Chapter X

Integration theory

Having made acquaintance in the last chapter with the fundamentals of measure theory, we will now turn to the theory of integration. In the first part of the chapter we study integrals over general measure spaces, while in the second half we take advantage of the special properties of the Lebesgue measure.

Integration with respect to arbitrary measures is not only important in many applications, but it will also be essential in the last chapter, when the underlying set is not "flat" but rather a manifold. This is why even an introductory text such as ours must deal with the subject.

In Section 1, we introduce μ-measurable functions and investigate their basic properties. A position of keen interest in analysis is held by natural measures with respect to which every continuous function is measurable. An example is the class of Radon measures, which we introduce in this section and which we will encounter again in Chapter XII.

In analysis, and not only there, it will be increasingly important to be able to deal with vector-valued functions, that is, maps with values in a Banach space. We have already worked along these lines in the first two volumes, and you will have noticed that the resulting exposition gains not only in elegance but, in many cases, also in simplicity. The same situation obtains regarding integration theory. Hence we have resolved from the outset to develop the theory in terms of vector-valued functions, and we therefore treat the Bochner–Lebesgue integral. This is possible with no significant extra effort. One of the few exceptions is the proof that a vector-valued function is μ-measurable if and only if it is measurable in the usual sense and is μ-almost separable valued. Of course, you could ignore this result and consider only scalar-valued functions. But this is not recommended, as it would cause you to miss out on an important and efficient addition to your toolkit.

Besides vector-valued maps, we will investigate in some detail functions with values in the extended number line $[0, \infty]$. This is primarily for technical reasons; in later sections it will save us from having to always single out special cases.

In Section 2, we introduce the general Bochner–Lebesgue integral, and do so via the $\mathcal{L}_1$-completion of the space of simple functions. This approach not only extends essentially unchanged to vector-valued functions, but also lays the foundation for the proof of Lebesgue's convergence theorem. We treat the latter, as well as other important convergence theorems, in Section 3.

Section 4 is devoted to the elementary theory of Lebesgue spaces. We prove their completeness and show that they become Banach spaces if we identify functions that agree almost everywhere. Because this identification is in our experience a source of difficulties for beginners, we make a meticulous distinction throughout the chapter between equivalence classes of functions and their respective representatives.

Although up to this point, we have considered integrals with respect to an arbitrary measure, we treat in Section 5 the special case of Lebesgue measure in $\mathbb{R}^n$. We show that the one-dimensional Lebesgue integral is an extension of the Cauchy–Riemann integral for absolutely integrable functions. This puts us in the position to bring what we learned about integrals in Volume II into the framework of the general theory. This is of particular significance in the context of Fubini's theorem, which gives a reduction procedure for evaluating higher-dimensional integrals.

Section 6 treats Fubini's theorem. We have decided not to prove it for arbitrary product measure spaces, but rather only for the Lebesgue measure space. This simplifies the presentation considerably and is in practice sufficient for all the needs of analysis — once strengthened by a extension to product manifolds, to be treated in Chapter XII.

The proof of Fubini's theorem in the vector-valued case requires delicate measurability arguments. For this reason, we study first the scalar case. We prove the vector-valued version at the end of Section 6 and exhibit some important applications. On first reading, this part may be skipped, because its results are not used in any essential way afterward, and also because the reader will probably become acquainted at some later point with the Hahn–Banach theorem of functional analysis: with its help Fubini's theorem for vector-valued functions is easily deduced from the scalar version.

Section 7 studies the convolution. This operation allows us to prove with extraordinary efficiency some fundamental approximation theorems, such as the theorem on smooth partitions of unity, which plays in important role in the final chapter. In the second half of the section we address the significance of the convolution and the approximation theorems in analysis and mathematical physics, offering a first glimpse of the very important generalization of the classical differential calculus known as the theory of distributions.

Besides the convergence theorems of Lebesgue and Fubini, the transformation theorem forms the third pillar of the entire integral calculus. It will be proved in Section 8, where we also discuss its more basic applications.

In the last section, we illustrate the power of the theory just developed by proving several basic facts about the Fourier transform. Like the second half of Section 7, this part affords a look at a related area of analysis which you may later encounter in more advanced studies.

1 Measurable functions

Suppose $(X, \mathcal{A}, \mu)$ is a measure space and $A \in \mathcal{A}$. An analogy with elementary geometrical constructions leads one to define the integral over X of the characteristic function χ_A with respect to the measure μ as $\int_X \chi_A \, d\mu := \mu(A)$. Obviously this only makes sense if A belongs to $\mathcal{A}$. The function $f = \chi_A$ must therefore be "compatible" in this sense with the underlying measure space $(\mathcal{A}, \mu)$. For more complicated functions, a suitable approximation argument makes it possible to generalize this notion of compatibility between functions and measures, leading to the concept of the measurability of functions.

In this section denote by

- $(X, \mathcal{A}, \mu)$ a complete, σ-finite measure space;
  $E = (E, |\cdot|)$ a Banach space.

Simple functions and measurable functions

Suppose E is a property that is either true or false of each point in X. We say that E holds **μ-almost everywhere**, or **for μ-almost every** $x \in X$, if there exists a μ-null set N such that $\mathsf{E}(x)$ is true for every $x \in N^c$. "Almost every" and "almost everywhere" are both abbreviated "a.e."

1.1 Examples **(a)** For $f, g \in \mathbb{R}^X$, we write "$f \geq g$ μ-a.e." if there is a μ-null set N such that $f(x) \geq g(x)$ for every $x \in N^c$.

(b) Suppose $f_j, f \in E^X$ for $j \in \mathbb{N}$. Then (f_j) converges to f μ-a.e. if and only if there is a μ-null set N such that $f_j(x) \to f(x)$ for $x \in N^c$.

(c) A function $f \in E^X$ is bounded μ-a.e. if and only if there is a μ-null set N and an $M \geq 0$ such that $|f(x)| \leq M$ for every $x \in N^c$.

(d) If E holds μ-a.e., the set $\{\, x \in X \;;\; \mathsf{E}(x) \text{ is not true} \,\}$ is μ-null.

Proof This follows from the completeness of $(X, \mathcal{A}, \mu)$. ■

(e) Suppose $(X, \mathcal{B}, \nu)$ is an incomplete measure space. Then there is a property E of X that holds ν-almost everywhere for which $\{\, x \in X \;;\; \mathsf{E}(x) \text{ is not true} \,\}$ is however not a ν-null set.

Proof Because $(X, \mathcal{B}, \nu)$ is not complete, there is a ν-null set N and an $M \subset N$ such that $M \notin \mathcal{B}$. If $f := \chi_M$, then $f = 0$ ν-almost everywhere, but $\{\, x \in X \;;\; f(x) \neq 0 \,\} = M$ is not a ν-null set. ■

We say $f \in E^X$ is **μ-simple**[1] if $f(X)$ is finite, $f^{-1}(e) \in \mathcal{A}$ for every $e \in E$, and $\mu\big(f^{-1}(E \backslash \{0\})\big) < \infty$. We denote by $\mathcal{S}(X, \mu, E)$ the set of all μ-simple functions

[1] If the identity of the measure space is clear, we call functions **simple** instead of μ-simple; similarly in the case of μ-measurable functions, about to be introduced.

from X to E.[2]

A function $f \in E^X$ is said to be **μ-measurable** if there is a sequence (f_j) in $\mathcal{S}(X,\mu,E)$ such that $f_j \to f$ μ-almost everywhere as $j \to \infty$. We set[3]

$$\mathcal{L}_0(X,\mu,E) := \{\, f \in E^X \ ;\ f \text{ is } \mu\text{-measurable}\,\}\ .$$

1.2 Remarks **(a)** We have the inclusions of vector subspaces

$$\mathcal{S}(X,\mu,E) \subset \mathcal{L}_0(X,\mu,E) \subset E^X\ .$$

(b) For $j = 0,\ldots,m$, where $m \in \mathbb{N}$, consider $e_j \in E$ and $A_j \in \mathcal{A}$ such that $\mu(A_j) < \infty$. Then $f := \sum_{j=0}^m e_j \chi_{A_j}$ belongs to $\mathcal{S}(X,\mu,E)$. We call this the **normal form** of f if

$$\begin{aligned} e_j &\neq 0 && \text{for } j = 0,\ldots,m\ ,\\ e_j &\neq e_k && \text{for } j \neq k\ ,\\ A_j \cap A_k &= \emptyset && \text{for } j \neq k\ . \end{aligned}$$

(c) Every simple function has a unique normal form, and[4]

$$\mathcal{S}(X,\mu,E) = \Bigl\{ \sum\nolimits_{j=1}^m e_j\chi_{A_j}\ ;\ m \in \mathbb{N},\ e_j \in E\backslash\{0\},\ A_j \in \mathcal{A},\ \mu(A_j) < \infty,\ A_j \cap A_k = \emptyset \text{ for } j \neq k \Bigr\}\ .$$

Proof Suppose $f \in \mathcal{S}(X,\mu,E)$. Then there is an $m \in \mathbb{N}$ and pairwise distinct elements $e_0,\ldots,e_m$ in E such that $f(X)\backslash\{0\} = \{e_0,\ldots,e_m\}$. Setting $A_j := f^{-1}(e_j)$, we have $A_j \in \mathcal{A}$ such that $\mu(A_j) < \infty$ and $A_j \cap A_k = \emptyset$ for $j \neq k$. One checks easily that

$$\sum_{j=0}^m e_j\chi_{A_j}$$

is the unique normal form of f. The second part now follows from (b). ∎

(d) Suppose $f \in E^X$ and $g \in \mathbb{K}^X$ are μ-simple [or μ-measurable]. Then $|f| \in \mathbb{R}^X$ and $gf \in E^X$ are also μ-simple [or μ-measurable]. In particular, $\mathcal{S}(X,\mu,\mathbb{K})$ and $\mathcal{L}_0(X,\mu,\mathbb{K})$ are subalgebras of $\mathbb{K}^X$.

(e) For $A \in \mathcal{A}$ and $f \in E^X$, form the restriction $\nu := \mu \,\big|\, (\mathcal{A}\,|\,A)$ (see Exercise IX.1.7). Then

$$\begin{aligned} f\,|\,A \in \mathcal{S}(A,\nu,E) &\iff \chi_A f \in \mathcal{S}(X,\mu,E)\ ,\\ f\,|\,A \in \mathcal{L}_0(A,\nu,E) &\iff \chi_A f \in \mathcal{L}_0(X,\mu,E)\ . \end{aligned}$$

Proof The simple verification is left to the reader. ∎

[2]We called the space of jump continuous functions $\mathcal{S}(I,E)$, but this will cause no confusion.

[3]Clearly the definition of measurability of functions is meaningful even on incomplete measure spaces.

[4]Compare the footnote to Exercise VI.6.8.

(f) Suppose $f \in \mathcal{L}_0(X,\mu,\mathbb{K})$ and $A := [f \neq 0]$. Also define $g \in \mathbb{K}^X$ through

$$g(x) := \begin{cases} 1/f(x) & \text{if } x \in A\ , \\ 0 & \text{if } x \notin A\ . \end{cases}$$

Then g is μ-measurable.

Proof The measurability of f implies the existence of a μ-null set N and a sequence (φ_j) in $\mathcal{S}(X,\mu,\mathbb{K})$ such that $\varphi_j(x) \to f(x)$ for $x \in N^c$. We set

$$\psi_j(x) := \begin{cases} 1/\varphi_j(x) & \text{if } \varphi_j(x) \neq 0\ , \\ 0 & \text{if } \varphi_j(x) = 0\ , \end{cases}$$

for $x \in X$ and $j \in \mathbb{N}$. By (c) and (d), $(\chi_A \psi_j)$ is a sequence in $\mathcal{S}(X,\mu,\mathbb{K})$, and one verifies easily that $(\chi_A\psi_j)(x) \to g(x)$ for every $x \in N^c$ (see Proposition II.2.6). ■

(g) Let $e \in E\backslash\{0\}$, and suppose $\mu(X) = \infty$. Then $e\chi_X$ belongs to $\mathcal{L}_0(X,\mu,E)$ but not to $\mathcal{S}(X,\mu,E)$.

Proof It is clear that $e\chi_X$ is not μ-simple. Since X is σ-finite, there is a sequence (A_j) in $\mathcal{A}$ such that $\bigcup_j A_j = X$ and $\mu(A_j) < \infty$ for $j \in \mathbb{N}$. For $j \in \mathbb{N}$, set $X_j := \bigcup_{k=0}^{j} A_k$ and $\varphi_j := e\chi_{X_j}$. Then (φ_j) is a sequence in $\mathcal{S}(X,\mu,E)$ that converges pointwise to $e\chi_X$. ■

A measurability criterion

A function $f \in E^X$ is said to be **$\mathcal{A}$-measurable** if the inverse images of open sets of E under f are measurable, that is, if $f^{-1}(\mathcal{T}_E) \subset \mathcal{A}$, where $\mathcal{T}_E$ is the norm topology on E. If there is a μ-null set N such that $f(N^c)$ is separable, we say f is **μ-almost separable valued**.

1.3 Remarks **(a)** Exercise IX.1.6 shows that the set of $\mathcal{A}$-measurable functions coincides with the set of $\mathcal{A}$-$\mathcal{B}(E)$-measurable functions.

(b) Every subspace of a separable metric space is separable.

Proof By Proposition IX.1.8, separability amounts to having a countable basis. But by restriction, a basis of a topological space yields a basis (of no greater cardinality) for any given subspace; see Proposition III.2.26. ■

(c) Suppose E is separable and $f \in E^X$. Then f is μ-almost separable valued.

Proof This follows from (b). ■

(d) Every finite-dimensional normed vector space is separable.[5] ■

The next result gives a characterization of μ-measurable functions, which, besides being of theoretical significance, is very useful in practice for determining measurability.

[5]Compare Example V.4.3(e).

1.4 Theorem *A function in E^X is μ-measurable if and only if it is $\mathcal{A}$-measurable and μ-almost separable valued.*

Proof "$\Rightarrow$" Suppose $f \in \mathcal{L}_0(X, \mu, E)$.

(i) There exist a μ-null set N and a sequence (φ_j) in $\mathcal{S}(X, \mu, E)$ such that

$$\varphi_j(x) \to f(x) \quad (j \to \infty) \quad \text{for } x \in N^c . \tag{1.1}$$

By Proposition I.6.8, $F := \bigcup_{j=0}^{\infty} \varphi_j(X)$ is countable and therefore the closure $\overline{F}$ is separable. Because of (1.1) we have $f(N^c) \subset \overline{F}$. Remark 1.3(b) now shows that f is μ-almost separable valued.

(ii) Let O be open in E and define $O_n := \{\, y \in O \ ; \ \operatorname{dist}(y, O^c) > 1/n \,\}$ for $n \in \mathbb{N}^\times$. Then O_n is open and $\overline{O}_n \subset O$. Also let $x \in N^c$. By (1.1), $f(x)$ belongs to O if and only if there exist $n \in \mathbb{N}^\times$ and $m = m(n) \in \mathbb{N}^\times$ such that $\varphi_j(x) \in O_n$ for $j \geq m$. Therefore

$$f^{-1}(O) \cap N^c = \bigcup_{m,n \in \mathbb{N}^\times} \bigcap_{j \geq m} \varphi_j^{-1}(O_n) \cap N^c . \tag{1.2}$$

But $\varphi_j^{-1}(O_n) \in \mathcal{A}$ for $n \in \mathbb{N}^\times$ and $j \in \mathbb{N}$, because φ_j is μ-simple. Hence (1.2) says that $f^{-1}(O) \cap N^c \in \mathcal{A}$.

Furthermore, the completeness of μ shows that $f^{-1}(O) \cap N$ is a μ-null set, and altogether we obtain

$$f^{-1}(O) = \bigl(f^{-1}(O) \cap N\bigr) \cup \bigl(f^{-1}(O) \cap N^c\bigr) \in \mathcal{A} .$$

"$\Leftarrow$" Suppose f is μ-almost separable valued and $\mathcal{A}$-measurable.

(iii) We consider first the case $\mu(X) < \infty$. Take $n \in \mathbb{N}$. By assumption, there is a μ-null set N such that $f(N^c)$ is separable. If $\{\, e_j \ ; \ j \in \mathbb{N} \,\}$ is a countable dense subset of $f(N^c)$, the collection $\{\, \mathbb{B}\bigl(e_j, 1/(n+1)\bigr) \ ; \ j \in \mathbb{N} \,\}$ covers the set $f(N^c)$, and thus

$$X = N \cup \bigcup_{j \in \mathbb{N}} f^{-1}\Bigl(\mathbb{B}\bigl(e_j, 1/(n+1)\bigr)\Bigr) .$$

Since f is $\mathcal{A}$-measurable, $X_{j,n} := f^{-1}\bigl(\mathbb{B}\bigl(e_j, 1/(n+1)\bigr)\bigr)$ belongs to $\mathcal{A}$ for every $(j, n) \in \mathbb{N}^2$. The continuity of μ from below and the assumption $\mu(X) < \infty$ then imply that there are $m_n \in \mathbb{N}^\times$ and $Y_n \in \mathcal{A}$ such that

$$\bigcup_{j=0}^{m_n} X_{j,n} = Y_n^c \quad \text{and} \quad \mu(Y_n) < \frac{1}{2^{n+1}} .$$

Now define $\varphi_n \in E^X$ through

$$\varphi_n(x) := \begin{cases} e_0 & \text{if } x \in X_{0,n} , \\ e_j & \text{if } x \in X_{j,n} \backslash \bigcup_{k=0}^{j-1} X_{k,n} \text{ for } 1 \leq j \leq m_n , \\ 0 & \text{otherwise} . \end{cases}$$

Obviously $\varphi_n \in \mathcal{S}(X,\mu,E)$ for $n \in \mathbb{N}$, and

$$|\varphi_n(x) - f(x)| < 1/(n+1) \quad \text{for } x \in Y_n^c \ .$$

The decreasing sequence $Z_n := \bigcup_{k=0}^{\infty} Y_{n+k}$ satisfies

$$\mu(Z_n) \le \sum_{k=0}^{\infty} \mu(Y_{n+k}) \le \frac{1}{2^n} \quad \text{for } n \in \mathbb{N} \ .$$

It therefore follows from the continuity of μ from above that $Z := \bigcap_{n\in\mathbb{N}} Z_n$ is μ-null. We now set

$$\psi_n(x) := \begin{cases} \varphi_n(x) & \text{if } x \in Z_n^c \ , \\ 0 & \text{if } x \in Z_n \ . \end{cases}$$

Then (ψ_n) is a sequence in $\mathcal{S}(X,\mu,E)$. Also there is for every $x \in Z^c = \bigcup_n Z_n^c$ an $m \in \mathbb{N}$ such that $x \in Z_m^c$. Since $Z_m^c \subset Z_n^c$ for $n \ge m$, it follows that

$$|\psi_n(x) - f(x)| = |\varphi_n(x) - f(x)| < 1/(n+1) \ .$$

Altogether, $\lim \psi_n(x) = f(x)$ for every $x \in Z^c$. Therefore f is μ-measurable.

(iv) Finally, we consider the case $\mu(X) = \infty$. Remark IX.2.4(c) shows there is a disjoint sequence (X_j) in $\mathcal{A}$ such that $\bigcup_j X_j = X$ and $\mu(X_j) < \infty$. By part (iii), there exist for each $j \in \mathbb{N}$ a sequence $(\varphi_{j,k})_{k\in\mathbb{N}}$ in $\mathcal{S}(X,\mu,E)$ and a μ-null set N_j such that $\lim_k \varphi_{j,k}(x) = f(x)$ for every $x \in X_j \cap N_j^c$. With $N := \bigcup_j N_j$ and

$$\varphi_k(x) := \begin{cases} \varphi_{j,k}(x) & \text{if } x \in X_j \ , \quad j \in \{0,\ldots,k\} \ , \\ 0 & \text{if } x \notin \bigcup_{j=0}^{k} X_j \end{cases}$$

for $k \in \mathbb{N}$, we have $\varphi_k \in \mathcal{S}(X,\mu,E)$ and $\lim_k \varphi_k(x) = f(x)$ for $x \in N^c$. The result follows because N is μ-null. ■

1.5 Corollary *Suppose E is separable and $f \in E^X$. The following statements are equivalent*:

(i) *f is μ-measurable.*

(ii) *f is $\mathcal{A}$-measurable.*

(iii) *$f^{-1}(\mathcal{S}) \subset \mathcal{A}$ for some $\mathcal{S} \subset \mathfrak{P}(E)$ such that $\mathcal{A}_\sigma(\mathcal{S}) = \mathcal{B}(E)$.*

(iv) *$f^{-1}(\mathcal{S}) \subset \mathcal{A}$ for any $\mathcal{S} \subset \mathfrak{P}(E)$ such that $\mathcal{A}_\sigma(\mathcal{S}) = \mathcal{B}(E)$.*

Proof This follows from Theorem 1.4, Remark 1.3(c), and Exercise IX.1.6. ■

1.6 Remark The proof of Theorem 1.4 and Remark 1.3(c) show that Corollary 1.5 remains true for incomplete measure spaces. ■

Without much effort, we obtain from Corollary 1.5 the following properties of μ-measurable functions.

1.7 Theorem

(i) *If E and F are separable Banach spaces and if we have maps $f \in \mathcal{L}_0(X,\mu,E)$ and $g \in C\big(f(X),F\big)$, then $g \circ f$ belongs to $\mathcal{L}_0(X,\mu,F)$. In particular, $|f| \in \mathcal{L}_0(X,\mu,\mathbb{R})$.*

(ii) *A map $f = (f_1,\ldots,f_n): X \to \mathbb{K}^n$ is μ-measurable if and only if each of its components f_j is.*

(iii) *Let $g, h \in \mathbb{R}^X$. Then $f = g + ih$ is μ-measurable if and only if g and h are.*

(iv) *If $f \in \mathcal{L}_0(X,\mu,E)$ and $g \in \mathcal{L}_0(X,\mu,F)$, then $(f,g) \in \mathcal{L}_0(X,\mu,E \times F)$.*

Proof (i) Let O be open in F. Since g is continuous, $g^{-1}(O)$ is open in $f(X)$. By Proposition III.2.26, there is an open subset U of E such that $g^{-1}(O) = f(X) \cap U$. Since f is Lebesgue measurable, $f^{-1}(U)$ belongs to $\mathcal{A}$ by Corollary 1.5. Because

$$(g \circ f)^{-1}(O) = f^{-1}\big(g^{-1}(O)\big) = f^{-1}\big(f(X) \cap U\big) = f^{-1}(U) ,$$

the claim follows from another application of Corollary 1.5.

(ii) The implication "$\Rightarrow$" follows from (i), because $f_j = \mathrm{pr}_j \circ f$ for $1 \le j \le n$.

"$\Leftarrow$" We consider first the case $\mathbb{K} = \mathbb{R}$. Take $I \in \mathbb{J}(n)$, and write it as $I = \prod_{j=1}^n I_j$, where $I_j \in \mathbb{J}(1)$ for $1 \le j \le n$. Because each $f_j^{-1}(I_j)$ belongs to $\mathcal{A}$, so does $f^{-1}(I) = \bigcap_{j=1}^n f_j^{-1}(I_j)$, that is, we have $f^{-1}\big(\mathbb{J}(n)\big) \subset \mathcal{A}$. Also, we know from Theorem IX.1.11 that $\mathcal{A}_\sigma\big(\mathbb{J}(n)\big) = \mathcal{B}^n$. Therefore Corollary 1.5 implies that f is μ-measurable.

Using the identification $\mathbb{C}^n = \mathbb{R}^{2n}$, the case $\mathbb{K} = \mathbb{C}$ follows immediately from what was just shown.

(iii) is a special case of (ii), and we leave (iv) as an exercise. ■

Measurable $\overline{\mathbb{R}}$-valued functions

In the theory of integration, it is useful to consider not only real-valued functions but also maps into the extended number line $\overline{\mathbb{R}}$. Such maps are called **$\overline{\mathbb{R}}$-valued** functions. An $\overline{\mathbb{R}}$-valued function $f: X \to \overline{\mathbb{R}}$ is said to be ***μ*-measurable** if $\mathcal{A}$ contains $f^{-1}(-\infty)$, $f^{-1}(\infty)$, and $f^{-1}(O)$ for every open subset O of $\mathbb{R}$. We denote the set of all μ-measurable $\overline{\mathbb{R}}$-valued functions on X by $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$.

1.8 Remarks **(a)** Any real-valued function $f: X \to \mathbb{R}$ can be regarded as an $\overline{\mathbb{R}}$-valued one. Thus there are in principle two notions of measurability that apply to f. But since $f^{-1}\big(\{-\infty,\infty\}\big) = \emptyset$, Corollary 1.5 implies that f is μ-measurable as a real-valued function if and only if it is μ-measurable as an $\overline{\mathbb{R}}$-valued function.

(b) Note that $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$ is *not* a vector space. ■

In the next result, we list simple measurability criteria for $\overline{\mathbb{R}}$-valued functions.

1.9 Proposition *For an $\overline{\mathbb{R}}$-valued function $f : X \to \overline{\mathbb{R}}$, the following statements are equivalent:*

(i) $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$.

(ii) $[f < \alpha] \in \mathcal{A}$ *for every* $\alpha \in \mathbb{Q}$ [*or* $\alpha \in \mathbb{R}$].

(iii) $[f \leq \alpha] \in \mathcal{A}$ *for every* $\alpha \in \mathbb{Q}$ [*or* $\alpha \in \mathbb{R}$].

(iv) $[f > \alpha] \in \mathcal{A}$ *for every* $\alpha \in \mathbb{Q}$ [*or* $\alpha \in \mathbb{R}$].

(v) $[f \geq \alpha] \in \mathcal{A}$ *for every* $\alpha \in \mathbb{Q}$ [*or* $\alpha \in \mathbb{R}$].

Proof "(i)⇒(ii)" The sets $f^{-1}(-\infty)$ and $f^{-1}\big((-\infty, \alpha)\big)$ with $\alpha \in \mathbb{Q}$ [or $\alpha \in \mathbb{R}$] belong to $\mathcal{A}$. Because

$$[f < \alpha] = f^{-1}\big([-\infty, \alpha)\big) = f^{-1}(-\infty) \cup f^{-1}\big((-\infty, \alpha)\big) \ ,$$

this is also true of $[f < \alpha]$.

The implications "(ii)⇒(iii)⇒(iv)⇒(v)" follow from the identities

$$[f \leq \alpha] = \bigcap_{j=1}^{\infty} [f < \alpha + 1/j] \ , \quad [f > \alpha] = [f \leq \alpha]^c \ , \quad [f \geq \alpha] = \bigcap_{j=1}^{\infty} [f > \alpha - 1/j] \ .$$

"(v)⇒(i)" Suppose O is open in $\mathbb{R}$. By Proposition IX.5.6, there exist $(\alpha_j), (\beta_j) \in \mathbb{Q}^{\mathbb{N}}$ such that $O = \bigcup_j [\alpha_j, \beta_j)$. Therefore

$$f^{-1}(O) = \bigcup_{j \in \mathbb{N}} f^{-1}\big([\alpha_j, \beta_j)\big) = \bigcup_{j \in \mathbb{N}} \big([f \geq \alpha_j] \cap [f < \beta_j]\big) \ ,$$

and because $[f < \alpha] = [f \geq \alpha]^c$, we conclude that $f^{-1}(O)$ belongs to $\mathcal{A}$. In addition

$$f^{-1}(-\infty) = \bigcap_{j \in \mathbb{N}} [f < -j] \quad \text{and} \quad f^{-1}(\infty) = \bigcap_{j \in \mathbb{N}} [f > j] \ .$$

Thus $f^{-1}(\pm\infty)$ also lies in $\mathcal{A}$. ■

The lattice of measurable $\overline{\mathbb{R}}$-valued functions

An ordered set $V = (V, \leq)$ is called a **lattice** if for every pair $(a, b) \in V \times V$, the infimum $a \wedge b$ and the supremum $a \vee b$ exist in V. A subset $U \subset V$ is a **sublattice** of V if U is a lattice when given the ordering induced by V. An ordered vector space that is also a lattice is called a **vector lattice**. If a vector subspace of a vector lattice is a sublattice, we call it a **vector sublattice**.

1.10 Examples **(a)** Suppose V is a lattice [or vector lattice]. Then V^X is a lattice [or vector lattice] with respect to the pointwise ordering.

(b) $\overline{\mathbb{R}}$ is a lattice, and $\mathbb{R}$ is a vector lattice.

(c) The vector lattice $\mathbb{R}^X$ satisfies

$$f \vee g = (f + g + |f - g|)/2 \ , \quad f \wedge g = (f + g - |f - g|)/2 \ .$$

(d) $B(X, \mathbb{R})$ is a vector sublattice of $\mathbb{R}^X$.

(e) Suppose X is a topological space. Then $C(X, \mathbb{R})$ is a vector sublattice of $\mathbb{R}^X$.

Proof This follows from (c) and the fact that $|f|$ is continuous if f is. ■

(f) $\mathcal{S}(X, \mu, \mathbb{R})$ and $\mathcal{L}_0(X, \mu, \mathbb{R})$ are vector sublattices of $\mathbb{R}^X$.

Proof The first statement is clear. The second follows from (c) and Theorem 1.7 or Remark 1.2(d). ■

(g) Suppose V is a vector lattice and $x, y, z \in V$. Then

$$\begin{aligned} (x \vee y) + z &= (x + z) \vee (y + z) \ , \\ (-x) \vee (-y) &= -(x \wedge y) \ , \\ x + y &= (x \vee y) + (x \wedge y) \ . \end{aligned}$$

Proof If $u \in V$ satisfies $u \geq x$ and $u \geq y$, then clearly $u + z \geq (x + z) \vee (y + z)$. Hence

$$(x \vee y) + z \geq (x + z) \vee (y + z) \ .$$

Suppose $v \geq (x + z) \vee (y + z)$. Then $v - z \geq x$ and $v - z \geq y$, and hence $v \geq (x \vee y) + z$. Because this holds for every upper bound v of $\{x + z, y + z\}$, it follows that

$$(x + z) \vee (y + z) \geq (x \vee y) + z \ .$$

This proves the first equality. The second is none other than the trivial relation

$$\sup\{-x, -y\} = \sup\bigl(-\{x, y\}\bigr) = -\inf\{x, y\} \ .$$

Using this, we now find

$$\begin{aligned} x \vee y &= \bigl(-y + (x + y)\bigr) \vee \bigl(-x + (x + y)\bigr) = \bigl((-y) \vee (-x)\bigr) + (x + y) \\ &= -(x \wedge y) + (x + y) \ , \end{aligned}$$

which proves the last claim. ■

(h) Suppose V is a vector lattice. For $x \in V$, we set

$$x^+ := x \vee 0 \ , \quad x^- := (-x) \vee 0 \ , \quad |x| := x \vee (-x) \ .$$

Then[6]

$$x = x^+ - x^- \ , \quad |x| = x^+ + x^- \ , \quad x^+ \wedge x^- = 0 \ .$$

[6] See footnote 8 in Section II.8.

Proof The first claim follows easily from (g). With this and (g), we find

$$x^+ + x^- = x + 2x^- = x + \big((-2x) \vee 0\big) = (-x) \vee x = |x| \ .$$

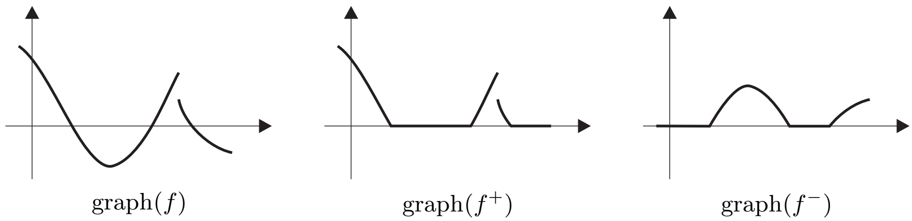
Analogously, we have

$$(x^+ \wedge x^-) - x^- = (x^+ - x^-) \wedge (x^- - x^-) = x \wedge 0 = -x^- \ ,$$

and therefore $x^+ \wedge x^- = 0$. ■

If V is a vector lattice and $x \in V$, we call x^+ the **positive part** and x^- the **negative part** of x, and $|x|$ is the **modulus**[7] of x. Clearly $x^+ \geq 0$, $x^- \geq 0$, and $|x| \geq 0$.

The figures below illustrate the positive and negative parts of an element f of the vector lattice $\mathbb{R}^X$.



Suppose $f \in \overline{\mathbb{R}}^X$. Then $f^+ := f \vee 0$ is called the **positive part** of f, and $f^- := 0 \vee (-f)$ the **negative part** of f. These terms are chosen in obvious analogy to the case of the vector lattice $\mathbb{R}^X$.[8] Here too we have

$$f^+ \geq 0 \ , \quad f^- \geq 0 \ , \quad f = f^+ - f^- \ , \quad |f| = f^+ + f^- \ .$$

The next result shows that $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$ is a sublattice of $\overline{\mathbb{R}}^X$ and that it is closed under countably many lattice operations.

1.11 Proposition *Suppose $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$, (f_j) is a sequence in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$, and $k \in \mathbb{N}$. Then each of the $\overline{\mathbb{R}}$-valued functions*

$$f^+ \ , \quad f^- \ , \quad |f| \ , \quad \max_{0 \leq j \leq k} f_j \ , \quad \min_{0 \leq j \leq k} f_j \ , \quad \sup_j f_j \ , \quad \inf_j f_j \ , \quad \varlimsup_j f_j \ , \quad \varliminf_j f_j$$

belongs to $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$.

Proof (i) Suppose $\alpha \in \mathbb{R}$. From Proposition 1.9, we know that $[f_j > \alpha]$ belongs to $\mathcal{A}$ for $j \in \mathbb{N}$. Therefore this is also true of

$$\big[\sup_j f_j > \alpha\big] = \bigcup_j [f_j > \alpha] \ ,$$

and Proposition 1.9 implies that $\sup_j f_j$ is μ-measurable.

[7] This is not to be confused with the norm of the vector x if V is a normed vector space. The modulus of $x \in V$ is always a vector in V, whereas the norm is a nonnegative number.

[8] Remember that $\overline{\mathbb{R}}^X$ is a lattice but not a vector lattice.

(ii) Because f_j belongs to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$, so does $-f_j$. It then follows from (i) that the function $\inf_j f_j = -\sup_j(-f_j)$ is μ-measurable.

(iii) For $j \in \mathbb{N}$, set

$$g_j := \begin{cases} f_j & \text{if } 0 \le j \le k \ , \\ f_k & \text{if } \quad j > k \ . \end{cases}$$

Because of (i), $\sup_j g_j = \max_{0 \le j \le k} f_j$ belongs to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$. Analogously, one shows that $\min_{0 \le j \le k} f_j$ is μ-measurable.

(iv) From (iii), it follows that f^+, f^-, and $|f|$ belong to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$.

(v) We have

$$\overline{\lim_j}\, f_j = \inf_j \sup_{k \ge j} f_k \quad \text{and} \quad \underline{\lim_j}\, f_j = \sup_j \inf_{k \ge j} f_k \ .$$

Therefore by (i) and (ii), $\overline{\lim}_j f_j$ and $\underline{\lim}_j f_j$ also belong to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$. ■

The positive cone $\mathcal{S}(X,\mu,\mathbb{R}^+)$ of $\mathcal{S}(X,\mu,\mathbb{R})$ is the set of all $f \in \mathcal{S}(X,\mu,\mathbb{R})$ such that $f(X) \subset \mathbb{R}^+$; see Remarks VI.4.7(b) and (d). Therefore it is natural to denote it by $\mathcal{S}(X,\mu,\mathbb{R}^+)$. Similarly, if $\overline{\mathbb{R}}^+ := [0,\infty]$ is the nonnegative part of the extended number line $\overline{\mathbb{R}}$, we denote by $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$ the set of all nowhere negative μ-measurable $\overline{\mathbb{R}}$-valued functions on X.

The set $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$ has an interesting characterization:

1.12 Theorem *For $f : X \to \overline{\mathbb{R}}^+$, the following statements are equivalent*:

(i) $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$.

(ii) *There is an increasing sequence (f_j) in $\mathcal{S}(X,\mu,\mathbb{R}^+)$ such that $f_j \to f$ for $j \to \infty$.*

Proof "(i)⇒(ii)" By the σ-finiteness of $(\mathcal{A},\mu)$, it suffices to consider the case $\mu(X) < \infty$ (compare part (iv) in the proof of Theorem 1.4). So for $j,k \in \mathbb{N}$, set

$$A_{j,k} := \begin{cases} \bigl[k2^{-j} \le f < (k+1)2^{-j}\bigr] & \text{if } k = 0,\ldots,j\,2^j - 1 \ , \\ [f \ge j] & \text{if } k = j\,2^j \ . \end{cases}$$

The sets $A_{j,k}$ are obviously disjoint for $k = 0,\ldots,j\,2^j$ and by Proposition 1.9 they lie in $\mathcal{A}$. Since $\mu(X) < \infty$, each $A_{j,k}$ has finite measure. By Remark 1.2(b), then,

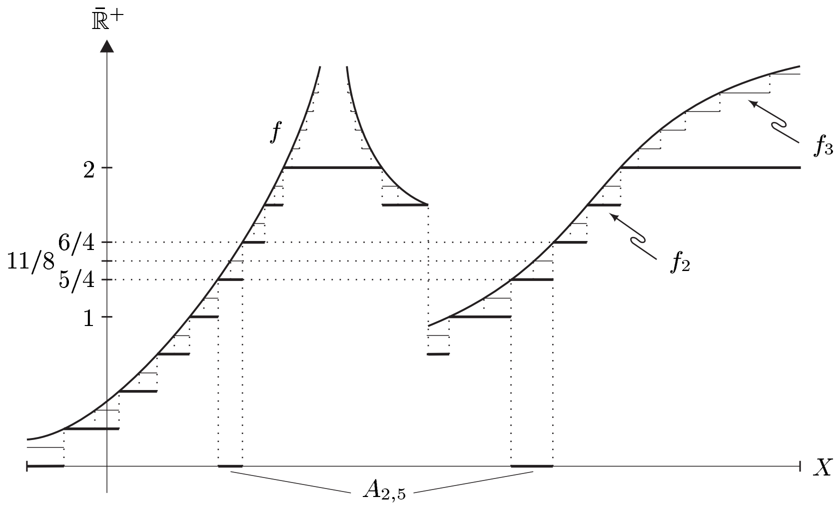
$$f_j := \sum_{k=0}^{j\,2^j} k2^{-j} \chi_{A_{j,k}} \quad \text{for } j \in \mathbb{N}$$

belongs to $\mathcal{S}(X,\mu,\mathbb{R})$. Further one verifies that $0 \le f_j \le f_{j+1}$ for $j \in \mathbb{N}$.

Now suppose $x \in X$. If $f(x) = \infty$, we have $f_j(x) = j$, so $\lim_j f_j(x) = f(x)$. On the other hand, if $f(x) < \infty$, then $f_j(x) \le f(x) < f_j(x) + 2^{-j}$ for $j > f(x)$, so $\lim_j f_j(x) = f(x)$ in this case as well. This shows (f_j) converges pointwise to f.

"(ii)⇒(i)" This follows from Proposition 1.11. ■

Here is an illustration of the construction of the $A_{j,k}$ in the proof of Theorem 1.12:



1.13 Corollary

(i) *For every* $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$, *there is a sequence* (f_j) *in* $\mathcal{S}(X, \mu, \mathbb{R})$ *such that* $f_j \to f$.

(ii) *Suppose* $f \in \mathcal{L}_0(X, \mu, \mathbb{R}^+)$ *is bounded. Then there is an increasing sequence* (f_j) *in* $\mathcal{S}(X, \mu, \mathbb{R}^+)$ *that converges uniformly to* f.

(iii) *Suppose* (f_j) *is a sequence in* $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$. *Then* $\sum_j f_j \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$.

Proof (i) In view of the decomposition $f = f^+ - f^-$, this follows from Theorem 1.12 and Remark 1.2(a).

(ii) Suppose $f \in \mathcal{L}_0(X, \mu, \mathbb{R}^+)$ is bounded. For the sequence (f_j) constructed in the proof of Theorem 1.12, we have

$$f_j(x) \le f(x) < f_j(x) + 2^{-j} \quad \text{for } j > \|f\|_\infty .$$

Thus (f_j) converges uniformly to f.

(iii) By Theorem 1.12, there is for every $j \in \mathbb{N}$ an increasing sequence $(\varphi_{j,k})_{k\in\mathbb{N}}$ in $\mathcal{S}(X, \mu, \mathbb{R}^+)$ such that $\varphi_{j,k} \uparrow f_j$ for $k \to \infty$. Set $s_{k,n} := \sum_{j=0}^k \varphi_{j,n}$ for $k, n \in \mathbb{N}$. Then $(s_{k,n})_{n\in\mathbb{N}}$ is an increasing sequence in $\mathcal{S}(X, \mu, \mathbb{R}^+)$ that converges to $s_k := \sum_{j=0}^k f_j$ as $n \to \infty$. By Theorem 1.12, then, (s_k) is a sequence in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ such that $\lim_k s_k = \sup_k s_k = \sum_{j=0}^\infty f_j$. The claim now follows from Proposition 1.11. ■

Pointwise limits of measurable functions

Let (f_j) be a pointwise convergent sequence in $\mathcal{L}_0(X,\mu,\mathbb{R})$. By Proposition 1.11, $f := \lim_j f_j$ is also in $\mathcal{L}_0(X,\mu,\mathbb{R})$. We will now derive an analogous statement for vector-valued sequences of functions.

1.14 Theorem *Suppose (f_j) is a sequence in $\mathcal{L}_0(X,\mu,E)$ and $f \in E^X$. If (f_j) converges μ-almost everywhere to f, then f is μ-measurable.*

Proof (i) We show first that f is μ-almost separable valued. By assumption, there is a μ-null set M such that $f_j(x) \to f(x)$ as $j \to \infty$, for any $x \in M^c$. For every $j \in \mathbb{N}$, there exists by Theorem 1.4 a μ-null set N_j such that $f_j(N_j^c)$ is separable, hence also a countable set B_j that is dense in $f_j(N_j^c)$:

$$B_j \subset f_j(N_j^c) \subset \overline{B}_j \quad \text{for } j \in \mathbb{N} .$$

With $B := \bigcup_j B_j$, we see from Corollary III.2.13(i) that $\bigcup_j \overline{B}_j \subset \overline{B}$, and we find

$$\bigcup_{j\in\mathbb{N}} f_j(N_j^c) \subset \bigcup_{j\in\mathbb{N}} \overline{B}_j \subset \overline{B} .$$

Finally let $N := M \cup \bigcup_j N_j$. Then N is a μ-null set satisfying, for any $k \in \mathbb{N}$, $N^c = M^c \cap \bigcap_j N_j^c \subset N_k^c$. Because $\lim_j f_j(x) = f(x)$ for $x \in M^c$, we thus have

$$f(N^c) \subset \overline{\bigcup_{j\in\mathbb{N}} f_j(N_j^c)} \subset \overline{\overline{B}} = \overline{B} .$$

Because B is countable, Remark 1.3(b) shows that $f(N^c)$ is separable.

(ii) Now we show that f is $\mathcal{A}$-measurable. Let O be open in E, and define $O_n := \{\, x \in O \ ;\ \operatorname{dist}(x, O^c) > 1/n \,\}$ for $n \in \mathbb{N}^\times$. As in (1.2), it follows that

$$f^{-1}(O) \cap M^c = \bigcup_{m,n\in\mathbb{N}^\times} \bigcap_{j\geq m} f_j^{-1}(O_n) \cap M^c .$$

By Theorem 1.4, $f_j^{-1}(O_n)$ belongs to $\mathcal{A}$ for every $j,n \in \mathbb{N}^\times$. Therefore this also applies to $f^{-1}(O)\cap M^c$. Moreover, the completeness of μ implies that $f^{-1}(O)\cap M$ is a μ-null set, and altogether we find

$$f^{-1}(O) = \bigl(f^{-1}(O) \cap M^c\bigr) \cup \bigl(f^{-1}(O) \cap M\bigr) \in \mathcal{A} .$$

The claim now follows from Theorem 1.4. ■

1.15 Remark Theorem 1.14 generally fails for incomplete measure spaces.

Proof Let C be the Cantor set. In the proof of Corollary IX.5.29, it was shown that C contains a Borel nonmeasurable subset $N \subset C$. We take $f_j := \chi_C$ for $j \in \mathbb{N}$ and $f := \chi_N$. Remark 1.2(b) and the compactness of C imply $\chi_C \in \mathcal{S}(\mathbb{R},\beta_1,\mathbb{R})$. Also $f_j(x) = f(x)$ for $x \in C^c \subset N^c$ and $j \in \mathbb{N}$. Because C has measure zero, (f_j) converges β_1-a.e. to f. However, because $[f > 0] = N \notin \mathcal{B}^1$, Proposition 1.9 says that f is not in $\mathcal{L}_0(\mathbb{R},\beta_1,\mathbb{R})$. ■

Radon measures

We conclude this section by exploring how measurability and continuity are related in vector-valued functions. Besides proving a simple measurability criterion, we prove Luzin's theorem, which exposes a surprisingly close connection between continuous and Borel measurable functions.

A metric space $X = (X, d)$ is said to be ***σ*-compact** if X is locally compact and there is a sequence $(X_j)_{j\in\mathbb{N}}$ of compact subsets of X such that $X = \bigcup_j X_j$.

Suppose X is a σ-compact metric space. A **Radon measure** on X is a regular, locally finite measure on a σ-algebra $\mathcal{A}$ over X such that $\mathcal{A} \supset \mathcal{B}(X)$. We say a Radon measure μ is **massive** if μ is complete and every nonempty open subset O of X satisfies $\mu(O) > 0$.

1.16 Remarks **(a)** Every σ-compact metric space is a σ-compact set in the sense of the definition of Section IX.5; however, a countable union of compact subsets of a metric space is not necessarily a σ-compact metric space.

Proof The first statement is clear. For the second, consider $\mathbb{Q} \subset \mathbb{R}$. ■

(b) Every Radon measure is σ-finite.

Proof This follows from Remark IX.5.3(b). ■

(c) Suppose X is a locally compact metric space. Then there is for every compact subset K of X a relatively compact[9] open superset of K.

Proof For every $x \in X$, we find a relatively compact open neighborhood $O(x)$ of x. Because K is compact, there are $x_0, \ldots, x_m \in K$ such that $O := \bigcup_{j=0}^m O(x_j)$ is an open superset of K. Corollary III.2.13(iii) implies $\overline{O} = \bigcup_{j=0}^m \overline{O(x_j)}$. Therefore $\overline{O}$ is compact. ■

(d) Every open subset of $\mathbb{R}^n$ is a σ-compact metric space.

Proof Let $X \subset \mathbb{R}^n$ be open and nonempty. For every $x \in X$, there is $r > 0$ such that $\overline{\mathbb{B}}(x, r) \subset X$. Because $\overline{\mathbb{B}}(x, r)$ is compact, X is then a locally compact metric space.

For $j \in \mathbb{N}^\times$, define[10]

$$U_j := \{\, x \in X \ ; \ \operatorname{dist}(x, U^c) > 1/j \,\} \cap \mathbb{B}(0, j) \ . \tag{1.3}$$

By Examples III.1.3(l) and III.2.22(c), the set U_j is open. Also $U_j \subset \overline{U}_j \subset U_{j+1}$, and $\bigcup_j \overline{U}_j \subset \bigcup_j U_j = X$. In particular, there exists $j_0 \in \mathbb{N}^\times$ such that $U_j \neq \emptyset$ for $j \geq j_0$. Because $\overline{U}_j$ is compact by the Heine–Borel theorem, the claim follows. ■

(e) For a locally compact metric space X, the following statements are equivalent:

(i) X is σ-compact.

(ii) X is the union of a sequence $(U_j)_{j\in\mathbb{N}}$ of relatively compact open subsets with $\overline{U}_j \subset U_{j+1}$ for $j \in \mathbb{N}$.

[9]A subset A of a topological space is said to be **relatively compact** if $\overline{A}$ is compact.

[10]$\operatorname{dist}(x, \emptyset) := \infty$.

(iii) X is a Lindelöf space.

(iv) X satisfies the second countability axiom.

(v) X is separable.

Proof "(i)⇒(ii)" Let $(X_j)_{j\in\mathbb{N}}$ be a sequence of compact sets in X such that $X = \bigcup_j X_j$. By (c), there is a relatively compact open superset U_0 of X_0. Inductively choose relatively compact open subsets U_j such that $U_j \supset \overline{U}_{j-1} \cup X_j$ for $j \geq 1$. Clearly $X = \bigcup_j U_j$.

"(ii)⇒(iii)" Suppose $\mathcal{O} := \{\, O_\alpha \ ;\ \alpha \in \mathsf{A} \,\}$ is an open cover of X. For every $j \in \mathbb{N}$, inductively choose $m(j) \in \mathbb{N}$ and $\alpha_0, \ldots, \alpha_{m(j)} \in \mathsf{A}$ such that $\overline{U}_j \subset \bigcup_{k=0}^{m(j)} O_{\alpha_k}$. Then $\{\, O_{\alpha_k} \ ;\ k = 0, \ldots, m(j),\ j \in \mathbb{N} \,\}$ is a countable subcover of $\mathcal{O}$ for X.

"(iii)⇒(i)" By assumption, there is a sequence (x_j) in X and relatively compact open neighborhoods $O(x_j)$ of x_j ($j \in \mathbb{N}$) such that $X = \bigcup_{j\in\mathbb{N}} O(x_j)$. So $X = \bigcup_{j\in\mathbb{N}} \overline{O(x_j)}$, showing that X is σ-compact.

The remaining equivalences follow from Proposition IX.1.8. ■

(f) Every locally finite Borel measure on a σ-compact metric space is regular and is therefore a Radon measure.

Proof This follows from (e) and Corollary VIII.1.12 in [Els99]. ■

(g) Finite Borel measures on (nonmetrizable) compact topological spaces need not be regular.

Proof See [Flo81, Example A4.5, S. 350]. ■

(h) Lebesgue n-measure, λ_n, is a massive Radon measure on $\mathbb{R}^n$.

Proof This follows from Theorems IX.5.1 and IX.5.4. ■

(i) The s-dimensional Hausdorff measure $\mathcal{H}^s$ is a Radon measure on $\mathbb{R}^n$ only when $s \geq n$. It is massive if and only if $s = n$.

Proof Every Borel set is $\mathcal{H}^s$-measurable, by Example IX.4.4(c) and Theorem IX.4.3. The regularity of $\mathcal{H}^s$ for $s > 0$ follows from Corollary IX.5.22 and Theorem IX.5.4.

Suppose O is open in $\mathbb{R}^n$ and nonempty. Because O has Hausdorff dimension n (Exercise IX.3.6), it follows that

$$\mathcal{H}^s(O) = \begin{cases} 0 & \text{if } s > n\ , \\ \infty & \text{if } s < n\ . \end{cases}$$

Therefore $\mathcal{H}^s$ cannot be a Radon measure on $\mathbb{R}^n$ if $s < n$. If $s > n$, on the other hand, $\mathcal{H}^s$ is a nonmassive Radon measure.

Lemma IX.5.21 shows that $\mathcal{H}^n$ is locally finite and therefore a Radon measure on $\mathbb{R}^n$. Finally, Corollary IX.5.22 implies $\mathcal{H}^n(O) > 0$, and we are done. ■

(j) Suppose $F : \mathbb{R} \to \mathbb{R}$ is measure generating, and denote by μ_F the Lebesgue–Stieltjes measure on $\mathbb{R}$ induced by F. Then μ_F is a Radon measure on $\mathbb{R}$, and is massive if only if F is strictly increasing.

Proof This follows from Example IX.4.4(b), Theorem IX.4.3, Exercise IX.5.19, and Proposition IX.3.5. ■

1.17 Theorem *Suppose μ is a complete Radon measure on X. Then $C(X,E)$ is a vector subspace of $\mathcal{L}_0(X,\mu,E)$.*

Proof Take $f \in C(X,E)$ and let (X_j) be a sequence of compact sets in X such that $X = \bigcup_j X_j$. According to Exercise IX.1.6(b), f is Borel measurable and therefore $\mathcal{A}$-measurable, where $\mathcal{A}$ is the domain of μ. By Remark 1.16(e), $f(X_j)$, being a compact subset of E, is separable. Therefore $f(X) = \bigcup_j f(X_j)$ is also separable, and the claim follows from Theorem 1.4. ■

1.18 Theorem (Luzin) *Suppose X is a σ-compact metric space, μ is a complete Radon measure on X, and $f \in \mathcal{L}_0(X,\mu,E)$. Then for every μ-measurable set A of finite measure and for every $\varepsilon > 0$, there is a compact subset K of X such that $\mu(A \backslash K) < \varepsilon$ and $f \,|\, K \in C(K,E)$.*

Proof (i) Because X is σ-compact, we can find a compact set $\widetilde{X}$ such that $\mu\bigl(A \setminus \widetilde{X}\bigr) < \varepsilon/2$. We set $\widetilde{f} := f \,\big|\, \widetilde{X}$ and $\widetilde{A} := A \cap \widetilde{X}$. Then $\mu\bigl(\widetilde{X}\bigr) < \infty$.

(ii) By Theorem 1.4, there is a μ-null set N of $\widetilde{X}$ such that $\widetilde{f}(N^c)$ is separable. Therefore by Proposition IX.1.8, there is a countable basis $\{\, \widetilde{V}_j \; ; \; j \in \mathbb{N} \,\}$ of $\widetilde{f}(N^c)$, and because of Proposition III.2.26, there exist open subsets V_j in E such that $\widetilde{V}_j = V_j \cap \widetilde{f}(N^c)$.

(iii) According to Theorem 1.4, $\widetilde{f}^{-1}(V_j)$ is μ-measurable for every $j \in \mathbb{N}$. Hence it follows from the regularity of μ and the finiteness of $\mu(\widetilde{X})$ that for every $j \in \mathbb{N}$ there exist a compact set K_j and an open set U_j with $K_j \subset \widetilde{f}^{-1}(V_j) \subset U_j$ and $\mu(U_j \backslash K_j) < \varepsilon 2^{-(j+3)}$. Putting $U := \bigcup_j (U_j \backslash K_j)$, we have $\mu(U) < \varepsilon/4$.

(iv) We set $Y := (U \cup N)^c$ and show that $\widetilde{f} \,\big|\, Y$ is continuous. To verify this, let V be open in E. Then there is a subset $\{\, V_{j_k} \; ; \; k \in \mathbb{N} \,\}$ of $\{\, V_j \; ; \; j \in \mathbb{N} \,\}$ such that $V \cap \widetilde{f}(N^c) = \bigcup_k V_{j_k} \cap \widetilde{f}(N^c)$. This implies

$$\widetilde{f}^{-1}(V) \cap N^c = \bigcup\nolimits_k \widetilde{f}^{-1}(V_{j_k}) \cap N^c \ .$$

Obviously $\widetilde{f}^{-1}(V_\ell) \cap Y \subset U_\ell \cap Y$ for $\ell \in \mathbb{N}$. Because

$$Y = U^c \cap N^c = \bigcap\nolimits_j (U_j^c \cup K_j) \cap N^c \subset \bigcap\nolimits_j \bigl(U_j^c \cup \widetilde{f}^{-1}(V_j)\bigr) \subset U_\ell^c \cup \widetilde{f}^{-1}(V_\ell) \ ,$$

it follows that $\widetilde{f}^{-1}(V_\ell) \cap Y = U_\ell \cap Y$, and we find

$$\bigl(\widetilde{f} \,|\, Y\bigr)^{-1}(V) = \widetilde{f}^{-1}(V) \cap N^c \cap U^c = \bigcup\nolimits_k U_{j_k} \cap Y \ .$$

Because $\bigcup_k U_{j_k}$ is open in X and therefore $\bigcup_k U_{j_k} \cap Y$ is open in Y, the continuity of $\widetilde{f} \,\big|\, Y$ follows.

(v) We apply once again the regularity of μ to deduce the existence of a compact subset K of the μ-measurable set Y such that $\mu(Y \backslash K) < \varepsilon/4$. Then

$\widetilde{f} \mid K$ belongs to $C(K, E)$, and

$$\mu(\widetilde{A} \setminus K) \le \mu(Y \setminus K) + \mu(Y^c \setminus K) \le \mu(Y \setminus K) + \mu(U) < \varepsilon/2 .$$

Because $\mu(A \setminus K) \le \mu(\widetilde{A} \setminus K) + \mu(A \setminus \widetilde{X}) < \varepsilon$, we are done. ■

Exercises

1 Suppose H is a separable Hilbert space. We say $f \in H^X$ is **weakly μ-measurable** if $(f \mid e)$ belongs to $\mathcal{L}_0(X, \mu, \mathbb{K})$ for every $e \in H$. Prove:

(a) If f is weakly μ-measurable, $|f|$ is μ-measurable.

(b) f is μ-measurable if and only if f is weakly μ-measurable.

(Hints: (a) Suppose $\{e_j \ ; \ j \in \mathbb{N}\}$ is a dense subset of $\overline{\mathbb{B}}_H$. Then

$$[\,|f| \le \alpha] = \bigcap_j [\,|(f \mid e_j)| \le \alpha] \quad \text{for } \alpha \in \mathbb{R} .$$

(b) "$\Leftarrow$" Using (a), we can construct as in the proof of Theorem 1.4 a sequence of a μ-simple functions that converge μ-a.e. to f.)

2 Denote by $\mathcal{S}(\mathbb{R}, E)$ the vector space of all E-valued admissible functions of $\mathbb{R}$ (see Section VI.8). Prove or disprove:

(a) $\mathcal{S}(\mathbb{R}, E) \subset \mathcal{L}_0(\mathbb{R}, \beta_1, E)$;

(b) $\mathcal{S}(\mathbb{R}, E) \supset \mathcal{L}_0(\mathbb{R}, \beta_1, E)$.

3 Prove the statement of Remark 1.2(e).

4 Show that every monotone $\overline{\mathbb{R}}$-valued function is Borel measurable.

5 Let $f, g \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$. Show that the sets $[f < g]$, $[f \le g]$, $[f = g]$, and $[f \ne g]$ belong to $\mathcal{A}$.

6 Suppose (f_j) is a sequence in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$. Show that

$$K := \{\, x \in X \ ; \ \lim_j f_j(x) \text{ exists in } \overline{\mathbb{R}} \,\}$$

is μ-measurable.

7 Suppose $f : X \to \overline{\mathbb{R}}$. Prove or disprove:

(a) $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}) \iff f^+, f^- \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$.

(b) $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}) \iff |f| \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$.

8 A nonempty subset B of $\overline{\mathbb{R}}^X$ is called a **Baire function space** if these statements hold:

(i) $\alpha \in \mathbb{R}$ and $f \in B$ imply $\alpha f \in B$.

(ii) If $f + g$ exists in $\overline{\mathbb{R}}^X$ for $f, g \in B$, then $f + g \in B$.

(iii) $\sup_j f_j$ belongs to B for every sequence (f_j) in B.

Prove:

(a) $\overline{\mathbb{R}}^X$ and $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$ are Baire function spaces.

(b) If $\{ B_\alpha \subset \overline{\mathbb{R}}^X \ ; \ \alpha \in \mathsf{A} \}$ is a family of Baire function spaces, then $\bigcap_{\alpha \in \mathsf{A}} B_\alpha$ is also a Baire function space.

9 For $C \subset \overline{\mathbb{R}}^X$, we call

$$\sigma(C) := \bigcap \{ B \subset \overline{\mathbb{R}}^X \ ; \ B \supset C, \ B \text{ is a Baire function space} \}$$

the **Baire function space generated by** C. By Exercise 8(b), $\sigma(C)$ is a well defined Baire function space. Show that

$$\sigma\big(\mathcal{S}(X,\mu,\mathbb{R})\big) = \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}) \ .$$

10 Prove that $\sigma\big(C(\mathbb{R}^n,\mathbb{R})\big) = \mathcal{L}_0(\mathbb{R}^n,\beta_n,\mathbb{R})$.

11 Show that the supremum of an uncountable family of measurable real-valued functions is generally not measurable.

12 A sequence (f_j) in E^X is said to be **μ-almost uniformly convergent** if for every $\delta > 0$ there is an $A \in \mathcal{A}$ with $\mu(A^c) < \delta$ such that the sequence $(f_j \,|\, A)$ converges uniformly.

(a) Suppose (f_j) is a μ-almost uniformly convergent sequence in $\mathcal{L}_0(X,\mu,E)$. Show there is an $f \in \mathcal{L}_0(X,\mu,E)$ such that $f_j \to f$ μ-a.e.

(b) Define $f_j(x) := x^j$ for $j \in \mathbb{N}$ and $x \in [0,1]$. Verify that (f_j) converges almost uniformly (with respect to Lebesgue measure), although there is no set $N \subset [0,1]$ of measure zero such that $(f_j \,|\, N^c)$ converges uniformly.

13 Suppose $(X,\mathcal{A},\mu)$ is a finite measure space and $f_j, f \in \mathcal{L}_0(X,\mu,E)$ with $f_j \to f$ μ-a.e. Prove:

(a) For $\varepsilon > 0$ and $\delta > 0$, there exist $k \in \mathbb{N}$ and $A \in \mathcal{A}$ such that $\mu(A^c) < \delta$ and $|f_j(x) - f(x)| < \varepsilon$ for $x \in A$ and $j \geq k$.

(b) The sequence (f_j) converges μ-almost uniformly to f (**Egorov's theorem**).

(c) Part (b) is generally false if $\mu(X) = \infty$.

(Hints: (a) Consider $K := [f_j \to f]$ and $K_k := \big[\, |f_j - f| < \varepsilon \ ; \ j \geq k \big]$, and apply the continuity of measures from above. (b) To obtain the A_j, choose $\varepsilon := 1/j$ and $\delta := \delta\, 2^{-j}$ in (a), and let $A := \bigcup_j A_j$. (c) Consider the measure space $(X,\mathcal{A},\mu) = \big(\mathbb{R}, \lambda_1, \mathcal{L}(1)\big)$ and set $f_j := \chi_{[j,j+1]}$.)

14 Suppose $(X,\mathcal{A},\mu)$ is a measure space and $f_j, f \in \mathcal{L}_0(X,\mu,E)$. We say (f_j) **converges in measure** to f if $\lim_{j\to\infty} \mu\big([\, |f_j - f| \geq \varepsilon]\big) = 0$ for every $\varepsilon > 0$.

Prove:

(a) $f_j \to f$ μ-almost uniformly $\Rightarrow f_j \to f$ in measure.

(b) If (f_j) converges in measure to f and to g, then $f = g$ μ-a.e.

(c) There is a sequence of Lebesgue measurable functions on $[0,1]$ that converges in measure, but does not converge pointwise anywhere.

(d) There is a sequence of Lebesgue measurable functions on $\mathbb{R}$ that converges pointwise but not in measure.

(Hints: (c) Set $f_j := \chi_{I_j}$, where the intervals $I_j \subset [0,1]$ are chosen so that $\lambda_1(I_j) \to 0$ and so that the sequence $\big(f_j(x)\big)$ has two cluster points for every $x \in [0,1]$. (d) Consider $f_j := \chi_{[j,j+1]}$.)

15 Suppose (f_j) is a sequence in $\mathcal{L}_0(X,\mu,E)$ converging in measure to $f \in \mathcal{L}_0(X,\mu,E)$. Show that (f_j) has a subsequence that converges μ-a.e. to f.
(Hint: There is an increasing sequence $(j_k)_{k\in\mathbb{N}}$ such that

$$\mu\big([\,|f_m - f_n| \geq 2^{-k}]\big) \leq 2^{-k} \quad \text{for } m,n \geq j_k \ .$$

With the help of $B_\ell := \bigcup_{k=\ell}^{\infty}[\,\big|f_{n_{k+1}} - f_{n_k}\big| \geq 2^{-k}]$, conclude that $(f_{j_k})_{k\in\mathbb{N}}$ converges μ-almost uniformly. Also note Exercises 12, 14(a) and 14(b).)

16 For $x = (x_j) \in \mathbb{K}^{\mathbb{N}}$ and $p \in [1,\infty]$, define[11]

$$\|x\|_p := \begin{cases} \big(\sum_{j=0}^{\infty}|x_j|^p\big)^{1/p} & \text{if } p \in [1,\infty) \ , \\ \sup_j |x_j| & \text{if } p = \infty \ , \end{cases}$$

and

$$\ell_p := \ell_p(\mathbb{K}) := \big(\{\, x \in \mathbb{K}^{\mathbb{N}} \ ; \ \|x\|_p < \infty \,\},\ \|\cdot\|_p\big) \ .$$

Prove:

(a) If $p \in [1,\infty)$, then ℓ_p is a separable normed vector space.

(b) ℓ_∞ is not separable.

17 Suppose $x \in \overline{\mathbb{R}}$. If $x \in \mathbb{R}$, we say $U \subset \overline{\mathbb{R}}$ is a **neighborhood in** $\overline{\mathbb{R}}$ of x if U contains a neighborhood *in* $\mathbb{R}$ of x. For $x \in \overline{\mathbb{R}}\backslash\mathbb{R}$, neighborhoods were defined in Section II.5 as those sets that contain a semi-infinite interval of the appropriate kind. Suppose $O \subset \overline{\mathbb{R}}$. We say O is **open in** $\overline{\mathbb{R}}$ if every $x \in O$ has a neighborhood U in $\overline{\mathbb{R}}$ such that $U \subset O$. Now define $\overline{\mathcal{T}} := \{\, O \subset \overline{\mathbb{R}} \ ; \ O \text{ is open in } \overline{\mathbb{R}} \,\}$. Prove:

(a) O is open in $\overline{\mathbb{R}}$ if and only if $O \cap \mathbb{R}$ is open in $\mathbb{R}$ and, in the case $\infty \in O$ [or $-\infty \in O$], there is an $a \in \mathbb{R}$ such that $(a,\infty] \subset O$ [or $[-\infty,a) \subset O$].

(b) $(\overline{\mathbb{R}},\overline{\mathcal{T}})$ is a compact topological space.

(c) $\mathcal{B}(\overline{\mathbb{R}}) = \big\{\, B \cup F \ ; \ B \in \mathcal{B}^1,\ F \subset \{-\infty,\infty\} \,\big\}$.

(d) $\mathcal{B}(\overline{\mathbb{R}}) \,|\, \mathbb{R} = \mathcal{B}^1$.

(e) An element of $\overline{\mathbb{R}}^X$ belongs to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$ if and only if it is $\mathcal{A}$-$\mathcal{B}(\overline{\mathbb{R}})$-measurable.

18 Check that, if S is a separable subset of E, the closure of its span is a separable Banach space.

19 For $f \in \mathbb{K}^X$, set

$$(\operatorname{sign} f)(x) := \begin{cases} f(x)/|f(x)| & \text{if } f(x) \neq 0 \ , \\ 0 & \text{if } f(x) = 0 \ . \end{cases}$$

Demonstrate that $f \in \mathcal{L}_0(X,\mu,\mathbb{K})$ implies $\operatorname{sign} f \in \mathcal{L}_0(X,\mu,\mathbb{K})$.

[11] See also Proposition IV.2.17.

2 Integrable functions

In this section, we define the general Bochner–Lebesgue integral and describe its basic properties. We also prove that the vector space of integrable functions is complete with respect to the seminorm induced by the integral.

As in the previous section, suppose

- $(X, \mathcal{A}, \mu)$ is a complete σ-finite measure space;
  $E = (E, |\cdot|)$ is a Banach space.

The integral of a simple function

In Remark 1.2(c), we learned that every simple function has a unique normal form. This form will prove to be useful in the sequel, and we work with it preferentially.

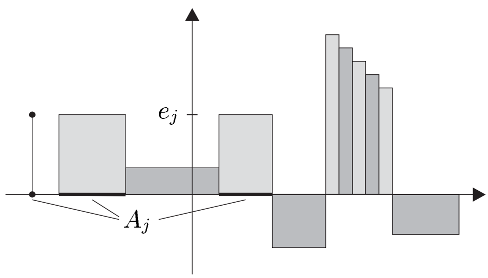
> Convention We will always represent μ-simple functions by their normal forms, unless we say otherwise. Further, we set[1]
>
> $$\infty \cdot 0_E := -\infty \cdot 0_E := 0_E \ , \tag{2.1}$$
>
> where 0_E is the zero vector of E.

For $\varphi \in \sum_{j=0}^m e_j \chi_{A_j} \in \mathcal{S}(X, \mu, E)$, we define **the integral of φ over X with respect to the measure μ** as the sum

$$\int_X \varphi \, d\mu := \int \varphi \, d\mu := \sum_{j=0}^m e_j \mu(A_j) \ .$$

If A is a μ-measurable set, we define the **integral of φ over A with respect to the measure μ** as

$$\int_A \varphi \, d\mu := \int_X \chi_A \varphi \, d\mu \ .$$



[1] Convention (2.1) is common in the theory of integration and makes it possible, for example, to integrate simple functions over their entire domains of definition. It is *not* to be understood in the case $E = \mathbb{R}$ as another calculation rule in $\overline{\mathbb{R}}$, but rather in the sense of "external" multiplication of $\infty, -\infty \in \overline{\mathbb{R}}$ by the zero vector of $\mathbb{R}$.

2.1 Remarks **(a)** For $\varphi \in \mathcal{S}(X,\mu,E)$ and $A \in \mathcal{A}$, the integral $\int_A \varphi\, d\mu$ is well defined.

Proof This follows from Remarks 1.2(c) and (d). ■

(b) Let $\varphi = \sum_{k=0}^n f_k \chi_{B_k}$, where $f_0,\ldots,f_n \in E$ are nonzero and $B_0,\ldots,B_n \in \mathcal{A}$ are pairwise disjoint, be a μ-simple function *not necessarily in normal form.* Then

$$\int_X \varphi\, d\mu = \sum_{k=0}^n f_k \mu(B_k) \ .$$

Proof Let $\sum_{j=0}^m e_j \chi_{A_j}$ be the normal form of φ. Set

$$A_{m+1} := \bigcap_{j=0}^m A_j^c \ , \quad B_{n+1} := \bigcap_{k=0}^n B_k^c \ , \quad e_{m+1} := 0 \ , \quad f_{n+1} := 0 \ . \tag{2.2}$$

Then $X = \bigcup_{j=0}^{m+1} A_j = \bigcup_{k=0}^{n+1} B_k$, so

$$A_j = \bigcup_{k=0}^{n+1} (A_j \cap B_k) \quad \text{and} \quad B_k = \bigcup_{j=0}^{m+1} (A_j \cap B_k) \quad \text{for } j=0,\ldots,m+1 \text{ and } k=0,\ldots,n+1 \ .$$

Because the sets $A_j \cap B_k$ are pairwise disjoint, we have

$$\mu(A_j) = \sum_{k=0}^{n+1} \mu(A_j \cap B_k) \quad \text{and} \quad \mu(B_k) = \sum_{j=0}^{m+1} \mu(A_j \cap B_k) \ .$$

If $A_j \cap B_k \neq \emptyset$, then $e_j = f_k$, and we find

$$\begin{aligned}\int_X \varphi\, d\mu &= \sum_{j=0}^m e_j \mu(A_j) = \sum_{j=0}^{m+1} e_j \sum_{k=0}^{n+1} \mu(A_j \cap B_k) = \sum_{k=0}^{n+1} f_k \sum_{j=0}^{m+1} \mu(A_j \cap B_k) \\ &= \sum_{k=0}^n f_k \mu(B_k) \ . \ ■\end{aligned}$$

(c) The integral $\int_X \cdot\, d\mu : \mathcal{S}(X,\mu,E) \to E$ is linear.

Proof Suppose $\varphi = \sum_{j=0}^m e_j \chi_{A_j}$ and $\psi = \sum_{k=0}^n f_k \chi_{B_k}$ are μ-simple functions and $\alpha \in \mathbb{K}$. One checks easily that $\int_X \alpha\varphi\, d\mu = \alpha \int_X \varphi\, d\mu$. With the relations (2.2), an argument like that in (b) implies

$$\chi_{A_j} = \sum_{k=0}^{n+1} \chi_{A_j \cap B_k} \quad \text{and} \quad \chi_{B_k} = \sum_{j=0}^{m+1} \chi_{A_j \cap B_k} \ ,$$

and thus[2]

$$\varphi + \psi = \sum_{j=0}^{m+1} \sum_{k=0}^{n+1} (e_j + f_k) \chi_{A_j \cap B_k} \ . \tag{2.3}$$

The claim now follows from (b). ■

[2]In general, (2.3) does *not* give the normal form of $\varphi + \psi$.

