Graduate Texts in Mathematics

Francis Hirsch
Gilles Lacombe

Elements of Functional Analysis

Springer

Graduate Texts in Mathematics 192

Springer Science+Business Media, LLC

Graduate Texts in Mathematics

1 TAKEUTI/ZARING. Introduction to Axiomatic Set Theory. 2nd ed.
2 OXTOBY. Measure and Category. 2nd ed.
3 SCHAEFER. Topological Vector Spaces.
4 HILTON/STAMMBACH. A Course in Homological Algebra. 2nd ed.
5 MAC LANE. Categories for the Working Mathematician. 2nd ed.
6 HUGHES/PIPER. Projective Planes.
7 SERRE. A Course in Arithmetic.
8 TAKEUTI/ZARING. Axiomatic Set Theory.
9 HUMPHREYS. Introduction to Lie Algebras and Representation Theory.
10 COHEN. A Course in Simple Homotopy Theory.
11 CONWAY. Functions of One Complex Variable I. 2nd ed.
12 BEALS. Advanced Mathematical Analysis.
13 ANDERSON/FULLER. Rings and Categories of Modules. 2nd ed.
14 GOLUBITSKY/GUILLEMIN. Stable Mappings and Their Singularities.
15 BERBERIAN. Lectures in Functional Analysis and Operator Theory.
16 WINTER. The Structure of Fields.
17 ROSENBLATT. Random Processes. 2nd ed.
18 HALMOS. Measure Theory.
19 HALMOS. A Hilbert Space Problem Book. 2nd ed.
20 HUSEMOLLER. Fibre Bundles. 3rd ed.
21 HUMPHREYS. Linear Algebraic Groups.
22 BARNES/MACK. An Algebraic Introduction to Mathematical Logic.
23 GREUB. Linear Algebra. 4th ed.
24 HOLMES. Geometric Functional Analysis and Its Applications.
25 HEWITT/STROMBERG. Real and Abstract Analysis.
26 MANES. Algebraic Theories.
27 KELLEY. General Topology.
28 ZARISKI/SAMUEL. Commutative Algebra. Vol.I.
29 ZARISKI/SAMUEL. Commutative Algebra. Vol.II.
30 JACOBSON. Lectures in Abstract Algebra I. Basic Concepts.
31 JACOBSON. Lectures in Abstract Algebra II. Linear Algebra.
32 JACOBSON. Lectures in Abstract Algebra III. Theory of Fields and Galois Theory.
33 HIRSCH. Differential Topology.
34 SPITZER. Principles of Random Walk. 2nd ed.
35 ALEXANDER/WERMER. Several Complex Variables and Banach Algebras. 3rd ed.
36 KELLEY/NAMIOKA et al. Linear Topological Spaces.
37 MONK. Mathematical Logic.
38 GRAUERT/FRITZSCHE. Several Complex Variables.
39 ARVESON. An Invitation to C^*-Algebras.
40 KEMENY/SNELL/KNAPP. Denumerable Markov Chains. 2nd ed.
41 APOSTOL. Modular Functions and Dirichlet Series in Number Theory. 2nd ed.
42 SERRE. Linear Representations of Finite Groups.
43 GILLMAN/JERISON. Rings of Continuous Functions.
44 KENDIG. Elementary Algebraic Geometry.
45 LOÈVE. Probability Theory I. 4th ed.
46 LOÈVE. Probability Theory II. 4th ed.
47 MOISE. Geometric Topology in Dimensions 2 and 3.
48 SACHS/WU. General Relativity for Mathematicians.
49 GRUENBERG/WEIR. Linear Geometry. 2nd ed.
50 EDWARDS. Fermat's Last Theorem.
51 KLINGENBERG. A Course in Differential Geometry.
52 HARTSHORNE. Algebraic Geometry.
53 MANIN. A Course in Mathematical Logic.
54 GRAVER/WATKINS. Combinatorics with Emphasis on the Theory of Graphs.
55 BROWN/PEARCY. Introduction to Operator Theory I: Elements of Functional Analysis.
56 MASSEY. Algebraic Topology: An Introduction.
57 CROWELL/FOX. Introduction to Knot Theory.
58 KOBLITZ. p-adic Numbers, p-adic Analysis, and Zeta-Functions. 2nd ed.
59 LANG. Cyclotomic Fields.
60 ARNOLD. Mathematical Methods in Classical Mechanics. 2nd ed.
61 WHITEHEAD. Elements of Homotopy Theory.

(continued after index)

Francis Hirsch
Gilles Lacombe

Elements of Functional Analysis

Translated by Silvio Levy

Springer

Francis Hirsch
Gilles Lacombe
Département de Mathématiques
Université d'Évry-Val d'Essonne
Boulevard des coquibus
Évry Cedex F-91025
France

Translator
Silvio Levy
Mathematical Sciences Research Institute
1000 Centennial Drive
Berkeley, CA 94720-5070
USA

Mathematics Subject Classification (1991): 46-01, 46Fxx, 47E05, 46E35

Library of Congress Cataloging-in-Publication Data
Hirsch, F. (Francis)
Elements of functional analysis / Francis Hirsch, Gilles Lacombe.
p. cm. — (Graduate texts in mathematics ; 192)
Includes bibliographical references and index.
ISBN 978-1-4612-7146-8 ISBN 978-1-4612-1444-1 (eBook)
DOI 10.1007/978-1-4612-1444-1
1. Functional analysis. I. Lacombe, Gilles. II. Title.
III. Series.
QA320.H54 1999
515.7—dc21 98-53153

Printed on acid-free paper.

French Edition: *Éléments d'analyse fonctionnelle* © Masson, Paris, 1997.

Production managed by A. Orrantia; manufacturing supervised by Jacqui Ashri.
Photocomposed copy prepared from the translator's PostScript files.

9 8 7 6 5 4 3 2 1

ISBN 978-1-4612-7146-8 SPIN 10675899

Preface

This book arose from a course taught for several years at the University of Évry–Val d'Essonne. It is meant primarily for graduate students in mathematics. To make it into a useful tool, appropriate to their knowledge level, prerequisites have been reduced to a minimum: essentially, basic concepts of topology of metric spaces and in particular of normed spaces (convergence of sequences, continuity, compactness, completeness), of "abstract" integration theory with respect to a measure (especially Lebesgue measure), and of differential calculus in several variables.

The book may also help more advanced students and researchers perfect their knowledge of certain topics. The index and the relative independence of the chapters should make this type of usage easy.

The important role played by exercises is one of the distinguishing features of this work. The exercises are very numerous and written in detail, with hints that should allow the reader to overcome any difficulty. Answers that do not appear in the statements are collected at the end of the volume.

There are also many simple application exercises to test the reader's understanding of the text, and exercises containing examples and counterexamples, applications of the main results from the text, or digressions to introduce new concepts and present important applications. Thus the text and the exercises are intimately connected and complement each other.

Functional analysis is a vast domain, which we could not hope to cover exhaustively, the more so since there are already excellent treatises on the subject. Therefore we have tried to limit ourselves to results that do not require advanced topological tools: all the material covered requires no more than metric spaces and sequences. No recourse is made to topological vector spaces in general, or even to locally convex spaces or Fréchet spaces. The Baire and Banach–Steinhaus theorems are covered and used only in some exercises. In particular, we have not included the "great" theorems of functional analysis, such as the Open Mapping Theorem, the Closed Graph Theorem, or the Hahn–Banach theorem. Similarly, Fourier transforms are dealt with only superficially, in exercises. Our guiding idea has been to limit the text proper to those results for which we could state significant applications within reasonable limits.

This work is divided into a prologue and three parts.

The prologue gathers together fundamentals results about the use of sequences and, more generally, of countability in analysis. It dwells on the notion of separability and on the diagonal procedure for the extraction of subsequences.

Part I is devoted to the description and main properties of fundamental function spaces and their duals. It covers successively spaces of continuous functions, functional integration theory (Daniell integration) and Radon measures, Hilbert spaces and L^p spaces.

Part II covers the theory of operators. We dwell particularly on spectral properties and on the theory of compact operators. Operators not everywhere defined are not discussed.

Finally, Part III is an introduction to the theory of distributions (not including Fourier transformation of distributions, which is nonetheless an important topic). Differentiation and convolution of distributions are studied in a fair amount of detail. We introduce explicitly the notion of a fundamental solution of a differential operator, and give the classical examples and their consequences. In particular, several regularity results, notably those concerning the Sobolev spaces $W^{1,p}(\mathbb{R}^d)$, are stated and proved. Finally, in the last chapter, we study the Laplace operator on a bounded subset of $\mathbb{R}^d$: the Dirichlet problem, spectra, etc. Numerous results from the preceding chapters are used in Part III, showing their usefulness.

Prerequisites. We summarize here the main post-calculus concepts and results whose knowledge is assumed in this work.

- *Topology of metric spaces:* elementary notions: convergence of sequences, lim sup and lim inf, continuity, compactness (in particular the Borel–Lebesgue defining property and the Bolzano–Weierstrass property), and completeness.
- *Banach spaces:* finite-dimensional normed spaces, absolute convergence of series, the extension theorem for continuous linear maps with values in a Banach space.
- *Measure theory:* measure spaces, construction of the integral, the Monotone Convergence and Dominated Convergence Theorems, the definition and elementary properties of L^p spaces (particularly the Hölder and Minkowski inequalities, completeness of L^p, the fact that convergence

of a sequence in L^p implies the convergence of a subsequence almost everywhere), Fubini's Theorem, the Lebesgue integral.

- *Differential calculus:* the derivative of a function with values in a Banach space, the Mean Value Theorem.

These results can be found in the following references, among others: For the topology and normed spaces, Chapters 3 and 5 of J. Dieudonné's *Foundations of Modern Analysis* (Academic Press, 1960); for the integration theory, Chapters 1, 2, 3, and 7 of W. Rudin's *Real and Complex Analysis*, McGraw-Hill; for the differential calculus, Chapters 2 and 3 of H. Cartan's *Cours de calcul différentiel* (translated as *Differential Calculus*, Hermann).

We are thankful to Silvio Levy for his translation and for the opportunity to correct here certain errors present in the French original.

We thankfully welcome remarks and suggestions from readers. Please send them by email to hirsch@lami.univ-evry.fr *or* lacombe@lami.univ-evry.fr.

Francis Hirsch
Gilles Lacombe

Contents

Notation

If A is a subset of X, we denote by A^c the complement of A in X. If $A \subset X$ and $B \subset X$, we set $A \setminus B = A \cap B^c$. The characteristic function of a subset A of X is denoted by 1_A. It is defined by

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

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ represent the nonnegative integers, the integers, the rationals, and the reals. If $\mathbb{E}$ is one of these sets, we write $\mathbb{E}^* = \mathbb{E} \setminus \{0\}$. We also write $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$. If $a \in \mathbb{R}$ we write $a^+ = \max(0, a)$ and $a^- = -\min(a, 0)$.

$\mathbb{C}$ denotes the complex numbers. As usual, if $z \in \mathbb{C}$, we denote by $\bar{z}$ the complex conjugate of z, and by $\operatorname{Re} z$ and $\operatorname{Im} z$ the real and imaginary parts of z.

If f is a function from a set X into $\mathbb{R}$ and if $a \in \mathbb{R}$, we write $\{f > a\} = \{x \in X : f(x) > a\}$. We define similarly the sets $\{f < a\}$, $\{f \geq a\}$, $\{f \leq a\}$, etc.

As usual, a number $x \in \mathbb{R}$ is *positive* if $x > 0$, and *negative* if $x < 0$. However, for the sake of brevity in certain statements, we adopt the convention that a real-valued function f is *positive* if it takes only nonnegative values (including zero), and we denote this fact by $f \geq 0$.

Let (X, d) be a metric space. If A is a subset of X, we denote by $\bar{A}$ and $\mathring{A}$ the closure and interior of A. If $x \in X$, we write $\mathscr{V}(x)$ for the set of neighborhoods of x (that is, subsets of X whose interior contains x). We set

$$B(x,r) = \{y \in X : d(x,y) < r\}, \quad \bar{B}(x,r) = \{y \in X : d(x,y) \leq r\}.$$

(We do not necessarily have $\bar{B}(x,r) = \overline{B(x,r)}$, but this equality does hold if, for example, X is a normed space with the associated metric.) If X is a normed vector space with norm $\|\cdot\|$, the closed unit ball of X is

$$\bar{B}(X) = \{x \in X : \|x\| \leq 1\}.$$

When no ambiguity is possible, we write $\bar{B}$ instead of $\bar{B}(X)$. If A is a subset of X, the diameter of A is

$$d(A) = \sup_{x,y \in A} d(x,y).$$

If $A \subset X$ and $B \subset X$, the distance between A and B is

$$d(A,B) = \inf_{(x,y) \in A \times B} d(x,y),$$

and $d(x,A) = d(\{x\},A)$ for $x \in X$.

We set $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. All vector spaces are over one or the other $\mathbb{K}$. If E is a vector space and A is a subset of E, we denote by $[A]$ the vector subspace generated by A. If E is a vector space, A, B are subsets of E, and $\lambda \in \mathbb{K}$, we write $A + B = \{x + y : x \in A,\ y \in B\}$ and $\lambda A = \{\lambda x : x \in A\}$.

Lebesgue measure over $\mathbb{R}^d$, considered as a measure on the Borel sets of $\mathbb{R}^d$, is denoted by λ_d. We also use the notations $d\lambda_d(x) = dx = dx_1 \ldots dx_d$. We omit the dimension subscript d if there is no danger of confusion.

If $x \in \mathbb{R}^d$, the euclidean norm of x is denoted by $|x|$.

Prologue: Sequences

Sequences play a key role in analysis. In this preliminary chapter we collect various relevant results about sequences.

1 Countability

This first section approaches sequences from a set-theoretical viewpoint.

A set X is **countably infinite** if there is a bijection φ from $\mathbb{N}$ onto X; that is, if we can order X as a sequence:

$$X = \{\varphi(0), \varphi(1), \ldots, \varphi(n), \ldots\},$$

where $\varphi(n) \neq \varphi(p)$ if $n \neq p$. The bijection φ can also be denoted by means of subscripts: $\varphi(n) = x_n$. In this case

$$X = \{x_0, x_1, \ldots, x_n, \ldots\} = \{x_n\}_{n \in \mathbb{N}}.$$

A set is **countable** if it is finite or countably infinite.

Examples

1. $\mathbb{N}$ is clearly countably infinite. So is $\mathbb{Z}$: we can write $\mathbb{Z}$ as the sequence

$$\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots, n, -n, \ldots\}.$$

Clearly, there can be no order-preserving bijection between $\mathbb{N}$ and $\mathbb{Z}$.

2. The set $\mathbb{N}^2$ is countable. For we can establish a bijection $\varphi : \mathbb{N} \to \mathbb{N}^2$ by setting, for every $p \geq 0$ and every $n \in \big[p(p+1)/2,\ (p+1)(p+2)/2\big)$,

$$\varphi(n) = \left(n - \frac{p(p+1)}{2}, \frac{p(p+3)}{2} - n\right).$$

This complicated expression means simply that we are enumerating $\mathbb{N}^2$ by listing consecutively the finite sets $A_p = \{(q,r) \in \mathbb{N}^2 : q+r=p\}$, each in increasing order of the first coordinate:

$$\mathbb{N}^2 = \{\, \overbrace{(0,0)},\ \overbrace{(0,1),\ (1,0)},\ \overbrace{(0,2),\ (1,1),\ (2,0)},\ (0,3),\ (1,2),\ \dots\}.$$

We see that explicitly writing down a bijection between $\mathbb{N}$ and a countable set X is often not at all illuminating. Fortunately, it is usually unnecessary as well, if the goal is to prove the countability of X. One generally uses instead results such as the ones we are about to state.

Proposition 1.1 *A nonempty set X is countable if and only if there is a surjection from $\mathbb{N}$ onto X.*

Proof. If X is countably infinite there is a bijection, and thus a surjection, from $\mathbb{N}$ to X. If X is finite with $n \geq 1$ elements, there is a bijection $\varphi : \{1, \dots, n\} \to X$. This can be arbitrarily extended to a bijection from $\mathbb{N}$ to X.

Conversely, suppose there is a surjection $\varphi : \mathbb{N} \to X$ and that X is infinite. Define recursively a sequence $(n_p)_p \in \mathbb{N}$ by setting $n_0 = 0$ and

$$n_{p+1} = \min\bigl\{n : \varphi(n) \notin \{\varphi(n_0), \varphi(n_1), \dots, \varphi(n_p)\}\bigr\} \quad \text{for } p \in \mathbb{N}.$$

This sequence is well-defined because X is infinite; by construction, the map $p \mapsto \varphi(n_p)$ is a bijection from $\mathbb{N}$ to X. □

Corollary 1.2 *If X is countable and there exists a surjection from X to Y, then Y is countable.*

Indeed, the composition of two surjections is surjective.

Corollary 1.3 *Every subset of a countable set is countable.*

Indeed, if $Y \subset X$, it is clear that there is a surjection from X to Y.

Corollary 1.4 *If Y is countable and there exists an injection from X to Y, then X is countable.*

Proof. An injection $f : X \to Y$ defines a bijection from X to $f(X)$. If Y is countable, so is $f(X)$, by the preceding corollary. Therefore X is countable. □

Corollary 1.5 *A set X is countable if and only if there is an injection from X to $\mathbb{N}$.*

Another important result about the preservation of countability is this:

Proposition 1.6 *If the sets $X_1, X_2, \dots, X_n$ are countable, the Cartesian product $X = X_1 \times X_2 \times \cdots \times X_n$ is countable.*

Proof. It is enough to prove the result for $n = 2$ and use induction. Suppose that X_1 and X_2 are countable, and let f_1, f_2 be surjections from $\mathbb{N}$ to X_1, X_2 (whose existence is given by Proposition 1.1). The map $(n_1, n_2) \mapsto \big(f_1(n_1), f_2(n_2)\big)$ is then a surjection from $\mathbb{N}^2$ to X. Since $\mathbb{N}^2$ is countable, the proposition follows by Corollary 1.2. □

We conclude with a result about countable unions of countable sets:

Proposition 1.7 *Let $(X_i)_{i\in I}$ be a family of countable sets, indexed by a countable set I. The set $X = \bigcup_{i\in I} X_i$ is countable.*

Proof. If, for each $i \in I$, we take a surjection $f_i : \mathbb{N} \to X_i$, the map $f : I \times \mathbb{N} \to X$ defined by $f(i,n) = f_i(n)$ is a surjection. But $I \times \mathbb{N}$ is countable. □

Note that a countable *product* of countable sets is not necessarily countable; see Example 5 below.

Examples and counterexamples

1. $\mathbb{Q}$ is countable. Indeed, the map $f : \mathbb{Z} \times \mathbb{N}^* \to \mathbb{Q}$ defined by $f(n,p) = n/p$ is surjective and $\mathbb{Z} \times \mathbb{N}^*$ is countable.
2. The sets $\mathbb{N}^n$, $\mathbb{Q}^n$, $\mathbb{Z}^n$, and $(\mathbb{Q} + i\mathbb{Q})^n$ are countable (see Proposition 1.6).
3. $\mathbb{R}$ is not countable. For assume it were; then so would be the subset $[0,1]$, that is, we would have $[0,1] = \{x_n\}_{n\in\mathbb{N}}$. We could then construct a sequence of subintervals $I_n = [a_n, b_n]$ of $[0,1]$ satisfying these properties, for all $n \in \mathbb{N}$:

$$I_{n+1} \subset I_n, \quad x_n \notin I_n, \quad d(I_n) = 3^{-n-1}.$$

   The construction is a simple recursive one: for $n = 0$ we choose I_0 as one of the intervals $\left[0, \frac{1}{3}\right]$, $\left[\frac{2}{3}, 1\right]$, subject to the condition $x_0 \notin I_0$; likewise, if $I_n = [a_n, b_n]$ has been constructed, we choose I_{n+1} as one of the intervals $[a_n,\, a_n + 3^{-n-1}]$, $[b_n - 3^{-n-1},\, b_n]$, not containing x_{n+1}. By construction, $\bigcap_{n\in\mathbb{N}} I_n = \{x\}$, where x is the common limit of the increasing sequence (a_n) and of the decreasing sequence (b_n). Clearly, $x \in [0,1]$, but $x \neq x_n$ for all $n \in \mathbb{N}$, which contradicts the assumption that $[0,1] = \{x_n\}_{n\in N}$.
   More generally, any complete space without an isolated point is uncountable; see, for example, Exercise 6 on page 16.
   Note also that if $\mathbb{R}$ were countable it would have Lebesgue measure zero, which is not the case.
4. The set $\mathscr{P}(\mathbb{N})$ of subsets of $\mathbb{N}$ is uncountable. Indeed, suppose there is a bijection $\varphi : \mathbb{N} \to \mathscr{P}(\mathbb{N})$, and set

$$A = \big\{n \in \mathbb{N} : n \notin \varphi(n)\big\} \in \mathscr{P}(\mathbb{N}).$$

Since φ is a surjection, A has at least one inverse image a under φ. We now see that a cannot be an element of A, since by the definition of A this would imply $a \notin \varphi(a) = A$, nor can it be an element of $\mathbb{N} \setminus A$, since this would imply $a \in \varphi(a)$ and hence $a \in A$. This contradiction proves the desired result.

This same reasoning can be used to prove that, if X is any set, there can be no surjection from X to $\mathscr{P}(X)$. This is called **Cantor's Theorem**.

5. The set $\mathscr{C} = \{0,1\}^{\mathbb{N}}$ of functions $\mathbb{N} \to \{0,1\}$ (sequences with values in $\{0,1\}$) is uncountable. Indeed, the map from $\mathscr{P}(\mathbb{N})$ into $\mathscr{C}$ that associates to each subset A of $\mathbb{N}$ the characteristic function 1_A is clearly bijective; its inverse is the map that associates to each function $\varphi : \mathbb{N} \to \{0,1\}$ the subset A of $\mathbb{N}$ defined by $A = \{n \in \mathbb{N} : \varphi(n) = 1\}$.
We remark that $\mathscr{C}$, and thus also $\mathscr{P}(\mathbb{N})$, is in bijection with $\mathbb{R}$ (see Exercise 3 on the next page).
6. The set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers is uncountable; otherwise $\mathbb{R}$ would be countable.
7. The set $\mathscr{P}_f(\mathbb{N})$ of finite subsets of $\mathbb{N}$ is countable; indeed, we can define a surjection f from $\{0\} \cup \bigcup_{p \in \mathbb{N}^*} \mathbb{N}^p$ (which is countable by Proposition 1.7) onto $\mathscr{P}_f(\mathbb{N})$, by setting
$$f(0) = \varnothing \qquad \text{and} \qquad f(n_1, \ldots, n_p) = \{n_1, \ldots, n_p\} \quad \text{for all } p \in \mathbb{N}^*.$$
8. The set $\mathbb{Q}[X]$ of polynomials in one indeterminate over $\mathbb{Q}$ is countable, because there is a surjective map from $\bigcup_{p \in \mathbb{N}^*} \mathbb{Q}^p$ (which is countable by Proposition 1.7) onto $\mathbb{Q}[X]$, defined by
$$f(q_1, \ldots, q_p) = q_1 + q_2 X + \cdots + q_p X^{p-1}.$$
We can show in an analogous way that the set $\mathbb{Q}[X_1, \ldots, X_n]$ of polynomials in n indeterminates over $\mathbb{Q}$ is countable.
9. If $\mathscr{A}$ is a family of nonempty, pairwise disjoint, open intervals in $\mathbb{R}$, then $\mathscr{A}$ is countable. Indeed, let φ be a bijection from $\mathbb{N}$ onto $\mathbb{Q}$. For $J \in \mathscr{A}$, let $n(J)$ be the first integer n for which $\varphi(n) \in J$. The map $\mathscr{A} \to \mathbb{N}$ that associates $n(J)$ to J is clearly injective, so $\mathscr{A}$ is countable by Corollary 1.5.

Exercises

1. Which, if any, of the following sets are countable?
   a. The set of sequences of integers.
   b. The set of sequences of integers that are zero after a certain point.
   c. The set of sequences of integers that are constant after a certain point.
2. Let A be an infinite set and B a countable set. Prove that there is a bijection between A and $A \cup B$.

3. Let $\mathscr{C} = \{0,1\}^{\mathbf{N}}$.

a. Let $f : \mathscr{C} \to [0,2]$ be the function defined by

$$f(x) = \sum_{n=0}^{+\infty} \frac{x_n}{2^n}.$$

Prove that f is surjective and that every element of $[0,2]$ has at most two inverse images under f. Find the set D of elements of $[0,2]$ that have two inverse images under f; prove that D and $f^{-1}(D)$ are countably infinite.

b. Construct a bijection between $\mathscr{C}$ and $[0,2]$, then a bijection between $\mathscr{C}$ and $\mathbb{R}$.

4. Let X be a connected metric space that contains at least two points. Prove that there exists an injection from $[0,1]$ into X. Deduce that X is not countable.

Hint. Let x and y be distinct points of X. Prove, that, for every $r \in \big[0,\, d(x,y)\big]$, the set

$$S_r = \{t \in X : d(x,t) = r\}$$

is nonempty.

5. Let A be a subset of $\mathbb{R}$ such that, for every $x \in A$, there exists $\eta > 0$ with $(x,\, x+\eta) \cap A = \varnothing$. Prove that A is countable.

Hint. Let x and y be distinct points of A. Prove that, given $\eta, \varepsilon > 0$, if the intervals $(x,\, x+\eta)$ and $(y,\, y+\varepsilon)$ do not intersect A, they do not intersect one another.

6. Let f be an increasing function from I to $\mathbb{R}$, where I is an open, nonempty interval of $\mathbb{R}$. Let S be the set of discontinuity points of f. If $x \in I$, denote by $f(x_+)$ and $f(x_-)$ the right and left limits of f at x (they exist since f is monotone).

a. Prove that $S = \{x \in I : f(x_-) < f(x_+)\}$.

b. For $x \in S$, write $I_x = \big(f(x_-), f(x_+)\big)$. By considering the family $(I_x)_{x \in S}$, prove that S is countable.

c. Conversely, let $S = \{x_n\}_{n \in \mathbf{N}}$ be a countable subset of I. Prove that there exists an increasing function whose set of points of discontinuity is exactly S.

Hint. Put $f(x) = \sum_{n=0}^{+\infty} 2^{-n} 1_{[x_n,+\infty)}(x)$.

7. More generally, a function on a nonempty, open interval I of $\mathbb{R}$ and taking values in a normed space is said to be *regulated* if it has a left and a right limit at each point of I. Let I be a regulated function from I to $\mathbb{R}$.

a. Let J be a compact interval contained in I. For $\varepsilon > 0$, write

$$J_\varepsilon = \big\{x \in J : \max\big(|f(x_+) - f(x)|,\, |f(x) - f(x_-)|\big) > \varepsilon\big\}.$$

Prove that J_ε has no cluster point.

Hint. Prove that at a cluster point of J_ε the function f cannot have both a right and a left limit.

b. Deduce that J_ε is finite.

c. Deduce that the number of points $x \in I$ where the function f is discontinuous is countable.

8. Let A and B be countable dense subsets of $(0, 1)$. We want to construct a strictly increasing bijection from A onto B.

a. Suppose first that A is the set

$$A = \{p2^{-q} : p, q \in \mathbb{N}^*, p < 2^q\}.$$

i. Prove that A is countable and that, if x is an element of A, there exists a unique pair (p, q) of integers such that $x = p2^{-q}$, with $q \in \mathbb{N}^*$ and $p < 2^q$ odd.

ii. Write $B = \{x_n : n \in \mathbb{N}\}$ and define the map $f : A \to B$ inductively, as follows:

- For $q = 1$, set $f\left(\frac{1}{2}\right) = x_0$.
- Suppose the values $f(p2^{-k})$ have been chosen for $1 \le k \le q$ and $1 \le p < 2^q$. We then define $f(p2^{-q-1})$, for $p < 2^{q+1}$ odd, by setting $f(p2^{-q-1}) = x_n$, where

$$n = \min\left\{m \in \mathbb{N} : f\left(\frac{p-1}{2^{q+1}}\right) < x_m < f\left(\frac{p+1}{2^{q+1}}\right)\right\}$$

(by convention, we have set $f(0) = 0$ and $f(1) = 1$).

Prove that $f(x)$ is well-defined for all $x \in A$; then prove that f is a strictly increasing bijection from A onto B.

iii. Deduce from this the case of arbitrary A.

9. *A bit of set theory*

a. Let I be an infinite set. The goal of this exercise is to prove, using the axiom of choice, that there exists a bijection from I to $I \times \mathbb{N}$.

Recall that a total order relation $\le$ on a set I is called a *well-ordering* if every nonempty subset of I has a least element for the order $\le$. Recall also that every set can be well-ordered; this assertion, called *Zermelo's axiom*, is equivalent to the axiom of choice. Let $\le$ be a well-ordering on I. The least element of I is denoted by 0. If $x \in I$, denote by $x + 1$ the *successor of* x, that is, the element of I defined by

$$x + 1 = \min\{y \in I : y > x\}.$$

Thus, every element of I, except possibly one, has a successor. A nonzero element of I that is not the successor of an element of I is called a *limit element.* If x is an element of I, we define (if possible) an element $x + n$, for integer n, by inductively setting $x + (n + 1) = (x + n) + 1$.

i. An example: suppose in this setting that $I = \mathbb{N}^2$ and that $\leq$ is the *lexicographical order* on $\mathbb{N}^2$:

$$(n,m) \leq (n',m') \iff (n < n') \text{ or } (n = n' \text{ and } m \leq m').$$

Check that this is a well-ordering. If $(n,m) \in I$, determine $(n,m)+1$. What are the limit elements of I?

ii. Let $x \in I$. Prove that x can be written in a unique way as $x = x' + n$, where $n \in \mathbb{N}$ and x' is 0 or a limit element.

iii. Let φ be a bijection from $\mathbb{N} \times \mathbb{N}$ onto $\mathbb{N}$. Define a map F from $I \times \mathbb{N}$ to I by $F(x,m) = x' + \varphi(n,m)$, where $x = x' + n$ is the decomposition given in the preceding item. Prove that F is a bijection.

b. Let X be a set and A a subset of X. Suppose there exists an injection $i : X \to A$. We wish to show that there is a bijection between X and A.

i. A subset Z of X is said to be *closed* (with respect to i) if $i(Z) \subset Z$. If Z is any subset of X, the *closure* $\bar{Z}$ of Z is the smallest closed subset of X containing Z. Prove that $\bar{Z}$ is well-defined for every $Z \subset X$.

ii. Set $Z = X \setminus A$. Let $\psi : X \to X$ be the map defined by

$$\psi(x) = \begin{cases} i(x) & \text{if } x \in \bar{Z}, \\ x & \text{if } x \in X \setminus \bar{Z}. \end{cases}$$

Prove that ψ is a bijection from X onto A.

c. *Cantor–Bernstein Theorem.* Let X and Y be sets. Suppose there is an injection $f : X \to Y$ and an injection $g : Y \to X$. Prove that there is a bijection between X and Y. (Note that this result does not require the axiom of choice.)
Hint. $f \circ g$ is an injection from Y to $f(X)$, and the latter is a subset of Y.

d. Let X and Y be sets. Suppose there is a surjection $f : X \to Y$ and a surjection $g : Y \to X$. Prove that there is a bijection between X and Y. (You can use the preceding result. Here it is necessary to use the axiom of choice.)

e. Let I be an infinite set, let $(J_i)_{i \in I}$ be a family of pairwise disjoint and nonempty countable sets, and set $J = \bigcup_{i \in I} J_i$. Prove that there exists a bijection between I and J.

2 Separability

We consider here a type of "topological countability" property, called separability. A metric space (X,d) is called **separable** if it contains a countable dense subset; that is, if there is a sequence of points (x_n) of X such that

$$\text{for all } x \in X \text{ and } \varepsilon > 0, \text{ there is } n \in \mathbb{N} \text{ such that } d(x_n, x) < \varepsilon.$$

It is easy to check that this condition is satisfied if and only if every nonempty open subset of X contains at least one point from the sequence (x_n). Thus, the notion of separability is topological: it does not depend on the metric d except insofar as d determines the family of open sets (the **topology**) of X.

Examples

1. Every finite-dimensional normed space is separable. Recall that on a finite-dimensional vector space, all norms are equivalent, that is, they determine the same topology. This reduces the problem to that of $\mathbb{R}^n$ or $\mathbb{C}^n$. But it is clear that $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$, and that $(\mathbb{Q} + i\mathbb{Q})^n$ is dense in $\mathbb{C}^n$.

2. *Compact metric spaces*

   Proposition 2.1 *Every compact metric space is separable.*

   Proof. If n is a strictly positive integer, the union of the balls $B\left(x, \frac{1}{n}\right)$, over $x \in X$, covers X. By the Borel–Lebesgue property, X can be covered by a finite number of such balls: $X = \bigcup_{j=1}^{J_n} B\left(x_j^n, \frac{1}{n}\right)$. It is then clear that the set

   $$D = \{x_j^n : n \in \mathbb{N}^*,\ 1 \le j \le J_n\}$$

   is dense in X. □

3. *σ-compact metric spaces.* A metric space is said to be ***σ*-compact** if it is the union of a countable family of compact sets.
   For example, every finite-dimensional normed space is σ-compact. Indeed, in such a space E any bounded closed set is compact, and $E = \bigcup_{n \in \mathbb{N}} \bar{B}(0, n)$. It will turn out later, as a consequence of the theorems of Riesz (page 49) and of Baire (page 22) that infinite-dimensional Banach spaces are no longer σ-compact; nonetheless, they can be separable.

   Proposition 2.2 *Every σ-compact metric space is separable.*

   This is an immediate consequence of Propositions 2.1 and 1.7.

Proposition 2.3 *If X is a separable metric space and Y is a subset of X, then Y is separable (in the induced metric).*

Proof. Let (x_n) be a dense sequence in X. Set

$$\mathcal{U} = \{(n, p) \in \mathbb{N} \times \mathbb{N}^* : B(x_n, 1/p) \cap Y \neq \varnothing\}.$$

For each $(n,p) \in \mathscr{U}$, choose a point $x_{n,p}$ of $B(x_n, 1/p) \cap Y$. We show that the family $D = \{x_{n,p},\ (n,p) \in \mathscr{U}\}$ (which is certainly countable) is dense in Y. To do this, choose $x \in Y$ and $\varepsilon > 0$. Let p be an integer such that $1/p < \varepsilon/2$; clearly there exists an integer $n \in \mathbb{N}$ such that $d(x, x_n) < 1/p$. But then $x \in B(x_n, 1/p) \cap Y$; therefore $(n,p) \in \mathscr{U}$ and $d(x, x_{n,p}) < 2/p < \varepsilon$. □

Example. The set $\mathbb{R} \setminus \mathbb{Q}$ of irrational numbers, with the usual metric, is separable. This can be seen either by applying the preceding proposition, or by observing that the set $D = \{q\sqrt{2} : q \in \mathbb{Q}\}$ is dense in $\mathbb{R} \setminus \mathbb{Q}$.

By reasoning as in Example 9 on page 4, one demonstrates the following proposition:

Proposition 2.4 *In a separable metric space, every family of pairwise disjoint nonempty open sets is countable.*

We will now restrict ourselves to the case of normed spaces. The metric will always be the one induced by the norm.

A subset D of a normed vector space E is said to be **fundamental** if it generates a dense subspace of E, that is, if, for every $x \in E$ and every $\varepsilon > 0$ there is a finite subset $\{x_1, \ldots, x_n\}$ of D and scalars $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that

$$\left\| x - \sum_{j=1}^{n} \lambda_j x_j \right\| < \varepsilon.$$

Proposition 2.5 *A normed space is separable if and only if it contains a countable fundamental family of vectors.*

Proof. The condition is certainly necessary, since a dense family of vectors is fundamental. Conversely, let D be a countable fundamental family of vectors in a normed space E. Let $\mathscr{D}$ be the set of linear combinations of elements of D with coefficients in the field $Q = \mathbb{Q}$ (if $\mathbb{K} = \mathbb{R}$) or $\mathbb{Q} + i\mathbb{Q}$ (if $\mathbb{K} = \mathbb{C}$). Then $\mathscr{D}$ is dense in E, because its closure contains the closure of the vector space generated by D, which is E. On the other hand, $\mathscr{D}$ is countable, because it is the image of the countable set $\bigcup_{n \in \mathbb{N}^*} (Q^n \times D^n)$ under the map f defined by

$$f(\lambda_1, \ldots, \lambda_n, x_1, \ldots, x_n) = \sum_{j=1}^{n} \lambda_j x_j. \qquad \square$$

Remark. Recall that in a normed space any finite-dimensional subspace is closed, since it is complete. It follows that a family of vectors whose span is finite-dimensional (in particular, a finite family) is fundamental if and only if its span is the whole space.

A free and fundamental family of vectors in a normed space E is called a **topological basis** for E.

Proposition 2.6 *A normed space is separable if and only if it has a countable topological basis.*

Proof. The "if" part follows immediately from the preceding proposition. To prove the converse, it is enough to consider an infinite-dimensional normed space E. By the preceding proposition, E has a fundamental sequence (x_n). Now define by induction

$$n_0 = \min\{n \in \mathbb{N} : x_n \neq 0\}$$

and, for every $p \in \mathbb{N}$,

$$n_{p+1} = \min\bigl\{n \in \mathbb{N} : x_n \notin [x_{n_0}, \ldots, x_{n_p}]\bigr\}.$$

Since E is infinite-dimensional by assumption, the sequence (n_p) is well-defined (see the preceding remark). By construction, the family $(x_{n_p})_{p\in\mathbb{N}}$ is free and generates the same subspace as $(x_n)_{n\in\mathbb{N}}$. Therefore it is fundamental. □

Exercises

1. Let X be a metric space. We say that a family of open sets $(U_i)_{i\in I}$ of X is a *basis of open sets* (or *open basis*) of X if, for every nonempty open subset U of X and for every $x \in U$, there exists $i \in I$ such that $x \in U_i \subset U$.
   a. Let $\mathscr{U}$ be an open basis of X. Prove that any open set U in X is the union of the elements of $\mathscr{U}$ contained in U.
   b. Prove that X is separable if and only if it has a countable open basis.
   Hint. If (x_n) is a dense sequence in X, the family
   $$\bigl(B(x_n, 1/(p+1))\bigr)_{n,p\in\mathbb{N}}$$
   is an open basis of X. Conversely, if (U_n) is an open basis of X, any sequence (x_n) with the property that $x_n \in U_n$ for every n is dense in X.
2. Let X be a separable metric space.
   a. Prove that there is an injection from X into $\mathbb{R}$.
   Hint. Let $(V_n)_{n\in\mathbb{N}}$ be a countable basis of open sets of X (see the preceding exercise). Consider the map from X into $\mathscr{P}(\mathbb{N})$ that takes $x \in X$ to $\{n \in \mathbb{N} : x \in V_n\}$.
   b. Prove that there is an injection from the set $\mathscr{U}$ of open sets of X into $\mathbb{R}$.
   Hint. Prove the injectivity of the map $U \to \mathscr{P}(\mathbb{N})$ that associates to each open set U in X the set $\{n \in \mathbb{N} : V_n \subset U\}$.
3. Let X be a separable metric space.

a. Let $f : X \to \mathbb{R}$ be a function, and let M be the set of points of X where f has a local extremum. Prove that $f(M)$ is countable.

Hint. Let M^+ be the set of points of X where f has a local maximum and let $\mathscr{U}$ be a countable open basis of X (see Exercise 1). Prove that there is an injection from $f(M^+)$ into $\mathscr{U}$.

b. Prove that a continuous function $f : \mathbb{R} \to \mathbb{R}$ that has a local extremum at every point is constant.

4. *Lindelöf's Theorem.* Prove that a metric space X is separable if and only if every open cover of X (that is, every family of open sets whose union is X) has a countable subcover (that is, some countable subset of the cover is still a cover).

Hint. "Only if": Let (V_n) be a countable basis of open sets of X (see Exercise 1) and let $(U_i)_{i\in I}$ be an open cover of X. Take $n \in \mathbb{N}$. If V_n is contained in some U_i, choose an element $i(n)$ of I such that $V_n \subset U_{i(n)}$; otherwise, choose $i(n) \in I$ arbitrarily. Prove that the family $(U_{i(n)})_{n\in\mathbb{N}}$ covers X. For the converse, one can work as in the proof of Proposition 2.1.

5. Let X be a separable metric space and let $\mathscr{U}$ be an uncountable family of open sets in X. Prove that there exists a point of X that belongs to uncountably many elements of $\mathscr{U}$.

6. *Theorem of Cantor and Bendixon.* Let X be a separable metric space. Prove that there is a closed subset E of X, with no isolated points, and a countable subset D of X such that $X = E \cup D$ and $E \cap D = \varnothing$.

Hint. One can choose for E the set of points of X that have no countable neighborhood.

7. Let $p \geq 1$ be a real number. Denote by ℓ^p the set of complex sequences $a = (a_n)$ such that the series $\sum |a_n|^p$ converges. Give ℓ^p the norm

$$\|a\|_p = \Bigl(\sum_{n\in\mathbb{N}} |a_n|^p\Bigr)^{1/p}.$$

Also, denote by ℓ^∞ the set of bounded complex sequences, with the norm

$$\|a\|_\infty = \sup_{n\in\mathbb{N}} |a_n|.$$

Finally, denote by c_0 the subset of ℓ^∞ consisting of sequences that tend to 0.

a. Prove that ℓ^p and ℓ^∞ are Banach spaces.

b. What is the closure in ℓ^∞ of the set of almost-zero sequences (those that have only finitely many nonzero terms)?

c. What is the closure of ℓ^p in ℓ^∞?

d. Prove that c_0, with the norm $\|\cdot\|_\infty$, is a separable Banach space.

e. Prove that ℓ^p is separable.

f. Prove that ℓ^∞ is not separable.
Hint. Check that $\{0,1\}^{\mathbb{N}} \subset \ell^\infty$ and that, if α, β are distinct elements of $\{0,1\}^{\mathbb{N}}$, then $\|\alpha - \beta\|_\infty = 1$. Then use Proposition 2.4 and the fact that $\{0,1\}^{\mathbb{N}}$ is uncountable.

g. Prove that the set of convergent sequences, with the $\|\cdot\|_\infty$ norm, is a separable Banach space.

8. Let I be a set. If $f : I \to [0, +\infty)$ is a map, denote by $\sum_{i\in I} f(i)$ the supremum of the set of all finite sums of the form $\sum_{i\in J} f(i)$, where $J \subset I$ is finite.

a. Prove that, if $\sum_{i\in I} f(i) < +\infty$, the set $J = \{i \in I : f(i) \neq 0\}$ is countable.
Hint. Check that $J = \bigcup_{n>0} E_n$, where, for each positive integer n, we set $E_n = \{i \in I : f(i) > 1/n\}$.

b. Let $p \geq 1$ be a real number. Denote by $\ell^p(I)$ the vector space consisting of functions $f : I \to \mathbb{C}$ such that $\sum_{i\in I} |f(i)|^p < +\infty$. We define on $\ell^p(I)$ a map $\|\cdot\|_p$ by setting

$$\|f\|_p = \left(\sum_{i\in I} |f(i)|^p\right)^{1/p}.$$

Prove that $\|\cdot\|_p$ is a norm, for which $\ell^p(I)$ is a Banach space.

c. Prove that $\ell^p(I)$ is separable if and only if I is countable.

3 The Diagonal Procedure

In this section we introduce a method for passing to subsequences, called the diagonal procedure, and present some of its applications. Recall that a subsequence of a given sequence $(x_n)_{n\in\mathbb{N}}$ is a sequence of the form $(x_{n_k})_{k\in\mathbb{N}}$, where $(n_k)_{k\in\mathbb{N}}$ is a strictly increasing sequence of integers. Such a sequence $k \mapsto n_k$ can also be considered as a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$. The subsequence (x_{n_k}) can then be written $(x_{\varphi(k)})_{k\in\mathbb{N}}$. Since the function φ is uniquely determined by its image $A = \varphi(\mathbb{N})$ (for $n \in \mathbb{N}$, the value of $\varphi(n)$ is the $(n+1)$-st term of A in the usual order of $\mathbb{N}$), the subsequence $(x_{\varphi(k)})_{k\in\mathbb{N}}$ is determined by the infinite set A; we can denote it by $(x_n)_{n\in A}$. We will use all three notations in the sequel.

Theorem 3.1 *Let $(X_p, d_p)_{p\in\mathbb{N}}$ be a sequence of metric spaces, and, for every $p \in \mathbb{N}$, let $(x_{n,p})_{n\in\mathbb{N}}$ be a sequence in X_p. If, for every $p \in \mathbb{N}$, the set $\{x_{n,p} : n \in \mathbb{N}\}$ is relatively compact in X_p, there exists a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ such that for every $p \in \mathbb{N}$ the sequence $(x_{\varphi(n),p})_{n\in\mathbb{N}}$ converges in X_p.*

Recall that a subset Y of a metric space X is called **relatively compact** in X if there exists a compact K of X such that $Y \subset K$, or, equivalently, if the closure of Y in X is compact. In terms of sequences, Y is relatively compact if and only if every sequence in Y has a subsequence that converges in X (though the limit may not be in Y).

The remarkable part of the theorem is that the function φ that defines the different subsequences does not depend on p.

Proof. Thanks to the assumption of relative compactness, one can inductively construct a decreasing subsequence (A_n) of infinite subsets of $\mathbb{N}$ such that, for every $p \in \mathbb{N}$, the sequence $(x_{n,p})_{n\in A_p}$ converges in X_p. The **diagonal procedure** consists in defining the map φ by setting

$$\varphi(p) = \text{the } (p+1)\text{-st element of } A_p.$$

Thus $\varphi(p+1)$ is strictly greater than the $(p+1)$-st element of A_{p+1}, which in turn is greater than the $(p+1)$-st element of A_p, which is $\varphi(p)$. Thus φ is strictly increasing. Moreover, for every $p \in \mathbb{N}$ the sequence $(x_{\varphi(n),p})_{n\geq p}$ is a subsequence of the sequence $(x_{n,p})_{n\in A_p}$, because, if $n \geq p$, we have $\varphi(n) \in A_n \subset A_p$. Therefore the sequence $(x_{\varphi(n),p})_{n\in\mathbb{N}}$ converges. □

Consider again a sequence $(X_p, d_p)_{p\in\mathbb{N}}$ of metric spaces (where d_p is the metric on X_p). Put

$$X = \prod_{p\in\mathbb{N}} X_p;$$

recall that this product is the set of sequences $x = (x_p)_{p\in\mathbb{N}}$ such that $x_p \in X_p$ for each $p \in \mathbb{N}$. It is easy to check that the expression

$$d(x,y) = \sum_{p=0}^{+\infty} 2^{-p} \min\bigl(d_p(x_p, y_p),\, 1\bigr)$$

defines a metric d on X; this is called the **product distance** on X. For this metric, a sequence $(x^n)_{n\in\mathbb{N}}$ of points in X converges to a point $x \in X$ if and only if $\lim_{n\to\infty} x_p^n = x_p$ for every $p \in \mathbb{N}$.

If the metric spaces (X_p, d_p) are all equal to the same space (Y, δ), we write $X = Y^{\mathbb{N}}$. Then X is the set of sequences in X, or, what is the same, the set of maps from $\mathbb{N}$ into Y, with the metric of pointwise convergence.

One can then rephrase Theorem 3.1 as follows:

Corollary 3.2 (Tychonoff's Theorem) *If $(X_p)_{p\in\mathbb{N}}$ is a sequence of compact metric spaces and $X = \prod_{p\in\mathbb{N}} X_p$ is the product space (with the product distance), X is compact.*

This follows immediately from the definition of the product metric, from Theorem 3.1, and from the characterization of compact sets by the Bolzano–Weierstrass property.

Example. The space $\mathscr{C} = \{0,1\}^{\mathbb{N}}$, with the product distance

$$d(x,y) = \sum_{n=0}^{+\infty} 2^{-n} |x_n - y_n|,$$

is compact. It is easy to see that the map $\mathscr{C} \to [0,1]$ defined by

$$f(x) = 2\sum_{n=0}^{+\infty} 3^{-n-1} x_n$$

is a continuous injection, whose image is the **Cantor set** (which is therefore homeomorphic to $\mathscr{C}$).

Precompactness

We now give another application of the diagonal procedure. We start with a definition. A subset A of a metric space is **precompact** if, for every $\varepsilon > 0$, there are finitely many subsets $A_1, A_2, \ldots, A_n$ of A, each of diameter at most ε, such that $A = \bigcup_{j=1}^n A_j$.

Remarks

1. Clearly, every precompact subset is bounded. The converse is false, as can be seen from the example of the unit ball in an infinite-dimensional normed vector space (compare Theorem 1.1 on page 49). Precompact sets are also called **totally bounded**.
2. Unlike relative compactness, which is a relative property, precompactness involves only the intrinsic (induced) metric of the subspace.
3. Unlike compactness, precompactness is not a topological notion. It depends crucially on the metric; see Exercise 2 below, for example.
4. Each of the following two properties is equivalent to the precompactness of a subset A of a metric space X:
   - For every $\varepsilon > 0$ there exist finitely many points $x_1, \ldots, x_n$ of A such that $A \subset \bigcup_{j=1}^n B(x_j, \varepsilon)$.
   - For every $\varepsilon > 0$ there exist finitely many points $x_1, \ldots, x_n$ of X such that $A \subset \bigcup_{j=1}^n B(x_j, \varepsilon)$.

   The proof is elementary.

Theorem 3.3 *Let X be a metric space. Every relatively compact subset of X is precompact. The converse is true if X is complete.*

Proof. The first statement follows directly from the definitions, from the Borel–Lebesgue property of compact sets, and from the fact that $A \subset X$ implies $\bar{A} \subset \bigcup_{x \in X} B(x, \varepsilon)$ for every $\varepsilon > 0$.

Now suppose that X is complete and that $A \subset X$ is precompact. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of points in A. To prove that it has a convergent subsequence, it is enough to find a Cauchy subsequence. For every $p \in \mathbb{N}$, let $A_1^p, \ldots, A_{N_p}^p$ be subsets of A of diameter at most $1/(p+1)$ and covering A. We will construct by induction a decreasing sequence $(B_p)_{p\in\mathbb{N}}$ of infinite subsets of $\mathbb{N}$ such that, for every $p \in \mathbb{N}$, there is an integer $j \le N_p$ for which $\{x_p\}_{p\in B_p} \subset A_j^p$.

Construction of B_0: since all terms of the sequence $(x_n)_{n\in\mathbb{N}}$ (of which there are infinitely many) are contained in A, which is the union of the finitely many sets $A_1^0, \ldots, A_{N_0}^0$, there is at least one of these sets, say $A_{j_0}^0$, containing infinitely many terms x_n. (This is the pigeonhole principle.) We then set $B_0 = \{n \in \mathbb{N} : x_n \in A_{j_0}^0\}$.

To construct B_{p+1} from B_p, the idea is the same: the terms of the subsequence $(x_n)_{n\in B_p}$ are all contained in the union of the finitely many sets $A_1^{p+1}, \ldots, A_{N_{p+1}}^{p+1}$; therefore at least one of the sets contains infinitely many terms of the subsequence. We define B_{p+1} as the set of indices of these terms.

Having constructed the B_p, we define a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ by setting

$$\varphi(p) = \text{the } (p+1)\text{-st element of } B_p.$$

Then, for every $p \in \mathbb{N}$ and every integer $n \ge p$, we have $\varphi(n) \in B_p$. By the construction of the B_p, we see that

$$d\bigl(x_{\varphi(n)}, x_{\varphi(n')}\bigr) \le \frac{1}{p+1} \quad \text{for all } n, n' \ge p.$$

Thus the sequence $(x_{\varphi(n)})$ is a Cauchy sequence. □

Exercises

1. Let $(X_p, d_p)_{p\in\mathbb{N}}$ be a sequence of nonempty metric spaces, and let X be the product space with the product metric.
   a. Prove that (X, d) is separable if and only if each space (X_p, d_p) is separable.
   b. If $n \in \mathbb{N}$, $x \in X$ and $r > 0$, write

   $$U(x, n, r) = \bigl\{y \in X : d_p(x_p, y_p) < r \text{ for all } p \le n\bigr\},$$

   and define $\mathscr{U} = \{U(x, n, r) : x \in X,\ n \in \mathbb{N},\ r > 0\}$.
   i. Show that all the sets $U(x, n, r)$ are open in X.
   ii. Take $x \in X$ and $r > 0$. Prove that if $0 < \rho < r/2$, there exists an integer $n \in \mathbb{N}$ such that $x \in U(x, n, \rho) \subset B(x, r)$.
   iii. Show that $\mathscr{U}$ is a basis of open sets of X (see Exercise 1 on page 10).

iv. Let D be a dense subset of (X, d). Prove that the set

$$\mathscr{U}_D = \{U(x, n, 1/q) : x \in D,\, n \in \mathbb{N},\, q \in \mathbb{N}^*\}$$

is a basis of open sets of X. Prove that, if D is infinite, there exists a surjection from D onto $\mathscr{U}_D$.
Hint. When D is uncountable, one must use Exercise 9a on page 6.

2. If x and y are real numbers, we write $d(x, y) = |x - y|$ and $\delta(x, y) = |\arctan x - \arctan y|$. Prove that δ is a metric on $\mathbb{R}$ equivalent to the usual metric d; that is, the two metrics define the same open sets. Show that $(\mathbb{R}, \delta)$ is precompact, but $(\mathbb{R}, d)$ is not.

3. Prove that every precompact metric space is separable.

4. Prove that a metric space X is precompact if and only if every sequence of elements in X has a Cauchy subsequence.

5. *Helly's Theorem.* Let (f_n) be a sequence of increasing functions from a nonempty interval $I \subset \mathbb{R}$ into $\mathbb{R}$, such that for every $x \in I$ the sequence $\big(f_n(x)\big)$ is bounded.

a. Prove that there is a subsequence $(f_{\varphi(n)})_{n\in\mathbb{N}}$ such that, for every $x \in \mathbb{Q} \cap I$, the sequence $(f_{\varphi(n)}(x))_{n\in\mathbb{N}}$ converges. For such values of x, set $g(x) = \lim_{n\to\infty} f_{\varphi(n)}(x)$.

b. Extend g to all of I by setting, for $x \in I \setminus \mathbb{Q}$,

$$g(x) = \sup\{g(y) : y \in \mathbb{Q} \cap I \text{ and } y < x\}.$$

Prove that $g(x)$ is well-defined for all $x \in I$ and that the function g is increasing on I.

c. Let C be the set of points of I where g is continuous. We know from Exercise 6 on page 5 that the set $D = I \setminus C$ is countable. Prove that, for every $x \in C$, the sequence $\big(f_{\varphi(n)}(x)\big)$ converges toward $g(x)$.
Hint. Let $x \in C$. Prove that, if $y, z \in \mathbb{Q} \cap I$ with $y < x < z$, we have

$$g(y) \le \liminf_{n\to\infty}\big(f_{\varphi(n)}(x)\big) \le \limsup_{n\to\infty}\big(f_{\varphi(n)}(x)\big) \le g(z).$$

d. Using the diagonal procedure again, prove that there exists a subsequence $(f_{\varphi(\psi(n))})$ such that, for every $x \in I$, the sequence $\big(f_{\varphi(\psi(n))}(x)\big)$ converges.

6. **a.** Let X be a complete metric space, nonempty and with no isolated points. We will show that X contains a subset that is homeomorphic to the set $\mathscr{C} = \{0, 1\}^{\mathbb{N}}$ with the product distance.

i. Let B be an open ball in X with radius $r > 0$. Prove that there exist disjoint closed balls B_1 and B_2, of positive radii at most $r/2$, and both contained in B.

ii. Let $\mathscr{C}_0 = \bigcup_{n\in\mathbb{N}^*} \{0,1\}^n$ be the set of finite sequences of 0s and 1s. Let $u = (u_0, u_1, \ldots, u_{n-1}) \in \{0,1\}^n$ and $v = (v_0, v_1, \ldots, v_{m-1}) \in \{0,1\}^m$ be elements of $\mathscr{C}_0$. We say that u is an *initial segment* of v if $n \leq m$ and $u_i = v_i$ for all $i < n$. We say that u and v are *incompatible* if u is not an initial segment of v and v is not an initial segment of u.
Prove that one can construct a map $u \mapsto B_u$ that associates to every $u \in \mathscr{C}_0$ a closed ball B_u of X, of positive radius, satisfying these properties:
- If u is an initial segment of v, then $B_v \subset \mathring{B}_u$.
- If u and v are incompatible, $B_u \cap B_v = \varnothing$.
- If u has length n, the radius of B_u is at most 2^{-n}.

Hint. One can start by defining $B_{(0)}$ and $B_{(1)}$, then work by induction on the length of the finite sequences: suppose the B_u have been constructed for all sequences u of length at most n, and give a procedure for constructing the B_u for sequences u of length $n+1$.

iii. If $\alpha \in \mathscr{C}$, define the set

$$X_\alpha = \bigcup_{\substack{u\in\mathscr{C}_0 \\ u \text{ an initial segment of } \alpha}} B_u.$$

(Naturally, we say that a finite sequence $(u_0, \ldots, u_{n-1})$ is an initial segment of α if $u_i = \alpha_i$ for all $i < n$.) Prove that X_α contains a single point, which we denote x_α.

iv. Prove that the map $x : \alpha \mapsto x_\alpha$ is a continuous (and even Lipschitz) injection from $\mathscr{C}$ into X.

v. Deduce that $\mathscr{C}$ and $x(\mathscr{C})$ are homeomorphic.

b. Prove that every complete separable space is either countable or in bijection with $\mathbb{R}$. In particular, this is the case for every closed subset of $\mathbb{R}$.
Hint. One can use Exercise 2 on page 10, the Cantor–Bendixon Theorem (Exercise 6 on page 11), Exercise 3 on page 5, and Exercise 9b on page 7.

7. Prove that the space $\mathscr{C} = \{0,1\}^{\mathbb{N}}$, with the product distance, is homeomorphic to $\mathscr{C} \times \mathscr{C}$.
Hint. One can show that the map

$$(x_n)_{n\in\mathbb{N}} \mapsto \big((x_{2n})_{n\in\mathbb{N}}, (x_{2n+1})_{n\in\mathbb{N}}\big)$$

is a continuous bijection between $\mathscr{C}$ and $\mathscr{C} \times \mathscr{C}$.

8. Let A be a subset of a normed vector space E. Prove that A is precompact if and only if A is bounded and, for every $\varepsilon > 0$, there exists a finite-dimensional vector subspace F_ε of E such that $d(x, F_\varepsilon) \leq \varepsilon$ for all $x \in A$.

9. Let E be a normed space.

a. Let A be a nonempty subset of E. Prove that there is a (unique) smallest closed convex set containing A. This set is called the *closed convex hull* of A, and we will denote it by $\bar{c}(A)$.

b. Let A be a precompact subset of E.

i. Set $M = \sup_{x \in A} \|x\|$ and, for every $\varepsilon > 0$, define a subset of E,

$$A_\varepsilon = \left\{ x \in E : \|x\| \leq M \text{ and } d(x, F_\varepsilon) \leq \varepsilon \right\},$$

where F_ε is a finite-dimensional vector space such that $d(x, F_\varepsilon) \leq \varepsilon$ for every $x \in A$ (see Exercise 8). Prove that, for every $\varepsilon > 0$, the set A_ε is a closed convex set containing A.

ii. Set $A_0 = \bigcap_{0<\varepsilon\leq 1} A_\varepsilon$. Prove that the set A_0 is convex, closed, and precompact. (Use Exercise 8.)

iii. Deduce that $\bar{c}(A)$ is precompact.

c. Suppose that E is a Banach space. Prove that if A is a relatively compact subset of E, then $\bar{c}(A)$ is compact.

4 Bounded Sequences of Continuous Linear Maps

We now use the denseness and separability results given earlier, together with consequences of the diagonal procedure, to study bounded sequences of continuous linear maps. We start with some notation.

Notation. Let E and F be normed vector spaces over the same field $\mathbb{K}$. We denote by $L(E, F)$ the space of continuous linear maps from E to F. In general, we use the same symbol $\|\cdot\|$ for the norms on E, on F and on $L(E, F)$. The latter norm assigns to $T \in L(E, F)$ the number

$$\|T\| = \sup\{\|Tx\| : x \in E \text{ and } \|x\| \leq 1\}.$$

Recall that, if F is a Banach space, so is $L(E, F)$. We use also the following notations: $L(E) = L(E, E)$, and $E' = L(E, \mathbb{K})$; we call E' the **topological dual** of E.

Recall also that in a normed space E, a subset A is said to be **bounded** if it is contained in a ball; that is, if the set of norms of elements of A is bounded.

The first proposition deals with the case where F is a Banach space.

Proposition 4.1 *Consider a normed space E, a fundamental family D in E, and a Banach space F. Consider also a bounded sequence $(T_n)_{n\in\mathbb{N}}$ of elements of $L(E, F)$. If, for every $x \in D$, the sequence $(T_n x)_{n\in\mathbb{N}}$ converges in F, there exists an operator $T \in L(E, F)$ such that*

$$\lim_{n\to+\infty} T_n x = Tx \quad \textit{for every } x \in E.$$

Proof. Let $M > 0$ be such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. It is clear that the sequence $(T_n x)$ converges for any element x of the vector space $[D]$ generated by D. Now take $x \in E$ and $\varepsilon > 0$. Since D is a fundamental family, there exists $y \in [D]$ such that $\|x-y\| \leq \varepsilon/(3M)$. The sequence $(T_n y)$ converges; therefore there is a positive integer N such that $\|T_n y - T_p y\| \leq \varepsilon/3$ for all $n, p \geq N$. By the triangle inequality we deduce that, for any $n, p \geq N$,

$$\|T_n x - T_p x\| \leq \|T_n x - T_n y\| + \|T_n y - T_p y\| + \|T_p y - T_p x\| \leq \varepsilon.$$

Thus $(T_n x)$ is a Cauchy sequence in F, and therefore convergent. For every $x \in E$ we then set $Tx = \lim_{n\to\infty} T_n x$. The map T thus defined is certainly linear, and, since $\|Tx\| \leq M\|x\|$ for all $x \in E$, it is also continuous. □

Corollary 4.2 (Banach–Alaoglu) *Let E be a separable normed space. For every bounded sequence $(T_n)_{n\in\mathbb{N}}$ in E', there are a subsequence $(T_{n_k})_{k\in\mathbb{N}}$ and a continuous linear form $T \in E'$ such that*

$$\lim_{k\to\infty} T_{n_k} x = Tx \quad \textit{for all } x \in E.$$

Warning: the sequence (T_{n_k}) does not necessarily converge in E'; that is, $\|T_{n_k} - T\|$ does not in general tend toward 0.

Proof. Choose $M > 0$ such that $\|T_n\| \leq M$ for every $n \in \mathbb{N}$, and let $(x_p)_{p\in\mathbb{N}}$ be a dense sequence in E. For every positive integer p, we have

$$|T_n x_p| \leq M\|x_p\| \quad \text{for all } n \in \mathbb{N}.$$

Therefore the set $\{T_n x_p\}_{n\in\mathbb{N}}$ is relatively compact in $\mathbb{K}$. By Theorem 3.1, there exists a subsequence (T_{n_k}) such that, for every p, the sequence of images $\left(T_{n_k} x_p\right)_{k\in\mathbb{N}}$ converges in $\mathbb{K}$. Now apply Proposition 4.1. □

This is not necessarily true if E is not separable; see, for example, Exercise 3 below.

A weaker result than Proposition 4.1 holds when F is any normed space:

Proposition 4.3 *Consider normed spaces E and F, a fundamental set D in E, a bounded sequence (T_n) in $L(E,F)$ and a map $T \in L(E,F)$. If the sequence $(T_n x)$ converges toward Tx for every point $x \in D$, it does also for every $x \in E$.*

Proof. By taking differences we can suppose that $T = 0$. Set

$$M = \sup_{n\in\mathbb{N}} \|T_n\|$$

and take $x \in E$. For every $y \in [D]$, we have

$$\|T_n x\| \leq M\|x-y\| + \|T_n y\|.$$

Since $T_n y \to 0$, we get $\limsup_{n\to\infty} \|T_n x\| \le M\|x-y\|$. This holds for every $y \in [D]$, and $[D]$ is dense in E; therefore

$$\lim_{n\to\infty} \|T_n x\| = 0. \qquad \square$$

Exercises

1. Consider normed spaces E and F, a bounded sequence $(T_n)_{n\in\mathbb{N}}$ in $L(E,F)$, and an element $T \in L(E,F)$. Prove that, if $\lim_{n\to+\infty} T_n x = Tx$ for every $x \in E$, the limit is uniform on any compact subset of E.
2. Consider a normed space E, a Banach space F, and a bounded sequence $(T_n)_{n\in\mathbb{N}}$ in $L(E,F)$. Prove that the set of points $x \in E$ for which the sequence $(T_n x)$ converges is a closed vector subspace of E.
3. Consider the space $E = \ell^\infty$ of Exercise 7 on page 11. Prove that the sequence (T_n) of E' defined by $T_n(x) = x_n$ has no pointwise convergent subsequence in E.
4. Let E be a separable normed vector space, and let $(x_p)_{p\in\mathbb{N}}$ be a dense sequence in E. Denote by B the unit ball of E', that is,

$$B = \{T \in E' : |T(x)| \le \|x\| \text{ for all } x \in E\}.$$

For T and S elements of B, we define the real number

$$d(T,S) = \sum_{p=0}^{+\infty} 2^{-p} \min\bigl(|T(x_p) - S(x_p)|,\, 1\bigr).$$

   a. Prove that d is a metric on B. If (T_n) is a sequence of elements of B and if $T \in B$, prove that

$$d(T_n, T) \to 0 \iff T_n(x) \to T(x) \text{ for all } x \in E.$$

   b. Prove that the metric space (B,d) is compact.

5. *Riemann integral of Banach-space valued functions.* Let $[a,b]$ be an interval in $\mathbb{R}$ and let E be a Banach space. We want to define the integral of a continuous function and, more generally, of a regulated function from $[a,b]$ into E.

   a. *Integral of staircase functions.* A *staircase function* from $[a,b]$ to E is one for which there is a subdivision $x_0 = a < x_1 < \cdots < x_n = b$ of $[a,b]$ and vectors $v_1, \ldots, v_{n-1}$ in E such that, for every $i \le n-1$ and every $x \in (x_i, x_{i+1})$, we have $f(x) = v_i$. The integral of such a function f over $[a,b]$ is defined by

$$I(f) = \int_a^b f(x)\,dx = \sum_{i=0}^{n-1} (x_{i+1} - x_i) v_i.$$

We denote by $\mathscr{E}$ the vector space of all staircase functions on $[a,b]$, with the uniform norm: $\|f\|_\infty = \sup_{x\in[a,b]} \|f(x)\|$. Check that I is a continuous linear map from $\mathscr{E}$ to E, with norm $b-a$. Check also that, if $f \in \mathscr{E}$, *Chasles's relation* holds for arbitrary $\alpha, \beta, \gamma \in [a,b]$:

$$\int_\alpha^\beta f(x)\,dx = \int_\alpha^\gamma f(x)\,dx + \int_\gamma^\beta f(x)\,dx,$$

where, by convention, we set

$$\int_u^v f(x)\,dx = -\int_v^u f(x)\,dx \quad \text{if } u > v.$$

b. Prove that a function from $[a,b]$ to E is regulated (Exercise 7 on page 5) if and only if it is the uniform limit of a sequence of staircase functions.

Hint. "Only if" part: Let f be a regulated function from $[a,b]$ to E, and choose $\varepsilon > 0$. Prove that there is a subdivision $a = x_0 < x_1 < \cdots < x_n = b$ of $[a,b]$ such that, for every i and every $x, y \in (x_i, x_{i+1})$, we have $\|f(x) - f(y)\| \le \varepsilon$. Deduce the existence of a staircase function g such that $\|f(x) - g(x)\| < \varepsilon$ for every $x \in [a,b]$.

"If" part: Since E is complete, f has a left limit at a point x if and only if, for every $\varepsilon > 0$, there exists $\eta > 0$ such that $\|f(y) - f(z)\| < \varepsilon$ for all $y, z \in (x-\eta,\, x)$.

c. **i.** Let $\mathscr{F}_b([a,b],\, E)$ be the space of bounded functions from $[a,b]$ into E, with the uniform norm: $\|f\|_\infty = \sup_{x\in[a,b]} \|f(x)\|$. Prove that $\mathscr{F}_b([a,b], E)$ is a Banach space.

ii. Let $\mathscr{R}$ be the set of regulated functions from $[a,b]$ into E. Prove that $\mathscr{R}$ is a closed subspace of $\mathscr{F}_b([a,b],\, E)$. Thus, $\mathscr{R}$ with the uniform norm is a Banach space.

d. *Integral of a regulated function.* Prove that I can be uniquely extended into a continuous linear map J on all of $\mathscr{R}$, of norm $b-a$. (One can use the theorem of extension of Banach-space-valued continuous linear maps.) For every $f \in \mathscr{R}$, the image of f under the map is of course denoted by

$$J(f) = \int_a^b f(x)\,dx.$$

e. Check that Chasles's relation (see item (a)) holds for all regulated functions. Check also that, if F is a continuous linear form on E and if $f \in \mathscr{R}$, then $F \circ f$ is a regulated function from $[a,b]$ into $\mathbb{K}$, and that

$$F(J(f)) = \int_a^b F(f(x))\,dx.$$

f. Prove that, for every function f in $\mathscr{R}$,

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \| f(x) \| \, dx.$$

g. If $\Delta = (x_0, \ldots, x_n)$ is a subdivision of $[a, b]$, and if $\xi = (\xi_0, \ldots, \xi_{n-1})$ is such that $\xi_j \in [x_j, x_{j+1}]$ for $0 \leq j \leq n-1$, we set

$$S(\Delta, \xi)(f) = \sum_{j=0}^{n-1} f(\xi_j)(x_{j+1} - x_j).$$

Prove that, if (Δ^p, ξ^p) is a sequence of subdivisions whose maximal step size tends to 0, and if f is any function in $\mathscr{R}$, then $S(\Delta^p, \xi^p)(f)$ converges to $\int_a^b f(x) \, dx$.

Hint. One can start with the case of a staircase function f, then use Proposition 4.3.

6. *The Baire and Banach–Steinhaus Theorems.* Let X be any metric space. Two players, Pierre and Paul, play the following "*game of Choquet*": Pierre chooses a nonempty open set U_1 in X, then Paul chooses a nonempty open set V_1 inside U_1, then Pierre chooses a nonempty open set U_2 inside V_1, and so on. At the end of the game, the two players have defined two decreasing sequences (U_n) and (V_n) of nonempty open sets such that

$$U_n \supseteq V_n \supseteq U_{n+1} \quad \text{for every } n \in \mathbb{N}.$$

Note that $\bigcap_{n \in \mathbb{N}} U_n = \bigcap_{n \in \mathbb{N}} V_n$; we denote this set by U. Pierre wins if U is empty, and Paul wins if U is nonempty. We say that one of the players has a winning strategy if he has a method that allows him to win whatever his opponent does. Therefore, the two players cannot both have a winning strategy; *a priori*, it is possible that neither does.

a. Prove that, if X has a nonempty open set O that is a countable union of closed sets F_n with empty interior, Pierre was a winning strategy.

Hint. Pierre starts with $U_1 = O$ and responds to each choice V_n of Paul's with $V_n \setminus F_n$.

b. Prove that, if X is complete, Paul has a winning strategy.

Hint. If (F_n) is a decreasing sequence of closed sets in X whose diameter tends to 0, the intersection of the F_n is nonempty.

c. *Application: Baire's Theorem.* Let X be a complete space. Prove that an open set of X cannot be the union of a countable family of closed sets with empty interior.

d. *Corollary: The Banach–Steinhaus Theorem.* Consider a Banach space E, a normed vector space F, and a family $(T_n)_{n \in \mathbb{N}}$ of elements of

$L(E, F)$ such that, for every $x \in E$, the set $\{\|T_n(x)\| : n \in \mathbb{N}\}$ is bounded. Prove that $\{\|T_n\| : n \in \mathbb{N}\}$ is bounded.
Hint. Show that there exists $k \in \mathbb{N}$ such that the set

$$F_k = \{x \in E : \|T_n(x)\| \le k \text{ for all } n \in \mathbb{N}\}$$

has nonempty interior, and therefore contains some open ball $B(a, r)$; then show that, for every $n \in \mathbb{N}$,

$$\|T_n\| \le \frac{1}{r}\Big(\sup_{m\in\mathbb{N}} \|T_m(a)\| + k\Big).$$

e. Prove that an infinite-dimensional Banach space cannot have a countable generating set. For example, $\mathbb{R}[X]$ cannot be made into a Banach space.
Hint. If this were not the case, the space would be a countable union of closed sets with empty interiors.

f. Let (T_n) be a sequence of continuous linear operators from a Banach space E into a normed vector space F, having the property that, for every $x \in E$, the sequence $(T_n(x))$ converges. Prove that the map $T : E \to F$ defined by $T(x) = \lim_{n\to\infty} T_n(x)$ is linear and continuous.

g. **i.** Let f be a function from $\mathbb{R}$ to $\mathbb{R}$. Prove that the set of points where f is continuous is a G_δ-set in $\mathbb{R}$, that is, a countable intersection of open sets in $\mathbb{R}$.
Hint. Define, for each $n \in \mathbb{N}^*$, the set C_n consisting of points $x \in \mathbb{R}$ for which there exists an open set V containing x and such that $|f(y) - f(z)| < 1/n$ for all $y, z \in V$. Prove that the sets C_n are open.

ii. Prove that $\mathbb{Q}$ is not a G_δ in $\mathbb{R}$.
Hint. If it were, $\mathbb{R}$ would be a countable union of closed sets with empty interior.

iii. Prove that there is no function from $\mathbb{R}$ to $\mathbb{R}$ that is continuous at every point of $\mathbb{Q}$ and discontinuous everywhere else.

iv. Prove that there exist functions from $\mathbb{R}$ to $\mathbb{R}$ that are discontinuous at every point of $\mathbb{Q}$ and continuous everywhere else.
Hint. Use Exercise 6c on page 5. More directly, if $\{x_n\}_{n\in\mathbb{N}}$ is an enumeration of $\mathbb{Q}$, the function f defined by $f(x) = 0$ if $x \notin \mathbb{Q}$ and $f(x_n) = 1/(n+1)$ for every $n \in \mathbb{N}$ has the desired properties.

7. An *invariant metric* on a vector space E is a metric d on E such that

$$d(x, y) = d(x-y,\, 0) \quad \text{for all } x, y \in E.$$

If d is an invariant metric on E, we set $|x| = d(x, 0)$ for $x \in E$. (Note that the map $|\cdot|$ thus defined is not necessarily a norm on E.) A vector space with an invariant metric d is said to have Property (F) if the metric space (E, d) is complete and, for every $k \in \mathbb{K}$, the map $x \mapsto kx$ is continuous from E to E. For example, every Banach space with the norm-induced metric has Property (F).

Let E be a vector space having an invariant metric with Property (F). Let F be a normed vector space, with norm $\|\cdot\|$.

a. Let H be a family of continuous linear maps from E to F such that, for every $x \in E$, the set $\{T(x)\}_{T \in H}$ is bounded. Prove that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|T(x)\| \le \varepsilon \quad \text{for all } x \in E \text{ with } |x| \le \delta \text{ and for all } T \in H;$$

in other words, $\lim_{x \to 0} T(x) = 0$ uniformly in $T \in H$.
Hint. Take $\varepsilon > 0$ and, for each $k \in \mathbb{N}^*$, set

$$F_k = \bigl\{x \in E : \|T(x/k)\| \le \varepsilon \text{ for all } T \in H\bigr\}.$$

Using Baire's Theorem (Exercise 6), prove that at least one of the F_k, say F_{k_0}, contains an open ball $B(a, r)$. Then use the fact that F_{k_0} is a symmetric convex set (*symmetry* here means that $-F_{k_0} = F_{k_0}$) and the continuity of the map $x \mapsto 2k_0 x$.

b. Let (T_n) be a sequence of continuous linear maps from E to F such that, for every $x \in E$, the sequence $(T_n(x))$ converges. Prove that the map from E to F defined by

$$T(x) = \lim_{n \to +\infty} T_n(x)$$

is linear and continuous. (This generalizes Exercise 6f above.)
We will be able to apply this result to sequences in $\mathfrak{M}_c(X)$ (Exercise 10 on page 92) or in L^p_c, for $1 < p \le \infty$ (Exercise 12 on page 168). See also Exercises 1 on page 147 and 1 on page 163.

Part I

FUNCTION SPACES AND THEIR DUALS

1

The Space of Continuous Functions on a Compact Set

Introduction and Notation

We will consider throughout this chapter a compact, nonempty metric space (X, d), and we will study the $\mathbb{K}$-vector space (for $\mathbb{K} = \mathbb{R}, \mathbb{C}$) of continuous functions from X to $\mathbb{K}$, which we denote by $C^{\mathbb{K}}(X)$, or simply $C(X)$ when no confusion is likely. We give $C(X)$ a commutative multiplication operation: for $f, g \in C(X)$ the product fg is defined by

$$(fg)(x) = f(x)g(x) \quad \text{for all } x \in X.$$

The constant function 1 is the unity element for this multiplication. We say that $C(X)$ is a **commutative algebra with unity**.

The space $C^{\mathbb{R}}(X)$ also has an order relation $\leq$, defined by

$$f \leq g \iff f(x) \leq g(x) \text{ for all } x \in X;$$

it is only a partial order, of course. For any $f, g \in C^{\mathbb{R}}(X)$, there exist a least upper bound and a greatest lower bound for f and g:

$$\left.\begin{aligned} \sup(f,g)(x) &= \max(f(x), g(x)) \\ \inf(f,g)(x) &= \min(f(x), g(x)) \end{aligned}\right\} \text{ for all } x \in X.$$

That the functions thus defined are continuous can be seen, for example, from the following equalities:

$$\sup(f,g) = \tfrac{1}{2}\bigl(f + g + |f - g|\bigr), \quad \inf(f,g) = \tfrac{1}{2}\bigl(f + g - |f - g|\bigr). \qquad (*)$$

We denote by $C^+(X)$ the set of continuous functions from X to $\mathbb{R}^+$. If $f \in C^{\mathbb{R}}(X)$, we write $f^+ = \sup(f, 0)$ and $f^- = -\inf(f, 0)$ (note that we use the same symbol for a constant function and its value). We therefore have

$$f^+(x) = \big(f(x)\big)^+, \quad f^-(x) = \big(f(x)\big)^-, \quad f = f^+ - f^-, \quad |f| = f^+ + f^-.$$

1 Generalities

We give $C(X)$ the **uniform norm** over X, denoted by $\|\cdot\|$ and defined by

$$\|f\| = \max_{x \in X} \big|f(x)\big|$$

The corresponding topology is called the **topology of uniform convergence**, since a sequence in $C(X)$ converges to $f \in C(X)$ in this norm if and only if it converges uniformly to f on X.

Clearly, $\|fg\| \le \|f\|\,\|g\|$ and $\|\,|f|\,\| = \|f\|$ for all $f, g \in C(X)$.

Proposition 1.1 *$C(X)$ is a separable Banach space.*

Proof. The reader can check that $C(X)$ is a Banach space. We show separability. Since X is precompact, for every $n \in \mathbb{N}^*$ there exist finitely many points $x_1^n, \dots, x_{N_n}^n$ of X such that $X = \bigcup_{j=1}^{N_n} B(x_j^n, 1/n)$. We therefore set, for $j \le N_n$,

$$\varphi_{n,j}(x) = \frac{\big(1/n - d(x, x_j^n)\big)^+}{\sum_{k=1}^{N_n} \big(1/n - d(x, x_k^n)\big)^+}.$$

From the choice of the points x_j^n, we see that the denominator does not vanish for any $x \in X$. Therefore, $\varphi_{n,j} \in C^+(X)$,

$$\sum_{j=1}^{N_n} \varphi_{n,j} = 1, \qquad \text{and} \qquad \varphi_{n,j}(x) = 0 \quad \text{if } d(x, x_j^n) \ge 1/n.$$

The set $\{\varphi_{n,j} : n \in \mathbb{N}^* \text{ and } 1 \le j \le N_n\}$ is certainly countable. We will show that it is a fundamental family in $C(X)$; this suffices by Proposition 2.5 on page 9.

Take $f \in C(X)$ and $\varepsilon > 0$. Since X is compact, the function f is uniformly continuous on X. Take $\eta > 0$ such that, for all $x, y \in X$ with $d(x, y) < \eta$, we have $|f(x) - f(y)| < \varepsilon$. Let $n \in \mathbb{N}$ be such that $1/n < \eta$. For every $x \in X$,

$$\begin{aligned}\left|f(x) - \sum_{j=1}^{N_n} f(x_j^n)\varphi_{n,j}(x)\right| &= \left|\sum_{j=1}^{N_n} \big(f(x) - f(x_j^n)\big)\varphi_{n,j}(x)\right| \\ &\le \sum_{j=1}^{N_n} \big|f(x) - f(x_j^n)\big|\, \varphi_{n,j}(x).\end{aligned}$$

Since $\varphi_{n,j}$ vanishes outside the ball $B(x_j^n, 1/n)$, and so outside $B(x_j^n, \eta)$, we see that, for every $x \in X$,

$$\left|f(x) - f(x_j^n)\right| \varphi_{n,j}(x) \leq \varepsilon \varphi_{n,j}(x).$$

Thus, for every $x \in X$,

$$\left|f(x) - \sum_{j=1}^{N_n} f(x_j^n)\varphi_{n,j}(x)\right| \leq \varepsilon \sum_{j=1}^{N_n} \varphi_{n,j}(x) = \varepsilon.$$

It follows that

$$\left\|f - \sum_{j=1}^{N_n} f(x_j^n)\varphi_{n,j}\right\| \leq \varepsilon,$$

which concludes the proof. □

We recall a sufficient criterion for uniform convergence (and therefore convergence in $C(X)$) that is often convenient:

Proposition 1.2 (Dini's Lemma) *Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in $C^{\mathbb{R}}(X)$* (this means that $f_n \leq f_{n+1}$ for all n). *If the sequence (f_n) converges pointwise to a function $f \in C(X)$, it also converges uniformly to f.*

Proof. Take $\varepsilon > 0$. For every $n \in \mathbb{N}$ we set $\Omega_n = \{x \in X : f_n(x) > f(x) - \varepsilon\}$. Clearly, (Ω_n) is an increasing sequence of open subsets in X whose union is X. By the Borel–Lebesgue property, there is an integer N such that $\Omega_N = X$, so that $f_N(x) > f(x) - \varepsilon$ for all $x \in X$. Thus, for every integer $n \geq N$, we have $f(x) - \varepsilon < f_n(x) \leq f(x)$ for all $x \in X$. This proves that $\|f - f_n\| \leq \varepsilon$. □

Remarks

1. Clearly, one can replace "increasing" by "decreasing" in the statement of Dini's Lemma.
2. The assumption that the pointwise limit f is continuous is essential. For example, the decreasing sequence (f_n) of continuous functions on $[0,1]$ given by $f_n(x) = x^n$ converges pointwise, but not uniformly, on $[0,1]$.

Example. Define by induction on n a sequence of polynomial functions (P_n) on $[-1,1]$, as follows:

$$P_0 = 0,$$
$$P_{n+1}(x) = P_n(x) + \tfrac{1}{2}(x^2 - P_n^2(x)) \quad \text{for all } n \in \mathbb{N}.$$

We check that, for every $n \in \mathbb{N}$, we have $0 \leq P_n(x) \leq P_{n+1}(x) \leq |x|$ for all $x \in [-1,1]$. For $n = 0$ this is clear; suppose by induction that it is true for some $n \geq 0$. Then, for all $x \in [-1,1]$,

$$0 \leq P_{n+1}(x) \leq P_{n+2}(x) = |x| - \big(|x| - P_{n+1}(x)\big)\big(1 - \tfrac{1}{2}(|x| + P_{n+1}(x))\big) \leq |x|.$$

Then the sequence $(P_n)_{n\in\mathbb{N}}$ is increasing and bounded, and therefore it converges pointwise to a function f. For $x \in [-1,1]$, we see that $0 \le f(x) \le |x|$ and $f^2(x) = x^2$, by taking to the limit the defining recursive relation of the P_n. Therefore $f(x) = |x|$, and Dini's Lemma applies. This proves that the polynomial sequence (P_n) converges uniformly to $|x|$ on $[-1,1]$.

We will generalize this result in the next section, demonstrating that every continuous function in $[-1,1]$ is the uniform limit of a sequence of polynomial functions (Weierstrass's Theorem).

Exercises

1. Show that there exists a sequence $(P_n)_{n\in\mathbb{N}}$ in $L(C(X))$ such that, for all $n \in \mathbb{N}$, the map P_n has finite rank (that is, $P_n(C(X))$ is a finite-dimensional vector space), has norm 1, is *positive* (that is, $P_n(f) \ge 0$ for all $f \ge 0$), and satisfies

$$\lim_{n\to+\infty} P_n f = f \quad \text{for all } f \in C(X).$$

2. Let p be a bounded, strictly increasing continuous function from $\mathbb{R}$ to $\mathbb{R}$. Set $p(-\infty) = \lim_{x\to-\infty} p(x)$ and $p(+\infty) = \lim_{x\to+\infty} p(x)$. Also set $X = [-\infty,+\infty] = \mathbb{R} \cup \{-\infty,+\infty\}$, and define a map $d_p : X^2 \to \mathbb{R}$ by $d_p(x,y) = \bigl|p(x) - p(y)\bigr|$. Prove that d_p is a metric on X, that the metric space (X, d_p) is compact, that d_p induces on $\mathbb{R}$ the usual topology, that $\mathbb{R}$ is dense in (X, d_p), and that $(\mathbb{R}, d_p)$ is precompact. Prove also that the topology thus defined on X (that is, the family of open sets defined by d_p) does not depend on p.
3. Let $(f_n)_{n\in\mathbb{N}^*}$ be a sequence of continuous functions on $\mathbb{R}^+$ defined by

$$f_n(x) = \begin{cases} (1 - x/n)^n & \text{if } x \le n, \\ 0 & \text{if } x > n. \end{cases}$$

Prove that the sequence (f_n) converges uniformly in $[0,+\infty)$ to the function $f : x \mapsto e^{-x}$.

Hint. Extend the functions f_n to have the value 1 on $[-\infty, 0]$ and the value 0 at $+\infty$. Then apply Dini's Lemma in the compact space $[-\infty,+\infty]$ introduced in Exercise 2.
4. *A generalization of Dini's Lemma.* Consider a compact metric space X, and elements f and $\{f_n\}_{n\in\mathbb{N}}$ of $C(X)$. Assume that there exists a constant $C > 0$ such that

$$|f - f_{p+q}| \le C\,|f - f_p| \quad \text{for all } p, q \in \mathbb{N}.$$

Prove that if the sequence (f_n) converges pointwise to f, it converges uniformly to f. (One can look at the proof of Dini's Lemma for inspiration.)

5. *Ideals in $C(X)$.* Let X be a compact metric space and J an ideal in the ring $(C(X), +, \cdot)$. Denote by Z the set of points x in X such that $g(x) = 0$ for all $g \in J$.

a. Prove that, if Z is empty, J contains a function g such that $g(x) > 0$ for all $x \in X$. Deduce that $J = C(X)$.

b. For $a \in X$, set $J_a = \{g \in C(X) : g(a) = 0\}$. Prove that J_a is a maximal ideal; that is, the only ideal that strictly contains J_a is $C(X)$.

c. Conversely, prove that, if J is a maximal ideal, there is a unique point a of X such that $J = J_a$.

d. Prove that $\bar{J} = \{f \in C(X) : f(x) = 0 \text{ for all } x \in Z\}$.

Hint. Let $f \in C(X)$ vanish everywhere in Z. To find an element of J that is 2ε-close to f, one can do this:

i. Let K be the set of points x of X for which $|f(x)| \geq \varepsilon$. Prove that there exists $g \in J$ such that $g(x) > 0$ for all $x \in K$ and $g(x) \geq 0$ for all $x \in X$.

ii. Prove that, for all large enough n, the function f_n defined by

$$f_n = f\frac{ng}{1+ng}$$

is in J, and that $\|f_n - f\| \leq 2\varepsilon$.

2 The Stone–Weierstrass Theorems

We now state denseness criteria for the subspaces of $C(X)$. These criteria are consequences of this fundamental lemma:

Lemma 2.1 *Suppose X has at least two elements. Let H be a subset of $C^{\mathbb{R}}(X)$ satisfying these two conditions:*

a. *For all $u, v \in H$, the functions* $\sup(u, v)$ *and* $\inf(u, v)$ *also lie in H.*

b. *If x_1, x_2 are distinct points in X and α_1, α_2 are real numbers, there exists $u \in H$ such that $u(x_1) = \alpha_1$ and $u(x_2) = \alpha_2$.*

Then H is dense in $C^{\mathbb{R}}(X)$.

Proof. Take $f \in C^{\mathbb{R}}(X)$ and $\varepsilon > 0$. We want to find an element of H that is ε-close to f. First fix $x \in X$. By assumption b, for every $y \neq x$ there exists $u_y \in H$ such that $u_y(x) = f(x)$ and $u_y(y) = f(y)$.

For $y \neq x$, set $O_y = \{x' \in X : u_y(x') > f(x') - \varepsilon\}$. This is an open set that contains y and x; therefore $X = \bigcup_{y \neq x} O_y$. By the Borel–Lebesgue property, X can be covered by finitely many sets O_y: $X = \bigcup_{j=1}^{r} O_{y_j}$, with $y_j \neq x$ for all j. Now set $v_x = \sup(u_{y_1}, \ldots, u_{y_r})$. A simple inductive argument, using assumption a, shows that $v_x \in H$. On the other hand,

$$v_x(x) = f(x) \quad \text{and} \quad v_x(x') > f(x') - \varepsilon \text{ for all } x' \in X.$$

Now make x vary and set, for each $x \in X$,

$$\Omega_x = \{x' \in X : v_x(x') < f(x') + \varepsilon\}.$$

Thus Ω_x is an open subset of X containing x; a new application of the Borel–Lebesgue property allows us to choose finitely many points $x_1, \ldots, x_p$ of X such that $\Omega_{x_1}, \ldots, \Omega_{x_p}$ cover X. Finally, set $v = \inf(v_{x_1}, \ldots, v_{x_p})$. Then $v \in H$ and $f - \varepsilon < v < f + \varepsilon$; that is, $\|f - v\| \leq \varepsilon$. □

A subset H of $C(X)$ is called **separating** if, for any two distinct points x, y of X, there exists $h \in H$ with $h(x) \neq h(y)$. A subset H of $C^{\mathbb{R}}(X)$ is called a **lattice** if, for any $f, g \in H$, the functions $\sup(f, g)$ and $\inf(f, g)$ also lie in H. Notice that a vector subspace of $C^{\mathbb{R}}(X)$ is a lattice if and only if, for every element h of H, the function $|h|$ is in H as well (the "only if" part follows from the relation $|h| = \sup(h, 0) - \inf(h, 0)$, and the "if" part from equations $(*)$ on page 27).

