Ulm University

Mathematical Foundations of Quantum Mechanics

Stephan Fackler

Version: May 6, 2015

Contents

1 A Crash Course in Measure Theory

In classical quantum mechanics (pure) a quantum mechanical system is described by some complex Hilbert space. For example, the (pure) states of a single one-dimensional particle can be described by elements in the Hilbert space $L^2(\mathbb{R})$ as introduced in introductory courses in quantum mechanics. A natural first attempt to mathematically define this space is the following:

$$L^2(\mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{C} : f_{|[-n,n]} \text{ Riemann-int. for } n \in \mathbb{N} \text{ and } \int_{-\infty}^{\infty} |f(x)|^2 \, dx < \infty \right\}.$$

However, there are several issues. First of all, the natural choice

$$\|f\|_2 := \left(\int_{-\infty}^{\infty} |f(x)|^2 \, dx \right)^{1/2}$$

does not define a norm on $L^2(\mathbb{R})$ as there exist functions $0 \neq f \in L^2(\mathbb{R})$ with $\|f\|_2 = 0$. This problem can easily be solved by identifying two functions $f, g \in L^2(\mathbb{R})$ whenever $\|f - g\|_2 = 0$. A more fundamental problem is that the above defined space is not complete, i.e. there exist Cauchy sequences in $L^2(\mathbb{R})$ which do not converge in $L^2(\mathbb{R})$. Therefore one has to replace $L^2(\mathbb{R})$ as defined above by its completion. This is perfectly legitimate from a mathematical point of view. However, this approach has a severe shortcoming: we do not have an explicit description of the elements in the completion. Even worse, we do not even know whether these elements can be represented as functions.

To overcome these issues, we now introduce an alternative way to integration, finally replacing the Riemann-integral by the so-called *Lebesgue-integral*. In order to be able to introduce the Lebesgue-integral we need first a rigorous method to measure the volume of subsets of $\mathbb{R}^n$ or more abstract sets which then can be used to define the Lebesgue integral.

The material covered in this chapter essentially corresponds to the basic definitions and results presented in an introductory course to measure theory. We just give the definitions with some basic examples to illustrate the concepts and then state the main theorems without proofs. More details and the proofs can be learned in any course on measure theory or from the many excellent text books, for example [Bar95] or [Rud87]. For further details we guide the interested reader to the monographs [Bog07].

1.1 Measure Spaces

For $n \in \mathbb{N}$ let $\mathcal{P}(\mathbb{R}^n)$ denote the set of all subsets of $\mathbb{R}^n$. The measurement of volumes can then be described by a mapping $m \colon \mathcal{P}(\mathbb{R}^n) \to \mathbb{R}_{\geq 0} \cup \{\infty\}$. In order to obtain a reasonable notion of volume one should at least require

(i) $m(A \cup B) = m(A) + m(B)$ for all $A, B \subset \mathbb{R}^n$ with $A \cap B = \emptyset$,

(ii) $m(A) = m(B)$ whenever $A, B \subset \mathbb{R}^n$ are congruent, i.e. B can be obtained from B by a finite combination of rigid motions.

Intuitively, this sounds perfectly fine. However, there is the following result published by S. Banach and A. Tarski in 1924.

Theorem 1.1.1 (Banach–Tarski paradox). *Let $n \geq 3$ and $A, B \subset \mathbb{R}^n$ be arbitrary bounded subsets with non-empty interiors. Then A and B can be partitioned into a finite number of disjoints subsets*

$$A = A_1 \cup \ldots \cup A_n \qquad \text{and} \qquad B = B_1 \cup \ldots \cup B_n$$

such that for all $i = 1, \ldots, n$ the sets A_i and B_i are congruent.

Using such paradoxical decompositions we see that m must agree for all bounded subsets of $\mathbb{R}^n$ with non-empty interiors. For example, by splitting a cube Q into two smaller parts, we see that $m(Q) \in (0, \infty)$ leads to a contradiction. Hence, it is impossible to measure the volume of arbitrary subsets of $\mathbb{R}^n$ in a reasonable way!

Remark 1.1.2. Of course, we all know that in physical reality such a paradox does not occur. Indeed, the decompositions given by the Banach–Tarski paradox are not constructive and therefore cannot be realized in the real world. More precisely in mathematical terms, the proof of the Banach–Tarski paradox requires some form of the axiom of choice.

