Piermarco Cannarsa
Teresa D'Aprile

Introduction to Measure Theory and Functional Analysis

UNITEXT

Springer

UNITEXT - La Matematica per il 3+2

Volume 89

More information about this series at http://www.springer.com/series/5418

Piermarco Cannarsa · Teresa D'Aprile

Introduction to Measure Theory and Functional Analysis

Springer

Piermarco Cannarsa
Department of Mathematics
Università degli Studi di Roma
"Tor Vergata"
Rome
Italy

Teresa D'Aprile
Department of Mathematics
Università degli Studi di Roma
"Tor Vergata"
Rome
Italy

Translated by the authors.

Translation from the Italian language edition: *Introduzione alla teoria della misura e all'analisi funzionale*, Piermarco Cannarsa and Teresa D'Aprile, © Springer-Verlag Italia, Milano 2008.

ISSN 2038-5722 ISSN 2038-5757 (electronic)
UNITEXT - La Matematica per il 3+2
ISBN 978-3-319-17018-3 ISBN 978-3-319-17019-0 (eBook)
DOI 10.1007/978-3-319-17019-0

Library of Congress Control Number: 2015937015

Springer Cham Heidelberg New York Dordrecht London

Cover design: Simona Colombo, Giochi di Grafica, Milano, Italy

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

To Daniele, Maria Cristina,
Raffaele, and Riccardo

Preface

*It was the quest for a meaning, the praise of doubt,
the wonderful fascination exerted by research in one
of my diaries that seized me. The author was
a mathematician from the school of Renato
Cacciopoli. He spent his entire life among numbers,
haunted by the febrile disquietude caused by the discovery
of infinities. Astronomers rub shoulders with them, they
seek them, they study them. Philosophers dream about
them, talk about them, invent them. Mathematicians
bring them to life, draw closer and closer
and eventually touch them.*

Walter Veltroni, *The discovery of dawn*[1]

This monograph aims at getting the reader acquainted with theories that play a central role in modern mathematics such as *integration* and *functional analysis*. Ultimately, these theories generalize notions that are treated in basic undergraduate courses—and even earlier, in high school—such as orthogonal vectors, linear transformations between Euclidean spaces, and the area delimited by the graph of a function of one real variable. Then, what is this generalization all about? It is about the more and more general nature of the environment in which these notions become meaningful: orthogonality in Hilbert spaces, linear transformations in Banach spaces, integration in measure spaces. These abstract structures are no longer restricted to a specific model like the real line or the Cartesian plane, but possess the least necessary properties to perform the operations we are interested in.

The reader should be warned that the above generalizations are not driven by mere search of abstraction or aesthetic pleasure. Indeed, on the one hand, this kind of procedure—typical in mathematics—allows to subsume a large body of results under few general theorems, the proof of which goes to the essence of the matter. On the other hand, in this way one discovers new phenomena and applications that

[1]Translated from Veltroni W., La scoperta dell'alba, p. 19. RCS Libri S.p.A., Milano (2006).

would be completely out of reach otherwise. In what follows, we have tried to share with the reader our interest for such an approach providing numerous examples, exercises, and some shortcuts to classical results, like our convolution-based proof of the Weierstrass approximation theorem for continuous functions.

We hope this textbook will be useful to graduate students in mathematics, who will find the basic material they will need in their future careers, no matter what they choose to specialize in, as well as researchers in other disciplines, who will be able to read this book without having to know a long list of preliminaries, such as Lebesgue integration in $\mathbb{R}^n$ or compactness criteria for families of continuous functions. The appendices at the end of the book cover a variety of topics ranging from the distance function to Ekeland's variational principle. This material is intended to render the exposition completely self-contained for whoever masters basic linear algebra and mathematical analysis.

Another aspect we would like to point out is that the two main subjects of this monograph, namely integration and functional analysis, are not treated as independent topics but as deeply intertwined theories. This feature is particularly evident in the large choice of problems we propose, the solution of which is often assisted with generous hints. Chapters 1–6 allow to cover both integration and functional analysis in a single course requiring a certain effort on the students' part.

If the material is split into two courses, then one can pick additional topics from the third part of the book, such as functions of bounded variation, absolutely continuous functions, signed measures, the Radon-Nikodym theorem, the characterization of the duals of Lebesgue spaces, and an introduction to set-valued maps. However, the two topics can be treated independently, as one is sometimes forced to do. In this case, Chaps. 1–4 provide the base for a course on integration theory for a broad range of students, not only for those with an interest in analysis. For instance, we have chosen an abstract approach to measure theory in order to quickly derive the extension theorem for countably additive set functions, which is a fundamental result of frequent use in probability. Chapters 5 and 6 are an essential introduction to functional analysis which highlights geometrical aspects of infinite-dimensional spaces. This part of the exposition is appropriate even for undergraduates once all examples requiring measure theory have been filtered out. Indeed, the new phenomena that occur in infinite-dimensional spaces are well exemplified in ℓ^p spaces, without need of any advanced measure-theoretical tools.

To conclude this preface, we would like to express our gratitude to all the people who made this work possible. In particular, we are deeply grateful to Giuseppe Da Prato who originated this monograph providing inspiration for both contents and methods. We would also like to thank our friend Ciro Ciliberto for encouraging us to turn our lecture notes into a book, getting us in touch with Francesca Bonadei who gave us all her valuable professional help and support. Many thanks are due to our students at the University of Rome "Tor Vergata", who read preliminary versions of our notes and solved most of the problems we propose in this textbook. We

wish to send special thanks, directly from our hearts, to Carlo Sinestrari and Francesca Tovena for standing by with their precious advice and invaluable patience. Finally, we would like to share with the reader our happiness for the increased set of names to whom this volume is dedicated, compared to the Italian edition. It is true that time does not go by in vain.

Rome, Italy
October 2014

Piermarco Cannarsa
Teresa D'Aprile

Contents

Part III Selected Topics

Part I
Measure and Integration

Chapter 1
Measure Spaces

The concept of *measure of a set* originates from the classical notion of volume of an interval in $\mathbb{R}^N$. Starting from such an intuitive idea, by a covering process one can assign to any set a nonnegative number which "quantifies its extent". Such an association leads to the introduction of a set function called *exterior measure*, which is defined for all subsets of $\mathbb{R}^N$. The exterior measure is monotone but fails to be additive. Following Carathéodory's construction, it is possible to select a family of sets for which the exterior measure enjoys further properties such as countable additivity. By restricting the exterior measure to such a family one obtains a *complete measure*. This is the procedure that allows to define the Lebesgue measure in $\mathbb{R}^N$. The family of all Lebesgue measurable sets is very large: sets that fail to be measurable can only be constructed by using the Axiom of Choice.

Although the Lebesgue measure was initially developed in euclidean spaces, this theory is independent of the geometry of the background space and applies to abstract spaces as well. This fact is essential for applications: indeed measure theory has been successfully applied to functional analysis, probability, dynamical systems, and other domains of mathematics.

In this chapter, we will develop measure theory from an abstract viewpoint, extending the procedure that leads to the Lebesgue measure in order to construct a large variety of measures on a generic space X. In the particular case of $X = \mathbb{R}^N$, a special role is played by Radon measures (of which the Lebesgue measure is an example) that have important *regularity* properties.

1.1 Algebras and σ-Algebras of Sets

1.1.1 Notation and Preliminaries

We shall denote by X a nonempty set, by $\mathscr{P}(X)$ the set of all parts (i.e., subsets) of X, and by $\varnothing$ the empty set.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_1

For any subset A of X we shall denote by A^c its complement, i.e.,

$$A^c = \{x \in X \mid x \notin A\}.$$

For any $A, B \in \mathscr{P}(X)$ we set $A \setminus B = A \cap B^c$.
Let $(A_n)_n$ be a sequence in $\mathscr{P}(X)$. The following *De Morgan* identity holds:

$$\left(\bigcup_{n=1}^{\infty} A_n\right)^c = \bigcap_{n=1}^{\infty} A_n^c.$$

We define[1]

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

If $L := \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$, then we set $L = \lim_{n\to\infty} A_n$, and we say that $(A_n)_n$ converges to L (in this sense we shall write $A_n \to L$).

Remark 1.1 (a) As is easily checked, $\limsup_{n\to\infty} A_n$ (resp., $\liminf_{n\to\infty} A_n$) consists of those elements of X that belong to infinitely many subsets A_n (resp., that belong to all but a finite number of subsets A_n). Therefore

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n.$$

(b) It is also immediate to check that if $(A_n)_n$ is increasing ($A_n \subset A_{n+1},\ n \in \mathbb{N}$), then

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n,$$

whereas, if $(A_n)_n$ is decreasing ($A_n \supset A_{n+1},\ n \in \mathbb{N}$), then

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

In the first case we shall write $A_n \uparrow L$, and in the second $A_n \downarrow L$.

[1] Observe the similarity with lim inf and lim sup for a sequence $(a_n)_n$ of real numbers. We have: $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k$ and $\liminf_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k$.

1.1.2 Algebras and σ-Algebras

Definition 1.2 A nonempty subset $\mathscr{A}$ of $\mathscr{P}(X)$ is called an *algebra* in X if

(a) $\varnothing, X \in \mathscr{A}$.
(b) $A, B \in \mathscr{A} \implies A \cup B \in \mathscr{A}$.
(c) $A \in \mathscr{A} \implies A^c \in \mathscr{A}$.

Remark 1.3 It is easy to see that if $\mathscr{A}$ is an algebra and $A, B \in \mathscr{A}$, then $A \cap B$ and $A \setminus B$ belong to $\mathscr{A}$. Therefore the symmetric difference

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

also belongs to $\mathscr{A}$. Moreover, $\mathscr{A}$ is stable under finite unions and intersections, that is,

$$A_1, \ldots, A_n \in \mathscr{A} \implies \begin{cases} A_1 \cup \cdots \cup A_n \in \mathscr{A}, \\ A_1 \cap \cdots \cap A_n \in \mathscr{A}. \end{cases}$$

Definition 1.4 An algebra $\mathscr{E}$ in X is called a σ-*algebra* if, for any sequence $(A_n)_n$ of elements of $\mathscr{E}$, we have that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{E}$. If $\mathscr{E}$ is a σ-algebra in X, the elements of $\mathscr{E}$ are called *measurable sets* and the pair $(X, \mathscr{E})$ is called a *measurable space*.

Exercise 1.5 Show that an algebra $\mathscr{E}$ in X is a σ-algebra if and only if, for any sequence $(A_n)_n$ of mutually disjoint elements of $\mathscr{E}$, we have $\bigcup_{n=1}^{\infty} A_n \in \mathscr{E}$.
Hint. Given a sequence $(A_n)_n$ of elements of $\mathscr{E}$, set $B_1 = A_1$ and $B_n = A_n \setminus (B_1 \cup \ldots \cup B_{n-1})$ for $n \geq 2$. Show that $(B_n)_n$ is a sequence of disjoint elements of $\mathscr{E}$ and $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n \in \mathscr{E}$.

We note that if $\mathscr{E}$ is a σ-algebra in X and $(A_n)_n \subset \mathscr{E}$, then $\bigcap_{n=1}^{\infty} A_n \in \mathscr{E}$ owing to the De Morgan identity. Moreover,

$$\liminf_{n\to\infty} A_n \in \mathscr{E}, \quad \limsup_{n\to\infty} A_n \in \mathscr{E}.$$

Example 1.6 The following examples explain the difference between algebras and σ-algebras.

1. Obviously, $\mathscr{P}(X)$ and $\mathscr{E} = \{\varnothing, X\}$ are σ-algebras in X. Moreover, $\mathscr{P}(X)$ is the largest σ-algebra in X, and $\mathscr{E}$ the smallest.
2. In $X = [0, 1)$, the class $\mathscr{A}$ consisting of $\varnothing$ and of all finite unions

$$A = \bigcup_{i=1}^{n} [a_i, b_i) \text{ with } 0 \leq a_i \leq b_i \leq a_{i+1} \leq 1 \tag{1.1}$$

is an algebra in $[0, 1)$. Indeed, for A as in (1.1), we have

$$A^c = [0, a_1) \cup [b_1, a_2) \cup \cdots \cup [b_n, 1) \in \mathscr{A}.$$

Moreover, in order to show that $\mathscr{A}$ is stable under finite unions it suffices to observe that the union of two (not necessarily disjoint) intervals $[a, b),\ [c, d) \subset [0, 1)$ belongs to $\mathscr{A}$.

3. In an infinite set X consider the class

$$\mathscr{A} = \{A \in \mathscr{P}(X) \mid A \text{ is finite, or } A^c \text{ is finite}\}.$$

Then $\mathscr{A}$ is an algebra in X. Indeed, the only point that needs to be checked is that $\mathscr{A}$ is stable under finite unions. Let $A, B \in \mathscr{A}$. If A and B are both finite, then so is $A \cup B$. In all other cases, $(A \cup B)^c$ is finite.

4. In an uncountable set X consider the class[2]

$$\mathscr{E} = \{A \in \mathscr{P}(X) \mid A \text{ is countable, or } A^c \text{ is countable}\}.$$

Then $\mathscr{E}$ is a σ-algebra in X. Indeed, $\mathscr{E}$ is stable under countable unions: let $(A_n)_n$ be a sequence in $\mathscr{E}$; if all A_n are countable, then so is $\cup_n A_n$; otherwise, $(\cup_n A_n)^c$ is countable.

Exercise 1.7 1. Show that the algebra $\mathscr{A}$ in Example 1.6(2) fails to be a σ-algebra.
2. Show that the algebra $\mathscr{A}$ in Example 1.6(3) fails to be a σ-algebra.
3. Give an example to show that the σ-algebra $\mathscr{E}$ in Example 1.6(4) is, in general, strictly smaller than $\mathscr{P}(X)$.
4. Let $\mathscr{K}$ be a subset of $\mathscr{P}(X)$. Show that the intersection of all σ-algebras in X including $\mathscr{K}$ is a σ-algebra in X (the minimal σ-algebra including $\mathscr{K}$).

Definition 1.8 Given a subset $\mathscr{K}$ of $\mathscr{P}(X)$, the intersection of all σ-algebras in X including $\mathscr{K}$ is called the *σ-algebra generated by $\mathscr{K}$*, and will be denoted by $\sigma(\mathscr{K})$.

Exercise 1.9 1. Show that if $\mathscr{E}$ is a σ-algebra in X, then $\sigma(\mathscr{E}) = \mathscr{E}$.
2. Find $\sigma(\mathscr{K})$ for $\mathscr{K} = \{\varnothing\}$ and $\mathscr{K} = \{X\}$.
3. Given $\mathscr{K},\ \mathscr{K}' \subset \mathscr{P}(X)$ with $\mathscr{K} \subset \mathscr{K}' \subset \sigma(\mathscr{K})$, show that

$$\sigma(\mathscr{K}') = \sigma(\mathscr{K}).$$

Example 1.10 1. Let X be a metric space. The σ-algebra generated by all open sets of X is called the *Borel σ-algebra* and is denoted by $\mathscr{B}(X)$. Obviously, $\mathscr{B}(X)$ coincides with the σ-algebra generated by all closed sets of X. The elements of $\mathscr{B}(X)$ are called *Borel sets*.

2. Let $X = \mathbb{R}$, and let $\mathscr{I}$ be the class of all half-closed intervals $[a, b)$ with $a < b$. Then $\sigma(\mathscr{I})$ coincides with $\mathscr{B}(\mathbb{R})$. Indeed, let us observe that every half-closed interval $[a, b)$ belongs to $\mathscr{B}(\mathbb{R})$ since

[2]In the following 'countable' stands for 'finite or countable'.

$$[a, b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}, b\right).$$

So $\sigma(\mathscr{I}) \subset \mathscr{B}(\mathbb{R})$. Conversely, let V be an open set in $\mathbb{R}$. Then, as is well known, V is the countable union of some family of open intervals.[3] Since any open interval (a, b) can be represented as

$$(a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b\right),$$

we conclude that $V \in \sigma(\mathscr{I})$. Thus, $\mathscr{B}(\mathbb{R}) \subset \sigma(\mathscr{I})$.

Exercise 1.11 Show that $\sigma(\mathscr{I}) = \mathscr{B}(\mathbb{R})$, where $\mathscr{I}$ is one of the following classes:

$$\mathscr{I} = \{(a, b) \,|\, a,\ b \in \mathbb{R},\ a < b\},$$

$$\mathscr{I} = \{(a, \infty) \,|\, a \in \mathbb{R}\},$$

$$\mathscr{I} = \{(-\infty, a] \,|\, a \in \mathbb{R}\}.$$

Exercise 1.12 Let $\mathscr{E}$ be a σ-algebra in X, and $X_0 \subset X$.

1. Show that $\mathscr{E}_0 = \{A \cap X_0 \mid A \in \mathscr{E}\}$ is a σ-algebra in X_0.
2. Show that if $\mathscr{E} = \sigma(\mathscr{K})$, then $\mathscr{E}_0 = \sigma(\mathscr{K}_0)$, where

$$\mathscr{K}_0 = \{A \cap X_0 \,|\, A \in \mathscr{K}\}.$$

Hint. The inclusion $\mathscr{E}_0 \supset \sigma(\mathscr{K}_0)$ follows from point 1. To prove the converse, show that

$$\mathscr{F} := \{A \in \mathscr{E} \,|\, A \cap X_0 \in \sigma(\mathscr{K}_0)\}$$

is a σ-algebra in X including $\mathscr{K}$.

1.2 Measures

1.2.1 Additive and σ-Additive Functions

Definition 1.13 Let $\mathscr{A}$ be an algebra in X and let $\mu : \mathscr{A} \to [0, \infty]$ be a function such that $\mu(\varnothing) = 0$.

---

[3] Indeed, each point $x \in V$ has an open interval $(p_x, q_x) \subset V$ with $x \in (p_x, q_x)$ and $p_x, q_x \in \mathbb{Q}$. Therefore V is contained in the union of all elements of the family $\{(p, q) \mid p, q \in \mathbb{Q},\ (p, q) \subset V\}$, and this family is countable.

- We say that μ is *additive* if, for any finite family $A_1, \ldots, A_n \in \mathscr{A}$ of mutually disjoint sets, we have
$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$
- We say that μ is σ-*additive* or *countably additive* if, for any sequence $(A_n)_n \subset \mathscr{A}$ of mutually disjoint sets such that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, we have
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$
- We say that μ is σ-*subadditive* (or *countably subadditive*) if, for any sequence $(A_n)_n \subset \mathscr{A}$ such that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, we have
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Remark 1.14 Let $\mathscr{A}$ be an algebra in X.

1. Any σ-additive function on $\mathscr{A}$ is also additive.
2. Any additive function μ on $\mathscr{A}$ is *monotone*. Indeed, if $A, B \in \mathscr{A}$ and $A \supset B$, then $\mu(A) = \mu(B) + \mu(A \setminus B)$. Therefore $\mu(A) \geq \mu(B)$.
3. Let μ be an additive function on $\mathscr{A}$, and let $(A_n)_n \subset \mathscr{A}$ be a sequence of mutually disjoint sets such that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$. Then
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \sum_{n=1}^{m} \mu(A_n) \quad \text{for all } m \in \mathbb{N}.$$
Therefore
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \sum_{n=1}^{\infty} \mu(A_n).$$
4. Any σ-additive function μ on $\mathscr{A}$ is also σ-*subadditive*. Indeed, let $(A_n)_n \subset \mathscr{A}$ be a sequence such that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, and define $B_1 = A_1$ and $B_n = A_n \setminus (B_1 \cup \ldots \cup B_{n-1})$ for $n \geq 2$. Then $(B_n)_n$ is a sequence of mutually disjoint sets of $\mathscr{A}$, $\cup_n A_n = \cup_n B_n \in \mathscr{A}$ and $\mu(B_n) \leq \mu(A_n)$ by the monotonicity of μ. Therefore $\mu(\cup_n A_n) = \mu(\cup_n B_n) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n)$.
5. In view of point 3 and 4, an additive function on $\mathscr{A}$ is σ-additive if and only if it is σ-*subadditive*.

Definition 1.15 An additive function μ on an algebra $\mathscr{A} \subset \mathscr{P}(X)$ is said to be:

- *finite* if $\mu(X) < \infty$.
- *σ-finite* if there exists a sequence $(A_n)_n \subset \mathscr{A}$ such that $\bigcup_{n=1}^{\infty} A_n = X$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.

Exercise 1.16 In $X = \mathbb{N}$, consider the algebra

$$\mathscr{A} = \{A \in \mathscr{P}(X) \mid A \text{ is finite, or } A^c \text{ is finite}\}$$

of Example 1.6. Show that:

- The function $\mu : \mathscr{A} \to [0, \infty]$ defined as

$$\mu(A) = \begin{cases} \#A & \text{if } A \text{ is finite,} \\ \infty & \text{if } A^c \text{ is finite} \end{cases}$$

 (where the symbol $\#A$ stands for the number of elements of A) is σ-additive.
- The function $\nu : \mathscr{A} \to [0, \infty]$ defined as

$$\nu(A) = \begin{cases} \displaystyle\sum_{n \in A} \frac{1}{2^n} & \text{if } A \text{ is finite,} \\ \infty & \text{if } A^c \text{ is finite} \end{cases}$$

 is additive but not σ-additive.

For an additive function, σ-additivity is equivalent to continuity in the sense of the following proposition.

Proposition 1.17 *Let μ be an additive function on an algebra $\mathscr{A}$. Then $(i) \Leftrightarrow (ii)$, where:*

(i) *μ is σ-additive.*
(ii) $(A_n)_n, \subset \mathscr{A}, \ A \in \mathscr{A}, \ A_n \uparrow A \implies \mu(A_n) \uparrow \mu(A)$.

Proof Let us first consider the implication (i) $\Rightarrow$ (ii). Let $(A_n)_n \subset \mathscr{A}$, $A \in \mathscr{A}$, $A_n \uparrow A$. Then

$$A = A_1 \cup \bigcup_{n=1}^{\infty} (A_{n+1} \setminus A_n),$$

the above union being disjoint. Since μ is σ-additive, we deduce that

$$\mu(A) = \mu(A_1) + \sum_{n=1}^{\infty} (\mu(A_{n+1}) - \mu(A_n)) = \lim_{n\to\infty} \mu(A_n),$$

and (ii) follows.

Let us pass to prove that (ii) $\Rightarrow$ (i). Let $(A_n)_n \subset \mathscr{A}$ be a sequence of mutually disjoint sets such that $A = \bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$. Define

$$B_n = \bigcup_{k=1}^{n} A_k.$$

Then $B_n \uparrow A$. So, in view of (ii), $\mu(B_n) = \sum_{k=1}^{n} \mu(A_k) \uparrow \mu(A)$. This implies (i). □

Proposition 1.18 *Let μ be a σ-additive function on an algebra $\mathscr{A}$. If $(A_n)_n \subset \mathscr{A}$, $A \in \mathscr{A}$, $\mu(A_1) < \infty$ and $A_n \downarrow A$, then $\mu(A_n) \downarrow \mu(A)$.*

Proof We have

$$A_1 = \bigcup_{n=1}^{\infty} (A_n \setminus A_{n+1}) \cup A,$$

the above union being disjoint. Consequently,

$$\mu(A_1) = \sum_{n=1}^{\infty} \big(\mu(A_n) - \mu(A_{n+1})\big) + \mu(A) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n) + \mu(A).$$

Since $\mu(A_1) < \infty$, the conclusion follows. □

Example 1.19 The conclusion of Proposition 1.18 may be false without assuming $\mu(A_1) < \infty$. This is easily checked taking $\mathscr{A}$ and μ as in Exercise 1.16 and $A_n = \{m \in \mathbb{N} \mid m \geq n\}$.

Exercise 1.20 Let μ be a finite additive function on an algebra $\mathscr{A}$. Show that if for any sequence $(A_n)_n \subset \mathscr{A}$ such that $A_n \downarrow A \in \mathscr{A}$ we have

$$\mu(A_n) \downarrow \mu(A),$$

then μ is σ-additive.

1.2.2 Measure Spaces

Definition 1.21 Let $\mathscr{E}$ be a σ-algebra in X.

- A σ-additive function $\mu : \mathscr{E} \to [0, \infty]$ is called a *measure* on $\mathscr{E}$.
- The triplet $(X, \mathscr{E}, \mu)$, where μ is a measure on $\mathscr{E}$, is called a *measure space*.
- A measure μ on $\mathscr{E}$ is called a *probability measure* if $\mu(X) = 1$.

Definition 1.22 A measure μ on a σ-algebra $\mathscr{E} \subset \mathscr{P}(X)$ is said to be

- *finite* if $\mu(X) < \infty$.
- *σ-finite* if there exists a sequence $(A_n)_n \subset \mathscr{E}$ such that $\bigcup_{n=1}^{\infty} A_n = X$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.
- *complete* if

$$A \in \mathscr{E}, \ B \subset A, \ \mu(A) = 0 \quad \Longrightarrow \quad B \in \mathscr{E}$$

  (and so $\mu(B) = 0$).
- *concentrated on a set* $A \in \mathscr{E}$ if $\mu(A^c) = 0$. In this case we say that A is a support of μ.

Example 1.23 Let $x \in X$. Define, for every $A \in \mathscr{P}(X)$,

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

Then δ_x is a measure on $\mathscr{P}(X)$, called the *Dirac measure at x*. Such a measure is concentrated on the singleton $\{x\}$.

Example 1.24 Let us define, for every $A \in \mathscr{P}(X)$,

$$\mu^{\#}(A) = \begin{cases} \#A & \text{if } A \text{ is finite} \\ \infty & \text{if } A \text{ is infinite} \end{cases}$$

(see Exercise 1.16). Then $\mu^{\#}$ is a measure on $\mathscr{P}(X)$, called the *counting measure*. It is easy to see that $\mu^{\#}$ is finite if and only if X is finite, and that $\mu^{\#}$ is σ-finite if and only if X is countable.

Exercise 1.25 Given a measure space $(X, \mathscr{E}, \mu)$, a set $A \in \mathscr{E}$ of measure zero is called a *null set* or *zero-measure set*. Show that a countable union of zero-measure sets is also a zero-measure set.

Definition 1.26 Given $(X, \mathscr{E}, \mu)$ a measure space and $A \in \mathscr{E}$, the *restriction* of μ to A (or μ *restricted* to A), written as $\mu \llcorner A$, is the set function[4]

$$(\mu \llcorner A)(B) = \mu(A \cap B) \qquad \forall B \in \mathscr{E}.$$

Exercise 1.27 In the same hypotheses of Definition 1.26, show that $\mu \llcorner A$ is a measure on $\mathscr{E}$.

Remark 1.28 Let us observe that, given a measure space $(X, \mathscr{E}, \mu)$, any subset $A \in \mathscr{E}$ can be naturally endowed with a measure space frame: more precisely, the new σ-algebra will be $\mathscr{E} \cap A$, namely the class of all measurable subsets of X which

[4] A *set function* is a function $\mathscr{D} \to [-\infty, +\infty]$, where $\mathscr{D} \subset \mathscr{P}(X)$ is a family including the empty set.

are contained in A, and the new measure, which we will continue to denote by μ, is identical to μ except for the restriction of its domain. The measure space $(A, \mathscr{E} \cap A, \mu)$ is called a *measure subspace* of $(X, \mathscr{E}, \mu)$.

As a corollary of Proposition 1.18 we have the following result.

Proposition 1.29 *Let μ be a finite measure on a σ-algebra $\mathscr{E}$. Then, for any sequence $(A_n)_n \subset \mathscr{E}$, we have*

$$\mu\Big(\liminf_{n\to\infty} A_n\Big) \le \liminf_{n\to\infty} \mu(A_n) \le \limsup_{n\to\infty} \mu(A_n) \le \mu\Big(\limsup_{n\to\infty} A_n\Big). \tag{1.2}$$

In particular, $A_n \to A \Longrightarrow \mu(A_n) \to \mu(A)$.

Proof Set $L = \limsup_{n\to\infty} A_n$. Then we can write $L = \bigcap_{n=1}^{\infty} B_n$, where $B_n = \bigcup_{k=n}^{\infty} A_k \downarrow L$. Now, by Proposition 1.18 it follows that

$$\mu(L) = \lim_{n\to\infty} \mu(B_n) = \inf_{n\in\mathbb{N}} \mu(B_n) \ge \inf_{n\in\mathbb{N}} \sup_{k\ge n} \mu(A_k) = \limsup_{n\to\infty} \mu(A_n).$$

We have thus proved that

$$\limsup_{n\to\infty} \mu(A_n) \le \mu\Big(\limsup_{n\to\infty} A_n\Big).$$

The remaining part of (1.2) can be proved similarly. □

1.2.3 Borel–Cantelli Lemma

The following result states a simple but useful property of measures.

Lemma 1.30 *Let μ be a measure on a σ-algebra $\mathscr{E}$. Then, for any sequence $(A_n)_n \subset \mathscr{E}$ satisfying $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, we have*

$$\mu\Big(\limsup_{n\to\infty} A_n\Big) = 0.$$

Proof Set $L = \limsup_{n\to\infty} A_n$. Then $L = \bigcap_{n=1}^{\infty} B_n$, where $B_n = \bigcup_{k=n}^{\infty} A_k \downarrow L$. Consequently, since μ is σ-subadditive, we have

$$\mu(L) \le \mu(B_n) \le \sum_{k=n}^{\infty} \mu(A_k)$$

for any $n \in \mathbb{N}$. As $n \to \infty$, we obtain $\mu(L) = 0$. □

1.3 The Basic Extension Theorem

A natural question arising both in theory and applications is the following.

Problem 1.31 Let $\mathscr{A}$ be an algebra in X, and let μ be an additive function on $\mathscr{A}$. Does there exist a σ-algebra $\mathscr{E}$ including $\mathscr{A}$, and a measure $\overline{\mu}$ on $\mathscr{E}$ that extends μ, i.e.,

$$\overline{\mu}(A) = \mu(A) \qquad \forall A \in \mathscr{A}?$$

Should the above problem have a solution, one could assume $\mathscr{E} = \sigma(\mathscr{A})$ since $\sigma(\mathscr{A})$ would be contained in $\mathscr{E}$ anyways. Moreover, for any sequence $(A_n)_n \subset \mathscr{A}$ of mutually disjoint sets such that $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$, we would have

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

Thus, for Problem 1.31 to have a positive answer, μ must be σ-additive. The following remarkable result shows that such a property is also sufficient for the existence of an extension, and more. We shall see an important application of this result to the construction of the Lebesgue measure later on in this chapter.

Theorem 1.32 *Let $\mathscr{A}$ be an algebra in X, and $\mu\colon \mathscr{A} \to [0, \infty]$ be a σ-additive function. Then μ can be extended to a measure on $\sigma(\mathscr{A})$. Moreover, such an extension is unique if μ is σ-finite.*

To prove the above theorem we need to develop suitable tools, namely Halmos' Monotone Class Theorem for uniqueness, and the concept of outer measure and additive set for existence.

1.3.1 Monotone Classes

Definition 1.33 A nonempty class $\mathscr{M} \subset \mathscr{P}(X)$ is called a *monotone class* if, for any sequence $(A_n)_n \subset \mathscr{M}$,

- $A_n \uparrow A \implies A \in \mathscr{M}$.
- $A_n \downarrow A \implies A \in \mathscr{M}$.

Remark 1.34 Clearly, any σ-algebra is a monotone class; the converse, however, may fail, as can be checked by considering the trivial example $\mathscr{M} = \{\varnothing\}$. On the other hand, if a monotone class $\mathscr{M}$ is also an algebra in X, then $\mathscr{M}$ is a σ-algebra in X. Indeed, given a sequence $(A_n)_n \subset \mathscr{M}$, we have $B_n := \cup_{k=1}^{n} A_k \in \mathscr{M}$ and $B_n \uparrow A := \cup_{k=1}^{\infty} A_k$. Therefore $A \in \mathscr{M}$.

Let us now prove the following result.

Theorem 1.35 (Halmos) *Let $\mathscr{A}$ be an algebra in X, and let $\mathscr{M}$ be a monotone class including $\mathscr{A}$. Then $\sigma(\mathscr{A}) \subset \mathscr{M}$.*

Proof Let $\mathscr{M}_0$ be the minimal monotone class[5] in X including $\mathscr{A}$. We are going to show that $\mathscr{M}_0$ is an algebra in X, and this will prove the theorem in view of Remark 1.34.

To begin with, we note that $\varnothing$ and X belong to $\mathscr{M}_0$. Define, for any $A \in \mathscr{M}_0$,

$$\mathscr{M}_A = \{B \in \mathscr{M}_0 \mid A \cup B,\ A \setminus B,\ B \setminus A \in \mathscr{M}_0\}.$$

We claim that $\mathscr{M}_A$ is a monotone class. Indeed, let $(B_n)_n \subset \mathscr{M}_A$ be an increasing sequence such that $B_n \uparrow B$. Then

$$A \cup B_n \uparrow A \cup B, \quad A \setminus B_n \downarrow A \setminus B, \quad B_n \setminus A \uparrow B \setminus A.$$

Since $\mathscr{M}_0$ is a monotone class, we deduce that

$$B,\ A \cup B,\ A \setminus B,\ B \setminus A \in \mathscr{M}_0.$$

Therefore $B \in \mathscr{M}_A$. By a similar argument one can check that

$$(B_n)_n \subset \mathscr{M}_A, \quad B_n \downarrow B \quad \Longrightarrow \quad B \in \mathscr{M}_A.$$

So $\mathscr{M}_A$ is a monotone class as claimed.

Next, let $A \in \mathscr{A}$. Then $\mathscr{A} \subset \mathscr{M}_A$ since any $B \in \mathscr{A}$ belongs to $\mathscr{M}_0$ and satisfies

$$A \cup B,\ A \setminus B,\ B \setminus A \in \mathscr{M}_0. \tag{1.3}$$

But $\mathscr{M}_0$ is the minimal monotone class including $\mathscr{A}$, so $\mathscr{M}_0 \subset \mathscr{M}_A$. Therefore $\mathscr{M}_0 = \mathscr{M}_A$ or, equivalently, (1.3) holds true for any $A \in \mathscr{A}$ and $B \in \mathscr{M}_0$.

Finally, let $A \in \mathscr{M}_0$. Since (1.3) is satisfied by any $B \in \mathscr{A}$, we deduce that $\mathscr{A} \subset \mathscr{M}_A$. Then $\mathscr{M}_A = \mathscr{M}_0$. This implies that $\mathscr{M}_0$ is an algebra. □

Proof of Theorem 1.32: *uniqueness* Let $\mathscr{E} = \sigma(\mathscr{A})$, and let μ_1, μ_2 be two measures extending μ to $\mathscr{E}$. We shall assume, first, that μ is finite and set

$$\mathscr{M} = \{A \in \mathscr{E} \mid \mu_1(A) = \mu_2(A)\}.$$

We claim that $\mathscr{M}$ is a monotone class including $\mathscr{A}$. Indeed, for any sequence $(A_n)_n \subset \mathscr{M}$, by Propositions 1.17 and 1.18 we have that

$$A_n \uparrow A \quad \Longrightarrow \quad \mu_1(A) = \lim_n \mu_i(A_n) = \mu_2(A) \quad (i = 1, 2),$$

$$A_n \downarrow A,\ \mu_1(X), \mu_2(X) < \infty \quad \Longrightarrow \quad \mu_1(A) = \lim_n \mu_i(A_n) = \mu_2(A) \quad (i = 1, 2).$$

---

[5]It is easy to see that the intersection of all monotone classes in X including $\mathscr{A}$ is also a monotone class.

Therefore, by Halmos' Theorem, $\mathscr{M} = \mathscr{E}$ and this implies that $\mu_1 = \mu_2$.

In the general case of a σ-finite function μ, we have that $X = \bigcup_{n=1}^{\infty} X_n$ for some $(X_n)_n \subset \mathscr{A}$ such that $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$. It is not restrictive to assume that the sequence $(X_n)_n$ is increasing. Now, define $\mu_n = \mu \llcorner X_n$, $\mu_{i,n} = \mu_i \llcorner X_n$ for $i = 1, 2$ (see Definition 1.26). Then, as is easily checked, μ_n is a finite σ-additive function on $\mathscr{A}$, and $\mu_{1,n}$, $\mu_{2,n}$ are measures extending μ_n to $\mathscr{E}$. So, by the conclusion of the first part of this proof, $\mu_{1,n} = \mu_{2,n}$. If $A \in \mathscr{E}$, then $A \cap X_n \uparrow A$, and therefore, again by Proposition 1.17, we obtain

$$\begin{aligned}\mu_1(A) &= \lim_{n\to\infty} \mu_1(A \cap X_n) = \lim_{n\to\infty} \mu_{1,n}(A) \\ &= \lim_{n\to\infty} \mu_{2,n}(A) = \lim_{n\to\infty} \mu_2(A \cap X_n) = \mu_2(A).\end{aligned}$$

The proof is thus complete. □

Example 1.36 The above extension may fail to be unique, in general, if the function μ is not σ-finite. Indeed, let us consider the algebra $\mathscr{A}$ of Example 1.6(2) and the σ-additive function μ on $\mathscr{A}$ defined by

$$\mu(A) = \begin{cases} 0 & \text{if } A = \varnothing, \\ \infty & \text{if } A \neq \varnothing. \end{cases} \tag{1.4}$$

By reasoning as in Example 1.10(2), it is easy to show that $\sigma(\mathscr{A}) = \mathscr{B}([0, 1))$. A trivial extension of μ to $\mathscr{B}([0, 1))$ is given by (1.4) itself. To construct a second one, let us consider an enumeration $(q_n)_{n\in\mathbb{N}}$ of $\mathbb{Q} \cap [0, 1)$ and set

$$\widehat{\mu}(A) = \sum_{n=1}^{\infty} \delta_{q_n}(A) \qquad \forall A \in \mathscr{B}([0, 1)),$$

where δ_x is the Dirac measure in x. Then $\widehat{\mu} = \mu$ on $\mathscr{A}$, but $\widehat{\mu}(\{q_1\}) = 1$ and $\mu(\{q_1\}) = \infty$. To prove that $\widehat{\mu}$ is σ-additive, let us first observe that $\widehat{\mu}$ is additive. Now, for any sequence $(A_k)_k \subset \mathscr{B}([0, 1))$, the σ-subadditivity of δ_{q_n} yields

$$\begin{aligned}\widehat{\mu}\left(\bigcup_{k=1}^{\infty} A_k\right) &= \sum_{n=1}^{\infty} \delta_{q_n}\left(\bigcup_{k=1}^{\infty} A_k\right) \le \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \delta_{q_n}(A_k) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N}\sum_{k=1}^{\infty} \delta_{q_n}(A_k) = \lim_{N\to\infty} \sum_{k=1}^{\infty}\sum_{n=1}^{N} \delta_{q_n}(A_k) \\ &\le \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} \delta_{q_n}(A_k) = \sum_{k=1}^{\infty} \widehat{\mu}(A_k).\end{aligned}$$

Therefore $\widehat{\mu}$ is σ-subadditive, and then $\widehat{\mu}$ is also σ-additive in view of Remark 1.14(5).

1.3.2 Outer Measures

Definition 1.37 A function $\mu^* : \mathscr{P}(X) \to [0, \infty]$ is called an *outer measure* on X if $\mu^*(\varnothing) = 0$, and μ^* is monotone and σ-subadditive, i.e.,

$$E_1 \subset E_2 \implies \mu^*(E_1) \le \mu^*(E_2),$$

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \le \sum_{n=1}^{\infty} \mu^*(E_n) \qquad \forall\, (E_n)_n \subset \mathscr{P}(X).$$

The following proposition studies an example of outer measure that will be essential for the proof of Theorem 1.32.

Proposition 1.38 *Let μ be a σ-additive function on an algebra $\mathscr{A} \subset \mathscr{P}(X)$. Define, for any $E \in \mathscr{P}(X)$,*

$$\mu^*(E) = \inf\left\{ \sum_{n=1}^{\infty} \mu(A_n) \;\middle|\; (A_n)_n \subset \mathscr{A},\ E \subset \bigcup_{n=1}^{\infty} A_n \right\}. \tag{1.5}$$

Then

1. *μ^* is finite whenever μ is finite.*
2. *μ^* is an extension of μ, that is,*

$$\mu^*(A) = \mu(A), \quad \forall\, A \in \mathscr{A}. \tag{1.6}$$

3. *μ^* is an outer measure on X.*

Proof The first assertion being obvious, let us proceed to check (1.6). Observe that the inequality $\mu^*(A) \le \mu(A)$ is immediate for any $A \in \mathscr{A}$. To prove the converse, let $(A_n)_n \subset \mathscr{A}$ be a countable covering of a set $A \in \mathscr{A}$. Then $(A_n \cap A)_n \subset \mathscr{A}$ is also a countable covering of A satisfying $\cup_{n=1}^{\infty}(A_n \cap A) = A \in \mathscr{A}$. Since μ is σ-subadditive, we get

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

Thus, taking the infimum as in (1.5), we conclude that $\mu^*(A) \ge \mu(A)$.

The monotonicity of μ^* follows from the definition (1.5) since, if $E_1 \subset E_2$, every countable covering of E_2 is also a countable covering of E_1.

It remains to show that μ^* is σ-subadditive. Let $(E_n)_n \subset \mathscr{P}(X)$, and set $E = \bigcup_{n=1}^{\infty} E_n$. The inequality $\mu^*(E) \le \sum_{n=1}^{\infty} \mu^*(E_n)$ is trivial if the right hand side is infinite. Therefore assume that all $\mu^*(E_n)$'s are finite. Then for any $n \in \mathbb{N}$ and any $\varepsilon > 0$ there exists $(A_{n,k})_k \subset \mathscr{A}$ such that

$$\sum_{k=1}^{\infty} \mu(A_{n,k}) < \mu^*(E_n) + \frac{\varepsilon}{2^n}, \quad E_n \subset \bigcup_{k=1}^{\infty} A_{n,k}.$$

Consequently,

$$\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu(A_{n,k}) \leq \sum_{n=1}^{\infty} \mu^*(E_n) + \varepsilon.$$

Since $E \subset \bigcup_{n,k} A_{n,k}$, we have[6]

$$\mu^*(E) \leq \sum_{(n,k)\in\mathbb{N}^2} \mu(A_{n,k}) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mu(A_{n,k}) \leq \sum_{n=1}^{\infty} \mu^*(E_n) + \varepsilon.$$

The conclusion follows from the arbitrariness of ε. □

Exercise 1.39 1. Let μ^* be an outer measure on X, and $Z \in \mathscr{P}(X)$. Show that

$$\nu^*(E) = \mu^*(Z \cap E) \qquad \forall E \in \mathscr{P}(X)$$

is an outer measure on X.

2. Let $(\mu_n^*)_n$ be a sequence of outer measures on X. Show that

$$\mu^*(E) = \sum_{n=1}^{\infty} \mu_n^*(E) \quad \text{and} \quad \mu_\infty^*(E) = \sup_{n\in\mathbb{N}} \mu_n^*(E) \qquad \forall E \in \mathscr{P}(X)$$

are outer measures on X.

Definition 1.40 Given an outer measure μ^* on X, a set $A \in \mathscr{P}(X)$ is said to be *additive* (or μ^**-measurable*) if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \forall E \in \mathscr{P}(X). \tag{1.7}$$

We denote by $\mathscr{G}$ the family of all additive sets.

Remark 1.41 (a) Notice that, since μ^* is σ-subadditive, (1.7) is equivalent to

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \forall E \in \mathscr{P}(X). \tag{1.8}$$

[6] Let us observe that if $(a_{n,k})_{n,k}$ is a sequence of real numbers such that $a_{nk} \geq 0$, then

$$\sum_{(n,k)\in\mathbb{N}^2} a_{n,k} = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} a_{n,k} = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} a_{n,k}.$$

(b) Since identity (1.7) is symmetric with respect to the exchange $A \leftrightarrow A^c$, we deduce that $A^c \in \mathscr{G}$ for any $A \in \mathscr{G}$.

Theorem 1.42 (Carathéodory) *Let μ^* be an outer measure on X. Then $\mathscr{G}$ is a σ-algebra in X, and μ^* is a measure on $\mathscr{G}$.*

Before proving Carathéodory's Theorem, let us use it to complete the proof of Theorem 1.32.

Proof of Theorem 1.32: existence Given a σ-additive function μ on an algebra $\mathscr{A}$, define the outer measure μ^* as in Proposition 1.38. Then $\mu^*(A) = \mu(A)$ for any $A \in \mathscr{A}$. Moreover, in light of Theorem 1.42, μ^* is a measure on the σ-algebra $\mathscr{G}$ of additive sets. So the proof will be complete if we show that $\mathscr{A} \subset \mathscr{G}$. Indeed, in this case, $\sigma(\mathscr{A})$ turns out to be contained in $\mathscr{G}$, and it suffices to take the restriction of μ^* to $\sigma(\mathscr{A})$ to obtain the required extension.

Now, let $A \in \mathscr{A}$ and $E \in \mathscr{P}(X)$. Assume $\mu^*(E) < \infty$ (otherwise (1.8) trivially holds), and fix $\varepsilon > 0$. Then there exists $(A_n)_n \subset \mathscr{A}$ such that $E \subset \bigcup_{n=1}^{\infty} A_n$ and

$$\begin{aligned}\mu^*(E) + \varepsilon > \sum_{n=1}^{\infty} \mu(A_n) &= \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \cap A^c) \\ &\geq \mu^*(E \cap A) + \mu^*(E \cap A^c).\end{aligned}$$

Since ε is arbitrary, we have $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Thus, by Remark 1.41(a) we deduce that $A \in \mathscr{G}$. □

We now proceed with the proof of Carathéodory's Theorem.

Proof of Theorem 1.42 We will split the proof into four steps.

1. $\mathscr{G}$ is an algebra.
   We note that $\varnothing$ and X belong to $\mathscr{G}$. In view of Remark 1.41(b) we already know that $A \in \mathscr{G}$ implies $A^c \in \mathscr{G}$. Let us now prove that if $A, B \in \mathscr{G}$, then $A \cup B \in \mathscr{G}$. For any $E \in \mathscr{P}(X)$ we have

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \\ &= \big(\mu^*(E \cap A) + \mu^*(E \cap A^c \cap B)\big) + \mu^*(E \cap (A \cup B)^c).\end{aligned} \tag{1.9}$$

   Since

$$(E \cap A) \cup (E \cap A^c \cap B) = E \cap (A \cup B),$$

   the subadditivity of μ^* implies that

$$\mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) \geq \mu^*(E \cap (A \cup B)).$$

So, by (1.9),

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c),$$

and $A \cup B \in \mathscr{G}$ as required.

2. μ^* is additive on $\mathscr{G}$.
Let us prove that if $A, B \in \mathscr{G}$ and $A \cap B = \varnothing$, then

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap A) + \mu^*(E \cap B) \quad \forall E \in \mathscr{P}(X). \tag{1.10}$$

Indeed, replacing E with $E \cap (A \cup B)$ in (1.7) we obtain

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap (A \cup B) \cap A) + \mu^*(E \cap (A \cup B) \cap A^c),$$

which is equivalent to (1.10) since $A \cap B = \varnothing$. In particular, taking $E = X$, it follows that μ^* is additive on $\mathscr{G}$.

3. $\mathscr{G}$ is a σ-algebra.
Let $(A_k)_k \subset \mathscr{G}$ be a sequence of mutually disjoint sets. We will show that $S := \bigcup_{k=1}^{\infty} A_k \in \mathscr{G}$. To this aim, set $S_n := \bigcup_{k=1}^{n} A_k$, $n \in \mathbb{N}$. By the σ-subadditivity of μ^*, for any $n \in \mathbb{N}$ we have

$$\begin{aligned}
\mu^*(E \cap S) + \mu^*(E \cap S^c) &\leq \sum_{k=1}^{\infty} \mu^*(E \cap A_k) + \mu^*(E \cap S^c) \\
&= \lim_{n\to\infty} \left(\sum_{k=1}^{n} \mu^*(E \cap A_k) + \mu^*(E \cap S^c) \right) \\
&= \lim_{n\to\infty} \big(\mu^*(E \cap S_n) + \mu^*(E \cap S^c) \big)
\end{aligned}$$

in view of (1.10). Since $S^c \subset S_n^c$, it follows that

$$\mu^*(E \cap S) + \mu^*(E \cap S^c) \leq \limsup_{n\to\infty} \big(\mu^*(E \cap S_n) + \mu^*(E \cap S_n^c) \big) = \mu^*(E).$$

Therefore $S \in \mathscr{G}$, and then, since $\mathscr{G}$ is an algebra, we deduce that $\mathscr{G}$ is a σ-algebra (see Exercise 1.5).

4. μ^* is σ-additive on $\mathscr{G}$.
Since μ^* is σ-subadditive and additive by Step 2, then Remark 1.14(5) gives the conclusion. □

Remark 1.43 Let us observe that any set with outer measure zero is additive. Indeed, for any $Z \in \mathscr{P}(X)$ with $\mu^*(Z) = 0$, and any $E \in \mathscr{P}(X)$, we have

$$\mu^*(E \cap Z) + \mu^*(E \cap Z^c) = \mu^*(E \cap Z^c) \leq \mu^*(E)$$

by the monotonicity of μ^*. Thus, $Z \in \mathscr{G}$. We deduce that the measure μ^* is *complete* on the σ-algebra $\mathscr{G}$ (see Definition 1.22).

Remark 1.44 Given a σ-additive function μ on an algebra $\mathscr{A}$, the σ-algebra $\mathscr{G}$ of all additive sets with respect to the outer measure μ^* defined in Proposition 1.38 satisfies the inclusions

$$\sigma(\mathscr{A}) \subset \mathscr{G} \subset \mathscr{P}(X). \tag{1.11}$$

We shall see later that the above inclusions are both strict, in general.

1.4 Borel Measures on $\mathbb{R}^N$

Definition 1.45 Let (X, d) be a metric space. A measure μ on $\mathscr{B}(X)$ is called a *Borel measure*. A Borel measure μ is called a *Radon measure* if $\mu(K) < \infty$ for every compact set $K \subset X$.

In this section we will study specific properties of Borel measures on $\mathbb{R}^N$. We begin by introducing the Lebesgue measure on the unit interval.

1.4.1 Lebesgue Measure on [0, 1)

Let $\mathscr{I}$ be the class of all half-closed intervals $[a, b)$ with $0 \le a \le b < 1$, and let $\mathscr{A}_0$ be the algebra of all finite disjoint unions of elements of $\mathscr{I}$ (see Example 1.6(2)). Then $\sigma(\mathscr{I}) = \sigma(\mathscr{A}_0) = \mathscr{B}([0, 1))$.

On $\mathscr{I}$, consider the set function

$$m([a, b)) := b - a, \quad 0 \le a \le b < 1. \tag{1.12}$$

If $a = b$, then $[a, b)$ reduces to the empty set, and we have $m([a, b)) = 0$.

Exercise 1.46 Let $[a, b)$ be contained in $[a_1, b_1) \cup \cdots \cup [a_n, b_n)$, with $-\infty < a \le b < \infty$ and $-\infty < a_i \le b_i < \infty$. Prove that

$$b - a \le \sum_{i=1}^{n} (b_i - a_i).$$

Proposition 1.47 *The set function m defined in* (1.12) *is σ-additive on $\mathscr{I}$, i.e., for any sequence $(I_k)_k$ of mutually disjoint sets in $\mathscr{I}$ such that $\cup_{k=1}^{\infty} I_k \in \mathscr{I}$, we have:*

$$m\left(\bigcup_{k=1}^{\infty} I_k\right) = \sum_{k=1}^{\infty} m(I_k).$$

Proof Let $(I_k)_k$ be a disjoint sequence in $\mathscr{I}$, with $I_k = [a_k, b_k)$, and suppose $I = [a_0, b_0) = \cup_{k=1}^{\infty} I_k \in \mathscr{I}$. Then, for any $n \in \mathbb{N}$, we have

$$\sum_{k=1}^{n} m(I_k) = \sum_{k=1}^{n} (b_k - a_k) \le b_0 - a_0 = m(I).$$

Therefore

$$\sum_{k=1}^{\infty} m(I_k) \le m(I).$$

To prove the reverse inequality, assume $a_0 < b_0$. For any $\varepsilon < b_0 - a_0$ we have

$$[a_0, b_0 - \varepsilon] \subset \bigcup_{k=1}^{\infty} \left(a_k - \varepsilon 2^{-k}, b_k\right).$$

Then the Heine–Borel Theorem implies that, for some $k_0 \in \mathbb{N}$,

$$[a_0, b_0 - \varepsilon) \subset [a_0, b_0 - \varepsilon] \subset \bigcup_{k=1}^{k_0} \left(a_k - \varepsilon 2^{-k}, b_k\right).$$

Consequently, thanks to the result in Exercise 1.46,

$$m(I) - \varepsilon = (b_0 - a_0) - \varepsilon \le \sum_{k=1}^{k_0} \left(b_k - a_k + \varepsilon 2^{-k}\right) \le \sum_{k=1}^{\infty} m(I_k) + \varepsilon.$$

The arbitrariness of ε gives

$$m(I) \le \sum_{k=1}^{\infty} m(I_k).$$

□

We now proceed to extend m to $\mathscr{A}_0$. For any set $A \in \mathscr{A}_0$ such that $A = \cup_{i=1}^{k} I_i$, where $I_1, \ldots, I_k$ are disjoint sets in $\mathscr{I}$, let us define

$$m(A) := \sum_{i=1}^{k} m(I_i). \tag{1.13}$$

It is easy to see that the above definition is independent of the representation of A as a finite disjoint union of elements of $\mathscr{I}$.

Exercise 1.48 Show that if $J_1, \ldots, J_h$ is another family of disjoint sets in $\mathscr{I}$ such that $A = \cup_{j=1}^{h} J_j$, then

$$\sum_{i=1}^{k} m(I_i) = \sum_{j=1}^{h} m(J_j).$$

Theorem 1.49 *m is σ-additive on $\mathscr{A}_0$.*

Proof Let $(A_n)_n \subset \mathscr{A}_0$ be a sequence of disjoint sets such that

$$A := \bigcup_{n=1}^{\infty} A_n \in \mathscr{A}_0.$$

Then

$$A = \bigcup_{i=1}^{k} I_i \qquad A_n = \bigcup_{j=1}^{k_n} I_{n,j} \qquad (\forall n \in \mathbb{N})$$

for some disjoint sets $I_1, \ldots, I_n$ and $I_{n,1}, \ldots, I_{n,k_n}$ in $\mathscr{I}$. Now, observe that, for any $i = 1, \ldots, k$,

$$I_i = I_i \cap A = \bigcup_{n=1}^{\infty} (I_i \cap A_n) = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{k_n} (I_i \cap I_{n,j}),$$

and, since $(I_i \cap I_{n,j})_{n,j}$ is a countable family of disjoint sets in $\mathscr{I}$, by applying Proposition 1.47 we obtain

$$m(I_i) = \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} m(I_i \cap I_{n,j}).$$

Hence,

$$m(A) = \sum_{i=1}^{k} m(I_i) = \sum_{i=1}^{k} \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} m(I_i \cap I_{n,j}) = \sum_{n=1}^{\infty} \sum_{i=1}^{k} \sum_{j=1}^{k_n} m(I_i \cap I_{n,j}).$$

Since disjoint union $\cup_{i=1}^{k} \cup_{j=1}^{k_n} I_i \cap I_{n,j}$ equals A_n, by definition (1.13) we get $m(A_n) = \sum_{i=1}^{k} \sum_{j=1}^{k_n} m(I_i \cap I_{n,j})$. □

Summing up, thanks to Theorem 1.32, we conclude that m can be uniquely extended to a measure on the σ-algebra $\mathscr{B}([0, 1))$. Such an extension is called the *Lebesgue measure* on $[0, 1)$.

1.4.2 Lebesgue Measure on $\mathbb{R}$

We now turn to the construction of the Lebesgue measure on $\mathbb{R}$. Usually, this is done by an intrinsic procedure, applying an extension result for σ-additive set functions on *half-rings*. In this book, we will follow a shortcut, based on the following simple observations; for a different approach we refer to [Br83, KF75, Ru74, Ru64, Wi62, WZ77], for instance.

Proceeding as in the previous section, one can define the Lebesgue measure on $[a, b)$ for any interval $[a, b) \subset \mathbb{R}$. Such a measure will be denoted by $m_{[a,b)}$. Let us begin by characterizing the associated Borel sets in $[a, b)$. The following general result holds.

Proposition 1.50 *Given $A \in \mathscr{B}(\mathbb{R}^N)$, then*

$$\mathscr{B}(A) = \{B \in \mathscr{B}(\mathbb{R}^N) \mid B \subset A\}.$$

Proof Consider the class $\mathscr{E} := \mathscr{B}(A) \cap \mathscr{B}(\mathbb{R}^N)$. It is immediate that $\mathscr{E}$ is a σ-algebra in A. Since $\mathscr{E}$ contains all the subsets of A which are open in the relative topology, we conclude that $\mathscr{B}(A) \subset \mathscr{E}$. This proves the inclusion $\mathscr{B}(A) \subset \mathscr{B}(\mathbb{R}^N)$.

Next, to prove the opposite inclusion, let $\mathscr{F} := \{B \in \mathscr{B}(\mathbb{R}^N) \mid B \cap A \in \mathscr{B}(A)\}$. Let us check that $\mathscr{F}$ is a σ-algebra in $\mathbb{R}^N$.

1. $\varnothing, \mathbb{R}^N \in \mathscr{F}$ by definition.
2. Let $B \in \mathscr{F}$. Since $B \cap A \in \mathscr{B}(A)$, we have $B^c \cap A = A \setminus (B \cap A) \in \mathscr{B}(A)$. Therefore $B^c \in \mathscr{F}$.
3. Let $(B_n)_n \subset \mathscr{F}$. Then $(\cup_{n=1}^{\infty} B_n) \cap A = \cup_{n=1}^{\infty}(B_n \cap A) \in \mathscr{B}(A)$. Therefore $\cup_{n=1}^{\infty} B_n \in \mathscr{F}$.

Since $\mathscr{F}$ contains all open sets in $\mathbb{R}^N$, we conclude that $\mathscr{B}(\mathbb{R}^N) \subset \mathscr{F}$. The proof is thus complete. □

Thus, for any pair of nested intervals $[a, b) \subset [c, d) \subset \mathbb{R}$, we have that $\mathscr{B}([a, b)) \subset \mathscr{B}([c, d))$. Moreover, a unique extension argument yields

$$m_{[a,b)}(A) = m_{[c,d)}(A) \qquad \forall A \in \mathscr{B}([a, b)). \tag{1.14}$$

Now, since $\mathbb{R} = \bigcup_{k=1}^{\infty} [-k, k)$, it is natural to define the Lebesgue measure on $\mathbb{R}$ as

$$m(A) := \lim_{k\to\infty} m_{[-k,k)}(A \cap [-k, k)) \qquad \forall A \in \mathscr{B}(\mathbb{R}). \tag{1.15}$$

Let us observe that, in view of (1.14), we have

$$\begin{aligned} m_{[-k,k)}(A \cap [-k, k)) &= m_{[-k-1,k+1)}(A \cap [-k, k)) \\ &\le m_{[-k-1,k+1)}(A \cap [-k-1, k+1)), \end{aligned}$$

by which we deduce that the function $k \mapsto m_{[-k,k)}(A \cap [-k, k))$ is nondecreasing; therefore for any $A \in \mathscr{B}(\mathbb{R})$ the limit in (1.15) is well defined (possibly infinite).

Our next exercise is intended to show that the definition of m would be the same if we took any other sequence of intervals covering $\mathbb{R}$.

Exercise 1.51 Let $(a_k)_k$ and $(b_k)_k$ be real sequences satisfying

$$a_k < b_k, \qquad a_k \downarrow -\infty, \qquad b_k \uparrow \infty.$$

Show that

$$m(A) = \lim_{k\to\infty} m_{[a_k,b_k)}(A \cap [a_k, b_k)) \qquad \forall A \in \mathscr{B}(\mathbb{R}).$$

In order to show that m is a measure on $\mathscr{B}(\mathbb{R})$, we still have to check σ-additivity.

Proposition 1.52 *The set function defined in* (1.15) *is σ-additive on $\mathscr{B}(\mathbb{R})$.*

Proof Let us first show that m is additive on $\mathscr{B}(\mathbb{R})$. Indeed, let $A_1, \ldots, A_n \in \mathscr{B}(\mathbb{R})$ be disjoint sets and let $A = \cup_{i=1}^n A_i$. Then, by the additivity of $m_{[-k,k)}$,

$$\begin{aligned} m(A) &= \lim_{k\to\infty} m_{[-k,k)}(A \cap [-k, k)) = \lim_{k\to\infty} \sum_{i=1}^n m_{[-k,k)}(A_i \cap [-k, k)) \\ &= \sum_{i=1}^n \lim_{k\to\infty} m_{[-k,k)}(A_i \cap [-k, k)) = \sum_{i=1}^n m(A_i). \end{aligned}$$

Now, let $(B_n)_n \subset \mathscr{B}(\mathbb{R})$ be a sequence of sets and let $B = \cup_{n=1}^\infty B_n$. Then, using the σ-subadditivity of $m_{[-k,k)}$,

$$\begin{aligned} m(B) &= \lim_{k\to\infty} m_{[-k,k)}(B \cap [-k, k)) \le \lim_{k\to\infty} \sum_{n=1}^\infty m_{[-k,k)}(B_n \cap [-k, k)) \\ &\le \sum_{n=1}^\infty m(B_n), \end{aligned}$$

since $m_{[-k,k)}(B_n \cap [-k, k)) \le m(B_n)$ for every n, k. This proves that m is σ-subadditive, and, consequently, σ-additive in view of Remark 1.14(5). □

Since m is bounded on bounded sets, the Lebesgue measure on $\mathbb{R}$ is a Radon measure. Another interesting property is *translation invariance*.

Proposition 1.53 *Let $A \in \mathscr{B}(\mathbb{R})$. Then, for every $x \in \mathbb{R}$,*

$$A + x := \{a + x \mid a \in A\} \in \mathscr{B}(\mathbb{R}), \tag{1.16}$$

$$m(A + x) = m(A). \tag{1.17}$$

Proof Define, for any $x \in \mathbb{R}$,

$$\mathscr{E}_x = \{A \in \mathscr{P}(\mathbb{R}) \mid A + x \in \mathscr{B}(\mathbb{R})\}.$$

Let us check that $\mathscr{E}_x$ is a σ-algebra in $\mathbb{R}$.

1. $\varnothing, \mathbb{R} \in \mathscr{E}_x$ by direct inspection.
2. Let $A \in \mathscr{E}_x$. Since $A^c + x = (A + x)^c \in \mathscr{B}(\mathbb{R})$, we deduce that $A^c \in \mathscr{E}_x$.
3. Let $(A_n)_n \subset \mathscr{E}_x$. Then $(\cup_{n=1}^{\infty} A_n)+x = \cup_{n=1}^{\infty}(A_n+x) \in \mathscr{B}(\mathbb{R})$. So $\cup_{n=1}^{\infty} A_n \in \mathscr{E}_x$.

Since $\mathscr{E}_x$ contains all open subsets in $\mathbb{R}$, $\mathscr{B}(\mathbb{R}) \subset \mathscr{E}_x$ for any $x \in \mathbb{R}$. This proves (1.16).

Let us prove (1.17). Fix $x \in \mathbb{R}$, and define

$$m_x(A) = m(A + x) \qquad \forall A \in \mathscr{B}(\mathbb{R}).$$

It is straightforward to check that m_x and m agree on the class

$$\mathscr{I}_{\mathbb{R}} := \big\{(-\infty, a) \mid -\infty < a \le \infty\big\} \bigcup \big\{[a, b) \mid -\infty < a \le b \le \infty\big\}.$$

Therefore m_x and m also agree on the algebra $\mathscr{A}_{\mathbb{R}}$ of all finite disjoint unions of elements of $\mathscr{I}_{\mathbb{R}}$. Since $\sigma(\mathscr{A}_{\mathbb{R}}) = \mathscr{B}(\mathbb{R})$, by the uniqueness result of Theorem 1.32 we conclude that $m_x(A) = m(A)$ for any $A \in \mathscr{B}(\mathbb{R})$. □

Exercise 1.54 Show that the set $\mathbb{Q}$ of all rational points in $\mathbb{R}$ is a Borel set, with Lebesgue measure zero.

1.4.3 Lebesgue Measure on $\mathbb{R}^N$

In Sects. 1.4.1 and 1.4.2 we constructed Lebesgue measure on $\mathbb{R}$, starting from a σ-additive function defined on the algebra of all finite disjoint unions of half-closed intervals $[a, b) \subset [0, 1)$. The same construction can be carried out, with few changes, in the case of a generic euclidean space $\mathbb{R}^N$ ($N \ge 1$), leading to the definition of the *Lebesgue measure m* on $\mathbb{R}^N$. More precisely, the half-closed intervals we used in the case $N = 1$ are now replaced by *half-closed N-dimensional rectangles* of the form

$$R = \prod_{i=1}^{N}[a_i, b_i) = \big\{(x_1, \dots, x_N) \,\big|\, a_i \le x_i < b_i,\ i = 1, \dots, N\big\}$$

where $a_i \le b_i$, $i = 1, \dots, N$. If the edge lengths $b_i - a_i$ are all equal, R is called a *N-dimensional half-closed cube*. Cubes will usually be denoted by the letter Q. By definition, the *Lebesgue measure* of a rectangle $R = \prod_{i=1}^{N}[a_i, b_i)$ is

$$m(R) = \prod_{i=1}^{N}(b_i - a_i).$$

Proceeding as in the previous sections, starting from the set function m defined on the class of all N-dimensional half-closed rectangles contained in the cube $[0, 1)^N :=$

$[0, 1) \times \cdots \times [0, 1)$, we can extend m by additivity to the algebra of all finite disjoint unions of such rectangles. Finally, using Theorem 1.32, we extend m to a measure on $\mathscr{B}([0, 1)^N)$, called the Lebesgue measure on $[0, 1)^N$. Analogously one can define the Lebesgue measure on R for any rectangle $R \subset \mathbb{R}^N$. Such a measure will be denoted by m_R. Then the Lebesgue measure m on $\mathbb{R}^N$ is defined as

$$m(A) := \lim_{k\to\infty} m_{[-k,k)^N}(A \cap [-k, k)^N) \qquad \forall A \in \mathscr{B}(\mathbb{R}^N).$$

As for the case of $N = 1$, the Lebesgue measure on $\mathbb{R}^N$ is a Radon measure and is translation invariant, as stated in the following reformulation of Proposition 1.53.

Proposition 1.55 *Let $A \in \mathscr{B}(\mathbb{R}^N)$. Then, for any $x \in \mathbb{R}^N$,*

$$A + x := \{a + x \mid a \in A\} \in \mathscr{B}(\mathbb{R}^N),$$
$$m(A + x) = m(A).$$

Definition 1.56 The elements of the σ-algebra $\mathscr{G}$ of all additive sets in $\mathbb{R}^N$ (with respect to the outer measure m^* defined in Proposition 1.38) are called *Lebesgue measurable sets* in $\mathbb{R}^N$.

Remark 1.57 The Lebesgue measure, which was defined only for Borel sets, can be extended to the σ-algebra $\mathscr{G}$ of all Lebesgue measurable sets. Such an extension is given by $m^*(A)$ for any $A \in \mathscr{G}$. This new measure is complete (see Remark 1.43) and continues to be called the *Lebesgue measure* on $\mathbb{R}^N$.

In what follows we shall use the notion of cube to obtain a basic decomposition of open sets in $\mathbb{R}^N$. For every $n \in \mathbb{N}$ let $\mathscr{Q}_n$ be the collection of cubes

$$\mathscr{Q}_n = \left\{ \prod_{i=1}^{N} \left[\frac{a_i}{2^n}, \frac{a_i + 1}{2^n} \right) \middle| a_i \in \mathbb{Z} \right\}.$$

In other words, $\mathscr{Q}_0$ is the collection of cubes with edge length 1 and vertices at points with integer coordinates. Bisecting each edge of a cube in $\mathscr{Q}_0$, we obtain from it 2^N subcubes of edge length $\frac{1}{2}$. The total collection of these subcubes forms the collection $\mathscr{Q}_1$ of cubes. If we continue bisecting, we obtain finer and finer collections $\mathscr{Q}_n$ of cubes such that each cube in $\mathscr{Q}_n$ has edge length 2^{-n} and is the union of 2^N disjoint cubes in $\mathscr{Q}_{n+1}$.

Definition 1.58 The cubes of the collection

$$\{Q \mid Q \in \mathscr{Q}_n, \; n = 0, 1, 2, \dots\}$$

are called *dyadic cubes*.

Remark 1.59 Dyadic cubes have the following properties:

(a) $\mathbb{R}^N = \cup_{Q \in \mathcal{Q}_n} Q$ with disjoint union for every n.
(b) If $Q \in \mathcal{Q}_n$ and $P \in \mathcal{Q}_k$ with $k \le n$, then $Q \subset P$ or $P \cap Q = \emptyset$.
(c) If $Q \in \mathcal{Q}_n$, then $m(Q) = 2^{-nN}$.

Lemma 1.60 *Every open set in $\mathbb{R}^N$ can be written as a countable union of disjoint dyadic cubes.*

Proof Let V be an open nonempty set in $\mathbb{R}^N$. Let $\mathcal{S}_0$ be the collection of all cubes in $\mathcal{Q}_0$ which lie entirely in V. Let $\mathcal{S}_1$ be those cubes in $\mathcal{Q}_1$ which lie in V but which are not subcubes of any cube in $\mathcal{S}_0$. More generally, for $n \ge 1$, let $\mathcal{S}_n$ be the cubes in $\mathcal{Q}_n$ which lie in V but which are not subcubes of any cube in $\mathcal{S}_0, \mathcal{S}_1, \ldots, \mathcal{S}_{n-1}$. If $\mathcal{S}$ is the total collection of cubes from all $\mathcal{S}_n$, then $\mathcal{S}$ is countable since each $\mathcal{Q}_n$ is countable, and the cubes in $\mathcal{S}$ are nonoverlapping by construction. Moreover, since V is open and the cubes in $\mathcal{Q}_n$ become arbitrarily small as $n \to \infty$, then by Remark 1.59(a) each point of V will eventually be caught in a cube of some $\mathcal{S}_n$. Hence, $V = \cup_{Q \in \mathcal{S}} Q$ and the proof is complete. □

Remark 1.61 Owing to Lemma 1.60 the collection of all open sets in $\mathbb{R}^N$ has the cardinality of the continuum. We claim that $\mathcal{B}(\mathbb{R}^N)$ has also the cardinality of the continuum. This follows by observing that each set in $\mathcal{B}(\mathbb{R}^N)$ can be constructed by a countable number of operations, starting from the family of all open sets, each of these operations consisting of countable union, countable intersection or taking the complement.

1.4.4 Examples

In this section we shall construct three examples of sets that are hard to visualize but possess very interesting properties.

Example 1.62 (Two unusual Borel sets) Let $(r_n)_n$ be an enumeration of $\mathbb{Q} \cap [0, 1]$. Given $\varepsilon > 0$, set

$$V = \bigcup_{n=1}^{\infty} \left(r_n - \frac{\varepsilon}{2^n}, r_n + \frac{\varepsilon}{2^n} \right).$$

Then $V \cap [0, 1]$ is *open* (with respect to the relative topology of $[0, 1]$) and dense in $[0, 1]$. By σ-subadditivity, we have $0 < m(V \cap [0, 1]) < 2\varepsilon$. Moreover, the *compact* set $K := [0, 1] \setminus V$ has no interior and measure *nearly* 1.

Example 1.63 (Cantor triadic set) To begin with, let us note that any $x \in [0, 1]$ has a triadic expansion of the form

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \qquad a_i = 0, 1, 2. \tag{1.18}$$

Such a representation is not unique due to the presence of periodic expansions. We can, however, choose a unique representation of the form (1.18) by picking the expansion[7] with fewer digits equal to 1. Now, observe that the set

$$C_1 := \left\{ x \in [0,1] \,\middle|\, x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_1 \neq 1 \right\}$$

is obtained from [0, 1] by removing the *'middle third'* $(\frac{1}{3}, \frac{2}{3})$. Therefore C_1 is the union of two closed disjoint intervals, each of which has measure $\frac{1}{3}$. More generally, for any $n \in \mathbb{N}$ the set

$$C_n := \left\{ x \in [0,1] \,\middle|\, x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_1, \ldots, a_n \neq 1 \right\}$$

is the union of 2^n closed disjoint intervals, each of which has measure $\left(\frac{1}{3}\right)^n$. So

$$C_n \downarrow C := \left\{ x \in [0,1] \,\middle|\, x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_i \neq 1 \ \forall i \in \mathbb{N} \right\},$$

where C is the so-called *Cantor set*. C is a closed set by construction, with measure zero since

$$m(C) \leq m(C_n) \leq \left(\frac{2}{3}\right)^n \qquad \forall n \in \mathbb{N}.$$

Nevertheless, C is *uncountable*. Indeed, the function

$$f\left(\sum_{i=1}^{\infty} \frac{a_i}{3^i}\right) = \sum_{i=1}^{\infty} a_i 2^{-(i+1)} \tag{1.19}$$

maps C *onto* [0, 1].

Exercise 1.64 Show that $f : C \to [0,1]$ defined by (1.19) is onto.

Remark 1.65 Since the Cantor set has measure zero, and recalling that the Lebesgue measure m on the σ-algebra $\mathscr{G}$ (constituted by all Lebesgue measurable sets in $\mathbb{R}$) is complete (see Remark 1.57), any subset of C is Lebesgue measurable:

$$\mathscr{P}(C) \subset \mathscr{G}.$$

[7] For instance, we choose the second of the following two triadic expansions for $x = \frac{1}{3}$:

$$\frac{1}{3} = \frac{1}{3} + \frac{0}{3^2} + \frac{0}{3^3} + \frac{0}{3^4} + \cdots, \qquad \frac{1}{3} = \frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \frac{2}{3^4} + \frac{2}{3^5} + \cdots.$$

Therefore

$$\#\mathscr{P}(C) \le \#\mathscr{G}.$$

Since C in uncountable, we deduce that the power of $\mathscr{G}$ is strictly greater than the power of the continuum. Using Remark 1.61 we conclude that the inclusion $\mathscr{B}(\mathbb{R}) \subset \mathscr{G}$ is *strict*.

Example 1.66 (A non-measurable set) We shall now show that $\mathscr{G}$ is also strictly contained in $\mathscr{P}(\mathbb{R})$. In $[0, 1)$, we define x and y to be equivalent if $x - y \in \mathbb{Q}$. By the Axiom of Choice, there exists a set $P \subset [0, 1)$ such that P consists of exactly one representative point from each equivalent class. We claim that P provides the required example of a set which fails to be measurable. Indeed, consider the countable family $(P_n)_n \subset \mathscr{P}(\mathbb{R})$, where $P_n = P + r_n$ and $(r_n)_n$ is an enumeration of $\mathbb{Q} \cap (-1, 1)$. Observe the following.

1. $(P_n)_n$ is a disjoint family. Indeed, suppose that $p, q \in P$ are such that $p + r_n = q + r_m$ with $n \neq m$; we have $p - q \in \mathbb{Q}$ and $p - q = r_m - r_n \neq 0$. Then P contains two distinct equivalent points, in contradiction with the definition of P.
2. $[0, 1) \subset \cup_{n=1}^{\infty} P_n \subset [-1, 2)$. Indeed, let $x \in [0, 1)$. Since x is equivalent to some element of P, we have $x - p = r$ for some $p \in P$ and some $r \in \mathbb{Q}$ satisfying $|r| < 1$. Then $r = r_n$ for some $n \in \mathbb{N}$, whence $x \in P_n$. The other inclusion is immediate.

If P were Lebesgue measurable, by monotonicity and σ-additivity of m it would follow that $1 = m([0, 1)) \le \sum_{n=1}^{\infty} m(P_n) \le m([-1, 2)) = 3$. But this is impossible since $m(P_n) = m(P)$ for every n, and therefore the sum $\sum_{n=1}^{\infty} m(P_n)$ is either 0 or ∞.

Exercise 1.67 Let us consider the following subset of $[0, 1]$ constructed by a recursive argument. As first step, we divide the interval $[0, 1]$ into five identical subintervals and we remove the *'middle fifth'*. For each of the remaining four intervals, we repeat the same procedure, namely we divide it into five identical subintervals and we remove the middle fifth. After iterating the procedure infinitely many times, the remaining set is of *Cantor type*. Show that such a set has measure zero.

Exercise 1.68 Given $\alpha_n \in (0, 1)$, let us construct the following *Cantor type* set. First, we remove from $[0, 1]$ an open interval of length α_1. Next, from each of the two remaining intervals, we remove an open interval of relative length α_2. Next, from each of the four remaining intervals, we remove an open interval of relative length α_3, and so on. The remaining set is of Cantor type. Show that such a set has measure zero if and only if $\sum_{n=1}^{\infty} \alpha_n = \infty$.

1.4.5 Regularity of Radon Measures

The aim of this section is to prove regularity properties of a Radon measure on $\mathbb{R}^N$. We begin by studying finite measures.

Proposition 1.69 *Let μ be a finite Borel measure on $\mathbb{R}^N$. Then for any $A \in \mathscr{B}(\mathbb{R}^N)$*

$$\mu(A) = \sup\{\mu(F) \mid F \subset A,\ F \text{ closed}\} = \inf\{\mu(V) \mid V \supset A,\ V \text{ open}\}. \quad (1.20)$$

Proof Let us first observe that, since μ finite, an equivalent formulation of (1.20) is the following:

$$\forall \varepsilon > 0\ \ \exists V \text{open},\ F \text{ closed s.t. } F \subset A \subset V \text{ and } \mu(V \setminus F) < \varepsilon. \quad (1.21)$$

Let us consider the set

$$\mathscr{E} = \{A \in \mathscr{B}(\mathbb{R}^N) \mid A \text{ verifies (1.21)}\}.$$

It is enough to show that $\mathscr{E}$ is a σ-algebra in $\mathbb{R}^N$ including all open sets. Obviously, $\mathscr{E}$ contains $\mathbb{R}^N$ and $\varnothing$. Moreover, it is immediate that if $A \in \mathscr{E}$, then its complement A^c belongs to $\mathscr{E}$.

Let us now prove the implication $(A_n)_n \subset \mathscr{E} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathscr{E}$. Since $A_n \in \mathscr{E}$, for any $n \in \mathbb{N}$ there exist an open set V_n and a closed set F_n such that

$$F_n \subset A_n \subset V_n, \qquad \mu(V_n \setminus F_n) \leq \frac{\varepsilon}{2^{n+1}}.$$

Now, define $V = \bigcup_{n=1}^{\infty} V_n$ and $S = \bigcup_{n=1}^{\infty} F_n$; we have $S \subset \bigcup_{n=1}^{\infty} A_n \subset V$ and, by σ-subadditivity,

$$\mu(V \setminus S) \leq \sum_{n=1}^{\infty} \mu(V_n - S) \leq \sum_{n=1}^{\infty} \mu(V_n - F_n) \leq \frac{\varepsilon}{2}.$$

However, V is open but S is not necessarily closed. To overcome this problem, let us approximate S by the sequence $S_n = \bigcup_{k=1}^{n} F_k$. For any $n \in \mathbb{N}$, S_n is obviously closed; moreover $S_n \uparrow S$ and so, by Proposition 1.17, $\mu(S_n) \uparrow \mu(S)$. Therefore there exists $n_\varepsilon \in \mathbb{N}$ such that $\mu(S \setminus S_{n_\varepsilon}) < \frac{\varepsilon}{2}$. The set $F := S_{n_\varepsilon}$ satisfies $F \subset \bigcup_{n=1}^{\infty} A_n \subset V$ and $\mu(V \setminus F) = \mu(V \setminus S) + \mu(S \setminus F) < \varepsilon$, by which $\bigcup_{n=1}^{\infty} A_n \in \mathscr{E}$. We have thus proved that $\mathscr{E}$ is a σ-algebra.

There remains to show that $\mathscr{E}$ contains all the open sets in $\mathbb{R}^N$. For this, let V be open, and set

$$F_n = \left\{x \in \mathbb{R}^N \;\middle|\; d_{V^c}(x) \geq \frac{1}{n}\right\},$$

where $d_{V^c}(x)$ is the distance of x from V^c. Since d_{V^c} is a continuous function, F_n is a closed set in $\mathbb{R}^N$ (see Appendix A). Moreover $F_n \uparrow V$. So, recalling that μ is finite, by applying Proposition 1.18 we conclude that $\mu(V \setminus F_n) \downarrow 0$. □

The following result is a straightforward consequence of Proposition 1.69.

Corollary 1.70 *Let μ and ν be finite Borel measures on $\mathbb{R}^N$ such that $\mu(F) = \nu(F)$ for any closed set F in $\mathbb{R}^N$. Then $\mu = \nu$.*

Now we will extend Proposition 1.69 to Radon measures.

Theorem 1.71 *Let μ be a Radon measure on $\mathbb{R}^N$, and let A be a Borel set. Then*

$$\mu(A) = \inf\{\mu(V) \mid V \supset A,\ V \text{ open}\}, \tag{1.22}$$

$$\mu(A) = \sup\{\mu(K) \mid K \subset A,\ K \text{ compact}\}. \tag{1.23}$$

Proof Since (1.22) is trivial if $\mu(A) = \infty$, we shall first assume that $\mu(A) < \infty$. For any $n \in \mathbb{N}$, denote by Q_n the cube $(-n, n)^N$, and consider the finite measures[8] $\mu \llcorner Q_n$. Fix $\varepsilon > 0$ and apply Proposition 1.69 to deduce that, for any $n \in \mathbb{N}$, there exists an open set $V_n \supset A$ such that

$$(\mu \llcorner Q_n)(V_n \setminus A) < \frac{\varepsilon}{2^n}.$$

Now, set $V := \cup_{n=1}^{\infty}(V_n \cap Q_n) \supset A$. V is obviously an open set and

$$\mu(V \setminus A) \le \sum_{n=1}^{\infty} \mu\big((V_n \cap Q_n) \setminus A\big) = \sum_{n=1}^{\infty} (\mu \llcorner Q_n)(V_n \setminus A) < \varepsilon,$$

which in turn implies (1.22).

Next, let us prove (1.23) for $\mu(A) < \infty$. Fix $\varepsilon > 0$ and apply Proposition 1.69 to the finite measures $\mu \llcorner \overline{Q}_n$ to obtain, for any $n \in \mathbb{N}$, a closed set $F_n \subset A$ satisfying

$$(\mu \llcorner \overline{Q}_n)(A \setminus F_n) < \varepsilon.$$

Consider the sequence of compact sets $K_n = F_n \cap \overline{Q}_n$. Since

$$\mu(A \cap \overline{Q}_n) \uparrow \mu(A),$$

for some $n_\varepsilon \in \mathbb{N}$ we have $\mu(A \cap \overline{Q}_{n_\varepsilon}) > \mu(A) - \varepsilon$. Therefore

$$\begin{aligned}
\mu(A \setminus K_{n_\varepsilon}) &= \mu(A) - \mu(K_{n_\varepsilon}) \\
&< \mu(A \cap \overline{Q}_{n_\varepsilon}) - \mu(F_{n_\varepsilon} \cap \overline{Q}_{n_\varepsilon}) + \varepsilon \\
&= (\mu \llcorner \overline{Q}_{n_\varepsilon})(A \setminus F_{n_\varepsilon}) + \varepsilon < 2\varepsilon.
\end{aligned}$$

If $\mu(A) = \infty$, then $A_n := A \cap Q_n \uparrow A$, and so $\mu(A_n) \to \infty$. Since $\mu(A_n) < \infty$, for any n there exists a compact set K_n such that $K_n \subset A_n$ and $\mu(K_n) > \mu(A_n) - 1$, by which $K_n \subset A$ and $\mu(K_n) \to \infty = \mu(A)$. □

[8] See Definition 1.26.

Remark 1.72 Properties (1.22) and (1.23) are called *external* and *internal regularity* of Radon measures on $\mathbb{R}^N$, respectively.

Exercise 1.73 Any Radon measure μ on $\mathbb{R}^N$ is clearly σ-finite. Conversely, is a σ-finite Borel measure on $\mathbb{R}^N$ necessarily Radon?
Hint. Consider $\mu = \sum_{n=1}^{\infty} \delta_{1/n}$ on $\mathscr{B}(\mathbb{R})$, where $\delta_{1/n}$ is the Dirac measure at $1/n$.

Exercise 1.74 Let μ be a Radon measure on $\mathbb{R}^N$.

- Show that if $K \subset \mathbb{R}^N$ is a compact set, then the function $f : x \in \mathbb{R}^N \mapsto \mu(K + x) \in \mathbb{R}$ is upper semicontinuous (see Appendix B). Give an example to show that f fails to be continuous, in general.
- Show that if $V \subset \mathbb{R}^N$ is an open set, then the function $f : x \in \mathbb{R}^N \mapsto \mu(V + x) \in [0, \infty]$ is lower semicontinuous. Give an example to show that f fails to be continuous, in general.

Next proposition characterizes all Radon measures having the property of translation invariance.

Proposition 1.75 *Let μ be a Radon measure on $\mathbb{R}^N$ such that μ is translation invariant, that is,*

$$\mu(A + x) = \mu(A) \quad \forall A \in \mathscr{B}(\mathbb{R}^N), \quad \forall x \in \mathbb{R}^N.$$

Then there exists $c \geq 0$ such that $\mu(A) = c\, m(A)$ for any $A \in \mathscr{B}(\mathbb{R}^N)$.

Proof Given $n \in \mathbb{N}$, by construction we have that $[0, 1)^N$ is the union of 2^{nN} disjoint dyadic cubes belonging to the collection $\mathscr{Q}_n$, and these cubes are identical up to a translation. Setting $c = \mu([0, 1)^N)$, and using the translation invariance of μ and m, for every $Q \in \mathscr{Q}_n$ we have

$$2^{nN} \mu(Q) = \mu([0, 1)^N) = c\, m([0, 1)^N) = 2^{nN} c\, m(Q).$$

Then μ and $c\, m$ coincide on the dyadic cubes. In view of Lemma 1.60, by σ-additivity we have that μ and $c\, m$ coincide on all open sets; finally, by (1.22), it follows that $\mu(A) = c\, m(A)$ for any $A \in \mathscr{B}(\mathbb{R}^N)$. □

Next theorem shows how the Lebesgue measure changes under nonsingular linear transformations.

Theorem 1.76 *Let $T : \mathbb{R}^N \to \mathbb{R}^N$ be a linear nonsingular transformation. Then*

(i) $T(A) \in \mathscr{B}(\mathbb{R}^N)$ *for any* $A \in \mathscr{B}(\mathbb{R}^N)$.
(ii) $m(T(A)) = |\det T|\, m(A)$ *for any* $A \in \mathscr{B}(\mathbb{R}^N)$.

Proof Consider the family

$$\mathscr{E} = \{A \in \mathscr{B}(\mathbb{R}^N) \mid T(A) \in \mathscr{B}(\mathbb{R}^N)\}.$$

Since T is nonsingular, then $T(\emptyset) = \emptyset$, $T(\mathbb{R}^N) = \mathbb{R}^N$, $T(E^c) = (T(E))^c$, $T(\cup_{n=1}^{\infty} E_n) = \cup_{n=1}^{\infty} T(E_n)$ for all $E, E_n \subset \mathbb{R}^N$. Hence $\mathscr{E}$ is a σ-algebra. Furthermore T maps open sets into open sets; so $\mathscr{E} = \mathscr{B}(\mathbb{R}^N)$ and (i) follows.

Next define

$$\mu(A) = m(T(A)) \quad \forall A \in \mathscr{B}(\mathbb{R}^N).$$

Since T maps compact sets into compact sets, we deduce that μ is a Radon measure. Moreover, if $A \in \mathscr{B}(\mathbb{R}^N)$ and $x \in \mathbb{R}^N$, since m is translation invariant, we have

$$\mu(A + x) = m(T(A + x)) = m(T(A) + T(x)) = m(T(A)) = \mu(A),$$

and so μ is also translation invariant. Proposition 1.75 implies that there exists $\Delta(T) \geq 0$ such that

$$\mu(A) = \Delta(T)m(A) \quad \forall A \in \mathscr{B}(\mathbb{R}^N). \tag{1.24}$$

It remains to show that $\Delta(T) = |\det T|$. To prove this, let $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ denote the standard basis in $\mathbb{R}^N$, i.e., $\mathbf{e}_i$ has the jth coordinate equal to 1 if $j = i$ and equal to 0 if $j \neq i$. We first consider the following elementary transformations:

(a) There exist $i \neq j$ such that $T(\mathbf{e}_i) = \mathbf{e}_j$, $T(\mathbf{e}_j) = \mathbf{e}_i$ and $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i, j$. In this case $T([0, 1)^N) = [0, 1)^N$ and $\det T = -1$. By taking $A = [0, 1)^N$ in (1.24), we deduce $\Delta(T) = 1 = |\det T|$.

(b) There exist $\alpha \neq 0$ and i such that $T(\mathbf{e}_i) = \alpha\mathbf{e}_i$ and $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i$. Assume $i = 1$. Then $T([0, 1)^N) = [0, \alpha) \times [0, 1)^{N-1}$ if $\alpha > 0$ and $T([0, 1)^N) = (\alpha, 0] \times [0, 1)^{N-1}$ if $\alpha < 0$. Therefore, by taking $A = [0, 1)^N$ in (1.24), we obtain $\Delta(T) = m(T([0, 1)^N)) = |\alpha| = |\det T|$.

