C := A \cap B$ and observe that $A \cap B^c = C^c \cap A$ and $A^c = C^c \cap A^c$. Therefore, for $Q \in \mathcal{P}(S)$ one has

$$\begin{aligned}
\alpha(Q) &= \alpha(Q\cap A) + \alpha(Q \cap A^c) && \text{(since } A \in \mathcal{M}_\alpha) \\
&= \alpha(Q\cap A \cap B) + \alpha(Q\cap A\cap B^c) + \alpha(Q \cap A^c) && \text{(since } B \in \mathcal{M}_\alpha) \\
&= \alpha(Q\cap C) + \alpha(Q\cap C^c \cap A) + \alpha(Q \cap C^c \cap A^c) && \text{(by the observation)} \\
&= \alpha(Q\cap C) + \alpha(Q \cap C^c) && \text{(since } A \in \mathcal{M}_\alpha).
\end{aligned}$$

Therefore, $A \cap B = C \in \mathcal{M}_\alpha$. By induction, we also obtain that finite unions and intersections of sets in $\mathcal{M}_\alpha$ are in $\mathcal{M}_\alpha$.

Step 2: Let $A, B \in \mathcal{M}_\alpha$ be disjoint. For any $Q \in \mathcal{P}(S)$ we have, using the definition of $A \in \mathcal{M}_\alpha$, that

$$\alpha\big(Q \cap (A \cup B)\big) = \alpha(Q \cap A) + \alpha(Q \cap B \cap A^c) = \alpha(Q\cap A) + \alpha(Q \cap B).$$

By induction, this identity also hold for finite unions of disjoint sets in $\mathcal{M}_\alpha$.

Step 3: To complete the proof, we will now show that for disjoint $A_1, A_2, \ldots \in \mathcal{M}_\alpha$ we have

$$A := \bigcup_{n=1}^\infty A_n \in \mathcal{M}_\alpha \qquad \text{and} \qquad \alpha(A) = \sum_{n=1}^\infty \alpha(A_n).$$

Fix an arbitrary $Q \in \mathcal{P}(S)$. By Lemma B.1, using only two sets, we have

$$\alpha(Q) \le \alpha(Q \cap A) + \alpha(Q \cap A^c).$$

For the converse inequality let $B_n = \bigcup_{j=1}^n A_j$ for $n \ge 1$. Note that $B_n \in \mathcal{M}_\alpha$ by Step 1. Therefore

$$\begin{aligned}
\text{(B.2)} \qquad \alpha(Q\cap A) + \alpha(Q\cap A^c) &\le \sum_{j=1}^\infty \alpha(Q \cap A_j) + \alpha(Q \cap A^c) && \text{(by Step 3)} \\
&= \lim_{n\to\infty} \sum_{j=1}^n \alpha(Q\cap A_j) + \alpha(Q\cap A^c) \\
&= \lim_{n\to\infty} \alpha(Q \cap B_n) + \alpha(Q \cap A^c) && \text{(by Step 2)} \\
&\le \lim_{n\to\infty} \alpha(Q\cap B_n) + \alpha(Q \cap B_n^c) && \text{(since } A^c \subseteq B_n^c) \\
&= \lim_{n\to\infty} \alpha(Q) = \alpha(Q) && \text{(since } B_n \in \mathcal{M}_\alpha).
\end{aligned}$$

Therefore, we have

$$\alpha(Q) = \alpha(Q\cap A) + \alpha(Q \cap A^c),$$

and thus $A \in \mathcal{M}_\alpha$. Moreover, the inequality in (B.2) is actually an equality[64], which for $Q = A$ yields

$$\alpha(A) = \sum_{n=1}^{\infty} \alpha(A_n).$$

finishing the proof. ☺

Applying Lemma B.3 to the outer measure $\overline{\mu}$ defined in (B.1), we have shown that $\mathcal{M}_{\overline{\mu}}$ is a σ-algebra and $\overline{\mu}\colon \mathcal{M}_{\overline{\mu}} \to [0,\infty]$ is a measure. Thus, to prove Theorem 3.1, it remains to show that $\sigma(\mathcal{R}) \subseteq \mathcal{M}_{\overline{\mu}}$ and $\overline{\mu}(A) = \mu(A)$ for all $A \in \mathcal{R}$.

Theorem B.4 (Carathéodory's extension theorem)**.** *Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring, $\mu : \mathcal{R} \to [0,\infty]$ be σ-additive and $\mu(\varnothing) = 0$. Then $\overline{\mu}$ defined in (B.1) satisfies the following properties:*

(i) $\mathcal{M}_{\overline{\mu}}$ is a σ-algebra and $\overline{\mu}\colon \mathcal{M}_{\overline{\mu}} \to [0,\infty]$ is a measure;
(ii) $\sigma(\mathcal{R}) \subseteq \mathcal{M}_{\overline{\mu}}$;
(iii) $\overline{\mu}(A) = \mu(A)$ for all $A \in \mathcal{R}$.

