MEASURE AND INTEGRATION

Lecture notes by Mark Veraar

Contents

Introduction 2
1. σ-algebras 3
2. Measures 6
3. Construction of measures 9
4. Measurable functions 13
5. Construction of the integral 18
5.1. Integral for simple functions 18
5.2. Integral for positive measurable functions 19
5.3. Integral for measurable functions 21
6. Convergence theorems and applications 25
6.1. The three main convergence results 25
6.2. Consequences and applications 26
7. L^p-spaces 30
7.1. Minkowski and Hölder's inequalities 30
7.2. Completeness of L^p 31
7.3. L^p-spaces on intervals 32
8. Applications to Fourier series 36
8.1. Fourier coefficients 36
8.2. Weierstrass' approximation result and uniqueness 37
8.3. Fourier series in $L^2(0, 2\pi)$ 39
8.4. Fourier series in $C([0, 2\pi])$ 41
Appendix A. Dynkin's lemma 46
Appendix B. Carathéodory's extension theorem 48
Appendix C. Non-measurable sets 51
References 52

Date: October 2, 2019.

Introduction

These notes have been created for the "Measure and integration theory" part of a course on real analysis at the TU Delft. Together with the first part of the course on metric spaces, these notes form the mathematical basis for several bachelor courses and master courses in applied mathematics at TU Delft.

I would like to thank Emiel Lorist and Bas Nieraeth for their support in the preparation of these notes. All the figures have been created by Emiel Lorist. I 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], [16] and [17]. A very complete treatment of measure theory is given in the impressive works [5].

In Sections 5, 6, 7 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], [15] for detailed treatments.

In Section 8 we give a brief introduction to the theory of Fourier series. More thorough treatments can be found in for example [9], [10], [13], [14] and [19]. A full course on Fourier Analysis is offered as 3rd elective course based on the lecture notes [10]. The theory of Fourier series will be used in the 2nd 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 [18]:

"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":

- The Riesz-Fischer theorems 7.5, 8.7, in its initial formulation primarily a theorem on Fourier series.
- The M. Riesz interpolation theorem.
- Structure of sets of measure 0.
- The theory of trigonometric series has become a workshop of new methods in analysis, a place where new methods are first discovered before they are generalized and applied in other contexts.
- The theory of Fourier series gave a fresh impulse to problems of the differentiability of functions (Sobolev spaces etc.).

1. σ-ALGEBRAS

For a set S we write $\mathcal{P}(S)$ for its power set.

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) $A, B \in \mathcal{R} \Longrightarrow B \setminus A \in \mathcal{R}$;
(iii) $n \in \mathbb{N}$, $A_1, A_2, \ldots, A_n \in \mathcal{R} \Longrightarrow \bigcup_{j=1}^{n} A_j \in \mathcal{R}$.

Remark 1.2.

(1) An equivalent definition is obtained if one replaces (iii) by $A, B \in \mathcal{R} \Longrightarrow A \cup B \in \mathcal{R}$. This follows by induction.
(2) If $\mathcal{R}$ is a ring, then for all $A, B \in \mathcal{R}$ one has $\mathcal{A} \cap B \in \mathcal{R}$. Indeed, this follows from the identity $A \cap B = A \setminus (A \setminus B)$.

Definition 1.3. *Let S be a set. A family $\mathcal{A} \subseteq \mathcal{P}(S)$ is called a **σ-algebra**[1] if*

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

*The sets $A \in \mathcal{A}$ are often called the **measurable subsets** of S.*

Remark 1.4.

(1) If $\mathcal{A}$ is a σ-algebra, then for all $A_1, A_2, \ldots, \in \mathcal{A}$ one has $\bigcap_{j=1}^{\infty} A_j \in \mathcal{A}$. This follows from the identity $\bigcap_{j=1}^{\infty} A_j = \Big(\bigcup_{j=1}^{\infty} A_j^c\Big)^c$.
(2) Every σ-algebra is a ring. This follows from the identity $B \setminus A = B \cap A^c$.

Example 1.5. 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 1.6. Let $S = \{1, 2, 3\}$.

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

Example 1.7. Let S be a set.

(a) The set $\mathcal{R} = \{A \subseteq S : A \text{ is finite}\}$ is a ring.
(b) Let[2] $\mathcal{A} = \{A \subseteq S : A \text{ is countable or } A^c \text{ is countable}\}$. Then $\mathcal{A}$ is a σ-algebra (see Exercise 1.2). It is called the countable-cocountable σ-algebra and is useful for counterexamples.

We continue with a more serious example which plays a crucial role in later constructions.

Example 1.8.

(a) Let $S = \mathbb{R}$. Let $\mathscr{I}^1$ be the collection of all sets of the form $(a, b]$ with $a \leq b$. These intervals will be called **half-open intervals**. Then $\mathscr{I}^1$ is not a ring since for instance $(0, 3] \setminus (1, 2] = (0, 1] \cup (2, 3]$ is not in $\mathscr{I}^1$.
(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 (thus of the form $(a, b]$ with $a \leq b$). Then $\mathcal{F}^1$ is not a σ-algebra for several reasons.[3] We 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).

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

[2] Recall that a set $A \subseteq S$ is countable if A is finite or there is a bijection $f : \mathbb{N} \to A$.

[3] For instance $\bigcup_{n=1}^{\infty} (0, 1 - \frac{1}{n}] = (0, 1)$ is not in $\mathcal{F}^1$.

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 $A = \bigcup\limits_{i=1}^{m} (a_i, b_i]$ and $B = \bigcup\limits_{j=1}^{n} (c_j, d_j]$ are in $\mathcal{F}^1$, then

$$B \setminus A = \bigcup_{j=1}^{n} (c_j, d_j] \setminus A = \bigcup_{j=1}^{n} \bigcap_{i=1}^{m} (c_j, d_j] \setminus (a_i, b_i].$$

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 $j \in \{1, \ldots, k\}$ the half-open rectangles are given by

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

Let $\mathcal{F}^d$ be the collection of sets which can be written as a finite unions of half-open rectangles. Then $\mathcal{F}^d$ is a ring (see Exercise 1.7).

The proof of the next result is Exercise 1.3.

Proposition 1.9 (Intersection of σ-algebras)**.** *Suppose that $\mathcal{A}_i$ is a σ-algebra on S for every $i \in \mathcal{I}$. Then $\bigcap\limits_{i \in \mathcal{I}} \mathcal{A}_i$ is a σ-algebra.*

Definition 1.10. *Let S be a set and let $\mathcal{F} \subseteq \mathcal{P}(S)$. We write $\sigma(\mathcal{F})$ for the smallest σ-algebra which contains $\mathcal{F}$. Then $\sigma(\mathcal{F})$ is called the σ-algebra* ***generated*** *by $\mathcal{F}$. More precisely:*[4]

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

Example 1.11. Let $S = \{1, 2, 3\}$, $\mathcal{F} = \{\{1, 2\}\}$ and $\mathcal{G} = \{\{2, 3\}, \{1, 2\}\}$.

(a) $\sigma(\mathcal{F}) = \{\varnothing, S, \{1, 2\}, \{3\}\}$.

(b) $\sigma(\mathcal{G}) = \mathcal{P}(S)$. 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 $\sigma(\mathcal{G})$. Therefore, the required result follows since we can form every subset of S by taking suitable finite unions.

Example 1.12. Let $S = \mathbb{N}$ and $\mathcal{F} = \{\{n\} : n \in \mathbb{N}\}$. Then $\sigma(\mathcal{F}) = \mathcal{P}(\mathbb{N})$.

Definition 1.13. *Let (S, d) be a metric space.*[5] *Let $\mathcal{B}(S)$ be the σ-algebra generated by the open sets in S. Thus*

$$\mathcal{B}(S) = \sigma\{\textit{open sets in } S\}.$$

The σ-algebra $\mathcal{B}(S)$ is called[6] *the* ***Borel σ-algebra*** *of S. The sets of $\mathcal{B}(S)$ are called the* ***Borel subsets*** *of S.*

Example 1.14. One of the most important σ-algebras is the Borel σ-algebra of $\mathbb{R}$ which is usually denoted by $\mathcal{B}(\mathbb{R})$. Later on we will show that $\mathcal{B}(\mathbb{R}) \neq \mathcal{P}(\mathbb{R})$.

The following lemma will be useful in some of the exercises on the Borel σ-algebra of $\mathbb{R}$ and $\mathbb{R}^d$.

Lemma 1.15 (Lindelöf)**.** *Let $A \subseteq \mathbb{R}^d$. Assume that for each $i \in I$ let $O_i \subseteq \mathbb{R}^d$ be open. If $A \subseteq \bigcup\limits_{i \in I} O_i$, then there exists a countable $J \subseteq I$ such that $A \subseteq \bigcup\limits_{i \in J} O_i$*

[4] Note that Proposition 1.9 ensures that $\sigma(\mathcal{F})$ is indeed a σ-algebra. The intersection makes sure we obtain the smallest possible one

[5] More generally one could take any topological space here

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

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, $\mathcal{F}$ has at most countably many sets. Indeed, this follows from the fact that it is a subset of $\{B(q, r) : q \in \mathbb{Q}^d, r \in \mathbb{Q} \cap (0, \infty)\}$ which is countable. 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 that $x \in O_{i_{x_n}} \subseteq \bigcup_{i \in J} O_i$ ☺

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. Prove that the collection in Example 1.7 (b) is a σ-algebra.

Exercise 1.3.
(a) Prove Proposition 1.9.
(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* 1.4. Let S be a set and let $\mathcal{F} = \{\{s\} : s \in S\}$ be the collection consisting of all sets which contain one element of S. Show that $\sigma(\mathcal{F})$ coincides with the countable-cocountable σ-algebra of Example 1.7.
Hint: Use the followsing (well-known) facts: The subset of a countable set is again countable; The countable union of countable sets is again countable.

Exercise 1.5. 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* 1.6. 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 1.15.
(c) Conclude that $\sigma(\mathcal{B}_0) = \mathcal{B}(\mathbb{R})$.

Exercise* 1.7. Let $\mathscr{I}^d \subseteq \mathcal{F}^d$ be the collection of half-open rectangles of Example 1.8. Prove the following assertions:
(a) If $I, J \in \mathscr{I}^d$ then $I \cap J \in \mathscr{I}^d$.
(b) If $I, J \in \mathscr{I}^d$, then $I \setminus J$ is the union of finitely many *disjoint* sets from $\mathscr{I}^d$, and thus $I \setminus J \in \mathcal{F}^d$.
Hint: Use induction on the dimension d. Use Example 1.8 (b) for $d = 1$.
(c) Each $A \in \mathcal{F}^d$ can be written as union of finitely many *disjoint* sets in $\mathscr{I}^d$.
Hint: Use induction on n to prove this for all sets of the form $A = \bigcup_{k=1}^{n} I_k$ with $I_1, \ldots, I_n \in \mathscr{I}^d$.
(d) $\mathcal{F}^d$ is a ring.

Exercise** 1.8. Prove that a σ-algebra is either finite or uncountable.[7]
Hint: Recall that $\mathcal{P}(\mathbb{N})$ is uncountable.

[7] This shows that σ-algebras are either easy finite sets or quite complicated

2. Measures

Definition 2.1. *Let S be a set and $\mathcal{R} \subseteq \mathcal{P}(S)$ a ring. Let $\mu : \mathcal{R} \to [0, \infty]$ be a function with*[8] $\mu(\varnothing) = 0$

(i) *μ is called **additive** if for all disjoint*[9] *$A_1, \ldots, A_n \in \mathcal{R}$ one has*

$$\mu\Big(\bigcup_{j=1}^{n} A_j\Big) = \sum_{j=1}^{n} \mu(A_j).$$

(ii) *μ is called **σ-additive** if for each disjoint sequence $(A_n)_{n\geq 1}$ in $\mathcal{R}$ which satisfies $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{R}$ it holds that*

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \sum_{j=1}^{\infty} \mu(A_j).$$

Remark 2.2.

(1) By an induction argument it suffices to consider $n = 2$ in the definition of additive. (See Exercise 2.1).
(2) If μ is σ-additive, then it is additive as follows by taking $A_m = \varnothing$ for $m \geq n + 1$.

Definition 2.3. *Let S be a set and let $\mathcal{A}$ be a σ-algebra on S.*

(i) *The pair $(S, \mathcal{A})$ is called a **measurable space**.*
(ii) *A function $\mu : \mathcal{A} \to [0, \infty]$ which satisfies $\mu(\varnothing) = 0$ and which is σ-additive on $\mathcal{A}$ is called a **measure**. In this case the triple $(S, \mathcal{A}, \mu)$ is called a **measure space**.*
(iii) *If additionally to (ii) $\mu(S) < \infty$, then μ is called a finite measure. If moreover, $\mu(S) = 1$, then μ is called a **probability measure** and $(S, \mathcal{A}, \mu)$ is called a **probability space**.* [10]

Example 2.4 (Counting measure). Let $S = \mathbb{N}$ 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 called the **counting measure**.

Example 2.5 (Dirac measure/Dirac's delta function). 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 $x \in \mathbb{R} \setminus A$. Then δ_x is a measure. It is usually called the **Dirac measure**[11] at x.

Example 2.6. Let S be a set and let $\mu : \mathcal{P}(S) \to [0, \infty]$ be given by $\mu(\varnothing) = 0$ and $\mu(A) = 1$ if $A \neq \varnothing$. If S contains at least two elements, then μ is not a measure.

Example 2.7 (Length of an interval). [12] Let $\mathcal{R} = \mathcal{F}^1$ as in Example 1.8. For $a \leq b$ let $\lambda((a, b]) = b - a$ (length of the interval). If $A = (a_1, b_1] \cup \ldots \cup (a_n, b_n]$ is a union of disjoint sets such that $a_j \leq b_j$ for $j = 1, \ldots, n$, then we let[13] $\lambda(A) = \sum_{j=1}^{n} b_j - a_j$. Then λ is additive. Later we will see that λ is σ-additive on $\mathcal{F}^1$ and has an extension to a measure on $\sigma(\mathcal{F}^1) = \mathcal{B}(\mathbb{R})$.

Theorem 2.8. *Let $\mathcal{R}$ be a ring and $\mu : \mathcal{R} \to [0, \infty]$ be additive. The following assertions hold:*

(i) *If $A, B \in \mathcal{R}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$ (monotonicity).*
(ii) *If $A_1, A_2, \ldots \in \mathcal{R}$ are disjoint and $\bigcup\limits_{j=1}^{\infty} A_j \in \mathcal{R}$, then*

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \geq \sum_{j=1}^{\infty} \mu(A_j).$$

[8] This assumption will always be made.
[9] Here we mean $A_i \cap A_j = \varnothing$ if $i \neq j$
[10] Measure theory is at the very heart of probability theory. See the third year elective course
[11] Named after the English theoretical physicist Paul Dirac 1902-1984. The "delta function" is actually not a function, but can be interpreted as a generalized function or as measure.
[12] See Section 3 for a further construction of the so-called Lebesgue measure
[13] One can check that this does not depend on the way we write the set A. See below Definition 3.6.

(iii) If $A_1, A_2, \ldots \in \mathcal{R}$ and $\bigcup_{j=1}^{\infty} A_j \in \mathcal{R}$ and μ is σ-additive on $\mathcal{R}$, then

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) \leq \sum_{j=1}^{\infty} \mu(A_j) \qquad (\sigma\text{-subadditivity}).$$

Proof. (i): 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): Let $n \in \mathbb{N}$. Then $\bigcup_{j=1}^{n} A_j \subseteq \bigcup_{j=1}^{\infty} A_j$ and therefore by additivity of μ

$$\sum_{j=1}^{n} \mu(A_j) = \mu\Big(\bigcup_{j=1}^{n} A_j\Big) \overset{(i)}{\leq} \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big).$$

The result follows by letting $n \to \infty$.