(d) For $A, B \in \mathcal{A}$ and $A \cap B = \emptyset$, we have

$$\int_{A \cup B} \varphi \, d\mu = \int_A \varphi \, d\mu + \int_B \varphi \, d\mu \quad \text{for } \varphi \in \mathcal{S}(X, \mu, E) \ .$$

Proof This follows from (c) and the equality $\chi_{A \cup B} \varphi = \chi_A \varphi + \chi_B \varphi$. ■

(e) For $\varphi \in \mathcal{S}(X, \mu, E)$ and $A \in \mathcal{A}$, we have

$$\Bigl| \int_A \varphi \, d\mu \Bigr| \leq \int_A |\varphi| \, d\mu \leq \|\varphi\|_\infty \, \mu(A) \ .$$

Proof This follows from Remark 1.2(d) and the triangle inequality. ■

(f) If $\varphi, \psi \in \mathcal{S}(X, \mu, \mathbb{R})$ satisfy $\varphi \leq \psi$, then $\int_A \varphi \, d\mu \leq \int_A \psi \, d\mu$.

Proof Clearly $\int_A \eta \, d\mu \geq 0$ for $\eta \in \mathcal{S}(X, \mu, \mathbb{R}^+)$. The claim now follows from (c). ■

The $\mathcal{L}_1$-seminorm

Suppose V is a vector space over $\mathbb{K}$. A map $p \colon V \to \mathbb{R}$ is called a **seminorm** on V if it satisfies these properties:

(i) $p(v) \geq 0$ for $v \in V$;

(ii) $p(\lambda v) = |\lambda| \, p(v)$ for $v \in V$ and $\lambda \in \mathbb{K}$;

(iii) $p(v + w) \leq p(v) + p(w)$ for $v, w \in V$.

For $v \in V$ and $r > 0$, we denote by

$$\mathbb{B}_p(v, r) := \{ \, w \in V \ ; \ p(v - w) < r \, \}$$

the **open ball** in (V, p) **around v with radius r**. A subset O of V is said to be **p-open** if, for every $v \in O$, there is an $r > 0$ such that $\mathbb{B}_p(v, r) \subset O$.

2.2 Remarks Suppose V is a vector space and p is a seminorm on V.

(a) The seminorm p is a norm if and only if $p^{-1}(0) = \{0\}$.

(b) Suppose $K \subset \mathbb{R}^n$ is compact, $k \in \mathbb{N} \cup \{\infty\}$, and

$$p_K(f) := \max_{x \in K} |f(x)| \quad \text{for } f \in C^k(\mathbb{R}^n, E) \ .$$

Then p_K is a seminorm on $C^k(\mathbb{R}^n, E)$, but not a norm.

Proof One verifies easily that p_K is a seminorm on $C^k(\mathbb{R}^n, E)$. Let U be an open neighborhood of K. Then Exercise VII.6.7 shows that there is an $f \in C^\infty(\mathbb{R}^n, \mathbb{R})$ such that $f(x) = 1$ for $x \in K$ and $f(x) = 0$ for $x \in U^c$. For $e \in E \backslash \{0\}$, we set $g := (\chi_{\mathbb{R}^n} - f)e$. Then g belongs to $C^\infty(\mathbb{R}^n, E)$, and we have $p_K(g) = 0$, but $g \neq 0$. Therefore p_K is not a norm on $C^k(\mathbb{R}^n, E)$. ■

(c) Let

$$\|\varphi\|_1 := \int_X |\varphi| \, d\mu \quad \text{for } \varphi \in \mathcal{S}(X,\mu,E) \ .$$

Then $\|\cdot\|_1$ is a seminorm on $\mathcal{S}(X,\mu,E)$. If there is a nonempty μ-null set in $\mathcal{A}$, then $\|\cdot\|_1$ is not a norm on $\mathcal{S}(X,\mu,E)$.

Proof It is clear that $\|\cdot\|_1$ is a seminorm on $\mathcal{S}(X,\mu,E)$. Letting N denote a nonempty μ-null set, we have $\|\chi_N\|_1 = 0$, but $\chi_N \neq 0$. ■

(d) $\mathcal{T}_p := \{\, O \subset V \ ; \ O \text{ is } p\text{-open} \,\}$ is a topology on V, the **topology generated by *p***.

Proof One easily checks that the argument used in the proof of Proposition III.2.4 transfers to this situation. ■

(e) The topology $\mathcal{T}_p$ is not necessarily Hausdorff. If it isn't, there is no metric on V that generates $\mathcal{T}_p$.

Proof We use the notation of (b) and set $K := \{0\}$. Further let $f \in C^k(\mathbb{R}^n, E)$ with $f(0) = 0$ and $f \neq 0$. Then $\mathbb{B}_{p_K}(f,\varepsilon) = \mathbb{B}_{p_K}(0,\varepsilon)$ for every $\varepsilon > 0$. Therefore $\mathcal{T}_{p_K}$ is not Hausdorff. The second statement follows from Proposition III.2.17. ■

(f) A linear map $A\colon V \to E$ is said to be **(*p*)-bounded** if there is an $M \geq 0$ such that $|Av| \leq Mp(v)$ for $v \in V$. For a linear map $A\colon V \to E$, these statements are equivalent:

(i) A is continuous.

(ii) A is continuous at 0.

(iii) A is bounded.

Proof This follows from the proof of Theorem VI.2.5, which used only the properties of a seminorm. ■

(g) $\int \cdot \, d\mu \colon \mathcal{S}(X,\mu,E) \to E$ is continuous.

Proof This follows from (c), (f), and Remark 2.1(c). ■

Suppose p is a seminorm on V. We know from Remark 2.2(e) that the topology of (V,p) may not be generated by a metric on V, in which case the metric notions of Cauchy sequence and completeness are not available. Accordingly, we need new definitions: A sequence $(v_j) \in V^{\mathbb{N}}$ is called a **Cauchy sequence in** (V,p) if for every $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $p(v_j - v_k) < \varepsilon$ for $j,k \geq N$. We call (V,p) **complete** if every Cauchy sequence in (V,p) converges.

2.3 Remarks **(a)** If (V,p) is a normed vector space, these notions agree with those of Section II.6.

(b) Suppose $(v_j) \in V^{\mathbb{N}}$ and $v \in V$. Then $v_j \to v$ if and only if $p(v - v_j) \to 0$. However, the limit of a convergent sequence is generally not uniquely determined: if p is not a norm, $v_j \to v$ implies $v_j \to w$ for every $w \in V$ such that $p(v - w) = 0$.

(c) The set of all Cauchy sequences in (V, p) forms a vector subspace of $V^{\mathbb{N}}$. ■

In the following, we always provide the space $\mathcal{S}(X, \mu, E)$ with the topology induced by $\|\cdot\|_1$. Then we may also call a Cauchy sequence in $\mathcal{S}(X, \mu, E)$ an $\boldsymbol{\mathcal{L}_1}$**-Cauchy sequence**.

A function $f \in E^X$ is called $\boldsymbol{\mu}$**-integrable** if f is a μ-a.e. limit of some $\mathcal{L}_1$-Cauchy sequence (φ_j) in $\mathcal{S}(X, \mu, E)$. We denote the set of E-valued, μ-integrable functions of X by $\mathcal{L}_1(X, \mu, E)$.

2.4 Proposition *In the sense of vector subspaces, we have the inclusions*

$$\mathcal{S}(X, \mu, E) \subset \mathcal{L}_1(X, \mu, E) \subset \mathcal{L}_0(X, \mu, E) \ .$$

Proof Clearly every μ-simple function is μ-integrable. That every μ-integrable function is μ-measurable follows from Remark 1.2(a) and Theorem 1.14. There remains to show that $\mathcal{L}_1(X, \mu, E)$ is a vector subspace of $\mathcal{L}_0(X, \mu, E)$. Take $f, g \in \mathcal{L}_1(X, \mu, E)$ and $\alpha \in \mathbb{K}$. There are $\mathcal{L}_1$-Cauchy sequences (φ_j) and (ψ_j) in $\mathcal{S}(X, \mu, E)$ such that $\varphi_j \to f$ and $\psi_j \to g$ μ-a.e. as $j \to \infty$. From the triangle inequality, it follows that $(\alpha\varphi_j + \psi_j)_{j\in\mathbb{N}}$ is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$ that converges μ-a.e. to $\alpha f + g$. Therefore $\alpha f + g$ is μ-integrable, as needed. ■

The Bochner–Lebesgue integral

Let $f \in \mathcal{L}_1(X, \mu, E)$. Then there is an $\mathcal{L}_1$-Cauchy sequence (φ_j) in $\mathcal{S}(X, \mu, E)$ such that $\varphi_j \to f$ μ-a.e. We will see that the sequence $(\int_X \varphi_j \, d\mu)_{j\in\mathbb{N}}$ converges in E. It is natural, then, to define the integral of f with respect to μ as the limit of this sequence of integrals. Of course we must check that the limit is independent of the approximating sequence of simple functions; that is, we must show that $\lim_j \int \varphi_j \, d\mu = \lim_j \int \psi_j \, d\mu$ if (ψ_j) is another Cauchy sequence in $\mathcal{S}(X, \mu, E)$ such that $\psi_j \to f$ μ-a.e.

2.5 Lemma *Suppose (φ_j) is a Cauchy sequence in $\mathcal{S}(X, \mu, E)$. Then there is a subsequence $(\varphi_{j_k})_{k\in\mathbb{N}}$ of (φ_j) and an $f \in \mathcal{L}_1(X, \mu, E)$ such that*

(i) *$\varphi_{j_k} \to f$ μ-a.e. as $k \to \infty$;*

(ii) *for every $\varepsilon > 0$, there exists $A_\varepsilon \in \mathcal{A}$ such that $\mu(A_\varepsilon) < \varepsilon$ and $(\varphi_{j_k})_{k\in\mathbb{N}}$ converges uniformly to f on A_ε^c.*

Proof (α) By assumption, there exists for any $k \in \mathbb{N}$ some $j_k \in \mathbb{N}$ such that the bound $\|\varphi_\ell - \varphi_m\|_1 < 2^{-2k}$ holds for $\ell, m \geq j_k$. Without loss of generality, we can choose the sequence $(j_k)_{k\in\mathbb{N}}$ to be increasing. With $\psi_k := \varphi_{j_k}$, we then get

$$\|\psi_\ell - \psi_m\|_1 < 2^{-2\ell} \quad \text{for } m \geq \ell \geq 0 \ .$$

(β) Set $B_\ell := \big[\,|\psi_{\ell+1} - \psi_\ell| \ge 2^{-\ell}\big]$ for $\ell \in \mathbb{N}$. Then B_ℓ belongs to $\mathcal{A}$, and we have $\mu(B_\ell) < \infty$ for $\ell \in \mathbb{N}$, because every ψ_m is μ-simple. Therefore χ_{B_ℓ} is also μ-simple, and Remark 2.1(f) implies

$$2^{-\ell}\mu(B_\ell) = 2^{-\ell}\int_X \chi_{B_\ell}\,d\mu \le \int_X |\psi_{\ell+1} - \psi_\ell|\,d\mu = \|\psi_{\ell+1} - \psi_\ell\|_1 < 2^{-2\ell}\ .$$

This leads to $\mu(B_\ell) < 2^{-\ell}$ for $\ell \in \mathbb{N}$.

Letting $A_n := \bigcup_{k=0}^\infty B_{n+k}$, we have $\mu(A_n) \le 2^{-n+1}$ for $n \in \mathbb{N}$, and we see that $A := \bigcap_{n=0}^\infty A_n$ is a μ-null set.

(γ) If x lies in $A_n^c = \bigcap_{k=0}^\infty B_{n+k}^c$, then

$$|\psi_{\ell+1}(x) - \psi_\ell(x)| < 2^{-\ell} \quad \text{for } \ell \ge n\ .$$

By the Weierstrass criterion (Theorem V.1.6), the series

$$\psi_0 + \sum(\psi_{\ell+1} - \psi_\ell)$$

on A_n^c converges uniformly in E. Now we set

$$f(x) := \begin{cases} \lim_k \psi_k(x) & \text{if } x \in A^c\ , \\ 0 & \text{if } x \in A\ . \end{cases}$$

Then $\varphi_{j_k} \to f$ μ-a.e. as $k \to \infty$. Further, there is for every $\varepsilon > 0$ an $n \in \mathbb{N}$ such that $\mu(A_n) \le 2^{-n+1} < \varepsilon$, and $(\varphi_{j_k})_{k\in\mathbb{N}}$ converges uniformly on A_n^c to f as $k \to \infty$. ■

2.6 Lemma *Suppose (φ_j) and (ψ_j) are $\mathcal{L}_1$-Cauchy sequences in $\mathcal{S}(X,\mu,E)$ that converge μ-a.e. to the same function. Then $\lim\|\varphi_j - \psi_j\|_1 = 0$.*

Proof (i) Take $\varepsilon > 0$ and set $\eta_j := \varphi_j - \psi_j$ for $j \in \mathbb{N}$. By Remark 2.3(c), (η_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X,\mu,E)$. Thus there is a natural number N such that $\|\eta_j - \eta_k\| < \varepsilon/8$ for $j,k \ge N$.

(ii) Because η_N is μ-simple, $A := [\eta_N \ne 0]$ belongs to $\mathcal{A}$ and $\mu(A) < \infty$. Moreover (η_j) converges μ-a.e. to zero. Then Lemma 2.5 says there exists $B \in \mathcal{A}$ such that $\mu(B) < \varepsilon/8(1 + \|\eta_N\|_\infty)$, and there is a subsequence $(\eta_{j_k})_{k\in\mathbb{N}}$ of (η_j) that converges uniformly to 0 on B^c. Hence there exists $K \ge N$ such that

$$|\eta_{j_K}(x)| \le \varepsilon/8\big(1 + \mu(A)\big) \quad \text{for } x \in A\backslash B\ .$$

This implies $\int_{A\backslash B} |\eta_{j_K}|\,d\mu \le \varepsilon/8$.

(iii) The properties of B and K imply

$$\begin{aligned}\int_B |\eta_{j_K}|\,d\mu &\le \int_B |\eta_{j_K} - \eta_N|\,d\mu + \int_B |\eta_N|\,d\mu \\ &\le \|\eta_{j_K} - \eta_N\|_1 + \|\eta_N\|_\infty\,\mu(B) < \varepsilon/4\ .\end{aligned}$$

From the definition of A, we have

$$\int_{A^c} |\eta_{j_K}|\, d\mu = \int_{A^c} |\eta_{j_K} - \eta_N|\, d\mu \le \|\eta_{j_K} - \eta_N\|_1 < \varepsilon/8 \ .$$

Altogether we obtain using Remark 2.1(d)

$$\|\eta_{j_K}\|_1 \le \int_{A^c \cup (A\setminus B)\cup B} |\eta_{j_K}|\, d\mu < \varepsilon/2 \ ,$$

and therefore $\|\eta_j\|_1 \le \|\eta_{j_K}\|_1 + \|\eta_j - \eta_{j_K}\|_1 < \varepsilon$ for $j \ge N$. Because $\varepsilon > 0$ was arbitrary, all is proved. ■

2.7 Corollary *Let (φ_j) and (ψ_j) be Cauchy sequences in $\mathcal{S}(X,\mu,E)$ converging μ-a.e. to the same function. The sequences $\left(\int_X \varphi_j\, d\mu\right)$ and $\left(\int_X \psi_j\, d\mu\right)$ converge in E, and*

$$\lim_j \int_X \varphi_j\, d\mu = \lim_j \int_X \psi_j\, d\mu \ .$$

Proof Because

$$\Bigl|\int_X \varphi_j\, d\mu - \int_X \varphi_k\, d\mu\Bigr| \le \|\varphi_j - \varphi_k\|_1 \quad \text{for } j,k \in \mathbb{N} \ ,$$

$(\int \varphi_j\, d\mu)_{j\in\mathbb{N}}$ is a Cauchy sequence in E; hence $\int \varphi_j\, d\mu \to e$ as $j \to \infty$, for some $e \in E$. Likewise there is $e' \in E$ such that $\int \psi_j\, d\mu \to e'$ for $j \to \infty$. Applying Lemma 2.6 and the continuity of the norm of E, we see that

$$|e - e'| = \lim_j \Bigl|\int_X \varphi_j\, d\mu - \int_X \psi_j\, d\mu\Bigr| \le \lim_j \int_X |\varphi_j - \psi_j|\, d\mu = \lim_j \|\varphi_j - \psi_j\|_1 = 0 \ ,$$

and we are done. ■

After these preparations, we define the integral of integrable functions in a natural way, extending the integral of simple functions. Suppose $f \in \mathcal{L}_1(X,\mu,E)$. Then there is an $\mathcal{L}_1$-Cauchy sequence (φ_j) in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j \to f$ μ-a.e. According to Corollary 2.7, the quantity

$$\int_X f\, d\mu := \lim_j \int_X \varphi_j\, d\mu$$

exists in E, and is independent of the sequence (φ_j). This is called the **Bochner–Lebesgue integral of $\boldsymbol{f}$ over $\boldsymbol{X}$ with respect to the measure $\boldsymbol{\mu}$**. Other notations are often used besides $\int_X f\, d\mu$, namely,

$$\int f\, d\mu \ , \quad \int_X f(x)\, d\mu(x) \ , \quad \int_X f(x)\, \mu(dx) \ .$$

Clearly, in the case of simple functions, the Bochner–Lebesgue integral agrees with the integral defined at the start of this section (page 80).

The completeness of $\mathcal{L}_1$

With the help of the integral, we now define a seminorm on $\mathcal{L}_1(X,\mu,E)$ and show that $\mathcal{L}_1(X,\mu,E)$ is complete with respect to this seminorm.

2.8 Lemma *If $f \in \mathcal{L}_1(X,\mu,E)$, then $|f|$ belongs to $\mathcal{L}_1(X,\mu,\mathbb{R})$. If (φ_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j \to f$ μ-a.e., then $\int |f|\,d\mu = \lim_j \int |\varphi_j|\,d\mu$.*

Proof The reverse triangle inequality (which clearly holds for seminorms too) gives

$$\big\| |\varphi_j| - |\varphi_k| \big\|_1 \le \|\varphi_j - \varphi_k\|_1 \quad \text{and} \quad \big| |\varphi_j| - |\varphi_k| \big| \le |\varphi_j - \varphi_k| \quad \text{for } j,k \in \mathbb{N}\ .$$

Thus $(|\varphi_j|)_{j\in\mathbb{N}}$ is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X,\mu,\mathbb{R})$ that converges μ-a.e. to $|f|$. Therefore $|f|$ belongs to $\mathcal{L}_1(X,\mu,\mathbb{R})$, and $\int |f|\,d\mu = \lim_j \int |\varphi_j|\,d\mu$. ■

2.9 Corollary *For $f \in \mathcal{L}_1(X,\mu,E)$, let $\|f\|_1 := \int_X |f|\,d\mu$. Then $\|\cdot\|_1$ is a seminorm on $\mathcal{L}_1(X,\mu,E)$, called the **$\mathcal{L}_1$-seminorm**.*

Proof Take $f,g \in \mathcal{L}_1(X,\mu,E)$ and let (φ_j) and (ψ_j) be $\mathcal{L}_1$-Cauchy sequences in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j \to f$ and $\psi_j \to g$ μ-a.e. From Lemma 2.8 and Remark 2.2(c) we have

$$\|f\|_1 = \int_X |f|\,d\mu = \lim_j \int_X |\varphi_j|\,d\mu = \lim_j \|\varphi_j\|_1 \ge 0\ ,$$

as well as

$$\|f+g\|_1 = \lim_j \|\varphi_j + \psi_j\|_1 \le \lim_j \big(\|\varphi_j\|_1 + \|\psi_j\|_1\big) = \|f\|_1 + \|g\|_1\ ,$$

and

$$\|\alpha f\|_1 = \lim_j \|\alpha\varphi_j\| = |\alpha| \lim_j \|\varphi_j\| = |\alpha|\, \|f\|_1$$

for any $\alpha \in \mathbb{K}$. ■

We will always give $\mathcal{L}_1(X,\mu,E)$ the topology induced by the seminorm $\|\cdot\|_1$.

2.10 Theorem

(i) *$\mathcal{S}(X,\mu,E)$ is dense in $\mathcal{L}_1(X,\mu,E)$.*

(ii) *The space $\mathcal{L}_1(X,\mu,E)$ is complete.*

Proof (i) Suppose $f \in \mathcal{L}_1(X,\mu,E)$, and let (φ_j) denote an $\mathcal{L}_1$-Cauchy sequence of simple functions such that $\varphi_j \to f$ μ-a.e. as $j \to \infty$. Also suppose $k \in \mathbb{N}$. Then $(\varphi_j - \varphi_k)_{j\in\mathbb{N}}$ is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X,\mu,E)$ such that $(\varphi_j - \varphi_k) \to (f - \varphi_k)$

μ-a.e. for $j \to \infty$. Because of Lemma 2.8,

$$\|f - \varphi_k\|_1 = \lim_j \|\varphi_j - \varphi_k\|_1 \quad \text{for } k \in \mathbb{N} .$$

Suppose $\varepsilon > 0$. Then there is an $N \in \mathbb{N}$ such that $\|\varphi_j - \varphi_k\|_1 < \varepsilon$ for $j, k \geq N$; taking the limit $j \to \infty$ we get $\|f - \varphi_N\|_1 \leq \varepsilon$. This shows that $\mathcal{S}(X, \mu, E)$ is dense in $\mathcal{L}_1(X, \mu, E)$.

(ii) Let (f_j) be a Cauchy sequence in $\mathcal{L}_1(X, \mu, E)$ and take $\varepsilon > 0$. Choose $M \in \mathbb{N}$ such that $\|f_j - f_k\|_1 < \varepsilon/2$ for $j, k \geq M$. We know from (i) that for any $j \in \mathbb{N}$ we can find $\varphi_j \in \mathcal{S}(X, \mu, E)$ such that $\|f_j - \varphi_j\|_1 < 2^{-j}$. Then

$$\|\varphi_j - \varphi_k\|_1 \leq \|\varphi_j - f_j\|_1 + \|f_j - f_k\|_1 + \|f_k - \varphi_k\|_1 < 2^{-j} + 2^{-k} + \varepsilon/2$$

for $j, k \geq M$. This shows that (φ_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$. By Lemma 2.5, therefore, there is a subsequence $(\varphi_{j_k})_{k \in \mathbb{N}}$ of (φ_j) and an f in $\mathcal{L}_1(X, \mu, E)$ such that $\varphi_{j_k} \to f$ μ-a.e. as $k \to \infty$. The proof of (i) shows that there exists an $N \geq M$ such that $\|f - \varphi_{j_N}\|_1 < \varepsilon/4$, and we get

$$\|f - f_j\|_1 \leq \|f - \varphi_{j_N}\|_1 + \|\varphi_{j_N} - f_{j_N}\|_1 + \|f_{j_N} - f_j\|_1 < \varepsilon \quad \text{for } j \geq N ,$$

that is, (f_j) converges in $\mathcal{L}_1(X, \mu, E)$ to f. ■

Elementary properties of integrals

We have seen that the integral on the space of simple functions is continuous, linear, and, for $E = \mathbb{R}$, also monotone — see Remarks 2.2(g), 2.1(c) and 2.1(f). We now show that these properties survive the extension of the integral from the space of simple functions to that of integrable functions.

2.11 Theorem

(i) $\int_X \cdot \, d\mu \colon \mathcal{L}_1(X, \mu, E) \to E$ *is linear and continuous, and*

$$\Bigl|\int_X f \, d\mu\Bigr| \leq \int_X |f| \, d\mu = \|f\|_1 .$$

(ii) $\int_X \cdot \, d\mu \colon \mathcal{L}_1(X, \mu, \mathbb{R}) \to \mathbb{R}$ *is a continuous, positive linear functional.*

(iii) *Suppose F is a Banach space and $T \in \mathcal{L}(E, F)$. Then*

$$Tf \in \mathcal{L}_1(X, \mu, F) \quad \text{and} \quad T \int_X f \, d\mu = \int_X Tf \, d\mu$$

for $f \in \mathcal{L}_1(X, \mu, E)$.

Proof (i) Proposition 2.4 showed that μ-integrable functions form a vector space. Take $f, g \in \mathcal{L}_1(X,\mu,E)$ and $\alpha \in \mathbb{K}$. Then there are $\mathcal{L}_1$-Cauchy sequences (φ_j) and (ψ_j) in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j \to f$ and $\psi_j \to g$ μ-a.e. By Remark 2.1(c),

$$\int_X (\alpha\varphi_j + \psi_j)\,d\mu = \alpha \int_X \varphi_j\,d\mu + \int_X \psi_j\,d\mu \quad \text{for } j \in \mathbb{N} .$$

The linearity of the integral on $\mathcal{L}_1(X,\mu,E)$ follows by taking the limit $j \to \infty$. By Corollary 2.9, $\|\cdot\|_1$ is a seminorm on $\mathcal{L}_1(X,\mu,E)$, and Remark 2.1(e) yields

$$\Bigl|\int_X \varphi_j\,d\mu\Bigr| \le \int_X |\varphi_j|\,d\mu = \|\varphi_j\|_1 \quad \text{for } j \in \mathbb{N} .$$

By Lemma 2.8, we can take the limit $j \to \infty$, and we find

$$\Bigl|\int_X f\,d\mu\Bigr| \le \int_X |f|\,d\mu = \|f\|_1 .$$

Continuity now follows from Remark 2.2(f).

The approach just used can be adapted without difficulty to the task of proving (ii) and (iii). This is left to the reader as an exercise. ■

2.12 Corollary

(i) *A map $f = (f_1,\ldots,f_n)\colon X \to \mathbb{K}^n$ is μ-integrable if and only if its every coordinate f_j is. In that case,*

$$\int_X f\,d\mu = \Bigl(\int_X f_1\,d\mu,\ldots,\int_X f_n\,d\mu\Bigr) .$$

(ii) *Suppose $g, h \in \mathbb{R}^X$ and define $f := g + i\,h$. Then f is in $\mathcal{L}_1(X,\mu,\mathbb{C})$ if and only if g and h are in $\mathcal{L}_1(X,\mu,\mathbb{R})$. In that case,*

$$\int_X f\,d\mu = \int_X g\,d\mu + i\int_X h\,d\mu .$$

(iii) *A function $f \in \mathbb{R}^X$ is μ-integrable if and only if f^+ and f^- are. In that case,*

$$\int_X f\,d\mu = \int_X f^+\,d\mu - \int_X f^-\,d\mu , \qquad \int_X |f|\,d\mu = \int_X f^+\,d\mu + \int_X f^-\,d\mu .$$

Proof (i) "$\Rightarrow$" Take $f \in \mathcal{L}_1(X,\mu,\mathbb{K}^n)$. Since $\mathrm{pr}_j \in \mathcal{L}(\mathbb{K}^n,\mathbb{K})$ for $j = 1,\ldots,n$, it follows from Theorem 2.11(iii) that $f_j = \mathrm{pr}_j \circ f$ belongs to $\mathcal{L}_1(X,\mu,\mathbb{K})$. Moreover $\int f_j\,d\mu = \mathrm{pr}_j \int f\,d\mu$, so

$$\int_X f\,d\mu = \Bigl(\int_X f_1\,d\mu,\ldots,\int_X f_n\,d\mu\Bigr) .$$

"$\Longleftarrow$" For $j = 1, \ldots, n$, we consider the map

$$b_j : \mathbb{K} \to \mathbb{K}^n \ , \quad y \mapsto (0, \ldots, 0, y, 0, \ldots, 0) \ ,$$

where y is in the j-th position on the right. Then

$$b_j \in \mathcal{L}(\mathbb{K}, \mathbb{K}^n) \quad \text{and} \quad f := \sum_{j=1}^{n} b_j \circ f_j \ .$$

The claim now follows from Theorem 2.11(i) and (iii).

(ii) This follows from (i) and the identification of $\mathbb{C}$ with $\mathbb{R}^2$.

(iii) For $f \in \mathbb{R}^X$, we have

$$f^+ = (f + |f|)/2 \ , \quad f^- = (|f| - f)/2 \ , \quad f = f^+ - f^- \ , \quad |f| = f^+ + f^- \ .$$

Hence Theorem 2.11(i) and Lemma 2.8 imply the conclusion. ■

2.13 Lemma *For $f \in \mathcal{L}_1(X, \mu, E)$ and $A \in \mathcal{A}$, we have $\chi_A f \in \mathcal{L}_1(X, \mu, E)$.*

Proof Suppose (φ_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$ that converges μ-a.e. to f. Then $\chi_A \varphi_j$ is μ-simple by Remark 1.2(d), and $(\chi_A \varphi_j)_{j \in \mathbb{N}}$ obviously converges μ-a.e. to $\chi_A f$. By Remark 2.1(f), we have

$$\int_X |\chi_A \varphi_j - \chi_A \varphi_k| \, d\mu = \int_X \chi_A \, |\varphi_j - \varphi_k| \, d\mu \le \int_X |\varphi_j - \varphi_k| \, d\mu \quad \text{for } j, k \in \mathbb{N} \ .$$

Therefore $(\chi_A \varphi_j)_{j \in \mathbb{N}}$ is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$. This shows that $\chi_A f$ is μ-integrable. ■

For $f \in \mathcal{L}_1(X, \mu, E)$ and $A \in \mathcal{A}$, we define the ***μ*-integral of *f* over *A*** by

$$\int_A f \, d\mu := \int_X \chi_A f \, d\mu \ .$$

This is well defined by Lemma 2.13.

2.14 Remarks Suppose $f \in \mathcal{L}_1(X, \mu, E)$ and $A \in \mathcal{A}$.

(a) $\int_A \cdot \, d\mu \colon \mathcal{L}_1(X, \mu, E) \to E$ is linear and continuous, and

$$\Bigl| \int_A f \, d\mu \Bigr| \le \int_A |f| \, d\mu = \|\chi_A f\|_1 \ .$$

(b) Suppose $\mathcal{B} := \mathcal{A} \,|\, A$ and $\nu := \mu \,|\, \mathcal{B}$. Then $\int_A f \, d\mu = \int_A f \,|\, A \, d\nu$.

Proof The proof is simple and is left to the reader (Exercise 1). ■

(c) If $E = \mathbb{R}$ and $f \geq 0$, the map

$$\mathcal{A} \to [0, \infty) \ , \quad A \mapsto \int_A f \, d\mu$$

is a finite measure (see Exercise 11). ■

2.15 Lemma *Suppose $f \in \mathcal{L}_1(X, \mu, E)$ and $g \in E^X$ satisfy $f = g$ μ-a.e. Then g also belongs to $\mathcal{L}_1(X, \mu, E)$, and $\int_X f \, d\mu = \int_X g \, d\mu$.*

Proof Suppose (φ_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$ such that $\varphi_j \to f$ μ-a.e. Also let M and N be μ-null sets such that $\varphi_j \to f$ on M^c and $f = g$ on N^c. Then (φ_j) converges μ-a.e. to g, because $\varphi_j(x) \to g(x)$ holds for $x \in (M \cup N)^c$. Therefore g belongs to $\mathcal{L}_1(X, \mu, E)$, and $\int g \, d\mu = \lim_j \int \varphi_j \, d\mu = \int f \, d\mu$. ■

2.16 Corollary

(i) *Suppose $f \in E^X$ vanishes μ-a.e. Then f is μ-integrable with $\int_X f \, d\mu = 0$.*

(ii) *Suppose $f, g \in \mathcal{L}_1(X, \mu, \mathbb{R})$ satisfy $f \leq g$ μ-a.e. Then $\int_X f \, d\mu \leq \int_X g \, d\mu$.*

Proof (i) This follows immediately from Lemma 2.15.

(ii) Theorem 2.11(ii) and Lemma 2.15 imply $0 \leq \int_X (g - f) \, d\mu$, and therefore $\int_X f \, d\mu \leq \int_X g \, d\mu$. ■

2.17 Proposition *For $f \in \mathcal{L}_1(X, \mu, E)$ and $\alpha > 0$, we have $\mu\big([\,|f| \geq \alpha]\big) < \infty$.*

Proof Lemma 2.5 shows there is an $\mathcal{L}_1$-Cauchy sequence (φ_j) in $\mathcal{S}(X, \mu, E)$ and a μ-measurable set A such that $\mu(A) \leq 1$ and (φ_j) converges uniformly on A^c to f. Because $|f|$ is μ-measurable, $B := A^c \cap [\,|f| \geq \alpha]$ belongs to $\mathcal{A}$. Also there is an $N \in \mathbb{N}$ such that $|\varphi_N(x) - f(x)| \leq \alpha/2$ for $x \in A^c$. Therefore

$$|\varphi_N(x)| \geq |f(x)| - |\varphi_N(x) - f(x)| \geq \alpha/2 \quad \text{for } x \in B \ .$$

In particular, B is contained in $[\varphi_N \neq 0]$. Thus $\mu(B) \leq \mu\big([\varphi_N \neq 0]\big) < \infty$, because φ_N is μ-simple. Since

$$[\,|f| \geq \alpha] = B \cup \big(A \cap [\,|f| \geq \alpha]\big) \subset B \cup A \ ,$$

it follows that $\mu\big([\,|f| \geq \alpha]\big) \leq \mu(B) + 1 < \infty$. ■

Convergence in $\mathcal{L}_1$

For every integrable function f, there is an $\mathcal{L}_1$-Cauchy sequence of simple functions converging to it almost everywhere. We show next that *every* Cauchy sequence in $\mathcal{L}_1(X, \mu, E)$ has in fact a subsequence that converges almost everywhere to the sequence's $\mathcal{L}_1$ limit.

2.18 Theorem *Let (f_j) be a sequence in $\mathcal{L}_1(X,\mu,E)$ converging to f in $\mathcal{L}_1(X,\mu,E)$.*

(i) *There is a subsequence $(f_{j_k})_{k\in\mathbb{N}}$ of (f_j) with the following properties:*

(α) *$f_{j_k} \to f$ μ-a.e. as $k\to\infty$.*

(β) *For every $\varepsilon > 0$, there is an $A_\varepsilon \in \mathcal{A}$ with $\mu(A_\varepsilon) < \varepsilon$ such that $(f_{j_k})_{k\in\mathbb{N}}$ converges uniformly on A_ε^c to f.*

(ii) *The integral $\int_X f_j \, d\mu$ converges to $\int_X f \, d\mu$ as $j\to\infty$.*

Proof (i) It suffices to treat the case $f = 0$ because, if $f \neq 0$, we may consider the sequence $(f_j - f)_{j\in\mathbb{N}}$.

As in the proof of Lemma 2.5, there is a subsequence (g_k) of (f_j) such that $\|g_\ell - g_m\|_1 < 2^{-2\ell}$ for $m \geq \ell \geq 0$. The limit $m\to\infty$ gives $\|g_\ell\|_1 \leq 2^{-2\ell}$ for $\ell\in\mathbb{N}$. We set $B_\ell := [\,|g_\ell| \geq 2^{-\ell}]$. By Lemma 2.8, Proposition 2.4, and Proposition 1.9, B_ℓ belongs to $\mathcal{A}$, and we find

$$2^{-\ell}\mu(B_\ell) \leq \int_{B_\ell} |g_\ell| \, d\mu \leq \int_X |g_\ell| \, d\mu = \|g_\ell\|_1 \leq 2^{-2\ell} \quad \text{for } \ell\in\mathbb{N}$$

(compare Theorem 2.11(ii)). Therefore $\mu(B_\ell) \leq 2^{-\ell}$ for $\ell\in\mathbb{N}$. With $A_n := \bigcup_{k=0}^\infty B_{n+k}$, we have $\mu(A_n) \leq 2^{-n+1}$, and we find that $A := \bigcap_{n=0}^\infty A_n$ is a μ-null set. We verify easily that (g_k) converges to 0 uniformly on A_n^c and pointwise on A^c (in this connection see the proof of Lemma 2.5).

(ii) From Theorem 2.11(i) it follows that

$$\Big|\int_X f_j \, d\mu - \int_X f \, d\mu\Big| \leq \int_X |f_j - f| \, d\mu = \|f_j - f\|_1 \quad \text{for } j\in\mathbb{N} \ ,$$

so the limit of the left-hand side as $j\to\infty$ is 0. ■

2.19 Corollary *For $f \in \mathcal{L}_1(X,\mu,E)$,*

$$\|f\|_1 = 0 \Longleftrightarrow f = 0 \ \ \mu\text{-a.e.}$$

Proof "$\Longrightarrow$" Because $\|f\|_1 = 0$, the sequence (f_j) with $f_j := 0$ for $j\in\mathbb{N}$ converges in $\mathcal{L}_1(X,\mu,E)$ to f. By Theorem 2.18 there is thus a subsequence $(f_{j_k})_{k\in\mathbb{N}}$ of (f_j) such that $f_{j_k} \to f$ μ-a.e. as $k\to\infty$. Therefore, $f = 0$ μ-a.e.

"$\Longleftarrow$" By assumption, $|f| = 0$ μ-a.e.; the claim follows from Corollary 2.16(i). ■

We conclude this section by illustrating its concepts and results in an especially simple situation.

2.20 Example (the space of summable series) Let X denote either $\mathbb{N}$ or $\mathbb{Z}$, and let $\mathcal{H}^0$ be its 0-dimensional Hausdorff measure, or counting measure. The topology induced by $\mathbb{R}$ clearly transforms X into a σ-compact metric space in which every one-point set is open. Hence the topology of X coincides with $\mathfrak{P}(X)$: every subset of X is open. Consequently, every map of X is continuous in E, that is, $C(X,E) = E^X$.

It follows further that $\mathcal{B}(X) = \mathcal{P}(X)$, and that $\mathcal{H}^0$ is a massive Radon measure on X. Thus, by Theorem 1.17,

$$\mathcal{L}_0(X, \mathcal{H}^0, E) = C(X,E) = E^X .$$

In addition, $\mathcal{H}^0$ has no nonempty null sets. Hence $\mathcal{H}^0$-a.e. convergence is the same as pointwise convergence.

For $\varphi \in E^X$, we define the **support** of φ as the set

$$\operatorname{supp}(\varphi) := \{\, x \in X \ ; \ \varphi(x) \neq 0 \,\} ,$$

and denote by $C_c(X,E)$ the space of continuous E-valued functions on X with compact support:

$$C_c(X,E) := \{\, \varphi \in C(X,E) \ ; \ \operatorname{supp}(\varphi) \text{ is compact} \,\} .$$

Clearly $\varphi \in C(X,E)$ belongs to $C_c(X,E)$ if and only if $\operatorname{supp}(\varphi)$ is a finite set. Also, $C_c(X,E)$ is a vector subspace of $C(X,E)$, and we verify easily that $C_c(X,E) = \mathcal{S}(X, \mathcal{H}^0, E)$.

For $\varphi \in C_c(X)$, it follows from Remark 2.1(b) that

$$\int_X \varphi \, d\mathcal{H}^0 = \sum_{x \in \operatorname{supp}(\varphi)} \varphi(x) . \tag{2.4}$$

We now set

$$\ell_1(X,E) := \{\, f \in E^X \ ; \ \textstyle\sum_{x\in X} |f(x)| < \infty \,\} .$$

For $f \in \ell_1(X,E)$ and $n \in \mathbb{N}$, let

$$\varphi_n(x) := \begin{cases} f(x) & \text{if } |x| \le n , \\ 0 & \text{if } |x| > n . \end{cases}$$

Then φ_n belongs to $C_c(X,E)$, and $\varphi_n \to f$ for $n \to \infty$. For $m > n$, we get from (2.4) that

$$\|\varphi_n - \varphi_m\|_1 = \sum_{n < |x| \le m} |f(x)| .$$

Therefore (φ_n) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mathcal{H}^0, E)$, which shows that f belongs to $\mathcal{L}_1(X, \mathcal{H}^0, E)$. Therefore $\ell_1(X,E) \subset \mathcal{L}_1(X, \mathcal{H}^0, E)$, and

$$\int_X f \, d\mathcal{H}^0 = \sum_{x \in X} f(x) \quad \text{for } f \in \ell_1(X,E) . \tag{2.5}$$

Now let $f \in \mathcal{L}_1(X, \mathcal{H}^0, E)$. Then there exists an $\mathcal{L}_1$-Cauchy sequence (ψ_j) in $\mathcal{S}(X, \mu, E)$ — and therefore in $C_c(X, E)$ — that converges pointwise to f. By Lemma 2.8, $|f|$ belongs to $\mathcal{L}_1(X, \mathcal{H}^0, \mathbb{R})$, and

$$\|f\|_1 = \int_X |f| \, d\mathcal{H}^0 = \lim_{j \to \infty} \int_X |\psi_j| \, d\mathcal{H}^0 = \lim_{j \to \infty} \sum_{x \in X} |\psi_j(x)| \ .$$

Hence there is a $k \in \mathbb{N}$ such that

$$\Bigl| \int_X |f| \, d\mathcal{H}^0 - \sum_{x \in X} |\psi_j(x)| \Bigr| \leq 1 \quad \text{for } j \geq k \ .$$

This implies

$$\sum_{x \in X} |\psi_j(x)| \leq 1 + \int_X |f| \, d\mathcal{H}^0 =: K < \infty \quad \text{for } j \geq k \ .$$

Therefore for every $m \in \mathbb{N}$ we have

$$\sum_{|x| \leq m} |\psi_j(x)| \leq K \quad \text{for } j \geq k \ ,$$

from which, because $\psi_j \to f$ as $j \to \infty$, we obtain

$$\sum_{|x| \leq m} |f(x)| \leq K \quad \text{for } m \in \mathbb{N} \ .$$

Now Theorem II.7.7 implies that f belongs to $\ell_1(X, E)$ (and satisfies $\|f\|_1 \leq K$). Therefore we have shown that

$$\mathcal{L}_1(X, \mathcal{H}^0, E) = \ell_1(X, E) \ ,$$

whereupon we obtain from (2.5) the relation[3]

$$\|f\|_1 = \sum_{x \in X} |f(x)| \ .$$

Finally, it follows from Theorem 2.10 and Remark 2.2(a) that

$$\ell_1(X, E) := \bigl(\ell_1(X, E), \| \cdot \|_1 \bigr)$$

is a Banach space, the **space of summable** (E-valued) **series**.

If $E = \mathbb{K}$, it is customary to write $\ell_1(\mathbb{Z})$ and $\ell_1(\mathbb{N})$ for $\ell_1(X, \mathbb{K})$, and the abbreviation $\ell_1 := \ell_1(\mathbb{N})$ is also common.[4] ■

[3] Note Theorem II.8.9.
[4] Compare Exercise II.8.6.

Exercises

1 Suppose $A \in \mathcal{A}$, $\mathcal{B} := \mathcal{A} \,|\, A$, and $\nu := \mu \,|\, \mathcal{B}$. Verify for $f \in E^X$ that

$$\chi_A f \in \mathcal{L}_1(X, \mu, E) \Longleftrightarrow f \,|\, A \in \mathcal{L}_1(A, \nu, E) \ .$$

For such an f, show that

$$\int_X \chi_A f \, d\mu = \int_A f \,|\, A \, d\nu \ .$$

2 Suppose (f_j) is a sequence in $\mathcal{L}_1(X, \mu, E)$ that converges uniformly to $f \in E^X$. Also suppose $\mu(X) < \infty$. Show that f belongs to $\mathcal{L}_1(X, \mu, E)$, that $f_j \to f$ in $\mathcal{L}_1(X, \mu, E)$, and that $\lim_j \int_X f_j \, d\mu = \int_X f \, d\mu$.

3 Verify that, for $f \in \mathcal{L}_1(X, \mu, \mathbb{R}^+)$,

$$\int_X f \, d\mu = \sup \Bigl\{ \int_X \varphi \, d\mu \ ; \ \varphi \in \mathcal{S}(X, \mu, \mathbb{R}^+), \ \varphi \le f \ \ \mu\text{-a.e.} \Bigr\} \ .$$

4 Suppose X is an arbitrary nonempty set, $a \in X$, and δ_a is the Dirac measure with support in a. Show that $\mathcal{L}_1(X, \delta_a, \mathbb{R}) = \mathbb{R}^X$, and calculate $\int f \, d\delta_a$ for $\mathbb{R}^X$.

5 Let μ_F be the Lebesgue–Stieltjes measure of Exercise IX.4.10. Determine $\mathcal{L}_1(\mathbb{R}, \mu_F, \mathbb{K})$ and calculate $\int f \, d\mu_F$ for $f \in \mathcal{L}_1(\mathbb{R}, \mu_F, \mathbb{K})$.

6 Prove statements (ii) and (iii) of Theorem 2.11.

7 Suppose that $f \in \mathcal{L}_0(X, \mu, E)$ is bounded μ-a.e. and that $\mu(X) < \infty$. Prove or disprove that f is μ-integrable.

8 Suppose (f_j) is an *increasing* sequence in $\mathcal{L}_1(X, \mu, \mathbb{R})$ such that $f_j \ge 0$, and suppose it converges μ-a.e. to $f \in \mathcal{L}_1(X, \mu, \mathbb{R})$. Then $\int_X f_j \, d\mu \uparrow \int_X f \, d\mu$. (This is known as the **monotone convergence theorem** in $\mathcal{L}_1$).
(Hint: Show that (f_j) is a Cauchy sequence in $\mathcal{L}_1(X, \mu, \mathbb{R})$, and identify its limit.)

9 Let (f_j) be a sequence in $\mathcal{L}_1(X, \mu, \mathbb{R})$ with $f_j \ge 0$ μ-a.e. and $\sum_{j=0}^\infty f_j \in \mathcal{L}_1(X, \mu, \mathbb{R})$. Show that $\sum_{j=0}^\infty \int f_j \, d\mu = \int \bigl(\sum_{j=0}^\infty f_j \bigr) \, d\mu$. (Hint: Exercise 8.)

10 Suppose that $f \in \mathcal{L}_1(X, \mu, \mathbb{R})$ satisfies $f > 0$ μ-a.e. Show that $\int_A f \, d\mu > 0$ for every $A \in \mathcal{A}$ such that $\mu(A) > 0$.

11 Given $f \in \mathcal{L}_1(X, \mu, \mathbb{R})$ with $f \ge 0$, define $\varphi_f(A) := \int_A f \, d\mu$ for $A \in \mathcal{A}$. Prove:

(a) $(X, \varphi_f, \mathcal{A})$ is a finite measure space.

(b) $\mathcal{N}_\mu \subset \mathcal{N}_{\varphi_f}$.

(c) $\mathcal{N}_\mu = \mathcal{N}_{\varphi_f}$ if $f > 0$ μ-a.e.

In particular, show that $(X, \mathcal{A}, \varphi_f)$ is a complete finite measure space if $f > 0$ μ-a.e.
(Hints: (a) Exercise 9. (b) Exercise 10.)

12 Suppose $f \in \mathcal{L}_1(X, \mu, \mathbb{R})$ satisfies $f > 0$ μ-a.e. and take $g \in \mathcal{L}_0(X, \mu, \mathbb{R})$. Show that g is φ_f-integrable if and only if gf is μ-integrable. In this case, show that

$$\int_X g \, d\varphi_f = \int_X fg \, d\mu \ .$$

13 For $f \in \mathcal{L}_1(X, \mu, \overline{\mathbb{R}}^+)$, prove the **Chebyshev inequality**:

$$\mu([f \geq \alpha]) \leq \frac{1}{\alpha} \int_X f \, d\mu \quad \text{for } \alpha > 0 \ .$$

14 Suppose $\mu(X) < \infty$ and let I be a perfect interval in $\mathbb{R}$. Also suppose $\varphi \in C^1(I, \mathbb{R})$ is convex. Prove **Jensen's inequality**, which says that if $f \in \mathcal{L}_1(X, \mu, \mathbb{R})$ satisfies $f(X) \subset I$ and $\varphi \circ f \in \mathcal{L}_1(X, \mu, \mathbb{R})$,

$$\varphi\Bigl(⨍_X f \, d\mu\Bigr) \leq ⨍_X \varphi \circ f \, d\mu \ , \qquad \text{where} \quad ⨍_X f \, d\mu := \frac{1}{\mu(X)} \int_X f \, d\mu \ .$$

(Hints: Show that $\alpha := ⨍ f \, d\mu$ lies in I, so the bound $\varphi(y) \geq \varphi(\alpha) + \varphi'(\alpha)(y - \alpha)$ applies).

15 Suppose $f \in \mathcal{L}_1(X, \mu, E)$. Show that for every $\varepsilon > 0$ there is a $\delta > 0$ such that $\bigl|\int_A f \, d\mu\bigr| < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$. (Hint: Consider Theorem 2.10.)

3 Convergence theorems

Lebesgue integration theory stands out in contrast to the Riemann theory of Chapter VI in that it contains very general and versatile criteria for the commutability of limit taking and integration. Thus the Bochner–Lebesgue integral is better suited to the needs of analysis than the (simpler) Riemann integral.

As usual, we suppose in the entire section that

- $(X, \mathcal{A}, \mu)$ is a complete σ-finite measure space;
  $E = (E, |\cdot|)$ is a Banach space.

Integration of nonnegative $\overline{\mathbb{R}}$-valued functions

In many applications of integration in mathematics, the natural sciences and other fields, real-valued functions play a prominent role. As a rule, one is interested in such cases in integrable functions, which is to say in finite integrals. However, we have already mentioned that the theory gains substantially in simplicity and elegance if it is made to encompass integrals over $\overline{\mathbb{R}}$-valued functions, ruling out infinite values neither for functions nor for integrals. As examples of results that gain from such an inclusive treatment we mention the monotone convergence theorem (Exercise 2.8 and Theorem 3.4) and the Fubini–Tonelli theorem on interchangeability of integrals (Theorem 6.11).

Because of the importance of the real-valued case, and because it offers useful additional results that rely on the total ordering of $\mathbb{R}$ and $\overline{\mathbb{R}}$, we will now develop, to complement the Bochner–Lebesgue integral, an integration theory for $\overline{\mathbb{R}}$-valued — in particular, real-valued — functions.[1]

According to Theorem 1.12, there is for every $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ an increasing sequence (f_j) in $\mathcal{S}(X, \mu, \mathbb{R}^+)$ that converges pointwise to f. It is natural to define the integral of f as the limit in $\overline{\mathbb{R}}^+$ of the increasing sequence $(\int_X f_j \, d\mu)_{j \in \mathbb{N}}$. This makes sense if we can ensure that the limit does not depend on the choice of (f_j).

3.1 Lemma *Suppose $\varphi_j, \psi \in \mathcal{S}(X, \mu, \mathbb{R}^+)$ for $j \in \mathbb{N}$. Also suppose (φ_j) is increasing and $\psi \le \lim_j \varphi_j$. Then*

$$\int_X \psi \, d\mu \le \lim_j \int_X \varphi_j \, d\mu \; .$$

Proof Let $\sum_{j=0}^m \alpha_j \chi_{A_j}$ be the normal form of ψ and fix $\lambda > 1$. For $k \in \mathbb{N}$, define $B_k := [\lambda \varphi_k \ge \psi]$. Because (φ_k) is increasing and $\lambda > 1$, we have $B_k \subset B_{k+1}$ for

[1] The theory of $\overline{\mathbb{R}}$-valued functions is the centerpiece of practically all textbooks on integration theory. It is in some ways simpler than the more general Bochner–Lebesgue theory, and suffices if one is only interested in real- and complex-valued functions, but it is inadequate for the needs of modern higher analysis, which is why we opted for a more general approach.

$k \in \mathbb{N}$ and $\bigcup_{k\in\mathbb{N}} B_k = X$. The continuity of measures from below then implies

$$\int_X \psi\, d\mu = \sum_{j=0}^{m} \alpha_j \mu(A_j) = \lim_k \sum_{j=0}^{m} \alpha_j \mu(A_j \cap B_k) = \lim_k \int_X \psi \chi_{B_k}\, d\mu \ .$$

By the definition of B_k, we have $\lambda\varphi_k \geq \psi\chi_{B_k}$, and we obtain

$$\int_X \psi\, d\mu = \lim_k \int_X \psi\chi_{B_k}\, d\mu \leq \lambda \lim_k \int_X \varphi_k\, d\mu \ .$$

Taking the limit $\lambda \downarrow 1$ now finishes the proof. ■

3.2 Corollary *Suppose* (φ_j) *and* (ψ_j) *are increasing sequences in* $\mathcal{S}(X, \mu, \mathbb{R}^+)$ *such that* $\lim_j \varphi_j = \lim_j \psi_j$. *Then*

$$\lim_j \int_X \varphi_j\, d\mu = \lim_j \int_X \psi_j\, d\mu \qquad \text{in } \overline{\mathbb{R}}^+ \ .$$

Proof By assumption, $\psi_k \leq \lim_j \psi_j = \lim_j \varphi_j$ for $k \in \mathbb{N}$. By Lemma 3.1, we get

$$\int_X \psi_k\, d\mu \leq \lim_j \int_X \varphi_j\, d\mu \quad \text{for } k \in \mathbb{N} \ ,$$

and, as $k \to \infty$,

$$\lim_k \int_X \psi_k\, d\mu \leq \lim_j \int_X \varphi_j\, d\mu \ .$$

Interchanging (φ_j) and (ψ_j), we obtain the opposite inequality, and hence the desired equality. ■

Suppose (φ_j) is an increasing sequence in $\mathcal{S}(X, \mu, \mathbb{R}^+)$ that converges pointwise to $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$. We call

$$\int_X f\, d\mu := \lim_j \int_X \varphi_j\, d\mu$$

the (**Lebesgue**) **integral of f over X with respect to the measure μ**. For $A \in \mathcal{A}$,

$$\int_A f\, d\mu := \int_X \chi_A f\, d\mu$$

is the (**Lebesgue**) **integral of f over the measurable set A**.

3.3 Remarks **(a)** $\int_A f\, d\mu$ is well defined for all $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ and $A \in \mathcal{A}$.

Proof This follows from Theorem 1.12 and Corollary 3.2. ■

(b) For $f, g \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ such that $f \leq g$ μ-a.e., we have $\int_X f\, d\mu \leq \int_X g\, d\mu$.

(c) For $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, these statements are equivalent:

(i) $\int_X f\,d\mu = 0$.

(ii) $[f>0]$ is a μ-null set.

(iii) $f = 0$ μ-a.e.

Proof "(i)⇒(ii)" We set $A := [f>0]$ and $A_j := [f > 1/j]$ for $j \in \mathbb{N}^\times$. Then (A_j) is an increasing sequence in $\mathcal{A}$ such that $A = \bigcup_j A_j$. Also $\chi_{A_j} \le jf$. It follows that

$$0 \le \mu(A_j) = \int_X \chi_{A_j}\,d\mu \le j\int_X f\,d\mu = 0 \quad \text{for } j \in \mathbb{N}^\times\ ,$$

and continuity from below implies $\mu(A) = \lim_j \mu(A_j) = 0$.

"(ii)⇒(iii)" is clear.

"(iii)⇒(i)" Let N be a μ-null set with $f(x) = 0$ for $x \in N^c$. Then[2] $f\chi_{N^c} = 0$ and $f\chi_N \le \infty\chi_N$. Together with the definition of the integral (see also (d) below), this yields

$$0 \le \int_X f\,d\mu = \int_X f\chi_N\,d\mu + \int_X f\chi_{N^c}\,d\mu \le \infty\mu(N) = 0\ . \ ■$$

(d) Suppose $f,g \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$ and $\alpha \in [0,\infty]$. Then

$$\int_X (\alpha f + g)\,d\mu = \alpha\int_X f\,d\mu + \int_X g\,d\mu\ .$$

Proof We consider the case $\alpha = \infty$ and $g = 0$. Letting $\varphi_j := j\chi_{[f>0]}$ for $j \in \mathbb{N}$, we have $f_j \uparrow \infty f$, and hence

$$\int_X (\infty f)\,d\mu = \begin{cases} 0 & \text{if } \mu\big([f>0]\big) = 0\ ,\\ \infty & \text{if } \mu\big([f>0]\big) > 0\ . \end{cases}$$

From (c), it now follows that $\int_X (\infty f)\,d\mu = \infty\int_X f\,d\mu$. The remaining statements follow easily from the definition of the integral and are left as exercises. ■

(e) (i) Suppose $f \in \mathcal{L}_0(X,\mu,\mathbb{R}^+)$ has a finite Lebesgue integral $\int_X f\,d\mu$. Then f belongs to $\mathcal{L}_1(X,\mu,\mathbb{R}^+)$ and the Lebesgue integral of f over X coincides with the Bochner–Lebesgue integral.

(ii) For $f \in \mathcal{L}_1(X,\mu,\mathbb{R}^+)$, the Lebesgue integral $\int_X f\,d\mu$ is finite and agrees with the Bochner–Lebesgue integral.

Proof (i) Theorem 1.12 guarantees the existence of a sequence (φ_j) in $\mathcal{S}(X,\mu,\mathbb{R}^+)$ such that $\varphi_j \uparrow f$. By assumption, there exists for every $\varepsilon > 0$ an $N \in \mathbb{N}$ such that $\int_X f\,d\mu - \int_X \varphi_j\,d\mu < \varepsilon$ for $j \ge N$. For $k \ge j \ge N$, the finiteness of $\int_X f\,d\mu$ now gives

$$\int_X |\varphi_k - \varphi_j|\,d\mu = \int_X (\varphi_k - \varphi_j)\,d\mu \le \int_X (f - \varphi_j)\,d\mu = \int_X f\,d\mu - \int_X \varphi_j\,d\mu < \varepsilon\ .$$

Therefore (φ_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X,\mu,\mathbb{R}^+)$. This shows that f belongs to $\mathcal{L}_1(X,\mu,\mathbb{R}^+)$. The second statement is a consequence of Exercise 2.8.

(ii) This follows from Theorem 1.12 and Exercise 2.8. ■

---

[2] We recall Convention (2.1).

(f) For every $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, we have

$$\int_X f\,d\mu = \sup\Big\{\int_X \varphi\,d\mu\ ;\ \varphi \in \mathcal{S}(X,\mu,\mathbb{R}^+) \text{ with } \varphi \le f\ \ \mu\text{-a.e.}\Big\}\ . \ ■$$

The monotone convergence theorem

We now prove a significant strengthening of the monotone convergence theorem stated in Exercise 2.8 for functions in $\mathcal{L}_1(X,\mu,\mathbb{R})$. We will see that for increasing sequences in $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, Lebesgue integration commutes with taking the limit.

3.4 Theorem (monotone convergence) *Suppose (f_j) is an increasing sequence in $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}+)$. Then*

$$\int_X \lim_j f_j\,d\mu = \lim_j \int_X f_j\,d\mu \quad \textit{in } \overline{\mathbb{R}}^+\ .$$

Proof (i) Set $f := \lim_j f_j$. By Proposition 1.11, f belongs to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, and $f_j \le f$ for $j \in \mathbb{N}$. By Remark 3.3(b), then, we have $\int f_j\,d\mu \le \int f\,d\mu$ for $j \in \mathbb{N}$, and hence $\lim_j \int f_j\,d\mu \le \int f\,d\mu$.

(ii) Suppose $\varphi \in \mathcal{S}(X,\mu,\mathbb{R}^+)$ with $\varphi \le f$. Take $\lambda > 1$ and set $A_j := [\lambda f_j \ge \varphi]$ for $j \in \mathbb{N}$. Then (A_j) is an increasing sequence in $\mathcal{A}$ with $\bigcup_j A_j = X$ and $\lambda f_j \ge \varphi\chi_{A_j}$. Moreover, $\varphi\chi_{A_j} \uparrow \varphi$, so

$$\int_X \varphi\,d\mu = \lim_j \int_X \varphi\chi_{A_j}\,d\mu \le \lambda \lim_j \int_X f_j\,d\mu\ .$$

Taking the limit $\lambda \downarrow 1$ we get $\int_X \varphi\,d\mu \le \lim_j \int_X f_j\,d\mu$ for every μ-simple function φ with $\varphi \le f$. By Remark 3.3(f), it follows that $\int_X f\,d\mu \le \lim_j \int_X f_j\,d\mu$, and we are done. ■

3.5 Corollary *Suppose (f_j) is a sequence in $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$. Then*

$$\sum_{j=0}^{\infty} \int_X f_j\,d\mu = \int_X \Big(\sum_{j=0}^{\infty} f_j\Big)\,d\mu \quad \textit{in } \overline{\mathbb{R}}^+\ .$$

Proof This follows from Corollary 1.13(iii) and Theorem 3.4. ■

3.6 Remarks **(a)** The conclusion of Theorem 3.4 can fail if the sequence is not increasing.

Proof Take $f_j := (1/j)\chi_{[0,j]}$ for $j \in \mathbb{N}^\times$. Then (f_j) is a (nonincreasing) sequence in $\mathcal{S}(\mathbb{R},\lambda_1,\mathbb{R}^+)$ that converges uniformly to 0. But the sequence $\int f_j\,d\lambda_1$ does not converge to 0, because $\int f_j\,d\lambda_1 = 1$ for $j \in \mathbb{N}^\times$. ■

(b) Suppose $a_{j,k} \in \mathbb{R}^+$ for $j, k \in \mathbb{N}$. Then

$$\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} a_{jk} = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_{jk} \ .$$

Proof We set $(X, \mu) := (\mathbb{N}, \mathcal{H}^0)$ and define $f_j : X \to \mathbb{R}^+$ by $f_j(k) := a_{jk}$ for $j, k \in \mathbb{N}$. Then (f_j) is a sequence in $\mathcal{L}_0(X, \mathcal{H}^0, \overline{\mathbb{R}}^+)$ (see Example 2.20), and the claim follows from Corollary 3.5. ■

For nonnegative double series, this result is stronger than Theorem II.8.10, because it is no longer assumed that $\sum_{jk} a_{jk}$ is summable.

Fatou's lemma

We now prove a generalization of the monotone convergence theorem for arbitrary (not necessarily increasing) sequences in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$.

3.7 Theorem (Fatou's lemma) *For every sequence (f_j) in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$, we have*

$$\int_X \Bigl(\varliminf_j f_j\Bigr)\, d\mu \le \varliminf_j \int_X f_j \, d\mu \quad \text{in } \overline{\mathbb{R}}^+ \ .$$

Proof Set $g_j := \inf_{k \ge j} f_k$. By Proposition 1.11, g_j belongs to $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$, and the increasing sequence (g_j) converges to $\varliminf_j f_j$. From Theorem 3.4 we then get $\lim_j \int g_j \, d\mu = \int \bigl(\varliminf_j f_j\bigr)\, d\mu$. Also $g_j \le f_k$, and therefore $\int g_j \, d\mu \le \int f_k \, d\mu$ for $k \ge j$. It follows that $\int g_j \, d\mu \le \inf_{k \ge j} \int f_k \, d\mu$, and taking the limit $j \to \infty$ finishes the proof. ■

3.8 Corollary *Suppose (f_j) is a sequence in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ and $g \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ satisfies $\int_X g \, d\mu < \infty$ with $f_j \le g$ μ-a.e. for $j \in \mathbb{N}$. Then*[3]

$$\varlimsup_j \int_X f_j \, d\mu \le \int_X \Bigl(\varlimsup_j f_j\Bigr)\, d\mu \quad \text{in } \overline{\mathbb{R}}^+ \ .$$

Proof Suppose N is a μ-null set such that $f_j(x) \le g(x)$ for $x \in N^c$ and $j \in \mathbb{N}$. Then $f_j \le g + \infty\chi_N$ on X, and $\int_X (g + \infty\chi_N)\, d\mu = \int_X g \, d\mu$ (see Remarks 3.3(c) and (d)). Therefore we can assume without losing generality that $f_j \le g$ for $j \in \mathbb{N}$. We set $g_j := g - f_j$ and obtain from Fatou's lemma that

$$\int_X \Bigl(\varliminf_j g_j\Bigr)\, d\mu = \int_X g \, d\mu - \int_X \Bigl(\varlimsup_j f_j\Bigr)\, d\mu \le \varliminf_j \int_X g_j \, d\mu = \int_X g \, d\mu - \varlimsup_j \int_X f_j \, d\mu \, .$$

The claim now follows because $\int_X g \, d\mu < \infty$. ■

[3] The assumption $\int_X g \, d\mu < \infty$ cannot be relaxed (see Exercise 1).

As a first application, we prove a fundamental characterization of integrable functions.

3.9 Theorem *For $f \in \mathcal{L}_0(X, \mu, E)$, the following are equivalent:*

(i) $f \in \mathcal{L}_1(X, \mu, E)$;

(ii) $|f| \in \mathcal{L}_1(X, \mu, \mathbb{R})$;

(iii) $\int_X |f|\, d\mu < \infty$.

If these conditions are satisfied, then $\left|\int_X f\, d\mu\right| \le \|f\|_1 < \infty$.

Proof "(i)⇒(ii)" follows from Lemma 2.8, and "(ii)⇒(iii)" is clear. "(iii)⇒(ii)" was proved in Remark 3.3(e).

"(ii)⇒(i)" Suppose (φ_j) is a sequence in $\mathcal{S}(X, \mu, E)$ converging to f μ-a.e. Set $A_j := [\,|\varphi_j| \le 2\,|f|\,]$ and $f_j := \varphi_j \chi_{A_j}$ for $j \in \mathbb{N}$. Theorem 1.7 and Proposition 1.9 show that A_j belongs to $\mathcal{A}$. Thus (f_j) is a sequence in $\mathcal{S}(X, \mu, E)$.

Take $N \in \mathcal{A}$ such that $\mu(N) = 0$ and $\varphi_j(x) \to f(x)$ for $x \in N^c$. If $f(x) \ne 0$ for some $x \in N^c$, there exists $k := k(x) \in \mathbb{N}$ such that $|\varphi_j(x) - f(x)| \le 3\,|f(x)|$ for $j \ge k$. Therefore $x \in N^c \cap [\,|f| > 0]$ belongs to A_j for $j \ge k(x)$. This implies $f_j(x) = \varphi_j(x)$ for $j \ge k(x)$, and therefore $f_j(x) \to f(x)$ for $x \in N^c \cap [\,|f| > 0]$. If $f(x) = 0$ for some $x \in N^c$, then likewise $f_j(x) \to f(x)$ for $j \to \infty$. Because x belongs to A_k for some $k \in \mathbb{N}$, we find $f_k(x) = \varphi_k(x) = 0$ because $|\varphi_k(x)| \le 2\,|f(x)| = 0$. For $x \notin A_k$, we likewise have $f_k(x) = \chi_{A_k}(x)\varphi_k(x) = 0$. This implies $|f - f_j| \to 0$ μ-a.e. Now clearly $|f - f_j| \le 3\,|f|$ for $j \in \mathbb{N}$, so Corollary 3.8 implies

$$\varlimsup_j \int_X |f - f_j|\, d\mu \le \int_X \varlimsup_j |f - f_j|\, d\mu = 0 \ .$$

Therefore we can find for every $\varepsilon > 0$ an $m \in \mathbb{N}$ such that $\int |f - f_j|\, d\mu < \varepsilon/2$ for $j \ge m$. It follows that, for $j, k \in \mathbb{N}$ with $j, k \ge m$,

$$\|f_j - f_k\|_1 = \int_X |f_j - f_k|\, d\mu \le \int_X |f_j - f|\, d\mu + \int_X |f - f_k|\, d\mu < \varepsilon \ .$$

Hence (f_j) is an $\mathcal{L}_1$-Cauchy sequence in $\mathcal{S}(X, \mu, E)$, and f is μ-integrable.

The last statement follows from Theorem 2.11(i). ■

3.10 Conclusions **(a)** Let $f \in \mathcal{L}_0(X, \mu, E)$, and suppose there is a sequence (f_j) in $\mathcal{L}_1(X, \mu, E)$ such that $f_j \to f$ μ-a.e. and $\varliminf_j \|f_j\|_1 < \infty$. Then f belongs to $\mathcal{L}_1(X, \mu, E)$, and $\|f\|_1 \le \varliminf_j \|f_j\|_1$.

Proof By Lemma 2.15, we can assume that (f_j) converges to f on all of X. Using Fatou's lemma, we obtain

$$\int_X |f|\, d\mu = \int_X \varliminf_j |f_j|\, d\mu \le \varliminf_j \int_X |f_j|\, d\mu < \infty \ ,$$

and the claim follows by Theorem 3.9. ■

(b) Let (f_j) be a sequence in $\mathcal{L}_1(X,\mu,\mathbb{R}^+)$. Suppose there is an $f \in \mathcal{L}_1(X,\mu,\mathbb{R})$ such that

$$f_j \to f \ \ \mu\text{-a.e.} \quad \text{and} \quad \int_X f_j \, d\mu \to \int_X f \, d\mu \ \ (j \to \infty) \ .$$

Then[4] (f_j) converges in $\mathcal{L}_1(X,\mu,\mathbb{R})$ to f.

Proof We can assume here too that (f_j) converges to f on all of X. Then $f \geq 0$ and $|f_j - f| \leq f_j + f$. From Theorem 3.7, it follows that

$$\begin{aligned} 2\int_X f \, d\mu &= \int_X \underline{\lim}_j (f_j + f - |f_j - f|) \, d\mu \leq \underline{\lim}_j \int_X (f_j + f - |f_j - f|) \, d\mu \\ &= 2\int_X f \, d\mu - \overline{\lim_j} \int_X |f_j - f| \, d\mu \ . \end{aligned}$$

According to Theorem 3.9, $\int_X f \, d\mu$ is finite, and we find $\lim_j \int_X |f_j - f| \, d\mu = 0$. ■

Integration of $\overline{\mathbb{R}}$-valued functions

The decomposition of an $\overline{\mathbb{R}}$-valued function into its positive and negative parts allows us also to extend the Lebesgue integral to measurable $\overline{\mathbb{R}}$-valued functions admitting negative values. We say that $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$ is **Lebesgue integrable with respect to μ** if $\int_X f^+ \, d\mu < \infty$ and $\int_X f^- \, d\mu < \infty$. In this case,

$$\int_X f \, d\mu := \int_X f^+ \, d\mu - \int_X f^- \, d\mu$$

is called the (**Lebesgue**) **integral over X with respect to the measure μ**.

3.11 Remarks **(a)** For $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}})$, these three statements are equivalent:

(i) f is Lebesgue integrable with respect to μ.

(ii) $\int_X |f| \, d\mu < \infty$;

(iii) There exists $g \in \mathcal{L}_1(X,\mu,\mathbb{R})$ such that $|f| \leq g$ μ-a.e.

Proof "(i)⇒(ii)" This is a consequence of $|f| = f^+ + f^-$.

"(ii)⇒(iii)" Theorem 3.9 says that $|f| \in \mathcal{L}_1(X,\mu,\mathbb{R})$. Hence (iii) holds with $g = |f|$.

"(iii)⇒(i)" This follows from $f^+ \vee f^- \leq |f| \leq g$ and Remark 3.3(b). ■

(b) Suppose $f \in \mathcal{L}_0(X,\mu,\mathbb{R})$. Then f is Lebesgue integrable with respect to μ if and only if f is μ-integrable. In that case, the Lebesgue integral of f over X equals the Bochner–Lebesgue integral. In other words, if we consider real-valued maps, the definition of Lebesgue integrability of $\overline{\mathbb{R}}$-valued functions is consistent with the definition from Section 2.[5]

Proof This follows from (a), Theorem 3.9, and Remark 3.3(e). ■

[4] Compare the statement of Theorem 2.18.

[5] See also Corollary 2.12(iii).

(c) If $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$ is Lebesgue integrable with respect to μ, then $[\,|f| = \infty]$ is a μ-null set.

Proof The assumption implies that $A := [\,|f| = \infty]$ is μ-measurable and also that $\int_X |f|\,d\mu < \infty$. Further we have $\infty\chi_A \le |f|$, and we find by Remarks 3.3(b) and (d) that

$$\infty\mu(A) = \int_X (\infty\chi_A)\,d\mu \le \int_X |f|\,d\mu < \infty\ .$$

Therefore $\mu(A) = 0$. ∎

Lebesgue's dominated convergence theorem

We now prove an extremely versatile and practical theorem about exchanging limits and integrals, proved by Henri Lebesgue. It is one of the cornerstones of Lebesgue integration theory and has countless applications.

3.12 Theorem (dominated convergence[6]) *Let (f_j) be a sequence in $\mathcal{L}_1(X, \mu, E)$ and suppose that there exists $g \in \mathcal{L}_1(X, \mu, \mathbb{R})$ such that*

(a) $$|f_j| \le g\ \ \mu\text{-a.e. for } j \in \mathbb{N}.$$

Suppose also that, for some $f \in E^X$,

(b) $$f_j \to f\ \ \mu\text{-a.e. for } j \to \infty.$$

Then f is μ-integrable, $f_j \to f$ in $\mathcal{L}_1(X, \mu, E)$, and $\int_X f_j\,d\mu \to \int_X f\,d\mu$ in E.