In particular, the restriction $\overline{\mu}\colon \sigma(\mathcal{R}) \to [0,\infty]$ is a measure.

Clearly, Theorem 3.1 follows from the above statement and actually shows that there is a further extension to the possibly larger σ-algebra $\mathcal{M}_{\overline{\mu}}$.

Proof. (i) follows by combining Lemma B.1 and Lemma B.3. It remains to prove (ii) and (iii). The final assertion follows from (i)-(iii), as the restriction of a measure to a smaller σ-algebra is a measure again.

Proof of (ii): Let $A \in \mathcal{R}$ and $Q \in \mathcal{P}(S)$. By the σ-subadditivity of $\overline{\mu}$ shown in Lemma B.1, we immediately get $\overline{\mu}(Q) \le \overline{\mu}(Q \cap A) + \overline{\mu}(Q \cap A^c)$. For the converse estimate the case $\overline{\mu}(Q) = \infty$ is trivial. In case $\overline{\mu}(Q) < \infty$, choose $B_1, B_2, \ldots, \in \mathcal{R}$ such that $Q \subseteq \bigcup\limits_{n=1}^{\infty} B_n$ and $\sum_{n=1}^{\infty} \mu(B_n) < \infty$. Then $B_n \cap A, B_n \cap A^c \in \mathcal{R}$ for all $n \ge 1$ and

$$Q \cap A \subseteq \bigcup_{n=1}^{\infty} B_n \cap A \quad \text{and} \quad Q \cap A^c \subseteq \bigcup_{n=1}^{\infty} B_n \cap A^c.$$

Therefore, using first the definition of $\overline{\mu}$ and then the additivity of μ on $\mathcal{R}$, we find

$$\overline{\mu}(Q \cap A) + \overline{\mu}(Q \cap A^c) \le \sum_{n=1}^{\infty} \mu(B_n \cap A) + \sum_{n=1}^{\infty} \mu(B_n \cap A^c) = \sum_{n=1}^{\infty} \mu(B_n).$$

Taking the infimum over all $(B_n)_{n\ge1}$ as above gives $\overline{\mu}(Q \cap A) + \overline{\mu}(Q \cap A^c) \le \overline{\mu}(Q)$. Combining both estimates we can conclude $A \in \mathcal{M}_{\overline{\mu}}$. Since $\mathcal{M}_{\overline{\mu}}$ is a σ-algebra, it follows that $\sigma(\mathcal{R}) \subseteq \mathcal{M}_{\overline{\mu}}$.

Proof of (iii): Let $A \in \mathcal{R}$. It is clear that $\overline{\mu}(A) \le \mu(A)$. Indeed, take $B_1 = A$ and $B_n = \varnothing$ for $n \ge 2$ in B.1. For the converse estimate the case $\overline{\mu}(A) = \infty$ is clear. Now suppose $\overline{\mu}(A) < \infty$ and let $B_1, B_2, \ldots, \in \mathcal{R}$ be such that $A \subseteq \bigcup\limits_{n=1}^{\infty} B_n$. Then by the σ-additivity of μ on $\mathcal{R}$ and Proposition 2.3 applied to $A = \bigcup\limits_{n=1}^{\infty} A \cap B_n$, we find

$$\mu(A) \le \sum_{n=1}^{\infty} \mu(A \cap B_n) \le \sum_{n=1}^{\infty} \mu(B_n).$$

Taking the infimum over all $(B_n)_{n\ge1}$ as above yields $\mu(A) \le \overline{\mu}(A)$. ☺

Exercises

Exercise B.1. We will show that $\mathcal{M}_{\overline{\mu}} \ne \mathcal{P}(S)$ in general.[65] Let $S = \{1,2,3\}$ and define a σ-algebra by $\mathcal{A} = \{\varnothing, S, \{1,2\}, \{3\}\}$. Assume μ is a measure satisfying $\mu(\{1,2\}) = \mu(\{3\}) = \frac{1}{2}$.

[64] Clearly, $x \le y \le z \le x$ enforces $x = y = z$.

[65] For the Lebesgue measure one also has $\mathcal{M}_{\overline{\mu}} \ne \mathcal{P}(\mathbb{R})$, but this is much harder to prove. See Appendix C.

(a) Show that $\overline{\mu}(\{1\}) = \overline{\mu}(\{2\}) = \frac{1}{2}$.
(b) Show that $\{1\}, \{2\} \notin \mathcal{M}_{\overline{\mu}}$.

Exercise B.2. Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring, $\mu : \mathcal{R} \to [0, \infty]$ be σ-additive on $\mathcal{R}$ and $\mu(\varnothing) = 0$. Suppose that $A \subseteq \mathcal{P}(S)$ satisfies $\overline{\mu}(A) = 0$. Show that $A \in \mathcal{M}_{\overline{\mu}}$.