(iii): Let

$$B_1 = A_1,\ B_2 = A_2 \setminus A_1,\ B_3 = A_3 \setminus (A_1 \cup A_2),\ \text{etc.} \tag{2.1}$$

Then $(B_n)_{n\geq 1}$ is a disjoint sequence and $\bigcup_{j=1}^{\infty} B_j = \bigcup_{j=1}^{\infty} A_j$. Therefore, by the σ-additivity of μ

$$\mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \mu\Big(\bigcup_{j=1}^{\infty} B_j\Big) = \sum_{j=1}^{\infty} \mu(B_j) \overset{(i)}{\leq} \sum_{j=1}^{\infty} \mu(A_j).$$

☺

Let $(a_n)_{n\geq 1}$ be a sequence of real numbers. 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 will now be extended to sets.

Definition 2.9. *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 $A = \bigcap_{n=1}^{\infty} A_n$.*

Theorem 2.10. *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: $\bigcup_{j=1}^{\infty} B_j = A$ and $\bigcup_{j=1}^{n} B_j = A_n$. Therefore, the σ-additivity of μ gives

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

(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 3.9.

Lemma 2.11 (Sufficient condition for σ-additivity)**.** *Let $\mathcal{R}$ be a ring and let $\mu : \mathcal{R} \to [0, \infty]$ be such that $\mu(\varnothing) = 0$ and μ is additive. Suppose that for each sequence $(A_n)_{n\geq 1}$ with $A_n \downarrow \varnothing$ one has $\mu(A_n) \to 0$. Then μ is σ-additive on $\mathcal{R}$.*

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

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

Let $A_n = \bigcup_{j=n}^{\infty} B_j = B \setminus (B_1 \cup \ldots \cup B_{n-1})$. Then $A_n \in \mathcal{R}$ and $A_n \downarrow \varnothing$. Now the assumption yields $\mu(A_n) \to 0$. On the other hand

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

Therefore, $\left|\mu(B) - \sum_{j=1}^{n-1} \mu(B_j)\right| = \mu(A_n) \to 0$ and (2.2) follows. ☺

Exercises

Exercise 2.1. Let $\mathcal{R}$ be a ring on a set S. Assume $\mu : \mathcal{R} \to [0,\infty]$ satisfies $\mu(\varnothing) = 0$ and $\mu(A \cup B) = \mu(A) + \mu(B)$ for all sets $A, B \in \mathcal{R}$ with $A \cap B = \varnothing$. Show that μ is additive.

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

(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) Show that for any $n \in \mathbb{N}$ and any sets $(A_j)_{j=1}^{n}$ in $\mathcal{R}$,
$$\text{finite subadditivity} \quad \mu(A_1 \cup \ldots \cup A_n) \leq \mu(A_1) + \ldots + \mu(A_n).$$

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

Exercise* 2.4.

(a) Prove Theorem 2.10 (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 $\mu(\limsup_{n\to\infty} A_n) = 0$.[14]

Exercise* 2.6. Let $\mathcal{A}$ be the σ-algebra from Example 1.7 (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.

[14] 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 $\mathbb{R}^d$ of which we have previously shown it is an additive mapping on the ring $\mathcal{F}^1$ in Example 2.7. To extend it to the Borel σ-algebra we use a deep result of Carathéodory.[15] 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.6 in the appendix, but it will do no harm if one takes the result for granted.[16]

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 $\mu : \mathcal{R} \to [0, \infty]$ is σ-additive on $\mathcal{R}$ and $\mu(\varnothing) = 0$. Then μ extends to a measure $\overline{\mu}$ on $\sigma(\mathcal{R})$. More precisely, there exists a measure $\overline{\mu}$ on $(S, \sigma(\mathcal{R}))$ such that $\overline{\mu}(A) = \mu(A)$ for all $A \in \mathcal{R}$.*

Remark 3.2. The measure $\overline{\mu}$ is often unique.[17] 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 $\mu(A) = \infty$ 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 which will be proved in the appendix.

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) Let $S = \mathbb{R}^d$. The half-open rectangles $\mathcal{I}^d$ are a π-system.

The following result will be proved in Proposition A.7.

Proposition 3.5 (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 $\sigma(\mathcal{E}) = \mathcal{A}$;
(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 $\mathcal{A}$.

We continue with the construction of the **Lebesgue** [18] **measure** λ. In Example 1.8 we introduced the half-open intervals $(a, b] \in \mathcal{I}^1$ with $a \leq b$. Also recall that $\mathcal{F}^1$ denotes the collection of all finite unions of half-open intervals. We have seen that $\mathcal{F}^1$ is a ring. Moreover, in of course every set in $\mathcal{F}^1$ can be written as a finite union of disjoint half-open intervals.

Definition 3.6 (on unions of half-open intervals)**.** *For $A \in \mathcal{F}^1$ of the form $A = (a_1, b_1] \cup \ldots \cup (a_n, b_n]$ with disjoint $((a_j, b_j])_{j=1}^n$ in $\mathcal{I}^1$ define $\lambda : \mathcal{F}^1 \to [0, \infty]$ as the sum of the lengths:*

$$\lambda(A) = \sum_{j=1}^{n} (b_j - a_j).$$

The above is well-defined. To see this assume $A = (c_1, d_1] \cup \ldots (c_m, d_m]$ is another representation of A as a union of disjoint intervals. Let $I_{ij} = (c_i, d_i] \cap (a_j, b_j]$. Then either I_{ij} is empty or a half open interval,

$$\bigcup_{i=1}^{m} I_{ij} = (a_j, b_j] \quad \text{and} \quad \bigcup_{j=1}^{n} I_{ij} = (c_j, d_j].$$

[15] Carathéodory 1873-1950 was a Greek mathematician working in Analysis, but also on Thermodynamics.
[16] The appendix is not part of the exam
[17] For instance when μ is a finite measure. See Proposition A.7.
[18] Henri Lebesgue 1875–1941 was a French mathematician well-known for his integration theory. See Section 5.

From the definition and the disjointness of the $(I_{ij})_{i,j=1}^{m,n}$ we obtain

$$\sum_{j=1}^{n}\lambda\big((a_j,b_j]\big)=\sum_{j=1}^{n}\lambda\Big(\bigcup_{i=1}^{m}I_{ij}\Big)=\sum_{j=1}^{n}\sum_{i=1}^{m}\lambda(I_{ij})=\sum_{i=1}^{m}\sum_{j=1}^{n}\lambda(I_{ij})=\sum_{i=1}^{m}\lambda\Big(\bigcup_{j=1}^{n}I_{ij}\Big)=\sum_{i=1}^{m}\lambda\big((c_i,d_i]\big),$$

which proves the well-definedness. Alternatively, one can observe that $\lambda(A)$ coincides with the Riemann integral of [19] $\mathbf{1}_A$. Indeed, fix an interval I such that $A\subseteq I$. By linearity of the Riemann integral

$$\int_I \mathbf{1}_A\,dx=\sum_{j=1}^{n}\int_I \mathbf{1}_{(a_j,b_j]}\,dx=\sum_{j=1}^{n}(b_j-a_j)=\lambda(A).$$

In Example 1.8 we introduced the half-open rectangles $(a,b]\in\mathcal{I}^d$ with $a=(\alpha_1,\ldots,\alpha_d)$ and $b=(\beta_1,\ldots,\beta_d)$ and $\alpha_j\le\beta_j$ for $j\in\{1,\ldots,d\}$. Also recall that $\mathcal{F}^d$ denotes the collection of all finite unions of half-open rectangles. By Exercise 1.7 (d), $\mathcal{F}^d$ is a ring. Moreover, in Exercise 1.7 (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 (on unions of half-open rectangles)**.** *For a half open rectangle $I=(a,b]\in\mathcal{I}^d$ with $a=(\alpha_1,\ldots,\alpha_d)$ and $b=(\beta_1,\ldots,\beta_d)$ let its **volume** be denoted 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 $(I_j)_{j=1}^n$ in $\mathcal{I}^d$ define $\lambda_d:\mathcal{F}^d\to[0,\infty]$ by

$$\lambda_d(A)=\sum_{j=1}^{n}|I_j|.$$

This is well-defined since for a rectangle $R\subseteq\mathbb{R}^d$ with $A\subseteq R$, we again have

$$\lambda_d(A)=\int_R \mathbf{1}_A\,dx,$$

where the latter is a d-dimensional Riemann integral.

Remark 3.8.

(1) When the dimension is fixed and there is no danger of confusion we will write λ for λ_d. It is common to write $|A|$ for $A\in\mathcal{F}^d$ as well.
(2) For $A\in\mathcal{F}^d$, $\lambda(A)=\lambda_d(A)=|A|$ equals the volume of A.

Next we want to extend λ_d to $\sigma(\mathcal{F}^d)=\mathcal{B}(\mathbb{R}^d)$ (see Exercise 3.1 for this identity). 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 2.11.

Lemma 3.9. *The function $\lambda:\mathcal{F}^d\to[0,\infty]$ is σ-additive on $\mathcal{F}^d$.*

Proof. By Lemma 2.11 it suffices to prove each sequence $(A_n)_{n\ge1}$ in $\mathcal{F}^d$ with $A_n\downarrow\varnothing$, satisfies $\mu(A_n)\to0$. Fix $\varepsilon>0$. We have to find $N\in\mathbb{N}$ such that $\lambda(A_n)<\varepsilon$ for all $n\ge N$.

Step 1: For each $n\in\mathbb{N}$ choose a $B_n\in\mathcal{F}^d$ such that $\overline{B_n}\subseteq A_n$ and $\lambda(A_n\setminus B_n)\le 2^{-n}\varepsilon$.[20] Since $\overline{B_n}\subseteq A_n$ also $\bigcap_{n=1}^{\infty}\overline{B_n}=\varnothing$. It follows that $\{(\overline{B_n})^c:n\in\mathbb{N}\}$ is an open cover of the set $\overline{A_1}$ which is compact by the Heine-Borel theorem. Therefore, there exists an N such that $\overline{A_1}\subseteq\bigcup_{n=1}^{N}(\overline{B_n})^c$. It follows that $\bigcap_{n=1}^{N}\overline{B_n}\subseteq A_1^c$. Since all for all $n\ge1$, $\overline{B_n}\subseteq A_1$, we must have that $\bigcap_{n=1}^{N}\overline{B_n}=\varnothing$.

[19] Recall that $\mathbf{1}_A(x)=1$ if $x\in A$ and $\mathbf{1}_A=0$ if $x\in A^c$.

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

Step 2: Let $C_n = \bigcap_{j=1}^{n} B_j$ for $n \geq 1$. For every $n \geq 1$, $A_n \setminus C_n = \bigcup_{j=1}^{n} (A_n \setminus B_j) \subseteq \bigcup_{j=1}^{n} (A_j \setminus B_j)$. Therefore, using Theorem 2.8 (i) in $(*)$ and Exercise 2.2 (c) in $(**)$, we find

$$\lambda(A_n \setminus C_n) \overset{(*)}{\leq} \lambda\Big(\bigcup_{j=1}^{n} (A_j \setminus B_j)\Big) \overset{(**)}{\leq} \sum_{j=1}^{n} \lambda(A_j \setminus B_j) \leq \sum_{j=1}^{n} 2^{-j}\varepsilon < \varepsilon.$$

Since $C_n = \varnothing$ for all $n \geq N$, we can conclude that $\lambda(A_n) = \lambda(A_n \setminus C_n) < \varepsilon$ for every $n \geq N$. ☺

We can now deduce the main result of this section.

Theorem 3.10 (Lebesgue measure)**.** *There exists a unique measure λ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that for all half-open rectangles I, one has $\lambda(I) = |I|$, where $|I|$ is the volume of I. Moreover, for all $h \in \mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$, $\lambda(A + h) = \lambda(A)$.*[21]

In the above $A + h := \{x + h : x \in A\}$.

Proof. Step 1: Existence. In Lemma 3.9 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 $\sigma(\mathcal{F}^d) = \mathcal{B}(\mathbb{R}^d)$ (see Exercise 3.1).

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 2.10 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 check that $\mathcal{A}_h$ is a σ-algebra. For each open set A one has $A + h$ is open and hence $A + h \in \mathcal{B}(\mathbb{R}^d)$. Therefore, $\mathcal{B}(\mathbb{R}^d) = \sigma(\{\text{open sets}\}) \subseteq \mathcal{A}_h$, and 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 proved the result. ☺

Remark 3.11. From Theorem B.6 one can actually see that for any $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\lambda(A) = \inf\Big\{\sum_{j=1}^{\infty} |I_j| : A \subseteq \bigcup_{j=1}^{\infty} I_j, \text{ where } (I_j)_{j\geq 1} \text{ are disjoint half-open rectangles}\Big\},$$

but we will not use this formula.

Exercises on the Lebesgue measure

If the dimension is fixed we write λ instead of λ_d for simplicity.

Exercise* 3.1. Let $\mathcal{F}^d$ be as in Example 1.8. Show that $\sigma(\mathcal{F}^d)$ is the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$.
Hint: Use the Lindelöf Lemma 1.15.

Exercise 3.2.

(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 $\lambda(\{x\}) = 0$.

[21]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 3.6.

Exercise* 3.3. 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 $j \in \{1, \ldots, k\}$ 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 2.10.

Exercise* 3.4 (Uncountable sets can have measure zero). Show that the Cantor set $C \subseteq [0, 1]$ is in $\mathcal{B}(\mathbb{R}^d)$ and satisfies $\lambda(C) = 0$.
Hint: What can you say about $\lambda(C_n)$?

Exercise* 3.5. For $A \subseteq \mathbb{R}$ and $t \geq 0$ 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 3.10.

Exercise* 3.6. Let $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ 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: [22]

(a) For each $x \geq 0$ 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 $x \geq 0$, $\mu\big((0, x]\big) = cx$.
(d) For each $a \leq b$, $\mu\big((a, b]\big) = c(b - a)$.
(e) For each $A \in \mathcal{B}(\mathbb{R})$, $\mu(A) = c\lambda(A)$.

Exercise** 3.7 (Lebesgue-Stieltjes[23] measure). 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 $a \leq b$ extends to a measure on $\mathcal{B}(\mathbb{R})$.

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) If $T \in \mathcal{A}$, show that $\mathcal{A}_T = \{A \subseteq T : A \in \mathcal{A}\}$.
(c) Let μ be a measure on $(S, \mathcal{A})$. If $T \in \mathcal{A}$ show that the restriction of μ to $\mathcal{A}_T$ is a measure again.
(d) If $\mathcal{A}$ is the Borel σ-algebra on metric space (S, d), then $\mathcal{A}_T$ coincides with the Borel σ-algebra on (T, d).

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 A.7?

---

[22] 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 $d \geq 2$ holds as well and can be proved in a similar way.

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

4. Measurable functions

One of the aims will be to integrate functions $f : S \to \mathbb{R}$ with respect to a measure μ on a measurable space $(S, \mathcal{A})$. A way to do this is to use discretization in the *range*[24] of f. So we would like to know the measure of for instance the set $A_{y,\varepsilon} = \{s \in S : f(s) \in [y, y+\varepsilon]\}$. Knowing this for all $y \in \mathbb{R}$ and all $\varepsilon > 0$ makes it possible estimate the "area" under f. Of course we do need the sets to be in $\mathcal{A}$ to make this work. This is the motivation of the definition of measurability. See Section 5 for details on integration.

The natural setting to introduce measurability of functions is a follows:

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*[25] *$f^{-1}(B) \in \mathcal{A}$.*

Remark 4.2.

(1) The composition of two measurable function is again measurable (see Exercise 4.1).
(2) Instead of $f^{-1}(B)$ or $\{s \in S : f(s) \in B\}$ one sometimes writes $\{f \in B\}$ for the same set.
(3) In probability theory measurable functions are called **random variables**.

It suffices to check measurability on a generating collection $\mathcal{F} \subseteq \mathcal{B}$:

Lemma 4.3. *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.