(c) There exist $i \neq j$ and $\alpha \neq 0$ such that $T(\mathbf{e}_i) = \mathbf{e}_i + \alpha\mathbf{e}_j$, $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i$. Assume $i = 1$ and $j = 2$ and set $R_\alpha = \{(x_1, \alpha x_2, x_3, \ldots, x_N) \,|\, 0 \leq x_i < 1\}$. Then we have

$$\begin{aligned}
T(R_\alpha) &= \left\{(x_1, \alpha(x_1 + x_2), x_3, \ldots, x_N) \,\middle|\, 0 \leq x_i < 1\right\} \\
&= \left\{(\xi_1, \alpha\xi_2, \xi_3, \ldots, \xi_N) \,\middle|\, \xi_1 \leq \xi_2 < \xi_1 + 1,\ 0 \leq \xi_i < 1 \text{ for } i \neq 2\right\} \\
&= E_1 \cup E_2
\end{aligned}$$

with disjoint union, where

$$\begin{aligned}
E_1 &= \left\{(\xi_1, \alpha\xi_2, \xi_3, \ldots, \xi_N) \,\middle|\, \xi_1 \leq \xi_2 < 1,\ 0 \leq \xi_i < 1 \text{ for } i \neq 2\right\}, \\
E_2 &= \left\{(\xi_1, \alpha\xi_2, \xi_3, \ldots, \xi_N) \,\middle|\, 1 \leq \xi_2 < \xi_1 + 1,\ 0 \leq \xi_i < 1 \text{ for } i \neq 2\right\}.
\end{aligned}$$

Observe that $E_1 \subset R_\alpha$ and $E_2 - \alpha\mathbf{e}_2 = R_\alpha \setminus E_1$; then

$$m(T(R_\alpha)) = m(E_1) + m(E_2) = m(E_1) + m\big(E_2 - \alpha\mathbf{e}_2\big) = m(R_\alpha).$$

By taking $A = R_\alpha$ in (1.24), we deduce $\Delta(T) = 1 = |\det T|$.

If $T = T_1 \cdot \ldots \cdot T_k$ with T_i elementary transformations of type (a)–(c), since $\Delta(T) = \Delta(T_1) \cdot \ldots \cdot \Delta(T_k)$ by (1.24), we have

$$\Delta(T) = |\det T_1| \cdot \ldots \cdot |\det T_k| = |\det T|.$$

Therefore the thesis will follow once we have proved the following claim: any non-singular linear transformation is the product of elementary transformations of type (a)–(c). We proceed by induction on the dimension N. The claim is trivially true for $N = 1$; assume that the claim holds for $N - 1$ and we pass to prove it for N. Set $T = (a_{i,j})_{i,j=1,\ldots,N}$, i.e.,

$$T(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \quad i = 1, \ldots, N.$$

For $k = 1, \ldots, N$, consider $T_k = (a_{i,j})_{j=1,\ldots,N-1,\, i=1,\ldots,N,\, i\neq k}$. Since $\det T = \sum_{k=1}^{N}(-1)^{k+N} a_{kN} \det T_k$, possibly exchanging two variables by a transformation of type (a), we may assume $\det T_N \neq 0$. Then, by induction, the transformation $S_1 : \mathbb{R}^N \to \mathbb{R}^N$ defined as

$$S_1(\mathbf{e}_i) = T_N(\mathbf{e}_i) = \sum_{j=1}^{N-1} a_{ij}\mathbf{e}_j \;\; i = 1, \ldots, N-1, \quad S_1(\mathbf{e}_N) = \mathbf{e}_N$$

is the product of elementary transformations. By applying $N - 1$ transformations of type (c) with triplets (i, j, α) equal to $(1, N, a_{1N}), \ldots, (N-1, N, a_{N-1,N})$ we arrive at $S_2 : \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$S_2(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \;\; i = 1, \ldots, N-1, \quad S_2(\mathbf{e}_N) = \mathbf{e}_N.$$

Next we compose S_2 with a transformation of type (b) to obtain

$$S_3(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \;\; i = 1, \ldots, N-1, \quad S_3(\mathbf{e}_N) = b\mathbf{e}_N,$$

where b will be chosen later. Now set $T_N^{-1} = (m_{ki})_{k,i=1,\ldots N-1}$. By applying again $N - 1$ transformations of type (c) with the triplets (i, j, α) equal to $(N, 1, \sum_{k=1}^{N-1} a_{Nk}m_{k1}), \ldots, (N, N-1, \sum_{k=1}^{N-1} a_{Nk}m_{k,N-1})$, we obtain

$$S_4(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \quad i = 1, \dots, N-1, \quad S_4(\mathbf{e}_N) = b\mathbf{e}_N + \sum_{i,k=1}^{N-1} a_{Nk}m_{ki} \sum_{j=1}^{N} a_{ij}\mathbf{e}_j.$$

Since $\sum_{i,k=1}^{N-1} a_{Nk}m_{ki} \sum_{j=1}^{N-1} a_{ij}\mathbf{e}_j = \sum_{k=1}^{N-1} a_{Nk}\mathbf{e}_k$, by choosing $b = a_{NN} - \sum_{i,k=1}^{N-1} a_{Nk}m_{ki}a_{iN}$ we conclude that $T = S_4$. □

Remark 1.77 As a corollary of Theorem 1.76 we obtain that the Lebesgue measure is *rotation invariant*.

References

[Br83] Brezis, H.: Analyse fonctionelle. Théorie et applications. Masson, Paris (1983)
[KF75] Kolmogorov, A.N., Fomin, S.V.: Introductory Real Analysis. Dover Publications Inc., New York (1975)
[Ru74] Rudin, W.: Real and Complex Analysis. McGraw-Hill, New York (1986)
[Ru64] Rudin, W.: Principles of Mathematical Analysis. McGraw-Hill, New York (1964)
[Wi62] Williamson, J.H.: Lebesgue Integration. Holt, Rinehart & Winston, New York (1962)
[WZ77] Wheeden, R.L., Zygmund, A.: Measure and Integration. Marcel Dekker Inc., New York (1977)

Chapter 2
Integration

The class of measurable, or Borel, functions $f : X \to \mathbb{R} \cup \{\pm\infty\}$ on a measurable space $(X, \mathscr{E}, \mu)$ can be defined in natural way using the notion of measurable sets. Such a class is stable under linear operations, product, and pointwise convergence. Moreover, if X is a topological space and $\mathscr{E}$ is the Borel σ-algebra, then every continuous function is Borel. In particular, for a Radon measure μ on $\mathbb{R}^N$, all Borel functions $f : \mathbb{R}^N \to \mathbb{R} \cup \{\pm\infty\}$ preserve the regularity properties of μ. A very useful consequence of this is the fact that measurable function can be approximated with continuous functions.

The class of Borel functions plays a crucial role in Lebesgue integration theory, which will be the object of the second part of this chapter. Lebesgue integral can be defined in several ways: our definition will be based on the notion of *archimedean integral* for the repartition function

$$t \geq 0 \mapsto \mu(\{f > t\}).$$

The central idea of all this theory is to make finer and finer partitions of the range of the function to integrate. Clearly, this approach relies on the definition of the integral of simple functions, that is, functions with a finite range. Since such a definition takes in no account the regularity of the function to integrate, the notion of integral can be given for quite a very large class of functions. The importance of Lebesgue integration is also revealed by the flexibility of limiting operations under the integral sign. Another advantage of Lebesgue's approach is that the construction of the integral is exactly the same for functions on a measure space as it is for functions on the real line.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_2

2.1 Measurable Functions

2.1.1 Inverse Image of a Function

Let X, Y be nonempty sets. For any map $f: X \to Y$ and any $A \in \mathscr{P}(Y)$ we set

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

$f^{-1}(A)$ is called the *inverse image* of A.

Let us recall some elementary properties of f^{-1}. The easy proofs are left to the reader as an exercise.

(i) $f^{-1}(A^c) = (f^{-1}(A))^c$ for every $A \in \mathscr{P}(Y)$.
(ii) If $A, B \in \mathscr{P}(Y)$, then $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$. In particular, if $A \cap B = \varnothing$, then $f^{-1}(A) \cap f^{-1}(B) = \varnothing$.
(iii) If $(A_n)_n \subset \mathscr{P}(Y)$, then

$$f^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(A_n).$$

Consequently, if $(Y, \mathscr{F})$ is a measurable space, then the family of parts of X

$$f^{-1}(\mathscr{F}) := \left\{f^{-1}(A) \,\middle|\, A \in \mathscr{F}\right\}$$

is a σ-algebra in X.

Exercise 2.1 Let $f: X \to Y$ and $A \in \mathscr{P}(X)$. Set

$$f(A) := \left\{f(x) \,\middle|\, x \in A\right\}.$$

Show that properties like (i), (ii) fail, in general, for $f(A)$.

2.1.2 Measurable Maps and Borel Functions

In what follows $(X, \mathscr{E})$ and $(Y, \mathscr{F})$ are given measurable spaces.

Definition 2.2 A map $f: X \to Y$ is said to be $\mathscr{E}$*-measurable* or simply *measurable* if $f^{-1}(\mathscr{F}) \subset \mathscr{E}$. If Y is a metric space and $\mathscr{F} = \mathscr{B}(Y)$, f is called a *Borel function*.

Proposition 2.3 *Let $\mathscr{I} \subset \mathscr{F}$ be such that $\sigma(\mathscr{I}) = \mathscr{F}$. Then $f: X \to Y$ is measurable if and only if $f^{-1}(\mathscr{I}) \subset \mathscr{E}$.*

Proof Clearly, if f is measurable, then $f^{-1}(\mathscr{I}) \subset \mathscr{E}$. Conversely, suppose $f^{-1}(\mathscr{I}) \subset \mathscr{E}$, and consider the family

$$\mathscr{G} := \left\{A \in \mathscr{F} \;\middle|\; f^{-1}(A) \in \mathscr{E}\right\}.$$

Using properties (i), (ii), and (iii) of f^{-1} from the previous section, one can easily show that $\mathscr{G}$ is a σ-algebra in Y including $\mathscr{I}$. So $\mathscr{G}$ coincides with $\mathscr{F}$ and the proof is complete. □

Proposition 2.4 *Let X, Y be metric spaces and $\mathscr{E} = \mathscr{B}(X)$, $\mathscr{F} = \mathscr{B}(Y)$. Then any continuous map $f : X \to Y$ is measurable.*

Proof Let $\mathscr{I}$ be the family of all open sets Y. Then $\sigma(\mathscr{I}) = \mathscr{B}(Y)$ and $f^{-1}(\mathscr{I}) \subset \mathscr{B}(X)$. So the conclusion follows from Proposition 2.3. □

Proposition 2.5 *Let $f : X \to Y$ be a measurable map, $(Z, \mathscr{G})$ a measurable space and $g : Y \to Z$ another measurable map. Then $g \circ f$ is measurable.*

Exercise 2.6 Given a measurable map $f : X \to Y$ and a measure μ on $\mathscr{E}$, let $f_\sharp\mu$ be defined by

$$f_\sharp\mu(A) = \mu\big(f^{-1}(A)\big) \qquad \forall A \in \mathscr{F}.$$

Show that $f_\sharp\mu$ is a measure on $\mathscr{F}$ (called the *push-forward* of μ under f).

Exercise 2.7 Let $f : X \to Y$ be such that $f(X)$ is countable. Show that f is measurable if, for every $y \in Y$, $f^{-1}(y) \in \mathscr{E}$.

Example 2.8 Let $f : X \to \mathbb{R}^N$. We regard $\mathbb{R}^N$ as a measurable space with the Borel σ-algebra $\mathscr{B}(\mathbb{R}^N)$. Denoting by f_i the components of f, that is, $f = (f_1, \ldots, f_N)$, let us show that

$$f \quad \text{is Borel} \quad \Longleftrightarrow \quad f_i \quad \text{is Borel} \quad \forall i \in \{1, \ldots, N\}. \tag{2.1}$$

Indeed, let $\mathscr{I}$ be the family of all rectangles of the form

$$R = \prod_{i=1}^{N} [y_i, y_i') = \{z = (z_1, \ldots, z_N) \in \mathbb{R}^N \mid y_i \le z_i < y_i' \ \forall i\},$$

where $y_i \le y_i'$, $i = 1, \ldots, N$. Observe that $\mathscr{B}(\mathbb{R}^N) = \sigma(\mathscr{I})$ to deduce, from Proposition 2.3, that f is Borel if and only if $f^{-1}(\mathscr{I}) \subset \mathscr{E}$. The following identity is easy to verify:

$$f^{-1}(R) = \bigcap_{i=1}^{N} \{x \in X \mid y_i \le f_i(x) < y_i'\} = \bigcap_{i=1}^{N} f_i^{-1}([y_i, y_i')).$$

This shows the '$\Leftarrow$' part of (2.1). To complete the argument, assume that f is Borel and let $i \in \{1, \ldots, N\}$ be fixed. Then for every $a \in \mathbb{R}$ we have

$$f_i^{-1}((-\infty, a]) = f^{-1}\big(\big\{(z_1, \ldots, z_N) \in \mathbb{R}^N \,\big|\, z_i \le a\big\}\big),$$

which implies $f_i^{-1}((-\infty, a]) \in \mathscr{E}$, and so, using Exercise 1.11, f_i is Borel.

Exercise 2.9 Let $f, g\colon X \to \mathbb{R}$ be Borel. Then $f+g,\ fg,\ \min\{f, g\}$ and $\max\{f, g\}$ are Borel.

Hint. Define $F(x) = (f(x), g(x))$ and $\varphi(y_1, y_2) = y_1 + y_2$. Then $F : X \to \mathbb{R}^2$ is a Borel map owing to Example 2.8, and $\varphi : \mathbb{R}^2 \to \mathbb{R}$ is a Borel function, since it is continuous. Thus, by Proposition 2.5, $f + g = \varphi \circ F$ is also Borel. The remaining assertions can be proved similarly.

Exercise 2.10 Let $f\colon X \to \mathbb{R}$ be Borel. Prove that the function

$$g : X \to \mathbb{R}, \quad g(x) = \begin{cases} \dfrac{1}{f(x)} & \text{if } f(x) \ne 0, \\ 0 & \text{if } f(x) = 0 \end{cases}$$

is also Borel.

Hint. Show, first, by a direct argument, that $\varphi : \mathbb{R} \to \mathbb{R}$ defined by

$$\varphi(x) = \begin{cases} \dfrac{1}{x} & \text{if } x \ne 0, \\ 0 & \text{if } x = 0 \end{cases}$$

is Borel.

When dealing with real valued functions defined on X, it is often convenient to allow for values in the extended space $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$. These are called *extended functions*. If $|f(x)| < \infty$ for all $x \in X$, f is said to be *finite* (or *finite valued*). We say that a mapping $f\colon X \to \overline{\mathbb{R}}$ is *Borel* if

$$f^{-1}(-\infty), \quad f^{-1}(\infty) \in \mathscr{E}$$

and $f^{-1}(A) \in \mathscr{E}$ for every $A \in \mathscr{B}(\mathbb{R})$. In what follows, for any $a, b \in \overline{\mathbb{R}}$, we shall often use the notation $\{f > a\}$, $\{f = a\}$, $\{a \le f < b\}$ etc. for the sets $f^{-1}((a, \infty])$, $f^{-1}(\{a\})$, $f^{-1}([a, b))$ etc.

Proposition 2.11 *A function $f : X \to \overline{\mathbb{R}}$ is Borel if and only if any of the following statements holds:*

(i) $\{f \le a\} \in \mathscr{E}$ *for all* $a \in \mathbb{R}$.
(ii) $\{f < a\} \in \mathscr{E}$ *for all* $a \in \mathbb{R}$.
(iii) $\{f \ge a\} \in \mathscr{E}$ *for all* $a \in \mathbb{R}$.

(iv) $\{f > a\} \in \mathscr{E}$ *for all* $a \in \mathbb{R}$.

Proof Since $\{f \le a\} = \{-\infty < f \le a\} \cup \{f = -\infty\}$, and since $(-\infty, a] \in \mathscr{B}(\mathbb{R})$, the measurability of f implies (i). Conversely, assume $\{f \le a\} \in \mathscr{E}$ for all $a \in \mathbb{R}$. Since $\{f > a\}$ is the complement of $\{f \le a\}$, we have $\{f > a\} \in \mathscr{E}$ for all $a \in \mathbb{R}$. Since $\{f = \infty\} = \cap_{k=1}^{\infty}\{f > k\}$ and $\{f = -\infty\} = \cap_{k=1}^{\infty}\{f \le -k\}$, we see that $\{f = \infty\}, \{f = -\infty\} \in \mathscr{E}$. Consequently, $\{a < f < \infty\} = \{f > a\} \setminus \{f = \infty\} \in \mathscr{E}$ for all $a \in \mathbb{R}$. Next consider the family

$$\mathscr{G} := \left\{A \in \mathscr{B}(\mathbb{R}) \;\middle|\; f^{-1}(A) \in \mathscr{E}\right\}.$$

Then $\mathscr{G}$ is a σ-algebra including all semi–infinite intervals (a, ∞). Exercise 1.11 implies that $\mathscr{G}$ coincides with $\mathscr{B}(\mathbb{R})$ and this proves that f is Borel if (i) holds. The proof of the other statements is similar. □

Proposition 2.12 *Let* $f_n : X \to \overline{\mathbb{R}}$ *be a sequence of Borel functions. Then the functions*

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

are Borel. In particular, if $\lim_{n\to\infty} f_n(x)$ *exists for every* $x \in X$*, then the function* $\lim_{n\to\infty} f_n$ *is itself Borel.*

Proof Let us set $\phi := \sup_{n\in\mathbb{N}} f_n$. For any $a \in \mathbb{R}$ we have

$$\{\phi \le a\} = \bigcap_{n=1}^{\infty}\{f_n \le a\} \in \mathscr{E}.$$

The conclusion follows from Proposition 2.11. In a similar way one can prove the other assertions. □

Exercise 2.13 Let $f, g : X \to \overline{\mathbb{R}}$ be Borel. Show that $\{x \in X \mid f = g\} \in \mathscr{E}$.

Exercise 2.14 Let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions. Show that $\{x \in X \mid \exists \lim_n f_n(x)\} \in \mathscr{E}$.

Exercise 2.15 Let $f : \mathbb{R}^N \to \overline{\mathbb{R}}$ be Borel. If $T : \mathbb{R}^N \to \mathbb{R}^N$ is a nonsingular linear transformation, show that $f \circ T : \mathbb{R}^N \to \mathbb{R}$ is Borel.
Hint. If $A_1 = \{f < a\}$ and $A_2 = \{f \circ T < a\}$, show that $A_2 = T^{-1}(A_1)$. Then the conclusion follows from Theorem 1.76.

Exercise 2.16 Let $f : X \to \overline{\mathbb{R}}$ be Borel and $A \in \mathscr{E}$. Show that the function $f_A : X \to \overline{\mathbb{R}}$ defined by

$$f_A(x) = \begin{cases} f(x) & \text{if } x \in A, \\ 0 & \text{if } x \notin A \end{cases}$$

is Borel.

Exercise 2.17 1. Any monotone function $f : \mathbb{R} \to \overline{\mathbb{R}}$ is Borel.
2. Let X be a metric space and $\mathscr{E} = \mathscr{B}(X)$. Then any lower semicontinuous function $f : X \to \mathbb{R} \cup \{\infty\}$ is Borel (see Appendix B).

Exercise 2.18 Let $\mathscr{G}$ be a σ-algebra in $\mathbb{R}$. Show that $\mathscr{G} \supset \mathscr{B}(\mathbb{R})$ if and only if any continuous function $f : \mathbb{R} \to \mathbb{R}$ is $\mathscr{G}$-measurable, that is, $f^{-1}(A) \in \mathscr{G}$ for every $A \in \mathscr{B}(\mathbb{R})$.

Exercise 2.19 Show that Borel functions $f : \mathbb{R} \to \mathbb{R}$ are the smallest class of functions which includes all continuous functions and is stable under pointwise convergence.

We note that the sum of two extended functions $f, g : X \to \overline{\mathbb{R}}$ is well defined wherever it is not of the form $\infty + (-\infty)$ or $-\infty + \infty$; thus we need to assume that at least one of the two functions is finite valued. As regards the product of extended functions, in addition to familiar conventions about the product of infinities, we adopt the convention $0 \cdot \pm\infty = \pm\infty \cdot 0 = 0$.

Exercise 2.20 Let $f, g : X \to \overline{\mathbb{R}}$ be Borel. Show that fg, $\min\{f, g\}$ and $\max\{f, g\}$ are Borel. Furthermore, if g is finite valued, then $f + g$ is Borel.

Definition 2.21 A Borel function $f : X \to \overline{\mathbb{R}}$ is said to be *simple* if its range $f(X)$ is a finite set. The class of all simple functions $f : X \to \overline{\mathbb{R}}$ is denoted by $\mathscr{S}(X)$.

It is immediate that the class $\mathscr{S}(X)$ is closed under the operations of sum (if well defined), product and lattice[1] $(\wedge, \vee)$.

Given $A \subset X$, the function $\chi_A : X \to \mathbb{R}$ defined by

$$\chi_A(x) = \begin{cases} 1 \text{ if } x \in A, \\ 0 \text{ if } x \notin A \end{cases}$$

is called the *characteristic function* of the set A. Clearly, $\chi_A \in \mathscr{S}(X)$ if and only if $A \in \mathscr{E}$.

Remark 2.22 1. We note that $f : X \to \overline{\mathbb{R}}$ is simple if and only if there exist $a_1, \ldots, a_n \in \overline{\mathbb{R}}$ and disjoint sets $A_1, \ldots, A_n \in \mathscr{E}$ such that

$$X = \bigcup_{i=1}^{n} A_i \quad \text{and} \quad f(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x) \quad \forall x \in X. \tag{2.2}$$

Indeed, any function of the form (2.2) is simple. Conversely, if f is simple, then

$$f(X) = \{a_1, \ldots, a_n\} \quad \text{with} \quad a_i \neq a_j \quad \text{if} \quad i \neq j.$$

So, taking $A_i := f^{-1}(a_i)$, $i \in \{1, \ldots, n\}$, we obtain a representation of f of type (2.2). Obviously, the choice of sets $A_1, \ldots, A_n \in \mathscr{E}$ and values $a_1, \ldots, a_n$ is far from being unique.

[1] By definition, $f \vee g = \max\{f, g\}$ and $f \wedge g = \min\{f, g\}$.

2. Given two simple functions f and g, they can always be represented as linear combinations of the characteristic functions of the same family of sets. To see this, let f be given by (2.2), and let

$$X = \bigcup_{j=1}^{m} B_j \quad \text{and} \quad g(x) = \sum_{j=1}^{m} b_j \chi_{B_j}(x), \quad \forall x \in X.$$

Since $A_i = \bigcup_{j=1}^{m} (A_i \cap B_j)$, we have that

$$\chi_{A_i}(x) = \sum_{j=1}^{m} \chi_{A_i \cap B_j}(x) \qquad i \in \{1, \ldots, n\}.$$

So

$$f(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \chi_{A_i \cap B_j}(x), \quad x \in X.$$

Similarly,

$$g(x) = \sum_{j=1}^{m} \sum_{i=1}^{n} b_j \chi_{A_i \cap B_j}(x), \quad x \in X.$$

Now, we show that any positive Borel function can be approximated by simple finite functions.

Proposition 2.23 *Let $f : X \to [0, \infty]$ be a Borel function. Define for any $n \in \mathbb{N}$*

$$f_n(x) = \begin{cases} \dfrac{i-1}{2^n} & \text{if } \dfrac{i-1}{2^n} \le f(x) < \dfrac{i}{2^n}, \quad i = 1, 2, \ldots, n2^n, \\ n & \text{if } f(x) \ge n. \end{cases} \tag{2.3}$$

Then $(f_n)_n \subset \mathscr{S}(X)$, $0 \le f_n \le f_{n+1}$ and $f_n(x) \uparrow f(x)$ for every $x \in X$. If, in addition, f is bounded, then the convergence is uniform.

Proof For every $n \in \mathbb{N}$ and $i = 1, \ldots, n2^n$ set

$$A_{n,i} = \left\{ \frac{i-1}{2^n} \le f < \frac{i}{2^n} \right\}, \quad B_n = \{f \ge n\}.$$

Since f is Borel, we have $A_{n,i},\ B_n \in \mathscr{E}$ and

$$f_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{A_{n,i}} + n\chi_{B_n}.$$

Then, by Remark 2.22, $f_n \in \mathscr{S}(X)$. Let $x \in X$ be such that $\frac{i-1}{2^n} \le f(x) < \frac{i}{2^n}$. So $\frac{2i-2}{2^{n+1}} \le f(x) < \frac{2i}{2^{n+1}}$ and we get

$$f_{n+1}(x) = \frac{2i-2}{2^{n+1}} \quad \text{or} \quad f_{n+1}(x) = \frac{2i-1}{2^{n+1}}.$$

In any case, $f_n(x) \le f_{n+1}(x)$. Now let $x \in X$ be such that $f(x) \ge n$; we have $f(x) \ge n+1$ or $n \le f(x) < n+1$. In the first case, $f_{n+1}(x) = n+1 > n = f_n(x)$. In the second case, consider $i = 1, \dots, (n+1)2^{n+1}$ such that $\frac{i-1}{2^{n+1}} \le f(x) < \frac{i}{2^{n+1}}$. Since $f(x) \ge n$, we deduce $\frac{i}{2^{n+1}} > n$, by which $i = (n+1)2^{n+1}$; therefore $f_{n+1}(x) = n + 1 - \frac{1}{2^{n+1}} > n = f_n(x)$. This proves that $f_n \le f_{n+1}$.

To prove convergence, fix $x \in X$ such that $f(x) \in [0, \infty)$ and let $n > f(x)$. Then

$$0 \le f(x) - f_n(x) < \frac{1}{2^n}. \tag{2.4}$$

So $f_n(x) \to f(x)$ as $n \to \infty$. On the other hand, if $f(x) = \infty$ then $f_n(x) = n \to \infty$. Finally, if $0 \le f(x) \le M$ for all $x \in X$ and some constant $M > 0$, then (2.4) holds for every $x \in X$ provided that $n > M$. Thus, $f_n \to f$ uniformly. □

2.2 Convergence Almost Everywhere

In this section we introduce a generalization of the ordinary notion of convergence for a sequence of functions. In the following $(X, \mathscr{E}, \mu)$ is a given measure space.

Definition 2.24 We say that a sequence of functions $f_n : X \to \overline{\mathbb{R}}$ converges to a function $f : X \to \overline{\mathbb{R}}$

- *almost everywhere* ($f_n \xrightarrow{a.e.} f$) if there exists a set $E \in \mathscr{E}$ of measure zero such that

$$\lim_{n\to\infty} f_n(x) = f(x) \qquad \forall x \in X \setminus E.$$

- *almost uniformly* ($f_n \xrightarrow{a.u.} f$) if f is finite and, for any $\varepsilon > 0$, there exists $E_\varepsilon \in \mathscr{E}$ such that $\mu(E_\varepsilon) < \varepsilon$ and $f_n \to f$ uniformly in $X \setminus E_\varepsilon$.

Exercise 2.25 Let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions.

1. Show that the pointwise limit of f_n, when it exists, is a Borel function.
2. Show that if $f_n \xrightarrow{a.u.} f$, then $f_n \xrightarrow{a.e.} f$.
3. Show that if $f_n \xrightarrow{a.e.} f$ and $f_n \xrightarrow{a.e.} g$, then $f = g$ except on a set of measure zero.
4. We say that $f_n \to f$ uniformly almost everywhere if there exists $E \in \mathscr{E}$ of measure zero such that $f_n \to f$ uniformly in $X \setminus E$. Show that almost uniform convergence does not imply uniform convergence almost everywhere, in general. *Hint*. Consider the sequence $f_n(x) = x^n$ defined on $[0, 1]$ with the Lebesgue measure.

Example 2.26 In contrast to Exercise 2.25(1), observe that the a.e. limit of Borel functions may fail to be Borel. Indeed, the trivial sequence $f_n \equiv 0$ defined on $(\mathbb{R}, \mathscr{B}(\mathbb{R}), m)$ (m denoting the Lebesgue measure) converges a.e. to χ_C, where C is the Cantor set (see Example 1.63), and also to χ_E where E is any subset of C which is not a Borel set. This is a consequence of the fact that the Lebesgue measure on $\mathscr{B}(\mathbb{R})$ is not complete. On the other hand, if the domain $(X, \mathscr{E}, \mu)$ of $(f_n)_n$ is such that μ is a *complete* measure on $\mathscr{E}$, then the a.e. limit of Borel functions is also a Borel function.

The following result establishes a surprising consequence of a.e. convergence on sets of finite measure.

Theorem 2.27 (Severini–Egorov) *Let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions. If $\mu(X) < \infty$ and f_n converges a.e. to a finite Borel function f, then $f_n \xrightarrow{a.u.} f$.*

Proof For any $k, n \in \mathbb{N}$ define

$$A_n^k = \bigcup_{i=n}^{\infty} \left\{ |f - f_i| > \frac{1}{k} \right\}.$$

Observe that $A_n^k \in \mathscr{E}$ since f_n and f are Borel functions. Moreover

$$A_n^k \downarrow \limsup_{n\to\infty} \left\{ |f - f_n| > \frac{1}{k} \right\} =: A^k \quad (n \to \infty).$$

So $A^k \in \mathscr{E}$. For any $x \in A^k$ we have $|f(x) - f_n(x)| > \frac{1}{k}$ for infinitely many indices n; thus, $\mu(A^k) = 0$ by our hypotheses. Recalling that μ is finite, by Proposition 1.18 we conclude that, for every $k \in \mathbb{N}$, $\mu(A_n^k) \downarrow 0$ as $n \to \infty$. Therefore, for any given $\varepsilon > 0$, there exists an increasing sequence of integers $(n_k)_k$ such that $\mu(A_{n_k}^k) < \frac{\varepsilon}{2^k}$ for all $k \in \mathbb{N}$. Let us set

$$E_\varepsilon := \bigcup_{k=1}^{\infty} A_{n_k}^k.$$

Then $\mu(E_\varepsilon) \le \sum_{k=1}^{\infty} \mu(A^k_{n_k}) < \varepsilon$. Moreover, for every $x \in X \setminus E_\varepsilon$, we have that

$$i \ge n_k \quad \Longrightarrow \quad |f(x) - f_i(x)| \le \frac{1}{k}$$

for all integers $k \ge 1$, namely $f_n \to f$ uniformly in $X \setminus E_\varepsilon$. □

Example 2.28 Theorem 2.27 is false, in general, if $\mu(X) = \infty$. For instance, consider $f_n = \chi_{[n,\infty)}$ defined on $\mathbb{R}$ with the Lebesgue measure m. Then $f_n \to 0$ pointwise, but $m(\{f_n = 1\}) = \infty$.

2.3 Approximation by Continuous Functions

The aim of this section is to prove that a Borel function $f : \mathbb{R}^N \to \mathbb{R}$ can be approximated in a measure theoretical sense by a continuous function, as shown by the following result known as *Lusin's theorem*.

Theorem 2.29 (Lusin) *Let μ be a Radon measure on $\mathbb{R}^N$, $f : \mathbb{R}^N \to \mathbb{R}$ a Borel function and $A \in \mathscr{B}(\mathbb{R}^N)$ such that*

$$\mu(A) < \infty \qquad \textit{and} \qquad f(x) = 0 \quad \forall x \notin A.$$

Then for any $\varepsilon > 0$ there exists a continuous function $f_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ with compact support[2] *such that*

$$\mu\big(\{f \ne f_\varepsilon\}\big) < \varepsilon, \tag{2.5}$$

$$\sup_{x\in\mathbb{R}^N} |f_\varepsilon(x)| \le \sup_{x\in\mathbb{R}^N} |f(x)|. \tag{2.6}$$

Proof We split the proof into five steps.

1. Assume that A is compact and $0 \le f < 1$. Let V be a bounded open set such that $A \subset V$. Consider the sequence $(f_n)_n \subset \mathscr{S}(X)$ defined in the statement of Proposition 2.23. We have

$$f_1 = \frac{1}{2}\chi_{A_1}, \qquad A_1 = \left\{f \ge \frac{1}{2}\right\}, \tag{2.7}$$

$$f_n - f_{n-1} = \frac{1}{2^n}\chi_{A_n}, \qquad A_n = \left\{f - f_{n-1} \ge \frac{1}{2^n}\right\} \quad \forall n \ge 2. \tag{2.8}$$

---

[2] Given a continuous function $f : \mathbb{R}^N \to \mathbb{R}$, the closure of the set $\{x \in \mathbb{R}^N \mid f(x) \ne 0\}$ is called the *support* of f, and is denoted by $\mathrm{supp}(f)$.

(2.7) is obvious; to prove (2.8), consider $x \in \mathbb{R}^N$ and $i = 1, \ldots, 2^{n-1}$ such that $\frac{i-1}{2^{n-1}} \le f(x) < \frac{i}{2^{n-1}}$. Then $f_{n-1}(x) = \frac{i-1}{2^{n-1}}$. Moreover

$$\frac{2i-2}{2^n} \le f(x) < \frac{2i-1}{2^n} \quad \text{or} \quad \frac{2i-1}{2^n} \le f(x) < \frac{2i}{2^n}.$$

In the first case, $x \notin A_n$ and $f_n(x) = \frac{2i-2}{2^n} = f_{n-1}(x)$; in the second case, $x \in A_n$ and $f_n(x) = \frac{2i-1}{2^n} = f_{n-1}(x) + \frac{1}{2^n}$. Therefore (2.8) follows. Since $f_n = f_1 + \sum_{i=2}^{n}(f_i - f_{i-1})$ for every $n \ge 2$, we deduce

$$f(x) = \lim_{n\to\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{1}{2^n}\chi_{A_n}(x) \tag{2.9}$$

where the series converges uniformly in $\mathbb{R}^N$. We observe that $A_n \in \mathscr{B}(\mathbb{R}^N)$ and $A_n \subset A$ for every $n \ge 1$.

Let us fix $\varepsilon > 0$. Owing to Theorem 1.71, for any n there exist a compact set K_n and an open set V_n such that

$$K_n \subset A_n \subset V_n \qquad \text{and} \qquad \mu(V_n \setminus K_n) < \frac{\varepsilon}{2^n}.$$

Possibly replacing V_n by $V_n \cap V$, we may assume $V_n \subset V$. Define[3]

$$g_n(x) = \frac{d_{V_n^c}(x)}{d_{K_n}(x) + d_{V_n^c}(x)} \qquad \forall x \in \mathbb{R}^N.$$

It is immediate to check that g_n is continuous and

$$0 \le g_n(x) \le 1 \quad \forall x \in \mathbb{R}^N \qquad \text{and} \qquad g_n \equiv \begin{cases} 1 & \text{in } K_n, \\ 0 & \text{in } V_n^c. \end{cases}$$

So, in some sense, g_n approximates χ_{A_n}. Now, let us set

$$f_\varepsilon(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} g_n(x) \qquad \forall x \in \mathbb{R}^N. \tag{2.10}$$

Since the series $\sum_{n=1}^{\infty} \frac{1}{2^n} g_n$ is totally convergent, we deduce that f_ε is continuous. Moreover,

$$\{f_\varepsilon \ne 0\} \subset \bigcup_{n=1}^{\infty} \{g_n \ne 0\} \subset \bigcup_{n=1}^{\infty} V_n \subset V,$$

[3] As usual, $d_S(x)$ denotes the distance between the set S and the point x (see Appendix A).

and so $\operatorname{supp}(f_\varepsilon) \subset \overline{V}$. Consequently, $\operatorname{supp}(f_\varepsilon)$ is compact. By (2.9) and (2.10) we have

$$\{f_\varepsilon \neq f\} \subset \bigcup_{n=1}^{\infty} \{g_n \neq \chi_{A_n}\} \subset \bigcup_{n=1}^{\infty} (V_n \setminus K_n)$$

which implies, in turn,

$$\mu\big(\{f_\varepsilon \neq f\}\big) \leq \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon.$$

Thus, conclusion (2.5) holds when A is compact and $0 \leq f < 1$.

2. Obviously, (2.5) also holds when A is compact and $0 \leq f < M$ for some constant $M > 0$ (it suffices to replace f by f/M). Moreover, if A is compact and f is bounded, then $|f| < M$ for some $M > 0$. So, in order to derive (2.5) in this case, it suffices to decompose $f = f^+ - f^-$, where $f^+ = \max\{f, 0\}$, $f^- = \max\{-f, 0\}$, and observe that $0 \leq f^+,\ f^- < M$.
3. We will now remove the compactness assumption for A. By Theorem 1.71, there exists a compact set $K \subset A$ such that $\mu(A \setminus K) < \varepsilon$. Let us set

$$\bar{f} = \chi_K f.$$

Since $\bar{f}$ vanishes outside K, from the previous steps we can approximate $\bar{f}$ by a continuous function with compact support, say f_ε. Then

$$\{f_\varepsilon \neq f\} \subset \{f_\varepsilon \neq \bar{f}\} \cup (A \setminus K).$$

Hence,

$$\mu\big(\{f_\varepsilon \neq f\}\big) < 2\varepsilon.$$

4. In order to remove the boundedness assumption for f, define Borel sets $(B_n)_n$ by

$$B_n = \{|f| \geq n\} \qquad n \in \mathbb{N}.$$

Clearly,

$$B_{n+1} \subset B_n \qquad \text{and} \qquad \bigcap_{n\in\mathbb{N}} B_n = \varnothing.$$

Since $\mu(A) < \infty$, Proposition 1.18 yields $\mu(B_n) \to 0$. Therefore, for some $\bar{n} \in \mathbb{N}$, we have $\mu(B_{\bar{n}}) < \varepsilon$. We define

$$\bar{f} = (1 - \chi_{B_{\bar{n}}}) f.$$

Since $\bar{f}$ is bounded (by $\bar{n}$), from the previous steps we can approximate $\bar{f}$ by a continuous function with compact support, that we again label f_ε. Then

$$\{f_\varepsilon \neq f\} \subset \{f_\varepsilon \neq \bar{f}\} \cup B_{\bar{n}},$$

by which

$$\mu(\{f_\varepsilon \neq f\}) < 2\varepsilon.$$

The proof of (2.5) is thus complete.

5. Finally, in order to prove (2.6), suppose $M := \sup_{\mathbb{R}^N} |f| < \infty$. Define

$$\theta_M : \mathbb{R} \to \mathbb{R} \qquad \theta_M(t) = \begin{cases} t & \text{if } |t| < M, \\ M\dfrac{t}{|t|} & \text{if } |t| \geq M \end{cases}$$

and $\bar{f}_\varepsilon = \theta_M \circ f_\varepsilon$ to obtain $|\bar{f}_\varepsilon| \leq M$. Since θ_R is continuous, so is $\bar{f}_\varepsilon$. Furthermore, $\mathrm{supp}(\bar{f}_\varepsilon) = \mathrm{supp}(f_\varepsilon)$ and

$$\{f_\varepsilon = f\} \subset \{\bar{f}_\varepsilon = f\}.$$

This completes the proof. □

It is useful to point out the following corollary of Lusin's Theorem.

Corollary 2.30 *Let μ be a Radon measure on $\mathbb{R}^N$, $A \subset \mathbb{R}^N$ a Borel set such that $\mu(A) < \infty$ and $f : A \to \mathbb{R}$ a Borel function. Then for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset A$ such that $f\big|_{K_\varepsilon} : K_\varepsilon \to \mathbb{R}$ is continuous and $\mu(A \setminus K_\varepsilon) < \varepsilon$.*

Proof Let us apply Lusin's Theorem to the function $\tilde{f}$ obtained by extending f to zero outside A: there exists a continuous function $f_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ such that, if we set $A_\varepsilon = \{x \in A \mid f(x) = f_\varepsilon(x)\}$, we have $\mu(A \setminus A_\varepsilon) \leq \frac{\varepsilon}{2}$. By Theorem 1.71 there exists a compact set $K_\varepsilon \subset A_\varepsilon$ such that $\mu(A_\varepsilon \setminus K_\varepsilon) \leq \frac{\varepsilon}{2}$. Therefore

$$\mu(A \setminus K_\varepsilon) = \mu(A \setminus A_\varepsilon) + \mu(A_\varepsilon \setminus K_\varepsilon) \leq \varepsilon.$$

□

2.4 Integral of Borel Functions

Let $(X, \mathscr{E}, \mu)$ be a given measure space. In this section we will define the integral of a Borel function $f : X \to \overline{\mathbb{R}}$ with respect to the measure μ. We will first consider the special case of positive functions, and then the case of functions with variable sign.

2.4.1 Integral of Positive Simple Functions

We begin with the definition of the integral in the class $\mathscr{S}_+(X)$ of positive simple functions, i.e.,

$$\mathscr{S}_+(X) = \{f : X \to [0, \infty] \,|\, f \in \mathscr{S}(X)\}.$$

Definition 2.31 Let $f \in \mathscr{S}_+(X)$. According to Remark 2.22(1) f has a representation of the form

$$f(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x) \qquad x \in X,$$

where $a_1, \ldots, a_n \in [0, \infty]$ and $A_1, \ldots, A_n$ are mutually disjoint sets in $\mathscr{E}$ such that $A_1 \cup \cdots \cup A_n = X$. Then, using the convention $0 \cdot \infty = 0$, the *(Lebesgue) integral of f over X with respect to the measure μ* is defined by

$$\int_X f(x)\, d\mu(x) = \int_X f\, d\mu = \sum_{i=1}^{n} a_i \mu(A_i).$$

Remark 2.32 It is easy to see that the above definition is independent of the representation of f. Indeed, given disjoint sets $B_1, \ldots, B_m \in \mathscr{E}$ with $B_1 \cup \cdots \cup B_m = X$ and numbers $b_1, \ldots, b_m \in [0, \infty]$ such that

$$f(x) = \sum_{j=1}^{m} b_j \chi_{B_j}(x) \qquad x \in X,$$

we have

$$A_i = \bigcup_{j=1}^{m} (A_i \cap B_j) \qquad B_j = \bigcup_{i=1}^{n} (A_i \cap B_j)$$

and

$$A_i \cap B_j \neq \varnothing \quad \Longrightarrow \quad a_i = b_j.$$

Therefore

$$\begin{aligned}\sum_{i=1}^{n} a_i \mu(A_i) &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \mu(A_i \cap B_j) \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} b_j \mu(A_i \cap B_j) = \sum_{j=1}^{m} b_j \mu(B_j).\end{aligned}$$

Proposition 2.33 *Let* $f, g \in \mathscr{S}_+(X)$ *and* $\alpha, \beta \in [0, \infty]$. *Then*

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

Proof Owing to Remark 2.22(2), f and g can be represented using the same family of mutually disjoint sets $A_1, \dots, A_n \in \mathscr{E}$ as

$$f = \sum_{i=1}^n a_i \chi_{A_i} \qquad g = \sum_{i=1}^n b_i \chi_{A_i}.$$

Then

$$\begin{aligned}\int_X (\alpha f + \beta g)\, d\mu &= \sum_{i=1}^n (\alpha a_i + \beta b_i)\mu(A_i) = \alpha \sum_{i=1}^n a_i \mu(A_i) + \beta \sum_{i=1}^n b_i \mu(A_i)\\ &= \alpha \int_X f\, d\mu + \beta \int_X g\, d\mu\end{aligned}$$

as required. □

We now proceed with what can rightfully be considered the central notion of Lebesgue integration.

2.4.2 Repartition Function

Let $f : X \to [0, \infty]$ be a Borel function. The *repartition function* M_f of f is defined by

$$M_f(t) := \mu(\{f > t\}) = \mu(f > t), \quad t \geq 0.$$

By definition, $M_f : [0, \infty) \to [0, \infty]$ is a decreasing[4] function; then M_f has a limit at ∞. Moreover, since

$$\{f = \infty\} = \bigcap_{n=1}^{\infty} \{f > n\},$$

we have

$$\lim_{t\to\infty} M_f(t) = \lim_{n\to\infty} M_f(n) = \lim_{n\to\infty} \mu(f > n) = \mu(f = \infty)$$

[4]That is,

$$t_1, t_2 \in [0, \infty),\ t_1 < t_2 \Longrightarrow M_f(t_1) \geq M_f(t_2).$$

whenever μ is finite. Other important properties of M_f are provided by the following result.

Proposition 2.34 *Let $f: X \to [0, \infty]$ be a Borel function and let M_f be its repartition function. Then the following properties hold:*

(i) For every $t_0 \geq 0$

$$\lim_{t \downarrow t_0} M_f(t) = M_f(t_0)$$

(that is, M_f is right continuous).

(ii) If $\mu(X) < \infty$, then for every $t_0 > 0$

$$\lim_{t \uparrow t_0} M_f(t) = \mu(f \geq t_0)$$

(that is, M_f possesses left limit).

Proof First observe that, since M_f is a decreasing function, then M_f has a left limit at any $t > 0$ and a right limit at any $t \geq 0$. Let us prove (i). We have

$$\lim_{t \downarrow t_0} M_f(t) = \lim_{n \to \infty} M_f\left(t_0 + \frac{1}{n}\right) = \lim_{n \to \infty} \mu\left(f > t_0 + \frac{1}{n}\right) = \mu(f > t_0) = M_f(t_0),$$

since

$$\left\{f > t_0 + \frac{1}{n}\right\} \uparrow \{f > t_0\}.$$

Now, to prove (ii), we note that

$$\left\{f > t_0 - \frac{1}{n}\right\} \downarrow \{f \geq t_0\}.$$

Thus, recalling that μ is finite, we obtain

$$\lim_{t \uparrow t_0} M_f(t) = \lim_{n \to \infty} M_f\left(t_0 - \frac{1}{n}\right) = \lim_{n \to \infty} \mu\left(f > t_0 - \frac{1}{n}\right) = \mu(f \geq t_0),$$

and (ii) follows. □

By Proposition 2.34 it follows that, when μ is finite, M_f is continuous at t_0 if and only if $\mu(f = t_0) = 0$.

Example 2.35 Let $f \in \mathscr{S}_+(X)$ and choose a representation of f of the form

$$f(x) = \sum_{i=0}^{n} a_i \chi_{A_i} \qquad x \in X,$$

with $0 = a_0 < a_1 < a_2 < \cdots < a_n = a \leq \infty$ and disjoint sets $A_0, A_1, \ldots, A_n \in \mathscr{E}$ such that $X = \cup_{i=0}^n A_i$. Then the repartition function M_f of f is given by

$$M_f(t) = \begin{cases} \mu(A_1) + \mu(A_2) + \cdots + \mu(A_n) = M_f(0) & \text{if } 0 \leq t < a_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \cdots\cdots\cdots \\ \mu(A_i) + \mu(A_{i+1}) + \cdots + \mu(A_n) = M_f(a_{i-1}) & \text{if } a_{i-1} \leq t < a_i, \\ \cdots\cdots\cdots\cdots & \cdots\cdots\cdots \\ \mu(A_n) = M_f(a_{n-1}) & \text{if } a_{n-1} \leq t < a, \\ 0 = M_f(a) & \text{if } t \geq a. \end{cases}$$

Thus we have

$$M_f(t) = \sum_{i=1}^n M_f(a_{i-1}) \chi_{[a_{i-1}, a_i)}(t) \qquad \forall t \geq 0$$

and $\mu(A_i) = M_f(a_{i-1}) - M_f(a_i)$. Therefore M_f is a simple function itself and a direct computation shows that

$$\begin{aligned} \int_X f\, d\mu &= \sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n a_i \big(M_f(a_{i-1}) - M_f(a_i)\big) \\ &= \sum_{i=1}^n M_f(a_{i-1})(a_i - a_{i-1}) = \int_{[0,\infty)} M_f(t)\, dm \end{aligned} \tag{2.11}$$

where m stands for the Lebesgue measure on $[0, \infty)$.

2.4.3 The Archimedean Integral

In order to be able to define the integral of f when f is a positive Borel function, we need to develop, first, the notion of *archimedean integral* of any decreasing function $F : [0, \infty) \to [0, \infty]$. For any $t \in (0, \infty)$ let us denote by $F(t^-)$ the left limit of F at t:

$$F(t^-) := \lim_{s \uparrow t} F(s).$$

We observe that $F(t^-) \geq F(t)$ and $t_1 < t_2 \Rightarrow F(t_1) \geq F(t_2^-)$.

Let Σ be the family of all finite sets $\{t_0, \ldots, t_n\}$, where $n \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_n < \infty$.

Definition 2.36 For any decreasing function $F\colon [0, \infty) \to [0, \infty]$ the *archimedean integral* of F is defined by

$$\int_0^\infty F(t)dt := \sup\{I_F(\sigma) : \ \sigma \in \Sigma\} \in [0, \infty]$$

where, for any $\sigma = \{t_0, t_1, \ldots, t_n\} \in \Sigma$, we set

$$I_F(\sigma) = \sum_{i=1}^{n} F(t_i^-)(t_i - t_{i-1}).$$

Exercise 2.37 Let $F,\ G\colon [0,\infty) \to [0,\infty]$ be decreasing functions. Show that:

1. If $\sigma, \zeta \in \Sigma$ and $\sigma \subset \zeta$, then $I_F(\sigma) \le I_F(\zeta)$.
2. If $F(t) \le G(t)$ for every $t > 0$, then $\int_0^\infty F(t)\,dt \le \int_0^\infty G(t)\,dt$.
3. If $F(t) = 0$ for every $t > 0$, then $\int_0^\infty F(t)\,dt = 0$.

Now we want to derive a crucial property of passage to the limit under the archimedean integral sign.

Proposition 2.38 *Let* $F_n\colon [0,\infty) \to [0,\infty]$ *be a sequence of decreasing functions such that*

$$F_n(t) \uparrow F(t) \qquad (n \to \infty) \qquad \forall t \ge 0.$$

Then

$$\int_0^\infty F_n(t)\,dt \ \uparrow \ \int_0^\infty F(t)\,dt.$$

Proof According to Exercise 2.37(2), since $F_n \le F_{n+1} \le F$, we obtain

$$\int_0^\infty F_n(t)\,dt \le \int_0^\infty F_{n+1}(t)\,dt \le \int_0^\infty F(t)\,dt$$

for every n. Then the inequality $\lim_{n\to\infty} \int_0^\infty F_n(t)\,dt \le \int_0^\infty F(t)\,dt$ is clear.

To prove the opposite inequality, let L be any number less than $\int_0^\infty F(t)\,dt$. Then there exists $\sigma = \{t_0, \ldots, t_N\} \in \Sigma$ such that

$$\sum_{i=1}^{N} F(t_i^-)(t_i - t_{i-1}) > L.$$

For $0 < \varepsilon < \min\{t_i - t_{i-1} \mid i = 1, \ldots, N\}$, let us set

$$t_0^\varepsilon = t_0 = 0, \quad t_i^\varepsilon = t_i - \varepsilon \quad \forall i = 1, \ldots, N.$$

Thus, $\sigma_\varepsilon = \{t_0^\varepsilon, \ldots, t_N^\varepsilon\} \in \Sigma$. Since $t_i^\varepsilon \uparrow t_i$ and $F(t_i^\varepsilon) \to F(t_i^-)$ for $\varepsilon \to 0^+$, choose ε sufficiently small such that

$$\sum_{i=1}^{N} F(t_i^\varepsilon)(t_i^\varepsilon - t_{i-1}^\varepsilon) > L.$$

Therefore, for n sufficiently large, say $n \geq n_L$,

$$\int_0^\infty F_n(t)\,dt \geq \sum_{i=1}^{N} F_n((t_i^\varepsilon)^-)(t_i^\varepsilon - t_{i-1}^\varepsilon) \geq \sum_{i=1}^{N} F_n(t_i^\varepsilon)(t_i^\varepsilon - t_{i-1}^\varepsilon) > L,$$

by which $\lim_{n\to\infty} \int_0^\infty F_n(t)\,dt \geq L$. The arbitrariness of L gives

$$\lim_{n\to\infty} \int_0^\infty F_n(t)\,dt \geq \int_0^\infty F(t)\,dt.$$

This concludes the proof. □

Exercise 2.39 Given a decreasing function $F\colon [0,\infty) \to [0,\infty]$, show that for any $a > 0$

$$\int_0^\infty F(t)\,dt \geq aF(a).$$

Remark 2.40 Let $F : [0,\infty) \to [0,\infty]$ be a simple decreasing function. Then F is a 'step function': more precisely, there exist $a_0, a_1, \ldots, a_n$ and $c_1, c_2, \ldots c_n$ such that

$$0 = a_0 < a_1 < \cdots < a_n = \infty, \qquad \infty \geq c_1 > c_2 > \cdots > c_n \geq 0$$

and

$$F\big|_{(a_{i-1},a_i)} = c_i \quad \forall i = 1,\ldots,n.$$

So $F \in \mathscr{S}_+([0,\infty))$ and therefore it makes sense to inquire whether the archimedean integral of F coincides with the integral of Definition 2.31 with respect to the Lebesgue measure on $[0,\infty)$, i.e.,

$$\int_0^\infty F(t)\,dt = \sum_{i=1}^{n} c_i(a_i - a_{i-1}). \tag{2.12}$$

Let us first assume $c_n = 0$. Given $\sigma \in \Sigma$, set $\sigma' = \sigma \cup \{a_0,\ldots,a_{n-1}\} \in \Sigma$. Then, if $\sigma' = \{t_0, t_1, \ldots, t_m\}$ with $0 = t_0 < t_1 < \cdots < t_m$, there exist k_i, $i = 0,\ldots,n-1$, such that $t_{k_i} = a_i$, and we have $0 = k_0 < \cdots < k_{n-1} \leq m$ and $F(t_j^-) = c_i$ for $k_{i-1} < j \leq k_i$. Since σ' is finer than σ, using Exercise 2.37(1), we deduce that $I_F(\sigma) \leq I_F(\sigma')$; moreover

$$\begin{aligned} I_F(\sigma') &= \sum_{j=1}^{m} F(t_j^-)(t_j - t_{j-1}) = \sum_{i=1}^{n-1} \sum_{j=k_{i-1}+1}^{k_i} F(t_j^-)(t_j - t_{j-1}) \\ &= \sum_{i=1}^{n-1} c_i \sum_{j=k_{i-1}+1}^{k_i} (t_j - t_{j-1}) = \sum_{i=1}^{n-1} c_i(a_i - a_{i-1}), \end{aligned}$$

and (2.12) is thus proved in the case $c_n = 0$.

If $c_n > 0$, then, using Exercise 2.39, for every $k > a_{n-1}$ we have $\int_0^\infty F(t)dt \geq kc_n$, and consequently $\int_0^\infty F(t)dt = \infty$, by which (2.12) follows.

Remark 2.41 Recalling Exercise 2.35, if $f \in \mathscr{S}_+(X)$, then its repartition function $M_f : [0, \infty) \to [0, \infty]$ is a simple decreasing function. Therefore, owing to Remark 2.40,

$$\int_0^\infty M_f(t)\, dt = \int_{[0,\infty)} M_f\, dm,$$

where m stands for the Lebesgue measure on $[0, \infty)$. Moreover, using (2.11), we deduce

$$\int_X f\, d\mu = \int_{[0,\infty)} M_f\, dm = \int_0^\infty M_f(t)\, dt = \int_0^\infty \mu(f > t)\, dt. \tag{2.13}$$

2.4.4 Integral of Positive Borel Functions

Using identity (2.13) obtained for simple functions, we can now extend the definition of the Lebesgue integral to positive Borel functions.

Definition 2.42 Given $f : X \to [0, \infty]$ a Borel function, the *(Lebesgue) integral of f over X with respect to the measure μ* is defined by

$$\int_X f\, d\mu = \int_X f(x)\, d\mu(x) : = \int_0^\infty \mu(f > t)\, dt,$$

where the integral in the right-hand side is the archimedean integral of the repartition function of f. If the integral of f is finite, f is said to be *μ-summable*.

Next result gives an estimate of the 'size' of f in terms of the integral of f.

Proposition 2.43 (Markov) *Let $f : X \to [0, \infty]$ be a Borel function. Then, for any $a \in (0, \infty)$,*

$$\mu(f > a) \leq \frac{1}{a} \int_X f\, d\mu.$$

Proof Recalling Exercise 2.39, for any $a \in (0, \infty)$ we have

$$\int_X f\, d\mu = \int_0^\infty \mu(f > t)\, dt \geq a\mu(f > a).$$

The conclusion follows. □

Markov's inequality has important consequences. Generalizing the notion of a.e. convergence (see Definition 2.24), we say that a property concerning the points of X

holds *almost everywhere*, or, in abbreviated form, *a.e.*, if it holds for all points of X except for a set $E \in \mathscr{E}$ with $\mu(E) = 0$.

Proposition 2.44 *Let $f : X \to [0, \infty]$ be a Borel function.*

(i) If f is μ-summable, then the set $\{f = \infty\}$ has measure zero, that is, f is a.e. finite.
(ii) The integral of f over X is zero if and only if f is a.e. equal to 0.

Proof (i) From Markov's inequality it follows that $\mu(f > a) < \infty$ for every $a > 0$ and

$$\lim_{a\to\infty} \mu(f > a) = 0.$$

Since

$$\{f > n\} \downarrow \{f = \infty\},$$

we have

$$\mu(f = \infty) = \lim_{n\to\infty} \mu(f > n) = 0.$$

(ii) If $f = 0$ a.e., we obtain $\mu(f > t) = 0$ for every $t > 0$. Then $\int_X f\, d\mu = \int_0^\infty \mu(f > t)\, dt = 0$ (see Exercise 2.37(3). Conversely, let $\int_X f d\mu = 0$. Then Markov's inequality implies $\mu(f > a) = 0$ for all $a > 0$. Since $\{f > \frac{1}{n}\} \uparrow \{f > 0\}$, we deduce

$$\mu(f > 0) = \lim_{n\to\infty} \mu\Big(f > \frac{1}{n}\Big) = 0.$$

The proof is thus complete. □

The following theorem, usually referred to as the *Monotone Convergence Theorem* or *Beppo Levi's Theorem*, is the first result that justifies passing to the limit under the integral sign.

Theorem 2.45 (Beppo Levi) *Let $f_n : X \to [0, \infty]$ be a sequence of Borel functions such that $f_n \le f_{n+1}$, and set*

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad \forall x \in X.$$

Then

$$\int_X f_n\, d\mu \ \uparrow \ \int_X f\, d\mu.$$

Proof Observe that, in consequence of the assumptions, we have

$$\{f_n > t\} \uparrow \{f > t\} \qquad \forall t > 0.$$

Therefore $\mu(f_n > t) \uparrow \mu(f > t)$ for any $t > 0$. The conclusion follows from Proposition 2.38. □

Combining Proposition 2.23 and Theorem 2.45 we deduce the following result.

Proposition 2.46 *Let $f : X \to [0,\infty]$ be a Borel function. Then there exists a sequence $f_n : X \to [0,\infty)$ such that $(f_n)_n \subset \mathscr{S}_+(X)$, $f_n(x) \uparrow f(x)$ for every $x \in X$ and*

$$\int_X f_n \, d\mu \uparrow \int_X f \, d\mu.$$

Let us state some basic properties of the integral.

Proposition 2.47 *Let $f, g : X \to [0,\infty]$ be Borel functions. Then the following properties hold:*

(i) If $\alpha, \beta \in [0,\infty]$, then $\int_X (\alpha f + \beta g)\, d\mu = \alpha \int_X f \, d\mu + \beta \int_X g \, d\mu$.
(ii) If $f \geq g$, then $\int_X f \, d\mu \geq \int_X g \, d\mu$.

Proof The conclusion of point (i) holds for $f, g \in \mathscr{S}_+(X)$, thanks to Proposition 2.33. To obtain it for Borel functions it suffices to apply Proposition 2.46.

To justify (ii), observe that the trivial inclusion $\{g > t\} \subset \{f > t\}$ implies $\mu(g > t) \leq \mu(f > t)$. The conclusion easily follows (see Exercise 2.37(2)). □

Proposition 2.48 *Let $f_n : X \to [0,\infty]$ be a sequence of Borel functions and let*

$$f(x) = \sum_{n=1}^{\infty} f_n(x) \quad \forall x \in X.$$

Then

$$\sum_{n=1}^{\infty} \int_X f_n \, d\mu = \int_X f \, d\mu.$$

Proof For every n set

$$g_n = \sum_{k=1}^{n} f_k.$$

Then $g_n(x) \uparrow f(x)$ for every $x \in X$. By applying the Monotone Convergence Theorem we get

$$\int_X g_n \, d\mu \to \int_X f \, d\mu.$$

On the other hand (i) of Proposition 2.47 implies

$$\int_X g_n \, d\mu = \sum_{k=1}^{n} \int_X f_k \, d\mu \to \sum_{k=1}^{\infty} \int_X f_k \, d\mu.$$

The thesis follows. □

The following basic result, known as *Fatou's Lemma*, provides a semicontinuity property of the integral.

Lemma 2.49 (Fatou) *Let $f_n : X \to [0, \infty]$ be a sequence of Borel functions and let $f = \liminf_{n\to\infty} f_n$. Then*

$$\int_X f\, d\mu \le \liminf_{n\to\infty} \int_X f_n\, d\mu. \tag{2.14}$$

Proof Setting $g_n(x) = \inf_{k\ge n} f_k(x)$, we have $g_n(x) \uparrow f(x)$ for every $x \in X$. Consequently, by the Monotone Convergence Theorem,

$$\int_X f\, d\mu = \lim_{n\to\infty} \int_X g_n\, d\mu = \sup_{n\in\mathbb{N}} \int_X g_n\, d\mu.$$

On the other hand, since $g_n \le f_k$ for every $k \ge n$, we get

$$\int_X g_n\, d\mu \le \inf_{k\ge n} \int_X f_k\, d\mu.$$

So

$$\int_X f\, d\mu \le \sup_{n\in\mathbb{N}} \inf_{k\ge n} \int_X f_k\, d\mu = \liminf_{n\to\infty} \int_X f_n\, d\mu.$$

The proof is thus complete. □

Corollary 2.50 *Let $f_n : X \to [0, \infty]$ be a sequence of Borel functions converging to f pointwise. If there exists $M \ge 0$ such that*

$$\int_X f_n\, d\mu \le M \qquad \forall n \in \mathbb{N},$$

then $\int_X f\, d\mu \le M$.

Remark 2.51 We can give a version of Theorem 2.45 and Corollary 2.50 that applies to a.e. convergence. In this case, the fact that the limit f is a Borel function is no longer guaranteed (see Example 2.26). This difficulty can be easily overcome by adding the assumption that f is Borel or, else, that the measure is complete.

Exercise 2.52 Taking into account of Remark 2.51, state and prove the analogue of Theorem 2.45 and Corollary 2.50 for a.e. convergence.

Exercise 2.53 Consider the measurable space $(X, \mathscr{P}(X), \delta_{x_0})$, where δ_{x_0} denotes the Dirac measure concentrated at $x_0 \in X$. Show that, for any function $f : X \to [0, \infty]$,

$$\int_X f\, d\delta_{x_0} = f(x_0).$$

Example 2.54 Consider the measurable space $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \mu^\#)$, where $\mu^\#$ denotes the counting measure. Then any sequence $(a_n)_n \subset \overline{\mathbb{R}}$ provides a Borel function $f : n \in \mathbb{N} \mapsto a_n \in \overline{\mathbb{R}}$. Assume $(a_n)_n \subset [0, \infty]$. Since $f(n) = \sum_{k=1}^{\infty} a_k \chi_{\{k\}}(n)$ for every $n \in \mathbb{N}$, by applying Propositions 2.47 and 2.48 we have

$$\int_{\mathbb{N}} f \, d\mu^\# = \sum_{k=1}^{\infty} a_k \int_{\mathbb{N}} \chi_{\{k\}} \, d\mu^\# = \sum_{k=1}^{\infty} a_k \mu^\#(\{k\}) = \sum_{k=1}^{\infty} a_k.$$

Exercise 2.55 Let $(a_{nk})_{n,k\in\mathbb{N}}$ be a sequence in $[0, \infty]$. Show that[5]

$$\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} a_{nk} = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} a_{nk}.$$

Hint. Consider the measure space $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \mu^\#)$ and set $f_k : n \mapsto a_{nk}$. Then $(f_k)_k$ is a sequence of positive Borel functions. Use Proposition 2.48 to conclude.

Exercise 2.56 Let $(a_{nk})_{n,k\in\mathbb{N}}$ be a sequence in $[0, \infty]$ such that, for every $n \in \mathbb{N}$,

$$h \le k \quad \Longrightarrow \quad a_{nh} \le a_{nk}. \tag{2.15}$$

Set, for any $n \in \mathbb{N}$,

$$\lim_{k\to\infty} a_{nk} =: \alpha_n \in [0, \infty]. \tag{2.16}$$

Show that

$$\lim_{k\to\infty} \sum_{n=1}^{\infty} a_{nk} = \sum_{n=1}^{\infty} \alpha_n.$$

Hint. Set $f_k : n \mapsto a_{nk}$ and use Monotone Convergence Theorem.

Example 2.57 The result of Exercise 2.56 can be proved directly by elementary computations. Suppose, first, $\sum_{n=1}^{\infty} \alpha_n < \infty$, and fix $\varepsilon > 0$. Then there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\sum_{n=n_\varepsilon+1}^{\infty} \alpha_n < \varepsilon.$$

Using (2.16), for k sufficiently large, say $k \ge k_\varepsilon$, we have $\alpha_n - \frac{\varepsilon}{n_\varepsilon} < a_{nk}$ for $n = 1, \ldots, n_\varepsilon$. So for every $k \ge k_\varepsilon$

$$\sum_{n=1}^{\infty} a_{nk} \ge \sum_{n=1}^{n_\varepsilon} \alpha_n - \varepsilon > \sum_{n=1}^{\infty} \alpha_n - 2\varepsilon.$$

Since $\sum_{n=1}^{\infty} a_{nk} \le \sum_{n=1}^{\infty} \alpha_n$, the thesis follows.

[5] See footnote 6 on page 17.

A similar argument applies to the case of $\sum_{n=1}^{\infty}\alpha_n = \infty$. The thesis is immediate if one of the values α_n is infinite. Thus, assume $\alpha_n < \infty$ for all n. Given $M > 0$, let $n_M \in \mathbb{N}$ be such that

$$\sum_{n=1}^{n_M}\alpha_n > 2M.$$

For k large enough, say $k \geq k_M$, we have $\alpha_n - \frac{M}{n_M} \leq a_{nk}$ for $n = 1, \ldots, n_M$. Then for every $k \geq k_M$

$$\sum_{n=1}^{\infty} a_{nk} \geq \sum_{n=1}^{n_M} a_{nk} \geq \sum_{n=1}^{n_M}\alpha_n - M > M.$$

Example 2.58 The monotonicity assumption in Exercise 2.56 is essential. Indeed, (2.15) fails for the sequence

$$a_{nk} = \delta_{nk} = \begin{cases} 1 & \text{if } n = k, \\ 0 & \text{if } n \neq k, \end{cases} \qquad \text{[Kronecker delta]}$$

since

$$\lim_{k\to\infty}\sum_{n=1}^{\infty} a_{nk} = 1 \neq 0 = \sum_{n=1}^{\infty}\lim_{k\to\infty} a_{nk}.$$

Exercise 2.59 Let $f, g : X \to [0,\infty]$ be Borel functions. Show that:

1. If $f \leq g$ a.e., then $\int_X f\,d\mu \leq \int_X g\,d\mu$.
2. If $f = g$ a.e., then $\int_X f\,d\mu = \int_X g\,d\mu$.

Exercise 2.60 Show that the monotonicity of the sequence $(f_n)_n$ is an essential hypothesis for Beppo Levi's Theorem.

Hint. Consider $f_n = \chi_{[n,n+1)}$ in $\mathbb{R}$ with the Lebesgue measure.

Exercise 2.61 Give an example to show that the inequality in Fatou's Lemma can be strict.

Hint. Consider $f_{2n} = \chi_{[0,1)}$ and $f_{2n+1}(x) = \chi_{[1,2)}$ in $\mathbb{R}$ with the Lebesgue measure.

Exercise 2.62 Let $(X, \mathscr{E}, \mu)$ be a measure space. Show that the following two statements are equivalent:

1. μ is σ-finite.
2. There exists a μ-summable function $f : X \to [0,\infty]$ such that $f(x) > 0$ for all $x \in X$.

Exercise 2.63 Show that if m denotes the Lebesgue measure on $[0,\infty)$ and $F : [0,\infty) \to [0,\infty]$ is a decreasing function, then

$$\int_0^{\infty} F(t)\,dt = \int_{[0,\infty)} F\,dm.$$

Hint. The result holds for simple functions (see Remark 2.40). For the general case use Proposition 2.23.

2.4.5 Integral of Functions with Variable Sign

Definition 2.64 A Borel function $f : X \to \overline{\mathbb{R}}$ is said to be μ-*summable* if there exist two μ-summable Borel functions $\varphi, \psi : X \to [0, \infty]$ such that

$$f(x) = \varphi(x) - \psi(x) \qquad \forall x \in X. \tag{2.17}$$

In this case, the number

$$\int_X f\, d\mu := \int_X \varphi\, d\mu - \int_X \psi\, d\mu \tag{2.18}$$

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

Remark 2.65 The integral of f is independent of the choice of the functions φ, ψ used to represent f as in (2.17). Indeed, let $\varphi_1, \psi_1 : X \to [0, \infty]$ be μ-summable Borel functions such that

$$f(x) = \varphi_1(x) - \psi_1(x) \qquad \forall x \in X.$$

Then, according to Proposition 2.44, φ, ψ, φ_1 and ψ_1 are a.e. finite, and

$$\varphi(x) + \psi_1(x) = \varphi_1(x) + \psi(x) \qquad \text{a.e.}$$

Therefore, owing to Exercise 2.59(2) and Proposition 2.47, we have

$$\int_X \varphi\, d\mu + \int_X \psi_1\, d\mu = \int_X \varphi_1\, d\mu + \int_X \psi\, d\mu.$$

Since the above integrals are all finite, we deduce

$$\int_X \varphi\, d\mu - \int_X \psi\, d\mu = \int_X \varphi_1\, d\mu - \int_X \psi_1\, d\mu$$

as claimed.

Remark 2.66 Let $f : X \to \overline{\mathbb{R}}$ be a μ-summable function.

1. The positive and negative parts of f

$$f^+(x) = \max\{f(x), 0\}, \quad f^-(x) = \max\{-f(x), 0\}$$

are Borel functions such that $f = f^+ - f^-$. Let $\varphi, \psi : X \to [0, \infty]$ be μ-summable functions verifying (2.17). If $x \in X$ is such that $f(x) \geq 0$, then $f^+(x) = f(x) \leq \varphi(x)$. So $f^+(x) \leq \varphi(x)$ for every $x \in X$ and, recalling Exercise 2.59(1), we deduce that f^+ is μ-summable. Similarly, one can show that f^- is μ-summable. Therefore

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

2. From the above remark we deduce that f is μ-summable if and only if f^+ and f^- are μ-summable. Since $|f| = f^+ + f^-$, it is also true that f is μ-summable if and only if $|f|$ is μ-summable. Moreover,

$$\left|\int_X f\,d\mu\right| \leq \int_X |f|\,d\mu. \tag{2.19}$$

Indeed,

$$\begin{aligned}\left|\int_X f\,d\mu\right| &= \left|\int_X f^+\,d\mu - \int_X f^-\,d\mu\right| \leq \\ &\leq \int_X f^+\,d\mu + \int_X f^-\,d\mu = \int_X |f|\,d\mu.\end{aligned}$$

Remark 2.67 The notion of integral can be further extended allowing infinite values. More precisely, the definition (2.18) does make sense if at least one of the two integrals $\int_X \varphi\,d\mu$, $\int_X \psi\,d\mu$ is finite, but not necessarily both of them. A Borel function $f\colon X \to \overline{\mathbb{R}}$ is said to be μ-*integrable* if at least one of the two functions f^+ and f^- is μ-summable. In this case, we define

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

Notice that $\int_X f\,d\mu \in \overline{\mathbb{R}}$, in general. It follows at once that any Borel function $f : X \to [0, \infty]$ is μ–integrable.

In order to state the analogue of Proposition 2.47 for functions with variable sign, we recall that the sum of two functions taking values in the extended space $\overline{\mathbb{R}}$ may fail to be well defined; thus, we need to assume that at least one of the two functions is finite.

Proposition 2.68 *Let $f, g : X \to \overline{\mathbb{R}}$ be μ-summable functions. Then the following properties hold:*

(i) If f is finite, then, for any $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ is μ-summable and

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

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

Proof (i) Assume first $\alpha, \beta \geq 0$. Since f is finite, so are f^+ and f^-. Then we have $\alpha f + \beta g = (\alpha f^+ + \beta g^+) - (\alpha f^- + \beta g^-)$ and so, by Definition 2.64,

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

The conclusion follows from Proposition 2.47(i). The case when α, β have different signs can be handled similarly.

(ii) Let $f \leq g$ a.e. It is immediate that $f^+ \leq g^+$ and $g^- \leq f^-$ a.e. Then, by Exercise 2.59, we obtain

$$\int_X g\,d\mu = \int_X g^+\,d\mu - \int_X g^-\,d\mu \geq \int_X f^+\,d\mu - \int_X f^-\,d\mu = \int_X f\,d\mu.$$

The proof is thus complete. □

We now proceed to define the integral on a measurable set.

Definition 2.69 Let $f\colon X \to \overline{\mathbb{R}}$ be μ-summable and let $A \in \mathscr{E}$. The *(Lebesgue) integral of f over A with respect to μ* is defined by

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

Remark 2.70 Observe that if $f\colon X \to \overline{\mathbb{R}}$ is μ-summable, so is $\chi_A f$ since $|\chi_A f| \leq |f|$. Taking into account that $f = \chi_A f + \chi_{A^c} f$, from Proposition 2.68(i). we obtain

$$\int_A f\,d\mu + \int_{A^c} f\,d\mu = \int_X f\,d\mu. \tag{2.20}$$

Remark 2.71 Recalling that any measurable set A is itself, in a natural way, a measure space with the σ-algebra $\mathscr{E} \cap A$ (see Remark 1.28), we deduce that it suffices to define the integral over the whole space X to have it automatically defined over any measurable subset A.

Exercise 2.72 Show that, for any μ-summable function $f : X \to \overline{\mathbb{R}}$,

$$\int_A f\,d\mu = \int_X f\,d\mu \llcorner A$$

where $\mu \llcorner A$ is the restriction of μ to A (see Definition 1.26).

If $A \in \mathscr{B}(\mathbb{R}^N)$, in the following we will denote by m the Lebesgue measure on A and we will write $\int_A f(x)\,dm(x)$ simply as

$$\int_A f(x)\,dx$$

or, equivalently, $\int_A f(y)\,dy$, $\int_A f(t)\,dt$ etc. in terms of the new dummy variable of integration y, t etc. If $N = 1$ and I is one of the sets (a, b), $(a, b]$, $[a, b)$, $[a, b]$, we will usually write $\int_I f(x)\,dm(x)$ as

$$\int_a^b f(x)\,dx.$$

Since the Lebesgue measure of a single point is zero, there is no need to specify which of the four sets the integral refers to. Owing to Exercise 2.63, this notation is consistent with the one of the archimedean integral.

Proposition 2.73 *Let $f : X \to \overline{\mathbb{R}}$ be a μ-summable function. Then the following properties hold:*

(i) f is a.e. finite, i.e., the set $\{|f| = \infty\}$ has measure zero.
(ii) If $f = 0$ a.e., then $\int_X f\,d\mu = 0$.
(iii) If $E \in \mathscr{E}$ has measure zero, then $\int_E f\,d\mu = 0$.
(iv) If $\int_A f\,d\mu = 0$ for every $A \in \mathscr{E}$, then $f = 0$ a.e.

Proof Parts (i), (ii) and (iii) follow immediately from Proposition 2.44. Let us prove (iv). Set $A = \{f^+ > 0\}$. Then we have

$$0 = \int_A f\,d\mu = \int_X f^+\,d\mu.$$

Proposition 2.44(ii) implies $f^+ = 0$ a.e. In a similar way we obtain $f^- = 0$ a.e. □

Remark 2.74 In view of the last proposition, the sets of measure zero are negligible in integration. Therefore it is natural to extend the definitions of measurability and summability to include functions f taking values in $\overline{\mathbb{R}}$ which are defined a.e. in X, by saying that such f is Borel if so is $\tilde{f}$, letting $\tilde{f}$ denote the extension of f to zero outside the subset where it is defined; similarly, we say that f is μ-summable if so is $\tilde{f}$. The (Lebesgue) integral of f over X with respect to μ is defined by

$$\int_X f\,d\mu := \int_X \tilde{f}\,d\mu.$$

For instance, one can give the following version of Proposition 2.68(i) that applies to a.e. defined functions: if f and g are μ-summable functions, defined a.e. in X, so is the sum[6] $\alpha f + \beta g$ for any $\alpha,\ \beta \in \mathbb{R}$; furthermore

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

The key result provided by the next proposition is referred to as the *absolute continuity* property of the integral.

Proposition 2.75 *Let* $f : X \to \overline{\mathbb{R}}$ *be* μ*-summable. Then for any* $\varepsilon > 0$ *there exists* $\delta_\varepsilon > 0$ *such that*

$$A \in \mathscr{E} \ \ \&\ \ \mu(A) < \delta_\varepsilon \quad \Longrightarrow \quad \int_A |f|\, d\mu \le \varepsilon. \tag{2.21}$$

Proof Without loss of generality, f may be assumed to be positive. Then

$$f_n(x) := \min\{f(x), n\} \uparrow f(x) \qquad \forall x \in X.$$

Therefore, by Beppo Levi's Theorem, $\int_X f_n\, d\mu \uparrow \int_X f\, d\mu$. So for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that

$$0 \le \int_X (f - f_n)\, d\mu < \frac{\varepsilon}{2} \qquad \forall n \ge n_\varepsilon.$$

Hence, if $\mu(A) < \frac{\varepsilon}{2n_\varepsilon}$, for all $n \ge n_\varepsilon$ we get

$$\int_A f\, d\mu \le \int_A f_{n_\varepsilon}\, d\mu + \int_X (f - f_{n_\varepsilon})\, d\mu < \varepsilon.$$

We have thus obtained the thesis with $\delta_\varepsilon = \frac{\varepsilon}{2n_\varepsilon}$. □

Exercise 2.76 Let $f : X \to \overline{\mathbb{R}}$ be μ-summable. Show that

$$\lim_{n\to\infty} \int_{\{|f|>n\}} |f|\, d\mu = 0.$$

Exercise 2.77 Let $(X, \mathscr{E})$ and $(Y, \mathscr{F})$ be measurable spaces. Given a measurable function $f : X \to Y$ and a measure μ on $\mathscr{E}$, let $f_\sharp\mu$ be the measure on $\mathscr{F}$ defined in Exercise 2.6. Show that if $\varphi : Y \to \overline{\mathbb{R}}$ is $f_\sharp\mu$-summable, then $\varphi \circ f$ is μ-summable and

$$\int_Y \varphi\, d(f_\sharp\mu) = \int_X (\varphi \circ f)\, d\mu.$$

[6] Observe that f and g are a.e. finite owing to Proposition 2.73(i); so the sum $\alpha f + \beta g$ is well defined a.e. in X.

2.5 Convergence of Integrals

We have already obtained two results that allow to take limits under the integral sign, namely Beppo Levi's Theorem and Fatou's Lemma. In this section, we will further analyze the problem. In the following $(X, \mathscr{E}, \mu)$ denotes a generic measure space.

2.5.1 Dominated Convergence

We begin with the following classical result, also known as the *Dominated Convergence Theorem* or *Lebesgue's Theorem*.

Proposition 2.78 (Lebesgue) *Let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions converging to f pointwise. Assume that there exists a μ-summable function $g : X \to [0, \infty]$ such that*

$$|f_n(x)| \le g(x) \qquad \forall x \in X,\ \forall n \in \mathbb{N}. \tag{2.22}$$

Then f_n, f are μ-summable and

$$\lim_{n\to\infty} \int_X f_n \, d\mu = \int_X f \, d\mu. \tag{2.23}$$

Proof We note that f_n, f are μ-summable since they are Borel and, in view of (2.22), $|f(x)| \le g(x)$ for all $x \in X$. Assume, first, $g : X \to [0, \infty)$. Since $g + f_n$ is positive, Fatou's Lemma yields

$$\int_X (g + f)\, d\mu \le \liminf_{n\to\infty} \int_X (g + f_n)\, d\mu = \int_X g \, d\mu + \liminf_{n\to\infty} \int_X f_n \, d\mu.$$

Consequently, since $\int_X g \, d\mu$ is finite, we deduce

$$\int_X f \, d\mu \le \liminf_{n\to\infty} \int_X f_n \, d\mu. \tag{2.24}$$

Similarly,

$$\int_X (g - f)\, d\mu \le \liminf_{n\to\infty} \int_X (g - f_n)\, d\mu = \int_X g \, d\mu - \limsup_{n\to\infty} \int_X f_n \, d\mu.$$

Hence,

$$\int_X f \, d\mu \ge \limsup_{n\to\infty} \int_X f_n \, d\mu. \tag{2.25}$$

The conclusion follows from (2.24) and (2.25).

In the general case $g : X \to [0, \infty]$, consider $E = \{g = \infty\}$. Then (2.23) holds over E^c and, by Proposition 2.44(i), we have $\mu(E) = 0$. Hence, using identity (2.20), we deduce

$$\int_X f_n \, d\mu = \int_{E^c} f_n \, d\mu \to \int_{E^c} f \, d\mu = \int_X f \, d\mu.$$

□

Exercise 2.79 Derive (2.23) when (2.22) is satisfied a.e. and $f_n \xrightarrow{a.e.} f$, with the additional restriction that f is Borel or else that μ is complete.

Exercise 2.80 Let $f, \ g : X \to \overline{\mathbb{R}}$ be Borel functions such that f is μ-summable and g is μ-integrable (see Remark 2.67). Assume that f or g is finite. Show that $f + g$ is μ-integrable and

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

Exercise 2.81 Let $f_n : X \to \overline{\mathbb{R}}$ be Borel functions satisfying, for some μ-summable function $g : X \to \overline{\mathbb{R}}$ and some (Borel) function f,

$$\left.\begin{array}{l} f_n(x) \geq g(x) \\ f_n(x) \uparrow f(x) \end{array}\right\} \qquad \forall x \in X.$$

Show that $f_n, \ f$ are μ-integrable and

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

Exercise 2.82 Let $f_n : X \to \overline{\mathbb{R}}$ be Borel functions satisfying, for some μ-summable function g and some (Borel) function f,

$$\left.\begin{array}{l} f_n(x) \geq g(x) \\ f_n(x) \to f(x) \end{array}\right\} \qquad \forall x \in X.$$

Show that $f_n, \ f$ are μ-integrable and

$$\int_X f \, d\mu \leq \liminf_{n \to \infty} \int_X f_n \, d\mu.$$

Exercise 2.83 Let $f_n : X \to \mathbb{R}$ be Borel functions. Show that if μ is finite and, for some constant M and some (Borel) function f,

$$\left.\begin{array}{l} |f_n(x)| \leq M \\ f_n(x) \to f(x) \end{array}\right\} \qquad \forall x \in X,$$

then f_n, f are μ-summable and

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

Exercise 2.84 Let $f : X \to \overline{\mathbb{R}}$ be a μ-summable function. Show that

$$\lim_{n\to\infty}\int_X |f|^{1/n}\,d\mu = \mu(f \neq 0).$$

Exercise 2.85 Let $f : X \to \overline{\mathbb{R}}$ be a μ-summable function such that $|f| \leq 1$. Show that

$$\lim_{n\to\infty}\int_X |f|^n\,dx = \mu(|f| = 1).$$

Exercise 2.86 Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} 0 & x \leq 0; \\ (x(|\log x| + 1))^{-\frac{1}{n}} & 0 < x \leq 1; \\ (x(\log x + 1))^{-n} & x > 1. \end{cases}$$

Show that:

(i) f_n is summable[7] for every $n \geq 2$.
(ii) $\lim_{n\to\infty}\int_{\mathbb{R}} f_n(x)\,dx = 1$.

Exercise 2.87 Let $f_n : (0, 1) \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{n}{x^{3/2}}\log\left(1 + \frac{x}{n}\right).$$

Show that:

(i) f_n is summable for every $n \geq 1$.
(ii) $\lim_{n\to\infty}\int_0^1 f_n(x)\,dx = 2$.

Exercise 2.88 Let $f_n : (0, \infty) \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{1}{x^{3/2}}\log\left(1 + \frac{x}{n}\right).$$

Show that:

(i) f_n is summable for every $n \geq 1$.
(ii) $\lim_{n\to\infty}\int_0^\infty f_n(x)\,dx = 0$.

---

[7] For simplicity, we often say 'summable' instead of 'm-summable', omitting explicit reference to the Lebesgue measure.

Exercise 2.89 Let $f_n : (0, \infty) \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{1}{x^{3/2}} \arctan \frac{x}{n}, \quad x > 0.$$

Show that:

(i) f_n is summable for every $n \geq 1$.
(ii) $\lim_{n\to\infty} \int_0^\infty f_n(x)dx = 0$.

Exercise 2.90 Let $f_n : (0, 1) \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{n\sqrt{x}}{1 + n^2x^2}.$$

Show that:

(i) $f_n(x) \leq \frac{1}{\sqrt{x}}$ for every $n \geq 1$.
(ii) $\lim_{n\to\infty} \int_0^1 f_n(x)\, dx = 0$.

Exercise 2.91 Let $f_n : (0, \infty) \to \mathbb{R}$ be defined by

$$f_n(x) = \frac{1}{x^{3/2}} \sin \frac{x}{n}.$$

Show that:

(i) f_n is summable for every $n \geq 1$.
(ii) $\lim_{n\to\infty} \int_0^\infty f_n(x)\, dx = 0$.

Exercise 2.92 Compute the limit

$$\lim_{n\to\infty} \int_0^\infty \frac{n}{1 + n\sqrt{x}} \left(\frac{\sin x}{x}\right)^n dx.$$

Exercise 2.93 Given a Borel measure μ on $\mathbb{R}$ and a μ-summable function $f : \mathbb{R} \to \overline{\mathbb{R}}$, set

$$\varphi : \mathbb{R} \to \mathbb{R}, \quad \varphi(x) = \int_{(x,\infty)} f\, d\mu.$$

(i) Show that if the measure μ is such that

$$\mu(\{x\}) = 0 \quad \forall x \in \mathbb{R}, \tag{2.26}$$

then φ is continuous.
(ii) Give an example to show that φ may fail to be continuous without the assumption (2.26), in general.

Exercise 2.94 Given a Borel measure μ on $\mathbb{R}$ and a Borel function $f : \mathbb{R} \to [0, \infty]$, show that the function

$$\varphi : \mathbb{R} \to \mathbb{R} \cup \{\infty\}, \quad \varphi(x) = \int_{(x,\infty)} f \, d\mu$$

is lower semicontinuous.

2.5.2 Uniform Summability

Definition 2.95 A sequence of μ-summable functions $f_n : X \to \overline{\mathbb{R}}$ is said to be *uniformly μ-summable* if it satisfies the following:

(a) For any $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that

$$\int_A |f_n| \, d\mu \leq \varepsilon \ \ \forall n \in \mathbb{N}, \quad \forall A \in \mathscr{E} \text{ with } \mu(A) < \delta_\varepsilon. \tag{2.27}$$

(b) For any $\varepsilon > 0$ there exists $B_\varepsilon \in \mathscr{E}$ such that

$$\mu(B_\varepsilon) < \infty \qquad \text{and} \qquad \int_{B_\varepsilon^c} |f_n| \, d\mu < \varepsilon \ \ \forall n \in \mathbb{N}. \tag{2.28}$$

Remark 2.96 A sequence $(f_n)_n$ satisfies (a) of Definition 2.95 if and only if

$$\lim_{\mu(A) \to 0} \int_A |f_n| \, d\mu = 0 \quad \text{uniformly with respect to } \ n.$$

Remark 2.97 Properties (a) and (b) of Definition 2.95 hold for a single μ-summable function f. Indeed, (a) follows directly from Proposition 2.75. To prove (b), observe that, by Markov's inequality, the sets $\{|f| > \frac{1}{n}\}$ have finite measure and, by Lebesgue's Theorem,

$$\int_{\{|f| \leq \frac{1}{n}\}} |f| \, d\mu = \int_X \chi_{\{|f| \leq \frac{1}{n}\}} |f| \, d\mu \to 0 \ \text{ as } n \to \infty.$$

The following theorem, due to Vitali, uses the notion of uniform summability to provide another sufficient condition to take limits under the integral sign.

Theorem 2.98 (Vitali) *Let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of uniformly μ-summable functions. If $(f_n)_n$ converges pointwise to an a.e. finite limit f, then f is μ-summable and*

$$\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu. \tag{2.29}$$

Proof Assume first that f and f_n are finite. Given $\varepsilon > 0$, let $\delta_\varepsilon > 0$, $B_\varepsilon \in \mathscr{E}$ be such that (2.27) and (2.28) hold. Since, by Theorem 2.27, $f_n \xrightarrow{a.u.} f$ in B_ε, there exists a measurable set $A_\varepsilon \subset B_\varepsilon$ such that $\mu(A_\varepsilon) < \delta_\varepsilon$ and

$$f_n \to f \text{ uniformly in } B_\varepsilon \setminus A_\varepsilon. \tag{2.30}$$

So

$$\begin{aligned}\int_{B_\varepsilon} |f_n - f|\, d\mu &= \int_{A_\varepsilon} |f_n - f|\, d\mu + \int_{B_\varepsilon \setminus A_\varepsilon} |f_n - f|\, d\mu \\ &\le \int_{A_\varepsilon} |f_n|\, d\mu + \int_{A_\varepsilon} |f|\, d\mu + \mu(B_\varepsilon) \sup_{B_\varepsilon \setminus A_\varepsilon} |f_n - f|.\end{aligned}$$

Notice that $\int_{A_\varepsilon} |f_n|\, d\mu \le \varepsilon$, $\int_{B_\varepsilon^c} |f_n|\, d\mu \le \varepsilon$ by (2.27) and (2.28). Also, owing to Corollary 2.50, $\int_{A_\varepsilon} |f|\, d\mu \le \varepsilon$, $\int_{B_\varepsilon^c} |f|\, d\mu \le \varepsilon$. Thus,

$$\begin{aligned}\int_X |f_n - f|\, d\mu &\le \int_{B_\varepsilon^c} |f|\, d\mu + \int_{B_\varepsilon^c} |f_n|\, d\mu + \int_{B_\varepsilon} |f_n - f|\, d\mu \\ &\le 4\varepsilon + \mu(B_\varepsilon) \sup_{B_\varepsilon \setminus A_\varepsilon} |f_n - f|.\end{aligned}$$

Since $\mu(B_\varepsilon) < \infty$, by (2.30) we deduce

$$\int_X |f_n - f|\, d\mu \to 0. \tag{2.31}$$

Then $f_n - f$ is μ-summable; consequently, since $f = (f - f_n) + f_n$, by Proposition 2.68(i) we deduce that f is μ-summable. The conclusion follows by (2.19) and (2.31).