Since we cannot measure the volume of arbitrary subsets of $\mathbb{R}^n$ in a consistent reasonable way, it is necessary to restrict the volume measurement to a subset of $\mathcal{P}(\mathbb{R}^n)$. This subset should be closed under basic set theoretic operations. This leads to the following definition which can be given for arbitrary sets Ω instead of $\mathbb{R}^n$.

Definition 1.1.3 (σ-algebra). Let Ω be a set. A subset $\Sigma \subset \mathcal{P}(\Omega)$ is called a *σ-algebra* if

(a) $\emptyset \in \Sigma$,

(b) $A^c \in \Sigma$ for all $A \in \Sigma$,

(c) $\cup_{n \in \mathbb{N}} A_n \in \Sigma$ whenever $(A_n)_{n \in \mathbb{N}} \subset \Sigma$.

The tuple (Ω, Σ) is called a *measurable space* and the elements of Σ are called *measurable*.

Note that it follows from the definition that for $A, B \in \Sigma$ one also has $A \cap B \in \Sigma$ and $B \setminus A \in \Sigma$. The closedness of Σ under countable unions may be the less intuitive of the above defining properties. It guarantees that σ-algebras behave well under limiting processes which lie at the hearth of analysis. We now give some elementary examples of σ-algebras.

Example 1.1.4. (i) Let Ω be an arbitrary set. Then the power set $\mathcal{P}(\Omega)$ is clearly a σ-algebra.

(ii) Let Ω be an arbitrary set. We define Σ as the set of subsets of Ω which are countable or whose complement is countable. One then can check that Σ is a σ-algebra. Here one has to use the fact that countable unions of countable sets are again countable. Note that Σ does in general not agree with $\mathcal{P}(\Omega)$. For example, if $\Omega = \mathbb{R}$, then the interval $[0,1]$ is not contained in Σ.

We now give an important and non-trivial example of a σ-algebra which will be frequently used in the following.

Example 1.1.5 (Borel σ-algebra). Let Ω be a subset of $\mathbb{R}^n$ for $n \in \mathbb{N}$, or more general a normed, metric or topological space. Then the smallest σ-algebra that contains all open sets $\mathcal{O}$ of Ω

$$\mathcal{B}(\Omega) = \bigcap_{\substack{\Sigma\ \sigma\text{-algebra:} \\ \Sigma \supset \mathcal{O}}} \Sigma$$

is called the *Borel σ-algebra* on Ω. One can show that $\mathcal{B}(\mathbb{R}^n)$ is the smallest σ-algebra that is generated by elements of the form $[a_1, b_1) \times \cdots [a_n, b_n)$ for $a_i < b_i$, i.e. by products of half-open intervals.

Recall that a function $f : \Omega_1 \to \Omega_2$ between two normed or more general metric or topological spaces is continuous if and only if the preimage of every open set under f is again open. This means that f preserves the topological structure. In the same spirit *measurable mappings* are compatible with the measurable structures on the underlying spaces.

Definition 1.1.6 (Measurable mapping). Let (Ω_1, Σ_1) and (Ω_2, Σ_2) be two measurable spaces. A map $f : \Omega_1 \to \Omega_2$ is called *measurable* if

$$f^{-1}(A) \in \Sigma_1 \qquad \text{for all } A \in \Sigma_2.$$

A function $f : \Omega_1 \to \Omega_2$ between two normed spaces (or more generally two metric or topological spaces) is called *measurable* if f is a measurable map between the measurable spaces $(\Omega_1, \mathcal{B}(\Omega_1))$ and $(\Omega_2, \mathcal{B}(\Omega_2))$.

It is often very convenient to consider functions $f\colon \Omega \to \overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ denotes the extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$. In this case one calls f measurable if and only if $X_\infty = \{x \in \Omega : f(x) = \infty\}$ and $X_{-\infty} = \{x \in \Omega : f(x) = -\infty\}$ are measurable and the restricted function $f\colon \Omega \setminus (X_\infty \cup X_{-\infty}) \to \mathbb{R}$ is measurable in the sense just defined above. If a real-valued function takes the values ∞ or $-\infty$, we will implicitly always work with this definition. We will often need the following sufficient conditions for a mapping to be measurable.