Now since $\mathcal{F} \subseteq \widetilde{\mathcal{B}}$, the claim yields $\mathcal{B} = \sigma(\mathcal{F}) \subseteq \widetilde{\mathcal{B}} \subseteq \mathcal{B}$. This implies $\widetilde{\mathcal{B}} = \mathcal{B}$. ☺

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

Proposition 4.4 (Continuous mappings are measurable)**.** *Let (X, d) and (Y, ρ) be metric spaces. If $f : X \to Y$ is continuous, then f is measurable.*[26]

Proof. By the continuity of f we find that for all open $O \subseteq Y$ the inverse image is $f^{-1}(O)$ open in X and hence in $\mathcal{B}(X)$. Since the open sets of Y generate the Borel σ-algebra, the result follows from Lemma 4.3 with $\mathcal{F} = \{O \subseteq Y : O \text{ is open}\}$. ☺

The most frequent case we will encounter is when $f : S \to \mathbb{R}$ and $(S, \mathcal{A})$ is a measurable space. Unless otherwise stated we consider the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$. The following characterization of measurability will be useful.

Proposition 4.5 (Real valued functions)**.** *Let $(S, \mathcal{A})$ be a measurable space. For $f : S \to \mathbb{R}$ the following are equivalent:*

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

---

[24] 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

[25] Recall that $f^{-1}(B)$ is called the inverse image of B by f and is defined by $f^{-1}(B) = \{s \in S : f(s) \in B\}$

[26] The same result holds for topological spaces and the proof is the same. In the setting of Borel-σ-algebras, measurable functions are often called Borel measurable.

Proof. (i) $\Rightarrow$ (ii) and (i) $\Rightarrow$ (iii) are trivial.

(iii) $\Rightarrow$ (i): Let $\mathcal{F} = \{(-\infty, r) : r \in \mathbb{R}\}$. In Exercise 1.6 we have seen that $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R})$. Therefore, (i) follows from Lemma 4.3.

(ii) $\Rightarrow$ (i): This can be proved as before by proving the required version of Exercise 1.6. ☺

Example 4.6. Let $S = \{1, 2, 3\}$ and $\mathcal{A} = \{S, \varnothing, \{1, 2\}, \{3\}\}$. Let $f : S \to \mathbb{R}$ be given by $f(s) = 2016$ if $s = 1$ and $f(s) = 0$ if $s \neq 1$. Then f is not measurable, because $f^{-1}(\{2016\}) = \{1\} \notin \mathcal{A}$. If we replace $\mathcal{A}$ by (for example) $\mathcal{A}' := \{S, \varnothing, \{1\}, \{2, 3\}\}$, then f becomes measurable.

Recall that $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$.

Theorem 4.7. *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 all 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 3.1 $\sigma(\text{half open rectangles}) = \sigma(\mathcal{F}^2) = \mathcal{B}(\mathbb{R}^2)$, we can use Lemma 4.3 to find that h is measurable.

To prove the statements we use that continuous functions are measurable (see Proposition 4.4) and the fact that the composition of measurable functions is again measurable (see Exercise 4.1). For instance let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ be given by $\varphi(x, y) = x + y$. Then φ is continuous and therefore measurable. Now writing $f + g = \varphi \circ h$ the required measurability follows.

The proofs for the difference, product, maximum, minimum are similar. Note that the maximum $x \vee y = \frac{1}{2}(x + y) + \frac{1}{2}|x - y|$ and this is a continuous function from $\mathbb{R}^2$ to $\mathbb{R}$. For the minimum there is an analogue formula.

The measurability of $f^{\pm}$, $|f|$, $\alpha \cdot f$ and $\frac{1}{f}$ all follow in the same way by rewriting them as $\phi \circ f$ for a suitable continuous function ϕ.

The case $\frac{1}{f}$ requires some comment. Let $\phi : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ be given by $\phi(x) = \frac{1}{x}$. Note that in this case $\mathbb{R} \setminus \{0\}$ is a metric space on which ϕ is continuous, and therefore measurable by Proposition 4.4. Now for all open sets $B \in \mathcal{B}(\mathbb{R})$, $C := \phi^{-1}(B)$ is open in $\mathbb{R} \setminus \{0\}$ and hence in $\mathcal{B}(\mathbb{R})$. Thus $(\phi \circ f)^{-1}(B) = f^{-1}(C) \in \mathcal{A}$. Therefore, the measurability of $\phi \circ f$ follows from Lemma 4.3. ☺

Let $\overline{\mathbb{R}} = [-\infty, \infty]$. It will be useful to introduce measurability of functions $f : S \to \overline{\mathbb{R}}$ as well. For this, we introduce an analogue of the Borel σ-algebra on $\overline{\mathbb{R}}$.

Definition 4.8. *Let $\mathcal{B}(\overline{\mathbb{R}})$ be the the σ-algebra generated by the sets $\{\infty\}, \{-\infty\}$ and $B \in \mathcal{B}(\mathbb{R})$. The σ-algebra $\mathcal{B}(\overline{\mathbb{R}})$ will be called the Borel σ-algebra of $\overline{\mathbb{R}}$.*

From Lemma 4.3 we see that a function $f : S \to \overline{\mathbb{R}}$ is measurable if and only if $\{s \in S : f(s) = \pm\infty\} \in \mathcal{A}$ and $f^{-1}(B) \in \mathcal{A}$ for each $B \in \mathcal{B}(\mathbb{R})$. We extend addition, multiplication, etc. to $\overline{\mathbb{R}}$ in the following way:

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

In this setting Theorem 4.7 remains true[27] for functions $f, g : S \to \overline{\mathbb{R}}$.

[27] 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 $\{\infty\}$, $\{-\infty\}$, $\{0\}$ are measurable. We leave this to the reader.

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

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

The limit of $y_k := \sup_{n\geq k} x_n$ exists[28] 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 $\liminf_{n\to\infty} x_n$. Recall that the $\limsup_{n\to\infty} x_n$ and $\liminf_{n\to\infty} x_n$ always exist. 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 4.9. *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:*[29]

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

Moreover, if $f_n \to f$ pointwise[30]*, then f is measurable again.*

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

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

Since $\sigma(\{[-\infty, r] : r \in \mathbb{R}\}) = \mathcal{B}(\overline{\mathbb{R}})$, the measurability of g follows from Lemma 4.3. The case of infima follows from $\inf_{n\geq 1} f_n = -\sup_{n\geq 1}(-f_n)$.

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 $\lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n$. ☺

Definition 4.10. *A function $f : S \to \mathbb{R}$ is called a* ***simple function***[31] *if f is measurable and takes only finitely many values.*

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

$$f = \sum_{j=1}^{n} x_j \cdot \mathbf{1}_{A_j}.$$

Of course if $x_j = 0$ for some $j \in \{1, \ldots, n\}$, we could leave it out from the sum.

Example 4.11. 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) $f = \mathbf{1}_{\mathbb{Q}}$.

Next we show that measurable functions can be written as limits of simple functions. For this we discretize in the range space in a suitable way. The result plays a crucial role in the integration theory in Section 5.

Theorem 4.12. *Let $(S, \mathcal{A})$ be a measurable space.* [32]

(i) Let $f : S \to [0, \infty]$ be measurable. Then there exists a sequence of simple functions $(f_n)_{n\geq 1}$ such that $0 \leq f_1(s) \leq f_2(s) \leq \ldots$ and $\lim_{n\to\infty} f_n(s) = f(s)$ for each $s \in S$.
(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 $\lim_{n\to\infty} f_n(s) = f(s)$ for all $s \in S$.

---

[28] Here we allow divergence to $\pm\infty$
[29] Here it is important what we work with countable suprema, infimum, etc.
[30] Here we allow divergence to $\pm\infty$
[31] In part of the literature this is called a step function, but we will use this name for a different class of functions
[32] In the following we allow divergence to $\pm\infty$

Proof. [33] (i): For each $n \in \mathbb{N}$ and $j \in \{0, 1, \dots, 4^n - 1\}$, let

$$A_{n,j} = \{s \in S : j2^{-n} \le f(s) < (j+1)2^{-n}\} \quad \text{and} \quad A_n = \{s \in S : f(s) \ge 2^n\}.$$

Then for each n, j one has $A_{n,j}, A_n \in \mathcal{A}$. Define[34]

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

It is clear that each f_n takes finitely many values. Moreover, by Exercise 4.5 and Theorem 4.7, each f_n is measurable. Thus each f_n is a simple function.

Now fix $s \in S$. We first prove $0 \le f_n(s) \le f_{n+1}(s)$ for each $n \in \mathbb{N}$. First assume $f(s) < 2^n$. Then selecting the unique $j \in \{0, \dots, 4^n - 1\}$ such that $s \in A_{n,j}$ we find that $f_n(s) = j2^{-n}$. Similarly, we can select $k \in \{0, \dots, 4^{n+1} - 1\}$ such that $s \in A_{n+1,k}$ and we find that $f_{n+1}(s) = k2^{-(n+1)}$. Since, $f(s) \ge j2^{-n} = 2j2^{-(n+1)}$, we find that $k \ge 2j$ and therefore

$$f_n(s) = j2^{-n} \le k2^{-(n+1)} = f_{n+1}(s).$$

The case $f(s) \ge 2^n$ 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 $f(s) < 2^N$ and $2^{-N} < \varepsilon$. Let $n \ge N$. Selecting $j \in \{0, \dots, 4^n - 1\}$ such that $s \in A_{n,j}$ we find that

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

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 $f = f^+ - f^-$. Then 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 $n \in \mathbb{N}$. Define the sets A_+ and A_- by $A_\pm = \{s \in S : \pm f(s) \in [0, \infty]\}$. Then $A_+ \cup A_- = S$. If $s \in A_+$, then

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

The case $s \in A_-$ is similar.

☺

Exercises

Exercise 4.1. Let $(S_j, \mathcal{A}_j)$ for $j = 1, 2, 3$ be measurable spaces. Assume $f : S_1 \to S_2$ and $g : S_2 \to S_3$ are both measurable. Show that the composition $g \circ f : S_1 \to S_3$ is measurable.

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. Let S be a set. For a function $f : S \to \mathbb{R}$ let $\mathcal{A}_f = \{f^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$. Show that $\mathcal{A}_f$ is a σ-algebra.[35]

Exercise 4.5. Let $(S, \mathcal{A})$ be a measurable space and let $A \subseteq S$. Show that $A \in \mathcal{A}$ if and only if the function $\mathbf{1}_A : S \to \mathbb{R}$ is measurable.

Exercise 4.6. 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* 4.7 (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 4.7.

---

[33] For the proof make a picture where you make a partition into intervals of length 2^{-n} on the y-axis. For the picture put the set S is on the x-axis

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

[35] The σ-algebra $\mathcal{A}_f$ is called the σ-algebra generated by f. It is the smallest σ-algebra for which f is measurable.

Exercise* 4.8 (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 4.5.

Exercise* 4.9 (Set of convergence). Let $(S, \mathcal{A})$ be a measurable space. Let $f_1, f_2, \ldots : S \to \mathbb{R}$ be measurable functions. Define

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

(a) Explain why $A = \{s \in S : (f_n(s))_{n\geq 1}$ 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 $A \in \mathcal{A}$.

5. Construction of the integral

In this section $(S, \mathcal{A}, \mu)$ is a measure space. Our goal is to construct an integral for measurable functions $f : S \to \overline{\mathbb{R}}$. Notation:

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

The 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, Σ, μ). Moreover, in the special case of the Lebesgue measure it extends the Riemann integral.

It will be convenient to use the following terminology.

Definition 5.1. *For measurable functions $f, g : S \to \overline{\mathbb{R}}$ we say that $f = g$* ***almost everywhere*** *if $\mu(\{s \in S : f(s) \neq g(s)\}) = 0$. Notation: $f = g$ a.e.*[36]

Similarly, one can define $f < \infty$ a.e., etc.

5.1. Integral for simple functions.

Definition 5.2 (Integral for simple functions)**.** *Let $f : S \to [0, \infty]$ be a simple function given by*

$$f = \sum_{j=1}^{n} x_j \cdot \mathbf{1}_{A_j}, \tag{5.1}$$

with $x_1, \ldots, x_n \in [0, \infty]$ and $(A_j)_{j=1}^n$ disjoint sets in $\mathcal{A}$. For $E \in \mathcal{A}$ let[37]

$$\int_E f \, \mathrm{d}\mu = \sum_{j=1}^{n} x_j \cdot \mu(E \cap A_j).$$

This is called the ***integral*** *of f over the set E.*

Remark 5.3. By using a common refinement one checks that if a different representation for the function f from (5.1) is used, then the integral over E gives the same value (see Lemma 5.4 (ii)). Clearly $\int_E f \, \mathrm{d}\mu \in [0, \infty]$ for every $E \in \mathcal{A}$.

We continue with some basic properties of the integral.

Lemma 5.4. *Let $f, g : S \to [0, \infty]$ 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 $E_1, E_2 \in \mathcal{A}$, $\int_{E_1 \cup E_2} f \, \mathrm{d}\mu = \int_{E_1} f \, \mathrm{d}\mu + \int_{E_2} f \, \mathrm{d}\mu$.
(vi) $\int_S f \, \mathrm{d}\mu = 0$ if and only if $f = 0$ almost everywhere.[38].

Proof. (i). This is immediate from the definition and the identity $\mathbf{1}_{E \cap A} = \mathbf{1}_E \cdot \mathbf{1}_A$.

For the proof of the remaining assertions we continue with a preliminary observation. Write $f = \sum_{i=1}^m x_i \cdot \mathbf{1}_{A_i}$ with $(A_i)_{i=1}^m$ disjoint sets in $\mathcal{A}$ with $\bigcup\limits_{i=1}^{m} A_i = S$, and $g = \sum_{j=1}^n y_j \cdot \mathbf{1}_{B_j}$ with

[36] In probability theory this is usually called almost surely and this is abbreviated as a.s.
[37] Here we use the convention $0 \cdot \infty = 0$
[38] See Definition 5.1

$(B_j)_{j=1}^n$ disjoint sets in $\mathcal{A}$ with $\bigcup_{i=1}^m B_i = S$. Let $C_{i,j} = A_i \cap B_j$ for all i and j. Then $(C_{i,j})_{i,j=1}^{m,n}$ are disjoint sets in $\mathcal{A}$. By additivity

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

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

(ii): For disjoint sets $A, B \subseteq S$ one has $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$ and hence

$$\mathbf{1}_E f = \sum_{i=1}^m \sum_{j=1}^n x_i \cdot \mathbf{1}_{E\cap C_{i,j}} \quad \text{and} \quad \mathbf{1}_E g = \sum_{j=1}^n \sum_{i=1}^m y_j \cdot \mathbf{1}_{E\cap C_{i,j}}. \tag{5.4}$$

Therefore, $x_i \leq y_j$ whenever i, j satisfy $E \cap C_{i,j} \neq \varnothing$. This together with (5.2), (5.3), yields (ii).

(iii): This follows from $\mu(E \cap A_i) \leq \mu(F \cap A_i)$ which is immediate from the monotonicity of μ.

(iv): By (5.4) with $E = S$, we can write $\alpha f + \beta g = \sum_{i=1}^m \sum_{j=1}^n (\alpha x_i + \beta y_j) \cdot \mathbf{1}_{C_{i,j}}$. Therefore,

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

where we used (5.2) and (5.3).

(v): Since $\mathbf{1}_{E_1\cup E_2} f = \mathbf{1}_{E_1} f + \mathbf{1}_{E_2} f$,

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

(vi): By leaving out some of those k for which $x_k = 0$, we can assume $x_i > 0$ for all i. Let $A = \{s \in S : f(s) > 0\}$ and observe that $A = \bigcup_{i=1}^n A_i$. Now $\int_S f \,\mathrm{d}\mu = \sum_{i=1}^m x_i \mu(A_i)$ with $\mu(A_i) \geq 0$ for each i. Therefore, if the integral is zero, then $\mu(A_i) = 0$ for each i and hence $\mu(A) = \sum_{i=1}^n \mu(A_i) = 0$. Conversely, if $f = 0$ a.e., monotonicity yields $\mu(A_i) \leq \mu(A) = 0$, and the result follows. ☺

Example 5.5. Let λ be the Lebesgue measure on $\mathbb{R}$. Since $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$ it follows from Exercises 1.5 and 4.5 that $\mathbf{1}_\mathbb{Q}$ is measurable and by Exercise 3.2 for any $E \in \mathcal{B}(\mathbb{R})$, $\int_E \mathbf{1}_\mathbb{Q} \,\mathrm{d}\lambda = \lambda(E \cap \mathbb{Q}) = 0$. 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 Example 5.16.