In the general case when f is a.e. finite and $f_n : X \to \overline{\mathbb{R}}$, we consider the sets

$$E_0 = \{|f| = \infty\} \qquad E_n = \{|f_n| = \infty\} \quad \forall n \ge 1.$$

Then $\mu(E_0) = 0$ by hypothesis and $\mu(E_n) = 0$ for all $n \ge 1$ owing to Proposition 2.73. Therefore $E = \cup_{n \ge 0} E_n$ is also a zero-measure set and (2.29) holds on E^c. So

$$\int_X f_n\, d\mu = \int_{E^c} f_n\, d\mu \to \int_{E^c} f\, d\mu = \int_X f\, d\mu,$$

and this proves the theorem in the general case. □

Exercise 2.99 Derive (2.29) when $f_n \xrightarrow{a.e.} f$ with the additional restriction that f is Borel or else that μ is complete.

Exercise 2.100 Give an example to show that the thesis of Theorem 2.98 may fail without the assumption that 'f is a.e. finite'.

Hint. In $\mathbb{N}$ with the counting measure $\mu^{\#}$ consider the sequence of functions $f_n : \mathbb{N} \to \mathbb{R}$, $f_n = n\chi_{\{1\}} - n\chi_{\{2\}}$. Show that $(f_n)_n$ is uniformly $\mu^{\#}$-summable, however its pointwise limit fails to be $\mu^{\#}$-summable.

For finite measures, (b) of Definition 2.95 is always satisfied by taking $B_\varepsilon = X$; hence we obtain the following corollary.

Corollary 2.101 *Assume $\mu(X) < \infty$ and let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of μ-summable functions satisfying (a) of Definition 2.95 and converging pointwise to an a.e. finite function f. Then f is μ-summable and*

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

Exercise 2.102 Give an example to show that when $\mu(X) = \infty$ the condition (b) of Definition 2.95 is essential to derive Vitali's Theorem.

Hint. Consider $f_n = \chi_{[n,n+1)}$ in $\mathbb{R}$ with the Lebesgue measure.

Remark 2.103 We point out that Vitali's Theorem can be regarded as a generalization of Lebesgue's Dominated Convergence Theorem. Indeed, let $f_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions satisfying (2.22) for some μ-summable function g. Since, by Remark 2.97, properties (a) and (b) of Definition 2.95 hold for a single function g, it immediately follows that $(f_n)_n$ is uniformly μ-summable. The converse is not true, in general, i.e., a uniformly μ-summable sequence may fail to be dominated. To see this, consider the sequence

$$f_n = n\chi_{[\frac{1}{n}, \frac{1}{n} + \frac{1}{n^2})}$$

defined in $\mathbb{R}$ with the Lebesgue measure. Since $\int_{\mathbb{R}} f_n \, dx = \frac{1}{n}$, then the sequence $(f_n)_n$ is uniformly summable. On the other hand

$$\sup_n f_n = g := \sum_{n=1}^{\infty} n\chi_{[\frac{1}{n}, \frac{1}{n} + \frac{1}{n^2})}$$

and

$$\int_{\mathbb{R}} g \, dx = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

Consequently, $(f_n)_n$ cannot be dominated by any summable function.

2.5.3 Integrals Depending on a Parameter

Let $(X, \mathscr{E}, \mu)$ be a measure space. In this section we shall see how to differentiate the integral over X of a function $f(x, y)$ depending on the extra variable y, which is called a parameter. We begin with a continuity result.

Proposition 2.104 *Let (Y, d) be a metric space, $y_0 \in Y$, U a neighborhood of y_0 and*

$$f : X \times Y \to \mathbb{R}$$

a function such that

(a) The map $x \mapsto f(x, y)$ is Borel for every $y \in Y$.
(b) The map $y \mapsto f(x, y)$ is continuous at y_0 for every $x \in X$.
(c) For some μ-summable function $g : X \to [0, \infty]$ we have

$$|f(x, y)| \le g(x) \qquad \forall x \in X, \ \forall y \in U.$$

Then $\Phi(y) := \int_X f(x, y)\, d\mu(x)$ is continuous at y_0.

Proof Let $(y_n)_n$ be a sequence in Y that converges to y_0. Suppose, further, $y_n \in U$ for every $n \in \mathbb{N}$. Then

$$\forall x \in X \qquad \begin{cases} f(x, y_n) \to f(x, y_0) & \text{as} \quad n \to \infty \\ |f(x, y_n)| \le g(x) & \forall n \in \mathbb{N}. \end{cases}$$

Therefore, by Lebesgue's Theorem,

$$\int_X f(x, y_n)\, d\mu(x) \longrightarrow \int_X f(x, y_0)\, d\mu(x) \quad \text{as} \quad n \to \infty,$$

and the conclusion follows from the arbitrariness of $(y_n)_n$. □

Exercise 2.105 Let $p > 0$ be fixed. For $t > 0$ define

$$f_t(x) = \frac{1}{t} x^p e^{-\frac{x}{t}} \quad x \in [0, 1].$$

For which values of p does each of the following statement hold true?

(a) $f_t \xrightarrow{a.e.} 0$ as $t \to 0$.
(b) $f_t \to 0$ uniformly in $[0, 1]$ as $t \to 0$.
(c) $\int_0^1 f_t(x)\, dx \longrightarrow 0$ as $t \to 0$.

For differentiability, we shall restrict the analysis to a real parameter.

Proposition 2.106 *Let $f : X \times (a, b) \to \mathbb{R}$ be a function such that*

(a) The map $x \mapsto f(x, y)$ is summable for every $y \in (a, b)$.
(b) The map $y \mapsto f(x, y)$ is differentiable in (a, b) for every $x \in X$.
(c) For some μ-summable function $g : X \to [0, \infty]$ we have

$$\sup_{a<y<b} \left|\frac{\partial f}{\partial y}(x, y)\right| \leq g(x) \qquad \forall x \in X.$$

Then $\Phi(y) := \int_X f(x, y)\, d\mu(x)$ is differentiable in (a, b) and

$$\Phi'(y) = \int_X \frac{\partial f}{\partial y}(x, y)\, d\mu(x) \qquad \forall y \in (a, b).$$

Proof We note, first, that the function $x \mapsto \frac{\partial f}{\partial y}(x, y)$ is Borel for every $y \in (a, b)$ because

$$\frac{\partial f}{\partial y}(x, y) = \lim_{n\to\infty} n\Big[f\Big(x, y + \frac{1}{n}\Big) - f(x, y)\Big] \qquad \forall (x, y) \in X \times (a, b).$$

Now, fix $y_0 \in (a, b)$ and let $(y_n)_n$ be a sequence in (a, b) converging to y_0. Then

$$\frac{\Phi(y_n) - \Phi(y_0)}{y_n - y_0} = \int_X \underbrace{\frac{f(x, y_n) - f(x, y_0)}{y_n - y_0}}_{\stackrel{n\to\infty}{\longrightarrow} \frac{\partial f}{\partial y}(x, y_0)} d\mu(x)$$

and

$$\left|\frac{f(x, y_n) - f(x, y_0)}{y_n - y_0}\right| \leq g(x) \qquad \forall x \in X,\ \forall n \in \mathbb{N}$$

thanks to the mean value theorem. Therefore Lebesgue's Theorem yields

$$\frac{\Phi(y_n) - \Phi(y_0)}{y_n - y_0} \quad\longrightarrow\quad \int_X \frac{\partial f}{\partial y}(x, y_0)\, d\mu(x) \quad \text{as} \quad n \to \infty.$$

Since $(y_n)_n$ is arbitrary, the conclusion follows. □

Remark 2.107 Note that assumption (b) of Proposition 2.106 must be satisfied on the whole interval (a, b) (not just a.e.) in order to be able to differentiate under the integral sign. Indeed, for $X = (a, b) = (0, 1)$, consider

$$f(x, y) = \begin{cases} 1 & \text{if } y \geq x, \\ 0 & \text{if } y < x. \end{cases}$$

Then $\frac{\partial f}{\partial y}(x, y) = 0$ for all $y \neq x$, but

$$\Phi(y) = \int_0^1 f(x, y)\, dx = y,$$

by which $\Phi'(y) = 1$.

Example 2.108 Let us compute the integral

$$\Phi(y) := \int_0^\infty e^{-x^2 - \frac{y^2}{x^2}}\, dx, \quad y \in \mathbb{R}.$$

Observe that

$$\begin{aligned} \left| \frac{\partial}{\partial y} e^{-x^2 - \frac{y^2}{x^2}} \right| &= \frac{2y}{x^2} e^{-x^2 - \frac{y^2}{x^2}} \\ &= \frac{2e^{-x^2}}{y} \underbrace{\frac{y^2}{x^2} e^{-\frac{y^2}{x^2}}}_{\leq 1/e} \leq \frac{2e^{-x^2}}{r} \quad \text{for } y \geq r > 0,\ \forall x > 0. \end{aligned}$$

So, for any $y > 0$,

$$\begin{aligned} \Phi'(y) &= -\int_0^\infty \frac{2y}{x^2} e^{-x^2 - \frac{y^2}{x^2}}\, dx \\ &\overset{t=y/x}{=} -2\int_0^\infty y\, \frac{t^2}{y^2} e^{-t^2 - \frac{y^2}{t^2}}\, \frac{y}{t^2}\, dt = -2\Phi(y). \end{aligned}$$

Since

$$\int_0^\infty e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2},$$

solving the Cauchy problem

$$\begin{cases} \Phi'(y) = -2\Phi(y),\ y > 0, \\ \lim\limits_{y \to 0^+} \Phi(y) = \dfrac{\sqrt{\pi}}{2}, \end{cases}$$

and recalling that Φ is an even function, we obtain

$$\Phi(y) = \frac{\sqrt{\pi}}{2} e^{-2|y|}.$$

Example 2.109 Applying Lebesgue's Theorem to the counting measure on $\mathbb{N}$, we shall compute

$$\lim_{n\to\infty} n \sum_{i=1}^{\infty} \sin\left(\frac{2^{-i}}{n}\right) = 1.$$

Indeed, observe that

$$f_n(i) := n \sin\left(\frac{2^{-i}}{n}\right)$$

satisfies $|f_n(i)| \le 2^{-i}$. Then by Lebesgue's Theorem we have

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} f_n(i) = \sum_{i=1}^{\infty} \lim_{n\to\infty} f_n(i) = \sum_{i=1}^{\infty} 2^{-i} = 1.$$

Exercise 2.110 Let us compute the limit

$$\lim_{R\to\infty} \int_0^R \frac{\sin x}{x}\, dx$$

proceeding as follows.

(i) Show that the above limit exists.

Hint. Observe that for every $R > \frac{\pi}{2}$ we have

$$\begin{aligned}\int_0^R \frac{\sin x}{x}\, dx &= \int_0^{\pi/2} \frac{\sin x}{x}\, dx - \frac{\cos R}{R} - \int_{\pi/2}^R \frac{\cos x}{x^2}\, dx \\ &\to \int_0^{\pi/2} \frac{\sin x}{x}\, dx - \int_{\pi/2}^{\infty} \frac{\cos x}{x^2}\, dx\end{aligned}$$

as $R \to \infty$, where the last convergence follows by Lebesgue's Theorem.

(ii) Show that

$$\Phi(t) := \int_0^{\infty} e^{-tx} \frac{\sin x}{x}\, dx$$

is differentiable for all $t > 0$.

Hint. Use that

$$e^{-tx} \le e^{-rx} \quad \forall t \ge r > 0,\ \forall x > 0.$$

(iii) Compute $\Phi'(t)$ for $t \in]0, \infty[$.

Hint. Proceed as in Example 2.108 using the following indefinite integral

$$\int e^{-tx}\sin x\,dx = -\frac{t\sin x+\cos x}{1+t^2}e^{-tx}+c, \quad c\in\mathbb{R}.$$

(iv) Compute $\Phi(t)$ for all $t\in]0,\infty[$.

(v) Setting $I=\lim_{R\to\infty}\int_0^R \frac{\sin x}{x}\,dx$, show that $\lim_{t\to 0^+}\Phi(t)=I$ and conclude that

$$I=\frac{\pi}{2}.$$

Hint. Observe that

$$\begin{aligned}\Phi(t) &= \int_0^{\pi/2} e^{-tx}\frac{\sin x}{x}\,dx - \int_{\pi/2}^{\infty}\frac{1+tx}{x^2}e^{-tx}\cos x\,dx\\ &\to \int_0^{\pi/2}\frac{\sin x}{x}\,dx - \int_{\pi/2}^{\infty}\frac{\cos x}{x^2}\,dx\end{aligned}$$

as $t\to 0^+$, where the last convergence follows by Lebesgue's Theorem.

2.6 Miscellaneous Exercises

Exercise 2.111 Given a Borel function $f:\mathbb{R}\to\mathbb{R}$, show that the function $g:\mathbb{R}\to\mathbb{R}$ defined by

$$g(x)=\begin{cases}\log(f(x)) & \text{if } f(x)>0\\ 0 & \text{if } f(x)\le 0\end{cases}$$

is also Borel.

Exercise 2.112 Show that the following functions $f:\mathbb{R}\to\mathbb{R}$ are Borel:

$$f(x)=\begin{cases}e^{\frac{1}{x}} & \text{if } x\neq 0\\ 0 & \text{if } x=0\end{cases},$$

$$f(x)=\begin{cases}\dfrac{1}{1+x^2} & \text{if } x\in\mathbb{Q}\\ 0 & \text{if } x\in\mathbb{R}\setminus\mathbb{Q}\end{cases},$$

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \in \bigcup_{n=1}^{\infty} \left[n, n+\frac{1}{2}\right] \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 2.113 Let $(X, \mathscr{E}, \mu)$ be a measure space and let $f : X \to \overline{\mathbb{R}}$ be a μ-summable function. Set

$$E_n = \{x \in X \mid n \le |f(x)| \le n+1\} \quad \forall n \ge 0.$$

Show that:

(i) $\frac{1}{n+1} \int_{E_n} |f| \, d\mu \le \mu(E_n) \le \frac{1}{n} \int_{E_n} |f| \, d\mu.$
(ii) $\lim_{n\to\infty} n\, \mu(E_n) = 0.$

Exercise 2.114 Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous and bijective function.

(i) Show that $f(E) \in \mathscr{B}(\mathbb{R})$ for every $E \in \mathscr{B}(\mathbb{R})$.
(*Hint*. First prove that $f(K) \in \mathscr{B}(\mathbb{R})$ for every compact set $K \subset \mathbb{R}$).
(ii) Setting $\mu(E) = m(f(E))$ for any $E \in \mathscr{B}(\mathbb{R})$, show that μ is a σ-finite measure on $\mathscr{B}(\mathbb{R})$.
(iii) Show that if $\varphi : \mathbb{R} \to \overline{\mathbb{R}}$ is a μ-summable function, then $\varphi \circ f^{-1}$ is summable and

$$\int_{\mathbb{R}} \varphi \, d\mu = \int_{\mathbb{R}} (\varphi \circ f^{-1}) \, dx.$$

Exercise 2.115 Compute the following limits:

$$\lim_{n\to\infty} \int_0^\infty n \frac{\sin^2 x}{x} e^{-nx} dx, \qquad \lim_{n\to\infty} \int_0^\infty \frac{\sin^n x}{x^3 + x^4} dx$$

$$\lim_{n\to\infty} \int_0^\infty \frac{1}{x + \sqrt{x}} e^{-\frac{x}{n}} dx, \qquad \lim_{n\to\infty} \int_0^\infty \frac{1}{\sqrt{x}} \frac{n}{1 + nx^2} dx,$$

$$\lim_{n\to\infty} \int_0^\infty \frac{x^n}{1 + x^n} \frac{1}{1 + x^2} dx, \qquad \lim_{n\to\infty} \int_0^1 \frac{\arctan(nx)}{x} dx,$$

$$\lim_{n\to\infty} \int_0^\infty e^{-nx} \frac{nx^2}{1 + nx} dx, \qquad \lim_{n\to\infty} \int_0^\infty \frac{n \log^4 x}{n + nx + x^2} dx,$$

$$\lim_{n\to\infty} \int_0^\infty \frac{1}{\sqrt{x}(1 + n^2 x^n)} dx, \qquad \lim_{n\to\infty} \int_{-\infty}^\infty \frac{n}{x(1 + x^2)} \sin \frac{x}{n} dx,$$

$$\lim_{n\to\infty} \int_0^\infty \frac{\arctan(nx)}{n^2 x + x^n} dx, \qquad \lim_{n\to\infty} \int_0^\infty \frac{x^n}{1 + x^{n+1}} e^{-\frac{x}{n}} dx,$$

$$\lim_{n\to\infty}\int_0^\infty \log^3 x \sin e^{-nx}\,dx, \qquad \lim_{n\to\infty}\int_0^\infty \frac{x}{\sqrt[n]{1+x^{3n}}}\,dx,$$

$$\lim_{n\to\infty}\int_{-\infty}^\infty \frac{1}{1+e^{nx}}e^{-\frac{x^2}{n}}\,dx, \qquad \lim_{n\to\infty}\int_0^\infty \frac{n}{1+n\sqrt{x}}\arctan\frac{n}{x^2}\,dx.$$

Exercise 2.116 Let $(X, \mathscr{E}, \mu)$ be a measure space and let $f, g : X \to \overline{\mathbb{R}}$ be μ-summable functions. Show that

$$\lim_{n\to\infty}\int_X \sqrt[n]{|f|^n + |g|^n}\,d\mu = \int_X \max\{|f|, |g|\}\,d\mu.$$

Exercise 2.117 Compute the following limit depending on the parameter $\alpha > 0$:

$$\lim_{n\to\infty}\int_0^\infty \sin\frac{1}{n+x^\alpha}dx.$$

Exercise 2.118 Let $\alpha > 0$ and let $f_n : (0, \infty) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \left(\sin\frac{1}{\sqrt{x}}\right)^n x^{n\alpha} \quad \forall x > 0.$$

Show that:

(a) If $\alpha \in (0, \frac{1}{2})$, then f_n is summable for n large enough.
(b) If $\alpha \in (0, \frac{1}{2})$, then $\lim_{n\to\infty}\int_0^\infty f_n(x)\,dx = 0$.
(c) f_n fails to be summable if $\alpha \geq \frac{1}{2}$.

Exercise 2.119 Compute the following limit depending on the parameter $\alpha \in \mathbb{R}$:

$$\lim_{n\to\infty}\int_0^{2\pi} \frac{n^2}{t^\alpha}\left(1 - \cos\frac{t}{n}\right)dt.$$

Exercise 2.120 Compute the following limit depending on the parameter $\alpha > 0$:

$$\lim_{n\to\infty}\int_0^\infty \frac{1}{x^\alpha + x^n}\,dx.$$

Chapter 3
L^p Spaces

As we observed in Chap. 2, the family of all μ-summable functions on a measure space $(X, \mathscr{E}, \mu)$ can be given the structure of a linear space. In this chapter, we study the so-called Lebesgue spaces, that are spaces of Borel functions $f : X \to \overline{\mathbb{R}}$ in which

$$d(f, g) = \int_X |f - g|^p \, d\mu \tag{3.1}$$

defines a distance, completeness being the crucial property we are interested in.

In the previous chapter, we defined several kinds of convergence for function sequences. We now complete the picture introducing convergence in measure and in the metric (3.1), and study the connections between different notions of convergence.

Among all L^p spaces, L^2 is the only one such that the product of any two of its elements is a summable function. Such a property makes of L^2 a Hilbert space—a functional analytic structure that will be studied in Chap. 5.

A more detailed analysis of L^p is possible when $X = \mathbb{R}^N$ and μ is a Radon measure. In this case, the special role played by continuous functions yields useful density results.

3.1 The Spaces $\mathscr{L}^p(X, \mu)$ and $L^p(X, \mu)$

Let $(X, \mathscr{E}, \mu)$ be a measure space. For any $p \in [1, \infty)$ and any Borel function $f : X \to \overline{\mathbb{R}}$ we define

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{1/p}.$$

Let $\mathscr{L}^p(X, \mu) = \mathscr{L}^p(X, \mathscr{E}, \mu)$ denote the class of all Borel functions f for which $\|f\|_p < \infty$.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_3

Remark 3.1 It is easy to check that $\mathscr{L}^p(X,\mu)$ is closed under the following operations: the sum of two functions (provided that at least one is finite valued) and multiplication of a function by a real number. Indeed,

$$\alpha \in \mathbb{R},\ f \in \mathscr{L}^p(X,\mu) \quad \Longrightarrow \quad \alpha f \in \mathscr{L}^p(X,\mu) \ \ \& \ \ \|\alpha f\|_p = |\alpha|\,\|f\|_p.$$

Moreover, if $f, g \in \mathscr{L}^p(X,\mu)$ and $f : X \to \mathbb{R}$, then we have[1]

$$|f(x)+g(x)|^p \leq 2^{p-1}(|f(x)|^p + |g(x)|^p) \qquad \forall x \in X,$$

and so $f+g \in \mathscr{L}^p(X,\mu)$.

Example 3.2 Consider the measure space $(\mathbb{N}, \mathscr{P}(\mathbb{N}), \mu^{\#})$, where $\mu^{\#}$ denotes the counting measure. Then we will use the notation ℓ^p for space $\mathscr{L}^p(\mathbb{N},\mu^{\#})$. Recalling Example 2.54, we have

$$\ell^p = \left\{ (x_n)_n \;\middle|\; x_n \in \mathbb{R},\ \sum_{n=1}^{\infty} |x_n|^p < \infty \right\},$$

and, for any sequence $(x_n)_n \in \ell^p$,

$$\|(x_n)_n\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}.$$

Observe that

$$1 \leq p \leq q \quad \Longrightarrow \quad \ell^p \subset \ell^q.$$

Indeed, let $(x_n)_n \in \ell^p$. Since $\sum_{n=1}^{\infty} |x_n|^p < \infty$, we get that $(x_n)_n$ is bounded, say $|x_n| \leq M$ for all $n \in \mathbb{N}$. Then $|x_n|^q \leq M^{q-p}|x_n|^p$. So $\sum_{n=1}^{\infty} |x_n|^q < \infty$.

Example 3.3 Consider the Lebesgue measure m on $(0,1]$. Let us set, for any $\alpha \in \mathbb{R}$,

$$f_\alpha(x) = x^\alpha \qquad \forall x \in (0,1].$$

Then $f_\alpha \in \mathscr{L}^p((0,1],m)$ if and only if $\alpha p + 1 > 0$. Thus, $\mathscr{L}^p((0,1],m)$ fails to be an algebra. For instance, $f_{-1/2} \in \mathscr{L}^1((0,1],m)$ but $f_{-1} = f_{-1/2}^2 \notin \mathscr{L}^1((0,1],m)$.

We have already observed that $\|\cdot\|_p$ is positively homogeneous of degree one. However, $\|\cdot\|_p$ is not a norm[2] on $\mathscr{L}^p(X,\mu)$, in general.

In order to construct a vector space on which $\|\cdot\|_p$ is a norm, let us consider the following equivalence relation in $\mathscr{L}^p(X,\mu)$:

$$f \sim g \quad \Longleftrightarrow \quad f(x) = g(x) \text{ a.e. in } X. \tag{3.2}$$

[1] Since $\varphi(t) = |t|^p$ is convex on $\mathbb{R}$, we have that $\left|\frac{a+b}{2}\right|^p \leq \frac{|a|^p+|b|^p}{2}$ for all $a, b \in \mathbb{R}$.

[2] See Definition 6.1.

We denote by $L^p(X,\mu) = L^p(X,\mathscr{E},\mu)$ the quotient space $\mathscr{L}^p(X,\mu)/\sim$. Thus the elements of $L^p(X,\mu)$ are equivalence classes of Borel functions. For any $f \in \mathscr{L}^p(X,\mu)$ let $\widetilde{f}$ be the equivalence class associated to f. It follows at once that $L^p(X,\mu)$ is a vector space. Indeed, the precise definition of the sum of two elements $\widetilde{f_1},\ \widetilde{f_2} \in L^p(X,\mu)$ is the following: let g_1, g_2 be two 'representatives' of $\widetilde{f_1}$ and $\widetilde{f_2}$, respectively (that is, $g_1 \in \widetilde{f_1}$, $g_2 \in \widetilde{f_2}$), such that $g_1,\ g_2$ are everywhere finite (such representatives exist by Proposition 2.73(i)). Then $\widetilde{f_1} + \widetilde{f_2}$ is the equivalence class of $g_1 + g_2$.

To introduce a norm on $L^p(X,\mu)$, we set

$$\|\widetilde{f}\|_p = \|f\|_p \qquad \forall \widetilde{f} \in L^p(X,\mu).$$

It is immediate to realize that the above definition is independent of the element f which is chosen in the class $\widetilde{f}$. Then, since the zero element of $L^p(X,\mu)$ is the class consisting of all functions vanishing almost everywhere, it is clear that $\|\widetilde{f}\|_p = 0$ if and only if $\widetilde{f} = 0$. To simplify notation, we will hereafter identify $\widetilde{f}$ with f and and we will talk about 'functions in $L^p(X,\mu)$' when there is no danger of confusion, with the understanding that we regard equivalent functions (i.e., functions differing only on a zero-measure set) as identical elements of the space $L^p(X,\mu)$.

In order to check that $\|\cdot\|_p$ is a norm on $L^p(X,\mu)$, we need only to verify that $\|\cdot\|_p$ is sublinear. First we derive two classical inequalities (Hölder's inequality and Minkowski's inequality) that play an essential role in real analysis.

Definition 3.4 Two numbers $p,\ p' \in (1,\infty)$ are called *conjugate exponents* if

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Note that $p' = \frac{p}{p-1}$ and that 2 is self-conjugate.

Proposition 3.5 (Hölder's inequality) *Let $p,\ p' \in (1,\infty)$ be conjugate exponents and $f,\ g : X \to \overline{\mathbb{R}}$ Borel functions. Then*[3]

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_{p'}.$$

Moreover, equality holds if and only if $|f|^p = \alpha |g|^{p'}$ a.e. for some $\alpha \geq 0$.

Proof The inequality is obvious if $\|f\|_p = 0$ or $\|g\|_{p'} = 0$; indeed in such a case $fg = 0$ a.e. by Proposition 2.44, and so $\|fg\|_1 = 0$. The inequality is also obvious if the right hand side is infinite. Thus, we may assume that $\|f\|_p$ and $\|g\|_{p'}$ are both finite and different from zero. Set

$$F(x) = \frac{|f(x)|}{\|f\|_p} \qquad G(x) = \frac{|g(x)|}{\|g\|_{p'}} \qquad \forall x \in X.$$

[3] As usual, in the following we will adopt the convention $0 \cdot \pm\infty = \pm\infty \cdot 0 = 0$.

Then, by Young's inequality (F.3),

$$F(x)G(x) \leq \frac{\big(F(x)\big)^p}{p} + \frac{\big(G(x)\big)^{p'}}{p'} \qquad \forall x \in X. \tag{3.3}$$

Integrating over X with respect to μ yields

$$\frac{\int_X |fg|\,d\mu}{\|f\|_p\,\|g\|_{p'}} = \int_X FG\,d\mu \leq \frac{1}{p}\int_X F^p\,d\mu + \frac{1}{p'}\int_X G^{p'}\,d\mu = 1. \tag{3.4}$$

Equality holds in (3.4) if and only if equality holds in (3.3) for almost every $x \in X$, i.e., recalling Example F.2, if $F^p = G^{p'}$ almost everywhere. □

Corollary 3.6 *Let $\mu(X) < \infty$. If $1 \leq p < q < \infty$, then*

$$L^q(X,\mu) \subset L^p(X,\mu)$$

and

$$\|f\|_p \leq (\mu(X))^{\frac{1}{p}-\frac{1}{q}}\|f\|_q \qquad \forall f \in L^q(X,\mu). \tag{3.5}$$

Proof Let $f \in L^q(X,\mu)$. Then $|f|^p \in L^{\frac{q}{p}}(X,\mu)$. Applying Hölder's inequality to $|f|^p$ and $g(x) = 1$ with exponents $\frac{q}{p}$ and $(1-\frac{p}{q})^{-1}$, respectively, we obtain

$$\int_X |f|^p\,d\mu \leq (\mu(X))^{1-\frac{p}{q}}\Big(\int_X |f|^q\,d\mu\Big)^{\frac{p}{q}}.$$

The conclusion follows. □

Next exercise provides a generalization of Hölder's inequality.

Exercise 3.7 Let $f_1, f_2, \ldots, f_k : X \to \overline{\mathbb{R}}$ be Borel functions and $p,\ p_1, \ldots, p_k \in (1,\infty)$ such that $f_i \in L^{p_i}(X,\mu)$ and

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_k}.$$

Then $f_1 f_2 \ldots f_k \in L^p(X,\mu)$ and

$$\|f_1 f_2 \ldots f_k\|_p \leq \|f_1\|_{p_1}\|f_2\|_{p_2}\ldots\|f_k\|_{p_k}.$$

Hint. Consider the functions $|f_i|^p \in L^{p_i/p}(X,\mu)$ and proceed by induction on k using Hölder's inequality.

Exercise 3.8 (*interpolation inequality*[4]) Let $1 \le p < r < q < \infty$ and $f \in L^p(X, \mu) \cap L^q(X, \mu)$. Then $f \in L^r(X, \mu)$ and

$$\|f\|_r \le \|f\|_p^\theta \, \|f\|_q^{1-\theta}$$

where $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$.

Hint. Apply the result of Exercise 3.7 to the functions $|f|^\theta$ and $|f|^{1-\theta}$ with exponents $\frac{p}{\theta}$ and $\frac{q}{1-\theta}$, respectively.

Exercise 3.9 Let $\mu(X) < \infty$ and $1 \le p < \infty$. Show that if $f : X \to \overline{\mathbb{R}}$ is a Borel function such that $fg \in L^1(X, \mu)$ for every $g \in L^p(X, \mu)$, then $f \in L^q(X, \mu)$ for all $q \in [1, p')$, where p' is the conjugate exponent[5] of p.
Hint. Observe that $f \in L^1(X, \mu)$ (why?). So, by taking $g = |f|^{1/p}$, we deduce that $|f|^{1+1/p} \in L^1(X, \mu)$. Iterate the argument.

Proposition 3.10 (Minkowski's inequality) *Let* $1 \le p < \infty$ *and* $f, g \in L^p(X, \mu)$. *Then* $f + g \in L^p(X, \mu)$ *and*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p. \tag{3.6}$$

Proof The thesis is immediate if $p = 1$. Assume $p > 1$. We have

$$\int_X |f + g|^p \, d\mu \le \int_X |f + g|^{p-1} |f| \, d\mu + \int_X |f + g|^{p-1} |g| \, d\mu.$$

Since $|f + g|^{p-1} \in L^{p'}(X, \mu)$, with $p' = \frac{p}{p-1}$, using Hölder's inequality we find

$$\int_X |f + g|^p \, d\mu \le \left(\int_X |f + g|^p \, d\mu \right)^{(p-1)/p} (\|f\|_p + \|g\|_p),$$

and the conclusion follows. □

From Minkowski's inequality it follows that $\|\cdot\|_p$ is a norm on $L^p(X, \mu)$ for any $1 \le p < \infty$.

We will often use the following notation: given a sequence $(f_n)_n \subset L^p(X, \mu)$ and $f \in L^p(X, \mu)$, we write

$$f_n \xrightarrow{L^p} f$$

to mean that $(f_n)_n$ converges to f in $L^p(X, \mu)$, that is, $\|f_n - f\|_p \to 0$ (as $n \to \infty$). Our next result shows that $L^p(X, \mu)$ is a Banach space.[6]

---

[4] For a more extended treatment of interpolation theory see [SW71].

[5] If $p = 1$, we set $p' = \infty$.

[6] See Definition 6.5.

Proposition 3.11 (Riesz–Fischer) *Let* $1 \le p < \infty$ *and let* $(f_n)_n$ *be a Cauchy sequence in the normed space* $L^p(X, \mu)$. *Then there exist a subsequence* $(f_{n_k})_k$ *and a function* $f \in L^p(X, \mu)$ *such that:*

(i) $f_{n_k} \xrightarrow{a.e.} f$.

(ii) $f_n \xrightarrow{L^p} f$.

Proof Since $(f_n)_n$ is a Cauchy sequence in $L^p(X, \mu)$, for any $i \in \mathbb{N}$ there exists $n_i \in \mathbb{N}$ such that

$$\|f_n - f_m\|_p < 2^{-i} \qquad \forall n, m \ge n_i. \tag{3.7}$$

Consequently, we can construct an increasing sequence of indices $(n_i)_i$ such that

$$\|f_{n_{i+1}} - f_{n_i}\|_p < 2^{-i} \qquad \forall i \in \mathbb{N}.$$

Next, let us define

$$g(x) = \sum_{i=1}^{\infty} |f_{n_{i+1}}(x) - f_{n_i}(x)|, \qquad g_k(x) = \sum_{i=1}^{k} |f_{n_{i+1}}(x) - f_{n_i}(x)|, \quad k \ge 1.$$

Minkowski's inequality implies that $\|g_k\|_p < 1$ for every k; since $g_k(x) \uparrow g(x)$ for every $x \in X$, the Monotone Convergence Theorem ensures that

$$\int_X |g|^p \, d\mu = \lim_{k\to\infty} \int_X |g_k|^p \, d\mu \le 1.$$

Then, owing to Proposition 2.44, g is finite a.e.; therefore the series

$$\sum_{i=1}^{\infty} (f_{n_{i+1}} - f_{n_i}) + f_{n_1} \tag{3.8}$$

converges a.e. in X; since

$$\sum_{i=1}^{k} (f_{n_{i+1}} - f_{n_i}) + f_{n_1} = f_{n_{k+1}},$$

we deduce that $(f_{n_k})_k$ converges a.e. in X. Let us set $f(x) = \lim_{k\to\infty} f_{n_k}(x)$ when such limit exists and $f(x) = 0$ in the remaining zero-measure set. Then f is a Borel function (see Exercise 2.14) and

$$f(x) = \lim_{k\to\infty} f_{n_k}(x) \text{ a.e. in } X.$$

Moreover, $|f(x)| \le g(x) + |f_{n_1}(x)|$ a.e., so $f \in L^p(X, \mu)$. This concludes the proof of point (i).

Next, to derive (ii), fix $\varepsilon > 0$; there exists $N \in \mathbb{N}$ such that

$$\|f_n - f_m\|_p \le \varepsilon \quad \forall n, m \ge N.$$

Taking $m = n_k$ and passing to the limit as $k \to \infty$, Fatou's Lemma yields

$$\int_X |f_n - f|^p \, d\mu \le \liminf_{k\to\infty} \int_X |f_n - f_{n_k}|^p \, d\mu \le \varepsilon^p \qquad \forall n \ge N.$$

The proof is thus complete. □

Example 3.12 The conclusion of point (i) in Proposition 3.11 only holds for a subsequence, in general. Indeed, given $k \in \mathbb{N}$, for $1 \le i \le k$ consider the function

$$f_i^k(x) = \begin{cases} 1 & \text{if } \frac{i-1}{k} \le x < \frac{i}{k}, \\ 0 & \text{otherwise,} \end{cases}$$

defined on the interval $[0, 1)$. The sequence

$$f_1^1, f_1^2, f_2^2, \dots, f_1^k, f_2^k, \dots, f_k^k, \dots$$

converges to 0 in[7] $L^p(0, 1)$ for all $1 \le p < \infty$, but it does not converge at any point whatsoever. Observe that the subsequence $f_1^k = \chi_{[0,\frac{1}{k})}$ converges a 0 a.e.

Exercise 3.13 Let $1 \le p < \infty$ and $f \in L^p(\mathbb{R})$ (with respect to the Lebesgue measure). Set

$$f_n(x) = \begin{cases} f(x) & \text{if } x \in [n, n+1], \\ 0 & \text{otherwise.} \end{cases}$$

Show that:

- $f_n \in L^q(\mathbb{R})$ for every $n \in \mathbb{N}$ and $q \in [1, p]$.
- $f_n \xrightarrow{L^q} 0$ for all $q \in [1, p]$.

Exercise 3.14 Generalize Exercise 2.76 showing that if $f_n, f \in L^1(X, \mu)$ and $f_n \xrightarrow{L^1} f$, then

$$\lim_{k\to\infty} \sup_{n\in\mathbb{N}} \int_{\{|f_n|\ge k\}} |f_n| \, d\mu = 0.$$

[7] If I denotes one of the sets (a, b), $(a, b]$, $[a, b)$, $[a, b]$, and m is the Lebesgue measure on I, we usually write $L^p(I, m)$ as $L^p(a, b)$. Since the Lebesgue measure of a single point is zero, there is no need to specify which of the four sets we refer to.

Hint. Observe that[8]

$$\int_{\{|f_n|\geq 2k\}} |f_n|\,d\mu \leq 2\int_{\{|f_n-f|\vee|f|\geq k\}} |f_n-f|\vee|f|\,d\mu$$
$$\leq 2\int_{\{|f_n-f|\geq k\}} |f_n-f|\,d\mu + 2\int_{\{|f|\geq k\}} |f|\,d\mu.$$

Example 3.15 Consider the Lebesgue measure m on $[0, 1)$ and set

$$\mu = m + \sum_{n=1}^{\infty} \delta_{1/n}$$

where $\delta_{1/n}$ denotes the Dirac measure concentrated on $\frac{1}{n}$. Then $f(x) := x$ is in $L^2([0, 1), \mu) \setminus L^1([0, 1), \mu)$ because

$$\int_{[0,1)} x^2\,d\mu = \frac{1}{3} + \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$
$$\int_{[0,1)} x\,d\mu = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

On the other hand

$$g(x) := \begin{cases} \frac{1}{\sqrt{x}} & \text{if} \quad x \in [0,1) \setminus \mathbb{Q}, \\ 0 & \text{if} \quad x \in [0,1) \cap \mathbb{Q} \end{cases}$$

belongs to $L^1([0, 1), \mu) \setminus L^2([0, 1), \mu)$, since

$$\int_{[0,1)} g(x)\,d\mu = \int_0^1 \frac{dx}{\sqrt{x}} = 2,$$
$$\int_{[0,1)} g^2(x)\,d\mu = \int_0^1 \frac{dx}{x} = \infty.$$

Exercise 3.16 Show that $L^p(0, \infty) \not\subset L^q(0, \infty)$ for $p \neq q$ $(1 \leq p, q < \infty)$.
Hint. Consider

$$f(x) = \frac{1}{|x(\log^2|x| + 1)|^{1/p}}$$

and show that $f \in L^p(0, \infty)$ but $f \notin L^q(0, \infty)$ for $q \neq p$.

[8] By definition, $|f_n - f| \vee |f| = \max\{|f_n - f|, |f|\}$.

Exercise 3.17 Let $(f_n)_n$ be a sequence in $L^1(X,\mu)$ such that

$$\sum_{n=1}^{\infty}\int_X |f_n|\,d\mu < \infty.$$

1. Show that $\sum_{n=1}^{\infty}|f_n(x)| < \infty$ for almost every $x \in X$.
2. Show that there exists a function $f \in L^1(X,\mu)$ such that $\sum_{n=1}^{\infty} f_n(x) = f(x)$ for almost every $x \in X$ and

$$\sum_{n=1}^{\infty}\int_X f_n\,d\mu = \int_X f\,d\mu.$$

Exercise 3.18 Let $1 \le p < \infty$. Show that if $f \in L^p(\mathbb{R}^N)$ (with respect to the Lebesgue measure) and f is uniformly continuous, then

$$\lim_{\|x\|\to\infty} f(x) = 0.$$

Hint. If, by contradiction, there exists $(x_n)_n \subset \mathbb{R}^N$ such that $\|x_n\| \to \infty$ and $|f(x_n)| \ge \varepsilon > 0$ for every n, then the uniform continuity of f implies the existence of $\eta > 0$ such that $|f(x)| \ge \frac{\varepsilon}{2}$ if $\|x_n - x\| \le \eta$. Show that this yields $\int_{\mathbb{R}^N} |f|^p\,dx = \infty$.

Exercise 3.19 Show that the result of Exercise 3.18 may fail if one assumes that f is just continuous.

Hint. Consider

$$f_n(x) = \begin{cases} \min\{n^2x+1,\, 1-n^2x\} & \text{if } -\frac{1}{n^2} \le x \le \frac{1}{n^2}, \\ 0 & \text{if } x \notin \left(-\frac{1}{n^2}, \frac{1}{n^2}\right), \end{cases}$$

defined on $\mathbb{R}$ and set $f(x) = \sum_{n=1}^{\infty} f_n(x-n)$.

3.2 The Space $L^\infty(X,\mu)$

Let $(X,\mathscr{E},\mu)$ be a measure space and $f\colon X \to \overline{\mathbb{R}}$ a Borel function. We define the *essential supremum* $\|f\|_\infty$ of f as follows: if $\mu(|f| > M) > 0$ for all $M \in \mathbb{R}$, let $\|f\|_\infty = \infty$; otherwise, let

$$\|f\|_\infty = \inf\{M \ge 0 \mid \mu(|f| > M) = 0\}. \tag{3.9}$$

We say that f is *essentially bounded* if $\|f\|_\infty < \infty$ and we denote by $\mathscr{L}^\infty(X, \mathscr{E}, \mu) = \mathscr{L}^\infty(X, \mu)$ the class of all essentially bounded functions.

Example 3.20 Consider the Lebesgue measure on $[0, 1)$ and define $f : [0, 1) \to \mathbb{R}$ by

$$f(x) = \begin{cases} 1 & \text{if } x \neq \frac{1}{n}, \\ n & \text{if } x = \frac{1}{n}. \end{cases}$$

Then f is essentially bounded and $\|f\|_\infty = 1$.

Example 3.21 Let $\mu^\#$ be the counting measure on $\mathbb{N}$. In the following we will use the notation ℓ^∞ for space $\mathscr{L}^\infty(\mathbb{N}, \mu^\#)$. We have

$$\ell^\infty = \Big\{(x_n)_n \;\Big|\; x_n \in \mathbb{R},\ \sup_{n \geq 1} |x_n| < \infty\Big\},$$

and, for every $(x_n)_n \in \ell^\infty$,

$$\|(x_n)_n\|_\infty = \sup_{n \geq 1} |x_n| < \infty.$$

Observe that

$$\ell^p \subset \ell^\infty \quad \forall p \in [1, \infty).$$

Remark 3.22 Recalling that the function $t \to \mu(|f| > t)$ is right continuous (see Proposition 2.34), we conclude that

$$M_n \downarrow M_0 \quad \& \quad \mu(|f| > M_n) = 0 \quad \Longrightarrow \quad \mu(|f| > M_0) = 0.$$

So the infimum in (3.9) is actually a minimum. In particular, for any $f \in \mathscr{L}^\infty(X, \mu)$,

$$|f(x)| \leq \|f\|_\infty \quad \text{a.e. in } X$$

and

$$\|f\|_\infty = \min\{M \geq 0 \,|\, |f(x)| \leq M \text{ a.e.}\}. \tag{3.10}$$

In order to construct a vector space on which $\|\cdot\|_\infty$ is a norm, we proceed as in the previous section defining $L^\infty(X, \mu)$ as the quotient space of $\mathscr{L}^\infty(X, \mu)$ modulo the equivalence relation introduced in (3.2). So $L^\infty(X, \mathscr{E}, \mu) = L^\infty(X, \mu)$ is obtained by identifying functions in $\mathscr{L}^\infty(X, \mu)$ that coincide almost everywhere.

Exercise 3.23 Show that $L^\infty(X, \mu)$ is a vector space and $\|\cdot\|_\infty$ is a norm on $L^\infty(X, \mu)$.

Hint. Use (3.10). For instance, for any $\alpha \neq 0$, we have $|\alpha f(x)| \le |\alpha|\,\|f\|_\infty$ for almost every $x \in X$. So $\|\alpha f\|_\infty \le |\alpha|\,\|f\|_\infty$. On the other hand, we also have

$$\|f\|_\infty = \left\|\frac{1}{\alpha}\alpha f\right\|_\infty \le \frac{1}{|\alpha|}\,\|\alpha f\|_\infty .$$

Thus, $\|\alpha f\|_\infty = |\alpha|\,\|f\|_\infty$.

Like in the case of $p < \infty$, given a sequence $(f_n)_n$ in $L^\infty(X,\mu)$ and a function $f \in L^\infty(X,\mu)$, in the following we will write

$$f_n \xrightarrow{L^\infty} f$$

to mean that $(f_n)_n$ converges to f in $L^\infty(X,\mu)$, or $\lim_{n\to\infty}\|f_n - f\|_\infty = 0$.

Exercise 3.24 Let $f_n,\ f \in L^\infty(X,\mu)$. Show that if $f_n \xrightarrow{L^\infty} f$, then $f_n \xrightarrow{a.e.} f$.

Proposition 3.25 *$L^\infty(X,\mu)$ is a Banach space.*

Proof For a given Cauchy sequence $(f_n)_n$ in $L^\infty(X,\mu)$, let us set, for any $n, m \in \mathbb{N}$,

$$\begin{aligned} A_n &= \{|f_n| > \|f_n\|_\infty\}\,,\\ B_{m,n} &= \{|f_n - f_m| > \|f_n - f_m\|_\infty\}\,. \end{aligned}$$

Observe that, in view of (3.10),

$$\mu(A_n) = 0 \quad \& \quad \mu(B_{m,n}) = 0 \qquad \forall m, n \in \mathbb{N}.$$

Therefore

$$X_0 := \left(\bigcup_{n=1}^\infty A_n\right) \cup \left(\bigcup_{m,n=1}^\infty B_{m,n}\right)$$

has measure zero and $(f_n)_n$ is a Cauchy sequence for uniform convergence in X_0^c. Thus, setting $f(x) = \lim_{n\to\infty} f_n(x)$ for $x \in X_0^c$ and $f(x) = 0$ for $x \in X_0$, we have that f is a bounded Borel function. So $f \in L^\infty(X,\mu)$ and $f_n \to f$ uniformly in X_0^c. The conclusion follows because convergence in $L^\infty(X,\mu)$ is equivalent to uniform convergence outside a zero-measure set. □

Exercise 3.26 Show that for any $1 \le p \le \infty$ we have

$$f \in L^p(X,\mu),\ g \in L^\infty(X,\mu) \quad \Longrightarrow \quad fg \in L^p(X,\mu)$$

and

$$\|fg\|_p \le \|f\|_p\,\|g\|_\infty .$$

Example 3.27 It is easy to realize that the spaces[9] $L^\infty(0,1)$ and ℓ^∞ fail to be separable.[10]

1. Set
$$f_t(x) = \chi_{(0,t)}(x) \qquad \forall t, x \in (0,1).$$
We have that
$$t \neq s \quad \Longrightarrow \quad \|f_t - f_s\|_\infty = 1.$$
Let $\mathscr{M}$ be a dense set in $L^\infty(0,1)$. Then $\mathscr{M}$ has the property that for every $t \in (0,1)$ there exists $g_t \in \mathscr{M}$ with $\|f_t - g_t\|_\infty < \frac{1}{2}$. For $t \neq s$ we have
$$\|g_t - g_s\|_\infty \geq \|f_t - f_s\|_\infty - \|f_t - g_t\|_\infty - \|f_s - g_s\|_\infty > 0.$$
Hence, $g_t \neq g_s$. Therefore $\mathscr{M}$ contains an uncountable number of functions.
2. Let $(x^{(n)})_n$ be a countable set in ℓ^∞. Let $x^{(n)} = (x_k^{(n)})_k$ for every n and define the sequence
$$x = (x_k)_k \quad x_k = \begin{cases} 0 & \text{if } |x_k^{(k)}| \geq 1, \\ 1 + x_k^{(k)} & \text{if } |x_k^{(k)}| < 1. \end{cases}$$
We have that $x \in \ell^\infty$ and $\|x\|_\infty \leq 2$. Furthermore, for every $n \in \mathbb{N}$,
$$\|x - x^{(n)}\|_\infty = \sup_{k \geq 1} |x_k - x_k^{(n)}| \geq |x_n - x_n^{(n)}| \geq 1.$$
Consequently, $(x^{(n)})_n$ is not dense in ℓ^∞.

Proposition 3.28 *Let* $1 \leq p < \infty$ *and* $f \in L^p(X,\mu) \cap L^\infty(X,\mu)$. *Then*
$$f \in \bigcap_{q \geq p} L^q(X,\mu) \qquad \& \qquad \lim_{q \to \infty} \|f\|_q = \|f\|_\infty.$$

Proof For $p \leq q < \infty$ we have
$$|f(x)|^q \leq \|f\|_\infty^{q-p} |f(x)|^p \quad \text{a.e. in } X.$$
So, by integrating,
$$\|f\|_q \leq \|f\|_p^{\frac{p}{q}} \, \|f\|_\infty^{1-\frac{p}{q}}.$$

[9] As for the case $p < \infty$ (see footnote 7), if I is one of the intervals (a,b), $(a,b]$, $[a,b)$, $[a,b]$, and m is the Lebesgue measure on I, we usually write $L^\infty(I,m)$ as $L^\infty(a,b)$.

[10] A metric space is said to be *separable* if it has a countable dense subset.

Consequently, $f \in \cap_{q \ge p} L^q(X, \mu)$ and

$$\limsup_{q\to\infty} \|f\|_q \le \|f\|_\infty. \tag{3.11}$$

Conversely, let $0 < a < \|f\|_\infty$ (for $\|f\|_\infty = 0$ the conclusion is trivial). By Markov's inequality we get

$$\mu(|f| > a) = \mu(|f|^q > a^q) \le a^{-q}\|f\|_q^q \quad \forall q \in [p, \infty).$$

Therefore

$$\|f\|_q \ge a\mu(|f| > a)^{1/q} \quad \forall q \in [p, \infty)$$

and so

$$\liminf_{q\to\infty} \|f\|_q \ge a$$

because $\mu(|f| > a) > 0$. Since a is any number less than $\|f\|_\infty$,

$$\liminf_{q\to\infty} \|f\|_q \ge \|f\|_\infty. \tag{3.12}$$

By (3.11) and (3.12) the conclusion follows. □

Corollary 3.29 *Let μ be a finite measure and let $f \in L^\infty(X, \mu)$. Then*

$$f \in \bigcap_{p\ge 1} L^p(X, \mu) \qquad \& \qquad \lim_{p\to\infty} \|f\|_p = \|f\|_\infty. \tag{3.13}$$

Proof For $1 \le p < \infty$ we have

$$\int_X |f(x)|^p \, d\mu \le \mu(X)\|f\|_\infty^p.$$

So $f \in \cap_{p\ge 1} L^p(X, \mu)$. The conclusion follows from Proposition 3.28. □

It is noteworthy that in general

$$\bigcap_{1\le p<\infty} L^p(X, \mu) \ne L^\infty(X, \mu).$$

Exercise 3.30 Show that

$$f(x) := \log x \qquad \forall x \in (0, 1]$$

belongs to $L^p(0, 1)$ for $1 \le p < \infty$, but $f \notin L^\infty(0, 1)$.

3.3 Convergence in Measure

We now discuss a kind of convergence for sequences of Borel functions which is of considerable importance in probability theory (see [Ha50]).

Definition 3.31 Let $f_n, f : X \to \mathbb{R}$ be Borel functions. Then $(f_n)_n$ is said to *converge in measure* to f if for any $\varepsilon > 0$:

$$\mu(|f_n - f| \geq \varepsilon) \to 0 \text{ as } n \to \infty.$$

Let us compare convergence in measure with other kinds of convergence.

Proposition 3.32 *Let $f_n, f : X \to \mathbb{R}$ be Borel functions. The following statements hold:*

1. *If $f_n \xrightarrow{a.e.} f$ and $\mu(X) < \infty$, then $f_n \to f$ in measure.*
2. *If $f_n \to f$ in measure, then there exists a subsequence $(f_{n_k})_k$ such that $f_{n_k} \xrightarrow{a.e.} f$.*
3. *If $1 \leq p \leq \infty$, $f_n, f \in L^p(X, \mu)$ and $f_n \xrightarrow{L^p} f$, then $f_n \to f$ in measure.*

Proof 1. Fix $\varepsilon, \eta > 0$. According to Theorem 2.27 there exists $E \in \mathscr{E}$ such that $\mu(E) < \eta$ and $f_n \to f$ uniformly in $X \setminus E$. Then, for n sufficiently large,

$$\{|f_n - f| \geq \varepsilon\} \subset E.$$

So

$$\mu(|f_n - f| \geq \varepsilon) \leq \mu(E) < \eta.$$

2. For every $k \in \mathbb{N}$ we have that

$$\mu\Big(|f_n - f| \geq \frac{1}{k}\Big) \to 0 \text{ as } n \to \infty.$$

Consequently, we can construct an increasing sequence $(n_k)_k$ of positive integers such that

$$\mu\Big(|f_{n_k} - f| \geq \frac{1}{k}\Big) < \frac{1}{2^k} \quad \forall k \in \mathbb{N}.$$

Now, set

$$A_k = \bigcup_{i=k}^{\infty} \Big\{|f_{n_i} - f| \geq \frac{1}{i}\Big\}, \quad A = \bigcap_{k=1}^{\infty} A_k.$$

Observe that $\mu(A_k) \leq \sum_{i=k}^{\infty} \frac{1}{2^i}$ for every $k \in \mathbb{N}$. Since $A_k \downarrow A$, Proposition 1.18 yields

$$\mu(A) = \lim_{k\to\infty} \mu(A_k) = 0.$$

For any $x \in A^c$ there exists $k \in \mathbb{N}$ such that $x \in A_k^c$, that is,

$$|f_{n_i}(x) - f(x)| < \frac{1}{i} \quad \forall i \geq k.$$

This shows that $f_{n_k}(x) \to f(x)$ for every $x \in A^c$.

3. Fix $\varepsilon > 0$. Assume first $1 \leq p < \infty$. Then Markov's inequality implies that

$$\mu(|f_n - f| > \varepsilon) \leq \frac{1}{\varepsilon^p} \int_X |f_n - f|^p \, d\mu \to 0 \text{ as } n \to \infty.$$

Let now $p = \infty$. For n large enough, we have that $|f_n - f| \leq \varepsilon$ a.e. in X, yielding $\mu(|f_n - f| > \varepsilon) = 0$. □

Exercise 3.33 Show that almost everywhere convergence does not imply convergence in measure if $\mu(X) = \infty$.

Hint. Consider $f_n = \chi_{[n,\infty)}$ in $\mathbb{R}$ with the Lebesgue measure.

Example 3.34 The sequence constructed in Example 3.12 converges to zero in $L^1(0, 1)$ and, consequently, in measure, but it does not converge at any point whatsoever. This shows that part 2 of Proposition 3.32 and part (i) of Proposition 3.11 only hold for a subsequence, in general.

Exercise 3.35 Give an example to show that convergence in measure does not imply convergence in $L^p(X, \mu)$.

Hint. Consider the sequence $f_n = n\chi_{(0,\frac{1}{n})}$ in $L^1(0, 1)$.

3.4 Convergence and Approximation in L^p

In this section we will exhibit techniques to derive convergence in L^p from almost everywhere convergence. Next, we will show that, if Ω is an open set in $\mathbb{R}^N$ and μ is a Radon measure on Ω, then all elements of $L^p(\Omega)$ can be approximated by continuous functions.

3.4.1 Convergence Results

In what follows, $(X, \mathscr{E}, \mu)$ denotes a given measure space.

Our next result is a direct consequence of Fatou's Lemma and Lebesgue's Dominated Convergence Theorem.

Proposition 3.36 *Let* $1 \leq p < \infty$, $(f_n)_n$ *a sequence in* $L^p(X, \mu)$ *and let* $f : X \to \overline{\mathbb{R}}$ *be a Borel function such that* $f_n \xrightarrow{a.e.} f$.

(i) *If* $(f_n)_n$ *is bounded*[11] *in* $L^p(X,\mu)$*, then* $f \in L^p(X,\mu)$ *and*

$$\|f\|_p \le \liminf_{n\to\infty} \|f_n\|_p.$$

(ii) *If there exists* $g \in L^p(X,\mu)$ *such that* $|f_n(x)| \le g(x)$ *for all* $n \in \mathbb{N}$ *and for almost every* $x \in X$*, then* $f \in L^p(X,\mu)$ *and* $f_n \xrightarrow{L^p} f$.

Exercise 3.37 Show that, for $p = \infty$, point (i) of Proposition 3.36 is still true, while (ii) fails in general.

Hint. Consider the sequence $f_n = \chi_{\left(\frac{1}{n},1\right)}$ in $L^\infty(0,1)$.

Exercise 3.38 Let $(f_n)_n$ be the sequence defined by

$$f_n(x) = \frac{\sqrt{n}}{1+\sqrt{nx}}, \quad x \in (0,1).$$

Show that:

- $(f_n)_n$ converges in $L^p(0,1)$ for every $p \in [1,2)$.
- $(f_n)_n$ is not bounded in $L^p(0,1)$ for every $p \in [2,\infty]$.

Now, observe that, since $|\,\|f_n\|_p - \|f\|_p\,| \le \|f_n - f\|_p$, the following holds:

$$f_n \xrightarrow{L^p} f \quad \Longrightarrow \quad \|f_n\|_p \to \|f\|_p.$$

So a necessary condition for convergence in $L^p(X,\mu)$ is convergence of L^p-norms. Our next result shows that if $f_n \xrightarrow{a.e.} f$, such a condition is also sufficient.

Proposition 3.39 *Given* $1 \le p < \infty$*, let* $f_n,\ f \in L^p(X,\mu)$ *be such that* $f_n \xrightarrow{a.e.} f$*. If* $\|f_n\|_p \to \|f\|_p$*, then* $f_n \xrightarrow{L^p} f$*.*

Proof [12] Consider the function $g_n \in L^1(X,\mu)$ defined by

$$g_n = \frac{|f_n|^p + |f|^p}{2} - \left|\frac{f_n - f}{2}\right|^p.$$

[11] A subset $\mathscr{M}$ of a normed linear space Y is said to be *bounded* if there exists a constant M such that $||y|| \le M$ for all $y \in \mathscr{M}$.

[12] This proof is due to Novinger [No72].

Since $p \geq 1$, a simple convexity argument shows that $g_n \geq 0$. Moreover, $g_n \xrightarrow{a.e.} |f|^p$. Therefore Fatou's Lemma yields

$$\int_X |f|^p \, d\mu \leq \liminf_{n\to\infty} \int_X g_n \, d\mu = \int_X |f|^p \, d\mu - \limsup_{n\to\infty} \int_X \left| \frac{f_n - f}{2} \right|^p d\mu.$$

So $\limsup_n \|f_n - f\|_p \leq 0$, that is, $f_n \xrightarrow{L^p} f$. □

The results below generalize Vitali's uniform summability property and give sufficient conditions for convergence in $L^p(X, \mu)$ for $1 \leq p < \infty$.

Corollary 3.40 *Let* $1 \leq p < \infty$, *let* $(f_n)_n \subset L^p(X, \mu)$, *and let* $f : X \to \overline{\mathbb{R}}$ *be a Borel function such that:*

(i) $f_n \xrightarrow{a.e.} f$.
(ii) For any $\varepsilon > 0$ *there exists* $\delta > 0$ *such that*

$$A \in \mathscr{E} \;\; \& \;\; \mu(A) < \delta \quad \Longrightarrow \quad \int_A |f_n|^p \, d\mu < \varepsilon \quad \forall n \in \mathbb{N}.$$

(iii) For any $\varepsilon > 0$ *there exists* $B_\varepsilon \in \mathscr{E}$ *such that*

$$\mu(B_\varepsilon) < \infty \qquad \& \qquad \int_{B_\varepsilon^c} |f_n|^p \, d\mu < \varepsilon \quad \forall n \in \mathbb{N}.$$

Then $f \in L^p(X, \mu)$ *and* $f_n \xrightarrow{L^p} f$.

Proof Let us set $g_n = |f_n|^p$. Then, by hypotheses (ii)–(iii), $(g_n)_n$ is uniformly μ-summable and converges to $|f|^p$ a.e. in X. Therefore Vitali's Theorem (Theorem 2.98) implies that $f \in L^p(X, \mu)$ and

$$\|f_n\|_p^p = \int_X g_n \, d\mu \quad \longrightarrow \quad \|f\|_p^p.$$

The conclusion now follows from Proposition 3.39. □

Remark 3.41 If μ is finite, then, by taking $B_\varepsilon = X$, we deduce that (iii) of Corollary 3.40 is always satisfied.

Corollary 3.42 *Assume* $\mu(X) < \infty$. *Let* $1 < p < \infty$ *and let* $(f_n)_n$ *be a bounded sequence in* $L^p(X, \mu)$ *converging a.e. to a Borel function* f. *Then* $f \in L^p(X, \mu)$ *and*

$$f_n \xrightarrow{L^q} f \quad \forall q \in [1, p).$$

Proof Let $M \geq 0$ be such that $\|f_n\|_p \leq M$ for every $n \in \mathbb{N}$. Part (i) of Proposition 3.36 implies that $f \in L^p(X, \mu)$. Consequently, by Corollary 3.6, $f_n, f \in \cap_{1 \leq q \leq p} L^q(X, \mu)$. Let now $1 \leq q < p$: by Hölder's inequality, for any $A \in \mathscr{E}$,

$$\int_A |f_n|^q \, d\mu \leq \left(\int_A |f_n|^p \, d\mu \right)^{\frac{q}{p}} (\mu(A))^{1-\frac{q}{p}} \leq M^q \, (\mu(A))^{1-\frac{q}{p}} .$$

The conclusion follows from Corollary 3.40. □

Corollary 3.43 *Assume $\mu(X) < \infty$. Let $(f_n)_n$ be a sequence in $L^1(X, \mu)$ converging a.e. to a Borel function f and suppose that*[13]

$$\int_X |f_n| \log^+ (|f_n|) \, d\mu \leq M \qquad \forall n \in \mathbb{N}$$

for some constant $M \geq 0$. Then $f \in L^1(X, \mu)$ and $f_n \xrightarrow{L^1} f$.

Proof Fix $\varepsilon \in (0, 1)$, $t \in X$, and apply inequality (F.4) with $x = \frac{1}{\varepsilon}$ and $y = \varepsilon|f_n(t)|$ to obtain

$$|f_n(t)| \leq \varepsilon |f_n(t)| \log(\varepsilon |f_n(t)|) + e^{\frac{1}{\varepsilon}} \leq \varepsilon |f_n(t)| \log^+(|f_n(t)|) + e^{\frac{1}{\varepsilon}} .$$

Consequently, for any $A \in \mathscr{E}$,

$$\int_A |f_n| \, d\mu \leq M\varepsilon + \mu(A) e^{\frac{1}{\varepsilon}} \qquad \forall n \in \mathbb{N}.$$

This implies that $(f_n)_n$ is uniformly μ-summable. The thesis follows from Corollary 3.40. □

Exercise 3.44 Show how Corollary 3.43 can be adapted to the case of $\mu(X) = \infty$ adding the assumption that $(f_n)_n$ satisfies (b) of Definition 2.95.

3.4.2 Dense Subsets in L^p

Let $\Omega \subset \mathbb{R}^N$ be an open set. The *support* of a continuous function $f : \Omega \to \mathbb{R}$, written as $\mathrm{supp}(f)$, is defined as the closure in $\mathbb{R}^N$ of the set $\{x \in \Omega \mid f(x) \neq 0\}$. If $\mathrm{supp}(f)$ is a compact subset of Ω, then f is said to be *of compact support*. The class of all continuous functions $f : \Omega \to \mathbb{R}$ with compact support is a linear space which will be denoted by $\mathscr{C}_c(\Omega)$.

[13] By definition, $\log^+(x) = \max\{\log x, 0\}$ for any $x > 0$.

Clearly, if μ is a Radon measure on Ω, then

$$\mathscr{C}_c(\Omega) \subset L^p(\Omega, \mu) \text{ for } 1 \le p \le \infty.$$

Theorem 3.45 *Let μ be a Radon measure on Ω. If $1 \le p < \infty$, then $\mathscr{C}_c(\Omega)$ is dense in $L^p(\Omega, \mu)$.*

Proof First consider the case $\Omega = \mathbb{R}^N$. We begin by proving the thesis under additional assumptions and split the argument into several steps, each of which will achieve a higher degree of generality.

1. Let us show how to approximate, by functions in $\mathscr{C}_c(\mathbb{R}^N)$, any function $f \in L^p(\mathbb{R}^N, \mu)$ that satisfies, for some $M, r > 0$,[14]

$$0 \le f(x \le M \quad \forall x \in \mathbb{R}^N, \tag{3.14}$$
$$f(x) = 0 \quad \forall x \in \mathbb{R}^N \setminus B_r. \tag{3.15}$$

Let $\varepsilon > 0$. Since μ is Radon, we have $\mu(B_r) < \infty$. Then, by Lusin's Theorem (Theorem 2.29), there exists a function $f_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ such that

$$\mu(f_\varepsilon \neq f) < \frac{\varepsilon}{(2M)^p} \qquad \& \qquad \|f_\varepsilon\|_\infty \le M.$$

Then

$$\int_{\mathbb{R}^N} |f - f_\varepsilon|^p \, d\mu \le (2M)^p \mu(f_\varepsilon \neq f) < \varepsilon.$$

2. We now proceed to remove assumption (3.15). Let $f \in L^p(\mathbb{R}^N, \mu)$ be a function satisfying (3.14) and fix $\varepsilon > 0$. Owing to Lebesgue's Theorem $f\chi_{B_n} \xrightarrow{L^p} f$. Then there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\|f - f\chi_{B_{n_\varepsilon}}\|_p < \varepsilon.$$

In view of step 1, there exists $g_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ such that $\|f\chi_{B_{n_\varepsilon}} - g_\varepsilon\|_p < \varepsilon$. Then we conclude that

$$\|f - g_\varepsilon\|_p \le \|f - f\chi_{B_{n_\varepsilon}}\|_p + \|f\chi_{B_{n_\varepsilon}} - g_\varepsilon\|_p < 2\varepsilon.$$

[14] Hereafter, $B_r = B_r(0) = \{x \in \mathbb{R}^N : |x| < r\}$.

3. Next, let us dispense with the upper bound in (3.14). Let $f \in L^p(\mathbb{R}^N, \mu)$ be such that $f \geq 0$ and set

$$0 \leq f_n(x) := \min\{f(x), n\} \quad \forall x \in \mathbb{R}^N;$$

by Lebesgue's Theorem we have that $f_n \xrightarrow{L^p} f$. Therefore there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\|f - f_{n_\varepsilon}\|_p < \varepsilon.$$

In view of step 2, there exists $g_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ such that $\|f_{n_\varepsilon} - g_\varepsilon\|_p < \varepsilon$. Then $\|f - g_\varepsilon\|_p \leq \|f - f_{n_\varepsilon}\|_p + \|f_{n_\varepsilon} - g_\varepsilon\|_p < 2\varepsilon$.

Finally, the extra assumption that $f \geq 0$ can be disposed of applying step 3 to f^+ and f^-. The proof is thus complete in the case of $\Omega = \mathbb{R}^N$.

Next, consider an open set $\Omega \subset \mathbb{R}^N$ and let $f \in L^p(\Omega, \mu)$. The function

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in \mathbb{R}^N \setminus \Omega \end{cases}$$

belongs to $L^p(\mathbb{R}^N, \tilde{\mu})$, where $\tilde{\mu}(A) = \mu(A \cap \Omega)$ for any Borel set $A \subset \mathbb{R}^N$. Since $\tilde{\mu}$ is a Radon measure on $\mathbb{R}^N$, then there exists $f_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ such that

$$\int_{\mathbb{R}^N} |\tilde{f} - f_\varepsilon|^p \, d\tilde{\mu} < \varepsilon.$$

Let $(V_n)_n$ be a sequence of open sets in $\mathbb{R}^N$ such that

$$\overline{V}_n \text{ is compact}, \quad \overline{V}_n \subset V_{n+1}, \quad \bigcup_{n=1}^{\infty} V_n = \Omega \tag{3.16}$$

(for instance, we can choose[15] $V_n = B_n \cap \{x \in \Omega \mid d_{\Omega^c}(x) > \frac{1}{n}\}$) and set

$$g_n(x) = f_\varepsilon(x) \frac{d_{V_{n+1}^c}(x)}{d_{V_{n+1}^c}(x) + d_{V_n}(x)}, \quad x \in \Omega.$$

We have $g_n = 0$ outside $\overline{V}_{n+1}$, so $g_n \in \mathscr{C}_c(\Omega)$. Furthermore, $g_n = f_\varepsilon$ in V_n and $V_n \uparrow \Omega$, which implies that $g_n(x) \to f_\varepsilon(x)$ for every $x \in \Omega$. Since $|g_n| \leq f_\varepsilon$, Lebesgue's Theorem yields $g_n \to f_\varepsilon$ in $L^p(\Omega, \mu)$. Therefore there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\int_\Omega |f_\varepsilon - g_{n_\varepsilon}|^p \, d\mu < \varepsilon.$$

[15] $d_{\Omega^c}(x)$ denotes the distance of the point x from the set Ω^c (see Appendix A).

Then

$$\left(\int_{\Omega} |f - g_{n_\varepsilon}|^p \, d\mu\right)^{\frac{1}{p}} \le \left(\int_{\Omega} |f - f_\varepsilon|^p \, d\mu\right)^{\frac{1}{p}} + \left(\int_{\Omega} |f_\varepsilon - g_{n_\varepsilon}|^p \, d\mu\right)^{\frac{1}{p}}$$
$$= \left(\int_{\mathbb{R}^N} |\tilde{f} - f_\varepsilon|^p \, d\tilde{\mu}\right)^{\frac{1}{p}} + \left(\int_{\Omega} |f_\varepsilon - g_{n_\varepsilon}|^p \, d\mu\right)^{\frac{1}{p}} < 2\varepsilon.$$

The proof is thus complete. □

Exercise 3.46 Explain why $\mathscr{C}_c(\Omega)$ is not dense in $L^\infty(\Omega)$ (with respect to the Lebesgue measure), and characterize the closure of $\mathscr{C}_c(\Omega)$ in $L^\infty(\Omega)$.

Hint. Show that the closure of $\mathscr{C}_c(\Omega)$ in $L^\infty(\Omega)$ coincides with the set $\mathscr{C}_0(\Omega)$ of all continuous functions $f : \Omega \to \mathbb{R}$ satisfying

$$\forall \varepsilon > 0 \ \ \exists K \subset \Omega \text{ compact such that } \sup_{x \in \Omega \setminus K} |f(x)| \le \varepsilon.$$

In particular, if $\Omega = \mathbb{R}^N$, we have

$$\mathscr{C}_0(\mathbb{R}^N) = \left\{ f : \mathbb{R}^N \to \mathbb{R} \,\middle|\, f \text{ continuous } \ \& \ \lim_{\|x\| \to \infty} f(x) = 0 \right\},$$

whereas, if Ω is bounded,

$$\mathscr{C}_0(\Omega) = \left\{ f : \Omega \to \mathbb{R} \,\middle|\, f \text{ continuous } \ \& \ \lim_{d_{\Omega^c}(x) \to 0} f(x) = 0 \right\}.$$

Proposition 3.47 *Let $A \subset \mathbb{R}^N$ be a Borel set and let μ be a Radon measure on A. Then $L^p(A, \mu)$ is separable for all $1 \le p < \infty$.*

Proof Assume first $A = \mathbb{R}^N$ and consider the class of the dyadic cubes in $\mathbb{R}^N$ (see Definition 1.58). Let $\mathscr{M}$ be the set of all (finite) linear combinations with rational coefficients of characteristic functions of these cubes. Then $\mathscr{M}$ is countable. We claim that $\mathscr{M}$ is dense in $L^p(\mathbb{R}^N, \mu)$ for $1 \le p < \infty$. Indeed, given $f \in L^p(\mathbb{R}^N, \mu)$ and $\varepsilon > 0$, according to Theorem 3.45 there exists $f_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ with $\|f - f_\varepsilon\|_p \le \varepsilon$. Setting

$$\eta_\varepsilon = \frac{\varepsilon}{(1 + \mu([-k, k)^N))^{1/p}},$$

where $k \in \mathbb{N}$ is such that $\operatorname{supp}(f) \subset [-k, k)^N$, by the uniform continuity of f_ε we get the existence of $\delta > 0$ such that

$$x, \ y \in \mathbb{R}^N \ \ \& \ \ \|x - y\| \le \delta \implies |f_\varepsilon(x) - f_\varepsilon(y)| < \eta_\varepsilon.$$

Next, let j be sufficiently large such that the cubes in $\mathscr{Q}_j$ have diameter less than δ and cover the cube $[-k,k)^N$ by a finite number of cubes $Q_1, \dots, Q_n \in \mathscr{Q}_j$. Choose $c_1, \dots, c_n \in \mathbb{Q}$ in such a way that

$$\inf_{Q_i} f_\varepsilon < c_i < \inf_{Q_i} f_\varepsilon + \eta_\varepsilon$$

and define

$$g_\varepsilon = \sum_{i=1}^n c_i \chi_{Q_i}.$$

It follows that $g_\varepsilon \in \mathscr{M}$ and $\|f_\varepsilon - g_\varepsilon\|_\infty \le \eta_\varepsilon$. So

$$\|f_\varepsilon - g_\varepsilon\|_p^p = \int_{[-k,k)^N} |f_\varepsilon - g_\varepsilon|^p \, d\mu \le \mu([-k,k)^N)\|f_\varepsilon - g_\varepsilon\|_\infty^p < \varepsilon^p,$$

yielding

$$\|f - g_\varepsilon\|_p \le \|f - f_\varepsilon\|_p + \|f_\varepsilon - g_\varepsilon\|_p < 2\varepsilon.$$

This completes the proof in the case $A = \mathbb{R}^N$.

To obtain the conclusion for an arbitrary Borel set A, let $\mathscr{M}'$ denote the restriction to A of the functions in $\mathscr{M}$. To see that $\mathscr{M}'$ is dense in $L^p(A,\mu)$, $1 \le p < \infty$, given $f \in L^p(A,\mu)$, set $\tilde{f} = f$ in A and $\tilde{f} = 0$ outside A. Then $\tilde{f} \in L^p(\mathbb{R}^N, \tilde{\mu})$, where $\tilde{\mu}(B) = \mu(B \cap A)$ for any Borel set $B \subset \mathbb{R}^N$. Since $\tilde{\mu}$ is a Radon measure on $\mathbb{R}^N$, given $\varepsilon > 0$ there exists $f_\varepsilon \in \mathscr{M}$ such that

$$\int_{\mathbb{R}^N} |\tilde{f} - f_\varepsilon|^p \, d\tilde{\mu} < \varepsilon.$$

Therefore $\int_A |f - f_\varepsilon|^p \, d\mu = \int_{\mathbb{R}^N} |\tilde{f} - f_\varepsilon|^p \, d\tilde{\mu} < \varepsilon$. This shows that $\mathscr{M}'$ is dense in $L^p(A,\mu)$ and completes the proof. □

Exercise 3.48 ℓ^p is separable for $1 \le p < \infty$.

Hint. Show that the set

$$\mathscr{M} = \left\{ (x_n)_n \,\middle|\, x_n \in \mathbb{Q}, \ \sup_{x_n \ne 0} n < \infty \right\}$$

is countable and dense in ℓ^p.

Our next result shows that the integral over $\mathbb{R}^N$ with respect to the Lebesgue measure is translation continuous.

Proposition 3.49 (translation continuity in L^p) *Let* $1 \leq p < \infty$ *and let* $f \in L^p(\mathbb{R}^N)$ (*with respect to the Lebesgue measure*). *Then*

$$\lim_{\|h\|\to 0} \int_{\mathbb{R}^N} |f(x+h) - f(x)|^p \, dx = 0.$$

Proof Assume first $f \in \mathscr{C}_c(\mathbb{R}^N)$ and let $K = \operatorname{supp}(f)$. Setting

$$\tilde{K} := \{x \in \mathbb{R}^N \mid d_K(x) \leq 1\}$$

we have that $\operatorname{supp}(f(x+h)) \subset \tilde{K}$ if $\|h\| \leq 1$. Hence, for $\|h\| \leq 1$, we get

$$\begin{aligned} \|f(x+h) - f(x)\|_p^p &= \int_{\tilde{K}} |f(x+h) - f(x)|^p \, dx \\ &\leq m(\tilde{K}) \sup_{\|x-y\|\leq\|h\|} |f(x) - f(y)|^p. \end{aligned}$$

Since f is uniformly continuous, $\sup_{\|x-y\|\leq\|h\|} |f(x) - f(y)| \to 0$ as $h \to 0$. The conclusion is thus proved when $f \in \mathscr{C}_c(\mathbb{R}^N)$.

In the general case, fix $f \in L^p(\mathbb{R}^N)$ and $\varepsilon > 0$. Theorem 3.45 implies the existence of $f_\varepsilon \in \mathscr{C}_c(\mathbb{R}^N)$ such that $\|f_\varepsilon - f\|_p < \varepsilon$. By the first part of the proof, we have that there exists $\delta > 0$ such that

$$\|f_\varepsilon(x+h) - f_\varepsilon(x)\|_p < \varepsilon \text{ for } \|h\| \leq \delta.$$

Then, using Minkowski's inequality and the translation invariance of the Lebesgue measure, if $\|h\| \leq \delta$ we deduce that

$$\begin{aligned} \|f(x+h) - f(x)\|_p &\leq \|f(x+h) - f_\varepsilon(x+h)\|_p + \|f_\varepsilon(x+h) - f_\varepsilon(x)\|_p \\ &\quad + \|f_\varepsilon(x) - f(x)\|_p \\ &= 2\|f_\varepsilon(x) - f(x)\|_p + \|f_\varepsilon(x+h) - f_\varepsilon(x)\|_p \leq 3\varepsilon. \end{aligned}$$

The proof is thus complete. □

3.5 Miscellaneous Exercises

Exercise 3.50 For each of the following sequences $(x_n)_n$ find the values $1 \leq p \leq \infty$ for which $(x_n)_n \in \ell^p$:

$$\frac{1}{n+\sqrt{n}}, \qquad \frac{1+n+\sqrt{n}}{n} \tan\frac{1}{\sqrt{n}}, \qquad \frac{n^2}{\sqrt{1+n}} \sin\frac{1}{n(2+n)},$$

$$\cos\frac{1}{n}, \qquad \frac{1}{\sqrt{n}\sqrt{1+n}}\cos\frac{1}{n}, \qquad \sin^3\Big(\frac{1}{1+\log^2 n}\Big),$$

$$\frac{n+1}{n(1+\log n)}, \qquad \sqrt{n^2+1+\log n}\,\sin\frac{1}{n}, \qquad \tan\frac{n+\log n}{1+n^2}.$$

Exercise 3.51 For each of the following functions f find the values $1 \le p \le \infty$ for which $f \in L^p(0,\infty)$:

$$\frac{\sin x}{x(1+x)}, \qquad \frac{1}{1+|\log x|}, \qquad \frac{\arctan x}{\sqrt{x^3+x^4}},$$

$$\sqrt{\frac{1+\log^2 x}{2+x}}, \qquad \frac{\arctan(x+x^2)}{x+e^x}, \qquad \tan\frac{1}{1+|\log x|}.$$

Exercise 3.52 Let $f_n : (1,\infty) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{n}{\sqrt{x}}e^{-nx}, \qquad x > 1.$$

Show that:

1. $f_n \in L^p(1,\infty)$ for every $1 \le p \le \infty$.
2. $f_n \to 0$ in $L^p(1,\infty)$ for every $1 \le p \le \infty$.

Exercise 3.53 Let $f_n : (0,1) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{n}{e^{n\sqrt{x}}-1}, \quad x \in (0,1).$$

Show that:

1. f_n is convergent in $L^p(0,1)$ for every $1 \le p < 2$.
2. $f_n \notin L^p(0,1)$ if $2 \le p \le \infty$.

Exercise 3.54 Let $f_n : (0,1) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{n\sqrt{x}\sin x}{1+nx^2}, \quad x \in (0,1).$$

Show that:

1. $f_n \in L^p(0,1)$ for every $1 \le p \le \infty$.
2. f_n is convergent in $L^p(0,1)$ for every $1 \le p < 2$.
3. f_n is not convergent in $L^p(0,1)$ if $2 \le p \le \infty$.

Exercise 3.55 Let $f_n : \mathbb{R} \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \begin{cases} \dfrac{\sin x}{1+x} & \text{if } x \in [n, n+1] \\ 0 & \text{otherwise} \end{cases}.$$

Show that:

1. $f_n \in L^p(\mathbb{R})$ for every $1 \leq p \leq \infty$.
2. $f_n \to 0$ in $L^p(\mathbb{R})$ for every $1 \leq p \leq \infty$.

Exercise 3.56 Let $f_n : (0, \infty) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{1 + \cos x}{\sqrt{x}} e^{-nx}, \quad x > 0.$$

Show that:

1. f_n is convergent in $L^p(0, \infty)$ for every $1 \leq p < 2$.
2. $f_n \notin L^p(0, \infty)$ if $2 \leq p \leq \infty$.

Exercise 3.57 Let $f_n : (0, 1) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{\sin \sqrt[3]{x}}{\sqrt{x}} \cos \frac{x}{n}, \quad x \in (0, 1).$$

Show that:

1. f_n is convergent in $L^p(0, 1)$ for every $1 \leq p < 6$.
2. $f_n \notin L^p(0, 1)$ if $6 \leq p \leq \infty$.

Exercise 3.58 Let $f_n : (0, \infty) \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \frac{1}{\sqrt{1+x}} \sin \frac{1}{1 + |\log x|^n}, \quad x > 0.$$

1. Show that $f_n \in L^p(0, \infty)$ for every $2 \leq p \leq \infty$.
2. Show that $f_n \notin L^p(0, \infty)$ if $p < 2$.
3. Find the values $2 \leq p \leq \infty$ for which f_n is convergent in $L^p(0, \infty)$.

Exercise 3.59 Let $1 \leq p < \infty$ and $f \in L^p(\mathbb{R})$. Consider the sets

$$A_n = \{x \in \mathbb{R} \mid |f_n(x)| \geq n\}.$$

Show that:

1. $m(A_n) \to 0$.
2. $f\chi_{A_n} \in L^q(\mathbb{R})$ for every $q \leq p$.
3. $f\chi_{A_n} \to 0$ in $L^q(\mathbb{R})$ for every $q \leq p$.

Exercise 3.60 Let $1 \le p < \infty$ and $f \in L^p(\mathbb{R})$. Consider the sets

$$B_n = \left\{ x \in \mathbb{R} \,\middle|\, |f(x)| \le \frac{1}{n} \right\}.$$

Show that:

1. $f\chi_{B_n} \in L^q(\mathbb{R})$ for every $p \le q \le \infty$.
2. $f\chi_{B_n} \to 0$ in $L^q(\mathbb{R})$ for every $p \le q \le \infty$.

Exercise 3.61 Find the values $1 \le p \le \infty$ for which the following sequence is convergent in $L^p(1, \infty)$

$$f_n(x) = \frac{n}{\sqrt{x}(n+x)}, \quad x \ge 1.$$

Exercise 3.62 Find the values $1 \le p \le \infty$ for which the following sequence is convergent in $L^p(0, \infty)$

$$f_n(x) = \begin{cases} \frac{\log x}{\sqrt[3]{2x^2 - x}} & \text{if } n \le x \le 2n \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 3.63 Find the values $1 \le p \le \infty$ for which the following sequence is convergent in $L^p(0, \infty)$

$$f_n(x) = \frac{e^{nx}}{x + e^{2nx}}, \quad x > 0.$$

Exercise 3.64 Find the values $1 \le p \le \infty$ for which the following sequence is convergent in $L^p(\mathbb{R})$

$$f_n(x) = \frac{\arctan nx}{1 + x^n}, \quad x \in \mathbb{R}.$$

Exercise 3.65 Find the values $1 \le p \le \infty$ for which the following sequence is convergent in $L^p(1, \infty)$

$$f_n(x) = \frac{n}{\sqrt{x}(n+x)}, \qquad x > 1.$$

References

[Ha50] Halmos, P.R.: Measure Theory. Van Nostrand, Princeton (1950)
[SW71] Stein, E.M., Weiss, G.: Introduction to Fourier Analysis on Euclidean Spaces. Princeton University Press, Princeton (1971)
[No72] Novinger, W.P.: Mean convergence in L^p spaces. Proc. Amer. Math. Soc. **34**, 627–628 (1972)

Chapter 4
Product Measures

On the Cartesian product of two measure spaces one can construct a measure—hence, an integral—which is directly connected with the measure on each factor. Then, the natural problem that arises is how to reduce a double (or multiple) integral to the computation of two (or more) simple integrals. Such a question plays a crucial role in Lebesgue integration.

The key results of the theory are Tonelli's Theorem and Fubini's Theorem, which provide sufficient conditions to compute a double integral by iterated integrations on the factors.

Such theorems have important consequences when applied to the product space $\mathbb{R}^{2N} = \mathbb{R}^N \times \mathbb{R}^N$. First, one can characterize the families of functions in $L^p(\mathbb{R}^N)$ with compact closure, thus obtaining an L^p-version of the Ascoli–Arzelà Theorem. Another important application of multiple integration is the study of the *convolution product* $f * g$ of two Lebesgue functions. Such an operation commutes with translation and derivation, and is a powerful tool to approximate the elements of $L^p(\mathbb{R}^N)$ by smooth functions.

4.1 Product Spaces

4.1.1 Product Measures

Let $(X, \mathscr{F})$ and $(Y, \mathscr{G})$ be measurable spaces. We will turn the Cartesian product $X \times Y$ into a measurable space in a canonical way. A set of the form $A \times B$, where $A \in \mathscr{F}$ and $B \in \mathscr{G}$, is called a *measurable rectangle*. Let us denote by $\mathscr{R}$ the family of all elementary sets, where by an *elementary set* we mean any finite disjoint union of measurable rectangles.

Proposition 4.1 *$\mathscr{R}$ is an algebra.*

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_4

Proof Clearly, $\varnothing$ and $X \times Y$ are measurable rectangles. It is also obvious that the intersection of any two measurable rectangles is again a measurable rectangle. Moreover, the intersection of any two elements of $\mathscr{R}$ stays in $\mathscr{R}$. Indeed, let[1] $\dot\cup_i (A_i \times B_i)$ and $\dot\cup_j (C_j \times D_j)$ be finite disjoint unions of measurable rectangles. Then

$$\Big(\dot\cup_i (A_i \times B_i)\Big) \bigcap \Big(\dot\cup_j (C_j \times D_j)\Big) = \dot\cup_{i,j}\Big((A_i \times B_i) \cap (C_j \times D_j)\Big) \in \mathscr{R}.$$

Let us show that the complement of any set $E \in \mathscr{R}$ is again in $\mathscr{R}$. This is true if $E = A \times B$ is a measurable rectangle since

$$E^c = (A^c \times B)\dot\cup(A \times B^c)\dot\cup(A^c \times B^c).$$

Now, proceeding by induction, let

$$E = \Big(\underbrace{\dot\bigcup_{i=1}^{n}(A_i \times B_i)}_{F}\Big)\dot\bigcup(A_{n+1} \times B_{n+1}) \in \mathscr{R}$$

and suppose $F^c \in \mathscr{R}$. Then $E^c = F^c \cap (A_{n+1} \times B_{n+1})^c \in \mathscr{R}$ because $(A_{n+1} \times B_{n+1})^c \in \mathscr{R}$ and we have already proved that $\mathscr{R}$ is closed under finite intersection. This completes the proof. □

Definition 4.2 The σ-algebra generated by $\mathscr{R}$ is called the *product σ-algebra* of $\mathscr{F}$ and $\mathscr{G}$ and is denoted by $\mathscr{F} \times \mathscr{G}$.

Exercise 4.3 Prove that:

- $\mathscr{B}\big([a,b) \times [c,d)\big) = \mathscr{B}\big([a,b)\big) \times \mathscr{B}\big([c,d)\big)$.
- If $N,\ N' \in \mathbb{N}$, then $\mathscr{B}(\mathbb{R}^{N+N'}) = \mathscr{B}(\mathbb{R}^N) \times \mathscr{B}(\mathbb{R}^{N'})$.

For any $E \in \mathscr{F} \times \mathscr{G}$, we define the *sections* of E as follows:

$$E_x = \{y \in Y : \ (x,y) \in E\}, \quad E^y = \{x \in X : \ (x,y) \in E\} \quad \forall x \in X\,, \ \forall y \in Y.$$

Proposition 4.4 *Let $(X, \mathscr{F}, \mu)$ and $(Y, \mathscr{G}, \nu)$ be σ-finite measure spaces. If $E \in \mathscr{F} \times \mathscr{G}$, then the following statements hold:*

(a) $E_x \in \mathscr{G}$ *and* $E^y \in \mathscr{F}$ *for any* $(x,y) \in X \times Y$.
(b) The functions

$$\begin{cases} X \to \mathbb{R} \\ x \mapsto \nu(E_x) \end{cases} \quad \textit{and} \quad \begin{cases} Y \to \mathbb{R} \\ y \mapsto \mu(E^y) \end{cases}$$

are Borel. Moreover,

$$\int_X \nu(E_x)\, d\mu = \int_Y \mu(E^y)\, d\nu.$$

[1] The symbol $\dot\cup$ denotes a disjoint union.