We can then deduce from Lemma 2.1 the following theorem:

Theorem 2.2 *If H is a separating vector subspace of $C^{\mathbb{R}}(X)$ that is a lattice and contains the constants, then H is dense in $C^{\mathbb{R}}(X)$.*

Proof. If X has a single element, the result is clear. Suppose X has at least two elements; we just need to check assumption b of the lemma. Let x_1 and x_2 be distinct elements of X. Since H is separating, there exists $h \in H$ such that $h(x_1) \neq h(x_2)$. If α_1 and α_2 are real numbers, the system of equations

$$\begin{cases} \lambda h(x_1) + \mu = \alpha_1 \\ \lambda h(x_2) + \mu = \alpha_2 \end{cases}$$

clearly has a unique solution $(\lambda, \mu) \in \mathbb{R}^2$. For such (λ, μ), we see that $(\lambda h + \mu)(x_1) = \alpha_1$ and $(\lambda h + \mu)(x_2) = \alpha_2$; moreover, $\lambda h + \mu \in H$, since H is a vector space containing constants. □

Example. Let H be the set of **Lipschitz functions** from X to $\mathbb{R}$, that is, the set of functions h from X to $\mathbb{R}$ for which there is a constant $C \geq 0$ (depending on h) such that $\left|h(x) - h(y)\right| \leq Cd(x, y)$ for all $(x, y) \in X^2$. Such a C is called a **Lipschitz constant** for h, and h is said to be **C-Lipschitz**. Clearly, H is a vector subspace of $C^{\mathbb{R}}(X)$ containing the constant functions. H is also a lattice: the absolute value of a Lipschitz function is Lipschitz as well, since

$$\left| |h(x)| - |h(y)| \right| \leq \left|h(x) - h(y)\right|.$$

Finally, H is separating since, for $x \neq y$, the function $h : z \mapsto d(x, z)$ is Lipschitz with constant 1 and satisfies $0 = h(x) \neq h(y)$. Therefore H is dense in $C^{\mathbb{R}}(X)$.

We now deduce from Theorem 2.2 another denseness criterion, where the assumption that H is a lattice is replaced by an assumption of closedness under multiplication. More precisely, we assume we have a vector subspace H of $C(X)$ that is a **subalgebra** of $C(X)$; this means that $fg \in H$ for $f, g \in H$. Since $fg = \frac{1}{2}\big((f+g)^2 - f^2 - g^2\big)$, this condition is equivalent to H being a vector space such that the square of every element of H is in H.

Theorem 2.3 (Stone–Weierstrass Theorem, real case) *Every separating subalgebra of $C^{\mathbb{R}}(X)$ containing the constant functions is dense in $C^{\mathbb{R}}(X)$.*

Proof. If H is a separating subalgebra of $C^{\mathbb{R}}(X)$ containing the constants, so is its closure $\bar{H}$. Therefore it suffices to show that $\bar{H}$ is a lattice and to apply Theorem 2.2. Thus, let f be a nonzero element of $\bar{H}$. We saw in the example on page 29 that there exists a sequence (P_n) of polynomials over $\mathbb{R}$ that converges uniformly on $[-1,1]$ to the function $x \mapsto |x|$. But then the sequence of functions $\big(P_n(f/\|f\|)\big)$ converges uniformly to $|f|/\|f\|$, so $|f|$ is the uniform limit of the sequence $\big(\|f\|\, P_n(f/\|f\|)\big)$. Since $\bar{H}$ is a subalgebra of $C^{\mathbb{R}}(X)$, all terms in this sequence are in $\bar{H}$; therefore so is their uniform limit $|f|$. This shows that $\bar{H}$ is a lattice. □

Examples

1. The set of Lipschitz functions from X to $\mathbb{R}$ satisfies the assumptions of Theorem 2.3.
2. Suppose X is a compact subset of $\mathbb{R}^d$, and let H be the set of polynomial functions (in d variables) from X to $\mathbb{R}$:

$$H = \big\{x \mapsto P(x) : P \in \mathbb{R}[X_1, \ldots, X_d]\big\}.$$

   Clearly, H is a subalgebra of $C^{\mathbb{R}}(X)$ containing the constants; on the other hand, if x and y are distinct points in X, they differ in at least one component: for example, $x_j \neq y_j$. But then the polynomial X_j takes different values at x and at y. Thus H is separating and hence dense in $C^{\mathbb{R}}(X)$.
   In the particular case where $d = 1$ and X is a compact interval $[a,b]$ in $\mathbb{R}$, this result is known as **Weierstrass's Theorem**. In fact, there are several explicit methods to associate to an element $f \in C^{\mathbb{R}}\big([a,b]\big)$ a sequence of polynomials (P_n) that converges uniformly to f on $[a,b]$; see, for example, Exercises 3 and 2 below.
   Note that, as a consequence of Weierstrass's Theorem, the set of monomials $\{1, x, x^2, \ldots, x^n, \ldots\}$, considered as functions on $[a,b]$ (for $a < b$) forms a topological basis of $C\big([a,b]\big)$. (We thus recover, in particular, the fact that $C^{\mathbb{R}}\big([a,b]\big)$ is separable.)

Remark. In the preceding theorem, one cannot replace $C^{\mathbb{R}}(X)$ by $C^{\mathbb{C}}(X)$, as the following example shows. Set $\mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$, and let H be the set of polynomial functions from $\mathbb{U}$ in $\mathbb{C}$:

$$H = \{z \mapsto P(z) : P \in \mathbb{C}[X]\}.$$

H is certainly a separating subalgebra of $C^{\mathbb{C}}(\mathbb{U})$ (the function $Z : z \mapsto z$ is an element of H, and $Z(z) \neq Z(z')$ if $z \neq z'$), and it contains the constants. But H is not dense in $C^{\mathbb{C}}(\mathbb{U})$. Indeed, since

$$\int_0^{2\pi} e^{in\theta} e^{i\theta} \, d\theta = 0$$

for every $n \in \mathbb{N}$, we get

$$\int_0^{2\pi} h(e^{i\theta}) e^{i\theta} \, d\theta = 0$$

for all $h \in H$. By taking uniform limits, we conclude that the same equality holds for $h \in \bar{H}$. On the other hand, the function $\bar{Z} : z \mapsto \bar{z}$ is an element of $C^{\mathbb{C}}(\mathbb{U})$, yet

$$\int_0^{2\pi} \bar{Z}(e^{i\theta}) e^{i\theta} \, d\theta = 2\pi.$$

Thus $\bar{Z} \notin \bar{H}$, and H is not dense in $C^{\mathbb{C}}(\mathbb{U})$.

Thus, in the complex case an additional assumption is necessary. We will suppose in this case that the subset H of $C^{\mathbb{C}}(X)$ is **self-conjugate**; this means that $h \in H$ implies $\bar{h} \in H$, where the conjugate $\bar{h}$ of h is defined by $\bar{h}(x) = \overline{h(x)}$.

Theorem 2.4 (Stone–Weierstrass Theorem, complex case) *Every separating subalgebra H of $C^{\mathbb{C}}(X)$ that is self-conjugate and contains the constant functions is dense in $C^{\mathbb{C}}(X)$.*

Proof. Set $H_{\mathbb{R}} = \{h \in H : h(x) \in \mathbb{R} \text{ for all } x \in X\}$. Clearly, $H_{\mathbb{R}}$ is a subalgebra of $C^{\mathbb{R}}(X)$ containing the constants. Now, if $f \in H$, the real and imaginary parts of f lie in $H_{\mathbb{R}}$, since H is self-conjugate and $\operatorname{Re} f = (f + \bar{f})/2$, $\operatorname{Im} f = (f - \bar{f})/(2i)$. If x_1 and x_2 are distinct points in X, there exists by assumption $h \in H$ such that $h(x_1) \neq h(x_2)$. Therefore there exists $g \in H_{\mathbb{R}}$ such that $g(x_1) \neq g(x_2)$: just take $g = \operatorname{Re} h$ or $g = \operatorname{Im} h$ as needed. It follows that $H_{\mathbb{R}}$ is separating, hence dense in $C^{\mathbb{R}}(X)$, by Theorem 2.3. Since $C^{\mathbb{C}}(X) = C^{\mathbb{R}}(X) + iC^{\mathbb{R}}(X)$ and H contains $H_{\mathbb{R}} + iH_{\mathbb{R}}$, the proof is complete. □

Examples

1. The set of Lipschitz functions from X to $\mathbb{C}$ is dense in $C^{\mathbb{C}}(X)$.
2. If X is compact in $\mathbb{R}^d$, the set of functions from X to $\mathbb{C}$ defined by complex polynomials in d variables is dense in $C^{\mathbb{C}}(X)$. In particular, if $[a, b]$ (with $a < b$) is a compact interval in $\mathbb{R}$, the set of restrictions to $[a, b]$ of the monomials $1, x, x^2, \ldots, x^n, \ldots$ forms a topological basis of $C^{\mathbb{C}}([a, b])$.

3. If X is a compact set in $\mathbb{C}^d$, the set H defined by

$$H = \{z \in X \mapsto P(z, \bar{z}) : P \in \mathbb{C}[X_1, \ldots, X_d, Y_1, \ldots, Y_d]\}$$

is dense in $C^{\mathbb{C}}(X)$.

In the particular case where $d = 1$ and $X = \mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$, we see that H is the vector space generated by the functions $Z^p : z \mapsto z^p$, with $p \in \mathbb{Z}$. Indeed, if $z \in \mathbb{U}$ we have $\bar{z} = z^{-1}$. Thus, the family $(Z^p)_{p \in \mathbb{Z}}$ (which is clearly free) is a topological basis of $C^{\mathbb{C}}(\mathbb{U})$.

4. Let $C^{\mathbb{K}}_{2\pi}$ be the set of continuous functions from $\mathbb{R}$ to $\mathbb{K}$ that are periodic of period 2π, with the uniform norm on $\mathbb{R}$, namely,

$$\|f\| = \max_{x \in \mathbb{R}} |f(x)| = \max_{x \in [0, 2\pi]} |f(x)|.$$

Lemma 2.5 *The map from $C^{\mathbb{C}}(\mathbb{U})$ to $C^{\mathbb{C}}_{2\pi}$ that associates to $\varphi \in C^{\mathbb{C}}(\mathbb{U})$ the function f given by $f(\theta) = \varphi(e^{i\theta})$ for every real θ is a surjective isometry.*

Proof. Only the surjectivity requires proof. For $z \in \mathbb{U}$, denote by $\arg z$ some real number such that $e^{i \arg z} = z$. We know that $\arg z$ is defined modulo 2π and that there exist choices of $\arg z$ that vary continuously in the neighborhood of a given point (for example, if $z_0 \in \mathbb{U}$ and $z \in \mathbb{U}$ with $|z - z_0| < 1$, we can take $\arg z = \arg z_0 +$ Arccos $\operatorname{Re}(z/z_0))$. Thus, if $f \in C^{\mathbb{C}}_{2\pi}$, the function φ defined by $\varphi(z) = f(\arg z)$ is well-defined and continuous in $\mathbb{U}$, and $f(\theta) = \varphi(e^{i\theta})$ for all $\theta \in \mathbb{R}$. □

It follows from the preceding example that the family $(e_n)_{n \in \mathbb{Z}}$ of elements of $C^{\mathbb{C}}_{2\pi}$ defined by $e_n(\theta) = e^{in\theta}$ is a topological basis of $C^{\mathbb{C}}_{2\pi}$. By taking the real and imaginary parts of the functions e_n, we deduce that the set $B = \{1\} \cup \{c_n, s_n\}_{n \in \mathbb{N}}$, with $c_n(x) = \cos nx$ and $s_n(x) = \sin nx$, forms a topological basis of $C^{\mathbb{R}}_{2\pi}$, and thus also of $C^{\mathbb{C}}_{2\pi}$. A linear combination of functions of B is called a **trigonometric polynomial**.

Note that one can explicitly determine a sequence of trigonometric polynomials that converges toward a given function $f \in C^{\mathbb{R}}_{2\pi}$ (see Exercise 2 below).

5. Let X and Y be compact metric spaces. We denote by $C(X) \otimes C(Y)$ the vector subspace of $C(X \times Y)$ generated by the functions $f \otimes g : (x, y) \mapsto f(x)g(y)$ with $f \in C(X)$ and $g \in C(Y)$. It is clear that $C(X) \otimes C(Y)$ is a subalgebra of $C(X \times Y)$ containing the constants and, when $\mathbb{K} = \mathbb{C}$, self-conjugate. It is also separating: if $(x_1, y_1) \neq (x_2, y_2)$ we have, say, $x_1 \neq x_2$, and then the function $d(\cdot, x_1) \otimes 1 : (x, y) \mapsto d(x, x_1)$ (where d is the metric on X) is an element of $C(X) \otimes C(Y)$ separating (x_1, y_1) and (x_2, y_2). Thus $C(X) \otimes C(Y)$ is dense in $C(X \times Y)$.

Exercises

1. Let D be a dense subset of $C^{\mathbb{R}}(X)$. Prove that, for all $f \in C^{\mathbb{R}}(X)$, there exists an increasing sequence of elements of D that converges uniformly to f.
Hint. For each positive integer n, prove that there is an element f_n of D such that $f - 2^{-n} \le f_n \le f - 2^{-n-1}$.

2. *Dirac sequences*

a. Let $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence of continuous functions from $\mathbb{R}^m$ to $\mathbb{R}$, with nonnegative values, and satisfying these properties:

- $\int_{\mathbb{R}^m} \varphi_n(x)\,dx = 1$ for every integer n.
- For every $\varepsilon > 0$, $\lim_{n\to+\infty} \int_{|x|\ge\varepsilon} \varphi_n(x)\,dx = 0$, where $|\cdot|$ denotes a norm on $\mathbb{R}^n$.

Let f be a bounded, continuous function on $\mathbb{R}^m$. Prove that the sequence $(\varphi_n * f)$ converges to f uniformly on every compact subset of $\mathbb{R}^m$. Recall that $\varphi_n * f$ is defined by

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

b. For each $n \in \mathbb{N}$, set $c_n = \int_{-1}^{1} (1-x^2)^n\,dx$, and let φ_n be the function from $\mathbb{R}$ to $\mathbb{R}$ defined by

$$\varphi_n(x) = \begin{cases} (1-x^2)^n/c_n & \text{if } |x| \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

i. Prove that the sequence (φ_n) satisfies the hypotheses of part a.

ii. Deduce that every continuous function on $[0,1]$ is the uniform limit on $[0,1]$ of a sequence of polynomial functions.
Hint. Deal first with the case of a function f satisfying $f(0) = f(1) = 0$, by showing that, if $\tilde{f}$ is the extension of f having the value 0 outside $[0,1]$, then $\varphi_n * \tilde{f}$ coincides in $[0,1]$ with a polynomial function.

c. *Fejér's Theorem.* Let f be a continuous function from $\mathbb{R}$ to $\mathbb{C}$, periodic of period 2π. Let D_n and K_n be the functions defined by

$$D_n(x) = \sum_{k=-n}^{n} e^{ikx}, \quad K_m(x) = \frac{1}{m} \sum_{n=0}^{m-1} D_n(x).$$

If $h, g \in C_{2\pi}^{\mathbb{C}}$, we write

$$h * g(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(x-y) g(y)\,dy$$

(this equals $g * h(x)$).

i. Prove that $K_n(2k\pi) = n$ for $k \in \mathbb{Z}$ and that, for all $x \notin 2\pi\mathbb{Z}$,

$$K_n(x) = \frac{1 - \cos nx}{n(1 - \cos x)}.$$

Show also that $\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(x)\,dx = 1$ and that, for all $\varepsilon \in (0, \pi)$,

$$\lim_{n\to+\infty} \int_{\varepsilon}^{\pi} K_n(x)\,dx = 0.$$

ii. Prove that the sequence of functions $(K_n * f)$ converges uniformly to f on $\mathbb{R}$.

iii. Express $D_n * f$, then $K_n * f$, in terms of the partial sums S_n of the Fourier series of f, which, as we recall, are given by

$$S_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}, \quad \text{with } c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) e^{-ikt}\,dt.$$

iv. Deduce that every continuous function periodic of period 2π is the uniform limit of a sequence of trigonometric polynomials.

3. *Another demonstration of Weierstrass's Theorem: Bernstein polynomials.* The functions in this exercise are real-valued ($\mathbb{K} = \mathbb{R}$).

a. *Korovkin's Theorem.* For $i \in \mathbb{N}$, we denote by X^i the element of $C([0,1])$ defined by $X^i(x) = x^i$. We also set $1 = X^0$ and $X = X^1$. Let (T_n) be a sequence of positive elements in $L\big(C([0,1])\big)$ (positivity here means that $f \geq 0$ implies $T_n(f) \geq 0$, or again that $f \leq g$ implies $T_n(f) \leq T_n(g)$). Assume that, for $i = 0, 1, 2$, the sequence of functions $(T_n(X^i))_{n\in\mathbb{N}}$ converges to X^i uniformly on $[0,1]$. We want to show that, for all $f \in C([0,1])$, the sequence $(T_n f)$ converges uniformly to f on $[0,1]$.

i. Let f be a continuous function on $[0,1]$. Define the *modulus of uniform continuity* of f as the function $\omega_f : \mathbb{R}^{+*} \to \mathbb{R}^+$ whose value at $\eta > 0$ is

$$\omega_f(\eta) = \sup_{\substack{(x,y)\in[0,1]^2 \\ |x-y|\leq\eta}} \big|f(x) - f(y)\big|.$$

Check that $\omega_f(\eta)$ is well-defined for all $\eta > 0$, and that $\omega_f(\eta)$ tends to 0 as η tends to 0. Now fix $\eta > 0$.

ii. Prove that, for all $x, y \in [0,1]$,

$$\big|f(x) - f(y)\big| \leq \omega_f(\eta) + 2(x-y)^2 \frac{\|f\|}{\eta^2}.$$

(One can deal separately with the cases $|x-y| \leq \eta$ and $|x-y| > \eta$.)

iii. If $x, y \in [0,1]$, set $g_y(x) = (x-y)^2$. Prove that, if $x, y \in [0,1]$, we have, for every $n \in \mathbb{N}$,

$$\left|(T_n f)(x) - f(y)T_n(1)(x)\right| \leq \omega_f(\eta)T_n(1)(x) + 2\frac{\|f\|}{\eta^2}(T_n g_y)(x).$$

iv. Set $h_n(x) = (T_n g_x)(x)$. Prove that the sequence of functions (h_n) converges uniformly to 0 in $[0,1]$.
Hint. $T_n g_x(x) = (T_n X^2 - 2XT_n X + X^2 T_n 1)(x)$.

v. Deduce that $\limsup_{n\to+\infty} \|T_n f - f\| \leq \omega_f(\eta)$. Wrap up the proof.

b. Let f be a function from $[0,1]$ to $\mathbb{R}$. For every integer $n \geq 1$, define the polynomial $B_n(f)$ by

$$B_n(f)(x) = \sum_{k=0}^{n} C_n^k f\left(\frac{k}{n}\right)x^k(1-x)^{n-k}.$$

i. Prove that

$$B_n(Xf) = XB_n(f) + \frac{X(1-X)}{n}B_n'(f),$$

where $B_n'(f)$ represents the derivative of the polynomial $B_n(f)$.

ii. Compute $B_n(1)$, $B_n(X)$, and $B_n(X^2)$ for every $n \in \mathbb{N}$.

iii. Prove that, for every $f \in C([0,1])$, the sequence $(B_n(f))$ converges uniformly to f.

4. *Another proof of Fejér's Theorem*

a. Let (T_n) be a sequence of positive elements of $L(C_{2\pi}^{\mathbb{R}})$ (see Exercise 3a for the definition of positivity) such that the sequence of functions $(T_n(f))_{n\in\mathbb{N}}$ converges to f uniformly on $\mathbb{R}$ when f is each of the three functions $x \mapsto 1$, $x \mapsto \cos x$, and $x \mapsto \sin x$. Prove that, for all $f \in C_{2\pi}^{\mathbb{R}}$, the sequence $(T_n f)$ converges uniformly to f.
Hint. Argue as in Exercise 3a, considering the interval $[-\pi, \pi]$ and replacing $(x-y)^2/\eta^2$ by $(1-\cos(x-y))/(1-\cos\eta)$.

b. Let (K_n) be the sequence of functions defined in Exercise 2c. Take $f \in C_{2\pi}^{\mathbb{R}}$. Derive from the preceding question another proof that the sequence $(K_n * f)$ converges uniformly to f on $\mathbb{R}$.

5. Let X be a compact interval in $\mathbb{R}$ and let H be the set of elements of $C(X)$ defined by polynomial functions with integer coefficients.

a. Prove that, if X and $\mathbb{Z}$ intersect, H is not dense in $C(X)$.

From now on in this exercise we assume that $X \subset (0,1)$. We denote by (p_n) the strictly increasing sequence of prime numbers and by (P_n) the sequence of elements of $C(X)$ defined by

$$P_n(x) = 1 - x^{p_n} - (1-x)^{p_n}.$$

b. Prove that, for every integer n, the function P_n/p_n is an element of H, and that $1/P_n$ belongs to $\bar{H}$.

c. Prove that, for every $k \in \mathbb{Z}^*$, the constant function $x \mapsto 1/k$ is an element of $\bar{H}$. (You might start with the case of k prime.)

d. Deduce that H is dense in $C(X)$.

6. *Equidistributed sequences and Weyl's Criterion*

a. Let E be the vector space generated by the functions from $[0,1]$ to $\mathbb{C}$ of the form $1_{[a,b]}$. Prove that every continuous function from $[0,1]$ to $\mathbb{C}$ is the uniform limit of functions in E.

b. A sequence $(u_p)_{p\in\mathbb{N}}$ of points in $[0,1]$ is called *equidistributed* if, for every $[a,b] \subset [0,1]$,

$$\lim_{n\to+\infty} \frac{\operatorname{Card}\{p \le n : u_p \in [a,b]\}}{n+1} = b - a.$$

Prove that, if $(u_p)_{p\in\mathbb{N}}$ is an equidistributed sequence of points in $[0,1]$ and $f : \mathbb{R} \to \mathbb{C}$ is any continuous function periodic of period 1, then

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{p=0}^{n} f(u_p) = \int_0^1 f(t)\,dt.$$

Hint. Check that this is true if $f \in E$, then use denseness (compare Proposition 4.3 on page 19).

c. Prove the converse.

Hint. One might start by showing that, if $[a,b] \subset [0,1]$ and $\varepsilon > 0$, there exist continuous functions f and g from $[0,1]$ to $\mathbb{R}$ such that $f(0) = f(1)$, $g(0) = g(1)$, $f \le 1_{[a,b]} \le g$ and

$$\int_0^1 \big(g(t) - f(t)\big)\,dt \le \varepsilon.$$

d. Deduce that a sequence $(u_p)_{p\in\mathbb{N}}$ of points in $[0,1]$ is equidistributed if and only if, for every $\lambda \in \mathbb{N}^*$, the *Weyl criterion* is satisfied:

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{p=0}^{n} e^{2i\pi\lambda u_p} = 0.$$

e. *Example.* Take $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and, for every $p \in \mathbb{N}$, set $u_p = \{p\alpha\} = p\alpha - \lfloor p\alpha \rfloor$, where $\lfloor p\alpha \rfloor$ denotes the integer part of $p\alpha$. Prove that the sequence (u_p) is equidistributed.

f. Same question with the sequence (u_p) defined by $u_p = \{p^\alpha\}$, where $\alpha \in (0,1)$.

Hint. Consider $I_n = \int_1^n e^{2i\pi\lambda x^\alpha}\,dx$, for λ a fixed positive integer. Prove, by change of variables and integration by parts, that $I_n = O(n^{1-\alpha})$. Next show that

$$\left| I_n - \sum_{p=0}^{n} e^{2i\pi\lambda p^\alpha} \right| = O(n^\alpha).$$

7. *Particular cases of the Tietze Extension Theorem*

a. Let Y be a metric space and X a nonempty compact subset of Y. Denote by $C_b(Y)$ the vector space consisting of continuous, bounded functions from Y to $\mathbb{K}$, with the norm defined by

$$\|f\| = \sup_{y\in Y} |f(y)|.$$

On $C(X)$ we take the uniform norm, also denoted by $\|\cdot\|$. Now consider the linear map $\Phi : C_b(Y) \to C(X)$ defined by restriction to X: $\Phi(f) = f_{|X}$ for every $f \in C_b(Y)$.

i. Prove that $C_b(Y)$ is a Banach space.

ii. Prove that, if $f \in C_b(Y)$, there exists $\tilde{f} \in C_b(Y)$ such that $\Phi(\tilde{f}) = \Phi(f)$ and $\|\tilde{f}\| = \|\Phi(f)\|$.
Hint. If $\Phi(f) \neq 0$, one can choose

$$\tilde{f} = \chi\left(\frac{f}{\|\Phi(f)\|}\right)\|\Phi(f)\|,$$

where $\chi : \mathbb{K} \to \mathbb{K}$ is defined by $\chi(x) = x/\max(|x|, 1)$.

iii. Prove that $\operatorname{im}\Phi$ is dense in $C(X)$.
Hint. Use the Stone–Weierstrass Theorem.

iv. Let g be an element of $C(X)$ that is the uniform limit of a sequence $(\Phi(f_n))$.

A. Prove that one can assume, after passing to a subsequence if necessary, that $\|\Phi(f_{n+1}) - \Phi(f_n)\| \le 2^{-n}$ for every n.

B. For $n \in \mathbb{N}$, choose $h_n \in C_b(Y)$ such that $\Phi(h_n) = \Phi(f_n - f_{n-1})$ and $\|h_n\| = \|\Phi(f_n - f_{n-1})\|$ (where $f_{-1} = 0$ by convention). The existence of the h_n was proved in ii above. Prove that the series $\sum_{n=0}^{\infty} h_n$ converges in $C_b(Y)$. Denote its sum by h.

C. Prove that $\Phi(h) = g$.

v. Deduce from the preceding facts that every function $g \in C(X)$ can be extended to a function $f \in C_b(Y)$ such that $\|f\| = \|g\|$.

b. Let (Y, d) be a metric space and let A be a nonempty subset of Y. Let f be a Lipschitz function from A to $\mathbb{R}$, with Lipschitz constant C. Set

$$g(y) = \inf_{x\in A}\big(f(x) + Cd(x, y)\big) \quad \text{for all } y \in Y.$$

Prove that g is a Lipschitz extension of f, also with constant C.

8. *Stone–Weierstrass Theorem in* $\mathbb{R}$. We denote by $C_0^{\mathbb{K}}(\mathbb{R})$ (or $C_0(\mathbb{R})$) the space of continuous functions f from $\mathbb{R}$ to $\mathbb{K}$ such that

$$\lim_{x\to-\infty} f(x) = \lim_{x\to+\infty} f(x) = 0.$$

We give this space the uniform norm: $\|f\| = \sup_{x\in\mathbb{R}} |f(x)|$. We again denote by $\mathbb{U}$ the set of complex numbers of absolute value 1, which is compact in the metric induced from $\mathbb{C}$.

a. **i.** Prove that $C_0(\mathbb{R})$ is a Banach space.

ii. Define a map φ from $\mathbb{R}$ onto $\mathbb{U}\setminus\{-1\}$ by setting $\varphi(x) = e^{2i \operatorname{Arctan} x}$. Prove that φ is a homeomorphism between $\mathbb{R}$ and $\mathbb{U}\setminus\{-1\}$, the inverse homeomorphism being $\psi(z) = \tan(\frac{1}{2} \operatorname{Arg} z)$, where $\operatorname{Arg} z$ denotes the argument of z in the interval $(-\pi, \pi)$. Check that $\lim_{x\to+\infty} \varphi(x) = \lim_{x\to-\infty} \varphi(x) = -1$.

iii. Prove that a function f on $\mathbb{R}$ belongs to $C_0(\mathbb{R})$ if and only if the function $\tilde{f}$ defined on $\mathbb{U}$ by

$$\tilde{f}(z) = \begin{cases} f(\psi(z)) & \text{if } z \neq -1, \\ 0 & \text{if } z = -1, \end{cases}$$

belongs to $C(\mathbb{U})$. Prove that the map $f \mapsto \tilde{f}$ defines an isometry between $C_0(\mathbb{R})$ and the set of elements of $C(\mathbb{U})$ that vanish at -1.

b. **i.** Let H be a vector subspace of $C_0(\mathbb{R})$ satisfying these conditions:

A. $f^2 \in H$ for all $f \in H$.

B. If x and y are distinct points of $\mathbb{R}$, there exists $f \in H$ such that $f(x) \neq f(y)$.

C. For any $x \in \mathbb{R}$, there exists $f \in H$ such that $f(x) \neq 0$.

D. In the complex case, H is self-conjugate (that is, $f \in H$ implies $\bar{f} \in H$).

Prove that H is dense in $C_0(\mathbb{R})$.

Hint. Apply Stone–Weierstrass to the compact space $\mathbb{U}$ and to the set $\tilde{H}$ consisting of functions of the form $\tilde{f} + a$, with $f \in H$ and $a \in \mathbb{K}$.

ii. Conversely, prove that every dense subset H of $C_0(\mathbb{R})$ satisfies conditions B and C above.

c. If $a \in \mathbb{C} \setminus \mathbb{R}$, we set $\varphi_a(x) = (a + x)^{-1}$. Prove that the family $\{\varphi_a\}_{a\in\mathbb{C}\setminus\mathbb{R}}$ is fundamental in $C_0^{\mathbb{C}}(\mathbb{R})$.

Hint. Prove first that $\varphi_a^2 = \lim_{h\to 0}(\varphi_a - \varphi_{a+h})/h$ in the sense of convergence in $C_0(\mathbb{R})$. Deduce that the closure of the vector space generated by the φ_a satisfies conditions A–D of part b above.

d. Let H be the set of functions from $\mathbb{R}$ to $\mathbb{R}$ of the form $x \mapsto e^{-x^2} P(x)$, with $P \in \mathbb{R}[X]$.

i. Take $r \in \mathbb{N}$ and $a \in (0, 1)$. For $n \in \mathbb{N}$ and $x \in \mathbb{R}$, set $R_n(x) = e^{-x^2} x^{2n+r} a^n / n!$. Prove that the sequence of functions (R_n) converges uniformly on $\mathbb{R}$ to the zero function.

Hint. Prove that if $u_n = \sup_{x\in\mathbb{R}} \lvert R_n(x)\rvert$, then

$$\lim_{n\to\infty} (u_{n+1}/u_n) = a.$$

ii. Deduce that the function $f_{a,r} : x \mapsto e^{-(1+a)x^2} x^r$ belongs to $\bar{H}$.

Hint. One can use Taylor's formula with integral remainder to approximate e^{-ax^2} by polynomials.

iii. Prove that the function $g_{a,r} : x \mapsto e^{-(1+a)^2x^2}x^r$ belongs to $\bar{H}$.
Hint. Write $g_{a,r} = (1+a)^{-r/2} f_{a,r}\big(\sqrt{1+a}\,x\big)$ and use part ii twice.

iv. Applying the facts above to $a = \sqrt{2} - 1$, show that H is dense in $C_0^{\mathbb{R}}(\mathbb{R})$.

e. Denote by $C_c^{\mathbb{K}}(\mathbb{R})$ or $C_c(\mathbb{R})$ the set of continuous functions f from $\mathbb{R}$ to $\mathbb{K}$ that vanish outside a compact interval in $\mathbb{R}$ (that depends on f). We assume in the sequel that $\mathbb{K} = \mathbb{C}$.

i. Prove that $C_c(\mathbb{R})$ is dense in $C_0(\mathbb{R})$. (Use part b above or give a direct proof.)

ii. For $\varphi \in C_c(\mathbb{R})$, set

$$\hat{\varphi}(x) = \int_{\mathbb{R}} e^{ixy}\varphi(y)\,dy.$$

Prove that $\hat{\varphi} \in C_0(\mathbb{R})$.
Hint. Show first that, if $a < b$, the function $x \mapsto \int_a^b e^{ixy}\,dy$ lies in $C_0(\mathbb{R})$. Then approximate φ by staircase functions.

iii. If $\varphi, \psi \in C_c(\mathbb{R})$, define

$$(\varphi * \psi)(x) = \int_{\mathbb{R}} \varphi(x-y)\psi(y)dy.$$

Prove that $\varphi * \psi \in C_c(\mathbb{R})$ and that $\widehat{\varphi * \psi} = \hat{\varphi}\,\hat{\psi}$.

iv. Deduce that the set $\{\hat{\varphi}\}_{\varphi \in C_c(\mathbb{R})}$ is dense in $C_0(\mathbb{R})$.
Hint. To check conditions B and C, one can compute the integral $\int_0^{+\infty} e^{ixy}e^{-y}\,dy$ and approximate the function

$$y \mapsto 1_{(0,+\infty)}(y)e^{-y}$$

in $L^1(dy)$ by functions in $C_c(\mathbb{R})$.

3 Ascoli's Theorem

In this section we present a criterion of relative compactness in $C(X)$.

Let x_0 be a point of X. A subset H of $C(X)$ is called **equicontinuous at** x_0 if, for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\big|h(x) - h(x_0)\big| < \varepsilon \quad \text{for all } h \in H \text{ and all } x \in X \text{ with } d(x, x_0) < \eta.$$

H is called **equicontinuous** if it is equicontinuous at every point of X. It is called **uniformly equicontinuous** if, for all $\varepsilon > 0$, there exists $\eta > 0$ such that

$$\big|h(x) - h(y)\big| < \varepsilon \quad \text{for all } h \in H \text{ and all } x, y \in X \text{ with } d(x, y) < \eta.$$

Since X has been assumed compact, these two notions are in fact equivalent:

Proposition 3.1 *A subset of $C(X)$ is equicontinuous if and only if it is uniformly equicontinuous.*

Proof. It is enough to show necessity. Let H be an equicontinuous subset of $C(X)$, and let $\varepsilon > 0$ be a real number. By assumption, for every $x \in X$ there exists $\eta_x > 0$ such that $\left|h(y) - h(x)\right| < \varepsilon/2$ whenever $h \in H$ and $d(x, y) < \eta_x$. By the Borel–Lebesgue property, we can choose finitely many points $x_1, \dots, x_r$ such that the balls $B(x_j, \eta_{x_j}/2)$ cover X. Now let η be the smallest of the $\eta_{x_j}/2$, and let x and y be points in X such that $d(x, y) < \eta$. Choosing j such that $x \in B(x_j, \eta_{x_j}/2)$, we see that $x, y \in B(x_j, \eta_{x_j})$, so

$$\left|h(y) - h(x)\right| \leq \left|h(y) - h(x_j)\right| + \left|h(x) - h(x_j)\right| < \varepsilon \quad \text{for all } h \in H. \qquad \square$$

Examples

1. Every finite subset of $C(X)$ is equicontinuous.
2. Every subset of an equicontinuous set is equicontinuous.
3. A finite union of equicontinuous sets is equicontinuous.
4. Any uniformly convergent sequence of functions in $C(X)$ consitutes an equicontinuous set (exercise).
5. If C is a positive real number, the set of C-Lipschitz functions from X to $\mathbb{K}$ is equicontinuous.

Proposition 3.2 *Let (f_n) be an equicontinuous sequence in $C(X)$ and let D be a dense subset of X. If, for all $x \in D$, the sequence of numbers $\big(f_n(x)\big)$ converges, the sequence of functions (f_n) converges uniformly to a function $f \in C(X)$.*

(Compare this result with Proposition 4.1 on page 18.)

Proof. It suffices to show that (f_n) is a Cauchy sequence in $C(X)$. To do this, take $\varepsilon > 0$. By assumption, there exists $\eta > 0$ such that, whenever $d(x, y) < \eta$,

$$\left|f_n(x) - f_n(y)\right| < \varepsilon/5 \quad \text{for all } n \in \mathbb{N}.$$

Since X is precompact, il can be covered by finitely many balls of radius η: $X = \bigcup_{j=0}^{r} B(x_j, \eta)$. Since D is dense, each ball $B(x_j, \eta)$ contains at least one point y_j from D. Since, by assumption, the sequences $(f_n(y_j))_{n \in \mathbb{N}}$ are Cauchy sequences, there exists a positive integer N such that, for any integer $j \leq r$,

$$\left|f_n(y_j) - f_p(y_j)\right| < \varepsilon/5 \quad \text{for all } n, p \geq N.$$

Now let x be a point in X, and let j be an integer such that $x \in B(x_j, \eta)$. Then, for $n, p \geq N$,

$$\begin{aligned} \left|f_n(x) - f_p(x)\right| \leq \left|f_n(x) - f_n(x_j)\right| + \left|f_n(y_j) - f_n(x_j)\right| + \left|f_n(y_j) - f_p(y_j)\right| \\ + \left|f_p(y_j) - f_p(x_j)\right| + \left|f_p(x) - f_p(x_j)\right| < \varepsilon. \end{aligned}$$

Thus, $\|f_n - f_p\| < \varepsilon$ for all $n, p \geq N$, and (f_n) is a Cauchy sequence. $\square$

We deduce from this our main result:

Theorem 3.3 (Ascoli) *A subset of $C(X)$ is relatively compact in $C(X)$ if and only if it is bounded and equicontinuous.*

Proof. For the "only if" part, let H be a relatively compact subset of $C(X)$. Then H is certainly bounded; this is true in any metric space. We must show it is equicontinuous. Fix $\varepsilon > 0$. Since H is precompact, we can choose finitely many elements $f_0, \dots, f_r$ in H such that the balls $B(f_j, \varepsilon/3)$ cover H. Since the finite family $(f_j)_{j\le r}$ is uniformly equicontinuous, there exists $\eta > 0$ such that $\left|f_j(x) - f_j(y)\right| < \varepsilon/3$ for all $j \le r$, whenever $d(x,y) < \eta$. It follows that, if $f \in H$ and $d(x,y) < \eta$, then

$$\left|f(x) - f(y)\right| \le \left|f(x) - f_j(x)\right| + \left|f_j(x) - f_j(y)\right| + \left|f_j(y) - f(y)\right| < \varepsilon,$$

where j is chosen so that $\|f - f_j\| < \varepsilon/3$. This shows that H is equicontinuous.

For the converse, suppose H is bounded and equicontinuous. X is compact, hence separable. Thus it contains a countable dense subset D. Let (f_n) be a sequence in H. For every point x in D, the sequence of numbers $(f_n(x))_{n\in\mathbb{N}}$ is bounded by $\sup_{h\in H}\|h\|$; thus, by Theorem 3.1 on page 12, there exists a subsequence $\left(f_{n_k}\right)_{k\in\mathbb{N}}$ such that $\left(f_{n_k}(x)\right)_{k\in\mathbb{N}}$ converges for all $x \in D$. By Proposition 3.2, we deduce that the sequence $(f_{n_k})_{k\in\mathbb{N}}$ converges in $C(X)$. □

Remark. The preceding proof also shows that, if H is an equicontinuous subset of $C(X)$, the following properties are equivalent:

- H is bounded.
- There is a dense subset D of X such that, for all $x \in D$, the set $\{f(x)\}_{f\in H}$ is a bounded subset of $\mathbb{K}$.

(This equivalence can also be proved directly.)

Example. Consider compact metric spaces X and Y, an element K of $C(X \times Y)$, and a Borel measure μ on Y having finite mass ($\mu(Y) < +\infty$). We define a linear operator T from $C(Y)$ to $C(X)$ by setting

$$Tf(x) = \int_Y K(x,y)f(y)\,d\mu(y) \quad \text{for all } f \in C(Y) \text{ and } x \in X.$$

Recall that $\bar{B}\big(C(Y)\big)$ denotes the closed unit ball in $C(Y)$:

$$\bar{B}\big(C(Y)\big) = \big\{f \in C(Y) : \|f\| \le 1\big\}.$$

Proposition 3.4 *The image under T of the closed unit ball of $C(Y)$ is a relatively compact subset of $C(X)$.*

We say that T is a **compact operator** from $C(Y)$ to $C(X)$ (see Chapter 6).

Proof. It is clear that $T\big(\bar{B}(C(Y))\big)$ is bounded by

$$M = \mu(Y) \max_{(x,y)\in X\times Y} \big|K(x,y)\big|.$$

On the other hand, K is uniformly continuous on $X \times Y$; in particular, for all ε, there exists $\eta > 0$ such that

$$\big|K(x_1,y) - K(x_2,y)\big| < \varepsilon \quad \text{for all } y \in Y \text{ and } x_1, x_2 \in X \text{ with } d(x_1,x_2) < \eta.$$

Thus, for all $f \in \bar{B}(C(Y))$, we have $\big|Tf(x_1) - Tf(x_2)\big| \leq \mu(Y)\varepsilon$. Therefore the subset $T\big(\bar{B}(C(Y))\big)$ of $C(X)$ is equicontinuous, and we can apply Ascoli's Theorem. □

Exercises

1. For each $n \in \mathbb{N}$, let f_n be the function from $[0,1]$ to $\mathbb{R}$ defined by $f_n(x) = x^n$. At what points in the interval $[0,1]$ is the family $\{f_n\}_{n\in\mathbb{N}}$ equicontinuous?
2. **a.** Let X be a metric space and (f_n) a sequence in $C(X)$. Prove that, if $\{f_n\}_{n\in\mathbb{N}}$ is equicontinuous at a point x of X, for any sequence (x_n) of X that converges to x the sequence $\big(f_n(x) - f_n(x_n)\big)$ converges to 0.

   b. Set $f_n(x) = \sin nx$. Prove that $\{f_n\}_{n\in\mathbb{N}}$ is not equicontinuous at any point x of $\mathbb{R}$.

   Hint. Consider the sequence (x_n) defined by $x_n = x + \pi/(2n)$.
3. Let X be a compact metric space. Prove that, if H is an equicontinuous subset of $C(X)$, the closure $\bar{H}$ of H in $C(X)$ is equicontinuous.
4. Let X be a compact metric space, and let H be an equicontinuous family of elements of $C(X)$.

   a. Prove that the set of points x of X such that the set $\{f(x) : f \in H\}$ is bounded is open and closed.

   b. Assume that X is connected. Prove that, if there exists a point $x \in X$ for which $\{f(x) : f \in H\}$ is bounded, H is a relatively compact subset of $C(X)$.
5. **a.** For $\alpha \in (0,1)$, let $C^\alpha\big([0,1]\big)$ be the set of functions f from $[0,1]$ to $\mathbb{R}$ such that

   $$|f|_\alpha = \sup_{\substack{0\leq x,y\leq 1\\ x\neq y}} \frac{\big|f(x) - f(y)\big|}{|x-y|^\alpha}$$

   is finite (such an f is called a *Hölder function of exponent* α). As usual, we denote by $\|\cdot\|$ the uniform norm.

i. Prove that $C^\alpha([0,1])$, with the norm $\|\cdot\|_\alpha = |\cdot|_\alpha + \|\cdot\|$, is a Banach space.

ii. Prove that $\bar{B}\bigl(C^\alpha([0,1])\bigr)$, the closed unit ball in $C^\alpha\bigl([0,1]\bigr)$, is a compact subset of $C\bigl([0,1]\bigr)$.

iii. Suppose $1 > \beta > \alpha > 0$.

A. Take $f \in C^\beta\bigl([0,1]\bigr)$. Prove that, for all $\eta > 0$,

$$|f|_\alpha \le \max\bigl(|f|_\beta\, \eta^{\beta-\alpha},\ 2\,\|f\|\,\eta^{-\alpha}\bigr).$$

Deduce that, if (f_n) is a bounded sequence in C^β that converges uniformly to $f \in C^\beta$, then $\|f_n - f\|_\alpha \to 0$.

B. Deduce that $\bar{B}\bigl(C^\beta([0,1])\bigr)$ is compact in $C^\alpha([0,1])$.

b. Let m be a nonnegative integer. We give $C^m([0,1])$ the norm defined by

$$\|f\|_m = \sum_{k=0}^{m} \sup_{x\in[0,1]} \bigl|f^{(k)}(x)\bigr|.$$

i. Prove that with this norm $C^m([0,1])$ is a Banach space.

ii. Prove that if m and n are nonnegative integers such that $m > n$, then $\bar{B}(C^m([0,1]))$ is a relatively compact subset of $C^n([0,1])$. (You might start with $m = 1$ and $n = 0$.) Is the ball $\bar{B}\bigl(C^m([0,1])\bigr)$ closed in $C^n([0,1])$?

c. Take $m \in \mathbb{N}$ and $\alpha \in (0,1)$. Denote by $C^{m+\alpha}([0,1])$ the vector space consisting of functions of $C^m([0,1])$ whose m-th derivative is an element of $C^\alpha([0,1])$, and define on this vector space a norm $\|\cdot\|_{m+\alpha}$ by setting $\|f\|_{m+\alpha} = \|f\|_m + |f^{(m)}|_\alpha$.

i. Prove that $C^{m+\alpha}([0,1])$, with the norm $\|\cdot\|_{m+\alpha}$, is a Banach space.

ii. Take $p, q \in \mathbb{R}$ such that $q > p \ge 0$. Prove that $\bar{B}\bigl(C^q([0,1])\bigr)$ is a relatively compact subset of $C^p([0,1])$.

6. *Ascoli's Theorem in* $\mathbb{R}$

a. Let f_n be the function defined for all $x \in \mathbb{R}$ by

$$f_n(x) = \begin{cases} \min(1, n/x) & \text{if } x \ne 0, \\ 1 & \text{if } x = 0. \end{cases}$$

Prove that the subset $\{f_n\}_{n\in\mathbb{N}}$ of $C_0(\mathbb{R})$ is bounded and equicontinuous (see Exercise 8 on page 40 for the definition of $C_0(\mathbb{R})$), but the sequence (f_n) has no uniformly convergent subsequence.

Hint. The sequence (f_n) converges pointwise but not uniformly to the constant function 1.

b. Let H be a subset of $C_0(\mathbb{R})$. Prove that H is relatively compact in $C_0(\mathbb{R})$ if and only if it is bounded and equicontinuous at every point of $\mathbb{R}$ and satisfies that condition that for any $\varepsilon > 0$ there exists $A > 0$ such that

$$\left|h(x)\right| < \varepsilon \quad \text{for all } h \in H \text{ and } x \in \mathbb{R} \text{ with } |x| \geq A.$$

Hint. Use Ascoli's Theorem in the space $C(\mathbb{U})$ (refer again to Exercise 8 on page 40).

7. *A particular case of Peano's Theorem.* Let f be a continuous function from $[0,1] \times \mathbb{R}$ to $\mathbb{R}$ for which there exists a constant $M > 0$ such that

$$\left|f(x,t)\right| \leq M\big(1 + |x|\big) \quad \text{for all } t \in [0,1] \text{ and } x \in \mathbb{R}.$$

a. Let n be a positive integer. We define points x_j^n, for $0 \leq j \leq n$, by setting $x_0^n = 0$ and

$$x_{j+1}^n = x_j^n + \frac{1}{n} f\Big(\frac{j}{n},\, x_j^n\Big) \quad \text{for } 0 \leq j \leq n-1.$$

i. Prove that $|x_j^n| \leq (1 + M/n)^j - 1 \leq e^M - 1$ for $0 \leq j \leq n$.

ii. Let φ_n be the continuous function on $[0,1]$ that is affine on each interval $\big[j/n,\, (j{+}1)/n\big]$ and satisfies $\varphi_n(j/n) = x_j^n$ for $0 \leq j \leq n$. That is, for $0 \leq j \leq n-1$ and $t \in \big[j/n,\, (j{+}1)/n\big]$ we have

$$\varphi_n(t) = x_j^n + \Big(t - \frac{j}{n}\Big) f\Big(\frac{j}{n},\, x_j^n\Big).$$

Prove that for $s, t \in [0,1]$ we have $\left|\varphi_n(t) - \varphi_n(s)\right| \leq M e^M |t - s|$.

iii. For $s \in [0,1]$, set

$$\psi_n(s) = \sum_{j=0}^{n-1} 1_{[j/n,\,(j+1)/n)}(s)\, f\bigg(\frac{j}{n},\, \varphi_n\Big(\frac{j}{n}\Big)\bigg).$$

Prove that $\varphi_n(t) = \displaystyle\int_0^t \psi_n(s)\, ds$ for all $t \in [0,1]$.

b. **i.** Show that there exists a subsequence $(\varphi_{n_k})_{k \in \mathbb{N}}$ that converges uniformly on $[0,1]$ to a function $\varphi \in C([0,1])$.

ii. Prove that the sequence $(\psi_{n_k})_{k \in \mathbb{N}}$ converges uniformly on $[0,1)$ to $f\big(s, \varphi(s)\big)$.

iii. Deduce that $\varphi(t) = \int_0^t f\big(s, \varphi(s)\big)\, ds$ for all $t \in [0,1]$; then prove that φ is of class C^1 on $[0,1]$ and satisfies the differential equation

$$\begin{cases} \varphi'(t) = f\big(t, \varphi(t)\big) & \text{for all } t \in [0,1], \\ \varphi(0) = 0. \end{cases} \tag{$*$}$$

Is the φ constructed above the only one that satisfies these conditions?

8. Let X be a compact metric space and H a subset of $C(X)$.

a. Suppose H is relatively compact. Prove that for all $\varepsilon > 0$ there exist constants $C > 0$ and $B > 0$ such that $d(f, L_C^B) \leq \varepsilon$ for all $f \in H$, where L_C^B denotes the set of C-Lipschitz functions on X with uniform norm at most B, and d is the metric associated with the same norm.

Hint. Use the fact that Lipschitz functions are dense in $C(X)$.

b. Show the converse.

Hint. Prove that L_C^B is precompact, and finally that so is H.

2

Locally Compact Spaces and Radon Measures

In this chapter we study a representation, in terms of measures, of positive linear forms on spaces of continuous functions; this representation leads to a description of the topological dual of such spaces. It is useful in applications to consider functions defined on metric spaces somewhat more general than compact spaces, namely, locally compact ones.

1 Locally Compact Spaces

A metric space (X, d) is called **locally compact** if every point in X has a compact neighborhood; equivalently, if for every $x \in X$ there exists a compact K of X whose interior contains x; equivalently, if for every $x \in X$ there exists $r > 0$ such that the closed ball $\bar{B}(x, r)$ is compact. Local compactness is clearly a topological notion.

Any compact space is obviously locally compact. The spaces $\mathbb{R}^d$ and $\mathbb{C}^d$, for $d \geq 1$, and more generally all normed spaces of finite nonzero dimension yield a first example of locally compact but noncompact spaces. The famous theorem of F. Riesz states that, conversely, the only locally compact normed spaces are those of finite dimension:

Theorem 1.1 (F. Riesz) *Let X be a normed space, with open unit ball B and closed unit ball $\bar{B}$. The following properties are equivalent:*

i. *X is finite-dimensional.*
ii. *X is locally compact.*

iii. *$\bar{B}$ is compact.*

iv. *B is precompact.*

Proof. Property i implies ii because closed balls in a finite-dimensional normed space are compact. If ii is true, there exists $r > 0$ such that $\bar{B}(0,r) = r\bar{B}$ is compact; this implies iii. That iii implies iv is obvious. Thus the only nontrivial part of the theorem is iv $\Longrightarrow$ i.

Suppose that B is precompact. Then there is a finite subset A of X such that

$$B \subset \bigcup_{x \in A} B\left(x, \tfrac{1}{2}\right) = A + \tfrac{1}{2}B.$$

Let Y be the (finite-dimensional) vector space generated by A; then $B \subset Y + 2^{-1}B$. One can easily show by induction that, for any integer $n \geq 1$, we have $B \subset Y + 2^{-n}B$, and therefore

$$B \subset \bigcap_{n \geq 1} (Y + 2^{-n}B).$$

In particular, if $x \in B$, there exists for all $n \geq 1$ a $y_n \in Y$ such that $\|x - y_n\| < 2^{-n}$. We deduce that $B \subset \bar{Y}$. Since Y is finite-dimensional, hence complete, hence closed in X, it follows that $B \subset Y$ and, by homogeneity, $X = Y$. □

We remark that any space with the discrete metric (defined by $d(x,y) = 1$ if $x \neq y$ and $d(x,x) = 0$) is locally compact.

Here is a simple but important consequence of the definition of local compactness.

Proposition 1.2 *If X is a locally compact space, there exists for every $x \in X$ and for every neighborhood V of x a real number $r > 0$ such that $\bar{B}(x,r)$ is compact and $\bar{B}(x,r) \subset V$.*

Proof. Just choose $r = \min(r', r'')$, where r' and r'' are such that $\bar{B}(x,r')$ is compact and $\bar{B}(x,r'') \subset V$. □

Corollary 1.3 *Let X be locally compact. If O is open in X and F is closed in X, the intersection $Y = O \cap F$ (with the induced metric) is locally compact.*

Proof. Take $x \in Y$. By the preceding proposition, there exists $r > 0$ such that $\bar{B}(x,r)$ is compact and contained in O. Then $\bar{B}(x,r) \cap Y = \bar{B}(x,r) \cap F$ is compact. □

In particular, every open set in a finite-dimensional normed space is locally compact.

Corollary 1.4 *Consider a locally compact space X, a compact subset K of X, and open subsets $O_1, \dots, O_n$ of X covering K. There exist compact sets $K_1, \dots, K_n$ with $K_j \subset O_j$ for each j and such that*

$$K \subset \bigcup_{j=1}^{n} \mathring{K}_j.$$

Proof. By Proposition 1.2, for all points x of K there exists $j \in \{1, \dots, n\}$ and a compact set K_x such that $x \in \mathring{K}_x \subset K_x \subset O_j$. By the Borel–Lebesgue property, K can be covered by finitely many of these interiors:

$$K \subset \bigcup_{i=1}^{p} \mathring{K}_{x_i}.$$

Now set $K_j = \bigcup_{K_{x_i} \subset O_j} K_{x_i}$ for $1 \le j \le n$. Then

$$\bigcup_{j=1}^{n} \mathring{K}_j \supset \bigcup_{j=1}^{n} \bigcup_{K_{x_i} \subset O_j} \mathring{K}_{x_i} = \bigcup_{i=1}^{p} \mathring{K}_{x_i} \supset K;$$

and, sure enough, $K_j \subset O_j$. □

The next result is about the separability of locally compact spaces.

Proposition 1.5 *Let X be a locally compact space. The following properties are equivalent:*

i. *X is separable.*
ii. *X is σ-compact.*
iii. *There exists a sequence (K_n) of compact sets covering X and such that $K_n \subset \mathring{K}_{n+1}$ for all $n \in \mathbb{N}$.*

Proof. It is clear that iii implies ii. The implication ii $\Longrightarrow$ i is a particular case of Proposition 2.2 on page 8.

Now suppose that X is separable and let (x_n) be a sequence dense in X. Set $A = \{(n,p) \in \mathbb{N} \times \mathbb{N}^* : \bar{B}(x_n, 1/p) \text{ is compact}\}$; we will show that the family $\mathscr{F} = (\bar{B}(x_n, 1/p))_{(n,p) \in A}$ covers X. Take $x \in X$ and let $r > 0$ be such that $\bar{B}(x, r)$ is compact. Then take $p \in \mathbb{N}^*$ such that $1/p < r/2$ and $n \in \mathbb{N}$ such that $d(x, x_n) < 1/p$. One sees that $x \in \bar{B}(x_n, 1/p) \subset B(x, 2/p) \subset \bar{B}(x, r)$. Therefore $\bar{B}(x_n, 1/p)$ is compact and x belongs to some element of $\mathscr{F}$. This shows that i implies ii.

Finally, we show that ii implies iii. Suppose that X is σ-compact and let (L_n) be a sequence of compact sets that cover X. We construct the sequence (K_n) by induction, as follows: set $K_0 = L_0$ and, for $n \ge 1$, choose K_n such $K_{n-1} \cup L_{n-1} \subset \mathring{K}_n$ (using Corollary 1.4). □

A sequence (K_n) of compact sets that covers X and satisfies $K_n \subset \mathring{K}_{n+1}$ for all n is said to **exhaust** X.

Proposition 1.6 *Let (K_n) be a sequence of compact sets that exhausts a metric space X. For every compact K of X there exists an integer n such that $K \subset \mathring{K}_n$.*

Proof. The open sets $\mathring{K}_n$ cover K. By the Borel–Lebesgue property, K is in fact contained in a finite union of sets $\mathring{K}_n$: but $\bigcup_{j\le n} \mathring{K}_j = \mathring{K}_n$. □

Continuous Functions on a Metric Space

We now introduce various spaces of continuous functions on a metric space (X, d).

We denote by $C_b^{\mathbb{K}}(X)$, or simply by $C_b(X)$, the vector space over $\mathbb{K}$ consisting of **bounded** continuous functions $f : X \to \mathbb{K}$; recall that f being bounded means that $\sup_{x\in X}|f(x)| < +\infty$. We give $C_b^{\mathbb{K}}(X)$ the uniform norm (or norm of uniform convergence), defined by

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

With this norm, $C_b^{\mathbb{K}}(X)$ is a Banach space.

We say that a function $f : X \to \mathbb{K}$ **tends to zero at infinity** if for all $\varepsilon > 0$ there exists a compact subset K of X such that $|f(x)| < \varepsilon$ for all $x \notin K$. We denote by $C_0^{\mathbb{K}}(X)$ or $C_0(X)$ the vector space over $\mathbb{K}$ consisting of continuous functions $X \to \mathbb{K}$ that tend to 0 at infinity. It is easy to check that $C_0(X)$ is a closed subspace of $C_b(X)$; therefore $C_0(X)$ with the uniform norm forms a Banach space.

We remark that Dini's Lemma (Proposition 1.2 on page 29) can be generalized to $C_0^{\mathbb{R}}(X)$:

Proposition 1.7 *Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in $C_0^{\mathbb{R}}(X)$, converging pointwise to a function $f \in C_0^{\mathbb{R}}(X)$. Then (f_n) converges uniformly to f.*

Proof. We show that the sequence (g_n) defined by $g_n = f - f_n$ converges uniformly to 0. Given $\varepsilon > 0$, there exists a compact K such that $g_0(x) \le \varepsilon$ for all $x \notin K$. By Dini's Lemma, there exists an integer n such that $g_n(x) \le \varepsilon$ for all $x \in K$. Since the sequence (g_n) is decreasing, this implies that for all $p \ge n$ and all $x \in X$ we have $0 \le g_p(x) \le \varepsilon$. □

The **support** of a function $f : X \to \mathbb{K}$, denoted $\operatorname{Supp} f$, is the closure of the set $\{x \in X : f(x) \ne 0\}$. Thus $\operatorname{Supp} f$ is the complement of the largest open set where f vanishes, this latter set being of course the interior of $f^{-1}(\{0\})$. We denote by $C_c^{\mathbb{K}}(X)$ or $C_c(X)$ the vector space over $\mathbb{K}$ consisting of the functions $X \to \mathbb{K}$ having compact support. Clearly $C_c(X)$ is a vector subspace of $C_0(X)$, but not in general a closed one; see Corollary 1.9 below, for example. Naturally, if X is compact we have $C_c(X) = C_0(X) = C_b(X) = C(X)$.

Finally, we denote by $C_b^+(X)$, $C_0^+(X)$, and $C_c^+(X)$ the subsets of $C_b^{\mathbb{R}}(X)$, $C_0^{\mathbb{R}}(X)$, and $C_c^{\mathbb{R}}(X)$ consisting of functions that take only positive values.

Proposition 1.8 (Partitions of unity) *Let X be locally compact. If K is a compact subset of X and $O_1, \dots, O_n$ are open subsets of X that cover K, there exist functions $\varphi_1, \dots, \varphi_n$ in $C_c^{\mathbb{R}}(X)$ such that $0 \le \varphi_j \le 1$ and $\operatorname{Supp} \varphi_j \subset O_j$ for each j and*

$$\sum_{j=1}^{n} \varphi_j(x) = 1 \quad \textit{for all } x \in K.$$

Proof. Let $K_1, \dots, K_n$ be the compact sets whose existence is granted by Corollary 1.4. We just have to set, for $x \in X$,

$$\varphi_j(x) = \frac{d(x, X \setminus \mathring{K}_j)}{d(x, K) + \sum_{k=1}^{n} d(x, X \setminus \mathring{K}_k)}.$$

In particular, $\operatorname{Supp} \varphi_j \subset K_j \subset O_j$. □

A family $(\varphi_1, \dots, \varphi_n)$ satisfying the conditions of the proposition is called a **partition of unity on K subordinate to the open cover $O_1, \dots, O_n$**.

Corollary 1.9 *If X is locally compact, $C_c(X)$ is dense in $C_0(X)$.*

Proof. Take $f \in C_0(X)$ and $\varepsilon > 0$. Let K be a compact such that $|f(x)| < \varepsilon$ for all $x \notin K$. Applying Proposition 1.8 with $n = 1$ and $O_1 = X$, we find a $\varphi \in C_c^{\mathbb{R}}(X)$ such that $0 \le \varphi \le 1$ and $\varphi = 1$ on K. Then $f\varphi \in C_c(X)$ and $\|f - f\varphi\| \le \varepsilon$. □

Corollary 1.10 *Let X be locally compact and separable and let O be open in X. There exists an increasing sequence (φ_n) of functions in $C_c^+(X)$, each with support contained in O, and such that $\lim_{n \to +\infty} \varphi_n(x) = 1_O(x)$ for all $x \in X$.*

Proof. O is a locally compact separable space, by Corollary 1.3 above and Proposition 2.3 on page 8. By Proposition 1.5 there exists a sequence of compact sets (K_n) such that $K_n \subset \mathring{K}_{n+1}$ for all n and $\bigcup_{n \in \mathbb{N}} K_n = O$. By Proposition 1.8 there exists for each n a map $\varphi_n \in C_c^{\mathbb{R}}(X)$ such that $0 \le \varphi_n \le 1$, $\varphi_n|_{K_n} = 1$, and $\operatorname{Supp} \varphi_n \subset \mathring{K}_{n+1}$. The sequence (φ_n) clearly satisfies the desired conditions. □

To conclude this section, we observe that $C_b(X)$ is a algebra with unity, that $C_c(X)$ and $C_0(X)$ are subalgebras of $C_b(X)$ (without unity if X is not compact), and that $C_b^{\mathbb{R}}(X)$, $C_0^{\mathbb{R}}(X)$ and $C_c^{\mathbb{R}}(X)$ are also lattices.

Exercises

1. **a.** Let X be a metric space. Prove that, if there exists a real number $r > 0$ such that all closed balls of radius r in X are compact, then X is complete.

b. Find a locally compact metric space X such that, for all $x \in X$, there is a compact closed ball of center x that is noncompact.

c. Find a locally compact metric space that is not complete.

d. Find a complete metric space that is not locally compact.

2. **a.** Let (X_1, d_1) and (X_2, d_2) be locally compact metric spaces. Prove that $X_1 \times X_2$, together with the product metric given by $d(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2)$, is locally compact.

b. Let $\big((X_p, d_p)\big)_{p\in\mathbb{N}}$ be a sequence of locally compact, nonempty metric spaces, and set $X = \prod_{p\in\mathbb{N}} X_p$, with the product metric d (see page 13).

i. Take $x \in X$ and $r \in (0, 1]$. Prove that if n and n' are integers satisfying $2^{-n} < r \le 2^{-n'}$, then

$$\prod_{p=0}^{n}\{x_p\} \times \prod_{p>n} X_p \subset B(x, r) \subset \prod_{p=0}^{n'} B_p(x_p, 2^p r) \times \prod_{p>n'} X_p,$$

where $B_p(\,\cdot\,,\cdot\,)$ and $B(\,\cdot\,,\cdot\,)$ represent open balls in (X_p, d_p) and (X, d), respectively.

ii. Prove that (X, d) is locally compact if and only if all but a finite number of factors (X_p, d_p) are compact.

3. Let X be a metric space and Y a subset of X.

a. Prove that $B(x,r)\cap\bar{Y} \subset \overline{B(x,r)\cap Y}$ for all $x \in Y$ and $r > 0$. Deduce that, if $\bar{B}(x,r) \cap Y$ is compact, then $B(x,r) \cap \bar{Y} \subset Y$.

b. Suppose that Y, with the induced metric, is locally compact. Show that there exists an open subset O of X such that $Y = O \cap \bar{Y}$. This gives a converse for Corollary 1.3.

4. Show that an infinite-dimensional Banach space cannot be σ-compact.
Hint. Use Baire's Theorem (Exercise 6 on page 22).

5. **a.** Prove that every metric space that can be exhausted by a sequence of compact sets is locally compact.

b. Find a σ-compact metric space that is not locally compact.

6. *Baire's Theorem, continued.* Let X be a metric space. Recall from Exercise 6 on page 22 the game of Choquet between Pierre and Paul.

a. Prove that Paul has a winning strategy if X is locally compact. Deduce that in X no open set can be a union of a countable family of closed sets with empty interior.

Hint. The intersection of a decreasing sequence of nonempty compact sets cannot be empty.

b. Take $X = \mathbb{R} \setminus \mathbb{Q}$, with the usual metric. Prove that X is neither complete nor locally compact (you can use Exercise 3 above, for example), but that Paul nonetheless has a winning strategy, so Baire's Theorem is valid in X.

Hint. Take an enumeration of the rationals, say $\mathbb{Q} = \{r_n\}_{n\in\mathbb{N}}$. Show that, whenever Pierre plays U_n, Paul can respond with $V_n = I_n \setminus \mathbb{Q}$, where I_n is a nonempty open interval in $\mathbb{R}$ such that $\bar{I}_n \setminus \mathbb{Q} \subset U_n$, $d(I_n) \le 1/n$, and $r_n \notin \bar{I}_n$.

7. *Alexandroff compactification.* Let (X, d) be a separable and locally compact metric space. Set $\hat{X} = X \cup \{\infty\}$, where ∞ is a point that does not belong to X. We wish to define on $\hat{X}$ a metric that extends the topology of X and that makes $\hat{X}$ compact. To do this, let $(V_n)_{n\in\mathbb{N}}$ be a countable basis of open sets in X (see Exercise 1 on page 10), and put

$$\mathscr{U} = \{(p,q) \in \mathbb{N}^2 : \bar{V}_p \subset V_q \text{ and } \bar{V}_p \text{ is compact }\}.$$

This set is countable; let $\mathscr{U} = \{(p_n, q_n)\}_{n\in\mathbb{N}}$ be an enumeration of it. For each n, let φ_n be an element of $C_c(X)$ such that $0 \le \varphi_n \le 1$ everywhere and $\varphi_n = 1$ on V_{p_n}, and whose support is contained in V_{q_n}. Put $\varphi_n(\infty) = 0$. Then, for $x, y \in \hat{X}$, define

$$\delta(x,y) = \sum_{n=0}^{+\infty} 2^{-n} \left|\varphi_n(x) - \varphi_n(y)\right|.$$

a. Prove that δ is a metric on $\hat{X}$.

b. Let $(x_j)_{j\in\mathbb{N}}$ be a sequence in $\hat{X}$. Prove that $\lim_{j\to+\infty} \delta(x_j, \infty) = 0$ if and only if, for any compact K in X, there is an integer J such that $x_j \notin K$ for $j \ge J$. (In this case we say that the sequence (x_j) *tends to infinity.*)

c. Let $(x_j)_{j\in\mathbb{N}}$ be a sequence in X and x a point in X. Prove that $\lim_{j\to+\infty} d(x_j, x) = 0$ if and only if $\lim_{j\to+\infty} \delta(x_j, x) = 0$. Together with the preceding result, this shows that the convergence of sequences in $\hat{X}$, and therefore the topology of $(\hat{X}, \delta)$, does not depend on the choice of d and δ.

d. Prove that $(\hat{X}, \delta)$ is a compact metric space.

e. Prove that X is compact if and only if ∞ is an isolated point of $\hat{X}$.

f. We now suppose that $X = \mathbb{R}^d$. Prove that $\hat{X}$ is homeomorphic to S_d, the (euclidean) unit sphere in $\mathbb{R}^{d+1}$, that is,

$$S_d = \left\{ x \in \mathbb{R}^{d+1} : \sum_{i=1}^{d+1} x_i^2 = 1 \right\},$$

with the distance induced by the euclidean norm in $\mathbb{R}^{d+1}$.

Hint. Use *stereographic projection*, the map $\varphi : S^d \to \widehat{\mathbb{R}^d}$ defined by $\varphi(0,\ldots,0,1) = \infty$ and

$$\varphi(x) = \left(\frac{x_j}{1-x_{d+1}}\right)_{1\le j\le d} \qquad \text{for } x \ne (0,\ldots,0,1).$$

g. Prove that $C_0(X)$ can be identified with the space of continuous functions on $\hat{X}$ that vanish at ∞.

h. Deduce that $C_0(X)$ is separable.

i. Prove that the Stone–Weierstrass Theorem, stated in Exercise 8b on page 41 for $\mathbb{R}$, generalizes to the case where $\mathbb{R}$ is replaced by X.

j. *Ascoli's Theorem in $C_0(X)$.* Prove that a subset H of $C_0(X)$ is relatively compact in $C_0(X)$ if and only if it is bounded and equicontinuous and satisfies the condition that for every $\varepsilon > 0$ there exists a compact subset K of X such that $|h(x)| < \varepsilon$ for every $x \in X \setminus K$ and every $h \in H$.

8. Let X be a locally compact space. Prove that X is separable if and only if $C_0(X)$ contains a function taking positive values everywhere.
Hint. X is separable if and only if it is σ-compact.

9. Let (X, d) be a metric space.

a. Prove that $C_b(X)$ and $C_0(X)$, with the uniform norm, are Banach spaces.

b. Prove that X is compact if and only if every continuous function from X into $\mathbb{R}$ is bounded.
Hint. Show that, if X is not compact, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in X having no convergent subsequence and a sequence $(r_n)_{n\in\mathbb{N}}$ of positive real numbers tending toward 0 and such that the balls $B(x_n, r_n)$ are pairwise disjoint. Then consider $\sum_{n\in\mathbb{N}} n\varphi_n$, where $\varphi_n(x) = \bigl(1 - d(x, x_n)/r_n\bigr)^+$.

c. Prove that $C_b(X)$ is separable if and only if X is compact.
Hint. Suppose that X is not compact and define, for each $\alpha \in \{0,1\}^{\mathbb{N}}$, a function f_α by setting $f_\alpha = \sum_{n\in\mathbb{N}} \alpha_n\varphi_n$, where the φ_n are as in part b. Prove that $f_\alpha \in C_b(X)$ and that $\|f_\alpha - f_\beta\| = 1$ if $\alpha \ne \beta$. Then use Proposition 2.4 on page 9. (Side question: Among the functions f_α, how many have compact support?)

10. *Tietze Extension Theorem, continued.* Let X be a locally compact space, K a compact subset of X, and f a continuous function $K \to \mathbb{K}$. Prove that there exits a function $\tilde{f} \in C_c(X)$ such that $\tilde{f}|_K = f$ and $\|\tilde{f}\| = \max_{x\in K} |f(x)|$.
Hint. Use Exercise 7 on page 40 and Proposition 1.8 above.

11. Extend the result of Exercise 1 on page 30 to the case where X is separable and locally compact and $C(X)$ is replaced by $C_0(X)$.
Hint. One can use Exercise 7 to reduce the problem to the one covered by the original result.

12. *Topology of uniform convergence on compact sets.* Let X be a separable, locally compact metric space and (K_n) an exhausting sequence of compact sets of X. Let $C(X)$ be the vector space consisting of continuous functions from X into $\mathbb{K}$. For each element f of $C(X)$ define a real number $q(f)$ as

$$q(f) = \sum_{n=0}^{+\infty} 2^{-n} \min\bigl(1, \|f\|_{K_n}\bigr),$$

where $\|\cdot\|_{K_n}$ represents the uniform norm on K_n.

a. Prove that the map $d : (f, g) \mapsto q(f-g)$ is a metric on $C(X)$.

b. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of elements of $C(X)$ and let f be an element of $C(X)$. Prove that (f_k) converges to f uniformly on every compact of X if and only if $\lim_{k\to+\infty} d(f_k, f) = 0$.

c. Prove that the metric space $(C(X), d)$ is complete.

d. For $n \in \mathbb{N}$, let $(\varphi_{n,p})_{p\in\mathbb{N}}$ be a dense sequence in $C(K_n)$. We know by Exercise 10 above that we can extend each $\varphi_{n,p}$ to a function $\tilde{\varphi}_{n,p} \in C_c(X)$. Prove that the family $(\tilde{\varphi}_{n,p})_{(n,p)\in\mathbb{N}^2}$ is dense in $(C(X), d)$.

e. Deduce that the metric space $(C(X), d)$ is separable and that $C_c(X)$ is dense in $(C(X), d)$.

f. Deduce that $(C_b(X), d)$ and $(C_0(X), d)$ are complete if and only if X is compact (see Exercise 9b above).

g. *Ascoli's Theorem in $C(X)$.* Let H be a subset of $C(X)$. Prove that H is relatively compact in $(C(X), d)$ if and only if it satisfies the following conditions:

- H is equicontinuous at every point of X.
- For every point x of X, the set $\{h(x)\}_{h\in H}$ is bounded.

Hint. Carry out the diagonal procedure using Ascoli's Theorem on each compact K_n.

2 Daniell's Theorem

This section approaches integration from a functional point of view. We assume the reader is familiar with the set-theoretical approach to integration, where a measure is defined as a σ-additive function on sets.

Notation. Let X be any nonempty set. We denote by $\mathscr{F}$ the vector space over $\mathbb{R}$ consisting of all functions from X to $\mathbb{R}$. This space, with the usual order relation, is a lattice: If f and g are elements of $\mathscr{F}$,

$$\bigl(\sup(f, g)\bigr)(x) = \max\bigl(f(x), g(x)\bigr) \quad \text{and} \quad \bigl(\inf(f, g)\bigr)(x) = \min\bigl(f(x), g(x)\bigr).$$

If (f_n) is a sequence in $\mathscr{F}$ and f is an element of $\mathscr{F}$, we write $f_n \nearrow f$ to mean that the sequence (f_n) is increasing and converges pointwise to f; the meaning of $f_n \searrow f$ is analogous.

As before, we use the same symbol for a constant function and its value.

If m is a measure on a σ-algebra of X, we denote by $\mathscr{L}^1(m)$ the subspace of $\mathscr{F}$ consisting of m-integrable functions. As usual, we denote by $L^1(m)$ the quotient vector space of $\mathscr{L}^1(m)$ by the equivalence relation given by equality m-almost everywhere, endowed with the norm defined by $\|f\| = \int |f|\, dm$ (we use the same symbol f for an equivalence class and one of its representatives). The normed space $L^1(m)$ is then a Banach space.

During the remainder of this section, we consider a vector subspace L of $\mathscr{F}$ that is a lattice (this is equivalent to saying that $f \in L$ implies $|f| \in L$) *and satisfies the following condition:*

$$\text{\textit{There exists a sequence} } (\varphi_n) \text{ \textit{in} } L \text{ \textit{such that} } \varphi_n \nearrow 1. \tag{$*$}$$

We will denote by $\sigma(L)$ the σ-algebra generated by L, that is, the smallest σ-algebra of X that makes all elements of L measurable. Finally, let $\mathscr{L}$ be the set of functions from X to $\mathbb{R}$ that are $\sigma(L)$-measurable.

Lemma 2.1 *$\mathscr{L}$ is the smallest subset of $\mathscr{F}$ that contains L and is closed under pointwise convergence* (the latter condition means that the pointwise limit of any sequence in $\mathscr{L}$ is also in $\mathscr{L}$).

Proof. It is clear that a minimal set satisfying these conditions exists. Call it $\mathscr{B}$.

- *$\mathscr{B}$ is a vector subspace of $\mathscr{F}$ and a lattice, and it contains the constants.*

  Proof. If $\lambda \in \mathbb{R}$, the set $\{f \in \mathscr{F} : \lambda f \in \mathscr{B}\}$ contains L and is closed under pointwise convergence, so it contains $\mathscr{B}$. Therefore $f \in \mathscr{B}$ and $\lambda \in \mathbb{R}$ imply $\lambda f \in \mathscr{B}$.
  Similarly, for every $g \in L$, the set $\{f \in \mathscr{F} : f + g \in \mathscr{B}\}$ contains $\mathscr{B}$, so the sum of an element of L and one of $\mathscr{B}$ is in $\mathscr{B}$. Using the same reasoning again we deduce from this that, for every $f \in \mathscr{B}$, the set $\{h \in \mathscr{F} : f + h \in \mathscr{B}\}$ contains $\mathscr{B}$. Thus the sum of two elements of $\mathscr{B}$ is in $\mathscr{B}$, and $\mathscr{B}$ is a vector space.
  Since L is a lattice we see by considering the set $\{f \in \mathscr{F} : |f| \in \mathscr{B}\}$ that $\mathscr{B}$ is a lattice as well. That $\mathscr{B}$ contains 1 and therefore all constants follows from condition $(*)$. □

- We now show that $\mathscr{B} = \mathscr{L}$. Set $\mathscr{T} = \{A \subset X : 1_A \in \mathscr{B}\}$. By the preceding paragraphs, $\mathscr{T}$ is a σ-algebra. If $f \in L$ and $a \in \mathbb{R}$, the characteristic function of the set $\{f > a\}$ is the pointwise limit of the sequence $(\inf(n(f-a)^+, 1))$, and so belongs to $\mathscr{B}$, and $\{f > a\} \in \mathscr{T}$. Thus the elements of L are $\mathscr{T}$-measurable, which implies that $\sigma(L) \subset \mathscr{T}$; in other words, $1_A \in \mathscr{B}$ for $A \in \sigma(L)$. Since every $\sigma(L)$-measurable function is the pointwise limit of $\sigma(L)$-measurable piecewise constant functions, we deduce that $\mathscr{L} \subset \mathscr{B}$ and, by the minimality of $\mathscr{B}$, that $\mathscr{L} = \mathscr{B}$. □

Example. Let X be a metric space. Recall that the **Borel σ-algebra** of X is the smallest σ-algebra of X that contains all open sets of X, and that the corresponding measurable functions are called **Borel functions**.

Proposition 2.2 *If X is a metric space, the set of Borel functions from X to $\mathbb{R}$ is the smallest subset of $\mathscr{F}$ that contains all continuous functions from X to $\mathbb{R}$ and is closed under pointwise convergence.*

Proof. Let L be the set of continuous functions from X to $\mathbb{R}$. Then L is a lattice and satisfies $(*)$, since $1 \in L$. On the other hand, let $\mathscr{B}$ be the Borel σ-algebra of X. Certainly every continuous function on X is $\mathscr{B}$-measurable, so $\sigma(L) \subset \mathscr{B}$. Conversely, every open set U of X is contained in $\sigma(L)$: to see this, note, for example, that U is the inverse image of the open set $\mathbb{R}^*$ under the continuous function f defined by $f(x) = d(x, U^c)$. Thus $\mathscr{B} \subset \sigma(L)$, which implies $\mathscr{B} = \sigma(L)$. Now apply Lemma 2.1. $\square$

Remark. One should not confuse $\mathscr{L}$ with the set of pointwise limits of sequences in L, which is generally strictly smaller that $\mathscr{L}$. In the situation of the preceding example, this smaller set is called the set of functions of first Baire class: see Exercise 4.

The rest of this section is devoted to the proof of the following result:

Theorem 2.3 (Daniell) *Let μ be a linear form on L satisfying these conditions:*

1. *μ is **positive**, that is, if $f \in L$ satisfies $f \geq 0$ then $\mu(f) \geq 0$.*
2. *If a sequence (f_n) in L satisfies $f_n \searrow 0$, then $\lim_{n\to+\infty} \mu(f_n) = 0$.*

Then there exists a unique measure m on the σ-algebra $\sigma(L)$ such that

$$L \subset \mathscr{L}^1(m) \qquad \text{and} \qquad \mu(f) = \int f\,dm \quad \text{for all } f \in L.$$

Uniqueness of m. Suppose that two measures m_1 and m_2 satisfy the stated properties. Let (φ_n) be a sequence satisfying condition $(*)$ on page 58. For every $n \in \mathbb{N}$ and every real $\lambda \geq 0$, the set

$$\left\{ f \in \mathscr{L} : \int \inf(f^+, \lambda\varphi_n)\,dm_1 = \int \inf(f^+, \lambda\varphi_n)\,dm_2 \right\}$$

equals $\mathscr{L}$, by the minimality of $\mathscr{L}$ (proved in Lemma 2.1) and the Dominated Convergence Theorem. Making n go to infinity, then λ, we conclude by the Monotone Convergence Theorem that $\int f^+\,dm_1 = \int f^+\,dm_2$ for all $f \in \mathscr{L}$. Therefore $m_1 = m_2$ on $\sigma(L)$. $\square$

Existence of m. The proof of existence is rather long and is carried out in several steps. First of all, let $\mathscr{U}$ be the set of functions from X into $\mathbb{R}$ that are pointwise limits of increasing sequences of elements of L. The measure

m is constructed by first extending the linear form μ to $\mathscr{U}$ (steps 1–3), then to the space L^1 defined in step 6 below. Some properties of L^1 and of μ are established in step 7, allowing us to conclude the proof in step 8.

1. *The set $\mathscr{U}$ contains the positive constants, is closed under addition and multiplication by nonnegative reals, and for any pair (f, g) of elements of $\mathscr{U}$, we have* $\sup(f, g) \in \mathscr{U}$ *and* $\inf(f, g) \in \mathscr{U}$. *Moreover $\mathscr{U}$ is closed under pointwise convergence of increasing sequences.*

Proof. Only the last assertion requires elaboration. Let (f_n) be an increasing sequence in $\mathscr{U}$ converging toward an element f of $\mathscr{F}$. By assumption, there exists, for any $n \in \mathbb{N}$, a sequence $(g_{n,m})_{m\in\mathbb{N}}$ in L that is increasing and converges to f_n. For each $m \in \mathbb{N}$, set $h_m = \sup_{0\leq n\leq m} g_{n,m}$. It is clear that $(h_m)_{m\in\mathbb{N}}$ is an increasing sequence in L and that $g_{n,m} \leq h_m \leq f_m$ if $m \geq n$. Making m go to infinity in this inequality, we get $f_n \leq \lim_{m\to+\infty} h_m \leq f$; then making n go to infinity, we get $h_m \nearrow f$, which shows that $f \in \mathscr{U}$. □

2. *Let (f_n) and (g_n) be increasing sequences in L, converging pointwise to elements f and g of $\mathscr{U}$, respectively. If $f \leq g$, then*

$$\lim_{n\to+\infty} \mu(f_n) \leq \lim_{n\to+\infty} \mu(g_n) \leq +\infty.$$

Proof. By linearity and positivity, the linear form μ is increasing on L ($f \leq g$ implies $L(f) \leq L(g)$). On the other hand, for each $n \in \mathbb{N}$, we have $\inf(f_n, g_m) \nearrow f_n$ as m goes to infinity, so $\lim_{m\to+\infty} \mu(\inf(f_n, g_m)) = \mu(f_n)$, by assumption 2 of the theorem applied to the sequence $(f_n - \inf(f_n, g_m))_m$. It follows that $\mu(f_n) \leq \lim_{m\to+\infty} \mu(g_m)$ for all $n \in \mathbb{N}$, and this shows the result. □

3. We extend μ to $\mathscr{U}$ by setting $\mu(f) = \lim_{n\to+\infty} \mu(f_n)$, where $f \in \mathscr{U}$ and (f_n) is an increasing sequence in L that converges to f pointwise. By step 2, μ is well-defined and increasing on $\mathscr{U}$, and it takes values in $(-\infty, +\infty]$. Moreover, μ is additive (that is, $\mu(f + g) = \mu(f) + \mu(g)$ for $f, g \in \mathscr{U}$) and, for all $f \in \mathscr{U}$ and every nonnegative real λ, we have $\mu(\lambda f) = \lambda\mu(f)$, with the usual convention $0\cdot\infty = 0$. Now, *if (f_n) is an increasing sequence in $\mathscr{U}$ that converges to $f \in \mathscr{F}$ pointwise,* $\mu(f) = \lim_{n\to+\infty} \mu(f_n)$.

Proof. By step 1, f is in $\mathscr{U}$. Using the same notation as in the proof of that step, we can write

$$\mu(f) = \lim_{m\to+\infty} \mu(h_m) \leq \lim_{m\to+\infty} \mu(f_m);$$

the reverse inequality is a consequence of the fact that μ is increasing in $\mathscr{U}$. □

4. We now extend μ to $-\mathscr{U}$ by setting $\mu(-f) = -\mu(f)$ for $f \in \mathscr{U}$. This gives no rise to inconsistencies: if $f \in \mathscr{U} \cap (-\mathscr{U})$, then $f + (-f) = 0$ and therefore $\mu(f) + \mu(-f) = 0$ and $\mu(f) = -\mu(-f)$. It is also clear that
$$\mu(g-h) = \mu(g) - \mu(h) \quad \text{if } g \in \mathscr{U} \text{ and } h \in -\mathscr{U}.$$
In particular, if $g \in \mathscr{U}$ and $h \in -\mathscr{U}$, then $h \leq g$ implies $\mu(h) \leq \mu(g)$.
5. Let $\mathscr{V}$ be the set consisting of elements $f \in \mathscr{F}$ such that there exist $g \in \mathscr{U}$ and $h \in -\mathscr{U}$ with $\mu(g)$ and $\mu(h)$ finite and $h \leq f \leq g$. For $f \in \mathscr{V}$, we put
$$\mu^*(f) = \inf\{\mu(g) : g \in \mathscr{U} \text{ and } g \geq f\} \in \mathbb{R},$$
$$\mu_*(f) = \sup\{\mu(h) : h \in -\mathscr{U} \text{ and } h \leq f\} \in \mathbb{R}.$$
The following properties follow easily from steps 3 and 4:
   - For every $f \in \mathscr{V}$ and every nonnegative real λ we have $\mu_*(f) \leq \mu^*(f)$, $\mu^*(-f) = -\mu_*(f)$, $\mu^*(\lambda f) = \lambda\mu^*(f)$, and $\mu_*(\lambda f) = \lambda\mu_*(f)$.
   - For every pair (f_1, f_2) of elements of $\mathscr{V}$, we have $\mu^*(f_1 + f_2) \leq \mu^*(f_1) + \mu^*(f_2)$ and $\mu_*(f_1 + f_2) \geq \mu^*(f_1) + \mu_*(f_2)$.
   - For every pair (f_1, f_2) of elements of $\mathscr{V}$ such that $f_1 \leq f_2$, we have $\mu^*(f_1) \leq \mu^*(f_2)$ and $\mu_*(f_1) \leq \mu_*(f_2)$.
6. We extend μ to the set $L^1 = \{f \in \mathscr{V} : \mu^*(f) = \mu_*(f)\}$ by putting $\mu(f) = \mu^*(f) = \mu_*(f)$, for $f \in L^1$. This definition is clearly consistent with the ones given in steps 3 and 4 for elements of $\mathscr{U}$ and $-\mathscr{U}$. Note that L^1 is a vector space containing L and that μ is a positive linear form on L^1.
7. *Some properties of L^1 and μ*
   a. *The vector space L^1 is a lattice.*

   Proof. Notice first that an element f of $\mathscr{F}$ belongs to L^1 if and only if for all $\varepsilon > 0$ there exist $g \in \mathscr{U}$ and $h \in -\mathscr{U}$ such that $h \leq f \leq g$ and $\mu(g) - \mu(h) = \mu(g-h) \leq \varepsilon$.

   Now take $f \in L^1$ and $\varepsilon > 0$, and choose $g \in \mathscr{U}$ and $h \in -\mathscr{U}$ as just described. Then g^+ and h^- are in $\mathscr{U}$, and g^- and h^+ are in $-\mathscr{U}$; furthermore, $h^+ + g^- \leq |f| \leq h^- + g^+$. On the other hand, $\mu(h^- + g^+) - \mu(h^+ + g^-) = \mu(g-h) \leq \varepsilon$. □

   b. *Let (f_n) be an increasing sequence in L^1 that converges pointwise to a function f. In order that $f \in L^1$, it is necessary and sufficient that $\lim_{n\to+\infty} \mu(f_n) < +\infty$ and that there be an element g of $\mathscr{U}$ such that $f \leq g$. If this is the case, $\mu(f) = \lim_{n\to+\infty} \mu(f_n)$.*

   Proof. The condition is clearly necessary; we show sufficiency. Since $f \geq f_0$, there exists $h \in -\mathscr{U}$ such that $\mu(h)$ is finite and $h \leq f$. At the same time, $\mu_*(f) \geq \lim_{n\to+\infty} \mu(f_n)$. Now take $\varepsilon > 0$. There exists a sequence (g_n) in $\mathscr{U}$ satisfying $f_0 \leq g_0$, $\mu(g_0) \leq \mu(f_0) + \varepsilon/2$ and, for all $n \in \mathbb{N}^*$,
$$f_n - f_{n-1} \leq g_n \quad \text{and} \quad \mu(g_n) \leq \mu(f_n - f_{n-1}) + 2^{-n-1}\varepsilon.$$

Set $l = \inf\left(\sum_{p\in\mathbb{N}} g_p, g\right)$. Then $l \in \mathscr{U}$ by step 1; also, $l \geq f$ and

$$\mu(l) \leq \sum_{p=0}^{+\infty} \mu(g_p) \leq \lim_{n\to+\infty} \mu(f_n) + \varepsilon$$

(see step 3). It follows that $f \in \mathscr{V}$ and

$$\mu^*(f) = \mu_*(f) = \lim_{n\to+\infty} \mu(f_n). \qquad \square$$

c. *Let (f_n) be a sequence in L^1 converging pointwise to f. If there exists an element g of $\mathscr{V}$ such that $|f_n| \leq g$ for all $n \in \mathbb{N}$, then $f \in L^1$ and $\mu(f) = \lim_{n\to+\infty} \mu(f_n)$.*

Proof. Clearly $f \in \mathscr{V}$ and, for all n, the function h_n defined by

$$h_n = \lim_{p\to+\infty} \inf_{n\leq k\leq p} f_k$$

belongs to $\mathscr{V}$. Moreover, $h_n \nearrow f$. We deduce, by an application of 7a and two of 7b, that $f \in L^1$ and $\mu(f) \leq \liminf_{k\to+\infty} \mu(f_k)$. One shows likewise that $\mu(f) \geq \limsup_{k\to+\infty} \mu(f_k)$. $\square$

d. *If $g \in L^1$ and $f \in \mathscr{L}$ satisfy $0 \leq f \leq g$, then $f \in L^1$.*

Proof. Assume $g \in L^1$ satisfies $g \geq 0$. The set

$$\{f \in \mathscr{F} : \inf(f^+, g) \in L^1\}$$

contains L, by steps 6 and 7a; by step 7c, it is closed under pointwise convergence. Therefore it contains $\mathscr{L}$, by Lemma 2.1. This implies the desired result: if $f \in \mathscr{L}$ and $0 \leq f \leq g$, then $f = f^+ = \inf(f^+, g) \in L^1$. $\square$

8. *Definition of the measure m.* For $A \in \sigma(L)$, we set $m(A) = \mu(1_A)$ if $1_A \in L^1$ and $m(A) = +\infty$ otherwise. All that remains to do is prove that m satisfies the properties stated in the theorem.

- σ-additivity of m. If A and B are disjoint elements of $\sigma(L)$, there are two possibilities: either 1_A and 1_B are both in L^1, in which case $m(A \cup B) = m(A) + m(B)$; or one of 1_A and 1_B is not in L^1, in which case neither is $1_{A\cup B}$ (by step 7d), and we still have $m(A \cup B) = m(A) + m(B)$. Now let (A_n) be an increasing sequence of elements of $\sigma(L)$, with union A. If all the 1_{A_n} are in L^1, we have $\lim_{n\to+\infty} m(A_n) = m(A)$ by step 7b; otherwise, by 7d, we have $1_A \notin L^1$ and $1_{A_n} \notin L^1$ for large enough n, and $\lim_{n\to+\infty} m(A_n) = +\infty = m(A)$.
- Finally, take $f \in L^1 \cap \mathscr{L}$ with $f \geq 0$. The function f is the pointwise limit of an increasing sequence of piecewise constant positive functions that belong to $\mathscr{L}$, and so also to L^1 by step 7d. By applying

the Monotone Convergence Theorem to the measure m and using property 7b for μ, we conclude that $f \in \mathscr{L}^1(m)$ and $\int f\,dm = \mu(f)$, and in fact that this equality holds for all $f \in L^1 \cap \mathscr{L}$ and so for $f \in L$ since $L \subset L^1 \cap \mathscr{L}$. This proves Theorem 2.3. $\square$

The next proposition follows quickly from the preceding proof.