5.2. **Integral for positive measurable functions.** In order to extend the definition of the integral to arbitrary measurable functions $f : S \to [0, \infty]$ we use the following lemma. In the sequel, we write $f_n \uparrow f$ if for all $s \in S$, $(f_n(s))_{n\geq 1}$ is increasing and $f_n(s) \to f(s)$.

Lemma 5.6 (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 $\int_E g \,\mathrm{d}\mu \leq \lim_{n\to\infty} \int_E f_n \,\mathrm{d}\mu$ for every $E \in \mathcal{A}$.*

Observe that $\lim_{n\to\infty} \int_E f_n \,\mathrm{d}\mu$ exists in $[0, \infty]$, because 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 5.4 in the estimates below, we find

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

It remains to prove $\int_{E_n} g \, \mathrm{d}\mu \to \int_E g \, \mathrm{d}\mu$. For this write $g = \sum_{i=1}^m x_i \cdot \mathbf{1}_{A_i}$ with $(A_i)_{i=1}^m$ in $\mathcal{A}$ disjoint and $x_1, \dots, x_m \geq 0$. Since $E_n \cap A_i \uparrow E \cap A_i$ as $n \to \infty$, Theorem 2.10 yields

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

☺

Definition 5.7 (Integral for positive functions)**.** *For a measurable function $f : S \to [0, \infty]$ and a sequence of simple functions $(f_n)_{n \geq 1}$ with $0 \leq f_n \uparrow f$ we define the **integral** of f over $E \in \mathcal{A}$ as*

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

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 4.12 we can always find 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 5.6 for each $m \geq 1$,

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

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.

Next we can extend the properties of the integral of Lemma 5.4.

Proposition 5.8. *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 $f \leq g$ 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 $E_1, E_2 \in \mathcal{A}$, $\int_{E_1 \cup E_2} f \, \mathrm{d}\mu = \int_{E_1} f \, \mathrm{d}\mu + \int_{E_2} f \, \mathrm{d}\mu$.
(vi) $\int_S f \, \mathrm{d}\mu = 0$ if and only if $f = 0$ almost everywhere.

Proof. (i): Let $(f_n)_{n \geq 1}$ be simple functions such that $0 \leq f_n \uparrow f$. Then by Lemma 5.4 (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 $(f_n)_{n \geq 1}$ such that $0 \leq f_n \uparrow f$. Then $f_n = 0$ a.e. for each $n \in \mathbb{N}$ and thus by Lemma 5.4 (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. Then $E = \{s \in S : f(s) > 0\}$ satisfies $\mu(E) > 0$. Letting $E_n = \{s \in S : f(s) \geq \frac{1}{n}\}$ we find $E_n \uparrow E$, so by Theorem 2.10, $\mu(E_n) \to \mu(E)$. Therefore, there exists an $n \in \mathbb{N}$ such that $\mu(E_n) > 0$ and thus

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

☺

Example 5.9 (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 $n \geq 1$, $f_n = \sum_{j=1}^n f(j) \mathbf{1}_{\{j\}}$. Since each f_n is a simple function and $0 \leq f_n \uparrow f$ we find

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

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

Definition 5.10. *A measurable function $f : S \to \overline{\mathbb{R}}$ is called **integrable** when both $\int_S f^+ \, d\mu$ and $\int_S f^- \, d\mu$ are finite. In this case the integral of f over $E \in \mathcal{A}$ is defined as*

$$\int_E f \, d\mu = \int_E f^+ \, d\mu - \int_E f^- \, d\mu.$$

Remark 5.11.
(1) If $f : S \to \overline{\mathbb{R}}$ is integrable, then by monotonicity $\int_E f^+ \, d\mu \in [0, \infty)$ and $\int_E f^- \, d\mu$ are finite for every $E \in \mathcal{A}$. Therefore, $\int_E f \, d\mu$ is a number in $\mathbb{R}$.
(2) A measurable function $f : S \to [0, \infty]$ is integrable if and only if $\int_S f \, d\mu < \infty$.
(3) If $f : S \to \overline{\mathbb{R}}$ is integrable, then $\mu(\{s \in S : f(s) = \pm\infty\}) = 0$ (see Exercise 5.1).

To check integrability the following characterization is very useful.

Proposition 5.12. *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 $E \in \mathcal{A}$,

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

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

$$\Big| \int_E f \, d\mu \Big| = |I^+ - I^-| \leq I^+ + I^- = \int_E f^+ + f^- \, d\mu = \int_E |f| \, d\mu.$$

☺

By considering the positive and negative part separately one can check Proposition 5.8 (i),(ii), (v) hold for all integrable functions $f, g : S \to \overline{\mathbb{R}}$ again (see Exercise 5.4). The following example shows that Parts (iii) and (vi) do not extend to this setting.

Example 5.13. 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

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

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

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

Proposition 5.14. *Let $f, g : S \to \overline{\mathbb{R}}$ be integrable functions and $\alpha, \beta \in \mathbb{R}$. Then*[39] *$\alpha f + \beta g$ is integrable, and for all $E \in \mathcal{A}$*

$$(\textit{linearity}) \qquad \int_E \alpha f + \beta g \, d\mu = \alpha \int_E f \, d\mu + \beta \int_E g \, d\mu.$$

Proof. Note that by Proposition 5.12 each of the following functions is integrable $f^\pm$, $g^\pm$, $|f|$, $|g|$. Therefore, $|\alpha| \, |f| + |\beta| \, |g|$ is integrable by Proposition 5.8 (iv). Since $|\alpha f + \beta g| \leq |\alpha| \, |f| + |\beta| \, |g|$, the function $\alpha f + \beta g$ is integrable as well (see Exercise 5.7).

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

It remains to prove the identity for each $E \in \mathcal{A}$. To do this we first consider the case $\alpha = \beta = 1$. Write $f + g = \phi - \psi$, where $\phi = f^+ + g^+$ and $\psi = (f^- + g^-)$. Then by Proposition 5.8 (iv) ϕ and ψ are integrable and Exercise 5.3 yields

$$\begin{aligned} \int_E f + g \,\mathrm{d}\mu &= \int_E \phi \,\mathrm{d}\mu - \int_E \psi \,\mathrm{d}\mu \\ &= \int_E f^+ \,\mathrm{d}\mu + \int_E g^+ \,\mathrm{d}\mu - \int_E f^- \,\mathrm{d}\mu - \int_E g^- \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu + \int_E g \,\mathrm{d}\mu. \end{aligned} \tag{5.6}$$

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

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

If $\alpha \geq 0$, then $\alpha f = (\alpha f)^+ - (\alpha f)^- = \alpha f^+ - \alpha f^-$. Proposition 5.8 (iv) yields

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

In case $\alpha < 0$, then by (5.6) and the previous case, we find

$$0 = \int_E \alpha f + (-\alpha) f \,\mathrm{d}\mu = \int_E \alpha f \,\mathrm{d}\mu + \int_E (-\alpha) f \,\mathrm{d}\mu = \int_E \alpha f \,\mathrm{d}\mu + (-\alpha) \int_E f \,\mathrm{d}\mu$$

and the result follows by subtracting the second term on both sides. ☺

Example 5.15. Every simple function $f : S \to \mathbb{R}$ given by $f = \sum_{j=1}^n x_j \mathbf{1}_{A_j}$ with $A_1, \ldots, A_n \in \mathcal{A}$ and $\mu(A_j) < \infty$ for all j, is integrable and by Proposition 5.14,

$$\int_E f \,\mathrm{d}\mu = \sum_{j=1}^n x_j \int_E \mathbf{1}_{A_j} \,\mathrm{d}\mu = \sum_{j=1}^n x_j \mu(E \cap A_j).$$

In the next example we show that for continuous functions the integral with respect to the Lebesgue measure coincides with the Riemann integral.[40]

Example 5.16. Let $a < b$ be real numbers. Let λ be the restriction of the Lebesgue measure to $\mathcal{B}([a, b])$ and let $f : [a, b] \to \mathbb{R}$ be continuous. Note that f is measurable by Proposition 4.4. Moreover, since there is a constant $M \geq 0$ such that $|f| \leq M$, we can use Exercise 5.7 to deduce that f is (Lebesgue) integrable. Since f is continuous, it is Riemann integrable. Below we show that $\mathrm{L}(f) = \mathrm{R}(f)$, where we used the abbreviations

$$\mathrm{L}(f) := \int_{[a,b]} f \,\mathrm{d}\lambda \ \ \text{(Lebesgue integral)} \quad \text{and} \quad \mathrm{R}(f) := \int_a^b f(x) \,\mathrm{d}x \ \ \text{(Riemann integral)}.$$

To prove this 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 $\frac{b-a}{n} < \delta$. Let $x_j = a + j\frac{b-a}{n}$ for $j = 0, \ldots, n$. Let $m_j = \min\{f(x) : x \in [x_{j-1}, x_j]\}$ and $M_j = \max\{f(x) : x \in [x_{j-1}, x_j]\}$. Let $g(x) = \mathbf{1}_{\{a\}} f(a) + \sum_{j=1}^n \mathbf{1}_{(x_{j-1}, x_j]} m_j$ and $G(x) = \mathbf{1}_{\{a\}} f(a) + \sum_{j=1}^n \mathbf{1}_{(x_{j-1}, x_j]} M_j$. Then g and G are integrable in the sense of Riemann and Lebesgue and one can check

$$\mathrm{L}(g) = \mathrm{R}(g) = \sum_{j=1}^n (x_j - x_{j-1}) m_j =: \alpha \ \text{ and } \ \mathrm{L}(G) = \mathrm{R}(G) = \sum_{j=1}^n (x_j - x_{j-1}) M_j =: \beta.$$

Since $g \leq f \leq G$, by monotonicity we find that $\mathrm{L}(f), \mathrm{R}(f) \in [\alpha, \beta]$. Therefore,

$$|\mathrm{R}(f) - \mathrm{L}(f)| \leq \beta - \alpha = \sum_{j=1}^n (x_j - x_{j-1})(M_j - m_j) \leq \sum_{j=1}^n \frac{b-a}{n} \frac{\varepsilon}{b-a} = \varepsilon,$$

Since $\varepsilon > 0$ is arbitrary, we find that $\mathrm{R}(f) = \mathrm{L}(f)$.

[40] It is even known that every Riemann integrable function f is Lebesgue integrable and the integrals coincide (see [1, Theorem 23.6]). The proof uses a similar method as ours. However, one should be aware that not every *improper* Riemann integral is Lebesgue integrable. See Theorem 6.12 and Exercise 6.10 for more details on improper Riemann integrals

Example 5.17. 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 2.5. Then for any $f : \mathbb{R} \to \mathbb{R}$.

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

Indeed, for simple functions f this is obvious. For $f : S \to [0, \infty]$ this follows by approximation from below by simple functions. For general $f : S \to [0, \infty]$ this follows by writing $f = f^+ - f^-$.

Example 5.18. [41] Consider the setting of the counting measure of Example 5.9. A function $f : \mathbb{N} \to \mathbb{R}$ is integrable if and only if $\sum_{j=1}^{\infty} |f(j)| < \infty$. In that case

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

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

Definition 5.19. *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 let*

$$\int_E f \, \mathrm{d}\mu = \int_E u \, \mathrm{d}\mu + i \int_E v \, \mathrm{d}\mu, \quad E \in \mathcal{A}.$$

Exercise 5.11 yields that Propositions 5.12 and 5.14 extend to the complex setting.

Exercises about general integrals

In the exercises below $(S, \mathcal{A}, \mu)$ denotes a measure space.

Exercise 5.1. Let $f : S \to [0, \infty]$ be an integrable function. Show that $f < \infty$ a.e.

Exercise 5.2.

(a) Prove Proposition 5.8 (iv).
  Hint: Approximate by simple functions as in the definition of the integral.
(b) Prove Proposition 5.8 (ii),
  Hint: Write $\mathbf{1}_E g = \mathbf{1}_E (g - f) + \mathbf{1}_E f$ and use Proposition 5.8 (iv).
(c) Prove Proposition 5.8 (iii).
(d) Prove Proposition 5.8 (v).

The following was used in the proof of Proposition 5.14.

Exercise 5.3. Assume $g, h : S \to [0, \infty]$ and $f : S \to \overline{\mathbb{R}}$ are all integrable functions and $f = g - h$. Show that

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

Hint: Note that $f^+ + h = f^- + g$ and use Proposition 5.8 (iv).

Exercise 5.4. Extend the following results to integrable functions $f, g : S \to \overline{\mathbb{R}}$:

(a) Proposition 5.8 (i). First show that $\mathbf{1}_E f$ is integrable for $E \in \mathcal{A}$.
(b) Proposition 5.8 (ii).
(c) Proposition 5.8 (v)

Exercise 5.5. 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: First try to prove this in the case f and g take values in $\mathbb{R}$. In this case apply Proposition 5.8 (vi) to $h = |f - g|$. For the case where f and g are $\overline{\mathbb{R}}$-valued one needs a careful analysis of positive and negative parts.

[41] For details see Exercise 5.8

Exercise 5.6. 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 $-\infty < a < b < \infty$

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

Is this true if we replace λ by an arbitrary measure on $\mathcal{B}(\mathbb{R})$?

Exercise* 5.7 (Domination). Let $f, g : S \to \mathbb{R}$ be measurable functions.

(a) Assume $|f| \le g$ 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 exercise you are asked to give the details of Example 5.18

Exercise* 5.8. Consider the setting of the counting measure of Example 5.9.

(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

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

Hint: Write $f = f^+ - f^-$ and approximate f^+ and f^- in a similar way as in Example 5.9.

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 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 5.14 for integrable functions $f, g : S \to \mathbb{C}$ and $\alpha, \beta \in \mathbb{C}$.
(c) Prove Proposition 5.12 for integrable functions $f : S \to \mathbb{C}$.
(d) Which parts of Proposition 5.8 remain true for integrable functions $f, g : S \to \mathbb{C}$?

6. Convergence theorems and applications

In this section $(S, \mathcal{A}, \mu)$ is a measure space.

One of the problems with the Riemann integral is that the cases where limit and integral can be interchanged are rather limited. 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.

We start with several simple examples which illustrate some difficulties.

Example 6.1. For instance if $(x_j)_{j\geq 1}$ is an enumerate of $\mathbb{Q}$, and $A_n = \{x_1, \ldots, x_n\}$, then $\mathbf{1}_{A_n} \to \mathbf{1}_{\mathbb{Q}}$ pointwise. Each $\mathbf{1}_{A_n}$ is Riemann integrable, but $\mathbf{1}_{\mathbb{Q}}$ is not.

Example 6.2. 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 $\int_{\mathbb{R}} f_n d\lambda \not\to 0$.

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);[42]
- Fatou's lemma;[43]
- Dominated Convergence Theorem (DCT).[44]

Theorem 6.3 (Monotone Convergence Theorem (MCT))**.** *Let $(f_n)_{n\geq 1}$ be measurable functions such that $0 \leq f_n \uparrow f$. Then f is measurable and*

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

Proof. By Theorem 4.9, $f : S \to [0, \infty]$ is measurable. By monotonicity of the integral

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

It remains to prove $\int_S f \, \mathrm{d}\mu \leq \alpha$. For this choose a sequence of simple functions $(g_m)_{m\geq 1}$ such that $0 \leq g_m \uparrow f$. We claim that $\int_S g_m \, \mathrm{d}\mu \leq \alpha$ for each $m \geq 1$. The required estimate follows from the claim by letting $m \to \infty$ and using the definition of $\int_S f \, \mathrm{d}\mu$. To prove the claim we repeat part of the argument of Lemma 5.6. Fix $m \in \mathbb{N}$ and write $g = g_m$.