Proof Suppose, first, that $E = \dot{\bigcup}_{i=1}^n (A_i \times B_i)$ belongs to $\mathscr{R}$. Then for $(x, y) \in X \times Y$ we have $E_x = \dot{\bigcup}_{i=1}^n (A_i \times B_i)_x$ and $E^y = \dot{\bigcup}_{i=1}^n (A_i \times B_i)^y$, where

$$(A_i \times B_i)_x = \begin{cases} B_i & \text{if } x \in A_i, \\ \varnothing & \text{if } x \notin A_i, \end{cases} \qquad (A_i \times B_i)^y = \begin{cases} A_i & \text{if } y \in B_i, \\ \varnothing & \text{if } y \notin B_i. \end{cases}$$

Consequently,

$$\nu(E_x) = \sum_{i=1}^n \nu\big((A_i \times B_i)_x\big) = \sum_{i=1}^n \nu(B_i)\chi_{A_i}(x),$$

$$\mu(E^y) = \sum_{i=1}^n \mu\big((A_i \times B_i)^y\big) = \sum_{i=1}^n \mu(A_i)\chi_{B_i}(y).$$

The conclusion follows and can be easily extended to elementary sets.

Now, let $\mathscr{E}$ be the family of all sets $E \in \mathscr{F} \times \mathscr{G}$ satisfying (a). Clearly, $\emptyset,\ X \times Y \in \mathscr{E}$. Furthermore, for any $E_n,\ E \in \mathscr{E}$ and $(x, y) \in X \times Y$ we have

$$(E^c)_x = (E_x)^c, \quad (E^c)^y = (E^y)^c,$$

$$\bigcup_{n=1}^\infty (E_n)_x = \left(\bigcup_{n=1}^\infty E_n\right)_x, \quad \bigcup_{n=1}^\infty (E_n)^y = \left(\bigcup_{n=1}^\infty E_n\right)^y.$$

Hence, $\mathscr{E}$ is a σ-algebra including $\mathscr{R}$ and, consequently, $\mathscr{E} = \mathscr{F} \times \mathscr{G}$.

We now prove (b). Assume first that μ and ν are finite and define

$$\mathscr{M} = \big\{E \in \mathscr{F} \times \mathscr{G} \,\big|\, E \text{ satisfies (b)}\big\}.$$

We claim that $\mathscr{M}$ is a monotone class. Indeed, consider $(E_n)_n \subset \mathscr{M}$ such that $E_n \uparrow E$. Then, for any $(x, y) \in X \times Y$,

$$(E_n)_x \uparrow E_x \quad \text{and} \quad (E_n)^y \uparrow E^y.$$

Thus,

$$\nu\big((E_n)_x\big) \uparrow \nu(E_x) \quad \text{and} \quad \mu\big((E_n)^y\big) \uparrow \mu(E^y).$$

Since the function $x \mapsto \nu\big((E_n)_x\big)$ is Borel for all $n \in \mathbb{N}$, so is $x \mapsto \nu(E_x)$. Similarly, $y \mapsto \mu(E^y)$ is Borel. Furthermore, by the Monotone Convergence Theorem,

$$\int_X \nu(E_x)\, d\mu = \lim_{n\to\infty} \int_X \nu\big((E_n)_x\big)\, d\mu = \lim_{n\to\infty} \int_Y \mu\big((E_n)^y\big)\, d\nu = \int_Y \mu(E^y)\, d\nu.$$

So $E \in \mathscr{M}$. Next, consider $(E_n)_n \subset \mathscr{M}$ such that $E_n \downarrow E$. Then a similar argument as above shows that, for every $(x, y) \in X \times Y$,

$$\nu\big((E_n)_x\big) \downarrow \nu(E_x) \quad \text{and} \quad \mu\big((E_n)^y\big) \downarrow \mu(E^y).$$

Consequently the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are Borel. Furthermore,

$$\nu\big((E_n)_x\big) \leq \nu(Y) \quad \forall x \in X, \qquad \mu\big((E_n)^y\big) \leq \mu(X) \quad \forall y \in Y,$$

and, since μ and ν are finite, all constant functions are summable. Then Lebesgue's Theorem yields

$$\int_X \nu(E_x)\, d\mu = \lim_{n\to\infty} \int_X \nu\big((E_n)_x\big)\, d\mu = \lim_{n\to\infty} \int_Y \mu\big((E_n)^y\big)\, d\nu = \int_Y \mu(E^y)\, d\nu,$$

which implies $E \in \mathscr{M}$. Therefore $\mathscr{M}$ is a monotone class as claimed. By the first part of the proof, we deduce that $\mathscr{R} \subset \mathscr{M}$. So Halmos' Theorem yields $\mathscr{M} = \mathscr{F} \times \mathscr{G}$, which proves the conclusion when μ and ν are finite.

Now, suppose μ and ν are σ-finite, so that $X = \cup_{n=1}^{\infty} X_n$, $Y = \cup_{n=1}^{\infty} Y_n$ for some increasing sequences $(X_n)_n \subset \mathscr{F}$ and $(Y_n)_n \subset \mathscr{G}$ with

$$\mu(X_n) < \infty, \quad \nu(Y_n) < \infty \quad \forall n \in \mathbb{N}. \tag{4.1}$$

Define $\mu_n = \mu \llcorner X_n$, $\nu_n = \nu \llcorner Y_n$ (see Definition 1.26) and fix $E \in \mathscr{F} \times \mathscr{G}$. For any $(x, y) \in X \times Y$

$$E_x \cap Y_n \uparrow E_x \quad \text{and} \quad E^y \cap X_n \uparrow E^y.$$

Thus

$$\nu_n(E_x) = \nu\big(E_x \cap Y_n\big) \uparrow \nu(E_x) \quad \text{and} \quad \mu_n(E^y) = \mu\big(E^y \cap X_n\big) \uparrow \mu(E^y).$$

Since μ_n and ν_n are finite measures, for all $n \in \mathbb{N}$ the function $x \mapsto \nu_n(E_x)$ is Borel; therefore $x \mapsto \nu(E_x)$ is also Borel. A similar argument proves that $y \mapsto \mu(E^y)$ is Borel. Furthermore, by the Monotone Convergence Theorem and Exercise 2.72,

$$\int_X \nu(E_x)\, d\mu = \lim_{n\to\infty} \int_{X_n} \nu_n(E_x)\, d\mu = \lim_{n\to\infty} \int_X \nu_n(E_x)\, d\mu_n.$$

$$\int_Y \mu(E^y)\, d\nu = \lim_{n\to\infty} \int_{Y_n} \mu_n(E^y)\, d\nu = \lim_{n\to\infty} \int_Y \mu_n(E^y)\, d\nu_n.$$

Since μ_n, ν_n are finite measures, then $\int_X \nu_n(E_x)\, d\mu_n = \int_Y \mu_n(E^y)\, d\nu_n$ for every n, and so $\int_X \nu(E_x)\, d\mu = \int_Y \mu(E^y)\, d\nu$. □

Theorem 4.5 *Let $(X, \mathscr{F}, \mu)$ and $(Y, \mathscr{G}, \nu)$ be σ-finite measure spaces. The set function $\mu \times \nu$ defined by*

$$(\mu \times \nu)(E) = \int_X \nu(E_x)\, d\mu = \int_Y \mu(E^y)\, d\nu \qquad \forall E \in \mathscr{F} \times \mathscr{G} \tag{4.2}$$

is a σ-finite measure on $\mathscr{F} \times \mathscr{G}$, called the product measure *of μ and ν. Moreover, if λ is any measure on $\mathscr{F} \times \mathscr{G}$ satisfying*

$$\lambda(A \times B) = \mu(A)\nu(B) \qquad \forall A \in \mathscr{F},\ \forall B \in \mathscr{G}, \tag{4.3}$$

then $\lambda = \mu \times \nu$.

Proof First, to check that $\mu \times \nu$ is σ-additive, let $(E_n)_n$ be a sequence of disjoint sets in $\mathscr{F} \times \mathscr{G}$. Then, for any $(x, y) \in X \times Y$, $((E_n)_x)_n$ and $((E_n)^y)_n$ are disjoint families in $\mathscr{G}$ and $\mathscr{F}$, respectively. Therefore, by Proposition 2.48,

$$\begin{aligned}(\mu \times \nu)\left(\bigcup_{n=1}^{\infty} E_n\right) &= \int_X \nu\left(\left(\cup_{n=1}^{\infty} E_n\right)_x\right) d\mu = \int_X \sum_{n=1}^{\infty} \nu\big((E_n)_x\big)\, d\mu \\ &= \sum_{n=1}^{\infty} \int_X \nu\big((E_n)_x\big)\, d\mu = \sum_{n=1}^{\infty} (\mu \times \nu)(E_n).\end{aligned}$$

To prove that $\mu \times \nu$ is σ-finite, observe that if $(X_n)_n \subset \mathscr{F}$ and $(Y_n)_n \subset \mathscr{G}$ are two increasing sequences such that $X = \cup_{n=1}^{\infty} X_n$, $Y = \cup_{n=1}^{\infty} Y_n$ and

$$\mu(X_n) < \infty, \quad \nu(Y_n) < \infty \quad \forall n \in \mathbb{N},$$

then, setting $Z_n = X_n \times Y_n$, we have that $Z_n \in \mathscr{F} \times \mathscr{G}$,

$$(\mu \times \nu)(Z_n) = \mu(X_n)\nu(Y_n) < \infty,$$

and $X \times Y = \cup_{n=1}^{\infty} Z_n$. Finally, if λ is a measure on $\mathscr{F} \times \mathscr{G}$ satisfying (4.3), then λ and $\mu \times \nu$ coincide on $\mathscr{R}$. So Theorem 1.32 ensures that λ and $\mu \times \nu$ coincide on $\sigma(\mathscr{R})$. □

The following result is a straightforward consequence of (4.2).

Corollary 4.6 *Under the same assumptions of Theorem 4.5, let $E \in \mathscr{F} \times \mathscr{G}$ be such that $(\mu \times \nu)(E) = 0$. Then $\mu(E^y) = 0$ for almost every $y \in Y$, and $\nu(E_x) = 0$ for almost every $x \in X$.*

Example 4.7 We note that $\mu \times \nu$ may fail to be a complete measure even when both μ and ν are complete. Indeed, let m denote the Lebesgue measure on $\mathbb{R}$ and $\mathscr{G}$ the σ-algebra of all Lebesgue measurable sets in $\mathbb{R}$ (see Definition 1.56). Let $A \in \mathscr{G}$ be a nonempty zero-measure set and let $B \subset \mathbb{R}$ be a set which is not Lebesgue measurable (see Example 1.66). Then $A \times B \subset A \times \mathbb{R}$ and $(m \times m)(A \times \mathbb{R}) = 0$. On the other hand, $A \times B \notin \mathscr{G} \times \mathscr{G}$, otherwise one would get a contradiction with Proposition 4.4(a).

4.1.2 Fubini-Tonelli Theorem

In this section we will reduce the computation of a double integral with respect to the product measure $\mu \times \nu$ to the computation of two simple integrals. The next two results are fundamental in multiple integration.

Theorem 4.8 (Tonelli) *Let $(X, \mathscr{F}, \mu)$ and $(Y, \mathscr{G}, \nu)$ be σ-finite measure spaces and let $F : X \times Y \to [0, \infty]$ be a Borel function. Then:*

(a) (i) For every $x \in X$ the function $F(x, \cdot) : y \mapsto F(x, y)$ is Borel.
(ii) For every $y \in Y$ the function $F(\cdot, y) : x \mapsto F(x, y)$ is Borel.

(b) (i) The function $x \mapsto \int_Y F(x, y)\, d\nu(y)$ is Borel.
(ii) The function $y \mapsto \int_X F(x, y)\, d\mu(x)$ is Borel.

(c) The following identities hold:

$$\int_{X \times Y} F(x, y)\, d(\mu \times \nu)(x, y) = \int_X \left[\int_Y F(x, y)\, d\nu(y) \right] d\mu(x) \tag{4.4}$$

$$= \int_Y \left[\int_X F(x, y)\, d\mu(x) \right] d\nu(y). \tag{4.5}$$

Proof Assume, first, that $F = \chi_E$ with $E \in \mathscr{F} \times \mathscr{G}$. Then

$$F(x, \cdot) = \chi_{E_x} \qquad \forall x \in X,$$
$$F(\cdot, y) = \chi_{E^y} \qquad \forall y \in Y.$$

So properties (a) and (b) follow from Proposition 4.4, while (c) reduces to formula (4.2) used to define the product measure. Consequently, the thesis holds true when F is a simple function. In the general case, owing to Proposition 2.46 we can approximate F pointwise by an increasing sequence of simple functions

$$F_n : X \times Y \to [0, \infty).$$

For every $x \in X$, $F_n(x, \cdot)$ is a sequence of Borel functions on Y such that

$$F_n(x, \cdot) \uparrow F(x, \cdot) \text{ pointwise as } n \to \infty.$$

So the function $F(x, \cdot)$ is Borel and (a)-(i) is proven. Moreover, by the first part of the proof, $x \mapsto \int_Y F_n(x, y)\, d\nu(y)$ is an increasing sequence of Borel functions satisfying, thanks to Monotone Convergence Theorem,

$$\int_Y F_n(x,y)\,d\nu(y) \quad \uparrow \quad \int_Y F(x,y)\,d\nu(y) \qquad \forall x \in X.$$

Hence, (b)-(i) holds true and, again by monotone convergence,

$$\int_X \left[\int_Y F_n(x,y)\,d\nu(y)\right] d\mu(x) \quad \uparrow \quad \int_X \left[\int_Y F(x,y)\,d\nu(y)\right] d\mu(x). \tag{4.6}$$

We also have

$$\int_{X\times Y} F_n(x,y)\,d(\mu\times\nu)(x,y) \quad \uparrow \quad \int_{X\times Y} F(x,y)\,d(\mu\times\nu)(x,y). \tag{4.7}$$

Since each F_n is a simple function, the left-hand sides of (4.6) and (4.7) are equal. Therefore the right-hand sides also coincide and this proves (4.4). By a similar argument one can show (a)-(ii), (b)-(ii), and (4.5). □

Theorem 4.9 (Fubini) *Let $(X,\mathscr{F},\mu)$, $(Y,\mathscr{G},\nu)$ be σ-finite measure spaces and let $F : X\times Y \to \overline{\mathbb{R}}$ be a $(\mu\times\nu)$-summable function. The the following statements hold:*

(a) (i) For almost every $x \in X$ the function $F(x,\cdot) : y \mapsto F(x,y)$ is ν-summable.
(ii) For almost every $y \in Y$ the function $F(\cdot,y) : x \mapsto F(x,y)$ is μ-summable.

(b) (i) The function[2] $x \mapsto \int_Y F(x,y)\,d\nu(y)$ is μ-summable.
(ii) The function $y \mapsto \int_X F(x,y)\,d\mu(x)$ is ν-summable.

(c) Identities (4.4) and (4.5) are valid.

Proof Let F^+ and F^- be the positive and negative parts of F. Then Theorem 4.8 applies to F^+ and F^-. In particular, since $\int_{X\times Y} F^{\pm}\,d(\mu\times\nu) \le \int_{X\times Y} |F|\,d(\mu\times\nu) < \infty$, identity (4.4) implies

$$\int_X \left[\int_Y F^{\pm}(x,y)\,d\nu(y)\right] d\mu(x) < \infty.$$

So the functions

$$x \mapsto \int_Y F^{\pm}(x,y)d\nu(y) \tag{4.8}$$

are μ-summable and, owing to Proposition 2.44(i), a.e. finite, that is,

$$\int_Y F^{\pm}(x,y)\,d\nu(y) < \infty \quad \text{for almost every } x \in X.$$

It follows that $F^{\pm}(x,\cdot)$ is ν-summable for almost every $x \in X$. Since

[2] Observe that the function $x \mapsto \int_Y F(x,y)\,d\nu(y)$ is defined a.e. in X, the exceptional zero-measure set consisting of those points where $F(x,\cdot)$ fails to be ν-summable. Similarly, the function $y \mapsto \int_X F(x,y)\,d\mu(x)$ is defined a.e. in Y. See Remark 2.74.

$$F^+(x,\cdot) - F^-(x,\cdot) = F(x,\cdot) \quad \forall x \in X, \tag{4.9}$$

we deduce (a)-(i).

We observe that, for every x such that $F(x,\cdot)$ is ν-summable, we can integrate identity (4.9) to obtain

$$\int_Y F^+(x,y)\,d\nu(y) - \int_Y F^-(x,y)\,d\nu(y) = \int_Y F(x,y)\,d\nu(y). \tag{4.10}$$

(b)-(i) holds true for F^+ and F^-, hence for F by (4.10). By interchanging the role of X and Y, one can prove (a)-(ii) and (b)-(ii).

Finally, identities (4.4) and (4.5) hold for $F^\pm$; by subtraction, we obtain the analogous identities for F. □

Example 4.10 Let $X = Y = [-1,1)$ with Lebesgue measure and consider the function

$$f(x,y) = \frac{xy}{(x^2+y^2)^2} \text{ for } (x,y) \neq (0,0).$$

By completing the definition of f arbitrarily in $(0,0)$, it follows at once that f is Borel. The iterated integrals exist and are equal; indeed

$$\int_{-1}^1 \left[\int_{-1}^1 f(x,y)\,dx\right] dy = \int_{-1}^1 \left[\int_{-1}^1 f(x,y)\,dy\right] dx = 0.$$

On the other hand the double integral fails to exist, since

$$\int_{[-1,1)^2} |f(x,y)|\,dxdy \geq \int_0^1 \left[\int_0^{2\pi} \frac{|\sin\theta\cos\theta|}{r}\,d\theta\right] dr = 2\int_0^1 \frac{dr}{r} = \infty.$$

This example shows that the existence of the iterated integrals does not imply the existence of the double integral, in general.

Example 4.11 Consider the spaces

$$([0,1], \mathscr{P}([0,1]), \mu^\#) \quad \text{and} \quad ([0,1], \mathscr{B}([0,1]), m),$$

where $\mu^\#$ and m denote the counting measure and Lebesgue measure on $[0,1]$. respectively. Let Δ be the diagonal of $[0,1]^2$, that is,

$$\Delta = \{(x,x) \mid x \in [0,1]\}.$$

For every $n \in \mathbb{N}$, set

$$R_n = \left[0, \frac{1}{n}\right]^2 \cup \left[\frac{1}{n}, \frac{2}{n}\right]^2 \cup \ldots \cup \left[\frac{n-1}{n}, 1\right]^2.$$

R_n is a finite union of measurable rectangles and $\Delta = \cap_{n=1}^{\infty} R_n$. So Δ belongs to $\mathscr{P}([0,1]) \times \mathscr{B}([0,1])$ and χ_Δ is measurable. Moreover,

$$\int_0^1 \left[\int_{[0,1]} \chi_\Delta(x,y)\, d\mu^\#(x)\right] dy = \int_0^1 1\, dy = 1,$$

$$\int_0^1 \left[\int_0^1 \chi_\Delta(x,y)\, dy\right] d\mu^\#(x) = \int_{[0,1]} 0\, d\mu^\# = 0,$$

which shows that the conclusion of Tonelli's Theorem may be false if μ is not σ-finite.

4.2 Compactness in L^p

We shall derive important results by using of Fubini's and Tonelli's Theorems. The first one is the characterization of all *relatively compact* subsets[3] of $L^p(\mathbb{R}^N)$ for any $1 \le p < \infty$, that is, all families of functions $\mathscr{M} \subset L^p(\mathbb{R}^N)$ with compact closure $\overline{\mathscr{M}}$.

First, we need the following lemma.

Lemma 4.12 *If $f : \mathbb{R}^N \to \overline{\mathbb{R}}$ is Borel, then the functions*

$$(x,y) \in \mathbb{R}^{2N} \mapsto f(x-y) \quad \textit{and} \quad (x,y) \in \mathbb{R}^{2N} \mapsto f(x+y)$$

are also Borel.

Proof Let $F_1 : (x,t) \in \mathbb{R}^{2N} \mapsto f(x)$. Since f is Borel, it follows that F_1 is Borel. Indeed, the set $\{(x,t) \mid F_1(x,t) > a\}$ coincides with the measurable rectangle $\{x \mid f(x) > a\} \times \mathbb{R}^N$. Given $(\xi,\eta) \in \mathbb{R}^{2N}$, consider the nonsingular linear transformation of $\mathbb{R}^{2N}$: $x = \xi - \eta$, $y = \xi + \eta$. Owing to Exercise 2.15, the function $F_2(\xi,\eta) = F_1(\xi-\eta, \xi+\eta)$ is Borel on $\mathbb{R}^{2N}$. Since $F_2(\xi,\eta) = f(\xi-\eta)$, the first part of the conclusion follows. The second one can be proved by a similar argument. □

Definition 4.13 Let $1 \le p < \infty$. For every $r > 0$ and $f \in L^p(\mathbb{R}^N)$ define $S_r f : \mathbb{R}^N \to \mathbb{R}$ by the Steklov formula

$$S_r f(x) = \frac{1}{\omega_N r^N} \int_{\|y\|<r} f(x+y)\, dy \qquad \forall x \in \mathbb{R}^N,$$

where ω_N is the volume of the unit ball $\mathbb{R}^N$.

[3] $L^p(\mathbb{R}^N) = L^p(\mathbb{R}^N, m)$ where m denotes the Lebesgue measure.

Proposition 4.14 *Let* $1 \leq p < \infty$ *and* $f \in L^p(\mathbb{R}^N)$*. Then for every* $r > 0$ $S_r f$ *is a continuous function. Furthermore* $S_r f \in L^p(\mathbb{R}^N)$ *and, using the notation* $\tau_h f(x) = f(x+h)$*, we have:*

$$|S_r f(x)| \leq \frac{1}{(\omega_N r^N)^{1/p}} \|f\|_p \quad \forall x \in \mathbb{R}^N; \tag{4.11}$$

$$|S_r f(x) - S_r f(x+h)| \leq \frac{1}{(\omega_N r^N)^{1/p}} \|f - \tau_h f\|_p \quad \forall x,\ h \in \mathbb{R}^N; \tag{4.12}$$

$$\|S_r f\|_p \leq \|f\|_p;$$

$$\|f - S_r f\|_p \leq \sup_{0 \leq \|h\| \leq r} \|f - \tau_h f\|_p. \tag{4.13}$$

Proof (4.11) can be derived using Hölder's inequality:

$$|S_r f(x)| \leq \frac{1}{(\omega_N r^N)^{1/p}} \left(\int_{\|y\|<r} |f(x+y)|^p \, dy \right)^{1/p}. \tag{4.14}$$

(4.12) follows from (4.11) applied to $f - \tau_h f$. Thus, (4.12) and Proposition 3.49 imply that $S_r f$ is a continuous function. By (4.14), using Lemma 4.12 and Tonelli's Theorem, we get

$$\begin{aligned}\int_{\mathbb{R}^N} |S_r f|^p \, dx &\leq \frac{1}{\omega_N r^N} \int_{|y|<r} \left[\int_{\mathbb{R}^N} |f(x+y)|^p \, dx \right] dy \\ &= \frac{\|f\|_p^p}{\omega_N r^N} \int_{\|y\|<r} dy = \|f\|_p^p.\end{aligned}$$

To obtain (4.13), observe that $(f - S_r f)(x) = \frac{1}{\omega_N r^N} \int_{\|y\|<r} (f(x) - f(x+y)) dy$. So

$$|(f - S_r f)(x)| \leq \frac{1}{(\omega_N r^N)^{1/p}} \left(\int_{\|y\|<r} |f(x) - f(x+y)|^p dy \right)^{1/p}.$$

Therefore Tonelli's Theorem yields

$$\begin{aligned}\int_{\mathbb{R}^N} |f - S_r f|^p \, dx &\leq \frac{1}{\omega_N r^N} \int_{\mathbb{R}^N} \left[\int_{\|y\|<r} |f(x) - f(x+y)|^p \, dy \right] dx \\ &= \frac{1}{\omega_N r^N} \int_{\|y\|<r} \left[\int_{\mathbb{R}^N} |f(x) - f(x+y)|^p \, dx \right] dy \\ &\leq \sup_{0 \leq \|h\| \leq r} \|f - \tau_h f\|_p^p \frac{\int_{\|y\|<r} dy}{\omega_N r^N} = \sup_{0 \leq \|h\| \leq r} \|f - \tau_h f\|_p^p.\end{aligned}$$

Inequality (4.13) is thus proved. □

The role played by the following theorem in the study of L^p-spaces is similar to the one played by the Ascoli–Arzelà Theorem (Theorem E.2) for continuous functions.

Theorem 4.15 (M. Riesz–Frechét–Kolmogorov) *Let* $1 \le p < \infty$ *and let* $\mathscr{M}$ *be a bounded set in* $L^p(\mathbb{R}^N)$. *Then* $\mathscr{M}$ *is relatively compact if and only if*

$$\sup_{f\in\mathscr{M}} \int_{\|x\|>R} |f|^p\,dx \to 0 \quad as \quad R\to\infty, \tag{4.15}$$

$$\sup_{f\in\mathscr{M}} \int_{\mathbb{R}^N} |f(x+h)-f(x)|^p\,dx \to 0 \quad as \quad h\to 0. \tag{4.16}$$

Proof To begin with, observe that (4.15) and (4.16) hold for a single element of $L^p(\mathbb{R}^N)$ ((4.15) follows from Lebesgue's Theorem; see Proposition 3.49 for (4.16)). Let us consider the balls in $L^p(\mathbb{R}^N)$:

$$B_r(f) := \left\{g\in L^p(\mathbb{R}^N) \mid \|f-g\|_p < r\right\} \qquad r>0,\ f\in L^p(\mathbb{R}^N).$$

If $\mathscr{M}$ is relatively compact, then for any $\varepsilon>0$ there exist functions $f_1,\dots,f_n\in\mathscr{M}$ such that $\mathscr{M}\subset B_\varepsilon(f_1)\cup\cdots\cup B_\varepsilon(f_n)$. As we have just recalled, each f_i satisfies (4.15) and (4.16). So there exist $R_\varepsilon, \delta_\varepsilon>0$ such that, for every $i=1,\dots,n$,

$$\int_{\|x\|>R_\varepsilon} |f_i|^p\,dx < \varepsilon^p \quad \& \quad \|f_i-\tau_h f_i\|_p<\varepsilon \text{ if } \|h\|<\delta_\varepsilon\,, \tag{4.17}$$

where $\tau_h f(x) = f(x+h)$ for any $x,h\in\mathbb{R}^N$. Let $f\in\mathscr{M}$. Then $f\in B_\varepsilon(f_i)$ for some $i=1,\dots,n$. By (4.17), using Minkowski's inequality, we have

$$\begin{aligned}\left(\int_{\|x\|>R_\varepsilon}|f|^p\,dx\right)^{1/p} &\le \left(\int_{\|x\|>R_\varepsilon}|f-f_i|^p\,dx\right)^{1/p} + \left(\int_{\|x\|>R_\varepsilon}|f_i|^p\,dx\right)^{1/p}\\ &\le \|f-f_i\|_p + \left(\int_{\|x\|>R_\varepsilon}|f_i|^p\,dx\right)^{1/p} < 2\varepsilon\end{aligned}$$

and, if $\|h\|\le\delta_\varepsilon$,

$$\|f-\tau_h f\|_p \le \|f-f_i\|_p + \|f_i-\tau_h f_i\|_p + \|\tau_h f_i - \tau_h f\|_p < 3\varepsilon.$$

The implication '⇒' is thus proved.

To prove the converse it suffices to show that $\mathscr{M}$ is *totally bounded*.[4] Let $\varepsilon>0$ be fixed. On account of assumption (4.15), we have that

[4]Given a metric space (X,d) and a subset $\mathscr{M}\subset X$, we say that $\mathscr{M}$ is *totally bounded* if for any $\varepsilon>0$ there exist finitely many balls of radius ε covering $\mathscr{M}$. A subset $\mathscr{M}$ of a complete metric space X is relatively compact if and only if it is totally bounded.

$$\exists R_\varepsilon > 0 \quad \text{such that} \quad \int_{\|x\|>R_\varepsilon} |f|^p\,dx < \varepsilon^p \qquad \forall f \in \mathscr{M}. \tag{4.18}$$

Also, recalling (4.13), assumption (4.16) implies

$$\exists \delta_\varepsilon > 0 \quad \text{such that} \quad \|f - S_{\delta_\varepsilon} f\|_p < \varepsilon \qquad \forall f \in \mathscr{M}, \tag{4.19}$$

where S_{δ_ε} is the Steklov operator of Definition 4.13. Moreover, properties (4.11) and (4.12) ensure that $\{S_{\delta_\varepsilon} f\}_{f\in\mathscr{M}}$ is a pointwise bounded and equicontinuous family on the compact set $K_\varepsilon := \{x \in \mathbb{R}^N : \|x\| \le R_\varepsilon\}$ and, consequently, is relatively compact in the space $\mathscr{C}(K_\varepsilon)$ thanks to Ascoli–Arzelà Theorem. Thus, there exists a finite number of continuous functions $g_1, \ldots, g_m : K_\varepsilon \to \mathbb{R}$ such that for each $f \in \mathscr{M}$ the function $S_{\delta_\varepsilon} f$ satisfies, for some j,

$$\left|S_{\delta_\varepsilon} f(x) - g_j(x)\right| < \frac{\varepsilon}{\left(\omega_N R_\varepsilon^N\right)^{1/p}} \qquad \forall x \in K_\varepsilon. \tag{4.20}$$

Set

$$f_j(x) := \begin{cases} g_j(x) & \text{if } \|x\| \le R_\varepsilon, \\ 0 & \text{if } \|x\| > R_\varepsilon. \end{cases}$$

Then $f_j \in L^p(\mathbb{R}^N)$ and, by (4.18)–(4.20),

$$\begin{aligned} \|f - f_j\|_p &= \left(\int_{\|x\|>R_\varepsilon} |f|^p\,dx\right)^{1/p} + \left(\int_{K_\varepsilon} |f - g_j|^p\,dx\right)^{1/p} \\ &< \varepsilon + \left(\int_{K_\varepsilon} |f - S_{\delta_\varepsilon} f|^p\,dx\right)^{1/p} + \left(\int_{K_\varepsilon} |S_{\delta_\varepsilon} f - g_j|^p\,dx\right)^{1/p} < 3\varepsilon. \end{aligned}$$

This shows that $\mathscr{M}$ is totally bounded and completes the proof. □

4.3 Convolution and Approximation

In this section we will develop a systematic procedure for approximating a L^p function by smooth functions. The operation of convolution[5] provides the tool to build such smooth approximations. In what follows, the measure space of interest is always $\mathbb{R}^N$ with the Lebesgue measure.

[5]The notion of convolution, extended to distributions (see [Ru73]), plays a fundamental role in the applications to differential equations.

4.3.1 Convolution Product

Definition 4.16 Let $f, g : \mathbb{R}^N \to \overline{\mathbb{R}}$ be Borel functions and $x \in \mathbb{R}^N$ such that the function

$$y \in \mathbb{R}^N \mapsto f(x-y)g(y) \tag{4.21}$$

is integrable.[6] The *convolution product* $(f * g)(x)$ is defined by

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

Remark 4.17 If $f, g : \mathbb{R}^N \to [0, \infty]$ are Borel, then the function (4.21) is positive and Borel for every $x \in \mathbb{R}^N$. It follows that the product $(f * g)(x)$ is well defined for every $x \in \mathbb{R}^N$; moreover $f * g : \mathbb{R}^N \to [0, \infty]$ is Borel by Tonelli's Theorem and Lemma 4.12.

Remark 4.18 By the change of variable $z = x - y$ and the translation invariance of the Lebesgue measure, we deduce that the function (4.21) is integrable if and only if the function $z \in \mathbb{R}^N \mapsto f(z)g(x-z)$ is integrable and $(f * g)(x) = (g * f)(x)$. This proves that convolution is commutative.

Our next result gives a sufficient condition to guarantee that the product $f * g$ is defined a.e. in $\mathbb{R}^N$.

Theorem 4.19 (Young) *Let $1 \le p, q, r \le \infty$ be such that*[7]

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1, \tag{4.22}$$

and let $f \in L^p(\mathbb{R}^N)$, $g \in L^q(\mathbb{R}^N)$. Then for almost every $x \in \mathbb{R}^N$ the function (4.21) is summable. Moreover[8] *$f * g \in L^r(\mathbb{R}^N)$ and*

$$\|f * g\|_r \le \|f\|_p \, \|g\|_q. \tag{4.23}$$

Finally, if $r = \infty$, then the function (4.21) is summable for every $x \in \mathbb{R}^N$ and $f * g$ is continuous on $\mathbb{R}^N$.

Proof Assume first $r = \infty$. Then $1/p + 1/q = 1$. By the translation invariance of the Lebesgue measure, for every $x \in \mathbb{R}^N$ the function $y \in \mathbb{R}^N \to f(x-y)$ belongs to $L^p(\mathbb{R}^N)$ and has the same L^p-norm as f. Hölder's inequality and Exercise 3.26 imply that for every $x \in \mathbb{R}^N$ the function (4.21) is summable and

[6] See Remark 2.67 for the definition of integrability.

[7] Hereafter we will adopt the convention $\frac{1}{\infty} = 0$.

[8] Observe that, in general, $f * g$ is defined a.e. in $\mathbb{R}^N$ (see Remark 2.74).

$$|(f * g)(x)| \le \|f\|_p \|g\|_q \quad \forall x \in \mathbb{R}^N. \tag{4.24}$$

Since at least one between p and q is finite and convolution is commutative we may assume, without loss of generality, $p < \infty$. Then, for any $x, h \in \mathbb{R}^N$, inequality (4.24) yields

$$|(f * g)(x+h) - (f * g)(x)| = |((\tau_h f - f) * g)(x)| \le \|\tau_h f - f\|_p \|g\|_q ,$$

where $\tau_h f(x) = f(x+h)$. Proposition 3.49 applies to f and implies that $\|\tau_h f - f\|_p \to 0$ as $h \to 0$; the continuity of $f * g$ follows. (4.23) can be derived immediately from (4.24).

Now, assume $r < \infty$ (so that $p, q < \infty$). We will get the conclusion in three steps.

1. Suppose $p = 1 = q$ (hence, $r = 1$). Then $|f| * |g| \in L^1(\mathbb{R}^N)$ and we have $\||f| * |g|\|_1 = \|f\|_1 \, \|g\|_1$.

   Indeed, according to Remark 4.17, $|f| * |g|$ is a Borel function and Tonelli's Theorem implies

$$\begin{aligned}\int_{\mathbb{R}^N} |f| * |g| \, dx &= \int_{\mathbb{R}^N} \left[\int_{\mathbb{R}^N} |f(x-y) g(y)| \, dy \right] dx \\ &= \int_{\mathbb{R}^N} |g(y)| \left[\int_{\mathbb{R}^N} |f(x-y)| \, dx \right] dy = \|f\|_1 \, \|g\|_1 .\end{aligned}$$

   Therefore the conclusion of step 1 follows.

2. We claim that, for every $f \in L^p(\mathbb{R}^N)$ and $g \in L^q(\mathbb{R}^N)$,

$$(|f| * |g|)^r(x) \le \|f\|_p^{r-p} \, \|g\|_q^{r-q} (|f|^p * |g|^q)(x) \qquad \forall x \in \mathbb{R}^N. \tag{4.25}$$

   Assume, first, $1 < p, q < \infty$ and let p' and q' be the conjugate exponents of p and q, respectively. Then

$$\frac{1}{p'} + \frac{1}{q'} + \frac{1}{r} = 2 - \frac{1}{p} - \frac{1}{q} + \frac{1}{r} = 1$$

   and

$$1 - \frac{p}{r} = p\left(1 - \frac{1}{q}\right) = \frac{p}{q'}, \quad 1 - \frac{q}{r} = q\left(1 - \frac{1}{p}\right) = \frac{q}{p'}.$$

   Using the above relations, for every $x, y \in \mathbb{R}^N$ we obtain

$$|f(x-y) g(y)| = \left(|f(x-y)|^p\right)^{1/q'} \left(|g(y)|^q\right)^{1/p'} \left(|f(x-y)|^p |g(y)|^q\right)^{1/r} .$$

Hence, applying the result of Exercise 3.7 with the exponents q', p', r,

$$(|f| * |g|)(x) \leq \|f\|_p^{p/q'} \ \|g\|_q^{q/p'} (|f|^p * |g|^q)^{1/r}(x) \qquad \forall x \in \mathbb{R}^N.$$

Raising the previous inequality to the rth power, (4.25) follows.
Inequality (4.25) is immediate for $p = 1 = q$. So let $p = 1$ and $1 < q < \infty$ (consequently, $r = q$). For every $x, y \in \mathbb{R}^N$ we have

$$|f(x-y)g(y)| = |f(x-y)|^{1/q'} \left(|f(x-y)||g(y)|^q\right)^{1/q}.$$

Thus, by Hölder's inequality we get

$$(|f| * |g|)(x) \leq \|f\|_1^{1/q'} \ (|f| * |g|^q)^{1/q}(x) \qquad \forall x \in \mathbb{R}^N.$$

Raising the previous inequality to the qth power we obtain (4.25).
The last case to study, namely $q = 1$ and $1 < p < \infty$, follows from the previous one since convolution is commutative.

3. Conclusion.

Owing to Remark 4.17, $|f| * |g|$ is a Borel function and

$$\int_{\mathbb{R}^N} (|f| * |g|)^r dx \underbrace{\leq \|f\|_p^{r-p} \ \|g\|_q^{r-q} \| \, |f|^p * |g|^q \, \|_1}_{\text{by (4.25)}} \underbrace{= \|f\|_p^r \ \|g\|_q^r}_{\text{by step 1}}. \tag{4.26}$$

Then $|f| * |g| \in L^r(\mathbb{R}^N)$, that is,

$$\int_{\mathbb{R}^N} \left(\int_{\mathbb{R}^N} |f(x-y)g(y)| \, dy \right)^r dx < \infty.$$

So the function $x \mapsto \left(\int_{\mathbb{R}^N} |f(x-y)g(y)| \, dy\right)^r$ is summable and, owing to Proposition 2.73(i), a.e. finite. It follows that $y \mapsto f(x-y)g(y)$ is summable for almost every $x \in \mathbb{R}^N$. Hence, $f * g$ is defined a.e. Using Remark 4.17 again, we obtain that $f^+ * g^+$, $f^- * g^-$, $f^+ * g^-$, $f^- * g^+$ are Borel functions; moreover, for every x such that (4.21) is integrable, we have

$$(f * g)(x) = (f^+ * g^+ + f^- * g^-)(x) - (f^+ * g^- + f^- * g^+)(x).$$

We deduce that $f * g$ is Borel and

$$\int_{\mathbb{R}^N} |f * g|^r dx \leq \int_{\mathbb{R}^N} (|f| * |g|)^r dx \underbrace{\leq \|f\|_p^r \ \|g\|_q^r}_{\text{by (4.26)}}$$

which completes the proof.

□

Remark 4.20 If $r = \infty$ and $1 < p, q < \infty$ in (4.22), then

$$\lim_{\|x\|\to\infty} (f * g)(x) = 0.$$

Indeed, for given $\varepsilon > 0$ let $R_\varepsilon > 0$ be such that

$$\int_{\|y\|\geq R_\varepsilon} |f(y)|^p \, dy < \varepsilon^p \quad \& \quad \int_{\|y\|\geq R_\varepsilon} |g(y)|^q \, dy < \varepsilon^q.$$

By Hölder's inequality we get

$$\begin{aligned}
|(f * g)(x)| &\leq \Big| \int_{\|y\|\geq R_\varepsilon} f(x-y)g(y)\, dy \Big| + \Big| \int_{\|y\|< R_\varepsilon} f(x-y)g(y)\, dy \Big| \\
&\leq \|f\|_p \Big(\int_{\|y\|\geq R_\varepsilon} |g(y)|^q \, dy \Big)^{1/q} + \|g\|_q \Big(\int_{\|x-z\|< R_\varepsilon} |f(z)|^p \, dz \Big)^{1/p}.
\end{aligned}$$

Therefore, for all $\|x\| \geq 2R_\varepsilon$, we have

$$|(f * g)(x)| \leq \varepsilon(\|f\|_p + \|g\|_q).$$

Remark 4.21 Observe that, when $q = 1$, Young's Theorem states that the convolution $f * g$ with a fixed $g \in L^1(\mathbb{R}^N)$ determines a transformation $f \mapsto f * g$ which maps functions in $L^p(\mathbb{R}^N)$ into the same $L^p(\mathbb{R}^N)$, and further

$$\|f * g\|_p \leq \|f\|_p \|g\|_1. \tag{4.27}$$

Remark 4.22 Taking $p = 1$ in Remark 4.21, we deduce that the operation of convolution

$$* : L^1(\mathbb{R}^N) \times L^1(\mathbb{R}^N) \to L^1(\mathbb{R}^N)$$

provides a multiplication structure for $L^1(\mathbb{R}^N)$. This operation is commutative (see Remark 4.18) and associative. Indeed, if $f, g, h \in L^1(\mathbb{R}^N)$, then, by the change of variable $z = t - y$ and Fubini's Theorem, we obtain

$$\begin{aligned}
((f * g) * h)(x) &= \int_{\mathbb{R}^N} (f * g)(x-y)h(y)\, dy \\
&= \int_{\mathbb{R}^N} h(y) \Big[\int_{\mathbb{R}^N} f(x-y-z)g(z)\, dz \Big] dy \\
&= \int_{\mathbb{R}^N} f(x-t) \Big[\int_{\mathbb{R}^N} g(t-y)h(y)\, dy \Big] dt \\
&= \int_{\mathbb{R}^N} f(x-t)(g * h)(t)\, dt = (f * (g * h))(x),
\end{aligned}$$

which proves associativity. Finally, it is clear that convolution obeys the distributive laws. However, there is no unit in $L^1(\mathbb{R}^N)$ under this multiplication. To see this, suppose there exists $g \in L^1(\mathbb{R}^N)$ such that $g * f = f$ for every $f \in L^1(\mathbb{R}^N)$. By the absolute continuity of the Lebesgue integral, there exists $\delta > 0$ such that

$$A \in \mathscr{B}(\mathbb{R}^N) \ \ \& \ \ m(A) \le \delta \Longrightarrow \int_A |g|\, dx < 1.$$

Let $\rho > 0$ be such that $m(\{\|y\| < \rho\}) < \delta$ and take $f = \chi_{\{\|y\|<\rho\}} \in L^1(\mathbb{R}^N)$. Then for every $x \in \mathbb{R}^N$ we have

$$\begin{aligned} |f(x)| = |(g * f)(x)| &\le \int_{\mathbb{R}^N} |g(x-y)|\,|f(y)|\, dy = \int_{\|y\|<\rho} |g(x-y)|\, dy \\ &= \int_{\|z-x\|<\rho} |g(z)|\, dz < 1, \end{aligned}$$

which contradicts the definition of f.

Exercise 4.23 Compute $f * g$ for $f(x) = \chi_{[-1,1]}(x)$ and $g(x) = e^{-|x|}$.

4.3.2 Approximation by Smooth Functions

Definition 4.24 A family $(\varphi_\varepsilon)_\varepsilon$ in $L^1(\mathbb{R}^N)$ is called an approximate identity if it satisfies the following:

$$\varphi_\varepsilon \ge 0, \quad \int_{\mathbb{R}^N} \varphi_\varepsilon(x)\, dx = 1 \quad \forall \varepsilon > 0, \tag{4.28}$$

$$\forall \delta > 0 : \quad \int_{\|x\|\ge\delta} \varphi_\varepsilon(x)\, dx \to 0 \ \text{ as } \varepsilon \to 0^+. \tag{4.29}$$

Remark 4.25 Properties (4.28) and (4.29) mean that taking smaller and smaller values of ε produces functions φ_ε with successively higher peaks concentrated in a smaller neighborhood of the origin.

Remark 4.26 A common way to produce approximate identities in $L^1(\mathbb{R}^N)$ is to take a function $\varphi \in L^1(\mathbb{R}^N)$ such that $\varphi \ge 0$ and $\int_{\mathbb{R}^N} \varphi(x)\, dx = 1$ and to define for $\varepsilon > 0$

$$\varphi_\varepsilon(x) = \varepsilon^{-N} \varphi(\varepsilon^{-1} x).$$

Conditions (4.28) and (4.29) are satisfied since, introducing the change of variable $y = \varepsilon^{-1} x$, we obtain

$$\int_{\mathbb{R}^N} \varphi_\varepsilon(x)\, dx = \int_{\mathbb{R}^N} \varphi(y)\, dy = 1$$

and, owing to Lebesgue's Dominated Convergence Theorem,

$$\int_{\|x\|\geq\delta} \varphi_\varepsilon(x)\,dx = \int_{\|y\|\geq\varepsilon^{-1}\delta} \varphi(y)\,dy \to 0 \text{ as } \varepsilon \to 0^+.$$

From property (4.29) one can guess that the effect of letting $\varepsilon \to 0$ in the formula $(f * \varphi_\varepsilon)(x) = \int f(x-y)\varphi_\varepsilon(y)\,dy$ will be to emphasize the values $f(x-y)$ corresponding to small $\|y\|$. Indeed, our next proposition shows that $f * \varphi_\varepsilon$ converges to f in various senses, if f is suitably chosen.

Proposition 4.27 *Let $(\varphi_\varepsilon)_\varepsilon \subset L^1(\mathbb{R}^N)$ be an approximate identity. Then the following holds:*

1. *If $f \in L^\infty(\mathbb{R}^N)$ is continuous in x_0, then $(f * \varphi_\varepsilon)(x_0) \to f(x_0)$ as $\varepsilon \to 0^+$.*
2. *If $f \in L^\infty(\mathbb{R}^N)$ is uniformly continuous, then $f * \varphi_\varepsilon \xrightarrow{L^\infty} f$ as $\varepsilon \to 0^+$.*
3. *If $1 \leq p < \infty$ and $f \in L^p(\mathbb{R}^N)$, then $f * \varphi_\varepsilon \xrightarrow{L^p} f$ as $\varepsilon \to 0^+$.*

Proof 1. By Young's Theorem we get that $f * \varphi_\varepsilon$ is continuous and $f * \varphi_\varepsilon \in L^\infty(\mathbb{R}^N)$. If f is continuous in x_0, then, given $\eta > 0$, there exists $\delta > 0$ such that

$$|f(x_0 - y) - f(x_0)| \leq \eta \quad \text{if } \|y\| \leq \delta. \tag{4.30}$$

Since $\int_{\mathbb{R}^N} \varphi_\varepsilon(y)\,dy = 1$, we have

$$\begin{aligned}
|(f * \varphi_\varepsilon)(x_0) - f(x_0)| &= \Big|\int_{\mathbb{R}^N} \big(f(x_0 - y) - f(x_0)\big)\varphi_\varepsilon(y)\,dy\Big| \\
&\leq \int_{\|y\|<\delta} \big|f(x_0 - y) - f(x_0)\big|\varphi_\varepsilon(y)\,dy \\
&\quad + \int_{\|y\|\geq\delta} \big|f(x_0 - y) - f(x_0)\big|\varphi_\varepsilon(y)\,dy \\
&\leq \eta \int_{\mathbb{R}^N} \varphi_\varepsilon(y)\,dy + 2\|f\|_\infty \int_{\|y\|\geq\delta} \varphi_\varepsilon(y)\,dy \\
&= \eta + 2\|f\|_\infty \int_{\|y\|\geq\delta} \varphi_\varepsilon(y)\,dy.
\end{aligned}$$

The conclusion follows from (4.29).

2. The proof is the same as in part 1, taking into account that now estimate (4.30) holds uniformly for $x_0 \in \mathbb{R}^N$.

3. According to Remark 4.21, we have $f * \varphi_\varepsilon \in L^p(\mathbb{R}^N)$ for all $\varepsilon > 0$. Since $\int_{\mathbb{R}^N} \varphi_\varepsilon(y)\,dy = 1$, for every $x \in \mathbb{R}^N$ we get

$$|(f * \varphi_\varepsilon)(x) - f(x)| \leq \int_{\mathbb{R}^N} |f(x-y) - f(x)|\varphi_\varepsilon(y)\,dy. \tag{4.31}$$

We claim that, for every $x \in \mathbb{R}^N$,

$$|(f * \varphi_\varepsilon)(x) - f(x)|^p \le \int_{\mathbb{R}^N} |f(x-y) - f(x)|^p \varphi_\varepsilon(y)\, dy. \tag{4.32}$$

(4.32) reduces to (4.31) when $p = 1$. If $1 < p < \infty$, by (4.31) we obtain

$$|(f * \varphi_\varepsilon)(x) - f(x)| \le \int_{\mathbb{R}^N} |f(x-y) - f(x)| (\varphi_\varepsilon(y))^{1/p} (\varphi_\varepsilon(y))^{1/p'}\, dy,$$

where $1/p + 1/p' = 1$. Applying Hölder's inequality and then raising both sides to the pth power, we conclude that

$$\begin{aligned} |(f * \varphi_\varepsilon)(x) - f(x)|^p &\le \left(\int_{\mathbb{R}^N} |f(x-y) - f(x)|^p \varphi_\varepsilon(y)\, dy \right) \left(\int_{\mathbb{R}^N} \varphi_\varepsilon(y)\, dy \right)^{p/p'} \\ &= \int_{\mathbb{R}^N} |f(x-y) - f(x)|^p \varphi_\varepsilon(y)\, dy. \end{aligned}$$

Hence, (4.32) holds for $1 \le p < \infty$. Then, taking the integral over $\mathbb{R}^N$ and changing the order of integration thanks to Tonelli's Theorem, we have

$$\|f * \varphi_\varepsilon - f\|_p^p \le \int_{\mathbb{R}^N} \|\tau_{-y} f - f\|_p^p \, \varphi_\varepsilon(y)\, dy$$

where $\tau_{-y} f(x) = f(x-y)$. Let us set $\Delta(y) = \|\tau_{-y} f - f\|_p^p$; the above inequality becomes

$$\|f * \varphi_\varepsilon - f\|_p^p \le (\Delta * \varphi_\varepsilon)(0).$$

By Proposition 3.49, Δ is continuous. Since $\Delta(y) \le 2^p \|f\|_p^p$, we have that $\Delta \in L^\infty(\mathbb{R}^N)$. The desired convergence follows noting that, by the first part of the proof, $(\Delta * \varphi_\varepsilon)(0) \to \Delta(0) = 0$. □

Before stating our next result, let us introduce some notation.

Let $\Omega \subset \mathbb{R}^N$ be an open set. $\mathscr{C}^0(\Omega) = \mathscr{C}(\Omega)$ is the space of continuous functions $f : \Omega \to \mathbb{R}$. For $k \in \mathbb{N}$, $\mathscr{C}^k(\Omega)$ denotes the space of all functions $f : \Omega \to \mathbb{R}$ which are k times continuously differentiable. Moreover, we set[9]

$$\mathscr{C}^\infty(\Omega) = \cap_{k=0}^\infty \mathscr{C}^k(\Omega),$$

$$\mathscr{C}_c^k(\Omega) = \mathscr{C}^k(\Omega) \cap \mathscr{C}_c(\Omega), \quad \mathscr{C}_c^\infty(\Omega) = \mathscr{C}^\infty(\Omega) \cap \mathscr{C}_c(\Omega).$$

[9] See Sect. 3.4.2 for the definition of $\mathscr{C}_c(\Omega)$.

If $f \in \mathscr{C}^k(\Omega)$ and $\alpha = (\alpha_1, \dots, \alpha_N)$ is a multi-index such that $|\alpha| := \alpha_1 + \cdots + \alpha_N \le k$, then we set

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_N^{\alpha_N}}.$$

If $\alpha = (0, \dots, 0)$, we set $D^0 f = f$.

Proposition 4.28 *Let $f \in L^1(\mathbb{R}^N)$ and let $g \in \mathscr{C}^k(\mathbb{R}^N)$ be such that $D^\alpha g$ belongs to $L^\infty(\mathbb{R}^N)$ for all $|\alpha| \le k$. Then $f * g \in \mathscr{C}^k(\mathbb{R}^N)$ and*

$$D^\alpha(f * g) = f * D^\alpha g \quad \textit{if } |\alpha| \le k.$$

Proof The continuity of $f * g$ follows from Young's Theorem. Let us show the conclusion when $k = 1$; the proof can easily be completed by an induction argument. Setting

$$F(x, y) = f(y)g(x - y),$$

we have

$$\Big|\frac{\partial F}{\partial x_i}(x, y)\Big| = |f(y)\frac{\partial g}{\partial x_i}(x - y)| \le \Big\|\frac{\partial g}{\partial x_i}\Big\|_\infty |f(y)|.$$

Since $(f * g)(x) = \int_{\mathbb{R}^N} F(x, y)dy$, Proposition 2.106 implies that $f * g$ is differentiable and

$$\frac{\partial(f * g)}{\partial x_i}(x) = \int_{\mathbb{R}^N} f(y)\frac{\partial g}{\partial x_i}(x - y)dy = \Big(f * \frac{\partial g}{\partial x_i}\Big)(x).$$

By hypothesis $\frac{\partial g}{\partial x_i} \in \mathscr{C}(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$ and so $f * \frac{\partial g}{\partial x_i} \in \mathscr{C}(\mathbb{R}^N)$ again by Young's Theorem. Hence, $f * g \in \mathscr{C}^1(\mathbb{R}^N)$. □

Thus, convolution with a smooth function produces a smooth function. This fact offers a powerful technique to prove a variety of density theorems.

Definition 4.29 For every $\varepsilon > 0$ let $\rho_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ be defined by

$$\rho_\varepsilon(x) = \begin{cases} C\varepsilon^{-N} \exp\Big(\dfrac{\varepsilon^2}{\|x\|^2 - \varepsilon^2}\Big) & \text{if } \|x\| < \varepsilon, \\ 0 & \text{if } \|x\| \ge \varepsilon, \end{cases}$$

where $\frac{1}{C} = \int_{\|x\|<1} \exp\big(\frac{1}{\|x\|^2-1}\big)dx$. The family $(\rho_\varepsilon)_\varepsilon$ is called the *standard mollifier*.

Lemma 4.30 *The standard mollifier $(\rho_\varepsilon)_\varepsilon$ satisfies*

$$\rho_\varepsilon \in \mathscr{C}_c^\infty(\mathbb{R}^N), \qquad \operatorname{supp}(\rho_\varepsilon) = \{\|x\| \le \varepsilon\} \quad \forall \varepsilon > 0;$$

$$(\rho_\varepsilon)_\varepsilon \textit{ is an approximate identity.}$$

Proof Let $f : \mathbb{R} \to \mathbb{R}$ be defined by:

$$f(t) = \begin{cases} \exp\left(\dfrac{1}{t-1}\right) & \text{if } t < 1, \\ 0 & \text{if } t \geq 1. \end{cases}$$

Then f is a $\mathscr{C}^\infty$ function. Indeed we only need to check the smoothness at $t = 1$. As $t \downarrow 1$ all the derivatives are zero. As $t \uparrow 1$ the derivatives are finite linear combinations of terms of the form $\frac{1}{(t-1)^l} \exp\left(\frac{1}{t-1}\right)$, l being an integer greater than or equal to zero, and these terms tend to zero as $t \uparrow 1$.

Observe that, for every $\varepsilon > 0$,

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^N} \rho_1\left(\frac{x}{\varepsilon}\right) = C \frac{1}{\varepsilon^N} f\left(\frac{\|x\|^2}{\varepsilon^2}\right).$$

Then $\rho_\varepsilon \in \mathscr{C}_c^\infty(\mathbb{R}^N)$ and $\operatorname{supp}(\rho_\varepsilon) = \{\|x\| \leq \varepsilon\}$. The definition of C implies $\int_{\mathbb{R}^N} \rho_1(x) dx = 1$. Remark 4.26 allows us to conclude. □

Lemma 4.31 *Let $f, g \in \mathscr{C}_c(\mathbb{R}^N)$. Then $f * g \in \mathscr{C}_c(\mathbb{R}^N)$ and*

$$\operatorname{supp}(f * g) \subset \operatorname{supp}(f) + \operatorname{supp}(g),$$

where the sum of two sets A and B in $\mathbb{R}^N$ is defined by:

$$A + B = \{x + y \,|\, x \in A,\ y \in B\}.$$

Proof By Proposition 4.28 we get $f * g \in \mathscr{C}(\mathbb{R}^N)$. Let $A = \operatorname{supp}(f)$ and $B = \operatorname{supp}(g)$. For every $x \in \mathbb{R}^N$ we have

$$(f * g)(x) = \int_{(x - \operatorname{supp}(f)) \cap \operatorname{supp}(g)} f(x - y) g(y)\, dy.$$

If $x \in \mathbb{R}^N$ is such that $(f * g)(x) \neq 0$, then $(x - \operatorname{supp}(f)) \cap \operatorname{supp}(g) \neq \emptyset$, that is, $x \in \operatorname{supp}(f) + \operatorname{supp}(g)$. □

Proposition 4.32 *Let $\Omega \subset \mathbb{R}^N$ be an open set. Then*[10]

- *$\mathscr{C}_c^\infty(\Omega)$ is dense in $\mathscr{C}_0(\Omega)$.*
- *$\mathscr{C}_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for every $1 \leq p < \infty$.*

[10] See Exercise 3.46 for the definition of $\mathscr{C}_0(\Omega)$.

Proof In view of Theorem 3.45 and Exercise 3.46 it is sufficient to prove that, given $f \in \mathscr{C}_c(\Omega)$, there exists a sequence $(f_n)_n \subset \mathscr{C}_c^\infty(\Omega)$ such that $f_n \xrightarrow{L^\infty} f$ and $f_n \xrightarrow{L^p} f$. To this aim, set

$$\tilde{f} = \begin{cases} f(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in \mathbb{R}^N \setminus \Omega. \end{cases}$$

Then $\tilde{f} \in \mathscr{C}_c(\mathbb{R}^N)$. Let $(\rho_\varepsilon)_\varepsilon$ be the standard mollifier and, for every n, define $f_n := f * \rho_{1/n}$. By Proposition 4.28 and Lemma 4.31 $f_n \in \mathscr{C}_c^\infty(\mathbb{R}^N)$. Next, let $K = \text{supp}(f)$ and[11] $\eta = \inf_{x \in K} d_{\partial\Omega}(x) > 0$. Then

$$\tilde{K} := \{x \in \mathbb{R}^N \mid d_K(x) \leq \frac{\eta}{2}\}$$

is a compact subset of Ω. By Lemma 4.31, if n is such that $1/n < \eta/2$, we have

$$\text{supp}(f_n) \subset K + \left\{ \|x\| \leq \frac{1}{n} \right\} = \left\{ x \in \mathbb{R}^N \,\middle|\, d_K(x) \leq \frac{1}{n} \right\} \subset \tilde{K}.$$

So $f_n \in \mathscr{C}_c^\infty(\Omega)$ for n sufficiently large. Since $\tilde{f}$ is uniformly continuous, Proposition 4.27(2) ensures that $f_n \to \tilde{f}$ in $L^\infty(\mathbb{R}^N)$, by which

$$f_n \to f \text{ in } L^\infty(\Omega). \tag{4.33}$$

Finally, for n large enough,

$$\int_\Omega |f_n - f|^p \, dx = \int_{\tilde{K}} |f_n - f|^p \, dx \leq m(\tilde{K}) \|f_n - f\|_\infty^p.$$

The conclusion follows recalling (4.33). □

An interesting consequence of smoothing properties of convolution is the following *Weierstrass Approximation Theorem.*[12]

Theorem 4.33 (Weierstrass) *Let $f \in \mathscr{C}_c(\mathbb{R}^N)$. Then there exists a sequence of polynomials $(P_n)_n$ such that $P_n \to f$ uniformly on all compact subsets of $\mathbb{R}^N$.*

Proof For every $\varepsilon > 0$ define

$$\varphi_\varepsilon(x) = (\varepsilon\sqrt{\pi})^{-N} \exp(-\|\varepsilon^{-1}x\|^2), \quad x \in \mathbb{R}^N.$$

[11] $d_{\partial\Omega}(x)$ denotes the distance between the set $\partial\Omega$ and the point x (see Appendix A).

[12] Weierstrass' Theorem is a particular case of a more general approximation result known as Stone-Weierstrass Theorem (see [Fo99]).

The well-known *Poisson formula*

$$\int_{\mathbb{R}^N} \exp(-\|x\|^2)dx = \pi^{N/2}$$

and Remark 4.26 imply that $(\varphi_\varepsilon)_\varepsilon$ is an approximate identity. Proposition 4.27 yields

$$\varphi_\varepsilon * f \xrightarrow{L^\infty} f \quad \text{as } \varepsilon \to 0. \tag{4.34}$$

Thus it is sufficient to show that, given $\varepsilon > 0$, there exists a sequence of polynomials $(P_n)_n$ such that

$$P_n \to \varphi_\varepsilon * f \text{ uniformly on all compact sets.} \tag{4.35}$$

To see this, observe that φ_ε is an analytic function, and so it can be approximated uniformly on any compact set by the partial sums $(Q_n)_n$ of its Taylor series which are, of course, polynomials. Next, set

$$P_n(x) = \int_{\mathbb{R}^N} Q_n(x-y) f(y)\, dy. \tag{4.36}$$

Since f is compactly supported, then the integrand in (4.36) is bounded by $|f| \sup_{y \in \text{supp}(f)} |Q_n(x-y)|$, which is summable for every $x \in \mathbb{R}^N$. Then P_n is well defined on $\mathbb{R}^N$. Moreover $Q_n(x-y)$ is a polynomial in the variables (x, y) and can be represented by a sum of the form $\sum_{k=1}^{K_n} s_k(x)t_k(y)$ with s_k, t_k polynomials in $\mathbb{R}^N$. Substituting in (4.36), we deduce that each P_n is also a polynomial. Let now $K \subset \mathbb{R}^N$ be a compact set. Then $\tilde{K} := K - \text{supp}(f)$ is also compact, and so $Q_n \to \varphi_\varepsilon$ uniformly in $\tilde{K}$. Hence, for every $x \in K$,

$$\begin{aligned} |P_n(x) - (\varphi_\varepsilon * f)(x)| &\le \int_{\text{supp}(f)} |Q_n(x-y) - \varphi_\varepsilon(x-y)||f(y)|dy \\ &\le \sup_{z \in \tilde{K}} |Q_n(z) - \varphi_\varepsilon(z)| \int_{\mathbb{R}^N} |f(y)|dy \end{aligned}$$

and (4.35) follows. □

Corollary 4.34 *Let $A \in \mathscr{B}(\mathbb{R}^N)$ be a bounded set and let $1 \le p < \infty$. Then the set $\mathscr{P}_A$ of all polynomials defined on A is dense in $L^p(A)$.*

Proof Consider $f \in L^p(A)$ and let $\tilde{f}$ be the extension of f by zero outside A. Then $\tilde{f} \in L^p(\mathbb{R}^N)$; given $\varepsilon > 0$, Proposition 4.32 implies the existence of $g \in \mathscr{C}_c(\mathbb{R}^N)$ such that

$$\int_A |f - g|^p\, dx \le \int_{\mathbb{R}^N} |\tilde{f} - g|^p\, dx \le \varepsilon^p.$$

Since $\bar{A}$ is a compact set, by Theorem 4.33 there exists a polynomial P such that $\sup_{x\in\bar{A}} |g(x) - P(x)| \le \varepsilon$. Then

$$\int_A |g - P|^p \, dx \le \sup_{x\in\bar{A}} |g(x) - P(x)|^p \, m(A) \le \varepsilon^p m(A).$$

So

$$\begin{aligned}\left(\int_A |f - P|^p \, dx\right)^{1/p} &\le \left(\int_A |f - g|^p \, dx\right)^{1/p} + \left(\int_A |g - P|^p \, dx\right)^{1/p} \\ &\le \varepsilon + \varepsilon (m(A))^{1/p}\end{aligned}$$

and the proof is complete. □

Remark 4.35 By Corollary 4.34 we deduce that if $A \in \mathscr{B}(\mathbb{R}^N)$ is bounded, then the set of all polynomials defined on A with rational coefficients is a countable dense subset of $L^p(A)$ for all $1 \le p < \infty$ (see Proposition 3.47).

References

[Fo99] Folland, G.B.: Real Analysis. Wiley, New York (1999)
[Ru73] Rudin, W.: Functional Analysis. McGraw Hill, New York (1973)

Part II
Functional Analysis

Chapter 5
Hilbert Spaces

With this chapter we begin the study of functional analysis, which represents the second main topic of this book. Just like in the first part of the book we have shown how to extend to an abstract environment fundamental analytical notions such as the integral of a real function, we now intend to explain how to generalize basic concepts from geometry and linear algebra to vector spaces with certain additional structures. We shall first examine Hilbert spaces, where the notion of orthogonal vectors can be defined thanks to the presence of a scalar product. In the next chapter, our analysis will move to the more general class of Banach spaces, where orthogonality no longer makes sense. One could go even further and consider topological vector spaces, but such a level of generality would exceed the scopes of this monograph.

Soon after giving the first definitions, we will set and solve the problem of finding the orthogonal projection of a point onto a closed convex set and, in particular, a closed subspace. Then, we shall study the space of all continuous linear functionals on a Hilbert space. Finally, we will investigate the possibility of representing any element of the space by its Fourier series, that is, as a linear combination of a countable set of orthogonal vectors. All these classical topics are treated in most introductory textbooks such as [Br83, Co90, Ko02, Ru73, Ru74, Yo65]. The reader is also referred to the above references for further developments of the theory of Hilbert spaces, such as the spectral theorem for compact self-adjoint operators, as well as other topics we will not even be able to mention.

Throughout this chapter, we will denote by H a real vector space. The theory of complex spaces is similar but requires adjustments that make notation somewhat heavier. In some of the above references, the reader will find adaptations of the results of this chapter to complex Hilbert spaces.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_5

5.1 Definitions and Examples

Let H be a vector space over $\mathbb{R}$.

Definition 5.1 A *scalar product* $\langle \cdot, \cdot \rangle$ on H is a mapping

$$\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$$

with the following properties:

1. $\langle x, x \rangle \geq 0$ for every $x \in H$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.
2. $\langle x, y \rangle = \langle y, x \rangle$ for every $x, y \in H$.
3. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$ for every $x, y, z \in H$ and $\alpha, \beta \in \mathbb{R}$.

A linear space H endowed with a scalar product is called a *pre-Hilbert space*.

Remark 5.2 Since, for any $y \in H$, $0\, y = 0$, we have

$$\langle x, 0 \rangle = 0\, \langle x, y \rangle = 0 \qquad \forall x \in H.$$

In a pre-Hilbert space $(H, \langle \cdot, \cdot \rangle)$, let us consider the function $\| \cdot \| : H \to \mathbb{R}$ defined by

$$\|x\| = \sqrt{\langle x, x \rangle} \qquad \forall x \in H. \tag{5.1}$$

The following fundamental inequality holds.

Proposition 5.3 (Cauchy–Schwarz) *Let $(H, \langle \cdot, \cdot \rangle)$ be a pre-Hilbert space. Then*

$$|\langle x, y \rangle| \leq \|x\|\, \|y\| \qquad \forall x, y \in H. \tag{5.2}$$

Moreover, equality holds if and only if x and y are linearly dependent.

Proof The conclusion is trivial if $y = 0$. So suppose $y \neq 0$. Then

$$0 \leq \left\| x - \frac{\langle x, y \rangle}{\|y\|^2} y \right\|^2 = \|x\|^2 - \frac{\langle x, y \rangle^2}{\|y\|^2}, \tag{5.3}$$

which implies (5.2). If x and y are linearly dependent, then it is clear that $|\langle x, y \rangle| = \|x\|\, \|y\|$. Conversely, if $\langle x, y \rangle = \pm \|x\|\, \|y\|$ and $y \neq 0$, then (5.3) yields

$$\left\| x - \frac{\langle x, y \rangle}{\|y\|^2} y \right\| = 0,$$

which implies x and y are linearly dependent. □

Exercise 5.4 Define

$$F(\lambda) = \|x + \lambda y\|^2 = \lambda^2 \|y\|^2 + 2\lambda\langle x, y\rangle + \|x\|^2 \qquad \forall \lambda \in \mathbb{R}.$$

Observing that $F(\lambda) \geq 0$ for every $\lambda \in \mathbb{R}$, give an alternative proof of (5.2).

Corollary 5.5 *Let $(H, \langle\cdot,\cdot\rangle)$ be a pre-Hilbert space. Then H is a normed linear space*[1] *with the norm defined by (5.1).*

Norm (5.1) is called the *norm induced by the scalar product* $\langle\cdot,\cdot\rangle$.

Proof It is sufficient to prove the triangle inequality, since all other properties easily follow from Definition 5.1. For any $x, y \in H$, we have

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2 \end{aligned}$$

by the Cauchy–Schwarz inequality. The conclusion follows. □

Remark 5.6 In a pre-Hilbert space $(H, \langle\cdot,\cdot\rangle)$ the function

$$d(x, y) = \|x - y\| \qquad \forall x, y \in H \tag{5.4}$$

is a metric, called the *metric associated* with the scalar product of H.

From now on we will often use the following notation: given H a pre-Hilbert space and $(x_n)_n \subset H$, $x \in H$, we will write

$$x_n \xrightarrow{H} x$$

to mean that $(x_n)_n$ converges to x in the metric (5.4), that is, $\|x_n - x\| \to 0$ (as $n \to \infty$).

Definition 5.7 A pre-Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is called a *Hilbert space* if it is complete with respect to the metric defined in (5.4).

Example 5.8 $\mathbb{R}^N$ is a Hilbert space with the scalar product

$$\langle x, y\rangle = \sum_{k=1}^{N} x_k y_k,$$

where $x = (x_1, \ldots, x_N)$, $y = (y_1, \ldots, y_N) \in \mathbb{R}^N$.

[1] See Definition 6.1.

Example 5.9 Let $(X, \mathscr{E}, \mu)$ be a measure space. Then $L^2(X, \mu)$, endowed with the scalar product

$$\langle f, g\rangle = \int_X f g \, d\mu, \quad f, \, g \in L^2(X, \mu),$$

is a Hilbert space (completeness follows from Proposition 3.11).

Example 5.10 Let ℓ^2 be the space of all sequences of real numbers $x = (x_k)_k$ such that[2]

$$\sum_{k=1}^{\infty} |x_k|^2 < \infty.$$

ℓ^2 is a linear space with the usual operations

$$a(x_k)_k = (a x_k)_k, \quad (x_k)_k + (y_k)_k = (x_k + y_k)_k, \quad a \in \mathbb{R}, \; (x_k)_k, (y_k)_k \in \ell^2.$$

The space ℓ^2, endowed with the scalar product

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k y_k, \quad x = (x_k)_k, \; y = (y_k)_k \in \ell^2,$$

is a Hilbert space. This is a special case of the above example, by taking $X = \mathbb{N}$ with the counting measure $\mu^{\#}$.

Exercise 5.11 Show that ℓ^2 is complete arguing as follows. Take a Cauchy sequence x^n, $n = 1, 2, \ldots$, in ℓ^2, and set $x^n = \left(x_k^n\right)_k$ for every $n \in \mathbb{N}$.

1. Show that, for every $k \in \mathbb{N}$, $\left(x_k^n\right)_n$ is a Cauchy sequence in $\mathbb{R}$, and deduce that the limit $x_k := \lim_{n\to\infty} x_k^n$ does exist.
2. Show that $x := (x_k)_k \in \ell^2$.
3. Show that $x^n \xrightarrow{\ell^2} x$ as $n \to \infty$.

Exercise 5.12 Let $H = \mathscr{C}([-1, 1])$ be the linear space of all continuous functions $f : [-1, 1] \to \mathbb{R}$. Show that:

1. H is a pre-Hilbert space with the scalar product

$$\langle f, g\rangle = \int_{-1}^{1} f(t) g(t) dt.$$

2. H is not a Hilbert space.
   Hint. Consider

[2]See Example 3.2.

$$f_n(t) = \begin{cases} 1 & \text{if } \ t \in \left[\frac{1}{n}, 1\right], \\ nt & \text{if } \ t \in \left(-\frac{1}{n}, \frac{1}{n}\right), \\ -1 & \text{if } \ t \in \left[-1, -\frac{1}{n}\right], \end{cases}$$

and show that $(f_n)_n$ is a Cauchy sequence in H. Observe that if $f_n \xrightarrow{H} f$, then

$$f(t) = \begin{cases} 1 & \text{if } \ t \in (0, 1], \\ -1 & \text{if } \ t \in [-1, 0). \end{cases}$$

Remark 5.13 Let $(H, \langle\cdot,\cdot\rangle)$ be a pre-Hilbert space. Then the scalar product $\langle\cdot,\cdot\rangle$ is itself expressible in terms of its associated norm:

$$\langle x, y\rangle = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big) \qquad \forall x, y \in H.$$

This is known as the *polarization identity*. Its validity is readily verified by direct simplification of the expression on the right hand side, using the properties of the scalar product. Similarly, one can prove the following *parallelogram identity*:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad \forall x, y \in H. \tag{5.5}$$

It can be shown that the parallelogram identity characterizes the norms associated with a scalar product. More precisely, one can prove that any norm satisfying (5.5) must be induced by a scalar product, as stated by the result of the following exercise (see also [Da73]).

Exercise 5.14 Let $\|\cdot\|$ be a norm on a linear space H verifying (5.5) and set

$$\langle x, y\rangle = \frac{1}{2}\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big) \qquad \forall x, y \in H.$$

Show that $\langle\cdot,\cdot\rangle$ is a scalar product on H inducing the norm $\|\cdot\|$.
Hint. Properties 1 and 2 of Definition 5.1 are clearly satisfied. Prove the validity of property 3 by arguing as follows. By using (5.5), show that

1. $\langle -x, y\rangle = -\langle x, y\rangle$ for every $x, y \in H$.
2. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ for every $x, y, z \in H$.

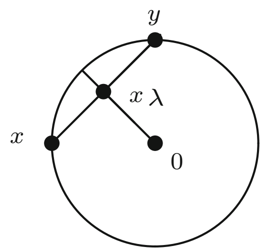
By step 2 deduce that

3. $\langle nx, y\rangle = n\langle x, y\rangle$ for every $x, y \in H$ and $n \in \mathbb{N}$,

and, consequently, using also step 1.

4. $\langle px, y\rangle = p\langle x, y\rangle$ for every $x, y \in H$ and $p \in \mathbb{Q}$.

Fig. 5.1 Uniform convexity



Finally, observe that, by (5.5),

$$\langle x, y\rangle - \langle x', y\rangle = \langle x - x', y\rangle = \frac{1}{2}\big(\|x - x' + y\|^2 - \|y\|^2\big) - \frac{1}{2}\|x - x'\|^2$$

and derive the continuity of the map $x \in H \mapsto \langle x, y\rangle \in \mathbb{R}$ from the continuity of $x \in H \mapsto \|x\| \in \mathbb{R}$ (which is a consequence of the inequality $|\|x\| - \|y\|| \le \|x - y\|$). So, approximating $\alpha \in \mathbb{R}$ by $p_n \in \mathbb{Q}$, by step 4 conclude that $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for every $x, y \in H$ and $\alpha \in \mathbb{R}$.

Exercise 5.15 Show that[3] $L^p(0,1)$ fails to be a Hilbert space for $p \neq 2$.

Hint. In view of the result of the previous exercise, it is sufficient to prove that identity (5.5) is not satisfied by taking the pair $f = \chi_{(0,\frac{1}{2})}$ and $g = \chi_{(\frac{1}{2},1)}$.

Exercise 5.16 Let H be a pre-Hilbert space and let $x, y \in H$ be linearly independent vectors such that $\|x\| = \|y\| = 1$. Show that

$$\|\lambda x + (1-\lambda)y\| < 1 \qquad \forall \lambda \in (0,1).$$

Hint. Observe that

$$\|\underbrace{\lambda x + (1-\lambda)y}_{x_\lambda}\|^2 = 1 + 2\lambda(1-\lambda)\big(\langle x, y\rangle - 1\big) \tag{5.6}$$

and use Cauchy–Schwarz inequality (see Fig. 5.1). Identity (5.6), written in the form $\|\lambda x + (1-\lambda)y\|^2 = 1 - \lambda(1-\lambda)\|x - y\|^2$, implies that any pre-Hilbert space is *uniformly convex*, see [Ko02].

5.2 Orthogonal Projection

Definition 5.17 Two vectors x and y of a pre-Hilbert space H are said to be *orthogonal* if $\langle x, y\rangle = 0$. In this case, we write $x \perp y$. Two subsets A, B of H are said to be *orthogonal* ($A \perp B$) if $x \perp y$ for every $x \in A$ and $y \in B$.

---

[3] $L^2(0,1) = L^2([0,1], m)$ where m is the Lebesgue measure on $[0,1]$. See footnote 7, p. 87.

The following is the *Pythagorean Theorem* in pre-Hilbert spaces.

Proposition 5.18 *If $x_1, \dots, x_n$ are pairwise orthogonal vectors in a pre-Hilbert space H, then*

$$\|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2.$$

Exercise 5.19 Prove Proposition 5.18.

Exercise 5.20 Show that if $x_1, \dots, x_n$ are pairwise orthogonal vectors in a pre-Hilbert space H, then $x_1, \dots, x_n$ are linearly independent.

5.2.1 Projection onto a Closed Convex Set

Definition 5.21 Given a pre-Hilbert space H, a set $K \subset H$ is said to be *convex* if, for any $x, y \in K$,

$$[x, y] := \{\lambda x + (1 - \lambda)y \mid \lambda \in [0, 1]\} \subset K.$$

Any subspace of H is convex, for instance. Similarly, for any $x_0 \in H$ and $r > 0$ the ball

$$B_r(x_0) = \big\{x \in H \mid \|x - x_0\| < r\big\} \tag{5.7}$$

is a convex set.

Exercise 5.22 Show that, if $(K_i)_{i \in I}$ is a family of convex subsets of a pre-Hilbert space H, then $\cap_{i \in I} K_i$ is also convex.

It is well known that, in a finite-dimensional space, a point x has a nonempty projection onto a nonempty closed set (see Proposition A.2). The following result extends such a property to convex subsets of a Hilbert space.

Theorem 5.23 *Let H be a Hilbert space and let $K \subset H$ be a nonempty closed convex set. Then for any $x \in H$ there is a unique element $y_x = p_K(x)$ in K, called the* orthogonal projection *of x onto K, such that*

$$\|x - y_x\| = \inf_{y \in K} \|x - y\|. \tag{5.8}$$

Moreover, y_x is the unique solution of the problem (see Fig. 5.2)

$$\begin{cases} y \in K, \\ \langle x - y, z - y\rangle \le 0 \qquad \forall z \in K. \end{cases} \tag{5.9}$$

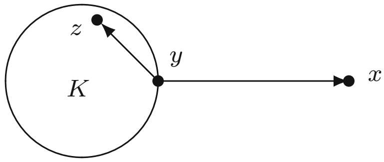


Fig. 5.2 Inequality (5.9) has a simple geometric meaning

Proof Let $d = \inf_{y\in K} \|x - y\|$. We shall split the proof into 4 steps.

1. Let $(y_n)_n \subset K$ be a minimizing sequence, that is,

$$\|x - y_n\| \to d \quad \text{as} \quad n \to \infty.$$

   Then $(y_n)_n$ is a Cauchy sequence. Indeed, for any $m, n \in \mathbb{N}$, parallelogram identity (5.5) yields

$$\|(x - y_n) + (x - y_m)\|^2 + \|(x - y_n) - (x - y_m)\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

   Since K is convex, we have that $\frac{y_n + y_m}{2} \in K$, and so

$$\begin{aligned}\|y_n - y_m\|^2 &= 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4\left\|x - \frac{y_n + y_m}{2}\right\|^2 \\ &\le 2\|x - y_n\|^2 + 2\|x - y_m\|^2 - 4d^2.\end{aligned}$$

   Hence $\|y_n - y_m\| \to 0$ as $m, n \to \infty$, as claimed.
2. Since H is complete and K is closed, $(y_n)_n$ converges to a point $y_x \in K$ satisfying $\|x - y_x\| = d$. The existence of y_x is thus proved.
3. We now proceed to show that (5.9) holds for any point $y \in K$ at which the infimum (5.8) is attained. Let $z \in K$ and let $\lambda \in (0, 1]$. Since $\lambda z + (1 - \lambda)y \in K$, we have that $\|x - y\| \le \|x - y - \lambda(z - y)\|$. So

$$0 \ge \frac{1}{\lambda}\left[\|x - y\|^2 - \|x - y - \lambda(z - y)\|^2\right] = 2\langle x - y, z - y\rangle - \lambda\|z - y\|^2.$$

   Taking the limit as $\lambda \downarrow 0$ we deduce (5.9).
4. We will complete the proof showing that problem (5.9) has at most one solution. Let y be another solution of problem (5.9). Then

$$\langle x - y_x, y - y_x\rangle \le 0 \quad \text{and} \quad \langle x - y, y_x - y\rangle \le 0.$$

   The above inequalities imply that $\|y - y_x\|^2 \le 0$, and so $y = y_x$.

□

Exercise 5.24 In the Hilbert space $H = L^2(0, 1)$, consider the set

$$H_+ = \{f \in H \mid f \ge 0 \text{ a.e.}\}.$$

1. Show that H_+ is a closed convex subset of H.
2. Given $f \in H$, show that $p_{H_+}(f) = f^+$, where $f^+ = \max\{f, 0\}$ is the positive part of f.

Example 5.25 In an infinite-dimensional Hilbert space the projection of a point onto a nonempty closed set may be empty (in absence of convexity). To see this, let Q be the set consisting of all sequences $x^n = \big(x_k^n\big)_k \in \ell^2$ defined by

$$x_k^n = \begin{cases} 0 & \text{if} \quad k \neq n, \\ 1 + \frac{1}{n} & \text{if} \quad k = n. \end{cases}$$

Then Q is closed. Indeed, since

$$n \neq m \quad \Longrightarrow \quad \|x^n - x^m\| > \sqrt{2},$$

Q has no limit points in ℓ^2. On the other hand, Q has no element of minimal norm (i.e., 0 has no projection onto Q), since

$$\inf_{n \geq 1} \|x^n\| = \inf_{n \geq 1} \left(1 + \frac{1}{n}\right) = 1,$$

but $\|x^n\| > 1$ for every $n \geq 1$.

Exercise 5.26 Let H be a Hilbert space and $K \subset H$ a nonempty closed convex set. Show that

$$\langle x - y, p_K(x) - p_K(y)\rangle \geq \|p_K(x) - p_K(y)\|^2 \qquad \forall x, y \in H.$$

Hint. Apply (5.9) to $z = p_K(x)$ and $z = p_K(y)$.

5.2.2 Projection onto a Closed Subspace

Theorem 5.23 applies, in particular, to subspaces of H. In this case, however, the variational inequality in (5.9) takes a special form.

Corollary 5.27 *Let M be a nonempty closed subspace of a Hilbert space H. Then, for every $x \in H$, $p_M(x)$ is the unique solution of problem*

$$\begin{cases} y \in M, \\ \langle x - y, v\rangle = 0 \qquad \forall v \in M. \end{cases} \tag{5.10}$$

Proof It suffices to show that problems (5.9) and (5.10) are equivalent when M is a subspace. If y is a solution of (5.10), then (5.9) follows taking $v = z - y$. Conversely, suppose that y satisfies (5.9). Then, taking $z = y + \lambda v$ with $\lambda \in \mathbb{R}$ and $v \in M$, we obtain

$$\lambda\langle x - y, v\rangle \leq 0 \qquad \forall \lambda \in \mathbb{R}.$$

Since λ is any real number, necessarily $\langle x - y, v\rangle = 0$. □

Exercise 5.28 Let H be a Hilbert space.

1. It is well known that any finite-dimensional subspace of H is closed (see Appendix C). Give an example to show that this fails, in general, for infinite-dimensional subspaces.

   Hint. Consider the set of all sequences $x = (x_k)_k \in \ell^2$ such that $x_k = 0$ except for a finite number of indices k, and show that this is a dense subspace of ℓ^2.
2. Show that, if M is a closed subspace of H and $M \neq H$, then there exists $x_0 \in H \setminus \{0\}$ such that $\langle x_0, y\rangle = 0$ for every $y \in M$.
3. Let L be a subspace of H. Show that $\overline{L}$ is also a subspace of H.
4. For any $A \subset H$ let us set

$$A^\perp = \{x \in H \mid x \perp A\}. \tag{5.11}$$

   Show that if $A, B \subset H$, then

   a. $A^\perp$ is a closed subspace of H and $\bar{A}^\perp = A^\perp$.
   b. $A \subset B \implies B^\perp \subset A^\perp$.
   c. $(A \cup B)^\perp = A^\perp \cap B^\perp$.

   $A^\perp$ is called the *orthogonal complement* of A in H.

Proposition 5.29 *Let M be a nonempty closed subspace of a Hilbert space H. Then the following statements hold (see Fig. 5.3):*

(i) For every $x \in H$ there exists a unique pair $(y_x, z_x) \in M \times M^\perp$ such that

$$x = y_x + z_x \tag{5.12}$$

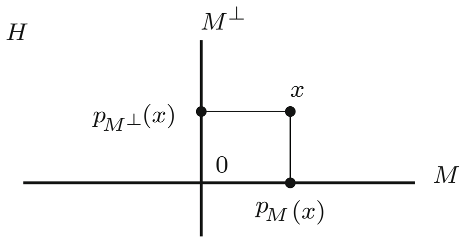


Fig. 5.3 Riesz orthogonal decomposition

(equality (5.12) *is called the* Riesz orthogonal decomposition *of the vector* x*). Moreover,*

$$y_x = p_M(x) \quad \text{and} \quad z_x = p_{M^\perp}(x).$$

(ii) $p_M : H \to H$ *is linear and* $\|p_M(x)\| \le \|x\|$ *for all* $x \in H$.

(iii) *(a)* $p_M \circ p_M = p_M$.

(b) $\ker p_M = M^\perp$.

(c) $p_M(H) = M$.

Proof Let $x \in H$.

(i) Define $y_x = p_M(x)$ and $z_x = x - y_x$; then by (5.10) it follows that $z_x \perp M$ and

$$\langle x - z_x, v\rangle = \langle y_x, v\rangle = 0 \qquad \forall v \in M^\perp.$$

So $z_x = p_{M^\perp}(x)$ in view of Corollary 5.27. Suppose $x = y + z$ for some $y \in M$ and $z \in M^\perp$. Then

$$y_x - y = z - z_x \in M \cap M^\perp = \{0\}.$$

(ii) For any $x_1, x_2 \in H$, $\alpha_1, \alpha_2 \in \mathbb{R}$ and $y \in M$, we have

$$\begin{aligned} &\langle(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 p_M(x_1) + \alpha_2 p_M(x_2)), y\rangle \\ &= \alpha_1\langle x_1 - p_M(x_1), y\rangle + \alpha_2\langle x_2 - p_M(x_2), y\rangle = 0. \end{aligned}$$

Then, by Corollary 5.27, $p_M(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 p_M(x_1) + \alpha_2 p_M(x_2)$. Moreover, since $\langle x - p_M(x), p_M(x)\rangle = 0$ for every $x \in H$, we obtain

$$\|p_M(x)\|^2 = \langle x, p_M(x)\rangle \le \|x\|\,\|p_M(x)\|.$$

(iii) Statement (a) follows from the fact that $p_M(x) = x$ for any $x \in M$. Statements (b) and (c) are consequences of (i). □

Exercise 5.30 In the Hilbert space $H = L^2(0, 1)$ consider the sets

$$M = \{u \in H \mid u \text{ is constant a.e. in } (0, 1)\}$$

and

$$N = \left\{u \in H \;\middle|\; \int_0^1 u(x)\,dx = 0\right\}.$$

1. Show that M and N are closed subspaces of H.
2. Showthat $N = M^\perp$.

3. Does the function $f(x) := 1/\sqrt[3]{x}$, $0 < x < 1$, belong to H? If so, find the Riesz orthogonal decomposition of f with respect to M and N.

Exercise 5.31 Given a Hilbert space H and $A \subset H$, show that the intersection of all closed subspaces including A is a closed subspace of H. Such a subspace, called the *closed subspace generated by* A, will be denoted by $\overline{\text{sp}}(A)$.

Given a Hilbert space H and $A \subset H$, we will denote by $\text{sp}(A)$ the *linear subspace generated by* A, that is,

$$\text{sp}(A) = \left\{ \sum_{k=1}^{n} c_k x_k \;\middle|\; n \geq 1\,,\; c_k \in \mathbb{R}\,,\; x_k \in A \right\}.$$

Exercise 5.32 Show that $\overline{\text{sp}}(A)$ is the closure of $\text{sp}(A)$, i.e., $\overline{\text{sp}}(A) = \overline{\text{sp}(A)}$.

Hint. Since $\overline{\text{sp}(A)}$ is a closed subspace including A, we have that $\overline{\text{sp}}(A) \subset \overline{\text{sp}(A)}$. Conversely, $\text{sp}(A) \subset \overline{\text{sp}}(A)$ yields $\overline{\text{sp}(A)} \subset \overline{\text{sp}}(A)$.

Corollary 5.33 *In a Hilbert space H the following properties hold:*

(i) If M is a closed subspace of H, then $(M^\perp)^\perp = M$.
(ii) For any $A \subset H$, $(A^\perp)^\perp = \overline{\text{sp}}(A)$.
(iii) If L is a subspace of H, then L is dense if and only if $L^\perp = \{0\}$.

Proof We will prove each step of the statement in sequence.