Proposition 2.4 *Under the same assumptions and with the same notation as in Theorem 2.3, the space L is dense in the Banach space $L^1(m)$.*

Proof. We maintain the same notation. It suffices to show that if A is in $\sigma(L)$ and $m(A)$ is finite then for every $\varepsilon > 0$ there exists an element φ of L such that $\mu(|1_A - \varphi|) < \varepsilon$. If $\varepsilon > 0$, there exists $\psi \in \mathscr{U}$ such that $1_A \leq \psi$ and $\mu(\psi) \leq \mu(1_A) + \varepsilon/2$. Now let $\varphi \in L$ be such that $\varphi \leq \psi$ and $\mu(\psi) \leq \mu(\varphi) + \varepsilon/2$. Since $|1_A - \varphi| \leq (\psi - 1_A) + (\psi - \varphi)$ and $\psi \in L^1$ by step 7b, the desired result follows. $\square$

Exercises

1. **a.** Let Ω be a set and Σ a σ-algebra on Ω (recall that the pair (Ω, Σ) is then called a *measure space*). Let L be a vector subspace of the space of real-valued Σ-measurable functions, such that L is a lattice, $\sigma(L) = \Sigma$, and L contains an increasing sequence that converges pointwise to 1.

   i. Let m_1 and m_2 be measures on (Ω, Σ). Prove that, if $L \subset \mathscr{L}^1(m_1) \cap \mathscr{L}^1(m_2)$ and $\int f\,dm_1 = \int f\,dm_2$ for all $f \in L$, then $m_1 = m_2$.

   ii. Let m be a measure on (Ω, Σ) and h a complex-valued Σ-measurable function such that, for all $f \in L$, the product fh is in $\mathscr{L}^1(m)$ and $\int fh\,dm = 0$. Prove that h vanishes m-almost everywhere.

   b. Assume that $\Omega = \mathbb{R}^d$ and that Σ is the Borel σ-algebra. Let Q be the set of subsets of $\mathbb{R}^d$ of the form $[a_1, b_1] \times \cdots \times [a_d, b_d]$, with $a_j, b_j \in \mathbb{R}$ and $a_j \leq b_j$. A Borel function h from $\mathbb{R}^d$ to $\mathbb{C}$ is called *locally integrable* if $\int_C |h(x)|\,dx < +\infty$ for all $C \in Q$, where dx is Lebesgue measure on $\mathbb{R}^d$. Prove that if a locally integrable function $h : \mathbb{R}^d \to \mathbb{C}$ satisfies $\int_C h(x)\,dx = 0$ for all $C \in Q$, it vanishes dx-almost everywhere.
   Hint. Prove that $\int f(x)h(x)\,dx = 0$ for all $f \in C_c^{\mathbb{R}}(\mathbb{R}^d)$.

   c. Let m be a Borel measure on $\mathbb{R}$ and let h be an m-integrable Borel function from $\mathbb{R}$ to $\mathbb{C}$. Prove that if

$$\int e^{ixy} h(y)\,dm(y) = 0 \quad \text{for all } x \in \mathbb{R},$$

   then h vanishes m-almost everywhere.

Hint. Prove, using Fubini's Theorem and Exercise 8e on page 42, that $\int f(y)h(y)\,dm(y) = 0$ for all $f \in C_0(\mathbb{R})$.

d. Prove likewise that, if m_1 and m_2 are Borel measures of finite mass on $\mathbb{R}$ such that

$$\int e^{ixy}\,dm_1(y) = \int e^{ixy}\,dm_2(y) \quad \text{for all } x \in \mathbb{R},$$

then $m_1 = m_2$.

2. *The monotone class theorem.* Let Ω be a set. A subset $\mathcal{T}$ of $\mathcal{P}(\Omega)$ is called a *monotone class* if it satisfies the following properties:

- $\Omega \in \mathcal{T}$.
- If $T, S \in \mathcal{T}$ and $T \subset S$, then $S \setminus T \in \mathcal{T}$.
- For every increasing sequence $(T_n)_{n\in\mathbb{N}}$ in $\mathcal{T}$, the set $\bigcup_{n\in\mathbb{N}} T_n$ is in $\mathcal{T}$.

Let $\mathcal{G}$ be a subset of $\mathcal{P}(\Omega)$ closed under finite intersections (this means that the intersection of two elements of $\mathcal{G}$ is in $\mathcal{G}$). Show that the smallest monotone class containing $\mathcal{G}$ is closed under finite intersections, and therefore is a σ-algebra.

Hint. Use for inspiration the proof of Lemma 2.1 on page 58. More precisely, denote by $\mathcal{T}$ the smallest monotone class containing $\mathcal{G}$; show first that the set of $T \in \mathcal{T}$ such that $T \cap A \in \mathcal{T}$ for all $A \in \mathcal{G}$ coincides with $\mathcal{T}$.

3. Let X be a locally compact and separable metric space.

a. Set $L = C_c^{\mathbb{R}}(X)$. Prove that L satisfies the assumptions of this section. In the sequel, as in the proof of Theorem 2.3, we will denote by $\mathcal{U}$ the set of pointwise limits of increasing sequences in L.

b. Take $f \in \mathcal{U}$. Prove that f is *lower semicontinuous* (which means that for all real a the set $\{f > a\}$ is open) and that the set $\{f < 0\}$ is relatively compact.

c. Let f be a lower semicontinuous function from X to $\mathbb{R}$ taking nonnegative values.

i. Prove that, for all point x of X,

$$f(x) = \sup_{\substack{\varphi \in C_c^+(X) \\ \varphi \le f}} \varphi(x).$$

ii. Let (K_n) be a sequence of compact sets exhausting X. Prove that for every $n \in \mathbb{N}^*$ there exists $\varphi_n \in C_c^+(X)$ such that $\varphi_n \le f$ and $\varphi_n(x) > f(x) - 1/n$ for all $x \in K_n$.

iii. Prove that the sequence (φ_n) converges pointwise to f; then prove that $f \in \mathcal{U}$.

d. Let f be a lower semicontinuous function from X to $\mathbb{R}$ such that the set $K = \overline{\{f < 0\}}$ is compact.

i. Prove that f is bounded below.

ii. By reducing to the case treated in the preceding question, deduce that $f \in \mathscr{U}$.

Hint. If K is nonempty, consider the function $f + \varphi$, with $\varphi \in C_c^+(X)$ such that $\varphi = -\inf_{x \in X} f(x)$ on K.

e. Deduce that $\mathscr{U} \cap (-\mathscr{U}) = C_c^{\mathbb{R}}(X)$.

4. *Functions of first Baire class.* A function from $\mathbb{R}$ to $\mathbb{R}$ is *of first (Baire) class* if it is the pointwise limit of a sequence of continuous functions from $\mathbb{R}$ to $\mathbb{R}$. We denote by $\mathscr{B}$ the set of such functions. If f is a function from $\mathbb{R}$ to $\mathbb{R}$, we write $\|f\| = \sup_{x \in \mathbb{R}} |f(x)|$ (so $\|f\|$ can be $+\infty$). We say that a function f from $\mathbb{R}$ to $\mathbb{R}$ is *F_σ-measurable* if, for every open subset U of $\mathbb{R}$, the set $f^{-1}(U)$ is an F_σ, that is, a union of countably many closed subsets of $\mathbb{R}$.

a. Prove that the uniform limit of a sequence of functions of first class is a function of first class.

Hint. Let (f_n) be a sequence of elements of $\mathscr{B}$ that converges uniformly to a function f. After passing to a subsequence if necessary, we may assume that $\|f - f_n\| \le 2^{-n}$ for every $n \in \mathbb{N}$. Thus f is the uniform limit of the series of functions $\sum_{n \in \mathbb{N}} (f_n - f_{n-1})$ (where by convention $f_{-1} = 0$). Prove that there exists, for each integer $n \ge 1$, a sequence $(\varphi_n^k)_{k \in \mathbb{N}}$ of continuous functions that converges pointwise to $f_n - f_{n-1}$ and satisfies $\|\varphi_n^k\| \le 2^{-n+2}$ for all $k \in \mathbb{N}$. Then prove that the sequence of functions (ψ_n) defined by

$$\psi_n = \varphi_1^n + \varphi_2^n + \cdots + \varphi_n^n$$

converges pointwise to $f - f_0$.

b. Prove that every function of first class is F_σ-measurable.

c. Prove that $\mathscr{B}$ is not closed under pointwise convergence.

Hint. Let $(f_m)_{m \in \mathbb{N}}$ be the sequence in $\mathscr{B}$ defined by

$$f_m(x) = \lim_{n \to +\infty} \cos(m!\,\pi x)^{2n}.$$

Prove that it converges pointwise to the function $1_{\mathbb{Q}}$; then use Exercise 6g-ii on page 23 to show that $1_{\mathbb{Q}} \notin \mathscr{B}$.

d. Let f be a function of first class from $\mathbb{R}$ to $\mathbb{R}$.

i. Let $(U_n)_{n \in \mathbb{N}}$ be a basis of open sets of $\mathbb{R}$ (see Exercise 1 on page 10) and, for each $n \in \mathbb{N}$, set $A_n = f^{-1}(U_n) \setminus \operatorname{Int}(f^{-1}(U_n))$. Prove that all the A_n are F_σ's having empty interior, and that the set of points where f is not continuous is $\bigcup_{n \in \mathbb{N}} A_n$.

ii. Deduce that the set of points where f is continuous is a G_δ (that is, the complement of an F_σ) and is dense in $\mathbb{R}$.

Hint. Use Baire's Theorem, Exercise 6 on page 22.

iii. Use this to give another proof that the function $1_{\mathbb{Q}}$ is not of first class.

e. Let $(U_k)_{k\in\mathbb{N}}$ be a sequence of open sets in $\mathbb{R}$ and set $G = \bigcap_{k\in\mathbb{N}} U_k$. Prove that there exists a function f of first class such that $G = f^{-1}(\{0\})$.

Hint. Prove that, for every $k \in \mathbb{N}$, there exists a continuous function f_k such that $U_k = f_k^{-1}(\mathbb{R}^*)$. Then, for $k \in \mathbb{N}$ and $x \in \mathbb{R}$, set $g_k(x) = \lim_{n\to+\infty} e^{-nf_k^2(x)}$. Prove that the function $f = \sum_{k=0}^{+\infty} 2^{-k} g_k$ satisfies the desired conditions.

f. Let f be a bounded and F_σ-measurable function from $\mathbb{R}$ to $\mathbb{R}$. We wish to show that f is of first class. Choose $(a,b) \in \mathbb{R}^2$ such that $a < b$ and $f(\mathbb{R}) \subset [a,b]$. Choose also $\varepsilon > 0$ and a subdivision $(a_0 = a, a_1, \dots, a_n = b)$ of $[a,b]$ with step at most ε (this means that $0 \le a_i - a_{i-1} \le \varepsilon$ for $1 \le i \le n$).

i. Prove that, for each $i \in \{1,\dots,n\}$, there exists $f_i \in \mathscr{B}$ such that $f_i^{-1}(\{0\}) = \{a_{i-1} \le f \le a_i\}$. In the sequel we will also write $f_0 = f_{n+1} = 1$.

ii. For each $i \in \{1,\dots,n\}$, set

$$\varphi_i = \prod_{j=0}^{i-1} f_j, \qquad \psi_i = \prod_{j=i+1}^{n+1} f_j, \qquad g_i = \frac{\varphi_i^2}{\varphi_i^2 + \psi_i^2}.$$

Prove that $g_i \in \mathscr{B}$. (Note that $\varphi_i^2 + \psi_i^2$ is never zero.)

iii. Set $g = a_0 + \sum_{i=1}^{n} (a_i - a_{i-1}) g_i$. Prove that $g \in \mathscr{B}$ and that $\|g - f\| \le \varepsilon$.

iv. Prove that $f \in \mathscr{B}$.

g. Prove that every F_σ-measurable function f is of first class.
Hint. If f is unbounded, consider $\tilde{f} = (1 + e^f)^{-1}$.

h. A function from $\mathbb{R}$ to $\mathbb{R}$ is *of second (Baire) class* if it is the pointwise limit of a sequence of functions of first class. (Earlier we saw an example of a function of second class that is not of first class). By working as in the preceding questions, prove that a function f is of second class if and only if the inverse image under f of every open set in $\mathbb{R}$ is a countable union of G_δ sets.

5. *Infinite product of measures, σ-compact case.* Let $X = \mathbb{R}^{\mathbb{N}}$ be the set of sequences $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$, endowed with the product distance (defined on page 13). Consider a measure μ on the Borel σ-algebra of $\mathbb{R}$ satisfying $\mu(X) = 1$ — in other words, a *probability measure* on $\mathbb{R}$. Denote by L the set of functions φ on X for which there exist an integer $n \in \mathbb{N}$ and a function $f \in C_b^{\mathbb{R}}(\mathbb{R}^{n+1})$ such that $\varphi(x) = f(x_0,\dots,x_n)$. Define a linear form Φ on L by setting, for $\varphi(x) = f(x_0,\dots,x_n)$,

$$\Phi(\varphi) = \int_{\mathbb{R}^{n+1}} f(x_0,\dots,x_n)\, d\mu(x_0)\dots d\mu(x_n).$$

a. Prove that Φ is well-defined on L (note that the representation $\varphi(x) = f(x_0, \dots, x_n)$ is not unique).

b. Prove that the set L satisfies the conditions of page 58.

c. Let $\mathscr{B}(X)$ be the Borel σ-algebra of the space X.

i. Let D be a countable and dense subset of X, and let $\mathscr{U}_D$ be the basis of open sets of X defined in Exercise 1b on page 15 (with $X_p = \mathbb{R}$ for all p). Prove that $\mathscr{U}_D \subset \sigma(L)$ and deduce that $\mathscr{B}(X) \subset \sigma(L)$ (use Exercise 1a on page 10).

ii. Prove that all elements of L are continuous functions on X and deduce that $\mathscr{B}(X) = \sigma(L)$.

d. We wish to show that condition 2 of Daniell's Theorem is satisfied.

i. Take $a \in (0,1)$. Prove that, for all $n \in \mathbb{N}$, there exists a compact K_n of $\mathbb{R}$ such that $\mu(K_n) \geq 1 - a^{n+1}$. Then put $K^{(n)} = \prod_{j=0}^{n} K_j$ and $K = \prod_{j=0}^{+\infty} K_j$. Thus, for each n, the set $K^{(n)}$ is compact in $\mathbb{R}^{n+1}$ and K is compact in X (by Tychonoff's Theorem).

ii. Prove that, for all $n \in \mathbb{N}$,

$$\int 1_{\mathbb{R}^{n+1} \setminus K^{(n)}}(x_0, \dots, x_n)\, d\mu(x_0) \dots d\mu(x_n) \leq \frac{a}{1-a}.$$

Hint. Check that the set $\mathbb{R}^n \setminus K^{(n)}$ is contained in the union of the sets $(\mathbb{R} \setminus K_0) \times \mathbb{R}^{n-1}$, $\mathbb{R} \times (\mathbb{R} \setminus K_1) \times \mathbb{R}^{n-2}$, $\mathbb{R}^2 \times (\mathbb{R} \setminus K_2) \times \mathbb{R}^{n-3}, \dots$.

iii. Let $(\varphi_k)_{k \in \mathbb{N}}$ be a decreasing sequence in L converging pointwise to 0. Prove that, for all $k \in \mathbb{N}$,

$$\Phi(\varphi_k) \leq \sup_{x \in K} \varphi_k(x) + \|\varphi_k\| \frac{a}{1-a},$$

where $\|\cdot\|$ denotes the uniform norm on X. Deduce that

$$\lim_{k \to +\infty} \Phi(\varphi_k) = 0.$$

(You might apply Dini's Lemma (see page 29) to the compact space K, then make a vary.)

e. Show that there exists a unique probability measure ν on X such that

$$\int_{\mathbb{R}^{n+1}} f(x_0, \dots, x_n)\, d\mu(x_0) \dots d\mu(x_n) = \int_X f(x_0, \dots, x_n)\, d\nu(x),$$

for all $n \in \mathbb{N}$ and $f \in C_b^{\mathbb{R}}(\mathbb{R}^{n+1})$. This measure is denoted $\nu = \mu^{\mathbb{N}}$.

f. More generally, let $(X_n)_{n \in \mathbb{N}}$ be a sequence of σ-compact metric spaces, each X_n having a probability measure μ_n. Prove that there exists a unique probability measure ν on the space $X = \prod_{n\in\mathbb{N}} X_n$ (endowed with the product distance) satisfying the equality

$$\int_{X^{(n)}} f(x_0,\ldots,x_n)\,d\mu_0(x_0)\ldots d\mu_n(x_n) = \int_X f(x_0,\ldots,x_n)\,d\nu(x)$$

for all $n \in \mathbb{N}$ and $f \in C_b^{\mathbb{R}}(X^{(n)})$ $\bigl($where $X^{(n)} = \prod_{j=0}^n X_j\bigr)$.

3 Positive Radon Measures

In all of this section we consider a locally compact and separable metric space X. We denote by $\mathscr{B}(X)$ the Borel σ-algebra of X. A **Borel measure** on X is a measure on $\mathscr{B}(X)$. If m is a Borel measure, the **mass** of m is, by definition, $m(X) = \int dm \le +\infty$. The measure m is **finite on compact sets** if $m(K)$ is finite for every compact K of X.

Proposition 3.1 *Let m be a Borel measure on X. There exists a largest open set O such that $m(O) = 0$.*

The complement of this set is called the **support** of m, written $\operatorname{Supp}(m)$.

Proof. Let $\mathscr{U}$ be the set of all open sets Ω of X such that $m(\Omega) = 0$. This set is nonempty since it contains $\varnothing$. Set $O = \bigcup_{\Omega\in\mathscr{U}} \Omega$; this is an open set, which we must prove has m-measure zero. If K is compact and contained in O, it can be covered by finitely many elements of $\mathscr{U}$. Each of these elements has measure zero, so $m(K) = 0$. But O is σ-compact (being locally compact and separable), so it too has measure zero, by the σ-additivity of m. □

Examples

1. For $a \in X$, the **Dirac measure** at a is the measure δ_a that assigns the value 1 to a Borel set A if it contains the point a, and the value 0 otherwise. The support of δ_a is clearly $\{a\}$.
2. Take $X = \mathbb{R}^d$ and let λ_d be Lebesgue measure on X (considered as a Borel measure). Naturally, the support of λ_d is $\mathbb{R}^d$.
3. Take $g \in C^+(\mathbb{R}^d)$ and let m be the Borel measure on $\mathbb{R}^d$ defined by $m(A) = \int g 1_A\, d\lambda_d$, for any Borel set A. Clearly, every Borel function f such that fg is Lebesgue-integrable is m-integrable, and

$$\int f\,dm = \int fg\,d\lambda_d.$$

We now check that the support of m equals the support of g. Using the continuity of g one shows easily that an open set Ω of $\mathbb{R}^d$ has m-measure zero if and only if $g = 0$ on Ω; this is equivalent to $\Omega \subset g^{-1}(\{0\})$. In the notation of Proposition 3.1, this implies that $O = \operatorname{Int}\bigl(g^{-1}(\{0\})\bigr)$, so the support of m is the same as that of g.

A **positive Radon measure** on X is a linear form on $C_c^{\mathbb{R}}(X)$ that assigns a nonnegative value to every $f \in C_c^{\mathbb{R}}(X)$ such that $f \geq 0$—in short, a positive linear form on $C_c^{\mathbb{R}}(X)$. We denote by $\mathfrak{M}^+(X)$ the set of positive Radon measures. This set is clearly closed under addition and multiplication by nonnegative scalars. On the other hand, by linearity, if $\mu \in \mathfrak{M}^+(X)$ and if $f, g \in C_c^{\mathbb{R}}(X)$ satisfy $f \leq g$, then $\mu(f) \leq \mu(g)$. As an immediate consequence we have:

Lemma 3.2 *If μ is a positive Radon measure on X,*

$$|\mu(f)| \leq \mu(|f|) \quad \textit{for all } f \in C_c^{\mathbb{R}}(X).$$

If K is compact in X, we denote by $C_K^{\mathbb{K}}(X)$ (or by $C_K(X)$, if no confusion can arise) the set of elements of $C_c^{\mathbb{K}}(X)$ whose support is contained in K. Clearly $C_K^{\mathbb{K}}(X)$ is a subspace of $C_b^{\mathbb{K}}(X)$, closed with respect to the uniform norm $\|\cdot\|$ on $C_b^{\mathbb{K}}(X)$. Henceforth these spaces $C_K^{\mathbb{K}}(X)$ will always be given this Banach space structure induced from the one on $C_b^{\mathbb{K}}(X)$.

Proposition 3.3 *Let μ be a positive Radon measure on X. For every compact set K in X, the restriction of μ to $C_K^{\mathbb{R}}(X)$ is continuous; that is, there exists a constant $C_K \geq 0$ such that*

$$|\mu(f)| \leq C_K \|f\| \quad \textit{for all } f \in C_K^{\mathbb{R}}(X).$$

(We say that μ is **continuous on** $C_c^{\mathbb{R}}(X)$.)

Proof. Let K be compact in X. By Proposition 1.8 on page 53, there exists $\varphi_K \in C_c^+(X)$ such that $0 \leq \varphi_K \leq 1$ and $\varphi_K = 1$ on K. Then, for all $f \in C_K^{\mathbb{R}}(X)$, we have $|f| \leq \|f\| \varphi_K$, and, by Lemma 3.2, $|\mu(f)| \leq \mu(|f|) \leq \|f\| \mu(\varphi_K)$. □

If m is a Borel measure on X finite on compact sets, one immediately checks that the map μ defined on $C_c^{\mathbb{R}}(X)$ by

$$\mu(f) = \int f\, dm \quad \text{for all } f \in C_c^{\mathbb{R}}(X)$$

is a positive Radon measure. The main theorem of this section states, among other things, that all positive Radon measures on X arise in this way:

Theorem 3.4 (Radon–Riesz) *For every positive Radon measure μ on X there exists a unique Borel measure m finite on compact sets and such that*

$$\mu(f) = \int f\, dm \quad \textit{for all } f \in C_c^{\mathbb{R}}(X).$$

The map $\mu \mapsto m$ thus defined is a bijection between $\mathfrak{M}^+(X)$ and the set of Borel measures finite on compact sets, and it commutes with addition and multiplication by nonnegative scalars.

Proof. This will follow as a particular case of Daniell's Theorem. Set $L = C_c^{\mathbb{R}}(X)$. This space satisfies the assumptions stated on page 58: in particular, property $(*)$ follows from Corollary 1.10 on page 53. Now take $\mu \in \mathfrak{M}^+(X)$; we will show that assumption 2 of Theorem 2.3 is satisfied. Let (f_n) be a decreasing sequence in L approaching 0 pointwise. Each f_n has support contained in the compact set $K = \operatorname{Supp} f_0$. Thus, by Dini's Lemma, (f_n) tends to 0 uniformly on K: in other words, $f_n \to 0$ in $C_K^{\mathbb{R}}(X)$. By Proposition 3.3, $\mu(f_n) \to 0$.

Next we check that $\sigma(L) = \mathscr{B}(X)$. Since every continuous function on X is a Borel function, the smallest σ-algebra that makes all elements of L measurable is certainly contained in $\mathscr{B}(X)$; that is, $\sigma(L) \subset \mathscr{B}(X)$. Conversely, $\mathscr{B}(X) \subset \sigma(L)$ because every open subset O of X is $\sigma(L)$-measurable. Indeed, with the notation of Corollary 1.10, an element $x \in X$ belongs to O if and only if there exists $n \in \mathbb{N}$ with $\varphi_n(x) > 0$. Thus O is the (countable) union of the sets $\varphi_n^{-1}\bigl((0, +\infty)\bigr)$, which are $\sigma(L)$-measurable since the functions φ_n are elements of L. Therefore O is $\sigma(L)$-measurable and we finally conclude that $\sigma(L) = \mathscr{B}(X)$.

Finally, we see that a Borel measure m on X is finite on compact sets if and only if $L \subset \mathscr{L}^1(m)$. It now suffices to apply Theorem 2.3 to derive the existence and uniqueness of m. The remaining statements of the theorem are easy to check. □

In the sequel we will often identify a positive Radon measure μ with the Borel measure m it defines. In particular, we use $\mathscr{L}_{\mathbb{K}}^1(\mu)$ or $\mathscr{L}_{\mathbb{K}}^1(m)$ interchangeably for the space of m-integrable $\mathbb{K}$-valued Borel functions, and $L_{\mathbb{K}}^1(\mu)$ or $L_{\mathbb{K}}^1(m)$ for the associated quotient Banach space. As usual, we omit the subscript $\mathbb{K}$ if no confusion is possible. Similarly, we can write $\operatorname{Supp} \mu$ for $\operatorname{Supp} m$, etc.

As a consequence of the preceding proof and of Proposition 2.4, we get:

Proposition 3.5 *Let μ be a positive Radon measure. The space $C_c^{\mathbb{R}}(X)$ is dense in the Banach space $L_{\mathbb{R}}^1(\mu)$.*

This of course implies that $C_c^{\mathbb{C}}(X)$ is dense in $L_{\mathbb{C}}^1(\mu)$.

We now look at positive linear forms on $C_0^{\mathbb{R}}(X)$. Denote by $\mathfrak{M}_f^+(X)$ the set of positive Radon measures μ of finite mass. Note first that a positive Radon measure μ of finite mass can immediately be extended to a linear form $\mathfrak{m}_\mu$ on $C_0^{\mathbb{R}}(X)$; just set, for all $f \in C_0^{\mathbb{R}}(X)$,

$$\mathfrak{m}_\mu(f) = \int f \, d\mu,$$

where, as announced earlier, we make no distinction between the Radon measure and the Borel measure it defines. The linear form $\mathfrak{m}_\mu$ thus defined makes sense (since every continuous function bounded over X is μ-integrable), and it is clearly continuous: its norm in the topological dual of

$C_0^{\mathbb{R}}(X)$ is at most $\mu(X)$. The next proposition asserts essentially that this process yields *all* positive linear forms on $C_0^{\mathbb{R}}(X)$.

Proposition 3.6 *For every positive linear form* $\mathfrak{m}$ *on* $C_0^{\mathbb{R}}(X)$ *there exists a unique positive Radon measure* μ *of finite mass and such that* $\mathfrak{m} = \mathfrak{m}_\mu$, *or equivalently such that*

$$\mathfrak{m}(f) = \int f\,d\mu \quad \textit{for all } f \in C_0^{\mathbb{R}}(X).$$

Thus the map $\mu \mapsto \mathfrak{m}_\mu$ *is a bijection between* $\mathfrak{M}_f^+(X)$ *and the set of positive linear forms on* $C_0^{\mathbb{R}}(X)$.

Proof. The uniqueness of μ clearly follows from the inclusion of $C_c^{\mathbb{R}}(X)$ in $C_0^{\mathbb{R}}(X)$. The important point is existence.

We first show that $\mathfrak{m}$ is continuous. If not, there exists a sequence (f_n) in $C_0^{\mathbb{R}}(X)$ such that, for all n, $\|f_n\| \le 1$ and $|\mathfrak{m}(f_n)| \ge n$. By replacing f_n by $|f_n|$, we can assume that $f_n \in C_0^+(X)$ (note that $\mathfrak{m}(|f_n|) \ge |\mathfrak{m}(f_n)| \ge n$ because $\mathfrak{m}$ is positive). Now set $f = \sum_{n=1}^{+\infty} f_n/n^2$; this function is in $C_0^+(X)$ because the series converges absolutely. But, for all integer $N \ge 1$,

$$\mathfrak{m}(f) \ge \sum_{n=1}^{N} \frac{\mathfrak{m}(f_n)}{n^2} \ge \sum_{n=1}^{N} \frac{1}{n},$$

so $\mathfrak{m}(f) = +\infty$, an impossibility. It follows that $\mathfrak{m}$ is continuous on $C_0^{\mathbb{R}}(X)$.

Its restriction to $C_c^{\mathbb{R}}(X)$ is a positive Radon measure μ. Let (φ_n) be an increasing sequence in $C_c^+(X)$ converging pointwise to 1. By the Monotone Convergence Theorem,

$$\int d\mu = \lim_{n\to+\infty} \int \varphi_n\,d\mu = \lim_{n\to+\infty} \mathfrak{m}(\varphi_n) \le \|\mathfrak{m}\|,$$

where $\|\mathfrak{m}\|$ is the norm of $\mathfrak{m}$ in the topological dual of $C_0^{\mathbb{R}}(X)$. Thus μ has finite mass and $\mathfrak{m}_\mu(f) = \mathfrak{m}(f)$ for all $f \in C_c^{\mathbb{R}}(X)$. Since $C_c^{\mathbb{R}}(X)$ is dense in $C_0^{\mathbb{R}}(X)$ and since $\mathfrak{m}_\mu$ and $\mathfrak{m}$ are continuous, we get $\mathfrak{m} = \mathfrak{m}_\mu$. □

Remark. The preceding proof also shows that the mass $\mu(X)$ of μ equals the norm of the linear form $\mathfrak{m}_\mu$ in $C_0^{\mathbb{R}}(X)'$.

The rest of this section is devoted to examples.

3A Positive Radon Measures on $\mathbb{R}$ *and the Stieltjes Integral*

Let α be an increasing function from $\mathbb{R}$ to $\mathbb{R}$. We will construct from α an integral — in other words, a positive linear form $f \mapsto \int f\,d\alpha$ — generalizing the Riemann integral (which will correspond to the case $\alpha(x) = x$).

First fix $a < b$. For $f \in C^{\mathbb{R}}([a,b])$ and $\Delta = \{x_0 = a < x_1 < \cdots < x_n = b\}$ a subdivision of $[a,b]$ with step $\delta(\Delta) = \max_{1 \le j \le n}(x_j - x_{j-1})$, we write

$$S_\Delta(f) = \sum_{j=0}^{n-1} \bigl(\alpha(x_{j+1}) - \alpha(x_j)\bigr) \max_{x \in [x_j,\, x_{j+1}]} f(x)$$

and

$$\mathfrak{S}_\Delta(f) = \sum_{j=0}^{n-1} \bigl(\alpha(x_{j+1}) - \alpha(x_j)\bigr) \min_{x \in [x_j,\, x_{j+1}]} f(x).$$

One checks easily the inequalities

$$0 \le S_\Delta(f) - \mathfrak{S}_\Delta(f) \le \bigl(\alpha(b) - \alpha(a)\bigr) \max_{\substack{|x-y| \le \delta(\Delta) \\ x,y \in [a,b]}} \bigl|f(x) - f(y)\bigr|,$$

so $\lim_{\delta(\Delta)\to 0}\bigl(S_\Delta(f) - \mathfrak{S}_\Delta(f)\bigr) = 0$ since f is uniformly continuous on $[a,b]$. Next, suppose Δ_1 and Δ_2 are subdivisions of $[a,b]$ with $\Delta_1 \subset \Delta_2$, by which we mean that every subdivision point of Δ_1 is a subdivision point of Δ_2. Then

$$\mathfrak{S}_{\Delta_1}(f) \le \mathfrak{S}_{\Delta_2}(f) \quad \text{and} \quad S_{\Delta_2}(f) \le S_{\Delta_1}(f).$$

It follows from all this that

$$\sup_\Delta \mathfrak{S}_\Delta(f) = \inf_\Delta S_\Delta(f) = \lim_{\delta(\Delta)\to 0} S_\Delta(f) = \lim_{\delta(\Delta)\to 0} \mathfrak{S}_\Delta(f).$$

The common value of these four expressions is denoted by $\int_a^b f\,d\alpha$. Thus,

$$\int_a^b f\,d\alpha = \lim_{\delta(\Delta)\to 0} \sum_{j=0}^{n-1} f(\xi_j)\bigl(\alpha(x_{j+1}) - \alpha(x_j)\bigr),$$

uniformly with respect to sequences $(\xi_0, \ldots, \xi_{n-1})$ such that $\xi_j \in [x_j,\, x_{j+1}]$ for $0 \le j \le n-1$. We deduce that the map from $C^{\mathbb{R}}([a,b])$ to $\mathbb{R}$ defined by $f \mapsto \int_a^b f\,d\alpha$ is a positive linear form.

If $a \le b \le c$ and $f \in C^{\mathbb{R}}([a,c])$, **Chasles's relation** is satisfied:

$$\int_a^c f\,d\alpha = \int_a^b f\,d\alpha + \int_b^c f\,d\alpha.$$

Therefore, if $f \in C_c^{\mathbb{R}}(\mathbb{R})$, the expression $\int_a^b f\,d\alpha$ does not depend on the choice of an interval $[a,b]$ containing the support of f. We denote this expression by $\int f\,d\alpha$. Thus, the map $f \mapsto \int f\,d\alpha$ is a positive Radon measure on $\mathbb{R}$. The associated Borel measure finite on compact sets (Theorem 3.4) is written $d\alpha$, and is called the **Stieltjes measure** associated with α.

Lemma 3.7 *Let α be an increasing function from $\mathbb{R}$ to $\mathbb{R}$. If a and b are real numbers with $a < b$, then*

$$d\alpha((a,b]) = \alpha(b_+) - \alpha(a_+),$$

where $\alpha(a_+)$ and $\alpha(b_+)$ denote the right limits of α at a and b.

Proof. Let $(\varphi_n)_{n\geq 1}$ be a sequence in $C_c^{\mathbb{R}}(X)$ such that $0 \leq \varphi_n \leq 1$, $\varphi_n = 1$ on $[a+1/n,\, b-1/n]$, and $\varphi_n = 0$ on $\mathbb{R} \setminus [a + 1/(n+1),\, b - 1/(n+1)]$. Then

$$\alpha\Big(b - \frac{1}{n}\Big) - \alpha\Big(a + \frac{1}{n}\Big) \leq \int \varphi_n \, d\alpha \leq \alpha\Big(b - \frac{1}{n+1}\Big) - \alpha\Big(a + \frac{1}{n+1}\Big).$$

By passing to the limit, we get

$$d\alpha((a,b)) = \alpha(b_-) - \alpha(a_+), \tag{$*$}$$

where $\alpha(x_-)$ is the left limit of α at x. This is true for any a and b with $a < b$. Applying it to the terms of the sequences (a_n), (b_n) defined by $a_n = b-1/n$, $b_n = b+1/n$ and taking the limit, we deduce that $d\alpha(\{b\}) = \alpha(b_+) - \alpha(b_-)$, which, together with $(*)$, yields the desired relation. □

This formula will allow us to demonstrate that, conversely, every positive Radon measure on $\mathbb{R}$ is a Stieltjes measure.

Theorem 3.8 *Let μ be a positive Radon measure on $\mathbb{R}$. There exists a unique increasing right-continuous function α with $\alpha(0) = 0$ and $\mu = d\alpha$.*

Proof. Uniqueness is clear since, by the preceding discussion, if α is right-continuous and vanishes at 0, it is determined everywhere:

$$\alpha(x) = \begin{cases} -\mu\big((x,0]\big) & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ \mu\big((0,x]\big) & \text{if } x > 0. \end{cases}$$

Conversely, define α by these relations. Then α is right-continuous and vanishes at 0. Also, for $a < b$ we have $\alpha(b) - \alpha(a) = \mu((a,b])$ (one checks the various possible situations of 0 with respect to a and b).

Now suppose $f \in C_c^{\mathbb{R}}(\mathbb{R})$ is supported within $[a,b]$, and let $\Delta = \{x_j\}_{0\leq j\leq n}$ be a subdivision of $[a,b]$. Then

$$\int f \, d\mu = \sum_{j=0}^{n-1} \int f 1_{(x_j,\, x_{j+1}]} d\mu$$

and so, since $\mu((x_j,\, x_{j+1}]) = \alpha(x_{j+1}) - \alpha(x_j)$,

$$\mathfrak{S}_\Delta(f) \leq \int f \, d\mu \leq S_\Delta(f).$$

By taking the limit we deduce that

$$\int f\,d\mu = \int_a^b f\,d\alpha = \int f\,d\alpha,$$

which concludes the proof. □

Remarks

1. By the same reasoning, if μ is a positive Radon measure of finite mass on $\mathbb{R}$, there exists a unique increasing, bounded, right-continuous function α such that $\lim_{x\to-\infty}\alpha(x) = 0$ and $\mu = d\alpha$. It is given by $\alpha(x) = \mu((-\infty, x])$. In this situation α is called the **distribution function** of the measure μ. For example, the distribution function of the Dirac measure δ_a is $Y_a = 1_{[a,+\infty)}$.
2. Suppose α is an increasing function of class C^1 on $\mathbb{R}$. Then

$$\int f\,d\alpha = \int f(x)\alpha'(x)\,dx \quad \text{for all } f \in C_c^{\mathbb{R}}(\mathbb{R}).$$

In short, $d\alpha = \alpha'\,dx$.
Indeed, suppose $f \in C_c^{\mathbb{R}}(\mathbb{R})$ is supported within $[a, b]$ and let $\Delta = \{x_j\}_{0\le j\le n}$ be a subdivision of $[a, b]$. By the Mean Value Theorem, for each $j \in \{0, \dots, n-1\}$ there exists $\xi_j \in [x_j, x_{j+1}]$ such that $\alpha(x_{j+1}) - \alpha(x_j) = \alpha'(\xi_j)(x_{j+1} - x_j)$. Therefore

$$\sum_{j=0}^{n-1} f(\xi_j)\bigl(\alpha(x_{j+1}) - \alpha(x_j)\bigr) = \sum_{j=0}^{n-1} f(\xi_j)\alpha'(\xi_j)(x_{j+1} - x_j).$$

Now it is enough to use the definition of the Stieltjes integral and that of the Riemann integral.

3B Surface Measure on Spheres in $\mathbb{R}^d$

For $r > 0$, we consider the sets

$$B_r = \{x \in \mathbb{R}^d : |x| < r\}, \quad S_r = \{x \in \mathbb{R}^d : |x| = r\}.$$

Here we will denote Lebesgue measure on $\mathbb{R}^d$ simply by λ.

Theorem 3.9 *There exists a unique family $(\sigma_r)_{r\in\mathbb{R}^{+*}}$ of positive Radon measures on $\mathbb{R}^d$ satisfying these conditions:*

1. $\operatorname{Supp}\sigma_r \subset S_r$ *for every* $r > 0$.
2. *For all $f \in C(\mathbb{R}^d)$ and $r > 0$,*

$$\int f(x)\,d\sigma_r(x) = r^{d-1}\int f(ru)\,d\sigma_1(u).$$

3. *For all $f \in C(\mathbb{R}^d)$ and $r > 0$,*

$$\int_{B_r} f(x)\, d\lambda(x) = \int_0^r \left(\int f(x)\, d\sigma_\rho(x) \right) d\rho.$$

We call σ_r, for each $r > 0$, the **surface measure** on S_r.

Proof. Uniqueness. If a family $(\sigma_r)_{r \in \mathbb{R}^{+*}}$ satisfies conditions 2 and 3, we must have, for all $f \in C(\mathbb{R}^d)$,

$$\frac{d}{dr} \int_{B_r} f(x)\, d\lambda(x) \bigg|_{r=1} = \int f(u)\, d\sigma_1(u),$$

which determines uniquely the Radon measure σ_1 and thus also the σ_r, by condition 2. (Note that conditions 2 and 3 are enough to prove uniqueness, so condition 1 is a consequence of 2 and 3.)

Existence. Let φ be the function from $\mathbb{R}^{+*} \times S_1$ to $(\mathbb{R}^d)^*$ defined by $\varphi(r, u) = ru$. Then φ is a homeomorphism and $\varphi^{-1}(x) = \big(|x|, x/|x|\big)$. If A is a Borel set in S_1, we write

$$\tilde{A} = \varphi\big((0,1) \times A\big) = \big\{x \in \mathbb{R}^d : 0 < |x| < 1 \text{ and } x/|x| \in A\big\}.$$

$\tilde{A}$ is a Borel set in $\mathbb{R}^d$. We then put

$$\sigma_1(A) = d \cdot \lambda(\tilde{A}).$$

Visibly σ_1 is a Borel measure of finite mass on S_1, and can also be regarded as a Borel measure on $\mathbb{R}^d$ with support contained in S_1. Next we define, for every Borel set A in S_r,

$$\sigma_r(A) = r^{d-1} \sigma_1(A/r).$$

Likewise, σ_r is a Borel measure supported within S_r. The family (σ_r) thus defined certainly satisfies conditions 1 and 2; we need only check 3.

Let A be a Borel set in S_1 and let r_1, r_2 be real numbers such that $0 < r_1 < r_2$. Then

$$\begin{aligned} \lambda\big(\varphi([r_1, r_2) \times A)\big) &= \lambda\big(\varphi((0, r_2) \times A)\big) - \lambda\big(\varphi((0, r_1) \times A)\big) \\ &= \lambda(r_2 \tilde{A}) - \lambda(r_1 \tilde{A}) = \frac{1}{d}(r_2^d - r_1^d)\sigma_1(A). \end{aligned}$$

On the other hand,

$$\begin{aligned} \int \left(\int 1_{\varphi([r_1, r_2) \times A)}(x)\, d\sigma_\rho(x) \right) d\rho &= \int_{r_1}^{r_2} \sigma_\rho(\rho A) d\rho = \int_{r_1}^{r_2} \rho^{d-1} \sigma_1(A)\, d\rho \\ &= \frac{1}{d}(r_2^d - r_1^d)\sigma_1(A). \end{aligned}$$

Therefore

$$\int 1_{[r_1,r_2)}(|x|)\, 1_A\Big(\frac{x}{|x|}\Big)\, d\lambda(x) = \int 1_{[r_1,r_2)}(\rho)\Big(\int 1_A(x/\rho)\, d\sigma_\rho(x)\Big)\, d\rho,$$

and this for all Borel sets A of S_1 and for any r_1, r_2 with $0 < r_1 < r_2$. It follows that if $0 < a < b$ we have, for all $f \in C([a,b])$ and all $g \in C(S_1)$,

$$\int_{\varphi([a,b]\times S_1)} (f \otimes g) \circ \varphi^{-1} d\lambda = \int_a^b \Big(\int (f \otimes g) \circ \varphi^{-1}\, d\sigma_\rho\Big)\, d\rho.$$

Since $C\big([a,b]\big) \otimes C(S_1)$ is dense in $C\big([a,b] \times S_1\big)$ (see Example 5 on page 35), we obtain, for all $f \in C\big(\varphi([a,b] \times S_1)\big)$,

$$\int_{\varphi([a,b]\times S_1)} f\, d\lambda = \int_a^b \Big(\int f\, d\sigma_\rho\Big)\, d\rho.$$

Since $\varphi\big([a,b] \times S_1\big) = \bar{B}_b \setminus B_a$, this proves condition 3. □

Remarks

1. Since λ is invariant under orthogonal linear transformations, so are the σ_r. In particular, the support of σ_r equals S_r. In fact, up to a multiplicative factor, σ_r is the unique measure supported within S_r and invariant under orthogonal transformations: see Exercise 17 below.
2. Property 3 generalizes to all positive Borel functions on $\mathbb{R}^d$: If f is such a function, then

$$\int f\, d\lambda = \int_0^{+\infty} \Big(\int f\, d\sigma_\rho\Big)\, d\rho = \int_0^{+\infty} \rho^{d-1}\Big(\int f(\rho x)\, d\sigma_1(x)\Big)\, d\rho \le +\infty.$$

   By taking $f = 1_{B_1}$, we obtain, in particular,

$$\int d\sigma_1 = d \cdot \lambda(B_1);$$

   this is the **area** of S_1. Indeed, by the preceding discussion,

$$\lambda(B_1) = \int f\, d\lambda = \int_0^1 \rho^{d-1}\Big(\int d\sigma_1(x)\Big)\, d\rho = \frac{1}{d}\int d\sigma_1(x).$$

   Also, for any nonnegative Borel function h on $\mathbb{R}^+$,

$$\int_{\mathbb{R}^d} h(|x|)\, dx = \Big(\int d\sigma_1\Big) \int_0^{+\infty} \rho^{d-1} h(\rho) d\rho \le +\infty$$

   since $\int h(|\rho x|)\, d\sigma_1(x) = h(\rho) \int d\sigma_1$.

Exercises

Unless otherwise stated, X is a locally compact separable metric space.

1. Let μ be a positive Radon measure on X. Show that $\operatorname{Supp}\mu$ is the complement of the largest open subset O of X such that any function f in $C_c^{\mathbb{R}}(X)$ with support contained in O satisfies $\mu(f) = 0$.
2. Prove that Proposition 3.1 holds when X is any separable metric space, not necessarily locally compact.

   Hint. Use the existence of a countable basis of open sets (Exercise 1 on page 10).
3. *A particular case of the Vitali–Carathéodory Theorem.* Let μ be a positive Radon measure on X. Prove that for every μ-integrable and bounded function f from X to $\mathbb{R}$ and for all $\varepsilon > 0$, there exists an upper semicontinuous function u and a lower semicontinuous function v such that $u \le f \le v$ and $\int (v-u)\,d\mu \le \varepsilon$. (We say that u is *upper semicontinuous* if $-u$ is lower semicontinuous.)

   Hint. Go over the proof of Daniell's Theorem (page 59) and use the result in Exercise 3 on page 64.
4. Let μ be a positive Radon measure on X and take $f \in L^1_{\mathbb{R}}(\mu)$. Prove that there exist μ-integrable and lower semicontinuous functions f_+ and f_- with values in $[0, +\infty]$, such that $f = f_+ - f_-$ μ-almost everywhere. (As in the case of real-valued functions (Exercise 3 on page 64), a function g with values in $[-\infty, +\infty]$ is called lower semicontinuous if the set $\{g > a\}$ is open for all $a \in \mathbb{R}$.)

   Hint. Show that there exists a sequence (φ_n) in $C_c^{\mathbb{R}}(\mathbb{R})$ that converges to f in $L^1_{\mathbb{R}}(\mu)$ and μ-almost everywhere and such that $\mu\bigl(|\varphi_n - \varphi_{n+1}|\bigr) \le 2^{-n}$ for all $n \in \mathbb{N}$. Then set $f_+ = \varphi_0^+ + \sum_{n=0}^{+\infty}(\varphi_{n+1} - \varphi_n)^+$ and $f_- = \varphi_0^- + \sum_{n=0}^{+\infty}(\varphi_{n+1} - \varphi_n)^-$.
5. *Regularity of Radon measures.* (This is a sequel to Exercise 3 on page 64.) Let μ be a positive Radon measure on X.

   a. Prove that, for every Borel set A of X,
   $$\mu(A) = \inf\left\{\int h\,d\mu \;:\; h \text{ is lower semicontinuous and } h \ge 1_A\right\}.$$

   b. Let A be a Borel set in X such that $\mu(A)$ is finite.

   i. Take $\varepsilon > 0$. Let h be a lower semicontinuous function such that $h \ge 1_A$ and $\int h\,d\mu \le \mu(A) + \varepsilon$, and set
   $$U = \left\{x \in X : h(x) > \frac{\mu(A)+\varepsilon}{\mu(A)+2\varepsilon}\right\}.$$
   Prove that $A \subset U$ and that $\mu(U) \le \mu(A) + 2\varepsilon$.

ii. Deduce that

$$\mu(A) = \inf\{\mu(U) : U \text{ is open and } U \supset A\}.$$

iii. Check that this is still true if $\mu(A) = \infty$ (this is obvious). A measure μ satisfying this equality for all Borel sets A is called *outer regular*.

c. Let U be an open subset of X. Prove that

$$\mu(U) = \sup\{\mu(K) : K \text{ is compact and } K \subset U\}.$$

Hint. U is σ-compact.

d. Let A be a Borel set of finite measure $\mu(A)$.

i. Let $\varepsilon > 0$. Justify the existence of:
- an open set U in X containing A and such that $\mu(U) \le \mu(A) + \varepsilon$;
- an open set V in X containing $U \setminus A$ and such that $\mu(V) \le 2\varepsilon$;
- a compact set K in X contained in U and such that $\mu(K) \ge \mu(U) - \varepsilon$.

Finally, set $C = K \setminus V$. Prove that $C \subset A$ and that $\mu(C) \ge \mu(A) - 3\varepsilon$.

ii. Deduce that

$$\mu(A) = \sup\{\mu(K) : K \text{ is compact and } K \subset A\}.$$

iii. Generalize to the case of an arbitrary Borel set A. A measure μ satisfying this equality for all Borel sets A is called *inner regular*. *Hint.* By exhausting X with a sequence of compact sets, prove that A is the union of an increasing sequence of Borel sets of finite measure.

e. i. Prove that for every Borel set A of X and all $\varepsilon > 0$ there exists an open set U in X such that $A \subset U$ and $\mu(U \setminus A) \le \varepsilon$.

ii. Prove that for every Borel set A of X and all $\varepsilon > 0$ there exists an open set U and a closed set F in X such that $F \subset A \subset U$ and $\mu(U \setminus F) \le \varepsilon$.
Hint. Apply the preceding result to A and $X \setminus A$.

6. *Lusin's Theorem.* Let m be a positive Radon measure on X.

a. Let f be a Borel function on X with values in $[0, 1]$. Prove that, for any open set O of finite measure and any $\varepsilon > 0$, there exists a compact $K \subset O$ such that $m(O \setminus K) < \varepsilon$ and the restriction $f|_K$ is continuous on K.
Hint. Use Proposition 3.5, Exercise 15 on page 155 and the fact that m is inner regular (see Exercise 5d).

b. Extend the preceding result to all Borel functions f from X to $\mathbb{K}$.
Hint. First reduce to the case where f takes values in $\mathbb{R}^+$, then consider $\tilde{f} = f/(1 + f)$.

c. Deduce that every Borel function f from X to $\mathbb{K}$ satisfies this property:

(L) For every $\varepsilon > 0$, there exists an open set ω in X such that $m(\omega) < \varepsilon$ and the restriction of f to $X \setminus \omega$ is continuous.

Hint. Consider an increasing sequence $(O_n)_{n\in\mathbb{N}}$ of relatively compact open sets that covers X. For each n, there exists a compact $K_n \subset O_n$ for which $m(O_n \setminus K_n) < \varepsilon 2^{-n-1}$ and $f|_{K_n}$ is continuous. Now set $\omega = \bigcup_n (O_n \setminus K_n)$. Prove that $(X \setminus \omega) \cap O_n \subset K_n$ for every n; then conclude the proof.

d. Show that a function f from X to $\mathbb{K}$ satisfies Property L if and only if there exists a Borel function that equals f m-almost everywhere.

Hint. To prove sufficiently, use the fact that m is outer regular (Exercise 5b).

7. a. Let μ be a positive Radon measure on X, with support F. Let $f \in C_c(X)$ be such that $f(x) = 0$ for all $x \in F$. Prove that $\int f\, d\mu = 0$.

b. Let $A = \{a_n\}_{n\le N}$ be a finite subset of X and μ a positive Radon measure on X. Prove that the support of μ equals A if and only if μ is a linear combination of Dirac measures δ_{a_n} with positive coefficients.

c. Let $A = \{a_n\}$ be a countable subset of X. For $f \in C_c(X)$ write

$$\mu(f) = \sum_{n\in\mathbb{N}} 2^{-n} f(a_n).$$

Prove that μ is a positive Radon measure on X whose support is the closure of A.

8. a. Let F be a closed subset of X. Prove that F is the support of a continuous function f from X to $\mathbb{R}$ if and only if F coincides with the closure of $\mathring{F}$.

b. Let μ be a positive Radon measure on X. We denote by $\mathscr{L}^1_{\text{loc}}(\mu)$ the space of *locally μ-integrable* functions on X, by which we mean Borel functions $\psi : X \to \mathbb{K}$ such that $1_K \psi \in \mathscr{L}^1(\mu)$ for any compact K of X. (For example, every continuous function on X is locally μ-integrable.) Fix a $\psi \in \mathscr{L}^1_{\text{loc}}(\mu)$ taking nonnegative values. For $f \in C_c(X)$, write

$$\nu(f) = \int \psi f\, d\mu.$$

Prove that ν is a positive Radon measure. Prove that

$$\operatorname{Supp} \nu \subset \overline{\{\psi \neq 0\} \cap \operatorname{Supp} \mu},$$

with equality if ψ is continuous.

c. For $f \in C_c(\mathbb{R}^2)$, write

$$\nu(f) = \int_{\mathbb{R}} f(x, x)\, dx.$$

Prove that ν is a positive Radon measure on $\mathbb{R}^2$ and determine its support.

Is there a continuous function ψ on $\mathbb{R}^2$ such that

$$\nu(f) = \int_{\mathbb{R}^2} f(x,y)\psi(x,y)\,dx\,dy$$

for all $f \in C_c(\mathbb{R}^2)$?

9. a. Let $\mathfrak{m}$ be a positive linear form on $C^{\mathbb{R}}(X)$. Show that there exists a compact K in X such that any $f \in C^{\mathbb{R}}(X)$ that vanishes on K satisfies $\mathfrak{m}(f) = 0$.

Hint. Exhaust X by a sequence (K_n) of compact sets. Show that, if there is no K as stated, there exists a sequence (f_n) of elements of $C^+(X)$ such that, for each $n \in \mathbb{N}$, the function f_n vanishes on K_n and $\mathfrak{m}(f_n) > 0$. Then consider $f = \sum_{n\in\mathbb{N}} f_n/\mathfrak{m}(f_n)$.

b. Let $\mathfrak{M}_c^+(X)$ be the set of positive Radon measures with compact support. To every $\mu \in \mathfrak{M}_c^+(X)$, associate the positive linear form $\mathfrak{m}_\mu$ on $C^{\mathbb{R}}(X)$ defined by

$$\mathfrak{m}_\mu(f) = \int f\,d\mu \quad \text{for } f \in C^{\mathbb{R}}(X).$$

Prove that the map $\mu \mapsto \mathfrak{m}_\mu$ is a bijection between $\mathfrak{M}_c^+(X)$ and the set of positive linear forms on $C^{\mathbb{R}}(X)$.

Hint. See the proof of Proposition 3.6 (page 71) for inspiration.

10. *Vague convergence.* We say that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of positive Radon measures on X *converges vaguely* to $\mu \in \mathfrak{M}^+(X)$ if

$$\mu_n(f) \to \mu(f) \quad \text{for all } f \in C_c(X).$$

a. *An example.* Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in X with no cluster point. Prove that the sequence $(\delta_{a_n})_{n\in\mathbb{N}}$ converges vaguely to 0.

b. *Another example.* Suppose $X = (0,1)$. Prove that the sequence (μ_n) defined by

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n-1} \delta_{k/n}$$

converges vaguely to Lebesgue measure on $(0,1)$.

c. Let (μ_n) be a sequence in $\mathfrak{M}^+(X)$ such that, for all $f \in C_c^+(X)$, the sequence $\big(\mu_n(f)\big)$ converges. Prove that the sequence (μ_n) is vaguely convergent.

d. Let μ be a positive Radon measure and A a relatively compact Borel set whose boundary has μ-measure zero. Prove that, if $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $\mathfrak{M}^+(X)$ that converges vaguely to μ, then

$$\lim_{n\to+\infty} \mu_n(A) = \mu(A).$$

Hint. Show the existence of an increasing sequence in $C_c^+(X)$ that converges pointwise to the characteristic function of $\mathring{A}$, and of a decreasing sequence in $C_c^+(X)$ that converges pointwise to the characteristic function of $\bar{A}$. Then consider the lim sup and lim inf of the sequence $(\mu_n(A))$.

e. Let (μ_n) be a sequence in $\mathfrak{M}^+(X)$ such that

$$\sup_{n\in\mathbb{N}} \int f\,d\mu_n < +\infty \quad \text{for all } f \in C_c^+(X).$$

(Check that this condition is satisfied if and only if $\sup_{n\in\mathbb{N}} \mu_n(K)$ is finite for every compact K of X.)
Prove that the sequence (μ_n) has a vaguely convergent subsequence.
Hint. Exhaust X by a sequence of compact sets (K_p) and apply Corollary 4.2 on page 19 to each of the separable Banach spaces $C_{K_p}(X)$.

11. a. Let (f_n) be a sequence of increasing functions from $\mathbb{R}$ to $\mathbb{R}$ such that the series $\sum f_n$ converges pointwise on $\mathbb{R}$ to a function f. Prove that the series $\sum_{n=0}^{+\infty} df_n$ converges vaguely to df (see Exercise 10).
Hint. Consider $\varphi \in C_c^+(\mathbb{R})$, a compact interval $[a, b]$ in $\mathbb{R}$ containing the support of φ, and a subdivision $\{x_j\}_{0\le j\le n}$ of $[a, b]$. Prove that, for every integer $l \in \mathbb{N}$,

$$\begin{aligned}\sum_{j=0}^{n-1} \varphi(x_j)\big(f(x_{j+1}) - f(x_j)\big) - \|\varphi\| \sum_{k=l+1}^{+\infty} \big(f_k(b) - f_k(a)\big)& \\ \le \sum_{k=0}^{l}\sum_{j=0}^{n-1} \varphi(x_j)\big(f_k(x_{j+1}) - f_k(x_j)\big)& \\ \le \sum_{j=0}^{n-1} \varphi(x_j)\big(f(x_{j+1}) - f(x_j)\big).&\end{aligned}$$

b. *Example.* Let (a_n) be a sequence in $\mathbb{R}$ and (c_n) a sequence in $\mathbb{R}^+$ such that $\sum_{n\in\mathbb{N}} c_n < +\infty$. Prove that the series of measures $\sum_{n\ge0} c_n\delta_{a_n}$ converges vaguely to a positive Radon measure whose distribution function is $f = \sum_{n=0}^{+\infty} c_n Y_{a_n}$, where $Y_{a_n} = 1_{[a_n,+\infty)}$.

12. *Narrow convergence.* We say that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of positive Radon measures of finite mass on X *converges narrowly* to $\mu \in \mathfrak{M}_f^+(X)$ if

$$\mu_n(f) \to \mu(f) \quad \text{for all } f \in C_b(X).$$

Every narrowly convergent sequence is vaguely convergent (Exercise 10).

a. *A counterexample.* Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in X with no cluster point. Prove that the sequence $(\delta_{a_n})_{n\in\mathbb{N}}$ does not converge narrowly to 0.

b. Let μ be a positive Radon measure of finite mass and A a Borel set whose boundary has μ-measure zero. Prove that, if $(\mu_n)_{n\in\mathbb{N}}$ is a sequence in $\mathfrak{M}_f^+(X)$ that converges narrowly to μ, then

$$\lim_{n\to+\infty} \mu_n(A) = \mu(A).$$

Hint. Work as in Exercise 10 above.

c. Let (μ_n) be a sequence in $\mathfrak{M}_f^+(X)$ and suppose $\mu \in \mathfrak{M}_f^+(X)$. Prove that the sequence (μ_n) converges narrowly to μ if and only if it converges vaguely to μ and $\lim_{n\to+\infty} \mu_n(X) = \mu(X)$.
Hint. For the "if" part, fix $f \in C_b^+(X)$ and $\varepsilon > 0$. Show that there exists a function $\alpha \in C_c^+(X)$ such that $\alpha \leq 1$ and $\int (1-\alpha)\, d\mu \leq \varepsilon$; then write

$$\mu_n(f) - \mu(f) = \mu_n(\alpha f) - \mu(\alpha f) + \mu_n\big((1-\alpha)f\big) - \mu\big((1-\alpha)f\big).$$

d. *Theorem of P. Lévy.* If ν is a positive Radon measure of finite mass on $\mathbb{R}$, we denote by $\hat{\nu}$ the function defined on $\mathbb{R}$ by

$$\hat{\nu}(x) = \int e^{itx}\, d\nu(t).$$

Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathfrak{M}_f^+(\mathbb{R})$ and μ an element of $\mathfrak{M}_f^+(\mathbb{R})$. Prove that (μ_n) converges narrowly to μ if and only if the sequence of functions $(\hat{\mu}_n)$ converges pointwise to $\hat{\mu}$.
Hint. Prove that if $(\hat{\mu}_n)$ converges pointwise to $\hat{\mu}$, then $(\int d\mu_n)$ converges to $\int d\mu$ and there exists a dense subspace H in $C_0^{\mathbb{C}}(\mathbb{R})$ such that

$$\lim_{n\to+\infty} \int h\, d\mu_n = \int h\, d\mu \quad \text{for all } h \in H$$

(see Exercise 8e on page 42). Conclude with Proposition 4.3 on page 19.

13. a. Let μ be a positive Radon measure on X. Suppose the support K of μ is compact. Show that there exists a sequence (μ_n) of Radon measures of finite support contained in K that converges narrowly to μ (see Exercise 12).
Hint. Take $n \in \mathbb{N}^*$. Construct a partition of K into finitely many nonempty Borel sets $(K_{n,p})_{p\leq P_n}$ of diameter at most $1/n$. Then, for each $p \leq P_N$, choose a point $x_{n,p}$ in $K_{n,p}$ and set

$$\mu_n = \sum_{p\leq P_n} \mu(K_{n,p})\delta_{x_{n,p}}.$$

b. Generalize to the case of any positive Radon measure of finite mass.

14. Let g be a Borel function on $\mathbb{R}$ taking nonnegative values and locally integrable (see Exercise 1b on page 63). Let a be a real number. Consider the function G on $\mathbb{R}$ defined by $G(x) = \int_a^x g(t)\, dt$.

a. Prove that

$$\int f\,dG = \int f(x)g(x)\,dx \quad \text{for all } f \in C_c^{\mathbb{R}}(\mathbb{R}),$$

where dx is Lebesgue measure on $\mathbb{R}$.

Hint. If $[a, b]$ is an interval containing the support of f and $\{x_j\}_{0\le j\le n}$ is a subdivision of $[a, b]$, and if we take for each $j \in \{0, \dots, n-1\}$ a point $\xi_j \in [x_j,\, x_{j+1}]$, then

$$\sum_{j=0}^{n-1} f(\xi_j)\bigl(G(x_{j+1}) - G(x_j)\bigr) = \int_a^b \Bigl(\sum_{j=0}^{n-1} f(\xi_j)1_{(x_j,x_{j+1}]}(x)\Bigr) g(x)\,dx.$$

Now use the Dominated Convergence Theorem.

b. Prove that the equality of the preceding question holds when f is any positive Borel function.

15. Recall that $\int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}$. For all real $t > 0$, put

$$\Gamma(t) = \int_0^{+\infty} x^{t-1}e^{-x}\,dx.$$

Let s_d be the area of the unit sphere in $\mathbb{R}^d$, that is, the mass of the surface measure of the unit sphere in $\mathbb{R}^d$. Prove that $s_d = 2\pi^{d/2}/\Gamma(d/2)$. Deduce the Lebesgue measure of the unit ball in $\mathbb{R}^d$.

Hint. Compute $\int_{\mathbb{R}^d} e^{-|x|^2}dx$ in two ways.

16. Let σ_1 be the surface measure of the unit sphere S_1 in $\mathbb{R}^d$.

a. Suppose $d = 2$. Prove that, for any Borel function f from $\mathbb{R}^2$ to $\mathbb{R}^+$,

$$\int f\,d\sigma_1 = \int_0^{2\pi} f(\cos\theta,\, \sin\theta)\,d\theta.$$

Hint. Use polar coordinates.

b. Suppose $d = 3$. Prove that, for any Borel function f from $\mathbb{R}^3$ to $\mathbb{R}^+$,

$$\int f\,d\sigma_1 = \int_0^{2\pi}\int_{-\pi/2}^{\pi/2} f(\cos\theta\cos\varphi,\, \sin\theta\cos\varphi,\, \sin\varphi)\cos\varphi\,d\theta\,d\varphi.$$

Hint. Use spherical coordinates.

17. Let σ be a positive Radon measure on $\mathbb{R}^d$ whose support is contained in the unit sphere S_1. Assume σ is invariant under orthogonal linear transformations; that is, for any orthogonal endomorphism O of $\mathbb{R}^d$ and any $f \in C(\mathbb{R}^d)$,

$$\int f(Ox)\,d\sigma(x) = \int f(x)\,d\sigma(x).$$

a. Show that there exists a function h_σ from $\mathbb{R}^+$ to $\mathbb{C}$ such that

$$\int e^{iu\cdot y} d\sigma(y) = h_\sigma(|u|) \quad \text{for all } u \in \mathbb{R}^d,$$

where $u \cdot y$ is the scalar product of u and y in $\mathbb{R}^d$. We define h_{σ_1} analogously, starting from the surface measure σ_1 on S_1.

b. Prove that, for all $t \in \mathbb{R}^+$,

$$\int h_\sigma(t|u|)\, d\sigma_1(u) = \int h_{\sigma_1}(t|y|)\, d\sigma(y),$$

and so that $h_\sigma(t) = h_{\sigma_1}(t)\big(\int d\sigma\big)/\big(\int d\sigma_1\big)$.

c. Deduce that

$$\sigma = \frac{\int d\sigma}{\int d\sigma_1}\sigma_1.$$

Hint. Generalize to $\mathbb{R}^d$ the result of Exercise 1d on page 64.

18. *Infinite product of measures, compact case.* Consider the space

$$X = [0,1]^{\mathbb{N}} = \big\{x = (x_n)_{n\in\mathbb{N}} : x_n \in [0,1] \text{ for all } n \in \mathbb{N}\big\},$$

and give it the product metric

$$d(x,y) = \sum_{n=0}^{\infty} 2^{-n}|x_n - y_n|.$$

With this metric, X is compact, by Tychonoff's Theorem. Consider also a sequence $(m_n)_{n\in\mathbb{N}}$ of probabitity measures—that is, Borel measures of mass 1—on $[0,1]$.

a. Show that, for each $n \in \mathbb{N}$, the function that maps $x \in X$ to $x_n \in [0,1]$ is continuous (in fact, Lipschitz).

b. For $n \in \mathbb{N}$, denote by F_n the set of functions from X to $\mathbb{R}$ of the form

$$x \mapsto f(x_0,\ldots,x_n),$$

with $f \in C^{\mathbb{R}}\big([0,1]^{n+1}\big)$. Prove the following facts:

i. F_n is a vector subspace of $C^{\mathbb{R}}(X)$ for all $n \in \mathbb{N}$.

ii. $F_n \subset F_{n+1}$ for all $n \in \mathbb{N}$.

iii. $F = \bigcup_{n\in\mathbb{N}} F_n$ is a dense vector subspace of $C^{\mathbb{R}}(X)$ with the uniform norm $\|\cdot\|$.

c. For each n, we define a linear form μ_n on F_n by associating to the element

$$\varphi : x \mapsto f(x_0,\ldots,x_n)$$

of F_n the real number

$$\mu_n(\varphi) = \int \cdots \int f(x_0,\ldots,x_n)\, dm_0(x_0)\ldots dm_n(x_n).$$

Prove that, if $\varphi \in F_n$, then $\mu_p(\varphi) = \mu_n(\varphi)$ for all $p \geq n$. Deduce the existence of a linear form μ on F such that

$$\mu(\varphi) = \mu_n(\varphi) \quad \text{for all } n \in \mathbb{N} \text{ and } \varphi \in F_n.$$

Then show that, for $\varphi \in F$, we have $|\mu(\varphi)| \leq \|\varphi\|$ and $\varphi \geq 0$ implies $\mu(\varphi) \geq 0$.

d. Prove that the linear form μ extends in a unique way to a positive Radon measure on X.

e. More generally, let $(X_n)_{n\in\mathbb{N}}$ be a sequence of compact metric spaces and, for each $n \in \mathbb{N}$, let m_n be a probability measure on X_n. Let $X = \prod_{n\in\mathbb{N}} X_n$ be the product space, with the product metric. By working as in the preceding questions, prove that there exists a unique probability measure μ on X satisfying

$$\int_{X^{(n)}} f(x_0, \ldots, x_n)\, dm_0(x_0) \ldots dm_n(x_n) = \int_X f(x_0, \ldots, x_n)\, d\mu(x)$$

for all $n \in \mathbb{N}$ and all $f \in C^{\mathbb{R}}(X^{(n)})$, where $X^{(n)} = \prod_{j=0}^{n} X_j$. (We thus recover the result of Exercise 5 on page 66 in this particular case.)

19. *Haar measure on a compact abelian group.* Let X be a compact metric space having an abelian group structure. We assume that addition is continuous as a map from X^2 to X.

We denote by B the set of continuous linear forms on $C^{\mathbb{R}}(X)$ of norm at most 1. We recall from Exercise 4 on page 20 that B can be given a metric d for which $d(\mu_n, \mu) \to 0$ if and only if

$$\lim_{n\to+\infty} \mu_n(f) = \mu(f) \quad \text{for all } f \in C(X),$$

and that the metric space (B, d) is compact. One can check that the set P of positive Radon measures of mass 1 on X is a nonempty, convex, closed subset of B, and that the topology induced by d on P is that of vague convergence.

a. *Markov–Kakutani Theorem.* Let K be a nonempty, compact, convex subset of (B, d).

i. Let φ be a continuous affine transformation from K to K (*affine* means that for any $(\mu, \mu') \in K^2$ and any $\alpha \in [0,1]$ we have $\varphi(\alpha\mu + (1-\alpha)\mu') = \alpha\varphi(\mu) + (1-\alpha)\varphi(\mu')$). Prove that φ has at least one fixed point in K — in other words, there is a point $\Lambda \in K$ such that $\varphi(\Lambda) = \Lambda$.

One can work as follows: Let μ be any element of K and, for any $n \in \mathbb{N}$, set

$$\mu_n = \frac{1}{n+1} \sum_{i=0}^{n} \varphi^i(\mu).$$

A. Check that $\mu_n \in K$ for each $n \in \mathbb{N}$.

B. Let (μ_{n_k}) be a subsequence of the sequence (μ_n) that converges (with respect to d) to $\Lambda \in K$. Prove that, for each integer k, we have $(1+n_k)\big(\varphi(\mu_{n_k}) - \mu_{n_k}\big) \in 2B$.

C. Deduce that $\varphi(\Lambda) = \Lambda$.

ii. Let $\mathscr{T}$ be a family of continuous affine transformations of K such that any two elements of $\mathscr{T}$ commute. For each $\varphi \in \mathscr{T}$ denote by F_φ the set of fixed points of φ.

A. Prove that all the F_φ are nonempty, compact, convex subsets of (B, d).

B. Suppose $\mathscr{T} = \{\varphi, \varphi'\}$. Prove that $\varphi'(F_\varphi) \subseteq F_\varphi$. Deduce that φ and φ' have a common fixed point.

C. Now make no assumption on $\mathscr{T}$. Prove that all the elements of $\mathscr{T}$ have at least one common fixed point. (Start with $\mathscr{T}$ finite, then use compactness.)

b. For $\mu \in \mathfrak{M}^+(X)$ and $x \in X$ we denote by $\tau_x\mu$ the positive Radon measure on X defined by $\tau_x\mu(f) = \int f(x+y)\,d\mu(y)$.

i. Prove that $\tau_x(P) \subset P$ for all $x \in X$. Deduce that there exists $\mu \in P$ such that $\tau_x\mu = \mu$ for all $x \in X$.

ii. Prove that there exists a Borel measure μ on X such that $\mu(X)=1$ and

$$\int f(t)\,d\mu(t) = \int f(x+t)\,d\mu(t) \quad \text{for all } f \in C(X) \text{ and } x \in X.$$

We call μ a *Haar measure* on X.

c. *Uniqueness of Haar measure.* Let μ and ν be Haar measures on X. Prove that $\mu = \nu$.

Hint. Take $f \in C(X)$. Using Fubini's Theorem, compute in two ways the integral

$$\iint f(x+y)\,d\mu(x)\,d\nu(y).$$

4 Real and Complex Radon Measures

The framework here is the same as in the previous section. A **real Radon measure** on X is by definition a linear form μ on $C_c^{\mathbb{R}}(X)$ whose restriction to each space $C_K^{\mathbb{R}}(X)$, for K compact in X, is continuous; that is, such that for any compact K of X there exists a real $C_K \geq 0$ such that

$$|\mu(f)| \leq C_K \|f\| \quad \text{for all } f \in C_K^{\mathbb{R}}(X).$$

We denote by $\mathfrak{M}^{\mathbb{R}}(X)$ the set of real Radon measures. We also call the elements of this set linear forms *continuous on* $C_c^{\mathbb{R}}(X)$; for an equivalent definition of this notion of continuity, see Exercise 5. By Proposition 3.3, $\mathfrak{M}^+(X) \subset \mathfrak{M}^{\mathbb{R}}(X)$. Conversely, every real Radon measure is the difference of two positive Radon measures:

Theorem 4.1 *Let μ be a real Radon measure on X. For each $f \in C_c^+(X)$, put*

$$\mu^+(f) = \sup\{\mu(g) : g \in C_c^+(X) \text{ and } g \leq f\},$$
$$\mu^-(f) = -\inf\{\mu(g) : g \in C_c^+(X) \text{ and } g \leq f\}.$$

Then μ^+ and μ^- can be uniquely extended to positive Radon measures and $\mu = \mu^+ - \mu^-$.

Proof

1. We first check that the definition of $\mu^+(f)$ given in the statement makes sense. If $f \in C_c^+(X)$ has support K, then for all $g \in C_c^+(X)$ such that $g \leq f$ we have $g \in C_K^{\mathbb{R}}(X)$, so
$$\mu(g) \leq |\mu(g)| \leq C_K \|g\| \leq C_K \|f\|.$$
Thus $\mu^+(f)$ is well-defined and $0 \leq \mu^+(f) \leq C_K \|f\|$. It is also clear that for λ real and nonnegative we have $\mu^+(\lambda f) = \lambda \mu^+(f)$.
2. The essential point is the additivity of μ^+ on $C_c^+(X)$. Take $f_1, f_2 \in C_c^+(X)$. That $\mu^+(f_1 + f_2) = \mu^+(f_1) + \mu^+(f_2)$ will follow from the set equality
$$\{g \in C_c^+(X) : g \leq f_1 + f_2\}$$
$$= \{g \in C_c^+(X) : g \leq f_1\} + \{g \in C_c^+(X) : g \leq f_2\}.$$
One of the inclusions is obvious and the other can be checked quickly: Suppose $g \in C_c^+(X)$ satisfies $g \leq f_1 + f_2$. Put $g_1 = \inf(g, f_1)$ and $g_2 = g - g_1 = \sup(0, g - f_1)$. We see that $0 \leq g_1 \leq f_1$, $0 \leq g_2 \leq f_2$, and $g = g_1 + g_2$.
3. The same properties hold for μ^-. On the other hand, if $f \in C_c^+(X)$,
$$\begin{aligned} \mu^+(f) - \mu(f) &= \sup\{\mu(g - f) : g \in C_c^+(X) \text{ and } g \leq f\} \\ &= -\inf\{\mu(f - g) : g \in C_c^+(X) \text{ and } g \leq f\} \\ &= -\inf\{\mu(h) : h \in C_c^+(X) \text{ and } h \leq f\} = \mu^-(f). \end{aligned}$$
Therefore $\mu(f) = \mu^+(f) - \mu^-(f)$.
4. We now extend μ^+ and μ^- to $C_c^{\mathbb{R}}(X)$ in the only possible way: Given $h \in C_c^{\mathbb{R}}(X)$ we take $f, g \in C_c^+(X)$ such that $h = f - g$ (for example, $f = h^+$ and $g = h^-$). Since μ^+ must be linear on $C_c^{\mathbb{R}}(X)$, we must set
$$\mu^+(h) = \mu^+(f) - \mu^+(g).$$

This definition does not depend on the choice of a decomposition for h. For if $h = f' - g'$ with $f', g' \geq 0$, then $f + g' = f' + g$ and, by the additivity of μ^+ on $C_c^+(X)$, we have $\mu^+(f) - \mu^+(g) = \mu^+(f') - \mu^+(g')$. One can easily see that the μ^+ defined in this way is indeed linear and so belongs to $\mathfrak{M}^+(X)$. We extend μ^- similarly, and we use item 3 to show that $\mu = \mu^+ - \mu^-$. □

Remarks

1. The decomposition $\mu = \mu^+ - \mu^-$ defined in Theorem 4.1 is minimal in the following sense: If $\mu = \mu_1 - \mu_2$ with $\mu_1, \mu_2 \in \mathfrak{M}^+(X)$, there exists a positive Radon measure ν on X such that $\mu_1 = \mu^+ + \nu$ and $\mu_2 = \mu^- + \nu$. Indeed, it is clear, in view of the definition of μ^+, that $\mu^+(f) \leq \mu_1(f)$ for all $f \in C_c^+(X)$. One easily deduces from this that the Radon measure on X defined by $\nu = \mu_1 - \mu^+$ is positive. (And of course $\nu = \mu_2 - \mu^-$ as well.)
2. Using the same construction, we obtain an analogous decomposition for continuous linear forms on a normed space E that has an order relation making it into a lattice and satisfying the following conditions, for all $f, g \in E$ and all $\lambda \in \mathbb{R}^{+*}$:
   - $0 \leq g \leq f$ implies $\|g\| \leq \|f\|$;
   - $f \geq 0$ implies $\lambda f \geq 0$;
   - $f \leq g$ if and only if $g - f \geq 0$.

A **bounded real Radon measure** on X is by definition a linear form μ on $C_c^{\mathbb{R}}(X)$ continuous with respect to the uniform norm on $C_c^{\mathbb{R}}(X)$; that is, one for which there exists a constant $C \geq 0$ such that

$$|\mu(f)| \leq C\|f\| \quad \text{for all } f \in C_c^{\mathbb{R}}(X).$$

We denote by $\mathfrak{M}_f^{\mathbb{R}}(X)$ the set of bounded real Radon measures on X; this is clearly a vector subspace of $\mathfrak{M}^{\mathbb{R}}(X)$.

Since $C_c^{\mathbb{R}}(X)$ is dense in the Banach space $C_0^{\mathbb{R}}(X)$ with the uniform norm, every bounded real Radon measure extends uniquely to a continuous linear form on $C_0^{\mathbb{R}}(X)$; this allows us to identify $\mathfrak{M}_f^{\mathbb{R}}(X)$ with the topological dual of $C_0^{\mathbb{R}}(X)$.

Proposition 4.2 *Every bounded real Radon measure is the difference of two positive Radon measures of finite mass. More precisely, if $\mu \in \mathfrak{M}_f^{\mathbb{R}}(X)$, the Radon measures μ^+ and μ^- defined in Theorem* 4.1 *have finite mass and*

$$\|\mu\| = \int d\mu^+ + \int d\mu^-,$$

where $\|\mu\|$ is the norm of μ in the dual of $C_0^{\mathbb{R}}(X)$.

Proof. We first see that, for any $f \in C_c^+(X)$,

$$\begin{aligned}\mu^+(f)+\mu^-(f) &= \sup\{\mu(g-h) : g,h \in C_c^+(X) \text{ and } g,h \le f\}\\ &= \sup\{\mu(\varphi) : \varphi \in C_c^{\mathbb{R}}(X) \text{ and } |\varphi| \le f\}.\end{aligned}$$

In particular, $\mu^+(f)+\mu^-(f) \le \|\mu\|\,\|f\|$. Applying this inequality to all terms of an increasing sequence of functions in $C_c^+(X)$ that converges pointwise to 1, we get $\int d\mu^+ + \int d\mu^- \le \|\mu\|$. Conversely, if $f \in C_c^{\mathbb{R}}(X)$, then

$$|\mu(f)| = |\mu^+(f)-\mu^-(f)| \le \mu^+(|f|)+\mu^-(|f|) \le \Bigl(\int d\mu^+ + \int d\mu^-\Bigr)\|f\|.$$

(Here we used Lemma 3.2.) □

Remark. The decomposition $\mu = \mu^+ - \mu^-$ with $\mu^+, \mu^- \in \mathfrak{M}^+(X)$ is unique if we insist that $\|\mu\| = \int d\mu^+ + \int d\mu^-$. Indeed, if $\mu = \mu_1 - \mu_2$ is a second decomposition of this form, the Radon measure $\nu = \mu_1 - \mu^+ = \mu_2 - \mu^-$ is positive (see Remark 1 above) and $\int d\mu_1 + \int d\mu_2 = \int d\mu^+ + \int d\mu^- + 2\int d\nu$.

Finally, we define **complex Radon measures** and **bounded complex Radon measures** by substituting $\mathbb{C}$ for $\mathbb{R}$ in the preceding definitions. We denote by $\mathfrak{M}^{\mathbb{C}}(X)$ and $\mathfrak{M}_f^{\mathbb{C}}(X)$ the corresponding spaces. In particular, $\mathfrak{M}_f^{\mathbb{C}}(X)$ can be identified with the topological dual of $C_0^{\mathbb{C}}(X)$. Since $C_c^{\mathbb{C}}(X) = C_c^{\mathbb{R}}(X) + iC_c^{\mathbb{R}}(X)$, a real Radon measure μ gives rise in a unique way to a complex Radon measure, which we also denote by μ, as follows:

$$\mu(f) = \mu(\operatorname{Re} f) + i\mu(\operatorname{Im} f) \quad \text{for all } f \in C_c^{\mathbb{C}}(X).$$

Then $\mathfrak{M}^{\mathbb{R}}(X) \subset \mathfrak{M}^{\mathbb{C}}(X)$ and $\mathfrak{M}_f^{\mathbb{R}}(X) \subset \mathfrak{M}_f^{\mathbb{C}}(X)$. Actually,

$$\mathfrak{M}^{\mathbb{C}}(X) = \mathfrak{M}^{\mathbb{R}}(X) + i\mathfrak{M}^{\mathbb{R}}(X), \qquad \mathfrak{M}_f^{\mathbb{C}}(X) = \mathfrak{M}_f^{\mathbb{R}}(X) + i\mathfrak{M}_f^{\mathbb{R}}(X).$$

For, if $\mu \in \mathfrak{M}^{\mathbb{C}}(X)$, we define $\operatorname{Re}\mu$ by setting

$$\operatorname{Re}\mu(f) = \operatorname{Re}(\mu(f)) \quad \text{for all } f \in C_c^{\mathbb{R}}(X),$$

and likewise for $\operatorname{Im}\mu$. Then $\mu = \operatorname{Re}\mu + i\operatorname{Im}\mu$. Such a decomposition is unique.

For $\mu \in \mathfrak{M}_f^{\mathbb{K}}(X)$, we define the integral of a bounded Borel function f on X as follows:

- If $\mathbb{K} = \mathbb{R}$, put $\int f\,d\mu = \int f\,d\mu^+ - \int f\,d\mu^-$.
- If $\mathbb{K} = \mathbb{C}$, put $\int f\,d\mu = \int f\,d(\operatorname{Re}\mu) + i\int f\,d(\operatorname{Im}\mu)$; that is,

$$\int f\,d\mu = \int f\,d(\operatorname{Re}\mu)^+ - \int f\,d(\operatorname{Re}\mu)^- + i\int f\,d(\operatorname{Im}\mu)^+ - i\int f\,d(\operatorname{Im}\mu)^-.$$

We define the Borel measure of a subset A of X as $\mu(A) = \int 1_A\,d\mu$.

Exercises

Throughout this set of exercises, X is a locally compact separable metric space.

1. Prove that $\mathfrak{M}^{\mathbb{R}}(X)$, with the order relation defined by

$$\mu \leq \nu \quad \Leftrightarrow \quad \nu - \mu \in \mathfrak{M}^+(X),$$

is a lattice.

Hint. Show first that if $\mu \in \mathfrak{M}^{\mathbb{R}}(X)$ and we write $|\mu| = \mu^+ + \mu^-$ in the notation of Theorem 4.1, then $|\mu| = \sup(\mu, -\mu)$.

2. a. Fix $\mu \in \mathfrak{M}^{\mathbb{K}}(X)$. Show that there exists a largest open set O such that any $f \in C_c^{\mathbb{K}}(X)$ whose support is contained in O satisfies $\mu(f) = 0$. (Use partitions of unity.) The complement of this largest open set is called the *support* of μ and is denoted $\operatorname{Supp} \mu$. By Exercise 1 on page 77, this definition coincides with the one introduced earlier for positive measures.

b. Prove that if $\mu \in \mathfrak{M}^{\mathbb{R}}(X)$ then $\operatorname{Supp} \mu = \operatorname{Supp} \mu^+ \cup \operatorname{Supp} \mu^-$, in the notation of Theorem 4.1, and that if $\mu \in \mathfrak{M}^{\mathbb{C}}(X)$ then

$$\operatorname{Supp} \mu = \operatorname{Supp}(\operatorname{Re} \mu) \cup \operatorname{Supp}(\operatorname{Im} \mu).$$

3. a. Fix $\mu \in \mathfrak{M}^+(X)$, and extend μ to a linear form on $C_c^{\mathbb{C}}(X)$. Prove that $|\mu(f)| \leq \mu(|f|)$ for all $f \in C_c^{\mathbb{C}}(X)$.

Hint. Let α be a complex number of absolute value 1 such that $\alpha\mu(f) = |\mu(f)|$. Prove that $|\mu(f)| = \mu(\operatorname{Re}(\alpha f))$.

b. Let μ be a bounded real Radon measure. By reasoning as in the previous question, show that μ has the same norm in the topological duals of $C_0^{\mathbb{R}}(X)$ and of $C_0^{\mathbb{C}}(X)$.

c. Fix $\mu \in \mathfrak{M}_f^{\mathbb{K}}(X)$. Prove that $|\mu(A)| \leq \|\mu\|$ for any Borel set A of X.

Hint. In the case $\mathbb{K} = \mathbb{C}$, put $\nu = (\operatorname{Re} \mu)^+ + (\operatorname{Re} \mu)^- + (\operatorname{Im} \mu)^+ + (\operatorname{Im} \mu)^-$ and consider a sequence $(f_n)_{n \in \mathbb{N}}$ of $C_c(X)$ that converges to 1_A in $L^1(\nu)$ and such that $0 \leq f_n \leq 1$ for all $n \in \mathbb{N}$. Prove that $\mu(A) = \lim_{n \to +\infty} \mu(f_n)$ and wrap up.

4. Let μ be a positive Radon measure on X and take $\psi \in L^1(\mu)$. Prove that the relation

$$\nu(f) = \int f\psi \, d\mu$$

defines a bounded Radon measure on X and that $\|\nu\| = \int |\psi| \, d\mu$.

Hint. Let s be a function defined on X such that $s\psi = |\psi|$ and $s = 0$ on $\{\psi = 0\}$. Prove that, for all $\varepsilon > 0$, there exists a $g \in C_c(X)$ such that $\int |\psi| \, |g - s| \, d\mu \leq \varepsilon$ and that, in addition, g can be chosen so that $\|g\| \leq 1$. Now estimate $\int |\psi| \, d\mu - \nu(g)$.

5. We say that a sequence (f_n) in $C_c(X)$ *converges in* $C_c(X)$ to $f \in C_c(X)$ if it converges uniformly to f and there exists a compact subset K of X containing the support of every f_n. Let μ be a linear form on $C_c^{\mathbb{K}}(X)$. Prove that $\mu \in \mathfrak{M}^{\mathbb{K}}(X)$ if and only if the image under μ of every sequence of functions in $C_c(X)$ that converges to 0 in $C_c(X)$ is a sequence that converges to 0 in $\mathbb{K}$.
6. We say a sequence (μ_n) in $\mathfrak{M}^{\mathbb{K}}(X)$ *converges vaguely* to $\mu \in \mathfrak{M}^{\mathbb{K}}(X)$ if

$$\lim_{n\to+\infty} \mu_n(f) = \mu(f) \quad \text{for all } f \in C_c^{\mathbb{K}}(X).$$

   a. Let (μ_n) be a sequence in $\mathfrak{M}^{\mathbb{K}}(X)$ such that, for all $f \in C_c^{\mathbb{K}}(X)$, the sequence $\big(\mu_n(f)\big)$ converges. Prove that the sequence (μ_n) converges vaguely.
   Hint. Let (K_p) be a sequence of compact sets that exhausts X. Apply to each space $C_{K_p}^{\mathbb{K}}(X)$ the result of Exercise 6f on page 23.
   b. Let (μ_n) be a sequence in $\mathfrak{M}(X)$ such that, for all $f \in C_c(X)$,

$$\sup_{n\in\mathbb{N}} \left| \int f\, d\mu_n \right| < +\infty.$$

   Prove that the sequence (μ_n) has a vaguely convergent subsequence.
   Hint. Work as in Exercise 10e on page 81, using the Banach–Steinhaus Theorem (Exercise 6d on page 22).
7. We say that a sequence (μ_n) in $\mathfrak{M}_f(X)$ *converges weakly* to $\mu \in \mathfrak{M}_f(X)$ if

$$\lim_{n\to+\infty} \int f\, d\mu_n = \int f\, d\mu \quad \text{for all } f \in C_0(X).$$

   a. Let (μ_n) be a sequence in $\mathfrak{M}_f(X)$. Prove that a sufficient condition for it to converge weakly is that, for all $f \in C_0(X)$, the sequence $(\int f\, d\mu_n)_{n\in\mathbb{N}}$ should converge.
   Hint. Use Exercise 6f on page 23.
   b. Prove that any bounded sequence (μ_n) in $\mathfrak{M}_f(X)$ $\big($one for which $\sup_{n\in\mathbb{N}} \|\mu_n\| < +\infty\big)$ has a weakly convergent subsequence.
   Hint. The space $C_0(X)$ is separable by Exercise 7h on page 56, so it is enough to use the Banach–Alaoglu Theorem, page 19.
   c. Prove that a sequence (μ_n) in $\mathfrak{M}_f(X)$ converges weakly if and only if it converges vaguely (see Exercise 6) and is bounded.
   d. Find a sequence (μ_n) in $\mathfrak{M}_f^+(X)$ that converges weakly but not narrowly (see Exercise 12 on page 81).
8. Let H be a relatively compact subset of $\mathfrak{M}_f(X)$ (we identify this space with the topological dual of $C_0(X)$). Prove that there exists a positive Radon measure λ of finite mass on X such that any $A \in \mathscr{B}(X)$ having λ-measure zero also has μ-measure zero for all $\mu \in H$. (The measures $\mu \in H$ are then said to be *absolutely continuous* with respect to λ.)

Hint. Define $\lambda = \sum_{n\in\mathbb{N}^*}\sum_{i=1}^{i_n} 2^{-n-i}v(\mu_i^n)$, where the μ_i^n are elements of H chosen so that, for every $n \in \mathbb{N}^*$, the balls $B(\mu_1^n, 1/n)$, $\dots$, $B(\mu_{i_n}^n, 1/n)$ cover H, and where we write, for $\mu \in \mathfrak{M}_f(X)$, $v(\mu) = \mu^+ + \mu^-$ if $\mathbb{K} = \mathbb{R}$ and $v(\mu) = (\operatorname{Re}\mu)^+ + (\operatorname{Re}\mu)^- + (\operatorname{Im}\mu)^+ + (\operatorname{Im}\mu)^-$ if $\mathbb{K} = \mathbb{C}$. You might use Exercise 3c.

9. Prove that the topological dual of $C_0(X)$ is separable if and only if X is countable.

Hint. Prove that, if $X = \{x_n\}_{n\in\mathbb{N}}$, the family $\{\delta_{x_n}\}_{n\in\mathbb{N}}$ is fundamental in $(C_0(X))'$. For the "only if" part, you might show that $\|\delta_a - \delta_b\| = 2$ for any two distinct points $a, b \in X$, and then use Proposition 2.4 on page 9.

10. Give $C(X)$ the metric d of uniform convergence on compact sets, defined in Exercise 12 on page 57. Prove that the topological dual of $(C(X), d)$ can be identified with the space $\mathfrak{M}_c(X)$ of Radon measures with compact support (the support of a Radon measure was defined in Exercise 2 above).

Hint. Argue as in Exercise 9 on page 80.

11. Let L be a continuous linear form on $C_0(X)$ and let (f_n) be a bounded sequence in $C_0(X)$. Prove that if (f_n) converges pointwise to $f \in C_0(X)$ then $\lim_{n\to+\infty} L(f_n) = L(f)$.

Hint. Use the Dominated Convergence Theorem.

12. Two Borel measures μ_1 and μ_2 of finite mass on X are called *mutually singular* if there exists a Borel set A in X such that $\mu_1(A) = \mu_1(X)$ and $\mu_2(A) = 0$. Let μ be a bounded real Radon measure on X and let μ_1 and μ_2 be positive Radon measures of finite mass on X such that $\mu = \mu_1 - \mu_2$.

a. Assume that μ_1 and μ_2 are mutually singular. Prove that $\|\mu\| = \mu_1(X) + \mu_2(X)$.