Proof Define

$$g_j := \sup_{k,\ell \ge j} |f_k - f_\ell|$$

for $j \in \mathbb{N}$. By Proposition 1.11, (g_j) is a sequence in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ that converges μ-a.e. to 0. Also $|f_k - f_\ell| \le 2g$ μ-a.e. for $k, \ell \in \mathbb{N}$, and hence $|g_j| \le 2g$ μ-a.e. for $j \in \mathbb{N}$. From Corollary 3.8 it follows that

$$0 \le \varlimsup_j \int_X g_j\,d\mu \le \int_X \varlimsup_j g_j\,d\mu = 0\ .$$

Therefore $\left(\int_X g_j\,d\mu\right)_{j \in \mathbb{N}}$ is a (decreasing) null sequence. This means that for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\int_X |f_k - f_\ell|\,d\mu \le \int_X \sup_{k,\ell \ge j} |f_k - f_\ell|\,d\mu < \varepsilon \quad \text{for } k, \ell \ge j \ge N\ .$$

Hence (f_j) is a Cauchy sequence in $\mathcal{L}_1(X, \mu, E)$, and the claim follows from the completeness of $\mathcal{L}_1(X, \mu, E)$ and Theorem 2.18. ∎

[6] Also referred to as "Lebesgue's theorem".

3.13 Remark The example in Remark 3.6(a) shows that the existence of an integrable dominating function g is essential for Theorem 3.12. ■

As a first application of the dominated convergence theorem, we prove a simple criterion for the integrability of a measurable function.

3.14 Theorem (integrability criterion) *Suppose* $f \in \mathcal{L}_0(X,\mu,E)$ *and* $g \in \mathcal{L}_1(X,\mu,\mathbb{R})$ *satisfy* $|f| \le g$ μ*-a.e. Then* f *belongs to* $\mathcal{L}_1(X,\mu,E)$.

Proof Let (φ_j) be a sequence in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j \to f$ μ-a.e. as $j \to \infty$. Set $A_j := [\,|\varphi_j| \le 2g]$ and $f_j := \chi_{A_j}\varphi_j$ for $j \in \mathbb{N}$. Then (f_j) is a sequence in $\mathcal{S}(X,\mu,E)$ that converges μ-a.e. to f (see the proof of Theorem 3.9). Because $|f_j| \le 2g$ for $j \in \mathbb{N}$, the claim follows from the dominated convergence theorem. ■

3.15 Corollary

(i) *Take* $f \in \mathcal{L}_1(X,\mu,E)$, $g \in \mathcal{L}_0(X,\mu,\mathbb{K})$, *and* $\alpha \in [0,\infty)$ *with* $|g| \le \alpha$ μ*-a.e. Then* gf *is* μ*-integrable, and*

$$\Bigl|\int_X gf\,d\mu\Bigr| \le \alpha\,\|f\|_1 \ .$$

(ii) *Take* $f \in \mathcal{L}_0(X,\mu,E)$ *and* $\alpha \in [0,\infty)$. *If* $|f| \le \alpha$ μ*-a.e. and* $\mu(X) < \infty$, *then* f *is* μ*-integrable with*

$$\Bigl|\int_X f\,d\mu\Bigr| \le \|f\|_1 \le \alpha\mu(X) \ .$$

(iii) *Let* X *be a* σ*-compact metric space and* μ *a complete Radon measure on* X. *Suppose that* $f \in C(X,E)$ *and that* $K \subset X$ *is compact. Then* $\chi_K f$ *belongs to* $\mathcal{L}_1(X,\mu,E)$, *and*

$$\Bigl|\int_K f\,d\mu\Bigr| \le \|\chi_K f\|_\infty\,\mu(K) \ .$$

Proof (i) By Remark 1.2(d), gf is μ-measurable. Also $|gf| \le \alpha\,|f|$ μ-a.e., and $\alpha\,|f|$ is μ-integrable. Hence Theorem 3.14 shows that gf is μ-integrable; Theorem 2.11(i) and Corollary 2.16(ii) imply

$$\Bigl|\int_X gf\,d\mu\Bigr| \le \int_X |gf|\,d\mu \le \int_X \alpha\,|f|\,d\mu = \alpha\,\|f\|_1 \ .$$

(ii) Since $\mu(X)$ is finite, χ_X belongs to $\mathcal{L}_1(X,\mu,\mathbb{R})$. By Theorem 1.7(i), $|f|$ is μ-measurable. Therefore (i) shows (with $g := |f|$ and $f := \chi_X$) that $|f|$ is

μ-integrable and that

$$\int_X |f|\,d\mu \le \alpha \left\|\chi_X\right\|_1 = \alpha\mu(X) < \infty \ .$$

The claim now follows from Theorem 3.9.

(iii) According to Theorem 1.17, f is μ-measurable. Moreover χ_K is μ-simple, because $\mu(K)$ is finite by Remark IX.5.3(a). Therefore $\chi_K f$ is μ-measurable, and the claim follows from (ii) with $\alpha := \max_{x\in K} |f(x)|$. ■

When dealing with a function not defined on all of X, it is occasionally useful in the theory of integration to extend its definition by setting it equal to 0 where it is not already defined. Measurability and integrability questions can then be explored with respect to the measure space $(X, \mathcal{A}, \mu)$. To that end, we set forth the following conventions.

For $f : \mathrm{dom}(f) \subset X \to E$, define the **trivial extension** $\widetilde{f} \in E^X$ of f to X by

$$\widetilde{f}(x) := \begin{cases} f(x) & \text{if } x \in \mathrm{dom}(f)\ , \\ 0 & \text{if } x \notin \mathrm{dom}(f)\ . \end{cases}$$

We say that f is **μ-measurable** or **μ-integrable** if $\widetilde{f}$ belongs to $\mathcal{L}_0(X, \mu, E)$ or $\mathcal{L}_1(X, \mu, E)$, respectively. If f is μ-integrable, we set $\int_X f\,d\mu := \int_X \widetilde{f}\,d\mu$.

3.16 Theorem (termwise integration of series) *Suppose (f_j) is a sequence in $\mathcal{L}_1(X, \mu, E)$ such that $\sum_{j=0}^{\infty} \int_X |f_j|\,d\mu < \infty$. Then $\sum_j f_j$ converges absolutely μ-a.e., $\sum_j f_j$ is μ-integrable, and*

$$\int_X \Bigl(\sum_{j=0}^{\infty} f_j\Bigr)\,d\mu = \sum_{j=0}^{\infty} \int_X f_j\,d\mu \ .$$

Proof (i) By Theorem 1.7(i) and Corollary 1.13(iii), the $\overline{\mathbb{R}}$-valued function $g := \sum_{j=0}^{\infty} |f_j|$ is μ-measurable. Corollary 3.5 implies

$$\int_X g\,d\mu = \sum_{j=0}^{\infty} \int_X |f_j|\,d\mu < \infty \ .$$

It therefore follows from Remarks 3.11(a) and (c) that $[g = \infty]$ is a μ-null set, which proves the absolute convergence of $\sum_j f_j$ for almost every $x \in X$.

(ii) Set $g_k := \sum_{j=0}^{k} f_j$ and $f(x) := \sum_{j=0}^{\infty} f_j(x)$ for $x \in [g < \infty]$. The sequence (g_k) converges μ-a.e. to f and we have the bounds $|g_k| \le \sum_{j=0}^{k} |f_j| \le g$. By the dominated convergence theorem, $\widetilde{f}$ belongs to $\mathcal{L}_1(X, \mu, E)$ and

$$\sum_{j=0}^{\infty} \int_X f_j\,d\mu = \lim_{k\to\infty} \int_X g_k\,d\mu = \int_X \lim_{k\to\infty} g_k\,d\mu = \int_X \Bigl(\sum_{j=0}^{\infty} f_j\Bigr)\,d\mu \ . \ ■$$

Parametrized integrals

As another application of the dominated convergence theorem, we investigate the continuity and differentiability of parametrized integrals.

3.17 Theorem (continuity of parametrized integrals) *Suppose M is a metric space and $f : X \times M \to E$ satisfies*

(a) $f(\cdot, m) \in \mathcal{L}_1(X, \mu, E)$ *for every* $m \in M$;

(b) $f(x, \cdot) \in C(M, E)$ *for μ-almost every* $x \in X$;

(c) *there exists* $g \in \mathcal{L}_1(X, \mu, E)$ *such that* $|f(x, m)| \le g(x)$ *for* $(x, m) \in X \times M$.

Then

$$F : M \to E , \quad m \mapsto \int_X f(x, m)\, \mu(dx)$$

is well defined and continuous.

Proof The first statement follows immediately from (a). Suppose $m \in M$, and let (m_j) be a sequence in M converging to m. We set $f_j := f(\cdot, m_j)$ for $j \in \mathbb{N}$. From (b), it follows that $f_j \to f$ μ-a.e. Therefore by (a) and (c), we can apply the dominated convergence theorem to the sequence (f_j), and we find

$$\lim_{j\to\infty} F(m_j) = \lim_{j\to\infty} \int_X f_j\, d\mu = \int_X \lim_{j\to\infty} f_j\, d\mu = \int_X f(x, m)\, \mu(dx) = F(m) .$$

The claim now follows from Theorem III.1.4. ■

3.18 Theorem (differentiability of parametrized integrals) *Suppose U is open in $\mathbb{R}^n$, or $U \subset \mathbb{K}$ is perfect and convex, and suppose $f : X \times U \to E$ satisfies*

(a) $f(\cdot, y) \in \mathcal{L}_1(X, \mu, E)$ *for every* $y \in U$;

(b) $f(x, \cdot) \in C^1(U, E)$ *for μ-almost every* $x \in X$;

(c) *there exists* $g \in \mathcal{L}_1(X, \mu, \mathbb{R})$ *such that*

$$\Bigl|\frac{\partial}{\partial y^j} f(x, y)\Bigr| \le g(x) \quad \textit{for } (x, y) \in X \times U \textit{ and } 1 \le j \le n .$$

Then

$$F : U \to E , \quad y \mapsto \int_X f(x, y)\, \mu(dx)$$

is continuously differentiable and

$$\partial_j F(y) = \int_X \frac{\partial}{\partial y^j} f(x, y)\, \mu(dx) \quad \textit{for } y \in U \textit{ and } 1 \le j \le n .$$

Proof Take $y \in U$ and $j \in \{1, \ldots, n\}$. Let (h_k) be a null sequence in $\mathbb{K}$ such that $h_k \neq 0$ and $y + h_k e_j \in U$ for $k \in \mathbb{N}$. Finally, set

$$f_k(x) := \frac{f(x, y + h_k e_j) - f(x, y)}{h_k} \quad \text{for } x \in X \text{ and } k \in \mathbb{N} \;,$$

The mean value theorem (Theorem VII.3.9) then gives

$$|f_k(x)| \leq \sup_{z \in U} \Bigl| \frac{\partial}{\partial y^j} f(x, z) \Bigr| \leq g(x) \quad \mu\text{-a.e.}$$

Because (f_k) converges μ-a.e. to $\partial f(\,\cdot\,, y)/\partial y^j$, it follows from Theorem 3.12 that

$$\lim_{k \to \infty} \frac{F(y + h_k e_j) - F(y)}{h_k} = \lim_{k \to \infty} \int_X f_k \, d\mu = \int_X \frac{\partial}{\partial y^j} f(x, y) \, \mu(dx) \;.$$

Therefore F is partially differentiable, and $\partial_j F(y) = \int_X \bigl(\partial/\partial y^j\bigr) f(x, y) \, \mu(dx)$. The result now follows from Theorems 3.17 and VII.2.10. ■

3.19 Corollary *Suppose U is open in $\mathbb{C}$, and $f : X \times U \to \mathbb{C}$ satisfies*

(a) *$f(\,\cdot\,, z) \in \mathcal{L}_1(X, \mu, \mathbb{C})$ for every $z \in U$;*

(b) *$f(x, \cdot\,) \in C^\omega(U, \mathbb{C})$ for μ-almost every $x \in X$;*

(c) *there is a $g \in \mathcal{L}_1(X, \mu, \mathbb{R})$ such that $|f(x, z)| \leq g(x)$ for $(x, z) \in X \times U$.*

Then

$$F : U \to \mathbb{C} \;, \quad z \mapsto \int_X f(x, z) \, \mu(dx)$$

is holomorphic, and

$$F^{(n)}(z) = \int_X \frac{\partial^n}{\partial z^n} f(x, z) \, \mu(dx) \tag{3.1}$$

for every $n \in \mathbb{N}$.

Proof Take $z_0 \in U$ and $r > 0$ such that $\overline{\mathbb{D}}(z_0, r) \subset U$. Cauchy's derivative formula (Corollary VIII.5.12) gives

$$\frac{\partial}{\partial z} f(x, z) = \frac{1}{2\pi i} \int_{\partial \mathbb{D}(z, r)} \frac{f(x, \zeta)}{(\zeta - z)^2} \, d\zeta \quad \text{for } \mu\text{-almost every } x \in X \text{ and } z \in \mathbb{D}(z_0, r) \;,$$

and we find from (c) and Proposition VIII.4.3(iv) that

$$\Bigl| \frac{\partial}{\partial z} f(x, z) \Bigr| \leq \frac{g(x)}{r} \quad \text{for } \mu\text{-almost every } x \in X \text{ and } z \in \mathbb{D}(z_0, r) \;.$$

Theorem 3.18 now shows that $F \,|\, \mathbb{D}(z_0, r)$ belongs to $C^1\bigl(\mathbb{D}(z_0, r), \mathbb{C}\bigr)$ and satisfies

$$F'(z) = \int_X \frac{\partial}{\partial z} f(x, z) \, \mu(dx) \quad \text{for } z \in \mathbb{D}(z_0, r) \;.$$

Holomorphy is a local property, so Theorem VIII.5.11 implies that F belongs to $C^\omega(U, \mathbb{C})$. The validity of (3.1) now follows from a simple induction argument. ■

Exercises

1 Find a measure space $(X, \mathcal{A}, \mu)$, a sequence (f_j) in $\mathcal{L}_0(X, \mu, \mathbb{R}^+)$, and a function g in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ such that

$$f_j \leq g \text{ for } j \in \mathbb{N} \quad \text{and} \quad \varlimsup_j \int_X f_j \, d\mu > \int_X (\varlimsup_j f_j) \, d\mu \ .$$

2 Suppose $f \in \mathcal{L}_1(X, \mu, E)$ and $\varepsilon > 0$. Show that there exists $A \in \mathcal{A}$ such that

$$\mu(A) < \infty \quad \text{and} \quad \Bigl|\int_X f \, d\mu - \int_B f \, d\mu\Bigr| < \varepsilon$$

for every $B \in \mathcal{A}$ with $B \supset A$.

3 Suppose (f_j) is a sequence in $\mathcal{L}_1(X, \mu, E)$ converging in measure to $f \in \mathcal{L}_0(X, \mu, E)$. Also suppose there is $g \in \mathcal{L}_1(X, \mu, \mathbb{R})$ such that $|f_j| \leq g$ μ-a.e. for all $j \in \mathbb{N}$. Then f belongs to $\mathcal{L}_1(X, \mu, E)$,

$$f_j \to f \text{ in } \mathcal{L}_1(X, \mu, E) \ , \quad \text{and} \quad \int_X f_j \, d\mu \to \int_X f \, d\mu \text{ in } E \ .$$

(Hint: If $\bigl(\int_X f_j \, d\mu\bigr)$ does not converge to $\int_X f \, d\mu$, there is a subsequence $(f_{j_k})_{k \in \mathbb{N}}$ and a $\delta > 0$ such that

$$\|f_{j_k} - f\|_1 \geq \delta \quad \text{for } k \in \mathbb{N} \ . \tag{3.2}$$

Use Exercise 1.15 and Theorem 3.12 to derive a contradiction from (3.2).

4 Let $f, g \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}})$ be Lebesgue integrable functions. Prove:

(i) If $f \leq g$ μ-a.e., then $\int_X f \, d\mu \leq \int_X g \, d\mu$.

(ii) $\displaystyle\Bigl|\int_X f \, d\mu\Bigr| \leq \int_X |f| \, d\mu$.

(iii) $f \wedge g$ and $f \vee g$ are Lebesgue integrable, and

$$-\int_X (|f| + |g|) \, d\mu \leq \int_X (f \wedge g) \, d\mu \leq \int_X (f \vee g) \, d\mu \leq \int_X (|f| + |g|) \, d\mu \ .$$

5 Suppose the sequence (f_j) in $\mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$ converges in measure to $f \in \mathcal{L}_0(X, \mu, \overline{\mathbb{R}}^+)$. Prove that

$$\int_X f \, d\mu \leq \varliminf_j \int_X f_j \, d\mu \ .$$

6 For $x \in \mathbb{R}^n \backslash \{0\}$, define

$$k_n(x) := \begin{cases} x^+ & \text{if } n = 1 \ , \\ \log|x| & \text{if } n = 2 \ , \\ |x|^{2-n} & \text{if } n \geq 3 \ . \end{cases}$$

Further suppose $U \subset \mathbb{R}^n$ is open and nonempty, that $A \in \mathcal{L}(n)$ satisfies $A \subset U^c$, and that $f \in C_c(\mathbb{R}^n)$.

(a) The map $A \to \mathbb{R}$, $x \mapsto f(x) k_n(|y - x|)$ is λ_n-integrable for every $y \in U$.

(b) The map $U \to \mathbb{R}$, $y \mapsto \int_A f(x) k_n(|y - x|) \, \lambda_n(dx)$ is smooth and harmonic.

7 Verify that

(i) $\mathcal{L}_1(\mathbb{R}^n, \lambda_n, E) \cap BC(\mathbb{R}^n, E) \subsetneq C_0(\mathbb{R}^n, E)$;

(ii) $\mathcal{L}_1(\mathbb{R}^n, \lambda_n, E) \cap BUC(\mathbb{R}^n, E) \subseteq C_0(\mathbb{R}^n, E)$.

4 Lebesgue spaces

We saw in Corollary VI.7.4 that the space of continuous $\mathbb{K}$-valued functions over a compact interval I is not complete with respect to the L_2 norm. The framework of Lebesgue integration theory now gives us the means to complete the inner product space $\bigl(C(I,\mathbb{K}),(\cdot\,|\,\cdot)_2\bigr)$: we will construct a vector space L_2 and an extension of $(\cdot\,|\,\cdot)_2$ onto $L_2\times L_2$ — also denoted by $(\cdot\,|\,\cdot)_2$ — such that $\bigl(L_2,(\cdot\,|\,\cdot)_2\bigr)$ is a Hilbert space containing $C(I,\mathbb{K})$ as a dense subspace.

This construction can be generalized in a natural way, leading to a new family of Banach spaces, the Lebesgue L_p-spaces. These are of great importance in many areas of mathematics.

In the following, we suppose that

- $(X,\mathcal{A},\mu)$ is a complete σ-finite measure space;
  $E=(E,|\cdot|)$ is a Banach space.

Essentially bounded functions

We say that a function $f\in\mathcal{L}_0(X,\mu,E)$ is ***μ*-essentially bounded** if there exists $\alpha\ge 0$ such that $\mu\bigl([\,|f|>\alpha]\bigr)=0$. The ***μ*-essential supremum** of f is then[1]

$$\|f\|_\infty := \operatorname*{ess\text{-}sup}_{x\in X}|f(x)| := \inf\bigl\{\,\alpha\ge 0\ ;\ \mu\bigl([\,|f|>\alpha]\bigr)=0\,\bigr\}\ .$$

4.1 Remarks **(a)** Let $f\in\mathcal{L}_0(X,\mu,E)$. There is equivalence between:

(i) f is μ-essentially bounded;

(ii) $\|f\|_\infty<\infty$;

(iii) f is bounded μ-a.e.

Proof "(i)⇒(ii)⇒(iii)" is clear.

"(iii)⇒(i)" Suppose N is a μ-null set and take $\alpha\ge 0$ such that $|f(x)|\le\alpha$ for $x\in N^c$. Then $[\,|f|>\alpha]\subset N$, and the completeness of μ implies that $\mu\bigl([\,|f|>\alpha]\bigr)=0$. ■

(b) Suppose $f\in\mathcal{L}_0(X,\mu,E)$. Then $|f|\le\|f\|_\infty$ μ-a.e.

Proof The case $\|f\|_\infty=\infty$ is clear. If $\|f\|_\infty<\infty$, then $[\,|f|>\|f\|_\infty+2^{-j}]$ is a μ-null set for every $j\in\mathbb{N}$, and hence so is the set $[\,|f|>\|f\|_\infty]=\bigcup_{j\in\mathbb{N}}[\,|f|>\|f\|_\infty+2^{-j}]$. ■

(c) Suppose f and g are μ-essentially bounded and $\alpha\in\mathbb{K}$. Then $\alpha f+g$ is also μ-essentially bounded, and

$$\|\alpha f+g\|_\infty\le|\alpha|\,\|f\|_\infty+\|g\|_\infty\ .$$

[1] Note that now $\|\cdot\|_\infty$ has two meanings, namely, the essential supremum of a measurable function and the supremum norm of a bounded function. The two values need *not* be the same; see (d) and (e) in Remark 4.1. When necessary we denote the supremum norm by $\|\cdot\|_{B(X,E)}$.

Proof By (a) and (b), there exist μ-null sets M and N such that $|f(x)| \le \|f\|_\infty$ for $x \in M^c$ and $|g(x)| \le \|g\|_\infty$ for $x \in N^c$. Therefore

$$|\alpha f(x) + g(x)| \le |\alpha|\,\|f\|_\infty + \|g\|_\infty \quad \text{for } x \in (M \cup N)^c = M^c \cap N^c \ .$$

Hence $\alpha f + g$ is μ-essentially bounded and $\|\alpha f + g\|_\infty \le |\alpha|\,\|f\|_\infty + \|g\|_\infty$. ∎

(d) Suppose $f \in \mathcal{L}_0(X, \mu, E)$ is bounded. Then $\|f\|_\infty \le \|f\|_{B(X,E)}$ (supremum norm). If N is a nonempty μ-null set, then $\|\chi_N\|_\infty = 0$ and $\|\chi_N\|_{B(X,E)} = 1$.

(e) Suppose X is σ-compact metric space and μ is a massive Radon measure on X. Then

$$\|f\|_\infty = \|f\|_{B(X,E)} \quad \text{for } f \in BC(X,E) \ .$$

Proof By Theorem 1.17, any $f \in BC(X,E)$ is μ-measurable, and by (d) we just have to show that $\|f\|_{B(X,E)} \le \|f\|_\infty$. Assume otherwise. Then there exists $x \in X$ such that

$$\|f\|_\infty < |f(x)| \le \|f\|_{B(X,E)} \ ,$$

and in view of the continuity of f there is an open neighborhood O of x in X such that $\|f\|_\infty < |f(y)|$ for $y \in O$. From (b) it follows that $\mu(O) = 0$, contradicting the assumption that μ is massive. ∎

The Hölder and Minkowski inequalities

Suppose $f \in \mathcal{L}_0(X, \mu, E)$. For $p \in (0, \infty)$, we set

$$\|f\|_p := \Bigl(\int_X |f|^p \, d\mu\Bigr)^{1/p}$$

with the convention that $\infty^{1/p} := \infty$. We define the **Lebesgue space over X with respect to the measure μ** as[2]

$$\mathcal{L}_p(X, \mu, E) := \bigl\{\, f \in \mathcal{L}_0(X, \mu, E) \ ; \ \|f\|_p < \infty \,\bigr\} \quad \text{for } p \in (0, \infty] \ .$$

For $p \in [1, \infty]$, we define the **dual exponent to p** as

$$p' := \begin{cases} \infty & \text{if } p = 1 \ , \\ p/(p-1) & \text{if } p \in (1, \infty) \ , \\ 1 & \text{if } p = \infty \ . \end{cases}$$

With this assignment, we obviously have

$$\frac{1}{p} + \frac{1}{p'} = 1 \quad \text{for } p \in [1, \infty] \ .$$

We are now in a position to state and prove two important inequalities.

[2] Theorem 3.9 shows that the notation $\mathcal{L}_p(X, \mu, E)$ is consistent in the case $p = 1$ with that of Section 2. In the following, we concentrate on the Lebesgue spaces $\mathcal{L}_p$ with $p \in [1, \infty]$. The case $p \in (0, 1)$ will be treated in Exercise 13.

4.2 Theorem *Suppose* $p \in [1, \infty]$.

(i) *For* $f \in \mathcal{L}_p(X, \mu, \mathbb{K})$ *and* $g \in \mathcal{L}_{p'}(X, \mu, \mathbb{K})$, *we have* $fg \in \mathcal{L}_1(X, \mu, \mathbb{K})$, *and*

$$\Bigl|\int_X fg\, d\mu\Bigr| \le \int_X |fg|\, d\mu \le \|f\|_p \, \|g\|_{p'} \qquad (\textbf{Hölder's}^3 \textbf{ inequality}).$$

(ii) *Suppose* $f, g \in \mathcal{L}_p(X, \mu, E)$. *Then* $f + g \in \mathcal{L}_p(X, \mu, E)$, *and*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p \qquad (\textbf{Minkowski's inequality}).$$

Proof (i) We consider first the case $p = 1$. By Remark 4.1(b), there is a μ-null set N such that $|g(x)| \le \|g\|_\infty$ for $x \in N^c$. It then follows from Remarks 1.2(d) and 3.3(b) and Lemma 2.15 that

$$\int_{N^c} |fg|\, d\mu \le \|g\|_\infty \int_{N^c} |f|\, d\mu = \|f\|_1 \, \|g\|_\infty < \infty \ .$$

Hence Remark 3.11(a), Theorem 3.9 and Lemma 2.15 result in fg being integrable, and Theorem 2.11(i) implies

$$\Bigl|\int_X fg\, d\mu\Bigr| \le \int_X |fg|\, d\mu = \int_{N^c} |fg|\, d\mu \le \|f\|_1 \, \|g\|_\infty \ .$$

Suppose now $p \in (1, \infty)$. If

$$f = 0 \ \mu\text{-a.e.} \quad \text{or} \quad g = 0 \ \mu\text{-a.e.} \ , \tag{4.1}$$

then fg also vanishes μ-a.e., and the claim follows from Corollary 2.16. On the other hand, if (4.1) does not apply, Corollary 2.19 gives $\|f\|_p > 0$ and $\|g\|_{p'} > 0$. We then set $\xi := |f| / \|f\|_p$, $\eta := |g| / \|g\|_{p'}$, and obtain from Young's inequality (Theorem IV.2.15) that

$$\frac{|fg|}{\|f\|_p \, \|g\|_{p'}} \le \frac{1}{p} \, \frac{|f|^p}{\|f\|_p^p} + \frac{1}{p'} \, \frac{|g|^{p'}}{\|g\|_{p'}^{p'}} \ .$$

It follows that

$$\begin{aligned} \int_X |fg|\, d\mu &\le \frac{1}{p} \, \|f\|_p^{1-p} \, \|g\|_{p'} \int_X |f|^p \, d\mu + \frac{1}{p'} \, \|f\|_p \, \|g\|_{p'}^{1-p'} \int_X |g|^{p'} \, d\mu \\ &= \|f\|_p \, \|g\|_{p'} \ , \end{aligned}$$

and we conclude using Theorem 3.9 that fg belongs to $\mathcal{L}_1(X, \mu, E)$. Therefore

$$\Bigl|\int_X fg\, d\mu\Bigr| \le \|fg\|_1 \le \|f\|_p \, \|g\|_{p'} \ .$$

The case $p = \infty$ is treated analogously to the case $p = 1$.

[3] For $p = 2$, this is the **Cauchy–Schwarz inequality**.

(ii) Because of Corollary 2.9 and Remark 4.1(c), it suffices to consider the case $p \in (1,\infty)$. In addition, we can assume without loss of generality that $\|f+g\|_p > 0$. We will first prove that $f+g$ belongs to $\mathcal{L}_p(X,\mu,E)$. Noting the inequality

$$|a+b|^p \le \bigl(2(|a| \vee |b|)\bigr)^p \le 2^p(|a|^p + |b|^p) \quad \text{for } a,b \in E \ , \tag{4.2}$$

we obtain

$$\int_X |f+g|^p \, d\mu \le 2^p \Bigl(\int_X |f|^p \, d\mu + \int_X |g|^p \, d\mu \Bigr) < \infty$$

because $f, g \in \mathcal{L}_p(X,\mu,E)$. Therefore $\|f+g\|_p < \infty$. Due to the equivalence

$$|f+g|^{p-1} \in \mathcal{L}_{p'}(X,\mu,\mathbb{R}) \Longleftrightarrow |f+g| \in \mathcal{L}_p(X,\mu,\mathbb{R}) \ ,$$

it follows from Hölder's inequality that

$$\int_X |h| \, |f+g|^{p-1} \, d\mu \le \|h\|_p \, \bigl\| |f+g|^{p-1} \bigr\|_{p'} = \|h\|_p \, \|f+g\|_p^{p/p'}$$

for $h \in \mathcal{L}_p(X,\mu,E)$, and we find

$$\begin{aligned} \int_X |f+g|^p \, d\mu &\le \int_X |f| \, |f+g|^{p-1} \, d\mu + \int_X |g| \, |f+g|^{p-1} \, d\mu \\ &\le \bigl(\|f\|_p + \|g\|_p \bigr) \, \|f+g\|_p^{p/p'} \ . \end{aligned} \tag{4.3}$$

The claim follows, because $\|f+g\|_p < \infty$ and $p/p' = p-1$. ■

4.3 Corollary *Suppose $p \in [1,\infty]$. Then $\mathcal{L}_p(X,\mu,E)$ is a vector subspace of $\mathcal{L}_0(X,\mu,E)$, and $\|\cdot\|_p$ is a seminorm on $\mathcal{L}_p(X,\mu,E)$.*

4.4 Remarks **(a)** Set $\mathcal{N} := \bigl\{ f \in \mathcal{L}_0(X,\mu,E) \ ; \ f = 0 \ \mu\text{-a.e.} \bigr\}$. For $f \in \mathcal{L}_0(X,\mu,E)$ the following statements are equivalent:

(i) $\|f\|_p = 0$ for all $p \in [1,\infty]$.

(ii) $\|f\|_p = 0$ for some $p \in [1,\infty]$.

(iii) $f \in \mathcal{N}$.

Proof "(i)⟹(ii)" is trivial. "(ii)⟹(iii)" follows from Corollary 2.19 and Remark 4.1(b).

"(iii)⟹(i)" For $p \in [1,\infty)$, use Lemma 2.15. The case $p = \infty$ is clear. ■

(b) $\mathcal{N}$ is a vector subspace of $\mathcal{L}_p(X,\mu,E)$ for every $p \in [1,\infty] \cup \{0\}$.

Proof The case $p = 0$ is clear; in particular, $\mathcal{N}$ is a vector space. For $p \in [1,\infty]$, the claim then follows from (a), "(iii)⟹(i)". ■

(c) For $p \in [1,\infty]$, we have these inclusions of vector subspaces:

$$\mathcal{S}(X,\mu,E) \subset \mathcal{L}_p(X,\mu,E) \subset \mathcal{L}_0(X,\mu,E) \ .$$

Proof It is clear that every μ-simple function is μ-essentially bounded. Take $p \in [1,\infty)$ and let $\varphi \in \mathcal{S}(X,\mu,E)$ have normal form $\sum_{j=0}^m e_j \chi_{A_j}$. Then $|\varphi|^p \le \sum_{j=0}^m |e_j|^p \, \chi_{A_j}$, so $\|\varphi\|_p < \infty$. The claim follows by Remark 1.2(a) and Corollary 4.3. ■

Lebesgue spaces are complete

We now generalize Theorem 2.10(ii), proving that all Lebesgue spaces $\mathcal{L}_p(X,\mu,E)$ with $p\in[1,\infty]$ are complete. For $p\in(1,\infty)$, this depends on the following lemma.

4.5 Lemma *Suppose V is a vector space and q is a seminorm on V. The following statements are equivalent*:

(i) *(V,q) is complete.*

(ii) *For every sequence $(v_j)\in V^{\mathbb{N}}$ such that $\sum_{j=0}^{\infty} q(v_j)<\infty$, the series $\sum_j v_j$ converges in V.*

Proof "(i)⇒(ii)" Suppose $(v_j)\in V^{\mathbb{N}}$ and $\sum_{j=0}^{\infty} q(v_j)<\infty$. For every $\varepsilon>0$ there exists $K\in\mathbb{N}$ such that $\sum_{j=\ell+1}^{\infty} q(v_j)<\varepsilon$ for $\ell\ge K$ (see Exercise II.7.4). We set $w_k:=\sum_{j=0}^{k} v_j$ for $k\in\mathbb{N}$ and get

$$q(w_m-w_\ell)=q\Big(\sum_{j=\ell+1}^{m} v_j\Big)\le\sum_{j=\ell+1}^{m} q(v_j)\le\sum_{j=\ell+1}^{\infty} q(v_j)<\varepsilon\quad\text{for } m>\ell\ge K\ .$$

Therefore (w_k) is a Cauchy sequence in V, and so converges to some $v\in V$ by the completeness of V. Hence the series $\sum_j v_j$ converges.

"(ii)⇒(i)" Let (v_j) be a Cauchy sequence in V. For $k\in\mathbb{N}$, take $j_k\in\mathbb{N}$ such that $q(v_{j_{k+1}}-v_{j_k})<2^{-(k+1)}$. Setting $w_k:=v_{j_{k+1}}-v_{j_k}$, we have $\sum_{k=0}^{\infty} q(w_k)\le 1$, and we can find by assumption a $v\in V$ such that $q\big(v-\sum_{k=0}^{\ell} w_k\big)\to 0$ as $\ell\to\infty$. Let $\varepsilon>0$ and $L\in\mathbb{N}$ be such that $q\big(v-\sum_{k=0}^{\ell} w_k\big)<\varepsilon/2$ for $\ell\ge L$. Because (v_j) is a Cauchy sequence in V, there exists $K\ge L$ such that $q(v_{j_{\ell+1}}-v_k)<\varepsilon/2$ for $k,\ell\ge K$. Finally setting $\widetilde{v}:=v+v_{j_0}$, we have for $k\ge K$ that

$$\begin{aligned}q(\widetilde{v}-v_k)&=q(v+v_{j_0}-v_{j_{K+1}}+v_{j_{K+1}}-v_k)\\&\le q\Big(v-\sum_{k=0}^{K} w_k\Big)+q(v_{j_{K+1}}-v_k)<\varepsilon\ .\end{aligned}$$

This shows that (v_k) converges to $\widetilde{v}$. ■

4.6 Theorem *For $p\in[1,\infty]$, $\mathcal{L}_p(X,\mu,E)$ is complete.*

Proof (i) Consider first the case $p\in(1,\infty)$. Let (f_j) be a sequence in $\mathcal{L}_p(X,\mu,E)$ such that $\sum_{j=0}^{\infty}\|f_j\|_p<\infty$. Set $g_k:=\sum_{j=0}^{k}|f_j|$ for $k\in\mathbb{N}$ and $g:=\sum_{k=0}^{\infty}|f_j|$. By Corollary 1.13(iii), g belongs to $\mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, and we have $|g_k|^p\to|g|^p$. Because

$$\|g_k\|_p\le\sum_{j=0}^{k}\|f_j\|_p\le\sum_{j=0}^{\infty}\|f_j\|_p<\infty\ ,$$

Conclusion 3.10(a) tells us that $g\in\mathcal{L}_p(X,\mu,\mathbb{R})$. By Remark 3.11(c), then, there is a μ-null set N such that $g(x)<\infty$ for $x\in N^c$. Therefore $f(x):=\sum_{j=0}^{\infty} f_j(x)$ is well defined for every $x \in N^c$ by the Weierstrass criterion (Theorem V.1.6). Also, since $|f|^p \le g^p$ μ-a.e. and $g \in \mathcal{L}_p(X,\mu,\mathbb{R})$, Theorem 3.14 implies that $\widetilde{f}$ belongs to $\mathcal{L}_p(X,\mu,E)$. Finally Fatou's lemma shows that

$$\Bigl\|\widetilde{f}-\sum_{j=0}^{k} f_j\Bigr\|_p^p = \int_X \Bigl|\lim_{\ell\to\infty}\sum_{j=k+1}^{\ell} f_j\Bigr|^p d\mu \le \varliminf_{\ell\to\infty}\int_X \Bigl|\sum_{j=k+1}^{\ell} f_j\Bigr|^p d\mu = \varliminf_{\ell\to\infty}\Bigl\|\sum_{j=k+1}^{\ell} f_j\Bigr\|_p^p ,$$

and we find

$$\Bigl\|\widetilde{f}-\sum_{j=0}^{k} f_j\Bigr\|_p \le \varliminf_{\ell\to\infty}\sum_{j=k+1}^{\ell}\|f_j\|_p = \sum_{j=k+1}^{\infty}\|f_j\|_p \quad \text{for } k\in\mathbb{N}\ .$$

Because $\sum_{j=0}^{\infty}\|f_j\|_p < \infty$, the sequence $\bigl(\sum_{j=k+1}^{\infty}\|f_j\|_p\bigr)_{k\in\mathbb{N}}$ converges to zero. Therefore so does $\bigl(\bigl\|\widetilde{f}-\sum_{j=0}^{k} f_j\bigr\|_p\bigr)_{k\in\mathbb{N}}$. Now it follows from Lemma 4.5 that $\mathcal{L}_p(X,\mu,E)$ is complete.

(ii) Now suppose (f_j) is a Cauchy sequence in $\mathcal{L}_\infty(X,\mu,E)$. We set

$$A_j := [\,|f_j| > \|f_j\|_\infty]\ ,\quad B_{k,\ell} := [\,|f_k-f_\ell| > \|f_k-f_\ell\|_\infty] \quad \text{for } j,k,\ell\in\mathbb{N}$$

and $N := \bigcup_j A_j \cup \bigcup_{k,\ell} B_{k,\ell}$. By Remarks 4.1(b) and IX.2.5(b), N is a null set and

$$|f_j(x)| \le \|f_j\|_\infty\ ,\quad |f_k(x)-f_\ell(x)| \le \|f_k-f_\ell\|_\infty \quad \text{for } j,k,\ell\in\mathbb{N}\ ,\ x\in N^c\ .$$

Therefore $(f_j\,|\,N^c)$ is a Cauchy sequence in the Banach space $B(N^c,E)$, and we can find an $f\in B(N^c,E)$ such that $(f_j\,|\,N^c)$ converges uniformly to f. Thus (f_j) converges μ-a.e. to $\widetilde{f}$. We know the function $\widetilde{f}$ is μ-essentially bounded because $\bigl[\bigl|wtf\bigr| > \|f\|_{B(N^c,E)}\bigr] = \emptyset$, and we have

$$\bigl|\widetilde{f}(x)-f_j(x)\bigr| \le \|f-f_j\,|\,N^c\|_{B(N^c,E)} \quad \text{for } x\in N^c \text{ and } j\in\mathbb{N}\ .$$

Hence (f_j) converges in $\mathcal{L}_\infty(X,\mu,E)$ to $\widetilde{f}$.

(iii) The case $p=1$ was dealt with in Theorem 2.10(ii). ■

4.7 Corollary *Let $p\in[1,\infty]$, and suppose $f_j, f\in\mathcal{L}_p(X,\mu,E)$ satisfy $f_j\to f$ in $\mathcal{L}_p(X,\mu,E)$.*

(i) *If $p=\infty$, then (f_j) converges μ-a.e. to f.*

(ii) *If $p\in[1,\infty)$, there is a subsequence $(f_{j_k})_{k\in\mathbb{N}}$ of (f_j) converging μ-a.e. to f.*

Proof Because (f_j) converges in $\mathcal{L}_p(X,\mu,E)$ to f, we know (f_j) is a Cauchy sequence in $\mathcal{L}_p(X,\mu,E)$. Statement (i) now follows immediately from the proof of Theorem 4.6.

If $p\in(1,\infty)$, choose a subsequence $(f_{j_k})_{k\in\mathbb{N}}$ of (f_j) such that $\|f_{j_{k+1}}-f_{j_k}\|_p < 2^{-(k+1)}$. Then the proof of Theorem 4.6 shows that there is a $g\in\mathcal{L}_p(X,\mu,E)$ such that $(f_{j_k}-f_{j_0})\to g$ in $\mathcal{L}_p(X,\mu,E)$ and $(f_{j_k}-f_{j_0})\to g$ μ-a.e. as $k\to\infty$. Because (f_j) converges in $\mathcal{L}_p(X,\mu,E)$ to f, we have $\|f-(g+f_{j_0})\|_p = 0$. Remark 4.4(a) implies $f = g+f_{j_0}$ μ-a.e., from which the claim follows.

The case $p=1$ was treated in Theorem 2.18. ■

4.8 Proposition *$\mathcal{S}(X,\mu,E)$ is dense in $\mathcal{L}_p(X,\mu,E)$ for $p\in[1,\infty)$.*[4]

Proof Suppose $f\in\mathcal{L}_p(X,\mu,E)$. Then f is μ-measurable by Remark 4.4(c). Thus there is a sequence (φ_j) in $\mathcal{S}(X,\mu,E)$ such that $\varphi_j\to f$ μ-a.e. as $j\to\infty$. We set $A_j:=[\,|\varphi_j|\le 2\,|f|\,]$ and $\psi_j:=\chi_{A_j}\varphi_j$. Then (ψ_j) is a sequence in $\mathcal{S}(X,\mu,E)$ that converges μ-a.e. to f. Moreover,

$$|\psi_j-f|^p\le(|\psi_j|+|f|)^p\le 3^p\,|f|^p\quad\text{for } j\in\mathbb{N}\ .$$

Because $3^p\,|f|^p$ belongs to $\mathcal{L}_1(X,\mu,\mathbb{R})$, we can apply the dominated convergence theorem, and we find

$$\|\psi_j-f\|_p^p=\int_X|\psi_j-f|^p\,d\mu\to 0\quad\text{as } j\to\infty\ ,$$

from which the claim follows. ■

L_p-spaces

We proved in Remark 4.4(b) that

$$\mathcal{N}:=\{\,f\in\mathcal{L}_0(X,\mu,E)\ ;\ f=0\ \mu\text{-a.e.}\,\}$$

is a vector subspace of $\mathcal{L}_p(X,\mu,E)$ for $p\in\{0\}\cup[1,\infty]$. Hence the quotient spaces

$$L_p(X,\mu,E):=\mathcal{L}_p(X,\mu,E)/\mathcal{N}\quad\text{for } p\in\{0\}\cup[1,\infty]$$

are well defined vector spaces over $\mathbb{K}$, by Example I.12.3(i). By Remark 4.4(c), we also have

$$L_p(X,\mu,E)\subset L_0(X,\mu,E)\quad\text{for } p\in[1,\infty]\ ,$$

in the sense of vector subspaces. Suppose $[f]\in L_0(X,\mu,E)$, and let g be a representative of $[f]$. Then $f-g\in\mathcal{N}$, that is, f and g agree μ-a.e. By Remark 4.4(a), the map

$$|||\cdot|||: L_0(X,\mu,E)\to\overline{\mathbb{R}}^+\ ,\quad [f]\mapsto\|f\|_p$$

is well defined for every $p\in[1,\infty]$, and for $[f]\in L_p(X,\mu,E)$, we have

$$|||\,[f]\,|||_p=\|f\|_p=0\iff f=0\ \mu\text{-a.e.}\iff [f]=0\ . \tag{4.4}$$

Since $|||\cdot|||_p$ obviously inherits the properties of the seminorm $\|\cdot\|_p$, (4.4) shows that $|||\cdot|||_p$ is a norm on $L_p(X,\mu,E)$. Therefore $L_p(X,\mu,E)$ is a normed vector space, whereas the space we constructed it from, $\mathcal{L}_p(X,\mu,E)$, is only seminormed. So limits in $\mathcal{L}_p(X,\mu,E)$ are generally not unique, but limits in $L_p(X,\mu,E)$ are.[5] The price we pay for the better topological structure of $L_p(X,\mu,E)$ is that its elements are not functions on X but rather cosets of the vector subspace $\mathcal{N}$ of $\mathcal{L}_p(X,\mu,E)$. In other words, we identify functions that coincide μ-a.e. Experience shows that the following simplified notation does not lead to misunderstandings.

[4]The statement can fail if $p=\infty$; see Exercise 8 (but also Exercise 9).
[5]See Remark 2.3(b).

Convention Suppose $p \in \{0\} \cup [1,\infty]$. Then we write the coset $[f] = f + \mathcal{N}$ in $L_p(X,\mu,E)$ as f and identify with each other functions that agree μ-a.e. Further, if $p \in [1,\infty]$, we denote the norm in $L_p(X,\mu,E)$ by $\|\cdot\|_p$ and set

$$L_p(X,\mu,E) := \bigl(L_p(X,\mu,E), \|\cdot\|_p\bigr) \quad \text{for } p \in [1,\infty] \ .$$

4.9 Remarks **(a)** For $f \in L_0(X,\mu,E)$ and $x \in X$, $f(x)$ is *undefined* if μ has nonempty null sets. That is, elements of $L_0(X,\mu,E)$ cannot be "evaluated pointwise". (Of course, if one chooses a representative $\overset{*}{f}$ of f, then $\overset{*}{f}(x)$ is defined.)

(b) For $p \in [1,\infty]$,

$$L_p(X,\mu,E) = \bigl\{\, f \in L_0(X,\mu,E) \ ; \ \|f\|_p < \infty \,\bigr\} \ .$$

Proof "$\subseteq$" Let $f \in L_p(X,\mu,E)$. Any representative $\overset{*}{f}$ of f lies in $\mathcal{L}_p(X,\mu,E)$, that is, it is μ-measurable and satisfies $\|\overset{*}{f}\|_p < \infty$. Hence f belongs to $L_0(X,\mu,E)$, and $\|f\|_p < \infty$.

"$\supseteq$" Consider $f \in L_0(X,\mu,E)$ with $\|f\|_p < \infty$. Every representative $\overset{*}{f}$ of f is μ-measurable, with $\|f\|_p = \|\overset{*}{f}\|_p < \infty$. Thus $\overset{*}{f}$ belongs to $\mathcal{L}_p(X,\mu,E)$, and so f belongs to $L_p(X,\mu,E)$. ∎

(c) Suppose $f,g \in L_0(X,\mu,\mathbb{R})$, and let $\overset{*}{f}, \overset{*}{g}$ be representatives of f,g. If we write

$$f \le g \ :\Longleftrightarrow \ \overset{*}{f} \le \overset{*}{g} \ \ \mu\text{-a.e.} \ ,$$

we obtain a well defined ordering $\le$ on $L_0(X,\mu,\mathbb{R})$, which makes this space into a vector lattice.

Proof We leave the simple proof as an exercise. ∎

(d) Suppose $(F,\le)$ is a vector lattice and $(F,\|\cdot\|)$ is a Banach space. If $|x| \le |y|$ implies $\|x\| \le \|y\|$, we call $(F,\le,\|\cdot\|)$ a **Banach lattice**.

(e) $\bigl(L_p(X,\mu,\mathbb{R}), \le, \|\cdot\|_p\bigr)$ is a Banach lattice for every $p \in [1,\infty]$.

Proof It is clear that $L_p(X,\mu,\mathbb{R})$ is a vector sublattice of $L_0(X,\mu,\mathbb{R})$. Also it follows immediately from the monotony of integrals and of the map $t \mapsto t^p$ that $L_p(X,\mu,\mathbb{R})$ is a Banach lattice in the case $p \in [1,\infty)$.

Suppose $f,g \in L_\infty(X,\mu,\mathbb{R})$ with $|f| \le |g|$, and let $\overset{*}{f}, \overset{*}{g}$ be representatives thereof. Then $|\overset{*}{f}| \le |\overset{*}{g}|$ μ-a.e. In addition, Remark 4.1(b) shows that $|\overset{*}{g}| \le \|g\|_\infty$ μ-a.e. Therefore $|\overset{*}{f}| \le \|g\|_\infty$ μ-a.e., and hence $\|f\|_\infty \le \|g\|_\infty$. ∎

4.10 Theorem

(i) *$L_p(X,\mu,E)$ is a Banach space for every $p \in [1,\infty]$.*

(ii) *If H is a Hilbert space, then so is $L_2(X,\mu,H)$ with respect to the scalar product*

$$(\cdot\,|\,\cdot)_2 : L_2(X,\mu,H) \times L_2(X,\mu,H) \to \mathbb{K} \ , \quad (f,g) \mapsto \int_X (f\,|\,g)_H \, d\mu \ .$$

Proof (i) Suppose $p \in [1, \infty]$. We already know that $L_p(X, \mu, E)$ is a normed vector space. Let (f_j) be a Cauchy sequence in $L_p(X, \mu, E)$, and $(\overset{*}{f}_j)$ a corresponding sequence of representatives. Then $(\overset{*}{f}_j)$ is a Cauchy sequence in $\mathcal{L}_p(X, \mu, E)$. By Theorem 4.6, there exists $\overset{*}{f} \in \mathcal{L}_p(X, \mu, E)$ such that $\|\overset{*}{f}_j - \overset{*}{f}\|_p \to 0$ as $j \to \infty$. Letting $f := \overset{*}{f} + \mathcal{N}$, we have $f \in L_p(X, \mu, E)$ and $\|f_j - f\|_p = \|\overset{*}{f}_j - \overset{*}{f}\|_p \to 0$. Therefore $L_p(X, \mu, E)$ is complete.

(ii) Using statements (i) and (iv) of Theorem 1.7 and Hölder's inequality, we easily prove that $(\cdot \,|\, \cdot)_2$ is a scalar product on $L_2(X, \mu, H)$ satisfying $|(f \,|\, f)_2| = \|f\|_2^2$ for $f \in L_2(X, \mu, H)$. The claim then follows from (i). ■

4.11 Corollary *$L_2(X, \mu, \mathbb{K})$ is a Hilbert space with respect to the scalar product*

$$(f \,|\, g)_2 = \int_X f\bar{g}\, d\mu \quad \text{for } f, g \in L_2(X, \mu, \mathbb{K}) \ .$$

Continuous functions with compact support

Let Y be a topological space. For $f \in E^Y$, we call

$$\operatorname{supp}(f) := \overline{\{\, x \in Y \ ; \ f(x) \neq 0 \,\}}$$

the **support** of f. Here, as usual, the bar denotes the closure (in Y). Continuous functions with compact support are particularly significant. We therefore define

$$C_c(Y, E) := \{\, f \in C(Y, E) \ ; \ \operatorname{supp}(f) \text{ is compact} \,\} \ .$$

4.12 Examples **(a)** For the Dirichlet function $\chi_{\mathbb{Q}} \in \mathbb{R}^{\mathbb{R}}$ of Example III.1.3(c), we have

$$\operatorname{supp}(\chi_{\mathbb{Q}}) = \operatorname{supp}(\chi_{\mathbb{R}-\mathbb{Q}}) = \mathbb{R} \ .$$

Proof This follows from Propositions I.10.8 and I.10.11. ■

(b) Suppose $X = \mathbb{Z}$ or $X = \mathbb{N}$, and provide X with the metric induced from $\mathbb{R}$. Let $\mathcal{H}^0$ be the counting measure on $\mathfrak{P}(X)$. Then[6]

$$C_c(X, E) = \mathcal{S}(X, \mathcal{H}^0, E) = \{\, \varphi \in E^X \ ; \ \operatorname{Num}[\varphi \neq 0] < \infty \,\} \ .$$

(c) Suppose X is a metric space. Then $C_c(X, E)$ is a vector subspace of $BC(X, E)$. If X is compact, then $C_c(X, E) = C(X, E) = BC(X, E)$.

Proof The first statement follows from Corollary III.3.7. The second is a consequence of Exercise III.3.2 and Corollary III.3.7. ■

[6] Compare Example 2.20.

4.13 Proposition *Suppose X is a metric space and A and B are closed, disjoint nonempty subsets of X. There exists $\varphi \in C(X)$ such that $0 \le \varphi \le 1$, $\varphi \,|\, A = 1$, and $\varphi \,|\, B = 0$. Such a function is a called a* **Urysohn function**.

Proof If $D \subset X$ is nonempty, Example III.1.3(l) shows that the distance function $d(\,\cdot\,, D)$ belongs to $C(X)$. If D is also closed, we have $d(x, D) = 0$ if and only if $x \in D$. Using these properties, we easily prove that the function defined by

$$\varphi(x) := \frac{d(x,B)}{d(x,A) + d(x,B)} \quad \text{for } x \in X \ ,$$

has the stated properties. ■

With help from Urysohn functions, we can now prove an important approximation theorem.

4.14 Theorem *Suppose X is a σ-compact metric space and μ is a Radon measure on X. Then $C_c(X, E)$ is a dense vector subspace of $\mathcal{L}_p(X, \mu, E)$ for $p \in [1, \infty)$.*

Proof Suppose $\varepsilon > 0$. According to Proposition 4.8, $\mathcal{S}(X, \mu, E)$ is dense in $\mathcal{L}_p(X, \mu, E)$. Thus, because of Theorem 1.17 and Minkowski's inequality (that is, the triangle inequality), it suffices to verify that for every μ-measurable set A of finite measure and every $e \in E\backslash\{0\}$, there exists $f \in C_c(X, E)$ such that $\|f - \chi_A e\|_p < \varepsilon$.

Suppose then that $A \in \mathcal{A}$ with $\mu(A) < \infty$. Because μ is regular, we can find a compact subset K and an open subset U of X such that $K \subset A \subset U$ and

$$\mu(U\backslash K) = \mu(U) - \mu(K) < (\varepsilon/|e|)^p \ .$$

Proposition 4.13 secures the existence of a Urysohn function φ on X with $\varphi \,|\, K = 1$ and $\varphi \,|\, U^c = 0$. Setting $f := \varphi e$, we get, as needed,

$$\|\chi_A e - f\|_p^p \le |e|^p \int_X \chi_{U\backslash K}\, d\mu \le |e|^p \, \mu(U\backslash K) < \varepsilon^p \ . \ ■$$

Embeddings

Suppose X and Y are topological spaces, and X is a subset of Y. Denoting by $j : X \to Y,\ x \mapsto x$ the inclusion[7] of X in Y, we say X is **continuously embedded** in Y if j is continuous.[8] In this case, we write $X \hookrightarrow Y$. We write $X \overset{d}{\hookrightarrow} Y$ if X is also a dense subset of Y. If X and Y are vector spaces, the notation $X \hookrightarrow Y$ (and the term "continuously embedded") will always mean in addition that X is a vector subspace of Y, not just any odd subset.

[7]See Example I.3.2(b).

[8]These notions become important when X is *not* provided with the topology induced by Y; see Remark 4.15(a).

4.15 Remarks **(a)** Suppose V and W are normed vector spaces. V is continuously embedded in W if and only if V is a vector subspace of W and there is an $\alpha > 0$ such that $\|v\|_W \leq \alpha \|v\|_V$ for $v \in V$, that is, if the norm of V is stronger than the norm induced from W on V.

If V carries the norm induced by W, then $V \hookrightarrow W$ always.

(b) Suppose X is open in $\mathbb{R}^n$. Then

$$BUC^k(X, E) \hookrightarrow BUC^\ell(X, E) \quad \text{for } k \geq \ell \ .$$

If X is bounded as well, then

$$BUC^k(X, \mathbb{K}) \overset{d}{\hookrightarrow} BUC(X, \mathbb{K}) \quad \text{for } k \in \mathbb{N} \ .$$

Proof The first statement is clear. The second follows from the Stone–Weierstrass approximation theorem (Corollary V.4.8) and then Application VI.2.2. ■

Simple examples (see Exercise 5.1) show that Lebesgue spaces are generally not contained in one another. Under suitable extra assumptions on the measure space $(X, \mathcal{A}, \mu)$, continuous embeddings exist for Lebesgue spaces. For example, if $\mathcal{H}^0$ is the counting measure on $\mathfrak{P}(\mathbb{N})$, the spaces ℓ_p introduced in Exercise 1.16 coincide with $\mathcal{L}_p(\mathbb{N}, \mathcal{H}^0, \mathbb{K})$ for $1 \leq p \leq \infty$, and we have the embeddings

$$\ell_1 \hookrightarrow \ell_p \hookrightarrow \ell_q \hookrightarrow \ell_\infty \quad \text{for } 1 \leq p \leq q \leq \infty \ ,$$

(see Exercise 11).

Finite measure spaces present an altogether different situation:

4.16 Theorem *Let $(X, \mathcal{A}, \mu)$ be a finite complete measure space. Then*

$$L_q(X, \mu, E) \overset{d}{\hookrightarrow} L_p(X, \mu, E) \quad \textit{for } 1 \leq p < q \leq \infty$$

and

$$\|f\|_p \leq \mu(X)^{1/p - 1/q} \, \|f\|_q \quad \textit{for } f \in L_q(X, \mu, E) \ . \tag{4.5}$$

Proof (i) Take $f \in L_q(X, \mu, E)$ and set $r := q/p$. Let $g \in \mathcal{L}_q(X, \mu, E)$ be a representative of f. Then $|g|^p$ belongs to $\mathcal{L}_r(X, \mu, \mathbb{R})$, and $1/r' = (q-p)/q$. Further, χ_X belongs to $\mathcal{L}_{r'}(X, \mu, \mathbb{R})$, because μ is a finite measure. Thus in the case $q < \infty$ Hölder's inequality gives

$$\|g\|_p^p = \int_X \chi_X \, |g|^p \, d\mu \leq \Bigl(\int_X \chi_X^{r'} \, d\mu\Bigr)^{1/r'} \Bigl(\int_X |g|^{pr} \, d\mu\Bigr)^{1/r} = \mu(X)^{(q-p)/q} \, \|g\|_q^p \ ,$$

and we find $\|g\|_p \leq \mu(X)^{1/p - 1/q} \, \|g\|_q$; this clearly also holds in the case $q = \infty$. Because g is an arbitrary representative of f, we see that f belongs to $L_p(X, \mu, E)$ and (4.5) holds. By Remark 4.15(a), it follows that $L_q(X, \mu, E) \hookrightarrow L_p(X, \mu, E)$.

(ii) $M := \{ [\varphi] \in L_0(X,\mu,E) \ ; \ \varphi \in \mathcal{S}(X,\mu,E) \}$ satisfies $M \subset L_q(X,\mu,E)$, and, because $p < \infty$, Proposition 4.8 implies that M is dense in $L_p(X,\mu,E)$. Therefore $L_q(X,\mu,E)$ is also dense in $L_p(X,\mu,E)$. ■

The next theorem shows that, in the case of a massive Radon measure μ, an element of $L_0(X,\mu,E)$ has at most one continuous representative. In this case, then, we can identify each function in $C(X,E)$ with the equivalence class it generates in $L_0(X,\mu,E)$, and regard $C(X,E)$ as a vector subspace of $L_0(X,\mu,E)$.

4.17 Proposition *Suppose μ is a massive Radon measure on a σ-compact space X. Then the map*

$$C(X,E) \to L_0(X,\mu,E) \ , \quad f \mapsto [f] \tag{4.6}$$

is linear and injective.

Proof Theorem 1.17 shows that the map (4.6) is well defined and linear.

Take $f, g \in C(X,E)$ with $[f] = [g]$. There exists $h \in \mathcal{N}$ such that $f - g = h$, that is, $f - g = 0$ μ-a.e. Assume for a contradiction that $f(x) \neq g(x)$ for some $x \in X$. By continuity, $(f-g)(y) \neq 0$ for all y in some open neighborhood U of x. But $\mu(U) > 0$, contrary to the assumption that $f - g = 0$ μ-a.e. Therefore $f = g$, which proves the asserted injectivity. ■

Convention Let μ be a massive Radon measure on a σ-compact space X. We identify $C(X,E)$ with its image in $L_0(X,\mu,E)$ under the injection (4.6) and so regard $C(X,E)$ as a vector subspace of $L_0(X,\mu,E)$. Then

$$\|f\|_{B(X,E)} = \|f\|_\infty \quad \text{for } f \in BC(X,E) \ .$$

The following result is a simple consequence of this convention.

4.18 Theorem *Let μ be a massive Radon measure on a σ-compact metric space X.*

(i) *$C_c(X,E)$ is a dense vector subspace of $L_p(X,\mu,E)$ for every $p \in [1,\infty)$.*

(ii) *$BC(X,E)$ is a closed vector subspace of $L_\infty(X,\mu,E)$.*

Proof The first statement follows from Theorem 4.14. The second is obvious. ■

Continuous linear functionals on L_p

For the rest of this section, we use for $p \in [1,\infty]$ the abbreviations

$$L_p(X) := L_p(X,\mu,\mathbb{K}) \quad \text{and} \quad L_p'(X) := \big(L_p(X)\big)' \ ,$$

the prime on the right indicating the dual space (Remark VII.2.13(a)). From Hölder's inequality, it follows that, for every $f \in L_{p'}(X)$, the map

$$T_f : L_p(X) \to \mathbb{K} \ , \quad g \mapsto \int_X fg \, d\mu$$

is a continuous linear functional on $L_p(X)$, that is, an element of $L_p'(X)$; it satisfies

$$\|T_f\|_{L_p'(X)} \leq \|f\|_{p'} \ . \tag{4.7}$$

In fact (4.7) holds with equality:

4.19 Proposition *The map*

$$T\colon L_{p'}(X) \to L_p'(X) \ , \quad f \mapsto T_f$$

is a linear isometry for every $p \in [1,\infty]$.

Proof (i) Clearly T is linear. Also, in view of (4.7), we need only show that for every $f \in L_{p'}(X)$ satisfying $f \neq 0$ and every $\varepsilon > 0$, there is $g \in L_p(X)$ such that

$$\|g\|_p = 1 \quad \text{and} \quad \|f\|_{p'} < \Bigl|\int_X fg\,d\mu\Bigr| + \varepsilon \ .$$

(ii) First assume $p \in (1,\infty)$, so $p' \in (1,\infty)$. Therefore

$$g := \overline{\operatorname{sign} f}\,\|f\|_{p'}^{1-p'}\,|f|^{p'-1}$$

is well defined and μ-measurable (see Exercise 1.19 and Theorem 1.7(i)). Also

$$\int_X |g|^p\,d\mu = \|f\|_{p'}^{p(1-p')} \int_X |f|^{p(p'-1)}\,d\mu = \|f\|_{p'}^{-p'}\,\|f\|_{p'}^{p'} = 1$$

and $fg = \|f\|_{p'}^{1-p'}\,|f|^{p'}$. Therefore $\|f\|_{p'} = \int_X fg\,d\mu$.

For $p = \infty$, we set $g := \overline{\operatorname{sign} f}$. Then

$$\|g\|_\infty = 1 \quad \text{and} \quad \|f\|_1 = \int_X fg\,d\mu \ .$$

(iii) Now suppose that $p = 1$. Suppose $0 < \varepsilon < \|f\|_\infty$ and set $\alpha := \|f\|_\infty - \varepsilon$. Because $[\,|f| > \alpha]$ has positive measure and μ is σ-finite, we can find $A \in \mathcal{A}$ such that $A \subset [\,|f| > \alpha]$ and $\mu(A) \in (0,\infty)$. Therefore $g := \overline{\operatorname{sign} f}\,\bigl(1/\mu(A)\bigr)\chi_A$ is well defined and μ-measurable. Clearly $\|g\|_1 = 1$ and

$$\int_X fg\,d\mu = \frac{1}{\mu(A)} \int_A |f|\,d\mu \geq \alpha = \|f\|_\infty - \varepsilon \ .$$

This concludes the proof. ■

4.20 Remarks **(a)** One can show that the map T of Proposition 4.19 is surjective for every $p \in [1,\infty)$, that is, every continuous linear functional on $L_p(X)$ can be represented is of the form T_f for an appropriate $f \in L_{p'}(X)$; see [Rud83, Theorem 6.1.6], for example. Consequently $T : L_{p'}(X) \to L'_p(X)$ is an isometric isomorphism for every $p \in [1,\infty)$. *This isomorphism allows us to identify $L_{p'}(X)$ with $L'_p(X)$ for $p \in [1,\infty)$.* The dual pairing $\langle\cdot,\cdot\rangle_{L_p}$: $L'_p(X) \times L_p(X) \to \mathbb{K}$ satisfies

$$\langle g, f\rangle_{L_p} = \int_X fg\,d\mu \quad \text{for } (g,f) \in L_{p'}(X) \times L_p(X)\ .$$

(b) In the case $p = \infty$, the map $T : L_1(X) \to L'_\infty(X)$ is generally not surjective; see [Fol99, S. 191].

(c) Denote by $\langle\cdot,\cdot\rangle_E : E' \times E \to \mathbb{K}$ the duality pairing between E and E'. Then the map

$$\kappa : E \to [E']'\ , \quad e \mapsto \langle\cdot, e\rangle_E$$

is linear and bounded. Its norm is at most 1.

Proof Clearly κ is linear. Suppose $e \in E$ with $\|e\|_E \le 1$. Then

$$\bigl|\bigl\langle \kappa(e), e'\bigr\rangle_{E'}\bigr| = |\langle e', e\rangle_E| \le \|e'\|_{E'} \quad \text{for } e' \in E'\ ,$$

and we find $\|\kappa(e)\|_{(E')'} \le 1$, from which the claim follows. ■

(d) With tools from functional analysis, one can show that κ is an isometry and therefore injective. We call κ the **canonical injection** of E into the **double dual space** $E'' := (E')'$ of E. If κ is surjective as well, and hence an isometric isomorphism, we say E is **reflexive**. In this case, the canonical isomorphism κ allows us to identify E with its double dual E''.

(e) $L_p(X)$ reflexive for $p \in (1,\infty)$.

Proof This follows from (a). ■

(f) The spaces $L_1(X)$ and $L_\infty(X)$ are generally not reflexive; see, for instance, [Ada75, Theorem 2.35]. ■

Exercises

1 Let $S(X,\mu,E) := \bigl\{\,[f] \in L_0(X,\mu,E)\ ;\ [f] \cap \mathcal{S}(X,\mu,E) \ne \emptyset\,\bigr\}$. Prove that $S(X,\mu,E)$ a dense vector subspace of $L_p(X,\mu,E)$ for $1 \le p < \infty$.

2 For $a \in \mathbb{R}^n$, we define $\tau_a : E^{(\mathbb{R}^n)} \to E^{(\mathbb{R}^n)}$, the **right translation by *a***, by

$$(\tau_a\varphi)(x) := \varphi(x-a) \quad \text{for } x \in \mathbb{R}^n\ , \quad \varphi \in E^{(\mathbb{R}^n)}\ .$$

Set $\tau_a[f] := [\tau_a f]$ for $[f] \in L_p$. Prove:

(i) $(\mathbb{R}^n,+) \to \bigl(\mathcal{L}\mathrm{aut}\bigl(L_p(\mathbb{R}^n,\lambda_n,E)\bigr),\circ\bigr)$, $a \mapsto \tau_a$ is a group homomorphism with $\|\tau_a\|_{\mathcal{L}(L_p)} = 1$ for every $p \in [1,\infty]$.

(ii) For $p \in [1,\infty)$ and $f \in L_p(\mathbb{R}^n, \lambda_n, E)$, we have $\lim_{a\to 0} \|\tau_a f - f\|_p = 0$.

(iii) If $\lim_{a\to 0} \|\tau_a f - f\|_\infty = 0$, there exists $g \in BUC(\mathbb{R}^n, E)$ such that $f = g$ μ-a.e.

3 Suppose μ is a complete Radon measure on a σ-compact space X, and let $(X_j)_{j\in\mathbb{N}}$ be a sequence of relatively compact open subsets of X covering X. For $p \in [1,\infty]$, set

$$q_{j,p}(f) := \|\chi_{X_j} f\|_p \quad \text{for } j \in \mathbb{N}\ , \quad f \in L_0(X,\mu,E)\ ,$$
$$L_{p,\mathrm{loc}}(X,\mu,E) := \bigl\{ f \in L_0(X,\mu,E)\ ;\ q_{j,p}(f) < \infty,\ j \in \mathbb{N} \bigr\}\ .$$

Finally, define

$$d_p(f,g) := \sum_{j=0}^{\infty} \frac{2^{-j} q_{j,p}(f-g)}{1 + q_{j,p}(f-g)} \quad \text{for } f,g \in L_{p,\mathrm{loc}}(X,\mu,E)\ .$$

(i) $L_{p,\mathrm{loc}}(X,\mu,E)$ is well defined, that is, independent of the particular sequence (X_j).

(ii) $\bigl(L_{p,\mathrm{loc}}(X,\mu,E), d_p\bigr)$ is a complete metric space.

(iii) $L_p(X,\mu,E) \overset{d}{\hookrightarrow} L_{p,\mathrm{loc}}(X,\mu,E) \overset{d}{\hookrightarrow} L_{1,\mathrm{loc}}(X,\mu,E)$.

(iv) The topology generated by d_p is independent of the sequence (X_j).

4 Suppose $p, q \in [1,\infty]$ and define

$$L_p \cap L_q := (L_p \cap L_q)(X,\mu,E) := L_p(X,\mu,E) \cap L_q(X,\mu,E)\ ,$$
$$L_p + L_q := (L_p + L_q)(X,\mu,E) := L_p(X,\mu,E) + L_q(X,\mu,E)\ .$$

Also set $\|f\|_{L_p\cap L_q} := \|f\|_p + \|f\|_q$ for $f \in L_p \cap L_q$, and put

$$\|f\|_{L_p+L_q} := \inf\bigl\{ \|g\|_p + \|h\|_q\ ;\ g \in L_p(X,\mu,E),\ h \in L_q(X,\mu,E) \text{ with } f = g + h \bigr\}$$

for $f \in L_p + L_q$.

(i) Check that the **interpolation inequality**

$$\|f\|_r \le \|f\|_p^{1-\theta}\, \|f\|_q^{\theta}\ , \qquad \text{where } \frac{1}{r} := \frac{1-\theta}{p} + \frac{\theta}{q}\ ,$$

holds for $f \in L_p \cap L_q$ and $\theta \in [0,1]$.

(ii) $(L_p \cap L_q, \|\cdot\|_{L_p\cap L_q})$ and $(L_p + L_q, \|\cdot\|_{L_p+L_q})$ are Banach spaces with

$$(L_p \cap L_q)(X,\mu,E) \hookrightarrow L_r(X,\mu,E) \hookrightarrow (L_p + L_q)(X,\mu,E) \hookrightarrow L_{1,\mathrm{loc}}(X,\mu,E)$$

for $1 \le p \le r \le q \le \infty$.

(Hints: (i) Hölder's inequality. (ii) Take $f \in L_p + L_q$ with $\|f\|_{L_p+L_q} = 0$. To show it vanishes, note that $L_r \hookrightarrow L_{1,\mathrm{loc}}$ for $r \in [1,\infty]$ (see Exercise 3). To prove the completeness of $L_p + L_q$ apply Lemma 4.5. The embedding $L_p \cap L_q \hookrightarrow L_r$ follows from (a).)

5 Suppose $p \in [1,\infty)$ and $f \in (L_p \cap L_\infty)(X,\mu,E)$. Prove that $\lim_{q\to\infty} \|f\|_q = \|f\|_\infty$.

6 Prove that the map

$$L_\infty(X,\mu,\mathbb{K}) \times L_p(X,\mu,E) \to L_p(X,\mu,E)\ , \quad \bigl([\varphi],[f]\bigr) \mapsto [\varphi f]$$

is bilinear and continuous and has norm at most 1.

7 Suppose $\mu(X) < \infty$, and for $f, g \in L_0(X, \mu, E)$ put

$$d_0(f,g) := \int_X \frac{|f-g|}{1+|f-g|}\, d\mu \ .$$

(i) $\bigl(L_0(X,\mu,E), d_0\bigr)$ is a metric space.

(ii) (f_j) converges to 0 in $\bigl(L_0(X,\mu,E), d_0\bigr)$ if and only if it converges to 0 in measure.

8 Let μ be a Radon measure on a σ-compact space X and let E be separable. Prove:

(i) $C_c(X, \mathbb{K})$ is separable.

(ii) $C_c(X, E)$ is separable.

(iii) $L_p(X, \mu, E)$ is separable for $p \in [1, \infty)$.

(iv) $L_\infty(X, \mu, E)$ is generally not separable.

(v) $\mathcal{S}(X, \mu, E)$ is generally not dense in $\mathcal{L}_\infty(X, \mu, E)$.

(Hints: (i) Corollary V.4.8 and Remark 1.16(e). (ii) Take $A \subset C_c(X, \mathbb{K})$ and let B be countable and dense in E. For $a \in A$ and $b \in B$, set $(a \otimes b)(x) := a(x)b$ for $x \in X$ and consider

$$\bigl\{ \textstyle\sum_{j=0}^m a_j \otimes b_j \ ; \ m \in \mathbb{N},\ (a_j, b_j) \in A \times B,\ j = 0, \dots, m \bigr\} \ .$$

(iii) Theorem 4.14. (iv) Find an uncountable subset A of L_∞ such that $\|f-g\|_\infty \ge 1$ for all distinct $f, g \in A$.)