Proposition 1.1.7. *Let Ω_1 and Ω_2 be two normed vector spaces or more generally metric or topological spaces. Then every continuous mapping $f\colon \Omega_1 \to \Omega_2$ is measurable. Further, every monotone function $f\colon \mathbb{R} \to \mathbb{R}$ is measurable.*

Furthermore, measurable functions are closed under the usual arithmetic operations and under pointwise limits.

Proposition 1.1.8. *Let (Ω, Σ, μ) be a measure space.*

(a) *Let $f, g\colon \Omega \to \mathbb{C}$ be measurable. Then $f+g$, $f-g$, $f \cdot g$ and f/g provided $g(x) \neq 0$ for all $x \in \Omega$ are measurable as well.*

(b) *Let $f_n\colon \Omega \to \mathbb{C}$ be a sequence of measurable functions such that $f(x) := \lim_{n\to\infty} f_n(x)$ exists for all $x \in \Omega$. Then f is measurable.*

We now assign a measure to a measurable space.

Definition 1.1.9 (Measure). Let (Ω, Σ) be a measurable space. A *measure* on (Ω, Σ) is a mapping $\mu\colon \Sigma \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ that satisfies

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

(ii) $\mu(\cup_{n\in\mathbb{N}} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ for all pairwise disjoint $(A_n)_{n\in\mathbb{N}} \subset \Sigma$.

The triple (Ω, Σ, μ) is a *measure space*. If $\mu(\Omega) < \infty$, then (Ω, Σ, μ) is called a *finite measure space*. If $\mu(\Omega) = 1$, one says that (Ω, Σ, μ) is a *probability space*.

One can deduce from the above definition that a measure satisfies $\mu(A) \leq \mu(B)$ for all measurable $A \subset B$ and $\mu(\cup_{n\in\mathbb{N}} B_n) \leq \sum_{n=1}^{\infty} \mu(B_n)$ for arbitrary $(B_n)_{n\in\mathbb{N}} \subset \Sigma$. Moreover, one has $\mu(A \setminus B) = \mu(A) - \mu(B)$ for measurable $B \subset A$ whenever $\mu(B) < \infty$. We begin with some elementary examples of measure spaces.

Example 1.1.10. (i) Consider $(\Omega, \mathcal{P}(\Omega))$ for an arbitrary set Ω and define $\mu(A)$ as the number of elements in A whenever A is a finite subset and $\mu(A) = \infty$ otherwise. Then μ is a measure on $(\Omega, \mathcal{P}(\Omega))$.

(ii) Let Ω be an arbitrary non-empty set and $a \in \Omega$. Define

$$\delta_a \colon \mathcal{P}(\Omega) \to \mathbb{R}_{\geq 0}$$
$$A \mapsto \begin{cases} 1 & \text{if } a \in A \\ 0 & \text{else} \end{cases}.$$

Then δ_a is a measure on $(\Omega, \mathcal{P}(\Omega))$ and is called the *Dirac measure in a*.

We now come to the most important example for our purposes.

Theorem 1.1.11 (Lebesgue Measure). *Let $n \in \mathbb{N}$. There exists a unique* Borel measure *λ, i.e. a measure defined on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, that satisfies*

$$\lambda([a_1, b_1) \times \cdots [a_n, b_n)) = \prod_{k=1}^{n} (b_k - a_k)$$

for all products with $a_i < b_i$. The measure λ is called the Lebesgue measure *on $\mathbb{R}^n$.*

Of course, one can also restrict the Lebesuge measure to $(\Omega, \mathcal{B}(\Omega))$ for subsets $\Omega \subset \mathbb{R}^n$. The uniqueness in the above theorem is not trivial, but essentially follows from the fact that the products of half-open intervals used in the above definition generate the Borel-σ-algebra and are closed under finite intersections. The existence is usually proved via Carathéodory's extension theorem.

1.2 The Lebesgue Integral

Given a measure space (Ω, Σ, μ), one can integrate certain functions $f \colon \Omega \to \mathbb{C}$ over the measure μ. One extends the integral step by step to more general classes of functions. A function $f \colon \Omega \to \mathbb{C}$ is a *simple function* if there exist finite measurable sets $A_1, \ldots, A_n \in \Sigma$ and $a_1, \ldots, a_n \in \mathbb{C}$ such that $f = \sum_{k=1}^{n} a_k \mathbb{1}_{A_k}$. Here $\mathbb{1}_{A_k}$ is the function defined by

$$\mathbb{1}_{A_k}(x) = \begin{cases} 1 & \text{if } x \in A_k \\ 0 & \text{if } x \notin A_k \end{cases}.$$

Definition 1.2.1 (Lebesgue integral). Let (Ω, Σ, μ) be a measure space.