(i) By point (i) of Proposition 5.29 we deduce that

$$p_{M^\perp} = I - p_M.$$

Similarly, $p_{(M^\perp)^\perp} = I - p_{M^\perp} = p_M$. Thus, owing to (iii) of the same proposition,

$$(M^\perp)^\perp = p_{(M^\perp)^\perp}(H) = p_M(H) = M.$$

(ii) Let $M = \overline{\text{sp}}(A)$. Since $A \subset M$, we have $A^\perp \supset M^\perp$ (recall Exercise 5.28(4)). Then $(A^\perp)^\perp \subset (M^\perp)^\perp = M$. Conversely, observe that A is contained in the closed subspace $(A^\perp)^\perp$. So $M \subset (A^\perp)^\perp$.

(iii) Observe that, since $\overline{L}$ is a closed subspace, $\overline{L} = \overline{\text{sp}}(L)$. So, in view of part (ii) above,

$$\overline{L} = H \iff (L^\perp)^\perp = H \iff L^\perp = \{0\}.$$

The proof is thus complete. □

Exercise 5.34 Using Corollary 5.33 show that

$$\ell^1 := \left\{ (x_n)_n \;\middle|\; x_n \in \mathbb{R}\,,\; \sum_{n=1}^{\infty} |x_n| < \infty \right\}$$

is a dense subspace of ℓ^2.

Exercise 5.35 Compute

$$\min_{a,b,c\in\mathbb{R}} \int_{-1}^{1} |x^3 - a - bx - cx^2|^2\, dx.$$

Exercise 5.36 Let H be the set of Borel functions $f : (0, \infty) \to \overline{\mathbb{R}}$ such that

$$\int_0^\infty f^2(x) e^{-x}\, dx < \infty.$$

Show that H is a Hilbert space with the scalar product

$$\langle f, g\rangle = \int_0^\infty f(x)g(x)e^{-x}\, dx.$$

Compute

$$\min_{a,b\in\mathbb{R}} \int_0^\infty |x^2 - a - bx|^2 e^{-x}\, dx.$$

Exercise 5.37 Let H be a Hilbert space, $x_0 \in H$ and $M \subset H$ a closed subspace. Show that

$$\min_{x\in M} \|x - x_0\| = \max\{|\langle x_0, y\rangle| \mid y \in M^\perp,\ \|y\| = 1\}.$$

5.3 Riesz Representation Theorem

5.3.1 Bounded Linear Functionals

Let H be a linear space over $\mathbb{R}$.

Definition 5.38 A linear map $F : H \to \mathbb{R}$ is called a *linear functional on* H.

Definition 5.39 A linear functional F on a pre-Hilbert space H is said to be *bounded* if there exists $C \geq 0$ such that

$$|F(x)| \leq C\|x\| \qquad \forall x \in H.$$

Proposition 5.40 *Let H be a pre-Hilbert space and F a linear functional on H. Then the following statements are equivalent:*

(a) F is continuous.
(b) F is continuous at 0.
(c) F is continuous at some point of H.
(d) F is bounded.

Proof The implications (a) $\Rightarrow$ (b) $\Rightarrow$ (c) and (d) $\Rightarrow$ (b) are trivial. So it suffices to show that (c) $\Rightarrow$ (a) and (b) $\Rightarrow$ (d).

(c) $\Rightarrow$ (a) Let F be continuous at x_0 and let $y_0 \in H$. For any sequence $(y_n)_n$ in H converging to y_0, we have

$$x_n = y_n - y_0 + x_0 \to x_0.$$

Then $F(x_n) = F(y_n) - F(y_0) + F(x_0) \to F(x_0)$. Therefore $F(y_n) \to F(y_0)$. So F is continuous at y_0.

(b) $\Rightarrow$ (d) By hypothesis, there exists $\delta > 0$ such that $|F(x)| < 1$ for every $x \in H$ satisfying $\|x\| < \delta$. Then for any $\varepsilon > 0$ and $x \in H$ we have

$$\left| F\left(\frac{\delta x}{\|x\| + \varepsilon} \right) \right| < 1.$$

So $|F(x)| < \frac{1}{\delta}(\|x\| + \varepsilon)$. Since ε is arbitrary, the conclusion follows. □

Definition 5.41 The family of all bounded linear functionals on a pre-Hilbert space H is called the (topological) *dual* of H and is denoted by H^*. For any $F \in H^*$ we set

$$\|F\|_* = \sup_{\|x\| \le 1} |F(x)|.$$

Exercise 5.42 Let H be a pre-Hilbert space.

1. Show that H^* is a linear space and $\| \cdot \|_*$ is a norm on H^*.
2. Show that for any $F \in H^*$ we have

$$\begin{aligned} \|F\|_* &= \min \left\{ C \ge 0 \;\middle|\; |F(x)| \le C\|x\| \; \forall x \in X \right\} \\ &= \sup_{\|x\|=1} |F(x)| = \sup_{x \ne 0} \frac{|F(x)|}{\|x\|} = \sup_{\|x\|<1} |F(x)| \,. \end{aligned}$$

5.3.2 Riesz Theorem

Example 5.43 Given H a pre-Hilbert space, for any $y \in H$ consider F_y the linear functional on H defined by

$$F_y(x) = \langle x, y \rangle \qquad \forall x \in H.$$

By the Cauchy–Schwarz inequality we get $|F_y(x)| \le \|y\| \, \|x\|$ for any $x \in H$. So $F_y \in H^*$ and $\|F_y\|_* \le \|y\|$. We have thus defined a map

$$\begin{cases} j : H \to H^*, \\ j(y) = F_y. \end{cases} \tag{5.13}$$

It is easy to check that j is linear. Moreover, since $|F_y(y)| = \|y\|^2$ for any $y \in H$, we deduce that $\|F_y\|_* = \|y\|$. Therefore j is a linear *isometry*.[4]

Our next result will show that the map j is also onto. So j is an isometric isomorphism,[5] called the *Riesz isomorphism.*

Theorem 5.44 (Riesz) *Let H be a Hilbert space and let F be a bounded linear functional on H. Then there exists a unique $y_F \in H$ such that*

$$F(x) = \langle x, y_F \rangle, \quad \forall x \in H. \tag{5.14}$$

Moreover, $\|F\|_* = \|y_F\|$.

Proof Suppose $F \neq 0$ (otherwise the conclusion is trivial by taking $y_F = 0$) and set $M = \ker F$. Since M is a closed proper[6] subspace of H, by Corollary 5.33(iii) there exists $y_0 \in M^\perp \setminus \{0\}$. Possibly substituting y_0 by $\frac{y_0}{F(y_0)} \in M^\perp \setminus \{0\}$, we can assume, without loss of generality, $F(y_0) = 1$. Thus, for any $x \in H$ we have that $F(x - F(x)y_0) = 0$, that is, $x - F(x)y_0 \in M$ (see Fig. 5.4). So $\langle x - F(x)y_0, y_0 \rangle = 0$, i.e.,

$$F(x)\|y_0\|^2 = \langle x, y_0 \rangle \qquad \forall x \in H.$$

This implies that $y_F := \frac{y_0}{\|y_0\|^2}$ satisfies (5.14). The uniqueness of y_F, as well as the equality $\|F\|_* = \|y_F\|$, follows from the fact that j in Example 5.43 is a linear isometry. □

---

[4]Given two linear normed spaces X, Y, a map $T : X \to Y$ is called an *isometry* if it satisfies $\|T(x)\| = \|x\|$ for every $x \in X$.

[5]Given two linear normed spaces X, Y, a *(topological) isomorphism* of X onto Y is a linear bijective mapping $T : X \to Y$ such that T and T^{-1} are continuous. If T is also an isometry, that is, $\|T(x)\| = \|x\|$ for every $x \in X$, then T is called an *isometric isomorphism* of X onto Y.

[6]That is, $M \neq H$.

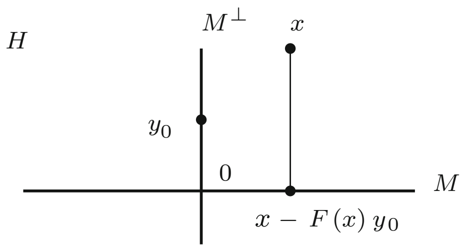


Fig. 5.4 Proof of Riesz Theorem

Example 5.45 If $H = L^2(X, \mu)$, where $(X, \mathscr{E}, \mu)$ is a measure space, from the above theorem we deduce that for every bounded linear functional $F : L^2(X, \mu) \to \mathbb{R}$ there exists a unique $g \in L^2(X, \mu)$ such that

$$F(f) = \int_X fg\, d\mu \qquad \forall f \in L^2(X, \mu).$$

Moreover, $\|F\|_* = \|g\|_2$.

Exercise 5.46 Let $F : L^2(0, 2) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^1 f(x)\, dx + \int_1^2 (x-1) f(x)\, dx.$$

Show that F is bounded and compute $\|F\|_*$.

Definition 5.47 Given a linear space H, a subset $\Pi \subset H$ is called an *affine manifold* if

$$\Pi = x_0 + \Pi_0 := \{x_0 + y \mid y \in \Pi_0\}$$

where x_0 is a fixed vector and Π_0 is a subspace of H. If Π_0 has codimension[7] 1, then the affine manifold Π is called a hyperplane in H.

Given a Hilbert space H and a linear functional $F \in H^*$, $F \neq 0$, for every $c \in \mathbb{R}$ let us set

$$\Pi_c = \{x \in H \mid F(x) = c\}.$$

By the Riesz Theorem we deduce that $\ker F = \Pi_0 = \{y_F\}^\perp$. So, owing to Corollary 5.33(ii), $\Pi_0^\perp = \{\lambda y_F \mid \lambda \in \mathbb{R}\}$. Then from the Riesz orthogonal decomposition it follows that Π_0 is a closed subspace of codimension 1. Moreover, for any $x_c \in \Pi_c$, we have that $\Pi_c = x_c + \Pi_0$. Therefore Π_c is a closed hyperplane in H.

The following result provides a sufficient condition for two convex sets to be 'strictly separated' by a closed hyperplane.

---

[7] We say that a subspace Π_0 of a linear space H has codimension n if there exist n linearly independent vectors $x_1, \ldots, x_n \in H$ such that $x_1, \ldots, x_n \notin \Pi_0$ and $H = \Pi_0 \oplus \mathbb{R}x_1 \oplus \cdots \oplus \mathbb{R}x_n$, where the symbol '$\oplus$' denotes the direct sum.

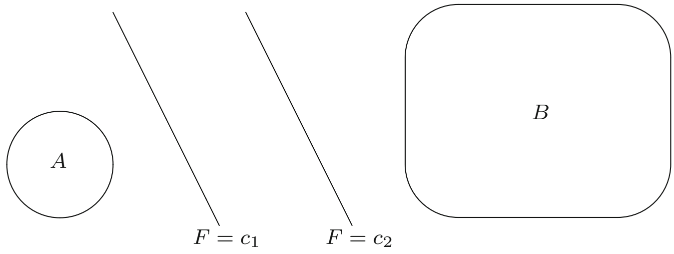


Fig. 5.5 Separation of convex sets

Proposition 5.48 *Let A and B be nonempty disjoint convex sets in a Hilbert space H. Suppose that A is compact and B is closed. Then there exist a functional $F \in H^*$ and two constants c_1, c_2 such that*

$$F(x) \le c_1 < c_2 \le F(y) \qquad \forall x \in A\,,\ \forall y \in B$$

(see Fig. 5.5).

Proof Let $C = B - A := \{y - x \mid x \in A\,,\ y \in B\}$. It is easy to verify that C is a nonempty convex set such that $0 \notin C$. We claim that C is closed. Let $(y_n - x_n)_n \subset C$ be a sequence such that $y_n - x_n \to z$. Since A is compact, there exists a subsequence $(x_{k_n})_n$ such that $x_{k_n} \to x \in A$. Therefore

$$y_{k_n} = y_{k_n} - x_{k_n} + \underbrace{x_{k_n} - x}_{\to 0} + x \to z + x$$

and so, since B is closed, $z + x \in B$. It follows that C is closed, as claimed. Then, thanks to Theorem 5.23, $z_0 := p_C(0)$ satisfies $z_0 \neq 0$ and

$$\langle 0 - z_0, y - x - z_0\rangle \le 0 \qquad \forall x \in A\,,\ \forall y \in B,$$

or, equivalently,

$$\langle x, z_0\rangle + \|z_0\|^2 \le \langle y, z_0\rangle \qquad \forall x \in A\,,\ \forall y \in B.$$

The conclusion follows taking

$$F = F_{z_0}\,, \qquad c_1 = \sup_{x\in A}\langle x, z_0\rangle, \qquad c_2 = \inf_{y\in B}\langle y, z_0\rangle. \qquad \square$$

Exercise 5.49 1. Given $N \geq 1$, define

$$F : \ell^2 \to \mathbb{R}, \quad F((x_k)_k) = x_N.$$

Show that $F \in (\ell^2)^*$ and find $y \in \ell^2$ such that $F = F_y$.

2. Show that, for any $x = (x_k)_k \in \ell^2$, the power series $\sum_{k=1}^{\infty} x_k z^k$ has radius of convergence at least 1.
3. For a given $z \in (-1, 1)$, set

$$F : \ell^2 \to \mathbb{R}, \quad F((x_k)_k) = \sum_{k=1}^{\infty} x_k z^k.$$

Show that $F \in (\ell^2)^*$ and find $y \in \ell^2$ satisfying $F = F_y$.

4. In ℓ^2 consider the sets[8]

$$A := \left\{(x_k)_k \in \ell^2 \mid k|x_k - k^{-2/3}| \leq x_1 \quad \forall k \geq 2\right\}$$

and

$$B := \left\{(x_k)_k \in \ell^2 \mid x_k = 0 \quad \forall k \geq 2\right\}.$$

a. Show that A and B are disjoint closed convex sets in ℓ^2.
b. Show that

$$A - B = \left\{(x_k)_k \in \ell^2 \mid \exists C \geq 0 \ : \ k|x_k - k^{-2/3}| \leq C \quad \forall k \geq 2\right\}.$$

c. Deduce that $A - B$ is *dense* in ℓ^2.
Hint. Given $x = (x_k)_k \in \ell^2$, let $(x^n)_n$ be the sequence in $A - B$ defined by

$$x_k^n = \begin{cases} x_k & \text{if } k \leq n, \\ k^{-2/3} & \text{if } k \geq n+1. \end{cases}$$

Then $x^n \xrightarrow{\ell^2} x$.

d. Show that A and B cannot be separated by a functional $F \in (\ell^2)^*$ as in Proposition 5.48. (This example shows that the compactness assumption on A cannot be dropped in Proposition 5.48.)
Hint. Otherwise $A - B$ would be contained in the half-space $\{F \leq 0\}$.

---

[8] See [Ko02, p. 14].

Example 5.50 (An unbounded functional) In ℓ^2, choose vectors $e_0 = (\frac{1}{n})_n$ and

$$e_k = (\overbrace{0, \dots, 0}^{k-1}, 1, 0, \dots) \qquad k = 1, 2, \dots.$$

Observe that the sequence $\{e_k \mid k \geq 0\}$ is a family of linearly independent vectors. So let

$$E = \{e_k \mid k \geq 0\} \cup \{f_i \mid i \in I\}$$

be a *Hamel basis*[9] of ℓ^2. Define the linear functional $\Phi : \ell^2 \to \mathbb{R}$ as follows: given $x \in \ell^2$, then x has a unique representation as a finite linear combination of vectors from the set E, say $x = \sum_{k=0}^{N} \lambda_k e_k + \sum_{i \in J} \mu_i f_i$ with $N \geq 0$ and $J \subset I$ finite. We set

$$\Phi\left(\sum_{k=0}^{N} \lambda_k e_k + \sum_{i \in J} \mu_i f_i\right) = \lambda_0.$$

Then $\{e_k \mid k \geq 1\} \subset \ker(\phi)$, so that $\ker(\Phi)$ is dense in ℓ^2. On the other hand, $e_0 \notin \ker(\Phi)$, hence $\ker(\Phi)$ is not closed. This shows that Φ fails to be continuous.

Definition 5.51 Let H be a pre-Hilbert space. A map $a : H \times H \to \mathbb{R}$ is called a *bilinear form* if it is linear in each argument separately:

$$a(\lambda_1 x_1 + \lambda_1 x_2, y) = \lambda_1 a(x_1, y) + \lambda_2 a(x_2, y) \quad \forall \lambda_1, \lambda_2 \in \mathbb{R}, \ \forall x_1, x_2, y \in H,$$

$$a(x, \lambda_1 y_1 + \lambda_1 y_2) = \lambda_1 a(x, y_1) + \lambda_2 a(x, y_2) \quad \forall \lambda_1, \lambda_2 \in \mathbb{R}, \ \forall x, y_1, y_2 \in H.$$

A bilinear form $a : H \times H \to \mathbb{R}$ is said to be

- *Bounded* if there exists $M > 0$ such that

$$a(x, y) \leq M\|x\|\|y\| \quad \forall x, y \in H.$$

- *Positive* (or *coercive*) if there exists $m > 0$ such that

$$a(x, x) \geq m\|x\|^2 \quad \forall x \in H.$$

---

[9]Given X a linear space, a maximal subset of X constituted by linearly independent vectors is called a Hamel basis. We recall that, by applying Zorn's Lemma, one can prove that every set of linearly independent vectors is contained in a Hamel basis. Moreover, if $(e_i)_{i \in I}$ is a Hamel basis in X, then the linear subspace generated by $(e_i)_{i \in I}$ coincides with X, i.e., $X = \{\sum_{j \in J} \lambda_j e_j \mid J \subset I \text{ finite}, \ \lambda_j \in \mathbb{R}\}$.

Theorem 5.52 (Lax–Milgram) *Let H be a Hilbert space and let $a : H \times H \to \mathbb{R}$ be a positive bounded bilinear form. In addition, let $F : H \to \mathbb{R}$ be a bounded linear functional. Then there exists a unique $y_F \in H$ such that*

$$F(x) = a(x, y_F) \quad \forall x \in H.$$

Proof For each fixed element $y \in H$, the mapping $x \in H \mapsto a(x, y)$ is a bounded linear functional on H. By the Riesz Representation Theorem we deduce that for every $y \in H$ there exists a unique element $Ay \in H$ such that

$$a(x, y) = \langle x, Ay \rangle \quad \forall x \in H.$$

We claim that the map $A : H \to H$ is a bounded linear operator (see Sect. 6.2). Indeed, linearity can be easily checked. Furthermore

$$\|Ax\|^2 = \langle Ax, Ax \rangle = a(Ax, x) \leq M\|Ax\|\|x\|,$$

by which $\|Ax\| \leq M\|x\|$. Owing to Proposition 6.10, A is a bounded linear operator. Hence, A is continuous. Moreover,

$$\|x\|\|Ax\| \geq \langle x, Ax \rangle = a(x, x) \geq m\|x\|^2 \quad \forall x \in H.$$

Thus,

$$\|Ax\| \geq m\|x\| \quad \forall x \in H. \tag{5.15}$$

Consequently, A is also injective and $R(A)$ is closed, where $R(A)$ stands for the range of A. Indeed, if $(Ax_n)_n \subset R(A)$ is such that $Ax_n \to y$ for some $y \in Y$, then by (5.15) we deduce that $m\|x_n - x_m\| \leq \|Ax_n - Ax_m\|$. So, being $(x_n)_n$ a Cauchy sequence in H, it converges to some $x \in X$; by continuity, $Ax_n \to Ax = y \in R(A)$.

We are going to prove that

$$R(A) = H. \tag{5.16}$$

If not, since $R(A)$ is closed, by Corollary 5.33(iii) there would exist a nonzero vector $w \in R(A)^\perp$. But this leads to a contradiction since it implies that $m\|w\|^2 \leq a(w, w) = \langle w, Aw \rangle = 0$.

Next, we observe that once more from Riesz Representation Theorem there exists $z_F \in H$ such that

$$F(x) = \langle x, z_F \rangle \quad \forall x \in H.$$

Then by (5.16) we find $y_F \in H$ satisfying $Ay_F = z_F$. Thus

$$F(x) = \langle x, z_F \rangle = \langle x, Ay_F \rangle = a(x, y_F) \quad \forall x \in H.$$

To prove the uniqueness of y_F, observe that if y and y' are such that $F(x) = a(x, y) = a(x, y')$ for all $x \in H$, then $a(x, y - y') = 0$ for all $x \in H$. Setting $x = y - y'$ we get $0 = a(y - y', y - y') \geq m\|y - y'\|^2$. □

Remark 5.53 If the bilinear form a is also symmetric, that is,

$$a(x, y) = a(y, x) \quad \forall x, y \in H,$$

then a much easier proof of Lax–Milgram Theorem can be provided by noting that $(x, y) \mapsto a(x, y)$ is a new scalar product on H which induces an equivalent norm, hence Riesz Representation Theorem directly applies.

5.4 Orthonormal Sequences and Bases

Definition 5.54 Let H be a pre-Hilbert space. A sequence $(e_k)_k \subset H$ is said to be *orthonormal* if

$$\langle e_h, e_k \rangle = \begin{cases} 1 & \text{if} \quad h = k, \\ 0 & \text{if} \quad h \neq k. \end{cases}$$

Example 5.55 The sequence of vectors

$$e_k = (\overbrace{0, \dots, 0}^{k-1}, 1, 0, \dots) \qquad k = 1, 2, \dots$$

is orthonormal in ℓ^2.

Example 5.56 Let $\{\varphi_k \mid k = 0, 1, \dots\}$ be the sequence of functions in $L^2(-\pi, \pi)$ defined by

$$\varphi_0(t) = \frac{1}{\sqrt{2\pi}},$$
$$\varphi_{2k-1}(t) = \frac{\sin(kt)}{\sqrt{\pi}}, \quad \varphi_{2k}(t) = \frac{\cos(kt)}{\sqrt{\pi}} \quad (k \geq 1).$$

Since for any $h, k \geq 1$ we have

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(ht)\sin(kt)\,dt = 0,$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} \sin(ht)\sin(kt)\,dt = \begin{cases} 0 & \text{if } h \neq k, \\ 1 & \text{if } h = k, \end{cases}$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} \cos(ht)\cos(kt)\,dt = \begin{cases} 0 & \text{if } h \neq k, \\ 1 & \text{if } h = k, \end{cases}$$

it is easy to check that $\{\varphi_k \mid k = 0, 1, \ldots\}$ is an orthonormal sequence in $L^2(-\pi, \pi)$. Such a sequence is called the *trigonometric system*.

5.4.1 Bessel's Inequality

Proposition 5.57 *Let H be a Hilbert space and let $(e_k)_k$ be an orthonormal sequence.*

1. *For every $N \in \mathbb{N}$,* Bessel's identity *holds:*

$$\left\| x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k \right\|^2 = \|x\|^2 - \sum_{k=1}^{N} \left| \langle x, e_k \rangle \right|^2 \quad \forall x \in H. \tag{5.17}$$

2. Bessel's inequality *holds:*

$$\sum_{k=1}^{\infty} \left| \langle x, e_k \rangle \right|^2 \leq \|x\|^2 \quad \forall x \in H.$$

   In particular, the series in the left-hand side is convergent.
3. *For any sequence $(c_k)_k \subset \mathbb{R}$ we have*[10]*:*

$$\sum_{k=1}^{\infty} c_k e_k \in H \iff \sum_{k=1}^{\infty} |c_k|^2 < \infty.$$

Proof Let $x \in H$. Bessel's identity can be easily checked by induction on N. For $N = 1$, (5.17) is true.[11] Suppose it holds for some $N \geq 1$. Then

$$\begin{aligned} &\left\| x - \sum_{k=1}^{N+1} \langle x, e_k \rangle e_k \right\|^2 \\ &= \left\| x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k \right\|^2 + \left| \langle x, e_{N+1} \rangle \right|^2 - 2\left\langle x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k, \langle x, e_{N+1} \rangle e_{N+1} \right\rangle \\ &= \|x\|^2 - \sum_{k=1}^{N} \left| \langle x, e_k \rangle \right|^2 - \left| \langle x, e_{N+1} \rangle \right|^2. \end{aligned}$$

[10] The statement '$\sum_{k=1}^{\infty} c_k e_k \in H$' means that the sequence of partial sums $\sum_{k=1}^{n} c_k e_k$ is convergent in the metric of H as $n \to \infty$.

[11] Indeed (5.17) for $N = 1$ has been used to prove Cauchy–Schwarz inequality (5.2).

So (5.17) holds for any $N \geq 1$. Moreover, Bessel's identity implies that all the partial sums of the series $\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2$ are bounded above by $\|x\|^2$, thus yielding Bessel's inequality. Finally, for every $n \in \mathbb{N}$ we have

$$\left\| \sum_{k=n+1}^{n+p} c_k e_k \right\|^2 = \sum_{k=n+1}^{n+p} |c_k|^2, \qquad p = 1, 2, \ldots.$$

Therefore the partial sums of the series $\sum_{k=1}^{\infty} c_k e_k$ is a Cauchy sequence in H if and only if the number series $\sum_{k=1}^{\infty} c_k^2$ is convergent. Since the space H is complete, the conclusion of point 3 follows. □

Definition 5.58 Under the same assumptions of Proposition 5.57, for every $x \in H$ the numbers $\langle x, e_k\rangle$ are called the *Fourier coefficients* of x and the series $\sum_{k=1}^{\infty} \langle x, e_k\rangle e_k$ is called the *Fourier series* of x.

Remark 5.59 Under the same assumptions of Proposition 5.57, fixed $n \in \mathbb{N}$ let us set $M_n := \mathrm{sp}\big(\{e_1, \ldots, e_n\}\big) = \mathrm{sp}(e_1, \ldots, e_n)$. Then

$$p_{M_n}(x) = \sum_{k=1}^{n} \langle x, e_k\rangle e_k \qquad \forall x \in H.$$

Indeed, for every $x \in H$ and $c_1, c_2, \ldots, c_n \in \mathbb{R}$, we have

$$\begin{aligned}\left\| x - \sum_{k=1}^{n} c_k e_k \right\|^2 &= \|x\|^2 - 2\sum_{k=1}^{n} c_k \langle x, e_k\rangle + \sum_{k=1}^{n} |c_k|^2 \\ &= \left(\|x\|^2 - \sum_{k=1}^{n} \big|\langle x, e_k\rangle\big|^2 \right) + \sum_{k=1}^{n} \big|c_k - \langle x, e_k\rangle\big|^2 \\ &= \left\| x - \sum_{k=1}^{n} \langle x, e_k\rangle e_k \right\|^2 + \sum_{k=1}^{n} \big|c_k - \langle x, e_k\rangle\big|^2\end{aligned}$$

owing to Bessel's identity (5.17).

5.4.2 Orthonormal Bases

Let us characterize situations where a vector $x \in H$ is given by the sum of its Fourier series.

Theorem 5.60 *Let $(e_k)_k$ be an orthonormal sequence in a Hilbert space H. Then the following properties are equivalent:*

(a) $\mathrm{sp}(e_k \,|\, k \in \mathbb{N})$ *is dense in H.*

(b) Every $x \in H$ is given by the sum of its Fourier series[12]*:*

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k.$$

(c) Every $x \in H$ satisfies Parseval's identity*:*

$$\|x\|^2 = \sum_{k=1}^{\infty} \left|\langle x, e_k \rangle\right|^2. \tag{5.18}$$

(d) If $x \in H$ is such that $\langle x, e_k \rangle = 0$ for every $k \in \mathbb{N}$, then $x = 0$.

Proof We will prove that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (a).

- (a) $\Rightarrow$ (b)
  For any $n \in \mathbb{N}$ let $M_n := \text{sp}(e_1, \ldots, e_n)$. Then by hypothesis for every $x \in H$ we have $d_{M_n}(x) \to 0$ as $n \to \infty$, where $d_{M_n}(x)$ denotes the distance of x from M_n. Thus, owing to Remark 5.59,

$$\left\| x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\|^2 = \|x - p_{M_n}(x)\|^2 = d^2_{M_n}(x) \to 0 \quad (n \to \infty).$$

  This yields (b).
- (b) $\Rightarrow$ (c) This part follows from Bessel's identity.
- (c) $\Rightarrow$ (d) Obvious.
- (d) $\Rightarrow$ (a)
  Let $L := \text{sp}(e_k \mid k \in \mathbb{N})$. Then by hypothesis $L^\perp = \{0\}$. So L is dense thanks to point (iii) of Corollary 5.33. □

Definition 5.61 An orthonormal sequence $(e_k)_k$ in a Hilbert space H is said to be *complete* if $\text{sp}(e_k \mid k \in \mathbb{N})$ is dense H (or if any of the four equivalent conditions of Theorem 5.60 holds). In this case, $(e_k)_k$ is called an *orthonormal basis* of H.

Exercise 5.62 Show that if a Hilbert space H possesses an orthonormal basis $(e_k)_k$, then H is *separable*, that is, H contains a dense countable set.

Hint. Consider the set of all finite linear combinations of the vectors e_k with rational coefficients.

[12] More exactly, the statement '$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$' means that the sequence of partial sums corresponding to the Fourier series of x converges to x in the metric of H, i.e., $\sum_{k=1}^{n} \langle x, e_k \rangle e_k \to x$ in H as $n \to \infty$.

Exercise 5.63 Let $(y_k)_k$ be a sequence in a Hilbert space H. Show that there exists an at most countable set of *linearly independent* vectors $\{x_j \mid j \in J\}$ in H such that

$$\mathrm{sp}(y_k \mid k \in \mathbb{N}) = \mathrm{sp}(x_j \mid j \in J).$$

Hint. For every $j \in \mathbb{N}$ let k_j be the first index $k \in \mathbb{N}$ such that

$$\dim \mathrm{sp}(y_1, \ldots, y_k) = j.$$

Set $x_j := y_{k_j}$. Then $\mathrm{sp}(x_1, \ldots, x_j) = \mathrm{sp}(y_1, \ldots, y_{k_j})$.

Exercise 5.64 Let $(e_k)_k$ be an orthonormal basis in a Hilbert space H. Show that

$$\langle x, y\rangle = \sum_{k=1}^{\infty} \langle x, e_k\rangle \langle y, e_k\rangle \qquad \forall x, y \in H.$$

Hint. Observe that

$$\langle x, y\rangle = \frac{\|x + y\|^2 - \|x\|^2 - \|y\|^2}{2}$$

and use Parseval's identity (5.18).

Next result shows the converse of the property described in Exercise 5.62.

Proposition 5.65 *Let H be an infinite-dimensional separable Hilbert space. Then H possesses an orthonormal basis.*

Proof Let $(y_k)_k$ be a dense sequence in H and let A be the set of at most countable linearly independent vectors constructed in Exercise 5.63. Then $\mathrm{sp}\big(A\big) = \mathrm{sp}(y_k \mid k \in \mathbb{N})$ is dense in H. We claim that A is infinite. Indeed, if not, then $\mathrm{sp}(A)$ would have finite dimension and, consequently, it would be a closed subspace of H (see Corollary C.4), which in turn implies $\mathrm{sp}(A) = H$, in contradiction with the assumption that H is infinite-dimensional. So we deduce that A is infinite and countable. Set $A = (x_k)_k$ and define[13]

$$e_1 = \frac{x_1}{\|x_1\|} \quad \text{and} \quad e_k = \frac{x_k - \sum_{j<k} \langle x_k, e_j\rangle e_j}{\Big\| x_k - \sum_{j<k} \langle x_k, e_j\rangle e_j \Big\|} \quad (k \geq 2).$$

Then $(e_k)_k$ is an orthonormal sequence by construction. Moreover, we have

$$\mathrm{sp}(e_1, \ldots, e_k) = \mathrm{sp}(x_1, \ldots, x_k) \qquad \forall k \geq 1. \tag{5.19}$$

Indeed, by induction it is easy to verify that $\{e_1, \ldots, e_k\} \subset \mathrm{sp}(x_1, \ldots, x_k)$, by which $\mathrm{sp}(e_1, \ldots, e_k) \subset \mathrm{sp}(x_1, \ldots, x_k)$. On the other hand, the vectors $e_1, \ldots, e_k$

[13]This procedure is known as *Gram–Schmidt orthonormalization*.

are linearly independent because they are orthogonal (see Exercise 5.20). Thus $\dim \operatorname{sp}(e_1, \dots, e_k) = k = \dim \operatorname{sp}(x_1, \dots, x_k)$, and (5.19) follows. Therefore $\operatorname{sp}(e_k \mid k \in \mathbb{N})$ is dense in H. □

Example 5.66 In $H = \ell^2$ it is immediate to verify that the orthonormal sequence $(e_k)_k$ of Example 5.55 is complete.

Remark 5.67 If H is not separable, we can also establish (using the Axiom of Choice) the existence of an uncountable orthonormal basis $\{e_i \mid i \in I\}$. Theorem 5.60 is still valid provided we substitute convergent series by *summable families* (see [Sh61]). For instance, let us consider an uncountable set A and, for every function $f : A \to [0, \infty)$, let us set

$$\sum_{\alpha \in A} f(\alpha) := \sup \left\{ \sum_{\alpha \in F} f(\alpha) : F \subseteq A, \quad F \text{ finite or countable} \right\}.$$

Observe that, since

$$\{\alpha \in A : f(\alpha) \neq 0\} = \bigcup_{n=1}^{\infty} \left\{ \alpha \in A : f(\alpha) \geq \frac{1}{n} \right\},$$

we deduce that

$$\sum_{\alpha \in A} f(\alpha) < \infty \implies A_f := \{\alpha \in A : f(\alpha) \neq 0\} \text{ is finite or countable.}$$

Next define

$$\ell^2(A) = \left\{ x : A \to \mathbb{R} : \|x\|_2^2 = \sum_{\alpha \in A} |x(\alpha)|^2 < \infty \right\}.$$

In other words, $\ell^2(A) = L^2(A, \mu^\#)$, where $\mu^\#$ denotes the counting measure on A. It follows that $\ell^2(A)$ is a Banach space. Set, moreover,

$$\langle x, y \rangle = \sum_{\alpha \in A_x \cup A_y} x(\alpha) y(\alpha) \quad \forall x, y \in \ell^2(A),$$

where $A_x \cup A_y$ is finite or countable. Then $\langle \cdot, \cdot \rangle$ is the scalar product associated with the norm $\| \cdot \|_2$. So $\ell^2(A)$ is a Hilbert space. Finally, if we define

$$x_\alpha(\beta) = \begin{cases} 1 & \text{if } \beta = \alpha, \\ 0 & \text{if } \beta \neq \alpha, \end{cases}$$

then $\{x_\alpha\}_{\alpha \in A}$ is an orthonormal family.

Proposition 1.75 guarantees that the Lebesgue measure on $\mathbb{R}^N$ is the unique Radon measure, up to multiplicative constants, which is translation invariant. The following exercise shows that in an infinite-dimensional Hilbert space there is no nontrivial measure with analogous properties.

Exercise 5.68 Let H be an infinite-dimensional separable Hilbert space. Show that if μ is a Borel measure on H, which is translation invariant and finite on all bounded subsets of H, then $\mu \equiv 0$.

Hint. Let μ be a Borel measure on H which is translation invariant and finite on bounded sets of H. Assume $\mu \neq 0$. By using the balls (5.7), $\mu(B_r(0)) > 0$ for some radius $r > 0$. Given an orthonormal complete sequence $(e_k)_k$, fix $R > r\sqrt{2}$. Then for $i \neq j$ we have $B_r(Re_i) \cap B_r(Re_j) = \varnothing$.

Exercise 5.69 Let $(e_k)_k$ and $(e'_k)_k$ be two orthonormal sequences in a Hilbert space H such that

$$\sum_{k=1}^{\infty} \|e_k - e'_k\|^2 < 1.$$

Show that:

- For every $x \in \{e'_k \mid k \in \mathbb{N}\}^{\perp} \setminus \{0\}$ we have $\displaystyle\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 < \|x\|^2$.
- $(e_k)_k$ is complete if and only if $(e'_k)_k$ is complete.

5.4.3 *Completeness of the Trigonometric System*

In this section we will show that the orthonormal sequence $\{\varphi_k \mid k = 0, 1, \ldots\}$ defined in Example 5.56 is an orthonormal basis in $L^2(-\pi, \pi)$.

To this aim we begin by constructing a sequence of trigonometric polynomials with special properties. We recall that a *trigonometric polynomial* $q(t)$ is a sum of the form

$$q(t) = a_0 + \sum_{k=1}^{n} \big(a_k \cos(kt) + b_k \sin(kt)\big) \qquad (n \in \mathbb{N})$$

with coefficients $a_k, b_k \in \mathbb{R}$, i.e., an element of $\operatorname{sp}(\varphi_k \mid k = 0, 1, \ldots)$. Any trigonometric polynomial q is a continuous 2π-periodic function.

Lemma 5.70 *There exists a sequence of trigonometric polynomials $(q_n)_n$ (see Fig. 5.6) such that*

(a) $q_n(t) \geq 0$ *for every* $t \in \mathbb{R}$ *and* $n \in \mathbb{N}$.

(b) $\frac{1}{2\pi}\int_{-\pi}^{\pi} q_n(t)\,dt = 1$ *for every* $n \in \mathbb{N}$.

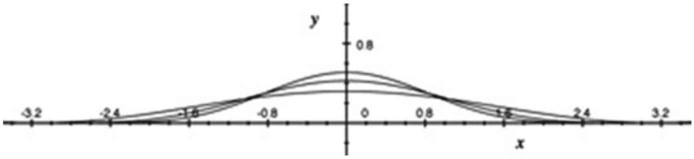


Fig. 5.6 The sequence q_n

(c) For any $\delta > 0$

$$\lim_{n\to\infty} \sup_{\delta \le |t| \le \pi} q_n(t) = 0.$$

Proof For every $n \in \mathbb{N}$, define

$$q_n(t) = c_n \Big(\frac{1+\cos t}{2}\Big)^n \qquad \forall t \in \mathbb{R},$$

where c_n is chosen in such a way that property (b) is satisfied. Recalling that

$$\cos(kt)\cos t = \frac{1}{2}\Big[\cos\big((k+1)t\big) + \cos\big((k-1)t\big)\Big],$$

it is easy to check that each q_n is a finite linear combination of elements $\cos(kt)$, $k \ge 0$. So q_n is a trigonometric polynomial.

Since property (a) is immediate, there only remains to check (c). Observe that, since q_n is even,

$$\begin{aligned} 1 = \frac{c_n}{\pi}\int_0^\pi \Big(\frac{1+\cos t}{2}\Big)^n dt &\ge \frac{c_n}{\pi}\int_0^\pi \Big(\frac{1+\cos t}{2}\Big)^n \sin t\, dt \\ &= \frac{c_n}{\pi(n+1)}\Big[-2\Big(\frac{1+\cos t}{2}\Big)^{n+1}\Big]_0^\pi = \frac{2c_n}{\pi(n+1)}, \end{aligned}$$

by which we deduce

$$c_n \le \frac{\pi(n+1)}{2} \qquad \forall n \in \mathbb{N}.$$

Now, fix $0 < \delta < \pi$. Since q_n is even in $[-\pi, \pi]$ and decreasing in $[0, \pi]$, using the above estimate for c_n, we obtain

$$\sup_{\delta \le |t| \le \pi} q_n(t) = q_n(\delta) \le \frac{\pi(n+1)}{2}\Big(\frac{1+\cos\delta}{2}\Big)^n \overset{n\to\infty}{\longrightarrow} 0,$$

which completes the proof. □

The next step is to derive the classical uniform approximation theorem by trigonometric polynomials.

Theorem 5.71 (Weierstrass) *Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous 2π-periodic function. Then there exists a sequence of trigonometric polynomials $(p_n)_n$ such that $\|f - p_n\|_\infty \to 0$ as $n \to \infty$.*

Proof [14] Let (q_n) be the sequence of trigonometric polynomials constructed in Lemma 5.70. For any $n \in \mathbb{N}$ and $t \in \mathbb{R}$, a simple periodicity argument shows that

$$
\begin{aligned}
p_n(t) :=& \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)q_n(s)\,ds \\
&= \frac{1}{2\pi}\int_{t-\pi}^{t+\pi} f(\tau)q_n(t-\tau)\,d\tau = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)q_n(t-\tau)\,d\tau.
\end{aligned}
$$

This implies that p_n is a trigonometric polynomial. Indeed, since

$$
q_n(t) = a_0 + \sum_{k=1}^{k_n} a_k \cos(kt),
$$

we have that

$$
\begin{aligned}
p_n(t) - \frac{a_0}{2\pi}\int_{-\pi}^{\pi} f(\tau)\,d\tau &= \frac{1}{2\pi}\sum_{k=1}^{k_n}\int_{-\pi}^{\pi} f(\tau)a_k \cos\big(k(t-\tau)\big)\,d\tau \\
= \frac{1}{2\pi}\sum_{k=1}^{k_n} a_k\Big[\cos(kt)&\int_{-\pi}^{\pi} f(\tau)\cos(k\tau)\,d\tau + \sin(kt)\int_{-\pi}^{\pi} f(\tau)\sin(k\tau)\,d\tau\Big].
\end{aligned}
$$

For any $\delta \in (0, \pi]$ let

$$
\omega_f(\delta) = \sup_{|x-y|<\delta} |f(x) - f(y)|.
$$

Using properties (a) and (b) of Lemma 5.70, for every $t \in \mathbb{R}$ we have

$$
\begin{aligned}
|f(t) - p_n(t)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\big[f(t) - f(t-s)\big]q_n(s)\,ds\right| \\
&\le \frac{1}{2\pi}\int_{-\pi}^{\pi}\big|f(t) - f(t-s)\big|q_n(s)\,ds \\
&\le \frac{1}{2\pi}\int_{-\delta}^{\delta}\omega_f(\delta)q_n(s)\,ds + \frac{1}{2\pi}\int_{\delta\le|s|\le\pi} 2\|f\|_\infty q_n(s)\,ds \\
&\le \omega_f(\delta) + 2\|f\|_\infty \sup_{\delta\le|s|\le\pi} q_n(s).
\end{aligned}
$$

[14] This proof, based on a *convolution* method, is due to de la Vallée Poussin.

Since f is uniformly continuous, we deduce that $\omega_f(\delta) \to 0$ as $\delta \to 0$. Given $\varepsilon > 0$, let $\delta_\varepsilon \in (0, \pi]$ be such that $\omega_f(\delta_\varepsilon) < \varepsilon$. Owing to point (c) of Lemma 5.70, there exists $n_\varepsilon \in \mathbb{N}$ such that $\sup_{\delta_\varepsilon \le |s| \le \pi} q_n(s) < \varepsilon$ for every $n \ge n_\varepsilon$. Then

$$\|f - p_n\|_\infty < (1 + 2\|f\|_\infty)\varepsilon \qquad \forall n \ge n_\varepsilon,$$

thus completing the proof. □

Remark 5.72 Weierstrass' Theorem can be reformulated as follows: any continuous function $f : [a, b] \to \mathbb{R}$ such that $f(a) = f(b)$ is the uniform limit of a sequence of trigonometric polynomials in $[a, b]$, where by a trigonometric polynomial in $[a, b]$ we mean a finite linear combination of elements of the system

$$1, \quad \cos\frac{2\pi kt}{b-a}, \quad \sin\frac{2\pi kt}{b-a} \qquad (k \ge 1).$$

Since the functions $\cos(kt)$ and $\sin(kt)$ are analytic, we deduce that any continuous function $f : [a, b] \to \mathbb{R}$ is the uniform limit of a sequence of algebraic polynomials.[15] For a direct proof see, for instance, [Ro68].

We are now ready to deduce the announced completeness of the trigonometric system. We recall that $\mathscr{C}_c(a, b) = \mathscr{C}_c\big((a, b)\big)$ denotes the space of all continuous functions $f : (a, b) \to \mathbb{R}$ with compact support (see Sect. 3.4.2).

Theorem 5.73 $\{\varphi_k \,|\, k = 0, 1, \ldots\}$ *is an orthonormal basis of* $L^2(-\pi, \pi)$.

Proof We will show that trigonometric polynomials are dense in $L^2(-\pi, \pi)$ and then the conclusion will follow from Theorem 5.60. Let $f \in L^2(-\pi, \pi)$ and fix $\varepsilon > 0$. Since $\mathscr{C}_c(-\pi, \pi)$ is dense in $L^2(-\pi, \pi)$ on account of Theorem 3.45, there exists $f_\varepsilon \in \mathscr{C}_c(-\pi, \pi)$ such that $\|f - f_\varepsilon\|_2 < \varepsilon$. Clearly, we can extend f_ε, by periodicity, to a periodic continuous function on the whole real line. Moreover, by the Weierstrass Theorem (Theorem 5.71) there exists a trigonometric polynomial p_ε such that $\|f_\varepsilon - p_\varepsilon\|_\infty < \varepsilon$. Then

$$\|f - p_\varepsilon\|_2 \le \|f - f_\varepsilon\|_2 + \|f_\varepsilon - p_\varepsilon\|_2 \le \varepsilon + \varepsilon\sqrt{2\pi}$$

and the conclusion follows. □

Remark 5.74 Let $f \in L^2(-\pi, \pi)$. According to Definition 5.58 the Fourier coefficients of f with respect to the trigonometric system are given by

$$\langle f, \varphi_k \rangle = \int_{-\pi}^{\pi} f(t)\varphi_k(t)\,dt := \hat{f}(k), \quad k = 0, 1, 2 \ldots.$$

[15] It suffices to write f as $f = (f - g) + g$, where $g = (x - a)\frac{f(b) - f(a)}{b - a}$, and apply Weierstrass Theorem to the function $f - g$ which satisfies $(f - g)(a) = (f - g)(b) = f(a)$.

Thus, the associated Fourier series is

$$\frac{\hat{f}(0)}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} \Big[\hat{f}(2k)\cos(kt) + \hat{f}(2k-1)\sin(kt) \Big], \tag{5.20}$$

whose partial sums are the following trigonometric polynomials

$$S_n(f) = S_n(f,t) = \frac{\hat{f}(0)}{\sqrt{2\pi}} + \frac{1}{\sqrt{\pi}} \sum_{k=1}^{n} \Big[\hat{f}(2k)\cos(kt) + \hat{f}(2k-1)\sin(kt) \Big].$$

Since the trigonometric system is an orthonormal basis of $L^2(-\pi,\pi)$, then by Theorem 5.60 we have that:

- f is given by the sum of its Fourier series with respect to the trigonometric system, that is,
$$S_n(f) \xrightarrow{L^2} f \text{ as } n \to \infty.$$
- Parseval's identity holds:
$$\|f\|_2^2 = \sum_{k=0}^{\infty} |\hat{f}(k)|^2. \tag{5.21}$$

We note that we have no information on the pointwise convergence of the Fourier series (5.20), except that there exists a subsequence $(S_{n_k})_k$ converging a.e. in $(-\pi,\pi)$ (see Theorem 3.11). Actually, one can prove that the Fourier series itself is convergent a.e. (see [Ka76, Mo71]).

Exercise 5.75 Applying (5.21) to the function

$$f(t) = t \qquad t \in [-\pi,\pi],$$

derive Euler's identity

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

5.5 Miscellaneous Exercises

Exercise 5.76 Determine the projection in ℓ^2 of the sequence $(\frac{1}{n!})_{n\geq 1}$ onto the subspace M defined by:

$$M = \left\{ \alpha\left(\frac{1}{2^n}\right)_{n\geq 1} + \beta\left(\frac{1}{3^n}\right)_{n\geq 1} \,\middle|\, \alpha,\ \beta \in \mathbb{R} \right\}.$$

Exercise 5.77 Let M be the subspace of ℓ^2 defined by

$$M := \left\{ \left(\frac{x_n}{n} \right)_n \middle| (x_n)_n \in \ell^2 \right\}.$$

Show that M is dense in ℓ^2 but $M \neq \ell^2$.

Exercise 5.78 Consider the subspace of $L^2(-\pi, \pi)$ defined by:

$$M = \{a + b \sin x + cx^2 \mid a, b, c \in \mathbb{R}\}.$$

1. Find an orthonormal basis of M.
2. Compute

$$\min_{f \in M} \int_{-\pi}^{\pi} |x - f|^2 \, dx.$$

Exercise 5.79 Compute

$$\min_{a,b \in \mathbb{R}} \int_1^{\infty} \left| \frac{1}{x^3} - \frac{a}{x} - \frac{b}{x^2} \right|^2 dx.$$

Exercise 5.80 Let $F : L^2(-1, 1) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^1 (f(x) - f(x-1)) \, dx.$$

Show that F is bounded and compute $\|F\|_*$.

Exercise 5.81 In the Hilbert space $L^2(\mathbb{R})$ consider the set

$$M = \{f \in H \ : \ f(x) = f(-x) \text{ a.e.}\}.$$

1. Show that M is a closed subspace of $L^2(\mathbb{R})$.
2. Show that the orthogonal projection onto M is given by

$$p_M(f)(x) = \frac{f(x) + f(-x)}{2}.$$

Exercise 5.82 Let $(e_n)_n$ be an orthonormal basis in a Hilbert space H.

1. Find all the functionals $F \in H^*$ such that

$$\sum_{n=1}^{\infty} |F(e_n)|^2 < \infty. \tag{5.22}$$

2. Let $F \in H^*$ satisfy (5.22). Find a sufficient assumption on $(e_n)_n$ to ensure that

$$\|F\|_*^2 = \sum_{n=1}^{\infty} |F(e_n)|^2.$$

Exercise 5.83 In the Hilbert space $L^2(0, \infty)$ consider the sequence

$$\phi_n(x) = \begin{cases} 1 & \text{if } x \in [n-1, n] \\ 0 & \text{if } x \in [0, \infty) \setminus [n-1, n] \end{cases} \qquad n \geqslant 1.$$

1. Show that $(\phi_n)_n$ is an orthonormal sequence.
2. Is $(\phi_n)_n$ an orthonormal basis?

Exercise 5.84 Let $(K_n)_n$ be a sequence of closed convex sets in a Hilbert space H such that $K_{n+1} \subset K_n$ and

$$K := \bigcap_{n=1}^{\infty} K_n \neq \emptyset .$$

Let $x \in H$.

1. Show that the sequence $(d_{K_n}(x))_n$ is convergent.
2. Show that $(p_{K_n}(x))_n$ is a Cauchy sequence, and therefore it converges to some $\bar{x} \in H$.

   Hint. Adapt the proof of Theorem 5.23.
3. Show that $d_{K_n}(x) \to d_K(x)$ and $\bar{x} = p_K(x)$.

Exercise 5.85 Let H be a Hilbert space.

1. A set $K \subseteq X$ is called a *cone* if for all $x \in K$ and $\lambda > 0$ we have that $\lambda x \in K$. Show that a cone K is convex if and only if

$$x, y \in K \text{ and } \lambda, \mu > 0 \implies \lambda x + \mu y \in K.$$

2. Let $K \neq \varnothing$ be a closed convex cone. Show that, for every $x \in X$,

$$p_K(x) = \bar{x} \iff \begin{cases} \bar{x} \in K \\ \langle x - \bar{x}, \bar{x} \rangle = 0 \\ \langle x - \bar{x}, y \rangle \leqslant 0 \quad \forall y \in K \end{cases} .$$

Exercise 5.86 Write the Fourier series of

$$f(x) = \left(\frac{\pi - |x|}{2}\right)^2, \quad x \in [-\pi, \pi].$$

Deduce that

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

Exercise 5.87 Let $f, g \in L^2(-\pi, \pi)$ and let $\hat{f}(n), \hat{g}(n)$ be their Fourier coefficients, respectively.

1. Show that $\hat{f}(n) \to 0$ as $n \to +\infty$.
2. Show that the following series are convergent

$$\sum_{n=0}^{\infty} \frac{\hat{f}(n)}{1+n}, \qquad \sum_{n=0}^{\infty} \hat{f}(n)\hat{g}(n).$$

References

[Br83] Brezis, H.: Analyse fonctionelle Théorie et applications. Masson, Paris (1983)
[Co90] Conway, J.B.: A Course in Functional Analysis, 2nd edn. Springer, New York (1990)
[Da73] Day, M.M.: Normed Linear Spaces. Springer, New York (1973)
[Ka76] Katznelson, Y.: An Introduction to Harmonic Analysis. Dover Books on Advanced Mathematics, New York (1976)
[Ko02] Komornik, V.: Précis d'analyse réelle. Analyse fonctionnelle, intégrale de Lebesgue, espaces fonctionnels (volume 2), Mathématiques pour le 2^e cycle, Ellipses, Paris (2002)
[Mo71] Mozzochi, C.J.: On the Pointwise Convergence of Fourier Series. Springer, Berlin (1971)
[Ro68] Royden, H.L.: Real Analysis. Macmillan, London (1968)
[Ru73] Rudin, W.: Functional Analysis. McGraw Hill, New York (1973)
[Ru74] Rudin, W.: Real and Complex Analysis. McGraw-Hill, New York (1986)
[Sh61] Shilov, G.E.: An Introduction to the Theory of Linear Spaces. Prentice-Hall Inc., Englewood Cliffs (1961)
[Yo65] Yosida, K.: Functional Analysis. Springer, Berlin (1980)

Chapter 6
Banach Spaces

In the previous chapter, we have seen how to associate a norm $\|\cdot\|$ with a scalar product $\langle\cdot,\cdot\rangle$ on a pre-Hilbert space H. We are now going to take a closer look at those vector spaces X that possess a norm $\|\cdot\|$, hence a metric, which is not necessarily associated with a scalar product. Such an extension is extremely useful because it allows for application to numerous examples of great relevance, such as $L^p(X,\mu)$ spaces with $p \neq 2$ or spaces of continuous functions.

Soon after the first definitions we shall introduce the notion of Banach space, that is, a normed space which is complete with respect to the associated metric. Then, we will study the space of all bounded linear maps between Banach spaces. Such a space enjoys important metric and topological properties, mostly discovered in the first half of the nineteenth century, that can ultimately be regarded as consequences of Baire's Lemma. Then, we will investigate the possibility of extending a bounded linear functional on a subspace to the whole space X via the Hahn-Banach Theorem, which has interesting geometric applications to the separation of convex sets. Finally, we will analyse the Bolzano-Weierstrass property in infinite dimension, which will lead us to introduce the notions of weak convergence and reflexive space.

Most of the examples in this chapter require the use of spaces of summable functions. On the other hand, all these examples make sense in the special case of ℓ^p spaces which can be treated without any knowledge of integration theory. In order to simplify the exposition, we shall often prove technical results in the latter special case. For instance, in this chapter we characterize the dual space of ℓ^p. The proof of the analogous characterization for $L^p(X,\mu)$ spaces needs a refined methodology and will be discussed in Sect. 8.4.

Once again, here we consider real Banach spaces only, even though most of the results of this chapter hold true for vector spaces over $\mathbb{C}$.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_6

6.1 Definitions and Examples

Let X be a linear space over $\mathbb{R}$.

Definition 6.1 A *norm* $\|\cdot\|$ on X is a map

$$\|\cdot\| : X \to [0, \infty)$$

with the following properties:

1. $\|x\| = 0$ if and only if $x = 0$.
2. $\|\alpha x\| = |\alpha|\|x\|$ for every $x \in X$ and $\alpha \in \mathbb{R}$.
3. $\|x + y\| \le \|x\| + \|y\|$ for every $x, y \in X$.

The pair $(X, \|\cdot\|)$ is called a *normed linear space*.

A function $X \to [0, \infty)$ satisfying the above properties except for 1 is called a *seminorm* on X.

As we already observed in Chap. 5, in a normed linear space $(X, \|\cdot\|)$ the function

$$d(x, y) = \|x - y\| \qquad \forall x, y \in X \tag{6.1}$$

is a metric.

Definition 6.2 Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on a linear space X are said to be *equivalent* if there exist two constants $C \ge c > 0$ such that

$$c\|x\|_1 \le \|x\|_2 \le C\|x\|_1 \qquad \forall x \in X.$$

Exercise 6.3 Given a linear space X, show that two norms on X are equivalent if and only if they induce the same topology in X.

Exercise 6.4 In $\mathbb{R}^N$, show that the following norms are equivalent

$$\|x\|_p = \left(\sum_{k=1}^{N} |x_k|^p\right)^{1/p} \qquad \text{and} \qquad \|x\|_\infty = \max_{1\le k\le N} |x_k|,$$

where $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and $p \ge 1$.

Definition 6.5 A normed linear space $(X, \|\cdot\|)$ is called a *Banach space* if it is complete with respect to the metric defined in (6.1).

Example 6.6 1. Any Hilbert space is a Banach space.

2. Given a set $S \neq \varnothing$, the family $\mathrm{B}(S)$ of all bounded functions $f : S \to \mathbb{R}$ is a linear space with the usual operations of sum and product defined by

$$\forall x \in S \quad \begin{cases} (f+g)(x) = f(x) + g(x), \\ (\alpha f)(x) = \alpha f(x), \end{cases}$$

for any $f, g \in \mathrm{B}(S)$ and $\alpha \in \mathbb{R}$. Moreover, $\mathrm{B}(S)$, equipped with the *uniform norm*

$$\|f\|_\infty = \sup_{x \in S} |f(x)| \qquad \forall f \in B(S),$$

is a Banach space (see, for instance, [Fl77, Proposition 2.13]).

3. Let (X, d) be a metric space. The family $\mathscr{C}_b(X)$ of all bounded continuous functions $f : X \to \mathbb{R}$ is a closed subspace of $\mathrm{B}(X)$. So $\big(\mathscr{C}_b(X), \|\cdot\|_\infty\big)$ is a Banach space.
4. Let $(X, \mathscr{E}, \mu)$ be a measure space. The spaces $L^p(X, \mu)$, with $1 \le p \le \infty$, introduced in Chap. 3 are some of the main examples of Banach spaces with the norm

$$\|f\|_p = \left(\int_X |f|^p d\mu\right)^{1/p} \qquad \forall f \in L^p(X, \mu), \quad 1 \le p < \infty$$

and

$$\|f\|_\infty = \inf\{m \ge 0 \mid \mu(|f| > m) = 0\} \qquad \forall f \in L^\infty(X, \mu).$$

We recall that, if $\mu^\#$ is the counting measure on $\mathbb{N}$, we will use the symbol ℓ^p to denote the space $L^p(\mathbb{N}, \mu^\#)$. In this case we have

$$\|x\|_p = \left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p} \qquad \forall x = (x_n)_n \in \ell^p, \quad 1 \le p < \infty$$

and

$$\|x\|_\infty = \sup_{n \ge 1} |x_n|, \qquad \forall x = (x_n)_n \in \ell^\infty.$$

The case $p = 2$ was studied in Chap. 5.

Exercise 6.7 1. Let (X, d) be a locally compact metric space. Show that the set $\mathscr{C}_0(X)$, consisting of all functions $f \in \mathscr{C}_b(X)$ such that for any $\varepsilon > 0$ the set $\big\{x \in X \mid |f(x)| \ge \varepsilon\big\}$ is compact, is a closed subspace of $\mathscr{C}_b(X)$ (so it is a Banach space).
Hint. Observe that, if $f_n \in \mathscr{C}_0(X)$ and $f_n \to f$ in $\mathscr{C}_b(X)$, for large n we have

$$\big\{x \in X \mid |f(x)| \ge \varepsilon\big\} \subset \big\{x \in X \mid |f_n(x)| \ge \varepsilon/2\big\}.$$

2. Show that the set

$$c_0 := \big\{(x_n)_n \in \ell^\infty \mid \lim_{n \to \infty} x_n = 0\big\} \tag{6.2}$$

is a closed subspace of ℓ^∞ (so it is a Banach space).

3. Show that the uniform norm $\|\cdot\|_\infty$ (in $B(S)$, $\mathscr{C}_b(M)$ or ℓ^∞) is not induced by a scalar product.
   Hint. Use parallelogram identity (5.5).

From now on we will often use the following notation: given a normed linear space X and $(x_n)_n \subset X$, $x \in X$, we will write

$$x_n \xrightarrow{X} x,$$

or, simply, $x_n \to x$, to mean that $(x_n)_n$ converges to x with respect to the metric (6.1), that is, $\|x_n - x\| \to 0$ (as $n \to \infty$). Observe that, thanks to the well-known inequality $|\,\|x\| - \|y\|\,| \le \|x - y\|$—which is a consequence of the triangle property of the norm—it follows that

$$x_n \xrightarrow{X} x \quad \Longrightarrow \quad \|x_n\| \longrightarrow \|x\|.$$

Exercise 6.8 In a Banach space X, let $(x_n)_n$ be a sequence such that $\sum_{n=1}^\infty \|x_n\| < \infty$. Show that the series $\sum_{n=1}^\infty x_n$ is convergent in X, that is, there exists $x \in X$ such that

$$\sum_{k=1}^n x_k \xrightarrow{X} x \quad \text{as} \quad n \to \infty.$$

Moreover,

$$\|x\| \le \sum_{n=1}^\infty \|x_n\|.$$

Hint. By property 3 of Definition 6.1 we deduce that

$$\left\| \sum_{k=n+1}^{n+p} x_k \right\| \le \sum_{k=n+1}^{n+p} \|x_k\| \qquad p = 1, 2, \dots,$$

by which it follows that the sequence of partial sums $(\sum_{k=1}^n x_k)_n$ is a Cauchy sequence.

6.2 Bounded Linear Operators

Let X, Y be two linear spaces. A *linear operator* from X to Y is a linear map $\Lambda : X \to Y$. If $Y = \mathbb{R}$, Λ is also called a *linear functional*.

In the following we will always consider normed linear spaces with their respective norms. To simplify notation, when there is no danger of confusion, we will denote each norm with the same symbol $\|\cdot\|$, by dropping the reference to the associated space.

Definition 6.9 Given two normed linear spaces X and Y, a linear operator $\Lambda : X \to Y$ is said to be *bounded* if there exists $C \geq 0$ such that

$$\|\Lambda x\| \leq C\|x\| \quad \forall x \in X.$$

The space of bounded linear operators from X to Y is denoted by $\mathscr{L}(X, Y)$. If $X = Y$, we will write $\mathscr{L}(X, X) = \mathscr{L}(X)$. If $Y = \mathbb{R}$, as in the Hilbert space case, $\mathscr{L}(X, \mathbb{R})$ is called the *topological dual* of X and is denoted by X^*. The elements of X^* are called *bounded linear functionals*.

Arguing exactly as in the proof of Proposition 5.40, one can prove the following result.

Proposition 6.10 *Given two normed linear spaces X, Y and a linear operator $\Lambda : X \to Y$, then the following properties are equivalent:*

(a) *Λ is continuous.*
(b) *Λ is continuous at* 0.
(c) *Λ is continuous at some point.*
(d) *Λ is bounded.*

As in Definition 5.41, let us set

$$\|\Lambda\| = \sup_{\|x\| \leq 1} \|\Lambda x\| \qquad \forall \Lambda \in \mathscr{L}(X, Y). \tag{6.3}$$

Then for any $\Lambda \in \mathscr{L}(X, Y)$, we have

$$\begin{aligned} \|\Lambda\| &= \min\left\{C \geq 0 \,\Big|\, \|\Lambda x\| \leq C\|x\| \ \ \forall x \in X\right\} \\ &= \sup_{\|x\|=1} \|\Lambda x\| = \sup_{\|x\|<1} \|\Lambda x\| = \sup_{x \neq 0} \frac{\|\Lambda x\|}{\|x\|} \end{aligned} \tag{6.4}$$

(see also Exercise 5.42). If $Y = \mathbb{R}$, (6.3) is called the *dual norm* and is also denoted by $\|\cdot\|_*$.

Exercise 6.11 Show that (6.3) is a norm on $\mathscr{L}(X, Y)$.

Proposition 6.12 *Let X, Y be two normed linear spaces. If Y is a Banach space, then $\mathscr{L}(X, Y)$ is also a Banach space. In particular, the topological dual X^* is a Banach space.*

Proof Let $(\Lambda_n)_n$ be a Cauchy sequence in $\mathscr{L}(X, Y)$. For every $x \in X$, since $\|\Lambda_n x - \Lambda_m x\| \leq \|\Lambda_n - \Lambda_m\| \, \|x\|$, we deduce that $(\Lambda_n x)_n$ is a Cauchy sequence in Y. Since Y is complete, then $(\Lambda_n x)_n$ converges to a point in Y that we label Λx. We have thus defined a mapping $\Lambda : X \to Y$. It is immediate to check that Λ is linear. Moreover, since $(\Lambda_n)_n$ is a bounded sequence in $\mathscr{L}(X, Y)$, say $\|\Lambda_n\| \leq M$ for every $n \in \mathbb{N}$, then

$$\|\Lambda_n x\| \leq M\|x\| \quad \forall n \in \mathbb{N}, \ \forall x \in X,$$

by which, taking the limit as $n \to \infty$, we have that $\|\Lambda x\| \leq M\|x\|$ for every $x \in X$. So $\Lambda \in \mathscr{L}(X, Y)$ and $\|\Lambda\| \leq M$. Finally, to show that $\Lambda_n \to \Lambda$ in $\mathscr{L}(X, Y)$, fix $\varepsilon > 0$ and choose $n_\varepsilon \in \mathbb{N}$ such that $\|\Lambda_n - \Lambda_m\| < \varepsilon$ for all $n, m \geq n_\varepsilon$. Then $\|\Lambda_n x - \Lambda_m x\| < \varepsilon\|x\|$ for every $x \in X$. Taking the limit as $m \to \infty$, we obtain $\|\Lambda_n x - \Lambda x\| \leq \varepsilon\|x\|$ for every $x \in X$. Hence, $\|\Lambda_n - \Lambda\| \leq \varepsilon$ for all $n \geq n_\varepsilon$ and the proof is complete. □

Exercise 6.13 1. Given a continuous function $f : [a, b] \to \mathbb{R}$, define[1] $\Lambda : L^1(a, b) \to L^1(a, b)$ by setting

$$\Lambda g(t) = f(t)g(t), \qquad t \in [a, b].$$

Show that Λ is a bounded linear operator and $\|\Lambda\| = \|f\|_\infty$.
Hint. By Exercise 3.26 it follows that $\|\Lambda\| \leq \|f\|_\infty$; to prove the equality, suppose $|f(x)| > \|f\|_\infty - \varepsilon$ for all $x \in [x_0, x_1] \subset [a, b]$ and let $g(x) = \chi_{[x_0,x_1]}$ be the characteristic function of the interval $[x_0, x_1]$; then estimate $\|\Lambda g\|_1$.

2. Let $\Lambda : \mathscr{C}([-1, 1]) \to \mathbb{R}$ be the linear functional defined by

$$\Lambda f = \int_{-1}^{1} f(x)\,\mathrm{sign}x\,dx.$$

Show that Λ is bounded and $\|\Lambda\|_* = 2$.
Hint. Consider the sequence $(f_n)_n$ of Exercise 5.12(2) and estimate Λf_n as $n \to \infty$.

Exercise 6.14 Let X be a Banach space.

1. Show that if $\Lambda, \Lambda' \in \mathscr{L}(X)$, then $\Lambda\Lambda' := \Lambda \circ \Lambda' \in \mathscr{L}(X)$ and $\|\Lambda\Lambda'\| \leq \|\Lambda\|\,\|\Lambda'\|$.
2. Show that if $\Lambda \in \mathscr{L}(X)$ satisfies $\|\Lambda\| < 1$, then $I - \Lambda$ is invertible and $(I - \Lambda)^{-1} \in \mathscr{L}(X)$.
Hint. Show that $(I - \Lambda)^{-1} = \sum_{n=0}^{\infty} \Lambda^n$ (for $n = 0$ set $\Lambda^0 = I$).
3. Show that the set of invertible operators $\Lambda \in \mathscr{L}(X)$ such that Λ^{-1} is continuous is open in $\mathscr{L}(X)$.
Hint. Observe that if $\Lambda_0^{-1} \in \mathscr{L}(X)$, then for every $\Lambda \in \mathscr{L}(X)$ such that $\|\Lambda - \Lambda_0\| < 1/\|\Lambda_0^{-1}\|$ we have that $\Lambda^{-1} = [I + \Lambda_0^{-1}(\Lambda - \Lambda_0)]^{-1}\Lambda_0^{-1}$.

If X is a normed linear space and $x_0 \in X$, in the following we will denote by $B_r(x_0)$ the open ball with center x_0 and radius $r > 0$, i.e.,

[1] $L^p(a, b) = L^p([a, b], m)$ where m is the Lebesgue measure on $[a, b]$. See footnote 7, p. 87.

$$B_r(x_0) = \{x \in X \mid \|x - x_0\| < r\},$$

whereas we will denote by $\overline{B}_r(x_0)$ the closed ball

$$\overline{B}_r(x_0) = \{x \in X \mid \|x - x_0\| \le r\}.$$

If $x_0 = 0$, we will write $B_r = B_r(0)$ and $\overline{B}_r = \overline{B}_r(0)$.

Exercise 6.15 Show that, in a normed linear space X, we have $\overline{B}_r(x) = \overline{B_r(x)}$ for every $r > 0$ and $x \in X$ (in contrast to what happens in a generic metric space, see Remark D.2).

Exercise 6.16 Let X be a normed linear space and $Y \subset X$ a subspace. Denote by X/Y the quotient space of X relative to Y and by Q the quotient map

$$Q : X \to X/Y,$$
$$Qx = x + Y.$$

Show that the map $\|\cdot\| : X/Y \to \mathbb{R}$ defined by

$$\|Qx\| = d_Y(x) = \inf_{y \in Y} \|x + y\| \tag{6.5}$$

is a seminorm on X/Y, where $d_Y(x)$ is the distance of x_0 from Y (see Appendix A). Under the additional assumption that Y is also closed, show that:

1. (6.5) is a norm on X/Y.
2. $\|Qx\| \le \|x\|$ for all $x \in X$ (so Q is continuous).
3. $W \subset X/Y$ open $\Longrightarrow Q^{-1}W$ open in X.
4. $U \subset X$ open $\Longrightarrow QU$ open in X/Y.
5. X Banach $\Longrightarrow X/Y$ Banach.

Finally, show that (6.5) fails to be a norm if Y is not closed.

Hint. To prove part 5, let $(Qx_n)_n$ be a Cauchy sequence in X/Y, that is, for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that

$$n, m \ge n_\varepsilon \Longrightarrow \|Qx_n - Qx_m\| = \inf_{y \in Y} \|x_n - x_m + y\| < \varepsilon.$$

Construct a subsequence $(x_{n_k})_k$ and a sequence $(y_k)_k \subset Y$ verifying

$$\|x_{n_k} + y_k - x_{n_{k+1}} - y_{k+1}\| \le \frac{1}{2^k} \quad \forall k \in \mathbb{N}.$$

So $(x_{n_k} + y_k)_k$ is a Cauchy sequence in X, hence it converges to some $x \in X$. By part 2 we get $Qx_{n_k} = Q(x_{n_k} + y_k) \to Qx$ in X/Y.

Exercise 6.17 Let H be a Hilbert space and M a closed subspace of H. Show that the quotient map $Q : H \to H/M$, restricted to $M^{\perp}$, becomes an isometric isomorphism.[2]

Example 6.18 (Volterra operator) Let $1 \le p \le \infty$ and $T > 0$ and for any $f \in L^p(0, T)$ set

$$V_p f(t) = \int_0^t f(s)ds \quad t \in (0, T).$$

1. Consider $1 \le p < \infty$. Denoting by p' the conjugate exponent of p, we have

$$|V_p f(t)| \le t^{1/p'} \Big(\int_0^t f(s)ds \Big)^{1/p}$$

with the convention $\frac{1}{\infty} = 0$. Thus,

$$\|V_p f\|_p^p \le \|f\|_p^p \int_0^T t^{p/p'} dt = \frac{T^p}{p} \|f\|_p^p.$$

So

$$V_p \in \mathscr{L}(L^p(0, T)) \text{ and } \|V_p\| \le \frac{T}{p^{1/p}}. \tag{6.6}$$

2. Consider $p = 1$. We claim that $\|V_1\| = T$.
Indeed, for any $n \in \mathbb{N}$ set $f_n = n\chi_{(0, \frac{1}{n})}$, which satisfies $\|f_n\|_1 = 1$. Then

$$V_1 f_n(t) = \min\{nt, 1\},$$

which implies

$$\|V_1\| \ge \|V_1 f_n\|_1 = T - \frac{1}{2n} \to T.$$

Combining this with (6.6) we get $\|V_1\| = T$ as claimed.
3. Consider $p = \infty$. Then

$$|V_\infty f(t)| \le \int_0^t |f(s)|ds \le t\|f\|_\infty.$$

Therefore

$$V_\infty \in \mathscr{L}(L^p(0, T)) \text{ and } \|V_\infty\| \le T.$$

Moreover, taking $f = 1$ yields $\|f\|_\infty = 1$ and $\|V_\infty\| \ge \|V_\infty f\|_\infty = T$. So

$$\|V_\infty\| = T.$$

[2]See footnote 5 at p. 147.

4. Consider $p = 2$ and set $V = V_2 \in \mathscr{L}(L^2(0,T))$. Then we have

$$\|V\|^2 = \sup_{0\neq f\in L^2(0,T)} \frac{\|Vf\|_2^2}{\|f\|_2^2} = \Big(\inf_{0\neq f\in L^2(0,T)} \frac{\|f\|_2^2}{\|Vf\|_2^2} \Big)^{-1}. \tag{6.7}$$

Let $R(V)$ be the range of V, which is given by the following subset of the space $AC(0,T)$ of absolutely continuous functions (see Chap. 7):

$$R(V) = \{u \in AC(0,T) \,|\, u(0) = 0,\ u' \in L^2(0,T)\}.$$

Hence

$$\inf_{0\neq f\in L^2(0,T)} \frac{\|f\|_2^2}{\|Vf\|_2^2} = \inf_{0\neq u\in R(V)} \frac{\|u'\|_2^2}{\|u\|_2^2}. \tag{6.8}$$

Suppose $T = \frac{\pi}{2}$ and for any $u \in R(V)$ consider the extension to $[0,\pi]$ by symmetry (i.e., $u(t) = u(\pi - t)$ for $t \in [\frac{\pi}{2}, \pi]$) and then the odd extension to $[-\pi,\pi]$. If we label by $\bar{u}$ the resulting extension to $[-\pi,\pi]$, then $\bar{u}$ satisfies $\bar{u}(0) = \bar{u}(-\pi) = \bar{u}(\pi) = 0$ and, denoting by $(a_k)_k$ and $(a')_k$ the Fourier coefficients of $\bar{u}$ and $\bar{u}'$, respectively, with respect to the trigonometric system (see Example 5.56), a direct computation gives

$$a_0 = a_0' = 0$$

$$a_{2k}' = k a_{2k-1}, \quad a_{2k-1}' = a_{2k} = 0 \quad \forall k \geq 1.$$

So Parseval's identity yields

$$\begin{aligned} 4\int_0^{\pi/2} |u(t)|^2 dt &= \int_{-\pi}^{\pi} |\bar{u}(t)|^2 dt = \sum_{k=1}^{\infty} a_{2k-1}^2 \leq \sum_{k=1}^{\infty} k^2 a_{2k-1}^2 \\ &= \int_{-\pi}^{\pi} |\bar{u}'(t)|^2 dt = 4\int_0^{\pi/2} |u'(t)|^2 dt. \end{aligned}$$

We deduce that $\|u'\|_2^2 \geq \|u\|_2^2$ for any $u \in R(V)$; on the other hand, by taking $u(t) = \sin t \in R(V)$, we obtain $\|u\|_2^2 = \|u'\|_2^2$, and so

$$\inf_{0\neq u\in R(V)} \frac{\|u'\|_2^2}{\|u\|_2^2} = 1.$$

Using (6.7) and (6.8), we conclude

$$\|V\| = 1 \quad \text{if } T = \frac{\pi}{2}.$$

In the general case $T > 0$, by an easy rescaling argument we get

$$\inf_{0\neq u\in R(V)} \frac{\|u'\|_2^2}{\|u\|_2^2} = \frac{\pi^2}{(2T)^2}$$

whence $\|V\| = \frac{2T}{\pi} < \frac{T}{\sqrt{2}}$.

6.2.1 The Principle of Uniform Boundedness

Our next result, usually ascribed to Banach and Steinhaus even though it was obtained by various authors in different formulations, is also known as *Principle of Uniform Boundedness*. Indeed, it allows to deduce uniform estimates for a family of bounded linear operators starting from pointwise estimates.

Theorem 6.19 (Banach-Steinhaus) *Let X be a Banach space, let Y be a normed linear space, and let $(\Lambda_i)_{i\in I} \subset \mathscr{L}(X, Y)$. Then*

either there exists $M \geq 0$ such that

$$\|\Lambda_i\| \leq M \qquad \forall i \in I, \tag{6.9}$$

or there exists a dense set $D \subset X$ such that

$$\sup_{i\in I} \|\Lambda_i x\| = \infty \qquad \forall x \in D. \tag{6.10}$$

Proof Define

$$\alpha(x) := \sup_{i\in I} \|\Lambda_i x\| \qquad \forall x \in X.$$

Since $\alpha : X \to [0, \infty]$ is a lower semicontinuous function (see Corollary B.6), for any $n \in \mathbb{N}$

$$V_n := \{x \in X \mid \alpha(x) > n\} \tag{6.11}$$

is an open set in X (see Theorem B.4). If all sets V_n are dense, then (6.10) holds on $D := \cap_{n=1}^{\infty} V_n$ and D is, in turn, a dense set owing to Baire's Lemma (see Proposition D.1). Now, suppose that one of these sets, say V_N, fails to be dense in X. Then there exists a closed ball $\overline{B}_r(x_0) \subset X\backslash\overline{V}_N$. Therefore

$$\|x\| \leq r \quad \Longrightarrow \quad x_0 + x \notin V_N \quad \Longrightarrow \quad \alpha(x_0 + x) \leq N.$$

Consequently, $\|\Lambda_i x\| \leq \|\Lambda_i x_0\| + \|\Lambda_i(x + x_0)\| \leq 2N$ for all $i \in I$ and $\|x\| \leq r$. So, for every $i \in I$,

$$\|\Lambda_i x\| = \frac{\|x\|}{r}\left\|\Lambda_i \frac{rx}{\|x\|}\right\| \le \frac{2N}{r}\|x\| \qquad \forall x \in X\setminus\{0\},$$

which yields (6.9) with $M = 2N/r$. □

Exercise 6.20 Give a direct proof (that is, a proof based only on the definition of the function α) of the fact that the sets V_n in (6.11) are open.

Corollary 6.21 *Let X be a Banach space, let Y be a normed linear space and let $(\Lambda_n)_n \subset \mathscr{L}(X,Y)$ be such that, for every $x \in X$, the sequence $(\Lambda_n x)_n$ is convergent. Then, setting $\Lambda x := \lim_{n\to\infty} \Lambda_n x$ for every $x \in X$, we have that $\Lambda \in \mathscr{L}(X,Y)$ and*

$$\|\Lambda\| \le \liminf_{n\to\infty} \|\Lambda_n\| < \infty\,.$$

Proof The Banach-Steinhaus Theorem ensures that

$$\sup_{n\in\mathbb{N}} \|\Lambda_n\| = M < \infty\,.$$

So $\liminf_{n\to\infty} \|\Lambda_n\| < \infty$. Moreover, for any $n \in \mathbb{N}$ we have that

$$\|\Lambda_n x\| \le M\|x\| \quad \forall x \in X.$$

Thus, taking the limit as $n \to \infty$, we obtain

$$\|\Lambda x\| \le M\|x\| \quad \forall x \in X\,.$$

Therefore, since it is immediate to verify that Λ is linear, we get $\Lambda \in \mathscr{L}(X,Y)$. Finally, taking the lim inf in the inequality $\|\Lambda_n x\| \le \|\Lambda_n\|\,\|x\|$, we deduce that

$$\|\Lambda x\| \le \liminf_{n\to\infty} \|\Lambda_n\|\,\|x\| \quad \forall x \in X,$$

which completes the proof. □

Exercise 6.22 Let $x = (x_n)_n \subset \mathbb{R}$ and let $1 \le p, p' \le \infty$ be conjugate exponents.[3] Show that if the series $\sum_{n=1}^{\infty} x_n y_n$ is convergent for every $y = (y_n)_n \in \ell^{p'}$, then $x \in \ell^p$.

Hint. Set

$$\Lambda_n : \ell^{p'} \to \mathbb{R}, \quad \Lambda_n y = \sum_{k=1}^{n} x_k y_k\,.$$

Show that $\Lambda_n \in (\ell^{p'})^*$, $\|\Lambda_n\|_* = (\sum_{k=1}^{n} |x_k|^p)^{1/p}$ and $\Lambda_n y \to \sum_{k=1}^{\infty} x_k y_k$ for every $y \in \ell^{p'}$. Then use Corollary 6.21.

[3]Two numbers $1 \le p, p' \le \infty$ are said to be *conjugate* if $\frac{1}{p} + \frac{1}{p'} = 1$, with the convention $\frac{1}{\infty} = 0$.

Exercise 6.23 [4] Given a σ-finite measure space $(X, \mathscr{E}, \mu)$, let $1 \le p, p' \le \infty$ be conjugate exponents and let $f : X \to \overline{\mathbb{R}}$ be a Borel function such that $fg \in L^1(X, \mu)$ for every $g \in L^{p'}(X, \mu)$. Show that $f \in L^p(X, \mu)$.

Hint. Let $(X_n)_n \subset \mathscr{E}$ be an increasing sequence such that $\mu(X_n) < \infty$ and $X_n \uparrow X$, and set

$$\Lambda_n : L^{p'}(X, \mu) \to \mathbb{R}, \quad \Lambda_n g = \int_{X_n} f \chi_{\{|f| \le n\}} g \, d\mu.$$

Show that $\Lambda_n \in (L^{p'}(X, \mu))^*$, $\|\Lambda_n\|_* = \|f \chi_{X_n \cap \{|f| \le n\}}\|_p$ and $\Lambda_n g \to \int_X f g \, d\mu$ for every $g \in L^{p'}(X, \mu)$. Then use Corollary 6.21.

6.2.2 The Open Mapping Theorem

Bounded linear operators between two Banach spaces enjoy topological properties—closely related one another—that are very useful for applications, for instance, to differential equations. The first and most relevant of these results is the so-called Open Mapping Theorem.

Theorem 6.24 (Schauder) *Let X, Y be Banach spaces and let $\Lambda \in \mathscr{L}(X, Y)$ be onto. Then Λ is an* open *mapping.*[5]

Proof We split the argument into four steps.

1. Let us show that there exists a radius $r > 0$ such that

$$B_{2r} \subset \overline{\Lambda(B_1)}. \tag{6.12}$$

Observe that, since Λ is onto, we have

$$Y = \bigcup_{k=1}^{\infty} \overline{\Lambda(B_k)}.$$

Therefore, by Proposition D.1 (Baire's Lemma), at least one of the closed sets $\{\overline{\Lambda(B_k)}\}_k$ has a nonempty interior, and therefore it contains a ball, say $B_s(y) \subset \overline{\Lambda(B_k)}$. Since $\Lambda(B_k)$ is a symmetric set with respect to the origin 0, we deduce that

$$B_s(-y) \subset -\overline{\Lambda(B_k)} = \overline{\Lambda(B_k)}$$

[4] Compare with Exercise 3.9.

[5] That is, Λ maps open sets of X into open sets of Y.

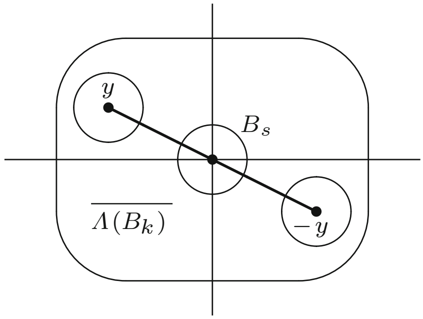


Fig. 6.1 The Open Mapping Theorem

(see Fig. 6.1). Consequently, for every $y' \in B_s$, we have $y' \pm y \in B_s(\pm y) \subset \overline{\Lambda(B_k)}$. Since $\overline{\Lambda(B_k)}$ is a convex set, we conclude

$$y' = \frac{(y'+y)+(y'-y)}{2} \in \overline{\Lambda(B_k)}.$$

Thus, $B_s \subset \overline{\Lambda(B_k)}$. Equation (6.12) follows by taking $r = s/2k$ and then rescaling. The argument goes as follows: let $z \in B_{2r} = B_{s/k}$; then $kz \in B_s$ and a sequence $(x_n)_n \subset B_k$ exists such that $\Lambda x_n \to kz$. So $x_n/k \in B_1$ and $\Lambda(x_n/k) \to z$, by which $z \in \overline{\Lambda(B_1)}$.

2. Observe that, by linearity, (6.12) yields the family of inclusions

$$B_{2^{1-n}r} \subset \overline{\Lambda(B_{2^{-n}})} \qquad \forall n \in \mathbb{N}. \tag{6.13}$$

3. We now proceed to show that

$$B_r \subset \Lambda(B_1). \tag{6.14}$$

Let $y \in B_r$. Applying (6.13) with $n = 1$, we can find a point

$$x_1 \in B_{2^{-1}} \quad \text{such that} \quad \left\|y - \Lambda x_1\right\| < \frac{r}{2}.$$

Thus, $y - \Lambda x_1 \in B_{2^{-1}r}$. Then, applying (6.13) with $n = 2$ we find a second point

$$x_2 \in B_{2^{-2}} \quad \text{such that} \quad \left\|y - \Lambda(x_1 + x_2)\right\| < \frac{r}{2^2}.$$

Iterating the above procedure gives a sequence $(x_n)_n$ in X such that

$$x_n \in B_{2^{-n}} \quad \text{and} \quad \left\|y - \Lambda(x_1 + \cdots + x_n)\right\| < \frac{r}{2^n}. \tag{6.15}$$

Since

$$\sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1,$$

recalling Exercise 6.8 we conclude that the series $\sum_{n=1}^{\infty} x_n$ converges to some point $x \in X$, that is, $\sum_{k=1}^{n} x_k \xrightarrow{X} x$. Moreover, $\|x\| \leq \sum_{n=1}^{\infty} \|x_n\| < 1$. By the continuity of Λ we have $\sum_{k=1}^{n} \Lambda x_k \xrightarrow{Y} \Lambda x$. On the other hand, by (6.15), $\sum_{k=1}^{n} \Lambda x_k \xrightarrow{Y} y$; then we get $y = \Lambda x \in \Lambda B_1$, and this proves (6.14).

4. Let $U \subset X$ be an open set and let $x \in U$. Then there exists $\rho > 0$ such that $B_\rho(x) \subset U$, whence $\Lambda x + \Lambda(B_\rho) \subset \Lambda(U)$. Therefore

$$B_{r\rho}(\Lambda x) = \underbrace{\Lambda x + B_{r\rho} \subset \Lambda x + \Lambda(B_\rho)}_{\text{by (6.14)}} \subset \Lambda(U).$$

This implies that $\Lambda(U)$ is an open set in Y.

The proof is thus complete. □

A first consequence of the above result is the following corollary, known as Inverse Mapping Theorem.

Corollary 6.25 (Banach) *Let X, Y be Banach spaces and let $\Lambda \in \mathscr{L}(X, Y)$ be bijective. Then $\Lambda^{-1} \in \mathscr{L}(Y, X)$. Consequently, the Banach spaces X and Y are isomorphic.*

Proof It is immediate that Λ^{-1} is linear. Moreover, for any open set $U \subset X$, we have that $(\Lambda^{-1})^{-1}(U) = \Lambda(U)$ is an open set in Y owing to the Open Mapping Theorem. It follows that Λ^{-1} is a continuous map, and so $\Lambda^{-1} \in \mathscr{L}(Y, X)$. □

Exercise 6.26 Let X, Y be Banach spaces and let $\Lambda \in \mathscr{L}(X, Y)$ be bijective. Show that there exists a constant $\lambda > 0$ such that

$$\|\Lambda x\| \geq \lambda \|x\| \qquad \forall x \in X.$$

Hint. Use Corollary 6.25 and apply Proposition 6.10 to Λ^{-1}.

Exercise 6.27 Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a linear space X. Suppose that X is a Banach space with respect to both $\|\cdot\|_1$ and $\|\cdot\|_2$. If there exists a constant $c > 0$ such that $\|x\|_2 \leq c\|x\|_1$ for any $x \in X$, then there also exists another constant $C > 0$ such that $\|x\|_1 \leq C\|x\|_2$ for any $x \in X$ (i.e., $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms).
Hint. It is sufficient to apply the result of Exercise 6.26 to the identity map $(X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$.

Example 6.28 (König-Witstock norm) Let X be a Banach space with norm $\|\cdot\|$ and let $f : X \to \mathbb{R}$ be a linear functional which is not bounded (see Example 5.50). We now exhibit a second norm $\|\cdot\|_f$ on X such that X is complete with respect to $\|\cdot\|_f$ but $\|\cdot\|_f$ is not equivalent to $\|\cdot\|$. Indeed, fix $p \in X$ such that $f(p) = 1$ and set

$$M = \mathbb{R}p = \{\lambda p \mid \lambda \in \mathbb{R}\}.$$

Let us define

$$\|x\|_f = |f(x)| + d_M(x) \quad \forall x \in X,$$

where $d_M(x)$ is the distance of x from M.

1. Let us show that $\| \cdot \|_f$ is a norm on X, which is known as König-Witstock norm ([KW92]). Indeed

   (i)
   $$\|x\|_f = 0 \Longleftrightarrow \begin{cases} f(x) = 0 \\ d_M(x) = 0 \end{cases}.$$

   Since M is closed, we get $x \in M$, and so $x = \lambda p$ for some $\lambda \in \mathbb{R}$. Hence, $f(x) = 0 = \lambda$ and $x = 0$.