Hint. Let $\varepsilon > 0$. Write $\varphi = 1_A - 1_{X\setminus A}$. Prove that there exists a function $f \in C_c^{\mathbb{R}}(X)$ such that $\|f - \varphi\|_{L^1(\mu_1+\mu_2)} \le \varepsilon$. Let $\tilde{f}$ be the function defined on X by

$$\tilde{f}(x) = \begin{cases} f(x) & \text{if } |f(x)| \le 1, \\ \operatorname{sign} f(x) & \text{otherwise.} \end{cases}$$

Check that $\tilde{f} \in C_c^{\mathbb{R}}(X)$, then show that $\mu(\tilde{f}) \ge \mu_1(X) + \mu_2(X) - \varepsilon$. Deduce that $\|\mu\| \ge \mu_1(X) + \mu_2(X)$. The opposite inequality is easy.

b. Prove the converse.

Hint. Suppose $\|\mu\| = \mu_1(X) + \mu_2(X)$. Let (f_n) be a sequence of elements of $C_c^{\mathbb{R}}(X)$ such that $\mu(f_n) \to \|\mu\|$ and $|f_n| \le 1$. Prove that

$$\int f_n^+ \, d\mu_1 \to \mu_1(X), \quad \int f_n^+ \, d\mu_2 \to 0.$$

Deduce the existence of a subsequence $(f_{n_k}^+)$ that converges μ_1-almost everywhere to 1 and μ_2-almost everywhere to 0. Conclude.

c. Let μ be a bounded real Radon measure on X. Show that there exists a unique pair (μ_1, μ_2) of mutually singular positive Radon measures of finite mass such that $\mu = \mu_1 - \mu_2$; show that $\mu_1 = \mu^+$ and $\mu_2 = \mu^-$.

13. *Functions of bounded variation.* Let f be a real-valued function on an interval $[a, b]$ of $\mathbb{R}$. If $\Delta = \{x_j\}_{0 \le j \le n}$ is a subdivision of $[a, b]$, we write

$$V(f, \Delta) = \sum_{j=0}^{n-1} \bigl|f(x_{j+1}) - f(x_j)\bigr|;$$

we also write $V(f, a, b) = \sup_\Delta V(f, \Delta)$. We say that f is of *bounded variation on* $[a, b]$ if $V(f, a, b)$ is finite. We say that a function $f : \mathbb{R} \to \mathbb{R}$ is of *bounded variation on* $\mathbb{R}$ if the expression

$$V(f) = \sup_{\substack{(a,b)\in\mathbb{R}^2 \\ a<b}} V(f, a, b)$$

is finite.

a. Let f be a monotone function on $[a, b]$. Prove that f is of bounded variation on $[a, b]$ and compute $V(f, a, b)$.

b. Prove that the set $BV(a, b)$ of functions of bounded variation on $[a, b]$ is a vector space and that $f \mapsto V(f, a, b)$ is a seminorm on $BV(a, b)$. Prove that for $f \in BV(a, b)$ we have $V(f, a, b) = 0$ if and only if f is constant on $[a, b]$.

c. Let $BV_0(a, b)$ be the space of functions f of bounded variation on $[a, b]$ such that $f(a) = 0$. Prove that $f \mapsto V(f, a, b)$ is a norm on $BV_0(a, b)$ with respect to which this space is complete.

d. Take $f \in BV(a, b)$. Prove that for $a \le c < d < e \le b$ we have
   i. $V(f, c, d) + V(f, d, e) = V(f, c, e)$,
   ii. $\bigl|f(c) - f(d)\bigr| \le V(f, c, d)$.

   Deduce that the functions $x \mapsto V(f, a, x)$ and $x \mapsto V(f, a, x) - f(x)$ are increasing functions from $[a, b]$ to $\mathbb{R}^+$.

e. Take $f \in BV(a, b)$. Prove that if f is right-continuous at a point $c \in [a, b)$, so is the function $x \mapsto V(f, c, x)$. Likewise, if f is left-continuous at $c \in (a, b]$, so is $x \mapsto V(f, c, x)$.

   Hint. If $x \mapsto V(f, c, x)$ is not right-continuous at c, there exists a real number $\eta > 0$ such that $V(f, c, x) > \eta$ for all $x \in (c, b]$. Now construct by induction a sequence (x_n) such that, for all $n \in \mathbb{N}$, $c < x_{n+1} < x_n < b$ and $V(f, x_{n+1}, x_n) > \eta$; then deduce that $V(f, c, b) = +\infty$, which is absurd.

f. Prove that a function f from $[a, b]$ to $\mathbb{R}$ is of bounded variation if and only if there exist two increasing functions g and h from $[a, b]$ to $\mathbb{R}^+$ such that $f = g - h$. Prove that if f is right-continuous at a point $c \in [a, b)$, then g and h can be chosen to satisfy the same condition. An analogous statement holds for left-continuous functions.

14. We resume the notation and terminology of Exercise 13. Let f and g be real- or complex-valued functions defined on an interval $[a,b]$ of $\mathbb{R}$. If $\Delta = \{x_j\}_{0\le j\le n}$ is a subdivision of $[a,b]$ and if $c = (c_0,\ldots,c_{n-1})$ is such that $c_i \in [x_i, x_{i+1}]$ for all $i \le n-1$, we write

$$S_{\Delta,c}(f,g) = \sum_{i=0}^{n-1} f(c_i)\big(g(x_{i+1}) - g(x_i)\big).$$

If, as $\delta(\Delta)$ approaches 0, the sequence $\big(S_{\Delta,c}(f,g)\big)$ has a limit uniform with respect to c, this limit is denoted by $\int_a^b f\,dg$.

a. Prove that, if f is continuous and g is increasing, $\int_a^b f\,dg$ is well-defined and coincides with the definition given on page 72. Prove that, if $g \in BV(a,b)$, the linear form L on $C([a,b])$ defined by $L(f) = \int_a^b f\,dg$ is continuous and has norm at most $V(g,a,b)$.

b. *Integration by parts.* Let f and g be real- or complex-valued functions from $[a,b]$ to $\mathbb{R}$ or $\mathbb{C}$. Prove that $\int_a^b f\,dg$ is defined if and only if $\int_a^b g\,df$ is, and that in this case

$$\int_a^b f\,dg + \int_a^b g\,df = f(b)g(b) - f(a)g(a)$$

(use summation by parts on the finite sums $S_{\Delta,c}(f,g)$).

c. *Second Mean Value Theorem.* Let f be an increasing function from $[a,b]$ to $\mathbb{R}^+$ and let g be a Lebesgue-integrable function from $[a,b]$ to $\mathbb{R}$. Show that there exists $\xi \in [a,b]$ such that

$$\int_a^b f(t)g(t)\,dt = f(b)\int_\xi^b g(t)\,dt.$$

This is called the Second Mean Value Theorem.
Hint. One can assume that $f(a) = 0$. Set $G(x) = \int_x^b g(t)\,dt$. Prove that

$$\int_a^b f(t)g(t)\,dt = -\int_a^b f\,dG = \int_a^b G\,df.$$

d. Let f be a function of bounded variation on $\mathbb{R}$. Suppose that $f(x)$ tends to 0 both as $x \to +\infty$ and as $x \to -\infty$. Show that there exists a constant $C > 0$ such that, for every nonzero real number t,

$$\left|\int_{-\infty}^{+\infty} f(x)e^{-itx}dx\right| \le \frac{C}{|t|}.$$

15. We continue with the notation and terminology of Exercises 13 and 14. A function f of bounded variation on $\mathbb{R}$ is called *normalized* if it is right-continuous and $\lim_{x\to-\infty} f(x) = 0$. We denote by $NBV(\mathbb{R})$ the vector space consisting of normalized functions of bounded variation from $\mathbb{R}$ to $\mathbb{R}$.

a. Prove that every element of $NBV(\mathbb{R})$ can be written as the difference of two increasing and right-continuous functions that approach 0 at $-\infty$.

b. Prove that the map $f \mapsto V(f)$ is a norm on $NBV(\mathbb{R})$.

c. If $f \in NBV(\mathbb{R})$, we define a linear form μ_f on $C_0(\mathbb{R})$ by

$$\mu_f(\varphi) = \lim_{a \to +\infty} \int_{-a}^{a} \varphi \, df \quad \text{for all } \varphi \in C_0(X).$$

Check that μ_f is well-defined, that $\mu_f \in \mathfrak{M}_f(\mathbb{R})$, and that $\|\mu_f\| \leq V(f)$, where $\|\mu_f\|$ is the norm of μ_f in $C_0(\mathbb{R})'$.

- **i.** Suppose $f, g \in NBV(\mathbb{R})$ satisfy $\mu_f = \mu_g$. Prove that $f = g$. *Hint.* Using part a above, prove that $f(a) = \mu_f\big((-\infty, a]\big)$ for all $a \in \mathbb{R}$.
- **ii.** Let $f \in NBV(\mathbb{R})$ be increasing. Prove that $V(f) = \|\mu_f\|$.
- **iii.** Take $f \in NBV(\mathbb{R})$. Prove that there exist bounded, increasing, right-continuous functions f_+ and f_- such that $f = f_+ - f_-$ and $\|\mu_f\| = V(f_+) + V(f_-)$. Deduce that $V(f) \leq \|\mu_f\|$.
- **iv.** Prove that the linear map $L : f \mapsto \mu_f$ is a bijective isometry from $NBV(\mathbb{R})$ onto the topological dual of $C_0(\mathbb{R})$.

d. Prove that $NBV(\mathbb{R})$ is a nonseparable Banach space. (That it is nonseparable is elementary: Consider the uncountable family consisting of functions $Y_a = 1_{[a,+\infty)}$, with $a \in \mathbb{R}$.)

3
Hilbert Spaces

This chapter is devoted to a class of normed spaces that is particularly important in both theory and applications.

1 Definitions, Elementary Properties, Examples

In all of this chapter we consider a vector space E over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. A **scalar product** on E is a map $(\cdot \mid \cdot)$ from $E \times E$ to $\mathbb{K}$ satisfying these conditions:

a. For all $y \in E$, the map $(\cdot \mid y) : E \to \mathbb{K}$ defined by $x \mapsto (x \mid y)$ is linear.
b. – If $\mathbb{K} = \mathbb{R}$: for all $x, y \in E$, $(y \mid x) = (x \mid y)$ (**symmetry**).
– If $\mathbb{K} = \mathbb{C}$: for all $x, y \in E$, $(y \mid x) = \overline{(x \mid y)}$ (**skew-symmetry**).
c. For all $x \in E$, $(x \mid x) \in \mathbb{R}^+$.
d. For all $x \in E$, $(x \mid x) = 0$ if and only if $x = 0$.

A map that satisfies the first three conditions but not necessarily the fourth is called a **scalar semiproduct**.

A space E endowed with a scalar product is called a **pre-Hilbert space** or **scalar product space**, further qualified as **real** if $\mathbb{K} = \mathbb{R}$ or **complex** if $\mathbb{K} = \mathbb{C}$. We leave out this qualification if no confusion is possible or if $\mathbb{K}$ need not be specified.

Remark. Suppose $(\cdot \mid \cdot)$ is a map from $E \times E$ to $\mathbb{K}$ that satisfies the first two conditions in the definition of a scalar product. Fix $x \in E$; if $\mathbb{K} = \mathbb{R}$, the map $(x\,|\,\cdot\,) : y \mapsto (x\,|\,y)$ is linear from E to $\mathbb{R}$. If $\mathbb{K} = \mathbb{C}$, the same map is **skew-linear**; that is, for all $x, y, z \in E$ and all $\lambda, \mu \in \mathbb{C}$,

$$(x\,|\,\lambda y + \mu z) = \bar{\lambda}(x\,|\,y) + \bar{\mu}(x\,|\,z).$$

Also, as a consequence of the first two conditions in the definition of a scalar product, we have, for $x, y \in E$:

- If $\mathbb{K} = \mathbb{R}$: $(x + y\,|\,x + y) = (x\,|\,x) + (y\,|\,y) + 2(x\,|\,y)$.
- If $\mathbb{K} = \mathbb{C}$: $(x + y\,|\,x + y) = (x\,|\,x) + (y\,|\,y) + 2\,\mathrm{Re}(x\,|\,y)$.

Examples

1. Let $E = \mathbb{R}^d$. If $a_1, \dots, a_d$ are nonnegative real numbers, the equation $(x\,|\,y) = \sum_{j=1}^d a_j x_j y_j$ defines on E a scalar semiproduct, which is a scalar product if and only if all the a_j are positive. If $a_j = 1$ for all j, this is called the **euclidean scalar product**, and E together with this scalar product is called **d-dimensional canonical euclidean space**. Similarly, if $E = \mathbb{C}^d$ and $a_1, \dots, a_d$ are nonnegative reals, a scalar semiproduct on E is defined by $(x\,|\,y) = \sum_{j=1}^d a_j x_j \bar{y}_j$, and this is a scalar product if all the a_j are positive. If $a_j = 1$ for all j, this is called the **hermitian scalar product**, and E together with this scalar product is called **d-dimensional canonical hermitian space**.
2. Let X be a locally compact separable metric space, μ a positive Radon measure on X, and $E = C_c^{\mathbb{K}}(X)$. The equations

$$(f\,|\,g) = \int f(x)g(x)\,d\mu(x) \quad \text{if } \mathbb{K} = \mathbb{R},$$
$$(f\,|\,g) = \int f(x)\overline{g(x)}\,d\mu(x) \quad \text{if } \mathbb{K} = \mathbb{C}$$

   define on E a scalar semiproduct, which is a scalar product if and only if $\operatorname{Supp}\mu = X$.
3. Fix $a > 0$, and let $E = C_a^{\mathbb{K}}$ be the set of continuous functions from $\mathbb{R}$ to $\mathbb{K}$ periodic of period a. The equations

$$(f\,|\,g) = \frac{1}{a}\int_0^a f(x)g(x)\,dx \quad \text{if } \mathbb{K} = \mathbb{R},$$
$$(f\,|\,g) = \frac{1}{a}\int_0^a f(x)\overline{g(x)}\,dx \quad \text{if } \mathbb{K} = \mathbb{C}$$

   define a scalar product on E.
4. Let m be a measure on a measure space $(\Omega, \mathscr{F})$ and let $\mathscr{E} = \mathscr{L}^2_{\mathbb{K}}(m)$ be the space of $\mathscr{F}$-measurable functions f from Ω to $\mathbb{K}$ that are square-integrable, that is, satisfy $\int |f|^2\,dm < +\infty$. (That this is a vector space

follows from the inequality $|f+g|^2 \le 2(|f|^2+|g|^2)$.) We give $\mathscr{E}$ a scalar semiproduct by setting

$$(f \mid g) = \int f g \, dm \quad \text{if } \mathbb{K} = \mathbb{R},$$

$$(f \mid g) = \int f \bar{g} \, dm \quad \text{if } \mathbb{K} = \mathbb{C}.$$

This scalar semiproduct induces a scalar product on the space $E = L^2_{\mathbb{K}}(m)$ defined as the quotient of $\mathscr{E}$ by the relation of equality m-almost everywhere.

5. An important particular case of the preceding situation is the following. Let I be any set and let $\mathscr{F} = \mathscr{P}(I)$ be the discrete σ-algebra on I — the one containing all subsets of I. On the measure space $(I, \mathscr{F})$ we take the **count measure** m, defined by $m(A) = \operatorname{Card}(A) \le +\infty$. (If I is countable, one can regard it as a locally compact separable metric space by giving it the discrete metric, defined by $d(x,y) = 1$ if $x \ne y$; then m is a positive Radon measure on I.) We generally use subscript notation for functions on I: thus $x = (x_i)_{i \in I}$. If x takes nonnegative values, we use the notation $\sum_{i \in I} x_i$ to denote $\int x \, dm \le +\infty$. One easily checks that
$$\sum_{i \in I} x_i = \sup_{J \in \mathscr{P}_f(I)} \sum_{i \in J} x_i \le +\infty,$$
where $\mathscr{P}_f(I)$ is the set of finite subsets of I. The space $\mathscr{L}^1_{\mathbb{K}}(m)$ in this case is denoted by $\ell^1_{\mathbb{K}}(I)$ and, for every $x \in \ell^1_{\mathbb{K}}(I)$, we write $\sum_{i \in I} x_i = \int x \, dm$. Similarly, we write $\ell^2_{\mathbb{K}}(I) = \mathscr{L}^2_{\mathbb{K}}(m)$. (See also Exercises 7 on page 11 and 8 on page 12.)
Since the only set of m-measure zero is the empty set, we have $L^2_{\mathbb{K}}(m) = \ell^2_{\mathbb{K}}(I)$; thus this space has a scalar product structure defined by

$$(x \mid y) = \sum_{i \in I} x_i y_i \quad \text{if } \mathbb{K} = \mathbb{R},$$

$$(x \mid y) = \sum_{i \in I} x_i \bar{y}_i \quad \text{if } \mathbb{K} = \mathbb{C}.$$

We omit I from the notation when $I = \mathbb{N}$.

Here is a fundamental property of scalar semiproducts.

Proposition 1.1 (Schwarz inequality) *Let E be a vector space with a scalar semiproduct $(\cdot \mid \cdot)$. For every $x, y \in E$,*

$$\bigl|(x \mid y)\bigr|^2 \le (x \mid x)(y \mid y).$$

Proof. One can assume $\mathbb{K} = \mathbb{C}$. If $x, y \in E$,

$$(x + ty \mid x + ty) = (x \mid x) + 2t \operatorname{Re}(x \mid y) + t^2 (y \mid y) \ge 0 \quad \text{for all } t \in \mathbb{R}.$$

Consider the expression on the left-hand side of this inequality as a polynomial in t, taking only nonnegative values. If $(y \mid y) = 0$, the polynomial is at most of first degree and must be constant, so $0 = (\operatorname{Re}(x \mid y))^2 \le (x \mid x)(y \mid y) = 0$. If $(y \mid y) \ne 0$, the polynomial is of second degree and must have negative or zero discriminant; again $(\operatorname{Re}(x \mid y))^2 \le (x \mid x)(y \mid y)$.

Now let u be a complex number of absolute value 1 such that

$$\bigl|(x \mid y)\bigr| = u(x \mid y) = (ux \mid y) = \operatorname{Re}(ux \mid y).$$

We see that $\bigl|(x \mid y)\bigr|^2 \le (ux \mid ux)(y \mid y) = (x \mid x)(y \mid y)$, since $u\bar{u} = 1$. □

Corollary 1.2 *Let E be a vector space with a scalar product $(\cdot \mid \cdot)$. The expression $\|x\| = (x \mid x)^{1/2}$ defines a norm on E.*

Proof. It is enough to check the triangle inequality. We have

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(x \mid y)\\ &\le \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = \bigl(\|x\| + \|y\|\bigr)^2.\end{aligned}$$ □

From now on, unless we specify otherwise, we will denote the scalar product on any space E by $(\cdot \mid \cdot)$, and the associated norm by $\|\cdot\|$. For example, if $E = L^2(m)$, as in Example 4 above,

$$\|f\| = \left(\int |f|^2\,dm\right)^{1/2}.$$

If $E = \ell^2(I)$,

$$\|x\| = \left(\sum_{i\in I} |x_i|^2\right)^{1/2}.$$

Note that, in any scalar product space, the scalar product can be recovered from the norm: If $\mathbb{K} = \mathbb{C}$, we have

$$\begin{aligned}\operatorname{Re}(x \mid y) &= \tfrac{1}{2}\bigl((\|x+y\|)^2 - \|x\|^2 - \|y\|^2\bigr),\\ \operatorname{Im}(x \mid y) &= \tfrac{1}{2}\bigl((\|x+iy\|)^2 - \|x\|^2 - \|y\|^2\bigr),\end{aligned}$$

and in the real case the first of these equalities holds.

Corollary 1.3 *Let E be a scalar product space. For every $y \in E$, the linear form $\varphi_y = (\cdot \mid y)$ is continuous and its norm in the topological dual E' of E equals $\|y\|$.*

Proof. By the Schwarz inequality, $\bigl|\varphi_y(x)\bigr| \le \|x\|\,\|y\|$ for all $x \in E$, so $\varphi_y \in E'$ and $\|\varphi_y\| \le \|y\|$. At the same time, $\varphi_y(y) = \|y\|^2$, so $\|\varphi_y\| = \|y\|$. □

Thus the map $y \mapsto \varphi_y$ is an isometry from E to E', linear if $\mathbb{K} = \mathbb{R}$ and skew-linear if $\mathbb{K} = \mathbb{C}$. We will see in Theorem 3.1 below that this isometry is bijective if the space E is complete.

Proposition 1.4 (Equality in the Schwarz inequality) *Two vectors x and y in a scalar product space satisfy $\bigl|(x\,|\,y)\bigr| = \|x\|\,\|y\|$ if and only if they are linearly dependent.*

Proof. The "if" part is obvious. To show the converse, suppose for example that $\mathbb{K} = \mathbb{C}$ and that $\bigl|(x\,|\,y)\bigr| = \|x\|\,\|y\|$. Let ε be a complex number of absolute value 1 such that $\operatorname{Re}\bigl(\varepsilon(x\,|\,y)\bigr) = \bigl|(x\,|\,y)\bigr|$. Then $\bigl\|\,\|x\|y - \varepsilon\|y\|x\bigr\|^2 = 0$ (expand the square), so $\|x\|y - \varepsilon\|y\|x = 0$. □

An immediate, but useful, consequence of the definition of the norm in a scalar product space is the **parallelogram identity**:

Proposition 1.5 *If x and y are elements of a scalar product space,*

$$\left\|\frac{x+y}{2}\right\|^2 + \left\|\frac{x-y}{2}\right\|^2 = \tfrac{1}{2}\bigl(\|x\|^2 + \|y\|^2\bigr).$$

Orthogonality

Two elements x and y of a scalar product space E are **orthogonal** if $(x\,|\,y) = 0$; in this case we write $x \perp y$. The orthogonality relation $\perp$ thus defined is of course symmetric. The **orthogonal space** to a subset A of E is, by definition, the set $A^\perp$ consisting of points orthogonal to all the elements of A. Thus, in the notation of Corollary 1.3,

$$A^\perp = \bigcap_{y\in A} \ker(\varphi_y).$$

It follows that $A^\perp$ is a closed vector subspace of E. At the same time, x belongs to $A^\perp$ if and only if $A \subset \ker\varphi_x$; since $\ker\varphi_x$ is closed, this inclusion is equivalent to $\overline{[A]} \subset \ker\varphi_x$, where $[A]$ is the span of A (the vector space consisting of linear combinations of elements of A). Thus

$$A^\perp = \bigl(\overline{[A]}\bigr)^\perp.$$

Two subsets A and B of E are called **orthogonal** if $x \perp y$ for any $x \in A$ and $y \in B$. The following relation between orthogonal vectors, called the **Pythagorean Theorem**, is immediate:

Proposition 1.6 *If x and y are orthogonal vectors in a scalar product space,*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

This result extends by induction to a finite number of pairwise orthogonal vectors $x_1, \ldots, x_n$: $\bigl\|\sum_{j=1}^n x_j\bigr\|^2 = \sum_{j=1}^n \|x_j\|^2$.

A scalar product space that is complete with respect to the norm defined by its scalar product is called a **Hilbert space**. . Here are the fundamental examples:

1. Every finite-dimensional scalar product space is a Hilbert space.
2. If m is a measure on a measure space $(\Omega, \mathscr{F})$, the space $L^2(m)$ with the scalar product defined in Example 4 above is a Hilbert space.

In particular, the space $\ell^2(I)$ of Example 5 above is a Hilbert space, for any set I. (This particular case is in fact the general case; see Theorem 4.4 below and Exercise 11 on page 133).

Exercises

1. Let E be a normed vector space over $\mathbb{C}$. Prove that the norm $\|\cdot\|$ comes form a scalar product if and only if it satisfies the parallelogram identity:

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big) \quad \text{for all } (x,y) \in E^2.$$

Prove that in this case the scalar product that defines $\|\cdot\|$ is

$$(x \,|\, y) = \tfrac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\big). \qquad (*)$$

Hint. To show sufficiency you might consider the map $(\cdot\,|\,\cdot)$ defined by $(*)$ and prove successively that it satisfies these properties:

a. $(x \,|\, x) = \|x\|^2$ for all $x \in E$.
b. $(x \,|\, y) = \overline{(y \,|\, x)}$ for all $(x,y) \in E^2$.
c. $(x+y \,|\, z) = 2(x \,|\, z/2) + 2(y \,|\, z/2)$ for all $(x,y,z) \in E^3$.
d. $(x+y \,|\, z) = (x \,|\, z) + (y \,|\, z)$ for all $(x,y,z) \in E^3$.
e. $(\lambda x \,|\, y) = \lambda(x \,|\, y)$ for all $(x,y) \in E^2$ and $\lambda \in \mathbb{C}$.

2. Assume that (x_n) and (y_n) are sequences contained in the unit ball of a scalar product space, and that $(x_n \,|\, y_n) \to 1$. Prove that $\|x_n - y_n\| \to 0$.

3. Let X be a compact metric space of infinite cardinality and let μ be a positive Radon measure on X, of support X. Give the space $E = C(X)$ the scalar product defined by $(f \,|\, g) = \int f\bar{g}\,d\mu$.

a. Let a be a cluster point of X. Prove that there exists a sequence of pairwise disjoint balls $\big(B(a_n, \varepsilon_n)\big)_{n\in\mathbb{N}}$ such that $\lim_{n\to+\infty} a_n = a$.
b. Prove that, for every integer $n \in \mathbb{N}$, there exists a continuous function φ_n on X supported inside $B(a_n, \varepsilon_n)$ and satisfying $|\varphi_n| \le 1$ and $\varphi_n(a_n) = (-1)^n$.
c. Prove that the series $\sum \varphi_n$ converges pointwise, uniformly on compact sets of $X\setminus\{a\}$, and in $L^2(\mu)$ to a continuous function on $X\setminus\{a\}$ that has no limit at the point a.
d. Deduce that E is not a Hilbert space.

4. Let Ω be an open subset of $\mathbb{C}$, considered with the euclidean metric. We denote by $H(\Omega)$ the space of holomorphic functions on Ω and by $H^2(\Omega)$ the subspace of $H(\Omega)$ consisting of holomorphic functions f on Ω such that $\int\int_\Omega |f(x+iy)|^2\,dx\,dy < +\infty$. We recall that $H(\Omega)$ is closed in $C(\Omega)$ with the metric of uniform convergence on compact sets of Ω. We give the space $H^2(\Omega)$ the scalar product defined by

$$(f \mid g) = \iint_\Omega f(x+iy)\,\overline{g(x+iy)}\,dx\,dy.$$

a. Take $f \in H(\Omega)$. Prove that, if $\bar{B}(z_0, r) \subset \Omega$,

$$f(z_0) = \frac{1}{\pi r^2} \iint_{\bar{B}(z_0,r)} f(x+iy)\,dx\,dy.$$

Deduce that, if $f \in H^2(\Omega)$,

$$|f(z_0)| \le \frac{1}{\sqrt{\pi}\, r} \|f\|,$$

where $\|\cdot\|$ denotes the norm coming from the scalar product.

b. Prove that, if K is a compact contained in Ω,

$$\sup_{z \in K} |f(z)| \le \frac{1}{\sqrt{\pi}\, d(K, \mathbb{C}\setminus\Omega)} \|f\|$$

for every $f \in H^2(\Omega)$.

c. Prove that $H^2(\Omega)$ is a Hilbert space.

5. Let I be a set and $x = (x_i)_{i\in I}$ a family of points in $\mathbb{K}$.

a. Suppose $x \in \ell^1_{\mathbb{K}}(I)$ and set $\xi = \sum_{i\in I} x_i$. Prove the following property:

(P) For every $\varepsilon > 0$, there exists a finite subset K of I such that, for any finite subset J of I containing K, $\left|\xi - \sum_{i\in J} x_i\right| \le \varepsilon$.

b. Conversely, suppose there exists $\xi \in \mathbb{K}$ such that Property (P) is satisfied. Prove that $x \in \ell^1_{\mathbb{K}}(I)$ and that $\xi = \sum_{i\in I} x_i$.
Hint. Assume first that $\mathbb{K} = \mathbb{R}$. Setting $I_1 = \{i \in I : x_i \ge 0\}$ and $I_2 = I\setminus I_1$, show that under the assumption of Property (P) we have $\sum_{i\in I_1} x_i < +\infty$ and $\sum_{i\in I_2}(-x_i) < +\infty$.

c. Suppose I is countably infinite. Prove that $x \in \ell^1_{\mathbb{K}}(I)$ if and only if, for any bijection $\varphi : \mathbb{N} \to I$, the series $\sum_{n=0}^{+\infty} x_{\varphi(n)}$ converges. Prove that in this case $\sum_{n=0}^{+\infty} x_{\varphi(n)} = \sum_{i\in I} x_i$.
Hint. To show that the condition is sufficient, reduce to the case $\mathbb{K} = \mathbb{R}$. Then prove that if either series $\sum_{i\in I_1} x_i$ or $\sum_{i\in I_2}(-x_i)$ diverges (I_1 and I_2 being defined as above), there exists a bijection $\varphi : I \to \mathbb{N}$ such that the series $\sum_{n=0}^{+\infty} x_{\varphi(n)}$ does not converge.

6. *Hilbert cube.* Take $c = (c_n)_{n\in\mathbb{N}} \in \ell^2$ and let C be the set of elements x of ℓ^2 such that $|x_n| \le |c_n|$ for all $n \in \mathbb{N}$. Prove that C is compact.
Hint. Use Tychonoff's Theorem.

7. If $a = (a_n)_{n\in\mathbb{N}}$ is a sequence of positive real numbers, we denote by ℓ^2_a the vector space consisting of sequences of complex numbers $u = (u_n)_{n\in\mathbb{N}}$ such that the series $\sum a_n |u_n^2|$ converges.

a. Prove that the formula

$$(u, v) = \sum_{n \in \mathbb{N}} a_n u_n \bar{v}_n$$

defines a scalar product on ℓ^2_a.

b. Prove that the map

$$i_a : (u_n)_n \longmapsto \left(\sqrt{a_n}\, u_n\right)_n$$

is a linear isometry from ℓ^2_a onto ℓ^2. Deduce that ℓ^2_a is a Hilbert space.

c. Let a and b be sequences of positive real numbers. Prove that if the sequence (a_n/b_n) tends to 0, the closed unit ball in ℓ^2_b is a compact subset of ℓ^2_a.

Hint. Use Exercise 8 on page 17.

d. If s is a real number, we define on $\mathbb{Z}$ a measure μ_s by setting

$$\mu_s(\{n\}) = (1 + n^2)^{s/2} \quad \text{for all } n \in \mathbb{Z},$$

and we put $H^s = L^2(\mu_s)$. Prove that for $r < s$ we have $H^s \subset H^r$ and the closed unit ball in H^s is a compact subset of H^r.

8. *Hilbert completion.* Let $\mathscr{E}$ be a vector space with a scalar semiproduct $(\cdot \,|\, \cdot)$. Write $p(x) = (x \,|\, x)^{1/2}$. By the Schwarz inequality, the map p satisfies the triangle inequality: $p(x + y) \le p(x) + p(y)$ for all $x, y \in \mathscr{E}$. In other words, p is a seminorm.

Consider the vector space $\hat{\mathscr{E}}$ consisting of sequences (x_n) that are Cauchy with respect to p (that is, satisfy $\lim_{n,m\to+\infty} p(x_n - x_m) = 0$). Define a relation $\mathscr{R}$ on $\hat{\mathscr{E}}$ by setting

$$(x_n)\, \mathscr{R}\, (y_n) \quad \Longleftrightarrow \quad \lim_{n\to+\infty} p(x_n - y_n) = 0.$$

$\mathscr{R}$ is easily seen to be an equivalence relation compatible with the vector space structure of $\hat{\mathscr{E}}$. We denote by E the quotient vector space of $\hat{\mathscr{E}}$ by $\mathscr{R}$, and by Φ the canonical map from $\hat{\mathscr{E}}$ to E (which associates to each element of $\hat{\mathscr{E}}$ its equivalence class modulo $\mathscr{R}$).

a. Let x and y be elements of E. Prove that if $\Phi\big((x_n)\big) = x$ and $\Phi\big((y_n)\big) = y$, the sequence $\big((x_n \,|\, y_n)\big)_{n\in\mathbb{N}}$ converges and its limit depends only on x and y.

b. Prove that the equation $\big(\Phi((x_n)) \,|\, \Phi((y_n))\big) = \lim_{n\to+\infty}(x_n \,|\, y_n)$ defines a scalar product on E. We denote by $\|\cdot\|$ the associated norm.

c. If x is an element of $\mathscr{E}$, we denote by $\hat{x}$ the image under Φ of the constant sequence equal to x. Prove that the map from $\mathscr{E}$ to E defined by $x \mapsto \hat{x}$ is linear and that $\|\hat{x}\| = p(x)$ for all $x \in \mathscr{E}$.

d. Prove that the set $E_0 = \{\hat{x} : x \in \mathscr{E}\}$ is dense in E.

e. Prove that E is a Hilbert space. (Show first that every sequence in E_0 that is Cauchy in the norm of E converges in E.)
The space E is called the *Hilbert completion* of $\mathscr{E}$. Such a space is unique in a sense to be made precise in the next question.

f. Let $(E^{\sim}, (\cdot\,|\,\cdot)^{\sim})$ be a Hilbert space such that there exists a linear map $L : \mathscr{E} \to E^{\sim}$ whose image is dense in $E^{\sim}$ and such that $\|L(x)\|^{\sim} = p(x)$ for all $x \in \mathscr{E}$. Prove that there exists a surjective isometry H from E onto $E^{\sim}$ such that $H(\hat{x}) = L(x)$ for all $x \in \mathscr{E}$.

2 The Projection Theorem

One of the main tools that make Hilbert spaces interesting is the Projection Theorem. We assume that E is a Hilbert space and we denote by $(\cdot\,|\,\cdot)$ its scalar product, by $\|\cdot\|$ its norm, and by d the metric defined by the norm.

Theorem 2.1 *Let C be a nonempty, closed, convex subset of E. For every point x of E, there exists a unique point y of C such that*

$$\|x - y\| = d(x, C).$$

This point, called the ***projection of** x **onto** C* *and denoted by $P_C(x)$, is characterized by the following property:*

$$y \in C \qquad \textit{and} \qquad \operatorname{Re}(x - y\,|\,z - y) \le 0 \quad \textit{for all } z \in C. \tag{$*$}$$

Proof. Fix $x \in E$. We first show the existence of the projection of x onto C. By the definition of $\delta = d(x, C)$, there exists a sequence (y_n) in C such that

$$\|x - y_n\|^2 \le \delta^2 + \frac{1}{n} \quad \text{for all } n \ge 1.$$

Applying the parallelogram identity to the vectors $x - y_n$ and $x - y_p$, for $n, p \ge 1$, we obtain

$$\left\|x - \frac{y_n + y_p}{2}\right\|^2 + \left\|\frac{y_n - y_p}{2}\right\|^2 = \tfrac{1}{2}\bigl(\|x - y_n\|^2 + \|x - y_p\|^2\bigr).$$

Since C is convex, $(y_n + y_p)/2$ is in C, so $\frac{1}{4}\|y_n - y_p\|^2 \le \frac{1}{2}(1/n + 1/p)$, which proves that (y_n) is a Cauchy sequence in C and so converges to an element y of C, which must certainly satisfy $\|x - y\|^2 = \delta^2$.

Now let y_1 and y_2 be points of C with $\|x - y_1\| = \|x - y_2\| = \delta$. By applying the parallelogram identity as before, we get $\|y_1 - y_2\|^2 \le 0$, which says that $y_1 = y_2$. This shows that $P_C(x)$ is unique.

Finally, we check that the point $y = P_C(x)$ satisfies property $(*)$. If $z \in C$ and $t \in (0, 1]$, the point $(1 - t)y + tz$ belongs to C (which is convex), so

$$\|x - (1 - t)y - tz\|^2 \ge \|x - y\|^2,$$

or, after expansion,

$$t^2\|y-z\|^2 + 2t\,\mathrm{Re}(x-y\,|\,y-z) \geq 0.$$

Dividing by t and making t approach 0, we get

$$\mathrm{Re}(x-y\,|\,z-y) \leq 0.$$

Conversely, suppose a point y of C satisfies $(*)$. Then, for all $z \in C$,

$$\begin{aligned}\|x-z\|^2 &= \|(x-y)+(y-z)\|^2 \\ &= \|x-y\|^2 + \|y-z\|^2 + 2\,\mathrm{Re}(x-y\,|\,y-z) \geq \|x-y\|^2,\end{aligned}$$

so $y = P_C(x)$. □

Remarks

1. In the case $\mathbb{K} = \mathbb{R}$, the characterization $(*)$—where Re disappears—says that $P_C(x)$ is the unique point y of C such that, for all $z \in C$, the angle between the vectors $x-y$ and $z-y$ is at least $\pi/2$.
2. The conclusion of the theorem remains true if we suppose only that E is a scalar product space and that the convex set C is complete with respect to the induced metric—for example, if C is contained in a finite-dimensional vector subspace of E. Indeed, this assumption suffices to ensure that the sequence (y_n) of the proof converges to a point of C.

Condition $(*)$ allows us to show that P_C is a contraction, and therefore continuous.

Proposition 2.2 *Under the assumptions of Theorem* 2.1,

$$\|P_C(x_1) - P_C(x_2)\| \leq \|x_1 - x_2\| \quad \textit{for all } x_1, x_2 \in E.$$

Proof. Set $y_1 = P_C(x_1)$ and $y_2 = P_C(x_2)$. First,

$$\begin{aligned}\mathrm{Re}(x_1-x_2\,|\,y_1-y_2) &= \mathrm{Re}(x_1-y_2\,|\,y_1-y_2)+\mathrm{Re}(y_2-x_2\,|\,y_1-y_2) \\ &= \mathrm{Re}(x_1-y_1\,|\,y_1-y_2)+\|y_1-y_2\|^2+\mathrm{Re}(y_2-x_2\,|\,y_1-y_2) \\ &\geq \|y_1-y_2\|^2.\end{aligned}$$

Thus, by the Schwarz inequality, $\|y_1-y_2\|^2 \leq \|x_1-x_2\|\,\|y_1-y_2\|$, and finally $\|y_1-y_2\| \leq \|x_1-x_2\|$. □

We now consider projections onto vector subspaces of E.

Proposition 2.3 *Let F be a closed vector subspace of E. Then P_F is a linear operator from E onto F. If $x \in E$, the image $P_F(x)$ is the unique element $y \in E$ such that*

$$y \in F \quad \textit{and} \quad x-y \in F^{\perp}.$$

Proof. Condition $(*)$ of Theorem 2.1 becomes

$$y \in F \qquad \text{and} \qquad \operatorname{Re}(x - y \mid z - y) \leq 0 \quad \text{for all } z \in F.$$

Now, if $y \in F$ and $\lambda \in \mathbb{C}^*$, the map $z' \mapsto z = y + \bar{\lambda} z'$ is a bijection from F onto F. Condition $(*)$ is therefore equivalent to

$$y \in F \qquad \text{and} \qquad \operatorname{Re}\bigl(\lambda(x - y \mid z')\bigr) \leq 0 \quad \text{for all } z' \in F \text{ and } \lambda \in \mathbb{C},$$

and this in turn is obviously equivalent to

$$y \in F \qquad \text{and} \qquad x - y \in F^{\perp}.$$

That P_F is linear follows easily. □

Corollary 2.4 *For every closed vector subspace F of E, we have*

$$E = F \oplus F^{\perp}$$

and the projection operator on F associated with this direct sum is P_F.

Proof. For $x \in E$, we can write $x = P_F(x) + \bigl(x - P_F(x)\bigr)$ and, by Proposition 2.3, $P_F(x) \in F$ and $x - P_F(x) \in F^{\perp}$. On the other hand, if $x \in F \cap F^{\perp}$, then $(x \mid x) = 0$ and so $x = 0$. □

Remark. Proposition 2.3 and Corollary 2.4 remain valid under the weaker assumption that E is a scalar product space and F is complete in the induced metric — in particular, if F is finite-dimensional (see Remark 2 on page 106).

Under the preceding assumptions, P_F is called the **orthogonal projection (operator)** or **orthogonal projector** from E onto F. The image $P_F(x)$, for $x \in E$, is the **orthogonal projection** of x onto F.

Corollary 2.5 *For every vector subspace F of E,*

$$E = \bar{F} \oplus F^{\perp}.$$

In particular, F is dense in E if and only if $F^{\perp} = \{0\}$.

Proof. Just recall that $F^{\perp} = \bar{F}^{\perp}$. □

This very useful denseness criterion is now applied, as an example, to prove a result that will be generalized in the next chapter by other methods.

Proposition 2.6 *Let μ be a positive Radon measure on a locally compact, separable metric space X. Then $C_c(X)$ is dense in $L^2(\mu)$.*

Proof. We write $F = C_c(X)$. If f is an element of $F^\perp$, then $\int \varphi \bar{f}\, d\mu = 0$ for all $\varphi \in C_c(X)$. Thus, for all $\varphi \in C_c^{\mathbb{R}}(X)$,

$$\int \varphi\,(\mathrm{Re}\, f)^+\, d\mu = \int \varphi\,(\mathrm{Re}\, f)^-\, d\mu,$$
$$\int \varphi\,(\mathrm{Im}\, f)^+\, d\mu = \int \varphi\,(\mathrm{Im}\, f)^-\, d\mu.$$

By the uniqueness part of the Radon–Riesz Theorem (page 69), these equalities hold for any nonnegative Borel function φ. Applying them to the characteristic functions of the sets $\{\mathrm{Re}\, f > 0\}$, $\{\mathrm{Re}\, f < 0\}$, $\{\mathrm{Im}\, f > 0\}$, and $\{\mathrm{Im}\, f < 0\}$, we conclude that $f = 0$ μ-almost everywhere; that is, $f = 0$ as an element of $L^2(\mu)$. We finish by using Corollary 2.5. ☐

We conclude this section with an alternate form of Corollary 2.5.

Corollary 2.7 *If E is a Hilbert space and F is a vector subspace of E, then $\bar{F} = F^{\perp\perp}$.*

Proof. Clearly $F \subset F^{\perp\perp}$. Therefore, since $F^{\perp\perp}$ is closed, $\bar{F} \subset F^{\perp\perp}$. On the other hand, we have $E = \bar{F} \oplus F^\perp$ and $E = F^{\perp\perp} \oplus F^\perp$. The result follows immediately. ☐

Exercises

1. Let E be a Hilbert space.

a. Let C_1 and C_2 be nonempty, convex, closed subsets of E such that $C_1 \subset C_2$. Prove that, for all $x \in E$,

$$\left\|P_{C_1}(x) - P_{C_2}(x)\right\|^2 \leq 2\big(d(x, C_1)^2 - d(x, C_2)^2\big).$$

Hint. Apply the parallelogram identity to the vectors $x - P_{C_1}(x)$ and $x - P_{C_2}(x)$.

b. Let (C_n) be an increasing sequence of nonempty, convex, closed sets and let C be the closure of their union.

i. Prove that C is closed and convex.

ii. Prove that $\lim_{n\to+\infty} P_{C_n}(x) = P_C(x)$ for all $x \in E$.
Hint. Start by showing that

$$\lim_{n\to+\infty} d(x, C_n) = d(x, C).$$

c. Let (C_n) be a decreasing sequence of nonempty, convex, closed sets and let C be their intersection.

i. Prove that, if C is nonempty,

$$\lim_{n\to+\infty} P_{C_n}(x) = P_C(x) \quad \text{for all } x \in E.$$

ii. Prove that, if C is empty,

$$\lim_{n\to+\infty} d(x, C_n) = +\infty \quad \text{for all } x \in E.$$

$\bigl($In particular, if one of the C_n is bounded, C is nonempty. This result is false if we only assume E to be a Banach space: take, for example, $E = C([0,1])$ and $C_n = \{f \in E : |f| \le 1,\ f(0) = 1,$ and $f(x) = 0$ for all $x \ge 1/n\}$.$\bigr)$

2. a. Let a be a nonzero element of a Hilbert space E. Prove that, for all $x \in E$,

$$d\bigl(x, \{a\}^\perp\bigr) = \frac{\bigl|(x \,|\, a)\bigr|}{\|a\|}.$$

b. Take $E = L^2([0,1])$ (see Example 2 on page 124) and let F be the vector subspace of E defined by

$$F = \biggl\{ f \in E : \int_0^1 f(x)\,dx = 0 \biggr\}.$$

Determine $F^\perp$. Compute the distance to F of the element f of E defined by $f(x) = e^x$.

3. Let m be a measure on a measure space $(\Omega, \mathscr{F})$ and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of measurable subsets of Ω that partitions Ω. For every $n \in \mathbb{N}$ define

$$E_n = \biggl\{ f \in L^2(m) : \int_{\Omega\setminus A_n} |f|\,dm = 0 \biggr\}.$$

Prove that the E_n are pairwise orthogonal and that their union spans a dense subspace in $L^2(m)$. For each $n \in \mathbb{N}$, write down explicitly the orthogonal projection from $L^2(m)$ onto E_n.

4. Let P be a continuous linear map from a Hilbert space E to itself.

a. Prove that P is an orthogonal projection (onto a closed subspace of E) if and only if $P^2 = P$ and $\|P\| \le 1$.

b. Prove that, if P is an orthogonal projection,

$$(Px \,|\, y) = (x \,|\, Py) = (Px \,|\, Py) \quad \text{for all } x, y \in E.$$

5. Let c_{00} be the set of sequences of complex numbers almost all of whose terms are zero, endowed with the scalar product

$$(x \,|\, y) = \sum_{i\in\mathbb{N}} x_i \bar{y}_i.$$

Let f be the linear form on c_{00} defined by

$$f(x) = \sum_{i\in\mathbb{N}} \frac{x_i}{i+1}.$$

a. Prove that f is continuous.

b. Set $F = \ker f$. Prove that F is a closed vector subspace strictly contained in c_{00} and that $F^{\perp} = \{0\}$. (Thus the assumption that E is complete cannot be omitted from the statement of Corollary 2.4.)

6. Let μ be a positive Radon measure on a compact metric space X, with support equal to X. Consider the scalar product space $E = C(X)$ with scalar product defined by $(f \mid g) = \int f\bar{g}\,d\mu$. If A is a closed subset of X, we write $E_A = \{f \in C(X) : f(x) = 0 \text{ for all } x \in A\}$.
Let A be a closed subset of X.

a. Prove that there exists an increasing sequence (f_n) of functions in E_A, each with support $\overline{X \setminus A}$, that converges pointwise to $1_{X\setminus A}$.

b. Prove that $(E_A)^{\perp} = E_{\overline{X\setminus A}}$.
Hint. Prove that, if $g \in (E_A)^{\perp}$, then $\int 1_{X\setminus A}|g|^2\,d\mu = 0$.

c. Take $g \in C(X)$. Prove that $d(g, E_A)^2 = \int 1_A |g|^2\,d\mu$. Deduce that E_A is dense in E if and only if $\mu(A) = 0$. Prove also that g admits a projection onto E_A if and only if it vanishes on the boundary of A.

d. Suppose X has no isolated points. Prove that there exists a closed subset A of X with empty interior and such that $\mu(A) > 0$. Check that, for such an A, $(E_A)^{\perp} = \{0\}$ but E_A is not dense in E.
Hint. If there exists $a \in X$ such that $\mu(\{a\}) > 0$, one can take $A = \{a\}$. Otherwise, consider a countable dense subset of X and use the fact that μ is regular (Exercise 5 on page 77).

7. Let m be a measure on a measure space $(\Omega, \mathcal{F})$. Suppose m is σ-finite; that is, Ω is a countable union of elements of $\mathcal{F}$ of finite m-measure. Define $L^2(m) \otimes L^2(m)$ as the vector space generated by functions of the form $(x, y) \mapsto f(x)g(y)$, with $f, g \in L^2(m)$. Prove that $L^2(m) \otimes L^2(m)$ is dense in $L^2(m \times m)$.
Hint. Let (A_n) be an increasing sequence of elements of $\mathcal{F}$ of finite measure and covering Ω. Let F be an element of the orthogonal space to $L^2(m) \otimes L^2(m)$ in $L^2(m \times m)$. Prove that, for all $n \in \mathbb{N}$, the set consisting of all $T \in \mathcal{F} \times \mathcal{F}$ such that

$$\iint_{T\cap(A_n\times A_n)} F(x, y)\,dm(x)\,dm(y) = 0$$

contains $\{A \times B : A, B \in \mathcal{F}\}$ and is a monotone class; this term is defined in Exercise 2 on page 64. Deduce from the same exercise that $F = 0$.

8. *The bipolar theorem.* Let E be a complex (say) Hilbert space. If A is a nonempty subset of E, the *polar* of A is defined as

$$A^0 = \{x \in E : \operatorname{Re}(x \mid y) \le 1 \text{ for all } y \in A\}.$$

The set A^{00} is called the *bipolar* of A.

a. Prove that the polar of any nonempty subset of E is a closed convex set containing 0.

b. Deduce that, if A is a nonempty subset of E, the closed convex hull of $A \cup \{0\}$ (see Exercise 9 on page 18) is contained in A^{00}.

c. We now want to show the reverse inclusion. Let C be the closed convex hull of $A \cup \{0\}$ and take $x \in A^{00}$.

  i. Prove that $\operatorname{Re}\big(x - P_C(x) \,|\, P_C(x)\big) \geq 0$.

  ii. Prove that, for all $\varepsilon > 0$,

$$\frac{1}{\varepsilon + \operatorname{Re}\big(x - P_C(x) \,|\, P_C(x)\big)} \big(x - P_C(x)\big) \in A^0.$$

  Deduce that $\big\|x - P_C(x)\big\|^2 \leq \varepsilon$, and so that $x \in C$.

d. Let A be a convex subset of E containing 0. Prove that $\bar{A} = A^{00}$.

e. Let A be a vector subspace of E. Prove that $A^0 = A^\perp$. (We thus recover the equality $\bar{A} = A^{\perp\perp}$.)

3 The Riesz Representation Theorem

We assume in this section that E is a Hilbert space. The Riesz Representation Theorem, which we now state, describes the topological dual of E.

Theorem 3.1 (Riesz) *The map from E to E' defined by $y \mapsto \varphi_y = (\cdot \,|\, y)$ is a surjective isometry. In other words, given any continuous linear form φ on E, there exists a unique $y \in E$ such that*

$$\varphi(x) = (x \,|\, y) \quad \textit{for all } x \in E,$$

and, furthermore, $\|\varphi\| = \|y\|$.

Proof. That this map is an isometry was seen in Corollary 1.3. We now show it is surjective. Take $\varphi \in E'$ such that $\varphi \neq 0$. We know from Corollary 2.4 that $E = \ker\varphi \oplus (\ker\varphi)^\perp$, since, φ being continuous, $\ker\varphi$ is closed. Now, φ is a nonzero linear form, so $\ker\varphi$ has codimension 1. The space $(\ker\varphi)^\perp$ therefore has dimension 1; it is generated by a vector e, which we can choose to have norm 1. Set $y = \overline{\varphi(e)}\, e$ if $\mathbb{K} = \mathbb{C}$, or $y = \varphi(e)e$ if $\mathbb{K} = \mathbb{R}$. Then $\varphi_y(e) = \varphi(e)$ and $\varphi_y = 0$ on $\ker\varphi$. It follows that φ_y and φ coincide on $(\ker\varphi)^\perp$ and on $\ker\varphi$, so $\varphi = \varphi_y$. □

We recall that this isometry is linear if $\mathbb{K} = \mathbb{R}$ and skew-linear if $\mathbb{K} = \mathbb{C}$.

The rest of this section is devoted to some important applications of Theorem 3.1.

3A Continuous Linear Operators on a Hilbert Space

Recall that $L(E)$ denotes the space of continuous linear maps (or operators) from E to E. We use the same symbol for the norm in E and the associated norm in $L(E)$. We denote by I the identity on E.

Proposition 3.2 *Given $T \in L(E)$, there exists a unique operator $T^* \in L(E)$ such that*

$$(Tx \mid y) = (x \mid T^*y) \quad \textit{for all } x, y \in E.$$

Moreover, $\|T^*\| = \|T\|$.

T^* is called the **adjoint** of T.

Proof. Take $y \in E$. The map $\varphi_y \circ T : x \mapsto (Tx \mid y)$ is an element of E', so by Theorem 3.1 there exists a unique element of E, which we denote by T^*y, such that

$$(Tx \mid y) = (x \mid T^*y) \quad \text{for all } x \in E;$$

moreover $\|T^*y\| = \|\varphi_y \circ T\| \le \|y\| \, \|T\|$. The uniqueness of such a T^*y easily shows that T^* is linear; at the same time, by the preceding inequality, $\|T^*\| \le \|T\|$. Moreover, if $x \in E$,

$$\|Tx\|^2 = (Tx \mid Tx) = (x \mid T^*Tx) \le \|x\| \, \|T^*\| \, \|Tx\|,$$

which implies that $\|Tx\| \le \|x\| \, \|T^*\|$, and so that $\|T\| \le \|T^*\|$. □

The properties in the next proposition are easily deduced from the definition of the adjoint.

Proposition 3.3 *The map from $L(E)$ to itself defined by $T \mapsto T^*$ is linear if $\mathbb{K} = \mathbb{R}$ and skew-linear if $\mathbb{K} = \mathbb{C}$. It is also an isometry and an involution (that is, $T^{**} = T$ for $T \in L(E)$). We have $I^* = I$ and $(TS)^* = S^*T^*$ for all $T, S \in L(E)$.*

Examples

1. Take $E = \mathbb{R}^d$ with the canonical euclidean structure. The space $L(E)$ can be identified with the space $M_d(\mathbb{R})$ of $d \times d$ matrices with real entries. Then T^* is the transpose of T. If $E = \mathbb{C}^d$ with the canonical hermitian structure, the space $L(E)$ can be identified with $M_d(\mathbb{C})$ and T^* is the conjugate of the transpose of T.
2. The next example can be regarded as an extension of the preceding one to infinite dimension. Let m be a measure on a measure space $(\Omega, \mathscr{F})$. Suppose m is **σ-finite**; that is, Ω is a countable union of elements of $\mathscr{F}$ of finite m-measure. This entails we can use Fubini's Theorem. We place

ourselves in the Hilbert space $E = L^2(m)$, and take $K \in L^2(m \times m)$. If $f \in E$, we define $T_K f(x)$ for m-almost every x by

$$T_K f(x) = \int K(x,y) f(y)\, dm(y).$$

Since, by the Schwarz inequality,

$$\int \Bigl(\int |K(x,y)|\, |f(y)|\, dm(y) \Bigr)^2 dm(x)$$
$$\leq \int |f(y)|^2\, dm(y) \iint |K(x,y)|^2\, dm(x)\, dm(y) < +\infty,$$

this expression defines an element $T_K f$ of E such that

$$\|T_K f\|^2 \leq \|f\|^2 \iint |K(x,y)|^2\, dm(x)\, dm(y),$$

which shows that T_K is a continuous linear operator on E whose norm is at most the norm of K in $L^2(m \times m)$. By Fubini's Theorem, if $f, g \in E$, we have, in the case $\mathbb{K} = \mathbb{C}$,

$$(T_K f \mid g) = \int f(y) \Bigl(\int K(x,y) \overline{g(x)}\, dm(x) \Bigr) dm(y) = (f \mid T_{K^*} g),$$

where we have put $K^*(x,y) = \overline{K(y,x)}$. Thus $T_K^* = T_{K^*}$. Naturally, in the case $\mathbb{K} = \mathbb{R}$, we get the same result with $K^*(x,y) = K(y,x)$.

The next property will be useful in the sequel.

Proposition 3.4 *For every $T \in L(E)$, we have $\|TT^*\| = \|T^*T\| = \|T\|^2$.*

Proof. Certainly $\|T^*T\| \leq \|T\|^2$. On the other hand,

$$\|Tx\|^2 = (Tx \mid Tx) = (x \mid T^*Tx) \leq \|x\|^2 \|T^*T\|,$$

which shows that $\|T\|^2 \leq \|T^*T\|$. Therefore $\|T^*T\| = \|T\|^2$ and, applying this result to T^*, we get $\|TT^*\| = \|T^*\|^2 = \|T\|^2$. □

An operator $T \in L(E)$ is called **selfadjoint** if $T = T^*$. We also call such operators **symmetric** if $\mathbb{K} = \mathbb{R}$ and **hermitian** if $\mathbb{K} = \mathbb{C}$. By the preceding proposition, if T is selfadjoint then $\|T^2\| = \|T\|^2$.

Examples

1. For every operator $T \in L(E)$, TT^* and T^*T are selfadjoint.
2. In Example 2 on the preceding page, T_K is selfadjoint if and only if, for $(m \times m)$-almost every (x,y), we have $K(x,y) = K(y,x)$ (if $\mathbb{K} = \mathbb{R}$) or $K(x,y) = \overline{K(y,x)}$ (if $\mathbb{K} = \mathbb{C}$). This condition is clearly sufficient and it is necessary by Exercise 7 on page 110.

3. Every orthogonal projection operator is selfadjoint (see Exercise 4 on page 109).

Note that, if T is a selfadjoint operator, $(Tx \mid x) \in \mathbb{R}$ for all $x \in E$. We say that $T \in L(E)$ is **positive selfadjoint** if

$$(Tx \mid x) \in \mathbb{R}^+ \quad \text{for all } x \in E.$$

Warning! If E is a function space, this notion of positivity has nothing to do with the condition $f \geq 0 \Longrightarrow Tf \geq 0$. In particular, in Example 2 above, T_K is positive selfadjoint if, for all $f \in L^2(m)$,

$$\iint K(x,y)\, f(x)\, \overline{f(y)}\, dm(x)\, dm(y) \geq 0,$$

and it is positive in the other sense if $K \geq 0$, which is altogether different.

One checks immediately that, for all $T \in L(E)$, the operators TT^* and T^*T are positive selfadjoint.

The last result of this section gives another expression of the norm of a selfadjoint operator.

Proposition 3.5 *Assume $E \neq \{0\}$. For every selfadjoint operator $T \in L(E)$,*

$$\|T\| = \sup\bigl\{|(Tx \mid x)| : x \in E \text{ and } \|x\| = 1\bigr\}.$$

Proof. Let γ be the right-hand side of the equality. Clearly $\gamma \leq \|T\|$ and, for all $x \in E$, $\bigl|(Tx \mid x)\bigr| \leq \gamma \|x\|^2$. Assume for example that $\mathbb{K} = \mathbb{C}$, and take $y, z \in E$ and $\lambda \in \mathbb{R}$. Then

$$\bigl|(T(y \pm \lambda z) \mid y \pm \lambda z)\bigr| = \bigl|(Ty \mid y) \pm 2\lambda \operatorname{Re}(Ty \mid z) + \lambda^2 (Tz \mid z)\bigr| \leq \gamma \|y \pm \lambda z\|^2.$$

We deduce, by combining the two inequalities, that

$$4\,|\lambda|\,\bigl|\operatorname{Re}(Ty \mid z)\bigr| \leq \gamma\bigl(\|y + \lambda z\|^2 + \|y - \lambda z\|^2\bigr) = 2\gamma\bigl(\|y\|^2 + \lambda^2 \|z\|^2\bigr),$$

and this holds for any real λ. We conclude that $\bigl|\operatorname{Re}(Ty \mid z)\bigr| \leq \gamma \|y\|\, \|z\|$, from the condition for a polynomial function on $\mathbb{R}$ of degree at most 2 to be nonnegative-valued. Now it is enough to choose $z = Ty$ to obtain $\|Ty\| \leq \gamma \|y\|$ for all $y \in E$, and hence $\|T\| \leq \gamma$. □

3B Weak Convergence in a Hilbert Space

We say that a sequence (x_n) in E **converges weakly** to $x \in E$ if

$$\lim_{n \to +\infty} (x_n \mid y) = (x \mid y) \quad \text{for all } y \in E.$$

In this case x is called the **weak limit** of the sequence (x_n). Clearly a sequence can have no more than one weak limit.

One deduces immediately from the Schwarz inequality that a sequence (x_n) of E that converges to a point x of E in the sense of the norm of E (one for which $\lim_{n\to+\infty} \|x_n - x\| = 0$) also converges weakly to x. The converse is generally false. For example, it is easy to check that the sequence (x_n) in $E = \ell^2$ defined by

$$(x_n)_j = \begin{cases} 1 & \text{if } j = n, \\ 0 & \text{otherwise} \end{cases}$$

converges weakly to 0, whereas $\|x_n\| = 1$ for all n. For this reason we sometimes call convergence in the sense of the norm **strong convergence**.

The next proposition pinpoints the relationship between weak and strong convergence.

Proposition 3.6 *Let (x_n) be a sequence in E that converges weakly to x. Then*

$$\liminf_{n\to+\infty} \|x_n\| \geq \|x\|.$$

Moreover, the following properties are equivalent:

1. *The sequence (x_n) converges (strongly) to x.*
2. $\limsup_{n\to+\infty} \|x_n\| \leq \|x\|$.
3. $\lim_{n\to+\infty} \|x_n\| = \|x\|$.

Proof. First,

$$\|x\|^2 = \lim_{n\to+\infty} \bigl|(x \mid x_n)\bigr| \leq \|x\| \liminf_{n\to+\infty} \|x_n\|,$$

which proves the first statement. At the same time, $\|x - x_n\|^2 = \|x\|^2 + \|x_n\|^2 - 2\,\mathrm{Re}(x_n \mid x)$, so

$$\limsup_{n\to+\infty} \|x - x_n\|^2 \leq \bigl(\limsup_{n\to+\infty} \|x_n\|\bigr)^2 - \|x\|^2,$$

which yields the equivalence between 1 and 2. The equivalence between 2 and 3 follows immediately from the first statement. □

The Riesz Representation Theorem enables us to prove the following version of the Banach–Alaoglu Theorem in a Hilbert space.

Theorem 3.7 *Any bounded sequence in E has a weakly convergent subsequence.*

Proof. Suppose first that E is separable. Let (x_n) be a bounded sequence in E. In the notation of Theorem 3.1, the Banach–Alaoglu Theorem (page 19) applied to the sequence (φ_{x_n}) guarantees the existence of a subsequence (x_{n_k}) and of a $\varphi \in E'$ such that

$$\lim_{k\to+\infty} \varphi_{n_k}(y) = \varphi(y) \quad \text{for all } y \in E.$$

By Theorem 3.1, there exists an element $x \in E$ such that $\varphi = \varphi_x$, which proves the theorem in the separable case.

We turn to the general case. Let (x_n) be a bounded sequence in E and let F be the closure of the vector subspace of E spanned by $\{x_n\}_{n\in\mathbb{N}}$. By construction, this is a separable Hilbert space. The first part of the proof says that there exists a subsequence (x_{n_k}) and a point $x \in F$ such that

$$\lim_{k\to+\infty} (x_{n_k} \,|\, y) = (x \,|\, y) \quad \text{for all } y \in F.$$

Since this equality obviously takes place also if $y \in F^\perp$, it suffices now to apply Corollary 2.4. □

The fact that any continuous linear operator has an adjoint allows us to prove the following property.

Proposition 3.8 *Let (x_n) be a sequence in E converging weakly to x. Then, for all $T \in L(E)$, the sequence (Tx_n) converges weakly to Tx.*

Proof. For every $y \in E$,

$$\lim_{n\to+\infty} (Tx_n \,|\, y) = \lim_{n\to+\infty} (x_n | T^* y) = (x | T^* y) = (Tx \,|\, y).$$ □

Exercises

1. *Theorem of Lax–Milgram. Galerkin approximation.* Let E be a real Hilbert space and a a bilinear form on E. Assume that a is continuous and *coercive*: this means that there exist constants $C > 0$ and $\alpha > 0$ such that

$$\begin{aligned} |a(x,y)| &\le C\|x\|\,\|y\| && \text{for all } x, y \in E,\\ a(x,x) &\ge \alpha\|x\|^2 && \text{for all } x \in E. \end{aligned}$$

a. **i.** Show there exists a continuous linear operator T on E such that

$$a(x,y) = (Tx \,|\, y) \quad \text{for all } x, y \in E.$$

ii. Prove that $T(E)$ is dense in E.

iii. Prove that $\|Tx\| \ge \alpha\|x\|$ for all $x \in E$. Deduce that T is injective and that $T(E)$ is closed.

iv. Deduce that T is an isomorphism from E onto itself.

b. Let L be a continuous linear form on E.

i. Deduce from the preceding questions that there exists a unique $u \in E$ such that

$$a(u,y) = L(y) \quad \text{for all } y \in E.$$

ii. Now suppose that the bilinear form a is symmetric and define, for $x \in E$,

$$\Phi(x) = \tfrac{1}{2}a(x,x) - L(x).$$

Prove that the point u is characterized by the condition

$$\Phi(u) = \min_{x \in E} \Phi(x).$$

c. We return to the notation and situation of question 1b-i. Let (E_n) be an increasing sequence of closed vector subspaces of E whose union is dense in E.

i. Prove that, for any integer $n \in \mathbb{N}$, there exists a unique $u_n \in E_n$ such that

$$a(u_n, y) = L(y) \quad \text{for all } y \in E_n.$$

Check, in particular, that if E_n has finite dimension d_n, determining u_n reduces to solving a linear system of the form $A_n U_n = Y_n$, where A_n is an invertible $d_n \times d_n$ matrix, which, moreover, is symmetric and positive definite if a is symmetric.

ii. Prove that, for any $n \in \mathbb{N}$,

$$\|u - u_n\| \leq \frac{C}{\alpha}\, d(u, E_n).$$

Deduce that the sequence (u_n) converges to u.
Hint. Take $y \in E_n$. Prove that

$$a(u - u_n,\ u - u_n) = a(u - u_n,\ u - y)$$

and deduce that $\alpha\|u - u_n\| \leq C\|u - y\|$.

2. *Lions–Stampacchia Theorem (symmetric case).* Consider a real Hilbert space E, a nonempty, closed, convex set C in E, a continuous and coercive (Exercise 1) bilinear symmetric form a on E, and a continuous linear form L on E. Let J be the function defined on E by

$$J(u) = a(u,u) - 2L(u) \quad \text{for all } u \in E.$$

Prove that there exists a unique $c \in C$ such that $J(c) \leq J(v)$ for all $v \in C$, and that c is characterized by the following condition:

$$a(c,\ v - c) \geq L(v - c) \quad \text{for all } v \in C.$$

Hint. By the Lax–Milgram Theorem (Exercise 1), there exists a unique element u of E such that $a(u,v) = L(v)$ for all $v \in E$. Check that $J(v) = a(v-u,\ v-u) - a(u,u)$, then work in the Hilbert space (E, a).

3. *Reproducing kernels.* Let X be a set and $\mathscr{F}$ the vector space of complex-valued functions on X.

a. Consider a vector subspace E of $\mathcal{F}$ endowed with a Hilbert space structure such that, for all $x \in X$, the linear form defined on E by $f \mapsto f(x)$ is continuous.

i. Prove that there exists a unique function K from X^2 to $\mathbb{C}$ satisfying these conditions:

- For all $y \in E$, the function $K(\cdot\,, y) : x \mapsto K(x, y)$ lies in E.
- For all $f \in E$ and $y \in X$, we have $\big(f \,|\, K(\cdot\,, y)\big) = f(y)$.

We call K the *reproducing kernel* of E.

ii. Prove:

A. For all $x, y \in E$, we have $\overline{K(x,y)} = K(y,x)$.

B. For all $n \in \mathbb{N}^*$, all $(\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$, and all $(x_1, \ldots, x_n) \in X^n$, we have

$$\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, x_j)\bar{\xi}_i\xi_j \geq 0.$$

iii. Prove that the family $\{K(\cdot\,, y)\}_{y \in X}$ is fundamental in E.

b. Conversely, consider a function K from X^2 to $\mathbb{C}$ satisfying properties A and B above.

i. Let $\mathcal{E}$ be the vector subspace of $\mathcal{F}$ spanned by $\{K(\cdot\,, y)\}_{y \in X}$. Prove that the relation

$$\left(\sum_{j=1}^{n} \lambda_j K(\cdot\,, x_j) \,\middle|\, \sum_{k=1}^{m} \mu_k K(\cdot\,, y_k)\right) = \sum_{j=1}^{n}\sum_{k=1}^{m} K(y_k, x_j)\lambda_j\bar{\mu}_k$$

defines a scalar semiproduct on $\mathcal{E}$. Check, in particular, that this expression does not depend on the representations involved.

ii. Let $(E^{\tilde{}}, (\cdot \,|\, \cdot)^{\tilde{}})$ be the Hilbert completion of $\mathcal{E}$ and let L be the associated canonical map from $\mathcal{E}$ to $E^{\tilde{}}$ (Exercise 8 on page 104). Define an application $\Psi : E^{\tilde{}} \to \mathcal{F}$ by

$$\Psi(\varphi)(x) = \big(\varphi \,|\, L(K(\cdot\,, x))\big)^{\tilde{}}.$$

Prove that Ψ is injective.

iii. Derive a Hilbert space structure for $E = \Psi(E^{\tilde{}})$, with respect to which K is the reproducing kernel.

c. Suppose $X = \mathbb{R}$ and fix a Borel measure μ of finite mass on $\mathbb{R}$. If $h \in L^2(\mu)$, denote by f_h the element of $\mathcal{F}$ defined by

$$f_h(x) = \int e^{itx} h(t)\, d\mu(t).$$

i. Prove that the map $h \mapsto f_h$ thus defined on $L^2(\mu)$ is injective (see Exercise 1c on page 63).

ii. Set $E = \{f_h : h \in L^2(\mu)\}$. For $h, k \in L^2(\mu)$, set

$$(f_h \mid f_k) = \int h\bar{k}\, d\mu.$$

Prove that E is a Hilbert space having as a reproducing kernel the function $K(x, y) = \int e^{it(x-y)}\, d\mu(t)$.

d. Let Ω be open in $\mathbb{C}$. Prove that the Hilbert space $H^2(\Omega)$ defined in Exercise 4 on page 102 has a reproducing kernel. This is called the *Bergman kernel.*

4. Let E be a Hilbert space over $\mathbb{C}$, distinct from $\{0\}$. If $T \in L(E)$, write

$$n(T) = \sup\bigl\{|(Tx \mid x)| : \|x\| = 1\bigr\}.$$

a. Prove that

$$n(T) \le \|T\| \le 2n(T) \quad \text{for all } T \in L(E). \tag{$*$}$$

Hint. For the second inequality, draw inspiration from the proof of Proposition 3.5 to show that, for every $x, y \in E$ and $S \in L(E)$,

$$\bigl|(Sx \mid y) + (Sy \mid x)\bigr| \le 2n(S)\,\|x\|\,\|y\|.$$

Then set $S = \lambda T$ and $y = \lambda Tx$, were λ is a complex number of absolute value 1 such that $\lambda^2(T^2x \mid x) \in \mathbb{R}^+$.

b. Prove that $(*)$ would be false if E were a Hilbert space over $\mathbb{R}$.

c. Prove that, if E has dimension at least 2, the constant 2 in $(*)$ cannot be replaced by a smaller real number.
Hint. Let u and v be orthogonal vectors in E, each of norm 1. Consider the operator defined on E by

$$T(\lambda u + \mu v + w) = \lambda v \quad \text{for all } \lambda, \mu \in \mathbb{K} \text{ and } w \in \{u, v\}^\perp.$$

d. Prove that the map $T \mapsto n(T)$ is a norm on $L(E)$ equivalent to the norm $\|\cdot\|$.

5. Let E be a Hilbert space over $\mathbb{C}$.

a. Take $T \in L(E)$. Prove that T is hermitian if and only if $(Tx \mid x) \in \mathbb{R}$ for all $x \in E$.
Hint. In the notation of Exercise 4, $T = T^*$ if and only if $n(T - T^*) = 0$.

b. Deduce that an operator T on E is positive hermitian if and only if $(Tx \mid x) \in \mathbb{R}^+$ for all $x \in E$.

6. Let T be a positive selfadjoint operator on a Hilbert space E.

a. Prove that

$$\bigl|(Tx \mid y)\bigr|^2 \le (Tx \mid x)(Ty \mid y) \quad \text{for all } x, y \in E.$$

Hint. Prove that $(x, y) \mapsto (Tx \mid y)$ is a scalar semiproduct on E and so satisfies the Schwarz inequality.

b. Derive another proof of Proposition 3.5 in this case.

7. Let P be a continuous linear operator on a Hilbert space E. We assume that P is a projection ($P^2 = P$). Prove that the following properties are equivalent:

- P is an orthogonal projection operator.
- P is selfadjoint: $P = P^*$.
- P is normal: $PP^* = P^*P$.
- $(Px \mid x) = \|Px\|^2$ for all $x \in E$.

8. Consider a Hilbert space E and an element $T \in L(E)$.

a. Prove that $\ker T^* = (\operatorname{im} T)^\perp$. Deduce that $\overline{\operatorname{im} T} = (\ker T^*)^\perp$, then that $\overline{\operatorname{im} T} = E$ if and only if T^* is injective.

b. Assume T is positive selfadjoint. Prove that an element $x \in E$ satisfies $Tx = 0$ if and only if $(Tx \mid x) = 0$ (use Exercise 6a above). Deduce that T is injective if and only if $(Tx \mid x) > 0$ for all $x \neq 0$.

9. *An ergodic theorem.* Consider a Hilbert space E and an element $T \in L(E)$ such that $\|T\| \leq 1$.

a. Prove that an element $x \in E$ satisfies $Tx = x$ if and only if $(Tx \mid x) = \|x\|^2$. (Use the fact that equality in the Schwarz inequality implies collinearity.) Deduce that $\ker(I - T) = \ker(I - T^*)$.

b. Show that $(\operatorname{im}(I - T))^\perp = \ker(I - T)$ (use Exercise 8a above) and deduce that

$$E = \ker(I - T) \oplus \overline{\operatorname{im}(I - T)}.$$

c. For $n \geq 1$, set

$$T_n = \frac{I + T + \cdots + T^n}{n + 1}.$$

Show that $\lim_{n \to +\infty} T_n x = Px$ for all $x \in E$, where P is the orthogonal projection onto $\ker(I - T)$.

Hint. Consider successively the cases $x \in \ker(I - T)$, $x \in \operatorname{im}(I - T)$, and $x \in \overline{\operatorname{im}(I - T)}$. In this last case, you might use Proposition 4.3 on page 19.

10. Let E be a Hilbert space.

a. Prove that every weakly convergent sequence in E is bounded.

Hint. Use the Banach–Steinhaus Theorem (Exercise 6d on page 22).

b. Let (x_n) and (y_n) be sequences in E. Prove that if (x_n) converges weakly to x and (y_n) converges strongly to y, the sequence $\big((x_n \mid y_n)\big)$ converges to $(x \mid y)$. What if we suppose only that (y_n) converges weakly to y ?

11. Let (x_n) be a sequence in a Hilbert space E. Prove that if, for all $y \in E$, the sequence $\big((x_n \mid y)\big)$ is convergent, the sequence (x_n) is weakly convergent.

Hint. Use Exercise 6f on page 23.

12. Let K be a compact subset of a Hilbert space E. Prove that every sequence in K that converges weakly also converges strongly.

13. Prove that in a finite-dimensional Hilbert space every weakly convergent sequence is strongly convergent. You might give a direct proof, not using Exercises 10 and 12.

14. Let D be a fundamental subset of a Hilbert space E. Prove that if (x_n) is a bounded sequence in E and if $\lim_{n\to+\infty}(x_n \mid y) = (x \mid y)$ for all $y \in D$, then (x_n) converges weakly to x. Prove that the assumption that (x_n) is bounded is necessary (see Exercise 10a above).

15. **a.** Let (x_n) be a weakly convergent sequence in a Hilbert space and let x be its weak limit. Prove that x lies in the closed convex hull of the set $\{x_n\}_{n\in\mathbb{N}}$.

Hint. Let C be the closed convex hull of the set $\{x_n\}_{n\in\mathbb{N}}$. Prove that $x = P_C x$.

b. Let C be a convex subset of a Hilbert space E. Prove that C is closed if and only if the weak limit of every weakly convergent sequence of points in C is an element of C.

16. *Banach–Saks Theorem.*

a. Let (x_n) be a sequence in a Hilbert space E converging weakly to $x \in E$. Prove that there exists a subsequence (x_{n_k}) such that the sequence (y_k) defined by

$$y_k = \frac{1}{k}\,(x_{n_1} + x_{n_2} + \cdots + x_{n_k})$$

converges (strongly) to x.

Hint. Reduce to the case where $x = 0$. Then construct (by induction) a strictly increasing sequence (n_k) of integers such that, for all $k \geq 2$,

$$\bigl|(x_{n_1} \mid x_{n_k})\bigr| \leq 1/k, \quad \bigl|(x_{n_2} \mid x_{n_k})\bigr| \leq 1/k, \quad \ldots, \quad \bigl|(x_{n_{k-1}} \mid x_{n_k})\bigr| \leq 1/k.$$

Then use Exercise 10a.

b. Deduce another demonstration of the result of Exercise 15.

17. *A particular case of the Browder Fixed-Point Theorem.* Let C be a nonempty, convex, closed and bounded subset of a Hilbert space E.

a. Let T be a map from C to C such that

$$\bigl\|T(x) - T(y)\bigr\| \leq \|x - y\| \quad \text{for all } x, y \in C.$$

i. Let a be a point of C. For every $n \in \mathbb{N}^*$ and $x \in C$, define

$$T_n(x) = \frac{1}{n}\,a + \frac{n-1}{n}\,T(x).$$

Show that there exists a unique point $x_n \in C$ such that $T_n(x_n) = x_n$.

Hint. The map T_n is strictly contracting.

ii. Let (x_{n_k}) be a weakly convergent subsequence of the sequence (x_n), tending to the weak limit x (see Theorem 3.7). Set $y_n = x_n - a$ and $y = x - a$. Prove that, for all $n \geq 2$,

$$\|y_n\|^2 \leq \frac{2n-2}{2n-1} \operatorname{Re}(y_n \mid y).$$

Deduce that the sequence (x_{n_k}) converges strongly to x, that $x \in C$, and that $T(x) = x$.

iii. Prove that the set $\{x \in C : T(x) = x\}$ is convex, closed, and nonempty.
Hint. To show convexity, take $x_0, x_1 \in C$ such that $T(x_0) = x_0$ and $T(x_1) = x_1$ and, for $t \in [0,1]$, set $x_t = tx_1 + (1-t)x_0$. Prove that

$$\|x_0 - x_1\| = \big\|T(x_t) - x_0\big\| + \big\|x_1 - T(x_t)\big\|.$$

Using the case of equality in the Schwarz inequality, deduce that $T(x_t) = x_t$.

b. Let $\mathscr{T}$ be a family of maps from C to C such that

- $T \circ S = S \circ T$ for all $T, S \in \mathscr{T}$, and
- $\big\|T(x) - T(y)\big\| \leq \|x - y\|$ for all $T \in \mathscr{T}$ and $x, y \in C$.

Suppose also that E is separable. Show that there exists a point $x \in C$ such that

$$T(x) = x \quad \text{for all } T \in \mathscr{T}.$$

Hint. Show first that there exists a metric that makes C compact; then argue as in Exercise 19a on page 85.

18. Consider a nonempty, convex, closed and bounded subset C of a real Hilbert space E, and a differentiable function J from E to $\mathbb{R}$. Recall that J is called *convex* on C if, for any pair (u, v) of points in C and any $\theta \in [0,1]$,

$$J\big(\theta u + (1-\theta)v\big) \leq \theta J(u) + (1-\theta)J(v).$$

By definition, the *gradient* of J at u, denoted by $\nabla J(u)$, is the element of E that the Riesz Representation Theorem associates to the derivative map $J'(u)$.

a. Prove that J is convex on C if and only if, for all $(u,v) \in C^2$,

$$J(v) \geq J(u) + \big(\nabla J(u) \mid v - u\big).$$

In particular, deduce that, if J is convex, it is bounded below over C.

b. Prove that if J is convex there exists at least one point $u_* \in C$ such that

$$J(u_*) = \inf_{u \in C} J(u).$$

You might proceed in the following way: Let m be the infimum on the right-hand side, and let (u_n) be a sequence in C such that $\lim_{n\to+\infty} J(u_n) = m$.

i. Prove that (u_n) has a weakly convergent subsequence (u_{n_k}).

ii. Let u_* be the weak limit of (u_{n_k}). Prove that $u_* \in C$ (see Exercise 15).

iii. Prove that $J(u_*) = m$.

c. Under the same hypotheses and with the same notation, prove that the set $C_0 = \{u_* \in C : J(u_*) = m\}$ is convex and closed. Prove also that $u \in C_0$ if and only if $(\nabla J(u) \mid v - u) \geq 0$ for all $v \in C$.

d. *An example of a convex function.* Take $T \in L(E)$ and $\Phi \in E'$, and set $J(u) = (Tu \mid u) + \Phi(u)$. Prove that J is convex on E if and only if the operator $T + T^*$ is positive selfadjoint.

4 Hilbert Bases

We consider a scalar product space E. A family $(X_i)_{i\in I}$ of elements of E is called **orthogonal** if $X_i \perp X_j$ whenever $i \neq j$. For such a family, the Pythagorean Theorem implies that, for any finite subset J of I,

$$\Bigl\|\sum_{i\in J} X_i\Bigr\|^2 = \sum_{i\in J} \|X_i\|^2.$$

Here is an immediate consequence of this:

Proposition 4.1 *An orthogonal family that does not include the zero vector is free.*

Proof. Let J be a finite subset of I and let $(\lambda_j)_{j\in J}$ be elements of $\mathbb{K}$ such that $\sum_{j\in J} \lambda_j X_j = 0$. Then

$$\Bigl\|\sum_{j\in J} \lambda_j X_j\Bigr\|^2 = \sum_{j\in J} |\lambda_j|^2 \, \|X_j\|^2 = 0,$$

which clearly implies that $\lambda_j = 0$ for all $j \in J$. □

An orthogonal family all of whose elements have norm 1 is called **orthonormal**. The preceding proposition shows that such a family is free. A fundamental orthonormal family in E is called a **Hilbert basis** of E. Thus a Hilbert basis is, in particular, a topological basis.

We give some fundamental examples.

Examples

1. Suppose $a > 0$ and let $C_a^{\mathbb{K}}$ be the space of continuous functions periodic of period a from $\mathbb{R}$ to $\mathbb{K}$, with the scalar product defined on page 98. For $n \in \mathbb{Z}$, we set
$$e_n(x) = e^{2i\pi nx/a}.$$
It is straightforward to show that the family $(e_n)_{n\in\mathbb{Z}}$ is orthonormal in $C_a^{\mathbb{C}}$. As in the particular case of Example 4 on page 35, this family is fundamental in $C_a^{\mathbb{C}}$ with the uniform norm. Since the norm associated with the scalar product never exceeds the uniform norm, the family $(e_n)_{n\in\mathbb{Z}}$ is a Hilbert basis of the scalar product space $C_a^{\mathbb{C}}$. It follows easily that the family
$$\left\{1,\ \sqrt{2}\cos\frac{2\pi}{a}x,\ \sqrt{2}\sin\frac{2\pi}{a}x,\ \dots,\ \sqrt{2}\cos\frac{2\pi n}{a}x,\ \sqrt{2}\sin\frac{2\pi n}{a}x,\ \dots\right\}$$
is a Hilbert basis of the scalar product space $C_a^{\mathbb{K}}$, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.
2. If A is a Borel set in $\mathbb{R}$, we denote by $L^2(A)$ the space $L^2(\lambda, A)$ associated with the restriction of Lebesgue measure to the Borel sets of A. Let $E = L^2((0,1))$. Clearly $L^2((0,1)) = L^2([0,1])$, since $\lambda(\{0\}) = \lambda(\{1\}) = 0$. We now set $e_n(x) = e^{2i\pi nx}$, for $n \in \mathbb{Z}$ and $x \in (0,1)$. Then $(e_n)_{n\in\mathbb{Z}}$ is an orthonormal family in $L^2_{\mathbb{C}}((0,1))$. We also know, by Proposition 2.6 on page 107, that $C_c((0,1))$ is dense in $L^2((0,1))$. Now, $C_c((0,1))$ can be identified with a subspace of C_1, the space of continuous functions periodic of period 1 (every element f of $C_c((0,1))$ extends uniquely to a continuous function periodic of period 1 on $\mathbb{R}$), and every element of C_1 is the uniform limit of a sequence of linear combinations of functions e_n extended to $\mathbb{R}$ by 1-periodicity (Example 4 on page 35). We deduce, by comparing norms as in the preceding example, that the family $(e_n)_{n\in\mathbb{Z}}$ is a Hilbert basis of $L^2_{\mathbb{C}}((0,1))$. As before, it follows that
$$\{1,\ \sqrt{2}\cos 2\pi x,\ \sqrt{2}\sin 2\pi x,\ \dots,\ \sqrt{2}\cos 2\pi nx,\ \sqrt{2}\sin 2\pi nx,\ \dots\}$$
is a Hilbert basis of $L^2_{\mathbb{K}}((0,1))$, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.
More generally, if $a, b \in \mathbb{R}$ and $a < b$, the family $(f_n)_{n\in\mathbb{Z}}$ defined by
$$f_n(x) = \frac{1}{\sqrt{b-a}}\, e^{2i\pi nx/(b-a)} \quad \text{for all } x \in (a,b)$$
is a Hilbert basis of $L^2_{\mathbb{C}}((a,b))$. One can also, in an analogous way, obtain a real Hilbert basis of $L^2_{\mathbb{K}}((a,b))$.
3. Consider the space $E = \ell^2(I)$ of Example 5 on page 99. For $j \in I$, we define an element e_j of E by setting $e_j(j) = 1$ and $e_j(i) = 0$ if $i \neq j$. The family $(e_j)_{j\in I}$ is obviously orthonormal. We now show that it is fundamental. To do this, take $x \in E$ and $\varepsilon > 0$. By the definition of the sum $\sum_{i\in I} |x_i|^2$, there exists a finite subset J of I such that

$$\sum_{i\in I\setminus J} |x_i|^2 = \sum_{i\in I} |x_i|^2 - \sum_{i\in J} |x_i|^2 \le \varepsilon^2.$$

But then

$$\Big\| x - \sum_{j\in J} x_j e_j \Big\|^2 = \sum_{i\in I\setminus J} |x_i|^2 \le \varepsilon^2.$$

Thus the family $(e_j)_{j\in I}$ is a Hilbert basis of E.

The main properties of orthonormal families follow from the next proposition, which is elementary.

Proposition 4.2 *Let $\{e_j\}_{j\in J}$ be a finite orthonormal family in E, spanning the vector subspace F. For every $x \in E$, the orthogonal projection $P_F(x)$ of x onto F is given by*

$$P_F(x) = \sum_{j\in J} (x \,|\, e_j) e_j.$$

As a consequence,

$$\|x\|^2 = \Big\| x - \sum_{j\in J} (x \,|\, e_j) e_j \Big\|^2 + \sum_{j\in J} \big|(x \,|\, e_j)\big|^2.$$

Proof. To prove the first statement, it is enough to show that the vector $y = \sum_{j\in J} (x \,|\, e_j) e_j$ satisfies the conditions characterizing $P_F(x)$ (see Proposition 2.3 and the remark on page 107). Now, it is clear that $y \in F$ and that $(x - y \,|\, e_j) = 0$ for all $j \in J$, which implies $x - y \in F^\perp$. The rest of the theorem follows immediately from the Pythagorean Theorem. $\square$

An important, though easy, first consequence is the **Bessel inequality**:

Proposition 4.3 *Let $(e_i)_{i\in I}$ be an orthonormal family in E. For all $x \in E$, we have*

$$\sum_{i\in I} \big|(x \,|\, e_i)\big|^2 \le \|x\|^2.$$

(In particular, the family $\big((x \,|\, e_i)\big)_{i\in I}$ lies in $\ell^2(I)$.)

The next result characterizes the case of equality in the Bessel inequality.

Theorem 4.4 (Bessel–Parseval) *Let $(e_i)_{i\in I}$ be an orthonormal family in E. The following properties are equivalent:*

1. *The family $(e_i)_{i\in I}$ is a Hilbert basis of E.*
2. $\|x\|^2 = \sum_{i\in I} \big|(x \,|\, e_i)\big|^2$ *for all $x \in E$* (**Bessel equality**).
3. $(x \,|\, y) = \sum_{i\in I} (x \,|\, e_i)(e_i \,|\, y)$ for all $x, y \in E$.

Thus, if $(e_i)_{i\in I}$ is a Hilbert basis of E, the map from E to $\ell^2(I)$ defined by $x \mapsto \big((x \mid e_i)\big)_{i\in I}$ is a linear isometry. This isometry is surjective if and only if E is a Hilbert space.

Proof

i. Assume property 1 holds. Then, for all $x \in E$ and all $\varepsilon > 0$, there exists a finite subset J of I such that the distance from x to the span of $\{e_j\}_{j\in J}$ is at most ε. By Proposition 4.2,

$$\sum_{j\in I} \big|(x \mid e_j)\big|^2 \geq \sum_{j\in J} \big|(x \mid e_j)\big|^2 \geq \|x\|^2 - \varepsilon^2.$$

By making ε go to 0 and taking Bessel's inequality into account, we obtain 2.

ii. Conversely, suppose property 2 holds. Then, for all $x \in E$ and all $\varepsilon > 0$, there exists a finite subset J of I such that $\sum_{j\in J}\big|(x \mid e_j)\big|^2 \geq \|x\|^2 - \varepsilon^2$; thus, by Proposition 4.2,

$$\Big\| x - \sum_{j\in J} (x \mid e_j) e_j \Big\| \leq \varepsilon.$$

This shows that the family $(e_i)_{i\in I}$ is fundamental, and so property 1.

iii. The equivalence between 2 and 3 can be derived immediately from the expression of the scalar product in terms of the norm, valid for any scalar product space (see the remark following Corollary 1.2).

iv. If the isometry is surjective, E is isometric to $\ell^2(I)$ and hence complete.

v. Finally, suppose E is a Hilbert space and let $(x_i)_{i\in I}$ be an element of $\ell^2(I)$. Set $a = \sum_{i\in I} |x_i|^2$. There exists then an increasing sequence (J_n) of finite subsets of I such that, for all $n \in \mathbb{N}$, $\sum_{i\in J_n} |x_i|^2 \geq a - 2^{-n}$ (we can assume that I is infinite, since the finite case is elementary). Put $u_n = \sum_{i\in J_n} x_i e_i$. Then, if $n < p$,

$$\|u_p - u_n\|^2 = \sum_{i\in J_p\setminus J_n} |x_i|^2 \leq 2^{-n}.$$

Since E is complete, we deduce that the sequence (u_n) converges to an element x of E. But

$$\sum_{i\in \bigcup_n J_n} |x_i|^2 = a.$$

Hence, for any $i \notin \bigcup_n J_n$, we have $x_i = 0$ and

$$(x \mid e_i) = \lim_{n\to+\infty} (u_n \mid e_i) = 0.$$

If $i \in \bigcup_n J_n$, then $(x \mid e_i) = \lim_{n\to+\infty}(u_n \mid e_i) = x_i$. Thus $(x \mid e_i) = x_i$ for all $i \in I$, which proves the surjectivity of the isometry. $\square$

Remark. More precisely, steps i and ii of the proof show that, if $(e_i)_{i\in I}$ is an orthonormal family in E, the equality

$$\|x\|^2 = \sum_{i\in I} |(x \mid e_i)|^2$$

characterizes those points x that belong to the closure of the span of the family $(e_i)_{i\in I}$.