Let $\varepsilon \in (0, 1)$ and for each $n \geq 1$, set $E_n = \{s \in S : (1 - \varepsilon)g(s) \leq f_n(s)\}$. Then using the indicated parts of Proposition 5.8 in the estimates below, we find

$$(1-\varepsilon)\int_{E_n} g \, \mathrm{d}\mu \overset{\text{(iv)}}{=} \int_{E_n} (1-\varepsilon) g \, \mathrm{d}\mu \overset{\text{(ii)}}{\leq} \int_{E_n} f_n \, \mathrm{d}\mu \overset{\text{(iii)}}{\leq} \int_S f_n \, \mathrm{d}\mu \overset{\text{(ii)}}{\leq} \alpha.$$

Finally, $\int_{E_n} g \, \mathrm{d}\mu \to \int_S g \, \mathrm{d}\mu$ follows from (5.5) with $E = S$ in Lemma 5.6. Therefore, we can conclude $(1 - \varepsilon)\int_S g \, \mathrm{d}\mu \leq \alpha$ and the claim follows since $\varepsilon \in (0, 1)$ was arbitrary. ☺

Lemma 6.4 (Fatou's Lemma)**.** *Let $(f_n)_{n\geq 1}$ be a sequence of measurable functions with values in $[0, \infty]$. 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 4.9. Fix $n \in \mathbb{N}$ and let $g_n = \inf\limits_{k\geq n} f_k$. For all $m \geq n$, we have $g_n \leq f_m$ and by the monotonicity of the 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

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

---

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

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

[44] This is due to Lebesgue (see footnote 18)

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

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

☺

Theorem 6.5 (Dominated Convergence Theorem (DCT))**.** *Let $(f_n)_{n\ge1}$ be a sequence of measurable functions such that $\lim_{n\to\infty} f_n = f$ pointwise. If there exists an integrable function $g : S \to [0,\infty]$ such that $|f_n| \le g$ for all $n \ge 1$, 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. The function f is measurable by Theorem 4.9. Moreover, since also $|f| \le g$, we see that f is integrable. The same holds for each f_n. Let $x_n := \int_S f_n \,\mathrm{d}\mu$ and $x := \int_S f \,\mathrm{d}\mu$. It suffices to show that $\limsup_{n\to\infty} x_n \le x \le \liminf_{n\to\infty} x_n$. It follows that

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

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

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

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 for all $s \in S \setminus A$, $f_n(s) \to f(s)$. For the details of the following remark we refer to Exercise 6.5.

Remark 6.6. Let $f, f_1, f_2, \ldots : S \to \overline{\mathbb{R}}$ be measurable functions.

(1) Theorem 6.3 also holds under the weaker assumption that $0 \le f_n \uparrow f$ a.e.
(2) Theorem 6.5 also holds under the weaker assumptions that $f_n \to f$ a.e. and $|f_n| \le g$ a.e.

Example 6.7. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be integrable with respect to the Lebesgue measure λ. 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 $x_n < x$ and $x_n \to x$, then

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

where the last step follows from Exercise 5.5 and the fact that $\mathbf{1}_{(-\infty,x)} f = \mathbf{1}_{(-\infty,x]} f$ almost everywhere. Similarly, one checks that $F(x_n) \to F(x)$ if $x_n \ge x$ and $x_n \to x$.

6.2. **Consequences and applications.** We continue with several consequences and applications. We start with a result on integration of series of positive functions.

Corollary 6.8 (Series and integrals)**.** *Let $f_1, f_2, \ldots : S \to [0,\infty]$ be measurable functions. Then*

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

Proof. This follows from the MCT (See Exercise 6.4). ☺

From a measure μ one can build many other measures in the following way:

Theorem 6.9 (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.[45] *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 4.12 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

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

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 6.10. Let $f : \mathbb{R} \to [0,\infty]$ be measurable. Define $\nu : \mathcal{B}(\mathbb{R}^d) \to [0,\infty)$ by

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

Then ν is a measure. For example one could take $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Then ν is the standard Gaussian measure on $\mathbb{R}$.

In Example 5.16 we have discussed a connection between the Riemann and Lebesgue integral. A similar connection holds for improper Riemann integrals.

Definition 6.11. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. We say that $\int_{-\infty}^{\infty} f(x) \,\mathrm{d}x$ exists as an **improper** Riemann integral if the limits $L_1 = \lim_{t\to\infty} \int_0^t f(x) \,\mathrm{d}x$ and $L_2 = \lim_{t\to-\infty} \int_t^0 f(x) \,\mathrm{d}x$ exist in $\mathbb{R}$. In that case define $\int_{-\infty}^{\infty} f(x) \,\mathrm{d}x = L_1 + L_2$.*

We show that there is a connection with the Lebesgue integral with respect λ.

Theorem 6.12. *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then the following are equivalent:*

(i) f is integrable;
(ii) $\int_{-\infty}^{\infty} |f(x)| \,\mathrm{d}x$ exists as an improper Riemann integral.

Moreover, in this case $\int_{-\infty}^{\infty} f \,\mathrm{d}x$ exists as am improper Riemann integral and

$$\int_{\mathbb{R}} f \,\mathrm{d}\lambda = \int_{-\infty}^{\infty} f \,\mathrm{d}x. \tag{6.5}$$

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

[45] The function f is usually called the density of ν.

Proof. (i)⇒(ii): First we make a general observation. Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous and integrable function. In Example 5.16 we have seen that $\int_0^t g(x)\,\mathrm{d}x = \int_{\mathbb{R}} \mathbf{1}_{[0,t]} g\,\mathrm{d}\lambda$. Setting $L_1 = \int_{\mathbb{R}} \mathbf{1}_{[0,\infty]} g\,\mathrm{d}\lambda$, we find

$$\Big|L_1 - \int_0^t g(x)\,\mathrm{d}x\Big| = \Big|L_1 - \int_{\mathbb{R}} \mathbf{1}_{[0,t]} g\,\mathrm{d}\lambda\Big| = \Big|\int_{\mathbb{R}} \mathbf{1}_{[t,\infty)} g\,\mathrm{d}\lambda\Big| \le \int_{\mathbb{R}} \mathbf{1}_{[t,\infty)}|g|\,\mathrm{d}\lambda.$$

Now the DCT yields that $\int_{\mathbb{R}} \mathbf{1}_{[t,\infty)}|g|\,\mathrm{d}\lambda \to 0$ as $t \to \infty$, and we may conclude that $\int_0^\infty g(x)\,\mathrm{d}x = L_1$ exists. The part on $(-\infty, 0]$ goes similarly and we find (6.5) with f replaced by g.

Now (i) ⇒ (ii) follows by letting $g = |f|$. Moreover, (6.5) for f follows by taking $g = f$.

(ii)⇒ (i): By Example 5.16, monotonicity of improper Riemann integrals,

$$\int_{\mathbb{R}} \mathbf{1}_{[-n,n]}|f|\,\mathrm{d}\lambda = \int_{-n}^n |f(x)|\,\mathrm{d}x \le \int_{-\infty}^\infty |f(x)|\,\mathrm{d}x =: M < \infty.$$

Therefore, by the MCT, $\int_{\mathbb{R}} |f|\,\mathrm{d}\lambda = \lim_{n\to\infty} \int_{\mathbb{R}} \mathbf{1}_{[-n,n]}|f|\,\mathrm{d}\lambda \le M$, and hence f is integrable. ☺

Example 6.13. 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 $0 \le \left(1 + \frac{x}{n}\right)^n \uparrow e^x$ for every $x \in [0,\infty)$ and thus $0 \le \mathbf{1}_{[0,n]} f_n \uparrow f$, where $f(x) = e^{-x}$. Therefore, by Example 5.16

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

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

Theorem 6.14 (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 $y_0 \in (a,b)$ there exists a $\delta > 0$ and integrable function $g : S \to [0,\infty]$ such that

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

(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 $n \ge 1$ be such that $h_n \to 0$. Let $\phi_n : (y_0 - \delta, y_0 + \delta) \times S \to \mathbb{R}$ and $F : (y_0 - \delta, y_0 + \delta) \to \mathbb{R}$ be given by

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

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 4.12 yields that $s \mapsto \phi_n(y,s)$ is measurable. From the mean value theorem[46] 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 $|\phi_n(y_0,s)| \le g(s)$ for all $s \in S$. It follows that

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

☺

Exercises

Exercise 6.1. Use convergence theorems to find $\lim_{n\to\infty} \int_{\mathbb{R}} f_n\,\mathrm{d}\lambda$ in each of the following cases:

(a) $f_n(x) = \mathbf{1}_{[4,32]}(x)\left(1 + \frac{\log(4x)}{n}\right)^n$ (use MCT; answer is 2016).

[46] See [11, Theorem 6.2.3]: 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) $f_n(x) = \mathbf{1}_{(1,\infty)}(x)\dfrac{\sin(nx)}{nx^2}$ (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. 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 for all $n \in \mathbb{N}$, $\int_S |f_n| \,\mathrm{d}\mu \leq M$
(ii) $f_n \to f$ pointwise.

Use Fatou's lemma to show that $\int_S |f| \,\mathrm{d}\mu \leq M$.

Exercise 6.3. Find $\displaystyle\lim_{n\to\infty}\sum_{m=1}^{\infty}\frac{1}{m^2+n^2}$.
Hint: Use the counting measure and the DCT.

Exercise 6.4. Deduce Corollary 6.8 from the MCT applied to $s_n = \sum_{j=1}^n f_j$ for $n \geq 1$.

Exercise* 6.5. Give the details of Remark 6.6 (use Exercise 5.5).

Exercise 6.6. Let $(f_n)_{n\geq 1}$ be a sequence of measurable functions with values in $\mathbb{R}$ or $\mathbb{C}$ and $f_n \to f$ 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 $n \to \infty$.
Hint: Apply the DCT in a suitable way.
(b) In the real case we know $\int_S f_n \,\mathrm{d}\mu \to \int_S f \,\mathrm{d}\mu$. Derive this in the complex case as well.
Hint: Use (a) and Exercise 5.11 (c).

Exercise* 6.7. Let $f : S \to \mathbb{R}$ be a measurable function. Define $\nu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ by $\nu(B) = \mu(f^{-1}(B))$. Prove the following:

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

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

Hint: Argue as in Theorem 6.9 step 2.
(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

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

Hint: Argue as in Theorem 6.9 step 3.

Exercise* 6.8. 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 and one of the convergence theorems.

Exercise** 6.9. Use induction and Theorem 6.14 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 the estimate $|\sin(x)| \geq \frac{1}{2}$ 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.[47]

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

7. L^p-spaces

In this section $(S, \mathcal{A}, \mu)$ is measure space. 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 $\mathbb{K} = \mathbb{C}$.

Definition 7.1. *For $p \in [1, \infty)$ let*

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

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 $f : S \to \mathbb{K}$.

If $\|f - g\|_p = 0$, then from Proposition 5.8 (vi) we can conclude $f = g$ a.e. However, we would like to have $f = g$. Therefore, we will identify f and g whenever $f = g$ a.e.[48] So 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 7.2. 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 $L^p(\mathbb{R})$.

7.1. Minkowski and Hölder's inequalities.

Note that for scalars $\alpha \in \mathbb{K}$, $\|\alpha f\|_p = |\alpha| \|f\|_p$. Therefore, the next result can be used to show that $L^p(S)$ is a normed vector space.

Proposition 7.3 (Minkowski's inequality [49]). *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. Observe that for all $a, b \in [0, \infty)$ one has (See Exercise 7.1)

$$(a + b)^p = \inf_{\theta \in (0,1)} \theta^{1-p} a^p + (1 - \theta)^{1-p} b^p. \tag{7.1}$$

It follows from (7.1) 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 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 (7.1). ☺

Another famous inequality which can be proved with the same method is the following.

Proposition 7.4 (Hölder's inequality[50]). *Let $p, q \in (1, \infty)$ satisfy[51] $\frac{1}{p} + \frac{1}{q} = 1$. 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 7.2. ☺

[48] 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.

[49] 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.

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

[51] These exponents are called conjugate exponents. If $p = 2$, then $q = 2$ and in this case the inequality is known as the Cauchy-Schwarz inequality.

7.2. Completeness of L^p.

Theorem 7.5 (Riesz–Fischer[52])**.** *Let $p \in [1, \infty)$. Then $(L^p(S), \|\cdot\|_p)$ is a Banach space.*[53]

Proof. Let $(f_k)_{k=1}^\infty$ be a Cauchy sequence with respect to the norm $\|\cdot\|_p$ of $L^p(S)$. By a standard argument it suffices to show that $(f_k)_{k=1}^\infty$ has a subsequence which is convergent in $L^p(S)$.

Recursively, we can find a subsequence $(f_{k_n})_{n\geq 1}$ such that

$$\|f_{k_{n+1}} - f_{k_n}\|_p \leq \frac{1}{2^n}, \quad n = 1, 2, \ldots \tag{7.2}$$

For notational convenience let $\phi_n = f_{k_n}$ for $n \in \mathbb{N}$. Also let $\phi_0 = 0$. We will show that there exists an $f \in L^p(S)$ such that $\|\phi_n - f\|_p \to 0$.

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

$$g := \sum_{n=0}^{\infty} |\phi_{n+1} - \phi_n| \ \text{ and } \ g_m := \sum_{n=0}^{m-1} |\phi_{n+1} - \phi_n|, \quad ,$$

By Minkowski's inequality we obtain for each $m \geq 1$,

$$\|g_m\|_p \leq \sum_{n=0}^{m-1} \|\phi_{n+1} - \phi_n\|_p = \sum_{n=0}^{\infty} \|\phi_{n+1} - \phi_n\|_p \overset{(7.2)}{\leq} \|\phi_1\|_p + \sum_{n\geq 1} 2^{-n} \leq \|\phi_1\|_p + 1.$$

Since $0 \leq g_m \uparrow g$, the MCT yields,

$$\int_S |g|^p \,\mathrm{d}\mu = \lim_{m\to\infty} \int_S |g_m|^p \,\mathrm{d}\mu = \lim_{m\to\infty} \|g_m\|_p^p \leq (\|\phi_1\|_p + 1)^p.$$

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

$$f = \sum_{n=0}^{\infty} \mathbf{1}_A(\phi_{n+1} - \phi_n),$$

where the series is absolutely convergent and f is measurable by Theorem 4.9. By a telescoping argument it follows that pointwise on S

$$f = \lim_{m\to\infty} \sum_{n=0}^{m-1} \mathbf{1}_A(\phi_{n+1} - \phi_n) = \lim_{m\to\infty} \mathbf{1}_A \phi_m.$$

Clearly,

$$|\phi_m| = \Big|\sum_{n=0}^{m-1} (\phi_{n+1} - \phi_n)\Big| \leq \sum_{n=0}^{m-1} |\phi_{n+1} - \phi_n| \leq |g|$$

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

$$|f - \phi_m|^p \leq (|f| + |\phi_m|)^p \leq (2|g|)^p = 2^p |g|^p.$$

Since $|f - \phi_m|^p \to 0$ a.e. and $2^p|g|^p$ is integrable, it follows from the DCT (see the a.e. version of Remark 6.6) that

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

☺

In the sequel we will say that $f_n \to f$ in $L^p(S)$ if $\|f_n - f\|_p \to 0$.

From the proof of Theorem 7.5 we deduce the following result.

Corollary 7.6. *Let $p \in [1, \infty)$. Suppose $f, f_1, f_2, \ldots \in L^p(S)$. If $f_k \to f$ in $L^p(S)$, then there exists a subsequence $(f_{k_n})_{n\geq 1}$ such that $f_{k_n} \to f$ a.e.*

---

[52] 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.

[53] 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].

In general one does not have $f_n \to f$ a.e. (see Exercise 7.7).