(i) For a simple function $f \colon \Omega \to \mathbb{R}_{\geq 0}$ given by $f = \sum_{k=1}^{n} a_k \mathbb{1}_{A_k}$ as above one defines the Lebesgue integral as

$$\int_\Omega f \, d\mu = \sum_{k=1}^{n} a_k \mu(A_k).$$

(ii) For a measurable function $f\colon \Omega \to \mathbb{R}_{\geq 0}$ the Lebesgue integral is defined as
$$\int_\Omega f\,d\mu = \sup_{\substack{g \text{ simple:}\\ 0\leq g\leq f}} \int_\Omega g\,d\mu.$$

(iii) A general measurable function $f\colon \Omega \to \mathbb{C}$ can be uniquely decomposed into for non-negative measurable functions $f\colon \Omega \to \mathbb{R}_{\geq 0}$ such hat $f = (f_1 - f_2) + i(f_3 - f_4)$. One says that f is *integrable* if $\int_\Omega f_i\,d\mu < \infty$ and writes $f \in \mathcal{L}^1(\Omega, \Sigma, \mu)$. In this case one sets the Lebesgue integral as
$$\int_\Omega f\,d\mu = \int_\Omega f_1\,d\mu - \int_\Omega f_2\,d\mu + i\left(\int_\Omega f_3\,d\mu - \int_\Omega f_4\,d\mu\right).$$

Moreover, for a measurable set $A \in \Sigma$ we use the short-hand notation
$$\int_A f\,d\mu \coloneqq \int_\Omega \mathbb{1}_A f\,d\mu$$
whenever the integral on the right hand side exists.

We will often use the following terminology. Let (Ω, Σ, μ) be a measure space and $P(x)$ a property for every $x \in \Omega$. We say that P holds almost everywhere if there exists a set $N \in \Sigma$ with $\mu(N) = 0$ such that $P(x)$ holds for all $x \notin N$. For example, on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ the function $f(x) = \cos(\pi x)$ satisfies $|f(x)| < 1$ almost everywhere because one can choose $N = \mathbb{Z}$ which has zero Lebesgue measure. In the following we will often make use of that the fact that the integrals over two measurable functions f and g agree whenever $f(x) = g(x)$ almost everywhere.

Notice that we can now integrate a function $f\colon [a,b] \to \mathbb{C}$ in two different ways by either using the Riemann or the Lebesgue integral. These two integrals however agree as soon as both make sense and the Lebesgue integral can be considered as a true extension of the Riemann integral (except for some minor measurability issues).

Theorem 1.2.2 (Lebesgue integral equals Riemann integral). *The Riemann and Lebesgue integral have the following properties.*

(a) *Let $f\colon [a,b] \to \mathbb{C}$ be a Riemann integrable function. Then there exists a measurable function $g\colon [a,b] \to \mathbb{C}$ with $f = g$ almost everywhere and $g \in \mathcal{L}^1([a,b], \mathcal{B}([a,b]), \lambda)$. Moreover, one has*
$$\int_a^b f(x)\,dx = \int_{[a,b]} g\,d\lambda.$$

(b) *Let $f: I \to \mathbb{C}$ for some interval $I \subset \mathbb{R}$ be Riemann integrable in the improper sense. If*

$$\sup_{\substack{K \subset I \\ \text{compact interval}}} \int_K |f(x)|\, dx < \infty,$$

then there exists a measurable function $g\colon I \to \mathbb{C}$ with $f = g$ almost everywhere and $g \in \mathcal{L}^1(I, \mathcal{B}(I), \lambda)$. Moreover, one has

$$\int_I f(x)\, dx = \int_I g\, d\lambda.$$

Moreover, if f is measurable (for example if f is continuous), one can choose g equal to f.

For an example of a Lebesgue-integrable function which is not Riemann-integrable, consider $f(x) = \mathbb{1}_{[0,1]\cap\mathbb{Q}}(x)$. Then f is not Riemann-integrable as on arbitrary fine partitions of $[0,1]$ the function takes both values 0 and 1, whereas the Lebesgue integral can be easily calculated as $\int_{[0,1]} f\, d\lambda = \lambda([0,1] \cap \mathbb{Q}) = 0$.