9 If μ finite and E is finite-dimensional, show that $\mathcal{S}(X, \mu, E)$ is dense in $\mathcal{L}_\infty(X, \mu, E)$.

10 Prove the statement of Remark 4.9(c).

11 Prove:

(i) $\ell_p = \mathcal{L}_p(\mathbb{N}, \mathcal{H}^0, \mathbb{K})$ for $1 \le p \le \infty$.

(ii) $\ell_p \hookrightarrow \ell_q$ with $\|\cdot\|_q \le \|\cdot\|_p$ if $1 \le p \le q \le \infty$.

(iii) $\ell_p \overset{d}{\hookrightarrow} \ell_q \overset{d}{\hookrightarrow} c_0 \hookrightarrow \ell_\infty$ if $1 \le p \le q < \infty$ (see Section II.2).

12 Suppose $p, q \in [1, \infty]$ with $1 \le p \le q \le \infty$. Prove:

(i) $L_\infty(X,\mu,E) \subset L_1(X,\mu,E) \Rightarrow L_q(X,\mu,E) \hookrightarrow L_p(X,\mu,E)$.

(ii) $L_1(X,\mu,E) \subset L_\infty(X,\mu,E) \Rightarrow L_p(X,\mu,E) \hookrightarrow L_q(X,\mu,E)$.

(iii) There exists a complete σ-finite measure space $(X, \mathcal{A}, \mu)$ [or $(Y, \mathcal{B}, \nu)$] realizing the embedding $L_\infty(X,\mu,\mathbb{R}) \hookrightarrow L_1(X,\mu,\mathbb{R})$ [or $L_1(Y,\nu,\mathbb{R}) \hookrightarrow L_\infty(Y,\nu,\mathbb{R})$].

(Hints: (i) Hölder's inequality. (ii) Show that $L_p \hookrightarrow L_\infty$ and apply Exercise 4(i).)

13 For $p \in (0, 1)$, prove:

(i) $\|f+g\|_p^p \le \|f\|_p^p + \|g\|_p^p$ for $f, g \in \mathcal{L}_0(X, \mu, E)$.

(ii) $\|f+g\|_p \le 2^{1/p-1}(\|f\|_p + \|g\|_p)$ for $f, g \in \mathcal{L}_0(X, \mu, E)$.

(iii) $\mathcal{L}_p(X, \mu, E)$ is a vector subspace of $\mathcal{L}_0(X, \mu, E)$.

(iv) $\mathcal{N} := \bigl\{ f \in \mathcal{L}_0(X,\mu,E) \ ; \ f = 0 \ \ \mu\text{-a.e.} \bigr\}$ is a vector subspace of $\mathcal{L}_p(X, \mu, E)$, and

$$\mathcal{N} = \bigl\{ f \in \mathcal{L}_p(X,\mu,E) \ ; \ \|f\|_p = 0 \bigr\} \ .$$

(v) Putting $\rho(f,g) := \|f-g\|_p^p$ induces a metric on

$$L_p(X,\mu,E) := \mathcal{L}_p(X,\mu,E)/\mathcal{N} \ .$$

(vi) $\bigl(L_p(X,\mu,E), \rho\bigr)$ is complete.

(vii) For $f, g \in \mathcal{L}_p(X,\mu,\mathbb{R})$ with $f \geq 0$ and $g \geq 0$, we have $\|f+g\|_p \geq \|f\|_p + \|g\|_p$.

(viii) The map

$$L_p(X,\mu,\mathbb{R}) \to \mathbb{R}^+ \ , \quad [f] \mapsto \|f\|_p$$

is *not* a norm.

(Hints: (i) For $a > 0$, the map $\bigl[t \mapsto a^p + t^p - (a+t)^p\bigr]$ is increasing on $\mathbb{R}^+$. (ii) For $a > 0$, examine $\bigl[t \mapsto (a^{1/p} + t^{1/p})/(a+t)^{1/p}\bigr]$. (vi) Adapt the proof of Lemma 4.5 and Theorem 4.6. (vii) Theorem 4.2.)

14 Suppose $p_j \in [1,\infty]$ for $j = 1,\ldots,m$; let $1/r := \sum_{j=1}^m 1/p_j$. For $f_j \in L_{p_j}(X,\mu,\mathbb{K})$, show that $\prod_{j=1}^m f_j$ belongs to $L_r(X,\mu,\mathbb{K})$ and that

$$\Bigl\|\prod_{j=1}^m f_j\Bigr\|_r \leq \prod_{j=1}^m \|f_j\|_{p_j} \ .$$

(Hint: Hölder's inequality.)

15 Suppose X is a metric space. The function $f \in E^X$ **vanishes at infinity** if for every $\varepsilon > 0$ there is a compact subset K of X such that $|f(x)| < \varepsilon$ for all $x \in K^c$. Verify that

$$C_0(X,E) := \{\, f \in C(X,E) \ ; \ f \text{ vanishes at infinity } \}$$

is the closure of $C_c(X,E)$ in $BUC(X,E)$.

16 For $f \in \mathcal{L}_0(X,\mu,E)$, set

$$\lambda_f(t) := \mu\bigl([\,|f| > t]\bigr) \quad \text{and} \quad f^*(t) := \inf\{\, s \geq 0 \ ; \ \lambda_f(s) \leq t \,\} \quad \text{for } t \in [0,\infty) \ .$$

We call $f^* : [0,\infty) \to [0,\infty]$ the **decreasing rearrangement** of f. Prove:

(i) λ_f and f^* are decreasing, continuous from the right, and Lebesgue measurable.

(ii) If $|f| \leq |g|$ for $g \in \mathcal{L}_0(X,\mu,E)$, then $\lambda_f \leq \lambda_g$ and $f^* \leq g^*$.

(iii) If (f_j) is an increasing sequence such that $|f_j| \uparrow |f|$, then $\lambda_{f_j} \uparrow \lambda_f$ and $f_j^* \uparrow f^*$.

(iv) For $p \in (0,\infty)$,

$$\int_X |f|^p \, d\mu = p\int_{\mathbb{R}^+} t^{p-1}\lambda_f(t)\,\lambda_1(dt) = \int_{\mathbb{R}^+} (f^*)^p \, d\lambda_1 \ .$$

(v) $\|f\|_\infty = f^*(0)$.

(vi) $\lambda_f = \lambda_{f^*}$.

(Hint for (iv): Consider first simple functions and then apply (iii) together with Theorems 1.12 and 3.4.)

17 For $j \in \mathbb{N}$, let $I_{j,k} := \bigl[k2^{-j}, (k+1)2^{-j}\bigr]$ for $k = 0,\ldots,2^{j-1}$. Further let $\{\, J_n \ ; \ n \in \mathbb{N} \,\}$ be a relabeling of $\{\, I_{j,k} \ ; \ j \in \mathbb{N}, \ k = 0,\ldots,2^{j-1} \,\}$ and set $f_n := \chi_{J_n}$ for $j \in \mathbb{N}$. Prove that (f_n) is a null sequence in $\mathcal{L}_p\bigl([0,1]\bigr)$ for every $p \in [1,\infty)$, even though $\bigl(f_n(x)\bigr)$ diverges for every $x \in [0,1]$.

18 Suppose (f_k) is a sequence in $L_p(X)$, where $1 \leq p < \infty$. We say that (f_k) **converges weakly** in $L_p(X)$ to $f \in L_p(X)$ if

$$\int_X f_k \varphi \, dx \to \int_X f\varphi \, dx \quad \text{for } \varphi \in L_{p'}(X) \ .$$

In this case, f is called a **weak limit** of (f_k) in $L_p(X)$.

Prove:

(i) Weak limits in $L_p(X)$ are unique.

(ii) Every convergent sequence in $L_p(X)$ converges weakly in $L_p(X)$.

(iii) If (f_k) converges weakly in $L_p(X)$ to f and converges μ-a.e. to $g \in L_p(X)$, then $f = g$.

(iv) If (f_k) converges weakly in $L_2(X)$ to f and $\|f_k\|_2 \to \|f\|_2$, then (f_k) converges in $L_2(X)$ to f.

(v) Let $\mathsf{e}_k(t) := (2\pi)^{-1/2} e^{ikt}$ for $0 < t < 2\pi$ and $k \in \mathbb{N}$. Then the sequence (e_k) converges weakly to 0 in $L_2\big((0,2\pi)\big)$, even though it diverges in $L_2\big((0,2\pi)\big)$.

(Hints: (i) For $f \in L_p(X)$ consider $\varphi(x) := \overline{f(x)}\,|f(x)|^{p/p'-1}$. (ii) Hölder's inequality. (iii) Show that $g \in L_p(X)$, so $[\,|g| = \infty]$ is a μ-null set. If $X_n := \big[\sup_{k \geq n} |f_k(x)| \geq n\big]$ then $\bigcap X_n$ is also a μ-null set. Now consider $\lim \int_{X_n^c} f_n \varphi \, dx$ for $\varphi \in L_{p'}(X)$. (iv) Apply the parallelogram identity in $L_2(X)$. (v) The first statement follows from Bessel's inequality, the second from (ii).)

5 The n-dimensional Bochner–Lebesgue integral

In this short section, we discuss the relationship between the Bochner–Lebesgue integral and the Cauchy–Riemann integral defined in Chapter VI. We show that every jump continuous function is Lebesgue measurable and that the corresponding integrals are equal. This connection will allow us to bring into Lebesgue integration theory the methods we developed for the Cauchy–Riemann integral.

We also show that a bounded scalar-valued function on a compact interval is Riemann integrable if and only if the set of its discontinuities has measure zero. From this it follows that there are Lebesgue integrable functions that are not Riemann integrable. Thus the Lebesgue integral is a proper extension of the Riemann integral — and therefore also of the Cauchy–Riemann integral.

In this entire section, suppose

- $X \subset \mathbb{R}^n$ is a λ_n-measurable set of positive measure;
  $E = (E, |\cdot|)$ is a Banach space.

Lebesgue measure spaces

From Exercise IX.1.7, we know that $\mathcal{L}_X := \mathcal{L}(n) \,|\, X$ is a σ-algebra over X. Thus the restriction $\lambda_n \,|\, X := \lambda_n \,|\, \mathcal{L}_X$ is a measure on X, called ***n*-dimensional Lebesgue measure** (or **Lebesgue *n*-measure**) on X. We denote this restriction by λ_n as well. We check easily that $(X, \mathcal{L}_X, \lambda_n)$ is a complete σ-finite measure space. If there is no danger of misunderstanding, we drop the qualifier "Lebesgue" (or "λ_n") from the words measurable, measure, integrable and so on.

If $f \in E^X$ is integrable, we call

$$\int_X f\, d\lambda_n := \int_X f\, d(\lambda_n \,|\, X) = \int_{\mathbb{R}^n} \widetilde{f} \chi_X \, d\lambda_n$$

the (***n*-dimensional**) (**Bochner–Lebesgue**) **integral** of f over X. The notations

$$\int_X f(x)\, d\lambda_n(x) \quad \text{and} \quad \int_X f(x)\, \lambda_n(dx)$$

are also common.

For short, we set

$$\mathcal{L}_p(X, E) := \mathcal{L}_p(X, \lambda_n, E) \text{ and } L_p(X, E) := L_p(X, \lambda_n, E) \;.$$

We also set $\mathcal{L}_p(X) := \mathcal{L}_p(X, \mathbb{K})$ and $L_p(X) := L_p(X, \mathbb{K})$ for $p \in [1, \infty] \cup \{0\}$.

The next theorem lists important properties of n-dimensional integrals.

5.1 Theorem *Suppose X is open in $\mathbb{R}^n$ or, in the case $n = 1$, a perfect interval. Then*:

(i) *λ_n is a massive Radon measure on X.*

(ii) *$C(X, E)$ is a vector subspace of $L_0(X, E)$.*

(iii) *$BC(X, E)$ is a closed vector subspace of $L_\infty(X, E)$.*

(iv) *$C_c(X, E)$ is a dense vector subspace of $L_p(X, E)$ for $p \in [1, \infty)$. If K is a compact subset of X, then*

$$\|f\|_p \le \lambda_n(K)^{1/p} \, \|f\|_\infty \quad \textit{for } f \in C_c(X, E) \ \textit{ such that } \ \operatorname{supp}(f) \subset K \ .$$

(v) *If X has finite measure and $1 \le p < q \le \infty$, then*

$$L_q(X, E) \overset{d}{\hookrightarrow} L_p(X, E)$$

and

$$\|f\|_p \le \lambda_n(X)^{1/p-1/q} \, \|f\|_q \quad \textit{for } f \in L_q(X, E) \ .$$

Proof (i) X is a σ-compact metric space — by Remark 1.16(e) if X is open, and for obvious reasons if X is an interval. Now the claim follows from Remark 1.16(h) and Exercise IX.5.21.

(ii) and (iii) are covered respectively by Proposition 4.17 and Theorem 4.18(ii).

(iv) The first statement is a consequence of Theorem 4.18(i), and the second is obvious.

(v) is a special case of Theorem 4.16. ■

5.2 Remark Suppose X is measurable and its boundary ∂X is a λ_n-null set. Then the Borel set $\mathring{X}$ belongs to $\mathcal{L}(n)$, and we have $\lambda_n(\mathring{X}) = \lambda_n(X)$. Further, one checks easily that the map

$$L_p(X, E) \to L_p(\mathring{X}, E) \ , \quad [f] \mapsto [f \,|\, \mathring{X}]$$

is a vector space isomorphism for $p \in [1, \infty] \cup \{0\}$. If $p \in [1, \infty]$, it is an isometry. Thus we can identify $L_p(X, E)$ and $L_p(\mathring{X}, E)$ for $p \in [1, \infty] \cup \{0\}$. In particular, for an interval X in $\mathbb{R}$ with endpoints $a := \inf X$ and $b := \sup X$, we have

$$L_p(X, E) = L_p([a, b], E) = L_p([a, b), E) = L_p((a, b], E) = L_p((a, b), E)$$

for $p \in [1, \infty] \cup \{0\}$.

The Lebesgue integral of absolutely integrable functions

We now show that every absolutely integrable function is Lebesgue integrable, and its integral in the sense of Section VI.8 equals the Lebesgue integral.

5.3 Theorem *Suppose $f:(a,b)\to E$ is absolutely integrable, where $a,b\in\overline{\mathbb{R}}$ and $a<b$. Then f belongs to $\mathcal{L}_1\big((a,b),E\big)$, and*

$$\int_{(a,b)} f\,d\lambda_1 = \int_a^b f\ .$$

Proof (i) Suppose $a<\alpha<\beta<b$. If $g:[\alpha,\beta]\to E$ is a staircase function, then g is obviously λ_1-simple and

$$\int_{(\alpha,\beta)} g\,d\lambda_1 = \int_\alpha^\beta g\ . \tag{5.1}$$

Now suppose $g:[\alpha,\beta]\to E$ is jump continuous. Then there is a sequence (g_j) of staircase functions that converges uniformly to g. Therefore g is measurable, and Remark VI.1.1(d) and Corollary 3.15(ii) show that g belongs to $\mathcal{L}_1\big((\alpha,\beta),E\big)$. Because g is bounded and the sequence (g_j) converges uniformly, there is an $M\geq 0$ such that $|g_j|\leq M$ for all $j\in\mathbb{N}$. Therefore it follows from Lebesgue's dominated convergence theorem that

$$\lim_{j\to\infty}\int_{(\alpha,\beta)} g_j\,d\lambda_1 = \int_{(\alpha,\beta)} g\,d\lambda_1$$

in E, and we conclude using (5.1) and the definition of the Cauchy–Riemann integral that

$$\int_\alpha^\beta g = \lim_{j\to\infty}\int_\alpha^\beta g_j = \lim_{j\to\infty}\int_{(\alpha,\beta)} g_j\,d\lambda_1 = \int_{(\alpha,\beta)} g\,d\lambda_1\ .$$

(ii) We fix $c\in(a,b)$ and choose a sequence (β_j) in (c,b) such that $\beta_j\uparrow b$. We also set[1]

$$g:=\chi_{[c,b)}f\quad,\quad g_j:=\chi_{[c,\beta_j]}f\quad\text{for } j\in\mathbb{N}\ .$$

By (i), (g_j) is a sequence in $\mathcal{L}_1(\mathbb{R},E)$. Obviously (g_j) converges pointwise to g and $(|g_j|)$ is an increasing sequence converging to $|g|$. Therefore g is measurable. From (i), it follows that

$$\int_{\mathbb{R}} |g_j|\,d\lambda_1 = \int_{(c,\beta_j)} |f|\,d\lambda_1 = \int_c^{\beta_j} |f|\ ,$$

and the absolute convergence of $\int_c^b f$ implies

$$\lim_{j\to\infty}\int_{\mathbb{R}} |g_j|\,d\lambda_1 = \lim_{j\to\infty}\int_c^{\beta_j}|f| = \int_c^b |f|\ . \tag{5.2}$$

[1]Here and in similar situations, we regard $\chi_{[c,b)}f$ as a function on $\mathbb{R}$. Writing $\chi_{[c,b)}\widetilde{f}$ would be more precise but cumbersome.

On the other hand, the monotone convergence theorem shows that

$$\int_{\mathbb{R}} |g|\, d\lambda_1 = \lim_{j\to\infty} \int_{\mathbb{R}} |g_j|\, d\lambda_1 \ ,$$

and we see from (5.2) that g belongs to $\mathcal{L}_1(\mathbb{R}, E)$. Therefore we can apply the dominated convergence theorem to the sequence (g_j), to get

$$\lim_{j\to\infty} \int_{\mathbb{R}} g_j\, d\lambda_1 = \int_{\mathbb{R}} g\, d\lambda_1 = \int_{[c,b)} f\, d\lambda_1$$

in E. Further, it follows from (i) that

$$\int_{\mathbb{R}} g_j\, d\lambda_1 = \int_{[c,\beta_j)} f\, d\lambda_1 = \int_c^{\beta_j} f \ ,$$

and hence, by Proposition VI.8.7,

$$\lim_{j\to\infty} \int_{\mathbb{R}} g_j\, d\lambda_1 = \lim_{j\to\infty} \int_c^{\beta_j} f = \int_c^b f$$

in E. Thus the limits $\int_{[c,b)} f\, d\lambda_1$ and $\int_c^b f$ are equal. In similar fashion, we show that $\chi_{(a,c]} f$ belongs to $\mathcal{L}_1(\mathbb{R}, E)$ and that $\int_{(a,c]} f\, d\lambda_1 = \int_a^c f$. This shows that f is Lebesgue integrable with $\int_{(a,b)} f\, d\lambda_1 = \int_a^b f$. ■

5.4 Corollary *For $-\infty < a < b < \infty$, we have $\mathcal{S}\big([a,b], E\big) \hookrightarrow \mathcal{L}_1\big([a,b], E\big)$ and*

$$\int_{[a,b]} f\, d\lambda_1 = \int_a^b f \quad \textit{for } f \in \mathcal{S}\big([a,b], E\big) \ .$$

Proof This follows from Theorem 5.3 and Proposition VI.8.3. ■

5.5 Remarks Fix $a, b \in \overline{\mathbb{R}}$ with $a < b$.

(a) Suppose $f : (a,b) \to E$ is admissible and $\int_a^b f$ exists as an improper integral. Then f need not belong to $\mathcal{L}_1\big((a,b), E\big)$.

Proof We define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) := \begin{cases} 0 & \text{if } x \in (-\infty, 0) \ , \\ (-1)^j/j & \text{if } x \in [j-1, j), \ \text{where } j \in \mathbb{N}^\times \ . \end{cases}$$

Obviously f is admissible, and $\int_{-\infty}^{\infty} f$ exists in $\mathbb{R}$, since

$$\int_{-\infty}^{\infty} f = \sum_{j=1}^{\infty} (-1)^j/j \ .$$

If f belonged to $\mathcal{L}_1(\mathbb{R})$, we would have $\int_{\mathbb{R}} |f|\,d\lambda_1 < \infty$, contradicting the monotone convergence theorem, which gives

$$\int_{\mathbb{R}} |f|\,d\lambda_1 = \lim_{k\to\infty} \int_{\mathbb{R}} \chi_{[0,k]}\,|f|\,d\lambda_1 = \lim_{k\to\infty} \sum_{j=1}^{k} 1/j = \infty\ . \ \blacksquare$$

(b) Suppose $f\colon (a,b) \to E$ is admissible and f belongs to $\mathcal{L}_1\bigl((a,b),E\bigr)$. Then f is absolutely integrable, and

$$\int_{(a,b)} f\,d\lambda_1 = \int_a^b f \qquad \text{in } E\ .$$

Proof Take $c \in (a,b)$ and let (α_j) be a sequence in (a,c) with $\alpha_j \to a$. Also let $f_j := \chi_{[\alpha_j,c]} f$. Then (f_j) converges pointwise to $\chi_{(a,c]} f$, and we have $|f_j| \le |f|$ for $j \in \mathbb{N}$. Because f is admissible, Proposition VI.4.3 shows that $|f| \,\big|\, [\alpha_j,c]$ belongs to $\mathcal{S}\bigl([\alpha_j,c],\mathbb{R}\bigr)$. Thus it follows from Corollary 5.4 and the dominated convergence theorem that

$$\int_{\alpha_j}^{c} |f| = \int_{[\alpha_j,c]} |f|\,d\lambda_1 \to \int_{(a,c]} |f|\,d\lambda_1\ .$$

Therefore $\int_a^c |f|$ exists. Analogously, we show the existence of $\int_c^b |f|$ and thus the absolute convergence of $\int_a^b f$. The second statement now follows from Theorem 5.3. ■

Suppose $f \in \mathcal{L}_1((a,b),E)$. Remark 5.5(b) shows that no misunderstanding should arise in this case if we denote $\int_{(a,b)} f\,d\lambda_1$ by $\int_a^b f$ or $\int_a^b f(x)\,dx$. From now on, we will usually write in the n-dimensional case

$$\int_X f\,dx := \int_X f\,d\lambda_n\ .$$

Theorem 5.3 and its corollary allow us to transfer the integration methods developed in Volume II to the framework of Lebesgue theory. In combination with the integrability criterion of Theorem 3.14 and the dominated convergence theorem, these provide very effective tools for proving the existence of integrals. This will be made clear in the remaining sections of this chapter, when we develop procedures for the concrete evaluation of "multidimensional" integrals.

A characterization of Riemann integrable functions

Theorem 5.3 showed that the Lebesgue integral is an extension of the Cauchy–Riemann integral. We now characterize Riemann integrable functions and show that this extension is proper.

5.6 Theorem *Let I be a compact interval, and let $f\colon I \to \mathbb{K}$ be bounded. Then f is Riemann integrable if and only if it is continuous λ_1-a.e. In this case, f is Lebesgue integrable, and the Riemann and Lebesgue integrals are equal.*

Proof (i) We can take without loss of generality the case $\mathbb{K} = \mathbb{R}$ and $I := [0,1]$. For $k \in \mathbb{N}$, let $\mathfrak{Z}_k := (\xi_{0,k}, \ldots, \xi_{2^k,k})$ be the partition of $[0,1]$ with $\xi_{j,k} := j\,2^{-k}$ for $j = 0, \ldots, 2^k$. Also suppose

$$I_{0,k} := [\xi_{0,k}, \xi_{1,k}] \quad , \quad I_{j,k} := (\xi_{j,k}, \xi_{j+1,k}] \quad \text{for } j = 1, \ldots, 2^k - 1 \ .$$

Finally, set $\alpha_{j,k} := \inf_{x \in \bar{I}_{j,k}} f(x)$, $\beta_{j,k} := \sup_{x \in \bar{I}_{j,k}} f(x)$, and

$$g_k := \sum_{j=0}^{2^k-1} \alpha_{j,k} \chi_{I_{j,k}} \ , \quad h_k := \sum_{j=0}^{2^k-1} \beta_{j,k} \chi_{I_{j,k}} \quad \text{for } k \in \mathbb{N} \ .$$

Then (g_k) is an increasing and (h_k) a decreasing sequence of λ_1-simple functions. Therefore their pointwise limits $g := \lim_k g_k$ and $h := \lim_k h_k$ are defined and λ_1-measurable, and $g \le f \le h$. Furthermore, we have

$$\int_{[0,1]} g_k \, d\lambda_1 = \underline{S}(f,k) \text{ and } \int_{[0,1]} h_k \, d\lambda_1 = \overline{S}(f,k) \ ,$$

where $\underline{S}(f,k)$ and $\overline{S}(f,k)$ stand for the lower and upper sums of f on $[0,1]$ with respect to the partition $\mathfrak{Z}_k$ (see Exercise VI.3.7). Denoting by $\underline{\int} f$ and $\overline{\int} f$ the lower and upper Riemann integrals of f on $[0,1]$, we find from the monotone convergence theorem that

$$\int_{[0,1]} (h - g) \, d\lambda_1 = \overline{\int} f - \underline{\int} f \ . \tag{5.3}$$

(ii) Let $R := \bigcup_{k \in \mathbb{N}} \{\xi_{0,k}, \ldots, \xi_{2^k,k}\}$ be the set of endpoints of the intervals $I_{j,k}$. Let C be the set of continuous points of f. Then

$$[g = h] \cap R^c \subset C \subset [g = h] \ . \tag{5.4}$$

To see this, take $\varepsilon > 0$. Suppose first that $x_0 \in R^c$ and $g(x_0) = h(x_0)$. We can find a $k \in \mathbb{N}$ such that $h_k(x_0) - g_k(x_0) < \varepsilon$ and a $j \in \{0, \ldots, 2^k - 1\}$ such that x_0 lies in the interval $(\xi_{j,k}, \xi_{j+1,k})$. For $x \in I_{j,k}$, we thus have

$$|f(x) - f(x_0)| \le \sup_{y \in \bar{I}_{j,k}} f(y) - \inf_{y \in \bar{I}_{j,k}} f(y) = h_k(x_0) - g_k(x_0) < \varepsilon \ ,$$

which proves the continuity of f at x_0.

Now suppose $x_0 \in C$. Take $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon/2$ for $x \in [x_0 - \delta, x_0 + \delta] \cap [0,1]$. Choose $k_0 \in \mathbb{N}$ with $2^{-k_0} \le \delta$ and take for every $k \ge k_0$ a $j \in \{0, \ldots, 2^k - 1\}$ such that $x_0 \in I_{j,k} \subset [x_0 - \delta, x_0 + \delta]$. Then

$$0 \le h_k(x_0) - g_k(x_0) = \sup_{x \in \bar{I}_{j,k}} \bigl(f(x) - f(x_0)\bigr) - \inf_{x \in \bar{I}_{j,k}} \bigl(f(x) - f(x_0)\bigr) < \varepsilon \ .$$

It follows that $h(x_0) - g(x_0) = \lim_k \bigl(h_k(x_0) - g_k(x_0)\bigr) = 0$. This proves (5.4).

(iii) If f is a Riemann integrable function, then $\underline{\int} f = \overline{\int} f = \int f$ (Exercise VI.3.10). Therefore (5.3) shows that

$$h = g = f \qquad \lambda_1\text{-a.e. }, \tag{5.5}$$

which implies the λ_1-measurability of f. Since f is bounded, $f \in \mathcal{L}_1\bigl([0,1]\bigr)$. We also have $|g_k| \le \|f\|_\infty$ λ_1-a.e. for $k \in \mathbb{N}$. Then Lebesgue's dominated convergence theorem results in

$$\int_{[0,1]} g\,d\lambda_1 = \lim_k \int_{[0,1]} g_k\,d\lambda_1 = \lim_k \underline{S}(f,k) = \int_0^1 f\ ,$$

where, in the last equality, we have once more used Exercise VI.3.10. From (5.5) and Lemma 2.15, it follows that $\int_{[0,1]} f\,d\lambda_1 = \int_0^1 f$. Finally (5.4), (5.5), and the countability of R imply that the discontinuous points of f form a set of Lebesgue measure zero.

(iv) Suppose conversely that C^c has measure zero. By (5.4), so does $[g \ne h]$, and the Riemann integrability of f follows from (5.3). This finishes the proof. ■

5.7 Corollary *Some Lebesgue integrable functions are not Riemann integrable. Thus the Lebesgue integral is a proper extension of the Riemann integral.*

Proof Consider the Dirichlet function

$$f\colon [0,1] \to \mathbb{R}\ , \quad f(x) := \begin{cases} 1 & \text{if } x \in \mathbb{Q}\ , \\ 0 & \text{if } x \notin \mathbb{Q}\ , \end{cases}$$

on $[0,1]$. By Lemma 2.15, f belongs to $\mathcal{L}_1\bigl([0,1]\bigr)$, since f vanishes almost everywhere. But we know from Example III.1.3(c) that f is nowhere continuous, hence not Riemann integrable by Theorem 5.6. ■

The equivalence class of maps that agree a.e. with the Dirichlet function contains Riemann integrable functions — for example, the null function. So this example is uninteresting from the viewpoint of L_1-spaces. However, in Exercise 13, it will be shown that there exists $f \in \mathcal{L}_1\bigl([0,1],\mathbb{R}\bigr)$ such that no $g \in [f]$ is Riemann integrable. This implies that the Riemann integral is inadequate for the theory of L_p-spaces.

Exercises

1 For $p, q \in [1,\infty]$ with $p \ne q$, show that $L_p(\mathbb{R},E) \not\subset L_q(\mathbb{R},E)$.

2 Suppose J is an open interval and $f \in C^1(J,E)$ has compact support. Then $\int_J f' = 0$.

3 Suppose $f \in \mathcal{L}_0\bigl([0,1],\mathbb{R}^+\bigr)$ is bounded. Show that

$$\underline{\int} f \le \int_{[0,1]} f\,d\lambda_1 \le \overline{\int} f\ .$$

4 Suppose I is a compact interval, and define the **space of functions of bounded variation** on I by

$$BV(I,E) := \{\, f : I \to E \ ;\ \mathrm{Var}(f,I) < \infty \,\} \ .$$

(a) In the sense of vector subspaces, we have the inclusions

$$C^{1\text{-}}(I,E) \subset BV(I,E) \subset B(I,E) \ .$$

(b) Let $\alpha := \inf I$ and $f \in \mathcal{L}_1(I,E)$. Then $F : I \to E,\ \ x \mapsto \int_\alpha^x f(t)\,dt$ belongs to $BV(I,E)$, and $\mathrm{Var}(F,I) \le \|f\|_1$.

(c) For every $f \in BV(I,\mathbb{R})$, there are increasing maps $s^\pm : I \to \mathbb{R}$ such that $f = s^+ - s^-$.

(d) $BV(I,\mathbb{R})$ is a vector subspace of the space $\mathcal{S}(I,\mathbb{R})$ of jump continuous functions $I \to \mathbb{R}$.

(e) Every monotone function belongs to $BV(I,\mathbb{R})$.

(Hint for (c): For $\alpha := \inf I$, consider the functions $s^+ := \bigl(x \mapsto \mathrm{Var}\bigl(f^+, [\alpha,x]\bigr)\bigr)$ and $s^- := s - f$.)

5 Suppose H is a separable Hilbert space. Show[2] that $BV\bigl([a,b],H\bigr)$ is a vector subspace of $\mathcal{L}_\infty\bigl([a,b],H\bigr)$ and that

$$\int_a^{b-h} \|f(t+h) - f(t)\|\,dt \le h\,\mathrm{Var}\bigl(f,[a,b]\bigr) \quad \text{for } 0 < h < b-a \ .$$

(Hints: Note Exercises 1.1 and 4(d). For $0 < h < b-a$ and $t \in [a,b-h]$, show that $\|f(t+h) - f(t)\| \le \mathrm{Var}\bigl(f,[a,t+h]\bigr) - \mathrm{Var}\bigl(f,[a,t]\bigr)$.)

6 Suppose $J \subset \mathbb{R}$ is a perfect interval. A function $f : J \to E$ is **absolutely continuous** if for every $\varepsilon > 0$ there is $\delta > 0$ such that

$$\sum_{k=0}^m |f(\beta_k) - f(\alpha_k)| < \varepsilon$$

for every finite family $\{\,(\alpha_k,\beta_k)\ ;\ k = 0,\dots,m\,\}$ of pairwise disjoint subintervals of J with $\sum_{k=0}^m (\beta_k - \alpha_k) < \delta$. We denote by $W_1^1(J,E)$ the set of all absolutely continuous functions in E^J. Prove:

(a) In the sense of vector subspaces, we have the inclusions

$$BC^1(J,E) \subset W_1^1(J,E) \subset C(J,E) \ .$$

(b) If J compact, then $W_1^1(J,E) \subset BV(J,E)$.

(c) The Cantor function (Exercise III.3.8) is continuous but not absolutely continuous.

(d) Set $\alpha := \inf J$ and take $f \in \mathcal{L}_1(J,E)$. Then $F : J \to E,\ x \mapsto \int_\alpha^x f(t)\,dt$ is absolutely continuous.

[2] One can show that the statement of Exercise 5 remains true if H is replaced by an arbitrary Banach space.

7 For $j = 1, 2$, define $f_j : [0,1] \to \mathbb{R}$ by

$$f_j(x) := \begin{cases} x^2 \sin(1/x^j) & \text{if } x \in (0,1] \ , \\ 0 & \text{if } x = 0 \ ; \end{cases}$$

compare Exercise IV.1.2. Prove:

(a) $f_1 \in BV\bigl([0,1],\mathbb{R}\bigr)$.

(b) $f_2 \notin BV\bigl([0,1],\mathbb{R}\bigr)$.

8 Let μ and ν be measures on a measurable space $(X,\mathcal{A})$. We say ν is **μ-absolutely continuous** if every μ-null set is also a ν-null set. In this case, we write $\nu \ll \mu$.

(a) Let $(X,\mathcal{A},\mu)$ be a σ-finite complete measure space. For $f \in \mathcal{L}_0(X,\mu,\overline{\mathbb{R}}^+)$, define

$$f \bullet \mu : \mathcal{A} \to [0,\infty] \ , \quad A \mapsto \int_A f \, d\mu \ .$$

Show that $f \bullet \mu$ is a complete measure on $(X,\mathcal{A})$ with $f \bullet \mu \ll \mu$.

(b) Let $\mathcal{A} := \mathcal{L}_{[0,1]}$, $\nu := \lambda_1$, and $\mu := \mathcal{H}^0$. Check:

(i) $\nu \ll \mu$.

(ii) there is no $f \in \mathcal{L}_0\bigl([0,1],\mathcal{A},\mu\bigr)$ such that $\nu = f \bullet \mu$.

9 Suppose $(X,\mathcal{A},\nu)$ is a finite measure space and μ is measure on $(X,\mathcal{A})$. The following statements are equivalent:

(i) $\nu \ll \mu$.

(ii) For every $\varepsilon > 0$ there is $\delta > 0$ such that $\nu(A) < \varepsilon$ for all $A \in \mathcal{A}$ with $\mu(A) < \delta$.

10 For $f \in \mathcal{L}_0(\mathbb{R},\lambda_1,\overline{\mathbb{R}}^+)$, let $F(x) := \int_{-\infty}^x f(t)\,dt$ for $x \in \mathbb{R}$, and denote by μ_F the Lebesgue–Stieltjes measure on $\mathbb{R}$ generated by F. Prove:

(a) $F \in W_1^1(\mathbb{R},\mathbb{R})$ implies $\mu_F \ll \lambda_1$.

(b) $\mu_F \ll \beta_1$ implies $F \in W_1^1(\mathbb{R},\mathbb{R})$ if μ_F is finite.

11 Let I is an interval and take $f \in \mathcal{L}_1(I,\mathbb{R}^n)$. For a fixed $a \in I$, suppose $\int_a^x f(t)\,dt = 0$ for $x \in I$. Show that $f(x) = 0$ for almost every $x \in I$.

12 Let $0 \le a < b < \infty$ and $I := (-b,-a) \cup (a,b)$, and suppose $f \in \mathcal{L}_1(I,E)$. Show that $\int_I f\,dx = 0$ if f is odd, and $\int_I f\,dx = 2\int_a^b f\,dx$ if f is even.

13 Define

$$\begin{aligned} K_0 &:= [0,1] \ , \\ K_1 &:= K_0 \backslash (3/8, 5/8) \ , \\ K_2 &:= K_1 \backslash \bigl((5/32, 7/32) \cup (25/32, 27/32)\bigr) \ , \ \ldots \end{aligned}$$

Generally, K_{n+1} is derived from K_n by the removal of open "middle fourths" of length $(1/4)^{n+1}$, rather than middle thirds as in the construction of the traditional Cantor set (Exercise III.3.8). Set $K := \bigcap K_n$ and $f := \chi_K$. Show that f belongs to $\mathcal{L}_1\bigl([0,1]\bigr)$ and that no $g \in [f]$ is Riemann integrable.

6 Fubini's theorem

The heart of this section is the proof that the Lebesgue integral of functions of multiple variables can be calculated iteratively and that this sequence of one-dimensional integrations can be performed in any order. Therefore multivariable integration reduces to integrating functions of only one variable. With the results of the previous section and the procedures developed in Volume II, multidimensional integrals can be calculated explicitly in many cases.

The method of iterative evaluation of integrals has wide-reaching theoretical applications, a few of which we will present.

Throughout this section, we suppose

- m, n are positive integers and E is a Banach space.

In addition, we will generally identify $\mathbb{R}^{m+n}$ with $\mathbb{R}^m \times \mathbb{R}^n$.

Maps defined almost everywhere

Suppose $(X, \mathcal{A}, \mu)$ is a measure space. We will often consider nonnegative $\overline{\mathbb{R}}$-valued functions that are only defined μ-a.e. For these, we shall simply write $x \mapsto f(x)$, without specifying the precise domain of definition. We say such a function $x \mapsto f(x)$ is **measurable** if there is a μ-null set N such that $f \,|\, N^c : N^c \to \overline{\mathbb{R}}^+$ is defined and μ-measurable. Therefore $\int_{N^c} f\,d\mu$ is defined. If M is another μ-null set such that $f \,|\, M^c : M^c \to \overline{\mathbb{R}}^+$ is defined and μ-measurable, the equalities $\mu(N) = \mu(M) = \mu(M \cup N) = 0$ and Remarks 3.3(a) and (b) imply that

$$\int_{N^c} f\,d\mu = \int_{M^c \cap N^c} f\,d\mu = \int_{M^c} f\,d\mu \ .$$

Therefore

$$\int_X f\,d\mu := \int_{N^c} f\,d\mu \tag{6.1}$$

is well defined and independent of the chosen null set N.

For an E-valued function $x \mapsto f(x)$ defined μ-a.e., we define measurability just as above. We say such an f is **integrable** if $f \,|\, N^c$ belongs to $\mathcal{L}_1(N^c, \mu, E)$. In this case, $\int_X f\,d\mu$ is also defined through (6.1), and Lemma 2.15 shows this definition is meaningful.

Consider for example $A \in \mathcal{L}(m+n)$, and assume that the cross section $A_{[x]}$ is λ_n-measurable for λ_m-almost every $x \in \mathbb{R}^m$. Then $x \mapsto \lambda_n(A_{[x]})$ is a nonnegative $\overline{\mathbb{R}}$-valued function defined λ_m-a.e. If $x \mapsto \lambda_n(A_{[x]})$ is measurable, the integral $\int_{\mathbb{R}^m} \lambda_n(A_{[x]})\,dx$ is well defined.

Cavalieri's principle

We denote by $\mathcal{C}(m,n)$ the set of all $A \in \mathcal{L}(m+n)$ for which

(i) $A_{[x]} \in \mathcal{L}(n)$ for λ_m-almost every $x \in \mathbb{R}^m$;

(ii) $x \mapsto \lambda_n(A_{[x]})$ is λ_m-measurable;

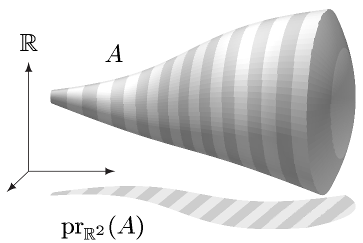
(iii) $\lambda_{m+n}(A) = \int_{\mathbb{R}^m} \lambda_n(A_{[x]})\,dx$.

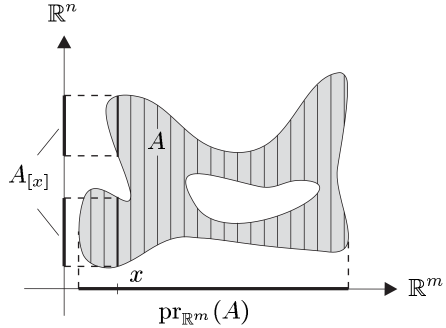
We want to show that $\mathcal{C}(m,n)$ agrees with $\mathcal{L}(m+n)$, but we need some preliminaries.

6.1 Remarks **(a)** Suppose $A \in \mathcal{C}(1,n)$ is bounded and $\mathrm{pr}_1(A)$ is an interval with endpoints a and b. Then

$$\lambda_{n+1}(A) = \int_a^b \lambda_n(A_{[x]})\,dx \ .$$

This statement is called **Cavalieri's principle** and makes precise the geometric idea that the measure (volume) of A can be determined by partitioning A into thin parallel slices and continuously summing (integrating) the volumes of these slices.





(b) $\mathcal{L}(m) \boxtimes \mathcal{L}(n) \subset \mathcal{C}(m,n)$.

(c) For every ascending sequence (A_j) in $\mathcal{C}(m,n)$, the union $\bigcup_j A_j$ belongs to $\mathcal{C}(m,n)$.

Proof (i) For $j \in \mathbb{N}$, let M_j be a λ_m-null set such that $A_{j,[x]} := (A_j)_{[x]} \in \mathcal{L}(n)$ for $x \in M_j^c$. Letting $A := \bigcup_j A_j$ and $M := \bigcup_j M_j$, we then have $A_{[x]} = \bigcup_j A_{j,[x]} \in \mathcal{L}(n)$ for $x \in M^c$. The continuity of λ_n from below implies $\lambda_n(A_{[x]}) = \lim_j \lambda_n(A_{j,[x]})$ for $x \in M^c$, and we conclude with the help of Proposition 1.11 that $x \mapsto \lambda_n(A_{[x]})$ is λ_m-measurable.

(ii) Because $A_j \in \mathcal{C}(m,n)$, we have

$$\int_{\mathbb{R}^m} \lambda_n(A_{j,[x]})\,dx = \lambda_{m+n}(A_j) \quad \text{for } j \in \mathbb{N} \ ,$$

and from the monotone convergence theorem, it follows that

$$\lim_j \int_{\mathbb{R}^m} \lambda_n(A_{j,[x]})\,dx = \int_{\mathbb{R}^m} \lambda_n(A_{[x]})\,dx \ . \tag{6.2}$$

The continuity of λ_{m+n} from below thus shows that

$$\lambda_{m+n}(A) = \lim_j \lambda_{m+n}(A_j) = \lim_j \int_{\mathbb{R}^m} \lambda_n(A_{j,[x]})\,dx = \int_{\mathbb{R}^m} \lambda_n(A_{[x]})\,dx\ .$$

Therefore A belongs to $\mathcal{C}(m,n)$. ■

(d) Suppose (A_j) is a descending sequence in $\mathcal{C}(m,n)$ and there is a $k \in \mathbb{N}$ such that $\lambda_{m+n}(A_k) < \infty$. Then $\bigcap_j A_j$ belongs to $\mathcal{C}(m,n)$.

Proof We set $A := \bigcap_j A_j$. The measurability of λ_m-almost all cross sections $A_{[x]}$ and of $x \mapsto \lambda_n(A_{[x]})$ follow as in (c). Next, Lebesgue's dominated convergence theorem shows that (6.2) is true in this case. The claim now follows as in (c). ■

(e) Suppose (A_j) is a disjoint sequence in $\mathcal{C}(m,n)$. Then $\bigcup_j A_j$ also belongs to $\mathcal{C}(m,n)$.

Proof Because of (c), it suffices to prove the statement for finite disjoint sequences. We leave this to the reader as an exercise. ■

(f) Every open set in $\mathbb{R}^{m+n}$ belongs to $\mathcal{C}(m,n)$.

Proof This follows from Proposition IX.5.6, (e) and (b). ■

(g) Every bounded G_δ-set in $\mathbb{R}^{m+n}$ belongs to $\mathcal{C}(m,n)$.

Proof This follows from (f) and (d). ■

(h) Suppose A is a λ_{m+n}-null set. Then A belongs to $\mathcal{C}(m,n)$, and there is a λ_m-null set M such that $A_{[x]}$ is a λ_n-null set for every $x \in M^c$.

Proof It suffices to verify there is a λ_m-null set M such that $\lambda_n(A_{[x]}) = 0$ for $x \in M^c$. So let $A_j := A \cap (j\mathbb{B}^{m+n})$ for $j \in \mathbb{N}$. Then (A_j) is an ascending sequence of bounded λ_{m+n}-null sets with $\bigcup_j A_j = A$. By Corollary IX.5.5, there is a sequence (G_j) of bounded G_δ-sets with $G_j \supset A_j$ and $\lambda_{m+n}(G_j) = 0$ for $j \in \mathbb{N}$. From (g), it therefore follows that

$$0 = \lambda_{m+n}(G_j) = \int_{\mathbb{R}^m} \lambda_n(G_{j,[x]})\,dx\ .$$

Hence, there is for every $j \in \mathbb{N}$ a λ_m-null set M_j such that $\lambda_n(G_{j,[x]}) = 0$ for $x \in M_j^c$ (see Remark 3.3(c)). Because

$$\bigcup_j G_{j,[x]} \supset \bigcup_j A_{j,[x]} = \Bigl(\bigcup_j A_j\Bigr)_{[x]} = A_{[x]} \quad \text{for } x \in \mathbb{R}^m\ ,$$

$M := \bigcup_j M_j$ has the desired property. ■

After these remarks, we can now show the equality of $\mathcal{C}(m,n)$ and $\mathcal{L}(m+n)$.

6.2 Proposition $\mathcal{C}(m,n) = \mathcal{L}(m+n)$.

Proof We need only check the inclusion $\mathcal{L}(m+n) \subset \mathcal{C}(m,n)$.

(i) Suppose $A \in \mathcal{L}(m+n)$ is bounded. By Corollary IX.5.5, there is a bounded G_δ-set G such that $G \supset A$ and $\lambda_{m+n}(G) = \lambda_{m+n}(A)$. Because A has finite measure, $G\backslash A$ is a bounded λ_{m+n}-null set by Proposition IX.2.3(ii), and we conclude using Remark 6.1(h) that $(G\backslash A)_{[x]} = G_{[x]}\backslash A_{[x]}$ is a λ_n-null set for λ_m-almost every $x \in \mathbb{R}^m$. By Remark 6.1(g), $G_{[x]}$ belongs to $\mathcal{L}(n)$ for λ_m-almost every $x \in \mathbb{R}^m$. Because

$$A_{[x]} = G_{[x]} \cap (G_{[x]}\backslash A_{[x]})^c \quad \text{for } x \in \mathbb{R}^m \ ,$$

this is also true of λ_m-almost every intersection $A_{[x]}$. In addition, $\lambda_n(A_{[x]}) = \lambda_n(G_{[x]})$ for λ_m-almost every $x \in \mathbb{R}^m$. We know by Remark 6.1(g) that G belongs to $\mathcal{C}(m,n)$. Therefore $x \mapsto \lambda_n(A_{[x]})$ is measurable, and

$$\lambda_{m+n}(G) = \int_{\mathbb{R}^m} \lambda_n(G_{[x]})\, dx = \int_{\mathbb{R}^m} \lambda_n(A_{[x]})\, dx \ .$$

Therefore A belongs to $\mathcal{C}(m,n)$.

(ii) If A is not bounded, we set $A_j := A \cap (j\mathbb{B}^{m+n})$ for $j \in \mathbb{N}$. Then (A_j) is an ascending sequence in $\mathcal{L}(m+n)$ with $\bigcup_j A_j = A$. The claim now follows from (i) and Remark 6.1(c). ■

6.3 Corollary *If $A \in \mathcal{L}(m+n)$ has finite measure, then $\lambda_n(A_{[x]}) < \infty$ for λ_m-a.e. $x \in \mathbb{R}^m$.*

Proof Because Proposition 6.2 implies

$$\int_{\mathbb{R}^m} \lambda_n(A_{[x]})\, dx = \lambda_{m+n}(A) < \infty \ ,$$

the claim follows from Remark 3.11(c). ■

For $A \in \mathcal{L}(m+n)$ and $x \in \mathbb{R}^m$, we have $\chi_A(x,\cdot\,) = \chi_{A_{[x]}}$, so Proposition 6.2 can also be formulated in terms of characteristic functions. It is then easy to apply the statement to linear combinations of characteristic functions and therefore to simple functions.

6.4 Lemma *Suppose $f \in \mathcal{S}(\mathbb{R}^{m+n}, E)$.*

(i) *$f(x,\cdot\,) \in \mathcal{S}(\mathbb{R}^n, E)$ for λ_m-almost every $x \in \mathbb{R}^m$.*

(ii) *the E-valued function $x \mapsto \int_{\mathbb{R}^n} f(x,y)\, dy$ is λ_m-integrable.*

(iii) *$\int_{\mathbb{R}^{m+n}} f\, d(x,y) = \int_{\mathbb{R}^m} \left[\int_{\mathbb{R}^n} f(x,y)\, dy\right] dx$.*

Proof (i) With $f = \sum_{j=0}^k e_j \chi_{A_j}$, we have $f(x,\cdot\,) = \sum_{j=0}^k e_j \chi_{A_{j,[x]}}$ for $x \in \mathbb{R}^m$. Then it follows easily from Proposition 6.2 and Corollary 6.3 that there is a λ_m-null set M such that $f(x,\cdot\,)$ belongs to $\mathcal{S}(\mathbb{R}^n, E)$ for every $x \in M^c$.

(ii) We set

$$g(x) := \int_{\mathbb{R}^n} f(x,y)\, dy = \sum_{j=0}^k e_j \lambda_n(A_{j,[x]}) \quad \text{for } x \in M^c \ . \tag{6.3}$$

Then Proposition 6.2 and Remark 1.2(d) show that $x \mapsto g(x)$ is λ_m-measurable. In addition, we have

$$\int_{\mathbb{R}^m} |g|\,dx \le \sum_{j=0}^{k} |e_j| \int_{\mathbb{R}^m} \lambda_n(A_{j,[x]})\,dx = \sum_{j=0}^{k} |e_j|\,\lambda_{m+n}(A_j) < \infty\ .$$

Therefore $x \mapsto g(x)$ is λ_m-integrable.

(iii) Finally, it follows from Proposition 6.2 and (6.3) that

$$\begin{aligned}\int_{\mathbb{R}^{m+n}} f\,d(x,y) &= \sum_{j=0}^{k} e_j \lambda_{m+n}(A_j) = \sum_{j=0}^{k} e_j \int_{\mathbb{R}^m} \lambda_n(A_{j,[x]})\,dx = \int_{\mathbb{R}^m} g\,dx\\ &= \int_{\mathbb{R}^m} \Big[\int_{\mathbb{R}^n} f(x,y)\,dy\Big]\,dx\ ,\end{aligned}$$

which completes the proof. ∎

6.5 Remark In the definition of the set $\mathcal{C}(m,n)$, we chose to single out the first m coordinates of $\mathbb{R}^{m+n}$. We could just as well have chosen the last n coordinates and made the same argument not with $\lambda_n(A_{[x]})$ but with $\lambda_m(A^{[y]})$ for λ_n-almost every $y \in \mathbb{R}^n$. With this definition of $\mathcal{C}(m,n)$, we would obviously have found that $\mathcal{C}(m,n) = \mathcal{L}(m+n)$. Thus the roles of x and y in Lemma 6.4 can be exchanged, and we conclude that, for $f \in \mathcal{S}(\mathbb{R}^{m+n}, E)$,

(i) *$f(\cdot\,,y) \in \mathcal{S}(\mathbb{R}^m, E)$ for λ_n-almost every $y \in \mathbb{R}^n$;*

(ii) *the E-valued function $y \mapsto \int_{\mathbb{R}^m} f(x,y)\,dx$ is λ_n-integrable;*

(iii) *$\int_{\mathbb{R}^{m+n}} f\,d(x,y) = \int_{\mathbb{R}^n} \big[\int_{\mathbb{R}^m} f(x,y)\,dx\big]\,dy$.*

In particular, we find

$$\int_{\mathbb{R}^m} \Big[\int_{\mathbb{R}^n} f(x,y)\,dy\Big]\,dx = \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^m} f(x,y)\,dx\Big]\,dy$$

for $f \in \mathcal{S}(\mathbb{R}^{m+n}, E)$. In other words, the integral $\int_{\mathbb{R}^{m+n}} f\,d(x,y)$ can be calculated iteratively in the case of simple functions, and the order in which the integrals are performed is irrelevant. ∎

Applications of Cavalieri's principle

The main result of this section is that the statement of Remark 6.5 about the iterative calculation of integrals remains true for arbitrary integrable functions f. Before we prove this theorem, we first give a few applications of Cavalieri's principle, meaning that we are working in the case $f = \chi_A$.

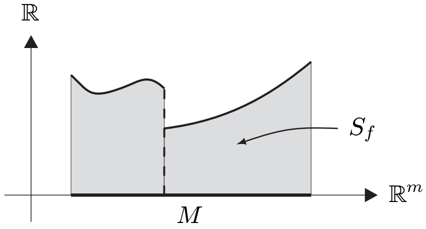
6.6 Examples **(a)** (geometric interpretation of the integral) For $M \in \mathcal{L}(m)$ and $f \in \mathcal{L}_0(M, \mathbb{R}^+)$, the set

$$S_f := S_{f,M} := \big\{ (x,y) \in \mathbb{R}^m \times \mathbb{R} \ ;\ 0 \le y \le f(x),\ x \in M \big\}$$

belongs to $\mathcal{L}(m+1)$, and

$$\int_M f\,dx = \lambda_{m+1}(S_f)\ ,$$

that is, the integral $\int_M f\,dx$ equals the $(m+1)$-dimensional Lebesgue measure of the set of points under the graph of f.[1]



Proof Set $f_1 := \mathrm{pr}_{\mathbb{R}}$ and $f_2 := f \circ \mathrm{pr}_{\mathbb{R}^m}$. Then f_1 and f_2 belong to $\mathcal{L}_0(M \times \mathbb{R}, \overline{\mathbb{R}}^+)$, and $S_f = [0 \le f_1 \le f_2]$. Therefore Proposition 1.9 implies the λ_{m+1}-measurability of S_f. Because $(S_f)_{[x]} = \big[0, f(x)\big]$ for $x \in M$, it follows that $\lambda_1\big((S_f)_{[x]}\big) = f(x)$, and hence

$$\lambda_{m+1}(S_f) = \int_{\mathbb{R}^m} \lambda_1\big((S_f)_{[x]}\big)\,dx = \int_M f\,dx\ ,$$

by Proposition 6.2. ■

(b) (substitution rule for linear maps) Suppose $T \in \mathcal{L}(\mathbb{R}^m)$, $a \in \mathbb{R}^m$ and $M \in \mathcal{L}(m)$. Also let $\varphi(x) := a + Tx$ for $x \in \mathbb{R}^m$ and $f \in \mathcal{L}_1(\varphi(M))$. Then $f \circ \varphi$ belongs to $\mathcal{L}_1(M)$, and

$$\int_{\varphi(M)} f\,dy = |\det T| \int_M (f \circ \varphi)\,dx\ . \tag{6.4}$$

In particular, the Lebesgue integral is **affine isometry invariant**, that is, for every affine isometry φ of $\mathbb{R}^m$, we have

$$\int_{\mathbb{R}^m} f = \int_{\mathbb{R}^m} f \circ \varphi \quad \text{for } f \in \mathcal{L}_1(\mathbb{R}^m)\ .$$

Proof (i) By Theorem IX.5.12, φ maps the σ-algebra $\mathcal{L}(m)$ into itself. Therefore $\varphi(M)$ belongs to $\mathcal{L}(m)$, and Theorem 1.4 implies that $f \circ \varphi$ lies in $\mathcal{L}_0(M)$. The decomposition $f = f_1 - f_2 + i\,(f_3 - f_4)$ with $f_j \in \mathcal{L}_1(\varphi(M), \mathbb{R}^+)$ shows that we can limit ourselves to the case of $f \in \mathcal{L}_1(\varphi(M), \mathbb{R}^+)$. Then (a) says that

$$\int_{\varphi(M)} f = \lambda_{m+1}(S_{f,\varphi(M)})\ , \qquad \int_M f \circ \varphi = \lambda_{m+1}(S_{f\circ\varphi, M})\ . \tag{6.5}$$

(ii) We set $\widehat{a} := (a, 0) \in \mathbb{R}^m \times \mathbb{R}$ and $\widehat{T}(x,t) := (Tx, t)$ for $(x,t) \in \mathbb{R}^m \times \mathbb{R}$. Then $\widehat{a} + \widehat{T}(S_{f\circ\varphi}) = S_f$ and $\det T = \det \widehat{T}$, because the representation matrix $\widehat{T}$ has the block

[1] Compare the introductory remarks to Section VI.3.

structure

$$[\widehat{T}\,] = \begin{bmatrix} [T] & 0 \\ 0 & 1 \end{bmatrix} .$$

Corollary IX.5.23 and Theorem IX.5.25 therefore imply

$$\lambda_{m+1}(S_f) = \lambda_{m+1}\big(\widehat{T}(S_{f\circ\varphi})\big) = |\det T|\,\lambda_{m+1}(S_{f\circ\varphi}) \ ,$$

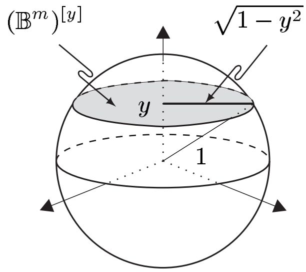
which, due to (6.5), proves (6.4). The integrability of $f\circ\varphi$ follows from Remark 3.11(a). ∎

(c) (the volume of the unit ball in $\mathbb{R}^m$) For $m \in \mathbb{N}^\times$, we have

$$\lambda_m(\mathbb{B}^m) = \frac{\pi^{m/2}}{\Gamma(1+m/2)} \ ;$$

in particular, $\lambda_1(\mathbb{B}^1) = 2$, $\lambda_2(\mathbb{B}^2) = \pi$, and $\lambda_3(\mathbb{B}^3) = 4\pi/3$.

Proof Setting $\omega_m := \lambda_m(\mathbb{B}^m)$, we obtain from Cavalieri's principle and Remarks IX.5.26(b) and 6.5 that



$$\begin{aligned}\omega_m &= \int_{-1}^{1} \lambda_{m-1}\big((\mathbb{B}^m)^{[y]}\big)\,dy \\ &= \int_{-1}^{1} \lambda_{m-1}\big(\sqrt{1-y^2}\,\mathbb{B}^{m-1}\big)\,dy \\ &= \omega_{m-1}\int_{-1}^{1} \big(\sqrt{1-y^2}\big)^{m-1}\,dy \ .\end{aligned}$$

To calculate the integral

$$B_m := \int_{-1}^{1} (1-y^2)^{(m-1)/2}\,dy = 2\int_0^1 (1-y^2)^{(m-1)/2}\,dy \quad \text{for } m \in \mathbb{N}^\times \ ,$$

we let $y = -\cos x$, so that $dy = \sin x\,dx$. This gives $B_m = 2\int_0^{\pi/2} \sin^m x\,dx$. It follows from the proof of Example VI.5.5(d) that

$$B_{2m} = \frac{(2m-1)(2m-3)\cdot\cdots\cdot 1}{2m(2m-2)\cdot\cdots\cdot 2}\cdot\pi \ , \quad B_{2m+1} = \frac{2m(2m-2)\cdot\cdots\cdot 2}{(2m+1)(2m-1)\cdot\cdots\cdot 1}\cdot 2 \ .$$

Thus we find $B_m B_{m-1} = 2\pi/m$ and

$$\omega_m = B_m\omega_{m-1} = B_m B_{m-1}\omega_{m-2} = \frac{2\pi}{m}\,\omega_{m-2} \ . \qquad (6.6)$$

Since $\omega_1 = 2$, we obtain $\omega_2 = B_2\omega_1 = 2B_2 = \pi$ and therefore, with (6.6),

$$\omega_{2m} = \frac{\pi^m}{m!} \ , \quad \omega_{2m+1} = \frac{(2\pi)^m}{1\cdot 3\cdot 5\cdot\cdots\cdot(2m+1)}\cdot 2 \ .$$

These two expressions can be unified with the help of the Gamma function, because

$$\Gamma(m+1) = m! \ , \quad \Gamma\Big(m+\frac{3}{2}\Big) = \frac{\sqrt{\pi}}{2^{m+1}}\cdot 1\cdot 3\cdot\cdots\cdot(2m+1)$$

(see Theorem VI.9.2 and Exercise VI.9.1). ∎

Tonelli's theorem

We now prove the advertised theorem that justifies the iterative calculation of integrals of nonnegative $\overline{\mathbb{R}}$-valued functions. This version, Tonelli's theorem, will give us an important integrability criterion in the case of E-valued functions.

6.7 Theorem (Tonelli) *For* $f \in \mathcal{L}_0(\mathbb{R}^{m+n}, \overline{\mathbb{R}}^+)$,

(i) $f(x, \cdot\,) \in \mathcal{L}_0(\mathbb{R}^n, \overline{\mathbb{R}}^+)$ *for* λ_m-a.a. $x \in \mathbb{R}^m$,
$f(\,\cdot\,, y) \in \mathcal{L}_0(\mathbb{R}^m, \overline{\mathbb{R}}^+)$ *for* λ_n-a.a. $y \in \mathbb{R}^n$;

(ii) $x \mapsto \int_{\mathbb{R}^n} f(x,y)\,dy$ *is* λ_m*-measurable,*
$y \mapsto \int_{\mathbb{R}^m} f(x,y)\,dx$ *is* λ_n*-measurable;*

(iii) $\int_{\mathbb{R}^{m+n}} f\,d(x,y) = \int_{\mathbb{R}^m} \bigl[\int_{\mathbb{R}^n} f(x,y)\,dy\bigr]\,dx = \int_{\mathbb{R}^n} \bigl[\int_{\mathbb{R}^m} f(x,y)\,dx\bigr]\,dy.$

Proof (i) By Theorem 1.12, there is an increasing sequence (f_j) in $\mathcal{S}(\mathbb{R}^{m+n}, \mathbb{R}^+)$ that converges to f. The monotone convergence theorem then gives

$$\lim_j \int_{\mathbb{R}^{m+n}} f_j\,d(x,y) = \int_{\mathbb{R}^{m+n}} f\,d(x,y) \qquad \text{in } \overline{\mathbb{R}}^+ \ . \tag{6.7}$$

Further, by Lemma 6.4, there is for every $j \in \mathbb{N}$ a λ_m-null set M_j such that $f_j(x, \cdot\,) \in \mathcal{S}(\mathbb{R}^n, \mathbb{R}^+)$ for $x \in M_j^c$. If we set $M := \bigcup_j M_j$, we then see from the monotone convergence theorem that

$$\int_{\mathbb{R}^n} f_j(x,y)\,dy \uparrow \int_{\mathbb{R}^n} f(x,y)\,dy \quad \text{for } x \in M^c \ . \tag{6.8}$$

Lemma 6.4(ii), Proposition 1.11, the fact that M has measure zero, and (6.8) imply that the $\overline{\mathbb{R}}$-valued function $x \mapsto \int_{\mathbb{R}^n} f(x,y)\,dy$ is λ_m-measurable. Next, it follows from (6.7), Lemma 6.4(iii), (6.8), and the monotone convergence theorem that

$$\begin{aligned}\int_{\mathbb{R}^{m+n}} f\,d(x,y) &= \lim_j \int_{\mathbb{R}^{m+n}} f_j\,d(x,y) = \lim_j \int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} f_j(x,y)\,dy\Bigr]\,dx \\ &= \int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} f(x,y)\,dy\Bigr]\,dx \ .\end{aligned}$$

The remaining statements are proved analogously (paying heed to Remark 6.5). ■

6.8 Corollary *For* $f \in \mathcal{L}_0(\mathbb{R}^{m+n}, E)$, *suppose* $f = 0$ λ_{m+n}-a.e. *Then there is a* λ_m*-null set* M *such that* $f(x, \cdot\,)$ *vanishes* λ_n-a.e. *for every* $x \in M^c$, *and a* λ_n*-null set* N *such that* $f(\,\cdot\,, y) = 0$ λ_m-a.e. *for every* $y \in N^c$.

Proof Clearly, it suffices to prove the existence of M (compare Remark 6.5). Tonelli's theorem gives

$$\int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} |f(x,y)|\,dy\Bigr]\,dx = \int_{\mathbb{R}^{m+n}} |f|\,d(x,y) = 0 \ .$$

Thus according to Remark 3.3(c) there is a λ_m-null set M such that

$$\int_{\mathbb{R}^n} |f(x,y)|\,dy = 0 \quad \text{for } x \in M^c \ ,$$

from which the claim follows again by Remark 3.3(c). ■

Fubini's theorem for scalar functions

It is now easy to extend Tonelli's to the case of integrable $\mathbb{K}$-valued functions, which is of particular interest for applications.

6.9 Theorem (Fubini) *For $f \in \mathcal{L}_1(\mathbb{R}^{m+n})$,*

(i) $f(x,\cdot\,) \in \mathcal{L}_1(\mathbb{R}^n)$ *for λ_m-almost every $x \in \mathbb{R}^m$,*
$f(\,\cdot\,,y) \in \mathcal{L}_1(\mathbb{R}^m)$ *for λ_n-almost every $y \in \mathbb{R}^n$;*

(ii) $x \mapsto \int_{\mathbb{R}^n} f(x,y)\,dy$ *is λ_m-integrable,*
$y \mapsto \int_{\mathbb{R}^m} f(x,y)\,dx$ *is λ_n-integrable;*

(iii) $\int_{\mathbb{R}^{m+n}} f\,d(x,y) = \int_{\mathbb{R}^m}\bigl[\int_{\mathbb{R}^n} f(x,y)\,dy\bigr]\,dx = \int_{\mathbb{R}^n}\bigl[\int_{\mathbb{R}^m} f(x,y)\,dx\bigr]\,dy.$

Proof (a) For $f \in \mathcal{L}_1(\mathbb{R}^{m+n},\mathbb{R}^+)$, the claim follows from Tonelli's theorem and Remark 3.3(e).

(b) Given the representation $f = f_1 - f_2 + i(f_3 - f_4)$, with $f_j \in \mathcal{L}_1(\mathbb{R}^{m+n},\mathbb{R}^+)$, the general case now follows by Corollary 2.12 and the linearity of the integral. ■

6.10 Corollary *Suppose $A \in \mathcal{L}(m)$ and $f \in \mathcal{L}_1(A)$. Let $(j_1,\ldots,j_m)$ denote a permutation of $(1,\ldots,m)$. Then*

$$\int_A f\,dx = \int_{\mathbb{R}}\Bigl(\int_{\mathbb{R}}\cdots\Bigl(\int_{\mathbb{R}} \widetilde{f}(x_1\ldots,x_m)\,dx_{j_1}\Bigr)\cdots\,dx_{j_{m-1}}\Bigr)\,dx_{j_m} \ .$$

Fubini's theorem guarantees that *integrable* functions can be integrated in any order. In combination with Tonelli's theorem, we obtain a simple, versatile, and extraordinarily important criterion for the integrability of functions of multiple variables, as well as a method for explicitly calculating integrals.

6.11 Theorem (Fubini–Tonelli) *Suppose $A \in \mathcal{L}(m+n)$ and $f \in \mathcal{L}_0(A)$.*

(i) *If one of the integrals*

$$\int_{\mathbb{R}^m}\Bigl[\int_{\mathbb{R}^n} |\widetilde{f}(x,y)|\,dy\Bigr]\,dx \ , \quad \int_{\mathbb{R}^n}\Bigl[\int_{\mathbb{R}^m} |\widetilde{f}(x,y)|\,dx\Bigr]\,dy \ , \quad \int_A |f|\,d(x,y)$$

is finite, then so is each of the others, and they are all equal. In that case, f is integrable, and the statement of Theorem 6.9 holds for $\widetilde{f}$.

(ii) *If* $\mathrm{pr}_{\mathbb{R}^m}(A)$ *is measurable*[2] *and* f *is integrable, then*

$$\int_A f\,d(x,y) = \int_{\mathrm{pr}_{\mathbb{R}^m}(A)} \Bigl[\int_{A_{[x]}} f(x,y)\,dy\Bigr]\,dx\ .$$

Proof Because $\widetilde{f}$ belongs to $\mathcal{L}_0(\mathbb{R}^{m+n})$, the first statement follows immediately from Tonelli's theorem. Then by Theorem 3.9, $\widetilde{f}$ is integrable, and hence so is f. The claim is now clear. ■

6.12 Remarks **(a)** We have lost no generality by choosing the first m coordinates, because by Corollary 6.10, this order can always be achieved by a permutation.

(b) Typically we omit the brackets in $\int_{\mathbb{R}^n}\bigl[\int_{\mathbb{R}^m} f(x,y)\,dx\bigr]\,dy$ and instead write, say,

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^m} f(x,y)\,dx\,dy\ . \tag{6.9}$$

In this notation, it is understood that the integrals are to be evaluated from the inside to the outside.[3] The **iterated integral** (6.9) is to be distinguished from the $(m+n)$-dimensional integral

$$\int_{\mathbb{R}^{m+n}} f\,d(x,y) = \int_{\mathbb{R}^{m+n}} f\,d\lambda_{m+n}\ .$$

(c) There exists $f \in \mathcal{L}_0(\mathbb{R}^2)\setminus\mathcal{L}_1(\mathbb{R}^2)$ such that

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dx\,dy = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dy\,dx = 0\ .$$

Therefore the existence and equality of the iterated integral does not imply that f is integrable.

Proof Define $f:\mathbb{R}^2\to\mathbb{R}$ by

$$f(x,y) := \begin{cases} \dfrac{xy}{(x^2+y^2)^2} & \text{if } (x,y)\neq(0,0)\ ,\\ 0 & \text{if } (x,y)=(0,0)\ . \end{cases} \tag{6.10}$$

Then f is λ_2-measurable. For every $y\in\mathbb{R}$, the improper Riemann integral $\int_{\mathbb{R}} f(x,y)\,dx$ converges absolutely. Also $f(\,\cdot\,,y)$ is odd. Hence $\int_{\mathbb{R}} f(x,y)\,dx = 0$ for every[4] $y\in\mathbb{R}$, and therefore, because $f(x,y)=f(y,x)$, we have

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dx\,dy = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\,dy\,dx = 0\ .$$

[2] As Remark IX.5.14(b) shows, this is not generally the case.

[3] That is, the integral $\int_{\mathbb{R}^m} f(x,y)\,dx$ is calculated for fixed y and the result is then integrated over y in $\mathbb{R}^n$.

[4] The case $y=0$ is covered in the given argument, although it follows more simply from the fact that $f(\,\cdot\,,0)=0$.