   (ii) Let $x \in X$ and $\alpha \in \mathbb{R} \setminus \{0\}$. Then
   $$\begin{aligned} \|\alpha x\|_f &= |\alpha|\,|f(x)| + \inf_{\lambda \in \mathbb{R}} \|\alpha x - \lambda p\| \\ &= |\alpha|\Big(|f(x)| + \inf_{\mu \in \mathbb{R}} \|x - \mu p\|\Big) = |\alpha|\,\|x\|_f. \end{aligned}$$

   (iii) Let $x, y \in X$. Given $\varepsilon > 0$, let us choose $\lambda_\varepsilon, \mu_\varepsilon \in \mathbb{R}$ such that
   $$\|x - \lambda_\varepsilon p\| < d_M(x) + \varepsilon, \quad \|y - \mu_\varepsilon p\| < d_M(y) + \varepsilon.$$

   Then
   $$\begin{aligned} \|x + y\|_f &= |f(x + y)| + d_M(x + y) \\ &\le |f(x)| + f(y)| + \|x - \lambda_\varepsilon p\| + \|y - \mu_\varepsilon p\| \\ &< \|x\|_f + \|y\|_f + 2\varepsilon. \end{aligned}$$

   So $\|x + y\|_f \le \|x\|_f + \|y\|_f$.

2. Let us show that X is complete with respect to the norm $\| \cdot \|_f$. Indeed, let $(x_n)_n \subset X$ be a Cauchy sequence with respect to the norm $\| \cdot \|_f$. Then for every $\varepsilon > 0$ there exists n_ε such that
   $$n, m \ge n_\varepsilon \Longrightarrow |f(x_n) - f(x_m)| + d_M(x_n - x_m) < \varepsilon.$$

   It follows that
   $$\begin{aligned} &(f(x_n))_n \text{ is a Cauchy sequence in } \mathbb{R}, \\ &(x_n + M)_n \text{ is a Cauchy sequence in } X/M, \end{aligned}$$

where X/M is the quotient space of X relative to M. Recalling that X/M is a Banach space owing to Exercise 6.16, we deduce that there exist $L \in \mathbb{R}$ and $\bar{x} \in X$ such that

$$f(x_n) \to L \quad \text{in } \mathbb{R},$$
$$x_n + M \to \bar{x} + M \quad \text{in } X/M.$$

Setting $\lambda = L - f(\bar{x})$ we have $L = f(\bar{x} + \lambda p)$. So

$$\|x_n - (\bar{x} + \lambda p)\|_f = |f(x_n) - L| + d_M(x_n - \bar{x}) \to 0.$$

3. Finally, let us show that $\|\cdot\|$ and $\|\cdot\|_f$ are not equivalent. Indeed, assume by contradiction that there exists a constant $C > 0$ such that

$$|f(x)| + d_M(x) = \|x\|_f \le C\|x\|.$$

Then $|f(x)| \le C\|x\|$, and this implies that f is continuous—a contradiction.

To introduce our next result, let us observe that the Cartesian product $X \times Y$ of two normed linear spaces X, Y is naturally equipped with the *product norm*

$$\|(x, y)\| := \|x\| + \|y\| \qquad \forall (x, y) \in X \times Y.$$

Exercise 6.29 Show that if X, Y are Banach spaces, then $\big(X \times Y, \|(\cdot,\cdot)\|\big)$ is also a Banach space.

We conclude with the so-called Closed Graph Theorem.

Corollary 6.30 (Banach) *Let X, Y be Banach spaces and let $\Lambda : X \to Y$ be a linear mapping. Then $\Lambda \in \mathscr{L}(X, Y)$ if and only if the* graph *of Λ, that is, the set*

$$\text{Graph}(\Lambda) := \big\{(x, y) \in X \times Y \;\big|\; y = \Lambda x\big\},$$

is closed in $X \times Y$.

Proof Suppose, first, that $\Lambda \in \mathscr{L}(X, Y)$. Then it is easy to see that

$$\Delta : X \times Y \to Y \qquad \Delta(x, y) = y - \Lambda x$$

is a continuous mapping. Therefore $\text{Graph}(\Lambda) = \Delta^{-1}(0)$ is a closed set.

Conversely, suppose that $\text{Graph}(\Lambda)$ is a closed set in $X \times Y$. Then $\text{Graph}(\Lambda)$ is in turn a Banach space with the product norm, since it is a closed subspace of the Banach space $X \times Y$. Moreover, the linear map

$$\Pi_\Lambda : \text{Graph}(\Lambda) \to X \qquad \Pi_\Lambda(x, \Lambda x) := x$$

is bounded and bijective. Therefore, owing to Corollary 6.25, the map

$$\Pi_\Lambda^{-1} : X \to \mathrm{Graph}(\Lambda) \qquad \Pi_\Lambda^{-1} x = (x, \Lambda x)$$

is continuous; since $\Lambda = \Pi_Y \circ \Pi_\Lambda^{-1}$, where

$$\Pi_Y : X \times Y \to Y \qquad \Pi_Y(x, y) := y,$$

we conclude that Λ is also continuous. □

Example 6.31 Consider the spaces

$$Y = \mathscr{C}([0,1]) = \{f : [0,1] \to \mathbb{R} \mid f \text{continuous}\}$$

and

$$X = \mathscr{C}^1([0,1]) = \{f : [0,1] \to \mathbb{R} \mid f \text{ differentiable and } f' \in \mathscr{C}([0,1])\}$$

both equipped with the norm $\|\cdot\|_\infty$. Define

$$\Lambda f(t) = f'(t) \qquad \forall f \in X,\ \forall t \in [0,1].$$

Then $\mathrm{Graph}(\Lambda)$ is a closed set in $X \times Y$ since

$$\begin{cases} f_n \xrightarrow{L^\infty} f \\ f_n' \xrightarrow{L^\infty} g \end{cases} \implies f \in \mathscr{C}^1([0,1]) \quad \& \quad f' = g.$$

On the other hand Λ fails to be a bounded operator. Indeed, taking

$$f_n(t) = t^n \qquad \forall t \in [0,1],$$

we have

$$f_n \in X, \qquad \|f_n\|_\infty = 1\,, \qquad \|\Lambda f_n\|_\infty = n \qquad \forall n \geq 1.$$

This shows the necessity of X being a Banach space in Corollary 6.30.

Exercise 6.32 Let X, Y be Banach spaces and let $\Lambda \in \mathscr{L}(X, Y)$. Show that the following properties are equivalent:

(a) There exists $c > 0$ such that $\|\Lambda x\| \geq c\|x\|$ for every $x \in X$.
(b) $\ker \Lambda = \{0\}$ and $\Lambda(X)$ is a closed set in Y.

Hint. For the implication (b) ⇒ (a) apply the result of Exercise 6.26 to the bijective operator $x \in X \mapsto \Lambda x \in \Lambda(X)$.

Exercise 6.33 Let H be a Hilbert space with scalar product $\langle\cdot,\cdot\rangle$ and let $A, B : H \to H$ be two linear operators such that

$$\langle Ax, y\rangle = \langle x, By\rangle \qquad \forall x, y \in H. \tag{6.16}$$

Show that[6] $A, B \in \mathscr{L}(H)$.
Hint. Use (6.16) to deduce that Graph(A) and Graph(B) are closed sets in $H \times H$; then apply Corollary 6.30.

Exercise 6.34 Let X be an infinite-dimensional separable Banach space and let $(e_i)_{i\in I}$ be a *Hamel basis*[7] of X such that $\|e_i\| = 1$ for every $i \in I$.

1. Show that I is uncountable.
   Hint. Suppose, by contradiction, $I = \mathbb{N}$ and use Baire's Lemma D.1 taking the closed sets $F_n = \mathbb{R}e_1 + \cdots + \mathbb{R}e_n = \{\sum_{i=1}^n \lambda_i e_i \,|\lambda_i \in \mathbb{R}\}$.
2. Show that the map $\|\cdot\|_1$ defined by

$$\|x\|_1 = \sum_{i\in J} |\lambda_i| \quad \text{if} \quad x = \sum_{i\in J} \lambda_i e_i, \quad J \subset I \text{ finite}$$

   is a norm in X and $\|x\| \le \|x\|_1$ for every $x \in X$.
3. Show that X is not complete with respect to the norm $\|\cdot\|_1$.
   Hint. If $(X, \|\cdot\|_1)$ were a Banach space, then $\|\cdot\|$ and $\|\cdot\|_1$ would be equivalent norms by Exercise 6.27, but, for any $i \neq j$, we have $\|e_i - e_j\|_1 = 2$, and this yields that $(X, \|\cdot\|_1)$ fails to be separable (as in Example 3.27).

6.3 Bounded Linear Functionals

In this section, we shall study a special class of bounded linear operators, namely $\mathbb{R}$-valued operators or—as we usually say—*bounded linear functionals*. We shall see, first, that functionals enjoy an important extension property described by the Hahn-Banach Theorem. Then we will derive useful analytic and geometric consequences of such a property. These results will be essential for the analysis of dual spaces that we shall develop in the next section. Finally, we will characterize the duals of the Banach spaces ℓ^p.

---

[6]This result dates from 1910 (see [Ko02, p.67]).
[7]See footnote 9 at p. 151.

6.3.1 Hahn-Banach Theorem

Consider the following extension problem: given a normed linear space X, a subspace $M \subset X$ (not necessarily closed) and a bounded linear functional $f : M \to \mathbb{R}$,

$$\text{find} \quad F \in X^* \quad \text{such that} \quad \begin{cases} F\big|_M = f, \\ \|F\|_* = \|f\|_* \end{cases} \tag{6.17}$$

(here we have used the same symbol to denote the dual norm of f, which is an element of M^*, and the dual norm of F, which is an element of X^*).

Remark 6.35 Observe that a bounded linear functional f defined on a subspace M can be uniquely extended to the closure $\overline{M}$ by a standard completeness argument. Indeed, let $\overline{x} \in \overline{M}$ and let $(x_n)_n \subset M$ be such that $x_n \to \overline{x}$. Since

$$|f(x_n) - f(x_m)| \leq \|f\|_* \|x_n - x_m\|,$$

$(f(x_n))_n$ is a Cauchy sequence in $\mathbb{R}$. So $(f(x_n))_n$ is convergent. Then it is easy to verify that $F(\overline{x}) := \lim_n f(x_n)$ is the required extension of f and $\|F\|_* = \|f\|_*$. Therefore the problem (6.17) has a *unique solution* when M is dense in X.

Remark 6.36 Problem (6.17) has a unique solution also when X is a Hilbert space. Indeed, let us still denote by f the extension of the given functional to the closure $\overline{M}$, obtained by the procedure described in Remark 6.35. Note that $\overline{M}$ is a Hilbert space. So, by the Riesz Theorem, there exists a unique vector $y_f \in \overline{M}$ such that $\|y_f\| = \|f\|_*$ and

$$f(x) = \langle x, y_f \rangle \qquad \forall x \in \overline{M}.$$

Define

$$F(x) = \langle x, y_f \rangle \qquad \forall x \in X.$$

Then $F \in X^*$, $F\big|_M = f$ and $\|F\|_* = \|y_f\| = \|f\|_*$. We claim that F is the *unique extension* of f with these properties. Indeed, let G be another solution of the problem (6.17) and let y_G be the vector in X associated with G in the Riesz representation. Consider the Riesz orthogonal decomposition of y_G, that is,

$$y_G = y_G' + y_G'' \quad \text{where} \quad y_G' \in \overline{M} \quad \text{and} \quad y_G'' \perp \overline{M}.$$

Then

$$\langle x, y_G' \rangle = G(x) = f(x) = \langle x, y_f \rangle \qquad \forall x \in \overline{M}.$$

So $y_G' = y_f$. Moreover

$$\|y_G''\|^2 = \|y_G\|^2 - \|y_G'\|^2 = \|G\|_*^2 - \|y_f\|^2 = \|f\|_*^2 - \|y_f\|^2 = 0.$$

In general, the following classical result ensures the existence of a solution for the problem (6.17) even though the uniqueness of the extension is no longer guaranteed.

Theorem 6.37 (Hahn-Banach) *Let X be a normed linear space, M a subspace of X, and $f : M \to \mathbb{R}$ a bounded linear functional. Then there exists $F \in X^*$ such that $F\big|_M = f$ and $\|F\|_* = \|f\|_*$.*

Proof To begin with, let us suppose $\|f\|_* \neq 0$ (otherwise one can take $F \equiv 0$ and the thesis immediately follows). We can also assume, without loss of generality, that $\|f\|_* = 1$. We will show, first, how to extend f to a subspace of X which *strictly* contains M. The general case will be treated later—in steps 2 and 3—using a maximality argument.

1. Suppose $M \neq X$ and let $x_0 \in X \setminus M$. Let us construct an extension of f to the subspace
$$M_0 := M + \mathbb{R}x_0 = \{x + \lambda x_0 \mid x \in M,\ \lambda \in \mathbb{R}\}.$$
Define
$$f_0(x + \lambda x_0) := f(x) + \lambda\alpha \qquad \forall x \in M,\ \forall \lambda \in \mathbb{R}, \tag{6.18}$$
where α is a real number to be chosen later. Clearly, f_0 is a linear functional on M_0 that extends f. We must find $\alpha \in \mathbb{R}$ such that the extended functional is bounded and has norm 1. This will occur if
$$|f_0(x + \lambda x_0)| \leq \|x + \lambda x_0\| \qquad \forall x \in M,\ \forall \lambda \in \mathbb{R}.$$
A simple rescaling argument allows to recast the above inequality as
$$|f_0(x_0 - y)| \leq \|x_0 - y\| \qquad \forall y \in M.$$
Therefore, replacing f_0 by its definition in (6.18), we deduce that α must satisfy $|\alpha - f(y)| \leq \|x_0 - y\|$ for every $y \in M$, or, equivalently,
$$f(y) - \|x_0 - y\| \leq \alpha \leq f(y) + \|x_0 - y\| \qquad \forall y \in M.$$
Now, such a choice of α is possible since
$$f(y) - f(z) = f(y - z) \leq \|y - z\| \leq \|x_0 - y\| + \|x_0 - z\| \quad \forall y, z \in M,$$
and so
$$\sup_{y\in M} \big\{f(y) - \|x_0 - y\|\big\} \leq \inf_{z\in M} \big\{f(z) + \|x_0 - z\|\big\}.$$

2. Denote by $\mathscr{P}$ the family of all pairs $(\widetilde{M}, \widetilde{f})$, where $\widetilde{M}$ is a subspace of X including M and $\widetilde{f}$ is a bounded linear functional extending f to $\widetilde{M}$ such that $\|\widetilde{f}\|_* = 1$.

$\mathscr{P} \neq \varnothing$ since it contains (M, f). Moreover, $\mathscr{P}$ is a partially ordered set with respect to the following order relation: for any $(M_1, f_1), (M_2, f_2) \in \mathscr{P}$,

$$(M_1, f_1) \leq (M_2, f_2) \iff \begin{cases} M_1 \text{ subspace of } M_2, \\ f_2 = f_1 \quad \text{on } M_1. \end{cases} \tag{6.19}$$

We claim that $\mathscr{P}$ is an inductive set, i.e., every totally ordered subset of $\mathscr{P}$ admits a supremum. To see this, let $\mathscr{Q} = \{(M_i, f_i)_{i \in I}\}$ be a totally ordered subset of $\mathscr{P}$. Then it is easy to check that, setting

$$\begin{cases} \widetilde{M} := \bigcup_{i \in I} M_i, \\ \widetilde{f}(x) := f_i(x) \quad \text{if} \quad x \in M_i, \end{cases}$$

the pair $(\widetilde{M}, \widetilde{f}) \in \mathscr{P}$ is an upper bound (actually, the supremum) of $\mathscr{Q}$.

3. By Zorn's Lemma, $\mathscr{P}$ has a maximal element, which we label $(\mathscr{M}, F)$. The thesis will follow if we prove that $\mathscr{M} = X$, because $F = f$ on M and $\|F\|_* = 1$ by construction. On the other hand, if $\mathscr{M}$ were a proper subspace of X, then the first step of the proof would imply the existence of a proper extension of $(\mathscr{M}, F)$, contradicting its maximality.

The theorem is thus proved. □

Example 6.38 In general, the extension provided by Hahn-Banach Theorem is not unique. For instance, consider the spaces

$$\tilde{c} := \left\{ x = (x_n)_n \in \ell^\infty \;\middle|\; \exists \lim_{n \to \infty} x_n \right\},$$

$$\tilde{c}' := \left\{ x = (x_n)_n \in \ell^\infty \;\middle|\; \exists \lim_{n \to \infty} x_{2n} \;\; \& \;\; \exists \lim_{n \to \infty} x_{2n+1} \right\}.$$

It is easy to see that $\tilde{c}, \tilde{c}'$ are closed subspaces of ℓ^∞. Clearly $\tilde{c} \subset \tilde{c}'$. Let $f \in (\tilde{c})^*$, $f_1, f_2 \in (\tilde{c}')^*$ be the continuous linear functionals defined by:

$$f(x) := \lim_{n \to \infty} x_n \qquad \forall x = (x_n)_n \in \tilde{c},$$

$$f_1(x) := \lim_{n \to \infty} x_{2n}, \quad f_2(x) := \lim_{n \to \infty} x_{2n+1} \qquad \forall x = (x_n)_n \in \tilde{c}'.$$

Then $\|f\|_* = \|f_1\|_* = \|f_2\|_* = 1$, $f_1 \equiv f_2 \equiv f$ on $\tilde{c}$, but $f_1 \not\equiv f_2$ on $\tilde{c}'$. Other examples of multiple extensions of continuous linear functionals are provided in Exercises 6.39 and 6.40.

Exercise 6.39 Let M be the closed subspace of ℓ^1:

$$M = \{x = (x_k)_k \in \ell^1 \mid x_k = 0 \;\; \forall k \geq 2\}$$

and define the functionals

$$f(x) = x_1 \quad \forall x = (x_k)_k \in M,$$

$$F(x) = \sum_{k=1}^{\infty} x_k, \quad F_n(x) = \sum_{k=1}^{n} x_k \quad \forall x = (x_k)_k \in \ell^1.$$

Show that, for every $n \geq 1$, $f \in M^*$, F, $F_n \in (\ell^1)^*$, $F_n\big|_M = F\big|_M = f$ and $\|F\|_* = \|F_n\|_* = \|f\|_* = 1$.

Exercise 6.40 In $\mathbb{R}^2$ with the norm

$$\|x\|_1 = |x_1| + |x_2| \quad \forall x = (x_1, x_2) \in \mathbb{R}^2,$$

consider the closed subspace

$$M = \{(x_1, 0) \mid x_1 \in \mathbb{R}\}$$

and the functionals:

$$f(x) = x_1 \quad \forall x = (x_1, 0) \in M,$$

$$F_1(x) = x_1, \quad F_2(x) = x_1 + x_2 \quad \forall x = (x_1, x_2) \in \mathbb{R}^2.$$

Show that $f \in M^*$, $F_1, F_2 \in (\mathbb{R}^2)^*$, $\|F_1\|_* = \|F_2\|_* = \|f\|_* = 1$, and $F_1\big|_M = F_2\big|_M = f$.

We shall now discuss some consequences of the Hahn-Banach Theorem.

Corollary 6.41 *Let X be a normed linear space, M a closed subspace of X and $x_0 \notin M$. Then there exists $F \in X^*$ such that:*

(a) $F(x_0) = 1$.
(b) $F(x) = 0$ *for every* $x \in M$.
(c) $\|F\|_* = 1/d_M(x_0)$*, where* $d_M(x_0)$ *is the distance of* x_0 *from* M *(see Appendix A).*

Proof Let $M_0 = M + \mathbb{R}x_0 = \{x + \lambda x_0 \mid x \in M, \ \lambda \in \mathbb{R}\}$. Define $f : M_0 \to \mathbb{R}$,

$$f(x + \lambda x_0) = \lambda \qquad \forall x \in M\,, \ \forall \lambda \in \mathbb{R}.$$

So $f(x_0) = 1$ and $f\big|_M = 0$. Moreover, since

$$\|x + \lambda x_0\| = |\lambda| \left\| \frac{x}{\lambda} + x_0 \right\| \geq |\lambda| d_M(x_0) \quad \forall x \in M, \ \forall \lambda \neq 0,$$

we have that $\|f\|_* \leq 1/d_M(x_0)$. Let $(x_n)_n \subset M$ be a sequence such that

$$\|x_n - x_0\| < \Big(1 + \frac{1}{n}\Big) d_M(x_0) \qquad \forall n \geq 1.$$

Then

$$\|f\|_* \|x_n - x_0\| \geq f(x_0 - x_n) = 1 > \frac{n}{n+1} \frac{\|x_n - x_0\|}{d_M(x_0)} \qquad \forall n \geq 1.$$

Therefore $\|f\|_* = 1/d_M(x_0)$. The existence of an extension $F \in X^*$ satisfying properties (a), (b), (c) follows from the Hahn-Banach Theorem. □

Corollary 6.42 *Let X be a normed linear space and $x_0 \in X \setminus \{0\}$. Then there exists $F \in X^*$ such that*

$$F(x_0) = \|x_0\| \quad \textit{and} \quad \|F\|_* = 1.$$

Proof Let $M = \{0\}$ and, given $x_0 \neq 0$, let $f \in X^*$ be the functional constructed in Corollary 6.41. Then, observing that $d_M(x_0) = \|x_0\|$, taking $F(x) = \|x_0\| f(x)$ yields the thesis. □

Exercise 6.43 Let $x_1, \ldots, x_n$ be linearly independent vectors in a normed linear space X and let $\lambda_1, \ldots, \lambda_n$ be real numbers. Show that there exists $f \in X^*$ such that

$$f(x_i) = \lambda_i \qquad \forall i = 1, \ldots, n.$$

Exercise 6.44 Let M be a subspace of a normed linear space X.

1. Show that a point $x \in X$ belongs to $\overline{M}$ if and only if $f(x) = 0$ for every $f \in X^*$ such that $f_{|M} = 0$.
2. Show that M is dense in X if and only if the unique functional $f \in X^*$ vanishing on M is $f \equiv 0$.

Exercise 6.45 Given a normed linear space X, show that X^* *separates the points* of X, i.e., for every $x_1, x_2 \in X$ with $x_1 \neq x_2$ there exists $f \in X^*$ such that $f(x_1) \neq f(x_2)$.

Exercise 6.46 Given a normed linear space X and $x \in X$, show that

$$\|x\| = \max\big\{f(x) \bigm| f \in X^*, \ \|f\|_* \leq 1\big\}.$$

Exercise 6.47 Let X, Y be two normed linear spaces and let $T : X \to Y$ be a bounded linear operator. The transpose $T^* : Y^* \to X^*$ is defined by $T^*\phi = \phi \circ T$ for all $\phi \in Y^*$. Show that $T^* \in \mathscr{L}(Y^*, X^*)$ and $\|T^*\| = \|T\|$. Moreover, if T is invertible and $T^{-1} \in \mathscr{L}(Y, X)$, show that T^* is also invertible and $(T^*)^{-1} = (T^{-1})^* \in \mathscr{L}(X^*, Y^*)$.

6.3.2 Separation of Convex Sets

Hahn-Banach Theorem has relevant geometric applications. Let us begin by extending our analysis to linear spaces.

Definition 6.48 A *sublinear functional* on a linear space X is a function $p : X \to \mathbb{R}$ such that:

(a) $p(\lambda x) = \lambda p(x)$ for every $x \in X$ and $\lambda > 0$.
(b) $p(x + y) \le p(x) + p(y)$ for every $x, y \in X$.

The Hahn-Banach Theorem can be generalized as follows.

Theorem 6.49 (Hahn-Banach: second analytic form) *Let p be a sublinear functional on a linear space X and let M be a subspace of X. If $f : M \to \mathbb{R}$ is a linear functional such that*

$$f(x) \le p(x) \qquad \forall x \in M, \tag{6.20}$$

then there exists a linear functional $F : X \to \mathbb{R}$ such that

$$\begin{cases} F\big|_M = f, \\ F(x) \le p(x) \;\; \forall x \in X. \end{cases} \tag{6.21}$$

Omitting the proof, we invite the reader to verify that the proof of Theorem 6.37 can be easily adapted to the above framework.

Theorem 6.50 (Hahn-Banach: first geometric form) *Let A, B be nonempty disjoint convex sets of a normed linear space X. If A is open, then there exists a functional $f \in X^*$ and a real number α such that*

$$f(x) < \alpha \le f(y) \qquad \forall x \in A, \;\; \forall y \in B. \tag{6.22}$$

Remark 6.51 Observe that (6.22) implies, in particular, $f \ne 0$. Given a functional $f \in X^* \setminus \{0\}$, for every $\alpha \in \mathbb{R}$ the set

$$\Pi_\alpha := f^{-1}(\alpha) = \{x \in X \mid f(x) = \alpha\} \tag{6.23}$$

is a closed hyperplane in X (see Definition 5.47). Indeed, since f is continuous, Π_α is a closed set. To show that Π_α is a hyperplane, let $y_0 \in X$ be such that $f(y_0) = 1$. Then for every $x \in X$, we have

$$x = \underbrace{x - f(x)y_0}_{\in \ker f} + f(x)y_0\,.$$

So $X = \ker f + \mathbb{R}y_0$, by which we deduce that $\ker f$ has codimension 1. It follows that $\Pi_\alpha = \ker f + \alpha y_0$ is a closed hyperplane in X. Therefore the conclusion of

Theorem 6.50 can by reformulated by stating that A and B can be separated by a closed hyperplane.

The proof of Theorem 6.50 is based on the following two lemmas.

Lemma 6.52 *Let C be an open convex set of a normed linear space X such that $0 \in C$. Then*

$$p_C(x) := \inf\{\tau > 0 \mid x \in \tau C\} \qquad \forall x \in X \tag{6.24}$$

is a sublinear functional on X called the Minkowski functional *or* gauge *of C. Moreover,*

(i) $\exists c > 0$ *such that* $0 \leq p_C(x) \leq c\|x\|$ *for every* $x \in X$.
(ii) $C = \{x \in X \mid p_C(x) < 1\}$.

Proof Let us observe, first, that C contains a ball B_R.

1. We begin by proving (i). For any $\varepsilon > 0$ and $x \in X$ we have

$$\frac{Rx}{\|x\| + \varepsilon} \in B_R \subset C.$$

From the arbitrariness of ε, it follows that $0 \leq p_C(x) \leq \|x\|/R$.

2. We now proceed with showing that p_C is a sublinear functional. Fix $\lambda > 0, x \in X$ and $\varepsilon > 0$. Let $0 < \tau_\varepsilon < p_C(x) + \varepsilon$ be such that $x \in \tau_\varepsilon C$. Then $\lambda x \in \lambda\tau_\varepsilon C$. So $p_C(\lambda x) \leq \lambda\tau_\varepsilon < \lambda(p_C(x) + \varepsilon)$. The arbitrariness of ε gives

$$p_C(\lambda x) \leq \lambda p_C(x) \qquad \forall\lambda > 0,\ \forall x \in X. \tag{6.25}$$

To obtain the opposite inequality, observe that, thanks to (6.25),

$$p_C(x) = p_C\Big(\frac{1}{\lambda}\,\lambda x\Big) \leq \frac{1}{\lambda}\,p_C(\lambda x).$$

Finally, let us check that p_C satisfy property (b) of Definition 6.48. Fix $x, y \in X$ and $\varepsilon > 0$. Let $0 < \tau_\varepsilon < p_C(x) + \varepsilon$ and $0 < \sigma_\varepsilon < p_C(y) + \varepsilon$ be such that $x \in \tau_\varepsilon C$ and $y \in \sigma_\varepsilon C$. Then $x = \tau_\varepsilon x_\varepsilon$ and $y = \sigma_\varepsilon y_\varepsilon$ for some points $x_\varepsilon, y_\varepsilon \in C$. Since C is convex, we deduce that

$$x + y = \tau_\varepsilon x_\varepsilon + \sigma_\varepsilon y_\varepsilon = (\tau_\varepsilon + \sigma_\varepsilon)\Big(\underbrace{\frac{\tau_\varepsilon}{\tau_\varepsilon + \sigma_\varepsilon}\,x_\varepsilon + \frac{\sigma_\varepsilon}{\tau_\varepsilon + \sigma_\varepsilon}\,y_\varepsilon}_{\in C}\Big).$$

Therefore

$$p_C(x + y) \leq \tau_\varepsilon + \sigma_\varepsilon < p_C(x) + p_C(y) + 2\varepsilon \qquad \forall\varepsilon > 0.$$

So $p_C(x + y) \leq p_C(x) + p_C(y)$.

3. Set $\widetilde{C} = \{x \in X \mid p_C(x) < 1\}$. Since C is convex and $0 \in C$, we have that $\tau C \subset C$ for every $\tau \in [0, 1]$, and so $\widetilde{C} \subset C$. Conversely, since C is open, each $x \in C$ belongs to some ball $\overline{B}_r(x) \subset C$. So, if $x \neq 0$, $(1 + \|x\|^{-1}r)x \in C$, whence $p_C(x) \leq 1/(1 + \|x\|^{-1}r) < 1$.

The lemma is thus completely proved. □

Lemma 6.53 *Let $C \neq \varnothing$ be an open convex set in a normed linear space X and let $x_0 \in X \setminus C$. Then there exists a functional $f \in X^*$ such that*

$$f(x) < f(x_0) \qquad \forall x \in C\,.$$

Proof We may assume, up to translation, $0 \in C$. Define $M := \mathbb{R}x_0 = \{\lambda x_0 \mid \lambda \in \mathbb{R}\}$ and $g : M \to \mathbb{R}$ by

$$g(\lambda x_0) = \lambda p_C(x_0) \qquad \forall \lambda \in \mathbb{R},$$

where p_C is the Minkowski functional of C. Observe that g satisfies condition (6.20) with respect to the sublinear functional p_C: for every $x = \lambda x_0 \in M$, the inequality

$$g(x) = \lambda p_C(x_0) \leq p_C(x),$$

which is obvious if $\lambda \leq 0$, follows from property (a) of Definition 6.48 if $\lambda > 0$. Then Theorem 6.49 guarantees the existence of a linear extension of g, which we label f, such that $f(x) \leq p_C(x)$ for every $x \in X$. Moreover, by property (i) of Lemma 6.52,

$$f(x) \leq c\|x\| \quad \text{and} \quad f(-x) \leq c\|x\| \qquad \forall x \in X,$$

so $f \in X^*$. Finally, once again thanks to Lemma 6.52,

$$f(x) \leq p_C(x) < 1 \leq p_C(x_0) = g(x_0) = f(x_0) \qquad \forall x \in C,$$

and this concludes the proof. □

Proof of Theorem 6.50. It is easy to check that

$$C := A - B = \{x - y \mid x \in A\,,\ y \in B\}$$

is a nonempty open convex set in X such that $0 \notin C$. Then by Lemma 6.53 there exists a linear functional $f \in X^*$ such that $f(z) < 0 = f(0)$ for every $z \in C$, that is, $f(x) < f(y)$ for every $x \in A$ and $y \in B$. So

$$\alpha := \sup_{x \in A} f(x) \leq f(y) \qquad \forall y \in B.$$

Let us show that $f(x) < \alpha$ for every $x \in A$ reasoning by contradiction: suppose that there exists $x_0 \in A$ such that $f(x_0) = \alpha$. Then the open set A contains a closed ball $\overline{B}_r(x_0)$ for some $r > 0$. So

$$f(x_0 + rx) \le \alpha \qquad \forall x \in \overline{B}_1.$$

If we choose $x_1 \in \overline{B}_1$ satisfying $f(x_1) > \|f\|_*/2$, we obtain the contradiction

$$f(x_0 + rx_1) = f(x_0) + rf(x_1) > \alpha + \frac{r\|f\|_*}{2}.$$

The conclusion follows. □

Next result deals with the separation in a 'strict sense' of two convex sets and generalizes to Banach spaces the analogous Proposition 5.48 for Hilbert spaces.

Theorem 6.54 (Hahn-Banach: second geometric form) *Let C and K be nonempty disjoint convex sets in a normed linear space X. If C is closed and K is compact, then there exists a functional $f \in X^*$ such that*

$$\sup_{x\in C} f(x) < \inf_{y\in K} f(y). \tag{6.26}$$

Proof Let us denote by d_C the distance function from C. Since C is closed and K is compact, the continuity of the function d_C implies that

$$\delta := \min_{y\in K} d_C(y) > 0\,. \tag{6.27}$$

Set

$$C_\delta := C + B_{\delta/2} = \{x + z \,|\, x \in C,\ z \in B_{\delta/2}\},$$

$$K_\delta := K + B_{\delta/2} = \{y + z \,|\, y \in K,\ z \in B_{\delta/2}\}.$$

It can be easily checked that C_δ and K_δ are nonempty open convex sets. They are also disjoint since, if $x + z = y + w$ for some choice of $x \in C$, $y \in K$ and $z, w \in B_{\delta/2}$, then we would have

$$d_C(y) \le \|x - y\| = \|w - z\| < \delta\,,$$

in contradiction with (6.27). By Theorem 6.50, there exist $f \in X^*$ and $\alpha \in \mathbb{R}$ such that

$$f\Big(x + \frac{\delta}{2}z\Big) < \alpha \le f\Big(y + \frac{\delta}{2}w\Big) \qquad \forall x \in C,\ \forall y \in K,\ \forall z, w \in B_1.$$

Recalling that $\|f\|_* > 0$ (see Remark 6.51) and $\|f\|_* = \sup_{\|x\|<1} |f(x)|$ by (6.4), let $z \in B_1$ be such that $f(z) > \|f\|_*/2$. Then

$$f(x) + \frac{\delta\|f\|_*}{4} < f\Big(x + \frac{\delta}{2}z\Big) \le \alpha \le f\Big(y - \frac{\delta}{2}z\Big) < f(y) - \frac{\delta\|f\|_*}{4}$$

for every $x \in C$ and $y \in K$. The conclusion follows. □

Corollary 6.55 *Let $C \neq \varnothing$ be a closed convex set in a normed linear space X, and let $x_0 \in X \setminus C$. Then there exists a functional $f \in X^*$ such that*

$$\sup_{x \in C} f(x) < f(x_0)\,.$$

Exercise 6.56 Let C be an open convex set in a normed linear space X such that $0 \in C$, and let $p_C(\cdot)$ be its Minkowski functional.

1. Show that if C does not contain any half-line of the form
$$\mathbb{R}_+ x_0 = \{\lambda x_0 \mid \lambda > 0\} \qquad x_0 \in X \setminus \{0\},$$
then $p_C(x) \neq 0$ for every $x \neq 0$.
2. Give an example to show that, in general, $p_C(\cdot)$ may vanish on vectors $x \neq 0$.
3. Show that if C is symmetric with respect to 0 (i.e., $x \in C \Leftrightarrow -x \in C$), then $p_C(\cdot)$ is a seminorm on X (see Sect. 6.1).
4. Deduce that if C is symmetric with respect to 0 and does not contain any half-line of the form $\mathbb{R}_+ x_0$ with $x_0 \neq 0$, then $p_C(\cdot)$ is a norm on X.
5. If C is bounded, it is obvious that C does not contain any half-line. Conversely, is it true that if C does not contain any half-line of the form $\mathbb{R}_+ x_0$ with $x_0 \neq 0$, then C is bounded?

6.3.3 The Dual of ℓ^p

In this section we will study the dual of the Banach spaces[8]

$$\ell^p = \Big\{ x = (x_k)_k \;\Big|\; \|x\|_p^p := \sum_{k=1}^{\infty} |x_k|^p < \infty \Big\} \qquad 1 \le p < \infty$$

and

$$c_0 = \Big\{ x = (x_k)_k \;\Big|\; \lim_{k \to \infty} x_k = 0 \Big\}.$$

These spaces, together with the Banach space

$$\ell^\infty = \Big\{ x = (x_k)_k \;\Big|\; \|x\|_\infty := \sup_{k \ge 1} |x_k| < \infty \Big\},$$

are of frequent use in this chapter because they provide simple examples of relevant new phenomena arising in infinite-dimensional settings in contrast with Euclidean spaces.

[8]See Example 6.6 and Exercise 6.7.

For $1 \le p \le \infty$, let us denote by p' the conjugate exponent of p, i.e.,

$$\frac{1}{p} + \frac{1}{p'} = 1 \quad \text{with the usual convention} \quad \frac{1}{\infty} = 0\,.$$

With any $y = (y_k)_k \in \ell^{p'}$ we can associate the linear map $f_y : \ell^p \to \mathbb{R}$ defined by

$$f_y(x) = \sum_{k=1}^{\infty} x_k y_k \qquad \forall x = (x_k)_k \in \ell^p\,.$$

Hölder's inequality and Exercise 3.26 ensure that, for $1 \le p \le \infty$,

$$|f_y(x)| \le \|y\|_{p'} \|x\|_p \qquad \forall x \in \ell^p\,. \tag{6.28}$$

Hence, $f_y \in (\ell^p)^*$ and $\|f_y\|_* \le \|y\|_{p'}$. Therefore the map

$$\boxed{1 \le p < \infty} \qquad \begin{cases} j_p : \ell^{p'} \to (\ell^p)^* \\ j_p(y) = f_y \end{cases}$$

is a bounded linear operator such that $\|j_p\| \le 1$. Moreover, for $y \in \ell^1$, (6.28) implies that f_y is a continuous linear functional on ℓ^∞, and, consequently, on c_0, since c_0 is a closed subspace of ℓ^∞. In the following, we will adopt this convention, considering $j_\infty(y) = f_y$ as an operator from ℓ^1 to $(c_0)^*$. Our next result contains the announced characterization of dual spaces.

Proposition 6.57 *For $1 \le p \le \infty$, set*

$$X_p = \begin{cases} \ell^p & \text{if} \quad 1 \le p < \infty, \\ c_0 & \text{if} \quad p = \infty\,. \end{cases}$$

Then the operator $j_p : \ell^{p'} \to (X_p)^*$ is an isometric isomorphism.[9]

Let us first prove the following lemma.

Lemma 6.58 *Let $1 \le p \le \infty$ and let $e_k \in X_p$ be the vectors*

$$e_k = (\overbrace{0, \ldots, 0}^{k-1}, 1, 0, \ldots) \qquad k = 1, 2 \ldots. \tag{6.29}$$

Then for every $x \in X_p$, we have

$$\sum_{k=1}^{n} x_k e_k \xrightarrow{X_p} x \qquad (n \to \infty)\,.$$

[9] See footnote 5 at p. 147.

Proof For $1 \le p < \infty$ we have, for any $x = (x_k)_k \in \ell^p$,

$$\Big\| x - \sum_{k=1}^{n} x_k e_k \Big\|_p^p = \sum_{k=n+1}^{\infty} |x_k|^p \to 0 \quad (n \to \infty).$$

Similarly, for any $x = (x_k)_k \in c_0$,

$$\Big\| x - \sum_{k=1}^{n} x_k e_k \Big\|_\infty = \max\{|x_k| \mid k > n\} \to 0 \quad (n \to \infty)$$

since $x_k \to 0$ by definition. The thesis follows. □

Remark 6.59 By Lemma 6.58 it follows that $\{\sum_{k=1}^{n} \lambda_k e_k \mid n \in \mathbb{N},\ \lambda_k \in \mathbb{Q}\}$ is a dense countable set in X_p for every $1 \le p \le \infty$. Consequently, c_0 and ℓ^p, for $1 \le p < \infty$, are separable spaces.

Remark 6.60 Observe that the conclusion of Lemma 6.58 is false for ℓ^∞: taking $x = (x_k)_k$ with $x_k = 1$ for every $n \in \mathbb{N}$, we have

$$\Big\| x - \sum_{k=1}^{n} x_k e_k \Big\|_\infty = 1 \not\to 0.$$

Indeed it is well-known that ℓ^∞ is not separable (see Example 3.27).

Proof of Proposition 6.57. Suppose, first, $1 < p < \infty$, whence $1 < p' < \infty$. Given $f \in (\ell^p)^*$, set

$$\begin{cases} y_k := f(e_k) & k \ge 1, \\ y := (y_k)_k , \end{cases} \tag{6.30}$$

where e_k is defined in (6.29). It suffices to show that

$$\boxed{y \in \ell^{p'}} \qquad \boxed{\|y\|_{p'} \le \|f\|_*} \qquad \boxed{f = f_y} \tag{6.31}$$

To this aim observe that, setting[10]

$$z^{(n)} = \sum_{k=1}^{n} |y_k|^{p'-2} y_k e_k \qquad \forall n \ge 1 ,$$

we have that $z^{(n)} \in \ell^p$, since all its components vanish except for a finite number, and

[10]Note that $|y_k|^{p'-2} y_k = 0$ if $y_k = 0$ since $p' > 1$.

$$\sum_{k=1}^{n} |y_k|^{p'} = f(z^{(n)}) \le \|f\|_* \, \|z^{(n)}\|_p = \|f\|_* \left(\sum_{k=1}^{n} |y_k|^{p'} \right)^{1/p} .$$

It follows that

$$\left(\sum_{k=1}^{n} |y_k|^{p'} \right)^{1/p'} \le \|f\|_* \qquad \forall n \ge 1.$$

This yields the first two assertions in (6.31). To obtain the third one, given $x = (x_k)_k \in \ell^p$, set

$$x^{(n)} := \sum_{k=1}^{n} x_k e_k$$

and observe that

$$f(x^{(n)}) = \sum_{k=1}^{n} x_k f(e_k) = \sum_{k=1}^{n} x_k y_k .$$

Since $x^{(n)} \to x$ in ℓ^p thanks to Lemma 6.58, by the continuity of f we have that $f(x^{(n)}) \to f(x)$. On the other hand, the series $\sum_{k=1}^{\infty} x_k y_k$ converges to $f_y(x)$. By uniqueness of limits, we conclude that $f = f_y$. This completes the analysis of the case $1 < p < \infty$. A similar argument applies to the cases $p = 1$ and $p = \infty$, see Exercise 6.61. □

Exercise 6.61 1. Prove Proposition 6.57 for $p = 1$.
Hint. Define y as in (6.30); the inequality $\|y\|_\infty \le \|f\|_*$ is immediate. To show that $f = f_y$ proceed as in the case $1 < p < \infty$.

2. Prove Proposition 6.57 for $p = \infty$.
Hint. Define y as in (6.30) and set

$$z^{(n)} = (z_k^{(n)})_k, \quad z_k^{(n)} = \begin{cases} \frac{y_k}{|y_k|} & \text{if} \quad k \le n \quad \text{and} \quad y_k \ne 0, \\ 0 & \text{if} \quad y_k = 0 \quad \text{or} \quad k > n. \end{cases}$$

Then $\|z^{(n)}\|_\infty \le 1$ and $\sum_{k=1}^{n} |y_k| = f(z^{(n)}) \le \|f\|_*$, by which it follows that $y \in \ell^1$ and $\|y\|_1 \le \|f\|_*$. To show that $f = f_y$ proceed as in the case $1 < p < \infty$.

Let $(X, \mathscr{E}, \mu)$ be a measure space and let $1 \le p \le \infty$. It is natural to ask whether the above analysis of the dual of ℓ^p can be generalized to the case $L^p(X, \mu)$. For every $g \in L^{p'}(X, \mu)$ let us define the linear functional $F_g : L^p(X, \mu) \to \mathbb{R}$

$$F_g(f) = \int_X f g \, d\mu \qquad \forall f \in L^p(X, \mu).$$

By Hölder's inequality and Exercise 3.26 it follows that

$$|F_g(f)| \le \|g\|_{p'} \|f\|_p \qquad \forall f \in L^p(X, \mu).$$

So $F_g \in (L^p(X, \mu))^*$ and $\|F_g\|_* \le \|g\|_{p'}$. We have thus defined the bounded linear operator

$$\begin{cases} L^{p'}(X, \mu) \to (L^p(X, \mu))^*, \\ g \mapsto F_g. \end{cases} \tag{6.32}$$

Proposition 6.62 *Let $(X, \mathscr{E}, \mu)$ be a measure space. If the following hypothesis holds*

$$1 < p < \infty \quad \textit{or} \quad p = 1 \ \& \ \mu \ \sigma\textit{-finite}, \tag{6.33}$$

then the bounded linear operator (6.32) is an isometric isomorphism.

For the proof we refer to Chap. 8 (Sect. 8.4).

Proposition 6.62 shows that, under assumption (6.33), any bounded linear functional on $L^p(X, \mu)$ can be represented as the integral with respect to a measure with density in $L^{p'}(X, \mu)$. The isometric isomorphism (6.32) allows to identify the dual of $L^p(X, \mu)$ with $L^{p'}(X, \mu)$. With this isometric isomorphism in mind, from now on it will be natural to make the identifications

$$(L^p(X, \mu))^* = L^{p'}(X, \mu) \ \text{ if } 1 < p < \infty, \tag{6.34}$$

$$(L^1(X, \mu))^* = L^\infty(X, \mu) \ \text{ if } \mu \text{ is } \sigma\text{-finite}. \tag{6.35}$$

In the particular case $X = \mathbb{N}$ with the counting measure $\mu = \mu^\#$, Proposition 6.57 allows to identify the spaces

$$(\ell^p)^* = \ell^{p'} \ \text{ if } 1 \le p < \infty, \qquad (c_0)^* = \ell^1. \tag{6.36}$$

Example 6.63 For $p = \infty$ the operator (6.32) is not onto, in general, as the following two examples show.

1. For instance, consider $L^\infty(-1, 1)$. Among the functionals of $(L^\infty(-1, 1))^*$, we find the extension—provided by Hahn-Banach Theorem—of the Dirac delta in the origin, which is a continuous linear functional on $\mathscr{C}([-1, 1])$:

$$\delta_0(f) = f(0) \quad \forall f \in \mathscr{C}([-1, 1]).$$

Let us label such an extension by T. Suppose by contradiction that there exists a function $g \in L^1(-1, 1)$ such that

$$T(f) = \int_{-1}^{1} fg\,dx \quad \forall f \in L^\infty(-1,1).$$

Set

$$f_n(t) = e^{-nt^2}, \quad t \in [-1,1].$$

An easy application of Dominated Convergence Theorem gives that

$$\int_{-1}^{1} f_n g\,dx \to 0;$$

on the other hand $T(f_n) = f_n(0) = 1$, and the contradiction follows. For a more extended treatment of the dual of $L^\infty(a,b)$ see [Yo65].

2. A similar example can be constructed in ℓ^∞. Indeed, let $\tilde{c}$ be the subspace defined in Example 6.38:

$$\tilde{c} := \{x = (x_k)_k \in \ell^\infty \mid \exists \lim_k x_k\},$$

and let f be the bounded linear functional on $\tilde{c}$ defined by

$$(x_k)_k \in \tilde{c} \mapsto \lim_{k\to\infty} x_k.$$

According to the Hahn-Banach Theorem there exists an extension $F \in (\ell^\infty)^*$ to the whole space ℓ^∞. Suppose by contradiction that there exists a sequence $y = (y_k)_k \in \ell^1$ such that

$$F((x_k)_k) = \sum_{k=1}^{\infty} x_k y_k \quad \forall (x_k)_k \in \ell^\infty.$$

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

$$x^{(n)} = (\underbrace{0, 0, \ldots, 0}_{n-1}, 1, 1, 1\ldots) \in \tilde{c}. \tag{6.37}$$

We get

$$\sum_{k=1}^{\infty} x_k^{(n)} y_k = \sum_{k=n+1}^{\infty} y_k \to 0 \quad \text{as } n \to \infty.$$

On the other hand, $F(x^{(n)}) = \lim_{k\to\infty} x_k^{(n)} = 1$, which gives a contradiction.

Example 6.64 If $p = 1$ and μ is not σ-finite, then the operator (6.32) may fail to be onto, as the following example shows. Indeed, consider the measure space

$$([0,1], \mathscr{B}([0,1]), \mu^\#)$$

where $\mathscr{B}([0,1])$ is the Borel σ-algebra on the interval $[0,1]$ and $\mu^{\#}$ is the counting measure. Now, let $E \subset [0,1]$ be a set which is not Borel (see Example 1.66). Observe that

$$f \in L^1([0,1],\mu^{\#}) \Longrightarrow A_f := \{x \in [0,1] : f(x) \neq 0\} \text{ is finite or countable.}$$

Consider the map $T : L^1([0,1],\mu^{\#}) \to \mathbb{R}$ defined by

$$f \in L^1([0,1],\mu^{\#}) \mapsto \sum_{x\in A_f\cap E} f(x).$$

Then

$$|T(f)| \leq \sum_{x\in A_f\cap E} |f(x)| \leq \sum_{x\in A_f} |f(x)| = \|f\|_1,$$

whence $T \in (L^1([0,1],\mu^{\#}))^*$. Suppose by contradiction that there exists a function $g \in L^\infty([0,1],\mu^{\#})$ such that

$$T(f) = \int_{[0,1]} fg\,d\mu^{\#} \quad \forall f \in L^1([0,1],\mu^{\#}).$$

Then for any $x \in [0,1]$, denoting by $\chi_{\{x\}}$ the characteristic function of the singleton $\{x\}$, we have $\chi_{\{x\}} \in L^1([0,1],\mu^{\#})$ and

$$T(\chi_{\{x\}}) = \chi_E(x).$$

On the other hand

$$\int_{[0,1]} \chi_{\{x\}} g\,d\mu^{\#} = g(x).$$

Thus g actually coincides with the characteristic function of the set E: so g fails to be a Borel function, in contrast with the assumption.

6.4 Weak Convergence and Reflexivity

Given a normed linear space X, an equivalent notation of frequent use to denote the action of a functional on X is

$$\langle f, x\rangle := f(x) \quad \forall f \in X^*,\ \forall x \in X.$$

Definition 6.65 The space $X^{**} = (X^*)^*$ is called the *bidual* of X.

Let $J_X : X \to X^{**}$ be the linear operator defined by

$$\langle J_X(x), f\rangle := \langle f, x\rangle \qquad \forall x \in X,\ \forall f \in X^*. \tag{6.38}$$

Then $|\langle J_X(x), f\rangle| \le \|f\|_* \|x\|$ by definition. So $\|J_X(x)\|_* \le \|x\|$. Moreover, by Corollary 6.42, for every $x \in X$ there exists a functional $f_x \in X^*$ such that $f_x(x) = \|x\|$ and $\|f_x\|_* = 1$. Thus, $\|x\| = |\langle J_X(x), f_x\rangle| \le \|J_X(x)\|_*$. It follows that $\|J_X(x)\|_* = \|x\|$ for every $x \in X$, that is, J_X is a *linear isometry*.

6.4.1 Reflexive Spaces

Since J_X is a linear operator, then $J_X(X)$ is a subspace of X^{**}. It is useful to single out the case where such a subspace coincides with the bidual.

Definition 6.66 A normed linear space X is said to be *reflexive* if the linear operator $J_X : X \to X^{**}$ defined by (6.38) is onto.

Recalling that J_X is a linear isometry, we deduce that any reflexive space X is isometrically isomorphic[11] to its bidual X^{**}. Since X^{**} is complete, like every dual space (Proposition 6.12), it follows that every reflexive space must also be complete.

Example 6.67
1. If H is a Hilbert space, then the Riesz isomorphism allows to identify H with H^*. Moreover, H^* is also a Hilbert space: indeed, the dual norm $\|\cdot\|_*$ is associated to the scalar product

$$\langle f, g\rangle = \langle y_f, y_g\rangle \quad \forall f,\ g \in H^*, \tag{6.39}$$

where y_f, y_g are the vectors in H related to f and g, respectively, according to Riesz representation. So H^*, as a Hilbert space, can be identified with its dual H^{**}. Furthermore, it is not difficult to check that the composition of the two Riesz isomorphisms coincides with the canonical embedding J_H defined in (6.38). So H is reflexive.

2. Let $1 < p < \infty$. Then, according to (6.36), $(\ell^p)^*$ can be identified by $\ell^{p'}$ under the isometric isomorphism j_p of Proposition 6.57, where p' is the conjugate exponent of p. Since $1 < p' < \infty$, then $(\ell^{p'})^* = \ell^p$ under the isomorphism $j_{p'}$. Consider the map $\ell^p \to (\ell^p)^{**}$, obtained by composing $j_{p'}$ with the transpose (see Exercise 6.47) of the inverse of j_p:

$$\ell^p \xrightarrow{j_{p'}} (\ell^{p'})^* \xrightarrow{(j_p^{-1})^*} (\ell^p)^{**}.$$

[11] See footnote 5 at p. 147.

This map coincides with the canonical embedding J_{ℓ^p} defined in (6.38) of ℓ^p into its bidual. Moreover, the map J_{ℓ^p} is onto, as composition of two onto maps, and this proves that the space ℓ^p is reflexive.

3. Let $(X, \mathscr{E}, \mu)$ be a measure space and let $1 < p < \infty$. Then, by (6.34), we have $(L^p(X,\mu))^* = L^{p'}(X,\mu)$ where p' is the conjugate exponent of p. Then, proceeding as for spaces ℓ^p, we deduce that $L^p(X,\mu)$ is reflexive.

Theorem 6.68 *Let X be a normed linear space. Then the following statements hold:*

(a) If X^* is separable, then X is separable.
(b) If X is complete and X^* is reflexive, then X is reflexive.

Proof (a) Let $(f_n)_n$ be a dense sequence in X^*. Then there exists a sequence $(x_n)_n$ in X such that

$$\|x_n\| = 1 \quad \text{and} \quad |\langle f_n, x_n\rangle| \ge \frac{\|f_n\|_*}{2} \qquad \forall n \ge 1.$$

Let M be the closed subspace generated by $(x_n)_n$, i.e., the closure of the set of all finite linear combinations of vectors x_n. By construction, M is separable (the finite linear combinations of vectors x_n with rational coefficients form a countable dense set in M). We claim that $M = X$. Indeed, suppose that there exists $x_0 \in X \setminus M$. Then, applying Corollary 6.41, we can find a functional $f \in X^*$ such that

$$\langle f, x_0\rangle = 1, \quad f\big|_M = 0, \quad \|f\|_* = \frac{1}{d_M(x_0)}.$$

So

$$\frac{\|f_n\|_*}{2} \le |\langle f_n, x_n\rangle| = |\langle f_n - f, x_n\rangle| \le \|f_n - f\|_*,$$

whence

$$\frac{1}{d_M(x_0)} = \|f\|_* \le \|f - f_n\|_* + \|f_n\|_* \le 3\|f - f_n\|_*,$$

in contradiction with the hypothesis that $(f_n)_n$ is dense in X^*.

(b) Observe that the linear operator $x \in X \mapsto J_X(x) \in J_X(X)$ is an isometric isomorphism of X onto $J_X(X)$. Therefore, if X is a Banach space, then $J_X(X)$ is also a Banach space and, consequently, a closed subspace of X^{**}. Suppose that there exists $\phi_0 \in X^{**} \setminus J_X(X)$. Then, by Corollary 6.41 applied to the bidual, there exists a bounded linear functional on X^{**} valued 1 at ϕ_0 and 0 on $J_X(X)$. Since X^* is reflexive, such a functional belongs to $J_{X^*}(X^*)$. So, for some $f \in X^*$,

$$\langle \phi_0, f \rangle = 1 \quad \text{and} \quad 0 = \langle J_X(x), f \rangle = \langle f, x \rangle \qquad \forall x \in X,$$

which yields a contradiction.

□

Remark 6.69 It is well-known that spaces ℓ^∞ and $L^\infty(a, b)$ are not separable (see Example 3.27), whereas ℓ^1 and $L^1(a, b)$ are separable (by Proposition 3.47 and Remark 6.59). Thanks to part (a) of Theorem 6.68, we deduce that $(\ell^\infty)^*$ and $(L^\infty(a, b))^*$ are not separable. So $(\ell^1)^{**} = (\ell^\infty)^*$ is not isomorphic to ℓ^1 and $(L^1(a, b))^{**} = (L^\infty(a, b))^*$ is not isomorphic to $L^1(a, b)$. So ℓ^1 and $L^1(a, b)$ fail to be reflexive. It follows that ℓ^∞ and $L^\infty(a, b)$ also fail to be reflexive, otherwise ℓ^1 and $L^1(a, b)$ would be reflexive by part (b) of Theorem 6.68.

Remark 6.70 The result of part (b) of Theorem 6.68 is an equivalence since the implication

$$X \quad \text{reflexive} \quad \Longrightarrow \quad X^* \quad \text{reflexive}$$

is trivial. On the contrary, the implication of part (a) cannot be reversed. Indeed, ℓ^1 is separable, whereas $(\ell^1)^*$ is not separable since it is isomorphic to ℓ^∞.

Corollary 6.71 *A Banach space X is reflexive and separable if and only if X^* is reflexive and separable.*

Proof The only part of the conclusion that needs to be justified is the fact that if X is reflexive and separable, then X^* is separable. But this follows by observing that X^{**} is separable, since it is isomorphic to X. So, by Theorem 6.68(a), X^* is separable. □

We conclude this section with the following result on the reflexivity of subspaces.

Proposition 6.72 *Let M be a closed subspace of a reflexive Banach space X. Then M is reflexive.*

Proof Let $\phi \in M^{**}$. Define a functional $\overline{\phi}$ on X^* by setting

$$\langle \overline{\phi}, f \rangle = \langle \phi, f_{|M} \rangle \qquad \forall f \in X^*.$$

Since $\overline{\phi} \in X^{**}$, by hypothesis we have that $\overline{\phi} = J_X(\overline{x})$ for some $\overline{x} \in X$. We split the remaining part of the proof into two steps.

1. We claim that $\overline{x} \in M$. Indeed, if $\overline{x} \in X \backslash M$, then by Corollary 6.41 there exists $\overline{f} \in X^*$ such that

$$\langle \overline{f}, \overline{x} \rangle = 1 \quad \text{and} \quad \overline{f}_{|M} = 0\,.$$

   This yields a contradiction since

$$1 = \langle \overline{\phi}, \overline{f} \rangle = \langle \phi, \overline{f}_{|M} \rangle = 0.$$

2. We claim that $\phi = J_M(\overline{x})$. Indeed, for any $f \in M^*$, let $\widetilde{f} \in X^*$ be an extension of f to X provided by the Hahn-Banach Theorem. Then

$$\langle \phi, f \rangle = \langle \overline{\phi}, \widetilde{f} \rangle = \langle \widetilde{f}, \overline{x} \rangle = \langle f, \overline{x} \rangle \qquad \forall f \in M^*.$$

So J_M is onto and M is reflexive. □

6.4.2 Weak Convergence and Bolzano-Weierstrass Property

It is well known that all closed bounded subsets of a finite-dimensional normed linear space are compact. Such a property is usually referred to as the *Bolzano-Weierstrass property*. One of the most interesting phenomena that occur in infinite dimensions is that the Bolzano-Weierstrass property is no longer true (see Appendix C). To surrogate such a property in infinite-dimensional spaces it is convenient to introduce a weaker notion of convergence in addition to the natural convergence associated with the norm.

Definition 6.73 Let X be a normed linear space. A sequence $(x_n)_n \subset X$ is said to *converge weakly* to a point $x \in X$ if

$$\lim_{n\to\infty} \langle f, x_n \rangle = \langle f, x \rangle \qquad \forall f \in X^*.$$

In this case we write $x_n \overset{X}{\rightharpoonup} x$, or, simply, $x_n \rightharpoonup x$.

Example 6.74 In the case of Hilbert spaces or spaces $L^p(X, \mu)$ we have constructed an isometric isomorphism which allows to characterize the abstract space X^*, and so to represent ‘practically’ the continuous linear functionals. Then the notion of weak convergence can be reformulated as follows:

- Let H be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and let $(x_n)_n \subset H, x \in H$. Then

$$x_n \rightharpoonup x \iff \langle x_n, y \rangle \to \langle x, y \rangle \quad \forall y \in H.$$

- For $1 \le p \le \infty$, set

$$X_p = \begin{cases} \ell^p & \text{if} \quad 1 \le p < \infty, \\ c_0 & \text{if} \quad p = \infty. \end{cases}$$

  Let $x^{(n)}$, $x \in X_p$, $n = 1, 2, \ldots$. Then, setting $x^{(n)} = (x_k^{(n)})_k$ and $x = (x_k)_k$, we have

$$x^{(n)} \rightharpoonup x \iff \sum_{k=1}^{\infty} x_k^{(n)} y_k \to \sum_{k=1}^{\infty} x_k y_k \quad \forall y = (y_k)_k \in \ell^{p'},$$

where p' is the conjugate exponent of p.

- Given a σ-finite measure space $(X, \mathscr{E}, \mu)$ and $1 \leq p < \infty$, consider functions $(f_n)_n \subset L^p(X, \mu)$ and $f \in L^p(X, \mu)$. Then

$$f_n \rightharpoonup f \iff \int_X f_n g \, d\mu \to \int_X f g \, d\mu \quad \forall g \in L^{p'}(X, \mu),$$

where p' is the conjugate exponent of p.

A sequence $(x_n)_n$ that converges in norm to x, namely $x_n \to x$, is also said to *converge strongly* to x. Since $|\langle f, x_n\rangle - \langle f, x\rangle| \leq \|f\|_* \|x_n - x\|$, it is immediate that

$$x_n \to x \Longrightarrow x_n \rightharpoonup x.$$

The converse is not true, in general, as the following example shows.

Example 6.75 Let $(e_n)_n$ be an orthonormal sequence in an infinite-dimensional Hilbert space H. Then, owing to Bessel's inequality, $\langle x, e_n\rangle \to 0$ as $n \to \infty$ for every $x \in H$. Therefore, recalling Example 6.74, $e_n \rightharpoonup 0$ as $n \to \infty$. But $\|e_n\| = 1$ for every n. So $(e_n)_n$ does not converge strongly to 0.

Proposition 6.76 *Let $(x_n)_n$, $(y_n)_n$ be sequences in a normed linear space X, and let $x, y \in X$.*

(a) If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$, then $x = y$.
(b) If $x_n \rightharpoonup x$ and $y_n \rightharpoonup y$, then $x_n + y_n \rightharpoonup x + y$.
(c) If $x_n \rightharpoonup x$, $(\lambda_n)_n \subset \mathbb{R}$, and $\lambda_n \to \lambda \in \mathbb{R}$, then $\lambda_n x_n \rightharpoonup \lambda x$.
(d) If $x_n \overset{X}{\rightharpoonup} x$ and $\Lambda \in \mathscr{L}(X, Y)$, then $\Lambda x_n \overset{Y}{\rightharpoonup} \Lambda x$.
(e) If $x_n \rightharpoonup x$, then $(x_n)_n$ is bounded.
(f) If $x_n \rightharpoonup x$, then $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.

Proof (a) By hypothesis we have $\langle f, x - y\rangle = 0$ for every $f \in X^*$. Then the conclusion follows recalling Exercise 6.45.

(b) For every $f \in X^*$ we have $\langle f, x_n\rangle \to \langle f, x\rangle$ and $\langle f, y_n\rangle \to \langle f, y\rangle$, and so $\langle f, x_n + y_n\rangle = \langle f, x_n\rangle + \langle f, y_n\rangle \to \langle f, x\rangle + \langle f, y\rangle = \langle f, x + y\rangle$.

(c) Since $(\lambda_n)_n$ is bounded, say $|\lambda_n| \leq C$, for any $f \in X^*$ we have

$$|\lambda_n \langle f, x_n\rangle - \lambda\langle f, x\rangle| \leq \underbrace{|\lambda_n|}_{\leq C} \, |\underbrace{\langle f, x_n - x\rangle}_{\to 0}| + \underbrace{|\lambda_n - \lambda|}_{\to 0} \, |\langle f, x\rangle|.$$

(d) Let $g \in Y^*$. Then $\langle g, \Lambda x_n\rangle = \langle g \circ \Lambda, x_n\rangle \to \langle g \circ \Lambda, x\rangle = \langle g, \Lambda x\rangle$ since $g \circ \Lambda \in X^*$.

(e) Consider the sequence $(J_X(x_n))_n$ in X^{**}. Since

$$\langle J_X(x_n), f\rangle = \langle f, x_n\rangle \to \langle f, x\rangle \qquad \forall f \in X^*,$$

we have $\sup_n |\langle J_X(x_n), f\rangle| < \infty$ for all $f \in X^*$. So the Banach-Steinhaus Theorem implies that

$$\sup_{n\geq 1} \|x_n\| = \sup_{n\geq 1} \|J_X(x_n)\|_* < \infty.$$

(f) Let $f \in X^*$ be such that $\|f\|_* \leq 1$. Then

$$\underbrace{|\langle f, x_n\rangle|}_{\to|\langle f,x\rangle|} \leq \|x_n\| \quad \Longrightarrow \quad |\langle f, x\rangle| \leq \liminf_{n\to\infty} \|x_n\|.$$

The conclusion follows from Exercise 6.46.

□

Example 6.77 Let $(e_n)_n$ be an orthonormal sequence in an infinite-dimensional Hilbert space H. Since $e_n \rightharpoonup 0$ in H (see Example 6.75), $(e_n)_n$ provides an example for which the inequality in Proposition 6.76(f) is strict.

Theorem 6.78 (Banach-Saks) *Let $(x_n)_n$ be a sequence in a Hilbert space H that converges weakly to $x \in H$. Then there exists a subsequence $(x_{n_k})_k$ such that the arithmetic means*

$$\frac{1}{N}\sum_{k=1}^{N} x_{n_k}$$

converge strongly to x.

Proof Observe that, without loss of generality, we may assume $x = 0$. Set $n_1 = 1$. Next, given $x_{n_1}, \dots, x_{n_k}$, since $\langle x_{n_h}, x_n\rangle \to 0$ as $n \to \infty$ for every $h = 1, \dots, k$, we define n_{k+1} as the first index $n > n_k$ such that

$$|\langle x_{n_h}, x_n\rangle| \leq \frac{1}{k} \quad \forall h = 1, \dots, n_k. \tag{6.40}$$

Recalling that $(x_n)_n$ is bounded, say $\|x_n\| \leq M$ for all $n \in \mathbb{N}$, by (6.40) we deduce that

$$\begin{aligned}\Big\|\frac{1}{N}\sum_{k=1}^{N} x_{n_k}\Big\|^2 &\leq \frac{1}{N^2}\sum_{k=1}^{N}\|x_{n_k}\|^2 + \frac{2}{N^2}\sum_{k=2}^{N}\sum_{h=1}^{k-1}|\langle x_{n_h}, x_{n_k}\rangle| \\ &\leq \frac{M^2}{N} + 2\frac{N-1}{N^2} \to 0\end{aligned}$$

as $N \to \infty$.

Exercise 6.79 Let $f_n : \mathbb{R} \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \begin{cases} \dfrac{1}{2^n} & \text{if } x \in [2^n, 2^{n+1}], \\ 0 & \text{otherwise.} \end{cases}$$

Show that:

- $f_n \to 0$ in $L^p(\mathbb{R})$ for all $1 < p \leq \infty$.
- $(f_n)_n$ does not converge weakly in $L^1(\mathbb{R})$.
  Hint. Consider $g := \sum_{n=1}^{\infty} (-1)^n \chi_{[2^n, 2^{n+1})}$ and estimate $\int_{\mathbb{R}} f_n g \, dx$.

Exercise 6.80 Given $1 < p < \infty$, let $(a_n)_n$ be a sequence of real numbers and let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} a_n & \text{if } x \in \big[n, n+1\big], \\ 0 & \text{otherwise.} \end{cases}$$

Show that $(a_n)_n$ is bounded if and only if $f_n \rightharpoonup 0$ in $L^p(\mathbb{R})$.

Exercise 6.81 Let $f \in L^p(\mathbb{R})$ and let $f_n(x) = f(x - n),\ n \geq 1$. Show that:

- $f_n \rightharpoonup 0$ in $L^p(\mathbb{R})$ if $1 < p < \infty$.
  Hint. Show, first, that $\int_{\mathbb{R}} f_n g \, dx \to 0$ for any function[12] $g \in \mathscr{C}_c(\mathbb{R})$.
- If $f \in L^1(\mathbb{R})$, then $(f_n)_n$ does not converge weakly in $L^1(\mathbb{R})$, in general.
  Hint. Consider $f = \chi_{[0,1]}$.

Exercise 6.82 Let $f \in L^1(\mathbb{R}^N)$ be such that $\int_{\mathbb{R}^N} f(x) \, dx = 1$ and set

$$f_n(x) = nf(nx), \quad n \geq 1.$$

Show that:

- $\int_{\mathbb{R}^N} f_n g \, dx \to g(0)$ for any $g \in \mathscr{C}_c(\mathbb{R}^N)$.
- $(f_n)_n$ does not converge weakly in $L^1(\mathbb{R}^N)$.

Exercise 6.83 Let $x^{(n)},\ x \in \ell^2$ $(n \in \mathbb{N})$ be such that

$$x^{(n)} \rightharpoonup x \text{ in } \ell^2.$$

Set $x^{(n)} = \big(x_k^{(n)}\big)_k$, $x = (x_k)_k$. Show that:

(a) $\lim_{n\to\infty} x_k^{(n)} = x_k$ for every $k \in \mathbb{N}$.

[12] We refer to p. 98 for the definition of the space $\mathscr{C}_c(\mathbb{R}^N)$.

(b) setting $y^{(n)} = \left(\frac{x_k^{(n)}}{k}\right)_k$, then

$$y^{(n)} \to y \text{ in } \ell^2$$

where $y = (\frac{x_k}{k})_k$.

Hint. Suppose $x = 0$ and observe that, if $\|x^{(n)}\|_2 \le C$ for all n, given $K \in \mathbb{N}$ we have

$$\sum_{k=1}^{\infty} |y_k^{(n)}|^2 \le \underbrace{\sum_{k=1}^{K} |y_k^{(n)}|^2}_{\to 0 \text{ by (a)}} + \frac{1}{K^2} \underbrace{\sum_{k=K+1}^{\infty} |x_k^{(n)}|^2}_{\le C^2}.$$

Exercise 6.84 Given a Hilbert space H with scalar product $\langle \cdot, \cdot \rangle$, let $(x_n)_n \subset H$ be a bounded sequence, $A \subset H$ a dense set and $x \in H$. Show that

$$x_n \rightharpoonup x \iff \langle x_n, y \rangle \to \langle x, y \rangle \;\; \forall y \in A.$$

Exercise 6.85 Given $1 < p < \infty$, let $x^{(n)}, x \in \ell^p$ ($n \in \mathbb{N}$) and suppose that $\|x^{(n)}\|_p \le C$ for all n. Then, setting $x^{(n)} = \left(x_k^{(n)}\right)_k$ and $x = (x_k)_k$, show that

$$x^{(n)} \rightharpoonup x \iff x_k^{(n)} \to x_k \;\; \forall k \in \mathbb{N} \;\; (\text{as } n \to \infty).$$

Hint. Concerning the implication '$\Longleftarrow$', suppose $x = 0$ and let $C > 0$ be such that $\|x^{(n)}\|_p \le C$ for all n. Given $\varepsilon > 0$ and $y = (y_k)_k \in \ell^{p'}$, where $p' \in (1, \infty)$ is the conjugate exponent of p, choose $k_\varepsilon \ge 1$ and $n_\varepsilon \ge 1$ such that

$$\left(\sum_{k=k_\varepsilon+1}^{\infty} |y_k|^{p'} \right)^{1/p'} < \varepsilon,$$

$$\left(\sum_{k=1}^{k_\varepsilon} |x_k^{(n)}|^p \right)^{1/p} < \varepsilon \qquad \forall n \ge n_\varepsilon.$$

Then, for every $n \ge n_\varepsilon$,

$$\left| \sum_{k=1}^{\infty} x_k^{(n)} y_k \right| = \left| \sum_{k=1}^{k_\varepsilon} x_k^{(n)} y_k \right| + \left| \sum_{k=k_\varepsilon+1}^{\infty} x_k^{(n)} y_k \right|$$
$$\le \underbrace{\left(\sum_{k=1}^{k_\varepsilon} |x_k^{(n)}|^p \right)^{\frac{1}{p}}}_{\le \varepsilon} \underbrace{\left(\sum_{k=1}^{k_\varepsilon} |y_k|^{p'} \right)^{\frac{1}{p'}}}_{\le \|y\|_{p'}} + \underbrace{\left(\sum_{k=k_\varepsilon+1}^{\infty} |x_k^{(n)}|^p \right)^{\frac{1}{p}}}_{\le C} \underbrace{\left(\sum_{k=k_\varepsilon+1}^{\infty} |y_k|^{p'} \right)^{\frac{1}{p'}}}_{\le \varepsilon}.$$

Exercise 6.86 Give an example to show that the results of Exercises 6.84 and 6.85 are false, in general, without the assumption that the sequence is bounded.
Hint. In ℓ^2 consider the sequence $x^{(n)} = n^2 e_n$, where e_n is the vector defined in (6.29). Then, for every $k \geq 1$, $x_k^{(n)} \to 0$ as $n \to \infty$. On the other hand, taking $y = (1/k)_k$ we have

$$y \in \ell^2 \quad \text{and} \quad \sum_{k=1}^{\infty} y_k x_k^{(n)} = n \to \infty.$$

Observe that if $\langle \cdot, \cdot \rangle$ denotes the scalar product in ℓ^2, then $\langle x^{(n)}, z \rangle \to 0$ for every $z \in A$, where A is the set of all finite linear combinations of the vectors e_n. Recall that A is dense in ℓ^2 (see Remark 6.59).

Exercise 6.87 Let $x^{(n)}, \ x \in c_0$ $(n \in \mathbb{N})$ and suppose that $\|x^{(n)}\|_\infty \leq C$ for every n. Then, setting $x^{(n)} = \big(x_k^{(n)}\big)_k$ and $x = (x_k)_k$, show that

$$x^{(n)} \rightharpoonup x \quad \Longleftrightarrow \quad x_k^{(n)} \to x_k \ \ \forall k \in \mathbb{N} \quad (\text{as } n \to \infty).$$

Hint. Proceed as in Exercise 6.85.

Exercise 6.88 Let $1 < p < \infty$ and let $x^{(n)}, \ x \in \ell^p$ $(n \in \mathbb{N})$. Show that

$$x^{(n)} \to x \quad \Longleftrightarrow \quad \begin{cases} x^{(n)} \rightharpoonup x, \\ \|x^{(n)}\|_p \to \|x\|_p. \end{cases}$$

Hint. Concerning the implication '$\Longleftarrow$' observe that, setting $x^{(n)} = \big(x_k^{(n)}\big)_k$ and $x = (x_k)_k$, thanks to Exercise 6.85, for every $k \geq 1$ we have $x_k^{(n)} \to x_k$ as $n \to \infty$. Then use Proposition 3.39 by taking $X = \mathbb{N}$ with the counting measure.

Exercise 6.89 Show that the result of Exercise 6.88 is false in c_0.
Hint. Consider the sequence $x^{(n)} = e_1 + e_n$ where e_1, e_n are the vectors defined in (6.29).

Exercise 6.90 Let H be a Hilbert space and let $x_n, \ x \in H$ $(n \in \mathbb{N})$. Show that

$$x_n \to x \quad \Longleftrightarrow \quad \begin{cases} x_n \rightharpoonup x, \\ \|x_n\| \to \|x\|. \end{cases} \tag{6.41}$$

Hint. Observe that $\|x_n - x\|^2 = \|x_n\|^2 + \|x\|^2 - 2\langle x_n, x \rangle$.

Remark 6.91 1. We say that a Banach space X has the *property of Radon-Riesz* if the equivalence (6.41) holds for every sequence $(x_n)_n$ in X. By the results of Exercises 6.88 and 6.90 we deduce that such a property holds in Hilbert spaces, in spaces ℓ^p for $1 < p < \infty$, whereas it fails in c_0 (Exercise 6.89).

The Radon-Riesz property is actually valid in a large variety of normed linear spaces, the so-called *uniformly convex spaces*, which include spaces $L^p(X, \mu)$ with $1 < p < \infty$ and $(X, \mathscr{E}, \mu)$ a generic measure space (see [Br83], [HS65], [Mo69]).

2. A surprising result, known as Schur's Theorem,[13] ensures that in ℓ^1 weak convergence entails strong convergence, that is, for any $x^{(n)}, x \in \ell^1$ $(n \in \mathbb{N})$ we have

$$x^{(n)} \to x \iff x^{(n)} \rightharpoonup x.$$

Then, owing to Schur's Theorem, ℓ^1 has the Radon-Riesz property. On the other hand, Schur's Theorem itself shows that the property described in Exercise 6.85 fails in ℓ^1. Indeed, the sequence $(e_n)_n$ defined in (6.29) does not converge strongly to 0 and, consequently, neither does weakly.

Exercise 6.92 Let M be a closed subspace of a normed linear space X and let $(x_n)_n \subset M$, $x \in X$. Show that if $x_n \rightharpoonup x$, then $x \in M$.
Hint. Use Corollary 6.41.

Exercise 6.93 Let C be a nonempty closed convex subset of a normed linear space X, and let $(x_n)_n \subset C$, $x \in X$. Show that if $x_n \rightharpoonup x$, then $x \in C$.
Hint. Use Corollary 6.55.

Besides strong and weak convergence, on a dual space X^* we can define another notion of convergence.

Definition 6.94 Given a normed linear space X, a sequence $(f_n)_n \subset X^*$ is said to *converge weakly−∗* to a functional $f \in X^*$ if

$$\langle f_n, x\rangle \to \langle f, x\rangle \quad \text{as} \quad n \to \infty \qquad \forall x \in X. \tag{6.42}$$

In this case we write

$$f_n \overset{*}{\rightharpoonup} f \quad (\text{as} \quad n \to \infty).$$

Remark 6.95 It is interesting to compare weak and weak−∗ convergence on a dual space X^*. By definition, a sequence $(f_n)_n \subset X^*$ converges weakly to $f \in X^*$ if and only if

$$\langle \phi, f_n\rangle \to \langle \phi, f\rangle \quad \text{as} \quad n \to \infty \tag{6.43}$$

for all $\phi \in X^{**}$, whereas $f_n \overset{*}{\rightharpoonup} f$ if and only if (6.43) holds for all $\phi \in J_X(X)$. Therefore

$$f_n \rightharpoonup f \implies f_n \overset{*}{\rightharpoonup} f.$$

Weak convergence is equivalent to weak−∗ convergence if X is reflexive but, in general, weak convergence is stronger than weak−∗ convergence, as we will show later.

[13] See, for instance, Proposition 2.19 in [Ko02].

Example 6.96 Using the identification (6.35) and (6.36), the notion of weak−∗ convergence can be reformulated as follows for the dual spaces $\ell^\infty = (\ell^1)^*$, $\ell^1 = (c_0)^*$ and $L^\infty(X,\mu) = (L^1(X,\mu))^*$ (if μ is a σ-finite measure).

- Let $x^{(n)}$, $x \in \ell^\infty$, $n = 1, 2, \ldots$. Then, setting $x^{(n)} = \big(x_k^{(n)}\big)_k$ and $x = (x_k)_k$, we have

$$x^{(n)} \overset{*}{\rightharpoonup} x \iff \sum_{k=1}^\infty x_k^{(n)} y_k \to \sum_{k=1}^\infty x_k y_k \quad \forall y = (y_k)_k \in \ell^1.$$

- Let $x^{(n)}$, $x \in \ell^1$, $n = 1, 2, \ldots$. Then, setting $x^{(n)} = \big(x_k^{(n)}\big)_k$ and $x = (x_k)_k$, we have

$$x^{(n)} \overset{*}{\rightharpoonup} x \iff \sum_{k=1}^\infty x_k^{(n)} y_k \to \sum_{k=1}^\infty x_k y_k \quad \forall y = (y_k)_k \in c_0.$$

- Let $(X, \mathscr{E}, \mu)$ be a σ-finite measure space, and let $(f_n)_n \subset L^\infty(X,\mu)$, $f \in L^\infty(X,\mu)$. Then

$$f_n \overset{*}{\rightharpoonup} f \iff \int_X f_n g\, d\mu \to \int_X f g\, d\mu \quad \forall g \in L^1(X,\mu).$$

Example 6.97 In $L^\infty(-1,1) = (L^1(-1,1))^*$ consider the sequence of functions

$$f_n(t) = e^{-nt^2}, \quad t \in [-1,1].$$

The Dominated Convergence Theorem ensures that

$$\int_{-1}^1 f_n g\, dx \to 0 \quad \forall g \in L^1(-1,1),$$

and this, thanks to Example 6.96, is equivalent to $f_n \overset{*}{\rightharpoonup} 0$. But $f_n \not\rightharpoonup 0$. Indeed, proceeding as in Example 6.63(1), among the functionals of $(L^\infty(-1,1))^*$ we find the extension of the Dirac delta at the origin, which is a bounded linear functional on $\mathscr{C}([-1,1])$:

$$\delta_0(f) = f(0) \quad \forall f \in \mathscr{C}([-1,1]).$$

If we label such an extension by T, we have $\langle T, f_n \rangle = f_n(0) = 1$.

Example 6.98 In $\ell^\infty = (\ell^1)^*$ consider the sequence $(x^{(n)})_n \subset \ell^\infty$ defined in (6.37). For every $y = (y_k)_k \in \ell^1$

$$\sum_{k=1}^\infty x_k^{(n)} y_k = \sum_{k=n+1}^\infty y_k \to 0 \quad (n \to \infty),$$

and this, thanks to Example 6.96, is equivalent to $x^{(n)} \overset{*}{\rightharpoonup} 0$. On the other hand $x^{(n)} \not\rightharpoonup 0$. To see this, we proceed as in Example 6.63(2): let $F \in (\ell^\infty)^*$ be an extension of the following bounded linear functional

$$f(x) := \lim_{k\to\infty} x_k \qquad \forall x = (x_k)_k \in \tilde{c}$$

where $\tilde{c} := \{x = (x_k)_k \in \ell^\infty \bigm| \exists \lim_k x_k\}$. So we have

$$\langle F, x^{(n)} \rangle = \lim_{k\to\infty} x_k^{(n)} = 1 \qquad \forall n \geq 1.$$

Exercise 6.99 Show that if X is a Banach space, then every sequence $(f_n)_n \subset X^*$ which converges weakly$-*$ is bounded.
Hint. Use the Banach-Steinhaus Theorem.

Exercise 6.100 Given a Banach space X, let $f_n,\ f \in X^*$ and $x_n,\ x \in X$ $(n \in \mathbb{N})$.

1. Show that if $x_n \rightharpoonup x$ and $f_n \to f$, then $\langle f_n, x_n \rangle \to \langle f, x \rangle$ as $n \to \infty$.
2. Show that if $x_n \to x$ and $f_n \overset{*}{\rightharpoonup} f$, then $\langle f_n, x_n \rangle \to \langle f, x \rangle$ as $n \to \infty$.

Exercise 6.101 Let $(a_n)_n$ be a sequence of real numbers and let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} a_n & \text{if } x \in \left[n, n + \dfrac{1}{n}\right], \\ 0 & \text{otherwise.} \end{cases}$$

Show that:

- If $\left(\frac{|a_n|^p}{n}\right)_n$ is bounded and $1 < p < \infty$, then $f_n \rightharpoonup 0$ in $L^p(\mathbb{R})$.
- If $a_n = n$, then $(f_n)_n$ does not converge weakly in $L^1(\mathbb{R})$.
- If $(a_n)_n$ is bounded, then $f_n \overset{*}{\rightharpoonup} 0$ in $L^\infty(\mathbb{R})$.

The following result yields a sort of weak$-*$ Bolzano-Weierstrass property in dual spaces.

Theorem 6.102 (Banach-Alaoglu) *Let X be a separable normed linear space. Then every bounded sequence $(f_n)_n \subset X^*$ has a weakly$-*$ convergent subsequence.*

Proof Let $(x_n)_n$ be a dense sequence in X and let $C \geq 0$ be such that $\|f_n\|_* \leq C$ for every $n \in \mathbb{N}$. Then $|\langle f_n, x_1 \rangle| \leq C\|x_1\|$. So, since the sequence $(\langle f_n, x_1 \rangle)_n$ is bounded in $\mathbb{R}$, there exists a subsequence of $(f_n)_n$, say $(f_{1,n})_n$, such that $\langle f_{1,n}, x_1 \rangle$ is convergent. Since $|\langle f_{1,n}, x_2 \rangle| \leq C\|x_2\|$, there exists a subsequence $(f_{2,n})_n \subset (f_{1,n})_n$, such that $\langle f_{2,n}, x_2 \rangle$ is convergent. Iterating this process, for any $k \geq 1$ we can construct nested subsequences

$$(f_{k,n})_n \subset (f_{k-1,n})_n \subset \cdots \subset (f_{1,n})_n \subset (f_n)_n$$

such that $\langle f_{k,n}, x_k \rangle$ is convergent as $n \to \infty$ for every $k \geq 1$. Define, for every $n \geq 1$, $g_n = f_{n,n}$. Then $(g_n)_n \subset (f_n)_n$ and $(\langle g_n, x_k \rangle)_n$ is convergent for every $k \geq 1$ since, for $n \geq k$, it is a subsequence of $(\langle f_{k,n}, x_k \rangle)_n$.

Let us complete the proof by showing that $(\langle g_n, x \rangle)_n$ is convergent for every $x \in X$. Fix $x \in X$ and $\varepsilon > 0$. Then there exist $k_\varepsilon, n_\varepsilon \geq 1$ such that

$$\begin{cases} \|x - x_{k_\varepsilon}\| < \varepsilon, \\ |\langle g_n, x_{k_\varepsilon} \rangle - \langle g_m, x_{k_\varepsilon} \rangle| < \varepsilon \quad \forall m, n \geq n_\varepsilon. \end{cases}$$

Therefore, for all $m, n \geq n_\varepsilon$,

$$\begin{aligned} |\langle g_n, x \rangle - \langle g_m, x \rangle| \leq & \underbrace{|\langle g_n, x \rangle - \langle g_n, x_{k_\varepsilon} \rangle| + |\langle g_m, x_{k_\varepsilon} \rangle - \langle g_m, x \rangle|}_{\leq 2C\|x - x_{k_\varepsilon}\|} \\ & + |\langle g_n, x_{k_\varepsilon} \rangle - \langle g_m, x_{k_\varepsilon} \rangle| \leq (2C + 1)\varepsilon. \end{aligned}$$

Thus, $(\langle g_n, x \rangle)_n$ is a Cauchy sequence satisfying $|\langle g_n, x \rangle| \leq C\|x\|$ for all $x \in X$. This implies that $f(x) := \lim_n \langle g_n, x \rangle$ is an element of X^*. □

The following result ensures that reflexive Banach spaces have the weak Bolzano-Weierstrass property.

Theorem 6.103 *Let X be a reflexive Banach space. Then every bounded sequence has a weakly convergent subsequence.*

Proof Let $(x_n)_n \subset X$ be a bounded sequence and let M be the closed subspace generated by x_n, that is, the closure of the set of all finite linear combinations of vectors x_n. By construction, M is separable (the finite linear combinations of vectors x_n with rational coefficients are a countable dense set in M). Moreover, in view of Proposition 6.72, M is reflexive. Therefore, by Corollary 6.71, M^* is also separable and reflexive. Consider the sequence $(J_M(x_n))_n \subset M^{**}$. Since J_M is an isometry, we have $\|J_M(x_n)\|_* = \|x_n\|$, and so the sequence $(J_M(x_n))_n$ is bounded in M^{**}. Applying the Banach-Alaoglu Theorem, there exists a subsequence $(x_{k_n})_n$ such that $J_M(x_{k_n}) \overset{*}{\rightharpoonup} \overline{\phi} \in M^{**}$ as $n \to \infty$. The reflexivity of M ensures that $\overline{\phi} = J_M(\overline{x})$ for some $\overline{x} \in M$. Therefore, for every $f \in M^*$,

$$\langle f, x_{k_n} \rangle = \langle J_M(x_{k_n}), f \rangle \to \langle J_M(\overline{x}), f \rangle = \langle f, \overline{x} \rangle \quad \text{as} \quad n \to \infty.$$

Finally, for any $F \in X^*$ we have $F_{\big|M} \in M^*$. Then

$$\langle F, x_{k_n} \rangle = \langle F_{\big|M}, x_{k_n} \rangle \to \langle F_{\big|M}, \overline{x} \rangle = \langle F, \overline{x} \rangle \quad \text{as} \quad n \to \infty.$$

So $x_{k_n} \rightharpoonup \overline{x}$. □

Example 6.104 If the space X is not reflexive, then the result of Theorem 6.103 is false, in general, as shown by the following two examples in $X = L^1(0, 1)$ and $X = \ell^1$, respectively.

1. Consider the sequence $f_n = n\chi_{(0,\frac{1}{n})}$ in $L^1(0, 1)$. $(f_n)_n$ is bounded since $\|f_n\|_1 = 1$. We are going to show that f_n does not converge weakly in $L^1(0, 1)$. Indeed, assume, by contradiction, that $f_n \rightharpoonup f$ for some $f \in L^1(0, 1)$. Then, recalling Example 6.74,
$$n\int_0^{1/n} g dt \to \int_0^1 f g dt \quad \forall g \in L^\infty(0, 1)$$
as $n \to \infty$. Observe that if we take $g \in \mathscr{C}([0, 1])$, then $n\int_0^{1/n} g(t) \to g(0)$, and so
$$\int_0^1 f g dt = g(0) \quad \forall g \in \mathscr{C}([0, 1]).$$
In particular $\int_0^1 f e^{-nt^2} dx = 1$ for all $n \geq 1$. On the other hand the Dominated Convergence Theorem implies
$$\int_0^1 f(t) e^{-nt^2} dx \to 0$$
and the contradiction follows. Since the above argument can be repeated for any subsequence, we conclude that $(f_n)_n$ does not admit a weakly convergent subsequence.
2. In ℓ^1 consider the sequence $(e_n)_n \subset \ell^1$ defined in (6.29), which is bounded since $\|e_n\|_1 = 1$. Assume, by contradiction, that $e_n \rightharpoonup x$ for some $x = (x_k)_k \in \ell^1$. Then, recalling Example 6.74,
$$y_n \to \sum_{k=1}^{\infty} x_k y_k \quad \forall y = (y_k)_k \in \ell^\infty$$
as $n \to \infty$. If $y = (y_k)_k$ is a convergent sequence, then $y \in \ell^\infty$ and $y_n \to \lim_{k\to\infty} y_k$, and so
$$\sum_{k=1}^{\infty} x_k y_k = \lim_{k\to\infty} y_k \quad \forall y \in \tilde{c}, \tag{6.44}$$
where $\tilde{c} = \{y = (y_k) \in \ell^\infty \mid \exists \lim_{k\to\infty} y_k\}$. In particular, if we consider the vectors $x^{(n)} \in \tilde{c}$ defined in (6.37) we obtain
$$\lim_{k\to\infty} x_k^{(n)} = 1 \quad \forall n \in \mathbb{N}.$$

On the other hand

$$\sum_{k=1}^{\infty} x_k x_k^{(n)} = \sum_{k=n+1}^{\infty} x_k \to 0$$

as $n \to \infty$ in contradiction with (6.44) applied to $y = x^{(n)}$. Since the above argument can be repeated for any subsequence, we conclude that $(e_n)_n$ does not admit a weakly convergent subsequence in ℓ^1.