Now suppose that one has given a sequence $f_n\colon \Omega \to \mathbb{C}$ of measurable functions such that $\lim_{n\to\infty} f_n(x)$ exists almost everywhere. Hence, there exists a measurable set N with $\mu(N) = 0$ such that the limit exists for all $x \notin N$. We now set

$$f(x) = \begin{cases} \lim_{n\to\infty} f_n(x) & \text{if this limit exists,} \\ 0 & \text{else.} \end{cases}$$

One can show that the set C of all $x \in \Omega$ for which the above limit exists is measurable. It follows easily from this fact the function $f\colon \Omega \to \mathbb{C}$ is measurable as well. Note further that because of $C \subset N$ one has $\mu(C) = 0$. Hence, the Lebesgue integral of f is independent of the concrete choice of the values at the non-convergent points and therefore the choice does not matter for almost all considerations. We make the agreement that we will always define the pointwise limit of measurable functions in the above way whenever the limit exists almost everywhere. This is particularly useful for the formulation of the following convergence theorems for the Lebesgue integral.

Theorem 1.2.3 (Monotone convergence theorem). *Let (Ω, Σ, μ) be a measure space and $f_n\colon \Omega \to \mathbb{R}$ a sequence of measurable functions with $f_{n+1}(x) \geq f_n(x) \geq 0$ almost everywhere. Suppose further that $f(x) = \lim_{n\to\infty} f_n(x)$ exists almost everywhere. Then*

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

Note that the monotonicity assumption is crucial for the theorem. In fact, in general one cannot switch the order of limits and integrals as the following example shows.

$$\lim_{n\to\infty} \int_{\mathbb{R}} \mathbb{1}_{[n,n+1]} \, d\lambda = 1 \neq 0 = \int_{\Omega} \lim_{n\to\infty} \mathbb{1}_{[n,n+1]} \, d\lambda.$$

However, the following result holds for non-positive and non-monotone sequences of functions.

Theorem 1.2.4 (Dominated convergence theorem). *Let (Ω, Σ, μ) be a measure space and $f_n \colon \Omega \to \mathbb{C}$ a sequence of measurable functions for which there exists an integrable function $g \colon \Omega \to \mathbb{R}$ such that for all $n \in \mathbb{N}$ one has $|f_n(x)| \le g(x)$ almost everywhere. Further suppose that $f(x) = \lim_{n\to\infty} f_n(x)$ exists almost everywhere. Then*

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

For the next result we need a finiteness condition on the underlying measure space.

Definition 1.2.5 (σ-finite measure space). A measure space (Ω, Σ, μ) is called *σ-finite* if there exists a sequence of measurable sets $(A_n)_{n\in\mathbb{N}} \subset \Sigma$ such that $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$ and

$$\Omega = \bigcup_{n=1}^{\infty} A_n.$$

For example, $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ together with the counting measure or the measure spaces $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ for $n \in \mathbb{N}$, where λ denotes the Lebesgue measure, are σ-finite. Moreover, every finite measure space and a fortiori every probability space is σ-finite. For an example of a non-σ-finite measure space consider $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ with the counting measure.

Definition 1.2.6 (Products of measure spaces). Consider the two measure spaces $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$.

(i) The σ-algebra on $\Omega_1 \times \Omega_2$ generated by sets of the form $A_1 \times A_2$ for $A_i \in \Sigma_i$ $(i = 1, 2)$ (i.e. the smallest σ-algebra that contains these sets) is called the *product σ-algebra* of Σ_1 and Σ_2 and is denoted by $\Sigma_1 \otimes \Sigma_2$.

(ii) A measure μ on the measurable space $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ is called a *product measure* of μ_1 and μ_2 if

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \qquad \text{for all } A_1 \in \Sigma_1, A_2 \in \Sigma_2.$$

Here we use the convention that $0 \cdot \infty = \infty \cdot 0 = 0$.