Now suppose f were integrable. Then, by Fubini's theorem, $x \mapsto \int_{\mathbb{R}} |f(x,y)|\,dy$ would also be integrable, which, because

$$\int_{\mathbb{R}} \frac{|xy|}{(x^2+y^2)^2}\,dy = \frac{1}{|x|} \quad \text{for } x \neq 0\ ,$$

cannot be true. ■

(d) There exists $g \in \mathcal{L}_0(\mathbb{R}^2) \setminus \mathcal{L}_1(\mathbb{R}^2)$ such that

$$0 < \Bigl|\int_{\mathbb{R}}\int_{\mathbb{R}} g(x,y)\,dx\,dy\Bigr| = \Bigl|\int_{\mathbb{R}}\int_{\mathbb{R}} g(x,y)\,dy\,dx\Bigr| < \infty\ .$$

Proof Let f be the function from (6.10) and take $h \in \mathcal{L}_1(\mathbb{R}^2)$ with $\int h\,d(x,y) > 0$. Then $g := f + h$ has the stated properties. ■

6.13 Examples **(a)** (multidimensional Gaussian integrals) For $n \in \mathbb{N}^\times$, we have

$$\int_{\mathbb{R}^n} e^{-|x|^2}\,dx = \pi^{n/2}\ .$$

Proof Using $|x|^2 = x_1^2 + \cdots + x_n^2$ and the properties of the exponential function, it follows from Tonelli's theorem that

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-|x|^2}\,dx &= \int_{\mathbb{R}} \cdot \cdots \cdot \int_{\mathbb{R}} e^{-x_1^2} e^{-x_2^2} \cdot \cdots \cdot e^{-x_n^2}\,dx_1 \cdot \cdots \cdot dx_n \\ &= \prod_{j=1}^{n} \int_{\mathbb{R}} e^{-x_j^2}\,dx_j = \Bigl(\int_{\mathbb{R}} e^{-t^2}\,dt\Bigr)^n\ .\end{aligned}$$

Now the claim follows from Application VI.9.7. ■

(b) (a representation of the beta function[5]) For $v, w \in [\operatorname{Re} z > 0]$,

$$\mathsf{B}(v,w) = \frac{\Gamma(v)\,\Gamma(w)}{\Gamma(v+w)}\ .$$

Proof Set $A := \bigl\{ (s,t) \in \mathbb{R}^2\ ;\ 0 < t < s \bigr\}$ and define $\gamma_{v,w} : A \to \mathbb{C}$ by $\gamma_{v,w}(s,t) := t^{v-1}(s-t)^{w-1}e^{-s}$ for $v, w \in [\operatorname{Re} z > 0]$. Setting $\gamma_v(t) := t^{v-1}e^{-t}$ for $t > 0$, we find from Tonelli's theorem that

$$\begin{aligned}\int_A |\gamma_{v,w}(s,t)|\,d(s,t) &= \int_0^\infty \int_t^\infty |\gamma_{v,w}(s,t)|\,ds\,dt \\ &= \Bigl(\int_0^\infty \gamma_{\operatorname{Re} v}(t)\,dt\Bigr)\Bigl(\int_0^\infty \gamma_{\operatorname{Re} w}(s)\,ds\Bigr) = \Gamma(\operatorname{Re} v)\,\Gamma(\operatorname{Re} w) < \infty\ .\end{aligned}$$

Therefore $\gamma_{v,w}$ is integrable, and Fubini's theorem analogously gives

$$\int_A \gamma_{v,w}(s,t)\,d(s,t) = \int_0^\infty \int_t^\infty \gamma_{v,w}(s,t)\,ds\,dt = \Gamma(v)\,\Gamma(w)\ . \tag{6.11}$$

[5]Compare Remark VI.9.12(a).

Since $\mathrm{pr}_1(A) = \mathbb{R}^+$ and $A_{[s]} = [0, s]$ for $s > 0$, we obtain from (6.11) and Theorem 6.11(ii)

$$\Gamma(v)\,\Gamma(w) = \int_0^\infty \Bigl(\int_0^s t^{v-1}(s-t)^{w-1}\,dt\Bigr)e^{-s}\,ds\ .$$

The substitution $r = t/s$ in the inner integral and the definition of the beta function give

$$\Gamma(v)\,\Gamma(w) = \int_0^\infty \Bigl(\int_0^1 r^{v-1}(1-r)^{w-1}\,dr\Bigr)s^{v+w-1}e^{-s}\,ds = \mathsf{B}(v,w)\,\Gamma(v+w)\ , \tag{6.12}$$

which completes the proof. ∎

Example (b) shows that complicated integrals can be simplified by a deft choice of integration order.

Fubini's theorem for vector-valued functions[6]

We now want to show that Fubini's theorem also holds for E-valued functions, and offer some applications. A few preliminary remarks will prove helpful.

Suppose $A \in \mathcal{L}(m+n)$ has finite measure. By Proposition 6.2 and Corollary 6.3, there is a λ_m-null set M such that $A_{[x]} \in \mathcal{L}(n)$ and $\lambda_n(A_{[x]}) < \infty$ for $x \in M^c$. We fix $q \in [1, \infty)$. Because $|\chi_{A_{[x]}}|^q = \chi_{A_{[x]}}$, we have

$$\int_{\mathbb{R}^n} |\chi_{A_{[x]}}(y)|^q\,dy = \int_{\mathbb{R}^n} \chi_{A_{[x]}}(y)\,dy = \lambda_n(A_{[x]}) < \infty\ .$$

If, as agreed to in Section 4, we identify $\chi_{A_{[x]}}$ with the equivalence class of all functions that coincide λ_n-a.e. with $y \mapsto \chi_{A_{[x]}}(y)$, we obtain the map

$$M^c \to F := L_q(\mathbb{R}^n)\ ,\quad x \mapsto \chi_{A_{[x]}}\ .$$

Because F is a Banach space, we can study its measurability and integrability properties.

6.14 Lemma *Suppose $A \in \mathcal{L}(m+n)$ has finite measure. Then the F-valued map $x \mapsto \chi_{A_{[x]}}$, which is defined λ_m-everywhere, is λ_m-measurable.*

Proof We denote by $\psi_A : \mathbb{R}^m \to F$ the trivial extension of $x \mapsto \chi_{A_{[x]}}$.

(i) Suppose A is a λ_{m+n}-null set. By Remark 6.1(h), there is a λ_m-null set M such that $A_{[x]}$ is a λ_n-null set for $x \in M^c$. Therefore $\psi_A(x) = 0$ in F for $x \in M^c$. The claim follows.

(ii) Now suppose A is an interval of the form $[a, b)$ with $a, b \in \mathbb{R}^{m+n}$. Set $J_1 := \prod_{j=1}^m [a_j, b_j)$ and $J_2 := \prod_{j=m+1}^{m+n} [a_j, b_j)$. Because $A = J_1 \times J_2$, we have

$$\chi_{A_{[x]}} = \chi_{J_1}(x)\,\chi_{J_2} \quad \text{for } x \in \mathbb{R}^m\ ,$$

and we see that in this case ψ_A belongs to $\mathcal{S}(\mathbb{R}^m, F)$.

[6]This section may be skipped on first reading.

(iii) Suppose $A \subset \mathbb{R}^{m+n}$ is open and (I_j) is a disjoint sequence of intervals of the form $[a, b)$ with $A = \bigcup_j I_j$ (see Proposition IX.5.6). We set

$$f_k := \sum_{j=0}^{k} \psi_{I_j} \quad \text{for } k \in \mathbb{N} .$$

By (ii) and Remark 1.2(a), (f_k) is a sequence in $\mathcal{S}(\mathbb{R}^m, F)$. Also, there is a set M of Lebesgue measure zero such that

$$\begin{aligned}\|\psi_A(x) - f_k(x)\|_F^q &= \int_{\mathbb{R}^n} \Big|\chi_{A_{[x]}}(y) - \Big(\sum_{j=0}^{k} \chi_{(I_j)_{[x]}}(y)\Big)\Big|^q \, dy \\ &= \lambda_n\Big(\bigcup_{j=k+1}^{\infty} (I_j)_{[x]}\Big) = \sum_{j=k+1}^{\infty} \lambda_n\big((I_j)_{[x]}\big)\end{aligned}$$

for $x \in M^c$. In addition, $A_{[x]}$ has finite measure by Corollary 6.3, and $\lambda_n(A_{[x]}) = \sum_{j=0}^{\infty} \lambda_n((I_j)_{[x]})$ for $x \in M^c$. Therefore (f_k) converges λ_m-a.e. to ψ_A in F, and we see that ψ_A belongs to $\mathcal{L}_0(\mathbb{R}^m, F)$.

(iv) Suppose A is a G_δ-set. The proof of Corollary IX.5.5 shows that there is a sequence (O_j) of open sets such that $\lambda_{m+n}(O_j) < \infty$ and $A = \bigcap O_j$. Set

$$f_k := \psi_{\bigcap_{j=0}^{k} O_j} , \qquad R_k := \bigcap_{j=0}^{k} O_j \backslash A \quad \text{for } k \in \mathbb{N} .$$

Then (f_k) is a sequence in $\mathcal{L}_0(\mathbb{R}^m, F)$ by (iii), and (R_k) is a descending sequence with $\bigcap_{k=0}^{\infty} R_k = \emptyset$ and $\lambda_{m+n}(R_0) < \infty$. Also, we have

$$\|f_k(x) - \psi_A(x)\|_F^q = \int_{\mathbb{R}^n} \big|\chi_{(\bigcap_{j=0}^{k} O_j)_{[x]}}(y) - \chi_{A_{[x]}}(y)\big|^q \, dy = \lambda_n\big((R_k)_{[x]}\big)$$

for λ_m-almost every $x \in \mathbb{R}^m$. The continuity of λ_n from above therefore implies that (f_k) converges λ_m-a.e. to ψ_A. From Theorem 1.14, it now follows that ψ_A belongs to $\mathcal{L}_0(\mathbb{R}^m, F)$.

(v) To conclude, consider $A \in \mathcal{L}(m + n)$ such that $\lambda_{m+n}(A)$ is finite. By Corollary IX.5.5, there is a G_δ-set G containing A and having the same measure as A. By Proposition IX.2.3(ii), $N := G\backslash A$ is a λ_{m+n}-null set with $\psi_A = \psi_G - \psi_N$ λ_m-a.e. Now the claim follows from (i) and (iv). ■

6.15 Corollary *Let $p, q \in [1, \infty)$, and suppose $\varphi \in \mathcal{S}(\mathbb{R}^{m+n}, E)$ has compact support. Then the $L_q(\mathbb{R}^n, E)$-valued function $x \mapsto [\varphi(x, \cdot)]$ is defined λ_m-a.e. and is L_p-integrable, that is,*

$$\int_{\mathbb{R}^m} \|\varphi(x, \cdot)\|_{L_q(\mathbb{R}^n, E)}^p \, dx < \infty .$$

If $p = q$, this holds for every $\varphi \in \mathcal{S}(\mathbb{R}^{m+n}, E)$.

Proof By Minkowski's inequality, it suffices to prove this for $\varphi := e\chi_A$ with $e \in E$ and $A \in \mathcal{L}(m+n)$, where A has finite measure if $p = q$, and A is bounded if $p \neq q$.

By Lemma 6.14, there is a λ_m-null set M such that the function

$$M^c \to L_q(\mathbb{R}^n) \ , \quad x \mapsto \chi_{A_{[x]}}$$

is λ_m-measurable. Because $\varphi(x,\cdot) = e\chi_{A_{[x]}}$, $(x \mapsto \varphi(x,\cdot)) \in \mathcal{L}_0(M^c, L_q(\mathbb{R}^n, E))$. From

$$\|\varphi(x,\cdot)\|_{L_q(\mathbb{R}^n,E)} = \Bigl(\int_{\mathbb{R}^n} |e|^q\, \chi_{A_{[x]}}(y)\, dy\Bigr)^{1/q} = |e| \, \bigl[\lambda_n(A_{[x]})\bigr]^{1/q} \quad \text{for } x \in M^c \ ,$$

we obtain

$$\int_{\mathbb{R}^m} \|\varphi(x,\cdot)\|^p_{L_q(\mathbb{R}^n,E)}\, dx = |e|^p \int_{\mathbb{R}^m} \lambda_n(A_{[x]})^{p/q}\, dx \ .$$

In the case $p = q$, Proposition 6.2 implies

$$\int_{\mathbb{R}^n} \lambda_n(A_{[x]})\, dx = \lambda_{m+n}(A) < \infty \ .$$

Suppose therefore $p \neq q$. Because φ has compact support, there are compact subsets $K \subset \mathbb{R}^m$ and $L \subset \mathbb{R}^n$ such that $A \subset K \times L$. Thus $A_{[x]} \subset L$, which implies $\lambda_n(A_{[x]}) \leq \lambda_n(L)$ for λ_m-almost every $x \in \mathbb{R}^m$. From this we deduce

$$\int_{\mathbb{R}^m} \lambda_n(A_{[x]})^{p/q}\, dx = \int_K \lambda_n(A_{[x]})^{p/q}\, dx \leq \lambda_n(L)^{p/q}\, \lambda_m(K) < \infty \ . \ \blacksquare$$

These preparations are more general than necessary for our current purpose, but will prove useful for further applications. We are ready to prove Fubini's theorem in the E-valued case.

6.16 Theorem (Fubini) *For $f \in \mathcal{L}_1(\mathbb{R}^{m+n}, E)$,*

(i) *$f(x,\cdot) \in \mathcal{L}_1(\mathbb{R}^n, E)$ for λ_m-almost every $x \in \mathbb{R}^m$;*
$f(\cdot,y) \in \mathcal{L}_1(\mathbb{R}^m, E)$ for λ_n-almost every $y \in \mathbb{R}^n$;

(ii) *$x \mapsto \int_{\mathbb{R}^n} f(x,y)\, dy$ is λ_m-integrable;*
$y \mapsto \int_{\mathbb{R}^m} f(x,y)\, dx$ is λ_n-integrable;

(iii) $\int_{\mathbb{R}^{m+n}} f\, d(x,y) = \int_{\mathbb{R}^m} \bigl[\int_{\mathbb{R}^n} f(x,y)\, dy\bigr]\, dx = \int_{\mathbb{R}^n} \bigl[\int_{\mathbb{R}^m} f(x,y)\, dx\bigr]\, dy.$

Proof (a) Let $f \in \mathcal{L}_1(\mathbb{R}^{m+n}, E)$. Then there is an $\mathcal{L}_1$-Cauchy sequence (f_j) in $\mathcal{S}(\mathbb{R}^{m+n}, E)$ and a λ_{m+n}-null set L such that $f_j(x,y) \to f(x,y)$ for $(x,y) \in L^c$. By Remark 6.1(h), there is a set M_1 of measure zero such that

$$f_j(x,\cdot) \to f(x,\cdot) \quad \lambda_n\text{-a.e.} \ , \tag{6.13}$$

for $x \in M_1^c$. We set $F := L_1(\mathbb{R}^n, E)$ and denote by φ_j the trivial extension of $x \mapsto f_j(x,\cdot)$. According to Corollary 6.15, (φ_j) is a sequence in $\mathcal{L}_1(\mathbb{R}^m, F)$ for which

$$\|\varphi_j - \varphi_k\|_1 = \int_{\mathbb{R}^m} \|\varphi_j(x) - \varphi_k(x)\|_F \, dx = \int_{\mathbb{R}^m}\int_{\mathbb{R}^n} |f_j(x,y) - f_k(x,y)| \, dy \, dx \ .$$

Further, Lemma 6.4 shows

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^n} |f_j(x,y) - f_k(x,y)| \, dy \, dx = \int_{\mathbb{R}^{m+n}} |f_j - f_k| \, d(x,y) = \|f_j - f_k\|_1 \ ,$$

and we see that (φ_j) is a Cauchy sequence in $\mathcal{L}_1(\mathbb{R}^m, F)$. By Theorems 2.10 and 2.18, there is thus a $\widehat{g} \in \mathcal{L}_1(\mathbb{R}^m, F)$, a λ_m-null set M_2, and a subsequence of (φ_j), which, for simplicity, we also denote by (φ_j), such that

$$\lim_{j\to\infty} \varphi_j(x) = \widehat{g}(x) \quad \text{for } x \in M_2^c \tag{6.14}$$

in F and $\varphi_j \to \widehat{g}$ in $\mathcal{L}_1(\mathbb{R}^m, F)$. For $x \in M_2^c$, let $g(x) \in \mathcal{L}_1(\mathbb{R}^n, E)$ be a representative of $\widehat{g}(x)$. Then there is a set $N(x)$ of Lebesgue measure zero and a subsequence of $(\varphi_j(x))$, which we also write as $(\varphi_j(x))$, such that, in E,

$$\lim_{j\to\infty} f_j(x,y) = \lim_{j\to\infty} \varphi_j(x)(y) = g(x)(y) \quad \text{for } x \in M_2^c \text{ and } y \in (N(x))^c \ .$$

Hence (6.13) implies that for every $x \in M_1^c \cap M_2^c$ the maps $f(x,\cdot)$, $g(x) : \mathbb{R}^n \to E$ are equal λ_n-a.e. Lemma 2.15 now shows that $f(x,\cdot)$ belongs to $\mathcal{L}_1(\mathbb{R}^n, E)$ and that

$$\int_{\mathbb{R}^n} g(x)(y) \, dy = \int_{\mathbb{R}^n} f(x,y) \, dy \quad \text{for } x \in M_1^c \cap M_2^c \ . \tag{6.15}$$

Furthermore, it follows from (6.13), (6.14), and Theorem 2.18(ii) that

$$\int_{\mathbb{R}^n} f_j(x,y) \, dy = \int_{\mathbb{R}^n} \varphi_j(x)(y) \, dy \to \int_{\mathbb{R}^n} g(x)(y) \, dy = \int_{\mathbb{R}^n} f(x,y) \, dy \tag{6.16}$$

for $x \in M_1^c \cap M_2^c$.

(b) For $\varphi \in F = L_1(\mathbb{R}^n, E)$, let $A\varphi := \int_{\mathbb{R}^n} \varphi \, dy$. By Theorem 2.11(i), A belongs to $\mathcal{L}(F, E)$. Theorem 2.11(iii) implies that $g_j := A\varphi_j$ defines a sequence in $\mathcal{L}_1(\mathbb{R}^m, E)$.

Because

$$g_j(x) = \int_{\mathbb{R}^n} \varphi_j(x)(y) \, dy = \int_{\mathbb{R}^n} f_j(x,y) \, dy \ , \tag{6.17}$$

we know from Theorem 2.11(i) that

$$|g_j(x) - g_k(x)| = \Big| \int_{\mathbb{R}^n} \big(f_j(x,y) - f_k(x,y)\big) \, dy \Big| \le \int_{\mathbb{R}^n} |f_j(x,y) - f_k(x,y)| \, dy \ .$$

Therefore Theorem 2.11(ii) gives

$$\int_{\mathbb{R}^m} |g_j - g_k| \, dx \le \int_{\mathbb{R}^m}\int_{\mathbb{R}^n} \big|\big(f_j(x,y) - f_k(x,y)\big)\big| \, dy \, dx = \|f_j - f_k\|_1 \ ,$$

where the last equality follows from Tonelli's theorem. Therefore (g_j) is a Cauchy sequence in $\mathcal{L}_1(\mathbb{R}^m, E)$, and by completeness there is some $h \in \mathcal{L}_1(\mathbb{R}^m, E)$ such that $g_j \to h$ in $\mathcal{L}_1(\mathbb{R}^m, E)$. Hence we can find a λ_m-null set M_3 and a subsequence of (g_j), which we also denote by (g_j), such that $g_j(x) \to h(x)$ for $x \in M_3^c$ and $j \to \infty$. In view of (6.17), it follows from (6.16) that

$$h(x) = \int_{\mathbb{R}^n} f(x,y)\,dy \quad \text{for } x \in M_1^c \cap M_2^c \cap M_3^c\ , \tag{6.18}$$

which proves the first statement of (ii).

(c) Since $g_j \to h$ in $\mathcal{L}_1(\mathbb{R}^m, E)$ and because of (6.17) and (6.18), Theorem 2.18(ii) implies

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^n} f_j(x,y)\,dy\,dx \to \int_{\mathbb{R}^m}\int_{\mathbb{R}^n} f(x,y)\,dy\,dx\ .$$

Finally, it follows from Lemma 6.4 that

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^n} f_j(x,y)\,dy\,dx = \int_{\mathbb{R}^{m+n}} f_j\,d(x,y)\ ,$$

and with $\int_{\mathbb{R}^{m+n}} f\,d(x,y) = \lim_j \int_{\mathbb{R}^{m+n}} f_j\,d(x,y)$, we have

$$\int_{\mathbb{R}^m}\int_{\mathbb{R}^n} f(x,y)\,dy\,dx = \int_{\mathbb{R}^{m+n}} f\,d(x,y)\ .$$

We have proved the first part of each of the statements (i) and (ii), and the first equality in (iii). The remaining claims follow by exchanging the roles of x and y. ■

6.17 Remark The analogues of the Fubini–Tonelli theorem and Corollary 6.10 clearly also hold in the E-valued case. ■

Minkowski's inequality for integrals

As an application we now prove a continuous version of Minkowski's inequality.

Fix $p, q \in [1, \infty)$. For $f \in \mathcal{L}_0(\mathbb{R}^{m+n}, E)$, Theorem 1.7(i) shows that $|f|^q$ belongs to $\mathcal{L}_0(\mathbb{R}^{m+n}, \mathbb{R}^+)$. Hence Tonelli's theorem implies that $|f(x,\cdot)|^q$ lies in $\mathcal{L}_0(\mathbb{R}^n, \mathbb{R}^+)$ for λ_m-almost every $x \in \mathbb{R}^m$ and that the $\overline{\mathbb{R}}^+$-valued function

$$x \mapsto \int_{\mathbb{R}^n} |f(x,y)|^q\,dy\ ,$$

which is defined λ_m-a.e., is λ_m-measurable. Therefore

$$\|f\|_{(p,q)} := \Bigl(\int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} |f(x,y)|^q\,dy\Bigr]^{p/q} dx\Bigr)^{1/p}$$

is defined in $\overline{\mathbb{R}}^+$. We easily check that

$$\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n}, E) := \bigl\{\, f \in \mathcal{L}_0(\mathbb{R}^{m+n}, E)\ ;\ \|f\|_{(p,q)} < \infty \,\bigr\}$$

is a vector subspace of $\mathcal{L}_0(\mathbb{R}^{m+n},E)$ and that $\|\cdot\|_{(p,q)}$ defines a seminorm on $\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n},E)$. Finally, we set

$$\mathcal{S}_c(\mathbb{R}^{m+n},E) := \{\, f\in \mathcal{S}(\mathbb{R}^{m+n},E)\ ;\ \mathrm{supp}(f) \text{ is compact} \,\}\ .$$

6.18 Lemma *$\mathcal{S}_c(\mathbb{R}^{m+n},E)$ is a dense vector subspace of $\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n},E)$.*

Proof (i) Take $f\in\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n},E)$ and let (g_k) be a sequence in $\mathcal{S}(\mathbb{R}^{m+n},E)$ such that $g_k\to f$ a.e. Set $A_k := [\,|g_k|\le 2\,|f|\,]\cap k\mathbb{B}^{m+n}$ and $f_k := \chi_{A_k} g_k$. Then (f_k) is a sequence in $\mathcal{S}_c(\mathbb{R}^{m+n},E)$, and there is a λ_{m+n}-null set L such that

$$f_k(x,y)\to f(x,y)\quad \text{for } (x,y)\in L^c\ . \tag{6.19}$$

Moreover

$$|f_k-f|\le |f_k|+|f|\le 3\,|f|\quad \text{for } k\in\mathbb{N}\ . \tag{6.20}$$

(ii) By (6.20), it follows from Tonelli's theorem and Theorem 3.9 that there is a λ_m-null set M_0 such that

$$|f(x,\cdot\,)-f_k(x,\cdot\,)|^q,\ |f(x,\cdot\,)|^q\in\mathcal{L}_1(\mathbb{R}^n)\quad \text{for } x\in M_0^c \text{ and } k\in\mathbb{N}\ . \tag{6.21}$$

Remark 6.1(h) says there is a λ_m-null set M_1 such that $L_{[x]}$ is a λ_n-null set for every $x\in M_1^c$. Set $M:=M_0\cup M_1$ and choose $x\in M^c$. From (6.19), we read off that $f_k(x,y)\to f(x,y)$ for $y\in(L_{[x]})^c$. By (6.20) and (6.21), we can apply Lebesgue's dominated convergence theorem to the sequence $\bigl(|f(x,\cdot\,)-f_k(x,\cdot\,)|^p\bigr)_{k\in\mathbb{N}}$, and we find

$$\lim_{k\to\infty}\int_{\mathbb{R}^n}|f(x,y)-f_k(x,y)|^q\,dy=0\quad \text{for } x\in M^c\ .$$

Now define

$$\varphi_k := \Bigl(x\mapsto\Bigl(\int_{\mathbb{R}^n}|f(x,y)-f_k(x,y)|^q\,dy\Bigr)^{p/q}\Bigr)^{\sim}\quad \text{for } k\in\mathbb{N}\ .$$

Then the sequence (φ_k) converges λ_m-a.e. to 0.

(iii) Finally, set

$$\varphi := \Bigl(x\mapsto 3^p\Bigl(\int_{\mathbb{R}^n}|f(x,y)|^q\,dy\Bigr)^{p/q}\Bigr)^{\sim}\ .$$

Because $f\in\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n},E)$, we know φ belongs to $\mathcal{L}_1(\mathbb{R}^m)$, and (6.20) implies $0\le\varphi_k\le\varphi$ λ_m-a.e. for $k\in\mathbb{N}$. Hence we can apply dominated convergence theorem to (φ_k) to see that $(\int_{\mathbb{R}^m}\varphi_k)_{k\in\mathbb{N}}$ is a null sequence in $\mathbb{R}^+$. The claim now follows because

$$\int_{\mathbb{R}^m}\varphi_k=\int_{\mathbb{R}^m}\Bigl[\int_{\mathbb{R}^n}|f(x,y)-f_k(x,y)|^q\,dy\Bigr]^{p/q}dx=\|f-f_k\|_{(p,q)}^p\ . \ \blacksquare$$

One easily checks that $\mathcal{N} := \{\, f \in \mathcal{L}_0(\mathbb{R}^{m+n}, E) \ ; \ f = 0 \text{ a.e.} \,\}$ is a vector subspace of $\mathcal{L}_{(p,q)}(\mathbb{R}^{m+n}, E)$ and that f belongs to $\mathcal{N}$ if and only if $\|f\|_{(p,q)} = 0$. Therefore

$$L_{(p,q)}(\mathbb{R}^{m+n}, E) := \mathcal{L}_{(p,q)}(\mathbb{R}^{m+n}, E)/\mathcal{N}$$

is a well defined vector space, and the assignment $[f] \mapsto \|f\|_{(p,q)}$ defines a norm on $L_{(p,q)}(\mathbb{R}^{m+n}, E)$, which we again denote by $\|\cdot\|_{(p,q)}$. In what follows, we always provide the space $L_{(p,q)}(\mathbb{R}^{m+n}, E)$ with the topology induced by $\|\cdot\|_{(p,q)}$.

We set

$$S_c(\mathbb{R}^{m+n}, E) := \{\, [f] \in L_0(\mathbb{R}^{m+n}, E) \ ; \ [f] \cap \mathcal{S}_c(\mathbb{R}^{m+n}, E) \neq \emptyset \,\} \ .$$

6.19 Remarks **(a)** $S_c(\mathbb{R}^{m+n}, E)$ is a dense vector subspace of $L_{(p,q)}(\mathbb{R}^{m+n}, E)$.

Proof This follows from Lemma 6.18. ■

(b) Let $f \in \mathcal{L}_0(\mathbb{R}^{m+n}, E)$. If $f(x, \cdot)$ belongs to $\mathcal{L}_q(\mathbb{R}^n, E)$ for almost every $x \in \mathbb{R}^m$ and

$$\Bigl[x \mapsto \Bigl(\int_{\mathbb{R}^n} |f(x,y)|^q \, dx\Bigr)^{1/q}\Bigr]^{\sim} \in \mathcal{L}_p(\mathbb{R}^n) \ ,$$

then $[f]$ belongs to $L_{(p,q)}(\mathbb{R}^{m+n}, E)$.

(c) $L_{(p,p)}(\mathbb{R}^{m+n}, E) = L_p(\mathbb{R}^{m+n}, E)$.

Proof This follows from Remark 4.9(b) and the Fubini–Tonelli theorem. ■

(d) $S_c(\mathbb{R}^n, E)$ is a dense vector subspace of $L_p(\mathbb{R}^n, E)$.

Proof This is a consequence of (a) and (c). ■

Consider $g \in \mathcal{S}_c(\mathbb{R}^{m+n}, E)$. By Corollary 6.15, $T_0 g := \bigl(x \mapsto \bigl[g(x, \cdot)\bigr]\bigr)^{\sim}$ belongs to $\mathcal{L}_p\bigl(\mathbb{R}^m, L_q(\mathbb{R}^n, E)\bigr)$. Denoting by $[T_0 g]$ the equivalence class of $T_0 g$ with respect to the vector subspace of all elements of $\mathcal{L}_0(\mathbb{R}^m, L_q(\mathbb{R}^n, E))$ that vanish λ_m-a.e., we have $[T_0 g] \in L_p(\mathbb{R}^m, L_q(\mathbb{R}^n, E))$. Further, it follows from Corollary 6.8 that $[T_0 g] = [T_0 h]$ if $g, h \in \mathcal{S}_c(\mathbb{R}^{m+n}, E)$ coincide λ_{m+n}-a.e. Thus

$$T \colon S_c(\mathbb{R}^{m+n}, E) \to L_p\bigl(\mathbb{R}^m, L_q(\mathbb{R}^n, E)\bigr) \ , \quad [g] \mapsto [T_0 g]$$

is a well defined linear map.

6.20 Lemma *There is a unique extension*

$$\overline{T} \in \mathcal{L}\bigl(L_{(p,q)}(\mathbb{R}^{m+n}, E), L_p\bigl(\mathbb{R}^m, L_q(\mathbb{R}^n, E)\bigr)\bigr)$$

of T, and $\overline{T}$ is an isometry with a dense image.

Proof (i) For $f \in S_c(\mathbb{R}^{m+n}, E)$, let $g \in f \cap \mathcal{S}_c(\mathbb{R}^{m+n}, E)$. Then

$$\int_{\mathbb{R}^m} \|Tf\|^p_{L_q(\mathbb{R}^n,E)}\, dx = \int_{\mathbb{R}^m} \Bigl(\int_{\mathbb{R}^n} |g(x,y)|^q\, dy\Bigr)^{p/q} dx = \|g\|^p_{(p,q)} = \|f\|^p_{(p,q)} \ .$$

Therefore $T \in \mathcal{L}\bigl(S_c(\mathbb{R}^{m+n}, E), L_p(\mathbb{R}^m, L_q(\mathbb{R}^n, E))\bigr)$ is an isometry. Now it follows from Theorem VI.2.6 and Remark 6.19(a) that there is a uniquely determined isometric extension $\overline{T}$ of T.

(ii) We set $F := L_q(\mathbb{R}^n, E)$ and choose $w \in \mathcal{L}_p(\mathbb{R}^m, F)$ and $\varepsilon > 0$. It follows from Remark 6.19(d) that there is a $\varphi \in \mathcal{S}_c(\mathbb{R}^m, F)$ such that $\|w - \varphi\|_p < \varepsilon/2$. Let $\sum_{j=0}^r \chi_{A_j} \widehat{f}_j$ be the normal form of φ. Then $\bigcup_{j=0}^r A_j$ is bounded in $\mathbb{R}^m$, and $\alpha := \sum_{j=0}^r \lambda_m(A_j)$ is finite. In the case $\alpha = 0$, we have

$$\|w\|_p = \|w - T0\|_p < \varepsilon/2 \ .$$

In the case $\alpha > 0$, we choose for every $j \in \{0, \ldots, r\}$ a representative f_j of $\widehat{f}_j$ and a $\psi_j \in \mathcal{S}_c(\mathbb{R}^n, E)$ such that

$$\|\psi_j - f_j\|_q < \alpha^{-1/p}(r+1)^{-1/q'}\varepsilon \ .$$

Also let

$$h(x,y) := \sum_{j=0}^r \chi_{A_j}(x)\psi_j(y) \quad \text{for } (x,y) \in \mathbb{R}^{m+n} \ .$$

With $\psi_j = \sum_{k_j=0}^{s_j} \chi_{B_{k_j}} e_{k_j}$ for $j \in \{0, \ldots, r\}$, we then have

$$h = \sum_{j=0}^r \sum_{k_j=0}^{s_j} \chi_{A_j}\chi_{B_{k_j}} e_{k_j} = \sum_{j=0}^r \sum_{k_j=0}^{s_j} \chi_{A_j \times B_{k_j}} e_{k_j} \ ,$$

and we see that h belongs to $\mathcal{S}_c(\mathbb{R}^{m+n}, E)$. Finally, let g be the equivalence class of h in $L_0(\mathbb{R}^{m+n}, E)$. Then g belongs to $S_c(\mathbb{R}^{m+n}, E)$, and $Tg = \sum_{j=0}^r [\chi_{A_j}\psi_j]$. From Hölder's inequality (for sums) and the equality $\chi_A^2 = \chi_A$, it follows that

$$\begin{aligned}
\int_{\mathbb{R}^m} \|Tg - \varphi\|^p_F &= \int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} \Bigl|\sum_{j=0}^r \chi_{A_j}(x)\bigl(\psi_j(y) - f_j(y)\bigr)\Bigr|^q dy\Bigr]^{p/q} dx \\
&\le (r+1)^{p/q'} \int_{\mathbb{R}^m} \Bigl[\int_{\mathbb{R}^n} \sum_{j=0}^r \chi_{A_j}(x)\, |\psi_j(y) - f_j(y)|^q\, dy\Bigr]^{p/q} dx \\
&= (r+1)^{p/q'} \int_{\mathbb{R}^m} \Bigl[\sum_{j=0}^r \chi_{A_j}(x)\, \|\psi_j - f_j\|^q_F\Bigr]^{p/q} dx \\
&\le (r+1)^{p/q'} \sum_{j=0}^r \lambda_m(A_j)\, \|\psi_j - f_j\|^p_F \ .
\end{aligned}$$

Therefore,

$$\|Tg-\varphi\|_p \le \alpha^{1/p}\,(r+1)^{1/q'} \max_j \|\psi_j - f_j\|_F < \varepsilon/2\ ,$$

and consequently $\|Tg-w\|_p < \varepsilon$. Because this holds for every choice of w and ε, we see that the image of T, and *a fortiori* that of $\overline{T}$, is dense. ■

As usual, we lighten the notation by writing T for $\overline{T}$. In addition, as stated in Section 4, our notation for elements of Lebesgue spaces does not distinguish between cosets and their representatives. This means that for $f \in L_{(p,q)}(\mathbb{R}^{m+n}, E)$ we may write $Tf(x)$ as $f(x,\cdot\,)$. With these conventions, Lemma 6.20 says that

$$T\colon L_{(p,q)}(\mathbb{R}^{m+n},E) \to L_p(\mathbb{R}^m, L_q(\mathbb{R}^n,E))\ , \quad f \mapsto (x \mapsto f(x,\cdot\,)) \tag{6.22}$$

is a linear isometry whose image is dense.

Now it is easy to prove our continuous Minkowski's inequality.

6.21 Proposition (Minkowski's inequality for integrals) *For $1 \le q < \infty$, we have:*

(i) $$\Bigl(\int_{\mathbb{R}^n}\Bigl[\int_{\mathbb{R}^m}|f(x,y)|\,dx\Bigr]^q dy\Bigr)^{1/q} \le \int_{\mathbb{R}^m}\Bigl[\int_{\mathbb{R}^n}|f(x,y)|^q dy\Bigr]^{1/q} dx$$

for $f \in \mathcal{L}_0(\mathbb{R}^{m+n},E)$.

(ii) $$\Bigl(\int_{\mathbb{R}^n}\Bigl|\int_{\mathbb{R}^m} f(x,y)\,dx\Bigr|^q\,dy\Bigr)^{1/q} \le \int_{\mathbb{R}^m}\Bigl[\int_{\mathbb{R}^n}|f(x,y)|^q\,dy\Bigr]^{1/q} dx < \infty$$

for $f \in \mathcal{L}_{(1,q)}(\mathbb{R}^{m+n},E)$.

Proof In case (i), we can assume without loss of generality that

$$\int_{\mathbb{R}^m}\Bigl[\int_{\mathbb{R}^n}|f(x,y)|^q\,dy\Bigr]^{1/q} dx < \infty\ .$$

Then $|f|$ belongs to $\mathcal{L}_{(1,q)}(\mathbb{R}^{m+n},\mathbb{R})$, and the claim is a special case of (ii), with f replaced by $|f|$ and E by $\mathbb{R}$. Suppose therefore that $f \in \mathcal{L}_{(1,q)}(\mathbb{R}^{m+n},E)$. It follows from Lemma 6.20 and Theorem 2.11(i) (with E replaced by $L_q(\mathbb{R}^n,E)$) that

$$\int_{\mathbb{R}^m} Tf\,dx = \int_{\mathbb{R}^m} f(x,\cdot\,)\,dx \in L_q(\mathbb{R}^n,E)$$

and

$$\begin{aligned}\Bigl(\int_{\mathbb{R}^n}\Bigl|\int_{\mathbb{R}^m} f(x,y)\,dx\Bigr|^q\,dy\Bigr)^{1/q} &= \Bigl\|\int_{\mathbb{R}^m} Tf\,dx\Bigr\|_{L_q(\mathbb{R}^n,E)} \le \int_{\mathbb{R}^m}\|Tf\|_{L_q(\mathbb{R}^n,E)}\,dx\\ &= \int_{\mathbb{R}^m}\Bigl(\int_{\mathbb{R}^n}|f(x,y)|^q\,dy\Bigr)^{1/q} dx\ .\end{aligned}$$

This completes the proof. ■

A characterization of $L_p(\mathbb{R}^{m+n}, E)$

As another consequence of Lemma 6.20, we obtain an often useful generalization and sharpening of Fubini's theorem.

6.22 Theorem *For $1 \le p < \infty$,*

$$L_p(\mathbb{R}^{m+n}, E) \to L_p(\mathbb{R}^m, L_p(\mathbb{R}^n, E)) , \quad f \mapsto (x \mapsto f(x, \cdot\,))$$

is an isometric isomorphism.

Proof Suppose $v \in L_p(\mathbb{R}^m, L_p(\mathbb{R}^n, E))$. By Lemma 6.20, there is a sequence (f_j) in $L_p(\mathbb{R}^{m+n}, E)$ such that $\lim_j Tf_j = v$ in $L_p(\mathbb{R}^m, L_p(\mathbb{R}^n, E))$. Because T is a linear isometry, it follows easily that (f_j) is a Cauchy sequence in $L_p(\mathbb{R}^{m+n}, E)$. Denoting by f its limit in $L_p(\mathbb{R}^{m+n}, E)$, we have $Tf = v$. Therefore T is surjective. This proves the claim. ■

By means of this isometric isomorphism, we can identify the Banach spaces $L_p(\mathbb{R}^{m+n}, E)$ and $L_p(\mathbb{R}^m, L_p(\mathbb{R}^n, E))$:

$$L_p(\mathbb{R}^{m+n}, E) = L_p(\mathbb{R}^m, L_p(\mathbb{R}^n, E)) \ .$$

6.23 Remarks **(a)** The statement of Theorem 6.22 is false for $p = \infty$, that is

$$L_\infty(\mathbb{R}^{m+n}, E) \ne L_\infty\bigl(\mathbb{R}^m, L_\infty(\mathbb{R}^n, E)\bigr) \ .$$

Proof Take $A := \bigl\{\, (x, y) \in \mathbb{R}^2 \ ; \ 0 \le y \le x \le 1 \,\bigr\}$ and $f := \chi_A$. Because A is Lebesgue measurable, f belongs to $L_\infty(\mathbb{R}^2)$. If we set

$$g(x) := f(x, \cdot\,) = \begin{cases} \chi_{[0,x]} & \text{if } 0 \le x \le 1 \ , \\ 0 & \text{otherwise} \ , \end{cases}$$

then $g(x)$ belongs to $L_\infty(\mathbb{R})$, and $\|g(x)\|_\infty \le 1$ for $x \in \mathbb{R}$. But g nevertheless does not belong to $L_\infty\bigl(\mathbb{R}, L_\infty(\mathbb{R})\bigr)$, because the map $g : \mathbb{R} \to L_\infty(\mathbb{R})$ is not λ_1-measurable. To see this, it suffices by Theorem 1.4 to show that g is not λ_1-almost separable-valued. To check this, note that

$$\|g(x) - g(r)\|_{L_\infty(\mathbb{R})} = 1 \quad \text{for } r \in \mathbb{R}\backslash\{x\} \tag{6.23}$$

for $x \in (0, 1]$. Were g λ_1-almost separable valued, there would be a λ_1-null set $N \subset \mathbb{R}$ and a sequence (r_j) in $\mathbb{R}$ such that

$$\inf_{j\in\mathbb{N}} \|g(x) - g(r_j)\|_\infty < 1/2 \quad \text{for } x \in N^c \ . \tag{6.24}$$

Because $\lambda_1\bigl((0, 1]\backslash N\bigr) = 1$, the set $(0, 1]\backslash N$ is uncountable. Hence it follows from (6.23) that (6.24) cannot hold, and g is not λ_1-measurable. ■

(b) Generalizing Theorem 6.22, one can show that for any $p, q \in [1, \infty)$, the map

$$L_{(p,q)}(\mathbb{R}^{m+n}, E) \to L_p(\mathbb{R}^m, L_q(\mathbb{R}^n, E)) , \quad f \mapsto (x \mapsto f(x, \cdot\,))$$

is an isometric isomorphism. Therefore $L_{(p,q)}(\mathbb{R}^{m+n}, E)$ is complete. ■

A trace theorem

From Example IX.5.2 and the invariance of the Lebesgue measure under isometries, it follows that every hyperplane Γ in $\mathbb{R}^n$ is a λ_n-null set. Hence for $u \in L_p(\mathbb{R}^n)$, the restriction $u \,|\, \Gamma$, or *trace* of u on Γ, is not defined, because u can be "arbitrarily changed" on Γ. As another application of Fubini–Tonelli, we now show that one can nevertheless define such a trace on Γ for elements of certain vector subspaces of $L_p(\mathbb{R}^n)$. Of course, this is trivially the case for the vector subspace $C_c^1(\mathbb{R}^n)$. The significance of what follows is that this space is given not the supremum norm, but rather the L_p norm, with derivatives thrown into the mix. In the next section, we will understand better the significance of these subspaces of $L_p(\mathbb{R}^n)$.

Consider the coordinate hyperplane $\Gamma := \mathbb{R}^{n-1} \times \{0\}$, which we identify with $\mathbb{R}^{n-1}$. For $u \in C(\mathbb{R}^n)$, we let $\gamma u := u \,|\, \Gamma$ be the **trace** of u on Γ:

$$(\gamma u)(x) := u(x,0) \quad \text{for } x \in \mathbb{R}^{n-1} .$$

Then $\gamma \colon C_c^1(\mathbb{R}^n) \to C_c(\mathbb{R}^{n-1}),\ \ u \mapsto \gamma u$ is a well defined linear map.

Now take $1 \le p < \infty$, and give $C_c^1(\mathbb{R}^n)$ the norm

$$\|u\|_{1,p} := \Bigl(\|u\|_p^p + \sum_{j=1}^n \|\partial_j u\|_p^p \Bigr)^{1/p} .$$

Further, set

$$\widehat{H}_p^1(\mathbb{R}^n) := \bigl(C_c^1(\mathbb{R}^n), \|\cdot\|_{1,p} \bigr) .$$

Since $C_c(\mathbb{R}^{n-1})$ is a vector subspace of $L_p(\mathbb{R}^{n-1})$,

$$\gamma \colon \widehat{H}_p^1(\mathbb{R}^n) \to L_p(\mathbb{R}^{n-1}) , \quad u \mapsto \gamma u$$

is a well defined linear map, the **trace operator** with respect to $\Gamma = \mathbb{R}^{n-1}$. The following trace theorem shows that γ is continuous.

6.24 Proposition *$\gamma \in \mathcal{L}(\widehat{H}_p^1(\mathbb{R}^n), L_p(\mathbb{R}^{n-1}))$ for $1 \le p < \infty$.*

Proof Define $h \in C^1(\mathbb{R})$ by $h(t) := |t|^{p-1}\, t$. For $v \in C_c^1(\mathbb{R}^n)$, it follows from the chain rule that $\partial_n h(v) = h'(v)\partial_n v$. Since v has compact support, the fundamental theorem of calculus then implies that

$$-h(v(x,0)) = \int_0^\infty \partial_n h(v)(x,y)\, dy = \int_0^\infty h'(v(x,y))\partial_n v(x,y)\, dy \quad \text{for } x \in \mathbb{R}^{n-1} .$$

Because $h'(t) = p\,|t|^{p-1}$, we find

$$\begin{aligned} |v(x,0)|^p = \bigl|h(v(x,0))\bigr| &\le \int_0^\infty \bigl|h'(v(x,y))\bigr|\, |\partial_n v(x,y)|\, dy \\ &= p \int_0^\infty |v(x,y)|^{p-1}\, |\partial_n v(x,y)|\, dy . \end{aligned}$$

Also, Young's inequality gives $\xi^{p-1}\eta \le \frac{p-1}{p}\xi^p + \frac{1}{p}\eta^p$ for $\xi, \eta \in [0,\infty)$, so

$$|v(x,0)|^p \le (p-1)\int_0^\infty |v(x,y)|^p\,dy + \int_0^\infty |\partial_n v(x,y)|^p\,dy\ .$$

With $c_p := \max\{p-1, 1\}$, it now follows from Fubini–Tonelli that

$$\begin{aligned}&\int_{\mathbb{R}^{n-1}} |v(x,0)|^p\,dx\\ &\qquad \le c_p\Bigl(\int_{\mathbb{R}^{n-1}\times\mathbb{R}} |v(x,y)|^p\,d(x,y) + \int_{\mathbb{R}^{n-1}\times\mathbb{R}} |\partial_n v(x,y)|^p\,d(x,y)\Bigr)\ .\end{aligned} \tag{6.25}$$

Therefore

$$\|\gamma v\|_{L_p(\mathbb{R}^{n-1})} \le c\,\|v\|_{\widehat{H}^1_p(\mathbb{R}^n)} \quad \text{for } v \in \widehat{H}^1_p(\mathbb{R}^n)\ ,$$

where $c := c_p^{1/p}$. This proves the theorem. ■

6.25 Remark Denote by $\mathbb{H}^n$ the upper half-space of $\mathbb{R}^n$:

$$\mathbb{H}^n := \mathbb{R}^{n-1}\times(0,\infty) = \bigl\{\,(x,y)\in\mathbb{R}^{n-1}\times\mathbb{R}\ ;\ y>0\,\bigr\}\ .$$

Then $\Gamma = \mathbb{R}^{n-1}\times\{0\} = \partial\mathbb{H}^n$. If we set

$$\widehat{H}^1_p(\mathbb{H}^n) := \bigl(\bigl\{\,u\,|\,\mathbb{H}^n\ ;\ u\in C^1_c(\mathbb{R}^n)\,\bigr\},\ \|\cdot\|_{1,p}\bigr)\ ,$$

then $\widehat{H}^1_p(\mathbb{H}^n)$ is a vector subspace of $L_p(\mathbb{H}^n)$, and from a statement analogous to (6.25), it follows that

$$\gamma \in \mathcal{L}(\widehat{H}^1_p(\mathbb{H}^n), L_p(\mathbb{R}^{n-1}))\ .$$

In this case, γu for $u \in \widehat{H}^1_p(\mathbb{R}^n)$ is the **trace** of u on **the boundary** $\partial\mathbb{H}^n$. ■

Exercises

1 Suppose $B \in \mathcal{L}(n)$ and $a \in \mathbb{R}^{n+1}$. Denote by

$$Z_a(B) := \bigl\{\,(x,0)+ta\in\mathbb{R}^{n+1}\ ;\ x\in B,\ t\in[0,1]\,\bigr\}$$

the cylinder with base B and edge a, and let

$$K_a(B) := \bigl\{\,(1-t)(x,0)+ta\in\mathbb{R}^{n+1}\ ;\ x\in B,\ t\in[0,1]\,\bigr\}$$

be the cone with base B and tip a. Prove:

(a) $\lambda_{n+1}\bigl(Z_a(B)\bigr) = |a_{n+1}|\,\lambda_n(B)$;

(b) $\lambda_{n+1}\bigl(K_a(B)\bigr) = |a_{n+1}|\,\lambda_n(B)/(n+1)$.

If one interprets $|a_{n+1}|$ as the height of the cylinder $Z_a(B)$ or the cone $K_a(B)$, then (b) says the volume of an n-dimensional cone is equal to the total volume of n cylinders with the same base and height.

2 For $0 < r < a$, let $V_{a,r}$ be the region in $\mathbb{R}^3$ enclosed by the 2-torus $\mathsf{T}^2_{a,r}$. Show that $V_{a,r} = 2\pi^2 ar^2$.

3 Suppose $J \subset \mathbb{R}$ is an interval with endpoints $a := \inf J$ and $b := \sup J$. Also let $f \in \mathcal{L}_0(J, \mathbb{R}^+)$, and denote by

$$R_f := \{\, (x,t) \in \mathbb{R}^n \times J \; ; \; |x| \le f(t) \,\}$$

the **solid of revolution** arising by rotation of the graph of f around the t-axis. Prove that

$$\lambda_{n+1}(R_f) = \omega_n \int_a^b (f(t))^n \, dt \ ,$$

where ω_n is the volume of $\mathbb{B}^n$. Interpret this formula geometrically in the case $n = 2$.

4 Suppose K is compact in $\mathbb{R}^n$ and $\rho_K := \int_K \rho(x)\, dx > 0$ for $\rho \in \mathcal{L}_1(K, \mathbb{R}^+)$. Then

$$S(K,\rho) := \frac{1}{\rho_K} \int_K x\rho(x)\, dx \in \mathbb{R}^n$$

is the **centroid of K with respect to the density ρ**. We set $S(K) := S(K, \mathbf{1})$. Now suppose $J := [a,b]$ is a perfect, compact interval in $\mathbb{R}$, and let $f \in \mathcal{L}_0(J, \mathbb{R}^+)$. Also put

$$A_f := \{\, (x,y) \in \mathbb{R}^2 \; ; \; 0 \le y \le f(x), \; x \in J \,\} \ ,$$

and denote by R_f the solid of revolution in $\mathbb{R}^3$ generated by f (by rotating about the x-axis). Prove:

(a) For $f \in \mathcal{L}_1(J, \mathbb{R}^+)$,

$$S(A_f) = \big(S_1(A_f), S_2(A_f)\big) = \frac{1}{\|f\|_1} \Big(\int_a^b x f(x)\, dx, \frac{1}{2} \int_a^b \big(f(x)\big)^2 \, dx \Big) \ .$$

(b) For $f \in \mathcal{L}_2(J, \mathbb{R}^+)$,

$$S(R_f) = \Big(\frac{1}{\|f\|_2^2} \int_a^b t\big(f(t)\big)^2 \, dt, 0, 0 \Big) \ .$$

(c) For $f \in \mathcal{L}_1(J, \mathbb{R}^+)$, we have **Guldin's first rule**

$$\lambda_3(R_f) = \pi \int_a^b (f(x))^2 \, dx = 2\pi S_2(A_f) \lambda_2(A_f) \ .$$

In words, *the volume of a solid of revolution is equal to the area of a meridional slice*[7] *times the circumference of the circle drawn by the centroid of that slice during a full revolution.*[8]

5 (a) For $\alpha \in [0, \pi/2)$, let $a := (\cos\alpha, 0, \sin\alpha)$. Determine the centroid of the cylinder $Z_a(\mathbb{B}_2)$ and the cone $K_a(\mathbb{B}_2)$ with respect to the density $\mathbf{1}$.

(b) Let $A_\lambda := \{\, (x,y) \in \mathbb{R}^2 \; ; \; 0 \le y \le e^{-\lambda x}, \; x \ge 0 \,\}$ for $\lambda > 0$. Show that $S(A_\lambda) \in A_\lambda$.

(c) Give an example where $S(A_f) \notin A_f$.

---

[7] That is, the intersection with a plane containing the rotation axis.

[8] Guldin's first rule also holds for solids of revolution not arising from the rotation of a graph; see Exercise XII.1.11.

6 Let $K \subset \mathbb{R}^n$ be convex and compact. Check that $S(K,\rho) \in K$ for $\rho \in \mathcal{L}_1(K,\mathbb{R}^+)$.

7 Denote by $\Delta_n := \bigl\{ x \in \mathbb{R}^n \ ;\ x_j \geq 0,\ \sum_{j=1}^n x_j \leq 1 \bigr\}$ the **standard simplex** in $\mathbb{R}^n$. Prove:

(a) $\lambda_n(\Delta_n) = 1/n!$.

(b) $S(\Delta_n) = \bigl(1/(n+1), 1/(n+1), \ldots, 1/(n+1)\bigr)$.

8 Given $f \in \mathcal{L}_1(\mathbb{R}^m,\mathbb{K})$, $g \in \mathcal{L}_1(\mathbb{R}^n, E)$, define $F(x,y) := f(x)g(y)$ for $(x,y) \in \mathbb{R}^m \times \mathbb{R}^n$. Show that F belongs to $\mathcal{L}_1(\mathbb{R}^{m+n}, E)$ and that

$$\int_{\mathbb{R}^{m+n}} F(x,y)\, d(x,y) = \int_{\mathbb{R}^m} f(x)\, dx \int_{\mathbb{R}^n} g(y)\ dy\ .$$

9 For $D := \bigl\{ (x,y) \in \mathbb{R}^2\ ;\ x,y \geq 0,\ x+y \leq 1 \bigr\}$, show that

$$\int_D x^m y^n\, d(x,y) = \frac{1}{n+1}\,\mathsf{B}(m+1, n+2) \quad \text{for } m,n \in \mathbb{N}\ .$$

10 Show that $\int_{[0,1]\times[0,1]} y/\sqrt{x}\ d(x,y) = 1$.

11 Show that $\int_{\mathbb{R}^n} \partial_j \varphi\, dx = 0$ for $\varphi \in C_c^1(\mathbb{R}^n, E)$ and $j \in \{1,\ldots,n\}$.

12 For each of the following maps $f\colon (0,1)\times(0,1) \to \mathbb{R}$, calculate

$$\int_0^1\!\!\int_0^1 f(x,y)\, dx\, dy\ ,\quad \int_0^1\!\!\int_0^1 f(x,y)\, dy\, dx\ ,\quad \int_0^1\!\!\int_0^1 |f(x,y)|\, dx\, dy\ ,\quad \int_0^1\!\!\int_0^1 |f(x,y)|\, dy\, dx\ .$$

(a) $f(x,y) := (x-y)/(x^2+y^2)^{3/2}$.

(b) $f(x,y) := 1/(1-xy)^\alpha$ for $\alpha > 0$.

13 Let $p,q \in [1,\infty]$. Prove:

(a) $L_p(\mathbb{R}^n) \not\subset L_q(\mathbb{R}^n)$ if $p \neq q$.

(b) if $X \subset \mathbb{R}^n$ is open and bounded, then $L_p(X) \subsetneq L_q(X)$ if $p > q$.

7 The convolution

In this section we use the translation invariance of the Lebesgue measure to introduce a new product on $L_1(\mathbb{R}^n)$, the convolution, which rests on the Lebesgue integral. We show that this operation is defined not only on $L_1(\mathbb{R}^n)$ but also on other function spaces, and that it has important smoothing properties. Among its applications are certain approximation theorems which we prove here for their great usefulness in later constructions.

We will consider mainly spaces of $\mathbb{K}$-valued functions defined on all of $\mathbb{R}^n$. For such spaces we omit the domain and image from the notation. In other words, if $\mathfrak{F}(\mathbb{R}^n) = \mathfrak{F}(\mathbb{R}^n, \mathbb{K})$ is a vector space of $\mathbb{K}$-valued functions on $\mathbb{R}^n$, we write simply $\mathfrak{F}$ if there is no risk of confusion. Thus L_p stands for $L_p(\mathbb{R}^n) = L_p(\mathbb{R}^n, \mathbb{K})$, and so on. Also $\int f\,dx$ will always mean $\int_{\mathbb{R}^n} f\,dx$.

Defining the convolution

Let F be a $\mathbb{K}$-vector space. For $f \in \mathrm{Funct}(\mathbb{R}^n, F)$, we define another function $\check{f} \in \mathrm{Funct}(\mathbb{R}^n, F)$ by $\check{f}(x) := f(-x)$, where $x \in \mathbb{R}^n$. The map $f \mapsto \check{f}$ is called **inversion** (about the origin).

Recall from IX.5.15 the definition of the translation group $\mathfrak{T} := \{\, \tau_a \,;\, a \in \mathbb{R}^n \,\}$. Now we define an action[1] of this group on $\mathrm{Funct}(\mathbb{R}^n, F)$ by

$$\mathfrak{T} \times \mathrm{Funct}(\mathbb{R}^n, F) \to \mathrm{Funct}(\mathbb{R}^n, F) \ , \quad (\tau_a, f) \mapsto \tau_a f \ , \tag{7.1}$$

where

$$\tau_a f(x) := f(x - a) \quad \text{for } a, x \in \mathbb{R}^n \ . \tag{7.2}$$

Therefore

$$\tau_a f = f \circ \tau_{-a} = (\tau_{-a})^* f \ ,$$

where $(\tau_{-a})^*$ is the pull back defined in Section VIII.3.

7.1 Remarks **(a)** For $f \in \mathrm{Funct} := \mathrm{Funct}(\mathbb{R}^n, \mathbb{K})$, we have $\check{f} = (-\mathrm{id}_{\mathbb{R}^n})^* f$.

(b) Inversion is an involutive[2] vector space isomorphism on Funct and on $\mathcal{L}_p$ for $p \in [1, \infty] \cup \{0\}$.

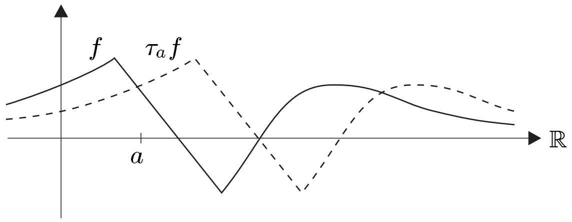
(c) Suppose $E \in \{\, BC^k, BUC^k, C_0 \ ;\ k \in \mathbb{N} \,\}$. Then inversion belongs to $\mathcal{L}\mathrm{aut}(E)$.

(d) For $f \in \mathrm{Funct}$ and $x \in \mathbb{R}^n$, we have

$$(\tau_{-x} f)\check{\ }(y) = \tau_x \check{f}(y) = f(x - y) \quad \text{for } y \in \mathbb{R}^n \ .$$

[1] See Exercise I.7.6.
[2] A map $f \in X^X$ is said to be **involutive** if $f \circ f = \mathrm{id}_X$.

(e) Suppose $n = 1$ and $a > 0$. Then $\tau_a : \mathbb{R} \to \mathbb{R},\ x \mapsto x + a$ is the **right translation** on $\mathbb{R}$ by a. Definition (7.2) means that τ_a also translates the graph of f to the right by a.



Therefore $\mathfrak{T}$ acts as a **right translation** on $\mathrm{Funct}(\mathbb{R}, F)$, which clarifies defining $\tau_a f$ as the pull back of the left translation τ_{-a} of $\mathbb{R}$. ■

Take $f, g \in \mathcal{L}_1$ and $x \in \mathbb{R}^n$, and let O be open in $\mathbb{K}$.

$$(\tau_{-x} f)^{-1}(O) = (f \circ \tau_x)^{-1}(O) = \tau_{-x}(f^{-1}(O)) .$$

Therefore it follows from Corollary 1.5 and Lemma IX.5.16 that $(\tau_{-x} f)^{-1}(O)$ is measurable. Hence, again by Corollary 1.5, $\tau_{-x} f$ belongs to $\mathcal{L}_0$. Now we deduce from Remark 1.2(d) and parts (b) and (d) of Remark 7.1 that $y \mapsto f(x - y)g(y)$ belongs to $\mathcal{L}_0$ for every $x \in \mathbb{R}^n$. If this function is integrable, we define the **convolution of *f* with *g* at *x*** by

$$f * g(x) := \int f(x - y)g(y)\, dy .$$

We say f and g are **convolvable** if $f * g(x)$ is defined for almost every $x \in \mathbb{R}^n$. In this case the a.e.-defined function

$$f * g := \big(x \mapsto f * g(x)\big)$$

is called the **convolution** of f with g. If f and g are convolvable and $(f * g)^p$ is integrable (or $f * g$ is essentially bounded for $p = \infty$), we write $f * g \in \mathcal{L}_p$, in a slight abuse of notation.[3])

We now show that every pair $(f, g) \in \mathcal{L}_p \times \mathcal{L}_1$ with $p \in [1, \infty]$ is convolvable. The following observation will be helpful.

7.2 Lemma *For $f \in \mathcal{L}_0$ and $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$, let*

$$F_1(x, y) := f(x) \quad \text{and} \quad F_2(x, y) := f(x - y) .$$

Then F_1 and F_2 belong to $\mathcal{L}_0(\mathbb{R}^{2n})$.

[3]We literally mean that the trivial extension of $f * g$ belongs to $\mathcal{L}_p$.

Proof (i) Suppose O is open in $\mathbb{K}$ and $A := f^{-1}(O)$. Then A belongs to $\mathcal{L}(n)$. Therefore Remark 6.1(b) and Proposition 6.2 show that $F_1^{-1}(O) = A \times \mathbb{R}^n$ is λ_{2n}-measurable. Now the claim for F_1 follows from Corollary 1.5.

(ii) Set $\varphi(x,y) := (x-y, y)$ for $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$. Then $\varphi \in \mathcal{L}\mathrm{aut}(\mathbb{R}^{2n})$ and $F_2 = F_1 \circ \varphi$. The claim then follows from (i) and Theorem IX.5.12. ∎

7.3 Theorem *Suppose $p \in [1,\infty]$ and $(f,g) \in \mathcal{L}_p \times \mathcal{L}_1$.*

(i) *f and g are convolvable.*

(ii) (Young's inequality) *$f * g \in \mathcal{L}_p$ and $\|f * g\|_p \le \|f\|_p \, \|g\|_1$.*

Proof (a) Suppose first that $p \in [1,\infty)$. By Lemma 7.2 and Remark 1.2(d), the map $(x,y) \mapsto f(x-y)g(y)$ belongs to $\mathcal{L}_0(\mathbb{R}^{2n})$. Using Hölder's inequality, we deduce that

$$\begin{aligned}\int |f(x-y)g(y)|\,dy &= \int |f(x-y)|\,|g(y)|^{1/p}\,|g(y)|^{1/p'}\,dy\\ &\le \Bigl(\int |f(x-y)|^p\,|g(y)|\,dy\Bigr)^{1/p}\Bigl(\int |g(y)|\,dy\Bigr)^{1/p'} .\end{aligned}$$

From this and Tonelli's theorem, we get

$$\begin{aligned}\int\Bigl(\int |f(x-y)g(y)|\,dy\Bigr)^p dx &\le \|g\|_1^{p/p'} \int\!\!\int |f(x-y)|^p\,|g(y)|\,dy\,dx\\ &= \|g\|_1^{p/p'} \int\!\!\int |f(x-y)|^p\;dx\,|g(y)|\,dy\\ &= \|g\|_1^{1+p/p'}\,\|f\|_p^p < \infty ,\end{aligned}$$

where in the last step we once more used the translation invariance of the Lebesgue integral. Thus we find[4]

$$\Bigl(\int\Bigl[\int |f(x-y)g(y)|\,dy\Bigr]^p dx\Bigr)^{1/p} \le \|f\|_p\,\|g\|_1 < \infty . \tag{7.3}$$

Now from Remark 3.11(c), we conclude that $\int |f(x-y)g(y)|\,dy < \infty$ for almost every $x \in \mathbb{R}^n$; by Remark 3.11(a), this suffices to show that f and g are convolvable. Part (ii) of the theorem now follows from (7.3).

(b) In the case $p = \infty$, we have

$$\int |f(x-y)g(y)|\,dy \le \|f\|_\infty\,\|g\|_1 < \infty \quad \text{for almost every } x \in \mathbb{R}^n ,$$

which immediately implies (i) and (ii). ∎

[4]Those readers who worked through the last part of the previous section will recognize that this bound can also be easily derived from Minkowski's inequality for integrals.

7.4 Corollary *Let* $\big([f],[g]\big) \in L_p \times L_1$ *with* $p \in [1,\infty]$*. Then*

$$f * g = \overset{*}{f} * \overset{*}{g} \qquad \text{a.e. in } \mathbb{R}^n$$

for $(\overset{*}{f}, \overset{*}{g}) \in \big([f],[g]\big)$.

Proof By Theorem 7.3, $f * g$, $\overset{*}{f} * \overset{*}{g}$, and $f * \overset{*}{g}$ are defined a.e. and belong to $\mathcal{L}_p$. Because

$$f * g - \overset{*}{f} * \overset{*}{g} = f * \big(g - \overset{*}{g}\big) + \big(f - \overset{*}{f}\big) * \overset{*}{g}\ ,$$

we obtain from Young's inequality that

$$\big\|f * g - \overset{*}{f} * \overset{*}{g}\big\|_p \le \big\|f\big\|_p \,\big\|g - \overset{*}{g}\big\|_1 + \big\|f - \overset{*}{f}\big\|_p \,\big\|\overset{*}{g}\big\|_1 = 0\ ,$$

from which the claim follows. ■

We can now define the convolution for elements of $L_p \times L_1$ with $p \in [1,\infty]$: indeed, Corollary 7.4 guarantees that the map

$$*\colon L_p \times L_1 \to L_p\ , \quad \big([f],[g]\big) \mapsto [f * g]$$

is well defined. We call this the $*$ the **convolution product** on $L_p \times L_1$, and $[f] * [g] := [f * g]$ **convolution of** $[f]$ with $[g]$. It is clear that the convolution can also be defined on $L_1 \times L_p$, and we use the symbol $*$ for this as well.

The translation group

To be able to better explore further properties of the convolution, we first gather some important definitions and facts about the representation of the translation group $(\mathbb{R}^n, +)$ on function spaces.

Let F be a $\mathbb{K}$-vector space and let V be a vector subspace of $\text{Funct}(\mathbb{R}^n, F)$ that is **invariant** under the action (7.1) of the translation group $\mathfrak{T}$ of $\mathbb{R}^n$, meaning that $\tau_a(V) \subset V$ for all $a \in \mathbb{R}^n$. By **restriction**, (7.1) induces an action

$$\mathfrak{T} \times V \to V\ , \quad (\tau_a, v) \mapsto \tau_a v$$

of the translation group $\mathfrak{T}$ on V. For every $a \in \mathbb{R}^n$, the map $T_a := (v \mapsto \tau_a v)$ is a linear map from V into itself. Because

$$\tau_a \tau_b v = \tau_{a+b} v \quad \text{and} \quad \tau_0 v = v\ ,$$

T_a is a vector space automorphism of V and $(T_a)^{-1} = T_{-a}$. Hence[5]

$$(\mathbb{R}^n, +) \to \text{Aut}(V)\ , \quad a \mapsto T_a$$

[5]See Remarks I.12.2(d) and I.7.6(e).

is a group homomorphism, a **linear representation** of the group $(\mathbb{R}^n, +)$ on V. In particular,

$$\mathfrak{T}_V := \{T_a \in \mathrm{Aut}(V) \ ; \ a \in \mathbb{R}^n\}$$

is a subgroup of $\mathrm{Aut}(V)$, called the **group of translations on V**. Instead of T_a, we tend to use the same symbol τ_a if there is no fear of misunderstanding. The invariance of V under (7.1) is also expressed by saying that $(\mathbb{R}^n, +)$ is **linearly representable** on V.

If V is a (semi)normed vector space, the group $\mathfrak{T}_V$ is said to be **strongly continuous** if $\lim_{a\to 0} \tau_a v = v$ for every $v \in V$.

7.5 Remarks **(a)** $(\mathbb{R}^n, +)$ is linearly representable on Funct and on $B := B(\mathbb{R}^n)$.

(b) $(\mathbb{R}^n, +)$ is linearly representable on $\mathcal{L}_\infty$, and $\|\tau_a f\|_\infty = \|f\|_\infty$ for $f \in \mathcal{L}_\infty$.

Proof Take $f \in \mathcal{L}_\infty$. For every $\alpha > \|f\|_\infty$ there is a set N of Lebesgue measure zero such that $|f(x)| \le \alpha$ for $x \in N^c$. By translation invariance (Theorem IX.5.17), $N_a := \tau_a(N)$ also has measure zero and

$$|\tau_a f(x)| = |f(x-a)| \le \alpha \quad \text{for } x \in N_a^c \ .$$

Therefore $\tau_a f$ is essentially bounded, and $\|\tau_a f\|_\infty \le \|f\|_\infty$. The claim follows since

$$\|f\|_\infty = \|\tau_{-a}(\tau_a f)\|_\infty \le \|\tau_a f\|_\infty \ . \ \blacksquare$$

(c) The translation groups $\mathfrak{T}_B$ and $\mathfrak{T}_{\mathcal{L}_\infty}$ are not strongly continuous.

Proof $\|\tau_a \chi_{\mathbb{B}^n} - \chi_{\mathbb{B}^n}\|_\infty = 1$ for $a \in \mathbb{R}^n \backslash \{0\}$. $\blacksquare$

(d) If $\mathfrak{T}_V$ is strongly continuous, then

$$(a \mapsto \tau_a f) \in C(\mathbb{R}^n, V) \quad \text{for } f \in V \ .$$

Proof This follows from $\tau_a f - \tau_b f = \tau_{a-b}(\tau_b f) - \tau_b f$ for $f \in V$ and $a, b \in \mathbb{R}^n$. $\blacksquare$

7.6 Theorem *Suppose $V = \mathcal{L}_p$ with $p \in [1, \infty)$ or $V = BUC^k$ with $k \in \mathbb{N}$. Then $(\mathbb{R}^n, +)$ is linearly representable on V, and the translation group $\mathfrak{T}_V$ is strongly continuous. Also $\|\tau_a f\|_V = \|f\|_V$ for $a \in \mathbb{R}^n$ and $f \in V$.*

Proof (i) We consider first the case $V = BUC^k$. Take $f \in BUC^k$, $a \in \mathbb{R}^n$, and $\varepsilon > 0$. Then there is $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in \mathbb{R}^n$ satisfying $|x - y| < \delta$. It follows that

$$|\tau_a f(x) - \tau_a f(y)| = |f(x-a) - f(y-a)| < \varepsilon \tag{7.4}$$

for $x, y \in \mathbb{R}^n$ such that $|x - y| < \delta$. Therefore $\tau_a f$ belongs to BUC, and because

$$\partial^\alpha \tau_a f = \tau_a \partial^\alpha f \quad \text{for } \alpha \in \mathbb{N}^n \text{ and } |\alpha| \le k \ , \tag{7.5}$$

we obtain $\tau_a f \in BUC^k$. Consequently $(\mathbb{R}^n, +)$ is linearly representable on BUC^k. From Remark 7.5(b) and (7.5), we find $\|\tau_a f\|_{BC^k} = \|f\|_{BC^k}$.

Now take $x \in \mathbb{R}^n$. If $|a| < \delta$, we can set $y = x + a$ in (7.4), and we get

$$|\tau_a f(x) - f(x)| < \varepsilon \quad \text{for } x \in \mathbb{R}^n ,$$

that is, $\|\tau_a f - f\|_\infty < \varepsilon$ for $a \in \delta\mathbb{B}^n$. Analogously, we can show with (7.5) that there is a $\delta_1 > 0$ such that $\|\tau_a f - f\|_{BC^k} < \varepsilon$ for $a \in \delta_1\mathbb{B}^n$. Therefore $\mathfrak{T}_{BUC^k}$ is strongly continuous.

(ii) Let $p \in [1, \infty)$ and $f \in \mathcal{L}_p$. The equality $\|\tau_a f\|_p = \|f\|_p$ follow from the translation invariance of the Lebesgue integral.

Now take $\varepsilon > 0$. By Theorem 4.14, there is a $g \in C_c$ such that $\|f - g\|_p < \varepsilon/3$. Because g has compact support, there is a compact subset K of $\mathbb{R}^n$ such that $\operatorname{supp}(\tau_a g) \subset K$ for $|a| \le 1$. Also, since g is uniformly continuous, there exists $\delta \in (0, 1]$ such that

$$\|\tau_a g - g\|_\infty < \varepsilon / 3\lambda_n(K)^{1/p} \quad \text{for } a \in \delta\mathbb{B}^n .$$

Suppose $a \in \delta\mathbb{B}^n$. Because $\operatorname{supp}(\tau_a g - g) \subset K$, Theorem 5.1(iv) implies that

$$\|\tau_a g - g\|_p < \varepsilon/3 \quad \text{for } a \in \delta\mathbb{B}^n .$$

Since

$$\|\tau_a f - f\|_p \le \|\tau_a f - \tau_a g\|_p + \|\tau_a g - g\|_p + \|g - f\|_p$$

and $\|\tau_a f - \tau_a g\|_p = \|\tau_a(f - g)\|_p = \|f - g\|_p$, we get $\|\tau_a f - f\|_p < \varepsilon$ for $a \in \delta\mathbb{B}^n$, and we are done. ■

We now define an action of $\mathfrak{T}$ on L_p for $p \in [1, \infty]$. By Remark 7.5(b) and Theorem 7.6, τ_a is an isometry of $\mathcal{L}_p$ for every $a \in \mathbb{R}^n$. Therefore the map

$$L_p \to L_p , \quad [f] \mapsto [\tau_a f]$$

is well defined for every $a \in \mathbb{R}^n$. We denote it by τ_a also, that is, we set

$$\tau_a[f] := [\tau_a f] \quad \text{for } f \in L_p \text{ and } a \in \mathbb{R}^n .$$

Then

$$\|\tau_a[f]\,\|_p = \|\,[\tau_a f]\,\|_p = \|\tau_a f\|_p = \|f\|_p = \|\,[f]\,\|_p . \tag{7.6}$$

Clearly

$$\mathfrak{T} \times L_p \to L_p , \quad (\tau_a, f) \mapsto \tau_a f$$

is an action of the translation group $\mathfrak{T}$ of $\mathbb{R}^n$ on L_p. By Remark 7.5(b) and Theorem 7.6, $T_a := (f \mapsto \tau_a f)$ is a linear isometry on L_p for every $a \in \mathbb{R}^n$. Again writing T_a as τ_a, we conclude that

$$(\mathbb{R}^n, +) \to \mathcal{L}\text{aut}(L_p) , \quad a \mapsto \tau_a$$

is a representation of the additive group of $\mathbb{R}^n$ by linear isometries on L_p. In particular, the **translation group on L_p**, namely

$$\mathfrak{T}_{L_p} := \left\{\, \tau_a \in \mathcal{L}\text{aut}(L_p) \;;\; a \in \mathbb{R}^n \,\right\} ,$$

is a subgroup of $\mathcal{L}\text{aut}(L_p)$ consisting of isometries.