We are now in a position to prove a fundamental theorem in the calculus of variations on the existence of minimum points, which is the infinite-dimensional version of the classical Weierstrass Theorem.

Theorem 6.105 *Let X be a reflexive Banach space and let $\varphi : X \to \mathbb{R}$ be a coercive*[14] *lower semicontinuous*[15] *convex function. Then φ has a minimum point in X.*

Proof Let $(x_n)_n \subset X$ be such that

$$\varphi(x_n) \to \inf_X \varphi.$$

The coecivity of φ implies that $(x_n)_n$ is bounded. Since X is a reflexive Banach space, Theorem 6.103 yields the existence of a subsequence $(x_{k_n})_n$ which converges weakly to a point $x_0 \in X$. Let α be any real number larger than $\inf_X \varphi$ and set $A_\alpha = \{x : \varphi(x) \leq \alpha\}$. Then the set A_α is convex (since φ is convex), closed (since φ is lower semicontinuous) and nonempty. We claim that $x_0 \in A_\alpha$.[16] Otherwise, by Corollary 6.55, there exists $f \in X^*$ such that $\sup_{x \in A_\alpha} \langle f, x \rangle < \langle f, x_0 \rangle$; so, since $x_{k_n} \in A_\alpha$ for large n, we have $\limsup_n \langle f, x_{k_n} \rangle < \langle f, x_0 \rangle$, in contraddiction with $\langle f, x_{k_n} \rangle \to \langle f, x_0 \rangle$. Therefore $x_0 \in A_\alpha$, i.e., $\varphi(x_0) \leq \alpha$. The arbitrariness of α yields $\inf_X \varphi > -\infty$ and $\varphi(x_0) = \inf_X \varphi$. □

6.5 Miscellaneous Exercises

Exercise 6.106 For any $f \in L^2(0, 1)$ let Tf be the following function:

$$Tf : x \in [0, 1] \mapsto \int_0^x f(t)dt.$$

1. Show that Tf is a continuous function on $[0, 1]$.

---

[14] That is, $\lim_{\|x\| \to \infty} \varphi(x) = \infty$.

[15] See Appendix B.

[16] This fact is also a direct consequence of Exercise 6.93: A_α is weakly closed and so, since $x_{k_n} \in A_\alpha$ for large n, we have $x_0 \in A_\alpha$.

2. Show that the linear operator $T : L^2(0, 1) \to \mathscr{C}([0, 1])$ is bounded, where $\mathscr{C}([0, 1])$ is equipped with the uniform norm.
3. Compute the norm of T.

Exercise 6.107 Let $(X, \mathscr{E}, \mu)$ be a finite measure space, $(a_n)_n$ a sequence of real numbers, $(E_n)_n \subset \mathscr{E}$. Let $1 < p < \infty$ and set

$$f_n(x) = a_n \chi_{E_n}(x), \quad x \in X.$$

1. Show that if $f_n \rightharpoonup 0$ and $f_n \not\to 0$ in $L^p(X, \mu)$, then $\mu(E_n) \to 0$.
2. Give an example to show that the conclusion of part 1 is false, in general, without the assumption $\mu(X) < \infty$.

Exercise 6.108 Let $T : L^1(1, \infty) \to L^1(1, \infty)$ be the linear operator defined by

$$Tf(x) = f(x) - \frac{1}{x} f(x) \quad \forall f \in L^1(1, \infty).$$

Show that T is bounded and compute $\|T\|$.

Exercise 6.109 Let $f_n : \mathbb{R} \to \mathbb{R}$ be the sequence defined by

$$f_n(x) = \begin{cases} nx & \text{if } -\frac{1}{n} < x < \frac{1}{n}, \\ -1 & \text{if } x \le -\frac{1}{n}, \\ 1 & \text{if } x \ge \frac{1}{n}. \end{cases}$$

Show that f_n does not converge strongly in $L^\infty(\mathbb{R})$ whereas $f_n \overset{*}{\rightharpoonup} \operatorname{sign} x$ in $L^\infty(\mathbb{R})$.

Exercise 6.110 1. Let (a, b) be a subinterval of $(-\pi, \pi)$. Show that

$$\lim_{n\to\infty} \int_a^b \cos nx \, dx = 0.$$

2. Deduce by part 1 that, if $E \subset (-\pi, \pi)$ is a Borel set, then

$$\lim_{n\to\infty} \int_E \cos nx \, dx = 0.$$

3. Conclude that $\cos nx$ converges weakly to zero in $L^p(-\pi, \pi)$ for every $1 \le p < \infty$ and $\cos nx \overset{*}{\rightharpoonup} 0$ in $L^\infty(\mathbb{R})$.
4. Show that $\cos nx$ does not converge weakly in $L^\infty(-\pi, \pi)$.

Exercise 6.111 For any $f \in L^1(\mathbb{R})$ set

$$Tf(x) = \int_0^x f(t) \arctan t \, dt, \quad x \in \mathbb{R}.$$

Show that:

1. $Tf \in \mathscr{C}(\mathbb{R}) \cap L^\infty(\mathbb{R})$ for every $f \in L^1(\mathbb{R})$.
2. The linear operator $T : L^1(\mathbb{R}) \to L^\infty(\mathbb{R})$ is bounded.
3. $\|T\| = \frac{\pi}{2}$.

Exercise 6.112 Let $(X, \mathscr{E}, \mu)$ be a finite measure space, $(a_n)_n$ a sequence of real numbers and $(E_n)_n \subset \mathscr{E}$. Set $f_n = a_n \chi_{E_n}$. Show that

$$f_n \overset{*}{\rightharpoonup} 0 \text{ in } L^\infty(X, \mu) \iff f_n \xrightarrow{L^\infty} 0.$$

Give an example to show that the above equivalence is false, in general, without assuming $\mu(X) < \infty$.

Exercise 6.113 Let X be a normed linear space.

1. Given $y \in X$ and $f \in X^*$ show that, setting

$$\Lambda x = \langle f, x \rangle \, y \qquad \forall x \in X,$$

   then Λ is a bounded linear operator from X into itself and $\|\Lambda\| = \|f\|_* \, \|y\|$.
2. Let $x_0 \in X$ be a nonzero vector. Show that for every $y \in X$ there exists $\Lambda_y \in \mathscr{L}(X)$ such that

$$\Lambda_y x_0 = y \qquad \text{and} \qquad \|\Lambda_y\| = \frac{\|y\|}{\|x_0\|}.$$

3. Show that if $\mathscr{L}(X)$ is complete, then X is also complete.

Exercise 6.114 For any $f \in L^1(0, 1)$ set

$$Tf(x) = \int_0^x t^2 f(t) \, dt, \quad x \in [0, 1].$$

Show that:

1. $Tf \in \mathscr{C}([0, 1])$ for every $f \in L^1(0, 1)$.
2. The linear operator $T : L^1(0, 1) \to \mathscr{C}([0, 1])$ is continuous, where $\mathscr{C}([0, 1])$ is equipped with the uniform norm.
3. $\|T\| = 1$.

Exercise 6.115 Let $x = (x_k)_k \in \ell^\infty$ and, for any $n \geq 1$, define $y^{(n)} = (y_k^{(n)})_k$ by

$$y_k^{(n)} = \begin{cases} 0 & \text{if } k \leq n, \\ x_{k-n} & \text{if } k > n. \end{cases}$$

Show that:

(i) $y^{(n)} \in \ell^\infty$ for all $n \geq 1$.
(ii) $y^{(n)} \overset{*}{\rightharpoonup} 0$ in ℓ^∞.
(iii) $y^{(n)}$ does not converge weakly in ℓ^∞, in general.

Exercise 6.116 Let $T : \ell^1 \to \ell^1$ be the linear operator defined by

$$T\big((x_n)_n\big) = \left(\frac{n^2}{e^n} x_n\right)_n \quad \forall (x_n)_n \in \ell^1.$$

Show that T is bounded and compute $\|T\|$.

Exercise 6.117 Let $1 \leq p \leq \infty$ and let $f_n : \mathbb{R}^N \to \mathbb{R}$ be a bounded sequence in $L^p(\mathbb{R}^N)$ such that

$$f_n \to 0 \text{ uniformly on compact sets of } \mathbb{R}^N.$$

1. Show that $f_n \rightharpoonup 0$ in $L^p(\mathbb{R}^N)$ if $1 < p < \infty$.
2. Show that $f_n \overset{*}{\rightharpoonup} 0$ in $L^\infty(\mathbb{R}^N)$ if $p = \infty$.
3. Give an example to show that f_n does not converge weakly if $p = 1$, in general.

Exercise 6.118 Let $T : \ell^1 \to \ell^1$ be the linear operator defined by

$$T\Big((x_n)_n\Big) = \Big(x_n \cos \frac{1}{n}\Big)_n \quad \forall (x_n)_n \in \ell^1.$$

Show that T is bounded and compute $\|T\|$.

Exercise 6.119 Let $(f_n)_n$ be a bounded sequence in $L^2(0, 1)$ such that

$$\int_0^x f_n \, dt \to 0 \quad \forall x \in (0, 1).$$

Show that

$$f_n \rightharpoonup 0 \text{ in } L^2(0, 1).$$

Exercise 6.120 Let $f_n, f \in L^2(0, 1)$, $g_n, g \in L^\infty(0, 1)$ be such that

$$f_n \rightharpoonup f \text{ in } L^2(0, 1), \qquad g_n \to g \text{ in } L^\infty(0, 1).$$

Show that $f_n g_n$ converges weakly to fg in $L^2(0, 1)$.

Exercise 6.121 Let $f \in L^1(a, b)$.

1. Show that there exists a sequence $(f_n)_n \subset \mathscr{C}([a, b])$ such that

$$|f_n| \leq 1, \quad m\left(f_n \neq \frac{|f|}{f}\chi_{\{f \neq 0\}}\right) \to 0,$$

where m denotes the Lebesgue measure on (a, b).
2. Let $\mathscr{C}([a, b])$ be equipped with the uniform norm. Deduce that the linear functional defined by

$$\Lambda : \mathscr{C}([a, b]) \to \mathbb{R}, \qquad \Lambda(g) = \int_a^b f g \, dx \quad \forall g \in \mathscr{C}([a, b])$$

is bounded and $\|\Lambda\|_* = \|f\|_1$.

Exercise 6.122 Let $1 \leq p < \infty$ and let f_n be a bounded sequence in $L^p(\mathbb{R}^N)$.

1. Show that $m(|f_n| \geq n) \to 0$.
2. Show that $f_n\chi_{\{|f_n| \geq n\}} \rightharpoonup 0$ in $L^p(\mathbb{R}^N)$ if $1 < p < \infty$.
3. Give an example to show that if $p = 1$, then $f_n\chi_{\{|f_n| \geq n\}}$ does not converge weakly in $L^1(\mathbb{R}^N)$, in general.

Exercise 6.123 For any $f \in L^\infty(0, \infty)$ set

$$Tf(x) = \int_0^x e^{y-x} f(y) dy \qquad x \geq 0\,.$$

1. Show that $Tf \in L^\infty(0, \infty)$ for all $f \in L^\infty(0, \infty)$.
2. Show that the linear operator $T : L^\infty(0, \infty) \to L^\infty(0, \infty)$ is bounded.
3. Compute $\|T\|$.
4. Is T injective and/or onto?

Exercise 6.124 Let $\mathscr{C}([0, 1])$ be equipped with the uniform norm and let $T : \mathscr{C}([0, 1]) \to \mathbb{R}$ be the linear functional defined by

$$T(f) = \int_0^1 f(x^2) dx \qquad \forall f \in \mathscr{C}([0, 1]).$$

1. Show that T is bounded.
2. Compute $\|T\|$.
3. Now consider $\mathscr{C}([0, 1])$ as a subspace of $L^4(0, 1)$ with the integral norm $\|\cdot\|_4$. Show that T has a unique extension to a bounded linear functional on $L^4(0, 1)$. Find the element $g_T \in L^{4/3}(0, 1)$, the dual of $L^4(0, 1)$, which represents such an extension.
Hint. Make the substitution $x^2 = y$.

Exercise 6.125 Let $T > 0$. Consider the linear operator $V : L^1(0, T) \to L^1(0, T)$ defined by

$$Vf(t) = \int_0^t f(s)ds \qquad t \in (0, T), \ \ \forall f \in L^1(0, T)\,.$$

1. Show that V is bounded and compute $\|V\|$.
2. Show that V maps weak convergence into strong convergence, i.e., for any $f_n, f \in L^1(0, T)$

$$f_n \rightharpoonup f \quad \Longrightarrow \quad Vf_n \to Vf\,.$$

   Hint. Prove, first, that

$$f_n \rightharpoonup f \quad \Longrightarrow \quad Vf_n(t) \to Vf(t) \quad \forall t \in [0, T]\,.$$

Exercise 6.126 Let $1 \le p \le \infty$. For any $f \in L^p(0, 1)$ set

$$A_p f(x) = xf(x) \qquad x \in (0, 1)\,.$$

1. Show that $A_p f \in L^p(0, 1)$ for every $f \in L^p(0, 1)$.
2. Show that $A_p \in \mathscr{L}\big(L^p(0, 1)\big)$.
3. Compute $\|A_p\|$.

Exercise 6.127 Let $1 \le p \le \infty$. For any $f \in L^p(0, \infty)$ set

$$\Lambda_p f(x) = \frac{f(x)}{1 + x^2}, \qquad x > 0\,.$$

1. Show that $\Lambda_p f \in L^p(0, \infty)$ for every $f \in L^p(0, \infty)$.
2. Show that $\Lambda_p \in \mathscr{L}\big(L^p(0, \infty)\big)$.
3. Compute $\|\Lambda_p\|$.

Exercise 6.128 Let $(e_n)_n$ be an orthonormal basis of a separable Hilbert space H.

1. Let $\bar{x}$ be a point of the unit sphere $S_1 = \{x \in H \ : \ \|x\| = 1\}$ and $\lambda \in [0, 1]$. For any $n \geqslant 1$ set

$$x_n(\lambda) = \lambda\bar{x} + (1 - \lambda)e_n.$$

   Compute the weak limit of $x_n(\lambda)$ in H and the limit of the sequence (of real numbers) $\|x_n(\lambda)\|$.
2. Analyse the weak convergence of the sequence $\frac{x_n(\lambda)}{\|x_n(\lambda)\|}$ in H.
3. Deduce that the weak closure of S_1 is given by the closed ball

$$\overline{B}_1 = \{x \in H \ : \ \|x\| \leqslant 1\}\,.$$

Exercise 6.129 Let X be a Banach space. A set $S \subset X$ is said to be (strongly) bounded if $\sup_{x \in S} \|x\| < \infty$. Similarly, S is said to be *weakly* bounded if

$$\sup_{x\in S}\big|\langle\phi, x\rangle\big| < \infty \qquad \forall\,\phi\in X^*\,.$$

1. Show that S is bounded if and only if S is weakly bounded.
   Hint. Consider the set $J_X(S)$ where $J_X : X \to X^{**}$ is the linear isometry defined in (6.38).
2. Does the above property hold in general for a normed linear space?

Exercise 6.130 Let X be a reflexive Banach space and let $A : X \to X$ be a linear operator which maps weak convergence into strong convergence, that is, for any $x_n, x \in X$

$$x_n \rightharpoonup x \Longrightarrow Ax_n \to Ax\,.$$

Show that

$$\|A\| = \max_{\|x\|=1}\|Ax\|\,.$$

Exercise 6.131 Let K be a nonempty convex closed set of a reflexive Banach space. Show that K has an element of minimum norm, that is, there exists $x \in K$ such that

$$\|x\| \le \|y\| \quad \forall y \in K\,.$$

Exercise 6.132 Let H be a Hilbert space and let $A \in \mathscr{L}(H)$. Show that A maps weak convergence into strong convergence (in the sense of Exercise 6.130) if and only if

$$x_n \rightharpoonup x \quad\Longrightarrow\quad \|Ax_n\| \to \|Ax\|\,.$$

Exercise 6.133 Let $\mathscr{C}([0,1])$ be equipped with the uniform norm and for any $u \in \mathscr{C}([0,1])$ set

$$Au(t) = \int_0^t e^{t-s}u(s)ds \qquad (t\in[0,1])\,.$$

1. Show that $Au \in \mathscr{C}([0,1])$ for every $u \in \mathscr{C}([0,1])$.
2. Show that $A : \mathscr{C}([0,1]) \to \mathscr{C}([0,1])$ is a bounded linear operator.
3. Compute $\|A\|$.
4. Show that A is a compact operator.[17]
5. Is A injective and/or onto?

Exercise 6.134 Let M be a closed subspace of $L^1(0,1)$ with the property that

$$\forall\, f\in M \quad \exists\, p>1 \quad \text{such that} \quad f\in L^p(0,1)\,.$$

[17] A linear operator $A : X \to Y$ is said to be compact if it maps any bounded subset of X into a relatively compact subset of Y. Clearly, any compact operator is also bounded.

1. Show that for every $n \geqslant 1$ the set

$$F_n = \left\{ f \in F \ : \ \|f\|_{1+\frac{1}{n}} \leqslant n \right\}$$

is closed in $L^1(0,1)$.
2. Using Baire's Lemma, show that $F \subset L^p(0,1)$ for some $p > 1$.

Exercise 6.135 Let $\mathscr{C}([0,1])$ be equipped with the uniform norm and set

$$A = \left\{x \in \mathscr{C}([0,1]) \ : \ x(t) > 0 \ \ \forall t \in [0,1]\right\}$$
$$B = \left\{x \in \mathscr{C}([0,1]) \ : \ x(t) \leqslant 0 \ \ \forall t \in [0,1]\right\}.$$

1. Show that there exists a functional $f \in X^*$ and a number $\alpha \in \mathbb{R}$ such that

$$\langle f, x\rangle > \alpha \geqslant \langle f, y\rangle \qquad \forall x \in A\,, \ \forall y \in B\,. \tag{6.45}$$

2. Give an example of $f \in X^*$ and $\alpha \in \mathbb{R}$ satisfying (6.45).
3. Show that (6.45) fails in its stronger form

$$\langle f, x\rangle \geqslant \alpha > \beta \geqslant \langle f, y\rangle \qquad \forall x \in A\,, \ \forall y \in B\,,$$

where $\alpha, \beta \in \mathbb{R}$.

Exercise 6.136 Let X and Y be Banach spaces and let $A : X \to Y$ be a compact operator.[18]

1. Show that if

$$\inf_{\|x\|=1} \|Ax\| > 0\,,$$

then the unit sphere $S = \left\{x \in X \ : \ \|x\| = 1\right\}$ is compact. Deduce that if $\dim X = \infty$, then

$$\inf_{\|x\|=1} \|Ax\| = 0\,.$$

2. Give an example where the above infimum in (6.136) is not a minimum.

Exercise 6.137 Let $(e_n)_n$ be an orthonormal basis of a Hilbert space H and let $\lambda = (\lambda_n)_n \in \ell^\infty$.

1. Show that

$$\sum_{n=1}^{\infty} \lambda_n \langle x, e_n\rangle e_n \ \in \ H \qquad \forall\, x \in H.$$

[18]See footnote 17.

2. Setting

$$\Lambda x = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \ \in \ H \qquad \forall\, x \in H,$$

show that $\Lambda \in \mathscr{L}(H)$ and compute $\|\Lambda\|$.
3. Show that if $\lambda \in c_0$ (i.e., $\lambda_n \to 0$ as $n \to \infty$), then Λ maps weak convergence into strong convergence (in the sense of Exercise 6.130).

Exercise 6.138 Let H be an infinite-dimensional separable Hilbert space. Given an orthonormal basis $(e_n)_n$ of H, for any $\lambda = (\lambda_n)_n \in \ell^\infty$ let $F_\lambda \in \mathscr{L}(H)$ be the operator defined in the previous exercise:

$$F_\lambda x := \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n \qquad \forall\, x \in H.$$

1. Show that the map $\lambda \mapsto F_\lambda$ is a linear isometry from ℓ^∞ to $\mathscr{L}(H)$.
2. Deduce that $\mathscr{L}(H)$ is *not* separable.

Exercise 6.139 Let $(e_n)_n$ be an orthonormal basis of a Hilbert space H and $(\alpha_n)_n$ a bounded sequence of real numbers. Setting

$$x_n = \frac{1}{n} \sum_{k=1}^{n} \alpha_k e_k,$$

show that:

1. $x_n \to 0$.
2. $\sqrt{n} x_n \rightharpoonup 0$.

Hint. To prove part (b), observe that, given n real numbers $a_1, \ldots, a_n$ and an integer $n_0 \in \{1, \ldots, n\}$, we have

$$\Big| \sum_{k=1}^{n} a_k \Big| \leqslant \sqrt{n_0} \Big(\sum_{k=1}^{n_0} a_k^2 \Big)^{\frac{1}{2}} + \sqrt{n - n_0} \Big(\sum_{k=n_0+1}^{n} a_k^2 \Big)^{\frac{1}{2}}.$$

Exercise 6.140 Let H be a Hilbert space and let $K \subset H$ be a nonempty convex closed bounded set. Given a sequence $(x_n)_n \subset H$, set

$$\overline{x}_n = p_K(x_n) \qquad (n \in \mathbb{N}),$$

where p_K denotes the orthogonal projection on K.

1. Show that $(\overline{x}_n)_n$ has a subsequence $(\overline{x}_{n_k})_k$ which converges weakly to a point $\overline{x} \in K$.
2. Show that if $(x_{n_k})_k$ converges strongly to x, then $\overline{x} = p_K(x)$.
   Hint. Recall the variational inequality which characterizes the projection.
3. Is identity $\overline{x} = p_K(x)$ still true if $(x_{n_k})_k$ converges weakly to x?
   Hint. Consider $H = L^2(-\pi, \pi)$ and take $K = \{f \in L^2(-\pi, \pi) \mid f \geq 0 \text{ a.e.}\}$ and $f_n(x) = \frac{\cos nx}{\sqrt{\pi}}$.

Exercise 6.141 Let $1 \leq p \leq \infty$ and let $F : L^p(0, \infty) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^\infty f(x)xe^{-x}dx.$$

1. Show that F is bounded.
2. Compute the norm of F for $p = 1$ and $p = \infty$.

Exercise 6.142 Let $1 \leq p < \infty$ and let $F : L^p(0, \infty) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^\infty \frac{f(x)}{1+x}dx.$$

Show that F is bounded and compute $\|F\|_*$.

Exercise 6.143 Let $1 \leq p \leq \infty$ and let $F : L^p(0, \infty) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_1^\infty e^{-x} f(x)dx + \int_0^1 f(x)dx.$$

Show that F is bounded and compute $\|F\|_*$.

Exercise 6.144 Let $F : L^1(0, \infty) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^1 xf(x)dx - \int_2^\infty \arctan x\, f(x)dx.$$

Show that F is bounded and compute $\|F\|_*$.

Exercise 6.145 Let $2 < p \leq \infty$ and let $F : L^p(0, \infty) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^{+\infty} \Big(\frac{f(x)}{\sqrt{x}}\chi_{(0,1)} + \frac{f(x)}{1+x^2}\Big)dx.$$

1. Show that F is bounded.
2. Compute $\|F\|_*$ for $p = \infty$.

Exercise 6.146 Let $1 \le p \le \infty$ and let $F : \ell^p \to \mathbb{R}$ be the linear functional defined by

$$F(x) = x_1 - 2x_2 + 6x_4 \quad \forall x = (x_n)_n \in \ell^p.$$

Show that F is bounded and compute $\|F\|_*$.

Exercise 6.147 Let $1 < p \le \infty$ and let $F : L^p(0,1) \to \mathbb{R}$ be the linear functional defined by

$$F(f) = \int_0^1 f(x) \log x \, dx \quad \forall f \in L^p(0,1).$$

1. Show that F is bounded.
2. Compute $\|F\|_*$ for $p = \infty$.

References

[Br83] Brezis, H.: Analyse fonctionelle: Théorie et applications. Masson, Paris (1983)

[Fl77] Fleming, W.: Functions of Several Variables. Springer, New York (1977)

[HS65] Hewitt, E., Stromberg, K.: Real and Abstract Analysis. Springer, New York (1965)

[Ko02] Komornik, V.: Précis d'analyse réelle. Analyse fonctionnelle, intégrale de Lebesgue, espaces fonctionnels, Mathématiques pour le 2^e cycle, vol. 2. Ellipses, Paris (2002)

[KW92] König, H., Witstock, G.: Non-equivalent complete norms and would-be continuous linear functionals. Expositiones Mathematicae **10**, 285–288 (1992)

[Mo69] Morawetz, C.: L^p inequalities. Bull. Am. Math. Soc. **75**, 1299–1302 (1969)

[Yo65] Yosida, K.: Functional Analysis. Springer, Berlin (1980)

Part III
Selected Topics

Chapter 7
Absolutely Continuous Functions

Let $f : [a, b] \to \mathbb{R}$ be a continuous function and let $F : [a, b] \to \mathbb{R}$ be continuously differentiable. Then the connection between derivation and integration is expressed by the well-known formulas

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x), \tag{7.1}$$

$$\int_a^x F'(t)\,dt = F(x) - F(a). \tag{7.2}$$

Thus, in Lebesgue's integration theory, it is natural to consider the following questions:

1. Is formula (7.1) still true almost everywhere for any function[1] $f \in L^1(a, b)$?
2. Can one characterize the largest class of functions verifying (7.2)?

In this chapter, we will answer the above questions. Let us observe that if f is positive, then Lebesgue's integral

$$\int_a^x f(t)\,dt, \quad x \in [a, b], \tag{7.3}$$

is an increasing function of the right end-point x. Moreover, since any summable function f is the difference of two positive summable functions, f^+ and f^-, the integral (7.3) is in turn the difference of two increasing functions. Therefore the study of (7.3) is strictly related to the study of monotone functions. Monotone functions enjoy several important properties that we now proceed to discuss.

[1] $L^1(a, b) = L^1([a, b], m)$ where m stands for the Lebesgue measure on $[a, b]$. See footnote 7 at p. 87.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_7

7.1 Monotone Functions

Definition 7.1 A function $f : [a, b] \to \mathbb{R}$ is said to be increasing if $a \leq x_1 \leq x_2 \leq b$ implies $f(x_1) \leq f(x_2)$ and decreasing if $a \leq x_1 \leq x_2 \leq b$ implies $f(x_1) \geq f(x_2)$. By a monotone function we mean a function which is either increasing or decreasing.

Definition 7.2 Given a monotone function $f : [a, b] \to \mathbb{R}$ and $x_0 \in [a, b)$, the limit

$$f(x_0^+) := \lim_{h \to 0,\, h > 0} f(x_0 + h)$$

is called the right-hand limit of f at the point x_0. Similarly, if $x_0 \in (a, b]$, the limit[2]

$$f(x_0^-) := \lim_{h \to 0,\, h > 0} f(x_0 - h)$$

is called the left-hand limit of f at x_0.

Remark 7.3 Let $f : [a, b] \to \mathbb{R}$ be an increasing function. If $a \leq x < y \leq b$, then

$$f(x^+) \leq f(y^-).$$

Similarly, if f is decreasing on $[a, b]$ and $a \leq x < y \leq b$, then

$$f(x^+) \geq f(y^-).$$

We now establish the basic properties of monotone functions.

Theorem 7.4 *Any monotone function $f : [a, b] \to \mathbb{R}$ is Borel and bounded, and hence summable.*

Proof Assume that f is increasing. Since $f(a) \leq f(x) \leq f(b)$ for all $x \in [a, b]$, f is clearly bounded. For any $c \in \mathbb{R}$ consider the set

$$E_c = \{x \in [a, b] \mid f(x) < c\}.$$

If E_c is empty, then E_c is (trivially) a Borel set. If E_c is nonempty, let y be the supremum of E_c. Then E_c is either the closed interval $[a, y]$, if $y \in E_c$, or the half-closed interval $[a, y)$, if $y \notin E_c$. In both cases, E_c is a Borel set; this proves that f is Borel. Finally, we have

$$\int_a^b |f(x)|\, dx \leq \max\{|f(a)|, |f(b)|\}(b - a),$$

by which it follows that f is summable. □

[2]Observe that such limits always exist and are finite.

Theorem 7.5 *Let $f : [a, b] \to \mathbb{R}$ be a monotone function. Then the set of all points of discontinuity of f is at most countable (i.e., countable or finite).*

Proof Suppose that f is increasing and let E be the set of points of discontinuity of f in (a, b). For $x \in E$ we have $f(x^-) < f(x^+)$; then to any point x of E we may associate a rational number $r(x)$ such that

$$f(x^-) < r(x) < f(x^+).$$

Since by Remark 7.3 $x_1 < x_2, x_1, x_2 \in E$, implies $f(x_1^+) \leq f(x_2^-)$, we deduce that $r(x_1) \neq r(x_2)$. We have thus established a bijective map between the set E and a subset of rational numbers. □

7.1.1 Differentiation of Monotone Functions

The aim of this section will be to show that a monotone function $f : [a, b] \to \mathbb{R}$ is differentiable almost everywhere on $[a, b]$. Before proving this result, due to Lebesgue, let us introduce some notation. For any $x \in (a, b)$ the following four quantities (which may take infinite values) always exist:

$$D'_L f(x) = \liminf_{h \to 0,\, h<0} \frac{f(x+h) - f(x)}{h}, \quad D''_L f(x) = \limsup_{h \to 0,\, h<0} \frac{f(x+h) - f(x)}{h},$$

$$D'_R f(x) = \liminf_{h \to 0,\, h>0} \frac{f(x+h) - f(x)}{h}, \quad D''_R f(x) = \limsup_{h \to 0,\, h>0} \frac{f(x+h) - f(x)}{h}.$$

These four quantities are called the *generalized derivatives* of f at x. It is clear that the following inequalities always hold

$$D'_L f(x) \leq D''_L f(x), \quad D'_R f(x) \leq D''_R f(x). \tag{7.4}$$

If $D'_L f(x)$ and $D''_L f(x)$ are equal and finite, their common value is the left-hand derivative of f at x. Similarly, if $D'_R f(x)$ and $D''_R f(x)$ are equal and finite, their common value is just the right-hand derivative of f at x. Moreover, f is differentiable at x if and only if all four generalized derivatives $D'_L f(x)$, $D''_L f(x)$, $D'_R f(x)$ and $D''_R f(x)$ are equal and finite.

Theorem 7.6 (Lebesgue) *Let $f : [a, b] \to \mathbb{R}$ be a monotone function. Then f is differentiable a.e. in $[a, b]$. Moreover,*[3] *$f' \in L^1(a, b)$ and*

$$\int_a^b |f'(t)|\, dt \leq |f(b) - f(a)|. \tag{7.5}$$

[3]Observe that, in general, f' is defined a.e. in $[a, b]$ (see Remark 2.74).

Proof We may assume, without loss of generality, that f is increasing, for, if f is decreasing, it suffices to apply the result to $-f$ which is obviously increasing. We begin by proving that the generalized derivatives of f are equal (possibly infinite) a.e. in $[a, b]$. It will be sufficient to show that the inequality

$$D'_L f(x) \geq D''_R f(x) \tag{7.6}$$

holds a.e. in $[a, b]$. Indeed, setting $f^*(x) = -f(-x)$, we get that f^* is increasing on $[-b, -a]$; moreover, it is easy to verify that

$$D'_L f^*(x) = D'_R f(-x), \quad D''_L f^*(x) = D''_R f(-x).$$

So, applying (7.6) to f^*, we deduce

$$D'_L f^*(x) \geq D''_R f^*(x)$$

or, equivalently,

$$D'_R f(x) \geq D''_L f(x).$$

Combining this inequality with (7.6), and using (7.4), we obtain

$$D''_R f \leq D'_L f \leq D''_L f \leq D'_R f \leq D''_R f,$$

and the a.e. equality of the four generalized derivatives is thus proved.

To show that (7.6) holds a.e., observe that, since the generalized derivatives are nonnegative, the set of points where $D'_L f < D''_R f$ can be represented as the union over $u,\ v \in \mathbb{Q}$ with $v > u > 0$ of the sets

$$E_{u,v} = \{x \in (a, b) \mid D''_R f(x) > v > u > D'_L f(x)\}.$$

So if we show that $m(E_{u,v}) = 0$ (where m denotes the Lebesgue measure on $[a, b]$), then it will follow that (7.6) is true a.e. Let $s = m(E_{u,v})$. Then, given $\varepsilon > 0$, thanks to Theorem 1.71 there exists an open set $V \subset (a, b)$ such that $E_{u,v} \subset V$ and $m(V) < s + \varepsilon$. For every $x \in E_{u,v}$ and $\delta > 0$, since $D'_L f(x) < u$, there exists $h_{x,\delta} \in (0, \delta)$ such that $[x - h_{x,\delta}, x] \subset V$ and

$$f(x) - f(x - h_{x,\delta}) < u h_{x,\delta}.$$

Since the family of closed intervals $([x - h_{x,\delta}, x])_{x \in E_{u,v},\, \delta > 0}$ is a fine cover of $E_{u,v}$, by Vitali's Covering Theorem[4] there exists a finite number of disjoint intervals of such a family, say

[4]See Appendix G.

$$I_1 := [x_1 - h_1, x_1], \dots, I_N := [x_N - h_N, x_N],$$

such that, setting $A = E_{u,v} \cap \bigcup_{i=1}^N (x_i - h_i, x_i)$, we have that

$$m(A) = m\left(E_{u,v} \cap \bigcup_{i=1}^N I_i\right) > s - \varepsilon.$$

Summing over all these intervals we obtain

$$\sum_{i=1}^N \big(f(x_i) - f(x_i - h_i)\big) < u \sum_{i=1}^N h_i \le u\, m(V) \le u(s + \varepsilon). \tag{7.7}$$

Let us argue as above using the inequality $D_R'' f(x) > v$: for every $y \in A$ and $\eta > 0$, since $D_R'' f(y) > v$, there exists $k_{y,\eta} \in (0, \eta)$ such that $[y, y + k_{y,\eta}] \subset I_i$ for some $i \in \{1, \dots, N\}$ and

$$f(y + k_{y,\eta}) - f(y) > v k_{y,\eta}.$$

Since the family of closed intervals $([y, y + k_{y,\eta}])_{y \in A,\, \eta > 0}$ is a fine cover of A, by Vitali's Covering Theorem there exists a finite number of disjoint intervals of such a family, say

$$J_1 := [y_1, y_1 + k_1], \dots, J_M := [y_M, y_M + k_M],$$

such that

$$m\left(A \cap \bigcup_{j=1}^M J_j\right) \ge m(A) - \varepsilon > s - 2\varepsilon.$$

Summing over all these intervals we deduce

$$\sum_{j=1}^M \big(f(y_j + k_j) - f(y_j)\big) > v \sum_{j=1}^M k_j = v\, m\left(\bigcup_{j=1}^M J_j\right) \ge v(s - 2\varepsilon). \tag{7.8}$$

For every $i \in \{1, \dots, N\}$, summing over all intervals J_j such that $J_j \subset I_i$, and, using the assumption that f is increasing, we obtain

$$\sum_{j,\, J_j \subset I_i} \big(f(y_j + k_j) - f(y_j)\big) \le f(x_i) - f(x_i - h_i).$$

Hence, summing over i and taking into account that every interval J_j is contained in some interval I_i,

$$\sum_{i=1}^{N}\big(f(x_i)-f(x_i-h_i)\big)\geq\sum_{i=1}^{N}\sum_{j,\,J_j\subset I_i}\big(f(y_j+k_j)-f(y_j)\big)$$
$$=\sum_{j=1}^{M}\big(f(y_j+k_j)-f(y_j)\big).$$

Owing to (7.7) and (7.8),

$$u(s+\varepsilon)\geq v(s-2\varepsilon).$$

The arbitrariness of ε implies $us\geq vs$; since $u<v$, then $s=0$. This proves that $m(E_{u,v})=0$, as claimed.

We have thus proved that the function

$$\Phi(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

is defined almost everywhere in $[a,b]$. Therefore f is differentiable at x if and only if $\Phi(x)$ is finite. Let

$$\Phi_n(x)=n\left(f\left(x+\frac{1}{n}\right)-f(x)\right)$$

where, to define Φ_n for every $x\in[a,b]$, we have set $f(x)=f(b)$ for $x\geq b$. Since f is summable on $[a,b]$, Φ_n is also summable. By integrating Φ_n we have

$$\int_a^b\Phi_n(x)dx=n\int_a^b\left(f\left(x+\frac{1}{n}\right)-f(x)\right)dx=n\left(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}}f(x)dx-\int_a^b f(x)dx\right)$$
$$=n\left(\int_b^{b+\frac{1}{n}}f(x)dx-\int_a^{a+\frac{1}{n}}f(x)dx\right)=f(b)-n\int_a^{a+\frac{1}{n}}f(x)dx$$
$$\leq f(b)-f(a),$$

where, in the last inequality, we have used the fact that f is increasing. By Fatou's Lemma it follows that

$$\int_a^b\Phi(x)\,dx\leq f(b)-f(a).$$

In particular, Φ is summable, and, consequently, almost everywhere finite. Then f is differentiable almost everywhere and $f'(x)=\Phi(x)$ for almost every $x\in[a,b]$. □

Example 7.7 It is easy to exhibit examples of monotone functions f for which (7.5) becomes a strict inequality. For instance, given $n+1$ points $a = x_0 < x_1 < \cdots < x_n = b$ and n numbers $h_1, h_2, \ldots, h_n$, consider the function

$$f(x) = \begin{cases} h_1 & \text{if } a \le x < x_1, \\ h_2 & \text{if } x_1 \le x < x_2, \\ \ldots & \\ h_n & \text{if } x_{n-1} \le x \le b. \end{cases}$$

A function of such a form is called a *step function*. If $h_1 < h_2 < \cdots < h_n$, then f is obviously increasing and

$$\int_a^b f'(x)\,dx = 0 < h_n - h_1 = f(b) - f(a).$$

Example 7.8 (**Cantor-Vitali function**) The function considered in the previous example is discontinuous. However, it is also possible to construct continuous increasing functions satisfying the strict inequality (7.5).

Consider the closed interval [0, 1] and delete the middle third

$$(a_1^1, b_1^1) = \left(\frac{1}{3}, \frac{2}{3}\right).$$

From the two remaining intervals $[0, \frac{1}{3}]$, $[\frac{2}{3}, 1]$ delete the middle thirds

$$(a_1^2, b_1^2) = \left(\frac{1}{9}, \frac{2}{9}\right), \quad (a_2^2, b_2^2) = \left(\frac{7}{9}, \frac{8}{9}\right);$$

from the four remaining intervals delete the middle thirds

$$(a_1^3, b_1^3) = \left(\frac{1}{27}, \frac{2}{27}\right), \quad (a_2^3, b_2^3) = \left(\frac{7}{27}, \frac{8}{27}\right),$$

$$(a_3^3, b_3^3) = \left(\frac{19}{27}, \frac{20}{27}\right), \quad (a_4^3, b_4^3) = \left(\frac{25}{27}, \frac{26}{27}\right)$$

and so on. Observe that the complement of the union of all intervals (a_k^h, b_k^h) is the Cantor set constructed in Example 1.63.

Let $f_0(x) = x$. For any $n \ge 1$ let $f_n : [0, 1] \to \mathbb{R}$ be the continuous function which satisfies

$$f_n(0) = 0, \quad f_n(1) = 1,$$

$$f_n(t) = \frac{2k-1}{2^h} \quad \text{if } t \in (a_k^h, b_k^h), \quad k = 1, \ldots, 2^{h-1}, \ h = 1, \ldots, n$$

and f_n increases linearly otherwise. For instance,

$$f_1(t) = \frac{1}{2} \quad \text{if } \frac{1}{3} < t < \frac{2}{3},$$

and f_1 increases linearly from 0 to $\frac{1}{2}$ in $[0, \frac{1}{3}]$ and from $\frac{1}{2}$ to 1 in $[\frac{2}{3}, 1]$. For f_2 we have

$$f_2(t) = \begin{cases} \frac{1}{4} & \text{if } \frac{1}{9} < t < \frac{2}{9}, \\ \frac{1}{2} & \text{if } \frac{1}{3} < t < \frac{2}{3}, \\ \frac{3}{4} & \text{if } \frac{7}{9} < t < \frac{8}{9}, \end{cases}$$

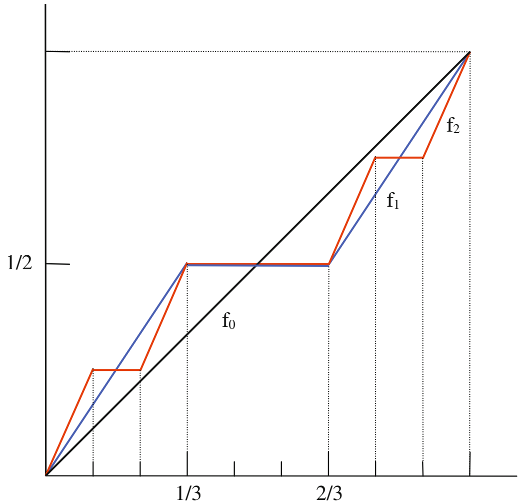
and f_2 increases linearly from 0 to $\frac{1}{4}$ in $[0, \frac{1}{9}]$, from $\frac{1}{4}$ to $\frac{1}{2}$ in $[\frac{2}{9}, \frac{1}{3}]$, from $\frac{1}{2}$ to $\frac{3}{4}$ in $[\frac{2}{3}, \frac{7}{9}]$, from $\frac{3}{4}$ to 1 in $[\frac{8}{9}, 1]$, and so on.

By construction, f_n is monotone, continuous, $f_n(0) = 0$, $f_n(1) = 1$ and $|f_n - f_{n+1}| \leq \frac{1}{2^{n+1}}$ (see Fig. 7.1). So, if $m > n$,

$$|f_m - f_n| \leq \sum_{k=n}^{m-1} |f_{k+1} - f_k| \leq \sum_{k=n}^{\infty} \frac{1}{2^{k+1}}.$$

Hence $(f_n)_n$ converges uniformly in $[0, 1]$. Let $f = \lim_n f_n$. Then f is continuous, monotone, and $f(t) = \frac{2k-1}{2^h}$ if $t \in (a_k^h, b_k^h)$. Such a function is the *Cantor-Vitali function*, also known as *Devil's staircase*. The derivative f' vanishes on every interval (a_k^h, b_k^h), and so $f'(x) = 0$ for almost every $x \in [0, 1]$, since the Cantor set has

Fig. 7.1 Graph of f_0, f_1, f_2



measure zero. It follows that

$$\int_0^1 f'(x)\,dx = 0 < 1 = f(1) - f(0).$$

7.2 Functions of Bounded Variation

Definition 7.9 A function $f : [a, b] \to \mathbb{R}$ is said to be of bounded variation if there exists a constant $C > 0$ such that

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \le C \tag{7.9}$$

for any partition

$$a = x_0 < x_1 < \cdots < x_n = b \tag{7.10}$$

of $[a, b]$. The *total variation* of f on $[a, b]$, denoted by $V_a^b(f)$, is the quantity:

$$V_a^b(f) = \sup \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \tag{7.11}$$

where the supremum is taken over all partitions (7.10) of the interval $[a, b]$.

Remark 7.10 By definition we have that, if $\alpha \in \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ is a function of bounded variation, then αf is also of bounded variation and

$$V_a^b(\alpha f) = |\alpha| V_a^b(f).$$

Example 7.11

1. If $f : [a, b] \to \mathbb{R}$ is a monotone function, then the left-hand side of (7.9) actually coincides with $|f(b) - f(a)|$ for any choice of partition. Then f is of bounded variation and $V_a^b(f) = |f(b) - f(a)|$.
2. If f is a *step function* of the type considered in Example 7.7, then, for any $h_1, \ldots, h_n \in \mathbb{R}$, f is of bounded variation and the total variation amounts to the sum of the sizes of the jumps, namely

$$V_a^b(f) = \sum_{k=1}^{n-1} |h_{k+1} - h_k|.$$

3. Let $f : [a, b] \to \mathbb{R}$ be a Lipschitz continuous function with Lipschitz constant K; then for any partition (7.10) of $[a, b]$ we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \leq K \sum_{k=0}^{n-1} |x_{k+1} - x_k| = K(b-a).$$

So f is of bounded variation and $V_a^b(f) \leq K(b-a)$.

Example 7.12 It is easy to exhibit examples of continuous functions which are not of bounded variation. Indeed, consider the function

$$f(x) = \begin{cases} x \sin \dfrac{1}{x} & \text{if } 0 < x \leq 1, \\ 0 & \text{if } x = 0 \end{cases}$$

and, fixed $n \in \mathbb{N}$, take the following partition of $[0, 1]$ associated to points $x_k = (\frac{\pi}{2} + k\pi)^{-1}$:

$$0,\ x_n,\ x_{n-1}, \ldots, x_1, x_0, 1.$$

The sum on the left-hand side of (7.9) for such a partition is given by

$$\frac{4}{\pi} \sum_{k=1}^{n} \frac{1}{2k+1} + \frac{2}{\pi} + \left| \sin 1 - \frac{2}{\pi} \right|.$$

Taking into account that $\sum_{k=1}^{\infty} \frac{1}{2k+1} = \infty$, we deduce that the supremum on the right-hand side of (7.11) taken over all partitions of $[0, 1]$ is infinite.

Proposition 7.13 *If $f, g : [a, b] \to \mathbb{R}$ are functions of bounded variation, then $f + g$ is also of bounded variation and*

$$V_a^b(f+g) \leq V_a^b(f) + V_a^b(g).$$

Proof For any partition of the interval $[a, b]$, we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) + g(x_{k+1}) - f(x_k) - g(x_k)|$$
$$\leq \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| + \sum_{k=0}^{n-1} |g(x_{k+1}) - g(x_k)| \leq V_a^b(f) + V_a^b(g).$$

Taking the supremum of the left-hand side over all partitions of $[a, b]$ we immediately get the conclusion. □

By Remark 7.10 and by Proposition 7.13 it follows that any finite linear combination of functions of bounded variation is itself a function of bounded variation. In other words, the set $BV([a, b])$ of all functions of bounded variation on the interval $[a, b]$ is a linear space (unlike the set of all monotone functions).

Proposition 7.14 *If $f : [a, b] \to \mathbb{R}$ is a function of bounded variation and $a < c < b$, then*

$$V_a^b(f) = V_a^c(f) + V_c^b(f).$$

Proof First we consider a partition of the interval $[a, b]$ such that c is one of the points of subdivision, say $x_r = c$. Then

$$\begin{aligned}\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \\ &= \sum_{k=0}^{r-1} |f(x_{k+1}) - f(x_k)| + \sum_{k=r}^{n-1} |f(x_{k+1}) - f(x_k)| \\ &\le V_a^c(f) + V_c^b(f).\end{aligned} \tag{7.12}$$

Now let us consider an arbitrary partition of $[a, b]$. It is clear that the sum $\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$ can never decrease by adding an extra point of subdivision to the partition. Therefore (7.12) holds for any partition of $[a, b]$, and so

$$V_a^b(f) \le V_a^c(f) + V_c^b(f).$$

On the other hand, fixed $\varepsilon > 0$, there exist partitions of the intervals $[a, c]$ and $[c, b]$, respectively, such that

$$\sum_i |f(x'_{i+1}) - f(x'_i)| > V_a^c(f) - \frac{\varepsilon}{2},$$

$$\sum_j |f(x''_{j+1}) - f(x''_j)| > V_c^b(f) - \frac{\varepsilon}{2}.$$

Combining all points of subdivision x'_i, x''_j, we obtain a partition of the interval $[a, b]$, with points of subdivision x_k, such that

$$\begin{aligned}V_a^b(f) \ge \sum_k |f(x_{k+1}) - f(x_k)| &= \sum_i |f(x'_{i+1}) - f(x'_i)| + \sum_j |f(x''_{j+1}) - f(x''_j)| \\ &> V_a^c(f) + V_c^b(f) - \varepsilon.\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows that $V_a^b(f) \ge V_a^c(f) + V_c^b(f)$. □

Corollary 7.15 *If $f : [a, b] \to \mathbb{R}$ is a function of bounded variation, then the function*

$$x \longmapsto V_a^x(f)$$

is increasing.

Proof Proposition 7.14 implies that

$$V_a^y(f) = V_a^x(f) + V_x^y(f) \geq V_a^x(f)$$

for all x, y satisfying $a \leq x < y \leq b$. □

Proposition 7.16 *A function $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if f can be represented as the difference of two increasing functions.*

Proof Since any monotone function is of bounded variation thanks to Example 7.11, and since the set $BV([a, b])$ is a linear space, we deduce that the difference between two increasing functions is of bounded variation. To prove the converse, set

$$g_1(x) = V_a^x(f), \quad g_2(x) = V_a^x(f) - f(x).$$

By Corollary 7.15 g_1 is an increasing function. We claim that g_2 is also increasing. Indeed, if $x < y$, then, using Proposition 7.14, we obtain

$$g_2(y) - g_2(x) = V_x^y(f) - (f(y) - f(x)). \tag{7.13}$$

By Definition 7.9 we have

$$|f(y) - f(x)| \leq V_x^y(f)$$

and so by (7.13) it follows $g_2(y) - g_2(x) \geq 0$. Writing $f = g_1 - g_2$, we get the desired representation of f as the difference between two increasing functions. □

Theorem 7.17 *Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation. Then the set of all points of discontinuity of f is at most countable. Moreover, f is differentiable a.e. in $[a, b]$, $f' \in L^1(a, b)$ and*

$$\int_a^b |f'(x)|\, dx \leq V_a^b(f). \tag{7.14}$$

Proof Invoking Theorems 7.5, 7.6 and Proposition 7.16 we conclude that f has at most countably many points of discontinuity, f is differentiable a.e. in $[a, b]$, and $f' \in L^1(a, b)$. Since, for all $a \leq x < y \leq b$,

$$|f(y) - f(x)| \leq V_x^y(f) = V_a^y(f) - V_a^x(f),$$

we deduce that

$$|f'(x)| \le (V_a^x(f))' \quad \text{a.e. in } [a, b].$$

Finally, by (7.5) we obtain

$$\int_a^b |f'(x)| dx \le \int_a^b (V_a^x(f))' dx \le V_a^b(f),$$

thus completing the proof. □

Remark 7.18 Step functions (Example 7.7) and the Cantor-Vitali function (Example 7.8) provide examples of functions of bounded variation verifying the strict inequality in (7.14).

Proposition 7.19 *A function $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if the cartesian curve*

$$y = f(x), \quad a \le x \le b,$$

is rectifiable.[5]

Proof For any partition of $[a, b]$ we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \le \sum_{k=0}^{n-1} \sqrt{(x_{k+1} - x_k)^2 + (f(x_{k+1}) - f(x_k))^2}$$
$$\le (b - a) + \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|.$$

By taking the supremum over all partitions we obtain the conclusion. □

Exercise 7.20 Show that if $f : [a, b] \to \mathbb{R}$ is a function of bounded variation, then $\sup_{x\in[a,b]} |f(x)| < \infty$. Show that, if $f, g : [a, b] \to \mathbb{R}$ are functions of bounded variation, then fg is also of bounded variation and

$$V_a^b(fg) \le V_a^b(f) \sup_{x\in[a,b]} |g(x)| + V_a^b(g) \sup_{x\in[a,b]} |f(x)|$$

[5] We recall that the length of a curve $y = f(x)$ $(a \le x \le b)$ is the supremum of the lengths of all inscribed polygonals, that is, the quantity

$$\sup \sum_{k=0}^{n-1} \sqrt{(x_{k+1} - x_k)^2 + (f(x_{k+1}) - f(x_k))^2},$$

where the supremum is taken over all partitions of $[a, b]$. A curve is said to be *rectifiable* if it has finite length.

Exercise 7.21 Let $(a_n)_n$ be a sequence of positive numbers and let

$$f(x) = \begin{cases} a_n & \text{if } x = \frac{1}{n},\ n \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Show that f is of bounded variation on [0, 1] if and only if $\sum_{n=1}^{\infty} a_n < \infty$.

Exercise 7.22 Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation such that

$$f(x) \geq c > 0 \quad \forall x \in [a, b].$$

Show that $\frac{1}{f}$ is of bounded variation and

$$V_a^b\left(\frac{1}{f}\right) \leq \frac{1}{c^2} V_a^b(f).$$

Exercise 7.23 Show that the function

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x^3} & 0 < x \leq 1, \\ 0 & x = 0 \end{cases}$$

is not of bounded variation on [0, 1].

Exercise 7.24 Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation such that

$$f(x) \geq 0 \quad \forall x \in [a, b].$$

(i) Show that if

$$f(x) \geq c > 0 \quad \forall x \in [a, b],$$

then $\sqrt{f}$ is also of bounded variation and

$$V_a^b(\sqrt{f}) \leq \frac{1}{2\sqrt{c}} V_a^b(f).$$

(ii) Give an example to show that $\sqrt{f}$ is not of bounded variation, in general.

7.3 Absolutely Continuous Functions

In order to address the problems we posed at the beginning of this chapter, we begin to study the largest class of functions for which formula (7.2) holds.

Definition 7.25 A function $f : [a, b] \to \mathbb{R}$ is said to be absolutely continuous if, given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon \tag{7.15}$$

for any finite family of disjoint subintervals

$$(a_k, b_k) \subset [a, b] \quad k = 1, \dots, n$$

of total length $\sum_{k=1}^{n}(b_k - a_k)$ less than δ.

Example 7.26 Let $f : [a, b] \to \mathbb{R}$ be a Lipschitz continuous function with Lipschitz constant K; then, choosing $\delta = \frac{\varepsilon}{K}$, we immediately obtain that f is absolutely continuous.

Remark 7.27 Any absolutely continuous function is uniformly continuous, as one can easily check by choosing a single subinterval $(a_1, b_1) \subset [a, b]$. However, a uniformly continuous function need not be absolutely continuous. For instance, the Cantor-Vitali function f constructed in Example 7.8 is continuous (hence, uniformly continuous) on [0, 1], but not absolutely continuous. Indeed, for any n consider the set

$$C_n = [0, 1] \setminus \bigcup_{h=1}^{n} \bigcup_{k=1}^{2^{h-1}} (a_k^h, b_k^h);$$

then C_n is the union of 2^n disjoint subintervals I_j of length $\frac{1}{3^n}$ (hence, the total length is $(\frac{2}{3})^n$). Since, by construction, the Cantor-Vitali function is constant on each subinterval (a_k^h, b_k^h), then the sum (7.15) associated to such a family I_j is equal to 1. So it is possible to find a finite disjoint family of subintervals of [0, 1] of arbitrarily small total length for which the sum (7.15) is equal to 1. The same example shows that a function of bounded variation needs not be absolutely continuous. On the other hand, any absolutely continuous function is necessarily of bounded variation owing to Proposition 7.28 below.

Proposition 7.28 *If $f : [a, b] \to \mathbb{R}$ is absolutely continuous, then f is of bounded variation.*

Proof Fixed $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

for any finite family of disjoint subintervals $(a_k, b_k) \subset [a, b]$ such that

$$\sum_{k=1}^{n}(b_k - a_k) < \delta.$$

Therefore, if $[\alpha, \beta]$ is any subinterval of length less than δ, then

$$V_\alpha^\beta(f) \leq \varepsilon.$$

Let $a = x_0 < x_1 < \cdots < x_N = b$ be a partition of $[a, b]$ into N subintervals $[x_k, x_{k+1}]$ of length less than δ. Then, by Proposition 7.14, $V_a^b(f) \leq N\varepsilon$. □

An immediate consequence of Definition 7.25 is the following proposition.

Proposition 7.29 *If $f : [a, b] \to \mathbb{R}$ is an absolutely continuous function, then αf is also absolutely continuous, where α is any constant. Moreover, if $f, g : [a, b] \to \mathbb{R}$ are absolutely continuous, $f + g$ is also absolutely continuous.*

Remark 7.30 By Proposition 7.29 and Remark 7.27 it follows that the set $AC([a, b])$ of all absolutely continuous functions on $[a, b]$ is a proper subspace of the linear space $BV([a, b])$ of all functions of bounded variation on $[a, b]$.

We now study the close connection between absolute continuity and the indefinite Lebesgue integral. We begin with the following result.

Lemma 7.31 *Let $g \in L^1(a, b)$ be such that $\int_I g(t)\,dt = 0$ for any subinterval $I \subset [a, b]$. Then $g = 0$ a.e. in $[a, b]$.*

Proof Using Lemma 1.60, any set V which is open in the relative topology of $[a, b]$ is a countable disjoint union of subintervals $I \subset [a, b]$. So $\int_V g(t)\,dt = 0$. Arguing by contradiction, suppose there is a Borel set $E \subset [a, b]$ such that $m(E) > 0$ and $g(x) > 0$ in E. By Theorem 1.71 there exists a compact set $K \subset E$ such that $m(K) > 0$. Then $V := [a, b] \setminus K$ is an open set in $[a, b]$. Hence,

$$0 = \int_a^b g(t)\,dt = \int_V g(t)\,dt + \int_K g(t)\,dt = \int_K g(t)\,dt > 0.$$

We have reached a contradiction thus completing the proof. □

As for the differentiation of an indefinite Lebesgue integral, in our next theorem we will compute the derivative (7.1), providing a positive answer to the first of the two questions posed at the beginning of the chapter.

Theorem 7.32 *Let $f \in L^1(a, b)$ and set*

$$F(x) = \int_a^x f(t)\,dt, \quad x \in [a, b].$$

Then F is absolutely continuous on $[a, b]$ and

$$F'(x) = f(x) \ \textit{for almost every } x \in [a, b]. \tag{7.16}$$

Proof Given a finite family of disjoint subintervals (a_k, b_k), we have

$$\sum_{k=1}^{n} |F(b_k) - F(a_k)| = \sum_{k=1}^{n} \left| \int_{a_k}^{b_k} f(t) dt \right| \le \sum_{k=1}^{n} \int_{a_k}^{b_k} |f(t)| dt = \int_{\bigcup_k (a_k, b_k)} |f(t)| dt.$$

By the absolute continuity of the Lebesgue integral, the last integral on the right-hand side tends to zero as the total length of the intervals (a_k, b_k) approaches zero. This proves that F is absolutely continuous on $[a, b]$. By Proposition 7.28 F is of bounded variation; consequently, thanks to Theorem 7.17, F is differentiable a.e. in $[a, b]$ and $F' \in L^1(a, b)$. To prove (7.16) assume, first, that $|f(x)| \le K$ for every $x \in [a, b]$ and some $K > 0$. Let

$$g_n(x) = n\left[F\left(x + \frac{1}{n}\right) - F(x)\right],$$

where, to define g_n for every $x \in [a, b]$, we have set

$$F(x) = F(b) \quad \text{for} \quad b < x \le b + 1.$$

Clearly,

$$\lim_{n \to \infty} g_n(x) = F'(x) \ \text{ a.e. in } [a, b].$$

Moreover,

$$|g_n(x)| = \left| n \int_x^{x+\frac{1}{n}} f(t)\, dt \right| \le K \quad \forall x \in [a, b].$$

Let $a \le c < d \le b$. By Lebesgue's Theorem we obtain

$$\begin{aligned} \int_c^d F'(x)\, dx &= \lim_{n \to \infty} \int_c^d g_n(x)\, dx = \lim_{n \to \infty} n\left[\int_{c+\frac{1}{n}}^{d+\frac{1}{n}} F(x)\, dx - \int_c^d F(x)\, dx\right] \\ &= \lim_{n \to \infty} n\left[\int_d^{d+\frac{1}{n}} F(x)\, dx - \int_c^{c+\frac{1}{n}} F(x)\, dx\right] = F(d) - F(c), \end{aligned}$$

where last equality follows from the mean value theorem. So we deduce that

$$\int_c^d F'(x)\, dx = F(d) - F(c) = \int_c^d f(t)\, dt.$$

Hence, appealing to Lemma 7.31, we conclude $F'(x) = f(x)$ a.e. in $[a, b]$.

We now remove the boundedness hypothesis on f. Without loss of generality, we may assume $f \geq 0$ (otherwise, we can argue separately for the positive part f^+ and the negative part f^-). Then F is an increasing function on $[a, b]$. Let us define f_n by:

$$f_n(x) = \begin{cases} f(x) & \text{if } 0 \leq f(x) \leq n, \\ n & \text{if } f(x) \geq n. \end{cases}$$

Since $f - f_n \geq 0$, the function

$$H_n(x) := \int_a^x (f(t) - f_n(t))dt$$

is increasing. So by Theorem 7.6 H_n is differentiable a.e. and $H_n'(x) \geq 0$. Since $0 \leq f_n \leq n$, using what we have shown in the first part of the proof we deduce that $\frac{d}{dx}\int_a^x f_n(t)dt = f_n(x)$ a.e.; therefore, for every $n \in \mathbb{N}$,

$$F'(x) = H_n'(x) + \frac{d}{dx}\int_a^x f_n(t)\,dt \geq f_n(x) \quad \text{a.e. in } [a, b].$$

This yields that $F'(x) \geq f(x)$ a.e., and so, after integration,

$$\int_a^b F'(x)\,dx \geq \int_a^b f(x)\,dx = F(b) - F(a).$$

On the other hand, since F is increasing on $[a, b]$, (7.5) yields

$$\int_a^b F'(x)dx \leq F(b) - F(a).$$

Consequently,

$$\int_a^b F'(x)\,dx = F(b) - F(a) = \int_a^b f(x)\,dx.$$

Hence,

$$\int_a^b (F'(x) - f(x))dx = 0.$$

Since $F'(x) \geq f(x)$ a.e., we conclude $F'(x) = f(x)$ a.e. in $[a, b]$. □

Lemma 7.33 *Let $f : [a, b] \to \mathbb{R}$ be an absolutely continuous function such that $f'(x) = 0$ for almost every $x \in [a, b]$. Then f is constant on $[a, b]$.*

Proof Fixed $c \in (a, b)$, we want to show that $f(c) = f(a)$. Let $E = \{x \in (a, c) \mid f'(x) = 0\}$. Then E is a Borel set and $m(E) = c - a$. Given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

for any finite family of disjoint subintervals $(a_k, b_k) \subset [a, b]$ such that

$$\sum_{k=1}^{n} (b_k - a_k) < \delta.$$

Let us fix $\eta > 0$. For every $x \in E$ and $\gamma > 0$, since $\lim_{y \to x} \frac{f(y)-f(x)}{y-x} = 0$, there exists $y_{x,\gamma} > x$ such that $[x, y_{x,\gamma}] \subset (a, c)$, $|y_{x,\gamma} - x| \leq \gamma$ and

$$|f(y_{x,\gamma}) - f(x)| \leq (y_{x,\gamma} - x)\eta. \tag{7.17}$$

The intervals $([x, y_{x,\gamma}])_{x \in E, \gamma > 0}$ are a fine cover of E; so, by Vitali's Covering Theorem, there exists a finite number of such disjoint intervals, which we label $I_1 = [x_1, y_1], \ldots, I_n = [x_n, y_n]$, where $x_k < x_{k+1}$, such that

$$m(E \setminus \cup_{k=1}^{n} I_k) < \delta.$$

Thus, we have

$$y_0 := a < x_1 < y_1 < x_2 < \cdots < y_n < c := x_{n+1}, \quad \sum_{k=0}^{n} (x_{k+1} - y_k) < \delta.$$

By the absolute continuity of f it follows that

$$\sum_{k=0}^{n} |f(x_{k+1}) - f(y_k)| < \varepsilon, \tag{7.18}$$

whereas, by (7.17),

$$\sum_{k=1}^{n} |f(y_k) - f(x_k)| \leq \eta \sum_{k=1}^{n} (y_k - x_k) \leq \eta(b - a). \tag{7.19}$$

Combining (7.18) and (7.19) we deduce that

$$|f(c) - f(a)| = \left| \sum_{k=0}^{n} (f(x_{k+1}) - f(y_k)) + \sum_{k=1}^{n} (f(y_k) - f(x_k)) \right| \leq \varepsilon + \eta(b - a).$$

Since ε and η are arbitrary, we conclude that $f(c) = f(a)$. □

Theorem 7.34 *If $f : [a, b] \to \mathbb{R}$ is absolutely continuous, then f is differentiable a.e. in $[a, b]$, $f' \in L^1(a, b)$, and*

$$f(x) = f(a) + \int_a^x f'(t)\,dt \quad \forall x \in [a, b]. \tag{7.20}$$

Proof Thanks to Proposition 7.28, f is of bounded variation. By Theorem 7.17, f is differentiable a.e. and $f' \in L^1(a, b)$. To prove (7.20), consider

$$g(x) = \int_a^x f'(t)\,dt.$$

Owing to Theorem 7.32, g is absolutely continuous and $g'(x) = f'(x)$ a.e. in $[a, b]$. Setting $\Phi = f - g$, Φ is absolutely continuous, since it is the difference between two absolutely continuous functions, and $\Phi'(x) = 0$ a.e. in $[a, b]$. From the previous lemma it follows that Φ is constant. So

$$\Phi(x) = \Phi(a) = f(a) - g(a) = f(a),$$

which yields in turn

$$f(x) = \Phi(x) + g(x) = f(a) + \int_a^x f'(t)\,dt \quad \forall x \in [a, b]$$

thus completing the proof. □

Remark 7.35 Using Theorems 7.32 and 7.34 we are now in a position to give a definite answer to the second question posed at the beginning of the chapter: formula

$$\int_a^x F'(t)\,dt = F(x) - F(a) \quad \forall x \in [a, b]$$

holds if and only if F is absolutely continuous on $[a, b]$.

Proposition 7.36 *Let $f : [a, b] \to \mathbb{R}$. The following properties are equivalent:*

(i) *f is absolutely continuous.*
(ii) *f is of bounded variation and*

$$\int_a^b |f'(t)|\,dt = V_a^b(f).$$

Proof We begin by proving the implication '(i) ⇒ (ii)'. For any partition $a = x_0 < x_1 < \cdots < x_n = b$ of $[a, b]$, Theorem 7.34 ensures that

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| = \sum_{k=0}^{n-1} \left| \int_{x_k}^{x_{k+1}} f'(t)\,dt \right|$$
$$\leq \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f'(t)|\,dt = \int_a^b |f'(t)|\,dt.$$

Therefore

$$V_a^b(f) \leq \int_a^b |f'(t)|\,dt.$$

Now, by Theorem 7.17, $\int_a^b |f'(t)|dt \leq V_a^b(f)$, and so $V_a^b(f) = \int_a^b |f'(t)|dt$.

Let us proceed to prove the implication '(ii) ⇒ (i)'. For every $x \in [a, b]$, by (7.14) we have

$$V_a^x(f) \geq \int_a^x |f'(t)|dt = \int_a^b |f'(t)|dt - \int_x^b |f'(t)|dt = V_a^b(f) - \int_x^b |f'(t)|dt$$
$$\geq V_a^b(f) - V_x^b(f) = V_a^x(f)$$

where the last equality follows from Proposition 7.14. Then we obtain

$$V_a^x(f) = \int_a^x |f'(t)|\,dt.$$

Since $f' \in L^1(a, b)$, Theorem 7.32 ensures that the function $x \mapsto V_a^x(f)$ is absolutely continuous. Given a family of disjoint subintervals $(a_k, b_k) \subset [a, b]$, we have

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| \leq \sum_{k=1}^{n} V_{a_k}^{b_k}(f) = \sum_{k=1}^{n} \big(V_a^{b_k}(f) - V_a^{a_k}(f)\big).$$

By the absolute continuity of the map $x \mapsto V_a^x(f)$, the last sum on the right-hand side tends to zero as the total length of the intervals (a_k, b_k) approaches zero. This proves that f is absolutely continuous. □

Applying the above proposition to the particular case of monotone functions, we obtain the following result.

Corollary 7.37 *Let $f : [a, b] \to \mathbb{R}$ be a monotone function. The following properties are equivalent:*

(i) *f is absolutely continuous.*
(ii) $\int_a^b |f'(t)|\,dt = |f(b) - f(a)|$.

Remark 7.38 Let $f,\ g : [a, b] \to \mathbb{R}$ be absolutely continuous functions. Then the following formula of integration by parts holds:

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx.$$

Indeed, by Tonelli's Theorem,

$$\iint_{[a,b]^2} |f'(x)g'(y)|\,dxdy = \int_a^b |f'(x)|\,dx \int_a^b |g'(y)|\,dy < \infty,$$

which yields $f'(x)g'(y) \in L^1([a, b]^2)$. Then consider the set

$$A = \{(x, y) \in [a, b]^2 \,|\, a \le x \le y \le b\}$$

and compute the integral

$$I = \iint_A f'(x)g'(y)\,dxdy$$

in two ways using Fubini's Theorem and formula (7.20). On the one hand,

$$\begin{aligned} I &= \int_a^b g'(y)\Big(\int_a^y f'(x)\,dx\Big)dy = \int_a^b g'(y)f(y)\,dy - f(a)\int_a^b g'(y)\,dy \\ &= \int_a^b g'(y)f(y)\,dy - f(a)\big(g(b) - g(a)\big). \end{aligned}$$

On the other hand, we have

$$\begin{aligned} I &= \int_a^b f'(x)\Big(\int_x^b g'(y)\,dy\Big)dx = g(b)\int_a^b f'(y)\,dy - \int_a^b f'(x)g(x)\,dx \\ &= g(b)\big(f(b) - f(a)\big) - \int_a^b f'(x)g(x)\,dx. \end{aligned}$$

Exercise 7.39 Show that if $f,\ g : [a, b] \to \mathbb{R}$ are absolutely continuous functions, then fg is also absolutely continuous.

Exercise 7.40 Let $(f_n)_n$ be a sequence of absolutely continuous functions on $[0, 1]$ converging pointwise to a function $f : [0, 1] \to \mathbb{R}$ and such that

$$\int_0^1 |f_n'(x)|\,dx \le M \quad \forall n \in \mathbb{N},$$

for some constant $M > 0$.

(i) Show that $\lim_{n\to\infty} \int_0^1 f_n(x)\,dx = \int_0^1 f(x)\,dx$.
(ii) Show that f is of bounded variation on $[0, 1]$.
(iii) Give an example to show that, in general, f fails to be absolutely continuous on $[0, 1]$.

Exercise 7.41 Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of absolutely continuous functions converging pointwise to a function $f : [a, b] \to \mathbb{R}$. Suppose that there exists $g \in L^1(a, b)$ such that

$$|f_n'| \leq g \quad \text{a.e. in } [a, b] \quad \forall n \in \mathbb{N}.$$

Show that f is absolutely continuous.

7.4 Miscellaneous Exercises

Exercise 7.42 Let $f_n : [a, b] \to \mathbb{R}$ be a sequence of absolutely continuous functions converging pointwise to a function $f : [a, b] \to \mathbb{R}$. Suppose that there exists $g \in L^1(a, b)$ such that

$$f_n' \rightharpoonup g \quad \text{in } L^1(a, b).$$

Show that f is absolutely continuous.

Exercise 7.43 Let $f \in BV([a, b])$, $f > 0$. Show that

$$\log f \in BV([a, b]) \iff \inf_{[a,b]} f > 0.$$

Exercise 7.44 Let $f \in AC([a, b])$ and let $x_0 \in (a, b)$. Show that

$$\lim_{\delta\to 0^+} V_{x_0}^{x_0+\delta}(f) = 0. \tag{7.21}$$

Give an example to show that (7.21) may fail if $f \in BV([a, b])$.

Exercise 7.45 Let $f \in AC([a, b])$ be such that $f(a) = 0$ and

$$f' \in L^p(a, b), \quad 1 < p < \infty. \tag{7.22}$$

1. Show that $|f(x)| \leq \|f'\|_p |x - a|^{\frac{p-1}{p}}$ for every $x \in [a, b]$.
2. Show that $\frac{f(x)}{x-a} \in L^1(a, b)$.
3. Give an example to show that, in general, $\frac{f(x)}{x-a} \notin L^1(a, b)$ if one drops assumption (7.22).
   Hint. Consider the function

$$f(x) = \begin{cases} \dfrac{1}{\log x} & \text{if } x \in \left(0, \dfrac{1}{2}\right] \\ 0 & \text{if } x = 0 \end{cases}.$$

Exercise 7.46 Let $f \in AC([a, b])$.

1. Show that $x^2 \sin^2 \frac{1}{x} \in AC([0, 1])$ and $x|\sin \frac{1}{x}| \notin AC([0, 1])$.
2. Deduce that if $\inf_{[a,b]} |f| = 0$, then $\sqrt{|f|} \notin AC([a, b])$, in general.
3. If $\inf_{[a,b]} |f| > 0$, show that $\sqrt{|f|} \in AC([a, b])$.

Chapter 8
Signed Measures

Given a measure space $(X, \mathscr{E}, \mu)$ and a function $\rho \in L^1(X, \mu)$, the so-called *Lebesgue indefinite integral*

$$\nu(E) = \int_E \rho \, d\mu \qquad (E \in \mathscr{E}) \tag{8.1}$$

defines a σ-additive set function, that is, if

$$E = \bigcup_n E_n \,,$$

with $E_n \in \mathscr{E}$ a family of disjoint sets, then

$$\nu(E) = \sum_n \nu(E_n).$$

Therefore, when $\rho \geq 0$, ν is a finite measure on $\mathscr{E}$ satisfying

$$E \in \mathscr{E} \;\; \& \;\; \mu(E) = 0 \quad \Rightarrow \quad \nu(E) = 0. \tag{8.2}$$

This raises the question whether all finite measures ν on $\mathscr{E}$ satisfying (8.2) can be represented as an indefinite integral of the form (8.1). Under suitable assumptions, a positive answer to this question is given by the Radon-Nikodym Theorem that we will prove using the so-called *Lebesgue decomposition*. Such a technique allows to represent a given measure ν as the sum of other two measures, one of which is absolutely continuous with respect to μ while the other one is singular, in the sense of Definition 8.1.

The properties of the indefinite integral, in turn, motivate the introduction of interesting generalizations. We will define and study *signed measures*, which subsume the familiar notion of positive measures considered in the first part of this monograph leading to further decomposition formulas.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_8

In the last section of this chapter, we will apply the above results to the characterization of the dual of $L^p(X, \mu)$.

8.1 Comparison Between Measures

Let $(X, \mathscr{E})$ be a measurable space. We recall that a measure μ on $\mathscr{E}$ is said to be *concentrated* on a set $A \in \mathscr{E}$ if $\mu(A^c) = 0$ or, equivalently, if

$$\mu(E) = \mu(A \cap E) \qquad \forall E \in \mathscr{E}.$$

Definition 8.1 Let μ and ν be two measures on $\mathscr{E}$.

- μ and ν are said to be *singular* if they are concentrated on disjoint sets. In this case we write $\mu \perp \nu$.
- ν is said to be *absolutely continuous* with respect to μ, and we write $\nu \ll \mu$, if

$$E \in \mathscr{E}, \quad \mu(E) = 0 \quad \Longrightarrow \quad \nu(E) = 0.$$

- μ and ν are said to be *equivalent*, and we write $\mu \sim \nu$, if $\nu \ll \mu$ and $\mu \ll \nu$.

Example 8.2 Let $\rho \in L^1(X, \mu)$ be such that $\rho \geq 0$ and set

$$\nu(E) = \int_E \rho(x)\, d\mu(x) \qquad \forall E \in \mathscr{E}.$$

It is easy to verify that ν is an additive function on $\mathscr{E}$. Moreover, if $(E_n)_n \subset \mathscr{E}$ is an increasing sequence converging to $E \in \mathscr{E}$, then by Monotone Convergence Theorem we have

$$\nu(E_n) = \int_X \rho(x)\chi_{E_n}(x)\, d\mu(x) \quad \uparrow \quad \int_X \rho(x)\chi_E(x)\, d\mu(x) = \nu(E).$$

So ν is a (finite) measure on $\mathscr{E}$ thanks to Proposition 1.17. Since the integral vanishes on sets of measure zero, it follows that $\nu \ll \mu$.

Exercise 8.3 Let m denote the Lebesgue measure on $\mathbb{R}$ and let $\rho : \mathbb{R} \to [0, \infty]$ be a Borel function, summable on all bounded subsets of $\mathbb{R}$. Define

$$\nu(E) = \int_E \rho(x)\, dx \qquad \forall E \in \mathscr{B}(\mathbb{R}).$$

Show that ν is a measure on $\mathscr{B}(\mathbb{R})$ and $\nu \ll m$.

Example 8.4 Let m denote the Lebesgue measure on $\mathbb{R}$ and let δ_{x_0} be the Dirac measure at $x_0 \in \mathbb{R}$. Then m is concentrated on $A := \mathbb{R} \setminus \{x_0\}$, whereas δ_{x_0} is concentrated on $B := \{x_0\}$. Therefore m and δ_{x_0} are singular.

Exercise 8.5 Show that the measures

$$\mu(E) = \int_E e^{-x^2} dx \qquad \forall E \in \mathscr{B}(\mathbb{R})$$

and

$$\nu(E) = \int_E e^{x^2} dx \qquad \forall E \in \mathscr{B}(\mathbb{R})$$

are equivalent.

Exercise 8.6 Let μ and ν be two measures on $\mathscr{E}$.

1. Show that if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad E \in \mathscr{E} \ \& \ \mu(E) < \delta \implies \nu(E) < \varepsilon, \tag{8.3}$$

   then $\nu \ll \mu$.
2. Show that if ν is finite, then $\nu \ll \mu$ implies property (8.3).
   Hint. Suppose, by contradiction, that there exist $\varepsilon > 0$ and $(A_n)_n \subset \mathscr{E}$ such that

$$\mu(A_n) < \frac{1}{2^n} \quad \text{and} \quad \nu(A_n) \geq \varepsilon \qquad \forall n \in \mathbb{N}.$$

   Then

$$B_n := \bigcup_{i \geq n} A_i \ \downarrow \ B = \limsup_{n \to \infty} A_n.$$

   So, using Proposition 1.18, $\mu(B_n) \downarrow \mu(B) = 0$, whereas $\nu(B_n) \downarrow \nu(B) \geq \varepsilon$.
3. Give an example to show that property (8.3) is false, in general, when $\nu \ll \mu$ but ν is σ-finite.
   Hint. On $\mathscr{B}((0, 1])$ consider the σ-finite measure

$$\nu(E) = \int_E \frac{dx}{x}.$$

   Then $\nu \ll m$ (denoting m the Lebesgue measure on $(0, 1]$), but (8.3) is false. Indeed, for every $\delta \in (0, 1]$, we have $\nu((0, \delta]) = \int_0^\delta \frac{dx}{x} = \infty$.

8.2 Lebesgue Decomposition

In this section we will prove two relevant results in measure theory, known as the *Lebesgue decomposition* and the *Radon-Nikodym derivative*. We will begin by analyzing the case of finite measures. In the following, $(X, \mathscr{E})$ denotes a generic measurable space.

8.2.1 The Case of Finite Measures

Theorem 8.7 *Let μ and ν be finite measures on $\mathscr{E}$. Then the following statements hold.*

(a) There exist two finite measures on $\mathscr{E}$, ν_a and ν_s, such that

$$\nu = \nu_a + \nu_s, \qquad \nu_a \ll \mu, \qquad \nu_s \perp \mu. \tag{8.4}$$

Moreover, such a decomposition is unique.

(b) There exists a unique function $\rho \in L^1(X, \mu)$ such that $\rho \geq 0$ and

$$\nu_a(E) = \int_E \rho(x)\, d\mu(x) \qquad \forall E \in \mathscr{E}. \tag{8.5}$$

Equation (8.4) *si called the* Lebesgue decomposition *of ν with respect to μ. The function ρ in* (8.5) *is called the* density *or the* Radon-Nikodym derivative *of ν_a with respect to μ, and is denoted by the symbol*

$$\rho = \frac{d\nu_a}{d\mu}.$$

Proof We split the proof into 6 steps.

1. *Construction of a bounded linear functional.*

Set

$$\lambda = \mu + \nu$$

and observe that $\mu \ll \lambda$, $\nu \ll \lambda$ and

$$L^2(X, \lambda) \subset \underbrace{L^2(X, \nu) \subset L^1(X, \nu)}_{(\nu(X)<\infty)}.$$

Therefore the following linear functional is well defined

$$F(\varphi) := \int_X \varphi(x)\, d\nu(x) \qquad \forall \varphi \in L^2(X, \lambda).$$

By Hölder's inequality we have

$$|F(\varphi)| \leq \sqrt{\nu(X)} \left(\int_X |\varphi(x)|^2 \, d\nu(x) \right)^{1/2} = \sqrt{\nu(X)} \, \|\varphi\|_2.$$

So F is bounded. Thanks to the Riesz Theorem, there exists a unique element f of $L^2(X, \lambda)$ such that

$$F(\varphi) = \int_X \varphi(x) \, d\nu(x) = \int_X f(x)\varphi(x) \, d\lambda(x) \quad \forall \varphi \in L^2(X, \lambda). \tag{8.6}$$

2. *Two estimates for* f.

Observe that, since λ is finite, χ_E belongs to $L^2(X, \lambda)$ for any $E \in \mathscr{E}$. Taking $\varphi = \chi_E$ in (8.6) we obtain

$$\nu(E) = \int_E f(x) \, d\lambda(x) \geq 0 \qquad \forall E \in \mathscr{E}.$$

Therefore $f \geq 0$ λ-a.e. and we may assume

$$f(x) \geq 0 \quad \forall x \in X.$$

Moreover, since $\int_X f\varphi \, d\lambda = \int_X f\varphi \, d\mu + \int_X f\varphi \, d\nu$, (8.6) can be rewritten in the form

$$\int_X \varphi(x)(1 - f(x)) \, d\nu(x) = \int_X f(x)\varphi(x) \, d\mu(x) \quad \forall \varphi \in L^2(X, \lambda). \tag{8.7}$$

Then, choosing $\varphi = \chi_E$ as before, we have

$$\int_E (1 - f(x)) \, d\nu(x) = \int_E f(x) \, d\mu(x) \geq 0 \qquad \forall E \in \mathscr{E},$$

by which it follows that $f \leq 1$ ν-a.e.

3. *Construction of* ν_a *and* ν_s.

Define the two Borel sets

$$A := \{x \in X \mid 0 \leq f(x) < 1\} \qquad B := X \setminus A = \{x \in X \mid f(x) \geq 1\},$$

and define[1]

$$\nu_a := \nu \llcorner A \qquad \nu_s := \nu \llcorner B.$$

[1] See Definition 1.26.

Then ν_a and ν_s are finite measures satisfying $\nu = \nu_a + \nu_s$. Taking $\varphi = \chi_B$ in (8.7), we deduce $\mu(B) = 0$. So μ is concentrated on A. Since ν_s is concentrated on B, it follows that $\mu \perp \nu_s$.

4. *Density of* ν_a.

Given $n \in \mathbb{N}$ and $E \in \mathscr{E}$, let us take

$$\varphi(x) = \big(1 + f(x) + \cdots + f^n(x)\big)\chi_{E\cap A}(x)$$

in (8.7), obtaining

$$\int_{E\cap A} \big(1 - f^{n+1}(x)\big)\, d\nu(x) = \int_{E\cap A} \big[f(x) + f^2(x) + \cdots + f^{n+1}(x)\big]\, d\mu(x).$$

Set

$$\rho(x) := \begin{cases} \lim_{n\to\infty} \big[f(x) + f^2(x) + \cdots + f^{n+1}(x)\big] = \dfrac{f(x)}{1 - f(x)} & \text{if } x \in A, \\ 0 & \text{if } x \in B. \end{cases}$$

The Monotone Convergence Theorem implies that

$$\nu_a(E) = \nu(E \cap A) = \int_{E\cap A} \rho(x)\, d\mu(x) = \int_E \rho(x)\, d\mu(x).$$

This proves (8.5). Moreover, taking $E = X$ in the above identity, we conclude that ρ is μ-summable. The fact that $\nu_a \ll \mu$ follows from Example 8.2.

5. *Uniqueness of the density.*

Let $\rho_1, \rho_2 \geq 0$ be two μ-summable functions satisfying (8.5). Then $\rho = \rho_1 - \rho_2$ is a μ-summable function such that

$$\int_E \rho(x)\, d\mu(x) = 0 \qquad \forall E \in \mathscr{E}.$$

Therefore $\rho = 0$ μ-a.e., so ρ_1 and ρ_2 are two identical elements of the space $L^1(X, \mu)$.

6. *Uniqueness of the Lebesgue decomposition.*

Let ν_a^i and ν_s^i, $i = 1, 2$, be finite measures satisfying

$$\nu = \nu_a^i + \nu_s^i \quad \text{with} \quad \nu_a^i \ll \mu \quad \text{and} \quad \nu_s^i \perp \mu.$$

Let A be a support of μ such that $\nu_s^1(A) = 0 = \nu_s^2(A)$. Then, for any $E \in \mathscr{E}$, we have

$$\begin{aligned}\nu_a^1(E) &= \nu_a^1(E \cap A) + \underbrace{\nu_a^1(E \cap A^c)}_{=0\ (\nu_a^1 \ll \mu)} \\ &= \nu_a^2(E \cap A) + \underbrace{\nu_s^2(E \cap A)}_{=0\ (\nu_s^2 \perp \mu)} - \underbrace{\nu_s^1(E \cap A)}_{=0\ (\nu_s^1 \perp \mu)} = \nu_a^2(E).\end{aligned}$$

□

The next result follows immediately from Theorem 8.7.

Theorem 8.8 (Radon-Nikodym) *Let μ and ν be finite measures on $\mathscr{E}$ such that $\nu \ll \mu$. Then there exists a unique function $\rho \geq 0$ in $L^1(X, \mu)$ such that*

$$\nu(E) = \int_E \rho(x)\, d\mu(x) \qquad \forall E \in \mathscr{E}.$$

8.2.2 The General Case

We now extend Lebesgue's decomposition to more general measures.

Theorem 8.9 *Let μ and ν be measures on $\mathscr{E}$. If μ is σ-finite and ν is finite, then the conclusions of Theorem 8.7 hold.*

Proof Let $(X_n)_n \subset \mathscr{E}$ be a sequence of disjoint sets such that $\mu(X_n) < \infty$ for every $n \in \mathbb{N}$ and $X = \cup_{n \geq 1} X_n$. Apply Theorem 8.7 to the finite measures

$$\mu_n := \mu \llcorner X_n \qquad \nu_n := \nu \llcorner X_n$$

and consider, for any $n \in \mathbb{N}$, the Lebesgue decomposition of ν_n with respect to μ_n, namely

$$\nu_n = (\nu_n)_a + (\nu_n)_s \quad \text{with} \quad (\nu_n)_a \ll \mu_n \quad \text{and} \quad (\nu_n)_s \perp \mu.$$

Thanks to (8.5) and Exercise 2.72,

$$(\nu_n)_a(E) = \int_E \rho_n(x)\, d\mu_n(x) = \int_{E \cap X_n} \rho_n(x)\, d\mu(x) \qquad \forall E \in \mathscr{E}$$

for some μ_n-summable functions $\rho_n \geq 0$. Define

$$\nu_a := \sum_{n=1}^{\infty} (\nu_n)_a \qquad \nu_s := \sum_{n=1}^{\infty} (\nu_n)_s$$

and

$$\rho(x) := \sum_{n=1}^{\infty} \rho_n(x) \chi_{X_n}(x) \qquad \forall x \in X.$$

Then ν_a and ν_s are finite measures such that

$$\nu = \sum_{n=1}^{\infty} \nu_n = \sum_{n=1}^{\infty} (\nu_n)_a + \sum_{n=1}^{\infty} (\nu_n)_s = \nu_a + \nu_s .$$

Moreover, for any $E \in \mathscr{E}$, Proposition 2.48 implies

$$\begin{aligned} \nu_a(E) &= \sum_{n=1}^{\infty} \int_{E \cap X_n} \rho_n(x)\, d\mu(x) \\ &= \int_E \sum_{n=1}^{\infty} \rho_n(x) \chi_{X_n}(x)\, d\mu(x) = \int_E \rho(x)\, d\mu(x). \end{aligned}$$

Taking $E = X$ in the above identity we deduce che ρ is μ-summable. Therefore $\nu_a \ll \mu$. To complete the proof, let $A_n, B_n \subset X_n$ be disjoint supports of μ_n and $(\nu_n)_s$, respectively. Then $A := \cup_n A_n$ and $B := \cup_n B_n$ are disjoint supports of μ and ν_s. It follows that $\nu_s \perp \mu$. The uniqueness of ρ and decomposition (8.4) can be recovered reasoning as in the proof of Theorem 8.7. □

Example 8.10 If measure μ is not σ-finite, then the conclusion of Theorem 8.7 is false, in general, even when ν is finite. For instance, on the Borel σ-algebra $\mathscr{B}([0, 1])$ consider the counting measure $\mu^{\#}$. Let m be the Lebesgue measure on $[0, 1]$. Then $m \ll \mu^{\#}$, but m does not admit any representation of the form

$$m(E) = \int_E f\, d\mu^{\#} \qquad \forall E \in \mathscr{B}([0, 1])$$

with $f : [0, 1] \to [0, \infty]$ $\mu^{\#}$-summable. Indeed, should such f exist, then we would obtain $m(\{x\}) = 0 = f(x)$ for every $x \in [0, 1]$, and so $f(x) = 0$ for every $x \in [0, 1]$. Taking $E = [0, 1]$, it would follow $m([0, 1]) = 0$.