For example, one has $\mathcal{B}(\mathbb{R}^n)\otimes\mathcal{B}(\mathbb{R}^m) = \mathcal{B}(\mathbb{R}^{n+m})$ which can be easily verified using the fact that products of half-open intervals generate $\mathcal{B}(\mathbb{R}^n)$. It follows from the characterizing property of the Lebesgue measure λ_n on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ that for all $n, m \in \mathbb{N}$ the measure λ_{n+m} is a product measure of λ_n and λ_m. One can show that there always exists a product measure for two arbitrary measure spaces. In most concrete situations there exists a uniquely determined product measure as the following theorem shows.

Theorem 1.2.7. *Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be two σ-finite measure spaces. Then there exists a unique product measure on $(\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2)$ which is denoted by $\mu_1 \otimes \mu_2$.*

It is now a natural question how integration over product measures is related to integration over the single measures. An answer is given by Fubini's theorem.

Theorem 1.2.8 (Fubini–Tonelli theorem). *Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be two σ-finite measure spaces and $f\colon (\Omega_1 \times \Omega_2, \Sigma_1 \otimes \Sigma_2) \to \mathbb{C}$ a measurable function. Then the functions*

$$y \mapsto \int_{\Omega_1} f(x,y)\, d\mu_1(x) \qquad \text{and} \qquad x \mapsto \int_{\Omega_2} f(x,y)\, d\mu_2(y)$$

are measurable functions $\Omega_2 \to \mathbb{C}$ respectively $\Omega_1 \to \mathbb{C}$. If one of the three integrals

$$\int_{\Omega_1}\int_{\Omega_2} |f(x,y)|\, d\mu_2(y)\, d\mu_1(x), \quad \int_{\Omega_2}\int_{\Omega_1} |f(x,y)|\, d\mu_1(x)\, d\mu_2(y) \quad \text{or} \quad \int_{\Omega_1\times\Omega_2} |f(x,y)|\, d\mu_1 \otimes \mu_2(x,y)$$

is finite, then one has for the product and iterated integrals

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f(x,y)\, d(\mu_1 \otimes \mu_2)(x,y) &= \int_{\Omega_1}\int_{\Omega_2} f(x,y)\, d\mu_2(y)\, d\mu_1(x) \\ &= \int_{\Omega_2}\int_{\Omega_1} f(x,y)\, d\mu_1(x)\, d\mu_2(y).\end{aligned}$$

Moreover, if f is a non-negative function, one can omit the finiteness assumption on the integrals and the conclusion is still valid (in this case all integrals can be infinite).

Note that there are also variants of Fubini's theorem (not in the above generality) for non σ-finite measure spaces. However, this case is more technical and rarely used in practice and therefore we omit it.

1.3 Lebesgue Spaces

We now come back to the motivation at the beginning of this chapter. After our preliminary work we can now define $L^2(\mathbb{R})$ or more generally $L^p(\Omega)$ over an arbitrary measure space (Ω, Σ, μ).

Definition 1.3.1 ($\mathcal{L}^p$-spaces). Let (Ω, Σ, μ) be a measure space. For $p \in [1, \infty)$ we set

$$\mathcal{L}^p(\Omega, \Sigma, \mu) := \left\{ f : \Omega \to \mathbb{K} \text{ measurable} : \int_\Omega |f|^p \, d\mu < \infty \right\},$$
$$\|f\|_p := \left(\int_\Omega |f|^p \, d\mu \right)^{1/p}.$$

For $p = \infty$ we set

$$\mathcal{L}^\infty(\Omega, \Sigma, \mu) := \{ f : \Omega \to \mathbb{K} \text{ measurable} : \exists C \ge 0 : |f(x)| \le C \text{ alm. everywhere}\}.$$
$$\|f\|_\infty := \inf\{C \ge 0 : |f(x)| \le C \text{ almost everywhere}\}.$$

Note that the space $\mathcal{L}^1(\Omega, \Sigma, \mu)$ agrees with the space $\mathcal{L}^1(\Omega, \Sigma, \mu)$ previously defined in Definition 1.2.1. One can show that $(\mathcal{L}^p(\Omega, \Sigma, \mu), \|\cdot\|_p)$ is a semi-normed vector space, i.e. $\|\cdot\|_p$ satisfies all axioms of a norm except for definiteness. Here, the validity of the triangle inequality, the so-called *Minkowski inequality*, is a non-trivial fact. If one identifies two functions whenever they agree almost everywhere, one obtains a normed space.