We shall see that the inverse image of $(x_i)_{i\in I}$ under the isometry $E \to \ell^2(I)$ of the preceding theorem can be considered as the sum $\sum_{i\in I} x_i e_i$ in a sense made precise in the following definition:

A family $(X_i)_{i\in I}$ in a normed vector space E is called **summable** in E if there exists $X \in E$, called the **sum** of the family $(X_i)_{i\in I}$, satisfying the following condition: For any $\varepsilon > 0$, there exists a finite subset J of I such that

$$\Bigl\| X - \sum_{i\in K} X_i \Bigr\| \le \varepsilon \quad \text{for any finite subset } K \subset I \text{ containing } J.$$

In this case we write

$$X = \sum_{i\in I} X_i.$$

It is easy to see that the sum of a summable family is unique. Observe that, in the case $E = \mathbb{K}$, a family $(x_i)_{i\in I}$ is summable in $\mathbb{K}$ if and only if $(x_i)_{i\in I} \in \ell^1_{\mathbb{K}}(I)$, and in this case the definition just given for the sum coincides with the one given in Example 5 on page 99 (see Exercise 5 on page 103). Naturally, if $I = \mathbb{N}$ and if the family $(X_i)_{i\in \mathbb{N}}$ is summable, the series $\sum_{i=0}^{+\infty} X_i$ converges in E, with $\sum_{i\in I} X_i = \sum_{i=0}^{+\infty} X_i$. The converse is false, even for $E = \mathbb{K}$: see Exercise 2 below.

Theorem 4.5 *Let $(e_i)_{i\in I}$ be a Hilbert basis of E. For any element x of E,*

$$x = \sum_{i\in I} (x \mid e_i) e_i.$$

Proof. By Proposition 4.2, we know that, for any finite subset J of I,

$$\Bigl\| x - \sum_{j\in J} (x \mid e_j) e_j \Bigr\|^2 = \|x\|^2 - \sum_{j\in J} |(x \mid e_j)|^2.$$

Now just apply the definitions and property 2 of Theorem 4.4. □

Example. Consider again the situation of Example 1 on page 124: the space $C^{\mathbb{C}}_{2\pi}$ with a Hilbert basis (e_n) defined by $e_n(x) = e^{inx}$. If $f \in C^{\mathbb{C}}_{2\pi}$ and $n \in \mathbb{Z}$, set

$$c_n(f) = (f \mid e_n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx.$$

The sequence $(c_n(f))_{n\in\mathbb{Z}}$ is the sequence of **complex Fourier coefficients** of f. Thus, for all $f \in C_{2\pi}^{\mathbb{C}}$,

$$\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2\,dx = \sum_{n\in\mathbf{Z}} |c_n(f)|^2.$$

At the same time we have

$$f = \sum_{n\in\mathbf{Z}} c_n(f)e_n \tag{$*$}$$

in the sense of summability in the space $C_{2\pi}^{\mathbb{C}}$ with the norm associated with the scalar product. Recall that in general the series of functions $\sum_{n\in\mathbb{Z}} c_n(f)e_n$ does not converge uniformly to f; therefore equality $(*)$ does not hold in general in $C_{2\pi}^{\mathbb{C}}$ with the uniform norm. (It holds when f is of class C^1, for example; see Exercise 15 below). On the other hand, since the scalar product space $C_{2\pi}^{\mathbb{C}}$ is not complete, the isometry from $C_{2\pi}^{\mathbb{C}}$ to $\ell^2(\mathbb{Z})$ defined by $f \mapsto (c_n(f))_{n\in\mathbb{Z}}$ is not surjective; hence not all elements of $\ell^2(\mathbb{Z})$ are sequences of Fourier coefficients of continuous functions.

Complex Fourier coefficients can be defined analogously for functions $f \in L^2((0,1))$, by setting $c_n(f) = \int_0^1 f(x)e^{-2i\pi nx}\,dx$ (see Example 2 on page 124). Bessel's equality remains valid in this case, as does equality $(*)$ in the sense of the norm of $L^2((0,1))$, which, unlike $C_{2\pi}^{\mathbb{C}}$, is complete. Thus the isometry from $L^2((0,1))$ to $\ell^2(\mathbb{Z})$ defined by $f \mapsto (c_n(f))_{n\in\mathbb{Z}}$ is surjective.

The rest of this section is devoted to the problem of existence and construction of Hilbert bases.

Proposition 4.6 (Schmidt orthonormalization process) *Suppose that $N \in \{1,2,3,\dots\} \cup \{+\infty\}$ and let $(f_n)_{0\le n<N}$ be a free family in E. There exists an orthonormal family $(e_n)_{0\le n<N}$ of E such that, for each nonnegative integer $n < N$, the families $(e_p)_{0\le p\le n}$ and $(f_p)_{0\le p\le n}$ span the same vector subspace of E.*

Such a family can be constructed by setting

$$e_0 = \frac{1}{\|f_0\|}f_0$$

and, for $0 \le n < N-1$,

$$x_{n+1} = f_{n+1} - P_n f_{n+1} \quad \textit{and} \quad e_{n+1} = \frac{1}{\|x_{n+1}\|}x_{n+1},$$

where P_n is the orthogonal projection onto the span of the family $(f_p)_{0\le p\le n}$.

Proof. We show that the sequence $(e_n)_{n\in\mathbb{N}}$ defined in the statement satisfies the desired conditions. First, since the family (f_n) is assumed to be free, it is clear that $x_n \neq 0$ for all n, and so that e_n is defined for all n. Let E_n and F_n be the vector subspaces of E spanned by, respectively, $(e_p)_{0\leq p\leq n}$ and $(f_p)_{0\leq p\leq n}$. Trivially, $E_0 = F_0$. Suppose that $E_n = F_n$ for $n < N-1$. Clearly $e_{n+1} \in F_{n+1}$, so $E_{n+1} \subset F_{n+1}$. Moreover $f_{n+1} \in E_{n+1}$, which shows the reverse inclusion. Hence, $E_n = F_n$ for all $0 \leq n < N$. At the same time, for each $n \geq 1$ the vector e_{n+1} is, by construction, orthogonal to F_n and thus to E_n. Therefore the family $(e_n)_{0\leq n<N}$ is orthonormal. □

Remark. The family $(e_n)_{0\leq n<N}$ can be recursively constructed using the following algorithm:

$$x_0 = f_0, \qquad e_0 = x_0/\|x_0\|,$$

$$x_{n+1} = f_{n+1} - \sum_{j=0}^{n} (f_{n+1}|e_j)e_j, \qquad e_{n+1} = x_{n+1}/\|x_{n+1}\|$$

(see Proposition 4.2).

Corollary 4.7 *A scalar product space is separable if and only if it has a countable Hilbert basis.*

Proof. According to Proposition 2.6 on page 10, the condition is sufficient. By the same proposition, separability implies the existence of a free and fundamental family $(f_n)_{n\in\mathbb{N}}$. Applying the Schmidt orthonormalization process to the family (f_n) we obtain a family (e_n) that is a Hilbert basis. □

Two scalar product spaces are called **isometric** if there exists a surjective isometry from one onto the other. Theorem 4.4 has the following consequence:

Corollary 4.8 *An infinite-dimensional Hilbert space is separable if and only if it is isometric to the Hilbert space ℓ^2.*

Exercises

1. Prove that every orthonormal sequence in a Hilbert space converges weakly to 0.

2. *Summable families in normed vector spaces.* Let $(X_i)_{i\in I}$ be a family in a normed vector space E.

   a. Suppose E is finite-dimensional. Show that $(X_i)_{i\in I}$ is summable if and only if $\sum_{i\in I} \|X_i\| < +\infty$.
   Hint. Reduce to the case $E = \mathbb{K}$ and use Exercise 5 on page 103.

   b. Make no assumptions on E, but suppose I is countably infinite.

      i. Prove that, if the family $(X_i)_{i\in I}$ is summable with sum X, we have, for any bijection φ from $\mathbb{N}$ onto I,

$$X = \sum_{n=0}^{+\infty} X_{\varphi(n)}.$$

ii. Suppose, conversely, that the series $\sum_{n=0}^{+\infty} X_{\varphi(n)}$ converges for any bijection $\varphi : \mathbb{N} \to I$. Prove that the family $(X_i)_{i\in I}$ is summable.

Hint. Let φ be a bijection from $\mathbb{N}$ onto I and set

$$X = \sum_{n=0}^{+\infty} X_{\varphi(n)}.$$

Prove that, if $\sum_{i\in I} X_i \neq X$, there exists $\varepsilon > 0$ with the following property: For any integer $n \in \mathbb{N}$, there exists a finite subset A of $\{n, n+1, n+2, \ldots\}$ such that $\bigl\|\sum_{k\in A} X_{\varphi(k)}\bigr\| \geq \varepsilon$. Deduce the existence of a sequence $(A_n)_{n\in\mathbb{N}}$ of pairwise disjoint finite subsets of $\mathbb{N}$ such that $\bigl\|\sum_{k\in A_n} X_{\varphi(k)}\bigr\| \geq \varepsilon$ for every $n \in \mathbb{N}$, then the existence of a bijection $\psi : \mathbb{N} \to I$ such that the series $\sum_{n=0}^{+\infty} X_{\psi(n)}$ does not satisfy the Cauchy criterion and so does not converge.

c. Suppose that E is a Hilbert space, I is arbitrary, and $(X_i)_{i\in I}$ is an orthogonal family. Show that the family $(X_i)_{i\in I}$ is summable if and only if $\sum_{i\in I} \|X_i\|^2 < +\infty$. (You might draw inspiration from the last part of the proof of Theorem 4.4.) Deduce that, in any infinite-dimensional Hilbert space, there exists a summable sequence $(X_n)_{n\in\mathbb{N}}$ such that $\sum_{n\in\mathbb{N}} \|X_n\|$ is infinite. (In fact, the *Dvoretzki–Rogers Theorem* asserts that there is such a sequence in any infinite-dimensional Banach space. The next question presents another simple example of this situation.)

d. Let X be an infinite metric space, and take $E = C_b(X)$, with the uniform norm, denoted $\|\cdot\|$.

i. Show that there exists in X a sequence $\bigl(B(a_n, r_n)\bigr)_{n\in\mathbb{N}}$ of pairwise disjoint nonempty open balls.

ii. Show that, for each integer $n \in \mathbb{N}$, there exists a continuous nonnegative-valued function f_n on X supported within $B(a_n, r_n)$ and having norm $\|f_n\| = 1/(n+1)$.

iii. Show that the sequence $(f_n)_{n\in\mathbb{N}}$ is summable in E and that the series $\sum_{n\in\mathbb{N}} \|f_n\|$ diverges.

3. Let A be a subset of $\mathbb{Z}$ and let E_A be the vector subspace of $L^2([0, 2\pi])$ defined by

$$E_A = \Bigl\{ f \in L^2([0, 2\pi]) : \int_0^{2\pi} f(x) e^{-inx}\, dx = 0 \text{ for all } n \in A \Bigr\}.$$

a. Show that E_A is closed and determine a Hilbert basis of E_A.

b. What is the orthogonal complement of E_A?

c. Write down explicitly the operator of orthogonal projection onto E_A.

4. *Legendre polynomials.* If n is a nonnegative integer, we define a polynomial P_n as

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}\big((x^2-1)^n\big).$$

   a. Show that the family $\big(\sqrt{n+\frac{1}{2}}\,P_n\big)_{n\in\mathbb{N}}$ is a Hilbert basis of the space $L^2([-1,1])$.
   b. Deduce an explicit expression for the orthogonal projection from $L^2([-1,1])$ onto the space $\mathbb{R}_n[X]$ of polynomial functions of degree at most n.

5. *Hermite polynomials.* Consider the Hilbert space $E = L^2(\mu)$, where μ is the positive Radon measure defined on $\mathbb{R}$ by

$$\mu(\varphi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(x) e^{-x^2/2}\,dx \quad \text{for all } \varphi \in C_c(\mathbb{R}).$$

   a. Show that, for every $n \in \mathbb{N}$, there exists a unique polynomial $\tilde{P}_n$ of degree n such that

$$\frac{d^n}{dx^n}\big(e^{-x^2/2}\big) = (-1)^n e^{-x^2/2} \tilde{P}_n(x).$$

   b. For each $n \in \mathbb{N}$, set $P_n = \tilde{P}_n/\sqrt{n!}$. Show that (P_n) is an orthonormal family in E.
   c. i. Take $\varphi \in C_c(\mathbb{R})$. Show that there exists a sequence of polynomials $(p_n)_{n\in\mathbb{N}}$ such that

$$\lim_{n\to+\infty} p_n(x) e^{-x^2/8} = \varphi(x) e^{-x^2/8}$$

      uniformly on $\mathbb{R}$.
      Hint. Use Exercise 8d on page 41 and perform a change of variables.
      ii. Deduce that $(p_n)_{n\in\mathbb{N}}$ converges to φ in E.
   d. Show that the family (P_n) is a Hilbert basis for E.

6. *Chebyshev polynomials.* Let μ be the positive Radon measure on $[-1,1]$ defined by

$$\mu(\varphi) = \int_{-1}^{1} \varphi(x)(1-x^2)^{-1/2}\,dx \quad \text{for all } \varphi \in C([-1,1]).$$

   For $x \in [-1,1]$, set $T_0(x) = \sqrt{1/\pi}$ and

$$T_n(x) = \sqrt{2/\pi}\cos(n \arccos x) \quad \text{for } n \geq 1.$$

   Show that, for every $n \in \mathbb{N}$, the function T_n is the restriction to $[-1,1]$ of a polynomial of degree n and that $(T_n)_{n\in\mathbb{N}}$ is a Hilbert basis for $L^2(\mu)$.

7. *Laguerre polynomials.* Let μ be the positive Radon measure on $\mathbb{R}^+$ defined by

$$\mu(\varphi) = \int_0^{+\infty} \varphi(x)e^{-x}\,dx \quad \text{for all } \varphi \in C_c(\mathbb{R}^+).$$

For each $n \in \mathbb{N}$, set

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(e^{-x}x^n).$$

a. Show that L_n is a polynomial of degree n for every $n \in \mathbb{N}$.

b. i. Compute the scalar product $(X^k \mid L_n)$, for $0 \le k \le n$, where $X^k : x \mapsto x^k$.

ii. Deduce that $(L_n)_{n\in\mathbb{N}}$ is an orthonormal family in the space $E = L^2(\mu)$.

c. Show that, if α is a nonnegative real number,

$$\sum_{n=0}^{+\infty}\left(\int_0^{+\infty} e^{-\alpha x}L_n(x)e^{-x}dx\right)^2 = \frac{1}{2\alpha+1}\,.$$

Deduce that the function $f_\alpha : x \mapsto e^{-\alpha x}$ lies in the closure in E of the vector space spanned by the sequence (L_n).

d. Show that the family $(f_n)_{n\in\mathbb{N}^*}$ is fundamental in $C_0(\mathbb{R}^+)$. (Use the Weierstrass Theorem and a change of variables, or the Stone–Weierstrass Theorem in $\mathbb{R}^+$: see Exercise 7i on page 56.) Deduce that $(L_n)_{n\in\mathbb{N}}$ is a Hilbert basis for E.

8. *Gaussian quadrature.* Let μ be a positive Radon measure on a compact interval $[a, b]$ in $\mathbb{R}$ (where $a < b$). Suppose the support of μ is not finite.

a. Show that there exists a Hilbert basis $(P_n)_{n\in\mathbb{N}}$ of $L^2_{\mathbb{R}}(\mu)$ such that, for every $n \in \mathbb{N}$, P_n is the restriction to $[a, b]$ of a real polynomial of degree n.

b. Show that, for $n \ge 1$, P_n has n distinct roots in (a, b).

Hint. Using the fact that $\int P_n\,d\mu = 0$, show that P_n has at least one root of odd multiplicity in (a, b). Now let $x_1, \dots, x_r$ be the roots of odd multiplicity of P_n in (a, b). By considering the integral $\int P_n(x)(x - x_1)\dots(x - x_r)\,d\mu(x)$, prove that $r = n$.

c. Fix $n \ge 1$ and let $x_1, \dots, x_n$ be the roots of P_n.

i. Show that there exists a unique n-tuple $(A_1, \dots, A_n)$ of real numbers such that, for every $k \in \{0, \dots, n-1\}$,

$$\int x^k\,d\mu(x) = \sum_{i=1}^{n} A_i x_i^k.$$

ii. Show that, for every polynomial P of degree at most $2n-1$,

$$\int P\,d\mu = \sum_{i=1}^{n} A_i P(x_i).$$

Hint. Write $P = Q + RP_n$, where R and Q are polynomials of degree at most $n-1$.

iii. Show that, for every $i \in \{1,\ldots,n\}$,

$$\int \prod_{j\neq i}(x-x_j)^2\,d\mu(x) = A_i \prod_{j\neq i}(x_i-x_j)^2.$$

Deduce that $A_i > 0$.

d. Now make n vary and denote by $x_1^{(n)},\ldots,x_n^{(n)}$ the roots of P_n and by $(A_1^{(n)},\ldots,A_n^{(n)})$ the coefficients determined in the preceding question. Show that, for every continuous function f on $[a,b]$,

$$\int f\,d\mu = \lim_{n\to+\infty} \sum_{i=1}^{n} A_i^{(n)} f(x_i^{(n)}).$$

Hint. Use Proposition 4.3 on page 19.

9. Let D be a dense subset and $(e_i)_{i\in I}$ an orthonormal family in a scalar product space E. Show that there exists a surjection from D onto I. Deduce that any orthonormal family in a separable scalar product space is countable.

10. Let $\mathscr{E}$ be the vector space spanned by the family of functions $(e_r)_{r\in\mathbb{R}}$ from $\mathbb{R}$ to $\mathbb{C}$ defined by $e_r(x) = e^{irx}$.

a. Show that, if f and g are elements of $\mathscr{E}$, the value

$$(f \mid g) = \lim_{T\to+\infty} \frac{1}{2T}\int_{-T}^{T} f(t)\overline{g(t)}\,dt$$

is well defined and that the bilinear form thus defined is a scalar product on $\mathscr{E}$.

b. Show that the family $(e_r)_{r\in\mathbb{R}}$ is a Hilbert basis of $\mathscr{E}$, and that $\mathscr{E}$ is not separable (see Exercise 9 above).

c. Let E be the Hilbert completion of $\mathscr{E}$ (Exercise 8 on page 104). Show that the family $(\hat{e}_r)_{r\in\mathbb{R}}$ (where we use the notation of Exercise 8 on page 104) is a Hilbert basis of E, and deduce that there exists a surjective isometry between E and $\ell^2(\mathbb{R})$.

11. *Hilbert bases in an arbitrary Hilbert space*

a. Let E be a scalar product space.

i. Show that E contains a maximal orthonormal family (that is, an orthonormal family that is not strictly contained in any other). *Hint.* Use *Zorn's Lemma* (which is apparently due to Kuratowski), one of the various equivalent forms of the axiom of choice: *Let $\prec$ be an order relation on a set $\mathscr{A}$, satisfying the following condition: Every subset of $\mathscr{A}$ that is totally ordered by $\prec$ has an upper bound. Then $\mathscr{A}$ has a maximal element.*

ii. Show that if E is a Hilbert space every maximal orthonormal family is a Hilbert basis for E. (Use Corollary 2.5.) Thus, with the axiom of choice, every Hilbert space has a Hilbert basis.

b. Let $(e_i)_{i\in I}$ and $(f_j)_{j\in J}$ be Hilbert bases of a Hilbert space E.

i. For $j \in J$ we write $I_j = \{i \in I : (e_i \mid f_j) \neq 0\}$. Show that all the sets I_j are nonempty and countable and that $I = \bigcup_{j\in J} I_j$.

ii. Deduce that there exists a bijection between I and J. *Hint.* Use Exercise 9 on page 6.

c. Show that two Hilbert spaces are isometric if and only if there is a bijection between their Hilbert bases. In particular, $\ell^2(I)$ and $\ell^2(J)$ are isometric if and only if there exists a bijection between I and J.

12. Let $\varphi \in L^2((0,1))$ be such that $\varphi(t) + \varphi\bigl(t+\frac{1}{2}\bigr) = 0$ for every $t \in \bigl(0, \frac{1}{2}\bigr)$. Extend φ to a function periodic of period 1 on $\mathbb{R}$ (also denoted φ). Then set $\varphi_0 = 1$ and, for every integer $n \geq 1$, set $\varphi_n(t) = \varphi(2^{n-1}t)$. Show that $(\varphi_n)_{n\in\mathbb{N}}$ is an orthogonal family in $L^2((0,1))$.

13. *Haar functions.* Consider the family of functions $(H_p)_{p\in\mathbb{N}}$ defined on $[0,1]$ by $H_0 = 1$ and, for $n \in \mathbb{N}$ and $1 \leq k \leq 2^n$,

$$H_{2^n+k-1}(x) = \begin{cases} \sqrt{2^n} & \text{if } x \in \bigl((2k-2)2^{-n-1},\, (2k-1)2^{-n-1}\bigr), \\ -\sqrt{2^n} & \text{if } x \in \bigl((2k-1)2^{-n-1},\, 2k \times 2^{-n-1}\bigr), \\ 0 & \text{otherwise.} \end{cases}$$

a. Show that $(H_p)_{p\in\mathbb{N}}$ is an orthonormal family in $L^2([0,1])$.

b. Let $f \in L^2([0,1])$ be such that $\int_0^1 f H_p\, dx = 0$ for every $p \in \mathbb{N}$. Set $F(y) = \int_0^y f(x)\, dx$.

i. Show that, for every $n \in \mathbb{N}$ and for every integer k such that $1 \leq k \leq 2^n$,

$$-F\left(\frac{2k-2}{2^{n+1}}\right) + 2F\left(\frac{2k-1}{2^{n+1}}\right) - F\left(\frac{2k}{2^{n+1}}\right) = 0.$$

ii. Deduce that $F = 0$. (Note that F is continuous.)

iii. Deduce that $f = 0$, then that $(H_p)_{p\in\mathbb{N}}$ is a Hilbert basis of $L^2([0,1])$.

In the sequel we will write, for each integrable function f on $[0,1]$ and each $p \in \mathbb{N}$,

$$\hat{f}_p = \int_0^1 f(x) H_p(x)\,dx, \qquad s_p(f)(x) = \sum_{q=0}^{p} \hat{f}_q H_q(x).$$

c. For $p \in \mathbb{N}$, we denote by $\mathscr{I}_p$ the set of maximal open intervals on which the functions H_q, with $q \leq p$, are constant: If $p = 2^n + k - 1$ with $n \in \mathbb{N}$ and $1 \leq k \leq 2^n$,

$$\mathscr{I}_p = \left\{\left(\frac{j-1}{2^{n+1}}, \frac{j}{2^{n+1}}\right)\right\}_{1 \leq j \leq 2k} \cup \left\{\left(\frac{2(j-1)}{2^{n+1}}, \frac{2j}{2^{n+1}}\right)\right\}_{k+1 \leq j \leq 2^n}.$$

(Note that $\mathscr{I}_p$ has $p+1$ elements.) Moreover, let F_p be the set of functions defined on $(0,1)$, constant on each interval $I \in \mathscr{I}_p$, and such that

$$f(x) = \tfrac{1}{2}\big(f(x_+) + f(x_-)\big) \quad \text{for all } x \in (0,1),$$

where $f(x_+)$ and $f(x_-)$ are the right and left limits of f at x. Show that $(H_q)_{q \leq p}$ is an orthonormal basis of F_p.

d. Suppose $f \in L^1([0,1])$.

i. Take $p \in \mathbb{N}$. Denote by f^* the element of F_p whose constant value on each interval $I \in \mathscr{I}_p$ of length $l(I)$ is

$$\frac{1}{l(I)} \int_I f(x)\,dx.$$

Show that, for every nonnegative integer $q \leq p$, we have $\hat{f}_q = \widehat{f^*_q}$. Deduce that $s_p(f) = s_p(f^*)$.

ii. Deduce that, for every integer $p \in \mathbb{N}$ and every interval $I \in \mathscr{I}_p$,

$$s_p(f)(t) = \frac{1}{l(I)} \int_I f(x)\,dx \quad \text{for all } t \in I.$$

e. Let $f \in C^{\mathbb{R}}([0,1])$.

i. Take $p \in \mathbb{N}$. Show that, for every $I \in \mathscr{I}_p$, there exists a point $x_I \in I$ such that

$$s_p(f)(t) = f(x_I) \quad \text{for all } t \in I.$$

ii. Deduce that, for every $p \in \mathbb{N}$,

$$\max_{x \in [0,1]} \big|s_p(f)(x) - f(x)\big| \leq \sup\big\{\big|f(x) - f(y)\big| : x, y \in [0,1],\ |x-y| \leq 2/p\big\}.$$

iii. Deduce that the series $\sum \hat{f}_q H_q$ converges uniformly to f in $[0,1]$.

14. *Rademacher functions.* For every integer $n \geq 1$ we define a function r_n on the interval $[0,1]$ by

$$r_n(x) = \begin{cases} 1 & \text{if } x \in ((k-1)2^{-n},\, k2^{-n}) \text{ with } 0 \leq k \leq 2^n,\ k \text{ odd}, \\ -1 & \text{if } x \in ((k-1)2^{-n},\, k2^{-n}) \text{ with } 2 \leq k \leq 2^n,\ k \text{ even}, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $r_n = \left(1/\sqrt{2^{n-1}}\right) \sum_{p=2^{n-1}}^{2^n-1} H_p$, where the H_p are the Haar functions defined in Exercise 13.

a. Show that $(r_n)_{n\geq 1}$ is an orthonormal family in $L^2([0,1])$. Deduce that, if $(a_n) \in \ell^2$, the series $\sum_{n\geq 1} a_n r_n$ converges in $L^2([0,1])$.

b. **i.** Prove that, if $\beta_1, \ldots, \beta_n$ are nonnegative integers whose sum is p, we have

$$\frac{(2p)!}{p!} \frac{(\beta_1)! \ldots (\beta_n)!}{(2\beta_1)! \ldots (2\beta_n)!} \leq p^p.$$

ii. Let $\alpha_1, \ldots, \alpha_n$ be nonnegative integers and

$$I = \int_0^1 r_1^{\alpha_1}(x) \ldots r_n^{\alpha_n}(x)\, dx.$$

Show that $I = 1$ if all the α_j are even and that $I = 0$ in any other case.

Hint. Observe that, for every $j \geq 1$, we have $r_j^2 = 1$ almost everywhere; this allows us to reduce to the case where all the α_j equal 0 or 1.

iii. Let $a_1, \ldots, a_n$ be real numbers and set $s_n = \sum_{j=1}^n a_j r_j$. Show that, for every $p \in \mathbb{N}$,

$$\int_0^1 s_n(x)^{2p}\, dx \leq p^p \left(\sum_{j=1}^n a_j^2 \right)^p.$$

c. Take $(a_n) \in \ell^2$ and let f be the sum in $L^2([0,1])$ of the series $\sum_{n\geq 1} a_n r_n$. Show that $f \in L^p([0,1])$ for every real $p \geq 1$.

d. Let F be the closure in $L^2([0,1])$ of the vector space spanned by the sequence $(r_n)_{n\geq 1}$.

i. Let G be the vector space spanned by the functions $f_\varepsilon : x \mapsto x^{-\varepsilon}$, where $\varepsilon < \frac{1}{2}$. Show that the projection $P_{F^\perp}$ is injective on G.

Hint. Use part c above and the equality

$$G \cap \left(\bigcap_{p\geq 1} L^p([0,1]) \right) = \{0\}.$$

ii. Deduce that $F^\perp$ has infinite dimension.

iii. Show that, for any finite family $(r'_n)_{n\leq N}$ in $L^2([0,1])$, the family $(r_n)_{n\geq 1} \cup (r'_n)_{n\leq N}$ is not fundamental in $L^2([0,1])$.

15. Let f be a function of class C^1 from $[0,1]$ to $\mathbb{C}$ such that $f(0) = f(1)$. For $n \in \mathbb{Z}$, set $c_n(f) = \int_0^1 f(x)e^{-2i\pi nx}\,dx$. Show that the series of functions $\sum_{n=-\infty}^{+\infty} c_n(f)e^{2i\pi nx}$ converges uniformly on $[0,1]$. Then show that, for every $x \in [0,1]$,

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n(f)e^{2i\pi nx}.$$

Hint. Show that $c_n(f') = 2i\pi nc_n(f)$ and, using Bessel's equality for f', deduce that

$$\sum_{n=-\infty}^{+\infty} \left|c_n(f)\right| < +\infty.$$

16. a. Let f be a function of class C^1 from $[0,1]$ to $\mathbb{C}$ such that $f(0) = f(1)$. Show that

$$\int_0^1 \left|f(x)\right|^2 dx - \left|\int_0^1 f(x)\,dx\right|^2 \leq \frac{1}{4\pi^2}\int_0^1 \left|f'(x)\right|^2 dx$$

and that equality takes place if and only if f is of the form $f(x) = \lambda + \mu e^{2i\pi x} + \nu e^{-2i\pi x}$, with $\lambda, \mu, \nu \in \mathbb{C}$.
Hint. Use Bessel's equality, considering the Hilbert basis of $L^2((0,1))$ defined in Example 2 on page 124.

b. Let f be a function of class C^1 from $[0,1]$ to $\mathbb{C}$. Show that

$$\int_0^1 \left|f(x)\right|^2 dx - \left|\int_0^1 f(x)\,dx\right|^2 \leq \frac{1}{\pi^2}\int_0^1 \left|f'(x)\right|^2 dx$$

and that equality takes place if and only if f is of the form $f(x) = \lambda + \mu\cos\pi x$, with $\lambda, \mu \in \mathbb{C}$.
Hint. Argue as in the preceding question, considering the even function from $[-1,1]$ to $\mathbb{C}$ that extends f.

c. Let f be a function of class C^2 from $[0,1]$ to $\mathbb{C}$ such that $f(0) = f(1) = 0$. Show that

$$\int_0^1 \left|f'(x)\right|^2 dx \leq \frac{1}{\pi^2}\int_0^1 \left|f''(x)\right|^2 dx$$

and that equality takes place if and only if f is of the form $f(x) = \lambda\sin\pi x$ with $\lambda \in \mathbb{C}$.

d. *Wirtinger's inequality.* Let f be a function of class C^1 from $[0,1]$ to $\mathbb{C}$ such that $f(0) = f(1) = 0$. Show that

$$\int_0^1 \left|f(x)\right|^2 dx \leq \frac{1}{\pi^2}\int_0^1 \left|f'(x)\right|^2 dx$$

and that equality takes place if and only if f is of the form $f(x) = \lambda \sin \pi x$, with $\lambda \in \mathbb{C}$.

Hint. Extend f into an odd function.

17. *Biorthogonal systems.* Let E be a Hilbert space. We say that two sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ in E form a *biorthogonal system* in E if, for every $i, j \in \mathbb{N}$,

$$(f_i \mid g_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Suppose that $(e_n)_{n\in\mathbb{N}}$ is a Hilbert basis of E and that $(f_n)_{n\in\mathbb{N}}$ is a sequence in E such that, for every finite sequence $(a_n)_{n\leq N}$ in $\mathbb{K}$,

$$\Bigl\| \sum_{n=0}^{N} a_n(e_n - f_n) \Bigr\|^2 \leq \theta^2 \sum_{n=0}^{N} |a_n|^2,$$

where θ is a real constant such that $0 \leq \theta < 1$.

a. Show that, for every $f \in E$, the series $\sum_{n=0}^{+\infty} (f \mid e_n)(e_n - f_n)$ converges in E. Denote its limit by Kf.

b. Show that the map K thus defined is a continuous linear operator on E, of norm at most θ.

c. Set $T = I - K$. Show that $Te_n = f_n$ for each $n \in \mathbb{N}$, and that T has a continuous inverse, which we denote by U.

d. For each $n \in \mathbb{N}$, set $g_n = U^* e_n$. Show that the sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ form a biorthogonal system in E.

e. Show that, for every $f \in E$,

$$f = \sum_{n\in\mathbb{N}} (f \mid g_n) f_n = \sum_{n\in\mathbb{N}} (f \mid f_n) g_n.$$

Deduce that the two families $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ are fundamental in E.

f. Show that, for every $f \in E$,

$$(1-\theta)\|f\| \leq \Bigl(\sum_{n\in\mathbb{N}} \bigl|(f \mid f_n)\bigr|^2 \Bigr)^{1/2} \leq (1+\theta)\|f\|,$$

$$(1+\theta)^{-1}\|f\| \leq \Bigl(\sum_{n\in\mathbb{N}} \bigl|(f \mid g_n)\bigr|^2 \Bigr)^{1/2} \leq (1-\theta)^{-1}\|f\|.$$

18. Suppose that E is a separable Hilbert space and let (e_n) be a Hilbert basis of E. For every pair (x, y) of points in the closed unit ball B of E, set

$$d(x, y) = \sum_{n=0}^{\infty} \frac{\bigl|(x-y \mid e_n)\bigr|}{2^n}.$$

a. Show that d is a metric on B and that a sequence (x_n) of points of B converges in (B,d) to a point $x \in B$ if and only if it converges weakly to x.

b. Show that the metric space (B,d) is compact.

19. *Gram matrices and Gram determinants.* Let E be a scalar product space over $\mathbb{R}$. If $x_1, \ldots, x_p$ are elements of E, the *Gram matrix* of $(x_1, \ldots, x_p)$ is by definition the $p \times p$ matrix $G(x_1, \ldots, x_p)$ whose (i,j) entry is $a_{i,j} = (x_i \mid x_j)$. The determinant of this matrix is called the *Gram determinant* of the p-tuple $(x_1, \ldots, x_p)$.

a. Show that the Gram determinant of a linearly dependent family of vectors in E vanishes.

b. Suppose that the family $(x_1, \ldots, x_p)$ is free. Let $\{e_1, \ldots, e_p\}$ be an orthonormal basis of the vector space spanned by $\{x_1, \ldots, x_p\}$. Let $M = (m_{i,j})$ be the matrix of change of basis $\bigl($thus $x_j = \sum_{i=1}^{p} m_{i,j} e_i$ if $1 \le j \le p\bigr)$. Show that $G(x_1, \ldots, x_p) = M^T M$, where M^T denotes the transpose of M. Deduce that $\det G(x_1, \ldots, x_p) > 0$.

c. Let $\{x_1, \ldots, x_p\}$ be a free family in E, spanning the subspace F. Show that, for every $x \in E$,

$$d^2(x,F) = \frac{\det G(x, x_1, \ldots, x_p)}{\det G(x_1, \ldots, x_p)}.$$

Hint. Let y be the orthogonal projection of x onto F. In the calculation of $\det G(x, x_1, \ldots, x_p)$, replace x by $(x-y)+y$ and use the fact that the determinant depends linearly on the first column.

d. **i.** Let $a_1, \ldots, a_p$ be positive reals and A the $p \times p$ matrix whose (i,j) entry is $a_{i,j} = 1/(a_i + a_j)$. Show that

$$\det A = 2^{-p} \prod_{j=1}^{p} \frac{1}{a_j} \prod_{1 \le j < k \le p} \left(\frac{a_j - a_k}{a_j + a_k} \right)^2.$$

Hint. Work by induction.

ii. Suppose that $E = L^2((0,1))$. For every nonnegative real number r, define an element f_r of E by $f_r(x) = x^r$. Let $r_1, \ldots, r_p$ be pairwise distinct nonnegative reals and let F be the vector space spanned by the functions $f_{r_1}, \ldots, f_{r_p}$. Show that, for every integer $n \in \mathbb{N}$,

$$d^2(f_n, F) = \frac{1}{2n+1} \prod_{j=1}^{p} \left(\frac{n - r_j}{n + r_j + 1} \right)^2.$$

20. *Müntz's Theorem.* Let $(r_p)_{p \in \mathbb{N}}$ be a strictly increasing sequence of nonnegative reals. For any real number $r \ge 0$, denote by f_r the function defined on $[0,1]$ by $f_r(x) = x^r$.

a. Consider the space $E = L^2((0,1))$ with its Hilbert space structure.

i. Show that the family $(f_{r_p})_{p\in\mathbb{N}}$ is fundamental in E if and only if, for every integer $n \in \mathbb{N}$,

$$\lim_{p\to+\infty} d(f_n, F_p) = 0,$$

where F_p is the vector space spanned by $(f_{r_j})_{0\le j\le p}$.
Hint. Start by showing that the family $(f_n)_{n\in\mathbb{N}}$ is fundamental.

ii. Show that the family $(f_{r_p})_{p\in\mathbb{N}}$ is fundamental in E if and only if

$$\sum_{p\ge 1} 1/r_p = +\infty.$$

Hint. Calculate $\log\bigl(d(f_n, F_p)\bigr)$ using Exercise 19d-ii.

b. We now place ourselves in the space $H = C^{\mathbb{R}}([0,1])$, considered with the uniform norm.

i. Suppose the family $(f_{r_p})_{p\in\mathbb{N}}$ is fundamental in H. Show that $\sum_{p\ge1} 1/r_p = +\infty$.

ii. Conversely, suppose that $\sum_{p\ge1} 1/r_p = +\infty$, $r_0 = 0$, and $r_1 \ge 1$.

A. Show that $\sum_{p\ge2} 1/(r_p - 1) = +\infty$. Deduce that the space of C^1 functions on $[0,1]$ vanishing at 0 is contained in the closure of the vector subspace of H spanned by the family $(f_{r_p})_{p\in\mathbb{N}^*}$.
Hint. Let f be a C^1 function vanishing at 0. Approximate f' in the space $L^2((0,1))$ by linear combinations of functions f_{r_p-1}.

B. Deduce that the family $(f_{r_p})_{p\in\mathbb{N}}$ is fundamental in H.

21. *Hilbert–Schmidt operators.* Let E be an infinite-dimensional separable Hilbert space.

a. i. Let $(e_n)_{n\in\mathbb{N}}$ and $(f_p)_{p\in\mathbb{N}}$ be Hilbert bases for E. Show that, for $T \in L(E)$,

$$\sum_{n=0}^{+\infty} \|Te_n\|^2 = \sum_{p=0}^{+\infty} \|T^* f_p\|^2 \le +\infty.$$

Deduce that

$$\sum_{n=0}^{+\infty} \|Te_n\|^2 = \sum_{p=0}^{+\infty} \|Tf_p\|^2.$$

We fix from now on a Hilbert basis $(e_n)_{n\in\mathbb{N}}$ for E and we denote by $\mathscr{H}(E)$ the vector space consisting of $T \in L(E)$ such that the expression $\|T\|_2 = \bigl(\sum_{n=0}^{+\infty} \|Te_n\|^2\bigr)^{1/2}$ is finite. Such a T is called *Hilbert–Schmidt operator* on E.

ii. Show that $\mathscr{H}(E) \neq L(E)$ and that $\|T\| \leq \|T\|_2$ for every $T \in \mathscr{H}(E)$. Show that $\|\cdot\|_2$ is a norm on $\mathscr{H}(E)$, with respect to which $\mathscr{H}(E)$ is a Hilbert space. This is called the *Hilbert–Schmidt norm*.

Show that any element T in $L(E)$ of finite rank (that is, such that $\operatorname{im} T$ is finite-dimensional) is Hilbert–Schmidt.

Hint. Consider a Hilbert basis of E that is the union of a basis of $\ker T$ and a basis of $(\ker T)^{\perp}$.

iii. Take $T \in \mathscr{H}(E)$. For $n \geq 0$, denote by P_n the operator of orthogonal projection onto the span of $\{e_j : 0 \leq j \leq n\}$. Show that, for every positive integer n, the composition TP_n belongs to $\mathscr{H}(E)$, and that $\lim_{n\to+\infty} \|T - TP_n\|_2 = 0$. Deduce that the set of operators of finite rank is dense in $\mathscr{H}(E)$.

b. Suppose that $E = L^2(m)$, where m is a σ-finite measure on a measure space $(\Omega, \mathscr{F})$ (such that $L^2(m)$ is separable). Choose a Hilbert basis $(e_n)_{n\in\mathbb{N}}$ for E.

i. Show that the family $(e_{n,p})_{n,p\in\mathbb{N}}$ defined by $e_{n,p} = e_n \otimes \bar{e}_p$ is a Hilbert basis for $L^2(m\times m)$. (Recall the notation $(e_n\otimes\bar{e}_p)(x, y) = e_n(x)\overline{e_p(y)}$.)

Hint. See Exercise 7 on page 110.

ii. Consider $K \in L^2(m \times m)$, and let T_K be the operator from E to E defined by

$$T_K f(x) = \int K(x, y) f(y)\, dm(y) \quad \text{for all } f \in E.$$

For $(n, p) \in \mathbb{N}^2$, set

$$k_{n,p} = (K \mid e_{n,p}) = (T_K e_p \mid e_n),$$

where we use the same notation for the scalar products in $L^2(m)$ and $L^2(m \times m)$. Show that

$$\|T_K\|_2 = \Bigl(\sum_{n,p\in\mathbb{N}} |k_{n,p}|^2\Bigr)^{1/2} = \|K\|_{L^2(m\times m)},$$

and so that $T_K \in \mathscr{H}(E)$.

iii. Conversely, take $T \in \mathscr{H}(E)$. For $n, p \in \mathbb{N}$ we write $k_{n,p} = (Te_p \mid e_n)$.

A. Show that $\sum_{n,p\in\mathbb{N}} |k_{n,p}|^2 < +\infty$.

B. Let K be the element of $L^2(m \times m)$ defined by

$$K = \sum_{n,p\in\mathbb{N}} k_{n,p} e_{n,p}.$$

Show that $T = T_K$. Hence, the map $K \mapsto T_K$ is a surjective isometry from $L^2(m \times m)$ onto $\mathscr{H}(L^2(m))$.

4

L^p Spaces

1 Definitions and General Properties

We first establish the notation and definitions that we will use throughout this chapter. The most basic results are recalled without proof; the reader can consult, for example, the first part of Chapter 3 of W. Rudin's *Real and Complex Analysis* (McGraw-Hill).

We consider a **measure space** $(X, \mathscr{F})$ — that is, a pair consisting of a set X and a σ-algebra $\mathscr{F}$ — and a measure m on $\mathscr{F}$. For every real p in the range $1 \le p < \infty$, we define $\mathscr{L}^p_{\mathbb{K}}(m)$ as the space of $\mathscr{F}$-measurable functions f from X to $\mathbb{K}$ such that $\int |f|^p \, dm < +\infty$. We denote by $\mathscr{L}^\infty_{\mathbb{K}}(m)$ the space of $\mathscr{F}$-measurable functions f from X to $\mathbb{K}$ for which there exists a nonnegative real number M (depending on f) such that $\left|f(x)\right| \le M$ m-almost everywhere. We can leave $\mathbb{K}$ and/or m out of the notation when there is no danger of confusion.

By extension, a function f with values in $\mathbb{K}$ and defined m-almost everywhere on X is said to belong to $\mathscr{L}^p_{\mathbb{K}}(m)$ if it equals m-almost everywhere some function of $\mathscr{L}^p_{\mathbb{K}}(m)$ in the original sense.

In the study of these spaces $\mathscr{L}^p$, an essential role is played by the Hölder inequality, a generalization of the Schwarz inequality (which corresponds to the case $p = p' = 2$).

Theorem 1.1 (Hölder inequality) *Suppose $p, p' \in (1, \infty)$ satisfy*

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

(We say that p and p' are **conjugate exponents**.) *If* $f \in \mathscr{L}^p$ *and* $g \in \mathscr{L}^{p'}$, *the product* fg *lies in* $\mathscr{L}^1$ *and*

$$\int |fg|\,dm \le \left(\int |f|^p\,dm\right)^{1/p} \left(\int |g|^{p'}\,dm\right)^{1/p'}.$$

We define the vector space $L^p_{\mathbb{K}}(m)$ as the quotient vector space of $\mathscr{L}^p_{\mathbb{K}}(m)$ by the equivalence relation $\mathscr{R}$ of equality m-almost everywhere (in other words, we identify in L^p functions that coincide m-almost everywhere). The vector space $L^p_{\mathbb{R}}(m)$ is a lattice. Except when explicitly stated otherwise, our notation will not distinguish between an element of $L^p(m)$ and its representatives in $\mathscr{L}^p(m)$.

If $f \in L^p_{\mathbb{K}}(m)$ with $1 \le p < \infty$, we define

$$\|f\|_p = \left(\int |f|^p\,dm\right)^{1/p};$$

if $f \in L^\infty_{\mathbb{K}}(m)$, we set

$$\|f\|_\infty = \min\bigl\{M \ge 0 : |f| \le M \ m\text{-almost everywhere}\bigr\}.$$

Obviously, these expressions do not depend on the representative chosen for f. One can show that, for $1 \le p \le \infty$, the map $\|\cdot\|_p$ thus defined is a norm on $L^p_{\mathbb{K}}(m)$.

By convention, 1 and ∞ are conjugate exponents. The Hölder inequality can be rephrased as follows:

Proposition 1.2 *Let p and p' be conjugate exponents with $1 \le p, p' \le \infty$. For every $f \in L^p_{\mathbb{K}}(m)$ and $g \in L^{p'}_{\mathbb{K}}(m)$ we have $fg \in L^1_{\mathbb{K}}(m)$ and*

$$\|fg\|_1 \le \|f\|_p\,\|g\|_{p'}.$$

Example. In the remainder of this chapter, we will say simply that "m is a **Radon measure**" to mean that X is a locally compact and separable metric space, $\mathscr{F}$ is the Borel σ-algebra on X, and m is a positive Radon measure on X, considered as a Borel measure. In this situation, for every $f \in C_b(X)$,

$$\|f\|_\infty = \sup\bigl\{|f(x)| : x \in \operatorname{Supp} m\bigr\}.$$

Suppose moreover that the support of m equals X. Then $\|f\|_\infty = \|f\|$ for every function $f \in C_b(X)$, where $\|f\|$, as usual, is the uniform norm of f on X. In other words, the map that associates to an element f of $C_b(X)$ (with the uniform norm) its class modulo $\mathscr{R}$ is an isometry (and in particular an injection) from $C_b(X)$ to $L^\infty(m)$ (with the norm $\|\cdot\|_\infty$). If f is a Borel function, there exists a greatest open set O of X such that $f(x) = 0$ for m-almost every x of O (to see this, one might reason as in the proof of Proposition 3.1 on page 68). The complement of O is called the **essential support** of f. If f is continuous, we see that the essential support of f is exactly the support of f (thanks to the assumption $\operatorname{Supp} m = X$). Moreover, by definition, two Borel functions that coincide almost everywhere have the same essential support. Hence we can define without ambiguity the essential support of a class modulo $\mathscr{R}$ as the essential support of any of its representatives. In the sequel, if f is a class of functions modulo $\mathscr{R}$, we will refer to the essential support of f as simply the **support** of f, and we will denote it by $\operatorname{Supp} f$ as well.

One fundamental property of the L^p spaces is completeness:

Theorem 1.3 (Riesz–Fischer) *If $1 \le p \le \infty$, the space $L^p_{\mathbb{K}}(m)$ with the norm $\|\cdot\|_p$ is a Banach space.*

Now suppose I is a set, $\mathscr{F} = \mathscr{P}(I)$ is the discrete σ-algebra on I, and m is the count measure on I (Example 5 on page 99). Then the space $L^p_{\mathbb{K}}(m)$ (with $X = I$) is denoted by $\ell^p(I)$, or more simply by ℓ^p if $I = \mathbb{N}$ (compare Exercises 7 on page 11 and 8 on page 12). In this case,

$$1 \le p \le q \le \infty \Longrightarrow \ell^p_{\mathbb{K}}(I) \subset \ell^q_{\mathbb{K}}(I)$$

and $\|x\|_q \le \|x\|_p$ for every $x \in \ell^p_{\mathbb{K}}(I)$.

By contrast, when m has finite mass ($m(X) < \infty$), the inclusions go in the opposite direction:

$$1 \le p \le q \le \infty \Longrightarrow L^q_{\mathbb{K}}(m) \subset L^p_{\mathbb{K}}(m)$$

and, for every $f \in L^q_{\mathbb{K}}(m)$,

$$\|f\|_p \le \|f\|_q \big(m(X)\big)^{(q-p)/qp},$$

as can be checked using the Hölder inequality.

More generally, we have the following interpolation result:

Proposition 1.4 *If $f \in L^1 \cap L^\infty$, then $f \in L^p$ for every $p \in (1, \infty)$, and*

$$\|f\|_p \le \|f\|_1^{1/p} \, \|f\|_\infty^{1-1/p}.$$

In addition, if $1 \le p < \infty$, $L^1 \cap L^\infty$ is dense in L^p.

Proof. If $f \in L^\infty$ and $1 < p < \infty$, we clearly have $|f|^p \le |f| \, \|f\|_\infty^{p-1}$ m-almost everywhere, which proves the first assertion of the proposition.

Now suppose that $1 \le p < \infty$ and that $f \in L^p$. Since $|f|^p$ is a positive integrable function, there exists an increasing sequence $(\varphi_n)_{n\in\mathbb{N}}$ of positive, integrable, piecewise constant functions that converges almost everywhere to $|f|^p$. Set

$$\alpha(x) = \begin{cases} f(x)/|f(x)| & \text{if } f(x) \ne 0, \\ 0 & \text{if } f(x) = 0. \end{cases}$$

Then the sequence $(\alpha\varphi_n^{1/p})$ is a sequence in $L^1 \cap L^\infty$ that converges almost everywhere to f, being bounded above in absolute value by $|f|$. By the Dominated Convergence Theorem, this sequence converges to f in L^p. □

Remark: denseness of piecewise constant functions in L^p. The preceding proof shows also that, if $p \in [1, +\infty)$, every positive element of L^p is the limit in L^p of an increasing sequence of positive, integrable, piecewise constant functions. By taking linear combinations, we deduce that integrable piecewise constant functions are dense in L^p for $p \in [1, +\infty)$. Note that this is false if $p = \infty$ and if m has infinite mass (see Exercise 8 below). Nonetheless, one sees easily that every positive element f of L^∞ is the limit in L^∞ of an increasing sequence of (not necessarily integrable) positive piecewise constant functions. For example, one can take the sequence $(f_n)_{n\in\mathbb{N}}$ defined by

$$f_n = M \sum_{k=0}^{2^n-1} k\,2^{-n}\,1_{\{Mk2^{-n} < f \le M(k+1)2^{-n}\}},$$

with $M = \|f\|_\infty$. It follows that the set of piecewise constant functions is dense in L^∞.

We now study other denseness results. We start with a convenient elementary lemma.

Lemma 1.5 *For each nonnegative real a, define a map $\Pi_a : \mathbb{K} \to \mathbb{K}$ by setting $\Pi_0(x) = 0$ and*

$$\Pi_a(x) = \frac{ax}{\max(a, |x|)} \quad \text{if } a > 0.$$

Then, for every $x \in \mathbb{K}$, we have $|\Pi_a(x)| \le \min(a, |x|)$ and, if $|x| \le a$, then $\Pi_a(x) = x$. Moreover,

$$|\Pi_a(x) - \Pi_a(y)| \le |x - y| \quad \text{for all } x, y \in \mathbb{K}.$$

Proof. It is clear that Π_a is exactly the projection map from the canonical euclidean space $\mathbb{R}$ (or the canonical hermitian space $\mathbb{C}$, as the case may be) onto $\bar{B}(0, a)$. The claims made are then obvious; the last of them can be seen as a particular case of Proposition 2.2 on page 106. □

The following theorem generalizes Proposition 2.6 on page 107, which represents the case $p = 2$.

Theorem 1.6 *If m is a Radon measure, the space $C_c(X)$ is dense in $L^p(m)$ for $1 \le p < +\infty$.*

Proof. The case $p = 1$ was proved in Chapter 2, Proposition 3.5 on page 70. Suppose $1 < p < \infty$. By Proposition 1.4, it suffices to approximate $f \in L^1 \cap L^\infty$ in the sense of the $\|\cdot\|_p$ norm. Thus, fix $f \in L^1 \cap L^\infty$ and let (φ_n) be a sequence in $C_c(X)$ that converges to f in L^1. Set $\psi_n = \Pi_{\|f\|_\infty}(\varphi_n)$, using the notation of Lemma 1.5. Then $\psi_n \in C_c(X)$ and

$$|f - \psi_n|^p \le |f - \psi_n| \big(2\|f\|_\infty\big)^{p-1}.$$

Now $f = \Pi_{\|f\|_\infty}(f)$ and so, since by Lemma 1.5 the maps Π_a are contracting,

$$|f - \psi_n|^p \le |f - \varphi_n| \big(2\|f\|_\infty\big)^{p-1},$$

which proves the result. □

Remark. If m is a Radon measure of support X, the closure of $C_c(X)$ in $L^\infty(m)$ is $C_0(X)$ (which is distinct from $L^\infty(m)$ if X is infinite).

Corollary 1.7 *If m is a Radon measure, the space $L^p(m)$ is separable for $1 \le p < \infty$.*

Proof. Let (K_n) be a sequence of compact sets exhausting X. Since

$$C_c(X) = \bigcup_{n\in\mathbb{N}} C_{K_n}(X),$$

it suffices, by the preceding theorem, to show that each $C_{K_n}(X)$ is separable with respect to the $\|\cdot\|_p$ norm. But $C_{K_n}(X)$ is separable with respect to the uniform norm $\|\cdot\|$, and $\|f\|_p \le \|f\|\, m(K_n)^{1/p}$ for every $f \in C_{K_n}(X)$. This proves the result. □

Remark. The assumption that X is separable is essential in Corollary 1.7. For example, if I is an uncountable set, the space $\ell^p(I)$ is not separable, by Exercise 8 on page 12.

Note also that the space L^∞ is not separable in general; see Exercise 10 below.

Exercises

We consider in these exercises a measure m on a measure space $(X, \mathscr{F})$.

1. *Spaces L^p for $0 < p < 1$.* Take $p \in (0,1)$. Define the space L^p as the set of equivalence classes (with respect to equality m-almost everywhere) of $\mathscr{F}$-measurable functions f from X to $\mathbb{K}$ for which the expression

$$|f|_p = \int |f|^p \, dm$$

is finite.

a. Show that L^p is a vector space and that the formula $d_p(f,g) = |f-g|_p$ defines on L^p a metric that makes L^p complete.

b. Suppose that X is an open set in $\mathbb{R}^d$, for $d \geq 1$, and that m is the restriction to X of Lebesgue measure on $\mathbb{R}^d$.

i. Show that bounded Borel functions with compact support are dense in (L^p, d_p).

ii. Let f be a bounded Borel function on X with compact support, and suppose that $r > 0$. Show that f lies in the closed convex hull $\bar{c}\big(B(0,r)\big)$ of the ball $B(0,r)$ of L^p (see Exercise 9 on page 18).
Hint. Let K be a parallelepiped in $\mathbb{R}^d$ containing the support of f. Write f in the form

$$f = \frac{1}{n}\sum_{i=1}^{n} n f \mathbf{1}_{K_i},$$

where $(K_i)_{1\leq i\leq n}$ is a partition of $K \cap X$ into n Borel subsets, each of measure at most $\lambda(K)/n$. Check that, for n large enough, all the functions $nf\mathbf{1}_{K_i}$ belong to $B(0,r)$.

iii. Deduce that $\bar{c}\big(B(0,r)\big) = L^p$ for every $r > 0$.

2. a. Let p, q, r be real numbers in $[1,\infty)$ satisfying $1/r = 1/p + 1/q$. Show that, if $f \in L^p$ and $g \in L^q$, then $fg \in L^r$ and

$$\|fg\|_r \leq \|f\|_p \, \|g\|_q.$$

b. Let f be an $\mathscr{F}$-measurable function from X to $\mathbb{K}$. Show that the set J defined by

$$J = \left\{ p \in [1,+\infty) \,:\, 0 < \int |f|^p \, dm < +\infty \right\}$$

is a (possibly empty) interval.
Hint. If $r \in [p,q]$ and $f \in L^p \cap L^q$, introduce the real number $x \in [0,1]$ such that $1/r = (x/p) + (1-x)/q$.

c. Let $(X, \mathscr{F})$ be $\mathbb{R}$ with its Borel σ-algebra, and let m be Lebesgue measure. For each $p \in [1,\infty]$, find an element of L^p that belongs to no other L^q, for $q \neq p$.

d. Show that the map from J to $\mathbb{R}$ defined by

$$p \mapsto \log\left(\int |f|^p \, dm\right)$$

is a convex function.

e. Show that, for every $q \in [1,\infty)$,

$$L^q \cap L^\infty \subset \bigcap_{q\leq p\leq\infty} L^p$$

and that, for every $f \in L^q \cap L^\infty$,

$$\lim_{p\to+\infty} \|f\|_p = \|f\|_\infty.$$

Hint. Show that $0 < a < \|f\|_\infty$ implies $a \le \liminf_{p\to\infty} \|f\|_p$.

3. Take $p \in (1,\infty)$ and let p' be the conjugate exponent of p. Let K be a nonnegative-valued Borel function on $(0,+\infty)^2$ satisfying these conditions:

- $xK(xy, xz) = K(y,z)$ for all $x,y,z \in (0,+\infty)$.
- $\displaystyle\int_0^{+\infty} K(1,z) z^{-1/p}\, dz = k < +\infty.$

a. Show that $\displaystyle\int_0^{+\infty} K(z,1) z^{-1/p'}\, dz = k.$

b. Show that the equation

$$Tf(x) = \int_0^{+\infty} K(x,y) f(y)\, dy$$

defines a continuous linear operator from $L^p((0,+\infty))$ to itself, of norm at most k.
Hint. First find an upper bound for $|Tf(x)|$, by writing

$$K(x,y) = K(x,y)^{1/p}\Big(\frac{y}{x}\Big)^{1/pp'} K(x,y)^{1/p'}\Big(\frac{x}{y}\Big)^{1/pp'}$$

and using the Hölder inequality.

c. Suppose in addition that $K(1,z) \le 1$ for every $z > 0$. If $\varepsilon > 0$, set

$$k_\varepsilon = \int_0^{+\infty} K(1,z) z^{-(1+\varepsilon)/p}\, dz,$$

$$f_\varepsilon(x) = 1_{\{x\ge 1\}}\, x^{-(1+\varepsilon)/p}, \qquad g_\varepsilon(x) = 1_{\{x\ge 1\}}\, x^{-(1+\varepsilon)/p'}.$$

Check that $f_\varepsilon \in L^p((0,+\infty))$ and $g_\varepsilon \in L^{p'}((0,+\infty))$; then show that, for every $\varepsilon < p/2p'$,

$$\int_0^{+\infty} Tf_\varepsilon(x) g_\varepsilon(x)\, dx \ge \big(k_\varepsilon - 2(p')^2\varepsilon\big) \|f_\varepsilon\|_p \|g_\varepsilon\|_{p'}.$$

Deduce that $\|T\| = k$.

d. Show that the maps K defined by $K(x,y) = 1/(x+y)$ and $K(x,y) = 1/\max(x,y)$ satisfy the assumptions above for every $p \in (1,+\infty)$. Compute the norm of the operator T in these two cases. Recall that, for $\alpha > 1$,

$$\int_0^{+\infty} \frac{dx}{1+x^\alpha} = \frac{\pi}{\alpha} \sin\frac{\pi}{\alpha}.$$

(See also Exercise 17 on page 228.)

4. Let m and n be σ-finite measures on measure spaces $(X, \mathscr{F})$ and $(Y, \mathscr{G})$, and let K be a nonnegative-valued function on $X \times Y$, measurable with respect to the product σ-algebra $\mathscr{F} \times \mathscr{G}$. Take $r, s \in [1, \infty)$ such that $s \leq r$. We wish to prove the following inequality:

$$\left(\int_Y \left(\int_X K(x,y)^s \, dm(x)\right)^{r/s} dn(y)\right)^{1/r}$$
$$\leq \left(\int_X \left(\int_Y K(x,y)^r \, dn(y)\right)^{s/r} dm(x)\right)^{1/s} \quad (\leq +\infty). \qquad (*)$$

a. Suppose that $s = 1 < r$, that K is bounded, and that m and n have finite mass. Put

$$a = \int_Y \left(\int_X K(x,y) \, dm(x)\right)^r dn(y) < +\infty,$$
$$b = \int_X \left(\int_Y K(x,y)^r \, dn(y)\right)^{1/r} dm(x) < +\infty.$$

i. Show that

$$a = \int_X \left(\int_Y K(x,y) \left(\int_X K(x',y) \, dm(x')\right)^{r-1} dn(y)\right) dm(x).$$

ii. Applying the Hölder inequality to the integral over Y, prove that $a \leq b \, a^{1/r'}$, where r' is the conjugate exponent of r.

iii. Deduce $(*)$ in this case.

b. Show that $(*)$ holds in general if $s = 1 \leq r$.

c. For s arbitrary, reduce to the preceding case by setting $\tilde{K} = K^s$ and $\tilde{r} = r/s$.

5. We suppose that m is σ-finite and fix $p \in [1, +\infty)$.

a. Let g be a measurable function on X such that $fg \in L^p$ for every $f \in L^p$. Show that $g \in L^\infty$.

Hint. Show that otherwise one can construct a sequence $(X_n)_{n \in \mathbb{N}}$ of pairwise disjoint measurable subsets of X, each with finite positive measure and such that $|g| > 2^n$ almost everywhere on X_n. Then consider the function f defined by

$$f = \sum_{n \in \mathbb{N}} 1_{X_n} 2^{-n} m(X_n)^{-1/p}.$$

Show that $f \in L^p$ and that $fg \notin L^p$.

b. For $g \in L^\infty$ we define a continuous operator T_g on L^p by $T_g(f) = gf$. Let T be a continuous operator on L^p that commutes with all the T_g, for $g \in L^\infty$. Show that there exists $h \in L^\infty$ such that $T = T_h$.

Hint. Construct a positive-valued function g such that $g \in L^p \cap L^\infty$. Let $h = T(g)/g$. Show that $T(f) = hf$ for every $f \in L^p \cap L^\infty$; then conclude.

6. Suppose that m is σ-finite. Let p and r be real numbers such that $1 \le r < p$, and let g be a measurable function such that $fg \in L^r$ for every $f \in L^p$.

a. Show that the map $\Phi : f \mapsto fg$ is continuous from L^p to L^r.
Hint. Show that otherwise there exists a sequence $(f_n)_{n\ge1}$ of positive functions of L^p such that, for every $n \ge 1$, $\|f_n\|_p \le 1$ and $\|f_n g\|_r \ge n$. Then prove, on the one hand, that the function $h = \sum_{n=1}^{+\infty} n^{-2} f_n^r$ is in $L^{p/r}$ and therefore that $f = h^{1/r}$ is in L^p, and on the other hand that $fg \notin L^r$.

b. Deduce that $g \in L^q$, where q is given by $1/r = 1/p + 1/q$.
Hint. Let $(A_n)_{n\in\mathbb{N}}$ be an increasing sequence of elements of $\mathscr{F}$ with finite measure and such that $\bigcup_{n\in\mathbb{N}} A_n = X$. Put

$$g_n = \bigl(\inf(|g|, n)\bigr) 1_{A_n}.$$

Show that

$$\left(\int g_n^q \, dm\right)^{1/r} \le \|\Phi\| \left(\int g_n^q \, dm\right)^{1/p}.$$

7. An ordered set $(E, \le)$ is called a *conditionally complete lattice* if every nonempty subset of E that has an upper bound has a supremum (least upper bound) in E, and every nonempty subset that has a lower bound has an infimum in E.
We consider the space $E = L^p_{\mathbb{R}}$, for $1 \le p \le \infty$, with the natural order defined by

$$f \le g \iff f(x) \le g(x) \ m\text{-almost everywhere.}$$

a. Suppose $p = 1$. Let $\mathscr{A}$ be a nonempty family in $L^1_{\mathbb{R}}$ bounded above, and let $\mathscr{U}$ be the set of its upper bounds.

i. Show that the expression $a = \inf\bigl\{\int f \, dm : f \in \mathscr{U}\bigr\}$ is finite.

ii. Show that there exists a decreasing sequence (f_n) in $\mathscr{U}$ such that

$$\lim_{n\to+\infty} \int f_n \, dm = a.$$

Let f be the almost-everywhere limit of (f_n). Show that $f \in \mathscr{U}$ and that $\int f \, dm = a$.

iii. Deduce that f is the supremum of $\mathscr{A}$ in $L^1_{\mathbb{R}}$, and so that $L^1_{\mathbb{R}}$ is a conditionally complete lattice.
Hint. If $g \in \mathscr{U}$, show that $\int \inf(f, g) \, dm = a$ and deduce that $f \le g$.

b. Suppose that $1 < p < \infty$. Show that $L^p_{\mathbb{R}}$ is a conditionally complete lattice.
Hint. If $\mathscr{A}$ is a nonempty family in $L^p_{\mathbb{R}}$ bounded above, the set $\{f\,|f|^{p-1} : f \in \mathscr{A}\}$ is contained in $L^1_{\mathbb{R}}$.

c. **i.** Show that if m is σ-finite $L^\infty_{\mathbb{R}}$ is a conditionally complete lattice.
Hint. Start by dealing with the case where m has finite mass (then $L^\infty \subset L^1$).

ii. Show that this result may be false if m is not σ-finite.
Hint. Take two uncountable disjoint sets A and B. Let X be their union, let $\mathscr{F}$ be the set of subsets of X that are countable or have countable complement, let m be the count measure on $\mathscr{F}$, and set $\mathscr{A} = (1_{\{x\}})_{x \in A}$.

d. Let E be the quotient of the space of $\mathscr{F}$-measurable real functions by the relation of equality m-almost everywhere. Give E the natural order defined earlier. Show that, if m is σ-finite, E is a conditionally complete lattice.
Is the space of $\mathscr{F}$-measurable real functions with the natural order a conditionally complete lattice?

8. Prove that the set of integrable piecewise constant functions is dense in L^∞ if and only if m has finite mass.
Hint. Take $f = 1$. If m has infinite mass, any integrable piecewise constant function s lies at a distance $\|s - f\|_\infty \geq 1$ from f.

9. Prove that $L^1 \cap L^\infty$ is dense in L^∞ if and only if m has finite mass.

10. Consider the following property:
(P) There exists an (infinite) sequence of $\mathscr{F}$-measurable, pairwise disjoint subsets of X of positive measure.

a. Show that, if (P) is satisfied, L^∞ is not separable.
Hint. You can use as inspiration the ℓ^∞ case in Exercise 7 on page 11.

b. Suppose (P) is not satisfied. Define an *atom* as any $\mathscr{F}$-measurable subset A of positive measure that does not contain any subset $B \in \mathscr{F}$ with $m(B) > 0$ and $m(A \setminus B) > 0$.

i. Show that every measurable subset of X with nonzero measure contains at least one atom.
Hint. Consider the relation $\leq$ defined on the set $\mathscr{A}$ of elements of $\mathscr{F}$ of nonzero m-measure by

$$A \leq B \iff m(B \setminus A) = 0.$$

Apply Zorn's Lemma (see Exercise 11 on page 133) to the order relation induced by $\leq$ on the quotient set $\mathscr{A}/_{\simeq}$, where $\simeq$ is equality almost everywhere:

$$A \simeq B \iff m(B \setminus A) = m(A \setminus B) = 0 \iff A \leq B \text{ and } B \leq A.$$

You might show, in particular, that every totally ordered subset of $\mathscr{A}/_{\simeq}$ has a greatest element.

ii. Show that there exists a finite sequence $(X_n)_{n \leq n_0}$ of atoms such that $m\big(X \setminus \bigcup_{n \leq n_0} X_n\big) = 0$ and

$$m(X_n \cap X_m) = 0 \quad \text{for } n \neq m.$$

iii. Show that every $\mathscr{F}$-measurable function coincides m-almost everywhere with a linear combination of functions 1_{X_n}, for $n \leq n_0$.

c. Show the equivalence of the following properties:

i. (P) is not satisfied.

ii. L^∞ has finite dimension.

iii. L^∞ is separable.

iv. Every $\mathscr{F}$-measurable function belongs to $\mathscr{L}^\infty$.

11. Let L be a vector subspace of $\mathscr{L}^1_{\mathbb{R}}(m) \cap \mathscr{L}^\infty_{\mathbb{R}}(m)$ satisfying these hypotheses:

- There exists an increasing sequence (φ_n) in L that converges to 1 m-almost everywhere.
- The σ-algebra $\sigma(L)$ generated by L equals $\mathscr{F}$.
- $f^2 \in L$ for all $f \in L$.

a. Give the space $L^1_{\mathbb{R}}(m) \cap L^\infty_{\mathbb{R}}(m)$ the norm $\|\cdot\|_1 + \|\cdot\|_\infty$ and denote by $\tilde{L}$ the closure of L in that space. Show that $f \in L$ implies $|f| \in \tilde{L}$. Deduce that $|f| \in \tilde{L}$ for all $f \in \tilde{L}$.

Hint. Use the example on page 29 and argue as in the proof of Theorem 2.3 on page 33.

b. Show that $\tilde{L}$ is dense in $L^1(m)$.

Hint. Apply Proposition 2.4 on page 63.

c. Deduce that L is dense in $L^p_{\mathbb{R}}(m)$ for $1 \leq p < \infty$.

Hint. If $f \in L^p_{\mathbb{R}}(m)$, you might show that, for every $n \in \mathbb{N}$, the function $\sup\bigl(\inf(f, n\varphi_n^+), -n\varphi_n^+\bigr)$ can be approximated in $L^p_{\mathbb{R}}(m)$ by a sequence in $\tilde{L}$.

12. Let m and μ be σ-finite measures on measurable spaces $(X, \mathscr{F})$ and $(Y, \mathscr{G})$, and suppose $p \in [1, \infty)$. We denote by $L^p(m) \otimes L^p(\mu)$ the vector subspace of $L^p(m \times \mu)$ spanned by the functions $(x, y) \mapsto f(x)g(y)$, with $f \in L^p(m)$ and $g \in L^p(\mu)$. Show that $L^p(m) \otimes L^p(\mu)$ is dense in $L^p(m \times \mu)$. This generalizes the result of Exercise 7 on page 110.

Hint. Apply the result of Exercise 11 above to the measure $m \times \mu$ and the space $L = \bigl(\mathscr{L}^1_{\mathbb{R}}(m) \cap \mathscr{L}^\infty_{\mathbb{R}}(m)\bigr) \otimes \bigl(\mathscr{L}^1_{\mathbb{R}}(\mu) \cap \mathscr{L}^\infty_{\mathbb{R}}(\mu)\bigr)$.

13. Assume m is σ-finite.

a. Suppose the σ-algebra $\mathscr{F}$ is *separable*, that is, generated by a countable family of subsets of X.

i. Show that there exists a countable family $\mathscr{B}$ of elements of $\mathscr{F}$ satisfying these conditions:

- $\sigma(\mathscr{B}) = \mathscr{F}$, where $\sigma(\mathscr{B})$ is the σ-algebra generated by $\mathscr{B}$.
- $A \cap B \in \mathscr{B}$ for all $A, B \in \mathscr{B}$.
- $m(A) < +\infty$ for all $A \in \mathscr{B}$.
- There exists an increasing sequence of elements of $\mathscr{B}$ whose union equals X.

ii. Show that the family $\{1_A\}_{A\in\mathscr{B}}$ is fundamental in L^p, for $1 \le p < \infty$.
Hint. Apply Exercise 11 above.

iii. Deduce that, if $1 \le p < \infty$, the space L^p is separable.

iv. Show that, if X is a separable metric space, the Borel σ-algebra $\mathscr{B}(X)$ is separable. Derive hence another proof for Corollary 1.7.

b. We say that a σ-algebra $\mathscr{F}$ is *almost separable* if there exists a separable σ-algebra $\mathscr{F}'$ contained in $\mathscr{F}$ such that, for all $A \in \mathscr{F}$, there exists $B \in \mathscr{F}'$ with

$$m(A \setminus B) = m(B \setminus A) = 0.$$

i. Show that, if $\mathscr{F}$ is almost separable, the space L^p is separable for every $p \in [1, \infty)$.
Hint. Use part a.

ii. Show that if there exists $p \in [1, \infty)$ such that L^p is separable, $\mathscr{F}$ is almost separable.
Hint. Consider the σ-algebra generated by a sequence of elements of $\mathscr{L}^p$ whose corresponding classes are dense in L^p.

iii. Show that $\mathscr{F}$ is almost separable if and only if there exists a sequence $(A_n)_{n\in\mathbb{N}}$ of measurable subsets of X of finite measure such that the sequence $(1_{A_n})_{n\in\mathbb{N}}$ is fundamental in L^1.

iv. Let $\mathscr{F}_f$ be the set of elements of $\mathscr{F}$ of finite measure, modulo the relation of equality m-almost everywhere. If $A, B \in \mathscr{F}_f$, we write $d(A, B) = m(A \,\Delta\, B)$, where $A \,\Delta\, B = (A \cup B) \setminus (A \cap B)$. Show that d makes $\mathscr{F}_f$ into a complete metric space, separable if and only if the σ-algebra $\mathscr{F}$ is almost separable.
Hint. $(\mathscr{F}_f, d)$ can be identified with the subset of L^1 consisting of (classes of) characteristic functions of elements in $\mathscr{F}$, with the metric defined by the norm $\|\cdot\|_1$.

14. Assume $p \in [1, \infty)$.

a. Let $\mathscr{P}$ be the set of finite families $(A_n)_{n\le n_0}$ in $\mathscr{F}$ such that

- $m(A_n \cap A_m) = 0$ if $n \ne m$, and
- $0 < m(A_n) < \infty$ for every $n \le n_0$.

If $\mathscr{A} = (A_n)_{n\le n_0}$ is an element of $\mathscr{P}$, we define an operator $T_{\mathscr{A}}$ on L^p by

$$T_{\mathscr{A}} f = \sum_{n\le n_0} \left(\frac{1}{m(A_n)} \int_{A_n} f\, dm \right) 1_{A_n}.$$

Show that $T_{\mathscr{A}}$ is a continuous linear operator on L^p, of norm at most 1.

b. If $\mathscr{A}$ and $\mathscr{B}$ are elements of $\mathscr{P}$, write $\mathscr{A} \sqsubseteq \mathscr{B}$ if every element of $\mathscr{B}$ is contained, apart from a set of measure zero, in an element of $\mathscr{A}$, and if every element of $\mathscr{A}$ is, apart from a set of measure zero, the union of the elements of $\mathscr{B}$ contained in it.

Let $\mathscr{A} = (A_n)_{n \le n_0}$ be an element of $\mathscr{P}$ and let f be a linear combination of functions 1_{A_n}, for $n \le n_0$. Show that, for every $\mathscr{B} \in \mathscr{P}$ such that $\mathscr{A} \sqsubseteq \mathscr{B}$, we have $T_{\mathscr{B}} f = f$. Deduce that, for every $\varepsilon > 0$ and every $f \in L^p$, there exists $\mathscr{A} \in \mathscr{P}$ such that

$$(\mathscr{B} \in \mathscr{P} \text{ and } \mathscr{A} \sqsubseteq \mathscr{B}) \Longrightarrow \|T_{\mathscr{B}} f - f\|_p \le \varepsilon.$$

Hint. Use the fact that the set of integrable piecewise constant functions is dense in L^p (see the remark on page 146).

c. Assume that m has finite mass and that there exists a sequence $(\mathscr{A}_n)$ of $\mathscr{P}$ increasing with respect to $\sqsubseteq$ and such that $\mathscr{A}_0 = \{X\}$. Assume also that $\bigcup_{n \in \mathbb{N}} \mathscr{A}_n$ generates $\mathscr{F}$ (you can check that there is such a sequence if the σ-algebra $\mathscr{F}$ is separable: see Exercise 13). Denote by $\mathscr{S}_n$ the set of piecewise constant functions that are constant on each element of $\mathscr{A}_n$. Show that $\bigcup_n \mathscr{S}_n$ is dense in L^p for $1 \le p < \infty$. (You could use Exercise 11, for example). Deduce that, for every $f \in L^p$, the sequence $(T_{\mathscr{A}_n} f)$ converges to f in L^p.

Example. Choose for X the interval $[0, 1]$, for m the Lebesgue measure on X, and for $\mathscr{F}$ the Borel σ-algebra of X. Find a sequence $(\mathscr{A}_n)$ satisfying the conditions stated above.

15. We say that a sequence $(f_n)_{n \in \mathbb{N}}$ of $\mathscr{F}$-measurable functions *converges in measure* to a $\mathscr{F}$-measurable function f if, for every $\varepsilon > 0$,

$$m\big(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}\big) \longrightarrow 0.$$

a. Assume $p \in [1, \infty)$.

i. *Bienaymé–Chebyshev inequality.* Take $f \in L^p$. Show that, for every $\delta > 0$,

$$m\big(\{x \in X : |f(x)| > \delta\}\big) \le \delta^{-p} \|f\|_p^p.$$

ii. Let (f_n) be a sequence of elements of L^p that converges in L^p to $f \in L^p$. Show that the sequence (f_n) converges to f in measure.

b. Let (f_n) be a sequence of measurable functions that converges in measure to a measurable function f.

i. Show that there exists a subsequence (f_{n_k}) such that, for every $k \in \mathbb{N}$,

$$m\big(\{x \in X : |f_{n_k}(x) - f(x)| > 2^{-k}\}\big) \le 2^{-k}.$$

ii. For each $k \in \mathbb{N}$, let Z_k be the subset of X defined by

$$Z_k = \bigcup_{j \ge k} \{x \in X : |f_{n_j}(x) - f(x)| > 2^{-j}\}.$$

Then set $Z = \bigcap_{k \in \mathbb{N}} Z_k$. Prove that $m(Z) = 0$.

iii. Deduce that the sequence (f_{n_k}) converges to f m-almost everywhere.

iv. Show also that, for every $\varepsilon > 0$, there exists a measurable subset A of X of measure at most ε and such that the sequence (f_{n_k}) converges uniformly to f on $X \setminus A$.
Hint. Choose $A = Z_k$, with k large enough.

c. Suppose $m(X) < +\infty$. Let (f_n) be a sequence of measurable functions that converges m-almost everywhere to a measurable function f. Show that the sequence (f_n) converges in measure to f.
Hint. Take $\varepsilon > 0$. For each integer $N \in \mathbb{N}$, put

$$A_N = \{x \in X : |f_n(x) - f(x)| \leq \varepsilon \text{ for all } n \geq N\}.$$

Show that there exists an integer $N \in \mathbb{N}$ for which $m(X \setminus A_N) < \varepsilon$, and therefore that $m(X \setminus A_n) < \varepsilon$ for every $n \geq N$.
Deduce that, for every integer $n \geq N$,

$$m(\{x \in E : |f_n(x) - f(x)| > \varepsilon\}) < \varepsilon.$$

16. Suppose $p \in [1, \infty]$. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in L^p such that the series $\sum_{n \in \mathbb{N}} \|f_n - f_{n+1}\|_p$ converges. Show that the sequence (f_n) converges almost everywhere and in L^p.
Hint. Suppose first that $m(X)$ is finite and prove that in this case

$$\int \sum_{n \in \mathbb{N}} |f_n - f_{n+1}| \, d\mu < +\infty.$$

If m is arbitrary and $p < \infty$, check that the set $\{x \in X : f_n(x) \neq 0$ for some $n \in \mathbb{N}\}$ is σ-finite.

17. *Equiintegrability.* Assume $p \in [1, \infty)$. A subset $\mathscr{H}$ of L^p is called *equiintegrable of order p* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for every measurable subset A of X of m-measure at most δ,

$$\int_A |f|^p \, dm \leq \varepsilon \quad \text{for all } f \in \mathscr{H}.$$

a. **i.** Show that every subset $\mathscr{H}$ of L^p for which

$$\lim_{n \to +\infty} \int_{\{|f| > n\}} |f|^p \, dm = 0 \text{ uniformly with respect to } f \in \mathscr{H} \qquad (*)$$

is equiintegrable of order p. Deduce that every finite subset of L^p is equiintegrable of order p.
Show that, conversely, every bounded subset $\mathscr{H}$ of L^p that is equiintegrable of order p satisfies $(*)$.

ii. Take $\mathscr{H} \subset L^p$. Suppose there exists an element $g \in L^p$, nonnegative m-almost everywhere, such that, for every $f \in \mathscr{H}$, we have $|f| \leq g$ m-almost everywhere. Show that $\mathscr{H}$ is equiintegrable of order p.

iii. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in L^p that converges in L^p to f. Show that the family $(f_n)_{n\in\mathbb{N}}$ is equiintegrable of order p.
Hint. You might check that, if A is a measurable subset of X, then

$$\left(\int_A |f_n|^p\,dm\right)^{1/p} \leq \|f - f_n\|_p + \left(\int_A |f|^p\,dm\right)^{1/p}.$$

b. We now assume that m has finite mass.

i. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in L^p and let $f \in L^p$. Show that the sequence $(f_n)_{n\in\mathbb{N}}$ converges to f in L^p if and only if these conditions are satisfied:

- The sequence $(f_n)_{n\in\mathbb{N}}$ converges in measure to f (see Exercise 15 above for definition).
- The family $\{f_n\}_{n\in\mathbb{N}}$ is equiintegrable of order p.

ii. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of elements of L^p that converges in measure to a function f. Assume that there exists $g \in L^p$ such that $|f_n| \leq |g|$ for every $n \in \mathbb{N}$. Show that $f \in L^p$ and that the sequence (f_n) converges to f in L^p.

iii. Let $(f_n)_{n\in\mathbb{N}}$ be a bounded sequence in L^p that converges almost everywhere to a function f. Check that $f \in L^p$. Then show that, for every real $q \in [1, p)$, we have $\lim_{n\to\infty} \|f_n - f\|_q = 0$.
Hint. Note that if A is a measurable subset of X and if $g \in L^p$, then $\int_A |g|^q\,dm \leq \|g\|_p^q\, m(A)^{1-q/p}$.

18. *Uniformly convex spaces.* A Banach space E is called *uniformly convex* if it has this property:
If (x_n) and (y_n) are sequences in the closed unit ball $\bar{B}(E)$ of E satisfying $\|x_n + y_n\| \to 2$, then $\|x_n - y_n\| \to 0$.

a. Show that every Hilbert space is uniformly convex.

b. Show that, for $n \geq 2$, the space $\mathbb{R}^n$ with the norm $\|\cdot\|_1$ or the norm $\|\cdot\|_\infty$ is not uniformly convex.

c. Let E be a uniformly convex space. Show that every nonempty convex closed subset of E contains a unique point of minimal norm.

d. Let E be a uniformly convex space.

i. Let f be a linear form on E of norm 1 and let (x_n) be a sequence of elements of E of norm 1. Show that, if $f(x_n) \to 1$, the sequence (x_n) converges.
Hint. You might show that (x_n) is a Cauchy sequence, using the fact that $f(x_n + x_m) \to 2$ when $n, m \to +\infty$.

ii. Deduce that the absolute value of any continuous linear form on E attains its maximum in the closed unit ball of E.

e. Assume $p \in (1, \infty)$.

i. Show that

$$\left|\frac{x+y}{2}\right|^p \leq \frac{|x|^p + |y|^p}{2} \quad \text{for all } x, y \in \mathbf{C}. \tag{$*$}$$

ii. Set $\bar{D} = \{z \in \mathbf{C} : |z| \leq 1\}$. Show that the function φ defined on $\bar{D}$ by

$$\varphi(z) = \frac{|1+z|^p}{1+|z|^p}$$

is continuous from $\bar{D}$ to $[0, 2^{p-1}]$ and that $\varphi(z) = 2^{p-1}$ if and only if $z = 1$. Deduce that, for every $\eta > 0$, there exists $\delta(\eta) > 0$ such that, for every $(x, y) \in \bar{D}^2$ with $|x - y| \geq \eta$,

$$\left|\frac{x+y}{2}\right|^p \leq \bigl(1 - \delta(\eta)\bigr) \frac{|x|^p + |y|^p}{2}.$$

iii. Take $\varepsilon > 0$ and let f and g be points in the closed unit ball of L^p such that $\|f - g\|_p \geq \varepsilon$. Set

$$E = \bigl\{x \in X : |f(x) - g(x)| \geq \varepsilon 2^{-2/p} \max(|f(x)|, |g(x)|)\bigr\}.$$

A. Show that $\displaystyle\int_{X \setminus E} |f - g|^p \, dm \leq \varepsilon^p/2$. Deduce that

$$\int_E \frac{|f|^p + |g|^p}{2} \, dm \geq \frac{\varepsilon^p}{2 \cdot 2^p}.$$

(You might use $(*)$ with $x = f$ and $y = -g$.)

B. Show that

$$\left\|\frac{f+g}{2}\right\|_p^p \leq 1 - \delta\left(\frac{\varepsilon}{2^{2/p}}\right) \frac{\varepsilon^p}{2^{p+1}},$$

where δ is as in part e-ii above.

Hint. Use $(*)$ in $X \setminus E$ and the conclusion of e-ii in E, taking $\eta = \varepsilon / 2^{2/p}$.

C. Deduce that L^p is uniformly convex (*Clarkson's Theorem*).

f. Let X be a metric space and give $E = C_b(X)$ the uniform norm $\|\cdot\|$. Suppose that X contains a point a that is not isolated, and fix a sequence (x_n) of pairwise distinct points in X that converges to a. For $f \in E$, put

$$L(f) = \sum_{n \in \mathbf{N}} \left(-\tfrac{1}{2}\right)^n f(x_n).$$

i. Show that L is a continuous linear form on E of norm 2 and that $|L(f)| < 2$ for all $f \in \bar{B}(E)$.

ii. Set $C = \{f \in E : L(f) = 2\}$. Show that C is a nonempty closed convex set in E, that $\|f\| > 1$ for all $f \in C$, and that $\inf_{f \in C} \|f\| = 1$.

19. Suppose m is a Radon measure. If $1 \le p \le \infty$, we denote by $L^p_{\text{loc}}(m)$, or, more simply, by L^p_{loc}, the set of equivalence classes of functions f such that, for every compact K in X, the function $1_K f$ lies in L^p. We denote by L^p_c the set of elements of L^p having compact support (the support of an element of L^p was defined on page 145).

a. Show that, if $1 \le p \le q \le \infty$, then $L^q_{\text{loc}} \subset L^p_{\text{loc}}$ and $L^q_c \subset L^p_c$.

b. Find a metric d on L^p_{loc} such that, for every sequence $(f_n)_{n \in \mathbb{N}}$ in L^p_{loc} and every $f \in L^p_{\text{loc}}$, the condition $\lim_{n \to +\infty} d(f, f_n) = 0$ is equivalent to the condition that $\lim_{n \to +\infty} \|1_K (f_n - f)\|_p = 0$ for every compact K of X. Show that L^p_{loc} is complete with this metric.
Hint. You might work as in Exercise 12 on page 57.

c. Show that the space L^p_c is dense in L^p_{loc} with the metric d.

2 Duality

We consider again in this section a measure space $(X, \mathscr{F})$ and a measure m on $\mathscr{F}$. We assume here that m is σ-finite. We will determine, for $1 \le p < \infty$, the topological dual $(L^p)'$ of the space L^p.

So fix $p \in [1, +\infty)$ and let p' be the conjugate exponent of p, so that $1/p + 1/p' = 1$. Note first that every element $g \in L^{p'}$ defines a linear form T_g on L^p, as follows:

$$T_g f = \int fg \, dm \quad \text{for all } f \in L^p. \tag{$*$}$$

As an immediate consequence of the Hölder inequality, the linear form T_g is continuous and its norm in $(L^p)'$ is at most that of g in $L^{p'}$. We will show that one obtains in this way *all* continuous linear forms on L^p.

Theorem 2.1 *If $1 \le p < \infty$, the linear map $g \mapsto T_g$ defined on $L^{p'}$ by $(*)$ is a surjective isometry from $L^{p'}$ onto $(L^p)'$.*

If $p = p' = 2$, this is of course an immediate consequence of the Riesz Representation Theorem (Theorem 3.1 on page 111) in the Hilbert space L^2. The basic scheme of the proof is to reduce the problem to this case. This can easily be done if $1 \le p < 2$, but we will give a proof that is valid for every $p \in [1, \infty)$, whose main idea goes back to J. von Neumann.

Proof. The proof of Theorem 2.1 will be carried out in several steps. The crucial point is the following lemma.

Lemma 2.2 *Suppose m has finite mass. Let T be a continuous linear form on $L^p_{\mathbb{R}}$. If T is positive (that is, if $Tf \geq 0$ for every $f \in L^p_{\mathbb{R}}$ such that $f \geq 0$), there exists a measurable function $g \geq 0$ such that, for every $f \in L^p_{\mathbb{R}}$,*

$$fg \in L^1_{\mathbb{R}} \quad \textit{and} \quad Tf = \int fg\,dm.$$

Proof. (All functions are assumed real-valued without further notice.) Since the linear form T is positive, we can define on $(X, \mathscr{F})$ a measure λ of finite mass by setting

$$\lambda(A) = T(1_A) \quad \text{for all } A \in \mathscr{F}. \tag{**}$$

That λ is σ-additive follows easily from the continuity and linearity of T (using the Dominated Convergence Theorem, which is allowed because m has finite mass). Then we set

$$\nu = \lambda + m. \tag{†}$$

Since T acts on classes of functions, we see that $m(A) = 0$ implies $\lambda(A) = 0$; thus, for $A \in \mathscr{F}$,

$$\nu(A) = 0 \iff m(A) = 0 \implies \lambda(A) = 0.$$

Hence the linear form $f \mapsto \int f\,d\lambda$ is well defined on $L^2(\nu)$ and we have, for every $f \in L^2(\nu)$,

$$\left|\int f\,d\lambda\right| \leq \left(\int |f|^2 d\lambda\right)^{1/2} (\lambda(X))^{1/2} \leq \|f\|_{L^2(\nu)} (\lambda(X))^{1/2}.$$

By the Riesz Representation Theorem (Theorem 3.1 on page 111) applied to the Hilbert space $L^2(\nu)$, there exists an element h in $L^2(\nu)$ such that

$$\int f\,d\lambda = \int fh\,d\nu \quad \text{for all } f \in L^2(\nu). \tag{‡}$$

In particular,

$$0 \leq \lambda(\{h < 0\}) = \int_{\{h<0\}} h\,d\nu \leq 0,$$

which implies that $h \geq 0$ ν-almost everywhere. Likewise,

$$\lambda(\{h \geq 1\}) = \int_{\{h \geq 1\}} h\,d\nu \geq \lambda(\{h \geq 1\}) + m(\{h \geq 1\}),$$

which implies that $h < 1$ m-almost everywhere and so ν-almost everywhere. Hence we can choose a representative of h such that $0 \leq h(x) < 1$ for every $x \in X$.

Now let f be an m-integrable piecewise constant function. By (**), (‡), and (†),

$$Tf = \int f\,d\lambda = \int fh\,dm + \int fh\,d\lambda.$$

At the same time, by approximating h with piecewise constant functions and using the continuity of T, we see easily that $\int fh\,d\lambda = T(fh)$. We deduce that

$$T(f(1-h)) = \int fh\,dm.$$

Since this holds for every m-integrable piecewise constant f, it also holds for every $f \in L^p(m)$ such that $f \geq 0$ (use an increasing approximating sequence; see the remark on page 146). Now let $f \in L^p(m)$ be such that $f \geq 0$. For every integer k, $\inf\bigl(f/(1-h),\, k\bigr) \in L^p(m)$, so

$$T\bigl(\inf(f,\, k(1-h))\bigr) = \int \inf\Bigl(\frac{f}{1-h},\, k\Bigr) h\,dm.$$

By making k approach infinity and using again the continuity of T, we get

$$Tf = \int \frac{fh}{1-h}\,dm.$$

Thus, $g = h/(1-h)$ serves our purposes. □

We now get, without having to assume that m has finite mass:

Lemma 2.3 *If $T \in (L^p)'$, there exists a measurable function g such that, for all $f \in L^p$,*

$$fg \in L^1 \quad \textit{and} \quad Tf = \int fg\,dm.$$

Proof. For $f \in L^p_{\mathbb{R}}$, set $T_1 f = \mathrm{Re}(Tf)$ and $T_2 f = \mathrm{Im}(Tf)$. Then T_1 and T_2 belong to $(L^p_{\mathbb{R}})'$. If Lemma 2.3 is true in the real case, we can apply it to T_1 and T_2 to obtain real functions g_1 and g_2, and clearly the function $g = g_1 + ig_2$ works for T. Therefore we can suppose we are in the real case.