7.7 Corollary *The translation group on L_p is strongly continuous for $1 \le p < \infty$.*

Proof This is an immediate consequence of Theorem 7.6 and (7.6). ■

Elementary properties of the convolution

After this digression about the translation group, we return to the convolution and derive its chief properties.

7.8 Theorem *Consider $(f, g) \in L_p \times L_1$ with $p \in [1, \infty]$.*

(i) *The convolution $f * g$ belongs to L_p, and satisfies* **Young's inequality**

$$\|f * g\|_p \le \|f\|_p \, \|g\|_1 \ .$$

(ii) $f * g = g * f$.

(iii) *If $p = \infty$, the convolution $f * g$ belongs to*[6] *BUC.*

(iv) *For $\varphi \in BC^k$, we have $\varphi * g \in BUC^k$,*

$$\partial^\alpha (\varphi * g) = \partial^\alpha \varphi * g \quad \text{for } \alpha \in \mathbb{N}^n \ , \quad |\alpha| \le k \ ,$$

and $\|\varphi * g\|_{BC^k} \le \|\varphi\|_{BC^k} \, \|g\|_1$.

Proof (i) follows from Theorem 7.3(ii) and Corollary 7.4.

(ii) Take $x \in \mathbb{R}$ and let $\overset{*}{f}$ and $\overset{*}{g}$ be representatives of f and g. Also set $\psi(y) := x - y$ for $y \in \mathbb{R}^n$. Then ψ is an involutive isometry of $\mathbb{R}^n$. It follows from Theorem 7.3(i) and Example 6.6(b) that

$$\begin{aligned} \overset{*}{f} * \overset{*}{g}(x) &= \int \overset{*}{f}(x-y) \overset{*}{g}(y) \, dy = \int \bigl(\overset{*}{f} \circ \psi\bigr) \bigl((\overset{*}{g} \circ \psi) \circ \psi\bigr) \, dy \\ &= \int (\overset{*}{g} \circ \psi) \overset{*}{f} \, dy = \int \overset{*}{g}(x-y) \overset{*}{f}(y) \, dy = \overset{*}{g} * \overset{*}{f}(x) \ . \end{aligned}$$

Therefore $f * g = g * f$ by Corollary 7.4.

(iii) The motion invariance of the Lebesgue integral yields $\|\breve{g}\|_1 = \|g\|_1$. From part (ii), and because the elements of $\mathfrak{T}_{L_1}$ are isometries, we then get

$$\begin{aligned} \bigl|\overset{*}{f} * \overset{*}{g}(x) - \overset{*}{f} * \overset{*}{g}(y)\bigr| &\le \int \bigl|\overset{*}{f}(z)\bigl(\overset{*}{g}(x-z) - \overset{*}{g}(y-z)\bigr)\bigr| \, dz \le \|\overset{*}{f}\|_\infty \, \|\tau_x \breve{g} - \tau_y \breve{g}\|_1 \\ &= \|f\|_\infty \, \|\tau_y(\tau_{x-y}\breve{g} - \breve{g})\|_1 = \|f\|_\infty \, \|\tau_{x-y}\breve{g} - \breve{g}\|_1 \end{aligned}$$

for $x, y \in \mathbb{R}^n$. Because $\breve{g} \in L_1$, the strong continuity of $\mathfrak{T}_{L_1}$ together with part (i) implies that $\overset{*}{f} * \overset{*}{g} \in BUC$. The claim follows.

[6] See Theorem 4.18.

(iv) In view of (iii), it suffices to consider the case $k \geq 1$. To this end we define $h(x,y) := \varphi(x-y)g(y)$ for $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$. Then h satisfies the assumptions of Theorem 3.18, and it follows that $\partial_j(\varphi * g) = \partial_j\varphi * g$ for $j \in \{1,\ldots,n\}$. By (iii) and Theorem VII.2.10, we have $\varphi * g \in BUC^1$. We now see inductively that $\varphi * g$ belongs to BUC^k and satisfies $\partial^\alpha(\varphi * g) = \partial^\alpha\varphi * g$ for every $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq k$. Finally, by (i), we have

$$\begin{aligned}\|\varphi * g\|_{BC^k} &= \max_{|\alpha|\leq k} \|\partial^\alpha(\varphi * g)\|_\infty = \max_{|\alpha|\leq k} \|(\partial^\alpha\varphi) * g\|_\infty \leq \bigl(\max_{|\alpha|\leq k} \|\partial^\alpha\varphi\|_\infty\bigr)\, \|g\|_1 \\ &= \|\varphi\|_{BC^k}\, \|g\|_1 \ . \ \blacksquare\end{aligned}$$

7.9 Corollary

(i) *Let* $p \in (1,\infty)$ *and* $k \in \mathbb{N}$. *The convolution satisfies*

$$* \in \begin{cases} \mathcal{L}^2_{\mathrm{sym}}(L_1, L_1) \ , \\ \mathcal{L}(L_p, L_1; L_p) \ , \\ \mathcal{L}(L_\infty, L_1; BUC) \ , \\ \mathcal{L}(BC^k, L_1; BUC^k) \ , \end{cases}$$

and all these maps have norm at most 1.

(ii) $(L_1, +, *)$ *is a commutative Banach algebra without a multiplicative identity.*

Proof (i) and the first statement of (ii) follow immediately from Theorem 7.8. We now assume there is $e \in L_1$ such that $e * f = f$ for every $f \in L_1$. We choose a representative $\overset{*}{e}$ of e and then find by Exercise 2.15 a $\delta > 0$ such that

$$\Bigl|\int_{\delta\mathbb{B}^n} \overset{*}{e}(x-y)\,dy\Bigr| = \Bigl|\int_{\mathbb{B}^n(x,\delta)} \overset{*}{e}(z)\,dz\Bigr| < 1 \quad \text{for } x \in \mathbb{R}^n \ .$$

Furthermore, there is a set N of Lebesgue measure zero such that $\chi_{\delta\mathbb{B}^n}(x) = \overset{*}{e} * \chi_{\delta\mathbb{B}^n}(x)$ for $x \in N^c$. However, for $x \in \delta\mathbb{B}^n \cap N^c$, we have

$$1 = \chi_{\delta\mathbb{B}^n}(x) = \overset{*}{e} * \chi_{\delta\mathbb{B}^n}(x) = \int_{\mathbb{R}^n} \overset{*}{e}(x-y)\chi_{\delta\mathbb{B}^n}(y)\,dy = \int_{\delta\mathbb{B}^n} \overset{*}{e}(x-y)\,dy < 1 \ ,$$

which is not possible. ■

7.10 Theorem (additivity of supports) *Suppose* $f, g \in \mathcal{L}_0$ *are convolvable and* f *has compact support. Then*

$$\mathrm{supp}(f * g) \subset \mathrm{supp}(f) + \mathrm{supp}(g) \ .$$

Proof (i) We can assume $f * g \neq 0$. For $x \in [f * g \neq 0]$, there is a $y \in \mathbb{R}^n$ such that $f(x-y)g(y) \neq 0$. It follows that $y \in \mathrm{supp}(g)$ and $x \in y + \mathrm{supp}(f)$, and thus x belongs to $\mathrm{supp}(f) + \mathrm{supp}(g)$. Hence $[f * g \neq 0] \subset \mathrm{supp}(f) + \mathrm{supp}(g)$.

(ii) We show that $\mathrm{supp}(f) + \mathrm{supp}(g)$ is closed. Let (x_k) be a sequence in $\mathrm{supp}(f) + \mathrm{supp}(g)$ such that $x_k \to x$ for some $x \in \mathbb{R}^n$. Then there are sequences (a_k) in $\mathrm{supp}(f)$ and (b_k) in $\mathrm{supp}(g)$ such that $x_k = a_k + b_k$ for $k \in \mathbb{N}$. Because $\mathrm{supp}(f)$ is compact, there is a subsequence $(a_{k_\ell})_{\ell\in\mathbb{N}}$ of (a_k) and an $a \in \mathrm{supp}(f)$ such that $a_{k_\ell} \to a$ as $\ell \to \infty$. Thus $b_{k_\ell} = x_{k_\ell} - a_{k_\ell} \to x - a$ as $k \to \infty$. Because $\mathrm{supp}(g)$ is closed, we know $x-a$ belongs to $\mathrm{supp}(g)$. Hence there exists $b \in \mathrm{supp}(g)$ such that $x = a+b$. This shows that $\mathrm{supp}(f)+\mathrm{supp}(g)$ is closed. The claim follows from Corollary III.2.13. ■

Approximations to the identity

We saw in Corollary 7.9 that the convolution algebra L_1 has no multiplicative identity. However, the next theorem secures the existence of "approximations to the identity", elements $\varphi \in L_1$ that satisfy $\|\varphi * f - f\|_1 < \varepsilon$ for every $f \in L_1$ (for a given $\varepsilon > 0$).

7.11 Theorem (approximation theorem) *Given $E \in \{\, L_p \ ;\ 1 \le p < \infty \,\}$ or $E \in \{\, BUC^k \ ;\ k \in \mathbb{N} \,\}$, set $\varphi \in \mathcal{L}_1$ and*

$$a := \int \varphi \, dx \ , \quad \varphi_\varepsilon(x) := \varepsilon^{-n}\varphi(x/\varepsilon) \quad \textit{for } x \in \mathbb{R}^n \ , \quad \varepsilon > 0 \ .$$

Then $\lim_{\varepsilon\to 0} \varphi_\varepsilon * f = af$ in E for $f \in E$.

Proof (i) Fix $\varepsilon > 0$. By the substitution rule — Example 6.6(b) — we know that $\varphi_\varepsilon \in L_1$ and $\int \varphi_\varepsilon \, dx = a$. Thus Theorem 7.8 shows that $\varphi_\varepsilon * f \in E$ for $f \in E$.

(ii) To prove the limit as $\varepsilon \to 0$, consider first the case $E = L_p$. Take $f \in \mathcal{L}_p$ and $\varepsilon > 0$. By Theorem 7.3(i) and the proof of Theorem 7.8(ii), and using the transformation $y \mapsto y/\varepsilon$ in Example 6.6(b), we obtain

$$\begin{aligned} \varphi_\varepsilon * f(x) - af(x) = f * \varphi_\varepsilon(x) - af(x) &= \int \bigl[f(x-y) - f(x)\bigr]\varphi_\varepsilon(y)\, dy \\ = \int \bigl[f(x-\varepsilon z) - f(x)\bigr]\varphi(z)\, dz &= \int \bigl[\tau_{\varepsilon z} f(x) - f(x)\bigr]\varphi(z)\, dz \end{aligned} \tag{7.7}$$

for almost every $x \in \mathbb{R}^n$. Corollary 7.7 and Remark 7.5(d) imply that

$$\bigl(z \mapsto (\tau_{\varepsilon z} f - f)\bigr) \in C(\mathbb{R}^n, E) \quad \text{for } \varepsilon > 0 \ , \tag{7.8}$$

and

$$\lim_{\varepsilon\to 0} \|\tau_{\varepsilon z} f - f\|_E = 0 \quad \text{for } z \in \mathbb{R}^n \ . \tag{7.9}$$

Now set

$$g^\varepsilon(z) := (\tau_{\varepsilon z} f - f)\varphi(z) \quad \text{for } z \in \mathbb{R}^n \text{ and } \varepsilon > 0 \ .$$

Then it follows from (7.8), Theorem 1.17, and Remark 1.2(d) that g^ε belongs to $\mathcal{L}_0(\mathbb{R}^n, E)$ for every $\varepsilon > 0$. Because $\|\tau_{\varepsilon z} f\|_E = \|f\|_E$, we also derive from the triangle inequality that

$$\|g^\varepsilon(z)\|_E \le 2\,\|f\|_E\,|\varphi(z)| \quad \text{for } z \in \mathbb{R}^n \text{ and } \varepsilon > 0\ .$$

Because $\varphi \in \mathcal{L}_1(\mathbb{R}^n)$, we therefore conclude that $g^\varepsilon \in \mathcal{L}_1(\mathbb{R}^n, E)$. Then (7.7) and Theorem 2.11(i) imply the bound[7]

$$\|\varphi_\varepsilon * f - af\|_E = \Big\|\int g^\varepsilon(z)\,dz\Big\|_E \le \int \|g^\varepsilon(z)\|_E\,dz\ .$$

Now the dominated convergence theorem shows that $\varphi_\varepsilon * f$ converges in E to af as $\varepsilon \to 0$, because, by (7.9), we have $\lim_{\varepsilon\to 0}\|g^\varepsilon(z)\|_E = 0$ for almost every $z \in \mathbb{R}^n$.

(iii) Now suppose $f \in BUC^k$. If $\varphi = 0$ λ_n-a.e., the claim is obviously true. So suppose $m := \int |\varphi|\,dx > 0$. From Theorem 7.8(ii) and (iv), it follows that

$$\partial^\alpha(\varphi_\varepsilon * f - af) = \varphi_\varepsilon * \partial^\alpha f - a\partial^\alpha f \quad \text{for } \alpha \in \mathbb{N}^n \text{ and } |\alpha| \le k\ .$$

Therefore it suffices to consider the case $k = 0$.

Let $\eta > 0$. Then there is a $\delta > 0$ such that

$$|f(x-y) - f(x)| \le \eta/2m \quad \text{for } x, y \in \mathbb{R}^n\ , \quad |y| < \delta\ ,$$

and we obtain

$$\begin{aligned} |\varphi_\varepsilon * f(x) - af(x)| &\le \int |f(x-y) - f(x)|\,|\varphi_\varepsilon(y)|\,dy \\ &\le \frac{\eta}{2m}\int_{[\,|y|<\delta]} |\varphi_\varepsilon(y)|\,dy + 2\,\|f\|_\infty \int_{[\,|y|\ge\delta]} |\varphi_\varepsilon(y)|\,dy \qquad (7.10) \\ &\le \frac{\eta}{2} + 2\,\|f\|_\infty \int_{[\,|y|\ge\delta]} |\varphi_\varepsilon(y)|\,dy \end{aligned}$$

for $x \in \mathbb{R}^n$. The substitution rule then gives

$$\int_{[\,|y|\ge\delta]} |\varphi_\varepsilon(y)|\,dy = \varepsilon^{-n}\int_{[\,|y|\ge\delta]} |\varphi(y/\varepsilon)|\,dy = \int_{[\,|z|\ge\delta/\varepsilon]} |\varphi(z)|\,dz\ .$$

By the dominated convergence theorem, then, there exists $\varepsilon_0 > 0$ such that

$$\int_{[\,|y|\ge\delta]} |\varphi_\varepsilon(y)|\,dy \le \frac{\eta}{4\,\|f\|_\infty} \quad \text{for } \varepsilon \in (0, \varepsilon_0]\ .$$

Now the claim follows from (7.10). ■

[7] This also follows from Minkowski's inequality for integrals.

Suppose $\varphi \in \mathcal{L}_1$ satisfies $\int \varphi\, dx = 1$ and set

$$\varphi_\varepsilon(x) := \varepsilon^{-n}\varphi(x/\varepsilon) \quad \text{for } x \in \mathbb{R}^n \text{ and } \varepsilon > 0\ . \tag{7.11}$$

The family $\{\, \varphi_\varepsilon\ ;\ \varepsilon > 0 \,\}$ is called an **approximating kernel** or an **approximation to the identity**. If

$$\varphi \in C^\infty(\mathbb{R}^n, \mathbb{R})\ , \quad \breve{\varphi} = \varphi\ , \quad \varphi \geq 0\ , \quad \operatorname{supp}(\varphi) \subset \overline{\mathbb{B}}^n\ , \quad \int \varphi\, dx = 1\ ,$$

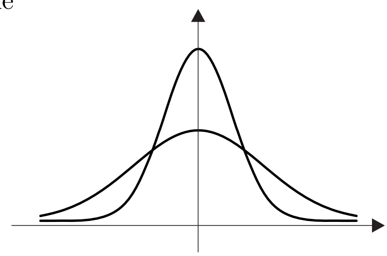
we call $\{\, \varphi_\varepsilon\ ;\ \varepsilon > 0 \,\}$ a **mollifier** or **smoothing kernel**. Every smoothing kernel obviously satisfies

$$\operatorname{supp}(\varphi_\varepsilon) \subset \varepsilon\overline{\mathbb{B}}^n \quad \text{for } \|\varphi_\varepsilon\|_1 = 1 \text{ and } \varepsilon > 0\ .$$

7.12 Examples[8] **(a)** The **Gaussian kernel** is the family $\{\, k_\varepsilon\ ;\ \varepsilon > 0 \,\}$ defined by

$$k(x) := (4\pi)^{-n/2}\, e^{-|x|^2/4} \quad \text{for } x \in \mathbb{R}^n\ .$$

It is an approximating kernel.



Proof From Example 6.13(a), we know that

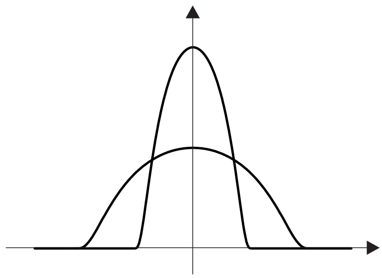
$$\int g(x)\, dx = 1$$

for $g(x) := \pi^{-n/2}\, e^{-|x|^2}$. Since $k(x) = 2^{-n} g(x/2)$ for $x \in \mathbb{R}^n$, it follows from the substitution rule that $\int k(x)\, dx = 1$. ■

(b) Let

$$\varphi(x) := \begin{cases} c\, e^{1/(|x|^2-1)} & \text{if } |x| < 1\ , \\ 0 & \text{if } |x| \geq 1\ , \end{cases}$$

where $c := \bigl(\int_{\mathbb{B}^n} e^{1/(|x|^2-1)}\, dx\bigr)^{-1}$ is chosen so that ϕ integrates to 1. Then the family $\{\, \varphi_\varepsilon\ ;\ \varepsilon > 0 \,\}$ is a smoothing kernel.



Proof Because $x \mapsto |x|^2 - 1$ is smooth on $\mathbb{R}^n$, Example IV.1.17 shows that φ belongs to $C^\infty(\mathbb{R}^n, \mathbb{R})$ (see Exercise VII.5.16). The claim follows easily. ■

Test functions

Let X be a metric space, and let A and B be subsets of X. We say A is **compactly contained in** B (in symbols: $A \subset\subset B$) if $\overline{A}$ is compact and is contained in the interior of B.

[8] In both examples, the area under the graphs is always 1, so smaller values of ε give correspondingly higher maxima.

If X is open in $\mathbb{R}^n$ and E is a normed vector space, we call

$$\mathcal{D}(X,E) := \{\, \varphi \in C^\infty(X,E) \ ;\ \mathrm{supp}(\varphi) \subset\subset X \,\}$$

the **space of (*E*-valued) test functions** on X. When $E = \mathbb{K}$, we write $\mathcal{D}(X) := \mathcal{D}(X,\mathbb{K})$, as usual. Clearly $\mathcal{D}(X,E)$ is a vector subspace of $C^\infty(X,E)$ and of $C_c(X,E)$, and $\mathcal{D}(X,E) = C^\infty(X,E) \cap C_c(X,E)$. Because the map

$$j : C_c(X,E) \to C_c(\mathbb{R}^n,E) \ , \quad g \mapsto \widetilde{g} \ ,$$

is linear and injective, we can identify $C_c(X,E)$ with a vector subspace of $C_c(\mathbb{R}^n,E)$ and regard (as needed) each element of the former as an element of the latter. Likewise, we identify $\mathcal{D}(X,E)$ with a vector subspace of $\mathcal{D}(\mathbb{R}^n,E)$. With these notations, we have the following inclusions of vector subspaces for every $p \in [1,\infty]$:

$$\mathcal{D}(X,E) \subset \mathcal{D}(\mathbb{R}^n,E) \subset C_c(\mathbb{R}^n,E) \subset L_p(\mathbb{R}^n,E) \ .$$

7.13 Theorem *Suppose X is open in $\mathbb{R}^n$ and $p \in [1,\infty)$. Then $\mathcal{D}(X)$ is a dense vector subspace of $L_p(X)$ and of $C_0(X)$.*

Proof (i) Take $g \in C_c(X)$ and $\eta > 0$. Also let $\{\, \varphi_\varepsilon \ ;\ \varepsilon > 0 \,\}$ be a smoothing kernel. By Theorem 7.8, $\varphi_\varepsilon * g$ belongs to BUC^k and therefore to BUC^∞ for every $k \in \mathbb{N}$. Because g has compact support, there is $\varepsilon_0 > 0$ such that[9] $\mathrm{dist}\big(\mathrm{supp}(g), X^c\big) \ge \varepsilon_0$. From Theorem 7.10, it follows that

$$\mathrm{supp}(\varphi_\varepsilon * g) \subset \mathrm{supp}(\varphi_\varepsilon) + \mathrm{supp}(g) \subset \mathrm{supp}(g) + \varepsilon\overline{\mathbb{B}}^n \quad \text{for } \varepsilon > 0 \ .$$

Then $\varphi_\varepsilon * g$ belongs to $\mathcal{D}(X)$ for $\varepsilon \in (0,\varepsilon_0)$. Finally by Theorem 7.11 we can find for every $q \in [1,\infty]$ some $\varepsilon_1 \in (0,\varepsilon_0)$ such that $\|\varphi_{\varepsilon_1} * g - g\|_q < \eta/2$.

(ii) Now suppose $f \in L_p(X)$. By Theorem 5.1, we can find $g \in C_c(X)$ such that $\|f - g\|_p < \eta/2$. By (i), there is $h \in \mathcal{D}(X)$ such that $\|f - h\|_p < \eta$.

(iii) For $f \in C_0(X)$, let K be a compact subset of X such that $|f(x)| < \eta/2$ for $x \in X\backslash K$. By Proposition 4.13, we can choose a $\varphi \in C_c(X)$ such that $0 \le \varphi \le 1$ and $\varphi \,|\, K = 1$. We set $g := \varphi f$. Because $f(x) = g(x)$ for $x \in K$, it follows that

$$|f(x) - g(x)| = |f(x)|\,|1 - \varphi(x)| < \eta/2 \quad \text{for } x \in X \ .$$

Therefore $\|f - g\|_\infty \le \eta/2$. The claim then follows from (i). ■

Smooth partitions of unity

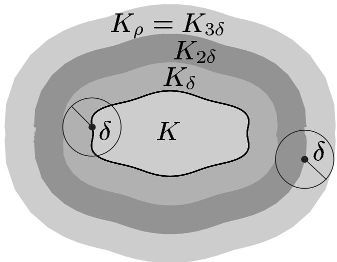
In Section 4, we proved the existence of continuous Urysohn functions in general metric spaces. This result can be distinctly improved in the special case of $\mathbb{R}^n$, where we can use mollifiers to actually construct smooth cutoff functions.

[9] $\mathrm{dist}\big(\mathrm{supp}(g), \emptyset\big) := \infty$.

7.14 Proposition (smooth cutoff functions) *Suppose $K \subset \mathbb{R}^n$ is compact, and set*

$$K_\rho := \{ x \in \mathbb{R}^n \ ; \ \mathrm{dist}(x, K) < \rho \} \quad \text{for } \rho > 0 \ .$$

Then for every $\alpha \in \mathbb{N}^n$ and every $\rho > 0$ there exist a positive constant $c(\alpha)$ and a map $\varphi \in \mathcal{D}(K_\rho)$ such that $0 \le \varphi \le 1$, $\varphi \,|\, K = 1$, and $\|\partial^\alpha \varphi\|_\infty \le c(\alpha)\rho^{-|\alpha|}$.



Proof Set $\{ \psi_\varepsilon \ ; \ \varepsilon > 0 \}$ be a smoothing kernel. Let $\delta := \rho/3$ and $\varphi := \psi_\delta * \chi_{K_\delta}$. Then φ belongs to BUC^∞, and it follows from Theorem 7.10 that

$$\begin{aligned}\mathrm{supp}(\varphi) &\subset \mathrm{supp}(\psi_\delta) + \overline{K}_\delta \subset \delta\overline{\mathbb{B}}^n + \overline{K}_\delta \\ &\subset \overline{K}_{2\delta} \subset K_{3\delta} = K_\rho \ .\end{aligned}$$

Therefore φ belongs to $\mathcal{D}(K_\rho)$. Moreover

$$\varphi(x) = \int \psi_\delta(x-y)\chi_{K_\delta}(y)\,dy \le \int \psi_\delta(x-y)\,dy = 1$$

for $x \in \mathbb{R}^n$, and hence $0 \le \varphi \le 1$. If x lies in K, then

$$\varphi(x) = \int \psi_\delta(y)\chi_{K_\delta}(x-y)\,dy = \int \psi_\delta(y)\,dy = 1 \ ,$$

and therefore $\varphi \,|\, K = 1$. Finally, since $\partial^\alpha \psi_\delta = \delta^{-|\alpha|}(\partial^\alpha \psi_1)_\delta$ for $\alpha \in \mathbb{N}^n$, we have from Theorem 7.8(iv) that

$$\partial^\alpha \varphi = \partial^\alpha(\psi_\delta * \chi_{K_\delta}) = \partial^\alpha \psi_\delta * \chi_{K_\delta} = \delta^{-|\alpha|}(\partial^\alpha \psi_1)_\delta * \chi_{K_\delta} \ .$$

Now $c(\alpha) := 3^{|\alpha|}\,\|\partial^\alpha \psi_1\|_1$ is independent of $\delta > 0$, and so it follows from Young's inequality that $\|\partial^\alpha \varphi\|_\infty \le c(\alpha)\rho^{-|\alpha|}$. ■

Let $K \subset \mathbb{R}^n$ be compact and denote by $\{ X_j \ ; \ 0 \le j \le m \}$ a finite open cover of K. If for every $j \in \{0, \ldots, m\}$, there is a $\varphi_j \in C^\infty(\mathbb{R}^n)$ such that

(i) $0 \le \varphi_j \le 1$,

(ii) $\mathrm{supp}(\varphi_j) \subset X_j$, and

(iii) $\sum_{j=0}^m \varphi_j(x) = 1$ for $x \in K$,

then $\{ \varphi_j \ ; \ 0 \le j \le m \}$ is called a **smooth partition of unity on *K*** subordinate to the cover $\{ X_j \ ; \ 0 \le j \le m \}$.

If X_0 is open in $\mathbb{R}^n$ and $K \subset X_0$, then $\mathrm{dist}(K, X_0^c) > 0$, and Proposition 7.14 (with $\rho := \mathrm{dist}(K, X_0^c)$) secures the existence of a smooth partition of unity on K subordinate to the one-element cover $\{X_0\}$ of K. To treat the general case of a finite cover, we need a technical result:

7.15 Lemma (shrinking lemma) *Let* $\{\, X_j \ ;\ 0 \le j \le m \,\}$ *be a finite open cover of a compact subset* K *of* $\mathbb{R}^n$*. Then there is an open cover* $\{\, U_j \ ;\ 0 \le j \le m \,\}$ *of* K *such that* $U_j \subset\subset X_j$ *for* $j \in \{0, \ldots, m\}$.

Proof Given $x \in K$, choose $j \in \{0, \ldots, m\}$ such that $x \in X_j$ and $r_x > 0$ such that $V_x := \mathbb{B}^n(x, r_x)$ is compact and contained in X_j. Then $\{\, V_x \ ;\ x \in K \,\}$ is an open cover of K, and there exist $k \in \mathbb{N}$ and $\{x_0, \ldots, x_k\} \subset K$ with $K \subset \bigcup_{i=0}^{k} V_{x_i}$. With $U_j := \bigcup \{\, V_{x_i} \ ;\ V_{x_i} \subset X_j \,\}$ for $j \in \{0, \ldots, m\}$, we have a family $\{\, U_j \ ;\ 0 \le j \le m \,\}$ having the desired properties. ■

7.16 Theorem (smooth partitions of unity) *If* K *is a compact subset of* $\mathbb{R}^n$*, every finite open cover of* K *has a smooth partition of unity subordinate to it.*

Proof Suppose $\{\, X_j \ ;\ 0 \le j \le m \,\}$ is a finite open cover of K. By Lemma 7.15, there is an open cover $\{\, U_j \ ;\ 0 \le j \le m \,\}$ such that $U_j \subset\subset X_j$ for $j \in \{0, \ldots, m\}$. We define $K_j := \overline{U}_j$. Then K_j is compact, and $\operatorname{dist}(K_j, X_j^c)$ is positive for every $j \in \{0, \ldots, m\}$. Proposition 7.14 now shows there is a $\psi_j \in \mathcal{D}(X_j)$ such that $0 \le \psi_j \le 1$ and $\psi_j \,|\, K_j = 1$. Defining

$$\varphi_0 := \psi_0 \quad \text{and} \quad \varphi_k := \psi_k \prod_{j=0}^{k-1} (1 - \psi_j) \quad \text{for } 1 \le k \le m \ ,$$

it is easy to check by induction that $\sum_{j=0}^{m} \varphi_j = 1 - \prod_{j=0}^{m} (1 - \psi_j)$. The claim now follows because $K \subset \bigcup_{j=0}^{m} K_j$. ■

We next present some simple applications of Theorem 7.16. Additional, more complicated situations will be described in succeeding chapters.

7.17 Applications **(a)** Suppose X is open in $\mathbb{R}^n$. Then for $f \in L_0(X)$ the following statements are equivalent:

(i) $f \in L_{1,\mathrm{loc}}(X)$;

(ii) $\varphi f \in L_1(X)$ for every $\varphi \in \mathcal{D}(X)$;

(iii) $\widetilde{f \,|\, K} \in L_1(X)$ for every $K = \overline{K} \subset\subset X$.

Proof Let $(U_j)_{j \in \mathbb{N}}$ be an ascending sequence of relatively compact open subsets of X with $X = \bigcup_j U_j$ (see Remarks 1.16(d) and (e)). Then (see Exercise 4.3)

$$L_{1,\mathrm{loc}}(X) = \bigl\{\, f \in L_0(X) \ ;\ \chi_{U_j} f \in L_1(X),\ j \in \mathbb{N} \,\bigr\} \ .$$

"(i)⇒(ii)" Let $\varphi \in \mathcal{D}(X)$. Since $K := \operatorname{supp}(\varphi)$ is compact and $(U_j)_{j \in \mathbb{N}}$ is ascending, there is a $k \in \mathbb{N}$ such that $K \subset U_k$. By virtue of Proposition 7.14, we find a $\psi \in \mathcal{D}(U_k)$ such that $0 \le \psi \le 1$ and $\psi \,|\, K = 1$. Then

$$\int_X |\varphi f| \, dx = \int_X |\varphi \psi f| \, dx \le \|\varphi\|_\infty \int_X |\psi f| \, dx \le \|\varphi\|_\infty \, \|\chi_{U_k} f\|_1 < \infty \ .$$

Therefore φf belongs to $L_1(X)$.

"(ii)⇒(iii)" Take $K = \overline{K} \subset\subset X$ and $\varphi \in \mathcal{D}(X)$ with $\varphi \,|\, K = 1$. Then

$$\int_K |f| = \int_K |\varphi f| \le \|\varphi f\|_1 < \infty \ ,$$

and therefore $\widetilde{f \,|\, K} \in L_1(X)$.

"(iii)⇒(i)" This implication is clear because every $\overline{U}_j$ is compact. ■

(b) Suppose X is open in $\mathbb{R}^n$. Then $C(X) \subset L_{1,\mathrm{loc}}(X)$.

Proof Take $f \in C(X)$ and $\varphi \in \mathcal{D}(X)$. Then φf belongs to $C_c(X)$. By Theorem 5.1, we have $\varphi f \in L_1(X)$, and the claim follows from (a). ■

(c) The linear representation of the group $(\mathbb{R}^n, +)$ in BUC^k is injective, hence a group isomorphism onto its image $\mathfrak{T}_{BUC^k}$.

Proof For $a \in \mathbb{R}^n$, suppose $\tau_a = \mathrm{id}_{BUC^k}$. We choose $r > |a|$ and a cutoff function $\varphi \in \mathcal{D}(\mathbb{R}^n)$ for $r\overline{\mathbb{B}}^n$. Then $f_j := \varphi \,\mathrm{pr}_j$ belongs to BUC^k, and we find

$$-a_j = \tau_a f_j(0) = f_j(0) = 0 \quad \text{for } j \in \{1, \ldots, n\} \ .$$

Therefore $a = 0$. This implies the injectivity of the representation $a \mapsto \tau_a$. ■

(d) Suppose X is open in $\mathbb{R}^n$ and bounded. Also let $\{ X_j \ ; \ 0 \le j \le m \}$ be a finite open cover of $\overline{X}$, and let $\{ \varphi_j \ ; \ 0 \le j \le m \}$ be a smooth partition of unity subordinate to it. Finally let $k \in \mathbb{N}$ and

$$|||u|||_{BC^k} := \sum_{j=0}^m \|\varphi_j u\|_{BC^k} \quad \text{for } u \in BC^k(X) \ .$$

Then $|||\cdot|||_{BC^k}$ is an equivalent norm on $BC^k(X)$.

Proof Take $u \in BC^k(X)$. Obviously

$$\|u\|_{BC^k} = \Big\| \sum_{j=0}^m \varphi_j u \Big\|_{BC^k} \le \sum_{j=0}^m \|\varphi_j u\|_{BC^k} = |||u|||_{BC^k} \ .$$

From Leibniz's rule (see Exercise VII.5.21), we obtain

$$\begin{aligned} |||u|||_{BC^k} &= \sum_{j=0}^m \max_{|\alpha| \le k} \|\partial^\alpha(\varphi_j u)\|_\infty = \sum_{j=0}^m \max_{|\alpha| \le k} \Big\| \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta \varphi_j \, \partial^{\alpha-\beta} u \Big\|_\infty \\ &\le \sum_{j=0}^m c_k \, \|\varphi_j\|_{BC^k} \, \|u\|_{BC^k} \le C \, \|u\|_{BC^k} \ , \end{aligned}$$

where we have set $c_k := \max_{|\alpha| \le k} \sum_{\beta \le \alpha} \binom{\alpha}{\beta}$ and $C := c_k \sum_{j=0}^m \|\varphi_j\|_{BC^k}$. ■

Convolutions of *E*-valued functions

A look back at preceding proofs shows that the convolution $f * g$ can also be defined when one of the two functions takes values in a Banach space F and the other is scalar-valued. All proofs carry through without change[10] so long as the substitution rule for isometries still holds for F-valued functions. This is indeed the case, as we shall show in the next section. In particular, the key approximation result in Theorem 7.11 remains true for the spaces $L_p(\mathbb{R}^n, F)$ and $BUC^k(\mathbb{R}^n, F)$ with $1 \leq p < \infty$ and $k \in \mathbb{N}$. One consequence of this is an analogue of Theorem 7.13 to the effect that $\mathcal{D}(X, F)$ is a dense vector subspace of $C_0(X, F)$ and of $L_p(X, F)$ for $1 \leq p < \infty$.

Distributions[11]

Suppose X is a nonempty open subset of $\mathbb{R}^n$. A scalar function on X, as is well known, is a rule for assigning a real or complex number to every point in X. But this definition is just an abstraction, since the individual points of X cannot in practice be discerned. If, for example, we want to determine the temperature distribution of some medium that occupies the set X — we must rely on an experimental probe. However, such a probe, being of nonzero size, can only determine values of f in an extended region; whatever value it assigns to $f(x_0)$ represents not the actual value at x_0 (if indeed such a thing is physically meaningful) but rather some kind of average around x_0: mathematically, an integral $\int_X \varphi f \, dx$, where φ is a "test function" that depends on the probe. Of course, the measurement will better approximate the exact value $f(x_0)$ the more the test function φ is concentrated about x_0, that is, the less the probe smears the data.

To claim complete knowledge of $f(x_0)$, one might imagine bringing to bear all conceivable probes, or in other words, determining the averages $\int_X \varphi f \, dx$ over all possible test functions φ. In mathematical terms, we'd be replacing the pointwise function $f : X \to \mathbb{K}$ by a functional defined on the space of all test functions, namely, the map

$$T_f : \mathcal{D}(X) \to \mathbb{K} , \quad \varphi \mapsto \int_X \varphi f \, dx . \tag{7.12}$$

Our choice of $\mathcal{D}(X)$ as the space of test functions is to a large extent arbitrary. For conceptual simplicity, we might want to consider $C_c(X)$ instead of $\mathcal{D}(X)$. At the same time, we would like to avoid performing more "measurements" than necessary; this warrants choosing a test space that is small in some sense. But the space must be large enough that the averages $\int_X \varphi f \, dx$ do determine f. That is, we want the equality of $\int_X \varphi f \, dx$ and $\int_X \varphi g \, dx$ for all test functions φ to imply $f = g$.

[10] Naturally, the commutativity formula $f * g = g * f$ must be interpreted correctly.

[11] The rest of this section is meant to provide glimpses of applications and more advanced theories; it can be skipped over on first reading.

The next theorem shows that this is indeed the case if we choose $\mathcal{D}(X)$ as the test space and work with "functions" in $L_{1,\mathrm{loc}}(X)$. (By Application 7.17(a), $L_{1,\mathrm{loc}}(X)$ is the largest vector subspace E of $L_0(X)$ such that $\int_X \varphi f\,dx$ is well defined for all $f \in E$ and all $\varphi \in \mathcal{D}(X)$.)

7.18 Theorem *Suppose $f \in L_{1,\mathrm{loc}}(X)$. If*

$$\int_X \varphi f\,dx = 0 \quad \textit{for } \varphi \in \mathcal{D}(X)\ , \tag{7.13}$$

then $f = 0$.

Proof Suppose $f \neq 0$, and let $\overset{*}{f} \in \mathcal{L}_{1,\mathrm{loc}}(X)$ be a representative of f. By regularity, there is a compact subset K of X of positive measure such that $\overset{*}{f}(x) \neq 0$ for $x \in K$. Take $\eta \in \mathcal{D}(X)$ with $\eta \,|\, K = 1$, and let $g := \eta \overset{*}{f}$. By Application 7.17(a), g belongs to $\mathcal{L}_1$. Also $g(x) \neq 0$ for $x \in K$. Let $\{\, \varphi_\varepsilon \ ;\ \varepsilon > 0 \,\}$ be a smoothing kernel. Then $\lim_{\varepsilon \to 0} \varphi_\varepsilon * g = g$ in $\mathcal{L}_1$. By Corollary 4.7, there is a null sequence (ε_j) and a set N of Lebesgue measure zero such that

$$\lim_{j \to \infty} \varphi_{\varepsilon_j} * g(x) = g(x) \quad \text{for } x \in N^c\ . \tag{7.14}$$

Given $x_0 \in K \cap N^c$, set $\psi_j := \eta \tau_{x_0} \varphi_{\varepsilon_j} \in \mathcal{D}(X)$ for $j \in \mathbb{N}$. Since $\check{\varphi}_\varepsilon = \varphi_\varepsilon$ by Remark 7.1(d), equality (7.13) gives

$$\begin{aligned}\varphi_{\varepsilon_j} * g(x_0) &= \int g(y)\, \varphi_{\varepsilon_j}(x_0 - y)\,dy = \int_X (\eta \overset{*}{f})(y)\, \varphi_{\varepsilon_j}(x_0 - y)\,dy \\ &= \int_X \overset{*}{f}(y)\, \psi_j(y)\,dy = 0\ .\end{aligned}$$

However, because of (7.14) this contradicts $g(x_0) \neq 0$. The claim follows because the representative $\overset{*}{f}$ of f was chosen arbitrarily. ■

Clearly the map T_f is a linear functional on $\mathcal{D}(X)$. For the interpretation of $T_f \varphi = \int_X \varphi f\,dx$ as a measurement value to be meaningful, $T_f \varphi$ must "depend continuously" on the measuring device; that is, small perturbations in the probe, hence in the test function φ, should cause only small changes in the measured value. Mathematically speaking, this means that T_f must be a continuous linear functional on $\mathcal{D}(X)$. So we must introduce a topology on $\mathcal{D}(X)$.

Since our treatment here is introductory, we will limit ourselves to stating what it means for a sequence to converge in $\mathcal{D}(X)$. This convergence should be compatible with the vector structure on $\mathcal{D}(X)$, so it suffices to consider the case where the limit is 0.

We say that a sequence (φ_j) **converges to 0** (or **is a null sequence**) **in $\mathcal{D}(X)$** if the following conditions are satisfied:

($\mathcal{D}_1$) There exists $K \subset\subset X$ such that $\mathrm{supp}(\varphi_j) \subset K$ for $j \in \mathbb{N}$.

($\mathcal{D}_2$) $\varphi_j \to 0$ in $BC^k(X)$ for every $k \in \mathbb{N}$.

Obviously $(\mathcal{D}_2)$ is equivalent to:

$$\text{The sequence } (\partial^\alpha \varphi_j)_{j\in\mathbb{N}} \text{ converges uniformly to 0 for every } \alpha \in \mathbb{N}^n. \tag{7.15}$$

So for φ_j to converge in $\mathcal{D}(X)$ to 0, not only must (7.15) hold, but the supports of the functions φ_j must all be contained in a fixed compact subset of X.

A linear functional $T\colon \mathcal{D}(X) \to \mathbb{K}$ is **continuous** if $T\varphi_j \to 0$ for every null sequence (φ_j) in $\mathcal{D}(X)$. A continuous linear functional on $\mathcal{D}(X)$ is also called a **Schwartz distribution**, or simply **distribution**, on X. The set of distributions on X is denoted by $\mathcal{D}'(X)$; it is clearly a vector subspace of $\operatorname{Hom}\bigl(\mathcal{D}(X), \mathbb{K}\bigr)$.

(In functional analysis — more precisely, in the theory of topological vector spaces — one shows that there is exactly one Hausdorff topology on $\mathcal{D}(X)$ that is locally convex,[12] compatible with the vector structure, and such that sequences converge to 0 in the sense above if and only if they converge to 0 in the topology. With respect to this topology, $\mathcal{D}'(X)$ is the dual of $\mathcal{D}(X)$, that is, the space of all continuous linear functionals on $\mathcal{D}(X)$. See, for example, [Sch66] or [Yos65].)

7.19 Examples **(a)** For every $f \in L_{1,\mathrm{loc}}(X)$, the linear functional T_f defined by (7.12) is a distribution on X.

Proof Let (φ_j) be a sequence in $\mathcal{D}(X)$ such that $\varphi_j \to 0$ in $\mathcal{D}(X)$. Then there is a compact subset K of X such that $\operatorname{supp}(\varphi_j) \subset X$ for $j \in \mathbb{N}$. It follows that

$$|T_f\varphi_j| = \Bigl|\int_X \varphi_j f\,dx\Bigr| \le \int_K |\varphi_j|\,|f|\,dx \le \|f\|_{L_1(K)}\,\|\varphi_j\|_\infty$$

for $j \in \mathbb{N}$. Because $\|f\|_{L_1(K)} < \infty$, we find that $T_f\varphi_j \to 0$ in $\mathbb{K}$, because $(\mathcal{D}_2)$ implies that $\|\varphi_j\|_\infty \to 0$. ■

(b) Let μ be a Radon measure on X. Then

$$\mathcal{D}(X) \to \mathbb{K}\ , \quad \varphi \mapsto \int_X \varphi\,d\mu$$

defines a distribution on X.

Proof Suppose (φ_j) is a sequence in $\mathcal{D}(X)$ such that $\varphi_j \to 0$ in $\mathcal{D}(X)$. Also suppose $K = \overline{K} \subset\subset X$ contains $\operatorname{supp}(\varphi_j)$ for all $j \in \mathbb{N}$. Then

$$\Bigl|\int_X \varphi_j\,d\mu\Bigr| \le \int_K |\varphi_j|\,d\mu \le \mu(K)\,\|\varphi_j\|_\infty \quad \text{for } j \in \mathbb{N}\ .$$

As in the proof of (a), this implies that φ is a distribution on X. ■

(c) Let δ be the Dirac measure on $\mathbb{R}^n$ with support at 0. Then

$$\varphi \mapsto \langle \delta, \varphi\rangle := \int_X \varphi\,d\delta = \varphi(0) \quad \text{for } \varphi \in \mathcal{D}(\mathbb{R}^n)\ ,$$

[12] This means the origin has an open neighborhood basis of convex sets.

is a distribution on $\mathbb{R}^n$, the **Dirac distribution**

$$\delta : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K} , \quad \varphi \mapsto \varphi(0) .$$

There is no $u \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$ such that $T_u = \delta$.

Proof The first statement is a special case of (b).

Suppose now that $u \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$ with $T_u = \delta$, that is,

$$\int_{\mathbb{R}^n} \varphi u \, dx = \varphi(0) \quad \text{for } \varphi \in \mathcal{D}(\mathbb{R}^n) . \tag{7.16}$$

Choosing only such $\varphi \in \mathcal{D}(\mathbb{R}^n)$ that $\mathrm{supp}(\varphi) \subset\subset X := \mathbb{R}^n \backslash \{0\}$, we have $\varphi(0) = 0$, and from Theorem 7.18, it follows that $u \,|\, X = 0$ in $L_{1,\mathrm{loc}}(X)$. But X and $\mathbb{R}^n$ differ only on a set of measure zero (a single point!), so $u = 0$ in $L_{1,\mathrm{loc}}(\mathbb{R}^n)$, contradicting (7.16). ■

(d) Let $\alpha \in \mathbb{N}^n$. Then

$$S_\alpha : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K} , \quad \varphi \mapsto \partial^\alpha \varphi(0)$$

defines a distribution. There is no $u \in L_{1,\mathrm{loc}}(\mathbb{R}^n)$ such that $T_u = S_\alpha$.

Proof Let (φ_j) be a sequence in $\mathcal{D}(\mathbb{R}^n)$ such that $\varphi_j \to 0$ in $\mathcal{D}(\mathbb{R}^n)$, and suppose $K = \overline{K} \subset\subset \mathbb{R}^n$ with $\mathrm{supp}(\varphi_j) \subset K$ for $j \in \mathbb{N}$. We can assume that 0 lies in K. Then we have the estimate

$$|\partial^\alpha \varphi_j(0)| \le \max_{x \in K} |\partial^\alpha \varphi_j(x)| \le \|\varphi_j\|_{BC^{|\alpha|}} \quad \text{for } j \in \mathbb{N} .$$

Thus $(\mathcal{D}_2)$ implies that $\partial^\alpha \varphi_j(0) \to 0$ in $\mathbb{K}$, which shows that S_α is a distribution. The second statement is proved as in (c). ■

The following key result is now a simple consequence of Theorem 7.18.

7.20 Theorem *The map*

$$L_{1,\mathrm{loc}}(X) \to \mathcal{D}'(X) , \quad f \mapsto T_f$$

is linear and injective.

Proof Example 7.19(a) shows the map is well defined. It is linear because integration is. It is injective by Theorem 7.18. ■

By Theorem 7.20, we can identify $L_{1,\mathrm{loc}}(X)$ with its image in $\mathcal{D}'(X)$. In other words, we can regard $L_{1,\mathrm{loc}}(X)$ as a vector subspace of the space of all Schwartz distributions, by identifying a function $f \in L_{1,\mathrm{loc}}(X)$ with the distribution

$$T_f = \Bigl(\varphi \mapsto \int_X \varphi f \, dx\Bigr) \in \mathcal{D}'(X) .$$

In this sense, every $f \in L_{1,\mathrm{loc}}(X)$ is a distribution. The elements of $L_{1,\mathrm{loc}}(X)$ are called **regular distributions**. All other distributions are **singular**. Examples 7.19(c) and (d) illustrate singular distributions.

The theory of distributions plays an important role in higher analysis, especially in the study of partial differential equations, and in theoretical physics. We cannot elaborate here, but see for example [Sch65], [RS72].

Linear differential operators

Let X be open in $\mathbb{R}^n$. Given functions $a_\alpha \in C^\infty(X)$, for each $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le m \in \mathbb{N}$, we set

$$\mathcal{A}(\partial)u := \sum_{|\alpha|\le m} a_\alpha \partial^\alpha u \quad \text{for } u \in \mathcal{D}(X) \ .$$

Obviously, $\mathcal{A}(\partial)$ is a linear map of $\mathcal{D}(X)$ onto itself; we say it is a **linear differential operator on $\boldsymbol{X}$** of order $\le m$ (with smooth coefficients). It has **order $\boldsymbol{m}$** if

$$\sum_{|\alpha|=m} \|a_\alpha\|_\infty \ne 0 \ ,$$

that is, if at least one coefficient a_α of the **leading part** $\sum_{|\alpha|=m} a_\alpha \partial^\alpha$ of $\mathcal{A}(\partial)$ does not vanish identically. We denote by $\mathcal{D}\mathrm{iffop}(X)$ the set of all linear differential operators on X; those of order $\le m$ are denoted by $\mathcal{D}\mathrm{iffop}_m(X)$.

A linear map $T : \mathcal{D}(X) \to \mathcal{D}(X)$ is said to be **continuous**[13] if $T\varphi_j \to 0$ in $\mathcal{D}(X)$ for every sequence (φ_j) in $\mathcal{D}(X)$ such that $\varphi_j \to 0$ in $\mathcal{D}(X)$. The set of all continuous endomorphisms of $\mathcal{D}(X)$ is a vector subspace of $\mathrm{End}\bigl(\mathcal{D}(X)\bigr)$, which we denote by $\mathcal{L}\bigl(\mathcal{D}(X)\bigr)$.[13]

7.21 Proposition *$\mathcal{D}\mathrm{iffop}(X)$ is a vector subspace of $\mathcal{L}\bigl(\mathcal{D}(X)\bigr)$, and $\mathcal{D}\mathrm{iffop}_m(X)$ is a vector subspace of $\mathcal{D}\mathrm{iffop}(X)$.*

Proof Let $m \in \mathbb{N}$, and take $\mathcal{A}(\partial) := \sum_{|\alpha|\le m} a_\alpha \partial^\alpha \in \mathcal{D}\mathrm{iffop}_m(X)$. Let (φ_j) be a null sequence in $\mathcal{D}(X)$, and let $K = \overline{K} \subset\subset X$ contain $\mathrm{supp}(\varphi_j)$ for all $j \in \mathbb{N}$. Then $\mathrm{supp}\bigl(\mathcal{A}(\partial)\varphi_j\bigr) \subset K$ for $j \in \mathbb{N}$. For $\beta \in \mathbb{N}^n$, the Leibniz rule gives

$$\begin{aligned}\|\partial^\beta(a_\alpha \partial^\alpha \varphi_j)\|_{C(K)} &= \Bigl\|\textstyle\sum_{\gamma\le\beta} \binom{\beta}{\gamma} \partial^\gamma a_\alpha \partial^{\beta-\gamma+\alpha}\varphi_j\Bigr\|_{C(K)} \\ &\le c(\alpha,\beta) \max_{\gamma\le\beta} \|\partial^\gamma a_\alpha\|_{C(K)} \, \|\partial^{\beta-\gamma+\alpha}\varphi_j\|_\infty \ .\end{aligned}$$

From this we derive for $k \in \mathbb{N}$ the inequality

$$\|\mathcal{A}(\partial)\varphi_j\|_{BC^k(X)} \le c(k) \sum_{|\alpha|\le m} \|a_\alpha\|_{BC^k(K)} \, \|\varphi_j\|_{BC^{k+m}(X)} \quad \text{for } j \in \mathbb{N} \ ,$$

[13] It is shown in functional analysis that these definitions are consistent with our previous definitions for continuity and $\mathcal{L}(E)$.

where the constant $c(k)$ is independent of j. Now $\mathcal{A}(\partial)\varphi_j \to 0$ in $BC^k(X)$ follows from $(\mathcal{D}_2)$. Because this is true for every $k \in \mathbb{N}$, we see that $\mathcal{A}(\partial)\varphi_j \to 0$ in $\mathcal{D}(X)$. This proves that $\mathcal{D}\mathrm{iffop}_m(X) \subset \mathcal{L}\bigl(\mathcal{D}(X)\bigr)$. The other statements are clear. ■

Let $(\cdot\,|\,\cdot)$ denote the inner product in $L_2(X)$, and suppose $\mathcal{A}(\partial)$ belongs to $\mathcal{D}\mathrm{iffop}(X)$. If there is a differential operator $\mathcal{A}^\sharp(\partial) \in \mathcal{D}\mathrm{iffop}(X)$ such that

$$\bigl(\mathcal{A}(\partial)u \bigm| v\bigr) = \bigl(u \bigm| \mathcal{A}^\sharp(\partial)v\bigr) \quad \text{for } u, v \in \mathcal{D}(X) ,$$

we say $\mathcal{A}^\sharp(\partial)$ is the **formal adjoint** of $\mathcal{A}(\partial)$. Because

$$\bigl(u \bigm| \mathcal{A}^\sharp(\partial)v\bigr) = \int_X u\overline{\mathcal{A}^\sharp(\partial)v}\,dx$$

and $\overline{\mathcal{A}^\sharp(\partial)v} \in \mathcal{D}(X) \subset L_{1,\mathrm{loc}}(X)$ for $v \in \mathcal{D}(X)$, it follows easily from Theorem 7.18 that $\mathcal{A}(\partial)$ has at most one formal adjoint. If $\mathcal{A}(\partial)$ has a formal adjoint $\mathcal{A}^\sharp(\partial)$ that coincides with $\mathcal{A}(\partial)$, then $\mathcal{A}(\partial)$ is **formally self-adjoint**.

We will now show that every $\mathcal{A}(\partial) \in \mathcal{D}\mathrm{iffop}(X)$ has a differential operator formally adjoint to it, and we derive an explicit form for $\mathcal{A}^\sharp(\partial)$. First we need this:

7.22 Proposition (integration by parts) *For $f \in C^1(X)$ and $g \in C^1_c(X)$,*

$$\int_X (\partial_j f)g\,dx = -\int_X f\partial_j g\,dx \quad \text{for } j \in \{1, \ldots, n\} .$$

Proof We need only consider the case $j = 1$; the general case will follow by permutation of coordinates, in view of Corollary 6.10. So write $x = (x_1, x') \in \mathbb{R} \times \mathbb{R}^{n-1}$. Since fg has compact support, it follows from integrating by parts that

$$\int_{-\infty}^{\infty} \bigl[\partial_1 f(x_1, x')\bigr] g(x_1, x')\,dx_1 = -\int_{-\infty}^{\infty} f(x_1, x')\partial_1 g(x_1, x')\,dx_1$$

for every $x' \in \mathbb{R}^{n-1}$. From Fubini's theorem, we now get

$$\begin{aligned}
\int_X (\partial_1 f)g\,dx &= \int_{\mathbb{R}^n} (\partial_1 f)g\,dx \\
&= \int_{\mathbb{R}^{n-1}} \Bigl(\int_{-\infty}^{\infty} \partial_1 f(x_1, x')g(x_1, x')\,dx_1\Bigr)\,dx' \\
&= -\int_{\mathbb{R}^{n-1}} \Bigl(\int_{-\infty}^{\infty} f(x_1, x')\partial_1 g(x_1, x')\,dx_1\Bigr)\,dx' \\
&= -\int_{\mathbb{R}^n} f\partial_1 g\,dx = -\int_X f\partial_1 g\,dx . \quad ■
\end{aligned}$$

7.23 Corollary *Suppose $f \in C^k(X)$ and $g \in C_c^k(X)$. Then*

$$\int_X (\partial^\alpha f) g \, dx = (-1)^{|\alpha|} \int_X f \partial^\alpha g \, dx$$

for $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le k$.

Integration by parts is also the core of the proof of the next result.

7.24 Proposition *Every differential operator*

$$\mathcal{A}(\partial) = \sum_{|\alpha| \le m} a_\alpha \partial^\alpha \in \mathcal{D}\mathrm{iffop}(X)$$

has a unique formal adjoint, which is explicitly given by

$$\mathcal{A}^\sharp(\partial) v = \sum_{|\alpha| \le m} (-1)^{|\alpha|} \partial^\alpha (\overline{a}_\alpha v) \quad \text{for } v \in \mathcal{D}(X) \ . \tag{7.17}$$

If $\mathcal{A}(\partial)$ has order m, then $\mathcal{A}^\sharp(\partial)$ is also an m-th order differential operator.

Proof We already know that there is at most one formal adjoint, so we need only prove existence and the validity of (7.17).

Take $u, v \in \mathcal{D}(X)$. Integrating by parts, we find

$$\begin{aligned} \big(\mathcal{A}(\partial) u \,\big|\, v\big) &= \int_X \big(\mathcal{A}(\partial) u\big) \overline{v} \, dx = \sum_{|\alpha| \le m} \int_X (a_\alpha \partial^\alpha u) \overline{v} \, dx \\ &= \sum_{|\alpha| \le m} (-1)^{|\alpha|} \int_X u \partial^\alpha (a_\alpha \overline{v}) \, dx = \int_X u \overline{\sum_{|\alpha| \le m} (-1)^{|\alpha|} \partial^\alpha (\overline{a}_\alpha v)} \, dx \ . \end{aligned}$$

Therefore

$$\big(\mathcal{A}(\partial) u \,\big|\, v\big) = \big(u \,\big|\, \mathcal{A}^\sharp(\partial) v\big) \quad \text{for } u, v \in \mathcal{D}(X)$$

if $\mathcal{A}^\sharp(\partial) v$ is as in (7.17). By Leibniz's rule, there exist $b_\alpha \in C^\infty(X)$ for $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m - 1$, such that

$$\mathcal{A}^\sharp(\partial) = (-1)^m \sum_{|\alpha| = m} \overline{a}_\alpha \partial^\alpha + \sum_{|\alpha| \le m-1} b_\alpha \partial^\alpha \ .$$

Therefore $\mathcal{A}^\sharp(\partial)$ belongs to $\mathcal{D}\mathrm{iffop}(X)$. The claim is now clear. ■

For differential operators that describe the time evolution of systems, it is usual to treat time as a distinguished variable. We recall, for instance, the wave operator $\partial_t^2 - \Delta_x$ and the heat operator $\partial_t - \Delta_x$ in the variables $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ (see Exercise VII.5.10). Another example is the **Schrödinger operator** $(1/i)\partial_t - \Delta_x$. All three operators are second order differential operators.

7.25 Examples **(a)** The wave operator and the Schrödinger operator are formally self-adjoint.[15]

(b) The heat operator has $(\partial_t - \Delta_x)^\sharp = -\partial_t - \Delta_x$ as its adjoint. It is therefore *not* formally self-adjoint.

(c) For $\mathcal{A}(\partial) := \partial_t - \sum_{j=1}^n \partial_j$, we have $\mathcal{A}^\sharp(\partial) = -\mathcal{A}(\partial)$.

(d) Suppose $a_{jk}, a_j, a_0 \in C^\infty(X, \mathbb{R})$ with

$$\sum_{j,k=1}^n \|a_{jk}\|_\infty \neq 0 , \quad a_{jk} = a_{kj} \quad \text{for } j,k \in \{1,\ldots,n\} .$$

Also define $\mathcal{A}(\partial) \in \mathcal{D}\mathrm{iffop}_2(X)$ by

$$\mathcal{A}(\partial)u := \sum_{j,k=1}^n \partial_j(a_{jk}\partial_k u) + \sum_{j=1}^n a_j \partial_j u + a_0 u \quad \text{for } u \in \mathcal{D}(X) .$$

Then we say $\mathcal{A}(\partial)$ is a **divergence form** operator.[16] In this case, we have

$$\mathcal{A}^\sharp(\partial)v = \sum_{j,k=1}^n \partial_j(a_{jk}\partial_k v) - \sum_{j=1}^n a_j \partial_j v + \Bigl(a_0 - \sum_{j=1}^n \partial_j a_j\Bigr) v \quad \text{for } v \in \mathcal{D}(X) .$$

Therefore the formal adjoint is also of divergence form, and $\mathcal{A}(\partial)$ is formally self-adjoint if and only if $a_j = 0$ for $j = 1, \ldots, n$.

Proof This follows easily from Proposition 7.22. ■

(e) The Laplace operator Δ is a formally self-adjoint second-order differential operator of divergence form.

Proof This follows from (d) by taking $a_{jk} = \delta_{jk}$ (the Kronecker delta). ■

Weak derivatives

We now explain briefly how the concept of derivative can be generalized so functions that are not differentiable in the classical sense can be assigned a generalized derivative.

Suppose X is open in $\mathbb{R}^n$. We say $u \in L_{1,\mathrm{loc}}(X)$ is **weakly differentiable** if there exists $u_j \in L_{1,\mathrm{loc}}(X)$ such that

$$\int_X (\partial_j \varphi) u \, dx = -\int_X \varphi u_j \, dx \quad \text{for } \varphi \in \mathcal{D}(X) \text{ and } 1 \le j \le n . \tag{7.18}$$

[15] These facts are of particular importance in mathematical physics.
[16] The reason for this language will be clarified in Section XI.6.

More generally, if $m \geq 2$ is an integer, we say $u \in L_{1,\mathrm{loc}}(X)$ is ***m*-times weakly differentiable on *X*** if there exists $u_\alpha \in L_{1,\mathrm{loc}}(X)$ such that

$$\int_X (\partial^\alpha \varphi) u \, dx = (-1)^{|\alpha|} \int_X \varphi u_\alpha \, dx \quad \text{for } \varphi \in \mathcal{D}(X) \,, \tag{7.19}$$

for all $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$. If this is the case, then it immediately follows from Theorem 7.18 that $u_\alpha \in L_{1,\mathrm{loc}}(X)$ is uniquely determined by u (and α). We call u_α the **α-th weak partial derivative** and set $\partial^\alpha u := u_\alpha$. In the case $m = 1$, we set $\partial_j u := u_j$. These notations are justified by the first of the following remarks.

7.26 Remarks **(a)** Suppose $m \in \mathbb{N}^\times$. Then every $u \in C^m(X)$ is m-times weakly differentiable, and the weak derivatives agree with the classical, or usual, partial derivative.

Proof This follows from Corollary 7.23. ■

(b) Let $W^m_{1,\mathrm{loc}}(X)$ be the set of all m-times weakly differentiable functions on X. Then $W^m_{1,\mathrm{loc}}(X)$ is a vector subspace of $L_{1,\mathrm{loc}}(X)$, and for every $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq m$, the map

$$W^m_{1,\mathrm{loc}}(X) \to W^{m-|\alpha|}_{1,\mathrm{loc}}(X) \,, \quad u \mapsto \partial^\alpha u$$

is well defined and linear.

Proof We leave the simple proof to the reader as an exercise. ■

(c) For $u \in W^m_{1,\mathrm{loc}}(X)$ and $\alpha, \beta \in \mathbb{N}^n$ with $|\alpha| + |\beta| \leq m$, we have $\partial^\alpha \partial^\beta u = \partial^\beta \partial^\alpha u$.

Proof This follows immediately from the defining equations (7.19) and the properties of smooth functions. ■

(d) Suppose $u \in L_{1,\mathrm{loc}}(\mathbb{R})$ is defined by $u(x) := |x|$ for $x \in \mathbb{R}$. Then u is weakly differentiable, and $\partial u = \mathrm{sign}$.

Proof First, the absolute value function $|\cdot|$ is smooth on $\mathbb{R}^\times$, and its derivative is $\mathrm{sign} \,|\, \mathbb{R}^\times$ there. Now suppose $\varphi \in \mathcal{D}(\mathbb{R})$. Integration by parts gives

$$\begin{aligned} \int_{\mathbb{R}} \varphi' u \, dx &= \int_0^\infty \varphi' u \, dx + \int_{-\infty}^0 \varphi' u \, dx \\ &= \varphi(x) x \big|_0^\infty - \int_0^\infty \varphi(x) \, dx - \varphi(x) x \big|_{-\infty}^0 + \int_{-\infty}^0 \varphi(x) \, dx \\ &= -\int_{\mathbb{R}} \varphi(x) \, \mathrm{sign}(x) \, dx \,. \end{aligned}$$

The claim follows since sign belongs to $L_{1,\mathrm{loc}}(\mathbb{R})$. ■

(e) The function sign belongs to $L_{1,\mathrm{loc}}(\mathbb{R})$ and is smooth on $\mathbb{R}^\times$. Nevertheless it is not weakly differentiable. Thus the absolute value function of item (d) is not twice weakly differentiable.

Proof For $\varphi \in \mathcal{D}(\mathbb{R})$, we have

$$\int_{\mathbb{R}} \varphi' \operatorname{sign} dx = \int_0^\infty \varphi'(x)\,dx - \int_{-\infty}^0 \varphi'(x)\,dx = -2\varphi(0)\ . \tag{7.20}$$

Were sign weakly differentiable, then there would be a $v \in L_{1,\mathrm{loc}}(\mathbb{R})$ such that

$$\int_{\mathbb{R}} \varphi v\,dx = 2\varphi(0) \quad \text{for } \varphi \in \mathcal{D}(\mathbb{R})\ ,$$

which is false: see Example 7.19(c). ■

In terms of the Dirac distribution δ, (7.20) assumes the form

$$\int_{\mathbb{R}} \varphi' \operatorname{sign} dx = -2\delta(\varphi) \quad \text{for } \varphi \in \mathcal{D}(X)\ .$$

Denoting the duality pairing as usual by

$$\langle\,\cdot\,,\cdot\,\rangle : \mathcal{D}'(X) \times \mathcal{D}(X) \to \mathbb{K}\ ,$$

so $\langle T, \varphi\rangle$ is the value of the continuous linear functional T on the element φ, we have

$$\langle \operatorname{sign}, \varphi'\rangle = -2\langle \delta, \varphi\rangle \quad \text{for } \varphi \in \mathcal{D}(\mathbb{R})\ , \tag{7.21}$$

where we have identified sign $\in L_{1,\mathrm{loc}}(\mathbb{R})$ with the regular distribution $T_{\mathrm{sign}} \in \mathcal{D}'(X)$, as discussed right after the proof of Theorem 7.20. A comparison of (7.19) and (7.21) suggests the following definition: Let $S, T \in \mathcal{D}'(X)$ and $\alpha \in \mathbb{N}^n$. Then S is called the α-th **distributional derivative** of T if

$$\langle T, \partial^\alpha \varphi\rangle = (-1)^{|\alpha|}\langle S, \varphi\rangle \quad \text{for } \varphi \in \mathcal{D}(X)\ .$$

In this case, S is clearly defined by T (and α), so we can set $\partial^\alpha T := S$. We see easily that every distribution has distributional derivatives of every order and that for every $\alpha \in \mathbb{N}^n$ the distributional derivative

$$\partial^\alpha : \mathcal{D}'(X) \to \mathcal{D}'(X)\ , \quad T \mapsto \partial^\alpha T$$

is a linear map.[17] In particular, (7.21) shows that, in the sense of distributions,

$$\partial(\operatorname{sign}) = 2\delta\ .$$

We cannot go any further here into the theory of distributions, but we want to briefly introduce Sobolev spaces. Suppose $m \in \mathbb{N}$ and $1 \le p \le \infty$. Because

[17] See Exercise 13.

$L_p(X) \subset L_{1,\mathrm{loc}}(X)$, every $u \in L_p(X)$ has distributional derivatives of all orders. We set[18]

$$W_p^m(X) := \left\{ u \in L_p(X) \ ; \ \partial^\alpha u \in L_p(X), \ |\alpha| \le m \right\} ,$$

where ∂^α denotes the α-th distributional derivative. Also let

$$\|u\|_{m,p} := \begin{cases} \Bigl(\sum\limits_{|\alpha|\le m} \|\partial^\alpha u\|_p^p \Bigr)^{1/p} , & 1 \le p < \infty , \\ \max\limits_{|\alpha|\le m} \|\partial^\alpha u\|_\infty , & p = \infty . \end{cases} \tag{7.22}$$

We verify easily that

$$W_p^m(X) := \bigl(W_p^m(X), \|\cdot\|_{m,p} \bigr)$$

is a normed vector space, called the **Sobolev space of order m**. In particular, $W_p^0(X) = L_p(X)$.

7.27 Theorem

(i) *$W_p^m(X)$ is continuously embedded in $L_p(X)$, and $u \in L_p(X)$ belongs to $W_p^m(X)$ if and only if u is m-times weakly differentiable and all weak derivatives of order $\le m$ belong to $L_p(X)$.*

(ii) *$W_p^m(X)$ is a Banach space.*

Proof (i) This is obvious.

(ii) Let (u_j) be a Cauchy sequence in $W_p^m(X)$. It follows immediately from (7.22) that $(\partial^\alpha u_j)_{j\in\mathbb{N}}$ is a Cauchy sequence in $L_p(X)$ for every $\alpha \in \mathbb{N}^n$ such that $|\alpha| \le m$. Because $L_p(X)$ is complete, there exists a unique $u_\alpha \in L_p(X)$ such that $\partial^\alpha u_j \to u_\alpha$ in $L_p(X)$ for $j \to \infty$ and $|\alpha| \le m$. We set $u := u_0$. Then it follows from (7.19) that, for all $j \in \mathbb{N}$,

$$\int_X (\partial^\alpha \varphi) u_j \, dx = (-1)^{|\alpha|} \int_X \varphi \partial^\alpha u_j \, dx \quad \text{for } \varphi \in \mathcal{D}(X) \text{ and } |\alpha| \le m . \tag{7.23}$$

From Hölder's inequality, we deduce

$$\Bigl| \int_X (\partial^\alpha \varphi) u_j \, dx - \int_X (\partial^\alpha \varphi) u \, dx \Bigr| = \Bigl| \int_X \partial^\alpha \varphi (u_j - u) \, dx \Bigr| \le \|\partial^\alpha \varphi\|_{p'} \, \|u_j - u\|_p ,$$

which shows that

$$\int_X (\partial^\alpha \varphi) u_j \, dx \to \int_X (\partial^\alpha \varphi) u \, dx \quad \text{for } \varphi \in \mathcal{D}(X) .$$

Analogously, we find that

$$\int_X \varphi \partial^\alpha u_j \, dx \to \int_X \varphi u_\alpha \, dx \quad \text{for } \varphi \in \mathcal{D}(X) .$$

[18] If X is an interval in $\mathbb{R}$, then one can show that $W_1^1(X)$ coincides with the space introduced in Exercise 5.6.

Thus it follows from (7.23) that

$$\int_X (\partial^\alpha \varphi) u \, dx = (-1)^{|\alpha|} \int_X \varphi u_\alpha \, dx \quad \text{for } \varphi \in \mathcal{D}(X) \ .$$

This shows that u_α is the α-th weak derivative of u, and we see that u is m-times weakly differentiable. Because $u_\alpha \in L_p(X)$ for $|\alpha| \leq m$, we also have $u \in W_p^m(X)$, and it is clear that $u_j \to u$ in $W_p^m(X)$. Therefore $W_p^m(X)$ is complete. ■

7.28 Corollary *$W_2^m(X)$ is a Hilbert space with the inner product*

$$(u \,|\, v)_m := \sum_{|\alpha| \leq m} (\partial^\alpha u \,|\, \partial^\alpha v) \quad \text{for } u, v \in W_2^m(X) \ .$$

We will conclude this section by proving the so-called trace theorem for Sobolev spaces. For $m \in \mathbb{N}$ and $1 \leq p < \infty$, set

$$\widehat{H}_p^m(X) := \bigl(\{ \, u \,|\, X \ ; \ u \in C_c^m(\mathbb{R}^n) \, \}, \ \|\cdot\|_{m,p} \bigr) \ .$$

Clearly $\widehat{H}_p^m(X)$ is a vector subspace of $W_p^m(X)$. If the boundary ∂X of X is sufficiently nice (for example, if $\overline{X} \subset \mathbb{R}^n$ is an n-dimensional submanifold with boundary,[19]) one can show that $\widehat{H}_p^m(X)$ is dense in $W_p^m(X)$. In particular, this is the case for $X := \mathbb{R}^n$ or $X := \mathbb{H}^n$.