Proof. Indeed, since $(f_n)_{n\geq 1}$ is convergent it is a Cauchy sequence in $L^p(S)$. In the proof of Theorem 7.5 there exists a subsequence $(f_{k_n})_{n\geq 1}$ and a function $\tilde f$ such that $f_{k_n} \to \tilde f$ in $L^p(S)$ and $f_{k_n} \to \tilde f$ a.e. Now Minkowski's inequality yields:

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

Thus $f = \tilde f$ a.e. and thus $f_{k_n} \to f$ a.e. ☺

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

$$\langle f, g\rangle = \int_S f(s)\overline{g(s)}\,\mathrm{d}\mu(s),$$

where $\overline{g(s)}$ stands for the complex conjugate of the number $g(s) \in \mathbb{K}$. Then $\langle \cdot, \cdot\rangle$ is an inner product and $\langle f, f\rangle = \|f\|_2^2$.[54]

The next theorem is immediate from Theorem 7.5.

Theorem 7.7 (Riesz–Fischer)**.** *$L^2(S)$ is a Hilbert space.*[55]

Example 7.8. Let $p \in [1, \infty)$. If $\mu = \tau$ is the counting measure on $\mathbb{N}$, then[56] $\ell^p := L^p(\mathbb{N})$ coincides with a space of sequences and

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

If $p \leq q < \infty$, then $\ell^p \subseteq \ell^q$ (see Exercise 7.8).

Example 7.9. Assume $\mu(S) < \infty$. If $1 \leq p \leq q < \infty$, then $L^q(S) \subseteq L^p(S)$ (see Exercise 7.3).

We end this section with a simple density result. Clearly, the definition of a simple function extends to the complex setting. The following result shows that any L^p-function can be approximated by simple functions.

Proposition 7.10 (Density of simple functions)**.** *Let $p \in [1, \infty)$. The set of simple functions is dense in $L^p(S)$.*

Proof. Write $f = u + iv$, where u and v are both real-valued. Write $u = g - h$, where $g = u^+$ and $h = u^-$. By Theorem 4.12 we can find simple functions g_n and h_n such that $0 \leq g_n \uparrow g$ and $0 \leq h_n \uparrow h$. Then $u_n := g_n - h_n$ is a simple function, $u_n \to u$ pointwise, and $|u_n| \leq g_n + h_n \leq g + h \leq |u|$. Similarly, one constructs simple functions v_n such that $v_n \to v$ pointwise, $|v_n| \leq |v|$. We can conclude that $f_n := u_n + iv_n$ is a simple function, $f_n \to f$ pointwise and $|f_n| \leq (|u_n|^2 + |v_n|^2)^{\frac{1}{2}} \leq (|u|^2 + |v|^2)^{\frac{1}{2}} \leq |f|$. Now since $|f_n - f|^p \leq (|f_n| + |f|)^p \leq 2^p|f|^p$ and the latter is integrable, it follows from the DCT that $\|f_n - f\|_p \to 0$ as $n \to \infty$. ☺

7.3. **L^p-spaces on intervals.** In the remaining part of this section we discuss $L^p(I)$, where I is an interval and $\mu = \lambda$ is the Lebesgue measure on I. These results will not be used in Section 8 or in the exercises.

For an interval $I \subseteq \mathbb{R}$ let the set of step functions $\mathrm{Step}(I)$ be defined by

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

One can check that 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 on L^p can be reduced to this special class by the following density result.

[54] Note that the conjugation is need otherwise the square of a complex number can be negative. Physicists often put the conjugation on $f(s)$ instead of $g(s)$. If $\mathbb{K} = \mathbb{R}$, then the conjugation does not play a role.

[55] Recall that a Hilbert space is a Banach space where $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$. 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 centuries.

[56] Sometimes we wish to use the counting measure on $\mathbb{Z}$ instead of $\mathbb{N}$. In this case we write $\ell^p(\mathbb{Z}) := L^p(\mathbb{Z})$.

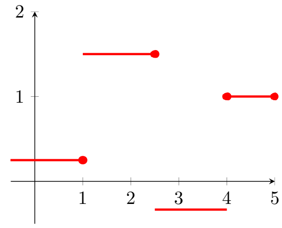

FIGURE 7.1. Example step function $\frac{1}{4}\mathbf{1}_{(-\frac{1}{2},1]} + \frac{3}{2}\mathbf{1}_{(1,\frac{5}{2}]} - \frac{1}{3}\mathbf{1}_{(\frac{5}{2},4)} + \mathbf{1}_{[4,5]}$

Theorem 7.11. *Let λ be the Lebesgue measure on $I = (a,b)$ with $-\infty \le a < b \le \infty$ and let $p \in [1,\infty)$. Then 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, $|f_n - f|^p \le |f|^p$. Therefore, the DCT yields $\|f_n - f\|_p \to 0$. Therefore, for $n \in \mathbb{N}$ large enough $\|f-g\|_p < \varepsilon$, where $g = f_n \in L^p(I)$.

Step 2: By Step 1 we can assume I is bounded, and moreover we can assume $I = (a,b]$ with $-\infty < a < b < \infty$. By Proposition 7.10 there exists a simple function $h : I \to \mathbb{K}$ such that $\|g-h\|_p < \varepsilon$. We can write $h = \sum_{j=1}^n \mathbf{1}_{A_j}x_j$ with $(A_j)_{j=1}^n$ in $\mathcal{B}(I)$ disjoint.

Step 3: By Exercise 7.5 there exist $F_j \in \mathcal{F}_{(a,b]}$ such that $\lambda(A_j \triangle F_j) < \frac{\varepsilon}{n(|x_j|+1)}$. Now let $\phi = \sum_{j=1}^n \mathbf{1}_{F_j}x_j$. Observe that $|\mathbf{1}_{A_j} - \mathbf{1}_{F_j}| = \mathbf{1}_{A_j\triangle F_j}$. Therefore, Minkowski's inequality yields that

$$\|h-\phi\|_p \le \sum_{j=1}^n |x_j|\|\mathbf{1}_{A_j\triangle F_j}\|_p = \sum_{j=1}^n |x_j|\mu(A_j \triangle F_j) \le \sum_{j=1}^n \frac{\varepsilon|x_j|}{n(|x_j|+1)} < \varepsilon.$$

Conclusion: Clearly, ϕ is a step function. Moreover, by Minkowski's inequality we find

$$\|f-\phi\|_p \le \|f-g\|_p + \|g-h\|_p + \|h-\phi\|_p \le 3\varepsilon.$$

☺

Corollary 7.12. *Let λ be the Lebesgue measure on $I = [a,b]$ with $-\infty < a < b < \infty$ 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 7.11 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 an arbitrary $\mathbf{1}_J$. Using Figure 7.2 and the DCT, the reader can easily convince him or herself that this can indeed be done. ☺

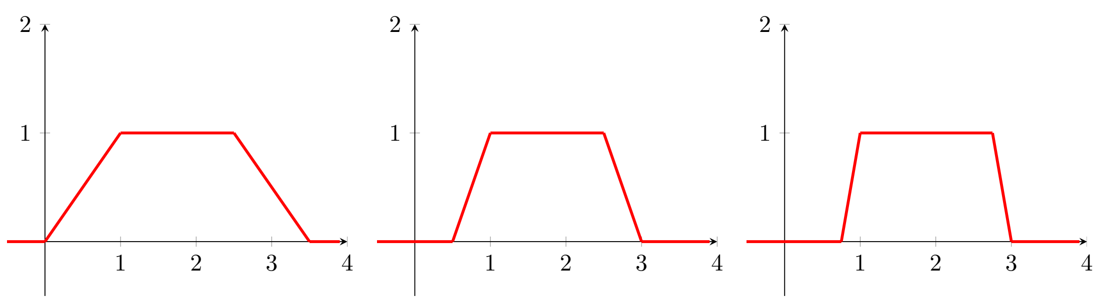

FIGURE 7.2. Approximation of $\mathbf{1}_{(1,\frac{5}{2}]}$ in L^p by continuous functions

Exercises

Exercise 7.1. Prove the identity (7.1).
Hint: First consider the case where $a, b > 0$ and minimize the right-hand side of (7.1).

Exercise* 7.2. Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$.

(a) Show that for all $a, b \geq 0$,
$$ab = \inf_{t>0} \Big(\frac{t^p a^p}{p} + \frac{b^q}{q t^q} \Big).$$
(b) Use the above identity to prove Proposition 7.4.
Hint: Argue as in Proposition 7.3, but use the identity from (a) instead.

Exercise 7.3. Assume $\mu(S) < \infty$ and $1 \leq q \leq p < \infty$. Show that $L^p(S) \subseteq L^q(S)$ and for all $f \in L^p(S)$,
$$\|f\|_q \leq \mu(S)^{\frac{1}{q} - \frac{1}{p}} \|f\|_p.$$
Hint: Apply Hölder's inequality to $|f|^q \cdot 1$.

Exercise 7.4. Let $p \in [1, \infty)$ and λ be the Lebesgue measure on $\mathbb{R}$. 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 $p \neq q$.

For sets $A, B \subseteq \mathbb{R}$, $A \triangle B = (A \setminus B) \cup (B \setminus A)$ denotes the **symmetric difference** of A and B.

Exercise* 7.5. Let λ be the Lebesgue measure restricted to $((a, b], \mathcal{B}((a, b]))$. Let $\mathcal{F}_{(a,b]}$ be the finite unions of half-open intervals in $(a, b]$.[57] Let
$$\mathcal{A} = \{A \in \mathcal{B}((a, b]) : \forall \varepsilon > 0 \ \exists F \in \mathcal{F}_{(a,b]} \ \text{ such that } \ \lambda(A \triangle F) < \varepsilon\}.$$

(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 3.1 and 3.8.

Exercise* 7.6. Let $-\infty < a < b < \infty$ and $f \in L^1((a, b])$ and assume $\int_{(a,t]} f \, \mathrm{d}\lambda = 0$ for all $t \in (a, b]$. We will derive that $f = 0$ in $L^1(\mathbb{R})$. By considering real and imaginary part separately, one can reduce to the case where f is real valued.

(a) Show that for all $A \in \mathcal{F}^1$ with $A \subseteq (a, b]$, $\int_A f \, \mathrm{d}\lambda = 0$.
(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 $A_k \subseteq (-a, b]$ and $\mathbf{1}_{A_n} \to \mathbf{1}_A$ a.e.
Hint: Use $\|\mathbf{1}_A - \mathbf{1}_B\|_1 = \lambda(A \triangle B)$ for all $A, B \in \mathcal{B}(\mathbb{R})$ and Exercise 7.5. Now apply Corollary 7.6 and the DCT.
(c) Show that for all $A \in \mathcal{B}(\mathbb{R})$ with $A \subseteq (a, b]$, $\int_A f \, \mathrm{d}\lambda = 0$.
Hint: Use (a) and (b).
(d) Derive that $f = 0$ in $L^1(\mathbb{R})$.
Hint: Consider $A = \{x \in \mathbb{R} : f(x) \geq 0\}$ and $A = \{x \in \mathbb{R} : f(x) \leq 0\}$.

[57] By Exercise 1.7 of the lecture notes we can take them disjoint

Exercise** 7.7. Let λ be the Lebesgue measure on $\mathbb{R}$. Observe that each $n \in \mathbb{N}$ can be uniquely written as $n = 2^k + j$ with $k \in \mathbb{N}_0$ and $j \in \{0, \dots, 2^k - 1\}$. 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 $f_n(x) = 1$.
Hint: First make a picture for $k = 1$ and $j = 0, 1$, and $k = 2$ and $j = 0, 1, 2, 3$.

Exercise** 7.8. Let $1 \le p \le q < \infty$).

(a) Prove $\ell^p \subseteq \ell^q$ and that for all $(a_n)_{n\ge1} \in \ell^p$ one has $\|(a_n)_{n\ge1}\|_q \le \|(a_n)_{n\ge1}\|_p$.
Hint: By homogeneity one can assume $\|(a_n)_{n\ge1}\|_{\ell^p} = 1$, and therefore $|a_n| \le 1$ 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 $(a_n)_{n\ge1} \in \ell^p$?

There is a natural limiting space of $L^p(S)$ for $p \to \infty$:

Exercise** 7.9. A measurable function $f : S \to \mathbb{K}$ is said to be in $L^\infty(S)$ if there exists an $M \ge 0$ such that $\mu(\{|f| > M\}) = 0$.[58] Define

$$\|f\|_\infty = \inf\{M \ge 0 : \mu(\{|f| > M\}) = 0\}.$$

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 all $f \in L^\infty(S)$ and $g \in L^1(S)$, $fg \in L^1(S)$ and

$$\|fg\|_1 \le \|f\|_\infty \|g\|_1.$$

(c) Assume $\mu(S) < \infty$ and $p \in [1, \infty)$. Show that $L^\infty(S) \subseteq L^p(S)$ and $\|f\|_p \le \mu(S)^{\frac{1}{p}} \|f\|_\infty$.
(d) Assume $S = \mathbb{N}$ with the counting measure and $p \in [1, \infty)$. Let $\ell^\infty := L^\infty(\mathbb{N})$. Show that $\ell^p \subseteq \ell^\infty$ and $\|(a_n)_{n\ge1}\|_\infty \le \|(a_n)_{n\ge1}\|_p$.
(e) Assume I is a finite interval and $\mu = \lambda$ is the Lebesgue measure. Show that the simple functions are dense in $L^\infty(I)$, but the step functions are not.

[58] In order words $|f| \le M$ a.e.

8. Applications to Fourier series

Fourier[59] 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[60]

$$\sum_{n\in\mathbb{Z}} c_n e^{inx}, \quad \text{or equivalently} \quad \frac{a_0}{2} + \sum_{n\geq 1} a_n\cos(nx) + b_n\sin(nx),$$

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 eigenfunctions 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.[61] 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], [13], [14] and [19]. In particular, very interesting but mostly elementary applications to geometry, ergodicity, number theory and PDEs can be found in [13].

8.1. **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 8.1. *Let $e_k : [0, 2\pi] \to \mathbb{C}$ be given by $e_k(x) = e^{ikx}$ for $k \in \mathbb{Z}$.*[62] *Let $f \in L^1(0, 2\pi)$.*

(i) *For $k \in \mathbb{Z}$ the k-**th order Fourier coefficient** is defined by*[63]

$$\widehat{f}(k) = \frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{e_k(x)}\,\mathrm{d}x.$$

(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*

$$s_n(f) = \sum_{|k|\leq n} \widehat{f}(k)e_k. \tag{8.1}$$

(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 8.2.

(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.
(2) Observe that by (the complex version of) Proposition 5.12 for each $k \in \mathbb{Z}$,
$$|\widehat{f}(k)| \leq \frac{1}{2\pi}\int_0^{2\pi} |f(x)e_k(x)|\,\mathrm{d}x = \frac{1}{2\pi}\|f\|_1. \tag{8.2}$$
In Exercise 8.4 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 integral.

[59] 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

[60] Recall Euler's formula: $e^{ix} = \cos(x) + i\sin(x)$ for $x \in \mathbb{R}$

[61] In Complex Function Theory these functions will be characterized as the so-called analytic functions

[62] Note that $e_0 = \mathbf{1}_{[0,2\pi]}$.

[63] To calculate Fourier transforms numerically one can use the so-called fast Fourier transform FFT (see [13, Section 7.1.3])

Example 8.3. 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

$$\sum_{|k|\leq n} c_k e_k = \sum_{k=-n}^{n} c_k e_k \quad \text{as } n \to \infty.$$

8.2. **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 8.5.

Theorem 8.4 (Weierstrass' approximation theorem[64] 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[65] $F_n = \frac{1}{n} \sum_{k=1}^{n} \sum_{|j|\leq k-1} e_j$ which is a periodic function as well. Define $P_n : [0, 2\pi] \to \mathbb{C}$ by $P_n(x) = \int_0^{2\pi} F_n(x - y) f(y)\, dy$. Then since $e_j(x - y) = e^{2\pi i x} e^{-2\pi i y}$ the following identity holds

$$P_n(x) = \frac{1}{2\pi} \int_0^{2\pi} F_n(x-y) f(y)\, dz = \frac{1}{2\pi n} \sum_{k=1}^{n} \sum_{|j|\leq k-1} e_j(x) \int_0^{2\pi} e_j(-y) f(y)\, dy = \frac{1}{n} \sum_{k=1}^{n} \sum_{|j|\leq k-1} e_j(x) \widehat{f}(j).$$

This shows that P_n is a trigonometric polynomial.