Exercise* B.3. Assume the conditions of Theorem B.4 and assume μ is σ-finite on $\mathcal{R}$, i.e. there exists a sequence $(A_n)_{n\geq 1}$ in $\mathcal{R}$ such that $A_n \uparrow S$ and $\mu(A_n) < \infty$ for all $n \geq 1$. Prove that the following are equivalent:
(a) $A \in \mathcal{M}_{\overline{\mu}}$;
(b) There exists a $B \in \sigma(\mathcal{R})$ such that $A \subseteq B$ and $\overline{\mu}(B \setminus A) = 0$.
Hint: First reduce to the case of a finite measure by intersecting with A_n.

Appendix C. Non-measurable sets

Let λ be the Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$. Let $\overline{\lambda} : \mathcal{P}(S) \to [0, \infty]$ be the outer measure associated with λ defined in (B.1). The σ-algebra $\mathcal{M}(\mathbb{R}^d) := \mathcal{M}_{\overline{\lambda}}$ is usually called the **Lebesgue σ-algebra**. It follows from Theorem B.4 that $\mathcal{F}^d \subseteq \mathcal{M}(\mathbb{R}^d)$ and thus also $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{M}(\mathbb{R}^d)$ and $\overline{\lambda}$ is a measure on $\mathcal{M}(\mathbb{R}^d)$. In the sequel we write λ again for this measure as it is just an extension of λ.

By Exercise B.3, for every $A \in \mathcal{M}(\mathbb{R}^d)$ there exists a $B \in \mathcal{B}(\mathbb{R}^d)$ such that $A \subseteq B$ and $\lambda(B \setminus A) = 0$. This shows that the Lebesgue σ-algebra is almost the same as the Borel σ-algebra up to sets of measure zero. Some strange things can happen with nonmeasurable sets.

The balls of Banach and Tarski

One can cut a ball of radius one in $\mathbb{R}^d$ in such a way that it can be used to form two balls of radius one. Of course something has to be nonmeasurable there. See:

https://en.wikipedia.org/wiki/Banach%E2%80%93Tarski_paradox

A set which Lebesgue measurable but not Borel measurable

There exist a set $A \in \mathcal{M}(\mathbb{R}^d)$ with $\lambda(A) = 0$, but $A \notin \mathcal{B}(\mathbb{R}^d)$. See [3, Appendix C] and [11, page 53]

http://onlinelibrary.wiley.com/doi/10.1002/9781118032732.app3/pdf

http://www.math3ma.com/mathema/2015/8/9/lebesgue-but-not-borel

You can also read about this in the Bachelor thesis of Gerrit Vos:

http://resolver.tudelft.nl/uuid:30d69b56-b846-435e-9d44-6a31b840a836

There exist sets which are not Lebesgue measurable A subset of $\mathbb{R}$ which is not in $\mathcal{M}(\mathbb{R}^d)$ is given by Vitali's example (see for example [4, Theorem 16.31]):

https://en.wikipedia.org/wiki/Vitali_set

References

[1] C. D. Aliprantis and O. Burkinshaw. *Principles of real analysis*. Academic Press, Inc., San Diego, CA, third edition, 1998.

[2] S. Banach. *Théorie des opérations linéaires*. Éditions Jacques Gabay, Sceaux, 1993. Reprint of the 1932 original.

[3] F. Burk. *Lebesgue measure and integration*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1998. An introduction, A Wiley-Interscience Publication.

[4] N. L. Carothers. *Real analysis*. Cambridge University Press, Cambridge, 2000.

[5] D. H. Fremlin. *Measure theory. Vol. 1, 2, 3, 4*. Torres Fremlin, Colchester, 2004-2006.

[6] D.J.H. Garling. *Inequalities: a journey into linear analysis*. Cambridge University Press, Cambridge, 2007.

[7] R. Haberman. *Applied Partial Differential Equations with Fourier Series and Boundary Valve Problems*. Pearson Higher Ed, 2012.

[8] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[9] Y. Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004.

[10] J Korevaar. Fourier analysis and related topics. https://staff.fnwi.uva.nl/j.korevaar/Foubook.pdf.

[11] H. Royden and P. Fitzpatrick. *Real analysis. 4th ed.* New York, NY: Prentice Hall, 4th ed. edition, 2010.

[12] E.M. Stein and R. Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.

[13] E.M. Stein and R. Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Princeton University Press, Princeton, NJ, 2005. Measure theory, integration, and Hilbert spaces.

[14] E.M. Stein and R. Shakarchi. *Functional analysis*, volume 4 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2011. Introduction to further topics in analysis.

[15] D. Werner. *Einführung in die höhere Analysis*. Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, corrected edition, 2009.

[16] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[17] A. Zygmund. The role of Fourier series in the development of analysis. In *Proceedings of the American Academy Workshop on the Evolution of Modern Mathematics (Boston, Mass., 1974)*, volume 2, pages 591–594, 1975.

[18] A. Zygmund. *Trigonometric series. Vol. I, II*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.