Definition 1.3.2 (L^p-spaces). Let (Ω, Σ, μ) be a measure space and $p \in [1, \infty]$. The space $L^p(\Omega, \Sigma, \mu)$ is defined as the space $\mathcal{L}^p(\Omega, \Sigma, \mu)$ with the additional agreement that two functions $f, g \colon \Omega \to \mathbb{K}$ are identified with each other whenever $f - g = 0$ almost everywhere.

As a consequence of the above identification $(L^p(\Omega, \Sigma, \mu), \|\cdot\|_p)$ is a normed vector space. In contrast to the variant using the Riemann integral these spaces are complete.

Definition 1.3.3 (Banach space). A normed vector space which is complete with respect to the given norm is called a *Banach space*.

Recall that a normed vector space or more generally a metric space is called *complete* if every Cauchy sequence converges to an element in the space. A sequence $(x_n)_{n \in \mathbb{N}}$ in a normed vector space $(V, \|\cdot\|)$ is called a *Cauchy sequence* if for all $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\|x_n - x_m\| \le \varepsilon$ for all $n, m \ge n_0$. Using this terminology we have

Theorem 1.3.4 (Riesz–Fischer). *Let (Ω, Σ, μ) be a measure space and $p \in [1, \infty]$. Then $L^p(\Omega, \Sigma, \mu)$ is a Banach space.*

Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^p(\Omega,\Sigma,\mu)$ with $f_n \to f$ in L^p. One often says that f_n converges to f in the *p-th mean* which gives the right visual interpretation for convergence in L^p-spaces. Note that the sequence $\mathbb{1}_{[0,1]}$, $\mathbb{1}_{[0,1/2]}$, $\mathbb{1}_{[1/2,1]}$, $\mathbb{1}_{[0,1/4]}$, $\mathbb{1}_{[1/4,1/2]}$ and so on converges in $L^p([0,1])$ for all $p \in [1,\infty)$ to the zero function although $f_n(x)$ diverges for all $x \in [0,1]$. Conversely, pointwise convergence in general does not imply convergence in L^p. For example, the sequence $f_n = \mathbb{1}_{[n,n+1]}$ does not converge in $L^p(\mathbb{R})$ although $f_n(x) \to 0$ for all $x \in \mathbb{R}$. In concrete situations one can often infer L^p-convergence from pointwise convergence with the help of the dominated convergence theorem. In the opposite direction one has the following useful result which actually follows directly from the proof of the Riesz–Fischer theorem.

Proposition 1.3.5. *Let (Ω,Σ,μ) be a measure space and $p \in [1,\infty)$. Further suppose that $f_n \to f$ in $L^p(\Omega,\Sigma,\mu)$. Then there exist a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and $g \in L^p(\Omega,\Sigma,\mu)$ such that*

(a) *$f_{n_k}(x) \to f(x)$ almost everywhere;*

(b) *$|f_{n_k}(x)| \le |g(x)|$ for all $n \in \mathbb{N}$ almost everywhere.*

We will later need some further properties of L^p-spaces. The following result is natural, but needs some effort to be proven rigorously.

Proposition 1.3.6. *Let $\Omega \subset \mathbb{R}^n$ be open and $p \in [1,\infty)$. Then $C_c(\Omega)$, the space of all continuous functions on Ω with compact support (in Ω), is a dense subspace of $L^p(\Omega)$.*

The Cauchy–Schwarz inequality for L^2-spaces generalizes to Hölder's inequality in the L^p-setting. In the following we use the agreement $1/\infty = 0$.

Proposition 1.3.7 (Hölder's inequality). *Let (Ω,Σ,μ) be a measure space. Further let $p \in [1,\infty]$ and $q \in [1,\infty]$ be its dual index given by $\frac{1}{p} + \frac{1}{q} = 1$. Then for $f \in L^p(\Omega,\Sigma,\mu)$ and $g \in L^q(\Omega,\Sigma,\mu)$ the product $f \cdot g$ lies in $L^1(\Omega,\Sigma,\mu)$ and satisfies*

$$\int_\Omega |fg|\,d\mu \le \left(\int_\Omega |f|^p\,d\mu\right)^{1/p} \left(\int_\Omega |g|^q\,d\mu\right)^{1/q}.$$

As an important and direct consequence of Hölder's inequality one has the following inclusions between L^p-spaces.