By Exercise 8.3 the following identity holds

$$F_n(z) = \frac{\sin^2(nz/2)}{n \sin^2(z/2)}, \quad z \in (0, 2\pi). \tag{8.3}$$

Therefore, $F_n(z) \geq 0$, and thus we can write

$$\|F_n\|_1 = \int_0^{2\pi} F_n(z)\, dz = \frac{1}{n} \sum_{k=1}^{n} \sum_{|j|\leq k-1} \int_0^{2\pi} e_j(z)\, dz = \frac{1}{n} \sum_{k=1}^{n} 2\pi = 2\pi.$$

[64]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 8.11

[65]This is called Féjer's kernel

Fix $x \in [0, 2\pi]$. It follows that

$$\begin{aligned} 2\pi|f(x) - P_n(x)| &= \Big|f(x)\int_0^{2\pi} F_n(x-y)\,\mathrm{d}y - \int_0^{2\pi} f(y)F_n(x-y)\,\mathrm{d}y\Big| \\ &= \Big|\int_0^{2\pi} F_n(x-y)(f(x)-f(y))\,\mathrm{d}y\Big| \\ &\le \int_0^{2\pi} |f(x)-f(y)|F_n(x-y)\,\mathrm{d}y\Big| \le T_1 + T_2, \end{aligned}$$

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

$$T_1 \le \frac{\varepsilon}{2}\int_I F_n(x-y)\,\mathrm{d}y \le \pi\varepsilon.$$

On $J := [0, x-\delta] \cup [x+\delta, 2\pi]$ we can estimate

$$T_2 \le 2\|f\|_\infty \int_J F_n(x-y)\,\mathrm{d}y = 2\|f\|_\infty \int_{B_\delta} F_n(z)\,\mathrm{d}z,$$

where we substituted $z := x - y$ and where $B_\delta = [\delta, 2\pi - \delta]$. Therefore, using (8.3) again and the fact that $|\sin(z/2)| \ge \sin(\delta/2)$ for $z \in B_\delta$ (recall that $\delta \le \pi$) we obtain

$$\int_{B_\delta} F_n(z)\,\mathrm{d}z \le \frac{|B_{\delta|}}{n\sin^2(\delta/2)} \le \frac{2\pi}{n\sin^2(\delta/2)}.$$

So choosing $n \ge 1$ 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 $|f(x) - P_n(x)| \le \frac{T_1+T_2}{2\pi} < \varepsilon$. 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 8.5 (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. [66] *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 8.4 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 $f = 0$.

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 6.7), and $F(0) = F(2\pi) = 0$.[67] By Exercise 8.6 (b) $\widehat{F}(k) = \frac{\widehat{f}(k)}{ik} = 0$ for all $k \ne 0$. 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 7.6 we find $f = 0$.

☺

[66] 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.

[67] Note that $F(2\pi) = (2\pi)\widehat{f}(0) = 0$.

8.3. **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 7.3). 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 7.3 for $f,g\in L^2(0,2\pi)$,

$$|\langle f,g\rangle| \leq \|f\|_2\|g\|_2 \quad \text{(Cauchy–Schwarz inequality).} \tag{8.4}$$

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[68]

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2. \tag{8.5}$$

Lemma 8.6 (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

$$\Big\|\sum_{j\in\mathbb{Z}} c_j e_j\Big\|_2 = (2\pi)^{\frac12}\Big(\sum_{j\in\mathbb{Z}}|c_j|^2\Big)^{\frac12}. \tag{8.6}$$

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

$$\Big\|\sum_{j\in\mathbb{Z}} c_j e_j\Big\|_2^2 = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}} c_j\overline{c_k}\langle e_j,e_k\rangle = 2\pi\sum_{j\in\mathbb{Z}}|c_j|^2.$$

☺

We extend this result to series using the completeness of $L^2(0,2\pi)$.

Theorem 8.7 (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

$$\|g\|_2 = (2\pi)^{\frac12}\|(c_n)_{n\in\mathbb{Z}}\|_{\ell^2} \qquad \textit{(Parseval's identity)} \tag{8.7}$$

(ii) If $f\in L^2(0,2\pi)$, then $(\widehat{f}(n))_{n\in\mathbb{Z}}$ in ℓ^2 and $f = \sum_{n\in\mathbb{Z}}\widehat{f}(n)e_n$ in $L^2(0,2\pi)$ and (8.7) holds with $g=f$ and $c_n = \widehat{f}(n)$ for $n\in\mathbb{Z}$.

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 8.8).

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

$$\Big(\sum_{|k|\geq N}|c_k|^2\Big)^{\frac12} < \frac{\varepsilon}{(2\pi)^{\frac12}}.$$

[68] This follows by writing out $\|f+g\|_2^2 = \langle f+g, f+g\rangle$.

Then for all integers $n > m \geq N$ by Lemma 8.6,

$$\|g_n - g_m\|_2 = \Big\| \sum_{m<|k|\leq n} c_k e_k \Big\|_2 = (2\pi)^{\frac{1}{2}} \Big(\sum_{m<|k|\leq n} |c_k|^2 \Big)^{\frac{1}{2}} \leq (2\pi)^{\frac{1}{2}} \Big(\sum_{|k|\geq N} |c_k|^2 \Big)^{\frac{1}{2}} < \varepsilon.$$

This proves that $(g_n)_{n\geq 1}$ is a Cauchy sequence. By completeness (see Theorem 7.7), $g := \lim_{n\to\infty} g_n$ exists in $L^2(0, 2\pi)$.

To check (8.7) note that by the continuity of $\|\cdot\|_2$ and Lemma 8.6,

$$\|g\|_2 = \lim_{n\to\infty} \|g_n\|_2 = \lim_{n\to\infty} (2\pi)^{\frac{1}{2}} \Big(\sum_{|k|\leq n} |c_k|^2 \Big)^{\frac{1}{2}} = (2\pi)^{\frac{1}{2}} \|(c_n)\|_{\ell^2}.$$

Finally, note that $\widehat{g}(k) = (2\pi)^{-1} \langle g, e_k \rangle = \lim_{n\to\infty} (2\pi)^{-1} \langle g_n, e_k \rangle = c_k$.[69]

(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 (8.5) yields

$$\|f\|_2^2 = \|f - s_n(f) + s_n(f)\|_2^2 = \|f - s_n(f)\|_2^2 + \|s_n(f)\|_2^2 \geq \|s_n(f)\|_2. \tag{8.8}$$

Let $c_k = \widehat{f}(k)$ for $k \in \mathbb{Z}$. Then (8.8) yields:

$$\|f\|_2^2 \geq \|s_n(f)\|_2 \overset{(8.6)}{=} \sum_{|k|\leq n} 2\pi |c_k|^2.$$

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 8.5 applied to $f - g$. ☺

For $f \in L^2(0, 2\pi)$ Theorem 8.7 yields that $f - s_n(f) = \sum_{|k|>n} \widehat{f}(k) e_k$ and by (8.7)

$$(L^2\text{-error estimate}) \qquad \|f - s_n(f)\|_2^2 = 2\pi \sum_{|k|>n} |\widehat{f}(k)|^2. \tag{8.9}$$

Moreover, since $s_n(f)$ is a trigonometric polynomial, we also find the following:[70]

Corollary 8.8. *The trigonometric polynomials are dense in $L^2(0, 2\pi)$.*

Example 8.9 (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 (see [11, Theorem 7.3.1]) and integration by parts,

$$\widehat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} (x - \pi) e^{-ikx} \,dx = \frac{1}{2\pi} \Big[\frac{(x - \pi) e^{-ikx}}{-ik} \Big]_0^{2\pi} - \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} \,dx = -\frac{1}{ik}. \tag{8.10}$$

Clearly, $\widehat{f}(0) = 0$. Therefore, Theorem 8.7 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 8.1) for plots of the partial sums).

The L^2-error can be estimated using (8.9):

$$\|f - s_n(f)\|_2^2 = 2\pi \sum_{|k|>n} |\widehat{f}(k)|^2 \leq 4\pi \sum_{k\geq n+1} \frac{1}{k^2} \leq 4\pi \int_n^{\infty} \frac{1}{x^2} \,dx = \frac{4\pi}{n}.$$

One can show that $s_n(f)$ will not converge to f uniformly (also see Figure 8.1). 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 (8.7) one can obtains a remarkable identity:

$$\|f\|_2^2 = 2\pi \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k^2}.$$

[69] Here we used $|\langle g - g_n, e_k \rangle| \leq \|g - g_n\|_2 \|e_k\|_2$ as follows from (8.5).

[70] With more advanced techniques one can show that the trigonometric polynomials are dense in any $L^p(0, 2\pi)$ with $p \in [1, \infty)$.

On the other hand, if we calculate $\|f\|_2^2$ with the fundamental theorem of calculus, we obtain

$$\|f\|_2^2 = \int_0^{2\pi} (x-\pi)^2 \,\mathrm{d}x = \left[\tfrac{1}{3}(x-\pi)^3\right]_0^{2\pi} = \frac{2}{3}\pi^3.$$

Combining both identities gives $\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{k^2} = \frac{1}{3}\pi^2$, and so we find $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{1}{6}\pi^2$.

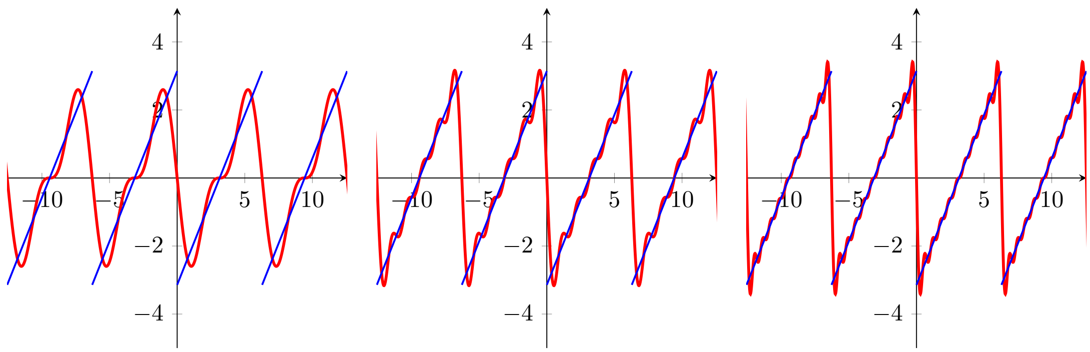

FIGURE 8.1. The Fourier series of the sawtooth function with $n = 2$, $n = 5$ and $n = 10$.

8.4. **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

$$\|f_n - f\|_2^2 \le \int_0^{2\pi} |f_n(x) - f(x)|^2 \,\mathrm{d}x \le \|f_n - f\|_\infty^2 \int_0^{2\pi} 1 \,\mathrm{d}x = 2\pi\|f_n - f\|_\infty^2. \tag{8.11}$$

From the above we see that convergence of Fouier 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 $L^2(0, 2\pi)$.[71]

$$\begin{array}{ccl} & & L^2\text{-conv.} \Longrightarrow L^1\text{-conv.} \Longrightarrow \text{a.e.-conv. subsequence.} \\ & \nearrow & \\ \text{uniform conv.} & & \\ & \searrow & \\ & & \text{pointwise conv.} \Longrightarrow \text{a.e.-conv.} \end{array}$$

L^1 and L^2-convergence are only implied by a.e. conv. 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.

Theorem 8.10 (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 $x \in [0, 2\pi]$,

$$\sum_{k\in\mathbb{Z}} |\widehat{f}(k)e_k(x)| = \sum_{k\in\mathbb{Z}} |\widehat{f}(k)| < \infty.$$

[71]For completeness we note that a.e.-conv. $\Longrightarrow$ conv. in measure $\Longrightarrow$ conv. in distribution.

Therefore, we can let $g = \sum_{k\in\mathbb{Z}} \widehat{f}(k)e_k$, where the series is absolutely convergent. Moreover,

$$\|g - s_n(f)\|_\infty \leq \sum_{|k|>n} |\widehat{f}(k)| \to 0,$$

and hence $g \in C([0, 2\pi])$ and $s_n(f) \to g$ uniformly. By (8.11) the convergence holds in $L^2(0, 2\pi)$ as well, and hence

$$\widehat{g}(k) = \langle g, e_k\rangle = \lim_{n\to\infty} \langle s_n(f), e_k\rangle = \widehat{f}(k)$$

and therefore, $g = f$ a.e. by Theorem 8.5. 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 $0 < \lambda(I) \leq \lambda([0, 2\pi] \setminus A) = \lambda([0, 2\pi]) - \lambda(A) = 0$. 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:

$$\text{(uniform error estimate)} \qquad \|f - s_n(f)\|_\infty \leq \sum_{|k|>n} |\widehat{f}(k)|. \tag{8.12}$$

The condition of Theorem 8.10 holds in the following situation:

Corollary 8.11. *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 6.7 that F is continuous on $[0, 2\pi]$. Moreover, for every $t \in [0, 2\pi]$,

$$|F(t)| \leq |c_0| + \int_0^{2\pi} |f(x)|\,\mathrm{d}x \leq |c_0| + \|f\|_1.$$

In particular, by monotonicity we see that $\|F\|_2 \leq (2\pi)^{\frac{1}{2}}(|c_0| + \|f\|_1)$.

By Exercise 8.6, $\widehat{F}(k) = \frac{\widehat{f}(k)}{ik}$ for $k \neq 0$. By Theorem 8.7, $\|\widehat{f}(k))_{k\in\mathbb{Z}}\|_{\ell^2} = \|f\|_2$. Therefore, by the Cauchy–Schwarz inequality (8.4)

$$\sum_{k\in\mathbb{Z}} |\widehat{F}(k)| = |\widehat{F}(0)| + \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{|\widehat{f}(k)|}{|k|} \leq |\widehat{F}(0)| + C\|(\widehat{f}(k))_{k\in\mathbb{Z}}\|_{\ell^2} \leq |\widehat{F}(0)| + \frac{C}{(2\pi)^{\frac{1}{2}}}\|f\|_2,$$

where used (8.7) in the last step. Therefore, the absolute and uniform convergence follows from Theorem 8.10. ☺

Example 8.12. 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 8.11.

Exercises

Exercise 8.1. 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 $k \neq 0$, and $\widehat{f}(k) = \frac{1}{\pi ik}$ for odd k, and $\widehat{f}(0) = \frac{1}{2}$.
(b) Write f as a series of sines as in Example 8.9 and give an estimate of L^2-error given by (8.9).
(c) Evaluate $\sum_{j\in\mathbb{Z}} \frac{1}{(2j+1)^2}$.
Hint: Argue as in Example 8.9.

Exercise 8.2. Let $f : [0, 2\pi) \to \mathbb{R}$ be given by $f(x) = |x - \pi|$. Example 8.12 yields that the Fourier series of f is absolutely and uniformly convergent. Calculate the Fourier series explicitly and give estimates for the L^2-error (8.9) and uniform error (8.12). One can also obtain the exact form of another famous series $\sum_{k\in\mathbb{Z}} |\widehat{f}(k)|^2$ using Parseval and the fundamental theorem of calculus.

Exercise 8.3 (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 $D_n(x) := \sum_{|k| \le n-1} e_k(x) = \dfrac{\sin((n-\frac{1}{2})x)}{\sin(\frac{1}{2}x)}$ for $n \ge 1$.

(b) (Féjer kernel) Show that $F_n(x) := \dfrac{1}{n} \sum_{j=1}^{n} D_j(x) = \dfrac{1}{n} \dfrac{\sin^2(n\frac{x}{2})}{\sin^2(\frac{1}{2}x)}$ for $n \ge 1$.

Exercise* 8.4 (Riemann-Lebesgue lemma).

(a) Show that for any step function $f : [0, 2\pi] \to \mathbb{C}$ (see Section 7.3) 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 7.11 and (a).

Exercise* 8.5.

(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 (8.7) 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* 8.6. Let $g \in C^1([0, 2\pi])$[72] 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 6.7, 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 7.12 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 $g = e_{-k}$.