Exercise 8.11 Let X be an uncountable set and let $\mathscr{E}$ be the σ-algebra which consists of all countable subsets of X and their complements. Show that if $\mu^{\#}$ is the counting measure on X and

$$\lambda(E) = \begin{cases} 0 & \text{if } E \text{ is countable,} \\ 1 & \text{if } E^c \text{ is countable,} \end{cases}$$

then $\lambda \ll \mu^{\#}$ but there is no $\mu^{\#}$-summable function f such that

$$\lambda(E) = \int_E f\, d\mu^{\#} \qquad \forall E \in \mathscr{E}.$$

Exercise 8.12 Adapting the proof of Theorem 8.9, show that if μ and ν are both σ-finite, then the conclusions of Theorem 8.7 are still true, with the difference that ρ is not necessarily μ-summable but only *locally μ-summable*, that is, there exists a sequence $(X_n)_n \subset \mathscr{E}$ such that $X_n \uparrow X$ and

$$\mu(X_n) < \infty, \quad \int_{X_n} \rho \, d\mu < \infty \qquad \forall n \in \mathbb{N}.$$

8.3 Signed Measures

Let $(X, \mathscr{E})$ be a measurable space.

Definition 8.13 A *signed measure* μ on $\mathscr{E}$ is a map $\mu : \mathscr{E} \to \mathbb{R}$ such that $\mu(\varnothing) = 0$ and, for any sequence $(E_n)_n \subset \mathscr{E}$ of disjoint sets,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n). \tag{8.8}$$

Example 8.14 Let μ_1 and μ_2 be finite measures on $\mathscr{E}$. Then the difference $\mu := \mu_1 - \mu_2$ is a signed measure on $\mathscr{E}$.

Remark 8.15 Let us observe that the series on the right-hand side of (8.8) must converge independently of the order of its terms (since the left-hand side is independent of such an order), so it must converge absolutely.

Remark 8.16 Definition 8.13 can be generalized including extended functions: more precisely, a function $\mu : \mathscr{E} \to \overline{\mathbb{R}}$ is called a signed measure if $\mu(\varnothing) = 0$ and μ satisfies (8.8). In such a case, however, μ cannot assume both the values ∞ and $-\infty$.

Exercise 8.17 Let $\mu : \mathscr{E} \to \mathbb{R}$ be an additive function such that $\mu(\varnothing) = 0$.

- Given a sequence $(E_n)_n \subset \mathscr{E}$, show that the following properties are equivalent:

  (a) $E_n \uparrow E \implies \mu(E_n) \to \mu(E)$.
  (b) $E_n \downarrow E \implies \mu(E_n) \to \mu(E)$.
  (c) $E_n \downarrow \varnothing \implies \mu(E_n) \to 0$.

- Show that μ is a signed measure on $\mathscr{E}$ if and only if one of the above properties holds.

Hint. Adapt the proof of Propositions 1.17 and 1.18.

8.3.1 Total Variation

Definition 8.18 Given $E \in \mathscr{E}$, a sequence $(E_n)_n \subset \mathscr{E}$ of disjoint sets such that $\bigcup_{n=1}^{\infty} E_n = E$ is called a *partition* of E.

Definition 8.19 Let μ be a signed measure on $\mathscr{E}$. The *total variation* of μ is the map $|\mu| : \mathscr{E} \to [0, \infty]$ defined by

$$|\mu|(E) = \sup \left\{ \sum_{n=1}^{\infty} |\mu(E_n)| \ : \ (E_n)_n \text{ partition of } E \right\} \qquad \forall E \in \mathscr{E}.$$

Proposition 8.20 *Let μ be a signed measure on $\mathscr{E}$. Then $|\mu|$ is a finite measure on $\mathscr{E}$.*

Proof We split the proof into 3 steps.

1. *Additivity.*

We claim that if $A, B \in \mathscr{E}$ are disjoint sets, then

$$|\mu|(A \cup B) = |\mu|(A) + |\mu|(B). \tag{8.9}$$

Indeed, consider $(E_n)_n$ a partition of $E := A \cup B$ and set

$$A_n = A \cap E_n, \quad B_n = B \cap E_n \qquad \forall n \in \mathbb{N}.$$

Then $(A_n)_n$ is a partition of A and $(B_n)_n$ is a partition of B. Moreover, since $E_n = A_n \cup B_n$ with disjoint union, we have $\mu(E_n) = \mu(A_n) + \mu(B_n)$ for every $n \in \mathbb{N}$. So

$$\sum_{n=1}^{\infty} |\mu(E_n)| \le \sum_{n=1}^{\infty} |\mu(A_n)| + \sum_{n=1}^{\infty} |\mu(B_n)| \le |\mu|(A) + |\mu|(B),$$

which in turn implies that $|\mu|(A \cup B) \le |\mu|(A) + |\mu|(B)$.

In order to prove the opposite inequality, let L and M be real numbers satisfying $L < |\mu|(A)$ and $M < |\mu|(B)$. Then there exist partitions $(A_n)_n$ of A and $(B_n)_n$ of B such that

$$\sum_{n=1}^{\infty} |\mu(A_n)| \ge L, \quad \sum_{n=1}^{\infty} |\mu(B_n)| \ge M.$$

Moreover, $(A_n)_n \cup (B_n)_n$ is a partition of $A \cup B$. Therefore

$$|\mu|(A \cup B) \ge \sum_{n=1}^{\infty} (|\mu(A_n)| + |\mu(B_n)|) \ge L + M.$$

Since L, M are arbitrary, we get $|\mu|(A \cup B) \ge |\mu|(A) + |\mu|(B)$.

2. *σ-additivity.*

Since $|\mu|$ is additive, it is sufficient to show that $|\mu|$ is σ-subadditive (see Remark 1.14). Consider a disjoint sequence $(E_n)_n \subset \mathscr{E}$ and set $E = \cup_{n=1}^{\infty} E_n$. Let $(F_i)_i$ be a partition of E. Then, for any given n, $(F_i \cap E_n)_i$ is a partition of E_n and, for any given i, $(F_i \cap E_n)_n$ is a partition of F_i. Therefore $\mu(F_i) = \sum_{n=1}^{\infty} \mu(F_i \cap E_n)$, which yields

$$\sum_{i=1}^{\infty} |\mu(F_i)| \leq \sum_{i=1}^{\infty}\sum_{n=1}^{\infty} |\mu(F_i \cap E_n)| = \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} |\mu(F_i \cap E_n)| \leq \sum_{n=1}^{\infty} |\mu|(E_n)\,.$$

So, by the arbitrariness of the partition $(F_i)_i$,

$$|\mu|(E) \leq \sum_{n=1}^{\infty} |\mu|(E_n).$$

3. $|\mu|(X) < \infty$.

Assuming $|\mu|(X) = \infty$, we will construct disjoint sets $A,\ B \in \mathscr{E}$ such that $X = A \cup B$ and

$$|\mu(A)| > 1 \quad \& \quad |\mu|(B) = \infty. \tag{8.10}$$

Later, we will show that (8.10) yields a contradiction.

Suppose $|\mu|(X) = \infty$. Then there exists a partition $(X_n)_n$ of X such that

$$\sum_{n=1}^{\infty} |\mu(X_n)| > 2(1 + |\mu(X)|).$$

Therefore one of the two sums

$$\sum_{n\geq 1,\, \mu(X_n)>0} |\mu(X_n)|, \qquad \sum_{n\geq 1,\, \mu(X_n)<0} |\mu(X_n)|$$

is greater than $1 + |\mu(X)|$. To fix ideas, assume we are in the first case: for some subsequence $(X_{n_k})_k$, we have

$$\sum_{k=1}^{\infty} \mu(X_{n_k}) > 1 + |\mu(X)|.$$

Set $A = \bigcup_{k=1}^{\infty} X_{n_k}$ and $B = A^c$. Then $|\mu(A)| > 1$ and

$$|\mu(B)| = |\mu(X) - \mu(A)| \geq |\mu(A)| - |\mu(X)| > 1.$$

Since

$$|\mu|(X) = |\mu|(A) + |\mu|(B) = \infty,$$

either $|\mu|(B) = \infty$ or $|\mu|(A) = \infty$. In both cases we obtain (8.10) exchanging, if necessary, the roles of A and B.

Finally, we claim (8.10) leads to a contradiction. Indeed, (replacing X by B and doing the same at each step) we construct a sequence $(A_n)_n$ of disjoint measurable sets such that $|\mu(A_n)| > 1$. Then, for some subsequence $(A_{n_k})_k$ of $(A_n)_n$, either $\mu(A_{n_k}) > 1$ or $\mu(A_{n_k}) < -1$ for every $k \in \mathbb{N}$. Therefore $\sum_k \mu(A_{n_k}) = \infty$ in the first case and $\sum_k \mu(A_{n_k}) = -\infty$ in the second case, in contradiction with $\mu(\cup_k A_{n_k}) \in \mathbb{R}$. □

Let us observe that if μ is a signed measure on $\mathscr{E}$, then

$$|\mu(E)| \leq |\mu|(E) \qquad \forall E \in \mathscr{E}. \tag{8.11}$$

Therefore, thanks to Proposition 8.20,

$$\mu^+ := \frac{1}{2}\,(|\mu| + \mu) \quad \text{and} \quad \mu^- = \frac{1}{2}\,(|\mu| - \mu) \tag{8.12}$$

are finite measures on $\mathscr{E}$, called the *positive part* and the *negative part* of μ, respectively. Moreover, the identity

$$\mu = \mu^+ - \mu^- \tag{8.13}$$

is called the *Jordan decomposition* of μ.

8.3.2 Radon-Nikodym Theorem

Let $(X, \mathscr{E}, \mu)$ be a measure space.

Definition 8.21 We say that a signed measure ν on $\mathscr{E}$ is *absolutely continuous* with respect to μ, and we write $\nu \ll \mu$, if

$$E \in \mathscr{E} \;\;\&\;\; \mu(E) = 0 \quad \Longrightarrow \quad |\nu|(E) = 0.$$

Remark 8.22 Let us note that, since $|\nu| = \nu^+ + \nu^-$,

$$\nu \ll \mu \quad \Longleftrightarrow \quad \nu^+ \ll \mu \;\;\&\;\; \nu^- \ll \mu\,.$$

Exercise 8.23 Given a signed measure ν on $\mathscr{E}$, show that $\nu \ll \mu$ if and only if

$$E \in \mathscr{E} \;\;\&\;\; \mu(E) = 0 \quad \Longrightarrow \quad \nu(E) = 0.$$

The following generalization of Radon-Nikodym Theorem holds.

Theorem 8.24 (Radon-Nikodym) *Let μ be a σ-finite measure on $\mathscr{E}$ and let ν be a signed measure on $\mathscr{E}$ such that $\nu \ll \mu$. Then there exists a unique function $\rho \in L^1(X, \mu)$ such that*

$$\nu(E) = \int_E \rho(x)\, d\mu(x) \qquad \forall E \in \mathscr{E}. \tag{8.14}$$

Proof By hypothesis ν^+ and ν^- are finite measures. They are also absolutely continuous with respect to μ thanks to Remark 8.22. Therefore Theorem 8.9 ensures that ν^+ and ν^- have derivatives

$$\rho_+ = \frac{d\nu^+}{d\mu} \quad \& \quad \rho_- = \frac{d\nu^-}{d\mu}.$$

Set $\rho := \rho_+ - \rho_-$. Then $\rho \in L^1(X, \mu)$ and (8.14) holds. The uniqueness property of ρ follows arguing as in the proof of Theorem 8.7. □

8.3.3 Hahn Decomposition

Our next result describes the structure of a signed measure. More precisely, it states that X is the union of two disjoint sets which are the supports of its positive and negative parts.

Theorem 8.25 *Let μ be a signed measure on $\mathscr{E}$ and let μ^+ and μ^- be its positive and negative parts, respectively. Then there exist disjoint sets $A, B \in \mathscr{E}$ such that $X = A \cup B$ and*

$$\mu^+(E) = \mu(A \cap E), \quad \mu^-(E) = -\mu(B \cap E) \qquad \forall E \in \mathscr{E}. \tag{8.15}$$

The pair (A, B) is called the Hahn decomposition of X *with respect to μ.*

Proof Observe, first, that $\mu \ll |\mu|$. So, applying Theorem 8.24, there exists a function $\rho \in L^1(X, |\mu|)$ such that

$$\mu(E) = \int_E \rho\, d|\mu| \qquad \forall E \in \mathscr{E}. \tag{8.16}$$

We now pass to show that $|\rho(x)| = 1$ $|\mu|$–a.e.

1. $|\rho| \leq 1$ $|\mu|$*–a.e.*

Set

$$E_1 = \{x \in X \mid \rho(x) > 1\}, \qquad E_2 = \{x \in X \mid \rho(x) < -1\}.$$

It suffices to show that $|\mu|(E_1) = |\mu|(E_2) = 0$. Suppose $|\mu|(E_1) > 0$. Then

$$\mu(E_1) = |\mu(E_1)| = \int_{E_1} \rho\, d|\mu| > |\mu|(E_1),$$

in contradiction with (8.11). Therefore $|\mu|(E_1) = 0$. Similarly we can prove that $|\mu|(E_2) = 0$.

2. $|\rho| = 1$ $|\mu|$ *–a.e.*

Set, for any $r \in (0, 1)$,

$$G_r = \{x \in X \mid 0 \leq \rho(x) < r\}, \qquad H_r = \{x \in X \mid -r < \rho(x) \leq 0\}.$$

As before, we will show that $|\mu|(G_r) = |\mu|(H_r) = 0$. Let $(G_{r,n})_n$ be a partition of G_r. Then

$$\mu(G_{r,n}) = |\mu(G_{r,n})| = \int_{G_{r,n}} \rho \, d|\mu| \leq r|\mu|(G_{r,n}).$$

So

$$\sum_{n=1}^{\infty} |\mu(G_{r,n})| \leq r|\mu|(G_r).$$

Since $(G_{r,n})_n$ is an arbitrary partition, we conclude that

$$|\mu|(G_r) \leq r|\mu|(G_r)\,.$$

Since $r \in (0, 1)$, necessarily $|\mu|(G_r) = 0$. Similarly, $|\mu|(H_r) = 0$.

3. *Conclusion.*

Thanks to the previous step we may assume $|\rho(x)| = 1$ for every $x \in X$. Let

$$A = \{x \in X \mid \rho(x) = 1\}, \qquad B = \{x \in X \mid \rho(x) = -1\}.$$

Then for any $E \in \mathscr{E}$ we have

$$\mu^+(E) = \frac{1}{2}\big(|\mu|(E) + \mu(E)\big) = \underbrace{\frac{1}{2}\int_E (1+\rho)\, d|\mu| = \int_{E\cap A} \rho\, d|\mu|}_{1+\rho(x)=0\ \forall x\in E\cap B} = \mu(E \cap A)$$

and

$$\mu^-(E) = \frac{1}{2}\big(|\mu|(E) - \mu(E)\big) = \underbrace{\frac{1}{2}\int_E (1-\rho)\, d|\mu| = \int_{E\cap B} \rho\, d|\mu|}_{1-\rho(x)=0\ \forall x\in E\cap A} = -\mu(E \cap B).$$

The proof is thus complete. □

Remark 8.26 A signed measure may admit more than one Hahn decomposition.

Exercise 8.27 Show that the positive part and the negative part of a signed measure μ are singular measures.

Exercise 8.28 Show that if μ is a signed measure on $\mathscr{E}$ and $\lambda_1,\ \lambda_2$ are two measures on $\mathscr{E}$ such that

$$\mu = \lambda_1 - \lambda_2,$$

then

$$\mu^+ \le \lambda_1, \quad \mu^- \le \lambda_2.$$

8.4 Dual of $L^p(X, \mu)$

Let $(X, \mathscr{E}, \mu)$ be a measure space. In this section we will characterize the dual of $L^p(X, \mu)$. Let $1 \le p \le \infty$ and let p' be the conjugate exponent of p, namely $1/p + 1/p' = 1$ with the usual convention $\frac{1}{\infty} = 0$. For any $g \in L^{p'}(X, \mu)$ let us define $F_g : L^p(X, \mu) \to \mathbb{R}$ by setting

$$F_g(f) = \int_X fg\, d\mu \qquad \forall f \in L^p(X, \mu). \tag{8.17}$$

Observe that, by Hölder's inequality,

$$|F_g(f)| \le \|f\|_p \|g\|_{p'} \qquad \forall f \in L^p(X, \mu).$$

Therefore $F_g \in \big(L^p(X, \mu)\big)^*$ and

$$\|F_g\|_* \le \|g\|_{p'}. \tag{8.18}$$

Then the map $g \mapsto F_g$ is a linear contraction $L^{p'}(X, \mu) \to (L^p(X, \mu))^*$. It is natural to ask whether all the bounded linear functionals on $L^p(X, \mu)$ have this form and if such a representation is unique. We will restrict our analysis to the case of σ-finite measures.

Theorem 8.29 *Let μ be a σ-finite measure on $\mathscr{E}$ and let $1 \le p < \infty$. Then the map $g \mapsto F_g$ defined by (8.17) is an isometric isomorphism[2] between $L^{p'}(X, \mu)$ and $\big(L^p(X, \mu)\big)^*$.*

Proof Let $F \in \big(L^p(X, \mu)\big)^*$. We will construct a function $g \in L^{p'}(X, \mu)$ such that $F = F_g$ and $\|g\|_{p'} \le \|F_g\|_*$. Consider, first, the case of $\mu(X) < \infty$. We split the proof into three steps.

1. $\exists g \in L^1(X, \mu)$ *such that* $F(f) = \int_X fg\, d\mu$ *for every* $f \in L^\infty(X, \mu)$.

Observe that, since μ is finite, $\chi_E \in L^p(X, \mu)$ for every $E \in \mathscr{E}$. Define

$$\nu(E) = F(\chi_E) \qquad \forall E \in \mathscr{E}.$$

[2] See footnote 5 at p. 147.

Since F is linear and $\chi_{A\cup B} = \chi_A + \chi_B$ if A and B are disjoint, we deduce that ν is additive and $\nu(\varnothing) = F(0) = 0$. Moreover, for any sequence $(E_n)_n \subset \mathscr{E}$ such that $E_n \uparrow E$, using Proposition 3.36, we have $\chi_{E_n} \xrightarrow{L^p} \chi_E$; the continuity of F implies that $\nu(E_n) \to \nu(E)$. So ν is a signed measure thanks to Exercise 8.17. It is easy to see that if $\mu(E) = 0$, then $\chi_E = 0$ in $L^p(X, \mu)$. Hence, $\nu(E) = 0$ and, by Exercise 8.23, we get $\nu \ll \mu$. Then the Radon-Nikodym Theorem (Theorem 8.24) ensures the existence of $g \in L^1(X, \mu)$ such that

$$F(\chi_E) = \int_E g\, d\mu \qquad \forall E \in \mathscr{E}.$$

By linearity, we have $F(f) = \int_X fg\, d\mu$ for any simple function $f : X \to \mathbb{R}$. Let now $f \in L^\infty(X, \mu)$. Applying Proposition 2.23 to f^+ and f^-, we construct a sequence of simple functions $f_n : X \to \mathbb{R}$ such that $|f_n| \le |f|$ and $f_n \xrightarrow{L^\infty} f$. By the Dominated Convergence Theorem we deduce that

$$F(f_n) = \int_X f_n g\, d\mu \to \int_X fg\, d\mu.$$

On the other hand, since μ is finite, thanks to Proposition 3.36 we conclude that $f_n \xrightarrow{L^p} f$. So $F(f_n) \to F(f)$.

2. *$g \in L^{p'}(X, \mu)$ and $\|g\|_{p'} \le \|F\|_*$.*

We distinguish two cases.

(2a) $1 < p < \infty$ (hence $1 < p' < \infty$). Given $k \in \mathbb{N}$, let $Y_k = \{x \in X \mid |g(x)| \le k\}$ and define

$$f_k(x) = \begin{cases} \chi_{Y_k}(x)\dfrac{|g(x)|^{p'}}{g(x)} & \text{if } g(x) \neq 0 \\ 0 & \text{if } g(x) = 0. \end{cases}$$

We have $f_k \in L^\infty(X, \mu)$; so $f_k \in L^p(X, \mu)$. Moreover $|f_k|^p = |g|^{p'}$ in Y_k; then

$$\int_{Y_k} |g|^{p'}\, d\mu = \int_X f_k g\, d\mu = F(f_k) \le \|F\|_* \|f_k\|_p$$
$$= \|F\|_* \left(\int_{Y_k} |g|^{p'}\, d\mu\right)^{\frac{1}{p}},$$

which yields

$$\left(\int_X \chi_{Y_k} |g|^{p'}\, d\mu\right)^{1/p'} \le \|F\|_*.$$

Passing to the limit as $k \to \infty$ and applying the Monotone Convergence Theorem we obtain $\|g\|_{p'} \le \|F\|_*$.

(2b) $p = 1$. For any $\varepsilon > 0$ set

$$A_\varepsilon = \{x \in X \mid g(x) \ge \|F\|_* + \varepsilon\}$$

and define $f_\varepsilon = \chi_{A_\varepsilon} \frac{g}{|g|}$. Then $f_\varepsilon \in L^1(X, \mu) \cap L^\infty(X, \mu)$ and $\|f_\varepsilon\|_1 = \mu(A_\varepsilon)$, and so

$$(\|F\|_* + \varepsilon)\mu(A_\varepsilon) \le \int_{A_\varepsilon} |g|\, d\mu = \int_X f_\varepsilon g\, d\mu = F(f_\varepsilon) \le \|F\|_* \mu(A_\varepsilon).$$

This implies $\mu(A_\varepsilon) = 0$ for any $\varepsilon > 0$, hence $\|g\|_\infty \le \|F\|_*$.

3. *Conclusion.*

For every $p \in [1, \infty)$ we have that $g \in L^{p'}(X, \mu)$ and $\|g\|_{p'} \le \|F\|_*$. Then F and F_g are bounded linear functionals coinciding on $L^\infty(X, \mu)$ which is dense in $L^p(X, \mu)$, so $F = F_g$. Moreover, recalling inequality (8.18),

$$\|g\|_{p'} \le \|F\|_* = \|F_g\|_* \le \|g\|_{p'}.$$

This complete the analysis of the case $\mu(X) < \infty$.

In the σ-finite case, consider a sequence $(X_k)_k \subset \mathscr{E}$ of disjoint sets such that $X = \cup_{k=1}^\infty X_k$. It is immediate that, for any $E \in \mathscr{E}$, the map

$$f \in L^p(E, \mu) \mapsto F(\tilde{f})$$

($\tilde{f}$ denotes the extension of f equal to zero outside E) is a continuous linear functional of norm less than or equal to $\|F\|_*$. Since, as we have just shown, the result holds true for the finite measure spaces $(X_k, \mathscr{E} \cap X_k, \mu)$ (see Remark 1.28), there exist functions $g_k \in L^{p'}(X_k, \mu)$ such that

$$F(\tilde{f}) = \int_{X_k} g_k f\, d\mu \quad \forall f \in L^p(X_k, \mu).$$

For every $x \in X$ set $g(x) = g_k(x)$ if $x \in X_k$ and let $Z_n = \cup_{k=1}^n X_k$. Since

$$F(\tilde{f}) = \int_{Z_n} g f\, d\mu \quad \forall f \in L^p(Z_n, \mu),$$

by the first part of the proof, $g \in L^{p'}(Z_n, \mu)$ and

$$\int_X |g|^{p'} \chi_{Z_n}\, d\mu = \int_{Z_n} |g|^{p'} d\mu \le \|F\|_*^{p'}.$$

Therefore, since $\chi_{Z_n} \uparrow 1$, by Fatou's Lemma $g \in L^{p'}(X, \mu)$ and $\|g\|_{p'} \leq \|F\|_*$. Finally, for any $f \in L^p(X, \mu)$,

$$F(\chi_{Z_n} f) = \int_{Z_n} gf \, d\mu = \int_X g\chi_{Z_n} f \, d\mu = F_g(\chi_{Z_n} f).$$

Since $f\chi_{Z_n} \xrightarrow{L^p} f$, we conclude that $F(f) = F_g(f)$ for every $f \in L^p(X, \mu)$. □

Remark 8.30 Theorem 8.29 actually holds for a generic measure space in the case $1 < p < \infty$, whereas it may fail for $p = 1$ when μ is not σ-finite (see Example 6.64). On the other hand, Theorem 8.29 is false, in general, for $p = \infty$: $L^1(X, \mu)$ does not provide all the bounded linear functionals on $L^\infty(X, \mu)$ (see Remark 6.63). The special case $p = p' = 2$ is already covered by the Riesz Representation Theorem, since $L^2(X, \mu)$ is a Hilbert space. Necessary and sufficient conditions on the space $(X, \mathscr{E}, \mu)$ to guarantee a characterization as in Theorem 8.29 are discussed in [Za67].

Reference

[Za67] Zaanen, A.C.: Integration. Noth-Holland, Amsterdam (1967)

Chapter 9
Set-Valued Functions

Motivated by applications to optimization and control theory, modern analysis has shown an increasing interest in set-valued maps, to which most of the known results for single-valued maps can be adapted. In this chapter, we provide a quick introduction to set-valued analysis aiming to deduce a classical theorem which guarantees the existence of a measurable selection.

9.1 Definitions and Examples

Given two integers $N, M \geq 1$, a *set-valued map* $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ is a map which associates to any $x \in \mathbb{R}^N$ a set $\Gamma(x) \subset \mathbb{R}^M$ (possibly empty). The set

$$D(\Gamma) = \{x \in \mathbb{R}^N \mid \Gamma(x) \neq \emptyset\}$$

is called the *domain* of Γ.

Example 9.1 1. Let $f : [a, b] \to \mathbb{R}$ be a lower semicontinuous function. Then

$$\Gamma(t) := \{x \in [a, b] \mid f(x) \leq t\}, \qquad t \in \mathbb{R},$$

is a set-valued map $\Gamma : \mathbb{R} \rightsquigarrow \mathbb{R}$ such that $D(\Gamma) = [\min f, \infty)$.

2. Given an integer $k \geq 1$, let $f : \mathbb{R}^N \times \mathbb{R}^k \to \mathbb{R}^M$ be a continuous function and let F be a closed nonempty subset of $\mathbb{R}^k$. Then

$$\Gamma(x) := f(x, F), \qquad x \in \mathbb{R}^N,$$

is a set-valued map $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ such that $D(\Gamma) = \mathbb{R}^N$.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0_9

Definition 9.2 Let $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ be a set-valued map. We say that Γ is:

(i) *closed* (*convex, compact*, respectively) if $\Gamma(x)$ is a closed (convex, compact, respectively) set for every $x \in \mathbb{R}^N$.
(ii) *Borel* if, for any open set $V \subset \mathbb{R}^M$, the inverse image

$$\Gamma^{-1}(V) := \{x \in \mathbb{R}^N \mid \Gamma(x) \cap V \neq \varnothing\}$$

is a Borel subset of $\mathbb{R}^N$.
(iii) *upper semicontinuous at a point* $x \in \mathbb{R}^N$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that[1]

$$\|x - x'\| < \delta \quad \Longrightarrow \quad \Gamma(x') \subset \Gamma(x) + B_\varepsilon \,.$$

(iv) *upper semicontinuous in a set* $E \subset \mathbb{R}^N$ if it is upper semicontinuous at every point of E.

Similarly, one can give a sense to lower semicontinuity and many other continuity properties for set-valued maps. In this chapter, however, we will confine ourselves to consider upper semicontinuous set-valued maps. For a more extended treatment of set-valued maps we refer to the monograph [AF90].

Exercise 9.3 Is the set-valued map $\Gamma : \mathbb{R} \rightsquigarrow \mathbb{R}$ of Example 9.1(1):

1. closed?
2. upper semicontinuous in $D(\Gamma)$?

Exercise 9.4 Let $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ be a closed set-valued map and let $x_0 \in \mathbb{R}^N$. The Kuratowski upper limit of Γ as $x \to x_0$ is defined by

$$\operatorname{Limsup}_{x\to x_0} \Gamma(x) = \left\{ y \in \mathbb{R}^M \;\middle|\; \begin{array}{lcl} \exists\, x_n \in D(\Gamma)\backslash\{x_0\} & : & x_n \to x_0 \\ \exists\, y_n \in \Gamma(x_n) & & y_n \to y \end{array} \right\}.$$

1. Show that if $x_0 \in \overline{D(\Gamma)}$, then[2]

$$\operatorname{Limsup}_{x\to x_0} \Gamma(x) = \left\{ y \in \mathbb{R}^M \;\middle|\; \liminf_{x\to x_0,\, x\in D(\Gamma)} d_{\Gamma(x)}(y) = 0 \right\}.$$

2. Show that if Γ is upper semicontinuous at x_0, then

$$\operatorname{Limsup}_{x\to x_0} \Gamma(x) \subset \Gamma(x_0). \tag{9.1}$$

[1] $\Gamma(x) + B_\varepsilon := \{y + z \mid y \in \Gamma(x),\ \|z\| < \varepsilon\}$.

[2] $d_{\Gamma(x)}(y)$ denotes the distance of the point y from the set $\Gamma(x)$.

3. Show that if

$$\exists\, r, R > 0 \quad \text{such that} \quad \|y\| \le R \quad \forall y \in \bigcup_{\|x-x_0\|<r} \Gamma(x), \tag{9.2}$$

then by (9.1) it follows that Γ is upper semicontinuous at x_0.

4. Does the above property hold without assuming (9.2)?

Hint. Consider $\Gamma : \mathbb{R} \rightsquigarrow \mathbb{R}$ defined by

$$\Gamma(x) = \begin{cases} \{n\} & \text{if} \quad x = \frac{1}{n}, \quad n \ge 1 \\ \emptyset & \text{otherwise} \end{cases}$$

(Γ fails to be upper semicontinuous at 0 but $\operatorname{Limsup}_{x\to 0}\Gamma(x) = \emptyset$).

Given a set-valued map $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$, the *graph* of Γ is defined by

$$\operatorname{Graph}(\Gamma) = \{(x, y) \in \mathbb{R}^N \times \mathbb{R}^M \mid y \in \Gamma(x)\}.$$

Proposition 9.5 *Given a set-valued map $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$, if* $\operatorname{Graph}(\Gamma)$ *is closed then Γ is closed and Borel.*

Proof The fact that Γ is closed is a direct consequence of the closure of $\operatorname{Graph}(\Gamma)$. We are going to prove that Γ is Borel.

First of all let us show that if $K \subset \mathbb{R}^M$ is compact, then $\Gamma^{-1}(K)$ is closed. Let $(x_n)_n \subset \Gamma^{-1}(K)$ be a sequence converging to a point $\bar{x} \in \mathbb{R}^N$. Then there exists a sequence $y_n \in \Gamma(x_n) \cap K$ and, by compactness, a subsequence y_{k_n} converging to a point $\bar{y} \in K$. Since $\operatorname{Graph}(\Gamma)$ is closed, the pair $(\bar{x}, \bar{y}) := \lim_n(x_{k_n}, y_{k_n})$ belongs to $\operatorname{Graph}(\Gamma)$. So $\bar{y} \in \Gamma(\bar{x})$ and, consequently, $\bar{x} \in \Gamma^{-1}(K)$. Therefore $\Gamma^{-1}(K)$ is closed, as claimed.

Let now $V \subset \mathbb{R}^M$ be open. Then $V = \cup_{n=1}^{\infty} K_n$ for some family $(K_n)_n$ of compact sets. So $\Gamma^{-1}(V) = \cup_{n=1}^{\infty} \Gamma^{-1}(K_n)$ is a countable union of closed sets. Hence, it is a Borel set. □

9.2 Existence of a Summable Selection

Definition 9.6 Given a set-valued map $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$, a *selection* of Γ on a nonempty set $S \subset D(\Gamma)$ is a function $\gamma : S \to \mathbb{R}^M$ such that $\gamma(x) \in \Gamma(x)$ for every $x \in S$.

The fact that any set-valued map admits at least one selection on $D(\Gamma)$ is a consequence of the Axiom of Choice. However, one is usually interested to know if there exist selections with suitable properties. In the sequel of the chapter we will provide sufficient conditions to guarantee the existence of a *summable selection* (with respect to the Lebesgue measure m) of a given set-valued map.

We say that $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ is *dominated by a summable function* on a Borel set $A \subset \mathbb{R}^N$ if there exists a function $g : A \to [0, \infty)$ such that[3] $g \in L^1(A)$ and, for every $x \in A$,

$$p \in \Gamma(x) \quad \Longrightarrow \quad \|p\| \leq g(x). \tag{9.3}$$

Theorem 9.7 *Let $\Gamma : \mathbb{R}^N \rightsquigarrow \mathbb{R}^M$ be closed, upper semicontinuous and dominated by a summable function on a nonempty Borel set $A \subset D(\Gamma)$ of finite measure. Then there exists*[4] $\gamma \in (L^1(A))^M$ *such that $\gamma(x) \in \Gamma(x)$ for almost every $x \in A$.*

Proof For the proof we need two technical steps.

1. Let us prove, first, that for every $f \in (L^1(A))^M$ there exists a Borel function $\phi(f) : A \to [0, \infty)$ such that

$$\phi(f)(x) = d_{\Gamma(x)}(f(x)) \ \text{ a.e. in } A$$

and

$$\phi(f)(x) \neq d_{\Gamma(x)}(f(x)) \ \Rightarrow \ \phi(f)(x) = 0.$$

To this aim, apply Corollary 2.30 of Lusin's Theorem to construct an increasing sequence of compact sets $K_n \subset A$ such that

$$\begin{cases} (a) \quad f_{\big|K_n} : K_n \to \mathbb{R}^M, \quad g_{\big|K_n} : K_n \to \mathbb{R}, \quad \text{are continuous} \quad \forall n \in \mathbb{N}, \\ (b) \quad m\big(A \setminus \cup_{n \geq 1} K_n\big) = 0, \end{cases}$$

and set

$$\phi(f) = \phi(f)(x) := \begin{cases} d_{\Gamma(x)}(f(x)) & \text{if } x \in \cup_{n \geq 1} K_n \\ 0 & \text{if } x \in A \setminus \cup_{n \geq 1} K_n. \end{cases} \tag{9.4}$$

Let us show that, for every $n \geq 1$, the restriction

$$\phi(f)_{\big|K_n} : K_n \to \mathbb{R}$$

is lower semicontinuous: given $n \in \mathbb{N}$ and $x_0 \in K_n$, let $x_j \in K_n$ and $p_j \in \Gamma(x_j)$ be sequences such that

[3] In what follows $L^1(A) = L^1(A, \mathscr{B}(A), m)$ where $\mathscr{B}(A)$ is the Borel σ-algebra and m denotes the Lebesgue measure on A.

[4] $(L^1(A))^M := \{(f_1, \ldots, f_M) \mid f_i \in L^1(A), \ \forall i = 1, \ldots, M\}$.

$$\liminf_{x \in K_j,\, x \to x_0} \phi(f)(x) = \lim_{j \to \infty} \phi(f)(x_j) = \lim_{j \to \infty} \|f(x_j) - p_j\|.$$

Observe that, owing to (9.3), $\|p_j\| \leq \max_{K_n} g$ for every $j \in \mathbb{N}$; so $(p_j)_j$ is bounded and we may suppose, up to a subsequence, that $p_j \to p_0$ as $j \to \infty$. Moreover, by part 2 of Exercise 9.4, we have $p_0 \in \Gamma(x_0)$. Therefore

$$\phi(f)(x_0) \leq \|f(x_0) - p_0\| = \lim_{j \to \infty} \|f(x_j) - p_j\| = \liminf_{x \in K_n,\, x \to x_0} \phi(f)(x).$$

Finally, setting

$$\phi_n : A \to \mathbb{R}, \quad \phi_n(x) = \begin{cases} \phi(f)\big|_{K_n}(x) & \text{if } \ x \in K_n \\ 0 & \text{if } \ x \in A \backslash K_n \end{cases} \qquad (n \in \mathbb{N}),$$

we get that ϕ_n is Borel in A and $\phi_n(x) \to \phi(f)(x)$ for every $x \in A$. So $\phi(f)$ is Borel.

2. Consider now the functional

$$J(f) := \int_A \phi(f)\, dx \qquad f \in (L^1(A))^M.$$

By the first step, J is well defined since the function $\phi(f)$ is positive and Borel. Moreover, thanks to hypothesis (9.3) and the Lipschitz continuity of the distance function, given $f \in (L^1(A))^M$, for almost every $x \in A$ we have

$$\phi(f)(x) = d_{\Gamma(x)}(f(x)) \leq \|f(x)\| + d_{\Gamma(x)}(0) \leq \|f(x)\| + g(x),$$

and so $J(f) < \infty$ for every $f \in (L^1(A))^M$. Moreover, if $f,\ g \in (L^1(A))^M$, for almost every $x \in A$

$$|\phi(f)(x) - \phi(g)(x)| = |d_{\Gamma(x)}(f(x)) - d_{\Gamma(x)}(g(x))| \leq |f(x) - g(x)|;$$

it follows that J is Lipschitz continuous with Lipschitz constant 1, hence continuous. We are going to show that J vanishes for at least an element of $(L^1(A))^M$. To begin with, apply Ekeland's Variational Principle (see Appendix H) to construct $\bar{f} \in (L^1(A))^M$ such that

$$J(f) > J(\bar{f}) - \frac{1}{3}\|f - \bar{f}\|_1 \qquad \forall f \in (L^1(A))^M \backslash \{\bar{f}\}. \tag{9.5}$$

Arguing by contradiction, suppose $J(\bar{f}) > 0$. Then

$$A_+ := \{x \in A \mid \phi(\bar{f}) > 0\}$$

is a Borel set and

$$\int_{A_+} \phi(\bar f)\,dx = \int_{A_+} d_{\Gamma(x)}(\bar f(x))\,dx > 0. \tag{9.6}$$

Given a dense sequence $(q_j)_{j\in\mathbb{N}}$ in $\mathbb{R}^M$, set

$$\begin{aligned} A_j &= \left\{x \in A_+ \;\middle|\; \|\bar f(x) - q_j\| < \frac{2}{3}\phi(f)(x),\ \phi(q_j)(x) < \frac{2}{3}\phi(\bar f)(x)\right\} \\ &= \left\{x \in A_+ \;\middle|\; \|\bar f(x) - q_j\| < \frac{2}{3}d_{\Gamma(x)}(\bar f(x)),\ \phi(q_j)(x) < \frac{2}{3}d_{\Gamma(x)}(\bar f(x))\right\}. \end{aligned}$$

Owing to the previous step, A_j is a Borel set for every $j \in \mathbb{N}$. Moreover, it is easy to prove that[5] $A_+ = \cup_{j\in\mathbb{N}} A_j$. Therefore by (9.6) it follows that, for at least an index j_0,

$$\int_{A_{j_0}} d_{\Gamma(x)}(\bar f(x))\,dx > 0.$$

Then, setting

$$\tilde f(x) = \begin{cases} q_{j_0} & \text{if } x \in A_{j_0}, \\ \bar f(x) & \text{if } x \in A\setminus A_{j_0}, \end{cases}$$

we have

$$\begin{aligned} \int_A \|\tilde f(x) - \bar f(x)\|\,dx &= \int_{A_{j_0}} \|q_{j_0} - \bar f(x)\|\,dx \\ &\le \frac{2}{3}\int_{A_{j_0}} d_{\Gamma(x)}(\bar f(x))\,dx. \end{aligned} \tag{9.7}$$

Moreover, by the definition of A_{j_0} we deduce

$$\begin{aligned} J(\tilde f) =& J(\bar f) - \int_{A_{j_0}} d_{\Gamma(x)}(\bar f(x))\,dx + \int_{A_{j_0}} \phi(q_{j_0})\,dx \\ &\le J(\bar f) - \frac{1}{3}\int_{A_{j_0}} d_{\Gamma(x)}(\bar f(x))\,dx < J(\bar f). \end{aligned}$$

[5]Indeed, let $x \in A_+$ and let $y \in \Gamma(x)$ be such that $\|\bar f(x) - y\| = d_{\Gamma(x)}(\bar f(x))$. Then, setting $z = \frac{1}{2}(\bar f(x) + y)$, z verifies $\|\bar f(x) - z\| = \frac{1}{2}d_{\Gamma(x)}(\bar f(x))$ and $d_{\Gamma(x)}(z) \le \|z - y\| = \frac{1}{2}d_{\Gamma(x)}(\bar f(x))$. So if $q_{n_j} \to z$, by the continuity of the distance function we have that $x \in A_{n_j}$ for large n.

Hence $\tilde{f} \neq \bar{f}$, and by (9.7) we conclude

$$J(\tilde{f}) \leq J(\bar{f}) - \frac{1}{2}\|\tilde{f} - \bar{f}\|_1,$$

in contrast with (9.5).

To conclude the proof it suffices to observe that the function $\bar{f}$ constructed in the previous step satisfies $d_{\Gamma(x)}(\bar{f}(x)) = 0$ for almost every $x \in A$. Thus,

$$\bar{f}(x) \in \Gamma(x) \text{ for almost every } x \in A.$$

□

Remark 9.8 If we modify the function γ of Theorem 9.7 on a set of measure zero, then, using the Axiom of Choice, the thesis of Theorem 9.7 can be reformulated as follows: there exists a selection $\tilde{\gamma}$ of Γ on A which coincides almost everywhere with a function in $L^1(A)$. Observe that this does not imply that $\tilde{\gamma}$ belongs to $L^1(A)$, since $\tilde{\gamma}$ may fail to be a Borel function. However, $\tilde{\gamma}$ is measurable with respect to the σ-algebra $\mathscr{G}$ of all Lebesgue measurable sets (Definition 1.56), since m is a complete measure on $\mathscr{G}$. Therefore, under the assumptions of Theorem 9.7, we deduce that Γ admits a selection $\tilde{\gamma}$ on A such that $\tilde{\gamma} \in L^1(A, \mathscr{G}, m)$.

Reference

[AF90] Aubin, J.-P., Frankowska, H.: Set-Valued Analysis. Birkhäuser, Boston (1990)

Appendix A
Distance Function

In this Appendix we will recall the basic properties of the distance function from a nonempty set $S \subset \mathbb{R}^N$.

Definition A.1 The *distance function* from S is the function $d_S : \mathbb{R}^N \to \mathbb{R}$ defined by

$$d_S(x) = \inf_{y \in S} \|x - y\| \qquad \forall x \in \mathbb{R}^N.$$

The *projection* of x onto S is the set consisting of those points (if any) at which the infimum defining $d_S(x)$ is attained. Such a set will be denoted by $\text{proj}_S(x)$.

Proposition A.2 *Let S be a nonempty subset of $\mathbb{R}^N$. Then the following properties hold:*

1. *d_S is Lipschitz continuous of rank 1, i.e., $|d_S(x) - d_S(x')| \le \|x - x'\|$ for any $x, x' \in \mathbb{R}^N$.*
2. *For any $x \in \mathbb{R}^N$ we have*
$$d_S(x) = 0 \iff x \in \overline{S}.$$
3. *S is closed $\iff \text{proj}_S(x) \ne \varnothing$ for every $x \in \mathbb{R}^N$.*

Proof 1. Let $x, x' \in \mathbb{R}^N$ and $\varepsilon > 0$ be fixed. Then there exists $y_\varepsilon \in S$ such that $\|x - y_\varepsilon\| < d_S(x) + \varepsilon$. Thus, by the triangle inequality for the Euclidean norm,

$$d_S(x') - d_S(x) \le \|x' - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon \le \|x' - x\| + \varepsilon.$$

Since ε is arbitrary, $d_S(x') - d_S(x) \le \|x' - x\|$. Exchanging the role of x and x' we conclude that $|d_S(x') - d_S(x)| \le \|x' - x\|$ as desired.

2. For any $x \in \mathbb{R}^N$ we have that $d_S(x) = 0$ if and only if there exists a sequence $(y_n)_n \subset S$ such that $\|x - y_n\| \to 0$ as $n \to \infty$, hence if and only if $x \in \overline{S}$.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

3. Let S be a closed set and let $x \in \mathbb{R}^N$ be fixed. Then

$$K := \{y \in S \mid \|x - y\| \leq d_S(x) + 1\}$$

is a nonempty compact set. Therefore any point $\widehat{x} \in K$ such that

$$\|x - \widehat{x}\| = \min_{y \in K} \|x - y\|$$

lies in $\text{proj}_S(x)$.
Conversely, assume $\text{proj}_S(y) \neq \emptyset$ for every $y \in \mathbb{R}^N$ and let $x \in \overline{S}$. Observe that, by part 2, $d_S(x) = 0$. Take $\widehat{x} \in \text{proj}_S(x)$. Then $\|x - \widehat{x}\| = 0$ and $x \in S$. □

Proposition A.3 *Given a nonempty closed set $S \subset \mathbb{R}^N$, let $x \in \mathbb{R}^N$ and $y \in S$. Then $y \in \text{proj}_S(x)$ if and only if*

$$(x - y) \cdot (y' - y) \leq \frac{1}{2}\|y' - y\|^2 \qquad \forall y' \in S. \tag{A.1}$$

Proof By definition, we have that $y \in \text{proj}_S(x)$ if and only if

$$\|x - y\|^2 \leq \|x - y'\|^2 \qquad \forall y' \in S.$$

Since $\|x - y'\|^2 = \|x - y\|^2 + \|y - y'\|^2 + 2(x - y) \cdot (y - y')$, the above inequality reduces to (A.1). □

Remark A.4 Let $S \subset \mathbb{R}^N$ be a nonempty closed set.

1. By applying (A.1) we easily get

$$y \in \text{proj}_S(x) \iff y \in \text{proj}_S(tx + (1 - t)y) \qquad \forall t \in [0, 1], \tag{A.2}$$

and so

$$y \in \text{proj}_S(x) \iff d_S(tx + (1 - t)y) = t\|x - y\| \qquad \forall t \in [0, 1]. \tag{A.3}$$

To justify the implication '$\Rightarrow$' in (A.2) (the $\Leftarrow$-part is immediate), let us fix $y \in \text{proj}_S(x)$. By (A.1) it follows that, for every $t \in [0, 1]$,

$$t(x - y) \cdot (y' - y) \leq \frac{t}{2}\|y' - y\|^2 \leq \frac{1}{2}\|y' - y\|^2 \qquad \forall y' \in S. \tag{A.4}$$

Thus, $y \in \text{proj}_S(tx + (1 - t)y)$ by (A.1) applied to the point $tx + (1 - t)y$.
2. Another interesting remark is the following:

$$y \in \text{proj}_S(x) \implies \text{proj}_S(tx + (1 - t)y) = \{y\} \qquad \forall t \in [0, 1). \tag{A.5}$$

Indeed, by (A.4) it follows that, for every $t \in [0, 1)$,

$$t(x-y)\cdot(y'-y) < \frac{1}{2}\|y'-y\|^2 \qquad \forall y' \in S \setminus \{y\}.$$

So, since $\|(tx+(1-t)y)-y\|^2 = \|(tx+(1-t)y)-y'\|^2 - \|y-y'\|^2 + 2t(x-y)\cdot(y'-y)$,

$$\|(tx+(1-t)y)-y\|^2 < \|(tx+(1-t)y)-y'\|^2 \qquad \forall y' \in S \setminus \{y\}.$$

The thesis (A.5) follows.

3. It is useful to observe that the set-valued map $\mathrm{proj}_S : \mathbb{R}^N \to \mathscr{P}(\mathbb{R}^N)$ is *upper semicontinuous*, i.e., for every $x \in \mathbb{R}^N$,

$$\limsup_{x'\to x} \mathrm{proj}_S(x') \subset \mathrm{proj}_S(x),$$

where the meaning of the above limit is the following:

$$\forall \varepsilon > 0\ \exists \delta > 0\ :\ \|x'-x\| < \delta \implies \mathrm{proj}_S(x') \subset \mathrm{proj}_S(x) + B_\varepsilon, \quad \text{(A.6)}$$

setting[1] $B_\varepsilon = \{x \in \mathbb{R}^N \mid \|x\| < \varepsilon\}$. To see this, it will be sufficient to prove that, given any two sequences $(x_n)_n$, $(y_n)_n$ in $\mathbb{R}^N$, we have

$$x_n \to x\ ,\ y_n \in \mathrm{proj}_S(x_n),\ y_n \to \overline{y} \implies \overline{y} \in \mathrm{proj}_S(x).$$

Indeed, $\overline{y} \in S$ since S is closed, and $\|x_n - y_n\| = d_S(x_n)$. Then the continuity of d_S implies that $\|x-\overline{y}\| = d_S(x)$, and so $\overline{y} \in \mathrm{proj}_S(x)$.

4. We point out, in particular, the fact that proj_S is *continuous* at all points $x \in \mathbb{R}^N$ for which $\mathrm{proj}_S(x)$ reduces to a singleton, that is,

$$\lim_{x'\to x} \mathrm{proj}_S(x') = \mathrm{proj}_S(x)$$

where the above limit means that proj_S satisfies (A.6) and also

$$\forall \varepsilon > 0\ \exists \delta > 0\ :\ \|x'-x\| < \delta \implies \mathrm{proj}_S(x) \subset \mathrm{proj}_S(x') + B_\varepsilon.$$

This is an immediate consequence of upper semicontinuity and the fact that $\mathrm{proj}_S(x)$ is a singleton.

A general principle is that geometric properties of the set S correspond to analytic properties of the function d_S. The following differentiability theorem is a case in point.

[1]The sum of two sets A and A' in $\mathbb{R}^N$ is defined by $A + A' := \{x + y \mid x \in A,\ y \in A'\}$.

Theorem A.5 *Let $S \subset \mathbb{R}^N$ be closed and nonempty. Then d_S is Fréchet differentiable at a point $x \in \mathbb{R}^N \setminus S$ if and only if* $\mathrm{proj}_S(x)$ *reduces to a singleton* $\{y\}$. *Moreover, in such a case,*

$$Dd_S(x) = \frac{x-y}{\|x-y\|}. \tag{A.7}$$

Proof Suppose, first, that d_S is Fréchet differentiable at $x \notin S$. Then, fixed $y \in \mathrm{proj}_S(x)$, the function $t \mapsto d_S(tx+(1-t)y)$ has left-hand derivative at $t=1$ which satisfies, by (A.5),

$$Dd_S(x)\cdot(x-y) = \frac{d}{dt^-}d_S(tx+(1-t)y)\Big|_{t=1} = \|x-y\|.$$

Moreover, since d_S is Lipschitz continuous of rank 1 owing to Proposition A.2, we get $\|Dd_S(x)\| \le 1$. So, by the Cauchy-Schwarz inequality,

$$1 = Dd_S(x)\cdot\frac{x-y}{\|x-y\|} \le 1.$$

(A.7) follows by recalling the cases when equality holds in Proposition 5.3. Furthermore, y is uniquely determined by (A.7) since

$$y = x - d_S(x)Dd_S(x).$$

Vice versa, suppose that $\mathrm{proj}_S(x) = \{y\}$. Then according to Remark A.4(4) because this condition.

$$\lim_{x'\to x} \mathrm{proj}_S(x') = \{y\}.$$

Consequently, the differentiability of d_S^2 (hence, of d_S) will follow once we have proved that, for every $x' \in \mathbb{R}^N$ and $y' \in \mathrm{proj}_S(x')$,

$$\begin{aligned}&\|x-x'\|^2 - 2\|x'-x\|\,\|y'-y\|\\ &\qquad\le d_S^2(x') - d_S^2(x) - 2(x-y)\cdot(x'-x) \le \|x-x'\|^2.\end{aligned}$$

To this aim observe that, in view of (A.1),

$$\begin{aligned}&d_S^2(x') - d_S^2(x) - 2(x-y)\cdot(x'-x)\\ &= \|x'-x\|^2 + \|y'-y\|^2 + 2(x'-x)\cdot(y-y') + 2(x-y)\cdot(y-y')\\ &\ge \|x'-x\|^2 + 2(x'-x)\cdot(y-y')\\ &\ge \|x-x'\|^2 - 2\|x'-x\|\,\|y'-y\|.\end{aligned} \tag{A.8}$$

Moreover,

$$
\begin{aligned}
&2(x-y)\cdot(y-y')+\|y'-y\|^2\\
&=2(x-y'+y'-y)\cdot(y-y')+\|y'-y\|^2\\
&=2(x-y')\cdot(y-y')-\|y'-y\|^2\\
&=2(x-x')\cdot(y-y')+2(x'-y')\cdot(y-y')-\|y'-y\|^2\\
&\le 2(x-x')\cdot(y-y')
\end{aligned}
\tag{A.9}
$$

again by (A.1) applied to x'. The desired inequalities are then immediate consequences of (A.8) and (A.9). □

Exercise A.6 Let $\Omega \subset \mathbb{R}^N$ be a nonempty bounded open set with boundary Γ. Show that there exists at least one point in Ω where d_Γ fails to be differentiable. *Hint.* Let $x \in \Omega$. If d_Γ is differentiable at x, then consider $x + tDd_\Gamma(x)$ for $t > 0\dots$

Exercise A.7 Let $S \subset \mathbb{R}^N$ be closed and nonempty.

1. Given $x \in \mathbb{R}^N \setminus S$ and $y \in \operatorname{proj}_S(x)$, show that d_S is Fréchet differentiable at every point of the open segment

$$\{tx + (1-t)y \mid t \in (0,1)\}.$$

2. Show that if S is convex, then d_S is a convex function on $\mathbb{R}^N$.
3. Prove the (semiconcavity) inequality

$$t d_S^2(x) + (1-t)d_S^2(x') - d_S^2(tx + (1-t)x') \le t(1-t)\|x - x'\|^2$$

for every $x, x' \in \mathbb{R}^N$ and $t \in [0,1]$. Deduce that the function

$$\phi_S(x) := \|x\|^2 - d_S^2(x), \quad x \in \mathbb{R}^N$$

is convex.

Appendix B
Semicontinuous Functions

Let (X, d) be a metric space. We now introduce the notion of semicontinuous function, which arises as a natural generalization of the concept of continuous function.

Definition B.1 A function $f : X \to \mathbb{R} \cup \{\infty\}$ is said to be *lower semicontinuous (lsc) at a point* $x_0 \in X$ if

$$f(x_0) \leq \liminf_{x \to x_0} f(x).$$

Similarly, a function $f : X \to \mathbb{R} \cup \{-\infty\}$ is said to be *upper semicontinuous (usc) at* x_0 if

$$f(x_0) \geq \limsup_{x \to x_0} f(x).$$

Remark B.2 1. f is lsc at x_0 if and only if $-f$ is usc at x_0.
2. If f_1, f_2 are lsc (respectively, usc) at x_0, then $f_1 + f_2$ is lsc (respectively, usc) at x_0.
3. If $\alpha > 0$ and f is lsc (respectively, usc) at x_0, then αf is lsc (respectively, usc) at x_0.
4. A function $f : X \to \mathbb{R}$ is continuous at x_0 if and only if f is both lsc and usc at x_0.

Example B.3 As simple examples of functions which are lsc everywhere in $\mathbb{R}$ but discontinuous at some x_0, we have

$$u_1(x) = \begin{cases} 0 & \text{if } x \leq x_0, \\ 1 & \text{if } x > x_0, \end{cases} \qquad u_2(x) = \begin{cases} 1 & \text{if } x \neq x_0, \\ 0 & \text{if } x = x_0. \end{cases}$$

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

Hence, $-u_1$ and $-u_2$ are usc everywhere in $\mathbb{R}$. The *Dirichlet function*

$$u(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is lsc at all irrational numbers and usc at all rational ones.

The next theorem characterizes lsc and usc functions.

Theorem B.4 *(i) A function $f : X \to \mathbb{R} \cup \{\infty\}$ is lsc*[2] *if and only if the sets $\{f \le a\}$ are closed (equivalently, the sets $\{f > a\}$ are open) for every $a \in \mathbb{R}$.*
(ii) A function $f : X \to \mathbb{R} \cup \{-\infty\}$ is usc if and only if the sets $\{f \ge a\}$ are closed (equivalently, the sets $\{f < a\}$ are open) for every $a \in \mathbb{R}$.

Proof Statements (i) and (ii) are equivalent since f is lsc if and only if $-f$ is usc. It is therefore enough to prove (i). Suppose, first, that f is lsc in X. Given $a \in \mathbb{R}$, let $x_0 \in X$ be a limit point of the set $\{f \le a\}$. Then there exists a sequence $(x_n)_n \subset X$ such that $x_n \to x_0$ and $f(x_n) \le a$. Since f is lsc at x_0, we have $f(x_0) \le \liminf_{n\to\infty} f(x_n)$. Therefore $f(x_0) \le a$, so that $x_0 \in \{f \le a\}$. This shows that $\{f \le a\}$ is closed.

Vice versa, let $x_0 \in X$. If f is not lsc at x_0, then there exist $M \in \mathbb{R}$ and $(x_n)_n \subset X$ such that $f(x_0) > M$, $x_n \to x_0$ and $f(x_n) \le M$. Hence, the set $\{f \le M\}$ is not closed since it does not include all its limit points. □

Corollary B.5 *If f is lsc (respectively, usc) in X, then f is Borel.*

Proof Let f be lsc in X. $\{f \le a\}$ is a Borel set, since it is closed, and the conclusion follows from Exercise 2.11. □

Corollary B.6 *If $(f_i)_{i\in I}$ is a family of lsc functions in X, then $\sup_{i\in I} f_i$ is lsc in X. If $(f_i)_{i\in I}$ is a family of usc functions in X, then $\inf_{i\in I} f_i$ is usc in X.*

Proof Since f is lsc if and only if $-f$ is usc and $\inf_{i\in I} f_i = -\sup_{i\in I}(-f_i)$, it is sufficient to prove the result for lsc functions. But this easily follows from Theorem B.4 and the fact that $\{\sup_{i\in I} f_i \le a\} = \cap_{i\in I}\{f_i \le a\}$. □

The next theorem generalizes to semicontinuous functions the analogous well-known result for continuous functions.

Theorem B.7 *Let X be a compact metric space.*

- *If $f : X \to \mathbb{R} \cup \{\infty\}$ is a lsc function, then f has a minimum point in X.*
- *If $f : X \to \mathbb{R} \cup \{-\infty\}$ is a usc function, then f has a maximum point in X.*

Proof Let $f : X \to \mathbb{R}\cup\{\infty\}$ be lsc. First of all, we will show that f is bounded from below. Indeed, suppose $\inf_X f = -\infty$. Then there exists a sequence $(y_n)_n \subset X$ such

[2] That is, lsc at every point of X.

that $f(y_n) < -n$ for every n. Recalling that X is compact, we get the existence of a subsequence $(y_{n_k})_k$ converging to a point $y \in X$. Since f is lsc at y, it follows that

$$-\infty < f(y) \leq \liminf_{k\to\infty} f(y_{n_k}) = -\infty,$$

which is a contradiction. So f is bounded from below and

$$\lambda := \inf_{x\in X} f(x) > -\infty.$$

A sequence $(x_n)_n \subset X$ exists such that

$$f(x_n) \leq \lambda + \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

The compactness of X implies the existence of a subsequence $(x_{n_k})_k$ which converges to a point $x_0 \in X$. The fact that f is lsc at x_0 gives

$$\lambda \leq f(x_0) \leq \liminf_{k\to\infty} f(x_{n_k}).$$

On the other hand, by construction we have $\liminf_{k\to\infty} f(x_{n_k}) \leq \lambda$. Thus, $f(x_0) = \lambda$.

The second statement follows by applying the first part to $-f$. □

Appendix C
Finite-Dimensional Linear Spaces

In the Euclidean space $\mathbb{R}^N$ let us consider the norm

$$\|\xi\| = \left(\sum_{i=1}^{N} |\xi_i|^2\right)^2 \quad \forall \xi = (\xi_1, \ldots, \xi_N) \in \mathbb{R}^N.$$

We will prove in this appendix that any normed linear space X of finite dimension N can be identified with $\mathbb{R}^N$; more precisely, X and $\mathbb{R}^N$ are topologically isomorphic in the sense of the following definition.

Definition C.1 Two normed linear spaces X and Y are said to be *topologically isomorphic* if there exists a bijective linear map $T : X \to Y$ such that T and T^{-1} are continuous.

Theorem C.2 *Let X and Y be two normed linear spaces such that* $\dim X = \dim Y = N$*. Then X and Y are topologically isomorphic.*

Proof Since the topological isomorphism is a transitive relation, it is sufficient to prove that X is topologically isomorphic to $\mathbb{R}^N$. Let $x_1, \ldots, x_N$ be a basis of X and define

$$T : \mathbb{R}^N \to X, \qquad T(\xi_1, \ldots, \xi_N) = \xi_1 x_1 + \cdots + \xi_N x_N.$$

Then T is a bijective linear map and, by Cauchy–Schwarz inequality,

$$\|T(x)\| \le \sum_{i=1}^{N} |\xi_i| \|x_i\| \le \left(\sum_{i=1}^{N} |\xi_i|^2\right)^{1/2} \left(\sum_{i=1}^{N} \|x_i\|^2\right)^{1/2} = M\|\xi\|$$

where we have set $M = \left(\sum_{i=1}^{N} \|x_i\|^2\right)^{1/2}$. So T is a bounded linear operator, hence it is continuous. There remains to show that T^{-1} is also continuous. Denote by S the unit sphere in $\mathbb{R}^N$, i.e., $S = \{\xi \in \mathbb{R}^N \mid \|\xi\| = 1\}$. Then $T(S)$ is compact, and,

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

consequently, closed in X. Since T is bijective, we deduce that $0 \notin T(S)$, therefore there exists $m > 0$ such that the ball $\|x\| < m$ is disjoint from $T(S)$, that is,

$$\|T(\xi)\| \geq m \quad \forall \xi \in S.$$

We have

$$\|T(\xi)\| \geq m\|\xi\| \quad \forall \xi \in \mathbb{R}^N,$$

or, equivalently,

$$\|T^{-1}(x)\| \leq m^{-1}\|x\| \quad \forall x \in X,$$

which implies that T^{-1} is continuous. □

It is apparent that if X and Y are topologically isomorphic and if X is complete, then Y is also complete. Since $\mathbb{R}^N$ is complete, we get the following result.

Theorem C.3 *Every finite-dimensional normed linear space is complete.*

Corollary C.4 *If X is a normed linear space, then any finite-dimensional subspace of X is closed.*

Corollary C.5 *Let X be a finite-dimensional normed linear space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on X. Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, i.e., there exist constants $m, M > 0$ such that*

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1 \quad \forall x \in X.$$

Proof Let us proceed like in the proof of Theorem C.2, denoting by $x_1, \ldots, x_N$ a basis X. For suitable constants $m_1, m_2, M_1, M_2 > 0$, we have that

$$m_1\|\xi\| \leq \left\| \sum_{i=1}^{N} \xi_i x_i \right\|_1 \leq M_1\|\xi\| \quad \forall \xi \in \mathbb{R}^N,$$

$$m_2\|\xi\| \leq \left\| \sum_{i=1}^{N} \xi_i x_i \right\|_2 \leq M_2\|\xi\| \quad \forall \xi \in \mathbb{R}^N.$$

Hence

$$\frac{m_2}{M_1}\|x\|_1 \leq \|x\|_2 \leq \frac{M_2}{m_1}\|x\|_1 \quad \forall x \in X.$$

□

It is well-known that a subset of $\mathbb{R}^N$ is compact if and only if it is closed and bounded (Bolzano–Weierstrass Theorem). Taking into account that the property of being closed, bounded, or compact is invariant under topological isomorphism, owing to Theorem C.2 such a characterization of compact sets holds also in finite-dimensional spaces.

Corollary C.6 *If X is a finite-dimensional normed linear space, then a subset of X is compact if and only if it is closed and bounded.*

The above property actually holds only in finite-dimensional spaces, as shown by the following result.

Theorem C.7 (F. Riesz) *Let X be a normed linear space such that the unit sphere*

$$S = \{x \in X \mid \|x\| = 1\}$$

is compact. Then X is finite-dimensional.

Proof Let us consider the open cover of S constituted by all the open balls having centers in S and radius $\frac{1}{2}$. Since S is compact, there exists a finite set $\{x_1, \dots, x_N\} \subset S$ such that S is covered by the union of the open balls having centers in $x_1, \dots, x_N$ and radius $\frac{1}{2}$. Let M be the N–dimensional (closed) subspace generated by $x_1, \dots, x_N$. We claim that $M = X$. Otherwise, let $x_0 \in X \setminus M$ and $d = \inf_{x \in M} \|x_0 - x\|$. Since M is closed, we have that $d > 0$. There exists $y \in M$ such that $\|x_0 - y\| < 2d$. Setting $\overline{x} = \frac{x_0 - y}{\|x_0 - y\|} \in S$, for every $x \in M$ we have

$$\|x - \overline{x}\| = \frac{1}{\|x_0 - y\|} \left\| \big(\|x_0 - y\| x + y\big) - x_0 \right\| \geq \frac{d}{\|x_0 - y\|} \geq \frac{1}{2}.$$

So $\overline{x}$ does not belong to any of the balls covering S. Then $M = X$ and X has finite dimension. □

Appendix D
Baire's Lemma

The following result is classical in topology and is usually referred to as Baire's Lemma.

Proposition D.1 (Baire) *Let (X, d) be a complete metric space. Then the following properties hold:*

(a) Any countable intersection of dense open sets V_n is dense.
(b) If X is the countable union of closed sets F_n, then at least one of the F_n's has nonempty interior.

Proof We shall use the closed balls

$$\overline{B}_r(x) := \{y \in X \mid d(x, y) \leq r\} \qquad r > 0\,,\ x \in X. \tag{D.1}$$

(a) Let us fix any ball $\overline{B}_{r_0}(x_0)$. We shall prove that $\left(\cap_{n=1}^{\infty} V_n\right) \cap \overline{B}_{r_0}(x_0) \neq \varnothing$. Since V_1 is dense, there exists a point $x_1 \in V_1 \cap B_{r_0}(x_0)$. Since V_1 is an open set, there also exists $0 < r_1 < 1$ such that

$$\overline{B}_{r_1}(x_1) \subset V_1 \cap B_{r_0}(x_0).$$

Since V_2 is dense, we can find a point $x_2 \in V_2 \cap B_{r_1}(x_1)$ and (since V_2 is open) a radius $0 < r_2 < 1/2$ such that

$$\overline{B}_{r_2}(x_2) \subset V_2 \cap B_{r_1}(x_1).$$

Iterating the above procedure, we can construct a decreasing sequence of closed balls $\overline{B}_{r_n}(x_n)$ such that

$$\overline{B}_{r_n}(x_n) \subset V_n \cap B_{r_{n-1}}(x_{n-1}) \quad \text{and} \quad 0 < r_n < \frac{1}{n}. \tag{D.2}$$

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

We claim that $(x_n)_n$ is a Cauchy sequence in X. Indeed, for any $h, k \geq n$ we have $x_h,\ x_k \in B_{r_n}(x_n)$ by construction. So $d(x_k, x_h) < 2r_n < 2/n$. Therefore, X being complete, $(x_n)_n$ converges to a point $x \in X$. Since x_k belongs to $\overline{B}_{r_n}(x_n)$ for $k > n$, we conclude that $x \in \overline{B}_{r_n}(x_n)$ for every n and, by (D.2), $x \in V_n$ for every n.

(b) Suppose, by contradiction, that all F_n's have empty interior. Applying part (a) to the open sets $V_n := X \setminus F_n$, we can find a point $x \in \cap_{n=1}^{\infty} V_n$. Then $x \in X \setminus \cup_{n=1}^{\infty} F_n$ in contrast with the fact that the F_n's cover X. □

Remark D.2 Recalling the closed balls (D.1) used in the above proof, we observe that, for such a family of closed sets,

$$\overline{B_r(x)} \subset \overline{B}_r(x).$$

The fact that, in general, the inclusion is strict can be verified by considering, in a set $X \neq \varnothing$, the discrete metric

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases} \qquad \forall x, y \in X.$$

Then we have, for every $x \in X$, $B_1(x) = \{x\} = \overline{B_1(x)}$ while $\overline{B}_1(x) = X$.

Appendix E
Relatively Compact Families of Continuous Functions

Let K be a compact topological space. We denote by $\mathscr{C}(K)$ the Banach space of all continuous functions $f : K \to \mathbb{R}$ endowed with the uniform norm

$$\|f\|_\infty = \max_{x\in K} |f(x)| \qquad \forall f \in \mathscr{C}(K).$$

In what follows, we shall use the open balls

$$B_r(f) := \big\{g \in \mathscr{C}(K) \mid \|f - g\|_\infty < r\big\} \qquad r > 0,\ f \in \mathscr{C}(K).$$

Definition E.1 A family $\mathscr{M} \subset \mathscr{C}(K)$ is said to be:

(i) *equicontinuous* if, for any $\varepsilon > 0$ and $x \in K$, there exists a neighbourhood V of x in K such that

$$|f(x) - f(y)| < \varepsilon \qquad \forall y \in V,\ \forall f \in \mathscr{M}.$$

(ii) *pointwise bounded* if, for any $x \in X$, $\{f(x) \mid f \in \mathscr{M}\}$ is a bounded subset of $\mathbb{R}$.

Theorem E.2 (Ascoli–Arzelà) *A family $\mathscr{M} \subset \mathscr{C}(K)$ is relatively compact*[3] *if and only if $\mathscr{M}$ is equicontinuous and pointwise bounded.*

Proof Since $\mathscr{M}$ be relatively compact, $\mathscr{M}$ is totally bounded in $\mathscr{C}(K)$—hence, pointwise bounded. So it suffices to show that $\mathscr{M}$ is equicontinuous. Fix $\varepsilon > 0$ and let $f_1, \cdots, f_n \in \mathscr{M}$ be such that $\mathscr{M} \subset B_\varepsilon(f_1) \cup \cdots \cup B_\varepsilon(f_n)$. Let $x \in K$. Since each function f_i is continuous at x, x possesses neighbourhoods $V_1, \ldots, V_n \subset K$ such that

$$|f_i(x) - f_i(y)| < \varepsilon \qquad \forall y \in V_i\,, \quad i = 1, \ldots, n.$$

[3]That is, the closure $\overline{\mathscr{M}}$ is compact.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

Set $V := V_1 \cap \cdots \cap V_n$ and fix $f \in \mathscr{M}$. Let $i \in \{1, \ldots, n\}$ be such that $f \in B_\varepsilon(f_i)$. Thus, for any $y \in V$,

$$|f(y) - f(x)| \leq |f(y) - f_i(y)| + |f_i(y) - f_i(x)| + |f_i(x) - f(x)| < 3\varepsilon.$$

This shows that $\mathscr{M}$ is equicontinuous.

Conversely, given a pointwise bounded equicontinuous family $\mathscr{M}$, since K is compact, for any $\varepsilon > 0$ there exist points $x_1, \ldots, x_m \in K$ and corresponding neighbourhoods $V_1, \ldots, V_m$ such that $K = V_1 \cup \cdots \cup V_m$ and

$$|f(x) - f(x_i)| < \varepsilon \qquad \forall f \in \mathscr{M}, \quad \forall x \in V_i, \quad i = 1, \ldots, m. \tag{E.1}$$

Since $\{(f(x_1), \ldots, f(x_m)) \mid f \in \mathscr{M}\}$ is a bounded set, hence relatively compact in $\mathbb{R}^m$, there exist functions $f_1, \ldots, f_n \in \mathscr{M}$ such that

$$\{(f(x_1), \ldots, f(x_m)) \mid f \in \mathscr{M}\} \subset \bigcup_{j=1}^{n} Q_j, \tag{E.2}$$

where $\{Q_j\}_{j=1}^n$ denotes the family of open cubes in $\mathbb{R}^m$ defined as

$$Q_j = \big(f_j(x_1) - \varepsilon, f_j(x_1) + \varepsilon\big) \times \cdots \times \big(f_j(x_m) - \varepsilon, f_j(x_m) + \varepsilon\big).$$

We claim that

$$\mathscr{M} \subset B_{3\varepsilon}(f_1) \cup \cdots \cup B_{3\varepsilon}(f_n), \tag{E.3}$$

which implies that $\mathscr{M}$ is totally bounded,[4] hence relatively compact. To obtain (E.3), let $f \in \mathscr{M}$ and let $j \in \{1, \ldots, n\}$ be such that

$$(f(x_1), \ldots, f(x_m)) \in Q_j.$$

Now, fix $x \in K$ and let $i \in \{1, \ldots, m\}$ be such that $x \in V_i$. Then, by (E.1),

$$|f(x) - f_j(x)| \leq |f(x) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(x)| < 3\varepsilon.$$

This proves (E.3) and completes the proof. □

Remark E.3 The compactness of K is essential in the Ascoli–Arzelà Theorem. Indeed, the sequence

$$f_n(x) := e^{-(x-n)^2} \qquad (x \in \mathbb{R}, \; n \in \mathbb{N})$$

[4] See footnote 4 of Chap. 4 at p. 117.

is a pointwise bounded equicontinuous family in $\mathscr{C}(\mathbb{R})$. On the other hand,

$$n \neq m \quad \Longrightarrow \quad \|f_n - f_m\|_\infty \geq 1 - \frac{1}{e}.$$

So $(f_n)_n$ fails to be relatively compact.

Exercise E.4 With reference to the proof of Theorem E.2, show how to construct $f_1, \ldots, f_n \in \mathscr{M}$ satisfying (E.2).

Appendix F
Legendre Transform

Let $f : \mathbb{R}^N \to \mathbb{R}$ be a convex function. The function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ defined by[5]

$$f^*(y) = \sup_{x \in \mathbb{R}} \{x \cdot y - f(x)\} \qquad \forall y \in \mathbb{R}^N$$

is called the *Legendre transform*[6] (and, sometimes, *Fenchel transform* or *convex conjugate*) of f. By the definition of f^* it follows that

$$x \cdot y \leq f(x) + f^*(y) \qquad \forall x, y \in \mathbb{R}^N. \tag{F.1}$$

We say that f has *superlinear growth* at ∞ if

$$\lim_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \infty. \tag{F.2}$$

The following proposition describes some of the main properties of the Legendre transform.

Proposition F.1 *Let $f : \mathbb{R}^N \to \mathbb{R}$ be a differentiable convex function with superlinear growth at ∞. Then the following properties hold:*

(a) For every $y \in \mathbb{R}^N$ there exists $x_y \in \mathbb{R}^N$ such that $f^*(y) = x_y y - f(x_y)$.
(b) f^* is finite valued, that is, $f^* : \mathbb{R}^N \to \mathbb{R}$.
(c) For every $x,\ y \in \mathbb{R}^N$,

$$y = Df(x) \iff f^*(y) + f(x) = x \cdot y.$$

[5] $x \cdot y$ denotes the scalar product between the vectors $x,\ y \in \mathbb{R}^N$.

[6] Legendre transform is a classical tool in convex analysis, see [Ro70].

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

(d) f^* *is convex.*
(e) f^* *has superlinear growth at* ∞.
(f) $f^{**} = f$.

Proof (a) Let $y \in \mathbb{R}^N$. Observe that the function $F_y(x) = x \cdot y - f(x)$ is continuous and verifies $\lim_{\|x\| \to \infty} F_y(x) = -\infty$ thanks to (F.2); consequently, F_y attains its maximum value at some point x_y.
(b) The conclusion follows immediately from (a).
(c) Observe that, since the function $F_y(x) = x \cdot y - f(x)$ is concave, then $DF_y(x) = 0$ if and only if x is a maximum point for F_y. So, given $x, y \in \mathbb{R}^N$, we have

$$y = Df(x) \Longleftrightarrow DF_y(x) = 0 \Longleftrightarrow F_y(x) = \sup_{z \in \mathbb{R}^N} F_y(z) = f^*(y).$$

(d) Let $y_1, y_2 \in \mathbb{R}^N$ and $t \in [0, 1]$, and let x_t be a point such that

$$f^*(ty_1 + (1-t)y_2) = \big(ty_1 + (1-t)y_2\big) \cdot x_t - f(x_t).$$

Since $f^*(y_i) \geq y_i \cdot x_t - f(x_t)$ for $i = 1, 2$, we conclude that

$$f^*(ty_1 + (1-t)y_2) \leq tf^*(y_1) + (1-t)f^*(y_2),$$

which says that f^* is convex.
(e) For every $M > 0$ and $y \in \mathbb{R}^N$ we have

$$f^*(y) \geq M\frac{y}{\|y\|} \cdot y - f\left(M\frac{y}{\|y\|}\right) \geq M\|y\| - \max_{\|x\|=M} f(x).$$

So

$$\liminf_{\|y\| \to \infty} \frac{f^*(y)}{\|y\|} \geq M.$$

The arbitrariness of M implies that f^* is superlinear.
(f) By (F.1), we obtain that $f(x) \geq x \cdot y - f^*(y)$ for every $x, y \in \mathbb{R}^N$. Therefore $f \geq f^{**}$. To prove the opposite inequality, let us fix $x \in \mathbb{R}^N$ and let $y_x = Df(x)$. Then, owing to step (c) and (F.1),

$$f(x) = x \cdot y_x - f^*(y_x) \leq f^{**}(x).$$

The conclusion follows.

□

Example F.2 (Young's inequality) Let us define, for $p > 1$,

$$f(x) = \frac{|x|^p}{p} \qquad \forall x \in \mathbb{R}.$$

Then f is a superlinear function of class $\mathscr{C}^1(\mathbb{R})$. Moreover,

$$f'(x) = |x|^{p-1}\operatorname{sign}(x)\,,$$

where

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

Thus, f' is an increasing function, and so f is convex.

On account of step (c) of Proposition F.1, in order to compute $f^*(y)$ it is sufficient to solve $y = Df(x)$, i.e., $y = |x|^{p-1}\operatorname{sign}(x)$. Now, since the solution is given by $x_y = |y|^{1/(p-1)}\operatorname{sign}(y)$, we obtain

$$f^*(y) = x_y y - f(x_y) = \frac{|y|^{p'}}{p'} \qquad \forall y \in \mathbb{R},$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. Then, thanks to (F.1), we obtain the following estimate:

$$xy \leq \frac{x^p}{p} + \frac{y^{p'}}{p'} \qquad \forall x, y \geq 0. \tag{F.3}$$

By using again step (c) of Proposition F.1, we conclude that equality holds in (F.3) if and only if $y = Df(x)$, that is, $y^{p'} = x^p$.

Exercise F.3 Let $f(x) = e^x,\ x \in \mathbb{R}$. Show that

$$f^*(y) = \sup_{x\in\mathbb{R}}\{xy - e^x\} = \begin{cases} \infty & \text{if } y < 0, \\ 0 & \text{if } y = 0, \\ y\log y - y & \text{if } y > 0. \end{cases}$$

Deduce the following estimate

$$xy \leq e^x + y\log y - y \qquad \forall x, y > 0. \tag{F.4}$$

Appendix G
Vitali's Covering Theorem

In this appendix, we prove a fundamental covering lemma due to Vitali. We refer the reader to [EG92] for generalizations and related results.

Definition G.1 A collection $\mathscr{F}$ of closed balls[7] in $\mathbb{R}^N$ is called a *fine cover* of a set $E \subset \mathbb{R}^N$ if

$$E \subset \bigcup_{B \in \mathscr{F}} B$$

and, for every $x \in E$,

$$\inf \Big\{ \operatorname{diam}(B) \mid B \in \mathscr{F} , \; x \in B \Big\} = 0, \tag{G.1}$$

where $\operatorname{diam}(B)$ denotes the diameter of the ball B.

Lemma G.2 (Vitali) *Let $E \subset \mathbb{R}^N$ be a Borel set such that*[8] $m(E) < \infty$ *and let $\mathscr{F}$ be a fine cover of E. Then for any $\varepsilon > 0$ there exists a finite collection of disjoint balls*[9] $B_1, \ldots, B_{n_\varepsilon} \in \mathscr{F}$ *such that*

$$m\Big(E \setminus \bigcup_{i=1}^{n_\varepsilon} B_n \Big) < \varepsilon. \tag{G.2}$$

Proof To begin with, observe that, without loss of generality, we can assume that all the balls of $\mathscr{F}$ are included in some open set V, containing E, such that $m(V) < \infty$

[7] A closed ball in $\mathbb{R}^N$ is a set of type $\{x \in \mathbb{R}^N \mid \|x - x_0\| \le r\}$ with $x_0 \in \mathbb{R}^N$ and $r > 0$.

[8] m denotes the Lebesgue measure on $\mathbb{R}^N$.

[9] Observe that the whole collection of balls depends on ε, not just their total number.

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

(such an open set exists by Proposition 1.69). Indeed, it suffices to replace $\mathscr{F}$ with the subfamily

$$\tilde{\mathscr{F}} := \{B \in \mathscr{F} \mid B \subset V\}, \tag{G.3}$$

which, owing to (G.1), is again a fine cover of E.

Consequently, it is not restrictive to assume that

$$\rho := \sup\big\{\operatorname{diam}(B) \mid B \in \mathscr{F}\big\} < \infty.$$

Given $\varepsilon > 0$, we now proceed to construct $B_1, B_2, \ldots, B_{n_\varepsilon}$ by an inductive method. Let B_1 be such that $\operatorname{diam}(B_1) > \rho/2$. Next, let $n \geq 1$ and suppose $B_1, \ldots, B_n$ are disjoint balls of $\mathscr{F}$ satisfying the following for $n > 1$: for every $i = 1, \ldots, n-1$,

$$0 < \frac{d_i}{2} < \operatorname{diam}(B_{i+1}) \leq d_i\,, \tag{G.4}$$

where

$$d_i = \sup\big\{\operatorname{diam}(B) \mid B \in \mathscr{F},\ \ B \cap B_j = \emptyset \ \ \forall j = 1, \ldots, i\big\}. \tag{G.5}$$

Then there are two possibilities: either

(a) $E \subset \cup_{i=1}^n B_i$, or
(b) there exists $\bar{x} \in E \setminus \cup_{i=1}^n B_i$.

In case (a), the conclusion (G.2) follows taking $n_\varepsilon = n$. Let us consider case (b) and denote by δ the (positive) distance of $\bar{x}$ from $\cup_{i=1}^n B_i$. Since $\mathscr{F}$ is a fine cover of E, there exists a ball $B \in \mathscr{F}$ such that $\bar{x} \in B$ and $\operatorname{diam}(B) < \frac{\delta}{2}$. Consequently, B is disjoint from $B_1, \ldots, B_n$ and there exists $B_{n+1} \in \mathscr{F}$ such that B_{n+1} is disjoint from $B_1, \ldots, B_n$ and $\operatorname{diam}(B_{n+1}) > d_n/2 > 0$. If the above process does not terminate, we get a sequence $B_1, B_2, \ldots, B_n, \ldots$ of disjoint balls in $\mathscr{F}$ such that

$$\frac{d_n}{2} < \operatorname{diam}(B_{n+1}) \leq d_n \qquad \forall n \geqslant 1.$$

Since $\cup_{n=1}^\infty B_n \subset V$, we have that $\sum_{n=1}^\infty m(B_n) \leq m(V) < \infty$. Then there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\sum_{n=n_\varepsilon+1}^\infty m(B_n) < \frac{\varepsilon}{5^N}.$$

We claim that

$$E \setminus \bigcup_{n=1}^{n_\varepsilon} B_n \subset \bigcup_{n=n_\varepsilon+1}^\infty B_n^*, \tag{G.6}$$

where B_n^* denotes the ball having the same center as B_n, but with radius five times as large. Indeed, let $x \in E \setminus \cup_{n=1}^{n_\varepsilon} B_n$. By reasoning as in case (b), one realizes that there exists a ball $B \in \mathscr{F}$ such that $x \in B$ and B is disjoint from $B_1, \ldots, B_{n_\varepsilon}$. Then B must intersect at least one of the B_n's with $n > n_\varepsilon$. Otherwise, for every $n > n_\varepsilon$, we would deduce from (G.4) and (G.5) that

$$\operatorname{diam}(B) \le d_n \le 2 \operatorname{diam}(B_{n+1}) \tag{G.7}$$

in contrast with the fact that $\sum_{n=1}^{\infty} m(B_n) < \infty$. Let j be the first index such that $B \cap B_j \neq \emptyset$. Then $j > n_\varepsilon$ and

$$\operatorname{diam}(B) \le d_{j-1} < 2 \operatorname{diam}(B_j).$$

Hence, B is contained in the ball which has the same center as B_j and five times the diameter of B_j, i.e., $B \subset B_j^*$. Then (G.6) holds true and so

$$m\Big(E \setminus \bigcup_{n=1}^{n_\varepsilon} B_n\Big) \le \sum_{n=n_\varepsilon+1}^{\infty} m(B_n^*) = 5^N \sum_{n=n_\varepsilon+1}^{\infty} m(B_n) \le \varepsilon\,,$$

which completes the proof. □

Appendix H
Ekeland's Variational Principle

The following result, which is surprising for its generality, has become a basic tool in analysis. It arises in different applications and has been generalized to various situations (see [AE84]).

Theorem H.1 (Ekeland) *Let* (X, d) *be a complete metric space and let*

$$f : X \to \mathbb{R} \cup \{\infty\}$$

be a lower semicontinuous function satisfying

$$\inf_X f > -\infty\,.$$

Let $x_0 \in X$ *be such that* $f(x_0) < \infty$ *and let* $\alpha > 0$. *Then there exists* $\bar{x} \in X$ *such that*

$$\begin{cases} (a) \quad f(\bar{x}) + \alpha d(\bar{x}, x_0) \le f(x_0), \\ (b) \quad f(\bar{x}) < f(x) + \alpha d(x, \bar{x}) \quad \forall x \in X \setminus \{\bar{x}\}. \end{cases}$$

Proof Given $\alpha > 0$, set

$$F(x) = \{\, y \in X \mid f(y) + \alpha d(x, y) \le f(x)\,\} \qquad x \in X.$$

Observe that, clearly, every $x \in X$ belongs to $F(x)$ and, since f is lower semicontinuous, $F(x)$ is closed. We are going to prove the thesis by showing that

$$\exists\, \bar{x} \in F(x_0) \;:\; F(\bar{x}) = \{\bar{x}\}. \tag{H.1}$$

1. Let us prove that, for every $x, y \in X$,

$$y \in F(x) \quad \Longrightarrow \quad F(y) \subset F(x). \tag{H.2}$$

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0

Let $y \in F(x)$ and $z \in F(y)$. Then

$$f(z) + \alpha d(x, z) \leq f(z) + \alpha d(y, z) + \alpha d(x, y) \leq f(y) + \alpha d(x, y) \leq f(x),$$

which in turn yields (H.2).

2. Starting from the given point $x_0 \in X$, let us construct the sequence $(x_n)_n \subset X$ as follows. Given $x_n \in X$ for any $n \in \mathbb{N}$, set

$$\lambda_n = \inf_{F(x_n)} f,$$

and let $x_{n+1} \in F(x_n)$ be such that $f(x_{n+1}) \leq \lambda_n + 2^{-n}$. Since $\lambda_0 \leq f(x_0) < \infty$, by induction it is easy to verify that $\lambda_n < \infty$ and, consequently, $f(x_n) < \infty$ for every $n \in \mathbb{N}$. Moreover, observe that, in view of (H.2), $(F(x_n))_n$ is decreasing. So $(\lambda_n)_n$ is an increasing sequence and we have, by construction,

$$f(x_{n+1}) \geq \lambda_n \geq \lambda_{n-1} \qquad \forall n \geq 1. \tag{H.3}$$

3. Let us show that $(x_n)_n$ is a Cauchy sequence: thanks to (H.3) we get, for every $n, p \geq 1$,

$$d(x_n, x_{n+p}) \leq \sum_{i=n}^{n+p-1} d(x_i, x_{i+1}) \leq \frac{1}{\alpha} \sum_{i=n}^{n+p-1} \big[f(x_i) - f(x_{i+1})\big]$$
$$\leq \frac{1}{\alpha} \sum_{i=n}^{n+p-1} \big[f(x_i) - \lambda_{i-1}\big] \leq \frac{1}{\alpha} \sum_{i=n}^{n+p-1} 2^{1-i} \overset{(n\to\infty)}{\longrightarrow} 0.$$

Therefore, since X is complete, $(x_n)_n$ is convergent. Setting $\bar{x} = \lim_n x_n$, we obtain that $\bar{x} \in F(x_0)$ by the fact that $x_n \in F(x_0)$, which is closed.

4. In order to complete the proof of (H.1), there remains to show that $F(\bar{x}) = \{\bar{x}\}$. Since $\bar{x} \in F(\bar{x})$, it will be sufficient to check that[10] diam $F(\bar{x}) = 0$. To this aim we observe that, by construction, $\bar{x} \in F(x_n)$ for every $n \in \mathbb{N}$. So, according to (H.2), we also have $F(\bar{x}) \subset F(x_n)$, by which

$$\operatorname{diam} F(\bar{x}) \leq \operatorname{diam} F(x_n) \qquad \forall n \in \mathbb{N}.$$

Moreover, for every $n \geq 1$ we deduce

$$\alpha d(x, x_n) \leq f(x_n) - f(x) < 2^{1-n} \qquad \forall x \in F(x_n).$$

Since $f(x) \geq \lambda_{n-1}$. It follows that diam $F(x_n) \leq \frac{2^{2-n}}{\alpha} \to 0$ as $n \to \infty$.

The proof of (H.1) is thus complete. □

[10] We recall that the diameter of a nonempty subset $S \subset X$ is defined by diam $S = \sup_{x,y \in S} d(x, y)$.

References

[AE84] Aubin, J.-P., Ekeland, I.: Applied Nonlinear Analysis. Wiley, New York (1984)
[EG92] Evans, L.C., Gariepy, R.F.: Measure Theory and Fine Properties of Functions. Studies in Advanced Mathematics. CRC Press, Ann Arbor (1992)
[Ro70] Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)

Index

P. Cannarsa and T. D'Aprile, *Introduction to Measure Theory and Functional Analysis*, UNITEXT - La Matematica per il 3+2 89,
DOI 10.1007/978-3-319-17019-0