7.29 Theorem (trace theorem) *Let $1 \leq p < \infty$ and $X = \mathbb{R}^n$ or $X = \mathbb{H}^n$. Then there is a unique* **trace operator** $\gamma \in \mathcal{L}\bigl(W_p^1(X), L_p(\mathbb{R}^{n-1})\bigr)$ *such that $\gamma u = u \,|\, \mathbb{R}^{n-1}$ for $u \in \mathcal{D}(\mathbb{R}^n)$ (more precisely, for $u \in \widehat{H}_p^1(X)$). Here $\mathbb{R}^{n-1}$ is identified with $\mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$.*

Proof Since $\widehat{H}_p^1(X)$ is dense in $W_p^1(X)$, the claim follows from Proposition 6.24, Remark 6.25, and Theorem VI.2.6. ■

This theorem says in particular that every element $u \in W_p^1(\mathbb{H}^n)$ has boundary values $\gamma u \in L_p(\partial \mathbb{H}^n)$. Because u is generally not continuous on $\overline{\mathbb{H}}^n$, γu cannot be simply determined by restriction.

The existence of a trace is the foundation for the treatment of boundary value problems in partial differential equations by the methods of functional analysis.

Exercises

1 For $a > 0$, calculate $\chi_{[-a,a]} * \chi_{[-a,a]}$ and $\chi_{[-a,a]} * \chi_{[-a,a]} * \chi_{[-a,a]}$.

2 Let $p, p' \in (1, \infty)$ satisfy $1/p + 1/p' = 1$. Prove:

---

[19] See Section XI.1.

(a) $f * g$ belongs to C_0 for $(f, g) \in \mathcal{L}_p \times \mathcal{L}_{p'}$, and $\|f * g\|_\infty \le \|f\|_p \|g\|_{p'}$.

(b) The convolution is a well defined, bilinear, continuous map from $L_p \times L_{p'}$ into C_0.

3 Let $p, q, r \in [1, \infty]$ with $1/p + 1/q = 1 + 1/r$. Verify that

$$* : \mathcal{L}_p \times \mathcal{L}_q \to \mathcal{L}_r , \quad (f, g) \mapsto f * g$$

is well defined, bilinear, and continuous. Also verify the **generalized Young inequality**

$$\|f * g\|_r \le \|f\|_p \|g\|_q \quad \text{for } (f, g) \in \mathcal{L}_p \times \mathcal{L}_q .$$

(Hint: The cases $r = 1$ and $r = \infty$ are covered by Theorem 7.3 and Exercise 2, respectively. For $r \in (1, \infty)$, consider

$$|f(x-y)g(y)| = |f(x-y)|^{1-p/r} \left(|f(x-y)|^p \, |g(y)|^q\right)^{1/r} |g(y)|^{1-q/r}$$

and apply Hölder's inequality.)

4 Show that $f * g$ belongs to C^k for $(f, g) \in C_c^k \times L_{1,\mathrm{loc}}$.

5 Suppose $f \in \mathcal{L}_{1,\mathrm{loc}}$ satisfies $\partial^\alpha f \in \mathcal{L}_{1,\mathrm{loc}}$ for a given $\alpha \in \mathbb{N}^n$. Verify that

$$\partial^\alpha (f * \varphi) = (\partial^\alpha f) * \varphi = f * \partial^\alpha \varphi \quad \text{for } \varphi \in BC^\infty .$$

6 Exhibit a vector subspace of Funct in which $(\mathbb{R}, +)$ is not linearly representable.

7 Given $p \in [1, \infty)$, suppose $K \subset L_p$ is compact. Prove that for every $\varepsilon > 0$ there is a $\delta > 0$ such that $\|\tau_a f - f\|_p < \varepsilon$ for all $f \in K$ and all $a \in \mathbb{R}^n$ with $|a| < \delta$. (Hint: Recall Theorem III.3.10 and Theorem 5.1(iv).)

8 Show that every nontrivial ideal of $(L_1, *)$ is dense in L_1.

9 Let $p \in [1, \infty]$, and denote by k the Gaussian kernel of Example 7.12(a). Prove:

(a) $\partial^\alpha k \in \mathcal{L}_p$ for $\alpha \in \mathbb{N}^n$.

(b) $k * u \in BUC^\infty$ for $u \in L_p$.

10 Let $f \in \mathcal{L}_1$, and suppose $\partial^\alpha f \in \mathcal{L}_1$ for some $\alpha \in \mathbb{N}^n$. Show that

$$\int (\partial^\alpha f) \varphi \, dx = (-1)^{|\alpha|} \int f \partial^\alpha \varphi \, dx \quad \text{for } \varphi \in BC^\infty .$$

11 Let $V \in \{\, \mathrm{Funct}, B, L_p \; ; \; 1 \le p \le \infty \,\}$. Show that the linear representation of $(\mathbb{R}^n, +)$ on V by translations is a group isomorphism.

12 For $f, g, h \in \mathcal{L}_0$, suppose f is convolvable with g and g with h. If $f * g$ is convolvable with h and f with $g * h$, show that $(f * g) * h = f * (g * h)$. Thus convolution is associative on L_1.

13 Show that the distributional derivative

$$\partial^\alpha : \mathcal{D}'(X) \to \mathcal{D}'(X) , \quad T \mapsto \partial^\alpha T$$

is a well defined linear map for every $\alpha \in \mathbb{N}^n$.

14 Show that $(f \mapsto fu) \in \mathcal{L}(BC^m(X), W_p^m(X))$ for $u \in W_p^m(X)$ with $1 \le p \le \infty$ and $m \in \mathbb{N}$.

15 Suppose (T_j) is a sequence in $\mathcal{D}'(X)$ and that $T \in \mathcal{D}'(X)$. We say (T_j) **converges in** $\mathcal{D}'(X)$ to T if

$$\lim_j \langle T_j, \varphi \rangle = \langle T, \varphi \rangle \quad \text{for } \varphi \in \mathcal{D}(X) \ .$$

Let $\{\, \varphi_\varepsilon \ ; \ \varepsilon > 0 \,\}$ be an approximation to the identity, and let (ε_j) be a null sequence. Show that $(\varphi_{\varepsilon_j})$ converges in $\mathcal{D}'(\mathbb{R}^n)$ to δ.

8 The substitution rule

In our treatment of the Cauchy–Riemann integral, we encountered the substitution rule of Theorem VI.5.1 as an essential tool for calculating integrals. Introducing new variables, that is, choosing appropriate coordinates, is a prominent technique also in higher dimensional integration. Unsurprisingly, the proof of the substitution rule in this case is more difficult. However, we have already laid a foundation in the form of the substitution rule for linear maps, which we derived in Theorem IX.5.25.

Besides proving the general substitution rule for n-dimensional Lebesgue integrals, this section will illustrate its significance by means of some important examples. The same theorem is also the cornerstone of the theory of integration on manifolds, the subject of our last chapter.

In the following, suppose

- X and Y are open subsets of $\mathbb{R}^n$;
  E is a Banach space.

Pulling back the Lebesgue measure

Let $(X, \mathcal{A})$ be a measurable space and $(Y, \mathcal{B}, \nu)$ a measure space. If $f : X \to Y$ is a *bijective* map that satisfies $f(\mathcal{A}) \subset \mathcal{B}$, that is, one whose inverse map is $\mathcal{B}$-$\mathcal{A}$-measurable, one easily verifies that

$$f^*\nu : \mathcal{A} \to [0, \infty] \ , \quad A \mapsto \nu\bigl(f(A)\bigr)$$

defines a measure on $\mathcal{A}$, the **pull back** (or the **inverse image**) of the measure ν by f. In the special case $(X, \mathcal{A}) = \bigl(\mathbb{R}^n, \mathcal{L}(n)\bigr)$ and $(Y, \mathcal{B}, \nu) = \bigl(\mathbb{R}^n, \mathcal{L}(n), \lambda_n\bigr)$, the particular case of the substitution rule covered in Theorem IX.5.25 describes the pull back of λ_n by automorphisms of $\mathbb{R}^n$:

$$\Phi^*\lambda_n = |\det \Phi|\, \lambda_n \quad \text{for } \Phi \in \mathcal{L}\text{aut}(\mathbb{R}^n) \ .$$

Using this result, we will now determine the pull back of the Lebesgue measure by arbitrary C^1-diffeomorphisms. A technical result is essential to that end:

8.1 Lemma *Suppose $\Phi \in \text{Diff}^1(X, Y)$. Then*

$$\lambda_n\bigl(\Phi(J)\bigr) \le \int_J |\det \partial\Phi|\, dx$$

for every interval $J \subset\subset X$ of the form $[a, b)$, where $a, b \in \mathbb{Q}^n$.

Proof (i) First consider a cube $J = \bigl[x_0 - (r/2)\mathbf{1}, x_0 + (r/2)\mathbf{1}\bigr)$ with center $x_0 \in X$ and edge length $r > 0$. Next set $\mathbb{R}^n_\infty := (\mathbb{R}^n, |\cdot|_\infty)$ and

$$K := \max_{x \in \bar{J}} \|\partial\Phi(x)\|_{\mathcal{L}(\mathbb{R}^n_\infty)} \ .$$

It follows from the mean value theorem that

$$|\Phi(x) - \Phi(x_0)|_\infty \le K\,|x - x_0|_\infty \quad \text{for } x \in J\ .$$

Therefore $\Phi(J)$ is contained in $\overline{\mathbb{B}}^n_\infty\big(\Phi(x_0), Kr/2\big)$, and we find

$$\lambda_n(\Phi(J)) \le (Kr)^n = K^n \lambda_n(J)\ . \tag{8.1}$$

(ii) Suppose $J \subset\subset X$ is of the form $[a, b)$, with $a, b \in \mathbb{Q}^n$. Take $\varepsilon > 0$ and let $M := \max_{x\in\bar J} \|[\partial\Phi(x)]^{-1}\|_{\mathcal{L}(\mathbb{R}^n_\infty)}$. Since $\partial\Phi$ is uniformly continuous on $\bar J$, there exists $\delta > 0$ such that

$$\|\partial\Phi(x) - \partial\Phi(y)\|_{\mathcal{L}(\mathbb{R}^n_\infty)} \le \varepsilon/M \tag{8.2}$$

for all $x, y \in \bar J$ such that $|x - y| < \delta$. Because $a, b \in \mathbb{Q}^n$, we can decompose J (by edge subdivision) into N disjoint cubes J_k of the form $[\alpha, \beta)^n$ with $0 < \beta - \alpha < \delta$. Now choose $x_k \in \bar J_k$ such that

$$|\det \partial\Phi(x_k)| = \min_{y\in \bar J_k} |\det \partial\Phi(y)|$$

and set $T_k := \partial\Phi(x_k)$ and $\Phi_k := T_k^{-1} \circ \Phi$. Because

$$\partial\Phi_k(y) = T_k^{-1}\partial\Phi(y) = 1_n + [\partial\Phi(x_k)]^{-1}\,[\partial\Phi(y) - \partial\Phi(x_k)]$$

it follows from (8.2) and the definition of M that

$$\max_{y\in\bar J_k} \|\partial\Phi_k(y)\|_{\mathcal{L}(\mathbb{R}^n_\infty)} \le 1 + \varepsilon \quad \text{for } k \in \{1, \ldots, N\}\ . \tag{8.3}$$

By the special case of the substitution rule treated in Theorem IX.5.25, we have

$$\lambda_n(\Phi(J_k)) = \lambda_n(T_k T_k^{-1}\Phi(J_k)) = |\det T_k|\,\lambda_n(\Phi_k(J_k))\ .$$

Thus (8.1) and (8.3) imply

$$\lambda_n(\Phi(J_k)) \le (1 + \varepsilon)^n\,|\det T_k|\,\lambda_n(J_k) \quad \text{for } k \in \{1, \ldots, N\}\ .$$

Taking into account the bijectivity of Φ and the choice of x_k, we find

$$\begin{aligned}
\lambda_n(\Phi(J)) &= \lambda_n\Big(\bigcup_{k=1}^N \Phi(J_k)\Big) = \sum_{k=1}^N \lambda_n(\Phi(J_k)) \\
&\le (1+\varepsilon)^n \sum_{k=1}^N |\det T_k|\,\lambda_n(J_k) \le (1+\varepsilon)^n \sum_{k=1}^N \int_{J_k} |\det \partial\Phi|\,dx \\
&= (1+\varepsilon)^n \int_J |\det \partial\Phi|\,dx\ .
\end{aligned}$$

The claim follows upon taking the limit $\varepsilon \to 0$. ■

8.2 Proposition *Suppose* $\Phi \in \mathrm{Diff}^1(X,Y)$. *Then*

$$\Phi^*\lambda_n(A) = \lambda_n\bigl(\Phi(A)\bigr) = \int_A |\det \partial\Phi|\,dx \quad \textit{for } A \in \mathcal{L}(n)\,|\,X\ .$$

Proof (i) From the monotone convergence theorem, it follows easily that

$$\mu_\Phi : \mathcal{L}(n)\,|\,X \to [0,\infty]\ , \quad A \mapsto \int_A |\det \partial\Phi|\,dx$$

is a complete measure (see Exercise 2.11).

(ii) Suppose U is open and compactly contained in X. By Proposition IX.5.6, there is a sequence (J_k) of disjoint intervals of the form $[a,b)$, with $a,b \in \mathbb{Q}^n$, such that $U = \bigcup_k J_k$. From (i) and Lemma 8.1, it follows that

$$\begin{aligned}\lambda_n(\Phi(U)) &= \lambda_n\Bigl(\bigcup\nolimits_k \Phi(J_k)\Bigr) = \sum\nolimits_k \lambda_n(\Phi(J_k)) \le \sum\nolimits_k \int_{J_k} |\det \partial\Phi|\,dx \\ &= \sum\nolimits_k \mu_\Phi(J_k) = \mu_\Phi\Bigl(\bigcup\nolimits_k J_k\Bigr) = \mu_\Phi(U) = \int_U |\det \partial\Phi|\,dx\ .\end{aligned}$$

(iii) Let U be open in X. By Remarks 1.16(d) and (e), there is a sequence (U_k) of open subsets of X such that $U_k \subset\subset U_{k+1}$ and $U = \bigcup_k U_k$. From (ii) and the continuity from below of the measures λ_n and μ_Φ, it follows that

$$\lambda_n(\Phi(U)) = \lim_k \lambda_n(\Phi(U_k)) \le \lim_k \mu_\Phi(U_k) = \mu_\Phi(U) = \int_U |\det \partial\Phi|\,dx\ .$$

(iv) Let $A \in \mathcal{L}(n)\,|\,X$ be bounded. Using Corollary IX.5.5, we find a sequence (U_k) of bounded open subsets of X such that $G := \bigcap_k U_k \supset A$ and $\lambda_n(G) = \lambda_n(A)$. From (iii) and the continuity from above of the measures λ_n and μ_Φ, we have

$$\begin{aligned}\lambda_n(\Phi(G)) &= \lim_k \lambda_n\Bigl(\Phi\Bigl(\bigcap_{j=0}^k U_j\Bigr)\Bigr) \le \lim_k \mu_\Phi\Bigl(\bigcap_{j=0}^k U_j\Bigr) \\ &= \mu_\Phi(G) = \int_G |\det \partial\Phi|\,dx\ .\end{aligned}$$

Noting that $A \subset G$ and $\lambda_n(A) = \lambda_n(G)$, we obtain

$$\lambda_n(\Phi(A)) \le \lambda_n(\Phi(G)) \le \int_G |\det \partial\Phi|\,dx = \int_A |\det \partial\Phi|\,dx\ .$$

(v) Take any $A \in \mathcal{L}(n)\,|\,X$, and set $A_k := A \cap k\mathbb{B}^n$ for $k \in \mathbb{N}$. From (iv) and the continuity of the measures from below, we obtain

$$\lambda_n(\Phi(A)) = \lim_k \lambda_n(\Phi(A_k)) \le \lim_k \mu_\Phi(A_k) = \mu_\Phi(A) = \int_A |\det \partial\Phi|\,dx\ .$$

(vi) Let $f \in \mathcal{S}(Y, \mathbb{R}^+)$ have normal form $f = \sum_{j=0}^{k} \alpha_j \chi_{A_j}$. By (v),

$$\begin{aligned}\int_Y f\,dy &= \sum_{j=0}^{k} \alpha_j \lambda_n(A_j) = \sum_{j=0}^{k} \alpha_j \lambda_n\bigl(\Phi(\Phi^{-1}(A_j))\bigr) \\ &\le \sum_{j=0}^{k} \alpha_j \int_{\Phi^{-1}(A_j)} |\det \partial\Phi|\,dx = \int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx\ .\end{aligned}$$

(vii) Suppose X is bounded. Given $f \in \mathcal{L}_0(Y, \mathbb{R}^+)$, let (f_k) be a sequence in $\mathcal{S}(Y, \mathbb{R}^+)$ such that $f_k \uparrow f$ (see Theorem 1.12). Then $f_k \circ \Phi$ belongs to $\mathcal{S}(X, \mathbb{R}^+)$. Because the sequence $(f_k \circ \Phi)_k$ converges increasingly to $f \circ \Phi$, we know that $(f \circ \Phi)\,|\det \partial\Phi|$ belongs to $\mathcal{L}_0(X, \mathbb{R}^+)$. Now (vi) and the monotone convergence theorem imply

$$\int_Y f\,dy = \lim_k \int_Y f_k\,dy \le \lim_k \int_X (f_k \circ \Phi)\,|\det \partial\Phi|\,dx = \int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx\ .$$

(viii) Let X be arbitrary and take $f \in \mathcal{L}_0(Y, \mathbb{R}^+)$. In view of Remarks 1.16(d) and (e), we can find an ascending sequence of relatively compact open subsets X_k of X such that $X = \bigcup_{k=0}^{\infty} X_k$. According to (vii), $g_k := \chi_{X_k} f\,|\det \Phi|$ belongs to $\mathcal{L}_0(X, \mathbb{R}^+)$, and we have $g_k \uparrow g := f\,|\det \Phi|$. Therefore $g \in \mathcal{L}_0(X, \mathbb{R}^+)$. Setting $Y_k := \Phi(X_k)$, we obtain from (vii) that

$$\int_{Y_k} f\,dy \le \int_{X_k} (f \circ \Phi)\,|\det \partial\Phi|\,dx\ .$$

Now $Y = \bigcup_{k=0}^{\infty} Y_k$ and the monotone convergence theorem yield

$$\int_Y f\,dy \le \int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx\ . \tag{8.4}$$

(ix) Suppose $A \in \mathcal{L}(n)\,|\,X$. We swap the roles of X and Y in (viii) and apply (8.4) to the C^1-diffeomorphism $\Phi^{-1} : Y \to X$ and the function $(\chi_{\Phi(A)} \circ \Phi)\,|\det \partial\Phi|$, which belongs to $\mathcal{L}_0(X, \mathbb{R}^+)$. Then

$$\begin{aligned}\int_X (\chi_{\Phi(A)} \circ \Phi)\,|\det \partial\Phi|\,dx &\le \int_Y \Bigl[\bigl((\chi_{\Phi(A)} \circ \Phi)\,|\det \partial\Phi|\bigr) \circ \Phi^{-1}\Bigr]\,|\det \partial\Phi^{-1}|\,dy \\ &= \int_Y \chi_{\Phi(A)}\,\bigl|\det\bigl[(\partial\Phi \circ \Phi^{-1})\partial\Phi^{-1}\bigr]\bigr|\,dy\ .\end{aligned}$$

Further noting that

$$1_n = \partial(\mathrm{id}_Y) = \partial(\Phi \circ \Phi^{-1}) = (\partial\Phi \circ \Phi^{-1})\partial\Phi^{-1} \tag{8.5}$$

and $\chi_{\Phi(A)} \circ \Phi = \chi_A$, we obtain

$$\int_A |\det \partial\Phi|\,dx \le \int_Y \chi_{\Phi(A)}\,dy = \lambda_n(\Phi(A))\ .$$

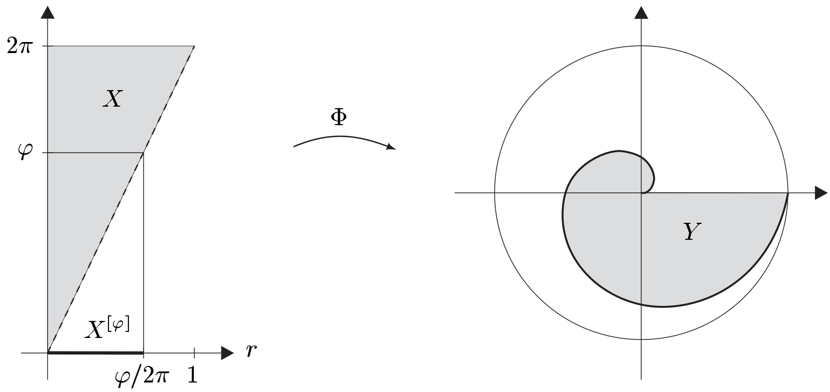
Because of (v), the claim follows. ■

8.3 Example Define $X := \{ (r,\varphi) \in \mathbb{R} \times (0, 2\pi) \; ; \; 0 < r < \varphi/2\pi \}$ and

$$\Phi : X \to \mathbb{R}^2 \; , \quad (r,\varphi) \mapsto (r\cos\varphi, r\sin\varphi) \; .$$

Then $Y := \Phi(X)$ is open in $\mathbb{R}^2$, and $\Phi \in \mathrm{Diff}^\infty(X,Y)$ satisfies

$$\bigl[\partial\Phi(r,\varphi)\bigr] = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix} \; .$$



Therefore $\det\partial\Phi(r,\varphi) = r$. Also $\mathrm{pr}_2(X) = (0, 2\pi)$, and $X^{[\varphi]} = (0, \varphi/2\pi)$ for $\varphi \in (0, 2\pi)$. By Proposition 8.2 and Tonelli's theorem, then,

$$\lambda_2(Y) = \int_X r \, d(r,\varphi) = \int_0^{2\pi} \int_0^{\varphi/2\pi} r \, dr \, d\varphi = \pi/3 \; .$$

The substitution rule: general case

After these preliminaries, it is no longer difficult to prove the substitution rule for diffeomorphisms. First we consider the scalar case, whose proof is accessible even to readers who skipped over the proof of Fubini's theorem for vector-valued functions. We treat the general case at the end of the section.

8.4 Theorem (substitution rule) *Suppose* $\Phi \in \mathrm{Diff}^1(X,Y)$.

(i) *For* $f \in \mathcal{L}_0(Y, \mathbb{R}^+)$,

$$\int_Y f \, dy = \int_X (f \circ \Phi) \, |\det\partial\Phi| \, dx \; . \tag{8.6}$$

(ii) *A function* $f : Y \to \mathbb{K}$ *is integrable if and only if* $(f \circ \Phi) \, |\det\partial\Phi|$ *belongs to* $\mathcal{L}_1(X)$. *In this case,* (8.6) *holds.*

Proof (i) Theorem IX.5.12 implies that $\Phi(\mathcal{L}_X) \subset \mathcal{L}_Y$. Hence $f \circ \Phi$ is measurable, by Corollary 1.5. Since $|\det\partial\Phi|$ is continuous, hence measurable, Remark 1.2(d) implies that $g := (f \circ \Phi)\,|\det \partial\Phi|$ is measurable also. From (8.5) we obtain $f = (g \circ \Phi^{-1})|\det \partial\Phi^{-1}|$. Now (8.4), with (Y, Φ^{-1}, g) in the role of (X, Φ, f), gives

$$\int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx \le \int_Y f\,dy\ .$$

Because of (8.4), this implies (8.6). Now (ii) follows from (i), parts (ii) and (iii) of Corollary 2.12, and Theorem 3.14. ■

In terms of the pull back of functions defined in Section VIII.3, the substitution rule (8.6) takes on the easily remembered form

$$\int_Y f\,d\lambda_n = \int_{\Phi^{-1}(Y)} (\Phi^* f)\,d(\Phi^* \lambda_n)\ .$$

This follows from Proposition 8.2 and Exercise 2.12.

For many applications, the assumption that Φ is a diffeomorphism is too restrictive. We weaken it somewhat in this simple yet important generalization of Theorem 8.4:[1]

8.5 Corollary *Let M be a measurable subset of X such that $M \backslash \overset{\circ}{M}$ has Lebesgue measure zero. Suppose $\Phi \in C^1(X, \mathbb{R}^n)$ is such that $\Phi \,|\, \overset{\circ}{M}$ is a diffeomorphism from $\overset{\circ}{M}$ onto $\Phi(\overset{\circ}{M})$.*

(i) *For every $f \in \mathcal{L}_0(M, \mathbb{R}^+)$,*

$$\int_{\Phi(M)} f\,dy = \int_M (f \circ \Phi)\,|\det \partial\Phi|\,dx\ . \tag{8.7}$$

(ii) A function $f : \Phi(M) \to \mathbb{K}$ *belongs to* $\mathcal{L}_1(\Phi(M))$ *if and only if* $(f \circ \Phi)\,|\det \partial\Phi|$ *belongs to* $\mathcal{L}_1(M)$. *In this case,* (8.7) *holds.*

Proof Because $\lambda_n(M \backslash \overset{\circ}{M}) = 0$, the set $\Phi(M) \backslash \Phi(\overset{\circ}{M}) \subset \Phi(M \backslash \overset{\circ}{M})$ also has measure zero, by Corollary IX.5.10. The claims then follow from Lemma 2.15 and Theorem 8.4. ■

It is clear that this corollary gives a (partial) generalization of the substitution rule of Theorem VI.5.1, though limited to diffeomorphisms. There is one obvious difference from the one-dimensional case considered before: now the derivative term (that is, the functional determinant) appears as an absolute value. The reason is that the prior result used the *oriented* integral.

[1] See Exercise 7 for a further generalization.

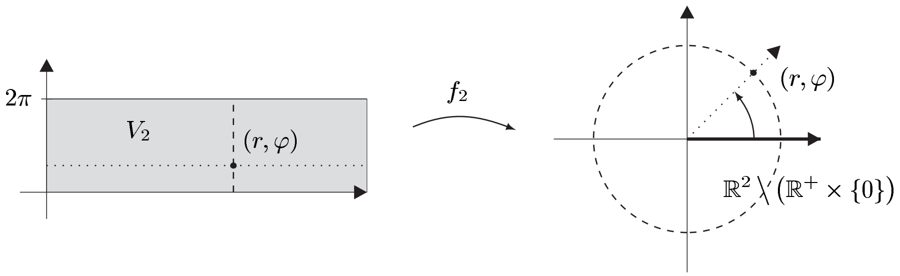
Plane polar coordinates

A special case of special importance in applications consists of diffeomorphisms induced by polar coordinates, which we now introduce. We begin with the two-dimensional case.

Let

$$f_2 : \mathbb{R}^2 \to \mathbb{R}^2 , \quad (r, \varphi) \mapsto (x, y) := (r \cos \varphi, r \sin \varphi)$$

be the (**plane**) **polar coordinate map**[2], and let $V_2 := (0, \infty) \times (0, 2\pi)$.



Then f_2 is smooth, and $\det \partial f_2(r, \varphi) = r$, as was shown in Example 8.3. Clearly $\overline{V}_2 \backslash V_2$ has measure zero; moreover

$$f_2(\overline{V}_2) = \mathbb{R}^2 , \quad f_2(V_2) = \mathbb{R}^2 \setminus (\mathbb{R}^+ \times \{0\}) \tag{8.8}$$

and

$$f_2 \,|\, V_2 \in \mathrm{Diff}^\infty(V_2, f_2(V_2)) . \tag{8.9}$$

Therefore Corollary 8.5 applies with $M := \overline{V}_2$:

8.6 Proposition (integration in polar coordinates)

(i) *For $g \in \mathcal{L}_0(\mathbb{R}^2, \mathbb{R}^+)$, we have*

$$\begin{aligned} \int_{\mathbb{R}^2} g(x, y)\, d(x, y) &= \int_0^{2\pi} \int_0^\infty g(r \cos \varphi, r \sin \varphi) r \, dr \, d\varphi \\ &= \int_0^\infty r \int_0^{2\pi} g(r \cos \varphi, r \sin \varphi) \, d\varphi \, dr . \end{aligned} \tag{8.10}$$

(ii) *The function $g : \mathbb{R}^2 \to \mathbb{K}$ is integrable if and only if the map*

$$(0, \infty) \times (0, 2\pi) \to \mathbb{K} , \quad (r, \varphi) \mapsto g(r \cos \varphi, r \sin \varphi) r$$

is integrable. Then (8.10) *holds.*

Proof This follows from Corollary 8.5 together with (8.8), (8.9), and the Fubini–Tonelli theorem. ■

[2]See Conclusion III.6.21(d).

These integrals simplify when f depends only on $|x|$, that is, on r. To illustrate, we present an elegant calculation of the Gaussian error integral, for which knowledge of the Γ-function is not required (compare Application VI.9.7).

8.7 Example $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$.

Proof Tonelli's theorem implies

$$\Bigl(\int_{-\infty}^{\infty} e^{-x^2}\,dx\Bigr)^2 = \int_{-\infty}^{\infty} e^{-x^2}\,dx \int_{-\infty}^{\infty} e^{-y^2}\,dy = \int_{\mathbb{R}}\Bigl(\int_{\mathbb{R}} e^{-(x^2+y^2)}\,dx\Bigr)\,dy$$
$$= \int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,d(x,y)\ .$$

Therefore Proposition 8.6(i) shows that

$$\Bigl(\int_{-\infty}^{\infty} e^{-x^2}\,dx\Bigr)^2 = \int_0^{2\pi}\int_0^{\infty} re^{-r^2}\,dr\,d\varphi = 2\pi\int_0^{\infty}\frac{d}{dr}\bigl[-e^{-r^2}/2\bigr]\,dr = \pi\ ,$$

and the claim follows. ■

Polar coordinates in higher dimensions

For $n \ge 1$, we define $h_n : \mathbb{R}^n \to \mathbb{R}^{n+1}$ recursively through

$$h_1(z) := (\cos z, \sin z) \quad \text{for } z \in \mathbb{R} \tag{8.11}$$

and

$$h_{n+1}(z) := \bigl(h_n(z')\sin z_{n+1}, \cos z_{n+1}\bigr) \quad \text{for } z = (z', z_{n+1}) \in \mathbb{R}^n \times \mathbb{R}\ . \tag{8.12}$$

Obviously h_n is smooth, and by induction, we verify that

$$|h_n(z)| = 1 \quad \text{for } z \in \mathbb{R}^n\ . \tag{8.13}$$

Now we define $f_n : \mathbb{R}^n \to \mathbb{R}^n$ for $n \ge 2$ by

$$f_n(y) := y_1 h_{n-1}(z) \quad \text{for } y = (y_1, z) \in \mathbb{R} \times \mathbb{R}^{n-1}\ . \tag{8.14}$$

Then f_n is also smooth, and we have

$$h_{n-1}(z) = f_n(1,z)\ , \quad |f_n(y)| = |y_1|\ . \tag{8.15}$$

We will usually follow convention by renaming the y-coordinates as

$$(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2}) := (y_1, y_2, y_3, \ldots, y_n)\ .$$

By induction, one checks easily that

$$f_n : \mathbb{R}^n \to \mathbb{R}^n\ , \quad (r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2}) \mapsto (x_1, x_2, x_3, \ldots, x_n) \tag{8.16}$$

is given by

$$\left.\begin{aligned} x_1 &= r\cos\varphi\sin\vartheta_1\sin\vartheta_2\cdots\sin\vartheta_{n-2}\ ,\\ x_2 &= r\sin\varphi\sin\vartheta_1\sin\vartheta_2\cdots\sin\vartheta_{n-2}\ ,\\ x_3 &= r\cos\vartheta_1\sin\vartheta_2\cdots\sin\vartheta_{n-2}\ ,\\ &\ \vdots\\ x_{n-1} &= r\cos\vartheta_{n-3}\sin\vartheta_{n-2}\ ,\\ x_n &= r\cos\vartheta_{n-2}\ . \end{aligned}\right\} \tag{8.17}$$

Thus f_2 coincides with the plane polar coordinate map, and f_3 is the **spherical coordinate map** of Example VII.9.11(a). In the general case, f_n is the ***n*-dimensional polar coordinate map**. From (8.12) and (8.14), the recursive relation

$$f_n(y) = \bigl(f_{n-1}(y')\sin y_n, y_1\cos y_n\bigr) \quad \text{for } y = (y', y_n) \in \mathbb{R}^{n-1}\times\mathbb{R} \tag{8.18}$$

follows for $n \geq 3$. For $n \geq 2$, we set

$$W_{n-1} := (0, 2\pi)\times(0,\pi)^{n-2}\ ,\quad V_n := (0,\infty)\times W_{n-1}\ , \tag{8.19}$$

and

$$V_n(r) := (0, r)\times W_{n-1} \quad \text{for } r > 0\ . \tag{8.20}$$

If we denote the closed $(n-1)$-dimensional half-space by

$$H_{n-1} := \mathbb{R}^+\times\{0\}\times\mathbb{R}^{n-2}\subset\mathbb{R}^n\ , \tag{8.21}$$

we find

$$h_{n-1}(W_{n-1}) = S^{n-1}\backslash H_{n-1}\ ,\quad f_n(V_n(r)) = r\mathbb{B}^n\backslash H_{n-1} \tag{8.22}$$

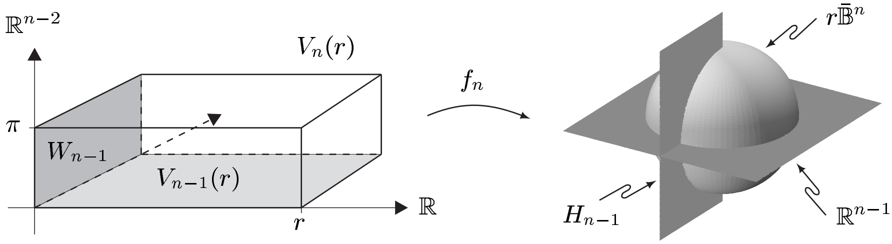
and

$$h_{n-1}\bigl(\overline{W}_{n-1}\bigr) = S^{n-1}\ ,\quad f_n\bigl(\overline{V_n(r)}\bigr) = r\bar{\mathbb{B}}^n\ . \tag{8.23}$$

Also

$$f_n(V_n) = \mathbb{R}^n\backslash H_{n-1}\ ,\quad f_n\bigl(\overline{V}_n\bigr) = \mathbb{R}^n\ . \tag{8.24}$$

In addition, the maps $h_{n-1}\,|\,W_{n-1}$ and $f_n\,|\,V_n$ are bijective onto their images.



These statements follow easily by induction.

8.8 Lemma *For $n \geq 3$ and $r > 0$, the map f_n is a C^∞ diffeomorphism from $V_n(r)$ onto $r\mathbb{B}^n \backslash H_{n-1}$ and from V_n onto $\mathbb{R}^n \backslash H_{n-1}$. Moreover*

$$\det \partial f_n(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2}) = (-1)^n r^{n-1} \sin \vartheta_1 \sin^2 \vartheta_2 \cdots \sin^{n-2} \vartheta_{n-2}$$

for $(r, \varphi, \vartheta_1, \ldots, \vartheta_{n-2}) \in \overline{V}_n$.

Proof In view of the foregoing, we need only calculate the value of the functional determinant $\det \partial f_n(y)$. We do this recursively. From (8.12) and (8.14), we have

$$\begin{aligned} \big[\partial f_{n+1}(y)\big] &= \left[\begin{array}{c:ccc} h_{n-1}(z') \sin z_n & & \big[r\partial_z\big(h_{n-1}(z') \sin z_n\big)\big] & \\ \hdashline \cos z_n & 0 & \cdots \quad 0 & -r \sin z_n \end{array}\right] \\ &= \left[\begin{array}{ccc:c} & & & * \\ & \big[\partial f_n(y') \sin z_n\big] & & \vdots \\ & & & * \\ \hdashline * & \cdots & * & -r \sin z_n \end{array}\right] , \end{aligned}$$

where $y = (r, z) = (y', z_n)$ and $z = (z', z_n) \in \mathbb{R}^n$. Expanding in the last row, we find

$$\det \partial f_{n+1}(y) = (-1)^n \cos z_n \det S - r \sin^{n+1} z_n \det \partial f_n(y') , \tag{8.25}$$

where $S := \big[r\partial_z\big(h_{n-1}(z') \sin z_n\big)\big]$. We can assume that $\sin z_n \neq 0$; otherwise the claim is trivial. In the last column of S we have $r h_{n-1}(z') \cos z_n$. This vector differs only by the factor $r \cot z_n$ from the first column vector, $h_{n-1}(z') \sin z_n$, of the matrix $T := \big[\partial f_n(y') \sin z_n\big]$. The first $n-1$ columns of S also agree with the last $n-1$ columns of T, in the same order. Therefore

$$\det S = (-1)^{n-1} r \cot z_n \det T = (-1)^{n-1} r \cos z_n \sin^{n-1} z_n \det \partial_n f(y') .$$

Thus it follows from (8.25) that

$$\det \partial f_{n+1}(y) = -r \sin^{n-1} z_n \det \partial f_n(y') .$$

The claim now follows because $\det \partial f_2(r, \varphi) = r$. ■

For short, let's set

$$w_n(\vartheta) := \sin \vartheta_1 \sin^2 \vartheta_2 \cdots \sin^{n-2} \vartheta_{n-2} , \quad \vartheta := (\vartheta_1, \ldots, \vartheta_{n-2}) \in [0, \pi]^{n-2} .$$

8.9 Proposition (integration in polar coordinates) *Suppose* $n \geq 3$.

(i) *For* $g \in \mathcal{L}_0(\mathbb{R}^n, \mathbb{R}^+)$, *we have*

$$\int_{\mathbb{R}^n} g\,dx = \int_{V_n} (g \circ f_n)(r, \varphi, \vartheta) r^{n-1} w_n(\vartheta)\, d(r, \varphi, \vartheta)\ . \tag{8.26}$$

(ii) *The map* $g \colon \mathbb{R}^n \to \mathbb{K}$ *is integrable if and only if*

$$V_n \to \mathbb{K}\ , \quad (r, \varphi, \vartheta) \mapsto (g \circ f_n)(r, \varphi, \vartheta) r^{n-1} w_n(\vartheta)$$

is integrable. Then (8.26) *holds.*

Proof Because $\lambda_n\big(\overline{V}_n \backslash V_n\big) = 0$, the claim follows from (8.24), Corollary 8.5, and Lemma 8.8. ■

8.10 Examples **(a)** For $g \in \mathcal{L}_0(\mathbb{R}^3, \mathbb{R}^+)$, we have

$$\begin{aligned}\int_{\mathbb{R}^3} & g(x, y, z)\, d(x, y, z) \\ &= \int_0^\infty \int_0^{2\pi} \int_0^\pi g(r \cos\varphi \sin\vartheta, r \sin\varphi \sin\vartheta, r\cos\vartheta) r^2 \sin\vartheta\, d\vartheta\, d\varphi\, dr\ .\end{aligned} \tag{8.27}$$

The integrals on the right side can be performed in any order.

Proof This follows from Proposition 8.9(i) and Tonelli's theorem. ■

(b) A map $g \colon \mathbb{R}^3 \to \mathbb{K}$ is integrable if and only if

$$V_3 \to \mathbb{K}\ , \quad (r, \varphi, \vartheta) \mapsto g(r \cos\varphi \sin\vartheta, r \sin\varphi \sin\vartheta, r \cos\vartheta)\, r^2 \sin\vartheta$$

is integrable. Such a map satisfies (8.27), and the integrals there can be performed in any order.

Proof This is a consequence of Proposition 8.9(ii) and the Fubini–Tonelli theorem. ■

(c) For $n \geq 3$, we have

$$2\pi \int_{[0,\pi]^{n-2}} w_n(\vartheta)\, d\vartheta = n\omega_n\ ,$$

where $\omega_n = \pi^{n/2}/\Gamma(1 + n/2)$ is the volume of $\mathbb{B}^n$.

Proof From (8.22), (8.23), Proposition 8.9, and Tonelli's theorem, it follows that

$$\begin{aligned}\omega_n &= \int_{\mathbb{B}^n} dx = \int_{\mathbb{B}^n} \mathbf{1}\, dx = \int_{V_n(1)} (\mathbf{1} \circ f_n)(r, \varphi, \vartheta) r^{n-1} w_n(\vartheta)\, d(r, \varphi, \vartheta) \\ &= \int_0^1 r^{n-1}\, dr \int_0^{2\pi} d\varphi \int_{[0,\pi]^{n-2}} w_n(\vartheta)\, d\vartheta = \frac{2\pi}{n} \int_{[0,\pi]^{n-2}} w_n(\vartheta)\, d\vartheta\ ,\end{aligned}$$

and we are done. ■

Integration of rotationally symmetric functions

Suppose $0 \le r_0 < r_1 \le \infty$ and set $\mathsf{R}(r_0, r_1) := \{ x \in \mathbb{R}^n ; r_0 < |x| < r_1 \}$. We say that a function $g : \mathsf{R}(r_0, r_1) \to E$ is **rotationally symmetric** if there is a map $\overset{\bullet}{g} : (r_0, r_1) \to E$ such that

$$g(x) = \overset{\bullet}{g}(|x|) \quad \text{for } x \in \mathsf{R}(r_0, r_1) \ .$$

This is the case if and only if g is constant on every sphere rS^{n-1} with $r_0 < r < r_1$. For such a function, $\overset{\bullet}{g}$ is uniquely determined by g (and vice versa).

As we saw in Example 8.7, integration problems simplify considerably for rotationally symmetric functions.

8.11 Theorem *Suppose* $0 \le r_0 < r_1 \le \infty$.

(i) *If* $g \in \mathcal{L}_0\big(\mathsf{R}(r_0, r_1), \mathbb{R}^+\big)$ *is rotationally symmetric, then*

$$\int_{\mathsf{R}(r_0,r_1)} g \, dx = n\omega_n \int_{r_0}^{r_1} \overset{\bullet}{g}(r) \, r^{n-1} \, dr \ , \tag{8.28}$$

where $\omega_n := \lambda_n(\mathbb{B}^n) = \pi^{n/2}/\Gamma(1 + n/2)$.

(ii) *A rotationally symmetric function* $g : \mathsf{R}(r_0, r_1) \to \mathbb{K}$ *is integrable if and only if*

$$(r_0, r_1) \to \mathbb{K} \ , \quad r \mapsto \overset{\bullet}{g}(r) \, r^{n-1}$$

is integrable. In this case (8.28) *holds.*

Proof The case $n = 1$ is clear (see Exercise 5.12). For $n \ge 2$, it follows from (8.15) and the rotational symmetry of g that

$$g \circ f_n(r, \varphi, \vartheta) = \overset{\bullet}{g}(r) \quad \text{for } r_0 < r < r_1 \text{ and } (\varphi, \vartheta) \in W_{n-1} \ .$$

Now the claim arises from Propositions 8.6 and 8.9 (applied to the trivial extension of g) and Example 8.10(c). ■

8.12 Examples **(a)** Suppose $f : \mathbb{R}^n \to \mathbb{K}$ is measurable and there are $c \ge 0$, $\rho > 0$, and $\varepsilon > 0$ such that

$$|f(x)| \le \begin{cases} c\,|x|^{-n+\varepsilon} & \text{if } 0 < |x| \le \rho \ , \\ c\,|x|^{-n-\varepsilon} & \text{if } \quad\ |x| \ge \rho \ . \end{cases}$$

Then f is integrable.

Proof Set

$$g(x) := c\big(|x|^{-n+\varepsilon} \, \chi_{\rho\bar{\mathbb{B}}^n}(x) + |x|^{-n-\varepsilon} \, \chi_{(\rho\mathbb{B}^n)^c}(x)\big) \quad \text{for } x \in \mathbb{R}^n \backslash \{0\} = \mathsf{R}(0, \infty) \ .$$

Then g is rotationally symmetric, and $|f(x)| \le g(x)$ for $x \in \mathsf{R}(0, \infty)$. By Examples VI.8.4(a) and (b), $r \mapsto \overset{\bullet}{g}(r) r^{n-1}$ belongs to $\mathcal{L}_1(\mathbb{R}^+)$. Hence Theorem 8.11 implies that g also belongs to $\mathcal{L}_1\big(\mathsf{R}(0, \infty)\big) = \mathcal{L}_1(\mathbb{R}^n)$. Now the claim follows from Theorem 3.14. ■

(b) Let $\mu \in \mathcal{L}_\infty(\mathbb{R}^n)$ have compact support. Also define

$$\frac{1}{r} : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}^+ \ , \quad x \mapsto \frac{1}{|x|} \ .$$

Then $(1/r)^\alpha * \mu$ exists for $\alpha < n$, and

$$\left(\frac{1}{r}\right)^\alpha * \mu(x) = \int_{\mathbb{R}^n} \frac{\mu(y)}{|x-y|^\alpha} \, dy \quad \text{for } x \in \mathbb{R}^n \ .$$

Proof Take $x \in \mathbb{R}^n$ and define $K := \operatorname{supp}(\mu)$ and $g_x(y) := \|\mu\|_\infty \, |y|^{-\alpha} \, \chi_{x-K}(y)$ for $y \neq 0$. Then $\widetilde{g}_x$ belongs to $\mathcal{L}_0(\mathbb{R}^n)$, and

$$|\mu(x-y)| \, |y|^{-\alpha} \leq g_x(y) \quad \text{for } y \neq 0 \ .$$

Because $\alpha < n$, part (a) shows that $\widetilde{g}_x$ is integrable. The claim now follows from Theorem 7.8(ii). ■

For $n \geq 3$, the function $u_n := (1/r)^{n-2} * \mu$, in the notation of (b), is called the **Newtonian** or **Coulomb potential** associated with the density μ. From Exercise 3.6 we know that u_n is smooth and harmonic in K^c, and (b) shows that u_n is defined on all of $\mathbb{R}^n$.

The substitution rule for vector-valued functions

We now prove the substitution formula of Theorem 8.4 for vector-valued functions.

8.13 Lemma *Let $f \in \mathcal{S}_c(Y, E)$ and $\Phi \in \operatorname{Diff}^1(X, Y)$. Then $(f \circ \Phi) \, |\det \partial\Phi|$ belongs to $\mathcal{L}_1(X, E)$, and*

$$\int_Y f \, dy = \int_X (f \circ \Phi) \, |\det \partial\Phi| \, dx \ .$$

Proof Because $\operatorname{supp}(f \circ \Phi) = \Phi^{-1}\big(\operatorname{supp}(f)\big)$, the support of $f \circ \Phi$ is compact. In particular, $f \circ \Phi$ belongs to $\mathcal{S}_c(X, E)$. It easily follows that $(f \circ \Phi) \, |\det \partial\Phi|$ is integrable. Also Theorem 2.11(iii) shows that, for $e \in E$ and $g \in \mathcal{L}_1(X, \mathbb{K})$, the function eg belongs to $\mathcal{L}_1(X, E)$ and $e \int_X g \, dx = \int_X eg \, dx$. Letting $\sum_{j=0}^m e_j \chi_{A_j}$ be the normal form of f, we see from Proposition 8.2 that

$$\begin{aligned}
\int_Y f \, dy &= \sum_{j=0}^m e_j \lambda_n(A_j) = \sum_{j=0}^m e_j \int_{\Phi^{-1}(A_j)} |\det \partial\Phi| \, dx \\
&= \sum_{j=0}^m \int_{\Phi^{-1}(A_j)} e_j \, |\det \partial\Phi| \, dx = \int_X (f \circ \Phi) \, |\det \partial\Phi| \, dx \ . \quad ■
\end{aligned}$$

8.14 Theorem (substitution rule) *Let $\Phi \in \mathrm{Diff}^1(X,Y)$ and $f \in E^Y$. Then f belongs to $\mathcal{L}_1(Y,E)$ if and only if $(f \circ \Phi)\,|\det \partial\Phi|$ belongs to $\mathcal{L}_1(X,E)$. In this case, we have*

$$\int_Y f\,dy = \int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx\ .$$

Proof (i) Let $f \in \mathcal{L}_1(Y,E)$, and take a sequence (f_j) in $\mathcal{S}_c(Y,E)$ converging a.e. in $\mathcal{L}_1(Y,E)$ to f and satisfying $\lim \int_Y f_j = \int_Y f$ (see Lemma 6.18, Remarks 6.19(a) and (c), and Theorem 2.18). Set $g_j := (f_j \circ \Phi)\,|\det \partial\Phi|$ for $j \in \mathbb{N}$. Thanks to Lemma 8.13, we know that (g_j) is a Cauchy sequence in $\mathcal{L}_1(X,E)$ and that $\int_Y f_j\,dy = \int_X g_j\,dx$. Because $\mathcal{L}_1(X,E)$ is complete, there exists $g \in \mathcal{L}_1(X,E)$ such that $g_j \to g$ in $\mathcal{L}_1(X,E)$. Also, it follows from Theorem 2.18 that $\int_X g_j\,dx$ converges to $\int_X g\,dx$ and that some subsequence $(g_{j_k})_{k\in\mathbb{N}}$ of (g_j) converges a.e. in X to g. Hence g and $(f \circ \Phi)\,|\det \partial\Phi|$ coincide a.e. in X. By Lemma 2.15, $(f \circ \Phi)\,|\det \partial\Phi|$ belongs to $\mathcal{L}_1(X,E)$, and $\int_X g = \int_X (f\circ\Phi)\,|\det \partial\Phi|$. It follows that

$$\int_Y f\,dy = \lim_j \int_Y f_j\,dy = \lim_j \int_X g_j\,dx = \int_X g\,dx = \int_X (f \circ \Phi)\,|\det \partial\Phi|\,dx\ .$$

(ii) For the converse, suppose $(f \circ \Phi)\,|\det \partial\Phi|$ belongs to $\mathcal{L}_1(X,E)$. From (8.5) we have

$$f = \bigl((f \circ \Phi)\,|\det \partial\Phi|\bigr) \circ \Phi^{-1}\,|\det \partial(\Phi^{-1})|\,,$$

so part (i) shows that f belongs to $\mathcal{L}_1(Y,E)$. ■

It is clear that Corollary 8.5 is also true for E-valued maps. From this it follows that Propositions 8.6(ii) and 8.9(ii) and Theorem 8.11(ii) also hold for E-valued functions.

Exercises

1 Let $G \in \mathbb{R}^{n\times n}$ be symmetric and positive definite. Prove that

$$\int_{\mathbb{R}^n} e^{-(Gx\,|\,x)}\,dx = \pi^{n/2}/\sqrt{\det G}\ .$$

(Hint: principal axis transformation.)

2 Show that for $p \in \mathbb{C}$ with $\mathrm{Re}\,p > n/2$, we have

$$\int_{\mathbb{R}^n} (1+|x|^2)^{-p}\,dx = \pi^{n/2}\Gamma(p-n/2)/\Gamma(p)\ .$$

(Hint: Look at Example 6.13(b).)

3 Suppose $D := \bigl\{\,(x,y) \in \mathbb{R}^2\ ;\ x,y \ge 0,\ x+y \le 1\,\bigr\}$ and $p,q \in (0,\infty)$. Show that for $f : (0,1) \to \mathbb{R}$, the function

$$D \to \mathbb{R}\ ,\quad (x,y) \mapsto x^{p-1}y^{q-1}f(x+y)$$

is integrable if and only if $s \mapsto s^{p+q-1} f(s)$ belongs to $\mathcal{L}_1((0,1))$. In this case, we have

$$\int_D x^{p-1} y^{q-1} f(x+y)\, d(x,y) = \mathsf{B}(p,q) \int_0^1 s^{p+q-1} f(s)\, ds \ .$$

(Hint: Consider $(s,t) \mapsto \bigl(s(1-t), st\bigr)$.)

4 Let $0 \le \alpha < \beta \le 2\pi$, and suppose $f : [\alpha, \beta] \to (0, \infty)$ is measurable. Show that

$$S(\alpha, \beta, f) := \bigl\{\, z \in \mathbb{C} \ ; \ \arg_N(z) \in [\alpha, \beta],\ |z| \le f\bigl(\arg_N(z)\bigr) \,\bigr\}$$

is Lebesgue measurable and that

$$\lambda_2\bigl(S(\alpha, \beta, f)\bigr) = \frac{1}{2} \int_\alpha^\beta \bigl[f(\varphi)\bigr]^2 \, d\varphi \ .$$

5 Suppose $g \in \mathcal{L}^2_{\mathrm{sym}}(\mathbb{R}^n)$ is positive definite. Calculate the volume of the solid ellipsoid $g^{-1}([0,1])$ enclosed by the surface $g^{-1}(1)$ (see Remark VII.10.18).

6 (Sard's lemma) Suppose $\Phi \in C^1(X, \mathbb{R}^n)$, and let $C := \bigl\{\, x \in X \ ; \ \partial\Phi(x) \notin \mathcal{L}\mathrm{aut}(\mathbb{R}^n) \,\bigr\}$ be the set of critical points of Φ. Show that $\Phi(C)$ has measure zero. (Hint: Because C is σ-compact, it suffices to check that $\Phi(C \cap J)$ is has measure zero for every compact n-dimensional cube J. Take $x_0 \in C$ and $r > 0$ such that $J_0 := \bigl[x_0 - (r/2)\mathbf{1}, x_0 + (r/2)\mathbf{1}\bigr]$ is compactly contained in X, and set

$$\rho(r) := \max_{x \in J_0} \int_0^1 \bigl\| \partial\Phi\bigl(x_0 + t(x - x_0)\bigr) \bigr\| \, dt \ .$$

Show that there is a $c_n > 0$ such that $\lambda_n(\Phi(J_0)) \le c_n r^n \rho(r)$. Because $\lim_{r \to 0} \rho(r) = 0$, the claim follows by subdividing the edges of J_0.)

7 Suppose $\Phi \in C^1(X, \mathbb{R}^n)$ and $C := \bigl\{\, x \in X \ ; \ \partial\Phi(x) \notin \mathcal{L}\mathrm{aut}(\mathbb{R}^n) \,\bigr\}$. Also suppose $\Phi \,|\, (X \backslash C)$ is injective. Prove:

(i) For $f \in \mathcal{L}_0(X, \mathbb{R}^+)$,

$$\int_{\Phi(X)} f \, dy = \int_X (f \circ \Phi)\, |\det \partial\Phi| \, dx \ . \tag{8.29}$$

(ii) The function $f : \Phi(X) \to E$ belongs to $\mathcal{L}_1\bigl(\Phi(X), E\bigr)$ if and only if $(f \circ \Phi)\, |\det \partial\Phi|$ lies in $\mathcal{L}_1(X, E)$. In this case, (8.29) holds.

9 The Fourier transform

To conclude this chapter, we introduce the most important integral transformation, called the Fourier transform.[1] The study of its fundamental properties is as it were a recapitulation of Lebesgue integration theory: we will encounter at every turn such cornerstones as the completeness of Lebesgue spaces, the dominated convergence theorem, and the Fubini–Tonelli theorem.

Particularly appealing is the interaction of the Fourier transform with the convolution and with the Hilbert space structure of L_2. We illustrate the former through Fourier multiplication operators and the second via Plancherel's theorem and applications of the position and momentum operators of quantum mechanics.

In this section, we exclusively consider spaces of complex-valued functions defined on all of $\mathbb{R}^n$. For this reason, as in Section 7, we omit $(\mathbb{R}^n, \mathbb{C})$ from our notation and write, for example, $\mathcal{L}_1$ for $\mathcal{L}_1(\mathbb{R}^n, \mathbb{C})$. In addition, $\int f\,dx$ always means $\int_{\mathbb{R}^n} f\,dx$, and we canonically identify $\mathbb{R}^n$ with its dual space, so that $\langle\cdot\,,\cdot\rangle$ coincides formally with the Euclidean inner product.

Definition and elementary properties

Let $f \in \mathcal{L}_1$. The map $\mathbb{R}^n \to \mathbb{C},\ x \mapsto e^{-i\langle x,\xi\rangle} f(x)$ belongs to $\mathcal{L}_1$ for every $\xi \in \mathbb{R}^n$. The map $\widehat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\widehat{f}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} f(x)\,dx \quad \text{for } \xi \in \mathbb{R}^n \tag{9.1}$$

is called the **Fourier transform of f**. The map $\mathcal{F} := \bigl(f \mapsto \widehat{f}\bigr)$ is also called the Fourier transform (or, if necessary to avoid confusion, the **Fourier transformation**).

Different conventions intervene in the definition just given; instead of (9.1), one often sees the Fourier transform being defined as[2]

$$\xi \mapsto \int e^{-i\langle x,\xi\rangle} f(x)\,dx \quad \text{or} \quad \xi \mapsto \int e^{-2\pi i\langle x,\xi\rangle} f(x)\,dx\ .$$

Obviously, these differences in normalization are immaterial to the underlying theory; however, they do cause powers of 2π to appear in some of the following expressions. One should be mindful of this when reading the literature. The normalization chosen here has the advantage that such factors appear only in a few places and that Plancherel's theorem takes on a particularly simple form.

9.1 Remarks **(a)** For $f \in L_1$, set $\mathcal{F}f := \widehat{f} := \mathcal{F}\overset{*}{f}$, where $\overset{*}{f}$ is an arbitrary representative of f. Then $\mathcal{F}f$ is well defined, and $\mathcal{F} \in \mathcal{L}(L_1, BC)$.

[1] The contents of this section will not be used in the rest of this book.
[2] See Section VIII.6.

Proof The first statement is obvious. Because

$$\left|\widehat{f}(\xi)\right| \le (2\pi)^{-n/2}\,\|f\|_1 \quad \text{for } \xi \in \mathbb{R}^n\ ,$$

the second follows easily from Theorem 3.17 (on the continuity of parameterized integrals) and Theorem VI.2.5. ■

(b) For $f \in L_1$, we have $\widehat{\widecheck{f}} = \widecheck{\widehat{f}}$. The function defined by

$$\mathbb{R}^n \to \mathbb{C}\ , \quad \xi \mapsto \widecheck{\widehat{f}}(\xi) = (2\pi)^{-n/2} \int e^{i\,\langle x,\xi\rangle} f(x)\,dx$$

is called the **inverse Fourier transform** of f, for reasons soon to become clear; and the map

$$\overline{\mathcal{F}} := (f \mapsto \widecheck{\widehat{f}}\,)$$

is the **inverse Fourier transform**(**ation**). Because inversion $(f \mapsto \widecheck{f}\,)$ is a continuous automorphism on $\mathcal{L}_1$, L_1, and BC, the inverse Fourier transform has the same continuity properties as the Fourier transform.

Proof This follows immediately from the substitution rule. ■

(c) For $\lambda > 0$, we denote by $\sigma_\lambda : \mathbb{R}^n \to \mathbb{R}^n,\ x \mapsto \lambda x$ the **dilation** by the factor λ. We define an action of the group $\bigl((0,\infty),\cdot\bigr)$ on $\mathrm{Funct} := \mathrm{Funct}(\mathbb{R}^n,\mathbb{C})$,

$$\bigl((0,\infty),\cdot\bigr) \times \mathrm{Funct} \to \mathrm{Funct}\ , \quad (\lambda, f) \mapsto \sigma_\lambda f\ , \tag{9.2}$$

by setting

$$\sigma_\lambda f := f \circ \sigma_{1/\lambda} = (\sigma_{1/\lambda})^* f\ .$$

If V is a vector subspace of Funct that is invariant under this action (meaning that $\sigma_\lambda(V) \subset V$ for $\lambda > 0$), the map

$$\sigma_\lambda : V \to V\ , \quad v \mapsto \sigma_\lambda v$$

is linear and satisfies $\sigma_\lambda \sigma_\mu = \sigma_{\lambda\mu}$ and $\sigma_1 = \mathrm{id}_V$ for $\lambda, \mu > 0$. Therefore σ_λ is a vector space automorphism of V, with $(\sigma_\lambda)^{-1} = \sigma_{1/\lambda}$ for $\lambda > 0$. This shows that

$$\bigl((0,\infty),\cdot\bigr) \to \mathrm{Aut}(V)\ , \quad \lambda \mapsto \sigma_\lambda$$

is a linear representation of the multiplicative group $\bigl((0,\infty),\cdot\bigr)$ on V. In particular, $\{\,\sigma_\lambda\ ;\ \lambda > 0\,\}$ is a subgroup of $\mathrm{Aut}(V)$, the **group of dilations** on V. Accordingly $\sigma_\lambda v$ is the **dilation of v** by the factor λ. As with the translation group, we say that $\bigl((0,\infty),\cdot\bigr)$ is **linearly representable** in V if V is invariant under (9.2).

Suppose $1 \le p \le \infty$. Then $\bigl((0,\infty),\cdot\bigr)$ is linearly representable on L_p, and

$$\|\sigma_\lambda f\|_p = \lambda^{n/p}\,\|f\|_p\ .$$

Proof This follows from the substitution rule. ■

(d) $\mathcal{F}\sigma_\lambda = \lambda^n \sigma_{1/\lambda}\mathcal{F}$ for $\lambda > 0$.

Proof Suppose $f \in \mathcal{L}_1$ and $\lambda > 0$. Then

$$\mathcal{F}\sigma_\lambda f(\xi) = (2\pi)^{-n/2} \int e^{-i\langle x,\xi\rangle} f(x/\lambda)\,dx = \lambda^n (2\pi)^{-n/2} \int e^{-\langle x/\lambda,\lambda\xi\rangle} f(x/\lambda)\lambda^{-n}\,dx$$

for $\xi \in \mathbb{R}^n$. But the substitution rule shows that the last expression is equal to $\lambda^n \widehat{f}(\lambda\xi)$. ■

(e) Suppose $a \in \mathbb{R}^n$. Then $\bigl(e^{i\langle a,\cdot\rangle} f\bigr)^\wedge = \tau_a \widehat{f}$ for $f \in \mathcal{L}_1$. ■

The space of rapidly decreasing functions

We now introduce a vector subspace of $\mathcal{L}_1$ where the Fourier transform is especially manageable. Using density arguments, we will then be able to broaden the results to larger function spaces.

We say $f \in C^\infty$ is **rapidly decreasing** if for every $(k,m) \in \mathbb{N}^2$, there is a $c_{k,m} > 0$ such that

$$(1+|x|^2)^k\,|\partial^\alpha f(x)| \le c_{k,m} \quad \text{for } x \in \mathbb{R}^n\,,\quad \alpha \in \mathbb{N}^n\,,\quad \text{and}\quad |\alpha| \le m\ .$$

In other words, $f \in C^\infty$ is rapidly decreasing if, as $|x| \to \infty$, every derivative $\partial^\alpha f$ goes to zero faster than any power of $1/|x|$.

We now set

$$q_{k,m}(f) := \max_{|\alpha|\le m} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^{k/2}\,|\partial^\alpha f(x)| \quad \text{for } f \in C^\infty \text{ and } k,m \in \mathbb{N}\ .$$

The space

$$\mathcal{S} := \bigl\{\, f \in C^\infty\ ;\ q_{k,m}(f) < \infty \text{ for } k,m \in \mathbb{N} \,\bigr\}$$

is called **Schwartz space** or the **space of rapidly decreasing functions**.

9.2 Remarks **(a)** $\mathcal{S}$ is a vector subspace of BUC^∞. Every $q_{k,m}$ is a norm on $\mathcal{S}$.

Proof Let $m \in \mathbb{N}$. Then $\mathcal{S}$ is a vector subspace of BC^m, since $q_{0,m}$ is the norm on BC^m. Let $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$. Then it follows easily from the mean values theorem that $\partial^\alpha f$ is uniformly continuous. The proves the first statement. The second is clear. ■

(b) For $(f,g) \in \mathcal{S} \times \mathcal{S}$, let

$$d(f,g) := \sum_{k,m=0}^{\infty} 2^{-(k+m)}\, \frac{q_{k,m}(f-g)}{1+q_{k,m}(f-g)}\ .$$

Then $(\mathcal{S},d)$ is a metric space.

Proof (i) Clearly the double series $\sum 2^{-(k+m)} q_{k,m}(f)/\bigl(1+q_{k,m}(f)\bigr)$ converges for every $f \in \mathcal{S}$. Thus $d : \mathcal{S}\times\mathcal{S} \to \mathbb{R}^+$ is well defined. Also d is symmetric and vanishes identically on the diagonal of $\mathcal{S} \times \mathcal{S}$.

(ii) Because $t \mapsto t/(1+t)$ is increasing on $\mathbb{R}^+$, we have, for $r, s, t \in \mathbb{R}^+$ with $r \le s+t$,

$$\frac{r}{1+r} \le \frac{s+t}{1+s+t} = \frac{s}{1+s+t} + \frac{t}{1+s+t} \le \frac{s}{1+s} + \frac{t}{1+t} .$$

Now it follows easily that d satisfies the triangle inequality. ■

(c) For $f \in \mathcal{S}$ and a sequence (f_j) in $\mathcal{S}$, there is equivalence between:

(i) $\lim f_j = f$ in $(\mathcal{S}, d)$;

(ii) $\lim(f - f_j) = 0$ in $(\mathcal{S}, d)$;

(iii) $\lim_j q_{k,m}(f - f_j) = 0$ for $k, m \in \mathbb{N}$.

Thus a sequence (f_j) converges in $\mathcal{S}$ to f if and only if $(f_j - f)$ converges to zero with respect to every seminorm $q_{k,m}$.

Proof "(i)⇒(ii)" This implication is clear.

"(ii)⇒(iii)" Take $\varepsilon \in (0, 1]$ and $k, m \in \mathbb{N}$. There exists an $N \in \mathbb{N}$ such that the inequality $d(f, f_j) < \varepsilon/2^{k+m+1}$ is satisfied for $j \ge N$. Thus

$$\frac{2^{-(k+m)} q_{k,m}(f - f_j)}{1 + q_{k,m}(f - f_j)} < \frac{\varepsilon}{2^{k+m+1}} ,$$

so $q_{k,m}(f - f_j) < \varepsilon$ for $j \ge N$.

"(iii)⇒(i)" Take $\varepsilon > 0$. There is an $N \in \mathbb{N}$ such that

$$\sum_{k+m=N+1}^{\infty} \frac{2^{-(k+m)} q_{k,m}(f - f_j)}{1 + q_{k,m}(f - f_j)} \le \sum_{\ell=N+1}^{\infty} 2^{-\ell} < \frac{\varepsilon}{2} .$$

By assumption, there is $M \in \mathbb{N}$ such that $q_{k,m}(f - f_j) \le \varepsilon/4$ for $j \ge M$ and $k + m \le N$. Therefore

$$d(f, f_j) \le \sum_{k,m=0}^{N} \frac{2^{-(k+m)} q_{k,m}(f - f_j)}{1 + q_{k,m}(f - f_j)} + \frac{\varepsilon}{2} \le \varepsilon \quad \text{for } j \ge M . \ ■$$

(d) $\mathcal{D}$ is a dense vector subspace of $\mathcal{S}$. The function $\mathbb{R}^n \to \mathbb{R}$, $x \mapsto e^{-|x|^2}$ belongs to $\mathcal{S}$ but not to $\mathcal{D}$.

Proof It is clear that $\mathcal{D}$ is a vector subspace of $\mathcal{S}$. Suppose $f \in \mathcal{S}$. We choose $\varphi \in \mathcal{D}$ such that $\varphi \,|\, \overline{\mathbb{B}}^n = 1$ and set

$$f_j(x) := f(x)\varphi(x/j) \quad \text{for } x \in \mathbb{R}^n , \quad j \in \mathbb{N}^\times .$$

Then f_j belongs to $\mathcal{D}$, and

$$f(x) - f_j(x) = f(x)\bigl(1 - \varphi(x/j)\bigr) \quad \text{for } x \in \mathbb{R}^n .$$

Therefore $\partial^\alpha(f - f_j)(x) = 0$ for $x \in j\overline{\mathbb{B}}^n$ and $\alpha \in \mathbb{N}^n$. Now Leibniz's rule shows that there is a $c = c(\varphi, m) > 0$ such that

$$\begin{aligned} |\partial^\alpha(f - f_j)(x)| &= \Bigl|\sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta f(x) \partial^{\alpha-\beta}(1 - \varphi)(x/j) j^{-|\alpha-\beta|}\Bigr| \le c \max_{\beta \le \alpha} |\partial^\beta f(x)| \\ &\le c\, q_{k+1,m}(f)(1 + |x|^2)^{-(k+1)/2} \end{aligned}$$

for $x \in \mathbb{R}^n$, $j \in \mathbb{N}^\times$, $k \in \mathbb{N}$, and $|\alpha| \le m$. Setting $C := c\, q_{k+1,m}(f)$, we find

$$\begin{aligned} q_{k,m}(f - f_j) &= \max_{|\alpha|\le m} \sup_{|x|\ge j} (1+|x|^2)^{k/2}\, |\partial^\alpha (f - f_j)(x)| \\ &\le c\, q_{k+1,m}(f) \sup_{|x|\ge j} (1+|x|^2)^{-1/2} \le C/j\ , \end{aligned}$$

and, as $j \to \infty$, the first claim follows from (c). The second one is clear. ■

(e) For $m \in \mathbb{N}$, we have $\mathcal{S} \hookrightarrow BUC^m$.

Proof This follows from (a) and (c). ■

(f) $\mathcal{S}$ is a dense vector subspace of C_0.

Proof Suppose $f \in \mathcal{S}$. Then it follows from (a) and because $|f(x)| \le q_{1,0}(f)(1 + |x|^2)^{-1/2}$ for $x \in \mathbb{R}^n$ that f belongs to C_0. Therefore $\mathcal{S}$ is a vector subspace of C_0. Theorem 7.13 shows $\mathcal{D}$ is a dense vector subspace of C_0, and therefore the claim follows from the inclusions $\mathcal{D} \subset \mathcal{S} \subset C_0$. ■

(g) For $k, m \in \mathbb{N}$, there are positive constants c and C such that

$$c \max_{\substack{|\alpha|\le m\\ |\beta|\le k}} \sup_{x\in\mathbb{R}^n} \bigl|\partial^\alpha \bigl(x^\beta f(x)\bigr)\bigr| \le q_{k,m}(f) \le C \max_{\substack{|\alpha|\le m\\ |\beta|\le k}} \sup_{x\in\mathbb{R}^n} |x^\beta \partial^\alpha f(x)| \quad \text{for } f \in \mathcal{S}\ .$$

Proof This follows easily from the Leibniz rule. ■

(h) Let $f \in \mathcal{S}$ and $\alpha, \beta \in \mathbb{N}^n$. Then $x \mapsto x^\alpha \partial^\beta f(x)$ belongs to $\mathcal{S}$.

Proof This is a consequence of (g). ■

(i) The inversion $f \mapsto \check{f}$ is a continuous automorphism of $\mathcal{S}$.

Proof This is obvious. ■

9.3 Theorem *Let $p \in [1, \infty)$. Then $\mathcal{S}$ is a dense vector subspace of L_p, and there is a $c = c(n,p) > 0$ such that*

$$\|f\|_p \le c\, q_{n+1,0}(f) \quad \textit{for } f \in \mathcal{S}\ . \tag{9.3}$$

Proof For $f \in \mathcal{S}$, we have

$$\begin{aligned} \int |f|^p\, dx &= \int |f(x)|^p\, (1+|x|^2)^{(n+1)p/2} (1+|x|^2)^{-(n+1)p/2}\, dx \\ &\le \bigl(q_{n+1,0}(f)\bigr)^p \int (1+|x|^2)^{-(n+1)p/2}\, dx\ . \end{aligned} \tag{9.4}$$

Further, by Theorem 8.11(i) and because $(n+1)p > n$, we have

$$\int_{[|x|\ge 1]} |x|^{-(n+1)p}\, dx = n\omega_n \int_1^\infty r^{-((n+1)p-n+1)}\, dr < \infty\ .$$

Therefore $\int (1+|x|^2)^{-(n+1)p/2}\, dx$ is also finite, and (9.3) follows from (9.4). In particular, f belongs to L_p, and we see that $\mathcal{S}$ is a vector subspace of L_p. By

Theorem 7.13, $\mathcal{D}$ is a dense vector subspace of L_p, and by Remark 9.2(d), it is contained in $\mathcal{S}$. The claim follows. ■

The convolution algebra $\mathcal{S}$

By Remark 9.2(a) and Theorem 9.3, $\mathcal{S}\times\mathcal{S}$ is contained in $BUC^\infty \times L_1$. Therefore the convolution is defined on $\mathcal{S}\times\mathcal{S}$, and by Corollary 7.9, we have

$$* : \mathcal{S}\times\mathcal{S} \to BUC^\infty \ . \tag{9.5}$$

The next result shows that $f*g$ is actually rapidly decreasing for $(f,g)\in\mathcal{S}\times\mathcal{S}$.