Exercise* 8.7. 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 8.6 (b).

(b) Show that $\|F\|_2 \le \|f\|_2$ and that equality holds if and only if $F = c_1e_1 + c_{-1}e_{-1}$ for $c_1, c_{-1} \in \mathbb{C}$.
Hint: Apply (8.7).

Exercise* 8.8. 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 $\phi \ne \psi$ 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 8.5

(c) Show that for every $(a_n)_{n\ge0}, (b_n)_{n\ge1} \in \ell^2$ the following series converges in $L^2(0, 2\pi)$.

$$g := \frac{a_0}{2} + \sum_{n\ge1} a_n \cos(n\cdot) + \sum_{n\ge1} b_n \sin(n\cdot)$$

Hint: Argue as in Theorem 8.7.

(d) Show that $\|g\|^2_{L^2(0,2\pi)} = \pi \sum_{n\ge0} |a_n|^2 + \pi \sum_{n\ge1} |b_n|^2$.
Hint: Argue as in Theorem 8.7.

[72]That means g is differentiable and its derivative is continuous on $[0, 2\pi]$

(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 8.7.

(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 8.7 or argue as in Theorem 8.7.

It follows from the previous exercise and (8.2) 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 8.4 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** 8.9. 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** 8.10. 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** 8.11 (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 8.4 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 8.10 (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 8.4 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**[73] 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 [74] $A_n \uparrow A$, 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)$\Rightarrow$(ii): This is Exercise A.1.

(ii)$\Rightarrow$(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 $A = \bigcup\limits_{j=1}^{\infty} A_j$ and $B_n = \bigcup\limits_{j=1}^{n} A_j$ for $n \geq 1$. Since $\mathcal{F}$ is a π-system it is closed under finite intersections. Therefore, using (D2) we obtain $B_n = \Big(\bigcap\limits_{j=1}^{n} A_j^c\Big)^c \in \mathcal{F}$. 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

$$\mathcal{D}_1 = \{D \in \mathcal{D}_0 : D \cap E \in \mathcal{D}_0 \quad \text{for each} \quad E \in \mathcal{E}\}.$$

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 $B \subseteq A$, 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 π-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 $E \in \mathcal{E}$, $A \cap E = \bigcup\limits_{n=1}^{\infty} (A_n \cap E) \in \mathcal{D}_0$ 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

$$\mathcal{D}_2 = \{D \in \mathcal{D}_0 : D \cap C \in \mathcal{D}_0 \quad \text{for each} \quad C \in \mathcal{D}_0\}.$$

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 $\sigma(\mathcal{E}) = \mathcal{A}$;

---
[73] Eugene Dynkin 1924–2014 was a Russian mathematician who worked on Algebra and Probability theory.
[74] See Definition 2.9 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 2.10, $\mu_j(A_n) \uparrow \mu_j(A)$ 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) $\Rightarrow$ (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 3.10.

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).

In order to state and prove Carathéodory extension Theorem 3.1 we first a new concept of measurability associated to a mapping $\alpha : \mathcal{P} \to [0, \infty]$ which satisfies $\alpha(\varnothing) = 0$.

Definition B.1. *Let $\alpha : \mathcal{P}(S) \to [0, \infty]$ be a mapping which satisfies $\alpha(\varnothing) = 0$. We say that $A \subseteq S$ is α-measurable if*

$$\alpha(Q) = \alpha(Q \cap A) + \alpha(Q \cap A^c), \quad \textit{for all } Q \in \mathcal{P}(S).$$

The collection of all α-measurable sets is denoted by $\mathcal{M}_\alpha$.

The mapping α could be rather general and it will not be additive in general. However, the cleverly defined collection $\mathcal{M}_\alpha$ turns out to be a ring on which α is additive.

Lemma B.2. *Let $\alpha : \mathcal{P}(S) \to [0, \infty]$ be a mapping which satisfies $\alpha(\varnothing) = 0$. Then $\mathcal{M}_\alpha$ is a ring and α is additive on $\mathcal{M}_\alpha$.*

Proof. In order to check that $\mathcal{M}_\alpha$ is a ring we check the following:
- (i) $\varnothing \in \mathcal{M}_\alpha$;
- (ii) $A \in \mathcal{M}_\alpha \implies A^c \in \mathcal{M}_\alpha$;
- (iii) $A, B \in \mathcal{M}_\alpha \implies A \cap B \in \mathcal{M}_\alpha$.

Given these properties it is straightforward to check that $\mathcal{M}_\alpha$ is a ring. Indeed, this follows from the formulas $B \setminus A = B \cap A^c$ and $A \cup B = (A^c \cap B^c)^c$.

Properties (i) and (ii) are clear. It remains to check (iii). Let $A, B \in \mathcal{M}_\alpha$ and write $C = A \cap B$. Let $Q \in \mathcal{P}(S)$ be arbitrary. Observation: $A \cap B^c = C^c \cap A$ and $A^c = C^c \cap A^c$. 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, $A \cap B = C \in \mathcal{M}_\alpha$ which yields (iii).

To check that α is additive fix two disjoint sets $A, B \in \mathcal{M}_\alpha$ and let $Q = A \cup B$. Then $Q \cap A = A$ and $Q \cap A^c = B$. Therefore, since $A \in \mathcal{M}_\alpha$ we find

$$\alpha(A \cup B) = \alpha(Q) = \alpha(Q \cap A) + \alpha(Q \cap A^c) = \alpha(A) + \alpha(B).$$

☺

In Lemma B.2 we have seen that α is additive. In order to obtain a measure we need a further condition on α.

Definition B.3. *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 each sequence $(A_n)_{n \ge 1}$ in $\mathcal{P}(S)$ one has* $\alpha\Big(\bigcup_{n=1}^{\infty} A_n\Big) \le \sum_{n=1}^{\infty} \alpha(A_n)$.

An outer measure is not necessarily a measure. For instance $\alpha(\varnothing) = 0$ and $\alpha(A) = 1$ if $A \subseteq S$ is non-empty is an example of an outer measure which is not a measure (if S contains at least two elements).

Next, we show that being an outer measure is the right additional ingredient to prove that α is a measure on $\mathcal{M}_\alpha$.

Lemma B.4. *Let α be an outer measure. Then $\mathcal{M}_\alpha$ is a σ-algebra and α is a measure on $(S, \mathcal{M}_\alpha)$.*

Proof. Lemma B.2 gives that $\mathcal{M}_\alpha$ is a ring and α is additive on $\mathcal{M}_\alpha$. It remains to check that for any disjoint sequence $(A_n)_{n\geq 1}$,

$$A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{M}_\alpha \quad \text{and} \quad \alpha(A) = \sum_{n=1}^{\infty} \alpha(A_n). \tag{B.1}$$

Let $B_n = \bigcup_{j=1}^{n} A_j$ for each $n \geq 1$. Fix an arbitrary $Q \subseteq S$. For every $n \in \mathbb{N}$ the following holds:

$$\begin{aligned} \textstyle\sum_{j=1}^n \alpha(Q\cap A_j) + \alpha(Q\cap A^c) &= \alpha(Q\cap B_n) + \alpha(Q\cap A^c) && \text{(by Lemma B.2)} \\ &\leq \alpha(Q\cap B_n) + \alpha(Q\cap B_n^c) && \text{(since } A^c \subseteq B_n^c) \\ &= \alpha(Q) && \text{(since } B_n \in \mathcal{M}_\alpha) \end{aligned}$$

Using first the σ-subadditivity of α and then the arbitrariness of $n \in \mathbb{N}$ in the above, we deduce

$$\alpha(Q\cap A) + \alpha(Q\cap A^c) \leq \sum_{j=1}^{\infty} \alpha(Q\cap A_j) + \alpha(Q\cap A^c) \leq \alpha(Q). \tag{B.2}$$

On the other hand by subadditivity also the converse estimate holds: $\alpha(Q) \leq \alpha(Q\cap A) + \alpha(Q\cap A^c)$, and hence $A \in \mathcal{M}_\alpha$. Moreover, all inequalities in (B.2) have to be identities[75], and hence (B.1) follows by taking $Q = A$ in (B.2). ☻

In the above results we have seen that with outer measures one can construct measures on certain σ-algebras. Our next aim is to show that there is a natural outer measure associated to an additive mapping μ on a ring $\mathcal{R}$.

Lemma B.5. *Let S be a set and $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring. Suppose $\mu : \mathcal{R} \to [0,\infty]$ is additive and satisfies $\mu(\varnothing) = 0$. For $A \subseteq S$ define*

$$\mu^*(A) = \inf\Big\{\sum_{j=1}^{\infty} \mu(B_j) : A \subseteq \bigcup_{j=1}^{\infty} B_j, \text{ where } B_j \in \mathcal{R} \text{ for } j \geq 1\Big\}, \tag{B.3}$$

where we let $\mu^*(A) = \infty$ if the above set is empty. Then μ^* is an outer measure

Proof. The mapping $\mu^* : \mathcal{P}(S) \to [0,\infty]$ clearly satisfies (i) and (ii). In order to check (iii) let $(A_n)_{n\geq 1}$ in $\mathcal{P}$ and let $\varepsilon > 0$. If $\mu^*(A_n) = \infty$ for some $n \geq 1$, then (iii) is trivial. Next assume $\mu^*(A_n) < \infty$ for all $n \geq 1$. Then by definition of μ^* for each fixed $n \geq 1$ we can find $B_{n,j} \in \mathcal{R}$ such that

$$A_n \subseteq \bigcup_{j=1}^{\infty} B_{n,j} \quad \text{and} \quad \mu^*(A_n) + 2^{-n}\varepsilon \geq \sum_{j=1}^{\infty} \mu(B_{n,j}).$$

Then $\bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n,j=1}^{\infty} B_{n,j}$ and again by the definition of μ^* we find

$$\mu^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n,j=1}^{\infty} \mu(B_n, j) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + 2^{-n}\varepsilon = \varepsilon + \sum_{n=1}^{\infty} \mu^*(A_n).$$

Since $\varepsilon > 0$ was arbitrary, this prove the required estimate. ☺

Theorem B.6 (Carathéodory's extension theorem)**.** *Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring and $\mu : \mathcal{R} \to [0,\infty]$ be σ-additive on $\mathcal{R}$ and $\mu(\varnothing) = 0$. Then μ^* defined by (B.3) satisfies the following properties:*

(i) μ^* is a measure on the σ-algebra $\mathcal{M}_{\mu^*}$;
(ii) $\mu^*(A) = \mu(A)$ for all $A \in \mathcal{R}$.
(iii) $\mathcal{R} \subseteq \mathcal{M}_{\mu^*}$;

In particular, $\overline{\mu} : \sigma(\mathcal{R}) \to [0,\infty]$ defined by $\overline{\mu}(A) = \mu^*(A)$ defines 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 $\mathcal{M}_{\mu^*}$.

---

[75] Clearly, $x \leq y \leq z \leq x$ enforces $x = y = z$

Proof. In Lemma B.5 we have proved that μ^* is an outer measure. Therefore, assertion (i) follows from Lemma B.4. In remains to prove (ii) and (iii). The assertion concerning $\overline{\mu}$ follows from (i)–(iii) as the restriction of a measure to a smaller σ-algebra is a measure again.

Step 1: Proof of (ii): Let $A \in \mathcal{R}$. It is clear that $\mu^*(A) \leq \mu(A)$. Indeed, take $B_1 = A$ and $B_n = \varnothing$ for $n \geq 2$ in (B.3). For the converse estimate the case $\mu^*(A) = \infty$ is clear. Now suppose $\mu^*(A) < \infty$ 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 Theorem 2.8 (iii) applied to $A = \bigcup\limits_{n=1}^{\infty} A \cap B_n$, we find

$$\mu(A) \leq \sum_{n=1}^{\infty} \mu(A \cap B_n) \leq \sum_{n=1}^{\infty} \mu(B_n).$$

Taking the infimum over all $(B_n)_{n\geq 1}$ as above yields $\mu(A) \leq \mu^*(A)$.

Step 2: Proof of (iii): Let $A \in \mathcal{R}$ and $Q \subseteq S$. Since μ^* is subadditive we find $\mu^*(Q) \leq \mu^*(Q \cap A) + \mu^*(Q \cap A^c)$. For the converse estimate the case $\mu^*(Q) = \infty$ is trivial. In case $\mu^*(Q) < \infty$, choose $B_1, B_2, \ldots, \in \mathcal{R}$ such that $Q \subseteq \bigcup\limits_{n=1}^{\infty} B_n$. Then $B_n \cap A, B_n \cap A^c \in \mathcal{R}$ for all $n \geq 1$ 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 μ^* and then the additivity of μ on $\mathcal{R}$, we find

$$\mu^*(Q \cap A) + \mu^*(Q \cap A^c) \leq \sum_{n=1}^{\infty} \mu(B_n \cap A) + \sum_{n=1}^{\infty} \mu(B_n \cap A^c) = \sum_{n=1}^{\infty} \mu(B_n)$$

Taking the infimum over all $(B_n)_{n\geq 1}$ as above gives $\mu^*(Q \cap A) + \mu^*(Q \cap A^c) \leq \mu^*(Q)$. Combining both estimates we can conclude $A \in \mathcal{M}_{\mu^*}$. ☺

Exercises

In the following exercise we show that $\mathcal{M}_{\mu^*} \neq \mathcal{P}(S)$ in general.[76]

Exercise B.1. Let $S = \{1, 2, 3\}$ and define a σ-algebra by $\mathcal{A} = \{\varnothing, S, \{1, 2\}, \{3\}\}$. Assume μ is a measure satisfying $\mu(\{1, 2\}) = \mu(\{3\}) = \frac{1}{2}$.
(a) Show that $\mu^*(\{1\}) = \mu^*(\{2\}) = \frac{1}{2}$.
(b) Show that $\{1\}, \{2\} \notin \mathcal{M}_{\mu^*}$.

Exercise B.2. Let $\alpha : \mathcal{P}(S) \to [0, \infty]$ be an outer measure and suppose that $A \subseteq \mathcal{P}(S)$ satisfies $\alpha(A) = 0$. Show that $A \in \mathcal{M}_\alpha$.

Exercise* B.3. Assume the conditions of Theorem B.6 and assume μ is σ-finite on $\mathcal{R}$, that means there exists a sequence $(S_n)_{n\geq 1}$ in $\mathcal{R}$ such that $\mu(S_n) < \infty$ for all $n \geq 1$ and $\bigcup\limits_{n=1}^{\infty} S_n = S$. Prove that the following are equivalent:
(a) $A \in \mathcal{M}_{\mu^*}$;
(b) There exists a $B \in \sigma(\mathcal{R})$ such that $A \subseteq B$ and $\mu^*(B \setminus A) = 0$
Hint: First reduce to the case of finite measure by intersecting with S_n. Use the definition of μ^* given in (B.3).

[76] For the Lebesgue measure one also has $\mathcal{M}_{\mu^*} \neq \mathcal{P}(\mathbb{R})$, but this is much harder to prove. See Appendix C

Appendix C. Non-measurable sets

Let λ be the Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$. Let $\lambda^* : \mathcal{P}(S) \to [0, \infty]$ be the outer measure associated with λ (see Lemma B.5). The σ-algebra $\mathcal{M}(\mathbb{R}^d) := \mathcal{M}_{\lambda^*}$ introduced in Definition B.1 is usually called the **Lebesgue σ-algebra**. It follows from Theorem B.6 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 λ^* 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 [12, 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] S.R. Lay. *Analysis. With an introduction to proof. 5th ed.* Pearson, 2015.

[12] H. Royden and P. Fitzpatrick. *Real analysis. 4th ed.* New York, NY: Prentice Hall, 4th ed. edition, 2010.

[13] E.M. Stein and R. Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003. An introduction.

[14] 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.

[15] 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.

[16] D. Werner. *Einführung in die höhere Analysis*. Springer-Lehrbuch. [Springer Textbook]. Springer-Verlag, Berlin, corrected edition, 2009.

[17] D. Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[18] 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.

[19] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002. With a foreword by Robert A. Fefferman.