In this case T can be written as the difference of two continuous and positive linear forms on $L^p_{\mathbb{R}}$ (apply Remark 2 on page 88 to the lattice $L^p_{\mathbb{R}}$). So we can in fact suppose that T is a positive continuous linear form on $L^p_{\mathbb{R}}$, and we do so.

Since the measure m is σ-finite, there exists a countable partition (K_n) of X consisting of elements of $\mathscr{F}$ of finite measure. For each integer n, let m_n be the restriction of m to K_n. If $f \in L^p_{\mathbb{R}}(m_n)$, denote by $\tilde{f}$ the extension of f to X taking the value 0 on $X \setminus K_n$. The linear form on $L^p_{\mathbb{R}}(m_n)$ defined by $f \mapsto T(\tilde{f})$ then satisfies on K_n the hypotheses of Lemma 2.2. Therefore there is a positive measurable function g_n on K_n such that, for all $f \in L^p_{\mathbb{R}}(m_n)$,

$$fg_n \in L^1_{\mathbb{R}}(m_n) \quad \text{and} \quad T(\tilde{f}) = \int fg_n\,dm_n.$$

Now let g be the measurable function on X whose restriction to each K_n is g_n. If $f \in L^p_{\mathbb{R}}(m)$ and $f \geq 0$, we have $f = \sum_{n=0}^{+\infty} f 1_{K_n}$, the series being convergent in $L^p_{\mathbb{R}}(m)$. By the continuity of T and monotone convergence in the integral, we deduce that

$$Tf = \sum_{n=0}^{+\infty} T(f 1_{K_n}) = \sum_{n=0}^{+\infty} \int_{K_n} fg\, dm = \int fg\, dm.$$

Thus g satisfies the necessary conditions. □

Lemma 2.4 *With the notation of Lemma 2.3, we have $g \in L^{p'}$ and $\|g\|_{p'} \leq \|T\|'_p$, where $\|\cdot\|'_p$ is the norm in $(L^p)'$.*

Proof. Since the measure m is σ-finite, there exists an increasing sequence (A_n) of elements of $\mathscr{F}$ of finite measure that cover X.

1. *Case $p = 1$.* Suppose the conclusion of the lemma is false. Then the set $\{|g| > \|T\|'_1\}$ has positive measure, so there exists $\varepsilon > 0$ such that the set $A = \{|g| > \|T\|'_1 + \varepsilon\}$ has positive measure. Let α be the function that equals $|g|/g$ on $\{g \neq 0\}$ and 1 on $\{g = 0\}$. Then, on the one hand,

$$T(\alpha 1_{A\cap A_n}) = \int_{A\cap A_n} |g|\, dm \geq \left(\|T\|'_1 + \varepsilon\right) m(A \cap A_n)$$

and, on the other,

$$T(\alpha 1_{A\cap A_n}) \leq \|T\|'_1 m(A \cap A_n).$$

There certainly exists an integer n for which $m(A \cap A_n) > 0$, so we deduce that $\|T\|'_1 + \varepsilon \leq \|T\|'_1$, which is absurd.

2. *Case $1 < p < \infty$.* Define α as in the preceding case and, for $n \in \mathbb{N}$, set $B_n = A_n \cap \{|g| \leq n\}$ and $f_n = 1_{B_n} \alpha |g|^{p'-1}$. Then, for every n,

$$\int_{B_n} |g|^{p'} dm = Tf_n \leq \|T\|'_p \left(\int_{B_n} |g|^{p'} dm\right)^{1/p},$$

so

$$\left(\int_{B_n} |g|^{p'} dm\right)^{1/p'} \leq \|T\|'_p,$$

whence we deduce the result by making n approach infinity. □

Thus we have proved the following fact: *For every $T \in (L^p)'$ there exists $g \in L^{p'}$ such that*

$$T = T_g \quad \textit{and} \quad \|g\|_{p'} = \|T\|'_p.$$

The proof of Theorem 2.1 will be complete if we show that the map $g \mapsto T_g$ is injective. Suppose that $g \in L^{p'}$ and $T_g = 0$. Defining a sequence

(A_n) and a function α as in the proof of Lemma 2.4, we see that, for every n, the function $g_n = \alpha 1_{A_n}$ is an element of $L^1 \cap L^\infty$, and so

$$T_g g_n = \int_{A_n} |g|\, dm = 0.$$

This proves that $g = 0$ m-almost everywhere. □

Remark. Theorem 2.1 is false for $p = \infty$. In general, L^1 is not isometric to the topological dual of L^∞, only to a proper subset thereof. (On this topic, see Exercises 3, 4, and 5 below.)

Exercises

In all the exercises, m denotes a σ-finite measure on a measure space $(X, \mathscr{F})$.

1. Suppose that X is an open set in $\mathbb{R}^d$ (with $d \geq 1$) and that m is the restriction to X of Lebesgue measure on $\mathbb{R}^d$. Fix $p \in (0,1)$. Let L be a continuous linear map from L^p to a normed vector space E, where we have given L^p the metric d_p defined in Exercise 1 on page 147. Show that $L = 0$. In particular, the topological dual of (L^p, d_p) is $\{0\}$.
   Hint. Show that, for every $\varepsilon > 0$, the inverse image under L of the closed ball $\bar{B}(0, \varepsilon)$ of E is a closed and convex neighborhood of 0 in L^p. Then use the result of Exercise 1 on page 147.
2. Set $X = \{0, 1\}$ and let ν be the measure on $\mathscr{P}(X)$ defined by $\nu(\{0\}) = 1$ and $\nu(\{1\}) = \infty$. Show that $L^\infty(\nu)$ is not isometric to the dual of $L^1(\nu)$.
3. Recall from Exercise 7 on page 11 that c_0 stands for the subspace of ℓ^∞ consisting of sequences that tend to 0 at infinity. Show that the map that associates to each element g of ℓ^1 the linear form on c_0 defined by

   $$T_g : f \mapsto \sum_{n \in \mathbb{N}} f_n g_n$$

   is a surjective isometry from ℓ^1 onto c_0'.
4. *A realization of the topological dual of $\ell^\infty(I)$.* Let I be an infinite set. Denote by $\Lambda(I)$ the set of *finitely additive* functions μ from $\mathscr{P}(I)$ to $[0, +\infty)$, that is, those satisfying

   $$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B) \quad \text{for all } A, B \in \mathscr{P}(I)$$

   and $\mu(\varnothing) = 0$.
   a. Take $\mu \in \Lambda(I)$. Define a linear form L_μ on the set of piecewise constant functions on I as follows: If $g = \sum_{k=1}^n g_k 1_{I_k}$, where $(I_k)_{1 \leq k \leq n}$ is a partition of I and $g_1, \ldots, g_n \in \mathbb{K}$, put

   $$L_\mu(g) = \sum_{k=1}^n g_k \mu(I_k).$$

i. Check that $L_\mu(g)$ is well defined for every piecewise constant function g and that $|L_\mu(g)| \le \mu(I)\|g\|_\infty$.

ii. Show that L_μ can be uniquely extended to a positive continuous linear form on $\ell^\infty(I)$ of norm $\mu(I)$, which we still denote by L_μ.

b. Show that, for every positive linear form L on $\ell^\infty(I)$, there is a unique $\mu \in \Lambda(I)$ such that $L = L_\mu$.

c. Describe the topological dual of $\ell^\infty(I)$.

d. **i.** If $f \in \ell^1(I)$ and $f \ge 0$, define a map μ_f on $\mathscr{P}(I)$ by setting

$$\mu_f(A) = \sum_{i\in A} f(i).$$

Show that $\mu_f \in \Lambda(I)$. Write down L_{μ_f} explicitly.

ii. It is a classical consequence of the axiom of choice that, given any infinite set E, there is a finitely additive function $\mathscr{P}(E) \to \{0,1\}$ that is not identically zero† and assigns to every finite subset of E the value 0. Let μ be such a function for the set I. Show that there exists no $f \in \ell^1(I)$ such that $f \ge 0$ and $\mu = \mu_f$. Deduce that there cannot be $f \in \ell^1(I)$ such that

$$L_\mu(g) = \sum_{i\in I} f(i)g(i) \quad \text{for all } g \in \ell^\infty(I).$$

5. *About the topological dual of* L^∞. We say that a linear form T on $L^\infty(m)$ satisfies Property (P) if, for every decreasing sequence (f_n) of $L^\infty_{\mathbb{R}}(m)$ that converges m-almost everywhere to 0, the sequence (Tf_n) converges to 0.

a. Take $g \in L^1(m)$. Define the linear form T_g on $L^\infty(m)$ by setting

$$T_g(f) = \int fg\,dm.$$

Show that T_g is continuous, that it has Property (P), and that its norm in $(L^\infty(m))'$ equals $\|g\|_1$.

b. Consider a continuous and positive linear form T on $L^\infty_{\mathbb{R}}(m)$ that has Property (P). Show that there exists a unique $g \ge 0$ in $L^1(m)$ such that $T = T_g$.

Hint. Define a measure λ of finite mass on $\mathscr{F}$ by $\lambda(A) = T(1_A)$. Then imitate the proof of Lemma 2.2, using the remark made on page 146 about the denseness of piecewise constant functions in $L^\infty(m)$.

c. Let T be a continuous linear form on $L^\infty_{\mathbb{R}}(m)$ that satisfies Property (P). Define T^+ and T^- according to the method of Theorem 4.1 on

---

†If μ is such a function, the set $\mathscr{U} = \mu^{-1}(\{1\})$ is called an *ultrafilter* on I.

page 87 (see Remark 2 on page 88). Show that T^+ and T^- belong to $(L^\infty_{\mathbb{R}}(m))'$ and satisfy (P).

Hint. Let (f_n) be an increasing sequence of positive functions in $L^\infty(m)$ that converges almost everywhere to $f \in L^\infty(m)$. Show that, if g is $\mathscr{F}$-measurable and $0 \le g \le f$, then

$$T(g) = \lim_{n\to+\infty} T\big(\inf(g, f_n)\big) \le \liminf_{n\to+\infty} T^+(f_n).$$

Deduce that $\lim_{n\to\infty} T^+(f_n) = T^+(f)$.

d. Deduce from the facts above that the map from $L^1(m)$ to $(L^\infty(m))'$ defined by $g \mapsto T_g$ is an isometry whose image consists of those elements of $(L^\infty(m))'$ that have Property (P).

6. *The Radon–Nikodým Theorem*

a. Let ν be a σ-finite measure on $\mathscr{F}$ such that any $A \in \mathscr{F}$ of m-measure zero has ν-measure zero. Show that there exists a positive measurable function g such that

$$\nu(A) = \int_A g\,dm \quad \text{for all } A \in \mathscr{F}.$$

Hint. Reduce to the case where ν has finite mass. Then show that the map $f \mapsto \int f\,d\nu$ defined on $L^\infty(m)$ is a continuous linear form satisfying Property (P) of Exercise 5, and use the result in the last question of that exercise.

b. Show that this result remains true if we assume that m is a positive Radon measure and ν is a bounded complex Radon measure on X, and do not require g to be positive, but merely in $L^1(m)$.

Hint. Apply the previous question to the positive measures $(\operatorname{Re}\nu)^+$, $(\operatorname{Re}\nu)^-$, $(\operatorname{Im}\nu)^+$, and $(\operatorname{Im}\nu)^-$, defined according to the notation of Theorem 4.1 on page 87 and the discussion on page 89.

7. *Conditional expectation in L^p.* Let $\mathscr{F}'$ be a σ-algebra contained in $\mathscr{F}$ and let m' be the restriction of m to $\mathscr{F}'$. Suppose m' is σ-finite.

a. Suppose $p \in (1, \infty]$. Show that, for every $f \in L^p(m)$, there exists a unique $\tilde f \in L^p(m')$ such that, for every element A of $\mathscr{F}'$ of finite measure,

$$\int_A f\,dm = \int_A \tilde f\,dm'.$$

Hint. Let p' be the conjugate exponent of p. Consider the linear form on $L^{p'}(m')$ defined by $g \mapsto \int gf\,dm$ and apply Theorem 2.1 on page 159.

b. Show that, for every $f \in L^1(m)$, there exists a unique $\tilde f \in L^1(m')$ such that, for every element A of $\mathscr{F}'$,

$$\int_A f\,dm = \int_A \tilde f\,dm'.$$

Hint. Argue as in the preceding question, using Exercise 5d above instead of Theorem 2.1.

c. Suppose $p \in [1, \infty]$. Show that the map T_p from $L^p(m)$ to $L^p(m')$ defined by $T(f) = \tilde{f}$ is linear and continuous, and that it satisfies $\|T_p\| \leq 1$,

$$f \geq 0 \implies T_p f \geq 0,$$

and $T_p f = f$ for all $f \in L^p(m')$, where we have identified $L^p(m')$ with a subspace of $L^p(m)$.

d. Show that T_2 is the operator of orthogonal projection from $L^2(m)$ onto $L^2(m')$.

e. Show that, if $1 \leq p, q \leq \infty$, then $T_p = T_q$ on $L^p(m) \cap L^q(m)$. Thus we can define an operator T on $\bigcup_{p \in [1,\infty]} L^p(m)$ whose restriction to each $L^p(m)$ is T_p. We call T the *operator of conditional expectation given $\mathscr{F}'$.*

8. Suppose $p, q \in [1, \infty)$. Let T be a continuous linear map from $L^p((0,1))$ to $L^q((0,1))$. Show that there exists a function K from $(0,1)^2$ to $\mathbb{K}$ with these properties: For every $x \in (0,1)$, the function $y \mapsto K(x,y)$ lies in $\mathscr{L}^{p'}((0,1))$ (where p' is the conjugate exponent of p), and

$$\int_0^x Tf(y)\,dy = \int_0^1 K(x,y)f(y)\,dy \quad \text{for all } f \in L^p((0,1)) \text{ and } x \in (0,1).$$

9. *Weak convergence in L^p spaces. Examples.* Let $p \in [1, \infty]$ and p' be conjugate exponents. We say that a sequence $(f_n)_{n \in \mathbb{N}}$ in $L^p(m)$ *converges weakly* to an element f of $L^p(m)$ if

$$\lim_{n \to +\infty} \int f_n g\,dm = \int fg\,dm \quad \text{for all } g \in L^{p'}(m).$$

To avoid confusion, when a sequence in $L^p(m)$ converges in the sense of the $L^p(m)$ norm we will say here that it *converges strongly* in $L^p(m)$.[†]

a. Prove that every sequence in $L^p(m)$ that converges strongly also converges weakly.

b. Show that a sequence (f_n) in $L^p(m)$ converges weakly to $f \in L^p(m)$ if and only if it is bounded and,

– if $p = 1$, $\displaystyle\lim_{n \to +\infty} \int_A f_n\,dm = \int_A f\,dm$ for all $A \in \mathscr{F}$;

[†] More generally, a sequence (f_n) in a normed vector space E is said to *converge weakly* to $f \in E$ if, for every $L \in E'$, the sequence (Lf_n) converges to Lf. A sequence (L_n) in E' is said to *converge weakly-** to $L \in E'$ if, for every $f \in E$, the sequence $(L_n f)$ converges to Lf. The definition given in the text for L^p spaces corresponds, in the case $p = \infty$, to weak-* convergence in L^∞, considered as the topological dual of L^1. If $1 < p < \infty$, L^p has $L^{p'}$ as its dual, but L^p can also be considered as the dual of $L^{p'}$ (Theorem 2.1). In this case, weak convergence and weak-* convergence coincide.

– if $p>1$, $\lim_{n\to+\infty} \int_A f_n \, dm = \int_A f \, dm$ for all $A \in \mathscr{F}$ with $m(A) < \infty$.

Hint. Use the Banach–Steinhaus Theorem (Exercise 6 on page 22), Proposition 4.3 on page 19, and the remark on page 146.

c. Suppose that $1 < p \leq \infty$. Show that a sequence (f_n) in ℓ^p converges weakly to $f \in \ell^p$ if and only if it is bounded and

$$\lim_{n\to+\infty} f_n(i) = f(i) \quad \text{for all } i \in \mathbb{N}.$$

d. *Schur's Lemma.* Show that a sequence in ℓ^1 converges weakly if and only if it converges strongly (to the same limit).

Hint. Suppose otherwise.

i. Show that there exists a sequence (f_n) of elements of ℓ^1 of unit norm that converges weakly to 0 and thus, in particular, such that $f_n(i) \to 0$ for every $i \in \mathbb{N}$.

ii. Construct by induction two strictly increasing sequences of integers (I_j) and (n_j) such that, for every integer j,

$$\sum_{i=0}^{I_{j-1}} \left|f_{n_j}(i)\right| \leq \tfrac{1}{5} \quad \text{and} \quad \sum_{i=I_j+1}^{+\infty} \left|f_{n_j}(i)\right| \leq \tfrac{1}{5}.$$

iii. Let $h : \mathbb{N} \to \mathbb{K}$ satisfy the following properties: If i is such that $I_{j-1} < i \leq I_j$, then $\left|h(i)\right| = 1$ and $f_{n_j}(i)h(i) = \left|f_{n_j}(i)\right|$. Show that, for every integer j,

$$\sum_{i=I_{j-1}+1}^{I_j} f_{n_j}(i)h(i) \geq \tfrac{3}{5},$$

and deduce that

$$\left|\sum_{i=0}^{+\infty} f_{n_j}(i)h(i)\right| \geq \tfrac{1}{5}.$$

iv. Deduce that the sequence (f_{n_j}) does not converge weakly to 0. Finish the proof.

e. Suppose that m is Lebesgue measure on the Borel σ-algebra of $\mathbb{R}^d$ and that $1 < p \leq \infty$. Let $f \in L^p$ vanish outside the unit ball of $\mathbb{R}^d$ and have norm 1 in L^p. For each $n \in \mathbb{N}$, set $f_n(x) = n^{d/p} f(nx)$. Show that the sequence (f_n) is a sequence of norm 1 in L^p that converges almost everywhere and weakly (but not strongly) to 0 in L^p.

f. Suppose m is Lebesgue measure on the interval $(0, 1)$. Show that the sequence (f_n) defined by $f_n(x) = e^{2i\pi nx}$ converges weakly (but not strongly) to 0 in every L^p, for $1 \leq p \leq \infty$, and that it does not converge almost everywhere.

Hint. You might start with the case $p=2$ (see Exercise 1 on page 129).

10. *Weak convergence in L^p spaces, continued.* Let $p \in (1, \infty]$ and p' be conjugate exponents.

a. Suppose $L^{p'}(m)$ is separable (or, which is the same, that $L^1(m)$ is separable: see Exercise 13b on page 154). Show that every bounded sequence in $L^p(m)$ has a weakly convergent subsequence.

Hint. Argue as in the first part of the proof of Theorem 3.7 on page 115.

b. Let (f_n) be a bounded sequence in $L^p(m)$.

i. Show that there exists a σ-algebra $\mathscr{F}'$ that is separable (in the sense of Exercise 13 on page 153), contained in $\mathscr{F}$, and satisfies these properties:

- For every $n \in \mathbb{N}$, f_n has a $\mathscr{F}'$-measurable representative.
- The restriction m' of the measure m to $\mathscr{F}'$ is σ-finite.

ii. Prove that, for every $g \in L^{p'}(m)$, there exists an element $g' \in L^{p'}(m')$ such that

$$\int fg\, dm = \int fg'\, dm' \quad \text{for all } f \in L^p(m').$$

Hint. Use the operator $T_{p'}$ defined in Exercise 7 on page 165.

iii. Show that the sequence (f_n) has a weakly convergent subsequence in $L^p(m)$.

Hint. By Exercise 13 on page 153, the space $L^{p'}(m')$ is separable.

11. Let $(c_n)_{n\in\mathbb{Z}}$ be a sequence of complex numbers. Define functions S_n and K_m by setting

$$S_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}, \qquad K_m(x) = \frac{1}{m}\sum_{n=0}^{m-1} S_n(x).$$

Show that, if the sequence (K_n) is bounded in $L^p((-\pi, \pi))$, with $1 < p \le \infty$, there exists an element f of $L^p((-\pi, \pi))$ such that

$$c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\, dx \quad \text{for all } n \in \mathbb{Z}.$$

Hint. Extract from the sequence (K_n) a subsequence that converges weakly in $L^p((-\pi, \pi))$ (see Exercise 10). The weak limit of this subsequence can be used for f.

12. We assume that m is a Radon measure and use the notation and definitions of Exercise 19 on page 159. Fix $p \in [1, \infty)$ and denote by p' the conjugate exponent of p.

a. For $g \in L_c^{p'}$, denote by T_g the linear form on L_{loc}^p defined by

$$T_g f = \int fg\, dm.$$

Show that this defines a linear isomorphism between $L_c^{p'}$ and the space $(L_{\mathrm{loc}}^p)'$ of continuous linear forms on L_{loc}^p (with the metric d).
Hint. To prove surjectivity, considerer a continuous linear form T on L_{loc}^p. Show that there exists $g \in L^{p'}$ such that $Tf = \int fg\,dm$ for every $f \in L^p$. Then show that the support of g is compact, and finish the proof.

b. A linear form T on L_c^p is said to be continuous if, for every compact K in X, the restriction of T to the space $\{f \in L^p : \operatorname{Supp} f \subset K\}$ with the norm $\|\cdot\|_p$ is continuous. We denote by $(L_c^p)'$ the set of continuous linear forms on L_c^p. If $g \in L_{\mathrm{loc}}^{p'}$, we denote by T_g the linear form on L_c^p defined by $T_g f = \int fg\,dm$. Show that this defines a linear isomorphism between $L_{\mathrm{loc}}^{p'}$ and $(L_c^p)'$.
Hint. Take $T \in (L_c^p)'$. Show that, for every compact K, there exists a unique $g_K \in L_c^{p'}$, supported within K and such that

$$T(1_K f) = \int 1_K f g_K\,dm \quad \text{for all } f \in L^p.$$

Then show that you can define $g \in L_{\mathrm{loc}}^{p'}$ by setting $1_K g = g_K$ for all K compact. Wrap up.

13. Assume that $m(X) < +\infty$ and that there exists a sequence $(A_n)_{n\in\mathbb{N}}$ of measurable subsets of X such that the sequence $(1_{A_n})_{n\in\mathbb{N}}$ is fundamental in $L^1(m)$ (see Exercise 13b on page 154). Show that the expression

$$|f| = \sum_{n\in\mathbb{N}} 2^{-n} \left| \int_{A_n} f\,dm \right|$$

defines a norm on $L^\infty(m)$ and that the subsets of $L^\infty(m)$ bounded with respect to the norm $\|\cdot\|_\infty$ are relatively compact with respect to $|\cdot|$. (Use Exercises 9 and 10.) Show that the space $\big(L^\infty(m), |\cdot|\big)$ is complete if and only if it has finite dimension.
Hint. Use Exercise 4 on page 54.

3 Convolution

Notation. In this section, the measure space $(X, \mathscr{F})$ under study will be the space $X = \mathbb{R}^d$ with its Borel σ-algebra $\mathscr{F} = \mathscr{B}(X)$, and the measure will be Lebesgue measure $m = \lambda = dx_1 \ldots dx_d$.

If f is a function from $\mathbb{R}^d$ to $\mathbb{K}$, we denote by $\check{f}$ the function on $\mathbb{R}^d$ defined by $x \mapsto f(-x)$; moreover, if $a \in \mathbb{R}^d$, we set $\tau_a f(x) = f(x-a)$. The function $\tau_a f$ thus defined is called the **translate** of f by a. The maps $f \mapsto \check{f}$ and $f \mapsto \tau_a f$ are linear and preserve measurability. Since Lebesgue measure is invariant under symmetries and translations, these operations are also defined on equivalence classes of functions modulo sets of Lebesgue measure zero.

If f is a function (or equivalence class of functions) and $a, b \in \mathbb{R}^d$, we clearly have

$$\tau_a(\tau_b f) = \tau_{a+b} f, \qquad \tau_a \check{f} = (\tau_{-a} f)\check{}, \qquad \tau_0 f = f.$$

Proposition 3.1 *If $1 \leq p \leq \infty$, the family $(\tau_a)_{a \in \mathbb{R}^d}$ forms an abelian group of isometries of L^p.*

If $1 \leq p < \infty$ and $f \in L^p$, the map Φ_f from $\mathbb{R}^d$ to L^p defined by $\Phi_f : a \mapsto \tau_a f$ is uniformly continuous.

Proof. The first assertion follows immediately from the remarks preceding the theorem (in particular, from the translation invariance of λ).

To prove the second assertion, since $\|\tau_a f - \tau_b f\|_p = \|\tau_{a-b} f - f\|_p$, it is enough to show that Φ_f is continuous at 0. Suppose first that $f \in C_c(\mathbb{R}^d)$. Then f is uniformly continuous on $\mathbb{R}^d$ and so, if $\varepsilon > 0$, there exists $\eta > 0$ such that $|y - y'| < \eta$ implies $\bigl|f(y) - f(y')\bigr| < \varepsilon$. Hence, if $|a| < \eta$,

$$\begin{aligned}\|\tau_a f - f\|_p &= \Bigl(\int |f(x-a) - f(x)|^p \, dx\Bigr)^{1/p} \\ &\leq \varepsilon\bigl(\lambda(a + \operatorname{Supp} f) + \lambda(\operatorname{Supp} f)\bigr)^{1/p};\end{aligned}$$

that is to say,

$$\|\tau_a f - f\|_p \leq \varepsilon\bigl(2\lambda(\operatorname{Supp} f)\bigr)^{1/p},$$

showing that Φ_f is continuous at 0 in this case.

Now, if f is any element of L^p, take a sequence (f_n) in $C_c(\mathbb{R}^d)$ converging to f in L^p (see Theorem 1.6 on page 146). The continuity of Φ_f at 0 then follows from the fact that the functions Φ_{f_n} converge uniformly to Φ_f (since $\|\Phi_{f_n}(a) - \Phi_f(a)\|_p = \|f_n - f\|_p$). ☐

When $f \in L^\infty$, the map $a \mapsto \tau_a f$ from $\mathbb{R}^d$ to L^∞ is continuous if and only if f has a uniformly continuous representative; see Exercise 6 below.

Let $p, p' \in [1, \infty]$ be conjugate exponents. If $f \in L^p$ and $g \in L^{p'}$, the **convolution** of f and g is, by definition, the function $f * g$ on $\mathbb{R}^d$ defined by

$$(f * g)(x) = \int f(x-y) g(y) \, dy.$$

For $x \in \mathbb{R}^d$, the function in the integrand is indeed integrable, being the product of $\tau_x \check{f} \in L^p$ and $g \in L^{p'}$. Thus $f * g$ is well-defined as a function on $\mathbb{R}^d$. Using the invariance of Lebesgue measure under translations and symmetries, one checks easily that

$$f * g = g * f.$$

Proposition 3.2 *Let $p, p' \in [1, \infty]$ be conjugate exponents and suppose $f \in L^p$ and $g \in L^{p'}$. Then $f * g$ is uniformly continuous and bounded, and*

$$\|f * g\|_\infty \leq \|f\|_p \|g\|_{p'}.$$

Moreover, if $1 < p < \infty$, we have $\lim_{|x| \to +\infty} f * g(x) = 0$; the same is true if $p = 1$ and g has compact support.

Proof. The Hölder inequality yields

$$\left|(f * g)(x) - (f * g)(x')\right| \leq \|\tau_x \check{f} - \tau_{x'} \check{f}\|_p \|g\|_{p'} \quad \text{for all } x, x' \in \mathbb{R}^d;$$

the uniform continuity of $f * g$ if $p < \infty$ follows because of Proposition 3.1. If $p = \infty$, we have $p' = 1$ and the property remains true since $f * g = g * f$.

We also have $\|f * g\|_\infty \leq \|f\|_p \|g\|_{p'}$, by the Hölder inequality and the fact that $\|\tau_x \check{f}\|_p = \|f\|_p$ for every x. This implies, in particular, that the bilinear map $(f, g) \mapsto f * g$ is continuous as a map from $L^p \times L^{p'}$ to $C_b(\mathbb{R}^d)$ with the uniform norm. Suppose that $f \in C_c(\mathbb{R}^d)$ and that $g \in L^\infty$ has compact support. We claim that

$$\operatorname{Supp}(f * g) \subset \operatorname{Supp} f + \operatorname{Supp} g;$$

indeed, $\operatorname{Supp} f + \operatorname{Supp} g$ is compact and for $x \notin \operatorname{Supp} f + \operatorname{Supp} g$ we have $\operatorname{Supp}(\tau_x \check{f}) \cap \operatorname{Supp} g = (x - \operatorname{Supp} f) \cap \operatorname{Supp} g = \varnothing$, so $(f * g)(x) = 0$. Since $\operatorname{Supp} f + \operatorname{Supp} g$ is compact, we conclude that $f * g \in C_c(\mathbb{R}^d)$. The last claim of the proposition follows, because $C_c(\mathbb{R}^d)$ is dense in L^p for $1 \leq p < \infty$ and because the uniform limit of a sequence of continuous functions with compact support tends to 0 at infinity. □

We will now extend the definition of the convolution product. Let f and g be (equivalence classes of) Borel functions. We say that f and g are **convolvable** if, for almost every $x \in \mathbb{R}^d$, the product $(\tau_x \check{f})g$ lies in L^1. If f and g are convolvable, the **convolution** of f and g is, by definition, the equivalence class of functions $f * g$ defined almost everywhere by

$$(f * g)(x) = \int f(x - y) g(y) \, dy.$$

Clearly, f and g are convolvable if and only if g and f are, and in this case $f * g = g * f$.

By reasoning as in the proof of Proposition 3.2, we obtain the following property.

Proposition 3.3 *If f and g are convolvable equivalence classes of functions,*

$$\operatorname{Supp}(f * g) \subset \overline{\operatorname{Supp} f + \operatorname{Supp} g}.$$

In particular, if f or g has compact support, we have

$$\operatorname{Supp}(f * g) \subset \operatorname{Supp} f + \operatorname{Supp} g$$

(since, if F is closed and K is compact, $F + K$ is closed). Thus, the convolution of two classes of functions with compact support has compact support.

The next theorem presents a sufficient criterion for the existence of the convolution. As usual, we set $1/\infty = 0$.

Theorem 3.4 (Young's inequality) *Suppose that $p, q \in [1, \infty]$ satisfy $1/p + 1/q \geq 1$, and let r be defined by $1/r = 1/p + 1/q - 1$. If $f \in L^p$ and $g \in L^q$, then f and g are convolvable, $f * g \in L^r$, and*

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Note that this applies, in particular, to $r = p = q = 1$.

Proof

1. We can assume that $r < \infty$, since $r = \infty$ corresponds to the case $q = p'$ treated in Proposition 3.2. Moreover, $r < \infty$ implies $p, q < \infty$ (if $p = \infty$, for example, then $q = 1$ and $r = \infty$). We can also assume that $f \geq 0$ and $g \geq 0$, by substituting $|f|$ and $|g|$ for f and g.
2. Consider first the case where $p = 1$, $1 \leq q < \infty$, and $r = q$. By applying the Hölder inequality to the measure $m = f\lambda$, we get

$$\int g(x-y) f(y)\, dy \leq \left(\int g^q(x-y) f(y)\, dy\right)^{1/q} \left(\int f(y)\, dy\right)^{1-1/q}$$

and

$$\int \left(\int g(x-y) f(y)\, dy\right)^q dx \leq \left(\iint g^q(x-y) f(y)\, dy\, dx\right) \left(\int f(y)\, dy\right)^{q-1}.$$

By Fubini's Theorem and the translation invariance of Lebesgue measure, the right-hand side of this inequality equals $\|g\|_q^q \|f\|_1^q$. We deduce that g and f are convolvable, that $g * f \in L^q$, and that $\|g * f\|_q \leq \|g\|_q \|f\|_1$. The case where $q = 1$, $1 \leq p < \infty$, and $r = p$ is analogous.
3. Finally, take the case $1 < p, q < \infty$, so that $\max(p, q) < r < \infty$. We continue to suppose, without loss of generality, that $f, g \geq 0$. Then

$$f(x-y) g(y) = f^{p/r}(x-y) g^{q/r}(y) f^{1-p/r}(x-y) g^{1-q/r}(y).$$

Using the Hölder inequality with the conjugate exponents r and $r' = r/(r-1)$, we obtain

$$\int f(x-y) g(y)\, dy$$
$$\leq \left(\int f^p(x-y) g^q(y)\, dy\right)^{1/r} \left(\int f^{\frac{r-p}{r-1}}(x-y) g^{\frac{r-q}{r-1}}(y)\, dy\right)^{1-1/r}.$$

In the second integral on the right-hand side, we use Hölder's inequality with the conjugate exponents $p(r-1)/(r-p)$ and $q(r-1)/(r-q)$ (to check conjugacy use the relation $1/r = 1/p + 1/q - 1$). We obtain

$$\int f^{\frac{r-p}{r-1}}(x-y)\,g^{\frac{r-q}{r-1}}(y)\,dy \le \left(\int f^p(y)\,dy\right)^{\frac{r-p}{p(r-1)}} \left(\int g^q(y)\,dy\right)^{\frac{r-q}{q(r-1)}},$$

which finally leads to

$$\int\left(\int f(x-y)g(y)\,dy\right)^r dx \le \left(\iint f^p(x-y)g^q(y)\,dx\,dy\right) \|f\|_p^{r-p}\,\|g\|_q^{r-q}.$$

The double integral in this expression equals $\|f\|_p^p\,\|g\|_q^q$, once more by Fubini's Theorem. We deduce that f and g are convolvable, that $f*g \in L^r$, and that $\|f*g\|_r \le \|f\|_p\,\|g\|_q$. □

Proposition 3.5 *Let $p,q,r \in [1,+\infty]$ be such that $1/p+1/q+1/r \ge 2$. If $f \in L^p$, $g \in L^q$, and $h \in L^r$, then $f*(g*h)$ and $(f*g)*h$ are well defined and belong to L^s, where s is given by $1/s = 1/p+1/q+1/r-2$. In addition,*

$$f*(g*h) = (f*g)*h.$$

Proof. That $f*(g*h)$ and $(f*g)*h$ are well defined and belong to L^s follows from Theorem 3.4. Next,

$$\begin{aligned}\bigl(f*(g*h)\bigr)(x) &= \iint f(x-y)g(y-z)h(z)\,dy\,dz \\ &= \iint f(x-y-z)g(y)h(z)\,dy\,dz = \bigl((f*g)*h\bigr)(x),\end{aligned}$$

which concludes the proof. (As an exercise, the reader might justify these formal calculations, especially the use of Fubini's Theorem.) □

Corollary 3.6 *The operations $+$ and $*$ make L^1 into a commutative ring.*

Proof. The convolution product is commutative and, by Theorem 3.4, L^1 is closed under it. Proposition 3.5 says it is also associative. The rest is obvious. □

In addition, L^1 is a Banach space and $*$ is a bilinear map from $L^1 \times L^1$ to L^1 such that

$$\|f*g\|_1 \le \|f\|_1\,\|g\|_1 \quad \text{for all } f,g \in L^1.$$

We say that the convolution product makes L^1 into a **commutative Banach algebra**.

Approximations of Unity

The ring $(L^1, +, *)$ has no unity (see Exercise 1 below). However, there are entities that behave under convolution approximately like unity, in a sense we now make precise.

By definition, an **approximation of unity** or **Dirac sequence** is any sequence $(\varphi_n)_{n\in\mathbb{N}}$ in L^1 satisfying these properties:

- For every $n \in \mathbb{N}$, we have $\varphi_n \geq 0$ and $\int \varphi_n(x)\,dx = 1$.
- For every $\varepsilon > 0$,

$$\lim_{n\to+\infty} \int_{\{|x|>\varepsilon\}} \varphi_n(x)\,dx = 0.$$

For example, one can start from any nonnegative-valued function $\varphi \in L^1$ such that $\int \varphi(x)\,dx = 1$, and set, for $n \geq 1$,

$$\varphi_n(x) = n^d \varphi(nx).$$

A change of variables shows that $\int \varphi_n(x)\,dx = 1$; moreover,

$$\int_{\{|x|>\varepsilon\}} \varphi_n(x)\,dx = \int_{\{|x|>n\varepsilon\}} \varphi(x)\,dx,$$

and this last expression tends to 0 as n tends to infinity, by the Dominated Convergence Theorem. (See also Exercise 2 on page 36.) If, in addition, φ is continuous and supported within $\bar{B}(0,1)$, the sequence (φ_n) constructed in this way is called a **normal Dirac sequence**.

The alternative name "approximation of unity" for Dirac sequences is explained by the next proposition.

Proposition 3.7 *Suppose $p \in [1,\infty)$ and let $(\varphi_n)_{n\in\mathbb{N}}$ be a Dirac sequence. If $f \in L^p$, then*

$$f * \varphi_n \in L^p \quad \text{and} \quad \|f * \varphi_n\|_p \leq \|f\|_p \quad \text{for every } n \in \mathbb{N},$$

and

$$\lim_{n\to+\infty} f * \varphi_n = f \quad \text{in } L^p.$$

Proof. That $f * \varphi_n \in L^p$ and $\|f * \varphi_n\|_p \leq \|f\|_p$ follows from Theorem 3.4. Further, for almost every x,

$$\begin{aligned}\big|f(x) - (f * \varphi_n)(x)\big| &\leq \int \big|f(x) - f(x-y)\big|\,\varphi_n(y)\,dy \\ &\leq \left(\int \big|f(x) - f(x-y)\big|^p \varphi_n(y)\,dy\right)^{1/p},\end{aligned}$$

the latter inequality being a consequence of Hölder's inequality applied to the measure $\varphi_n(y)\,dy$. We deduce that

$$\|f - f * \varphi_n\|_p^p \leq \int \|f - \tau_y f\|_p^p \varphi_n(y)\,dy.$$

Now, for every $\varepsilon > 0$, we can write

$$\int \|f - \tau_y f\|_p^p \varphi_n(y)\, dy \le \sup_{|y|<\varepsilon} \|f - \tau_y f\|_p^p + \left(2\|f\|_p\right)^p \int_{\{|y|>\varepsilon\}} \varphi_n(y)\, dy,$$

by breaking $\mathbb{R}^d$ into the disjoint union of $\{|y| \le \varepsilon\}$ and $\{|y| > \varepsilon\}$. It follows that

$$\limsup_{n\to+\infty} \|f - f * \varphi_n\|_p \le \sup_{|y|\le\varepsilon} \|f - \tau_y f\|_p.$$

Now it suffices to apply Proposition 3.1. □

Remark. If we assume in addition that, for every $n \in \mathbb{N}$, the function φ_n lies in L^∞ and has compact support, Proposition 3.2 implies that $f * \varphi_n \in C_0(\mathbb{R}^d)$ for every $n \in \mathbb{N}$. This happens, in particular, when (φ_n) is a normal Dirac sequence. In this particular case, we see from the preceding calculations that, for any $p \in [1, \infty)$, any $f \in L^p$, and any $n \in \mathbb{N}$,

$$\|f - f * \varphi_n\|_p \le \sup_{|y|\le 1/n} \|f - \tau_y f\|_p.$$

This will lead to a criterion of relative compactness in L^p.

Relative Compactness in L^p

Theorem 3.8 *Suppose $p \in [1, \infty)$ and let H be a subset of L^p. In order for H to be relatively compact in L^p, it is necessary and sufficient that the following three properties be satisfied:*

- **i.** *H is bounded in L^p.*
- **ii.** $\displaystyle\lim_{R\to+\infty} \int_{\{|x|>R\}} |f(x)|^p\, dx = 0$ *uniformly with respect to $f \in H$.*
- **iii.** $\lim_{a\to 0} \tau_a f = f$ *in L^p, uniformly with respect to $f \in H$.*

Proof. Since L^p is complete, H is relatively compact if and only if it is precompact (Theorem 3.3 on page 14).

Suppose H is precompact. Take $\varepsilon > 0$ and let $f_1, \dots, f_k$ be elements of L^p such that the balls $B(f_1, \varepsilon), \dots, B(f_k, \varepsilon)$ cover H. In particular, property i of the theorem is satisfied. By the Dominated Convergence Theorem, there exists $R_0 > 0$ such that, for any $R \ge R_0$ and any $j \in \{1, \dots, k\}$,

$$\left(\int_{\{|x|>R\}} |f_j(x)|^p\, dx\right)^{1/p} < \varepsilon;$$

thus, for any $R \ge R_0$ and any $f \in H$,

$$\left(\int_{\{|x|>R\}} |f(x)|^p\, dx\right)^{1/p} < 2\varepsilon.$$

Similarly, by Proposition 3.1, there exists $\eta > 0$ such that, for any a with $|a| < \eta$ and any $j \in \{1, \ldots, k\}$,

$$\|\tau_a f_j - f_j\|_p < \varepsilon,$$

and so, for any a with $|a| < \eta$ and any $f \in H$,

$$\|\tau_a f - f\|_p < 3\varepsilon.$$

Thus, if H is precompact, properties i–iii of the theorem are satisfied.

Suppose, conversely, that those three properties are satisfied, and fix $\varepsilon > 0$. By property ii, there exists $R > 0$ such that

$$\left(\int_{\{|x|>R\}} |f(x)|^p \, dx\right)^{1/p} < \varepsilon \quad \text{for all } f \in H.$$

Let (φ_n) be a normal Dirac sequence. As we saw in the remark preceding the theorem, we have, for any $n \geq 1$ and any $f \in L^p$,

$$\|f - f * \varphi_n\|_p \leq \sup_{|y| \leq 1/n} \|f - \tau_y f\|_p.$$

Hence, by property iii, there exists an integer $N \in \mathbb{N}$ such that

$$\|f - f * \varphi_N\|_p < \varepsilon \quad \text{for all } f \in H.$$

Now, by Hölder's inequality, for any $x, x' \in \mathbb{R}^d$ we have

$$\bigl|(f * \varphi_N)(x) - (f * \varphi_N)(x')\bigr| \leq \|\tau_x \check{f} - \tau_{x'} \check{f}\|_p \, \|\varphi_N\|_{p'} \quad \text{for all } f \in L^p,$$

where p' is the conjugate exponent of p; whereas the invariance properties of the Lebesgue measure imply that

$$\|\tau_x \check{f} - \tau_{x'} \check{f}\|_p = \|\tau_{x-x'} f - f\|_p.$$

Thus, for every $f \in H$ and every $x, x' \in \mathbb{R}^d$,

$$\bigl|(f * \varphi_N)(x) - (f * \varphi_N)(x')\bigr| \leq \|\tau_{x-x'} f - f\|_p \, \|\varphi_N\|_{p'}$$

and

$$\bigl|(f * \varphi_N)(x)\bigr| \leq \|f\|_p \, \|\varphi_N\|_{p'}.$$

Then it follows from assumptions i and iii and from the Ascoli Theorem (page 44) that the subset of $C(\bar{B}(0, R))$ consisting of the restrictions to $\bar{B}(0, R)$ of the functions (continuous on $\mathbb{R}^d$) $f * \varphi_N$, with $f \in H$, is relatively compact and so precompact in $C(\bar{B}(0, R))$. Hence there exists a finite sequence $(f_1, \ldots, f_k)$ of elements of H such that, for every $f \in H$, there exists $j \in \{1, \ldots, k\}$ such that

$$\bigl|(f * \varphi_N)(x) - (f_j * \varphi_N)(x)\bigr| \leq \varepsilon \lambda(\bar{B}(0, R))^{-1/p} \quad \text{for all } x \in \bar{B}(0, R),$$

and so

$$\|f - f_j\|_p \leq \left(\int_{\{|x|>R\}} |f(x)|^p \, dx\right)^{1/p} + \left(\int_{\{|x|>R\}} |f_j(x)|^p \, dx\right)^{1/p}$$
$$+ \|f - f * \varphi_N\|_p + \|f_j - f_j * \varphi_N\|_p$$
$$+ \lambda(\bar{B}(0,R))^{1/p} \sup_{x \in \bar{B}(0,R)} \bigl|(f * \varphi_N)(x) - (f_j * \varphi_N)(x)\bigr|,$$

this last result being obtained via the triangle inequality starting from

$$|f - f_j| \leq 1_{\{|x|>R\}} |f| + 1_{\{|x|>R\}} |f_j|$$
$$+ |f - f * \varphi_N| + |f_j - f_j * \varphi_N| + 1_{\{|x| \leq R\}} |f * \varphi_N - f_j * \varphi_N|.$$

Pulling everything together we obtain $\|f - f_j\|_p \leq 5\varepsilon$, which shows that H is precompact. □

Exercises

1. **a.** Let $(\varphi_n)_{n \in \mathbb{N}}$ be a normal Dirac sequence. Show that (φ_n) converges almost everywhere to 0. Deduce that it does not converge in L^1.
   b. Deduce that the algebra L^1 does not have a unity; that is, there is no element g of L^1 such that $f * g = f$ for all $f \in L^1$.
   c. More generally, show that, if $p \in [1, \infty]$, there is no element g of L^1 such that $f * g = f$ for all $f \in L^p$.
2. *Hardy's inequality.* Let $p \in (1, \infty)$ and p' be conjugate exponents. If f is a function or equivalence class of functions on $(0, +\infty)$, define $\hat{f}$ on $\mathbb{R}$ by
$$\hat{f}(x) = e^{x/p} f(e^x).$$
   Finally, if $f \in L^p((0, +\infty))$, define
$$Tf(x) = \frac{1}{x} \int_0^x f(t) \, dt \quad \text{for } x > 0.$$
   a. Show that $f \in L^p((0, +\infty))$ if and only if $\hat{f} \in L^p(\mathbb{R})$ and that, in this case, $\|f\|_{L^p((0,+\infty))} = \|\hat{f}\|_{L^p(\mathbb{R})}$.
   b. Let g be the function defined on $\mathbb{R}$ by
$$g(x) = e^{-x/p'} 1_{[0,+\infty)}(x).$$
   Show that $g \in L^1(\mathbb{R})$ and that, if $f \in L^p((0, +\infty))$, we have $\widehat{Tf} = \hat{f} * g$. Deduce that T is a continuous linear operator from $L^p((0, +\infty))$ to itself, of norm at most p'.

c. For $n \in \mathbb{N}$, let f_n be the function defined on $(0, +\infty)$ by

$$f_n(t) = (nt)^{-1/p} 1_{[1,e^n]}(t).$$

Show that, for every $n \in \mathbb{N}$, we have $f_n \in L^p((0, +\infty))$ and

$$\|f_n\|_{L^p((0,+\infty))} = 1.$$

Show also that $\lim_{n\to+\infty} \|Tf_n\|_{L^p((0,+\infty))} = p'$. Deduce that T has norm p'. Show likewise that, for all $x > 0$, $\lim_{n\to+\infty} Tf_n(x) = 0$. (See also Exercise 17 on page 228.)

3. *The convolution product in $\ell^p(\mathbb{Z})$.* We say that two functions f and g from $\mathbb{Z}$ to $\mathbb{C}$ are *convolvable* if

$$\sum_{k\in\mathbb{Z}} |f(n-k)|\,|g(k)| < +\infty \quad \text{for all } n \in \mathbb{Z}.$$

If this is the case, the *convolution* $f * g$ is defined by

$$(f * g)(n) = \sum_{k\in\mathbb{Z}} f(n-k)g(k) \quad \text{for all } n \in \mathbb{Z}.$$

a. Show that f and g are convolvable if and only if g and f are, and that in this case $f * g = g * f$.

b. Let $p, q \in [1, \infty]$ be such that $1/p + 1/q \geq 1$, and suppose $f \in \ell^p(\mathbb{Z})$ and $g \in \ell^q(\mathbb{Z})$. Show that f and g are convolvable, that $f*g \in \ell^r(\mathbb{Z})$, where $1/r = 1/p + 1/q - 1$, and that $\|f * g\|_r \leq \|f\|_p \|g\|_q$.

c. Show that the normed space $\ell^1(\mathbb{Z})$ with the operation $*$ is a commutative Banach algebra with unity.

d. **i.** For $m \in \mathbb{Z}$, we denote by δ_m the function on $\mathbb{Z}$ defined by $\delta_m(n) = 1$ if $n = m$ and $\delta_m(n) = 0$ otherwise. Show that, for $m, p \in \mathbb{Z}$, $\delta_m * \delta_p = \delta_{m+p}$.

ii. Let $\mathscr{M}$ be the set of continuous linear forms Φ on $\ell^1(\mathbb{Z})$ that are not identically zero and satisfy

$$\Phi(f * g) = \Phi(f)\Phi(g) \quad \text{for all } f, g \in \ell^1(\mathbb{Z}).$$

Let $\mathbb{U}$ be the set of complex numbers of absolute value 1. If $u \in \mathbb{U}$, prove that the linear form defined on $\ell^1(\mathbb{Z})$ by

$$\Phi_u(f) = \sum_{k\in\mathbb{Z}} u^k f(k)$$

belongs to $\mathscr{M}$.

iii. Show that the map $u \mapsto \Phi_u$ thus defined is a bijection between $\mathbb{U}$ and $\mathscr{M}$.

Hint. If $\Phi \in \mathscr{M}$, there exists $\varphi \in \ell^\infty(\mathbb{Z})$ such that

$$\Phi(f) = \sum_{k \in \mathbb{Z}} f(k)\varphi(k) \quad \text{for all } f \in \ell^1(\mathbb{Z}).$$

Now show, using part i, that $\varphi(n+m) = \varphi(n)\varphi(m)$ for every $n, m \in \mathbb{Z}$; deduce that φ is of the form $\varphi(n) = u^n$, with $u \in \mathbb{U}$.

4. We denote by $\cdot$ the scalar product on $\mathbb{R}^d$.

a. *Riemann–Lebesgue Lemma.* Show that, if $f \in L^1$,

$$\lim_{|\xi| \to \infty} \int e^{ix\cdot\xi} f(x)\,dx = 0.$$

Hint. Show that, if $\xi \neq 0$,

$$F(f,\xi) = \int e^{ix\cdot\xi} f(x)\,dx = -F\big(\tau_{\pi\xi/|\xi|^2} f, \xi\big).$$

Deduce that $|2F(f,\xi)| \leq \|f - \tau_{\pi\xi/|\xi|^2} f\|_1$.

b. For $f \in L^1$, we define a map $\hat{f}$ by

$$\hat{f}(\xi) = \int e^{ix\cdot\xi} f(x)\,dx \quad \text{for all } \xi \in \mathbb{R}^d.$$

Prove that $\hat{f} \in C_0(\mathbb{R}^d)$ and that the uniform norm of $\hat{f}$ is at most $\|f\|_1$.

c. Show that the map $\Phi : L^1 \to C_0(\mathbb{R}^d)$ defined by $\Phi(f) = \hat{f}$ is a continuous linear map and that

$$\Phi(f * g) = \Phi(f)\Phi(g) \quad \text{for all } f, g \in L^1.$$

The map Φ is called a *morphism of Banach algebras* from $(L^1, *)$ to $C_0(\mathbb{R}^d)$ (where the latter space is considered with its ordinary multiplication).

5. *The spectrum of the algebra L^1.* The goal of this exercise is to characterize the spectrum of the algebra L^1, that is, the set $\mathscr{M}$ of nonzero continuous linear forms Φ on L^1 such that

$$\Phi(f * g) = \Phi(f)\Phi(g) \quad \text{for all } f, g \in L^1.$$

Once more we denote by $\cdot$ the scalar product on $\mathbb{R}^d$.

a. Show that, for every $\xi \in \mathbb{R}^d$, the linear form Φ_ξ defined by

$$\Phi_\xi(f) = \int e^{i\xi\cdot x} f(x)\,dx \quad \text{for all } f \in L^1$$

belongs to $\mathscr{M}$.

b. Let φ be a bounded continuous function from $\mathbb{R}^d$ to $\mathbb{C}$, not identically zero and such that

$$\varphi(s+t) = \varphi(s)\varphi(t) \quad \text{for all } s, t \in \mathbb{R}^d.$$

i. Show that $\varphi(0) = 1$.

ii. Show that, for every $\varepsilon > 0$,

$$\int_{t_1}^{t_1+\varepsilon} \cdots \int_{t_d}^{t_d+\varepsilon} \varphi(s)\, ds = \left(\int_{[0,\varepsilon]^d} \varphi(s)\, ds \right) \varphi(t) \quad \text{for all } t \in \mathbb{R}^d.$$

Deduce that φ is of class C^1, and then that

$$\frac{\partial \varphi}{\partial t_j}(t) = \frac{\partial \varphi}{\partial t_j}(0)\varphi(t) \quad \text{for all } j \in \{1, \ldots, d\} \text{ and } t \in \mathbb{R}^d.$$

iii. Deduce that there exists $\xi \in \mathbb{R}^d$ such that $\varphi(t) = e^{i\xi \cdot t}$ for every $t \in \mathbb{R}^d$.

Hint. Set $a_j = (\partial\varphi/\partial t_j)(0)$. Show that the function $t \mapsto \varphi(t)e^{-a\cdot t}$ is constant.

c. Let Φ be an element of $\mathscr{M}$.

i. Show that there exists $\varphi \in L^\infty$ such that

$$\Phi(f) = \int f(x)\varphi(x)\, dx \quad \text{for all } f \in L^1.$$

ii. Show that, for every element f of L^1,

$$\Phi(\tau_a f) = \Phi(f)\varphi(a) \quad \text{for almost every } a \in \mathbb{R}^d.$$

Hint. Show that, for every $g \in L^1$,

$$\begin{aligned} \int \Phi(f)\varphi(a)g(a)\, da &= \Phi(f)\Phi(g) = \Phi(f * g) \\ &= \int \left(\int f(x-a)g(a)\, da \right) \varphi(x)\, dx \\ &= \int \Phi(\tau_a f) g(a)\, da. \end{aligned}$$

iii. Deduce that φ has a representative in $C_b(\mathbb{R}^d)$ (which we still denote by φ) satisfying

$$\Phi(\tau_a f) = \Phi(f)\varphi(a) \quad \text{for all } f \in L^1 \text{ and } a \in \mathbb{R}^d.$$

iv. Then show that

$$\varphi(a+b) = \varphi(a)\varphi(b) \quad \text{for all } a, b \in \mathbb{R}^d$$

and deduce that there exists $\xi \in \mathbb{R}^d$ such that $\varphi(t) = e^{i\xi\cdot t}$ for every $t \in \mathbb{R}^d$.

d. Show that the map $\xi \mapsto \Phi_\xi$ is a bijection between $\mathbb{R}^d$ and $\mathscr{M}$.

6. Suppose $f \in L^\infty$.

a. Show that, if f admits a uniformly continuous representative, the map $\mathbb{R}^d \to L^\infty$ given by $a \mapsto \tau_a f$ is continuous.

b. Conversely, suppose the map $a \mapsto \tau_a f$ from $\mathbb{R}^d$ to L^∞ is continuous.

i. Show that, for almost every x in $\mathbb{R}^d$,

$$\big|f(x) - f(x-y)\big| \le \|\tau_y f - f\|_\infty \quad \text{for almost every } y \in \mathbb{R}^d.$$

Hint. Use Fubini's Theorem.

ii. Let (φ_n) be a Dirac sequence. Show that

$$\|f - f * \varphi_n\|_\infty \le \int \|\tau_y f - f\|_\infty \varphi_n(y)\, dy.$$

Deduce that

$$\lim_{n\to\infty} \|f - f * \varphi_n\|_\infty = 0.$$

iii. Show that f has a uniformly continuous representative.

7. Let $(\varphi_n)_{n\in\mathbb{N}}$ be a normal Dirac sequence. Show that, for every continuous function f on $\mathbb{R}^d$, the sequence $(f*\varphi_n)_{n\in\mathbb{N}}$ converges to f uniformly on every compact of $\mathbb{R}^d$.

8. *Convolution semigroups.* Consider a family $(p_t)_{t\in\mathbb{R}^{+*}}$ of positive elements of L^1 satisfying these conditions:

- $\int p_t(x)\, dx = 1$ for all $t > 0$.
- $p_{t+s} = p_t * p_s$ for all $t, s > 0$.
- $\lim_{t\to 0} \int_{\{|x|>\varepsilon\}} p_t(x)\, dx = 0$ for all $\varepsilon > 0$.

Such a family will be called a *convolution semigroup* in the sequel.

a. Suppose $p \in [1, \infty)$. For every $f \in L^p$, set $P_t f = p_t * f$. Show the following facts:

i. For every $t > 0$, P_t is a continuous linear map of norm 1 from L^p to L^p.

ii. $P_t P_s = P_{t+s}$ for all $t, s > 0$.

iii. $\lim_{t\to 0} P_t f = f$ in L^p for all $f \in L^p$.

iv. For all $f \in L^p$, the map $t \mapsto P_t f$ from $\mathbb{R}^{+*}$ to L^p is continuous.

b. *The Gaussian semigroup.* Show that the family (p_t) defined by

$$p_t(x) = \frac{1}{(2\pi t)^{d/2}}\, e^{-|x|^2/2t}$$

satisfies the conditions for a convolution semigroup.
Hint. Recall that $\int_{-\infty}^{+\infty} e^{-x^2}\, dx = \sqrt{\pi}$. To prove that $p_t * p_s = p_{t+s}$, use the fact that Lebesgue measure is translation invariant.

c. *The Cauchy semigroup.* Now assume that $d = 1$. Show that the family (p_t) defined by

$$p_t(x) = \frac{1}{\pi} \frac{t}{t^2 + x^2}$$

satisfies the conditions for a convolution semigroup.
Hint. To show that $p_t * p_s = p_{t+s}$, start by checking that

$$\frac{1}{t^2+(x-y)^2} \frac{1}{s^2+y^2} = \frac{1}{(x^2+(t+s)^2)(x^2+(t-s)^2)} \times \left(\frac{2x(x-y)+(x^2+s^2-t^2)}{t^2+(x-y)^2} + \frac{2xy+(x^2+t^2-s^2)}{s^2+y^2} \right).$$

d. Suppose $p = \infty$. Show that properties i and ii are still satisfied, and that properties iii and iv are satisfied for $f \in L^\infty$ if and only if f has a uniformly continuous representative.

e. Show that the result of part a is still true if L^p is replaced everywhere by the space $C_0(\mathbb{R}^d)$ with the uniform norm, or by the space $C_{u,b}(\mathbb{R}^d)$ of uniformly continuous bounded functions with the uniform norm.

9. We adopt the definitions and notation of Exercise 19 on page 159, in the special case where m is Lebesgue measure on $\mathbb{R}^d$.

a. **i.** Suppose $p \in [1, \infty)$ and let H be a subset of L^p satisfying conditions i and iii of Theorem 3.8. Show that H is relatively compact in L^p_{loc} with the metric d.
Hint. Revisit the proof of Theorem 3.8.

ii. Let $p, q, r \in [1, +\infty)$ be such that $1/r = 1/p + 1/q - 1$. Show that, if $G \in L^p$, the set

$$\{G * f : f \in L^q \text{ and } \|f\|_q \le 1\}$$

is relatively compact in (L^r_{loc}, d).

b. Let $p, q, r \in [1, +\infty)$ be such that $1/r = 1/p+1/q-1$. Show that any function $f \in L^p_{\text{loc}}$ can be convolved with any $g \in L^q_c$, and that for such functions we have $f*g \in L^r_{\text{loc}}$ and $\operatorname{Supp}(f*g) \subset \operatorname{Supp} f + \operatorname{Supp} g$.

c. Show that, if $p, p' \in [1, \infty]$ are conjugate exponents, the convolution of a function $f \in L^p_{\text{loc}}$ and a function $g \in L^{p'}_c$ belongs to $C(\mathbb{R}^d)$.

d. Suppose $m \in \mathbb{N}^* \cup \{\infty\}$. Show that, if $f \in L^1_{\text{loc}}$ and g is a function of class C^m with compact support, $f * g$ is of class C^m and, for every $(p_1, \dots, p_d) \in \mathbb{N}^d$ such that $|p| = p_1 + \dots + p_d \le m$, we have

$$\frac{\partial^{|p|}}{\partial x_1^{p_1} \dots \partial x_d^{p_d}} (f * g) = f * \left(\frac{\partial^{|p|} g}{\partial x_1^{p_1} \dots \partial x_d^{p_d}} \right).$$

e. Show that this equation remains true if we assume that $f \in L^1_c$ and that g is of class C^m with arbitrary support.

10. *Compactness in $\ell^p(I)$, for $0 < p < \infty$.* Let I be a set.

a. Suppose $p \in [1, \infty)$. Show that a subset H in $\ell^p(I)$ is relatively compact if and only if it is bounded and there exists, for every $\varepsilon > 0$, a finite subset J of I such that

$$\|1_{I \setminus J} f\|_p < \varepsilon \quad \text{for all } f \in H.$$

(Compare with Theorem 3.8.)
Hint. Use Exercise 8 on page 17.

b. Suppose $p \in (0, 1)$. Consider the space $\ell^p(I)$ with the metric d_p defined in Exercise 1 on page 147. Show that the result of the preceding question remains valid if we replace $\|\cdot\|_p$ by $|\cdot|_p = d_p(\,\cdot\,, 0)$.
Hint. Use Exercise 1a on page 148 to adapt the method above.

Part II

OPERATORS

5

Spectra

1 Operators on Banach Spaces

We fix here a Banach space E over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and we wish to study the (noncommutative) Banach algebra $L(E)$ of continuous linear maps from E to E, the product operation being composition. We use the same notation $\|\cdot\|$ for the norm on E and the associated norm on $L(E)$, and we denote by I the identity map on E. Thus, I is the unity of the algebra $L(E)$. An element $T \in L(E)$ is called **invertible** if it has an inverse in $L(E)$; that is, if there exists a continuous linear map S such that $TS = ST = I$. Because composition is associative, T has an inverse in $L(E)$ if and only if it has a right inverse (an element U such that $TU = I$) and a left inverse (an element V such that $VT = I$) in $L(E)$. Clearly, if T is invertible, it is bijective and its inverse in $L(E)$ is unique and equals the inverse map T^{-1}. Thus, for $T \in L(E)$, the following properties are equivalent:

- T is invertible.
- T is bijective and T^{-1} is continuous.
- $\ker T = \{0\}$, $\operatorname{im} T = E$, and T^{-1} is continuous.

In fact, the map inverse to a bijective continuous linear operator from E onto E is always continuous; this follows directly from the **Open Mapping Theorem**, itself a consequence of Baire's Theorem (Exercise 6 on page 22). We will not make use of this result here.

Finally, we note that, if T and S are invertible elements of $L(E)$, the composition TS is also invertible and $(TS)^{-1} = S^{-1}T^{-1}$.

We make the convention that $T^0 = I$ for $T \in L(E)$.

Proposition 1.1 *The set $\mathscr{I}$ of invertible elements in $L(E)$ is an open subset of $L(E)$ containing I. The map $T \mapsto T^{-1}$ from $\mathscr{I}$ to $\mathscr{I}$ is continuous.*

More precisely, if $T_0 \in \mathscr{I}$ and $\|T - T_0\| < \|T_0^{-1}\|^{-1}$, then $T \in \mathscr{I}$ and

$$T^{-1} = \sum_{n=0}^{+\infty} (I - T_0^{-1}T)^n T_0^{-1} = \sum_{n=0}^{+\infty} T_0^{-1}(I - TT_0^{-1})^n.$$

Proof. Take $T_0 \in \mathscr{I}$.

1. First,

$$\|I - T_0^{-1}T\| = \|T_0^{-1}(T_0 - T)\| \le \|T_0^{-1}\| \, \|T - T_0\|$$

and

$$\|I - TT_0^{-1}\| = \|(T_0 - T)T_0^{-1}\| \le \|T - T_0\| \, \|T_0^{-1}\|.$$

Thus, if $\|T - T_0\| < \|T_0^{-1}\|^{-1}$, the series

$$\sum_{n=0}^{+\infty} (I - T_0^{-1}T)^n T_0^{-1} \quad \text{and} \quad \sum_{n=0}^{+\infty} T_0^{-1}(I - TT_0^{-1})^n$$

converge absolutely and so converge. At the same time, one easily sees by induction that, for all $n \in \mathbb{N}$,

$$(I - T_0^{-1}T)^n T_0^{-1} = T_0^{-1}(I - TT_0^{-1})^n :$$

the equality is certainly true for $n = 0$ and, if it holds for $n \in \mathbb{N}$, we have

$$\begin{aligned}(I - T_0^{-1}T)^{n+1} T_0^{-1} &= (I - T_0^{-1}T)^n (T_0^{-1} - T_0^{-1}TT_0^{-1}) \\ &= (I - T_0^{-1}T)^n T_0^{-1}(I - TT_0^{-1}) \\ &= T_0^{-1}(I - TT_0^{-1})^{n+1}.\end{aligned}$$

Thus, the two series are equal. Let S be their sum.

2. We check that S is indeed the inverse of T.

$$\begin{aligned}ST &= ST_0\big((T_0^{-1}T - I) + I\big) \\ &= -\sum_{n=0}^{+\infty} (I - T_0^{-1}T)^{n+1} + \sum_{n=0}^{+\infty} (I - T_0^{-1}T)^n = I.\end{aligned}$$

(These manipulations are justified because the product is a continuous bilinear map from $L(E) \times L(E)$ to $L(E)$ and because the series converge.) Likewise,

$$\begin{aligned}TS &= \big((TT_0^{-1} - I) + I\big)T_0 S \\ &= -\sum_{n=0}^{+\infty} (I - TT_0^{-1})^{n+1} + \sum_{n=0}^{+\infty} (I - TT_0^{-1})^n = I.\end{aligned}$$

Thus, if $\|T - T_0\| < \|T_0^{-1}\|^{-1}$, the element T is invertible and T^{-1} is indeed given by the series in question.

3. In particular, if $\|T - T_0\| < \|T_0^{-1}\|^{-1}$,

$$\|T^{-1} - T_0^{-1}\| \leq \sum_{n=1}^{+\infty} \|I - T_0^{-1}T\|^n \, \|T_0^{-1}\|$$
$$= \frac{\|T_0^{-1}\| \, \|I - T_0^{-1}T\|}{1 - \|I - T_0^{-1}T\|} \leq \frac{\|T_0^{-1}\|^2 \, \|T - T_0\|}{1 - \|T - T_0\| \, \|T_0^{-1}\|},$$

which shows that the map $T \mapsto T^{-1}$ is continuous at T_0. □

Remark. According to the proof, the map $T \mapsto T^{-1}$ from $\mathcal{I}$ to $L(E)$ has a local series expansion everywhere. This would allow us to show that this map is in fact of class C^∞.

Definitions and Notation. Suppose $T \in L(E)$. A **spectral value** of T is any element $\lambda \in \mathbb{K}$ such that $\lambda I - T$ is not invertible. The set of spectral values of T is called the **spectrum** of T and is denoted by $\sigma(T)$. Any point of $\mathbb{K}$ that is not a spectral value of T is called a **regular value** or **resolvent value** of T. The set $\rho(T) = \mathbb{K} \setminus \sigma(T)$ of regular values of T is called the **resolvent set** of T.

An **eigenvalue** of T is any element $\lambda \in \mathbb{K}$ such that $\lambda I - T$ is not injective (so that $\ker(\lambda I - T) \neq \{0\}$). Thus, every eigenvalue of T is a spectral value of T, but the converse is generally false (unless of course if E has finite dimension or, more generally, if T has finite rank; see Exercise 13 below).

If λ is an eigenvalue of T, the space $\ker(\lambda I - T)$ is called the **eigenspace** associated with λ. We denote by $\operatorname{ev}(T)$ the set of eigenvalues of T.

Example. Take $E = C([0,1])$ and let T be the operator that associates to $f \in E$ the function Tf defined by

$$Tf(x) = \int_0^x f(t)\,dt. \tag{$*$}$$

One sees right away that $\ker T = \{0\}$ and $\operatorname{im} T = \{g \in C^1([0,1]) : g(0) = 0\}$. Thus T is injective but not surjective: that is, $0 \notin \operatorname{ev}(T)$ but $0 \in \sigma(T)$. We now show that 0 is the only spectral value of T.

To this end, take $\lambda \neq 0$ and $g \in E$. If $f \in E$ satisfies

$$\lambda f - Tf = g, \tag{$**$}$$

the function $h = Tf$ is an element of $C^1([0,1])$ such that

$$h(0) = 0 \quad \text{and} \quad \lambda h' - h = g. \tag{$\dagger$}$$

Conversely, if $h \in C^1([0,1])$ satisfies (†), the function $f = h'$ is a solution of (∗∗). Now, it is easy to check that the differential equation (†) has as its unique solution

$$h(x) = \frac{e^{x/\lambda}}{\lambda} \int_0^x g(t) e^{-t/\lambda}\, dt.$$

Therefore (∗∗) is satisfied if and only if

$$f(x) = \frac{1}{\lambda}\Bigl(g(x) + \frac{e^{x/\lambda}}{\lambda} \int_0^x g(t) e^{-t/\lambda} dt \Bigr),$$

whence we deduce that λ is a regular value of T and that

$$\bigl((\lambda I - T)^{-1} g\bigr)(x) = \frac{1}{\lambda}\Bigl(g(x) + \frac{e^{x/\lambda}}{\lambda} \int_0^x g(t) e^{-t/\lambda}\, dt \Bigr).$$

To summarize, $\operatorname{ev}(T) = \varnothing$, $\sigma(T) = \{0\}$, and $\rho(T) = \mathbb{K} \setminus \{0\}$.

Proposition 1.2 *Suppose $T \in L(E)$. The limit $\lim_{n\to\infty} \|T^n\|^{1/n}$ exists and*

$$\lim_{n\to\infty} \|T^n\|^{1/n} = \inf_{n\in\mathbb{N}^*} \|T^n\|^{1/n}.$$

This value is denoted by $r(T)$. Moreover, the spectrum $\sigma(T)$ is a compact subset of $\mathbb{K}$ and

$$|\lambda| \le r(T) \quad \textit{for all } \lambda \in \sigma(T).$$

In particular, we see that $r(T) \le \|T\|$ and so

$$|\lambda| \le \|T\| \quad \text{for all } \lambda \in \sigma(T).$$

Proof

1. Set $a = \inf_{n\in\mathbb{N}^*} \|T^n\|^{1/n}$. Certainly we have

$$a \le \liminf_{n\to+\infty} \|T^n\|^{1/n}.$$

Take $\varepsilon > 0$ and let $n_0 \in \mathbb{N}^*$ be such that $\|T^{n_0}\|^{1/n_0} \le a + \varepsilon$. Given $n \in \mathbb{N}^*$, we can write, by dividing with remainder, $n = p(n)n_0 + q(n)$, with $p(n) \in \mathbb{N}$, $q(n) \in \mathbb{N}$ and $0 \le q(n) < n_0$. Thus

$$\|T^n\| \le \|T^{n_0}\|^{p(n)} \, \|T\|^{q(n)}.$$

Since $\lim_{n\to+\infty} q(n)/n = 0$ and $\lim_{n\to+\infty} p(n)/n = 1/n_0$, we deduce that

$$\limsup_{n\to+\infty} \|T^n\|^{1/n} \le \|T^{n_0}\|^{1/n_0} \le a + \varepsilon.$$

This holds for all $\varepsilon > 0$, so $\lim_{n\to+\infty} \|T^n\|^{1/n} = a$.

2. The map $\lambda \mapsto (\lambda I - T)$ from $\mathbb{K}$ to $L(E)$ is clearly continuous. Therefore, by Proposition 1.1, $\rho(T)$ is open and $\sigma(T)$ is closed. All that remains to show is that $\sigma(T)$ is bounded by $r(T)$.
3. Take $\lambda \in \mathbb{K}$ such that $|\lambda| > r(T)$, and consider $r \in \big(r(T), |\lambda|\big)$. Since $r > r(T)$, there exists an integer $n_0 \in \mathbb{N}^*$ such that

$$\|T^n\| \le r^n \quad \text{for all } n \ge n_0.$$

The series $\sum_{n=0}^{+\infty} \lambda^{-n-1}T^n$ converges absolutely in $L(E)$ (since $r < |\lambda|$) and it is easy to see that

$$(\lambda I - T)\Big(\sum_{n=0}^{+\infty} \lambda^{-n-1}T^n\Big) = \Big(\sum_{n=0}^{+\infty} \lambda^{-n-1}T^n\Big)(\lambda I - T) = I,$$

and so that $\lambda \in \rho(T)$. Since this holds for all $|\lambda| > r(T)$, the proof is complete. □

We take up again the operator T on $E = C([0,1])$ defined by equation $(*)$ on page 189. Clearly, $\|T\| = 1$. On the other hand, an easy inductive computation shows that, for every $n \in \mathbb{N}^*$,

$$T^n f(x) = \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} f(t)\,dt,$$

so that $\|T^n\| \le 1/n!$, which implies that $r(T) = 0$. Here, then, $r(T) < \|T\|$.

For $T \in L(E)$ and $\lambda \in \rho(T)$, write

$$R(\lambda, T) = (\lambda I - T)^{-1}.$$

Proposition 1.3 *Suppose $T \in L(E)$. For all $\lambda, \mu \in \rho(T)$, we have*

$$R(\lambda,T) - R(\mu,T) = (\mu - \lambda)R(\lambda,T)R(\mu,T) = (\mu - \lambda)R(\mu,T)R(\lambda,T).$$

(This is called the **resolvent equation.**) *Moreover, the map $\lambda \mapsto R(\lambda,T)$ from the open subset $\rho(T)$ of $\mathbb{K}$ to $L(E)$ is differentiable and*

$$\frac{d}{d\lambda}R(\lambda,T) = -\big(R(\lambda,T)\big)^2.$$

Proof. First,

$$\begin{aligned} R(\lambda,T) - R(\mu,T) &= R(\lambda,T)\big((\mu I - T) - (\lambda I - T)\big)R(\mu,T) \\ &= (\mu - \lambda)R(\lambda,T)R(\mu,T), \end{aligned}$$

which proves the resolvent equation. In particular,

$$\frac{1}{h}\big(R(\lambda + h, T) - R(\lambda, T)\big) = -R(\lambda,T)R(\lambda + h, T),$$

with $h \in \mathbb{K}^*$ and $\lambda, \lambda+h \in \rho(T)$. By the continuity of the map $\lambda \mapsto R(\lambda, T)$ (an immediate consequence of Proposition 1.1) and the continuity of the product in $L(E)$, we obtain

$$\lim_{h\to 0} \frac{1}{h}\big(R(\lambda+h,T) - R(\lambda,T)\big) = -\big(R(\lambda,T)\big)^2,$$

which concludes the proof. □

We know that, if E is finite-dimensional, the spectrum of T can be empty if $\mathbb{K} = \mathbb{R}$ but not if $\mathbb{K} = \mathbb{C}$, since d'Alembert's Theorem (the Fundamental Theorem of Algebra) guarantees that the characteristic polynomial of T has at least one complex root. We shall show that this is also the case in infinite dimension.

Theorem 1.4 *Suppose $T \in L(E)$. If $\mathbb{K} = \mathbb{C}$, the spectrum $\sigma(T)$ of T is nonempty, and*

$$r(T) = \max\big\{|\lambda| : \lambda \in \sigma(T)\big\}.$$

In contrast, T may have no eigenvalues, even when $\mathbb{K} = \mathbb{C}$, as shown by the example on page 189.

The real number $r(T)$ is called the **spectral radius** of T.

Proof

1. For $z \in \rho(T)$, set $R_z = R(z,T)$. By step 3 in the proof of Proposition 1.2, we know that $|z| > r(T)$ implies that

$$R_z = \sum_{n=0}^{+\infty} z^{-n-1} T^n,$$

the series converging absolutely in $L(E)$. We deduce that, for every $t \in (r(T), +\infty)$,

$$R_{te^{i\theta}} = \sum_{n=0}^{+\infty} e^{-i(n+1)\theta} t^{-n-1} T^n,$$

the series converging uniformly with respect to $\theta \in \mathbb{R}$ in $L(E)$. Multiplying by $(te^{i\theta})^{p+1}$, with $p \in \mathbb{N}$, and integrating the result from 0 to 2π, we obtain, by the continuity of the Riemann integral with values in $L(E)$ (see Exercise 5 on page 20, for instance),

$$\int_0^{2\pi} (te^{i\theta})^{p+1} R_{te^{i\theta}}\, d\theta = \sum_{n=0}^{+\infty} \int_0^{2\pi} (te^{i\theta})^{p-n} T^n \, d\theta = 2\pi T^p.$$

Thus, for every $p \in \mathbb{N}$ and $t > r(T)$,

$$T^p = \frac{1}{2\pi} \int_0^{2\pi} (te^{i\theta})^{p+1} R_{te^{i\theta}}\, d\theta.$$

2. We now prove that the spectrum of T is nonempty. Assume the contrary. Applying the preceding equality in the case $p = 0$, we have

$$I = \frac{1}{2\pi}\int_0^{2\pi} te^{i\theta} R_{te^{i\theta}}\, d\theta \quad \text{for all } t > r(T).$$

But, if we suppose that $\rho(T) = \mathbb{C}$, the function J_0 given by

$$J_0(t) = \frac{1}{2\pi}\int_0^{2\pi} te^{i\theta} R_{te^{i\theta}}\, d\theta$$

is defined and continuous on $[0, +\infty)$ and is of class C^1 on $(0, +\infty)$; moreover

$$\frac{dJ_0}{dt}(t) = \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial}{\partial t}\big(te^{i\theta} R_{te^{i\theta}}\big)\, d\theta \quad \text{for all } t > 0.$$

(In what concerns differentiation under the integral sign, the Riemann integral of functions with values in a Banach space behaves as that of scalar functions.) But

$$\frac{\partial}{\partial t}\big(te^{i\theta} R_{te^{i\theta}}\big) = e^{i\theta}\frac{d}{dz}(zR_z)\Big|_{z=te^{i\theta}}$$

and

$$\frac{\partial}{\partial \theta}\big(te^{i\theta} R_{te^{i\theta}}\big) = ite^{i\theta}\frac{d}{dz}(zR_z)\Big|_{z=te^{i\theta}},$$

since we saw in Proposition 1.3 that the map $z \mapsto R_z$ from $\rho(T)$ to $L(E)$ is differentiable (holomorphic). Thus

$$\frac{dJ_0}{dt}(t) = \frac{1}{2i\pi t}\int_0^{2\pi} \frac{\partial}{\partial \theta}\big(te^{i\theta} R_{te^{i\theta}}\big)\, d\theta = 0 \quad \text{for all } t > 0.$$

We deduce, using the Mean Value Theorem for Banach-space valued functions, that J_0 is constant on $[0, +\infty)$, which cannot be the case since $J_0(0) = 0$ and $J_0(t) = I$ for $t > r(T)$. This contradiction shows that $\sigma(T)$ is nonempty.

3. Set $\rho = \max\{|\lambda| : \lambda \in \sigma(T)\}$. We know by Proposition 1.2 that $\rho \le r(T)$. For $n \in \mathbb{N}^*$ and $t > \rho$, set

$$J_n(t) = \frac{1}{2\pi}\int_0^{2\pi} (te^{i\theta})^{n+1} R_{te^{i\theta}}\, d\theta.$$

As before, we see that $dJ_n/dt = 0$ on $(\rho, +\infty)$. Thus $J_n(t) = T^n$ for every $t > \rho$. Now write $M_t = \max\big\{\|R_{te^{i\theta}}\| : \theta \in [0, 2\pi]\big\}$. Then

$$\|T^n\| \le t^{n+1} M_t \quad \text{for all } n \in \mathbb{N}^* \text{ and } t > \rho,$$

which implies that $r(T) \le t$ for every $t > \rho$, and so that $r(T) \le \rho$. $\square$

Now fix $T \in L(E)$. To every polynomial $P = a_0 + a_1X + \cdots + a_nX^n$ with coefficients in $\mathbb{K}$, we can associate the operator $P(T) \in L(E)$ defined by

$$P(T) = a_0I + a_1T + \cdots + a_nT^n.$$

Clearly, for any $\lambda, \mu \in \mathbb{K}$ and $P, Q \in \mathbb{K}[X]$,

$$(\lambda P + \mu Q)(T) = \lambda P(T) + \mu Q(T), \qquad PQ(T) = P(T)Q(T), \qquad 1(T) = I.$$

In other words, the map $P \mapsto P(T)$ from $\mathbb{K}[X]$ to $L(E)$ is a morphism of algebras with unity. We will compare the spectrum of $P(T)$ with the image under P of the spectrum of T.

Theorem 1.5 (spectral image) *If $T \in L(E)$ and $P \in \mathbb{K}[X]$, we have*

$$P(\sigma(T)) \subset \sigma(P(T)),$$

with equality if $\mathbb{K} = \mathbb{C}$.

Proof

1. Take $\lambda \in \mathbb{K}$. Since λ is a root of the polynomial $P - P(\lambda)$, there exists a polynomial $Q_\lambda \in \mathbb{K}[X]$ such that $P - P(\lambda) = (X - \lambda)Q_\lambda$. Then

$$P(T) - P(\lambda)I = (T - \lambda I)Q_\lambda(T) = Q_\lambda(T)(T - \lambda I).$$

Suppose that $P(\lambda) \notin \sigma(P(T))$, and set $S = \big(P(\lambda)I - P(T)\big)^{-1}$. Then

$$(\lambda I - T)Q_\lambda(T)S = SQ_\lambda(T)(\lambda I - T) = I,$$

showing that $(\lambda I - T)$ is invertible, with inverse $SQ_\lambda(T) = Q_\lambda(T)S$; thus $\lambda \notin \sigma(T)$. Thus $\lambda \in \sigma(T)$ implies $P(\lambda) \in \sigma(P(T))$, which is to say $P(\sigma(T)) \subset \sigma(P(T))$.

2. Suppose that $\mathbb{K} = \mathbb{C}$ and that P has degree at least 1 (if P is constant, the result is trivial). Take $\mu \in \sigma(P(T))$. Write the polynomial $P - \mu$ as a product of factors of degree 1:

$$P - \mu = C(X - \lambda_1)\dots(X - \lambda_n),$$

with $C \neq 0$. Then

$$P(T) - \mu I = C(T - \lambda_1 I)\dots(T - \lambda_n I).$$

Since, by assumption, $P(T) - \mu I$ is not invertible, one of the factors $T - \lambda_j I$ is not invertible. Then, for this value of j, we have $\lambda_j \in \sigma(T)$. Since $P(\lambda_j) = \mu$, this shows that $\mu \in P(\sigma(T))$. □

Remark. In most of this section, we haven't really needed the fact that we are dealing with operators; all we've used is the structure of $L(E)$ as a Banach algebra with unity. These results extend to any Banach algebra with unity.

Exercises

1. Let T be a continuous operator on a Banach space E. Show that the inequality $|\lambda| > \|T\|$ implies

$$\|(\lambda I - T)^{-1}\| \leq \frac{1}{|\lambda| - \|T\|}.$$

2. Let T be a continuous operator on a Banach space E and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $\rho(T)$ converging to $\lambda \in \mathbb{K}$. Show that, if the sequence $(R(\lambda_n, T))$ is bounded in $L(E)$, then $\lambda \in \rho(T)$.
Hint. Show that the sequence $(R(\lambda_n, T))$ converges in $L(E)$. Let S be its limit. Show that $S(\lambda I - T) = (\lambda I - T)S = I$.

3. Let X be a metric space. Take $E = C_b(X)$ and let T be a positive operator on E (recall that this means that $Tf \geq 0$ for any $f \in E$ with $f \geq 0$.)

a. Show that $|Tf| \leq T|f|$ for every $f \in E$.
Hint. Take $x \in X$ and let α be a complex number of absolute value 1 such that $|Tf(x)| = \alpha Tf(x)$. Show that $\alpha Tf(x) = T(\mathrm{Re}(\alpha f))(x)$.

b. Take $\lambda \in \mathbb{K}$ such that $|\lambda| > r(T)$. Show that

$$\|R(\lambda, T)\| \leq \|R(|\lambda|, T)\|.$$

Hint. Show that, for every $f \in E$,

$$|R(\lambda, T)f| \leq R(|\lambda|, T)|f|.$$

c. Deduce that $r(T) \in \sigma(T)$.
Hint. Take $\lambda \in \sigma(T)$ such that $|\lambda| = r(T)$. Consider a sequence $(\lambda_n)_{n\in\mathbb{N}}$ converging to λ and such that $|\lambda_n| > r(T)$ for every $n \in \mathbb{N}$. Then use Exercise 2.

4. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of complex numbers and p a real number in the range $[1, +\infty)$. Define an operator T on ℓ^p by setting

$$(Tu)(n) = \lambda_n u(n) \quad \text{for all } n \in \mathbb{N}.$$

a. Show that T is continuous if and only if the sequence (λ_n) is bounded.

b. When T is continuous, compute its eigenvalues and spectrum.

5. Suppose $p \in [1, \infty]$. Define an operator S on ℓ^p by setting

$$(Su)(n) = u(n+1) \quad \text{for all } n \in \mathbb{N}.$$

We call S the *left shift*.

a. If $p < \infty$, show that $\mathrm{ev}(S) = \{\lambda \in \mathbb{K} : |\lambda| < 1\}$. If $p = \infty$, show that $\mathrm{ev}(S) = \{\lambda \in \mathbb{K} : |\lambda| \leq 1\}$.

b. Deduce that $\sigma(S) = \{\lambda \in \mathbb{K} : |\lambda| \leq 1\}$ in both cases.

6. *Spectrum of an isometry.* Let E be a Banach space and T an isometry of E (recall that this means $T \in L(E)$ and $\|Tx\| = \|x\|$ for all $x \in E$). Set $D = \{\lambda \in \mathbb{K} : |\lambda| < 1\}$, $C = \{\lambda \in \mathbb{K} : |\lambda| = 1\}$, and $\bar{D} = D \cup C$.

a. Show that $\operatorname{ev}(T) \subset C$, that $\sigma(T) \subset \bar{D}$, and that, if $\lambda \in D$,

$$\operatorname{im}(\lambda I - T) = E \iff \lambda \in \rho(T).$$

b. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence in $D \cap \rho(T)$ converging to $\lambda \in D$. Show that $\lambda \in \rho(T)$.

Hint. Show that $\|R(\lambda_n, T)\| \leq 1/(1 - |\lambda_n|)$ for every $n \in \mathbb{N}$; then use Exercise 2.

c. Show that $D \cap \rho(T)$ is open and closed in D. Deduce that $D \cap \rho(T)$ is either empty or equal to D.

d. Show that the spectrum of T is either contained in C or equal to $\bar{D}$. Show that the first case occurs if and only if T is surjective.

e. Assume that $E = \ell^p$, with $p \in [1, \infty]$, and that T is defined by $(Tu)(0) = 0$ and

$$(Tu)(n) = u(n-1) \quad \text{for all } n \in \mathbb{N}^*.$$

(T is called the *right shift.*) Show that the spectrum of T equals $\bar{D}$, and that T has no eigenvalues.

7. *Spectrum of a projection.* Let E be a Banach space and let $P \in L(E)$ be such that $P^2 = P$, $P \neq 0$, and $P \neq I$. Show that $\operatorname{ev}(P) = \sigma(P) = \{0, 1\}$. (The converse holds if P is assumed hermitian: see Exercise 13 on page 212.)

8. Let S and T be continuous operators on a Banach space E.

a. Show that ST and TS have the same nonzero spectral values.

Hint. If U is the inverse of $\lambda I - ST$, consider $V = I + TUS$.

b. Show that, if S or T is invertible, then $\sigma(ST) = \sigma(TS)$. What happens in the general case? (You might consider the operators S and T introduced in Exercises 5 and 6e above.)

9. Let X be a compact metric space and take $\varphi \in C(X)$. Let T be the operator defined on $C(X)$ by

$$Tf = \varphi f \quad \text{for all } f \in C(X).$$

Show that $\sigma(T) = \varphi(X)$ and that

$$\operatorname{ev}(T) = \bigl\{\lambda \in \mathbb{K} : \{\varphi = \lambda\} \text{ has nonempty interior}\bigr\}.$$

What if we consider T as an operator from $L^p(m)$ to itself, where m is a positive Radon measure on X and $p \in [1, \infty]$?

10. Let T be the operator defined on $C([0,1])$ by

$$T(f)(x) = \begin{cases} \dfrac{\pi}{2} f(0) & \text{if } x = 0, \\ \displaystyle\int_0^x \frac{f(y)}{\sqrt{x^2 - y^2}}\, dy & \text{if } x \neq 0. \end{cases}$$

Show that T is a continuous operator from $C([0,1])$ to itself and that $\|T\| = \pi/2$. Show that every point in the interval $(0, \pi/2]$ is an eigenvalue of T. Compute the spectral radius of T.

11. Suppose $p \in [1, \infty]$ and let S be the operator on $L^p((0,1))$ defined by

$$Sf(x) = \int_0^1 x^2 y^2 f(y)\, dy.$$

Solve the equation

$$\lambda f - Sf = g$$

for f, as a function of $\lambda \in \mathbb{K}^*$ and $g \in L^p((0,1))$. Determine the eigenvalues and spectral values of S.

Hint. If $Sf = \lambda f$, with $f \in L^p((0,1))$ and $\lambda \in \mathbb{C}^*$, then f is of the form $f(x) = ax^2$.

12. Same questions for the operator T defined on $L^p((0,1))$ by

$$Tf(x) = \int_0^1 xy(1 - xy) f(y)\, dy.$$

13. *Spectrum of a finite-rank operator.* Consider a Banach space E and an element $T \in L(E)$ of finite rank, which means that the image of T is finite-dimensional (see, for example, Exercises 11 and 12).

a. Set $F = \operatorname{im} T$ and let T_F be the operator on F given by restriction of T to F. Clearly, $T_F \in L(F)$. Show that T and T_F have the same nonzero eigenvalues.

b. Take $\lambda \in \mathbb{K}^*$ and put $S = \lambda I_F - T_F \in L(F)$, where I_F is the identity on F. Assume that S is invertible. Show that $\lambda \in \rho(T)$.

Hint. Show that $\lambda I - T$ is injective. Then compute

$$(\lambda I - T)(I + S^{-1} T)$$

and deduce that $\lambda I - T$ is bijective and that its inverse is continuous.

c. **i.** Show that $\sigma(T) \cap \mathbb{K}^* = \operatorname{ev}(T) \cap \mathbb{K}^*$.

ii. Show that, if E is infinite-dimensional, then $0 \in \operatorname{ev}(T)$.

iii. Show that $\sigma(T) = \operatorname{ev}(T)$.

14. Let E be a Banach space and take $T \in L(E)$. Denote by F the closure of $\operatorname{im} T$. If $S \in L(E)$ and $S(F) \subset F$, denote by S_F the element of $L(F)$ that is the restriction of S to F.

a. Suppose $\lambda \in \rho(T)$. Show that $R(\lambda, T)(F) \subset F$ and deduce that $\lambda \in \rho(T_F)$ and $R(\lambda, T_F) = (R(\lambda, T))_F$.

b. Suppose $\lambda \in \rho(T_F) \setminus \{0\}$. Show that $(\lambda I - T)$ is injective and that $(\lambda I - T)(I + R(\lambda, T_F)T) = \lambda I$. Deduce that $\lambda \in \rho(T)$ and that

$$R(\lambda, T) = \frac{1}{\lambda}\big(I + R(\lambda, T_F)T\big).$$

c. Deduce from the preceding results that

$$\sigma(T) \cap \mathbb{K}^* \subset \sigma(T_F) \subset \sigma(T).$$

d. Show directly that

$$r(T) = r(T_F).$$

Hint. $(T_F)^n = (T^n)_F$ and $T^n = (T_F)^{n-1}T$.

15. *Volterra operators.* Suppose $K \in C([0,1]^2)$ and let T be the operator on $C([0,1])$ defined by

$$T(f)(x) = \int_0^x K(x,y) f(y)\, dy.$$

a. Show that, for every positive integer n and every $f \in C([0,1])$,

$$\big|T^n f(x)\big| \le \|f\|\, \|K\|^n \frac{x^n}{n!},$$

where $\|\cdot\|$ is the uniform norm in $C([0,1])$ and in $C([0,1]^2)$.

b. Determine the spectral radius and then the spectrum of T.