Proposition 1.3.8. *Let (Ω,Σ,μ) be a finite measure space, i.e. $\mu(\Omega) < \infty$. Then for $p \ge q \in [1,\infty]$ one has the inclusion*

$$L^p(\Omega,\Sigma,\mu) \subset L^q(\Omega,\Sigma,\mu).$$

Proof. We only deal with the case $p \in (1, \infty)$ (the other cases are easy to show). It follows from Hölder's inequality because of $p/q \geq 1$ that

$$\left(\int_\Omega |f|^q \, d\mu\right)^{1/q} = \left(\int_\Omega |f|^q \mathbb{1} \, d\mu\right)^{1/q} \leq \left(\int_\Omega |f|^p \, d\mu\right)^{1/p} \left(\int_\Omega \mathbb{1} \, d\mu\right)^{(1-q/p)\cdot 1/q}$$
$$= \mu(\Omega)^{1/q-1/p} \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}. \qquad \square$$

A second application of Hölder's inequality is the next important estimate on convolutions of two functions.

Definition 1.3.9. Let $f, g \in L^1(\mathbb{R}^n)$. We define the *convolution* of f and g by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y) g(x-y) \, dy.$$

Note that it is not clear that $f * g$ exists under the above assumptions. This is indeed the case as the following argument shows. Note that the function $(x,y) \mapsto f(y)g(x-y)$ is measurable as a map $\mathbb{R}^{2n} \to \mathbb{R}$ by the definition of product σ-algebras and the fact that the product and the composition of measurable functions is measurable. It follows from Fubini's theorem that the function $x \mapsto (f * g)(x)$ is measurable and satisfies

$$\int_{\mathbb{R}^n} |f * g|(x) \, dx \leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)||g(x-y)| \, dy \, dx = \int_{\mathbb{R}^n} |f(y)| \int_{\mathbb{R}^n} |g(x-y)| \, dx \, dy$$
$$= \int_{\mathbb{R}^n} |f(y)| \int_{\mathbb{R}^n} |g(x)| \, dx \, dy = \|f\|_1 \|g\|_1.$$

Hence, the function $f * g$ is finite almost everywhere. Moreover, we have shown that $f * g \in L^1(\mathbb{R}^n)$ and that the pointwise formula in the definition holds with finite values almost everywhere after taking representatives.

Proposition 1.3.10 (Minkowski's inequality for convolutions). *For some $p \in [1, \infty]$ let $g \in L^p(\mathbb{R}^n)$ and $f \in L^1(\mathbb{R}^n)$. Then one has*

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p.$$

Proof. We only deal with the cases $p \in (1, \infty)$ as the boundary cases are simple to prove. We apply Hölder's inequality to the functions $|g(x-y)|$ and $\mathbb{1}$ for the measure $\mu = |f(y)| \, dy$ (i.e. $\mu(A) = \int_A |f(y)| \, dy$) and obtain

$$|(f * g)(x)| \leq \left(\int_{\mathbb{R}^n} |g(x-y)|^p |f(y)| \, dy\right)^{1/p} \left(\int_{\mathbb{R}^n} |f(y)| \, dy\right)^{1/q},$$

where $1/p + 1/q = 1$. Taking the L^p-norm in the above inequality, we obtain the desired inequality

$$\begin{aligned}\|f * g\|_p &\leq \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |g(x-y)|^p |f(y)|\, dy\, \|f\|_1^{p/q}\, dx\right)^{1/p} \\ &= \|f\|_1^{1/q} \left(\int_{\mathbb{R}^n} |f(y)| \int_{\mathbb{R}^n} |g(x-y)|^p\, dx\, dy\right)^{1/p} = \|f\|_1^{1/q}\, \|f\|_1^{1/p}\, \|g\|_p \\ &= \|f\|_1\, \|g\|_p\,. \end{aligned}$$

□

Bibliography

[Bar95] Robert G. Bartle. *The elements of integration and Lebesgue measure*. Wiley Classics Library, vol. Containing a corrected reprint of the 1966 original [ıt The elements of integration, Wiley, New York; MR0200398 (34 #293)], A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995, pp. xii+179.

[Bog07] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007, Vol. I: xviii+500 pp., Vol. II: xiv+575.

[Rud87] Walter Rudin. *Real and complex analysis*. Third. McGraw-Hill Book Co., New York, 1987, pp. xiv+416.