9.4 Proposition *The convolution $\mathcal{S}\times\mathcal{S}$ is a continuous and bilinear map into $\mathcal{S}$.*

Proof (i) We verify next that the convolution $\mathcal{S}\times\mathcal{S}$ maps into $\mathcal{S}$. So suppose $(f,g)\in\mathcal{S}\times\mathcal{S}$ and $k,m\in\mathbb{N}$. By (9.5), it suffices to check that $q_{k,m}(f*g)$ is finite. Because

$$|x|^k \le (|x-y|+|y|)^k = \sum_{j=0}^{k}\binom{k}{j}|x-y|^j\,|y|^{k-j} \quad \text{for } x,y\in\mathbb{R}^n \ ,$$

there is a $c_k>0$ such that

$$\begin{aligned}|x|^k\,|f*g(x)| &\le \int\sum_{j=0}^{k}\binom{k}{j}|x-y|^j\,|f(x-y)|\,|y|^{k-j}\,|g(y)|\,dy\\ &\le c_k q_{k,0}(f)\int(1+|y|^2)^{k/2}\,|g(y)|\,dy \ .\end{aligned}$$

Noting that $\widetilde{c}_n := \int(1+|y|^2)^{-(n+1)/2}\,dy$ is finite, we find

$$|x|^k\,|f*g(x)| \le c_k\widetilde{c}_n\, q_{k,0}(f)q_{k+n+1,0}(g) \ .$$

Thus by Remark 9.2(g), there is a $c=c(k,n)\ge 1$ such that

$$q_{k,0}(f*g) \le c\, q_{k,0}(f)q_{k+n+1,0}(g) \ . \tag{9.6}$$

Finally by Theorem 7.8(iv), we have

$$q_{k,m}(f*g) = \max_{|\alpha|\le m} q_{k,0}\big(\partial^\alpha(f*g)\big) = \max_{|\alpha|\le m} q_{k,0}\big((\partial^\alpha f)*g\big) \ ,$$

and (9.6) implies

$$q_{k,m}(f*g) \le c\max_{|\alpha|\le m} q_{k,0}(\partial^\alpha f)q_{k+n+1,0}(g) = c\, q_{k,m}(f)q_{k+n+1,0}(g) \ . \tag{9.7}$$

(ii) It is clear that the convolution is bilinear. Suppose $(f,g)\in\mathcal{S}\times\mathcal{S}$ and $((f_j,g_j))_{j\in\mathbb{N}}$ is a sequence in $\mathcal{S}\times\mathcal{S}$ such that $(f_j,g_j)\to(f,g)$ in $\mathcal{S}\times\mathcal{S}$ as $j\to\infty$.

Also let

$$\alpha := c\big(q_{k,m}(f) + q_{k+n+1,0}(g) + 1\big)$$

where c is the constant from (9.7). Let $\varepsilon \in (0,1]$. By Remark 9.2(c), there is an $N \in \mathbb{N}$ such that

$$q_{k,m}(f - f_j) < \varepsilon/\alpha \ , \quad q_{k+n+1,0}(g - g_j) < \varepsilon/\alpha \quad \text{for } j \geq N \ .$$

Because

$$f * g - f_j * g_j = (f - f_j) * g + (f_j - f) * (g - g_j) + f * (g - g_j)$$

it follows from (9.7) that

$$\begin{aligned} q_{k,m}(f * g - f_j * g_j) \leq c\big(&q_{k,m}(f - f_j)q_{k+n+1,0}(g) + q_{k,m}(f - f_j)q_{k+n+1,0}(g - g_j) \\ &+ q_{k,m}(f)q_{k+n+1,0}(g - g_j)\big) < \varepsilon \end{aligned}$$

for $j \geq N$. Thus we are done. ■

9.5 Corollary $(\mathcal{S}, +, *)$ *is a subalgebra of the commutative algebra* $(L_1, +, *)$.

Proof This follows from Proposition 9.4 and Theorem 9.3. ■

Calculations with the Fourier transform

We now derive some rules for the Fourier transformation of derivatives and the derivatives of Fourier transforms. It will simplify the presentation of these formulas to set $\Lambda(x) := (1 + |x|^2)^{1/2}$ for $x \in \mathbb{R}^n$ and

$$D_j := -i\partial_j \ , \quad j \in \{1, \ldots, n\} \quad \text{for } D^\alpha := D_1^{\alpha_1} \cdots D_n^{\alpha_n} \ , \quad \alpha \in \mathbb{N}^n \ ,$$

where i is the imaginary unit. As usual, the polynomial function induced by the polynomial $p \in \mathbb{C}[X_1, \ldots, X_n]$ will also be denoted by p.

9.6 Proposition *Suppose* $f \in \mathcal{L}_1$.

(i) *For* $\alpha \in \mathbb{N}^n$, *suppose* $D^\alpha f$ *exists and belongs to* $\mathcal{L}_1$. *Then* $X^\alpha \widehat{f} = \widehat{D^\alpha f}$.

(ii) *For* $m \in \mathbb{N}$, *suppose* $\Lambda^m f$ *belongs to* $\mathcal{L}_1$. *Then* $\widehat{f}$ *belongs to* BC^m, *and*

$$D^\alpha \widehat{f} = (-1)^{|\alpha|}\widehat{X^\alpha f} \quad \text{for } \alpha \in \mathbb{N}^n \ , \quad |\alpha| \leq m \ .$$

Proof (i) Suppose $\{\, \varphi_\varepsilon \ ; \ \varepsilon > 0 \,\}$ is a smoothing kernel. By integration by parts (see Exercise 7.10), it follows that

$$\begin{aligned} \int \xi^\alpha e^{-i\langle x,\xi\rangle}(f * \varphi_\varepsilon)(x)\, dx &= (-1)^{|\alpha|} \int D_x^\alpha(e^{-i\langle x,\xi\rangle})(f * \varphi_\varepsilon)(x)\, dx \\ &= \int e^{-i\langle x,\xi\rangle}\big((D^\alpha f) * \varphi_\varepsilon\big)(x)\, dx \ . \end{aligned} \tag{9.8}$$

Theorem 7.11 and Theorem 2.18(ii) imply that

$$\lim_{\varepsilon\to 0}(2\pi)^{-n/2}\int \xi^\alpha e^{-i\langle x,\xi\rangle}(f*\varphi_\varepsilon)(x)\,dx = \xi^\alpha \widehat{f}(\xi)$$

and

$$\lim_{\varepsilon\to 0}(2\pi)^{-n/2}\int e^{-i\langle x,\xi\rangle}\big((D^\alpha f)*\varphi_\varepsilon\big)(x)\,dx = \widehat{D^\alpha f}(\xi)$$

for $\xi\in\mathbb{R}^n$. Using (9.8), this proves the claim.

(ii) We set $h(x,\xi):=e^{-i\langle x,\xi\rangle}f(x)$ for $(x,\xi)\in\mathbb{R}^n\times\mathbb{R}^n$. Then $h(\cdot,\xi)$ belongs to $\mathcal{L}_1$ for every $\xi\in\mathbb{R}^n$, and $h(x,\cdot)$ belongs to C^∞ for every $x\in\mathbb{R}^n$. Further, we have

$$D_\xi^\alpha h(x,\xi)=(-1)^{|\alpha|}x^\alpha h(x,\xi)\quad\text{for } (x,\xi)\in\mathbb{R}^{2n}\ ,\quad \alpha\in\mathbb{N}^n\ ,$$

and thus

$$|D_\xi^\alpha h(x,\xi)|\le(1+|x|^2)^{|\alpha|/2}\,|h(x,\xi)|=\Lambda^{|\alpha|}(x)\,|f(x)|\ . \tag{9.9}$$

It then follows from the theorem on the differentiation of parametrized integrals that $\widehat{f}$ belongs to C^m and that

$$\begin{aligned}D^\alpha\widehat{f}(\xi)&=(2\pi)^{-n/2}\int D_\xi^\alpha h(x,\xi)\,dx=(2\pi)^{-n/2}(-1)^{|\alpha|}\int x^\alpha h(x,\xi)\,dx\\&=(-1)^{|\alpha|}\widehat{X^\alpha f}(\xi)\end{aligned}$$

for $\xi\in\mathbb{R}^n$ and $\alpha\in\mathbb{N}^n$ with $|\alpha|\le m$. Finally (9.9) shows

$$\big|D^\alpha\widehat{f}(\xi)\big|\le(2\pi)^{-n/2}\int|D_\xi^\alpha h(x,\xi)|\,dx\le(2\pi)^{-n/2}\,\|\Lambda^{|m|}f\|_1<\infty\quad\text{for }\xi\in\mathbb{R}^n\ .$$

Thus $\widehat{f}$ belongs to BC^m. ■

9.7 Proposition *The Fourier transformation maps $\mathcal{S}$ continuously and linearly into itself.*

Proof (i) Suppose $f\in\mathcal{S}$ and $m\in\mathbb{N}$. Then

$$\begin{aligned}\int\Lambda^m(x)\,|f(x)|\,dx&=\int(1+|x|^2)^{(m+n+1)/2}\,|f(x)|\,(1+|x|^2)^{-(n+1)/2}\,dx\\&\le q_{m+n+1,0}(f)\int(1+|x|^2)^{-(n+1)/2}\,dx<\infty\ .\end{aligned}$$

We find using Proposition 9.6(ii) that $\widehat{f}$ belongs to BC^m and thus to BC^∞.

(ii) Suppose $k,m\in\mathbb{N}$ and $\alpha,\beta\in\mathbb{N}^n$ with $|\alpha|\le m$ and $|\beta|\le k$. Also suppose $f\in\mathcal{S}$. Then it follows from Remark 9.2(h) and Theorem 9.3 that $\Lambda^m f$

and $D^\beta(X^\alpha f)$ belong to $\mathcal{L}_1$. Therefore Proposition 9.6 implies

$$\xi^\beta D^\alpha \widehat{f}(\xi) = (-1)^{|\alpha|}\xi^\beta \widehat{X^\alpha f}(\xi) = (-1)^{|\alpha|}\bigl(D^\beta(X^\alpha f)\bigr)\widehat{\ }(\xi) \quad \text{for } \xi \in \mathbb{R}^n . \tag{9.10}$$

Remark 9.2(g) shows there is a $c > 0$ such that

$$\begin{aligned}\bigl|\xi^\beta D^\alpha \widehat{f}(\xi)\bigr| &\le (2\pi)^{-n/2}\int |D^\beta(X^\alpha f)(x)|\,(1+|x|^2)^{(n+1)/2}(1+|x|^2)^{-(n+1)/2}\,dx\\ &\le c\,q_{m+n+1,k}(f)\end{aligned}$$

for $|\alpha| \le m$ and $|\beta| \le k$. Hence there is a $C > 0$ such that

$$q_{k,m}\bigl(\widehat{f}\,\bigr) \le C q_{m+n+1,k}(f) . \tag{9.11}$$

Therefore $\widehat{f}$ belongs to $\mathcal{S}$. The continuity of the Fourier transformation now follows easily from (9.11) and Remark 9.2(c). Thus we are done. ■

9.8 Corollary *For $f \in \mathcal{S}$ and $\alpha \in \mathbb{N}^n$, we have*

$$\widehat{D^\alpha f} = X^\alpha \widehat{f} \quad \textit{and} \quad \widehat{X^\alpha f} = (-1)^{|\alpha|} D^\alpha \widehat{f} .$$

Proof These are special cases of (9.10). ■

Proposition 9.6 and Corollary 9.8 show that the Fourier transformation maps differentiation into multiplication by functions, and conversely. This fact underlies much of its great utility.

It is now easy to improve the statement of Remark 9.1(a) to one saying that the image of L_1 under $\mathcal{F}$ already lies in C_0.

9.9 Proposition[3] (Riemann–Lebesgue) $\mathcal{F} \in \mathcal{L}(L_1, C_0)$.

Proof Proposition 9.7 and $\mathcal{S} \subset C_0$ imply $\mathcal{F}(\mathcal{S}) \subset C_0$. From Theorem 9.3, we know that $\mathcal{S}$ is a dense vector subspace of L_1, and Remark 9.1(a) guarantees that $\mathcal{F}$ maps the space L_1 continuously into BC. The claim now follows because C_0 is a closed vector subspace of BC. ■

9.10 Examples **(a)** For $g := g_n : \mathbb{R}^n \to \mathbb{R},\ x \mapsto e^{-|x|^2/2}$, we have $\widehat{g} = g$.

Proof (i) A property of the exponential implies

$$g_n(x) = g_1(x_1)\cdot \cdots \cdot g_1(x_n) \quad \text{for } x = (x_1,\dots,x_n) \in \mathbb{R}^n .$$

[3] Also known as the Riemann–Lebesgue lemma.

For clarity, we denote by $\mathcal{F}_n$ the Fourier transformation on $\mathbb{R}^n$. Then it follows from the Fubini–Tonelli theorem that

$$\begin{aligned}\mathcal{F}_n(g_n)(\xi) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-i\,\langle x,\xi\rangle}e^{-|x|^2/2}\,dx = (2\pi)^{-n/2}\int_{\mathbb{R}^n}\prod_{j=1}^n e^{-i\,x_j\xi_j}e^{-x_j^2/2}\,dx\\ &= \prod_{j=1}^n (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-i\,x_j\xi_j}e^{-x_j^2/2}\,dx_j = \prod_{j=1}^n \mathcal{F}_1(g_1)(\xi_j)\ .\end{aligned}$$

This shows that it suffices to treat the one-dimensional case.

(ii) Suppose therefore $n = 1$. For $f := \widehat{g}$, we have from Example 8.7 that

$$f(0) = \widehat{g}(0) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = 1\ .$$

Because $xe^{-x^2/2} = -\partial(e^{-x^2/2})$, that is, because $Xg = -\partial g = -i\,Dg$, Corollary 9.8 gives

$$\partial f = \partial\widehat{g} = i\,D\widehat{g} = -i\,\widehat{Xg} = -\widehat{Dg} = -X\widehat{g} = -Xf\ .$$

Therefore f solves the linear initial value problem $y'(t) = -t\,y(t)$ with $y(0) = 1$ on $\mathbb{R}$; its unique solution is g. ■

(b) With the notation of (a) and (7.11), we have

$$\widehat{g(\varepsilon\,\cdot\,)}(\xi) = g_\varepsilon(\xi) \quad \text{for } \xi \in \mathbb{R}^n\ , \quad \varepsilon > 0\ .$$

Proof Because $g(\varepsilon\,\cdot\,) = \sigma_{1/\varepsilon}g$, this follows from (a) and Remark 9.1(d). ■

(c) Suppose

$$\varphi(x) := (2\pi)^{-n/2}e^{-|x|^2} \quad \text{for } x \in \mathbb{R}^n\ ,$$

and let $\varepsilon > 0$. Then $\widehat{\varphi(\varepsilon\,\cdot\,)} = k_\varepsilon$, where $k_1 = k$, is the **Gaussian kernel**.

Proof From $\varphi = (2\pi)^{-n/2}\sigma_{1/\sqrt{2}}\,g$ and Remark 9.1(c), it follows that

$$\varphi(\varepsilon\,\cdot\,) = \sigma_{1/\varepsilon}\varphi = (2\pi)^{-n/2}\sigma_{1/\sqrt{2}\,\varepsilon}\,g = (2\pi)^{-n/2}g\big(\sqrt{2}\,\varepsilon\,\cdot\,\big)\ .$$

Thus we get from (b) that

$$\widehat{\varphi(\varepsilon\,\cdot\,)}(x) = (2\pi)^{-n/2}g_{\sqrt{2}\,\varepsilon}(x) = \varepsilon^{-n}(4\pi)^{-n/2}e^{-|x|^2/4\varepsilon^2} = k_\varepsilon(x)$$

for $x \in \mathbb{R}^n$. ■

The Fourier integral theorem

To prepare for more in-depth study of the Fourier transformation on L_1, we provide the following results.

9.11 Proposition *Suppose $f, g \in L_1$. Then $\widehat{f} g$ and $f \widehat{g}$ belong to L_1, and*

$$\int \widehat{f} g \, dx = \int f \widehat{g} \, dx \ .$$

Proof From Proposition 9.9, it follows easily that $\widehat{f} g$ and $f \widehat{g}$ belong to L_1. Let $\overset{*}{f}$ and $\overset{*}{g}$ be representatives of f and g, respectively. Then Lemma 7.2 shows that

$$h : \mathbb{R}^{2n} \to \mathbb{C} \ , \quad (x,y) \mapsto e^{-i\,\langle x,y\rangle} \overset{*}{f}(x) \overset{*}{g}(y) \tag{9.12}$$

is measurable. Because

$$\int \int |h(x,y)| \, dx \, dy = \|f\|_1 \, \|g\|_1 \ , \tag{9.13}$$

we can apply the Fubini–Tonelli theorem to h, and we find

$$\begin{aligned} \int \widehat{\overset{*}{f}}(y) \overset{*}{g}(y) \, dy &= \int (2\pi)^{-n/2} \int e^{-i\,\langle x,y\rangle} \overset{*}{f}(x) \, dx \, \overset{*}{g}(y) \, dy \\ &= \int (2\pi)^{-n/2} \int e^{-i\,\langle x,y\rangle} \overset{*}{g}(y) \, dy \, \overset{*}{f}(x) \, dx = \int \widehat{\overset{*}{g}}(x) \overset{*}{f}(x) \, dx \ . \end{aligned}$$

Then claim now follows after noting $\widehat{f} = \widehat{\overset{*}{f}}$ and $\widehat{g} = \widehat{\overset{*}{g}}$. ■

We now prove theorems about the inverse of the Fourier transformation for various assumptions on the function and its transform.

9.12 Theorem *For $f \in L_1$, these statements are true:*

(i) $$\lim_{\varepsilon \to 0} (2\pi)^{-n/2} \int e^{i\,\langle\,\cdot\,,\xi\rangle} \widehat{f}(\xi) e^{-\varepsilon^2\,|\xi|^2} \, d\xi = f \qquad \textit{in } L_1 \ .$$

(ii) (Fourier integral theorem for L_1) *If $\widehat{f}$ belongs to L_1, then $f = \overline{\mathcal{F}}\big(\widehat{f}\,\big)$, where $\overline{\mathcal{F}}$ is the Fourier cotransformation.*

Proof (i) We use the notation of Example 9.10 and set

$$\varphi^{\varepsilon}(\xi,y) := e^{i\,\langle \xi,y\rangle} \varphi(\varepsilon\xi) = (2\pi)^{-n/2} e^{i\,\langle y,\xi\rangle} e^{-\varepsilon^2\,|\xi|^2}$$

for $\xi, y \in \mathbb{R}^n$ and $\varepsilon > 0$. We let $\widehat{\varphi^{\varepsilon}}(\,\cdot\,,y)$ be the Fourier transform of $\xi \mapsto \varphi^{\varepsilon}(\xi,y)$ for $y \in \mathbb{R}^n$. From Example 9.10(c) and Remark 9.1(e), it follows that

$$\widehat{\varphi^{\varepsilon}}(x,y) = k_{\varepsilon}(y-x) \quad \text{for } x,y \in \mathbb{R}^n \ .$$

Therefore Proposition 9.11 implies

$$\begin{aligned} (2\pi)^{-n/2} \int \widehat{f}(\xi) e^{i\,\langle y,\xi\rangle} e^{-\varepsilon^2\,|\xi|^2} \, d\xi &= \int \widehat{f}(\xi) \varphi^{\varepsilon}(\xi,y) \, d\xi \\ &= \int f(x) \widehat{\varphi^{\varepsilon}}(x,y) \, dx = k_{\varepsilon} * f(y) \end{aligned}$$

for $y \in \mathbb{R}^n$. The claim now follows from Theorem 7.11 and Example 7.12(a).

(ii) If $\widehat{f}$ belongs to L_1, the dominated convergence theorem shows that

$$\lim_{\varepsilon\to 0}\int e^{i\langle y,\xi\rangle}\widehat{f}(\xi)e^{-\varepsilon^2|\xi|^2}\,d\xi=\int e^{i\langle y,\xi\rangle}\widehat{f}(\xi)\,d\xi=(2\pi)^{n/2}\mathcal{F}\bigl(\widehat{f}\,\bigr)^{\vee}(y)$$

for $y\in\mathbb{R}^n$. Thus (i), Remark 9.1(b), and Theorem 2.18(i) finish the proof. ■

9.13 Corollary

(i) (Fourier integral theorem for $\mathcal{S}$) *The Fourier transformation is a continuous automorphism of $\mathcal{S}$. Its inverse is the Fourier cotransformation.*

(ii) *The Fourier transformation maps L_1 continuously and injectively into C_0 and has a dense image.*

(iii) *For $f\in L_1\cap BUC$, the equality*[4]

$$f(x)=\lim_{\varepsilon\to 0}(2\pi)^{-n/2}\int e^{i\langle x,\xi\rangle}\widehat{f}(\xi)e^{-\varepsilon^2|\xi|^2}\,d\xi$$

holds uniformly with respect to $x\in\mathbb{R}^n$.

(iv) *For $f\in L_1\cap BUC$, suppose $\widehat{f}$ belongs to L_1. Then*

$$f(x)=(2\pi)^{-n/2}\int e^{i\langle x,\xi\rangle}\widehat{f}(\xi)\,d\xi\quad\text{for } x\in\mathbb{R}^n\ .$$

Proof (i) As in the case of normed vector spaces, we denote by $\mathcal{L}(\mathcal{S})$ the vector space of all continuous endomorphisms of $\mathcal{S}$; similarly, we let $\mathcal{L}\mathrm{aut}(\mathcal{S})$ be the automorphisms of $\mathcal{S}$. Then it follows from Remark 9.2(i) and Proposition 9.7 that $\mathcal{F}$ and $\overline{\mathcal{F}}$ belong to $\mathcal{L}(\mathcal{S})$. Because $\mathcal{S}\subset L_1$, Theorem 9.12(ii) therefore shows that $\overline{\mathcal{F}}$ is a left inverse of $\mathcal{F}$ in $\mathcal{L}(\mathcal{S})$. It then follows from $\check{\widehat{u}}=\widehat{\check{u}}$ that $\mathcal{F}\overline{\mathcal{F}}f=\mathcal{F}(\mathcal{F}f)^{\vee}=\check{\mathcal{F}}\mathcal{F}f=\overline{\mathcal{F}}\mathcal{F}f=f$ for $f\in\mathcal{S}$. Therefore $\overline{\mathcal{F}}$ is also a right inverse of $\mathcal{F}$ in $\mathcal{L}(\mathcal{S})$, which proves $\mathcal{F}\in\mathcal{L}\mathrm{aut}(\mathcal{S})$.

(ii) If $\widehat{f}=0$ for $f\in L_1$, then $f=0$ follows from Theorem 9.12(ii). Therefore $\mathcal{F}$ is injective on L_1, and from the Riemann–Lebesgue lemma, we know that $\mathcal{F}$ belongs to $\mathcal{L}(L_1,C_0)$. Because (i) and $\mathcal{S}\subset L_1$, we have $\mathcal{S}=\mathcal{F}(\mathcal{S})\subset\mathcal{F}(L_1)$. It then follows from Remark 9.2(f) that $\mathcal{F}(L_1)$ is dense in C_0.

(iii) follows from the proof of Theorem 9.12(i) and Theorem 7.11.

(iv) is now clear. ■

9.14 Remarks **(a)** For $f\in\mathcal{S}$, we have $\widehat{\widehat{f}}=\check{f}$.

(b) One can show that L_1 does not have a closed image in C_0 under the Fourier transformation (see [Rud83]). Hence $\mathcal{F}\in\mathcal{L}(L_1,C_0)$ is not surjective. ■

[4]One can show that (iii) and (iv) remain true for $f\in L_1\cap C$.

Convolutions and the Fourier transform

We now study what happens to convolutions under the Fourier transformations. So we first introduce another space of smooth functions; these will turn out to be particularly significant in the next subsection.

Suppose $\varphi \in C^\infty$. If to every $\alpha \in \mathbb{N}^n$, there are constants $c_\alpha > 0$ and $k_\alpha \in \mathbb{N}$ such that

$$|\partial^\alpha \varphi(x)| \le c_\alpha (1+|x|^2)^{k_\alpha} \quad \text{for } x \in \mathbb{R}^n \ ,$$

then we say φ is **slowly increasing**. We denote by $\mathcal{O}_M$ the set of all functions with this property, the **space of slowly increasing functions**.

9.15 Remarks **(a)** In the sense of vector subspaces, we have the inclusions $\mathcal{S} \subset \mathcal{O}_M \subset C^\infty$ and $\mathbb{C}[X_1,\ldots,X_n] \subset \mathcal{O}_M$.

(b) $(\mathcal{O}_M, +, \cdot\,)$ is a commutative algebra with unity.

(c) Suppose $(\varphi, f) \in \mathcal{O}_M \times \mathcal{S}$. Then φf belongs to $\mathcal{S}$, and to every $m \in \mathbb{N}$, there are $c = c(\varphi, m) > 0$ and $k' = k'(\varphi, m) \in \mathbb{N}$ such that $q_{k,m}(\varphi f) \le c\, q_{k+k',m}(f)$ for $k \in \mathbb{N}$.

Proof Suppose $m \in \mathbb{N}$. Then there are $c = c(\varphi, m) > 0$ and $k' = k'(\varphi, m) \in \mathbb{N}$ such that

$$|\partial^\alpha \varphi(x)| \le c(1+|x|^2)^{k'/2} \quad \text{for } x \in \mathbb{R}^n \ , \quad \alpha \in \mathbb{N}^n \ , \quad |\alpha| \le m \ .$$

Now it follows from the Leibniz rule that

$$\begin{aligned} q_{k,m}(\varphi f) &= \max_{|\alpha|\le m} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^{k/2} \Big| \sum_{\beta \le \alpha} \binom{\alpha}{\beta} \partial^\beta \varphi(x) \partial^{\alpha-\beta} f(x) \Big| \\ &\le c \max_{|\alpha|\le m} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^{(k+k')/2} \, |\partial^\alpha f(x)| = c\, q_{k+k',m}(f) \end{aligned}$$

for $f \in \mathcal{S}$ and $k \in \mathbb{N}$. ■

(d) Suppose $\varphi \in \mathcal{O}_M$. Then $f \mapsto \varphi f$ is a linear and continuous map of $\mathcal{S}$ into itself.

Proof This follows from (c) and Remark 9.2(c). ■

(e) For every $s \in \mathbb{R}$, Λ^s belongs to $\mathcal{O}_M$. ■

We can now prove another important property of the Fourier transformation.

9.16 Theorem (convolution theorem)

(i) $(f * g)\widehat{\ } = (2\pi)^{n/2} \widehat{f}\,\widehat{g}$ *for* $(f,g) \in \mathcal{L}_1 \times \mathcal{L}_1$.

(ii) $\widehat{\varphi} * \widehat{f} = (2\pi)^{n/2} \widehat{\varphi f}$ *for* $(\varphi, f) \in \mathcal{S} \times \mathcal{L}_1$.

Proof (i) By (9.12) and (9.13), we see that the Fubini–Tonelli theorem can be applied. It then follows from Corollary 7.9 that

$$\begin{aligned}(f*g)\widehat{\ }(\xi) &= (2\pi)^{-n/2}\int e^{-i\langle x,\xi\rangle}\int f(x-y)g(y)\,dy\,dx\\ &= (2\pi)^{-n/2}\int g(y)\int e^{-i\langle x,\xi\rangle}f(x-y)\,dx\,dy\ .\end{aligned}$$

Because

$$\int e^{-i\langle x,\xi\rangle}f(x-y)\,dx = e^{-i\langle y,\xi\rangle}\int e^{-i\langle z,\xi\rangle}f(z)\,dz = e^{-i\langle y,\xi\rangle}(2\pi)^{n/2}\widehat{f}(\xi)\ ,$$

we then get

$$(f*g)\widehat{\ }(\xi) = (2\pi)^{-n/2}\int (2\pi)^{n/2}\widehat{f}(\xi)e^{-i\langle y,\xi\rangle}g(y)\,dy = (2\pi)^{n/2}\widehat{f}(\xi)\widehat{g}(\xi)\ .$$

(ii) Suppose $(\varphi,f)\in\mathcal{S}\times\mathcal{L}_1$. By Theorem 9.3, we find a sequence (f_j) in $\mathcal{S}$ such that $f_j\to f$ in $\mathcal{L}_1$. Propositions 9.4 and 9.7 imply that $\widehat{\varphi}*\widehat{f_j}$ belongs to $\mathcal{S}$. By Remark 9.15(c), φf_j also belongs to $\mathcal{S}$, so it follows from (i) and Remark 9.14(a) that

$$\big(\widehat{\varphi}*\widehat{f_j}\big)\widehat{\ } = (2\pi)^{n/2}\widehat{\widehat{\varphi}}\widehat{\widehat{f_j}} = (2\pi)^{n/2}(\varphi f_j)\check{\ }\quad\text{for } j\in\mathbb{N}\ .$$

By Theorem 9.12(ii), we then get

$$\widehat{\varphi}*\widehat{f_j} = (2\pi)^{n/2}\widehat{\varphi f_j}\quad\text{for } j\in\mathbb{N}\ . \tag{9.14}$$

Because $f_j\to f$ in L_1, it follows from Remark 9.1(a) that $\widehat{f_j}\to\widehat{f}$ in BC. Therefore Corollary 7.9 implies, because $\widehat{\varphi}\in\mathcal{S}\subset L_1$, that the sequence $(\widehat{\varphi}*\widehat{f_j})$ converges in BC to $\widehat{\varphi}*\widehat{f}$. Because $\varphi f_j\to\varphi f$ clearly holds in L_1, we deduce from Proposition 9.9 that the sequence $(\widehat{\varphi f_j})$ converges in BC to $\widehat{\varphi f}$. Then the claim follows from Remark 9.1(a). ■

As an application of the convolution theorem, we prove a lemma, which forms the basis for the L_2-theory of the Fourier transformation.

9.17 Lemma *For $f\in\mathcal{L}_1\cap\mathcal{L}_2$, $\widehat{f}$ belongs to $C_0\cap\mathcal{L}_2$, and $\|f\|_2=\|\widehat{f}\|_2$.*

Proof Suppose $f\in\mathcal{L}_1\cap\mathcal{L}_2$. Because $\widehat{f}$ belongs to C_0 by the Riemann–Lebesgue lemma, it suffices to verify $\|f\|_2=\|\widehat{f}\|_2$. So we set $g:=f*\overline{\overline{f}}$. By Theorem 7.3(ii) and Exercise 7.2, we know g belongs to $\mathcal{L}_1\cap C_0$, and

$$g(0)=\int f(y)\overline{\overline{f}}(0-y)\,dy=\int f\overline{f}=\|f\|_2^2\ .$$

From Corollary 9.13(iii), it follows that

$$\|f\|_2^2 = g(0) = \lim_{\varepsilon\to 0}(2\pi)^{-n/2}\int \widehat{g}(\xi)e^{-\varepsilon^2\,|\xi|^2}\,d\xi\ . \tag{9.15}$$

Now we note

$$\widehat{\overline{\widehat{f}}} = (2\pi)^{-n/2}\int e^{-i\langle x,\xi\rangle}\overline{f}(-x)\,dx = (2\pi)^{-n/2}\overline{\int e^{-i\langle -x,\xi\rangle}f(-x)\,dx} = \overline{\widehat{f}}\ ,$$

which follows from the Euclidean invariance of integrals. Then Theorem 9.16(i) shows

$$\widehat{g} = \bigl(f * \overline{\overline{f}}\,\bigr)^{\widehat{}} = (2\pi)^{n/2}\widehat{f}\,\widehat{\overline{\overline{f}}} = (2\pi)^{n/2}\,\bigl|\widehat{f}\,\bigr|^2\ .$$

In particular $\widehat{g}$, is not negative. Therefore (9.15) and monotone convergence theorem imply $\|f\|_2 = \|\widehat{f}\,\|_2$. ■

Fourier multiplication operators

To illustrate the significance of the mapping properties of the Fourier transformation, we now consider linear differential operators with constant coefficients and show that they are represented "in the Fourier domain" by multiplication operators.

For $m \in \mathbb{N}$, we denote by $\mathbb{C}_m[X_1,\ldots,X_n]$ the vector subspace $\mathbb{C}[X_1,\ldots,X_n]$ consisting of all polynomials of degree $\leq m$. For

$$p = \sum_{|\alpha|\leq m} a_\alpha X^\alpha \in \mathbb{C}_m[X_1,\ldots,X_n]\ ,$$

we let

$$p(D) := \sum_{|\alpha|\leq m} a_\alpha D^\alpha\ ,$$

which is a linear differential operator of order $\leq m$ with *constant coefficients*. Here p is called the **symbol** of $p(D)$. In the following, we set

$$\mathcal{D}\mathrm{iffop}^0 := \bigl\{\, p(D)\ ;\ p \in \mathbb{C}[X_1,\ldots,X_n]\,\bigr\}\ ,$$

and $\mathcal{D}\mathrm{iffop}_m^0$ is the subset of all constant-coefficient, linear differential operators of order not higher than m.

9.18 Remarks **(a)** $p(D) \in \mathcal{D}\mathrm{iffop}^0$ is a linear and continuous map of $\mathcal{S}$ into itself, that is, $p(D) \in \mathcal{L}(\mathcal{S})$.

Proof This follows from Remarks 9.2(c) and (h). ■

(b) The map

$$\mathbb{C}[X_1,\ldots,X_n] \to \mathcal{L}(\mathcal{S})\ ,\quad p \mapsto p(D) \tag{9.16}$$

is linear and injective.

Proof The linearity is obvious. Suppose $p = \sum_{|\alpha|\le m} a_\alpha X^\alpha \in \mathbb{C}[X_1,\dots,X_n]$ and that $p(D)f = 0$ for all $f \in \mathcal{S}$. We choose a $\varphi \in \mathcal{D}$ such that $\varphi \,|\, \mathbb{B}^n = 1$. For $\beta \in \mathbb{N}^n$, it follows from the Leibniz rule that

$$D^\alpha(\varphi X^\beta) = \varphi D^\alpha X^\beta + \sum_{\gamma<\alpha} \binom{\alpha}{\gamma} D^{\alpha-\gamma}\varphi D^\gamma X^\beta \ .$$

Because $\varphi(x) = 1$ for $|x| < 1$, we then derive

$$D^\alpha(\varphi X^\beta)(0) = D^\alpha X^\beta(0) = \begin{cases} \beta! & \text{if } \alpha = \beta \ , \\ 0 & \text{otherwise} \end{cases}$$

for $\alpha \in \mathbb{N}^n$. Because $\varphi X^\beta \in \mathcal{D} \subset \mathcal{S}$, we thus find $0 = p(D)(\varphi X^\beta) = \beta!\, a_\beta$ for $\beta \in \mathbb{N}^n$ with $|\beta| \le m$; therefore $p = 0$. This proves the claimed injectivity. ■

(c) $p(D)$ is formally self-adjoint if and only if p has real coefficients.

Proof Letting

$$\mathcal{A}(\partial) := p(D) = \sum_{|\alpha|\le m} a_\alpha (-i)^{|\alpha|}\, \partial^\alpha$$

it follows from Proposition 7.24 that

$$\mathcal{A}^\sharp(\partial) = \sum_{|\alpha|\le m} (-1)^{|\alpha|} \overline{a_\alpha (-i)^{|\alpha|}}\, \partial^\alpha = \sum_{|\alpha|\le m} (-i)^{|\alpha|} \overline{a}_\alpha\, \partial^\alpha \ ,$$

which finishes the proof. ■

By Remark 9.18(b), we can identify $\mathcal{D}\mathrm{iffop}^0$ [or $\mathcal{D}\mathrm{iffop}^0_m$] with the image of $\mathbb{C}[X_1,\dots,X_n]$ [or $\mathbb{C}_m[X_1,\dots,X_n]$] under the map (9.16). In other words, in the sense of vector subspaces, we have

$$\mathcal{D}\mathrm{iffop}^0_m \subset \mathcal{D}\mathrm{iffop}^0 \subset \mathcal{L}(\mathcal{S}) \quad \text{for } m \in \mathbb{N} \ .$$

For $a \in \mathcal{O}_M$ and $f \in \mathcal{S}$, it follows from Corollary 9.13(i) and Remark 9.15(d) that $\big(f \mapsto a\widehat{f}\big) \in \mathcal{L}(\mathcal{S})$. Then it follows again from Corollary 9.13(i) that

$$a(D) := \mathcal{F}^{-1} a \mathcal{F} : \mathcal{S} \to \mathcal{S} \ , \quad f \mapsto \mathcal{F}^{-1}\big(a\widehat{f}\big)$$

is a well defined element of $\mathcal{L}(\mathcal{S})$, a **Fourier multiplication operator** with **symbol** a. We set

$$\mathrm{Op} := \big\{ a(D) \in \mathcal{L}(\mathcal{S}) \ ; \ a \in \mathcal{O}_M \big\} \ .$$

9.19 Proposition Op *is a commutative algebra of* $\mathcal{L}(\mathcal{S})$ *with unity, and the map*

$$ev\colon (\mathcal{O}_M, +, \cdot\,) \to \mathrm{Op}\ , \quad a \mapsto a(D)$$

is an algebra isomorphism.

Proof It is clear that $\mathcal{O}_M := (\mathcal{O}_M, +, \cdot\,)$ is a commutative subalgebra with unity of the algebra $\mathbb{C}^{(\mathbb{R}^n)}$. It is also easy to verify that ev maps the vector space $\mathcal{O}_M$ linearly in $\mathcal{L}(\mathcal{S})$.

For $a, b \in \mathcal{O}_M$ and $f \in \mathcal{S}$, we have

$$(ab)(D)f = \mathcal{F}^{-1}\bigl(ab\widehat{f}\,\bigr) = \mathcal{F}^{-1}\bigl(a\mathcal{F}\mathcal{F}^{-1}\bigl(b\widehat{f}\,\bigr)\bigr) = \mathcal{F}^{-1}\bigl(a\widehat{b(D)f}\bigr) = a(D) \circ b(D)f\ .$$

Therefore ev is a surjective algebra homomorphism.

Finally let $a, b \in \mathcal{O}_M$ with $a(D) = b(D)$. Let $\xi \in \mathbb{R}^n$, and denote by $\varphi \in \mathcal{D}$ a cutoff function for $\overline{\mathbb{B}}^n(\xi, 1)$. Then $f := \mathcal{F}^{-1}\varphi$ belongs to $\mathcal{S}$ with $\widehat{f}(\xi) = 1$, and it follows from Corollary 9.13(i) that

$$a(\xi) = \bigl(a\widehat{f}\,\bigr)(\xi) = \mathcal{F}\bigl(a(D)f\bigr)(\xi) = \mathcal{F}\bigl(b(D)f\bigr)(\xi) = \bigl(b\widehat{f}\,\bigr)(\xi) = b(\xi)\ .$$

Because this is true for every $\xi \in \mathbb{R}^n$, we have $a = b$. Therefore ev is injective. ■

9.20 Corollary

(i) *For* $a, b \in \mathcal{O}_M$*, we have* $ab(D) = a(D)b(D) = b(D)a(D)$.

(ii) $\mathbf{1}(D) = 1_{\mathcal{L}(\mathcal{S})}$.

(iii) $\mathcal{D}\mathrm{iffop}^0$ *is the image of* $\mathbb{C}[X_1, \ldots, X_n]$ *under ev. In particular,* $\mathcal{D}\mathrm{iffop}^0$ *is a commutative subalgebra of* Op *with unity.*

Proof (i) and (ii) are special cases of Proposition 9.19.

(iii) For $p \in \mathbb{C}[X_1, \ldots, X_n] \subset \mathcal{O}_M$ with $p = \sum_{|\alpha| \le m} a_\alpha X^\alpha$, we get from Proposition 9.6(i) that

$$\begin{aligned} ev(p)f = \mathcal{F}^{-1}p\mathcal{F}f = \mathcal{F}^{-1}\bigl(p\widehat{f}\,\bigr) &= \sum_{|\alpha| \le m} a_\alpha \mathcal{F}^{-1}\bigl(X^\alpha \widehat{f}\,\bigr) \\ &= \sum_{|\alpha| \le m} a_\alpha \mathcal{F}^{-1}\bigl(\widehat{D^\alpha f}\bigr) = \sum_{|\alpha| \le m} a_\alpha D^\alpha f \end{aligned}$$

for $f \in \mathcal{S}$. Therefore $ev(p) = \sum_{|\alpha| \le m} a_\alpha D^\alpha$, from which the claim follows. ■

This corollary implies that the Fourier transformation can be used to solve linear differential equations with constant coefficients by reducing them to simple algebraic equations. This fact is part of the fundamental significance of the Fourier transformation. The following examples give a first glimpse into these methods.

9.21 Examples **(a)** Suppose the polynomial $p \in \mathbb{C}[X_1, \ldots, X_n]$ has no real zeros. Then $p(D) \in \mathcal{L}(\mathcal{S})$ is an automorphism of $\mathcal{S}$, and $[p(D)]^{-1} = (1/p)(D)$.

Proof We see easily that $1/p$ belongs to $\mathcal{O}_M$. Now we deduce from Corollary 9.20 that

$$1_{\mathcal{L}(\mathcal{S})} = \mathbf{1}(D) = (p \cdot 1/p)(D) = p(D)(1/p)(D) = (1/p)(D)p(D)\ .$$

Because $a(D) \in \mathcal{L}(\mathcal{S})$ for $a \in \mathcal{O}_M$, this proves the claim. ■

(b) $1 - \Delta \in \mathcal{L}\mathrm{aut}(\mathcal{S})$, and $(1 - \Delta)^{-1} = \Lambda^{-2}(D)$.

Proof Because $1 - \Delta = \Lambda^2(D)$, this follows from (a). ■

Example 9.21(b) says that the partial differential equation

$$-\Delta u + u = f \tag{9.17}$$

has a unique solution $u \in \mathcal{S}$ for every $f \in \mathcal{S}$ and that u depends continuously on f in the topology of $\mathcal{S}$. Also we can obtain the solution $u \in \mathcal{S}$ of (9.17) by first "Fourier transforming" this equation. This, according to Proposition 9.6, gives the equation $(|\xi|^2 + 1)\widehat{u}(\xi) = \Lambda^2(\xi)\widehat{u}(\xi) = \widehat{f}(\xi)$ for $\xi \in \mathbb{R}^n$. This equation can then be solved for $\widehat{u}$, giving $\widehat{u} = \Lambda^{-2}\widehat{f}$, and then "reverse Fourier transformed", giving $u = \mathcal{F}^{-1}\bigl(\Lambda^{-2}\widehat{f}\bigr) = \Lambda^{-2}(D)f$. This "method of Fourier transformation" plays a prominent role in the theory of partial differential equations. Note that $\Lambda^{-2}(D)$ or, more generally, $(1/p)(D)$, is not a differential operator.

Plancherel's theorem

To conclude this chapter, we show that the Fourier transformation can also be defined on L_2, and we explain a few consequences of this fact.

Suppose H is a Hilbert space. We say $T : H \to H$ is **unitary** if T is an isometric isomorphism.

9.22 Remarks Suppose H is a (real or complex) Hilbert space and $T : H \to H$ is linear.

(a) If T is unitary, then T belongs to $\mathcal{L}\mathrm{aut}(H)$, and

$$(Tx \,|\, Ty) = (x \,|\, y) \quad \text{for } x, y \in H\ .$$

Proof The first statement is clear. Because T is an isometry, we have

$$4\,\mathrm{Re}(Tx \,|\, Ty) = \|T(x+y)\|^2 - \|T(x-y)\|^2 = \|x+y\|^2 - \|x-y\|^2 = 4\,\mathrm{Re}(x \,|\, y)\ ,$$

and therefore $\mathrm{Re}(Tx \,|\, Ty) = \mathrm{Re}(x \,|\, y)$ for $x, y \in H$. Replacing y in this identity by iy, we get

$$\mathrm{Im}(Tx \,|\, Ty) = \mathrm{Re}(Tx \,|\, Tiy) = \mathrm{Re}(x \,|\, iy) = \mathrm{Im}(x \,|\, y)\ ,$$

and thus $(Tx \,|\, Ty) = (x \,|\, y)$ for $x, y \in H$. ■

(b) If H is finite-dimensional, then the following statements are equivalent:

(i) T is unitary.

(ii) $(Tx \,|\, Ty) = (x \,|\, y)$ for $x, y \in H$.

(iii) $T^*T = \mathrm{id}_H$.

Proof "(i)⇒(ii)" is a consequence of (a).

"(ii)⇒(iii)" Let $\{b_1, \ldots, b_m\}$ be an orthonormal basis of H. Then every $y \in H$ can be expanded as $y = \sum_{j=1}^m (y \,|\, b_j) b_j$ (see Exercise II.3.12 and Theorem VI.7.14). From Exercise VII.1.5 and (ii), it follows that

$$T^*Tx = \sum_{j=1}^m (T^*Tx \,|\, b_j) b_j = \sum_{j=1}^m (Tx \,|\, Tb_j) b_j = \sum_{j=1}^m (x \,|\, b_j) b_j = x$$

for every $x \in H$.

"(iii)⇒(i)" Because $T^*T = \mathrm{id}_H$, we know T is injective and is therefore also surjective by the rank formula of linear algebra. For $x \in H$, we also have

$$\|Tx\|^2 = (Tx \,|\, Tx) = (T^*Tx \,|\, x) = (x \,|\, x) = \|x\|^2 \ .$$

Therefore T is an isometry. ■

9.23 Theorem (Plancherel) *The Fourier transformation has a unique extension from $L_1 \cap L_2$ to a unitary operator on L_2.*

Proof Denote by X_2 the vector subspace $L_1 \cap L_2$ of the Hilbert space L_2. Then it follows from Lemma 9.17 that $\mathcal{F}$ belongs to $\mathcal{L}(X_2, L_2)$ and is an isometry. Because X_2 contains the space $\mathcal{S}$, Theorem 9.3 and VI.2.6 imply the existence of a unique isometric extension $\mathfrak{F} \in \mathcal{L}(L_2)$. As an isometry, $\mathfrak{F}$ has a closed image, which by Corollary 9.13(i) contains the space $\mathcal{S}$. Therefore Proposition V.4.4 implies that $\mathfrak{F}$ is surjective and therefore unitary. ■

As usual, we reuse the symbol $\mathcal{F}$ for the unique continuous extension of $\mathfrak{F}$ and likewise call it the **Fourier transformation**.[5]

The next proposition describes the Fourier transform $\mathcal{F}f$ for an arbitrary $f \in L_2$.

9.24 Proposition *For $f \in L_2$, we have*

$$\mathcal{F}f = \lim_{R\to\infty} \mathcal{F}(\chi_{R\mathbb{B}^n} f) = \lim_{R\to\infty} (2\pi)^{-n/2} \int_{[\,|x|\le R]} e^{-i\,\langle x,\cdot\,\rangle} f(x)\, dx \qquad \text{in } L_2\ .$$

[5] On L_2, the Fourier transformation will sometimes also be called the Fourier–Plancherel, or Plancherel, transformation.

Proof For $R>0$, the element $f_R := \chi_{R\bar{\mathbb{B}}^n} f$ belongs to $L_1 \cap L_2$, and the dominated convergence theorem implies

$$\int |f - f_R|^2 \, dx = \int |f|^2 \, (1 - \chi_{R\bar{\mathbb{B}}^n})^2 \, dx \to 0 \quad (R \to \infty) \; .$$

Therefore $\lim_{R\to\infty} f_R = f$ in L_2. Then by Plancherel's theorem, $\mathcal{F} f_R$ converges in L_2 to $\mathcal{F} f$. Because

$$\mathcal{F}(f_R)(\xi) = (2\pi)^{-n/2} \int_{[\,|x|\le R\,]} e^{-i\,\langle x,\xi\rangle} f(x)\, dx \quad \text{for } \xi \in \mathbb{R}^n \; ,$$

the claim follows. ■

9.25 Example Suppose $n=1$ and $a>0$. Also let $f := \chi_{[-a,a]} \in \mathcal{L}_1(\mathbb{R})$. Then

$$\widehat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} e^{-i x \xi} \, dx = \frac{-1}{\sqrt{2\pi}\; i\xi} \big(e^{-i\xi a} - e^{i\xi a}\big) = \sqrt{\frac{2}{\pi}}\; a\, \frac{\sin(a\xi)}{a\xi}$$

for $\xi \in \mathbb{R}$. Because $\int |f|^2 \, dx = 2a$, Plancherel's theorem gives

$$\int_{-\infty}^{\infty} \Big[\frac{\sin(ax)}{ax}\Big]^2 \, dx = \frac{\pi}{a} \quad \text{for } a > 0 \; .$$

Note that $x \mapsto \sin(x)/x$ does not belong to $\mathcal{L}_1(\mathbb{R})$. ■

Symmetric operators

Suppose E is a Banach space over $\mathbb{K}$. By a **linear operator A in E**, we mean a map $A \colon \mathrm{dom}(A) \subset E \to E$ such that $\mathrm{dom}(A)$ is a vector subspace of E and such that A is linear. For linear operators $A_j : \mathrm{dom}(A_j) \subset E \to E$ and $\lambda \in \mathbb{K}^{\times}$, we define $A_0 + \lambda A_1$ by

$$\mathrm{dom}(A_0 + \lambda A_1) := \mathrm{dom}(A_0) \cap \mathrm{dom}(A_1) \; , \quad (A_0 + \lambda A_1)x := A_0 x + \lambda A_1 x \; .$$

The **product** $A_0 A_1$ is defined by

$$\mathrm{dom}(A_0 A_1) := \big\{ x \in \mathrm{dom}(A_1) \; ; \; A_1 x \in \mathrm{dom}(A_0) \big\} \; , \quad (A_0 A_1)x := A_0(A_1 x) \; .$$

Finally, the operator defined by

$$\mathrm{dom}\big([A_0, A_1]\big) := \mathrm{dom}(A_0 A_1 - A_1 A_0) \; , \quad [A_0, A_1]x := (A_0 A_1 - A_1 A_0)x$$

is called the **commutator** of A_0 and A_1. Obviously $A_0 + \lambda A_1$, $A_0 A_1$, and $[A_0, A_1]$ are linear operators in E, for which

$$A_0 + \lambda A_1 = \lambda A_1 + A_0 \; , \quad \lambda A_0 = A_0(\lambda \mathrm{id}_E) \; , \quad [A_0, A_1] = -[A_1, A_0] \; .$$

Suppose now H is a Hilbert space, and $A\colon \mathrm{dom}(A) \subset H \to H$ is a linear operator on H. If

$$(Au \,|\, v) = (u \,|\, Av) \quad \text{for } u, v \in \mathrm{dom}(A) \ ,$$

we say A is **symmetric**.

9.26 Remarks **(a)** Suppose H is a complex Hilbert space and A is a linear operator on H. Then these statements are equivalent:

(i) A is symmetric.

(ii) $(Au \,|\, u) \in \mathbb{R}$ for $u \in \mathrm{dom}(A)$.

Proof "(i)⇒(ii)" Because A is symmetric, it follows that

$$(Au \,|\, u) = (u \,|\, Au) = \overline{(Au \,|\, u)} \quad \text{for } u \in \mathrm{dom}(A) \ ,$$

and therefore $\mathrm{Im}(Au \,|\, u) = 0$.

"(ii)⇒(i)" For $u, v \in \mathrm{dom}(A)$, we have

$$\bigl(A(u+v) \bigm| u+v\bigr) = (Au \,|\, u) + (Av \,|\, u) + (Au \,|\, v) + (Av \,|\, v) \ . \tag{9.18}$$

Because of (ii), it follows that $\mathrm{Im}(Au \,|\, v) = -\,\mathrm{Im}(Av \,|\, u)$, and therefore

$$\mathrm{Im}(Au \,|\, v) = -\,\mathrm{Im}(Av \,|\, u) = -\,\mathrm{Im}\,\overline{(u \,|\, Av)} = \mathrm{Im}(u \,|\, Av) \ .$$

Replacing u in (9.18) by iu, we get

$$\mathrm{Re}(Au \,|\, v) = \mathrm{Im}\bigl(A(iu) \bigm| v\bigr) = \mathrm{Im}(iu \,|\, Av) = \mathrm{Re}(u \,|\, Av) \ .$$

Therefore $(Au \,|\, v) = (v \,|\, Au)$. ∎

(b) Suppose $p \in \mathbb{C}[X_1, \ldots, X_n]$ and P is the linear operator on L_2 such that $\mathrm{dom}(P) = \mathcal{S}$ and $Pu := p(D)u$ for $u \in \mathcal{S}$. Then these statements are equivalent:

(i) P is symmetric.

(ii) $p(D)$ is formally self-adjoint.

(iii) p has real coefficients.

Proof "(i)⇒(ii)" That P is symmetric implies

$$\bigl(p(D)u \bigm| v\bigr) = (Pu \,|\, v) = (u \,|\, Pv) = \bigl(u \bigm| p(D)v\bigr) \quad \text{for } u, v \in \mathcal{D} \ ,$$

which in turn implies (ii) by the uniqueness of formally adjoint operators.

"(ii)⇒(iii)" Remark 9.18(c).

"(iii)⇒(i)" Suppose $p = \sum_{|\alpha| \le m} a_\alpha X^\alpha$. Then Corollary 9.20(iii) and Plancherel's theorem imply

$$(Pu \,|\, u) = \bigl(p(D)u \bigm| u\bigr) = (p\widehat{u} \,|\, \widehat{u}) = \sum_{|\alpha| \le m} a_\alpha \int \xi^\alpha \, |\widehat{u}(\xi)|^2 \, d\xi \quad \text{for } u \in \mathcal{S} \ .$$

Therefore $(Pu \,|\, u)$ is real, and the claim follows from (a). ∎

(c) With $\mathcal{S}$ as their domain, the Laplace, wave, and Schrödinger operators are symmetric in L_2.

Proof This follows from (b) and Examples 7.25(a) and (e). ∎

The Heisenberg uncertainty relation

As another application of Plancherel's theorem, we close this section by discussing several important properties of the position and momentum operators of quantum mechanics. So we fix $j \in \{1, \dots, n\}$ and set

$$\operatorname{dom}(A_j) := \{\, u \in L_2 \ ;\ X_j \widehat{u} \in L_2 \,\} \ , \quad \operatorname{dom}(B_j) := \{\, u \in L_2 \ ;\ X_j u \in L_2 \,\} \ .$$

Then we define linear operators in L_2, the **momentum operator** A_j and the **position operator** B_j (for the j-th coordinate), by

$$A_j u := \mathcal{F}^{-1}(X_j \widehat{u}) \quad \text{and} \quad B_j v := X_j v \quad \text{for } u \in \operatorname{dom}(A_j) \ , \quad v \in \operatorname{dom}(B_j) \ .$$

9.27 Remarks **(a)** We have $\mathcal{S} \subset \operatorname{dom}(A_j)$, and

$$A_j u = X_j(D) u = D_j u = -i \partial_j u \quad \text{for } u \in \mathcal{S} \ .$$

Proof This follows from Proposition 9.7 and Corollary 9.8. ■

(b) We have $\mathcal{F}\bigl(\operatorname{dom}(A_j)\bigr) = \operatorname{dom}(B_j)$ and a commutative diagram:

$$\begin{array}{ccc} \operatorname{dom}(A_j) & \xrightarrow{\ A_j\ } & L_2 \\ {\scriptstyle \mathcal{F}}\downarrow & & \downarrow{\scriptstyle \mathcal{F}} \\ \operatorname{dom}(B_j) & \xrightarrow{\ B_j\ } & L_2 \end{array}$$

In particular,

$$A_j u = \mathcal{F}^{-1} B_j \mathcal{F} u \ , \quad u \in \operatorname{dom}(A_j) \quad \text{and} \quad B_j u = \mathcal{F} A_j \mathcal{F}^{-1} u \ , \quad u \in \operatorname{dom}(B_j) \ .$$

Proof These are consequences of Plancherel's theorem. ■

(c) The position and momentum operators of quantum mechanics are symmetric.

Proof Let $u \in \operatorname{dom}(A_j)$. Then (b) and Plancherel's theorem imply

$$(A_j u \,|\, u) = (\mathcal{F}^{-1} B_j \mathcal{F} u \,|\, u) = (B_j \widehat{u} \,|\, \widehat{u}) = \int \xi_j \, |\widehat{u}(\xi)|^2 \, d\xi \ ,$$

Now the claim follows from Remark 9.26(a). ■

(d) For $u \in \operatorname{dom}\bigl([A_j, B_j]\bigr)$, we have $\bigl([A_j, B_j] u \bigm| u\bigr) = 2i \operatorname{Im}(A_j B_j u \,|\, u)$.

Proof By (b), (c), and Plancherel's theorem, we get for $u \in \operatorname{dom}\bigl([A_j, B_j]\bigr)$ that

$$\begin{aligned} \bigl([A_j, B_j] u \bigm| u\bigr) &= (A_j B_j u - B_j A_j u \,|\, u) \\ &= (\mathcal{F}^{-1} B_j \mathcal{F} B_j u - B_j \mathcal{F}^{-1} B_j \mathcal{F} u \,|\, u) \\ &= (\mathcal{F} B_j u \,|\, B_j \mathcal{F} u) - (B_j \mathcal{F} u \,|\, \mathcal{F} B_j u) \\ &= 2i \operatorname{Im}(\mathcal{F} B_j u \,|\, B_j \mathcal{F} u) = 2i \operatorname{Im}(\mathcal{F}^{-1} B_j \mathcal{F} B_j u \,|\, u) \\ &= 2i \operatorname{Im}(A_j B_j u \,|\, u) \ . \quad ■ \end{aligned}$$

(e) The operator $i\,[A_j, B_j]$ is symmetric in L_2.

Proof This follows from (d). ■

(f) We have $\mathcal{S} \subset \operatorname{dom}\bigl([A_j, B_j]\bigr)$, and $[A_j, B_j]u = -i\,u$ on $u \in \mathcal{S}$.

Proof The first statement follows easily from Proposition 9.7 and Remark 9.2(h). Also (a) shows that

$$[A_j, B_j]u = D_j(X_j u) - X_j D_j u = (D_j X_j)u = -i\,u$$

for $u \in \mathcal{S}$. ■

(g) (Heisenberg uncertainty relation for $\mathcal{S}$) For $j \in \{1, \ldots, n\}$, we have

$$\|u\|_2^2 \le 2\,\|\partial_j u\|_2\,\|X_j u\|_2 \quad \text{for } u \in \mathcal{S}\ .$$

Proof Let $u \in \mathcal{S}$. By (d) and (f), we have

$$-i\,\|u\|_2^2 = -i\,(u \,|\, u) = \bigl([A_j, B_j]u \bigm| u\bigr) = 2i\,\operatorname{Im}(A_j B_j u \,|\, u)\ .$$

The Cauchy–Schwarz inequality therefore gives

$$\|u\|_2^2 = 2\,\bigl|\operatorname{Im}(A_j B_j u \,|\, u)\bigr| \le 2\,\bigl|(A_j B_j u \,|\, u)\bigr| = 2\,\bigl|(B_j u \,|\, A_j u)\bigr| \le 2\,\|A_j u\|_2\,\|B_j u\|_2\ ,$$

and thus the claim follows because of (a). ■

We conclude this section by extending the validity of the Heisenberg uncertainty relation on $\mathcal{S}$ to $\operatorname{dom}(A_j) \cap \operatorname{dom}(B_j)$. We first need a lemma.

9.28 Lemma *For every $u \in \operatorname{dom}(A_j) \cap \operatorname{dom}(B_j)$, there is a sequence (u_m) in $\mathcal{S}$ such that*

$$\lim_{m\to\infty}(u_m, A_j u_m, B_j u_m) = (u, A_j u, B_j u) \qquad \text{in } L_2^3\ .$$

Proof (i) Suppose $u \in \operatorname{dom}(A_j) \cap \operatorname{dom}(B_j)$, and let $\{\, k_\varepsilon \ ;\ \varepsilon > 0 \,\}$ be the Gaussian approximation kernel. We set $u^\varepsilon := k_\varepsilon * u$. By Exercise 8(iv), u^ε belongs to $\mathcal{S}$, and Theorem 7.11 shows $\lim_{\varepsilon\to 0} u^\varepsilon = u$ in L_2.

(ii) Because $\check{k} = k$, it follows from Example 9.10(c) that

$$\widehat{k}_\varepsilon(\xi) := \widehat{\check{k}}_\varepsilon(\xi) = \mathcal{F}^{-1}k_\varepsilon(\xi) = \varphi(\varepsilon\xi) = (2\pi)^{-n/2}e^{-\varepsilon^2\,|\xi|^2} \quad \text{for } \xi \in \mathbb{R}^n\ .$$

According to Theorem 9.3, we can find a sequence (v_m) in $\mathcal{S}$ such that $\lim_m v_m = u$ in L_2. The convolution theorem therefore shows

$$(k_\varepsilon * v_m)\widehat{\ }(\xi) = (2\pi)^{n/2}\,\widehat{k}_\varepsilon(\xi)\widehat{v}_m(\xi) = e^{-\varepsilon^2\,|\xi|^2}\,\widehat{v}_m(\xi) \quad \text{for } \xi \in \mathbb{R}^n\ .$$

The limit $m \to \infty$ then gives $\widehat{u^\varepsilon} = e^{-\varepsilon^2\,|\cdot|^2}\,\widehat{u}$ (see Corollary 7.9 and Theorem 9.23). Because

$$\|A_j u - A_j u^\varepsilon\|_2^2 = \|X_j\widehat{u} - X_j\widehat{u^\varepsilon}\|_2^2 = \int |\xi_j\widehat{u}(\xi)|^2\bigl(1 - e^{-\varepsilon^2\,|\xi|^2}\bigr)^2\,d\xi\ ,$$

it follows from the dominated convergence theorem that $\lim_{\varepsilon\to 0} A_j u^\varepsilon = A_j u$ in L_2.

(iii) Let $\overset{*}{u}$ be a representative of u. We set

$$d_\varepsilon(x,z) := x_j\big(\overset{*}{u}(x) - \overset{*}{u}(x-\varepsilon z)\big) \ , \quad g_\varepsilon(x,z) := d_\varepsilon(x,z)k(z)$$

for $\varepsilon > 0$ and $(x,z) \in \mathbb{R}^n \times \mathbb{R}^n$. Then it follows, as in (7.7) (or from the Minkowski inequality for integrals), that

$$\|X_j u - X_j u^\varepsilon\|_2 \le \Big(\int \Big[\int |g_\varepsilon(x,z)|\, dz\Big]^2 dx\Big)^{1/2} \le \int \|d_\varepsilon(\,\cdot\,,z)\|_2\, k(z)\, dz \ , \quad (9.19)$$

where for the last inequality we used $g_\varepsilon = \big(d_\varepsilon\sqrt{k}\,\big)\sqrt{k}$ and $\int k\, dx = 1$ together with the Cauchy–Schwarz inequality. Further noting

$$d_\varepsilon(\,\cdot\,,z) = X_j\overset{*}{u} - \tau_{\varepsilon z}(X_j\overset{*}{u}) - \varepsilon z_j \tau_{\varepsilon z}\overset{*}{u} \ ,$$

it follows from the strong continuity of the translation group on L_2 and the translation invariance of integrals that

$$\lim_{\varepsilon\to 0} \|d_\varepsilon(\,\cdot\,,z)\|_2\, k(z) = 0 \quad \text{for } z \in \mathbb{R}^n \ ,$$

and

$$\|d_\varepsilon(\,\cdot\,,z)\|_2\, k(z) \le 2\max\big\{\|X_j u\|_2, \|u\|_2\big\}(1+|z_j|)k(z) \quad \text{for } \varepsilon \in (0,2] \ , \quad z \in \mathbb{R}^n \ .$$

Because $z \mapsto (1+|z_j|)k(z)$ belongs to $\mathcal{L}_1$, the claim is implied by (9.19) and the dominated convergence theorem. ■

9.29 Corollary (Heisenberg uncertainty relation) *For $1 \le j \le n$, we have*

$$\|u\|_2^2 \le 2\, \|A_j u\|_2\, \|B_j u\|_2 \quad \textit{for } u \in \operatorname{dom}(A_j) \cap \operatorname{dom}(B_j) \ .$$

Proof This follows from Remarks 9.27(a) and (g) and Lemma 9.28. ■

From Remark 9.27(a) and Lemma 9.28 it easily follows, as in the proof of Theorem 7.27, that the distributional derivative $\partial_j u$ belongs to L_2 for $u \in \operatorname{dom}(A_j)$ and is therefore a weak L_2-derivative. Also $A_j u = -i\,\partial_j u$. Consequently, we can also write the Heisenberg uncertainty relation for $u \in \operatorname{dom}(A_j) \cap \operatorname{dom}(B_j)$ in the form

$$\Big(\frac{1}{2}\int |u|^2\, dx\Big)^2 \le \int |\partial_j u|^2\, dx \int |X_j u|^2\, dx$$

if we interpret $\partial_j u$ in the weak sense. The significance of this broadened interpretation of the operators A_j and B_j is clarified in the theory of unbounded self-adjoint operators on Hilbert spaces, as developed in functional analysis. Self-adjoint operators built from the position and momentum operators, in particular the Schrödinger operators, are used in the mathematical construction of quantum mechanics (for example [RS72]). For an interpretation of the Heisenberg uncertainty relation, we refer you to the physics literature.

Exercises

1 Let $a > 0$. Determine the Fourier transform of

$$\text{(i) } \sin(ax)/x\ , \qquad \text{(ii) } 1/(a^2 + x^2)\ , \qquad \text{(iii) } e^{-a\,|x|}\ ,$$
$$\text{(iv) } (1 - |x|/a)\chi_{[-a,a]}(x)\ , \qquad \text{(v) } \bigl(\sin(ax)/x\bigr)^2\ .$$

(Hint: See Section VIII.6.)

2 Let $f(x) := (\sin(x)/x)^2$ and $g(x) := e^{2i\,x} f(x)$ for $x \in \mathbb{R}^\times$. Then show $f * g = 0$. (Hint: Apply Exercise 1 and Theorem 9.16.)

3 Show that if $f \in \mathcal{L}_1$ satisfies either $f * f = f$ or $f * f = 0$, then $f = 0$.

4 Let $\{\, \varphi_\varepsilon\ ;\ \varepsilon > 0 \,\}$ be an approximation to the identity, and let (ε_j) be a null sequence. Show that $(\mathcal{F}(\varphi_{\varepsilon_j}))$ converges in $\mathcal{D}'(\mathbb{R}^n)$ to $(2\pi)^{-n/2}\mathbf{1}$.

5 For $a, f \in \mathcal{S}$, show $a(D)f = \widehat{a} * f$.

6 For $s \geq 0$, define $H^s := \{\, u \in L_2\ ;\ \Lambda^s \widehat{u} \in L_2 \,\}$ and $(u \,|\, v)_{H^s} := (\Lambda^s \widehat{u} \,|\, \widehat{v})_{L_2}$ for $u, v \in H^s$.
Show

(i) $H^s := \bigl(H^s\ ;\ (\cdot \,|\, \cdot)_{H^s}\bigr)$ is a Hilbert space with $H^0 = L_2$, and

$$\mathcal{S} \overset{d}{\hookrightarrow} H^s \overset{d}{\hookrightarrow} H^t \overset{d}{\hookrightarrow} L_2 \quad \text{for } s > t > 0\ ;$$

(ii) $H^m = W_2^m$ for $m \in \mathbb{N}$.

7 For $s > n/2$, show

(i) $\mathcal{F}(H^s) \subset L_1$ and

(ii) $H^2 \overset{d}{\hookrightarrow} C_0$ (**Sobolev embedding theorem**).

(Hints: (i) Apply the Cauchy–Schwarz inequality to $\Lambda^s\,|\widehat{u}|\,\Lambda^{-s}$.
(ii) The Riemann–Lebesgue theorem.)

8 Suppose $\sigma \geq 0$, and let $\{\, k_\varepsilon\ ;\ \varepsilon > 0 \,\}$ be the Gaussian approximating kernel. Prove:

(i) $T(t) := [f \mapsto k_{\sqrt{t}} * f]$ belongs to $\mathcal{L}(H^\sigma)$ for every $t > 0$.

(ii) $T(t + s) = T(t)T(s),\ \ s, t > 0$.

(iii) $\lim_{t\to 0} T(t)f = f$ for $f \in H^\sigma$.

(iv) $T(t)(L_2) \subset \mathcal{S},\ \ t > 0$.

(v) For $f \in L_2 \cap C$, let $u(t, x) := T(t)f(x)$ for $(t, x) \in [0, \infty) \times \mathbb{R}^n$. Show that u solves the initial value problem of the heat equation in $\mathbb{R}^n$, that is,

$$\partial_t u - \Delta u = 0 \text{ in } (0, \infty) \times \mathbb{R}^n \quad \text{and} \quad u(0, \cdot\,) = f \text{ on } \mathbb{R}^n\ , \qquad (9.20)$$

in the sense that $u \in C^\infty\bigl((0, \infty) \times \mathbb{R}^n\bigr) \cap C(\mathbb{R}^+ \times \mathbb{R}^n)$ and that u satisfies (9.20) pointwise.

Remark Let $T(0) := \mathrm{id}_{H^\sigma}$. Then $\{ T(t) \; ; \; t \geq 0 \}$ is called the **Gauss–Weierstrass semigroup** (of H^σ).
(Hint: (v) To get an initial value problem for an ordinary differential equation, apply to (9.20) the Fourier transformation with respect to $x \in \mathbb{R}^n$.)

9 Let $n = 1$ and $p_y(x) := \sqrt{2/\pi}\, y/(x^2 + y^2)$ for $(x, y) \in \mathbb{H}^2$. Also let $\sigma \geq 0$. Prove these statements:

(i) $P(y) := [f \mapsto p_y * f]$ belongs to $\mathcal{L}(H^\sigma)$ for every $t > 0$.

(ii) $P(y + z) = P(y)P(z)$ for $y, z > 0$.

(iii) $\lim_{y \to 0} P(y)f = f$ for $f \in H^\sigma$.

(iv) $P(y)(L_2) \subset \mathcal{S}$.

(v) For $f \in L_2 \cap C$, let

$$u(x, y) := \bigl(P(y)f\bigr)(x) \quad \text{for } (x, y) \in \mathbb{H}^2 \; .$$

Then u belongs to $C^2(\mathbb{H}^2) \cap C(\overline{\mathbb{H}}^2)$ and solves the Dirichlet boundary value problem for the half plane given by

$$\Delta u = 0 \text{ in } \mathbb{H}^2 \quad \text{and} \quad u(\,\cdot\,, 0) = f \text{ on } \mathbb{R} \; .$$

Remark With $P(0) := \mathrm{id}_{H^\sigma}$, we call $\{ P(y) \; ; \; y \geq 0 \}$ the **Poisson semigroup** (of H^σ).
(Hints: (ii) Exercise 1. (v) Example 9.21(b).)

10 Suppose X is open in $\mathbb{R}^n$ and (X_k) is an ascending sequence of relatively compact open subsets of X with $X = \bigcup_k X_k$ (see Remarks 1.16(d) and (e)). Also let

$$q_k(f) := \max_{|\alpha| \leq k} \|\partial^\alpha f\|_{\infty, \overline{X}_k} \quad \text{for } f \in C^\infty(X) \text{ and } k \in \mathbb{N} \; ,$$

and

$$d(f, g) := \sum_{k=0}^{\infty} 2^{-k} \, \frac{q_k(f - g)}{1 + q_k(f - g)} \quad \text{for } f, g \in C^\infty(X) \; .$$

Show that $(C^\infty(X), d)$ is a complete metric space. (Hint: To prove the completeness, apply the diagonal sequence principle (Remark III.3.11(a)).)

11 Show that $\mathcal{D} \overset{d}{\hookrightarrow} C^\infty$ and $\mathcal{S} \overset{d}{\hookrightarrow} C^\infty$.
(Hint: Consider $\varphi(\varepsilon\,\cdot\,)$ with a cutoff function φ for $\overline{\mathbb{B}}^n$).

12 For $f \in \mathcal{D}$, let

$$F(z) := \int e^{-i (z \,|\, x)_{\mathbb{C}^n}} f(x)\, dx \quad \text{for } z \in \mathbb{C} \; .$$

Show then that F belongs to $C^\omega(\mathbb{C}, \mathbb{C})$.
(Hint: With Remark V.3.4(c) in mind, apply Corollary 3.19.)

13 Show that $\widehat{f}$ does not belong to $\mathcal{D}$ for $f \in \mathcal{D} \backslash \{0\}$. (Hint: Recall Exercise 12 and the identity theorem for analytic functions (Theorem V.3.13).)