16. a. Let E be a Banach space endowed with an order relation $\le$ satisfying these conditions:

- for any $f, g \in E$, $f \le g$ if and only if $g - f \ge 0$;
- for any $f \in E$ and $\lambda \in \mathbb{R}^+$, $f \ge 0$ implies $\lambda f \ge 0$;
- for any $f, g \in E$, $0 \le f \le g$ implies $\|f\| \le \|g\|$.

(For example, all the function spaces studied in the preceding chapters, such as L^p, $C_b(X)$, $C_0(X)$, and so on, have these properties when given the natural order relation.) Let $T \in L(E)$ be a positive operator (recall that this means $Tf \ge 0$ for all $f \in E$ with $f \ge 0$), and suppose that $\lambda \in \mathbb{R}^+$. Show that, if there exists a nonzero element f in E such that

$$f \ge 0 \quad \text{and} \quad Tf \ge \lambda f,$$

then $r(T) \ge \lambda$.

Hint. Show that $\|T^n f\| \ge \lambda^n \|f\|$ for every $n \in \mathbb{N}$.

b. Let φ be a continuous map from $[0,1]$ to $[0,1]$ and K a continuous map from $[0,1]^2$ to $\mathbb{R}^+$. Define an operator $T \in L(C[0,1]))$ by setting

$$Tf(x) = \int_0^{\varphi(x)} K(x,y) f(y)\, dy \quad \text{for all } f \in C([0,1]) \text{ and } x \in [0,1].$$

i. Prove that, if $\varphi(x) \le x$ for every $x \in [0,1]$, then $r(T) = 0$ (see Exercise 15).

ii. Suppose there is a point $x_0 \in (0,1)$ such that

$$K(x_0, x_0) > 0 \quad \text{and} \quad \varphi(x_0) > x_0.$$

Show that $r(T) > 0$.
Hint. By assumption, there exists $\delta > 0$ such that, for every $x, y \in [0,1]$,

$$|x - x_0| \le \delta \text{ and } |y - x_0| \le \delta \quad \Longrightarrow \quad K(x,y) \ge \delta \text{ and } \varphi(x) \ge x + \delta.$$

Now consider the element f in $C([0,1])$ defined by

$$f(x) = (\delta - |x - x_0|)^+$$

and show that $|x - x_0| \le \delta$ implies $Tf(x) \ge \delta^3/2$. Deduce that $Tf \ge \delta^2 f/2$.

17. Let T be a continuous operator on a Banach space E for which the sequence $(\|T^n\|)_{n\in\mathbb{N}}$ converges to 0. Show that $I - T$ is invertible, that the series $\sum_{n=0}^{+\infty} T^n$ is absolutely convergent in $L(E)$, and that its sum is $(I-T)^{-1}$.
Hint. Show that $r(T) < 1$.

18. Consider a compact space X and a linear operator T on $C(X)$. Assume that T is positive (if $f \in C(X)$ satisfies $f \ge 0$, then $Tf \ge 0$).

a. Show that T is continuous and that $\|T\| = \|T1\|$, where the right-hand side is the norm in $C(X)$ of $T1$, the image under T of the constant function 1 on X.

Now suppose that there exists a constant $C \ge 0$ such that, for all $n \in \mathbb{N}$ and all $x \in X$, we have

$$0 \le \sum_{j=0}^{n} (T^j 1)(x) \le C.$$

b. Show that, given any pair (p,q) of nonnegative integers, we have $T^{p+q}1 \le C\, T^p 1$. Show also that, for every point x in X, the sequence $((T^n 1)(x))_{n\in\mathbb{N}}$ converges to 0.

c. Deduce that the sequence of functions $(T^n 1)_{n\in\mathbb{N}}$ converges uniformly on X to 0. (You might use Exercise 4 on page 30.)

d. Deduce that $r(T) < 1$ and that the series $\sum_{n=0}^{\infty} T^n 1$ converges absolutely in $C(X)$ (see Exercise 17).

19. Let T be a continuous operator on a Banach space E. Show that, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\sigma(S) \subset \{\lambda \in \mathbb{K} : d(\lambda, \sigma(T)) < \varepsilon\} \quad \text{for every } S \in L(E) \text{ with } \|T - S\| < \delta.$$

Hint. Set $M = \sup\{\|(\lambda I - T)^{-1}\| : d(\lambda, \sigma(T)) \geq \varepsilon\}$. Show that M is finite (see Exercise 1) and that $\delta = 1/M$ works.

20. *Approximate eigenvalues.* Let T be a continuous operator on a Banach space E. By definition, and *approximate eigenvalue* of T is any $\lambda \in \mathbb{K}$ for which there exists a sequence $(x_n)_{n \in \mathbb{N}}$ of elements in E of norm 1 such that $\lim_{n \to +\infty} Tx_n - \lambda x_n = 0$. We denote by $\operatorname{aev}(T)$ the set of approximate eigenvalues of T.

a. Suppose $\lambda \in \mathbb{K}$ and write $\alpha(\lambda) = \inf_{\|x\|=1} \|\lambda x - Tx\|$. Show that λ is an approximate eigenvalue of T if and only if $\alpha(\lambda) = 0$. Show also that the map $\lambda \mapsto \alpha(\lambda)$ from $\mathbb{K}$ to $\mathbb{R}^+$ is continuous (in fact, 1-Lipschitz).

b. Show that $\operatorname{aev}(T)$ is compact and that

$$\overline{\operatorname{ev}(T)} \subset \operatorname{aev}(T) \subset \sigma(T). \tag{$*$}$$

c. Show that $\operatorname{aev}(T)$ contains the boundary of $\sigma(T)$, that is, the set $\sigma(T) \cap \overline{\rho(T)}$. In particular, $\operatorname{aev}(T)$ is nonempty if $\mathbb{K} = \mathbb{C}$.

Hint. Use Exercise 2 above.

d. **i.** Suppose $S \in L(E)$ is not invertible. Show that, if there is $C > 0$ such that

$$\|x\| \leq C\|Sx\| \quad \text{for all } x \in E,$$

the image of S is not dense in E.

Hint. The assumption implies that the map $x \mapsto Sx$ from E to $\operatorname{im} S$ has a continuous inverse U. If $\operatorname{im} S$ is dense in E, then U can be extended to a continuous linear map from E to E.

ii. Suppose that $\lambda \in \sigma(T)$. Show that, if $\operatorname{im}(\lambda I - T)$ is dense in E, then λ is an approximate eigenvalue of T. Is the converse true?

e. Suppose that T is an isometry (see Exercise 6 above). Show that

$$\operatorname{aev}(T) = \sigma(T) \cap \{\lambda \in \mathbb{K} : |\lambda| = 1\}.$$

Hint. One inclusion is obvious. To prove the other, you might use Exercise 6.

f. Find operators T for which the inclusions $(*)$ are strict.

21. *Continuous one-parameter groups.* Let E be a Banach space.

a. Suppose $A \in L(E)$. For $t \in \mathbb{R}$, put

$$P(t) = \exp(tA) = \sum_{n=0}^{+\infty} \frac{t^n}{n!} A^n.$$

Show the following facts:

A. P is a continuous function from $\mathbb{R}$ to $L(E)$.
B. $P(0) = I$ and $P(t+s) = P(t)P(s)$ for all $t, s \in \mathbb{R}$.
C. P is of class C^1 and $dP/dt = AP$.

b. Conversely, consider a function P from $\mathbb{R}$ to $L(E)$ satisfying properties A and B above; we call the family $(P(t))_{t\in\mathbb{R}}$ a *continuous!one-parameter group of operators*.

i. Show that there exists $h \in \mathbb{R}^{+*}$ such that $\int_0^h P(s)\,ds$ is invertible. Fix such an h for now on, and put

$$A = (P(h) - I)\left(\int_0^h P(s)\,ds\right)^{-1}.$$

ii. Show that

$$\left(\int_0^h P(s)\,ds\right)P(t) = \int_t^{t+h} P(s)\,ds \quad \text{for every } t \in \mathbb{R},$$

and deduce that P satisfies property C above.

iii. Compute

$$\frac{d}{dt}\big(P(t)\exp(-tA)\big)$$

and deduce that $P(t) = \exp(tA)$ for every $t \in \mathbb{R}$.

2 Operators in Hilbert Spaces

In this section, we consider the particular case where E is a Hilbert space not equal to $\{0\}$. We make heavy use of the results established in Section 3A of Chapter 3 (pages 112 and following). To simplify the notation we assume that $\mathbb{K} = \mathbb{C}$, but all results in this section remain true for $\mathbb{K} = \mathbb{R}$ (see Exercise 1 below). We first give a simple result that links the spectral properties of an operator $T \in L(E)$ with those of its adjoint T^*, defined on page 112.

Proposition 2.1 *Suppose $T \in L(E)$. Then:*

i. $\ker T = (\operatorname{im} T^*)^\perp$.
ii. $\overline{\operatorname{im} T} = (\ker T^*)^\perp$.

iii. *T is invertible if and only if T^* is, and in this case*

$$(T^*)^{-1} = (T^{-1})^*.$$

Proof. For $x \in E$, we have $x \in \ker T$ if and only if

$$(Tx \mid y) = (x \mid T^*y) = 0 \quad \text{for all } y \in E,$$

which proves the first assertion. The second is a consequence of the first, in view of Corollary 2.7 on page 108 and of the equality $T^{**} = T$. Finally, if T is invertible, we have $TT^{-1} = T^{-1}T = I$ and, by Proposition 3.3 on page 112, $(T^{-1})^*T^* = T^*(T^{-1})^* = I$. Therefore T^* is invertible and $(T^*)^{-1} = (T^{-1})^*$. □

The next result follows immediately.

Corollary 2.2 *If $T \in L(E)$, then*

$$\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}.$$

If $\lambda \in \rho(T)$, then $\bar{\lambda} \in \rho(T^*)$ and

$$R(\bar{\lambda}, T^*) = (R(\lambda, T))^*.$$

In contrast, there is generally no relation between the eigenvalues of T and those of T^* (part ii of Proposition 2.1 allows us to say only that λ is an eigenvalue of T^* if and only if the image of $\bar{\lambda}I - T$ is not dense). For example, if $E = \ell^2$ and T is the right shift of Exercise 6e on page 196, defined by $(Tu)(0) = 0$ and

$$(Tu)(n) = u(n-1) \quad \text{for all } n \in \mathbb{N}^*,$$

there are no eigenvalues. But it is easy to see that the adjoint of T is none other than the left shift of Exercise 5 on page 195, defined by

$$(T^*u)(n) = u(n+1) \quad \text{for all } n \in \mathbb{N};$$

thus $\operatorname{ev}(T^*) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$.

Recall that an operator $T \in L(E)$ is called hermitian if it coincides with its adjoint T^*.

Proposition 2.3 *The spectral radius and the norm of a hermitian operator on E coincide.*

Proof. If T is hermitian, Proposition 3.4 on page 113 says that $\|T^2\| = \|T\|^2$. Iterating this property, which we can do because the square of a hermitian operator is hermitian, we obtain

$$\|T^{2^n}\| = \|T\|^{2^n} \quad \text{for all } n \in \mathbb{N}.$$

We conclude that

$$r(T) = \lim_{n\to+\infty} \|T^{2^n}\|^{2^{-n}} = \|T\|,$$

since the limit of the sequence $(\|T^n\|^{1/n})_{n\in\mathbb{N}}$ equals the limit of any of its subsequences. □

We can now deduce immediately from Proposition 3.4 on page 113 the following corollary:

Corollary 2.4 *For $T \in L(E)$,*

$$\|T\| = \sqrt{r(TT^*)} = \sqrt{r(T^*T)}.$$

2A Spectral Properties of Hermitian Operators

Proposition 2.5 *Every hermitian operator T on E has the following properties:*

i. *The eigenvalues of T are real.*
ii. *For every $\lambda \in \mathbb{C}$, we have $\overline{\operatorname{im}(\lambda I - T)} = (\ker(\bar{\lambda} I - T))^{\perp}$.*
iii. *The eigenspaces of T associated with distinct eigenvalues are orthogonal.*

Proof. Suppose that λ is an eigenvalue of T, and let $x \in E$ be an associated nonzero eigenvector, so that $Tx = \lambda x$ and $x \neq 0$. Then

$$\lambda \|x\|^2 = (\lambda x \,|\, x) = (Tx \,|\, x).$$

Since the operator T is selfadjoint, we have $(Tx \,|\, x) \in \mathbb{R}$ and so $\lambda \in \mathbb{R}$, which proves the first part of the proposition.

The second part is an immediate consequence of the equality $\overline{\operatorname{im} S} = (\ker S^*)^{\perp}$, valid for all $S \in L(E)$ by Proposition 2.1.

Finally, if λ and μ are distinct eigenvalues of T and if x and y are corresponding eigenvectors, we have

$$\lambda (x \,|\, y) = (Tx \,|\, y) = (x \,|\, Ty) = \mu (x \,|\, y),$$

since $\mu \in \mathbb{R}$. Therefore $(x \,|\, y) = 0$. □

The next theorem states, in particular, that the spectrum of a hermitian operator T is also contained in $\mathbb{R}$.

Theorem 2.6 *Let T be a hermitian operator on E. Put*

$$m = \inf\bigl\{(Tx \,|\, x) : x \in E \text{ with } \|x\| = 1\bigr\},$$
$$M = \sup\bigl\{(Tx \,|\, x) : x \in E \text{ with } \|x\| = 1\bigr\}.$$

Then $\sigma(T) \subset [m, M]$, $m \in \sigma(T)$, and $M \in \sigma(T)$.

In other words, $[m, M]$ is the smallest interval containing the spectrum of T.

Proof

1. Take $\lambda \in \mathbb{C}$ and a nonzero element x in E. Then

$$(\lambda x - Tx \mid x) = \left(\lambda - \left(T\left(\frac{x}{\|x\|}\right) \,\Big|\, \frac{x}{\|x\|}\right)\right)\|x\|^2.$$

Denote by $d(\lambda)$ the distance from λ to the interval $[m, M]$:

$$d(\lambda) = \min\bigl\{|\lambda - t| : t \in [m, M]\bigr\}.$$

Then, by the Schwarz inequality and the definition of m and M,

$$\|\lambda x - Tx\|\,\|x\| \geq \bigl|(\lambda x - Tx \mid x)\bigr| \geq d(\lambda)\,\|x\|^2.$$

It follows that

$$\|\lambda x - Tx\| \geq d(\lambda)\,\|x\| \quad \text{for all } x \in E. \tag{$*$}$$

Suppose that $\lambda \notin [m, M]$. Then $d(\lambda) > 0$ and, by $(*)$, $\lambda I - T$ is injective. We now prove that $\operatorname{im}(\lambda I - T)$ is closed. If $(y_n)_{n\in\mathbb{N}}$ is a sequence in $\operatorname{im}(\lambda I - T)$ converging to $y \in E$, with $y_n = \lambda x_n - Tx_n$ for each n, equation $(*)$ implies that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and so converges to some $x \in E$, which clearly satisfies $\lambda x - Tx = y$. Thus $y \in \operatorname{im}(\lambda I - T)$. We then deduce from Proposition 2.5 that

$$\operatorname{im}(\lambda I - T) = (\ker(\bar{\lambda} I - T))^{\perp}.$$

But, since $\bar{\lambda}$ does not belong to $[m, M]$ either, the operator $\bar{\lambda} I - T$ is also injective. We deduce that $\lambda I - T$ is a bijection from E onto itself. Since, by $(*)$, the inverse of this map is continuous (and has norm at most $1/d(\lambda)$), we get $\lambda \in \rho(T)$. Therefore $\sigma(T) \subset [m, M]$.

2. We prove, for example, that $m \in \sigma(T)$. (That $M \in \sigma(T)$ follows by interchanging T and $-T$.) Set $S = T - mI$. By the definition of m, S is a positive hermitian operator. The map $(x, y) \mapsto (Sx \mid y)$ is therefore a scalar semiproduct on E. Applying the Schwarz inequality to this scalar semiproduct, we get

$$\bigl|(Sx \mid y)\bigr|^2 \leq (Sx \mid x)(Sy \mid y) \quad \text{for all } x, y \in E. \tag{$**$}$$

At the same time, by the definition of m, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ of elements in E of norm 1 such that $\lim_{n\to+\infty}(Sx_n \mid x_n) = 0$. By $(**)$,

$$\|Sx_n\|^2 \leq (Sx_n \mid x_n)^{1/2}(S^2x_n \mid Sx_n)^{1/2} \leq (Sx_n \mid x_n)^{1/2}\,\|S\|^{1/2}\,\|Sx_n\|,$$

so that

$$\|Sx_n\| \leq \|S\|^{1/2} (Sx_n \mid x_n)^{1/2},$$

which implies that $\lim_{n\to+\infty} Sx_n = 0$. If m were not a spectral value of T, the operator S would be invertible in $L(E)$ and $x_n = S^{-1}Sx_n$ would tend to 0, which is absurd. Therefore $m \in \sigma(T)$. □

Remark. The second part of this proof did not use the completeness of E. Thus m and M are spectral values for any hermitian operator T, even if the underlying space E is not complete. In particular, the spectrum of any hermitian operator on any scalar product space is nonempty.

Suppose T is hermitian. Recall that $\|T\| = \max(|m|, |M|)$ (see Proposition 3.5 on page 114), and that T is called positive hermitian if $m \geq 0$ (see page 114). The next corollary is an immediate consequence of the preceding results.

Corollary 2.7 *A hermitian operator T on E is positive hermitian if and only if its spectrum $\sigma(T)$ is contained in $\mathbb{R}^+$. If this is the case, $\|T\| \in \sigma(T)$.*

2B Operational Calculus on Hermitian Operators

We saw in Section 1 (page 194) that each element T in $L(E)$ defines a morphism of algebras $P \mapsto P(T)$ from $\mathbb{C}[X]$ to $L(E)$. Now, for T hermitian, we will extend this morphism and define $f(T)$ for every continuous complex-valued map f defined on the spectrum of T.

Let T be a hermitian operator on E. If $P = a_0+a_1X+\cdots+a_nX^n \in \mathbb{C}[X]$, we write $\bar{P} = \bar{a}_0 + \bar{a}_1X + \cdots + \bar{a}_nX^n$ and $|P|^2 = P\bar{P}$.

Proposition 2.8 *For every $P \in \mathbb{C}[X]$, we have $(P(T))^* = \bar{P}(T)$ and*

$$\|P(T)\| = \max_{t\in\sigma(T)} |P(t)|.$$

Proof. The first assertion is an immediate consequence of the fact that T is hermitian (see Proposition 3.3 on page 112). Next, for $P \in \mathbb{C}[X]$, Proposition 3.4 on page 113 gives

$$\|P(T)\| = \|P(T)P(T)^*\|^{1/2} = \||P|^2(T)\|^{1/2}.$$

But, since $|P|^2(T)$ is positive hermitian,

$$\||P|^2(T)\| = \max \sigma(|P|^2(T)),$$

by Corollary 2.7. By the Spectral Image Theorem (page 194), we have $\sigma(|P|^2(T)) = |P|^2(\sigma(T))$, so

$$\|P(T)\| = \Bigl(\max_{t\in\sigma(T)} |P|^2(t)\Bigr)^{1/2} = \max_{t\in\sigma(T)} |P(t)|,$$

which concludes the proof. □

Theorem 2.9 *The map $P \mapsto P(T)$ defined earlier from $\mathbb{C}[X]$ to $L(E)$ extends uniquely to a linear isometry $f \mapsto f(T)$ from $C(\sigma(T))$ to $L(E)$. Moreover:*

i. $(fg)(T) = f(T)g(T)$ *for all* $f, g \in C(\sigma(T))$.
ii. $(f(T))^* = \bar{f}(T)$ *for all* $f \in C(\sigma(T))$.
iii. $\sigma(f(T)) = f(\sigma(T))$ *for all* $f \in C(\sigma(T))$ (**spectral image**).

Proof

1. Let Π be the subset of $C(\sigma(T))$ consisting of restrictions of polynomial functions to $\sigma(T)$. By Proposition 2.8, two polynomials P and Q that have the same restriction to $\sigma(T)$ must satisfy $P(T) = Q(T)$, since $\|P(T) - Q(T)\| = \max_{t \in \sigma(T)} |P(t) - Q(t)| = 0$. Therefore the map $P \mapsto P(T)$ defines an isometry from Π to $L(E)$. By the Stone–Weierstrass Theorem, Π is dense in $C(\sigma(T))$ (see Example 2 on page 34). Using the fact that $L(E)$ is a Banach space, we can apply the Extension Theorem and extend this isometry in a unique way to a linear isometry on $C(\sigma(T))$, which must satisfy the first two properties of the theorem since it extends a map that does.
2. If $\lambda \notin f(\sigma(T))$, the function $1/(\lambda - f)$ is continuous on $\sigma(T)$, and clearly
$$(\lambda I - f(T))^{-1} = \left(\frac{1}{\lambda - f}\right)(T)$$
and $\lambda \in \rho(f(T))$ (the norm of the operator $(\lambda I - f(T))^{-1}$ being the inverse of the distance from λ to $f(\sigma(T))$). Thus $\sigma(f(T)) \subset f(\sigma(T))$.
3. Now take $f \in C^{\mathbb{R}}(\sigma(T))$, $f \geq 0$, with $f(T)$ invertible. We wish to show that $0 \notin f(\sigma(T))$. Since $\sigma(f(T)) \subset f(\sigma(T)) \subset \mathbb{R}^+$, it follows that $-1/n$ is a regular value of $f(T)$ for any $n \in \mathbb{N}^*$, and, by the preceding discussion,
$$R(-1/n, f(T)) = \left(\frac{1}{-1/n - f}\right)(T).$$
Now, the function $\lambda \mapsto R(\lambda, f(T))$ is continuous on $\rho(f(T))$, so
$$\lim_{n \to +\infty} R(-1/n, f(T)) = R(0, f(T)) = -(f(T))^{-1}.$$
At the same time, the map $f \mapsto f(T)$ is isometric from $C(\sigma(T))$ (considered with the uniform norm, still denoted by $\|\cdot\|$) to $L(E)$; therefore
$$\|R(-1/n, f(T))\| = \left\|\frac{1}{-1/n - f}\right\|.$$
If f vanished anywhere in $\sigma(T)$, the value of $\|R(-1/n, f(T))\|$ would go to infinity as $n \to +\infty$, which is a contradiction. Therefore f does not vanish on $\sigma(T)$, which is to say $0 \notin f(\sigma(T))$.

4. Finally, take $f \in C(\sigma(T))$. Suppose that $\lambda \in \rho(f(T))$. Then the operator $\lambda I - f(T)$ is invertible, as are its adjoint $\bar{\lambda} I - \bar{f}(T)$ and hence the product $(\lambda I - f(T))(\bar{\lambda} I - \bar{f}(T)) = |\lambda - f|^2(T)$. Since $|\lambda - f|^2$ is a positive function, we can apply step 3 to it. This implies that the function $|\lambda - f|^2$ does not vanish on $\sigma(T)$; therefore the same is true of $\lambda - f$. Thus shows that $\lambda \notin f(\sigma(T))$, and so that $f(\sigma(T)) \subset \sigma(f(T))$. □

Corollary 2.10 *Let f be a continuous function from $\sigma(T)$ to $\mathbb{C}$. The operator $f(T)$ is hermitian if and only if f is real-valued. It is positive hermitian if and only if $f \geq 0$.*

Proof. The first assertion follows from part ii of Theorem 2.9. The second follows from part iii of the same theorem and from Corollary 2.7. □

Example. If T is a positive hermitian operator and if $\alpha \in (0, +\infty)$, we can define T^α, which is a positive hermitian operator. Then

$$\begin{aligned} T^\alpha T^\beta &= T^{\alpha+\beta} && \text{for all } \alpha, \beta > 0, \\ \sigma(T^\alpha) &= \{t^\alpha : t \in \sigma(T)\} && \text{for all } \alpha > 0. \end{aligned}$$

Moreover, the map $\alpha \mapsto T^\alpha$ is continuous from $(0, +\infty)$ to $L(E)$.

Exercises

1. Let E be a *real* Hilbert space and T a symmetric operator on E.

a. Show that the proof of Theorem 2.6, and so also the theorem itself, remain valid. Deduce that, if there is a constant $C > 0$ such that

$$(Tx \mid x) \geq C\|x\|^2 \quad \text{for all } x \in E,$$

T is invertible.

b. Let $P = X^2 + aX + b$ be a real polynomial having no real roots. Show that $P(T)$ is invertible.
Hint. P can be written as $P = (X+\alpha)^2 + \beta^2$, with $\alpha \in \mathbb{R}$ and $\beta > 0$. But then, for every $x \in E$, we have $(P(T)x \mid x) \geq \beta^2\|x\|^2$.

c. Show that for any $P \in \mathbb{R}[X]$ we have $P(\sigma(T)) = \sigma(P(T))$. (Thus the spectral image property is valid for symmetric operators when $\mathbb{K} = \mathbb{R}$.)
Hint. Imitate the proof of Theorem 1.5, using a factorization of the polynomial $P - \mu$ over $\mathbb{R}$ and the previous question.

d. Show that $r(T) = \|T\| = \max\{|\lambda| : \lambda \in \sigma(T)\}$.
Hint. For the second equality, one might use part a of this exercise and Proposition 3.5 on page 114.

e. Show that the results of Section 2B remain valid when $\mathbb{K} = \mathbb{R}$.
Hint. In view of parts a–d, one can use the same proofs.

2. Let A and B be complementary orthogonal subspaces in a Hilbert space E, and suppose $T \in L(E)$. Assume that T leaves A and B invariant, that is, $T(A) \subset A$ and $T(B) \subset B$. Show that

$$\sigma(T) = \sigma(T_{|A}) \cup \sigma(T_{|B}).$$

(You might show the corresponding equality involving the resolvent set.)
Example. Determine the spectrum of the operator T defined on ℓ^2 by

$$(Tu)(n) = u(n+2) + \frac{1+(-1)^n}{2}\, u(n) \quad \text{for all } n \in \mathbb{N}.$$

(You might use Exercise 5 on page 195.)

3. Let E be a Hilbert space and take $T \in L(E)$. Denote by $\operatorname{aev}(T)$ the set of approximate eigenvalues of T (see Exercise 20 on page 200). Also put

$$i(T) = \{(Tx \,|\, x) : \|x\| = 1\}.$$

a. Show that the spectrum of T equals $\operatorname{aev}(T) \cup \{\bar{\lambda} : \lambda \in \operatorname{ev}(T^*)\}$. In particular, $\sigma(T) = \operatorname{aev}(T)$ if T is hermitian.
Hint. Use Exercise 20d-ii on page 200.

b. Show that $\operatorname{aev}(T) \subset \overline{i(T)}$.

c. Deduce that $\sigma(T) \subset \overline{i(T)}$. (This generalizes the first part of Theorem 2.6.)

d. Deduce that, if $\mathbb{K} = \mathbb{C}$,

$$r(T) \le \sup_{\|x\|=1} |(Tx \,|\, x)| \le \|T\|.$$

4. Let E be a Hilbert space over $\mathbb{C}$. An operator T on E is said to be *normal* if $TT^* = T^*T$.

a. **i.** We assume (in this subitem only) that $E = \ell^2$. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a bounded sequence on $\mathbb{C}$ and let T be the operator on E defined by

$$Tf(n) = \lambda_n f(n).$$

Show that T is normal. Recall from Exercise 4 on page 195 that the spectrum of T equals the closure of the set $\{\lambda_n\}_{n\in\mathbb{N}}$.

ii. Deduce that, if E is infinite-dimensional, every nonempty compact subset of $\mathbb{C}$ is the spectrum of a normal operator on E.
Hint. Use the previous result to handle the case where E is separable; then handle the general case using Exercise 2.

b. **i.** Let $T \in L(E)$. Show that T is normal if and only if $\|Tx\| = \|T^*x\|$ for every $x \in E$.

ii. Let T be a normal operator on E. Show that, for every $\lambda \in \mathbb{C}$,

$$\ker(\lambda I - T) = \ker(\bar{\lambda} I - T^*).$$

Deduce, in particular, that $\lambda \in \operatorname{ev}(T)$ if and only if $\bar{\lambda} \in \operatorname{ev}(T^*)$. Show also that eigenspaces of T associated with distinct eigenvalues are orthogonal. (Work as in the proof of Proposition 2.5.)

c. Let T be a normal operator on E.

i. Show that $\|T\| = r(T)$.
Hint. Start by proving that $r(TT^*) \le r(T)^2$.

ii. Deduce that

$$\|T\| = \sup_{\|x\|=1} \bigl|(Tx \mid x)\bigr|.$$

(Use Exercise 3.)

5. Let T be a continuous operator on a separable Hilbert space. Show that if T is hermitian it has countably many eigenvalues. Show that this conclusion still holds if T is only assumed normal (see Exercise 4), but not if we make no assumptions on T.

6. Let (T_n) be a bounded sequence of positive hermitian operators on a Hilbert space E satisfying, for every $n \in \mathbb{N}$, the condition $T_{n+1} \ge T_n$ (that is, $T_{n+1} - T_n$ is positive hermitian). Set $M = \sup_{n\in\mathbb{N}} \|T_n\|$.

a. Take $n, m \in \mathbb{N}$ such that $m < n$. Show that $T_{n,m} = T_n - T_m$ is a positive hermitian operator of norm at most M. Using equation (**) on page 204 with $S = T_{n,m}$, deduce that, for every $x \in E$,

$$\|T_n x - T_m x\|^4 \le M^3\bigl((T_n x \mid x) - (T_m x \mid x)\bigr)\,\|x\|^2.$$

b. Deduce that for every $x \in E$ the sequence $(T_n x)$ converges and that the map T defined by $Tx = \lim_{n\to+\infty} T_n x$ is a positive hermitian operator.

7. Define an operator T on the Hilbert space $E = L^2((0,+\infty))$ by setting

$$Tf(x) = \int_0^{+\infty} \frac{f(y)}{x+y}\,dy \quad \text{for all } f \in E \text{ and } x \in (0,+\infty).$$

It was shown in Exercise 3 on page 149 that

$$\|T\| = \int_0^{+\infty} \frac{z^{-1/2}}{1+z}\,dz = \pi.$$

a. Let L be the operator on E defined by

$$Lf(x) = \int_0^{+\infty} e^{-xy} f(y)\,dy \quad \text{for all } f \in E \text{ and } x \in (0,+\infty).$$

L is called the *Laplace transform operator* on $L^2((0,+\infty))$. Show that L is a hermitian operator and that $L^2 = T$. Deduce that T is a positive hermitian operator and that $\|L\| = \sqrt{\pi}$.

Hint. To prove that L is continuous, one can write, for $f \in E$ nonnegative-valued,

$$(Lf(x))^2 = \int_0^{+\infty} e^{-xy} f(y)\, dy \int_0^{+\infty} e^{-xy'} f(y')\, dy'.$$

b. Show that $\operatorname{im}(T) \subset C((0, +\infty))$ and deduce that 0 is a spectral value of T. Show that 0 is *not* an eigenvalue of T. (Start by showing that L is injective.)

c. Show that $[0, \pi]$ is the smallest interval containing the spectrum of T (in fact the two sets coincide).

8. Let T be a hermitian operator on a Hilbert space E. For $f \in C^{\mathbb{R}}(\sigma(T))$ and $g \in C(f(\sigma(T)))$, show that

$$(g \circ f)(T) = g(f(T)).$$

In particular, if $\alpha, \beta > 0$ and T is positive hermitian, $(T^\alpha)^\beta = T^{\alpha\beta}$.

9. *Explicit construction of the square root of a positive hermitian operator.* (This exercise is meant to be solved without recourse to the results of Section 2B.) Let T be a positive hermitian operator on a Hilbert space E.

a. Suppose in this item that $\|T\| \le 1$, and consider the sequence of hermitian operators (S_n) defined by $S_0 = 0$ and

$$S_{n+1} = \tfrac{1}{2}(I - T + S_n^2) \quad \text{for all } n \ge 0.$$

i. Show by induction on n that $0 \le S_n \le S_{n+1} \le I$ for every integer $n \in \mathbb{N}$, where $U \ge V$ means that $U - V$ is positive hermitian.
Hint. Set $U = I - T$. Show by induction that, for every integer $n \in \mathbb{N}$, the operators S_n and $S_{n+1} - S_n$ can be expressed as polynomials in U with positive coefficients.

ii. Deduce that there exists a positive hermitian operator S such that $\lim_{n\to+\infty} S_n x = Sx$ for every $x \in E$. (Use Exercise 6 above.)

iii. Set $R = I - S$. Show that $R^2 = T$.

iv. Show that R commutes with every operator on E that commutes with T.

b. Now make no assumption on the norm of T. Show that there exists a hermitian operator R such that $R^2 = T$ and that commutes with every operator that commutes with T.

10. Let E be a Hilbert space.

a. Let T be a hermitian operator on E. Show that, if $f \in C(\sigma(T))$, the operator $f(T)$ commutes with every operator on E that commutes with T.

b. *Uniqueness of the square root of a positive hermitian operator.* Let T be a positive hermitian operator on E and set $R = T^{1/2}$. Let R' be a positive hermitian operator such that $(R')^2 = T$.

  i. Show that $RR' = R'R$.

  ii. Let X and X' be positive hermitian operators such that $X^2 = R$ and $(X')^2 = R'$. Show that, for every $x \in E$,

$$\|Xy\|^2 + \|X'y\|^2 = 0, \quad \text{where } y = (R-R')x.$$

  iii. Deduce that $\|(R-R')x\|^2 = 0$ for every $x \in E$, and so that $R = R'$.

c. Let T and S be positive hermitian operators on E such that $ST = TS$.

  i. Show that ST is a positive hermitian operator.
  Hint. One might introduce $U = S^{1/2}$.

  ii. Show that, if $S \le T$ (that is, if $T - S$ is positive hermitian), then $S^2 \le T^2$.
  Hint. Note that $T^2 - S^2 = (T+S)(T-S)$.

11. *Polar decomposition.* Let T be a continuous operator on a Hilbert space E, and set $P = (T^*T)^{1/2}$.

a. Show that $\ker P = \ker T$ and that $\overline{\operatorname{im} P} = (\ker T)^\perp$.

b. Show that there exists a unique operator $U \in L(E)$ such that
- $\|Ux\| = \|x\|$ for every $x \in (\ker T)^\perp$,
- $Ux = 0$ for every $x \in \ker T$, and
- $T = UP$.

Hint. If $x \in \operatorname{im} P$ and $x = Pz$, we must have $Ux = Tz$.

c. Show that U^*U is the orthogonal projection operator onto $(\ker T)^\perp$.

d. Show that if T is normal ($TT^* = T^*T$), then $UP = PU$.
Hint. One can use the fact that an operator commutes with P^2 if and only if it commutes with P (see Exercise 10a, for example).

e. *Example.* Determine the operators U and P when where $E = L^2(m)$ (m being a measure on a measure space $(X, \mathscr{F})$) and T is defined by

$$Tf = af \quad \text{for all } f \in E,$$

for a fixed $a \in L^\infty(m)$.

12. Let T be a hermitian operator on a Hilbert space E. Show that every isolated point of the spectrum of T is an eigenvalue of T.
Hint. Let λ be an isolated point of $\sigma(T)$. Define a function f on $\sigma(T)$ by

$$f(t) = \begin{cases} 1 & \text{if } t = \lambda, \\ 0 & \text{otherwise.} \end{cases}$$

Then f is continuous on $\sigma(T)$ and $f(T) \neq 0$. Show that $(T-\lambda I)f(T) = 0$ and conclude. (You can also prove that $f(T)$ is the orthogonal projection onto $\ker(T - \lambda I)$.)

13. Let T be a hermitian operator on a Hilbert space and suppose $\sigma(T) = \{0, 1\}$. Show that T is an orthogonal projection operator.
Hint. The function f defined by $f(x) = x^2 - x$ vanishes on the spectrum of T.

14. Let m be a measure on a measure space $(X, \mathscr{F})$ and take $\varphi \in L^\infty_{\mathbb{R}}(m)$. Define an operator T on $L^2(m)$ by

$$Tu = \varphi u \quad \text{for all } u \in L^2(m).$$

Determine the operator $f(T)$, for each continuous function f.

15. *Spectral measure.* Let T be a hermitian operator on a Hilbert space E and set $X = \sigma(T)$.

a. Suppose $u, v \in E$. Show that there exists a complex Radon measure $\mu_{u,v}$ on X such that

$$\big(f(T)u \,|\, v\big) = \mu_{u,v}(f) \quad \text{for all } f \in C(X).$$

Show that, for every $u \in E$, the measure $\mu_{u,u}$ is positive.

b. Let $\mathscr{B}$ be the space of bounded Borel functions on X, and suppose $f \in \mathscr{B}$. Show that the map

$$(u, v) \mapsto \int f \, d\mu_{u,v}$$

is a sesquilinear, skew-symmetric, continuous form on E. (Sesquilinear means linear in the first argument and skew-linear in the second.) Deduce that there exists a continuous operator on E, which we denote by $f(T)$, such that

$$\big(f(T)u \,|\, v\big) = \int f \, d\mu_{u,v} \quad \text{for all } u, v \in E.$$

Check that $\|f(T)\| \le \sup_{x \in X} |f(x)|$.
Hint. Approximate f by a sequence of functions in $C(X)$ bounded by $\sup_{x \in X} |f(x)|$; then use the Dominated Convergence Theorem.

c. Show that the map from $\mathscr{B}$ to $L(E)$ taking f to $f(T)$ is a morphism of algebras and that $(f(T))^* = \bar{f}(T)$ for all $f \in \mathscr{B}$.

d. Let (f_n) be a bounded sequence in $\mathscr{B}$ that converges pointwise to a function f. Show that $\lim_{n \to +\infty} f_n(T)(u) = f(T)(u)$ for every $u \in E$.
Hint. Show that $\lim_{n \to +\infty} \big(|f_n - f|^2(T)(u) \,|\, u\big) = 0$.

e. Suppose $a \in X$ and let f_a be the restriction to X of the function $1_{(-\infty, a]}$. Show that $a \le b$ implies $f_a(T) \le f_b(T)$ (this notation means that $f_b(T) - f_a(T)$ is positive hermitian). Show also that $f_a(T)$ and $f_b(T)$ are orthogonal projection operators, as is $f_b(T) - f_a(T)$ if $a < b$.

6

Compact Operators

1 General Properties

Consider two normed spaces E and F over the same field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. As usual, we denote by $L(E, F)$ the space of continuous linear maps from E to F, and use the same notation $\|\cdot\|$ for the norm in E, in F, and in $L(E, F)$. Thus, if $T \in L(E, F)$, we have $\|T\| = \sup \{\|Tx\| : x \in E \text{ with } \|x\| \leq 1\}$.

We say that an element T in $L(E, F)$ is a **compact operator** if the image of the closed unit ball $\bar{B}(E)$ of E is a relatively compact subset of F. We denote by $\mathscr{K}(E, F)$ the set of compact operators from E to F, and we write $\mathscr{K}(E) = \mathscr{K}(E, E)$.

Clearly, an element T of $L(E, F)$ is a compact operator if and only if the image under T of every bounded subset of E is relatively compact in F.

Note that the Riesz Theorem (page 49) can be expressed as follows: The identity map on E is a compact operator from E to E if and only if E is finite-dimensional.

Examples

1. Every finite-rank operator T from E to F is compact. (Recall that an operator is said to have finite rank if its image has finite dimension, and infinite rank otherwise. The dimension of the image of a finite-rank operator is called its rank.) Indeed, T maps $\bar{B}(E)$ to a bounded, and therefore relatively compact, subset of $\operatorname{im} T$. Since any compact set in $\operatorname{im} T$ is compact in F, the image $T(\bar{B}(E))$ is relatively compact in F.

2. Consider compact metric spaces X and Y, a function $K \in C(X \times Y)$, and a (possibly complex) Radon measure μ on Y. We define an operator T_K from $C(Y)$ to $C(X)$ by

$$T_K f(x) = \int K(x,y) f(y)\, d\mu(y) \quad \text{for all } f \in C(Y) \text{ and } x \in X.$$

(In this situation the map K is called the **kernel** of the operator T_K.) The operator T_K is compact: this was proved on page 44 when μ is a positive Radon measure, and the proof can be immediately adapted to the case where μ is not necessarily positive.

3. Let a and b be real numbers such that $a < b$, and suppose $K \in C([a,b]^2)$. Let α and β be continuous functions from $[a,b]$ to $[a,b]$. For $f \in C([a,b])$ and $x \in [a,b]$, we put

$$Tf(x) = \int_{\alpha(x)}^{\beta(x)} K(x,y) f(y)\, dy.$$

The operator T thus defined from $C([a,b])$ to itself is compact.

Proof. Let $\|K\|$ be the uniform norm of K. Then, for every $f \in C([a,b])$,

$$\|Tf\| \le \|K\|\, \|f\|.$$

Therefore $T(\bar{B}(E))$ is a bounded subset of $C([a,b])$. On the other hand, if $x_1, x_2 \in [a,b]$ and $f \in C([a,b])$,

$$\begin{aligned}
&\left|Tf(x_1) - Tf(x_2)\right| \\
&\quad \le \|f\| \times \Big(\|K\| \left(|\beta(x_2) - \beta(x_1)| + |\alpha(x_2) - \alpha(x_1)|\right) \\
&\qquad\qquad + (b-a) \sup_{y \in [a,b]} \left|K(x_1,y) - K(x_2,y)\right|\Big).
\end{aligned}$$

Since K is a uniformly continuous function on $[a,b]^2$, this shows that $T(\bar{B}(E))$ is an equicontinuous subset of $C([a,b])$. The result now follows from the Ascoli Theorem (page 44). □

In particular, the integration operator

$$Tf(x) = \int_0^x f(t)\, dt$$

is a compact operator from $C([0,1])$ to itself.

4. Other examples of compact operators have been seen in the exercises: between Hölder spaces (Exercise 5 on page 45), the map $f \mapsto f$ from $C^p([0,1])$ to $C^q([0,1])$, with $q > p > 0$; and between discrete Sobolev spaces (Exercise 7d on page 104), the map $f \mapsto f$ from H^s to H^r with $r < s$.

We now study certain closure properties of compact operators.

Proposition 1.1 *$\mathscr{K}(E,F)$ is a vector subspace of $L(E,F)$.*

Proof. Consider compact operators T and S from E to F and elements $\lambda, \mu \in \mathbb{K}$. Then

$$(\lambda T + \mu S)(\bar{B}(E)) \subset \lambda \overline{T(\bar{B}(E))} + \mu\, \overline{S(\bar{B}(E))}.$$

But, if K_1 and K_2 are compact sets in F, the set $\lambda K_1 + \mu K_2$, being the image of the compact $K_1 \times K_2$ under the continuous map $(x, y) \mapsto \lambda x + \mu y$, is also compact. □

Proposition 1.2 *Let R be a compact operator from E to F. If E_1 and F_1 are normed spaces and if $T \in L(E_1, E)$ and $S \in L(F, F_1)$ are arbitrary, the composition SRT is a compact operator from E_1 to F_1.*

Proof. Indeed,

$$SRT\big(\bar{B}(E_1)\big) \subset \|T\|\, S\big(\overline{R(\bar{B}(E))}\big).$$

Since a continuous image of a compact set is compact, the result follows. □

Corollary 1.3 *$\mathscr{K}(E)$ is a two-sided ideal of the algebra $L(E)$.*

Proposition 1.4 *If F is complete, the limit in $L(E,F)$ of every convergent sequence of compact operators from E to F is a compact operator.*

Proof. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of compact operators from E to F that converges to T in $L(E,F)$. By Theorem 3.3 on page 14, it suffices to show that $T(\bar{B}(E))$ is precompact. Choose $\varepsilon > 0$ and let $n \in \mathbb{N}$ be such that $\|T - T_n\| \le \varepsilon/3$. We can cover $T_n(\bar{B}(E))$ with a finite number k of balls $B(T_n f_j, \varepsilon/3)$, where $f_1, \dots, f_k \in \bar{B}(E)$. Suppose $f \in \bar{B}(E)$ and let $j \le k$ be such that $\|T_n f - T_n f_j\| < \varepsilon/3$. By the triangle inequality, $\|Tf - Tf_j\| < \varepsilon$. Therefore

$$T(\bar{B}(E)) \subset \bigcup_{j=1}^{k} B(Tf_j, \varepsilon),$$

and $T(\bar{B}(E))$ is precompact. □

The result of the proposition can fail if F is not complete: see Exercise 8 on page 222.

Since every finite-rank operator is compact, as we saw in Example 1 on page 213, Proposition 1.4 has the following important consequence:

Corollary 1.5 *If F is complete, every limit in $L(E,F)$ of finite-rank operators is a compact operator.*

This provides a frequently useful criterion for proving that an operator is compact.

Examples

1. Let $(X, \mathscr{F})$ and $(Y, \mathscr{G})$ be measure spaces endowed with σ-finite measures m and μ, respectively. Let $p \in [1, +\infty)$ and p' be conjugate exponents, and suppose $K \in L^p(m \times \mu)$. We define an operator T_K from $L^{p'}(\mu)$ to $L^p(m)$ by setting, for every $f \in L^{p'}(\mu)$ and m-almost every $x \in X$,

$$T_K f(x) = \int K(x,y) f(y)\, d\mu(y).$$

(As in Example 2 on page 213, the map K is called the **kernel** of the operator T_K.) Then T_K is a compact operator.

Proof. We use the same notation $\|\cdot\|$ for the norms in $L^p(m \times \mu)$ and in $L(L^{p'}(\mu), L^p(m))$. We deduce easily from Hölder's inequality and Fubini's Theorem that T_K is continuous and that

$$\|T_K\| \leq \|K\|. \qquad (*)$$

Suppose that K is an element of $L^p(m) \otimes L^p(\mu)$, the vector subspace of $L^p(m \times \mu)$ spanned by the elements $f \otimes g : (x,y) \mapsto f(x)g(y)$ for $f \in L^p(m)$ and $g \in L^p(\mu)$; that is, suppose

$$K(x,y) = \sum_{j=1}^{k} f_j(x) g_j(y).$$

Then the image of T_K is contained in the span of the family $\{f_1, \ldots, f_k\}$, so T_K has finite rank.
Now, if $K \in L^p(m \times \mu)$ is arbitrary, K is the limit in $L^p(m \times \mu)$ of a sequence $(K_n)_{n \in \mathbb{N}}$ in $L^p(m) \otimes L^p(\mu)$ (see Exercise 12 on page 153). But then, by $(*)$, the sequence $(T_{K_n})_{n \in \mathbb{N}}$ converges to T_K, showing that T_K is compact by Corollary 1.5. □

Notice that the compactness of the operator considered in Example 2 on page 213 could be proved by the same method, using Example 5 on page 35.

2. *Hilbert–Schmidt operators.* Let E be an infinite-dimensional separable Hilbert space. If $(e_n)_{n \in \mathbb{N}}$ is a Hilbert basis of E, we say that an operator $T \in L(E)$ is a **Hilbert–Schmidt operator** if the series of numbers $\sum_{n=0}^{+\infty} \|Te_n\|^2$ converges. One can show (Exercise 21 on page 140) that this definition does not depend on the Hilbert basis considered. Now let P_n be the orthogonal projection from E onto the span of the family $(e_j)_{0 \leq j \leq n}$. One can show that, if T is a Hilbert–Schmidt operator, the sequence $(TP_n)_{n \in \mathbb{N}}$ converges in $L(E)$ to T (see Exercise 21 on page 140 again). Thus, every Hilbert–Schmidt operator is a compact operator.
In the case $E = L^2(m)$, where m is a σ-finite measure on a measure space $(X, \mathscr{F})$ (still assuming E separable), the Hilbert–Schmidt operators on

E are exactly the operators of the form T_K defined in the preceding example, with $K \in L^2(m \times m)$ (see Exercise 21 on page 140 once more).

We observe that, for many (but not all) Banach spaces F, Corollary 1.5 has a converse: Every compact operator from E to F is the limit of a sequence of operators of finite rank. See Exercise 24 on page 232.

1A Spectral Properties of Compact Operators

Consider again an arbitrary normed space E. We do not assume that E is complete, but we use nonetheless the notions and notation introduced in Chapter 5 (page 189): spectral values, regular values, eigenvalues, spectrum, and so on.

Proposition 1.6 *Let T be a compact operator from E to E.*

1. *The kernel of the operator $I - T$ has finite dimension.*
2. *The image of $I - T$ is closed.*
3. *The operator $I - T$ is invertible in $L(E)$ if and only if it is injective.*

Proof

1. Write $F = \ker(I - T)$. Then F is a closed subspace of E and
$$\bar{B}(F) = T(\bar{B}(F)) \subset \overline{T(\bar{B}(E))} \cap F,$$
which is compact. By the Riesz Theorem (page 49), F is finite-dimensional.
2. Take $y \in \overline{\operatorname{im}(I - T)}$ and let (x_n) be a sequence in E such that
$$\lim_{n \to +\infty} (x_n - Tx_n) = y.$$
First case: the sequence (x_n) is bounded. Since T is compact, we can assume, by passing to a subsequence if necessary, that the sequence (Tx_n) converges to some point $z \in E$. Then $\lim_{n \to +\infty} x_n = y + z$ and, by the continuity of T, we get $z = T(y+z)$, which implies that $y = (y+z) - T(y+z) \in \operatorname{im}(I - T)$.
Second case: the sequence (x_n) is not bounded. For every $n \in \mathbb{N}$, set $d_n = d(x_n, \ker(I - T))$. Since $\ker(I - T)$ is finite-dimensional by part 1, there exists a point $z_n \in \ker(I - T)$ such that $\|x_n - z_n\| = d_n$ (indeed, the continuous function $x \mapsto d(x_n, x)$ must achieve its minimum over the nonempty compact set $\bar{B}(x_n, \|x_n\|) \cap \ker(I - T)$).
If the sequence (d_n) is bounded, we can replace x_n by $x_n - z_n$ to reduce to the first case; thus $y \in \operatorname{im}(I - T)$.
Otherwise, by taking a subsequence, we can assume that the sequence $(d_n)_{n \in \mathbb{N}}$ tends to $+\infty$. Since the sequence $\big((x_n - z_n)/d_n\big)$ is bounded, we

can assume, again by passing to a subsequence, that $T\big((x_n - z_n)/d_n\big)$ converges to a point $u \in E$ (since T is compact). We deduce that

$$\lim_{n\to+\infty} d_n^{-1}(x_n - z_n) = u + \lim_{n\to+\infty} d_n^{-1}y = u,$$

which implies two things: that $Tu = u$ (by the continuity of T), so that $u \in \ker(I - T)$; and that, for n large enough, $\|x_n - z_n - d_n u\| < d_n$. But this contradicts the definition of d_n. Therefore the sequence (d_n) is bounded and $y \in \operatorname{im}(I - T)$, which proves part 2.

3. We now assume that the operator $I - T$ is injective. To prove its surjectivity, we will use a general lemma.

Lemma 1.7 *If F is a proper closed subspace of a normed vector space G, there exists $u \in G$ such that $\|u\| = 1$ and $d(u, F) \geq \frac{1}{2}$.*

Proof. Take $v \in G \setminus F$ and set $\delta = d(v, F) > 0$. Certainly there exists $w \in F$ such that $\|v - w\| < 2\delta$. Then the point $u = \|v - w\|^{-1}(v - w)$ works: if $z \in F$, we have

$$\|u-z\| = \|v-w\|^{-1}\big\|v-w-\|v-w\|z\big\| \geq \frac{1}{2\delta}\,\delta = \frac{1}{2},$$

proving the lemma. □

We now argue by contradiction. Set $E_1 = \operatorname{im}(I - T)$ and suppose that $E_1 \neq E$. For every $n \in \mathbb{N}$, set $E_n = \operatorname{im}(I - T)^n$ (and set $E_0 = E$). We show by induction that, for every $n \in \mathbb{N}$, the subspace E_n is closed, $E_n \supset E_{n+1}$, and $E_n \neq E_{n+1}$.

The claim holds for $n = 0$ by assumption. Suppose it holds for $n \in \mathbb{N}$. Clearly, $T(E_n) \subset E_n$; thus T induces an operator $T_n \in L(E_n)$. The set $T_n(\bar{B}(E_n))$ is contained in $\overline{T(\bar{B}(E))} \cap E_n$, which is compact since E_n is closed. Therefore T_n is a compact operator on E_n. Since $E_{n+1} = (I_n - T_n)(E_n)$, where I_n is the identity on E_n, part 2 above applied to T_n implies that E_{n+1} is closed in E_n and so in E. It is also clear that $E_{n+1} \supset E_{n+2}$. Finally, because we assumed $I - T$ to be injective, the subspaces $E_{n+1} = (I - T)(E_n)$ and $E_{n+2} = (I - T)(E_{n+1})$ cannot be equal since $E_n \neq E_{n+1}$. This completes the induction step.

By applying Lemma 1.7, we now obtain a sequence $(u_n)_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$,

$$u_n \in E_n, \quad \|u_n\| = 1, \quad \text{and} \quad d(u_n, E_{n+1}) \geq \tfrac{1}{2}.$$

Then, for $n < m$,

$$Tu_n - Tu_m = u_n - v_{n,m} \quad \text{with} \quad v_{n,m} = Tu_m + (I - T)u_n \in E_{n+1}.$$

It follows that

$$\|Tu_n - Tu_m\| \geq \tfrac{1}{2} \quad \text{for all } n \neq m.$$

Since every point of the sequence $(u_n)_{n\in\mathbb{N}}$ lies in $\bar{B}(E)$, this contradicts the relative compactness of $T(\bar{B}(E))$ (no subsequence of $(Tu_n)_{n\in\mathbb{N}}$ is a Cauchy sequence). This contradiction proves that $I - T$ is surjective.
There remains to show the continuity of $(I-T)^{-1}$. Here again we argue by contradiction, by assuming that there is a sequence $(x_n)_{n\in\mathbb{N}}$ that does not tend to 0 and such that $\lim_{n\to+\infty}(x_n - Tx_n) = 0$ (this condition is equivalent to $(I-T)^{-1}$ not being continuous at 0). By passing to a subsequence if necessary, we can assume that $\|x_n\| \geq \varepsilon$, for every $n \in \mathbb{N}$ and a fixed $\varepsilon > 0$. Now put $u_n = x_n/\|x_n\|$. Since T is a compact operator, we can assume, again by passing to a subsequence, that the sequence $(Tu_n)_{n\in\mathbb{N}}$ converges to a point $v \in E$. But then $\lim_{n\to\infty} u_n = v$, which implies that $\|v\| = 1$ and, by the continuity of T, that $v = Tv$, contradicting the injectivity of $I - T$. □

We can now state our main theorem, which shows that, as far as spectral properties are concerned, compact operators behave almost like operators of finite rank (see Exercise 13 on page 197).

Theorem 1.8 *Let T be a compact operator from E to E.*

1. *If E is infinite-dimensional, 0 is a spectral value of T.*
2. *Every nonzero spectral value of T is an eigenvalue of T and has a finite-dimensional associated eigenspace.*
3. *The spectrum of T is countable. If it is infinite, its nonzero elements can be arranged in a sequence $(\lambda_n)_{n\in\mathbb{N}}$ such that, for all $n \in \mathbb{N}$,*

$$|\lambda_{n+1}| \leq |\lambda_n| \quad \textit{and} \quad \lim_{n\to+\infty} \lambda_n = 0.$$

Proof

1. Suppose that 0 is not a spectral value of T. Then $I = TT^{-1}$ is a compact operator by Proposition 1.2. By the Riesz Theorem (page 49), this implies that E is finite-dimensional.
2. Take $\lambda \in \mathbb{K}^*$. Then λ is an eigenvalue of T if and only if $I - T/\lambda$ is not injective, and $\ker(\lambda I - T) = \ker(I - T/\lambda)$. On the other hand, λ is a spectral value of T if and only if $I - T/\lambda$ is not invertible in $L(E)$. Thus it suffices to apply Proposition 1.6 to prove assertion 2.
3. For assertion 3, it is enough to show that, for every $\varepsilon > 0$, there is only a finite number (perhaps 0) of spectral values λ of T such that $|\lambda| \geq \varepsilon$. Suppose, on the contrary, that, for a certain $\varepsilon > 0$, there exists a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of pairwise distinct spectral values of T such that $|\lambda_n| \geq \varepsilon$ for every $n \in \mathbb{N}$. By part 2, all the λ_n are eigenvalues of T. Thus there exists a sequence (e_n) of elements of E of norm 1 such that $Te_n = \lambda_n e_n$ for every $n \in \mathbb{N}$. Since the eigenvalues λ_n are pairwise distinct, it is easy to see (and it is a classical result) that the family $\{e_n\}_{n\in\mathbb{N}}$ is linearly independent. For each $n \in \mathbb{N}$, let E_n be the span of

the $n+1$ first vectors $e_0, \dots, e_n$. The sequence $(E_n)_{n\in\mathbb{N}}$ is then a strictly increasing sequence of finite-dimensional spaces. By Lemma 1.7, there exists a sequence $(u_n)_{n\in\mathbb{N}}$ of vectors of norm 1 such that, for every integer $n \in \mathbb{N}$,

$$u_n \in E_{n+1} \quad \text{and} \quad d(u_n, E_n) \geq \tfrac{1}{2}$$

(in fact, since E_n has finite dimension, we could replace $\frac{1}{2}$ by 1 here). Define $v_n = \lambda_{n+1}^{-1} u_n$. The sequence (v_n) is bounded by $1/\varepsilon$. Moreover, if $n > m$,

$$Tv_n - Tv_m = u_n - v_{n,m} \quad \text{with} \quad v_{n,m} = Tv_m + \frac{1}{\lambda_{n+1}}(\lambda_{n+1} I - T)u_n.$$

But $Tv_m \in E_{m+1} \subset E_n$ and $(\lambda_{n+1} I - T)(E_{n+1}) \subset E_n$. Thus $v_{n,m} \in E_n$ and $\|Tv_n - Tv_m\| \geq \frac{1}{2}$, contradicting the compactness of T (the sequence $(v_n)_{n\in\mathbb{N}}$ is bounded and its image under T has no Cauchy subsequence, hence no convergent subsequence). □

Example. We now discuss a compact operator whose spectrum is countably infinite, and we determine this spectrum explicitly. Consider the operator T on the space $C([0,1])$ (with the uniform norm) defined by

$$Tf(x) = \int_0^{1-x} f(t)\,dt \quad \text{for all } f \in C([0,1]).$$

We know from Example 3 on page 214 that T is a compact operator. By Theorem 1.8, zero is a spectral value of T, but clearly it is not an eigenvalue. To determine the spectrum explicitly, it is enough to find the eigenvalues. Let λ be an eigenvalue of T and let $g \in C([0,1])$ be a corresponding nonzero eigenvector, so that

$$\lambda g(x) = \int_0^{1-x} g(t)\,dt \quad \text{for all } x \in [0,1].$$

Since λ is nonzero, g is necessarily of class C^1 in $[0,1]$; moreover $g(1) = 0$ and

$$\lambda g'(x) = -g(1-x) \quad \text{for all } x \in [0,1].$$

It follows that g is of class C^2 in $[0,1]$ and that

$$g \neq 0, \quad g(1) = 0, \quad g'(0) = 0, \quad \lambda g'(1) = -g(0), \tag{$*$}$$

$$\lambda g''(x) = -g(x)/\lambda \quad \text{for all } x \in [0,1]. \tag{$**$}$$

The solutions of the differential equation $(**)$ satisfying $g'(0) = 0$ are the functions $g(x) = A\cos(x/\lambda)$. In order for such a function to satisfy conditions $(*)$, it is necessary that $\cos(1/\lambda) = 0$ and $\sin(1/\lambda) = 1$, which is to say $1/\lambda = \pi/2 + 2k\pi$, with $k \in \mathbb{Z}$, or yet

$$\lambda = \frac{1}{\pi/2 + 2k\pi}, \quad \text{with } k \in \mathbb{Z}.$$

Conversely, if $\lambda = 1/(\pi/2 + 2k\pi)$ with $k \in \mathbb{Z}$, one easily checks that the function g defined by $g(x) = \cos(x/\lambda)$ is an eigenvector of T associated with λ. Thus

$$\sigma(T) = \{0\} \cup \left\{\frac{1}{\pi/2 + 2k\pi} : k \in \mathbb{Z}\right\}.$$

We also see that all the eigenspaces of T have dimension 1 and that the spectral radius of T is $2/\pi$.

Exercises

1. Let E be an infinite-dimensional Banach space and F any normed vector space. Let T be an operator from E to F for which there exists a constant $\alpha > 0$ such that $\|Tx\| \geq \alpha\|x\|$ for every $x \in E$. Show that T is not compact.

2. Let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of complex numbers and let T be the operator on ℓ^p (where $p \in [1, +\infty)$) defined by

$$Tf(n) = \lambda_n f(n) \quad \text{for all } f \in \ell^p \text{ and } n \in \mathbb{N}.$$

We know from Exercise 4 on page 195 that T is continuous if and only if the sequence $(\lambda_n)_{n\in\mathbb{N}}$ is bounded.

a. Show that T is compact if and only if $\lim_{n\to+\infty} \lambda_n = 0$.
Hint. You might use Exercise 10 on page 183, for example.

b. Suppose $p = 2$. Show that T is a Hilbert–Schmidt operator if and only if

$$\sum_{n\in\mathbb{N}} |\lambda_n|^2 < +\infty.$$

c. Let S be the right shift in ℓ^p, where $p \in [1, +\infty)$ (see Exercise 6e on page 196). Is S a compact operator?

d. Suppose that the sequence $(\lambda_n)_{n\in\mathbb{N}}$ tends to 0. Determine the eigenvalues and the spectral values of TS.

3. Let X be a compact metric space and suppose $\varphi \in C(X)$. Show that the operator T on $C(X)$ defined by $Tf = \varphi f$ is compact if and only if φ vanishes on every cluster point of X.
Hint. Suppose that T is compact and that $|\varphi(x)| > 0$ at a point $x \in X$. Then there exists a closed neighborhood Y of x on which $|\varphi| > 0$. Show that the restriction of T to $C(Y)$ is an invertible compact operator in $L(C(Y))$ (to show compactness you will probably need Tietze's Extension Theorem, Exercise 7a on page 40). Deduce that Y is finite. For the converse, use Ascoli's Theorem, page 44.

4. Let P be a polynomial not vanishing at 0 and let T be a linear operator on an infinite-dimensional normed space E. Assume $P(T) = 0$. Show that T is not compact.

5. Let E be a Hilbert space and suppose $T \in L(E)$. Show that T is a compact operator if and only if T^* is one.

Hint. Let (x_n) be a bounded sequence in E. Put $M = \sup_n \|x_n\|$ and define $y_n = T^* x_n$ for each integer n. Show that, for every $n, m \in \mathbb{N}$,

$$\|y_n - y_m\|^2 \leq 2M \|Ty_n - Ty_m\|.$$

Deduce that T^* is compact.

6. **a.** Let T be a continuous operator on a Hilbert space E. Show that T is compact if and only if the image under T of every sequence in E that converges weakly to 0 is a sequence that converges (strongly) to 0.

Hint. For the "if" part, use Exercise 12 on page 121 and Proposition 3.8 on page 116. For the converse, use Theorem 3.7 on page 115.

b. Show that this result remains true if $E = L^p(m)$, where m is a σ-finite measure on a measure space $(X, \mathscr{F})$ and $p \in (1, +\infty)$. (Weak convergence in $L^p(m)$ was defined in Exercise 9 on page 166. You can also use Exercise 10 on page 168.)

c. Show that this result is false if $E = \ell^1$.

Hint. Use Exercise 9d on page 167.

7. Let μ be a positive Radon measure on a compact metric space X, with support equal to X. Suppose $K \in C(X \times X)$. Fix $p \in [1, \infty)$ and denote by E_p the space $C(X)$ with the norm induced by that of $L^p(\mu)$. Define an operator T from E_p to itself by

$$Tf(x) = \int K(x, y) f(y)\, d\mu(y) \quad \text{for all } x \in X.$$

Show that T is compact, and deduce that the spectrum of T does not depend on p.

8. Let E be the space $C^1([0,1])$ with the norm $\|\cdot\|_E$ defined by

$$\|f\|_E = \|f\| + \|f'\|,$$

where $\|\cdot\|$ denotes the uniform norm on $[0,1]$. Let F be the same space $C^1([0,1])$ with the uniform norm on $[0,1]$. Let T be the operator from E to F defined by

$$Tf = f \quad \text{for all } f \in C^1([0,1]).$$

a. Show that the norm of T equals 1.

b. Show that T is not compact.

c. Let (T_n) be the sequence in $L(E, F)$ defined by

$$T_n f = B_n f \quad \text{for all } f \in E,$$

where B_n is the Bernstein operator defined in Exercise 3 on page 37. Show that each T_n has finite rank and that the sequence (T_n) converges to T in $L(E, F)$.

Hint. Using the estimates from Exercise 3 on page 37, show that $\|T - T_n\| \le (2n)^{-1/3}$.

d. Deduce that the hypothesis that F is complete cannot be omitted from Proposition 1.4 or Corollary 1.5.

9. Suppose $p \in [1, \infty]$. Define an operator T on the space $L^p([0,1])$ by setting

$$Tf(x) = \int_0^{1-x} f(t)\,dt \quad \text{for all } f \in L^p([0,1]) \text{ and } x \in [0,1].$$

Show that T is compact and determine its spectrum.

Hint. Notice that any eigenvector associated with a nonzero eigenvalue must be a continuous map. Therefore the eigenvalues can be determined as in the text; see page 220. (In particular, the spectrum of T does not depend on p.)

10. Let $\hat{E}$ and $\hat{F}$ be Banach spaces and let E and F be dense subspaces of $\hat{E}$ and $\hat{F}$, respectively. Consider a compact operator T from E to F.

a. Show that T can be extended in a unique way to a continuous operator $\hat{T}$ from $\hat{E}$ to $\hat{F}$. Show that $\hat{T}$ is compact and that $\operatorname{im} \hat{T} \subset F$. Deduce that $\hat{T}$ is also compact, when considered as an operator from $\hat{E}$ to F.

b. Assume $\hat{E} = \hat{F}$ and $E = F$. Show that T and $\hat{T}$ have the same nonzero eigenvalues and that the eigenspace associated with a given nonzero eigenvalue is the same for T and $\hat{T}$.

c. Apply this to Exercise 7 above in order to show that the study of the spectrum of the operator T on E_p is reducible to the study of a compact operator $\hat{T}$ on $L^p(\mu)$.

11. Let E be one of $C([0,1])$ or $L^p([0,1])$, where $p \in [1, \infty]$. Determine the spectrum of the operator T from E to itself defined by

$$Tf(x) = \int_0^1 \min(x,y) f(y)\,dy = \int_0^x y f(y)\,dy + x \int_x^1 f(y)\,dy.$$

Hint. Note that T is compact and that an eigenvector f of T associated with a nonzero eigenvalue is a differentiable function and satisfies $f(0) = f'(1) = 0$.

12. Let T be the linear operator on $L^2((0,1))$ defined by

$$Tf(x) = \int_0^1 e^{-|x-y|} f(y)\,dy \quad \text{for all } f \in L^2((0,1)) \text{ and } x \in [0,1].$$

a. Show that T is a compact hermitian operator and that $\|T\| \le 1$.

b. Suppose $f \in C([0,1])$ and put $g = Tf$. Show that $g \in C^2([0,1])$ and that $g(0) = g'(0)$, $g(1) = -g'(1)$, and

$$g''(x) - g(x) = -2f(x) \quad \text{for all } x \in [0,1].$$

c. Conversely, suppose $g \in C^2([0,1])$ satisfies $g(0) = g'(0)$ and $g(1) = -g'(1)$. Put $f = -(g'' - g)/2$. Show that $g = Tf$.
Hint. Consider $h = g - Tf$.

d. Show that $\operatorname{im} T$ is dense in $L^2((0,1))$ and deduce that 0 is not an eigenvalue of T. Is 0 a spectral value of T?
Hint. For denseness, note that, by part c, $\operatorname{im} T$ contains the space $C_c^2((0,1))$ of C^2 functions with compact support in $(0,1)$.

e. Show that, if $f \in C([0,1])$ and $g = Tf$, then

$$(Tf \mid f) = \frac{1}{2}\biggl(\int_0^1 \bigl(|g(x)|^2 + |g'(x)|^2\bigr)\,dx + |g(1)|^2 + |g(0)|^2\biggr).$$

Deduce that, for every $f \in L^2((0,1))$,

$$(Tf \mid f) \ge \tfrac{1}{2}\,\|Tf\|^2.$$

f. Show that $\sigma(T) \subset [0,1]$.

g. For $\lambda \in (0,1]$, set $a_\lambda = \sqrt{(2-\lambda)/\lambda}$. Show that $\lambda \in \sigma(T)$ if and only if

$$(1 - a_\lambda^2)\sin a_\lambda + 2a_\lambda \cos a_\lambda = 0.$$

Deduce that $\sigma(T) = \{0\} \cup \{\lambda_n\}_{n\in\mathbb{N}}$, where, for every $n \in \mathbb{N}$,

$$\frac{2}{1 + (\pi/2 + n\pi)^2} < \lambda_n < \frac{2}{1 + (n\pi)^2}.$$

13. *A Sturm–Liouville problem.* Suppose $g \in C([0,1])$, and consider the differential equation

$$(pf')' - qf = g, \tag{E$_g$}$$

on the interval $[0,1]$, with boundary conditions

$$\alpha_0 f'(0) - \alpha_1 f(0) = 0, \quad \beta_0 f'(1) - \beta_1 f(1) = 0. \tag{BC}$$

Here q is a continuous function on $[0,1]$ and p is a function of class C^1 on $[0,1]$ taking positive values only; in addition we assume that $(\alpha_0, \alpha_1) \ne (0,0)$ and $(\beta_0, \beta_1) \ne (0,0)$. By definition, a solution of the problem $(\mathrm{E})_g$ + (BC) is a function f of class C^2 on the interval $[0,1]$ satisfying conditions $(\mathrm{E})_g$ and (BC).

a. Suppose for now that the boundary value problem $(\mathrm{E})_0$ + (BC) has only the trivial solution (identically zero).

i. Take a nontrivial solution f_1 of (E)$_0$ with $\alpha_0 f_1'(0) - \alpha_1 f_1(0) = 0$ and a nontrivial solution f_2 of (E)$_0$ with $\beta_0 f_2'(1) - \beta_1 f_2(1) = 0$. Justify the existence of f_1 and f_2 and prove that the expression

$$W = \big(f_1'(x) f_2(x) - f_1(x) f_2'(x)\big) p(x)$$

is constant and nonzero on $[0, 1]$.

ii. Define a function G on $[0, 1]^2$ by

$$G(x, y) = \begin{cases} -\dfrac{f_2(y) f_1(x)}{W} & \text{if } 0 \le x \le y \le 1, \\ -\dfrac{f_1(y) f_2(x)}{W} & \text{if } 0 \le y \le x \le 1. \end{cases}$$

(G is the *Green's function* associated with the problem (E) + (BC).) Let T be the operator from $C([0, 1])$ to itself defined by

$$Tf(x) = \int_0^1 G(x, y) f(y)\, dy.$$

Show that T is compact and that, if $g \in C([0, 1])$, the function $f = Tg$ is the unique solution of (E)$_g$ + (BC).

iii. A. Show that $\operatorname{im} T$ equals the set of functions of class C^2 on $[0, 1]$ that satisfy (BC).

B. Take $\lambda \in \mathbb{K}^*$. Show that $\ker(\lambda I - T)$ equals the set of solutions on $[0, 1]$ of the equation

$$(py')' - (q + \lambda^{-1}) y = 0$$

that satisfy (BC). Deduce that $\ker(\lambda I - T)$ has dimension at most 1.

b. Suppose that $\alpha_0 \alpha_1 = \beta_0 \beta_1 = 0$, that q is nonnegative-valued, and that, if $\alpha_1 = \beta_1 = 0$, then q is not identically zero. Show that the problem (E)$_0$ + (BC) has only the trivial solution.
Hint. Let f be a solution of (E)$_0$ + (BC). Show that

$$\int_0^1 q(t)\, |f(t)|^2\, dt + \int_0^1 p(t)\, |f'(t)|^2\, dt = 0.$$

c. Study the particular case $p = 1$, $q = 0$, $\alpha_0 = \beta_1 = 0$. Write down the corresponding function G. Compare with Exercise 11.

d. Suppose that $\alpha_0 = \beta_0 = 0$ and that $q(x) > 0$ for every $x \in (0, 1)$.

i. Show that $g \ge 0$ implies $-Tg \ge 0$; thus $-T$ is a positive operator.
Hint. Suppose $g \ge 0$ and write $f = Tg$. Check that f is real-valued. Suppose next that there exists a point $x \in [0, 1]$ such that $f(x) > 0$, and work with a point of $[0, 1]$ where f achieves its maximum.

ii. Deduce that $G(x,y) \le 0$ for every $(x,y) \in [0,1]^2$.

iii. Show that this remains true if we assume only that $q \ge 0$.
Hint. Approximate q by $q+\varepsilon$ and prove that the kernel G_ε corresponding to $q+\varepsilon$ converges to G.

14. *A particular case of Krein–Rutman Theorem.* Let X be a compact metric space. Consider $E = C^{\mathbb{C}}(X)$ and let T be a positive compact operator from E to itself. We wish to show, among other things, that if T has a positive spectral radius $r(T)$, it has a nonzero, nonnegative-valued eigenvector associated with the eigenvalue $\lambda = r(T)$.
Denote by E^+ the set of $f \in E$ such that $f \ge 0$, and define $E^{+*} = E^+ \setminus \{0\}$.

a. For $f \in E^{+*}$, we put

$$r(f) = \max\{\rho \in \mathbb{R}^+ : \rho f \le Tf\}$$

and

$$r = \sup\{r(f) : f \in E^{+*}\}.$$

Show that r is well defined and that $r = r(T)$.
Hint. You might have to use Exercise 3a on page 195.

b. We suppose for now that if $f \in E^{+*}$ then $Tf(x) > 0$ for all $x \in X$.

i. Show that $r > 0$.

ii. Show that there exists an element $g \in E^{+*}$ such that $r = r(g)$.
Hint. Check that there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of elements of E^{+*} of norm 1 such that $\lim_{n\to+\infty} r(f_n) = r$ and that, by passing to a subsequence, one can assume that the sequence $(Tf_n)_{n\in\mathbb{N}}$ converges to some element g of E. Show that $g \in E^{+*}$ and that $r(g) \ge r$. Wrap up.

iii. Show that $Tg = rg$.
Hint. Show that, if $Tg \ne rg$, we have $rTg(x) < T(Tg)(x)$ for every x; then finish.

iv. We will show that the eigenspace E_r associated with the eigenvalue r has dimension 1.

A. Show that, if $h \in E_r$, the functions $(\operatorname{Re} h)^+$, $(\operatorname{Re} h)^-$, $(\operatorname{Im} h)^+$, and $(\operatorname{Im} h)^-$ belong to E_r.
Hint. Work as in part b-iii. Observe that, for example,

$$T((\operatorname{Re} h)^+) \ge (T(\operatorname{Re} h))^+ = r(\operatorname{Re} h)^+.$$

B. Let $h \in E_r$ be such that $h \ge 0$. Show that there exists $\rho \ge 0$ such that $h = \rho g$.
Hint. Consider $\rho = \max\{\lambda \ge 0 : \lambda g \le h\}$. If $h - \rho g \ne 0$, we have $h(x) - \rho g(x) > 0$ for all $x \in X$, which leads to a contradiction.

C. Deduce that E_r is spanned by g.

c. Let μ be a positive Radon measure on X, of support equal to X. (Why is there such a measure?) For $\varepsilon > 0$, put $T_\varepsilon f = Tf + \varepsilon \int f\,d\mu$.

i. Show that T_ε is a compact operator in E and that $T_\varepsilon f(x) > 0$ for every $f \in E^{+*}$ and $x \in X$.

ii. Show that, if $r > 0$, there exists $g \in E^{+*}$ such that $Tg = rg$.
Hint. Let r_ε be the positive real number associated with T_ε and take $g_\varepsilon \in E^{+*}$ of norm 1 and such that $T_\varepsilon g_\varepsilon = r_\varepsilon g_\varepsilon$. Show that $r_\varepsilon \geq r$; then that there exists a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ approaching 0 and such that g_{ε_n} converges to g. (Observe that in this case the eigenspace associated to r need not have dimension 1.)

15. Let m be a measure of finite mass on a measure space $(X, \mathscr{F})$, and take $K \in L^\infty(m \times m)$. Show that, for every $p, q \in (1, +\infty)$, the operator T defined from $L^p(m)$ to $L^q(m)$ by $Tf(x) = \int K(x,y) f(y)\,dm(y)$ is compact.
Hint. Use Example 1 on page 216 and the fact that, if $s \geq r$, the canonical injection $f \mapsto f$ from $L^s(m)$ to $L^r(m)$ is continuous.

16. Take $p \in [1, \infty]$. Consider a σ-finite measure m on a measure space $(X, \mathscr{F})$ and a map $K : X^2 \to \mathbb{K}$ that is measurable (with respect to the product σ-algebra on X^2) and such that the expression

$$C_K = \max\Bigl(\sup_{x\in X} \int |K(x,y)|\,dm(y), \sup_{y\in X} \int |K(x,y)|\,dm(x)\Bigr)$$

is finite.

a. Show that the equation

$$T_K f(x) = \int K(x,y) f(y)\,dm(y)$$

defines a continuous operator T_K from $L^p(m)$ to itself of norm at most C_K.
Hint. Write $|K(x,y)| = |K(x,y)|^{1/p} |K(x,y)|^{1/p'}$, where p' is the conjugate exponent of p.

b. Suppose that m is Lebesgue measure on the Borel σ-algebra of $X = [0,1]$ and that $K(x,y) = |x-y|^{-\alpha}$, with $\alpha \in (0,1)$.

i. Check that K satisfies the assumptions of part a.

ii. For each $n \in \mathbb{N}$, set $K_n = \inf(K, n)$. Show that the operators T_{K_n} from $L^p([0,1])$ to itself are compact.
Hint. Note that $K_n \in C([0,1]^2)$.

iii. Show that, for every $n \in \mathbb{N}^*$,

$$C_{K-K_n} \leq \frac{2}{1-\alpha}\, n^{-(1-\alpha)/\alpha}.$$

iv. Deduce that the operator T_K from $L^p([0,1])$ to itself is compact. (See also Exercise 21e below.)

17. *Two examples of noncompact kernel operators.*

a. We consider the operator T defined on $L^p((0,+\infty))$ in Exercise 3 on page 149 and we maintain the assumptions and notation of part c of that exercise. Set $\varphi_\varepsilon = \varepsilon^{1/p} f_\varepsilon$. Show that $\|\varphi_\varepsilon\|_p = 1$,

$$\lim_{\varepsilon \to 0} \|T\varphi_\varepsilon\|_p = k,$$

and that, for every $x > 0$,

$$\lim_{\varepsilon \to 0} T\varphi_\varepsilon(x) = 0.$$

Deduce that, unless T is the zero operator, it cannot be a compact operator on $L^p((0,+\infty))$.

b. Let T be the operator defined on $L^p((0,+\infty))$, with $p \in (1,\infty)$, by

$$Tf(x) = \frac{1}{x}\int_0^x f(y)\,dy.$$

Using the last part of Exercise 2 on page 177, prove that T is not a compact operator on $L^p((0,+\infty))$.

18. For $r \in [0,1)$, we define an operator T_r on the Hilbert space ℓ^2 by

$$(T_r u)(n) = r^n u(n).$$

a. Show that T_r is compact for any $r \in [0,1)$ (see Exercise 2).

b. Consider a sequence (r_n) in $[0,1)$ converging to 1 and a bounded sequence $(u^{(n)})$ in ℓ^2 converging weakly to u. Show that the sequence

$$(u^{(n)} - T_{r_n} u^{(n)})_{n \in \mathbb{N}}$$

converges weakly to 0.

Hint. Show first that, for every $v \in \ell^2$, the sequence $(T_{r_n} v)_{n\in\mathbb{N}}$ converges (strongly) to v in ℓ^2.

c. Deduce that, if T is a compact operator from ℓ^2 to itself, then

$$\lim_{r \to 1^-} \|TT_r - T\| = 0.$$

Hint. Reason by contradiction and use Exercise 6a.

d. Show that, if T is a compact operator from ℓ^2 to itself, we have

$$\lim_{r \to 1^-} T_r T = \lim_{r \to 1^-} T_r T T_r = T$$

in $L(\ell^2)$.

Hint. Show first that $\lim_{r\to 1^-} T_r T = T$, using Exercise 1 on page 20.

19. *Hankel operators.* For $f \in L^\infty([0,1])$, we set

$$c_n(f) = \int_0^1 f(t) e^{-2i\pi nt}\, dt \quad \text{for all } n \in \mathbb{N}.$$

We associate with f a linear map T_f on ℓ^2 by setting

$$(T_f u)(p) = \sum_{n=0}^{+\infty} u(n) c_{n+p}(f) \quad \text{for all } p \in \mathbb{N}.$$

a. Suppose $u \in \ell^2$. We denote by $\tilde{u}$ the sum in $L^2([0,1])$ of the series $\sum_{n=0}^{+\infty} u(n) e^{-2i\pi nt}$. Show that, for every integer $p \in \mathbb{N}$,

$$(T_f u)(p) = \int_0^1 f(t) \tilde{u}(t) e^{-2i\pi pt}\, dt.$$

Deduce that the operator T_f from ℓ^2 to itself is continuous and that its norm is at most $\|f\|_\infty$.

b. Show that, if there exists $N \in \mathbb{N}$ such that

$$c_n(f) = 0 \quad \text{for all } n \geq N,$$

the operator T_f has finite rank. Deduce that, if f is continuous on $[0,1]$ and $f(0) = f(1)$, then T_f is compact (as an operator from ℓ^2 to itself).

c. If $f \in L^\infty([0,1])$ and $r \in [0,1)$, put

$$f_r(t) = \sum_{n=0}^{+\infty} r^n c_n(f) e^{2i\pi nt}.$$

i. Show that this series converges uniformly, that f_r is continuous on $[0,1]$, and that $f_r(0) = f_r(1)$.

ii. Show that if T_f is compact we have

$$\lim_{r \to 1^-} \|T_{f_r} - T_f\| = 0.$$

Hint. Use Exercise 18d.

iii. Show that, if T_f is compact, there exists a sequence (φ_n) in the span of the functions $t \mapsto e^{2i\pi kt}$ (where $k \in \mathbb{N}$) such that the sequence (T_{φ_n}) converges to T_f in $L(\ell^2)$.

20. Let E be a normed space having an order relation $\leq$ compatible with addition and multiplication by positive scalars, and such that, for all $f, g \in E$, the condition $0 \leq f \leq g$ implies $\|f\| \leq \|g\|$. Suppose also that the set of nonnegative elements is closed in E. (For example, all the function spaces studied in the preceding chapters, such as L^p, $C_b(X)$, and $C_0(X)$, satisfy these properties when given the natural order.) Let T be a positive compact operator on E (positive means that $Tf \geq 0$ for all $f \in E$ such that $f \geq 0$), and suppose $\lambda \in \mathbb{R}^{+*}$.

a. Take $h \in E$. Suppose that there exist elements f_0 and g_0 of E such that
$$f_0 \le g_0, \quad Tf_0 \ge \lambda f_0 + h, \quad \text{and} \quad Tg_0 \le \lambda g_0 + h.$$
Show that the sequences $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ defined by
$$f_{n+1} = \frac{Tf_n - h}{\lambda} \quad \text{and} \quad g_{n+1} = \frac{Tg_n - h}{\lambda}$$
for all $n \in \mathbb{N}$ converge to two (not necessarily equal) solutions f_∞ and g_∞ of the equation
$$Tf = \lambda f + h$$
satisfying $f_0 \le f_\infty \le g_\infty \le g_0$. (In particular, if $h = 0$ and the inequalities $f_0 \le 0$ and $g_0 \ge 0$ are not both true, λ is an eigenvalue of T. Compare with Exercise 16 on page 198.)

b. Take $E = C([0,1])$, define T by
$$Tf(x) = \int_0^1 K(x,y) f(y)\, dy \quad \text{for all } f \in C([0,1]) \text{ and } x \in [0,1],$$
where K is a continuous map on $[0,1]^2$ with values in $\left[0, \frac{1}{2}\right]$, and let k be an element of $C^{\mathbb{R}}([0,1])$ taking values in $[0,1]$. Show that the two sequences (f_n) and (g_n) defined as above with $f_0 = 0$, $g_0 = 2$, $h = -k$, and $\lambda = 1$ converge to the unique solution f of the equation
$$f(x) - \int_0^1 K(x,y) f(y)\, dy = k(x) \quad \text{for all } x \in [0,1].$$

21. Let X and Y be compact metric spaces.

a. Let $\mu : y \mapsto \mu_y$ be a map from Y to the space $\mathfrak{M}^{\mathbb{K}}(X)$ of Radon measures on X. Assume that μ is *weakly continuous* in the following sense: for every $f \in C(X)$, the map $y \mapsto \int f\, d\mu_y$ from Y to $\mathbb{K}$ is continuous. (You might check that μ is weakly continuous if and only if it takes convergent sequences in Y to weakly convergent sequences of measures on X; see exercise 7 on page 91.) For $m \in \mathfrak{M}(X)$, denote by $\|m\|$ the norm of m, considered as an element of the topological dual of $C(X)$.

i. Define $|\mu| = \sup_{y\in Y} \|\mu_y\|$. Show that $|\mu| < +\infty$.
Hint. Use the Banach–Steinhaus Theorem, page 22.

ii. Show that the equation
$$T_\mu f(y) = \int f(x)\, d\mu_y(x) \quad \text{for all } f \in C(X) \text{ and } y \in Y$$
defines a continuous linear operator T_μ from $C(X)$ to $C(Y)$, of norm $|\mu|$.

b. Conversely, prove that, for every continuous linear operator T from $C(X)$ to $C(Y)$, there exists a weakly continuous map μ from Y to $\mathfrak{M}(X)$ such that $T = T_\mu$.

c. Let μ be a weakly continuous map from Y to $\mathfrak{M}(X)$. Show that the operator T_μ is compact if and only if μ is continuous as a map from Y to the Banach space $\mathfrak{M}(X) = C(X)'$.
Hint. Use Ascoli's Theorem, page 44.

d. Let T be a continuous linear operator from $C(X)$ to $C(Y)$. Show that T is compact if and only if there exists a map K from $Y \times X$ to $\mathbb{K}$ and a positive Radon measure m on X such that

$$(*) \quad Tf(y) = \int K(y,x) f(x)\, dm(x) \quad \text{for all } f \in C(X) \text{ and } y \in Y,$$

the map K being required to satisfy the following conditions:
- For every $y \in Y$, the map $K_y : x \mapsto K(y,x)$ belongs to $L^1(m)$.
- The map $y \mapsto K_y$ from Y to $L^1(m)$ is continuous.

Show also that, in this case,

$$\|T\| = \sup_{y \in Y} \int \bigl|K(y,x)\bigr|\, dm(x).$$

Hint. For necessity, use Exercise 8 on page 91, then the Radon–Nikodým Theorem (Exercise 6 on page 165), and Exercise 4 on page 90.

e. Take $\alpha \in (0,1)$. Show that the operator T from $C([0,1])$ to itself defined by

$$Tf(x) = \int_0^1 |x-y|^{-\alpha} f(y)\, dy$$

is compact. Find its norm. (See also Exercise 16b above.)

22. Let X be a compact metric space and m a σ-finite measure on a measure space $(\Omega, \mathscr{F})$. Let $p \in (1,\infty)$ and p' the conjugate exponent.

a. Let K be a function from $X \times \Omega$ to $\mathbb{K}$ satisfying these conditions:

(H1) For every $x \in X$, the function $K_x : s \mapsto K(x,s)$ belongs to $L^{p'}(m)$.

(H2) The map $x \mapsto K_x$ takes convergent sequences in X to weakly convergent sequences in $L^{p'}(m)$.

(Weak convergence in $L^{p'}$ is defined in Exercise 9 on page 166.)

i. Check that (H2) is equivalent to the following property:

(H2)′ For every $f \in L^p$, the map $x \mapsto \displaystyle\int K(x,s) f(s)\, dm(s)$ from X to $\mathbb{K}$ is continuous.

ii. Define $|K| = \sup_{x\in X} \|K_x\|_{p'}$. Show that $|K|$ is finite.
Hint. Consider $(K_x)_{x\in X}$ as a family of continuous linear forms on $L^p(m)$ and use the Banach–Steinhaus theorem, page 22.

iii. For $f \in L^p(m)$ and $x \in X$, put

$$T_K f(x) = \int K(x,s) f(s)\, dm(s).$$

Show that the linear operator T_K from $L^p(m)$ to $C(X)$ thus defined is continuous and has norm $|K|$.

b. Conversely, prove that, for every continuous linear operator from $L^p(m)$ to $C(X)$, there exists a function K from $X \times \Omega$ to $\mathbb{K}$ satisfying conditions (H1) and (H2) and such that $T = T_K$.
Hint. Use Theorem 2.1 on page 159.

c. Let K be a function from $X \times \Omega$ to $\mathbb{K}$ satisfying conditions (H1) and (H2). Show that the operator T_K is compact if and only if the map $x \mapsto K_x$ from X to $L^{p'}(m)$ is continuous.
Hint. Use Ascoli's Theorem, page 44.

23. Let E be a normed space. Suppose there exists a sequence $(P_n)_{n\in\mathbb{N}}$ in $L(E)$ consisting of finite-rank operators of norm at most 1 and such that $\lim_{n\to+\infty} P_n x = x$ for every $x \in E$. (We know that this is the case for $E = C_0(X)$ when X is a locally compact separable metric space (see Exercises 1 on page 30 and 11 on page 56), and also when $E = L^p(m)$, if $p \in [1, +\infty)$ and m is a measure of finite mass on a measure space $(X, \mathscr{F})$ whose σ-algebra is separable (Exercise 14c on page 155).)

a. Show that E is separable.

b. Show that every separable scalar product space has the property that we are assuming about E.
Hint. Let $(e_n)_{n\in\mathbb{N}}$ be a Hilbert basis. Take for P_n the projection onto the finite-dimensional vector space spanned by $(e_i)_{i\le n}$.

c. Show that every compact operator from a normed space F to E is the limit in $L(F,E)$ of a sequence of operators of finite rank.
Hint. If T is a compact operator from F to E, consider $T_n = P_n T$ and use Exercise 1 on page 20.

24. (This exercise generalizes the preceding one to the case of nonseparable normed spaces.) A normed space E is said to have the *approximation property* if, for every compact K in E, there exists a sequence $(P_n)_{n\in\mathbb{N}}$ in $L(E)$ consisting of operators of finite rank that converges to the identity I uniformly on K; in symbols,

$$\lim_{n\to+\infty} \sup_{x\in K} \|P_n x - x\| = 0.$$

a. Let E be a normed space having this property. Show that every compact operator from a normed space F to E is the limit in $L(F,E)$ of a sequence of finite-rank operators.

b. Show that every scalar product space satisfies the approximation property.
Hint. If K is compact, the vector space spanned by K is separable.

c. Show that, for every $p \in [1, +\infty)$ and every measure m on a measure space $(X, \mathscr{F})$, the space $L^p(m)$ has the approximation property.
Hint. For any integer $n \in \mathbb{N}$, a compact K in $L^p(m)$ can be covered by finitely many balls $B(f_1^n, 1/n), \ldots, B(f_{j_n}^n, 1/n)$. Now apply the result of Exercise 14b on page 154 to each f_j^n.

25. Let T be a compact operator on a normed space E. Let λ be a nonzero eigenvalue of T, and put $S = T - \lambda I$.

a. Show that, for every integer n,

$$\ker S^n \subseteq \ker S^{n+1}, \quad S(\ker S^{n+1}) \subseteq \ker S^n, \quad T(\ker S^n) \subseteq \ker S^n.$$

b. Deduce that there exists an integer n for which $\ker S^n = \ker S^{n+1}$.
Hint. Assuming otherwise, prove that one can construct a sequence (x_n) such that, for every n,

$$x_n \in \ker S^{n+1}, \quad \|x_n\| \leq 1, \quad d(x_n, \ker S^n) \geq \tfrac{1}{2}.$$

Show that for any two distinct integers m, n, we have $\|Tx_n - Tx_m\| \geq |\lambda|/2$, which is absurd.

In the sequel n will denote the smallest integer for which $\ker S^n = \ker S^{n+1}$. This integer is called the *index* of the eigenvalue λ.

c. Show that $\ker S^n = \ker S^{n+k}$ for every integer $k \in \mathbb{N}$.

d. Show that $\ker S^n$ and $\operatorname{im} S^n$ are closed and that $\ker S^n \cap \operatorname{im} S^n = \{0\}$.

e. Show that the restrictions of S and S^n to $\operatorname{im} S^n$ are invertible elements of $L(\operatorname{im} S^n)$.

f. Deduce from the preceding results that $E = \ker S^n \oplus \operatorname{im} S^n$ and that the projection operators associated with this direct sum are continuous. Show also that $\ker S^n$ is finite-dimensional.

g. Let μ be an eigenvalue of T distinct from λ, having index m. Show that

$$\ker(T - \mu I)^m \subseteq \operatorname{im}(T - \lambda I)^n.$$

Hint. By Bezout's Theorem, there exist polynomials P and Q such that

$$P(T)(T - \lambda I)^n + Q(T)(T - \mu I)^m = I.$$

h. Let (λ_k) be the sequence of nonzero eigenvalues of T and (n_k) the sequence of their indexes. For $n \in \mathbb{N}$, denote by F_n and H_n the vector subspaces of E defined by

$$F_n = \sum_{k=0}^{n} \ker(T - \lambda_k)^{n_k}, \quad H_n = \operatorname{im}\biggl(\prod_{k=0}^{n} (T - \lambda_k)^{n_k}\biggr).$$

Show that F_n and H_n are closed, that F_n is finite-dimensional, that $E = F_n \oplus H_n$, and that the projection operators associated with this direct sum are continuous.
Hint. Work by induction on n.

2 Compact Selfadjoint Operators

A classical theorem of linear algebra says that any normal operator (one that commutes with its adjoint) on a complex finite-dimensional Hilbert space is diagonalizable with respect to an orthonormal basis. Here we will see how this result generalizes to infinite dimension. We will restrict our study to compact selfadjoint operators, but the results extend almost without change to compact normal operators on a complex Hilbert space (see Exercise 8 below). In contrast, the compactness assumption is essential. For instance, one can easily check that the operator T on the Hilbert space $L^2([0,1])$ defined by

$$Tf(x) = xf(x) \quad \text{for all } f \in L^2([0,1])$$

is selfadjoint and has no eigenvalues.

In all of this section we consider a scalar product space E over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ and a compact selfadjoint operator T on E. Since we are not assuming that E is complete, the general definition of the adjoint (page 112) does not work; selfadjointness here means that

$$(Tx \mid y) = (x \mid Ty) \quad \text{for all } x, y \in E.$$

Suppose that T has finite rank. Note that, for every $x \in E$,

$$Tx = 0 \quad \iff \quad (Tx \mid y) = 0 \text{ for all } y \in E \quad \iff \quad x \in (\operatorname{im} T)^\perp;$$

thus $\ker T = (\operatorname{im} T)^\perp$ and, since $\operatorname{im} T$ is finite-dimensional, we have $E = \operatorname{im} T \oplus \ker T$ (see Corollary 2.4 on page 107 and the remark after it). The operator T then induces on the finite-dimensional space $\operatorname{im} T$ an invertible selfadjoint operator whose eigenvalues equal the nonzero eigenvalues of T (this much is clear). Using the standard diagonalization results for hermitian and symmetric operators in finite dimension, we deduce that $\operatorname{im} T$ is the orthogonal direct sum of the eigenspaces of T associated with nonzero eigenvalues, and finally that

$$E = \bigoplus_{\lambda \in \operatorname{ev}(T)} \ker(\lambda I - T).$$

We now generalize this diagonalization property to the case where T is any compact selfadjoint operator. *We assume from now on that T does not have finite rank.* The argument is based on the following fundamental lemma:

Lemma 2.1 *Let S be a compact selfadjoint operator on a scalar product space F not equal to $\{0\}$. Then S has at least one eigenvalue and*

$$\max\bigl\{|\lambda| : \lambda \in \operatorname{ev}(S)\bigr\} = \|S\|.$$

Proof. Clearly, if λ is an eigenvalue of S, then $|\lambda| \le \|S\|$. On the other hand, we know from the remark following Theorem 2.6 on page 203 that there exists a spectral value λ of S such that $|\lambda| = \sup_{\|x\|=1} \bigl|(Sx \mid x)\bigr|$, which equals $\|S\|$ by Proposition 3.5 on page 114 (whose proof did not use the completeness of E). We can assume $S \neq 0$ (else the result is trivial), so λ is nonzero and must be an eigenvalue, by Theorem 1.8 on page 219. □

Theorem 2.2 *Let Λ be the set of eigenvalues of T. Write $\Lambda^* = \Lambda \setminus \{0\}$ and, for each eigenvalue λ, let E_λ be the eigenspace of T associated with λ.*

- *Λ is a countable, infinite, bounded subset of $\mathbb{R}$ whose only cluster point is 0.*
- *The eigenspace associated with any nonzero eigenvalue of T has finite dimension.*
- *Eigenspaces of T associated with distinct eigenvalues are orthogonal.*
- *For each nonzero eigenvalue λ of T, let P_λ be the orthogonal projection operator onto E_λ. Then*

$$T = \sum_{\lambda \in \Lambda^*} \lambda P_\lambda,$$

in the sense of a summable family in $L(E)$.

The definition of a summable family in a normed vector space was given on page 127.

We remark also that the orthogonal projection onto a finite-dimensional vector subspace of a scalar product space E is well defined, even when E is not complete; see the remark following Corollary 2.4 on page 107.

Proof

1. That all eigenvalues are real and that eigenspaces associated with distinct eigenvalues are orthogonal comes from parts i and iii of Proposition 2.5 on page 203, whose proof did not use the completeness of E. That eigenspaces associated with nonzero eigenvalues are finite-dimensional comes from Theorem 1.8 on page 219.
2. We prove that Λ^* is infinite. By Lemma 2.1, there exists an eigenvalue λ of T such that $|\lambda| = \|T\|$. Since T is nonzero (recall that T has infinite rank), we deduce that $\lambda \neq 0$ and so that Λ^* is nonempty. Suppose that T has finitely many nonzero eigenvalues: $\Lambda^* = \{\lambda_1, \dots, \lambda_k\}$. Set $G = \bigoplus_{j=1}^{k} E_{\lambda_j}$ and $F = G^\perp$. Since G is finite-dimensional, $E = F \oplus G$ (once more by the remark following Corollary 2.4 on page 107). It is clear that $T(G) \subset G$. Since T is selfadjoint, we quickly deduce that $T(F) \subset F$. The operator T therefore induces an operator T_F from F to itself, and we easily check that T_F is compact, because F is closed. Naturally, T_F is a selfadjoint operator on F, and it is nonzero ($T_F = 0$ would imply $\operatorname{im} T \subset G$, contradicting the fact that T has infinite rank).

By Lemma 2.1, T_F has a nonzero eigenvalue λ. We see then that λ is a nonzero eigenvalue of T distinct from all the λ_j, for $1 \le j \le k$, since one of its associated eigenvectors lies in F and thus not in G. This is a contradiction. It follows that Λ^* is infinite and, by Theorem 1.8 on page 219, Λ is countable and has 0 as its only cluster point.

3. Let J be a finite subset of Λ^* and put $G_J = \bigoplus_{\lambda\in J} E_\lambda$ and $F_J = G_J^\perp$. Arguing as above and using Lemma 2.1, we see that T induces on F_J a compact selfadjoint operator T_{F_J} whose norm equals $\|T_{F_J}\| = \max_{\lambda\in \mathrm{ev}(T_{F_J})} |\lambda|$. Now observe that, as before, every eigenvalue λ of T_{F_J} is an eigenvalue of T (this is clear) but does not belong to J, since, by construction, F_J intersects trivially all the eigenspaces E_μ, for $\mu \in J$. Therefore $\mathrm{ev}(T_{F_J}) \subset \Lambda \setminus J$. Conversely, if $\lambda \in \Lambda \setminus J$, the orthogonality property of eigenspaces implies that $E_\lambda \subset G_J^\perp = F_J$, so λ is an eigenvalue of T_{F_J}. Therefore $\mathrm{ev}(T_{F_J}) = \Lambda \setminus J$ and

$$\|T_{F_J}\| = \max_{\lambda\in\Lambda\setminus J} |\lambda|.$$

Meanwhile, the operator of orthogonal projection onto G_J is $\sum_{\lambda\in J} P_\lambda$. Thus, for every $x \in E$, we have $x - \sum_{\lambda\in J} P_\lambda x \in F_J$ and

$$\Bigl\|T\Bigl(x - \sum_{\lambda\in J} P_\lambda x\Bigr)\Bigr\| = \Bigl\|T_{F_J}\Bigl(x - \sum_{\lambda\in J} P_\lambda x\Bigr)\Bigr\| \le \Bigl\|x - \sum_{\lambda\in J} P_\lambda x\Bigr\| \max_{\lambda\in\Lambda\setminus J} |\lambda|.$$

By orthogonality and the Pythagorean Theorem, we have

$$\Bigl\|x - \sum_{\lambda\in J} P_\lambda x\Bigr\| \le \|x\|,$$

so we conclude that

$$\Bigl\|T - \sum_{\lambda\in J} TP_\lambda\Bigr\| \le \max_{\lambda\in\Lambda\setminus J} |\lambda|.$$

By the definition of P_λ, we have $TP_\lambda = \lambda P_\lambda$, so

$$\Bigl\|T - \sum_{\lambda\in J} \lambda P_\lambda\Bigr\| \le \max_{\lambda\in\Lambda\setminus J} |\lambda|.$$

Now take $\varepsilon > 0$. Since 0 is the only cluster point of Λ, the set K of eigenvalues λ with absolute value at least ε is finite. But then, for every finite subset J of Λ^* containing K,

$$\Bigl\|T - \sum_{\lambda\in J} \lambda P_\lambda\Bigr\| \le \max_{\lambda\in\Lambda\setminus J} |\lambda| \le \max_{\lambda\in\Lambda\setminus K} |\lambda| \le \varepsilon,$$

which proves the third assertion of the theorem. □

Remark. More precisely, the preceding reasoning shows that, for every finite subset J of Λ^*,

$$\Bigl\| T - \sum_{\lambda \in J} \lambda P_\lambda \Bigr\| \le \max_{\lambda \in \Lambda \setminus J} |\lambda|.$$

Corollary 2.3 *In the notation of Theorem 2.2,*

$$\overline{\operatorname{im} T} = \overline{\bigoplus_{\lambda \in \Lambda^*} E_\lambda}.$$

Proof. We know that $Tx = \sum_{\lambda \in \Lambda^*} \lambda P_\lambda x$ for every $x \in E$. It follows that

$$\operatorname{im} T \subset \overline{\bigoplus_{\lambda \in \Lambda^*} E_\lambda}, \quad \text{and hence} \quad \overline{\operatorname{im} T} \subset \overline{\bigoplus_{\lambda \in \Lambda^*} E_\lambda}.$$

On the other hand, if $\lambda \in \Lambda^*$, we clearly have $E_\lambda \subset \operatorname{im} T$, proving the reverse inclusion. □

Theorem 2.2 and Corollary 2.3 can be expressed as follows:

Corollary 2.4

- *The space $\overline{\operatorname{im} T}$ has a countable Hilbert basis $(f_n)_{n \in \mathbb{N}}$ consisting of eigenvectors of T associated with nonzero eigenvalues.*
- *The sequence $(\mu_n)_{n \in \mathbb{N}}$ of eigenvalues associated with the vectors f_n tends to 0 and*

$$Tx = \sum_{n \in \mathbb{N}} \mu_n (x \mid f_n) f_n \quad \textit{for all } x \in E.$$

The Hilbert basis (f_n) is obtained simply by taking the union of all the finite Hilbert bases of the eigenspaces of T associated with nonzero eigenvalues. Note that in the sequence (μ_n) each nonzero eigenvalue λ of T appears d_λ times, where d_λ is the dimension of the eigenspace associated with E_λ.

The first assertion of Corollary 2.4 says in particular that

$$x = \sum_{n \in \mathbb{N}} (x \mid f_n) f_n \quad \text{for all } x \in \overline{\operatorname{im} T},$$

which is to say:

Corollary 2.5 *For every $x \in \overline{\operatorname{im} T}$,*

$$x = \sum_{\lambda \in \Lambda^*} P_\lambda x.$$

Corollary 2.6 *Suppose that E is complete. Let P_0 be the operator of orthogonal projection onto $E_0 = \ker T$. Then*

$$x = \sum_{\lambda \in \Lambda} P_\lambda x \quad \textit{for all } x \in E$$

and

$$E = \overline{\bigoplus_{\lambda \in \Lambda} E_\lambda}.$$

Proof. Since T is selfadjoint, we have $E_0 = \ker T = \overline{\operatorname{im} T}^{\perp}$. Therefore, if E is complete, $E = E_0 \oplus \overline{\operatorname{im} T}$ by Corollary 2.4 on page 107. □

If, moreover, E is separable, so is $\ker T$. Thus $\ker T$ has a countable Hilbert basis, by Corollary 4.7 on page 129. Taking the union of such a basis with the Hilbert basis of $\overline{\operatorname{im} T}$ given by Corollary 2.4, we obtain the following diagonalization result:

Corollary 2.7 *If E is a separable Hilbert space, it has a Hilbert basis consisting of eigenvectors of T.*

This is still true if E is an arbitrary Hilbert space, but then we have to use the axiom of choice in order to guarantee the existence of a Hilbert basis for $\ker T$ and so for E (see Exercise 11 on page 133).

2A Operational Calculus and the Fredholm Equation

We assume here that E is complete and we consider a compact selfadjoint operator T on E. If λ is an eigenvalue of T, we denote as above by $E_\lambda = \ker(\lambda I - T)$ the eigenspace of T associated with λ and by P_λ the orthogonal projection onto E_λ.

Let f be a bounded function on the set $\operatorname{ev}(T)$. We define an operator $f(T)$ on E by

$$f(T)x = \sum_{\lambda \in \operatorname{ev}(T)} f(\lambda) P_\lambda x \quad \text{for all } x \in E.$$

Since the eigenspaces E_λ are pairwise orthogonal, we deduce from the Bessel equality that

$$\|f(T)x\|^2 = \sum_{\lambda \in \operatorname{ev}(T)} |f(\lambda)|^2 \|P_\lambda x\|^2 \quad \text{and} \quad \|x\|^2 = \sum_{\lambda \in \operatorname{ev}(T)} \|P_\lambda x\|^2,$$

the second equality being a consequence of Corollary 2.6. We deduce that

$$\|f(T)\| = \sup_{\lambda \in \operatorname{ev}(T)} |f(\lambda)|.$$

Therefore, for a compact selfadjoint operator, the operational calculus thus defined extends to all bounded functions the calculus defined on page 205 for continuous functions and in Exercise 15 on page 212 for bounded Borel functions. In particular, if $\mu \in \mathbb{K}^*$ is not an eigenvalue of T, we have

$$(\mu I - T)^{-1}x = \sum_{\lambda \in \mathrm{ev}(T)} (\mu - \lambda)^{-1} P_\lambda x \quad \text{for all } x \in E. \tag{$*$}$$

Suppose to the contrary that μ is a nonzero eigenvalue of T (so $\mu \in \mathbb{R}^*$). Then $\mathrm{im}(\mu I - T)$ is closed, by Proposition 1.6 on page 217, and so equal to $E_\mu^\perp$, by Proposition 2.1 on page 201 applied to the hermitian operator $\mu I - T$. The operator T induces on $E_\mu^\perp$ a compact hermitian operator whose set of eigenvalues is $\mathrm{ev}(T) \setminus \{\mu\}$, and we can apply $(*)$ to this induced operator. We deduce, for $x \in E_\mu^\perp$, the following equivalence valid for all $\tilde{y} \in E_\mu^\perp$:

$$\mu\tilde{y} - T\tilde{y} = x \quad \Longleftrightarrow \quad \tilde{y} = \sum_{\substack{\lambda \in \mathrm{ev}(T) \\ \lambda \neq \mu}} (\mu - \lambda)^{-1} P_\lambda x.$$

Next, if $x \in E_\mu^\perp$ and $y \in E$, we can write $y = \tilde{y} + z$, with $\tilde{y} \in E_\mu^\perp$ and $z \in E_\mu$. It follows that $\mu y - Ty = x$ if and only if there exists $z \in E_\mu$ such that

$$y = z + \sum_{\substack{\lambda \in \mathrm{ev}(T) \\ \lambda \neq \mu}} (\mu - \lambda)^{-1} P_\lambda x.$$

To summarize, if we consider the **Fredholm equation**

$$\mu y - Ty = x, \tag{$**$}$$

with $\mu \in \mathbb{K}^*$ and $x \in E$, there are two possible cases:

- μ is not an eigenvalue of T. Then the equation $(**)$ has a unique solution y, given by

$$y = \sum_{\lambda \in \mathrm{ev}(T)} (\mu - \lambda)^{-1} P_\lambda x.$$

- μ is an eigenvalue of T. Then the equation $(**)$ has infinitely many solutions if $x \in (\ker(\mu I - T))^\perp$ and no solutions otherwise. In the first case, the solutions are given by

$$y = z + \sum_{\substack{\lambda \in \mathrm{ev}(T) \\ \lambda \neq \mu}} (\mu - \lambda)^{-1} P_\lambda x,$$

  with $z \in \ker(\mu I - T)$.

This dichotomy is called the **Fredholm Alternative Theorem**. .

2B Kernel Operators

We study here the particular case of the Hilbert space $E = L^2(m)$, where m is a σ-finite measure on a measure space $(X, \mathscr{F})$. Suppose in addition that E is separable.

Consider a kernel $K \in L^2(m \times m)$ such that $K(x,y) = \overline{K(y,x)}$ for $(m \times m)$-almost every (x,y). The operator $T = T_K$ associated to this kernel by the equation

$$T_K f(x) = \int K(x,y) f(y)\, dm(y)$$

is a compact selfadjoint operator (see Examples 1 and 2 on page 216). If λ is a nonzero eigenvalue of T, let d_λ be the dimension of the eigenspace associated with $E_\lambda = \ker(\lambda I - T)$. *We assume in the sequel that T does not have finite rank* and, as in Corollary 2.4, we denote by $(f_n)_{n\in\mathbb{N}}$ a Hilbert basis of $\overline{\operatorname{im} T}$ consisting of eigenvectors of T and by $(\mu_n)_{n\in\mathbb{N}}$ the sequence of corresponding (nonzero) eigenvalues.

Proposition 2.8 *With the notation and hypotheses above,*

$$\iint |K(x,y)|^2\, dm(x)\, dm(y) = \sum_{n=0}^{+\infty} \mu_n^2 = \sum_{\substack{\lambda\in \operatorname{ev}(T)\\ \lambda\neq 0}} d_\lambda \lambda^2.$$

Proof. Take $u \in \ker T$. For almost every y, the function $K_y : x \mapsto K(x,y)$ lies in E and

$$(K_y \mid u) = \int K(x,y)\bar{u}(x)\, dm(x) = \overline{Tu}(y) = 0.$$

The second of these equalities is true for almost every y: more precisely, for every y not in a subset A_u of X of measure zero, and which a priori may depend on u. But, since E is separable, $\ker T$ is also separable. Let $(u_n)_{n\in\mathbb{N}}$ be a dense subset of $\ker T$. Then, for every y not belonging to the set $A = \bigcup_{n\in\mathbb{N}} A_{u_n}$ of measure zero, we have $(K_y \mid u_n) = 0$ for every $n \in \mathbb{N}$ and, because of denseness, $(K_y \mid u) = 0$ for every $u \in \ker T$. It follows that $K_y \in (\ker T)^\perp = \overline{\operatorname{im} T}$ for almost every y. At the same time, for each $n \in \mathbb{N}$,

$$(K_y \mid f_n) = \overline{Tf_n}(y) = \overline{\mu_n f_n}(y) = \mu_n \overline{f_n}(y) \qquad (*)$$

for almost every y. We then deduce from the Bessel equality that, for almost every y,

$$\int |K(x,y)|^2\, dm(x) = \|K_y\|^2 = \sum_{n=0}^{+\infty} \mu_n^2\, |f_n(y)|^2. \qquad (**)$$

Now just integrate with respect to y to obtain the desired result. ☐

Remark. The preceding proposition is also a direct consequence of Exercise 21 on page 140.

We know from Exercise 7 on page 110 that the space $L^2(m) \otimes L^2(m)$ is dense in $L^2(m \times m)$. From the preceding proof, we obtain an explicit approximation of the kernel K by elements of $L^2(m) \otimes L^2(m)$.

Proposition 2.9 *We have*

$$K(x, y) = \sum_{n=0}^{+\infty} \mu_n f_n(x) \overline{f_n}(y),$$

the series being convergent in $L^2(m \times m)$.

Proof. Set $K_N(x, y) = \sum_{n=0}^{N} \mu_n f_n(x) \overline{f_n}(y)$. By equality $(*)$ above, for almost every y, we have

$$K_y = \sum_{n=0}^{+\infty} \mu_n \overline{f_n}(y) f_n$$

in the sense of convergence in $L^2(m)$. Thus, still by Bessel's equality,

$$\int \left|K(x, y) - K_N(x, y)\right|^2 dm(x) = \left\|(K_N)_y - K_y\right\|^2 = \sum_{n=N+1}^{+\infty} \mu_n^2 \left|f_n(y)\right|^2.$$

We deduce, integrating this equality with respect to y, that

$$\|K - K_N\|^2 = \sum_{n=N+1}^{+\infty} \mu_n^2,$$

where $\|\cdot\|$ represents the norm in $L^2(m \times m)$. This proves the result. $\square$

Proposition 2.10 *Suppose that $\Phi : x \mapsto \int |K(x, y)|^2 \, dm(y)$ belongs to $L^\infty(m)$. Then, for every $n \in \mathbb{N}$, we have $f_n \in L^\infty(m)$ and*

$$f = \sum_{n=0}^{+\infty} (f \mid f_n) f_n \quad \textit{for every } f \in \operatorname{im} T,$$

the convergence of the series taking place in $L^\infty(m)$.

In particular, (f_n) is a fundamental family in the space $\operatorname{im} T$ considered with the norm of $L^\infty(m)$. Recall that the convergence in $L^2(m)$ of the series $\sum_{n=0}^{+\infty} (f \mid f_n) f_n$ and the fact that the sum equals f are consequences of Corollary 2.4.

Proof. For every $n \in \mathbb{N}$ we have

$$f_n(x) = \frac{1}{\mu_n} \int K(x, y) f_n(y) \, dm(y).$$

Therefore $f_n \in L^\infty(m)$ and $\|f_n\|_\infty \le \mu_n^{-1}\sqrt{L}$, where $L = \|\Phi\|_\infty$. We show that the series $\sum_{n=0}^{+\infty}(f \mid f_n) f_n$ satisfies Cauchy's criterion on $L^\infty(m)$. Let $f = Tg$ be an element of $\operatorname{im} T$. For every $n \in \mathbb{N}$, $(f \mid f_n) = (Tg \mid f_n) = (g \mid Tf_n) = \mu_n (g \mid f_n)$. If $k \le l$,

$$\left|\sum_{n=k}^{l}(f \mid f_n) f_n(x)\right| = \left|\sum_{n=k}^{l}\mu_n (g \mid f_n) f_n(x)\right|$$
$$\le \left(\sum_{n=k}^{l}|(g|f_n)|^2\right)^{1/2}\left(\sum_{n=0}^{+\infty}\mu_n^2|f_n(x)|^2\right)^{1/2},$$

by the Schwarz inequality. Now, by an earlier calculation (see equality (**) on page 240), we have

$$\sum_{n=0}^{+\infty}\mu_n^2\left|f_n(x)\right|^2 = \int\left|K(x,y)\right|^2 dm(y) \le L,$$

which finally implies that

$$\left\|\sum_{n=k}^{l}(f \mid f_n) f_n\right\|_\infty \le \left(\sum_{n=k}^{l}\left|(g \mid f_n)\right|^2\right)^{1/2} L^{1/2},$$

which proves the result, since the series $\sum_{n=0}^{+\infty}\left|(g \mid f_n)\right|^2$ converges by the Bessel inequality and so satisfies Cauchy's criterion. □

Example. An important special case in which the hypothesis of Proposition 2.10 is satisfied is when m is a Radon measure on a compact space X and K is continuous on $X \times X$. In this case, for every $f \in E$, the image Tf is a continuous function: indeed, if $x, x' \in X$,

$$\left|Tf(x) - Tf(x')\right| \le \sup_{y \in X}\left|K(x,y) - K(x',y)\right| m(X)^{1/2}\|f\|.$$

Thus it is enough to use the uniform continuity of K on the compact set $X \times X$. Therefore each f_n is a continuous function and we deduce from Proposition 2.10 that, if $\operatorname{Supp} m = X$, we have

$$g = \sum_{n=0}^{+\infty}(g \mid f_n) f_n \quad \text{for every } g = Tf \in \operatorname{im} T,$$

the series converging uniformly on X; that is, for every $f \in E$,

$$Tf = \sum_{n=0}^{+\infty}\mu_n (f \mid f_n) f_n$$

(since $(g \mid f_n) = (Tf \mid f_n) = (f \mid Tf_n) = \mu_n (f \mid f_n)$), the series converging in the space $C(X)$ considered with the uniform norm.

Exercises

1. Let E be a Hilbert space, $(f_n)_{n\in\mathbb{N}}$ an orthonormal family in E, and $(\mu_n)_{n\in\mathbb{N}}$ a real-valued sequence that tends to 0. Show that the equation

$$Tx = \sum_{n=0}^{+\infty} \mu_n (x \mid f_n) f_n \quad \text{for all } x \in E$$

defines on E a compact selfadjoint operator T. (Thus, the property stated in Corollary 2.4 characterizes compact selfadjoint operators whose rank is not finite.)

2. Let T be a compact selfadjoint operator on an infinite-dimensional Hilbert space E. Let f be a continuous function on the set $\sigma(T)$. Show that $f(T)$ is compact if and only if $f(0) = 0$. (In particular, if T is positive, $T^{1/2}$ is compact.)
Hint. For sufficiency use Exercise 1, for example.

3. Let E be a scalar product space and T a compact selfadjoint operator on E. Let λ be a nonzero eigenvalue of T. Show that $\ker(\lambda I - T) = (\operatorname{im}(\lambda I - T))^{\perp}$. Deduce that

$$\ker(\lambda I - T) \cap \operatorname{im}(\lambda I - T) = \{0\},$$

then that $\ker(\lambda I - T)^2 = \ker(\lambda I - T)$. This shows, in particular, that the eigenvalue λ has index 1 (see Exercise 25 on page 233), and therefore that

$$E = \ker(\lambda I - T) \oplus \operatorname{im}(\lambda I - T).$$

4. Let T be the operator defined on $L^2([0,1])$ by

$$Tf(x) = \int_0^x f(y)\, dy.$$

a. Show that the adjoint of T is given by

$$T^* f(x) = \int_x^1 f(y)\, dy \quad \text{for all } f \in L^2([0,1]).$$

b. Show that TT^* is the operator T introduced in Exercise 11 on page 223. Deduce the spectral radius of TT^*, then the norm of T.

5. Deduce from Exercise 11 on page 223 that

$$\sum_{n=0}^{+\infty} \frac{1}{(2n+1)^4} = \frac{\pi^4}{96},$$

then that

$$\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

6. Take $E = L^2([0,1])$ and define a function K on $[0,1]^2$ by

$$K(x,y) = \begin{cases} 1 & \text{if } x+y \leq 1, \\ 0 & \text{if } x+y > 1. \end{cases}$$

a. Write down explicitly the operator T on E defined by the kernel K. Check that T is the extension to $L^2([0,1])$ of the operator on $C([0,1])$ defined in the example on page 220.

b. Show that $E = \overline{\operatorname{im} T}$. Use this to find a Hilbert basis consisting of eigenvectors of T.

c. Deduce that, if $g \in C^1([0,1])$ and $g(1) = 0$, then

$$\begin{aligned} g(x) &= 2\sum_{n\in\mathbb{Z}} \left(\int_0^1 g(t)\cos((\pi/2 + 2n\pi)t)\,dt\right)\cos((\pi/2+2n\pi)x) \\ &= 2\sum_{n=0}^{+\infty} \left(\int_0^1 g(t)\cos((2n+1)\pi t/2)\,dt\right)\cos((2n+1)\pi x/2), \end{aligned}$$

the series converging uniformly on $[0,1]$.

(This result can be obtained using Section 2B above, or using the theory of Fourier series by extending g to an even periodic function of period 4 such that $g(2-x) = -g(x)$ for every $x \in [0,1]$.)

7. Let T be a compact selfadjoint operator on a Hilbert space E. For every nonzero eigenvalue λ of T, denote by P_λ the orthogonal projection onto the eigenspace $E_\lambda = \ker(\lambda I - T)$. Let x be an element of E. Show that the equation

$$Ty = x \tag{$*$}$$

has a solution if and only if $x \in (\ker T)^\perp$ and

$$\sum_{\substack{\lambda\in \mathrm{ev}(T)\\ \lambda\neq 0}} \frac{\|P_\lambda x\|^2}{\lambda^2} < +\infty,$$

and that in this case all the solutions are given by

$$y = z + \sum_{\substack{\lambda\in \mathrm{ev}(T)\\ \lambda\neq 0}} \frac{P_\lambda x}{\lambda} \quad \text{with } z \in \ker T.$$

8. *Diagonalization of normal compact operators.* Let E be a Hilbert space over $\mathbb{C}$. A continuous operator T on E is called *normal* if $TT^* = T^*T$. You might recall, for subsequent use, the result of Exercise 4b on page 208.

a. Let T be a normal compact operator on E. Show that T has at least one eigenvalue $\lambda \in \mathbb{C}$ of absolute value $\|T\|$.

Hint. Put $\mu = \|T^*T\| = \|T\|^2$. Show that μ is an eigenvalue of T^*T, that the associated eigenspace $F = \ker(\mu I - T^*T)$ is finite-dimensional and is invariant under T and T^*, and that T induces on F a normal operator T_F. Then show that T_F has at least one eigenvalue λ and that $|\lambda|^2 = \mu$. (Note that here the fact that the base field is $\mathbb{C}$ is essential.)

b. Show that all the results of the preceding section, from page 234 to the Fredholm Alternative Theorem, remain true without change for a normal operator T (on a complex Hilbert space), with the only exception that the eigenvalues of T need not be real in this case (see Exercise 4a-ii on page 208).

c. Let T be a compact operator on E. Show that T is normal if and only if

$$E = \overline{\bigoplus_{\lambda \in \mathrm{ev}(T)} \ker(\lambda I - T)},$$

the direct sum being orthogonal. (See also Exercise 1.)

d. *An example.* Let G be an element of $L^2([0,1])$, and extend it to a periodic function of period 1 on $\mathbb{R}$. Consider the operator T on $L^2([0,1])$ defined by

$$Tf(x) = \int_0^1 G(x-y)f(y)\,dy \quad \text{for all } f \in L^2([0,1]).$$

i. Show that T is a normal compact operator.

ii. Show that the eigenvalues of T are the Fourier coefficients of G, namely, the numbers $c_n(G)$ defined for $n \in \mathbb{Z}$ by

$$c_n(G) = \int_0^1 G(x)e^{-2in\pi x}\,dx.$$

Show that the corresponding eigenvectors are the vectors of the Hilbert basis $(e_n)_{n\in\mathbb{Z}}$ defined by $e_n(x) = e^{2in\pi x}$.

iii. Show that, for every $f \in L^2([0,1])$,

$$Tf(x) = \sum_{n\in\mathbb{N}} c_n(G)c_n(f)e^{2i\pi nx},$$

the series converging uniformly (and absolutely) on $[0,1]$.

9. Let T be a compact selfadjoint operator on a separable Hilbert space E. For each nonzero eigenvalue λ of T, let d_λ be the dimension of the associated eigenspace $E_\lambda = \ker(\lambda I - T)$. Show that T is a Hilbert–Schmidt operator if and only if

$$\sum_{\substack{\lambda\in\mathrm{ev}(T)\\ \lambda\neq 0}} d_\lambda \lambda^2 < +\infty.$$

(See also Exercise 10.)

10. *Singular values of a compact operator.* Let E be a Hilbert space and suppose $T \in L(E)$. For each $n \in \mathbb{N}$, define a nonnegative real number $\sigma_n(T)$ by

$$\sigma_n(T) = \inf \big\{ \|T - R\| : R \in L(E),\ \operatorname{rank}(R) \le n \big\}.$$

We know from Exercise 24 on page 232 that T is compact if and only if the sequence $(\sigma_n(T))_{n\in\mathbb{N}}$ tends to 0. In what follows, we suppose that T is compact.

a. Show that the operator $P = (T^*T)^{1/2}$ is selfadjoint and compact (see Exercise 2).

We denote by $(\mu_n)_{n\in\mathbb{N}}$ the sequence of nonzero eigenvalues of P, in decreasing order and counted with multiplicity (that is, each nonzero eigenvalue λ appears d_λ times in the sequence (μ_n), where d_λ is the dimension of $\ker(\lambda I - P)$). The entries of this sequence (μ_n) are called the *singular values* of the operator T.

We denote by $(f_n)_{n\in\mathbb{N}}$ a Hilbert basis of $\overline{\operatorname{im} P}$ such that, for every $x \in E$,

$$Px = \sum_{n=0}^{+\infty} \mu_n (x \mid f_n) f_n$$

(see Corollary 2.4). If T has finite rank N, we use the convention that $\mu_n = 0$ and $f_n = 0$, for $n \ge N$, the Hilbert basis of $\overline{\operatorname{im} P}$ being the finite family $(f_0, \dots, f_{N-1})$.

b. Check that, if T is selfadjoint, its singular values equal the absolute values of the eigenvalues of T.

c. *Schmidt decomposition of the operator T.* Show that there exists an orthonormal family $(g_n)_{n\in\mathbb{N}}$ in E such that

$$Tx = \sum_{n=0}^{+\infty} \mu_n (x \mid f_n) g_n \quad \text{for all } x \in E.$$

Hint. Put $g_n = Uf_n$, where U is the operator such that $T = UP$ defined in Exercise 11 on page 211.

d. **i.** Let $R \in L(E)$ be an operator of rank at most n. Show that $\|T - R\| \ge \mu_n$.

Hint. If F_n is the vector space spanned by the family $(f_j)_{0\le j\le n}$, check that $F_n \cap \ker R$ contains a nonzero element x; then show that $\|(T - R)x\| \ge \mu_n \|x\|$.

ii. *Allakhverdief's Lemma.* Deduce that

$$\mu_n = \sigma_n(T) \quad \text{for all } n \in \mathbb{N}$$

and that, in the definition of $\sigma_n(T)$, we can replace inf by min.

e. Suppose from now on that E is separable and fix a Hilbert basis $(e_n)_{n\in\mathbb{N}}$ of E. Denote by $\|T\|_H$ the (possibly infinite) Hilbert–Schmidt norm of T, defined in Exercise 21 on page 140:

$$\|T\|_H^2 = \sum_{n\in\mathbb{N}} \|Te_n\|^2.$$

Show that

$$\|T\|_H^2 = \sum_{n\in\mathbb{N}} \sigma_n(T)^2.$$

In particular, T is a Hilbert–Schmidt operator if and only if

$$\sum_{n\in\mathbb{N}} \sigma_n(T)^2 < +\infty.$$

f. An operator T is called *nuclear* if

$$\sum_{n\in\mathbb{N}} \sigma_n(T) < +\infty.$$

i. Show that, if T is the product of two Hilbert–Schmidt operators, then T is nuclear.
Hint. If $T = AB$, where A and B are Hilbert–Schmidt operators, prove that $\mu_n = (Bf_n \mid A^*g_n)$ for every $n \in \mathbb{N}$.

ii. Conversely, prove that, if T is nuclear, it is the product of two Hilbert–Schmidt operators.
Hint. Take the polar decomposition $T = UP$ of T defined in Exercise 11 on page 211, and show that T being nuclear implies that $UP^{1/2}$ and $P^{1/2}$ are Hilbert–Schmidt operators.

11. *Calculation of the eigenvalues: the Courant–Fischer formulas.* Let E be a Hilbert space distinct from $\{0\}$ and let T be a compact positive selfadjoint operator on E. Order the nonzero eigenvalues of T as $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_n \geq \cdots$, where the number of times each eigenvalue appears is the dimension of the associated eigenspace. For every $p \in \mathbb{N}$, denote by $\mathscr{V}_p$ the set of p-dimensional subspaces of E. The goal of this exercise is to prove the formulas

$$\left.\begin{aligned} \mu_n &= \min_{W\in\mathscr{V}_n} \max_{x\in W^\perp\setminus\{0\}} \frac{(Tx \mid x)}{\|x\|^2}, \\ \mu_n &= \max_{W\in\mathscr{V}_{n+1}} \min_{x\in W\setminus\{0\}} \frac{(Tx \mid x)}{\|x\|^2}. \end{aligned}\right\} \qquad (*)_n$$

(In this context, recall Proposition 3.5 on page 114.) Let $(f_n)_{n\in\mathbb{N}}$ be a Hilbert basis of $\overline{\operatorname{im} T}$ such that $Tf_n = \mu_n f_n$ for every $n \in \mathbb{N}$.

a. Show that, if F is a closed subspace of E distinct from $\{0\}$, there exists an element x of F of norm 1 such that

$$(Tx \mid x) = \sup_{y \in F,\ \|y\|=1} (Ty \mid y).$$

(In particular, we can replace max by sup in the first formula $(*)_n$.)

b. If $n \in \mathbb{N}$, let W_n be the vector space spanned by $f_0, \dots, f_{n-1}$ (with $W_0 = \{0\}$). Show that

$$\max_{x \in W_n^\perp \setminus \{0\}} \frac{(Tx \mid x)}{\|x\|^2} = \mu_n.$$

Hint. Consider the restriction of T to $W_n^\perp$.

c. Take $W \in \mathscr{V}_n$. Show that $W^\perp \cap W_{n+1}$ is distinct from $\{0\}$ and that, for every nonzero element x of W_{n+1},

$$\frac{(Tx \mid x)}{\|x\|^2} \geq \mu_n.$$

Deduce from these results the first equality $(*)_n$. (You should check in particular that the minimum is attained by the space W_n.)

d. Take $W \in \mathscr{V}_{n+1}$. Show that $W \cap W_n^\perp$ is distinct from $\{0\}$ and deduce that there exists a nonzero element x of W such that

$$\frac{(Tx \mid x)}{\|x\|^2} \leq \mu_n.$$

Then show the second equality $(*)_n$. (You should check in particular that the maximum is attained by the space W_{n+1}.)

e. *Application.* Let S and T be compact positive selfadjoint operators on E such that $S \leq T$ (that is, $T - S$ is positive selfadjoint). Show that $\mu_n(S) \leq \mu_n(T)$ for every $n \in \mathbb{N}$.

12. *Sturm–Liouville problem, continued.* Let p be a function of class C^1 on $[0,1]$ taking positive values. Let q be a continuous real-valued function on $[0,1]$, and suppose $\varepsilon_0, \varepsilon_1 \in \{0,1\}$. For $\lambda \in \mathbb{R}$, consider the differential equation on $[0,1]$ given by

$$(py')' - (q + \lambda)y = 0, \tag{E$_\lambda$}$$

with boundary conditions

$$\varepsilon_0 y(0) + (1 - \varepsilon_0) y'(0) = 0, \quad \varepsilon_1 y(1) + (1 - \varepsilon_1) y'(1) = 0. \tag{BC}$$

a. Suppose (in this item only) that q is positive-valued. Let $T_{p,q}$ be the operator on $C([0,1])$ defined in Exercise 13 on page 224 and characterized by the fact that, for every $g \in C([0,1])$, $T_{p,q}g$ is the unique solution on $[0,1]$ of the equation

$$(py')' - qy = g$$

satisfying (BC). Show that $T_{p,q}$ is negative selfadjoint (that is, $-T_{p,q}$ is positive selfadjoint) and compact on the scalar product space $C([0,1])$ considered with the norm induced by $L^2([0,1])$.

b. Take $a > \max_{x\in[0,1]}(-q(x))$. Show that the set Λ of real numbers λ for which (E_λ) + (BC) has a non identically zero solution forms a sequence $(\lambda_n)_{n\in\mathbb{N}}$ such that

$$a > \lambda_0 > \lambda_1 > \cdots > \lambda_n > \cdots$$

and $\lim_{n\to+\infty} \lambda_n = -\infty$ (more precisely, the series $\sum_{n=0}^{+\infty}(a-\lambda_n)^{-2}$ converges). The constants λ_n, for $n \in \mathbb{N}$, are called *critical values* of the problem (E_λ) + (BC).

Hint. We have $\lambda \in \Lambda$ if and only if $1/(\lambda - a)$ is an eigenvalue of $T_{p,q+a}$.

c. Show that, for every $n \in \mathbb{N}$, there exists a solution φ_n of (E_{λ_n})+(BC) such that $\int_0^1 |\varphi_n|^2(t)\,dt = 1$ and that φ_n is, up to a multiplicative factor, the unique solution of (E_{λ_n}) + (BC). Show that the family $(\varphi_n)_{n\in\mathbb{N}}$ is a Hilbert basis of $L^2([0,1])$ and that, if $f \in C^2([0,1])$ satisfies (BC), the series $\sum_{n=0}^{+\infty}(f \mid \varphi_n)\varphi_n$, where $(\cdot \mid \cdot)$ is the scalar product in $L^2([0,1])$, converges uniformly to f.

d. Suppose that $p = 1$ and $q = 0$. Determine the sequences (λ_n) and (φ_n) in the following cases:

i. $\varepsilon_0 = 0,\ \varepsilon_1 = 0$;

ii. $\varepsilon_0 = 0,\ \varepsilon_1 = 1$;

iii. $\varepsilon_0 = 1,\ \varepsilon_1 = 1$.

e. Suppose $\varepsilon_0 = \varepsilon_1 = 1$. Show that the function φ_0 does not take the value 0 in the interval $(0,1)$, and that no other function φ_n has this property.

Hint. Show first that $\varphi_0 \geq 0$ or $\varphi_0 \leq 0$, using Exercises 13d on page 225 and 14 on page 226. Deduce that, if $\varphi_0(\xi) = 0$ with $\xi \in (0,1)$, we must have $\varphi_0'(\xi) = 0$ and therefore $\varphi_0 = 0$, since φ_0 is a solution of (E_{λ_0}). But this is impossible.

13. *Legendre's equation.* Let E be the space $C([-1,1])$ with the scalar product induced by $L^2([-1,1])$.

We define on E a kernel operator T by

$$Tf(x) = \int_{-1}^{1} K(x,y)f(y)\,dy,$$

where

$$K(x,y) = \begin{cases} \frac{1}{2} - \log 2 + \frac{1}{2}\log\bigl((1-y)(1+x)\bigr) & \text{if } -1 < y \le x, \\ \frac{1}{2} - \log 2 + \frac{1}{2}\log\bigl((1+y)(1-x)\bigr) & \text{if } \ \ 1 > y \ge x. \end{cases}$$

a. Show that T is a compact hermitian operator from E to itself.

b. Consider on $[-1,1]$ the differential equation

$$\bigl((1-x^2)y'\bigr)' = g, \tag{E_g}$$

with $g \in E$. By definition, a solution of (E_g) is a function of class C^1 on the interval $[-1,1]$ satisfying the equation (E_g) on $[-1,1]$. Show that (E_g) has a solution in E if and only if $\int_{-1}^{1} g(x)\,dx = 0$ and that, in this case, all solutions of (E_g) are given by

$$y = Tg + C \quad \text{with } C \in \mathbb{K},$$

the function $f = Tg$ being the unique solution of (E_g) such that

$$\int_{-1}^{1} f(x)\,dx = 0.$$

c. Show that $\ker T$ equals the set of constant functions on $[-1,1]$ and that $\overline{\operatorname{im} T}$ is the set of elements of E whose integral over $[-1,1]$ is zero.

d. Show that the operator $-T$ is positive hermitian.
Hint. Check that, if $g \in E$,

$$(Tg \mid g) = -\int_{-1}^{1} \bigl|(Tg)'(x)\bigr|^2 (1-x^2)\,dx.$$

e. Let $(P_n)_{n\in\mathbb{N}}$ be the sequence of Legendre polynomials defined in Exercise 4 on page 131. Show that, for every $n \in \mathbb{N}$,

$$\bigl((1-x^2)P_n'\bigr)' = -n(n+1)P_n.$$

Use this to find the eigenvalues and eigenvectors of T. Derive another proof that $-T$ is positive hermitian.

14. *About the zeros of a Bessel function.* For $k \in \mathbb{N}$, the *Bessel function* J_k is defined by

$$J_k(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n (x/2)^{k+2n}}{n!\,(n+k)!}.$$

a. Consider on $(0,1]$ the differential equation

$$x^2 y'' - \tfrac{3}{4} y = 0. \tag{$*$}$$

i. Find all solutions of the form $y = x^\lambda$.
ii. Use this to find all solutions of $(*)$.
iii. Prove that the only solution y of $(*)$ satisfying

$$\lim_{x\to 0} y(x) = 0 \quad \text{and} \quad y'(1) + \tfrac{1}{2}y(1) = 0 \tag{**}$$

is the zero solution.

b. For $x, t \in (0,1]$, define

$$K(x,t) = \sqrt{xt}\exp\bigl(-|\log(x/t)|\bigr),$$

and set $K(x,t) = 0$ if $x = 0$ or $t = 0$. Define an operator T from $L^2([0,1])$ to itself by

$$Tf(x) = \int_0^1 K(x,t)f(t)\,dt \quad \text{for all } x \in [0,1].$$

Show that T is a compact hermitian operator.

c. Take $f \in C([0,1])$ and set $F = Tf$.

i. Show that, for every $x \in (0,1]$,

$$F(x) = x^{-1/2}\int_0^x t^{3/2}f(t)\,dt + x^{3/2}\int_x^1 t^{-1/2}f(t)\,dt$$

and that $F(0) = 0$. Deduce that $F \in C^1([0,1])$, $F'(0) = 0$, and $F'(1) + F(1)/2 = 0$.

ii. Show that $F \in C^2(]0,1])$ and that F satisfies on $(0,1]$ the equation

$$F'' - \tfrac{3}{4}x^{-2}F = -2f. \tag{†}$$

iii. Show that F is the unique function of class C^2 on $(0,1]$ satisfying $(**)$ and (†).

d. Deduce from all this that the image of T contains the space of functions of class C^2 on $(0,1)$ with compact support. Then show that $\operatorname{im} T$ is dense in $L^2([0,1])$, then that T is injective.

e. Show that, if $f \in C([0,1])$,

$$\int_0^1 Tf(t)\overline{f(t)}\,dt = \tfrac{1}{4}\bigl|Tf(1)\bigr|^2 + \tfrac{1}{2}\int_0^1 \bigl|(Tf)'(t)\bigr|^2\,dt + \tfrac{3}{8}\int_0^1 t^{-2}\bigl|Tf(t)\bigr|^2\,dt.$$

Deduce that T is a positive hermitian operator.

f. Show that a real $\lambda > 0$ is an eigenvalue of T if and only if the equation

$$y'' + \left(\frac{2}{\lambda} - \frac{3}{4x^2}\right)y = 0 \tag{E_λ}$$

has a solution in $(0,1]$ that does not vanish identically and that satisfies $(**)$.

g. Take $\lambda > 0$. Study the solutions y of (E_λ) of the form $y(x) = x^\alpha f(x)$, where f has a power series expansion at 0. (*Partial answer:* $\alpha = \frac{3}{2}$ and $\alpha = -\frac{1}{2}$.) Deduce that (E_λ) has a unique (up to a multiplicative factor) solution H_λ such that $\lim_{x\to 0} H_\lambda(x) = 0$, and that this solution is given by

$$H_\lambda(x) = x^{1/2} J_1\bigl(x\sqrt{2/\lambda}\bigr).$$

h. Show that, for every x,

$$xJ_0(x) = xJ_1'(x) + J_1(x).$$

Deduce that the eigenvalues of T are the numbers $\lambda > 0$ for which $J_0\bigl(\sqrt{2/\lambda}\bigr) = 0$.

i. Show that J_0 has a sequence of positive roots

$$0 < \mu_0 < \mu_1 < \cdots < \mu_n < \cdots$$

and that

$$\sum_{n=0}^{+\infty} \frac{1}{\mu_n^4} = \frac{1}{32}.$$

j. For $n \in \mathbb{N}$ and $x \in [0,1]$, put $\varphi_n(x) = x^{1/2} J_1(\mu_n x)$. Show that $(\varphi_n)_{n\in\mathbb{N}}$ is a fundamental orthogonal family in $L^2([0,1])$ and that, if $f \in C_c^2((0,1))$, there exist coefficients $c_n(f)$ such that the series $\sum_{n=0}^{+\infty} c_n(f)\varphi_n$ converges uniformly on $[0,1]$, with sum f.

Remark. An analogous study can be made of the zeros of the function J_k, by considering the kernel

$$K_k(x,t) = \sqrt{xt}\,\exp\bigl(-(k+1)\,|\log(x/t)|\bigr).$$

15. *Approximate calculation of an eigenvalue of a compact positive selfadjoint operator.* Let T be a compact selfadjoint operator in a Hilbert space E satisfying the condition that $(Tx \mid x) > 0$ for every $x \neq 0$. Let x_0 be a nonzero element of E. For each $n \in \mathbb{N}$, we set

$$x_n = T^n x_0, \quad \beta_n = \frac{\|x_n\|}{\|x_{n+1}\|}, \quad \alpha_n = \frac{(x_{n+1}|x_n)}{\|x_{n+1}\|^2}.$$

a. We wish to show that the sequences α_n and β_n converge to the inverse of an eigenvalue of T (the same for both sequences).

i. Show that $0 < \alpha_n \le \beta_n$ and that the sequence (β_n) is decreasing.

ii. Let $(f_k)_{k\in\mathbb{N}}$ be a Hilbert basis of E consisting of eigenvectors of T, and denote by μ_k the eigenvalue associated with f_k; we assume that the μ_k are arranged in nonincreasing order. Let k_0 be the smallest integer k such that $(x_0 \mid f_k) \neq 0$. Show that

$$\lim_{n\to+\infty} \|x_n\|^{1/n} = \mu_{k_0}.$$

Deduce that the sequence $(\beta_n)_{n\in\mathbb{N}}$ converges to $1/\mu_{k_0}$.
Hint. Note that

$$\|x_n\|^2 = \sum_{j=k_0}^{+\infty} \left|(x_0 \,|\, T^n f_j)\right|^2.$$

iii. Show that $\alpha_n \geq 1/\mu_{k_0}$ for every $n \in \mathbb{N}$. Deduce that the sequence $(\alpha_n)_{n\in\mathbb{N}}$ converges to $1/\mu_{k_0}$.

b. *Stopping criterion.* Show that, for every integer $n \in \mathbb{N}$, there exists an eigenvalue λ of T such that

$$\sqrt{\beta_n^2 - \alpha_n^2} \geq \left|\alpha_n - \frac{1}{\lambda}\right|.$$

Hint. Observe that $\|x_n - \alpha_n x_{n+1}\|^2 = \|x_{n+1}\|^2(\beta_n^2 - \alpha_n^2)$ and use Bessel's equality.

Part III

DISTRIBUTIONS

7

Definitions and Examples

Distributions, as we shall see, are objects that generalize locally integrable functions and Radon measures on $\mathbb{R}^d$. One of the main attractions of the theory of distributions, apart from its unifying power, is the construction of an extension of the usual differential calculus in such a way that every distribution is differentiable infinitely often. This theory has become an essential tool, particularly in the study of partial differential equations. It has also allowed the precise mathematical modeling of numerous physical phenomena.

The fundamental idea of the theory is to define distributions by means of their action on a space of functions, called **test functions**. Note that this idea already appears in the definition of measures by Daniell's method (Chapter 2), and in particular in the definition of Radon measures.

In the first section of this chapter, we introduce the various test function spaces. We will be working in an open subset Ω of $\mathbb{R}^d$. We will often omit the symbol Ω from the notation when $\Omega = \mathbb{R}^d$.

1 Test Functions

1A Notation

- If $m \in \mathbb{N}$, $\mathscr{E}^m(\Omega)$ denotes the space of complex-valued functions on Ω of class C^m, and $\mathscr{E}(\Omega)$ the space of those of class C^∞. By convention, $\mathscr{E}^0(\Omega) = C(\Omega)$.

- An element $p \in \mathbb{N}^d$ is called a **multiindex**. If $p = (p_1, \ldots, p_d)$ is a multiindex, we define the **length** of p to be the sum $|p| = p_1 + \cdots + p_d$, and we put $p! = p_1! \ldots p_d!$. We give $\mathbb{N}^d$ the product order: if p and q are two multiindices, we write $p \le q$ if $p_1 \le q_1, \ldots, p_d \le q_d$. If $p, q \in \mathbb{N}^d$ and $q \le p$, we put

$$\binom{p}{q} = \prod_{j=1}^{d} \binom{p_j}{q_j} = \frac{p!}{q!\,(p-q)!},$$

where, as usual, $\binom{p_j}{q_j}$ represents the binomial coefficient $\dfrac{p_j!}{q_j!\,(p_j - q_j)!}$.
- If $1 \le j \le d$, we often use D_j to denote $\dfrac{\partial}{\partial x_j}$. Then, if p is a multiindex, we write

$$D^p = D_1^{p_1} \ldots D_d^{p_d} = \frac{\partial^{|p|}}{\partial x_1^{p_1} \ldots \partial x_d^{p_d}}.$$

The differentiation operator D^p is also denoted by

$$\frac{\partial^{|p|}}{\partial x^p} \quad \text{or} \quad \partial_x^p.$$

By convention, D_j^0 (differentiation of order 0 with respect to any index) is the identity map.
We see that each operator D^p, where $p \in \mathbb{N}^d$, acts on the spaces $\mathscr{E}^m(\Omega)$, for $|p| \le m$. We recall the following classical result:

Proposition 1.1 (Leibniz's formula) *Suppose $f, g \in \mathscr{E}^m(\Omega)$. For each multiindex p such that $|p| \le m$,*

$$D^p(fg) = \sum_{q \le p} \binom{p}{q} D^{p-q} f D^q g.$$

- If K is a compact subset of $\mathbb{R}^d$ contained in Ω (equivalently, if K is a compact subset of Ω) and if $m \in \mathbb{N}$, we write

$$\mathscr{D}_K^m(\Omega) = \{f \in \mathscr{E}^m(\Omega) : \operatorname{Supp} f \subset K\}.$$

We observe that, since K is closed, the property $\operatorname{Supp} f \subset K$ is equivalent to $\{f \neq 0\} \subset K$, or again to "$f = 0$ on $\Omega \setminus K$".
Denote by $\mathscr{K}(\Omega)$ the set of compact subsets of Ω. Put

$$\mathscr{D}^m(\Omega) = \bigcup_{K \in \mathscr{K}(\Omega)} \mathscr{D}_K^m(\Omega).$$

In other words, $\mathscr{D}^m(\Omega)$ is the space of functions of class C^m having compact support in Ω. In particular, $\mathscr{D}^0(\Omega) = C_c(\Omega)$.

Clearly, $m' \geq m$ implies $\mathscr{D}^{m'}(\Omega) \subset \mathscr{D}^m(\Omega)$. Now put

$$\mathscr{D}(\Omega) = \bigcap_{m\in\mathbb{N}} \mathscr{D}^m(\Omega).$$

Thus $\mathscr{D}(\Omega)$ is the space of functions of class C^∞ having compact support in Ω; such functions are called **test functions** on Ω. Finally, if K is a compact subset of Ω, we denote by $\mathscr{D}_K(\Omega)$ the space of functions of class C^∞ having support contained in K:

$$\mathscr{D}_K(\Omega) = \bigcap_{m\in\mathbb{N}} \mathscr{D}_K^m(\Omega) = \{f \in \mathscr{E}(\Omega) : \operatorname{Supp} f \subset K\}.$$

Thus

$$\mathscr{D}(\Omega) = \bigcup_{K\in\mathscr{K}(\Omega)} \mathscr{D}_K(\Omega).$$

Clearly, a function in $\mathscr{D}^m(\Omega)$ or $\mathscr{D}(\Omega)$, when extended with the value 0 on $\mathbb{R}^d \setminus \Omega$, becomes an element of $\mathscr{D}^m(\mathbb{R}^d)$ or $\mathscr{D}(\mathbb{R}^d)$, respectively. Thus, $\mathscr{D}^m(\Omega)$ and $\mathscr{D}(\Omega)$ can be considered as subspaces of $\mathscr{D}^m(\mathbb{R}^d)$ and $\mathscr{D}(\mathbb{R}^d)$, respectively. We will often make this identification without saying so explicitly. Conversely, an element f in $\mathscr{D}^m(\mathbb{R}^d)$ or $\mathscr{D}(\mathbb{R}^d)$ belongs to all the spaces $\mathscr{D}^m(\Omega)$ or $\mathscr{D}(\Omega)$ such that $\Omega \supset \operatorname{Supp} f$.

1B Convergence in Function Spaces

We will not need to give the function spaces just introduced a precise topological structure. It will suffice to define the notion of convergence of sequences.

Convergence in $\mathscr{D}_K^m(\Omega)$ and $\mathscr{D}_K(\Omega)$

Let K be a compact subset of Ω. We say that a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{D}_K^m(\Omega)$ converges to $f \in \mathscr{D}_K^m(\Omega)$ in $\mathscr{D}_K^m(\Omega)$ if, for every multiindex $p \in \mathbb{N}^d$ such that $|p| \leq m$, the sequence $(D^p f_n)_{n\in\mathbb{N}}$ converges uniformly to $D^p f$. An analogous definition applies with the replacement of $\mathscr{D}_K^m(\Omega)$ by $\mathscr{D}_K(\Omega)$, where now there is no restriction on the multiindex $p \in \mathbb{N}^d$.

The convergence thus defined on $\mathscr{D}_K^m(\Omega)$ clearly corresponds to convergence in the norm $\|\cdot\|^{(m)}$ defined on $\mathscr{D}_K^m(\Omega)$ by

$$\|f\|^{(m)} = \sum_{|p|\leq m} \|D^p f\|,$$

where $\|\cdot\|$ denotes the uniform norm. In contrast, no norm on $\mathscr{D}_K(\Omega)$ yields the notion of convergence we have defined in that space.

Convergence in $\mathscr{D}^m(\Omega)$ and $\mathscr{D}(\Omega)$

We say that a sequence $(\varphi_n)_{n\in\mathbb{N}}$ in $\mathscr{D}^m(\Omega)$ converges to $\varphi \in \mathscr{D}^m(\Omega)$ in $\mathscr{D}^m(\Omega)$ if there exists a compact subset K of Ω such that

$$\operatorname{Supp}\varphi \subset K \quad \text{and} \quad \operatorname{Supp}\varphi_n \subset K \quad \text{for all } n \in \mathbb{N}$$

and such that the sequence $(\varphi_n)_{n\in\mathbb{N}}$ converges to φ in $\mathscr{D}^m_K(\Omega)$. An analogous definition applies with the replacement of $\mathscr{D}^m_K(\Omega)$ and $\mathscr{D}^m(\Omega)$ by $\mathscr{D}_K(\Omega)$ and $\mathscr{D}(\Omega)$.

Convergence in $\mathscr{E}^m(\Omega)$ and $\mathscr{E}(\Omega)$

We say that a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{E}^m(\Omega)$ converges to $f \in \mathscr{E}^m(\Omega)$ if, for every multiindex p such that $|p| \le m$ and for every compact K in Ω, the sequence $(D^p f_n)_{n\in\mathbb{N}}$ converges to $D^p f$ uniformly on K. An analogous definition applies with the replacement of $\mathscr{E}^m(\Omega)$ by $\mathscr{E}(\Omega)$, where now there is no restriction on the multiindex $p \in \mathbb{N}^d$.

For $m = 0$, the convergence in $\mathscr{E}^0(\Omega)$ thus defined coincides with uniform convergence on compact subsets (defined in Exercise 12 on page 57).

We remark that the definitions of convergence of sequences just made extend immediately to families (φ_λ), where λ runs over a subset in $\mathbb{R}$ and $\lambda \to \lambda_0$, with $\lambda_0 \in [-\infty, +\infty]$.

It is possible to give the spaces $\mathscr{D}_K(\Omega)$, $\mathscr{E}^m(\Omega)$, and $\mathscr{E}(\Omega)$ complete metric structures for which convergence of sequences coincides with the notions just defined (see Exercise 7 on page 265). In contrast, one can show that the convergence we have defined in $\mathscr{D}^m(\Omega)$ and $\mathscr{D}(\Omega)$ cannot come from a metric structure.

In fact, the only topological notions that we will use in connection with these function spaces are continuity and denseness, and these notions, in the case of metric spaces, can always be expressed in terms of sequences. *In the sequel, denseness and continuity in the function spaces just introduced—in particular, in $\mathscr{D}^m(\Omega)$ and $\mathscr{D}(\Omega)$—will be defined in terms of sequences.* For example, a subset H of $\mathscr{D}^m(\Omega)$ will be called dense in $\mathscr{D}^m(\Omega)$ if, for every $\varphi \in \mathscr{D}^m(\Omega)$, there exists a sequence $(\varphi_n)_{n\in\mathbb{N}}$ in H converging to φ in $\mathscr{D}^m(\Omega)$. Likewise, a function F on $\mathscr{D}(\Omega)$ and taking values in a metric space or in one of the spaces just introduced will be called continuous if, for every sequence $(\varphi_n)_{n\in\mathbb{N}}$ in $\mathscr{D}(\Omega)$ that converges to φ in $\mathscr{D}(\Omega)$, the sequence $(F(\varphi_n))_{n\in\mathbb{N}}$ converges to $F(\varphi)$ in the space considered. One easily checks that this is equivalent to saying that the restriction of F to each metric space $\mathscr{D}_K(\Omega)$, where K is a compact subset of Ω, is continuous.

For example, the **canonical injection** from $\mathscr{D}^m(\Omega)$ into $\mathscr{E}^m(\Omega)$ — that is, the map that associates to each function $\varphi \in \mathscr{D}^m(\Omega)$ the same φ considered as an element of $\mathscr{E}^m(\Omega)$ — is continuous. This means simply that every sequence in $\mathscr{D}^m(\Omega)$ that converges in $\mathscr{D}^m(\Omega)$ also converges in $\mathscr{E}^m(\Omega)$ (to the same limit). Similarly, the canonical injections from $\mathscr{D}(\Omega)$ into $\mathscr{E}(\Omega)$ and into $\mathscr{D}^m(\Omega)$ are continuous.

1C Smoothing

We start by showing the existence of nontrivial elements of $\mathscr{D}$ (recall our convention that $\mathscr{D} = \mathscr{D}(\mathbb{R}^d)$). First take the function ρ on $\mathbb{R}$ defined by

$$\rho(x) = \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0. \end{cases}$$

Then $\rho \in \mathscr{E}(\mathbb{R})$. Indeed, one shows easily by induction that, for every integer $n \in \mathbb{N}$, ρ is of class C^n and $\rho^{(n)}$ is of the form

$$\rho^{(n)}(x) = \begin{cases} H_n(1/x)e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

where H_n is a polynomial function.

Next, for $x \in \mathbb{R}^d$, we set $\varphi(x) = \rho(1 - |x|^2)$, where, as usual, $|x|$ means the canonical euclidean norm of x in $\mathbb{R}^d$: $|x|^2 = x_1^2 + \cdots + x_d^2$. Finally, put $a = \int \varphi(x)\,dx > 0$ and $\chi = \varphi/a$. One then checks that the function χ satisfies the following properties:

$$\chi \in \mathscr{D}(\mathbb{R}^d), \quad \chi \geq 0, \quad \int \chi(x)\,dx = 1, \quad \operatorname{Supp}\chi = \bar{B}(0,1).$$

In particular, if we put $\chi_n(x) = n^d \chi(nx)$ for $n \in \mathbb{N}^*$, the sequence $(\chi_n)_{n\in\mathbb{N}^*}$ is a normal Dirac sequence (see page 174) consisting of functions of class C^∞. Such a sequence is also called a **smoothing Dirac sequence**.

Now fix a smoothing Dirac sequence (χ_n).

Proposition 1.2 *Assume $\varphi \in \mathscr{D}^m$, for some $m \in \mathbb{N}$. For every integer $n \geq 1$, the convolution $\varphi * \chi_n$ belongs to $\mathscr{D}$ and*

$$\lim_{n\to+\infty} \varphi * \chi_n = \varphi \quad \textit{in } \mathscr{D}^m.$$

Proof. Since the functions φ and χ_n have compact support, so does $\varphi * \chi_n$. More precisely,

$$\operatorname{Supp}(\varphi * \chi_n) \subset \operatorname{Supp}\varphi + \operatorname{Supp}\chi_n \subset \operatorname{Supp}\varphi + \bar{B}(0, 1/n) \subset \operatorname{Supp}\varphi + \bar{B}(0,1).$$

At the same time, a classical theorem about differentiation under the integral sign easily implies, on the one hand, that $\varphi * \chi_n$ is of class C^∞ and so $\varphi * \chi_n \in \mathscr{D}$, and, on the other, that $D^p(\varphi * \chi_n) = (D^p\varphi) * \chi_n$ for $|p| \leq m$. Now, since the support of χ_n is contained in $\bar{B}(0, 1/n)$ and $\int \chi_n(y)\,dy = 1$, we get

$$(D^p\varphi) * \chi_n(x) - (D^p\varphi)(x) = \int_{|y|\leq 1/n} \big(D^p\varphi(x-y) - D^p\varphi(x)\big)\chi_n(y)\,dy$$

and

$$\sup_{x\in\mathbb{R}^d} \left|(D^p\varphi) * \chi_n(x) - (D^p\varphi)(x)\right| \le \sup_{\substack{x,z\in\mathbb{R}^d \\ |z-x|\le 1/n}} \left|D^p\varphi(z) - D^p\varphi(x)\right|.$$

Since $D^p\varphi$ is uniformly continuous (being continuous and having compact support), we deduce that the sequence $(D^p(\varphi*\chi_n))_{n\in\mathbb{N}}$ converges uniformly to $D^p\varphi$. □

Corollary 1.3 *For every $n \in \mathbb{N}$, the space $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}^m(\Omega)$. In particular, $\mathscr{D}(\Omega)$ is dense in $C_c(\Omega)$.*

Proof. If $\varphi \in \mathscr{D}^m(\Omega)$, we can consider φ as an element of $\mathscr{D}^m$ (by extending it with the value 0 on $\mathbb{R}^d \setminus \Omega$). Now

$$\operatorname{Supp}(\varphi * \chi_n) \subset \operatorname{Supp}\varphi + \bar{B}(0, 1/n);$$

therefore $\operatorname{Supp}(\varphi*\chi_n) \subset \Omega$ for n large enough — say $n > 1/d(\operatorname{Supp}\varphi, \mathbb{R}^d\backslash\Omega)$. Then, by the preceding proposition, $\varphi * \chi_n$ belongs to $\mathscr{D}(\Omega)$ for n large enough, and $\lim_{n\to+\infty} \varphi * \chi_n = \varphi$ in $\mathscr{D}^m(\Omega)$. □

Remark. The approximating sequence just constructed preserves positivity. Therefore, if φ is a positive element of $\mathscr{D}^m(\Omega)$, there exists a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of positive elements of $\mathscr{D}(\Omega)$ that converges to φ in $\mathscr{D}^m(\Omega)$ (namely, $\varphi_n = \varphi * \chi_n$).

1D C^∞ Partitions of Unity

We now sharpen Proposition 1.8 on page 53 in the case of $\mathbb{R}^d$.

Proposition 1.4 *If K is a compact subset of $\mathbb{R}^d$ and $O_1, \dots, O_n$ are open sets in $\mathbb{R}^d$ such that $K \subset \bigcup_{j=1}^n O_j$, there exist functions $\varphi_1, \dots, \varphi_n$ in $\mathscr{D}$ such that*

$$0 \le \varphi_j \le 1 \quad \textit{and} \quad \operatorname{Supp}\varphi_j \subset O_j \qquad \textit{for } j \in \{1, \dots, n\},$$

and such that $\sum_{j=1}^n \varphi_j(x) = 1$ for every $x \in K$.

Proof. Set $d = d(K, \mathbb{R}^d\backslash O)$, with $O = \bigcup_{j=1}^n O_j$ (the metric being the canonical euclidean metric in $\mathbb{R}^d$). Set $K' = \{x : d(x, K) \le d/2\}$. The set K' is compact and, since $d > 0$,

$$\mathring{K}' \supset \{x : d(x, K) < d/2\} \supset K.$$

Thus $K \subset \mathring{K}' \subset K' \subset O$. By Proposition 1.8 on page 53, there exist functions $h_1, \dots, h_n$ in C_c such that

$$0 \le h_j \le 1 \quad \text{and} \quad \operatorname{Supp} h_j \subset O_j \qquad \text{for } j \in \{1, \dots, n\},$$

and such that $\sum_{j=1}^{n} h_j(x) = 1$ for every $x \in K'$. Define $\delta = d(K, \mathbb{R}^d \backslash \mathring{K}')$, $\eta_j = d(\operatorname{Supp} h_j, \mathbb{R}^d \backslash O_j)$ for $1 \le j \le n$, and

$$\varepsilon = \tfrac{1}{2}\min(\delta, \eta_1, \ldots, \eta_n).$$

Let χ be the function defined on page 261 and let u be defined by

$$u(x) = \varepsilon^{-d}\chi(x/\varepsilon).$$

Then $u \in \mathscr{D}$, $u \ge 0$, $\int u(x)\,dx = 1$, and $\operatorname{Supp} u = \bar{B}(0,\varepsilon)$.

For $1 \le j \le n$, set $\varphi_j = h_j * u$. Then φ_j is of class C^∞ (this follows immediately from the theorem on differentiation under the integral sign) and

$$\operatorname{Supp}\varphi_j \subset \operatorname{Supp} h_j + \bar{B}(0,\varepsilon) \subset O_j.$$

In particular, $\varphi_j \in \mathscr{D}$. Moreover, $0 \le \varphi_j \le 1$. Finally, if $x \in K$ and $y \in \bar{B}(0,\varepsilon)$, we have $x - y \in K'$ and so

$$\sum_{j=1}^{n} h_j(x-y)u(y) = u(y).$$

Integrating we obtain

$$\sum_{j=1}^{n} \varphi_j(x) = \int u(y)\,dy = 1 \quad \text{for all } x \in K. \qquad \square$$

We deduce the following denseness result:

Proposition 1.5 *The space $\mathscr{D}(\Omega)$ is dense in $\mathscr{E}(\Omega)$ and in $\mathscr{E}^m(\Omega)$, for every $m \in \mathbb{N}$.*

Proof. Let $(K_n)_{n\in\mathbb{N}}$ be a sequence of compact subsets of Ω exhausting Ω. By the previous proposition, there exists, for every integer $n \in \mathbb{N}$, an element $\varphi_n \in \mathscr{D}(\Omega)$ such that

$$0 \le \varphi_n \le 1, \qquad \varphi_n = 1 \text{ on } K_n, \qquad \operatorname{Supp}\varphi_n \subset K_{n+1}.$$

If $f \in \mathscr{E}(\Omega)$, we have $f\varphi_n \in \mathscr{D}(\Omega)$ for every $n \in \mathbb{N}$. If K is a compact subset of Ω, there exists $N \in \mathbb{N}$ such that $K \subset \mathring{K}_N$ (see Proposition 1.6 on page 52); thus, for every $n \ge N$ and every $p \in \mathbb{N}^d$, we have $D^p(f\varphi_n) = D^p f$ on K. By the definition of convergence in $\mathscr{E}(\Omega)$, we deduce that $\lim_{n\to+\infty}(f\varphi_n) = f$ in $\mathscr{E}(\Omega)$.

Using the same reasoning, one shows that $\mathscr{D}^m(\Omega)$ is dense in $\mathscr{E}^m(\Omega)$. Moreover, as we saw in Corollary 1.3, $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}^m(\Omega)$. Thus every element of $\mathscr{D}^m(\Omega)$ is the limit of a sequence of elements of $\mathscr{D}(\Omega)$ in the sense of convergence in $\mathscr{E}^m(\Omega)$ (since the canonical injection from $\mathscr{D}^m(\Omega)$ into $\mathscr{E}^m(\Omega)$ is continuous: see page 260). This implies, finally, that $\mathscr{D}(\Omega)$ is dense in $\mathscr{E}^m(\Omega)$ (because $\mathscr{E}^m(\Omega)$ is a metric space: see Exercise 7 below). $\square$

Remark. This proof also shows that every positive element of $\mathscr{E}^m(\Omega)$ (or $\mathscr{E}(\Omega)$) is the limit in $\mathscr{E}^m(\Omega)$ (or in $\mathscr{E}(\Omega)$, respectively) of a sequence of positive elements of $\mathscr{D}(\Omega)$.

Exercises

Throughout the exercises, Ω stands for an open subset of $\mathbf{R}^d$. Many of the exercises use the result of Exercise 1.

1. *Taylor's formula with integral remainder.* Let f be an element of $\mathscr{E}^n(\Omega)$ (where $n \geq 1$) and let $x \in \Omega$. Take $h \in \mathbf{R}^d$ such that $[x, x+h] \subset \Omega$. Show that

$$f(x+h) = f(x) + \sum_{k=1}^{n-1} \left(\frac{1}{k!} \sum_{|p|=k} D^p f(x)\, h^p \right)$$
$$+ \frac{1}{(n-1)!} \int_0^1 \sum_{|p|=n} D^p f(x+th)\, h^p (1-t)^{n-1}\, dt,$$

where, for $p = (p_1, \ldots, p_d) \in \mathbf{N}^d$ and $h = (h_1, \ldots, h_d) \in \mathbf{R}^d$, we have written

$$h^p = h_1^{p_1} \ldots h_d^{p_d}.$$

2. Take $h \in \mathscr{E}(\mathbf{R})$. Show that the function f defined by

$$f(x,y) = \frac{h(x) - h(y)}{x - y}$$

can be extended by continuity to a function in $\mathscr{E}(\mathbf{R}^2)$.
If we assume merely that $h \in \mathscr{E}^n(\mathbf{R})$, with $n \geq 1$, how smooth is in general the function obtained in this way?

3. Let $h \in \mathscr{E}(\mathbf{R})$ be such that $h(0) = h'(0) = \cdots = h^{(n)}(0) = 0$. Show that the function $f(x) = x^{-n-1} h(x)$ can be extended to an element of $\mathscr{E}(\mathbf{R})$. What is the value of this new function at 0?

4. Take $f \in \mathscr{E}(\mathbf{R}^d)$. Show that f satisfies

$$D^p f(0) = 0 \quad \text{for all } p \in \mathbf{N}^d \text{ with } |p| \leq m$$

if and only if there exists a family $(\varphi_j)_{j \in \mathbf{N}^d,\, |j| = m+1}$ of elements of $\mathscr{E}(\mathbf{R}^d)$ such that

$$f(x) = \sum_{|j|=m+1} \varphi_j(x) x^j$$

(where $x^j = x_1^{j_1} \ldots x_d^{j_d}$).

5. **a.** Let E be a closed subset of Ω. Show that there exists a positive function $f \in \mathscr{E}(\Omega)$ such that $E = f^{-1}(0)$.
You could work as follows:
   i. First show the result assuming that E is the complement in Ω of an open ball (in the euclidean metric).
   ii. Let (f_n) be a countable family of functions in $\mathscr{E}(\Omega)$. Show that there exist positive real numbers μ_n such that the series of functions $\sum \mu_n f_n$ converges in $\mathscr{E}(\Omega)$.

iii. Wrap up using the fact that $\Omega \setminus E$ is a countable union of open balls.

b. Let E and F be disjoint closed subsets of Ω. Show that there exists a function f in $\mathscr{E}(\Omega)$ such that $0 \le f \le 1$, $E = f^{-1}(0)$, and $F = f^{-1}(1)$.
Hint. Let φ and ψ be positive functions in $\mathscr{E}(\Omega)$ such that $\varphi^{-1}(0) = E$ and $\psi^{-1}(0) = F$. Check that $f = \varphi / \sqrt{\varphi^2 + \psi^2}$ satisfies the desired conditions.

c. Let E be a closed subset of Ω. Prove that E is the support of a function in $\mathscr{E}(\Omega)$ if and only if E equals (in Ω) the closure of its interior.

6. *Borel's Theorem.* Let (a_n) be an arbitrary sequence of complex numbers. Show that there exists a function $f \in \mathscr{E}(\mathbb{R})$ such that $f^{(k)}(0) = a_k$ for every integer k.
Some hints:

a. Let $\varphi \in \mathscr{D}(\mathbb{R})$ be such that $\varphi = 1$ in $[-1, 1]$. For $n \in \mathbb{N}$, set

$$f_n(x) = \frac{a_n}{n!}\, x^n \varphi(\mu_n x).$$

Show that one can choose the μ_n in such a way that $\|f_n\|^{(n-1)} \le 2^{-n}$ for every $n \ge 1$.

b. Show that the series $\sum f_n$ converges to a function having the desired property.

7. *Topologizing spaces of smooth functions*

a. Let K be a compact subset of Ω and take $m \in \mathbb{N}$. Show that the space $\mathscr{D}_K^m(\Omega)$ with the norm $\|\cdot\|^{(m)}$ is a Banach space.

b. If $f, g \in \mathscr{D}_K(\Omega)$, define

$$d(f, g) = \sum_{m=0}^{\infty} 2^{-m} \min\bigl(\|f - g\|^{(m)}, 1\bigr).$$

Show that d is a complete metric on $\mathscr{D}_K(\Omega)$ and that a sequence converges in this metric if and only if it converges in $\mathscr{D}_K(\Omega)$ (in the sense defined in the text).

c. Take $m \in \mathbb{N}$ and let $(K_n)_{n \in \mathbb{N}}$ be an exhausting sequence of compact subsets of Ω (see page 52). If $f, g \in \mathscr{E}^m(\Omega)$, define

$$\delta_\Omega^m(f, g) = \sum_{n=0}^{\infty} 2^{-n} \min\bigl(\|f - g\|_{K_n}^{(m)}, 1\bigr),$$

with

$$\|f\|_{K_n}^{(m)} = \sum_{|p| \le m} \sup_{x \in K_n} \bigl|D^p f(x)\bigr|.$$

Show that δ_Ω^m is a complete metric on $\mathscr{E}^m(\Omega)$ and that a sequence converges in this metric if and only if it converges in $\mathscr{E}^m(\Omega)$ (in the sense defined in the text).

d. Same questions for the metric δ_Ω defined on $\mathscr{E}(\Omega)$ by

$$\delta_\Omega(f,g) = \sum_{m=0}^{\infty} 2^{-m} \min\bigl(\delta_\Omega^m(f,g), 1\bigr).$$

8. a. Let P be the linear operator on $\mathscr{D}^n(\Omega)$ defined by

$$Pf(x) = \sum_{|p|\le m} a_p(x) D^p f(x),$$

where $m \le n$ and where each function a_p belongs to $\mathscr{E}^{n-m}(\Omega)$. Show that P is a continuous operator from $\mathscr{D}^n(\Omega)$ to $\mathscr{D}^{n-m}(\Omega)$.

b. Suppose the functions a_p lie in $\mathscr{E}(\Omega)$. Show that P defines a continuous linear operator from $\mathscr{D}(\Omega)$ to $\mathscr{D}(\Omega)$ and from $\mathscr{E}(\Omega)$ to $\mathscr{E}(\Omega)$.

9. Suppose $u \in \mathbb{R}^d$ and $\varphi \in \mathscr{D}(\mathbb{R}^d)$. If $h \in \mathbb{R}^*$, define an element φ_h of $\mathscr{D}(\mathbb{R}^d)$ by setting

$$\varphi_h(x) = \frac{\varphi(x+hu) - \varphi(x)}{h}.$$

Show that the sequence $(\varphi_{1/n})_{n\in\mathbb{N}^*}$ converges in $\mathscr{D}(\mathbb{R}^d)$. Find its limit.

10. Let $\varphi \in \mathscr{D}$ be nonzero. If $n \in \mathbb{N}^*$, set

$$\varphi_n(x) = \frac{1}{n}\varphi(x/n) \quad \text{for } x \in \mathbb{R}^d.$$

Show that the sequence $(\varphi_n)_{n\in\mathbb{N}^*}$ converges to 0 in $\mathscr{E}$ but not in $\mathscr{D}$.

11. Let $O_1, \dots, O_n$ be open subsets of $\mathbb{R}^d$ such that $\Omega = \bigcup_{j=1}^n O_j$, and take $\varphi \in \mathscr{D}(\Omega)$. Show that there exist functions $\varphi_1, \dots, \varphi_n$ in $\mathscr{D}(\Omega)$ satisfying

$$\operatorname{Supp} \varphi_j \subset O_j \quad \text{for all } j \in \{1, \dots, n\}$$

and such that $\sum_{j=1}^n \varphi_j = \varphi$. Check also that, if $\varphi \ge 0$, the functions φ_j can be chosen to be positive.

12. a. Let f be a real-valued element of $\mathscr{E}^m(\Omega)$ or $\mathscr{E}(\Omega)$. Show that there exist positive-valued functions f_1, f_2 in $\mathscr{E}^m(\Omega)$ or $\mathscr{E}(\Omega)$, respectively, such that $f = f_1 - f_2$.

Hint. Take $f_1 = f^2 + 1$ and $f_2 = f^2 - f + 1$.

b. Show that analogous results hold for $\mathscr{D}^m(\Omega)$ and $\mathscr{D}(\Omega)$ instead of $\mathscr{E}^m(\Omega)$ and $\mathscr{E}(\Omega)$.

13. Suppose $f \in C(\Omega)$ satisfies

$$\int f(x)\varphi(x)\,dx = 0 \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

Show that $f = 0$.

14. Let $(O_j)_{j\in J}$ be a locally finite relatively compact open cover of Ω; that is, each O_j is a relatively compact open subset of Ω, the union $\bigcup_{j\in J} O_j$ equals Ω, and, for every compact K in Ω, the set $\{j \in J : O_j \cap K \neq \varnothing\}$ is finite. Show that there exists a family $(\varphi_j)_{j\in J}$ of elements of $\mathscr{D}(\mathbb{R}^d)$ such that

$$\operatorname{Supp}\varphi_j \subset O_j \quad\text{and}\quad 0 \le \varphi_j \le 1 \quad\text{for all } j \in J$$

and

$$\sum_{j\in J} \varphi_j(x) = 1 \quad\text{for all } x \in \Omega.$$

Hint. Let $(K_n)_{n\in\mathbb{N}}$ be an exhausting sequence of compact subsets of Ω such that $K_0 = \varnothing$. For each $n \in \mathbb{N}^*$, put

$$J_n = \bigl\{j \in J : O_j \cap (\mathring{K}_{n+2}\backslash K_{n-1}) \neq \varnothing\bigr\}.$$

Consider a C^∞ partition of unity $(\varphi_j^n)_{j\in J_n}$ with respect to the compact $K_{n+1} \setminus \mathring{K}_n$ and to the finite open family $\bigl\{O_j \cap (\mathring{K}_{n+2}\backslash K_{n-1})\bigr\}_{j\in J_n}$. For each $j \in \bigcup_{n\in\mathbb{N}^*} J_n$, define

$$\varphi_j = \frac{\displaystyle\sum_{n\in\mathbb{N}_j} \varphi_j^n}{\displaystyle\sum_{n\in\mathbb{N}}\sum_{k\in J_n} \varphi_k^n},$$

where $\mathbb{N}_j = \{n \in \mathbb{N}^* : j \in J_n\}$.

2 Distributions

2A Definitions

By definition, a **distribution** on Ω is a continuous linear form on $\mathscr{D}(\Omega)$. Thus, by what we saw in Section 1B, a linear form T on $\mathscr{D}(\Omega)$ is a distribution if, for every sequence $(\varphi_n)_{n\in\mathbb{N}}$ that converges to 0 in $\mathscr{D}(\Omega)$, the sequence $(T(\varphi_n))_{n\in\mathbb{N}}$ tends to 0 (in $\mathbb{C}$); equivalently, if, for every compact subset K of Ω, the restriction of T to the metric space $\mathscr{D}_K(\Omega)$ defined in Exercise 7 on page 265 is continuous. We denote by $\mathscr{D}'(\Omega)$ the set of distributions on Ω; of course $\mathscr{D}'(\Omega)$ is a vector space.

If $\Omega = \mathbb{R}^d$, we will sometimes use the simplified notation $\mathscr{D}' = \mathscr{D}'(\mathbb{R}^d)$. Also, if T is a distribution on Ω and $\varphi \in \mathscr{D}(\Omega)$, we denote by

$$T(\varphi) = \langle T, \varphi\rangle$$

the result of evaluating the distribution T at the test function φ.

Proposition 2.1 *Let T be a linear form on $\mathscr{D}(\Omega)$. Then T is a distribution on Ω if and only if, for every compact K in Ω, there exist $m \in \mathbb{N}$ and $C \geq 0$ such that*

$$\left|T(\varphi)\right| \leq C\|\varphi\|^{(m)} \quad \textit{for all } \varphi \in \mathscr{D}_K(\Omega).$$

Proof. The "if" part follows easily from the definitions. Conversely, suppose that the criterion is not satisfied. Then there exists a compact subset K of Ω and a sequence $(\varphi_n)_{n\in\mathbb{N}}$ in $\mathscr{D}_K(\Omega)$ such that

$$\left|T(\varphi_n)\right| > n \left\|\varphi_n\right\|^{(n)} \quad \text{for every } n \in \mathbb{N}.$$

For every $n \geq 1$, set

$$\psi_n = \frac{1}{n\left\|\varphi_n\right\|^{(n)}}\,\varphi_n.$$

Obviously, $\psi_n \in \mathscr{D}_K(\Omega)$; moreover, for every $m \in \mathbb{N}$,

$$\left\|\psi_n\right\|^{(m)} \leq \left\|\psi_n\right\|^{(n)} \leq 1/n \quad \text{for all } n > m.$$

Thus the sequence $(\psi_n)_{n\in\mathbb{N}}$ converges to 0 in $\mathscr{D}_K(\Omega)$. Now $|T(\psi_n)| \geq 1$ for every $n \in \mathbb{N}$, so the sequence $(T(\psi_n))_{n\in\mathbb{N}}$ does not converge to 0. Therefore T is not distribution on Ω. □

Order of a distribution

A distribution T on Ω is said to have **finite order** if there exists an integer $m \in \mathbb{N}$ with the following property:

For any compact subset K of Ω there exists $C \geq 0$ such that $\left|T(\varphi)\right| \leq C\|\varphi\|^{(m)}$ for all $\varphi \in \mathscr{D}_K(\Omega)$. $(*)$

In other words, T has finite order if the integer m that appears in Proposition 2.1 can be made independent of the compact $K \subset \Omega$. If T has finite order, the **order** of T is, by definition, the smallest integer m for which $(*)$ is satisfied.

2B First Examples

Locally integrable functions

Let $L^1_{\text{loc}}(\Omega)$ be the space of equivalence classes (with respect to Lebesgue measure) of locally integrable functions f on Ω; "locally integrable" means that, for every compact subset K of Ω, $1_K f$ lies in $L^1(\Omega)$, the L^1-space corresponding to Lebesgue measure restricted to Ω. (See Exercise 19 on page 159.) If $f \in L^1_{\text{loc}}(\Omega)$, we define a distribution $[f]$ by

$$\langle [f], \varphi\rangle = \int_\Omega \varphi(x) f(x)\,dx \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

One easily checks that $[f]$ is a distribution of order 0 on Ω: Given a compact subset K of Ω, just take $C = \int_K |f(x)|\,dx$ in order to get the inequality in $(*)$, with $m = 0$.

Proposition 2.2 *Two locally integrable functions on Ω define the same distribution if and only if they coincide almost everywhere.*

Proof. Take $f \in L^1_{\mathrm{loc}}(\Omega)$ such that $[f] = 0$. Because $\mathscr{D}(\Omega)$ is dense in $C_c(\Omega) = \mathscr{D}^0(\Omega)$ (Corollary 1.3), we see that

$$\int_\Omega g(x) f(x)\,dx = 0 \quad \text{for all } g \in C_c(\Omega).$$

Thus, for every $g \in C_c^{\mathbb{R}}(\Omega)$,

$$\int_\Omega g(x)\big(\operatorname{Re} f(x)\big)^+\,dx = \int_\Omega g(x)\big(\operatorname{Re} f(x)\big)^-\,dx\,,$$
$$\int_\Omega g(x)\big(\operatorname{Im} f(x)\big)^+\,dx = \int_\Omega g(x)\big(\operatorname{Im} f(x)\big)^-\,dx\,.$$

By the uniqueness part of the Radon–Riesz Theorem (page 69), these equalities are valid for any positive Borel function g. Applying them to the characteristic functions of the sets $\{\operatorname{Re} f > 0\}$, $\{\operatorname{Re} f < 0\}$, $\{\operatorname{Im} f > 0\}$ and $\{\operatorname{Im} f < 0\}$, we deduce that $f = 0$ almost everywhere. □

Thus, the map that associates to each $f \in L^1_{\mathrm{loc}}(\Omega)$ the distribution $[f] \in \mathscr{D}'(\Omega)$ is injective. By identifying f with $[f]$, we can write $L^1_{\mathrm{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$. It is in this sense that distributions are "generalized functions".

From now on we will omit the brackets from the notation if there is no danger of confusion, and we will normally not distinguish between a locally integrable function and the distribution defined thereby.

Radon measures

More generally, every complex Radon measure μ on Ω defines a distribution T, as follows:

$$\langle T, \varphi\rangle = \int \varphi\,d\mu \quad \text{for all } \varphi \in \mathscr{D}(\Omega). \tag{$*$}$$

By the very definition of a complex Radon measure, we see that the linear form T thus defined is a distribution of order 0. Because $\mathscr{D}(\Omega)$ is dense in $C_c(\Omega)$ (Corollary 1.3), the map $\mu \mapsto T$ defined in this way is injective. Thus we can identify a Radon measure with the distribution it defines, and we can write $\mathfrak{M}(\Omega) \subset \mathscr{D}'(\Omega)$.

If μ is a positive Radon measure, the distribution T it defines is **positive**, that is,

$$\langle T, \varphi\rangle \geq 0 \quad \text{for any positive } \varphi \in \mathscr{D}(\Omega).$$

We now show the converse.

Proposition 2.3 *Every positive distribution has order* 0.

Proof. Let T be a positive distribution on Ω. Let K be a compact subset of Ω and let $\rho \in \mathscr{D}(\Omega)$ be such that $0 \le \rho \le 1$ and $\rho = 1$ on K. For every $\varphi \in \mathscr{D}_K(\Omega)$, we have $|\varphi| = |\varphi\rho| \le \|\varphi\|\,\rho$, where $\|\varphi\|$ denotes the uniform norm of φ. If φ is real-valued, this means that

$$-\|\varphi\|\,\rho \le \varphi \le \|\varphi\|\,\rho.$$

We then deduce from the linearity and the positivity of T that $|T(\varphi)| \le \|\varphi\|\,T(\rho)$. When we no longer assume φ to be real-valued, the decomposition $\varphi = \operatorname{Re}\varphi + i\operatorname{Im}\varphi$ leads to the inequality $|T(\varphi)| \le 2T(\rho)\,\|\varphi\|$, which proves that T has order 0. □

We will see later, as a particular case of Proposition 3.1, that in fact every distribution of order 0 can be obtained from a Radon measure by means of (∗) on the previous pages. Positive distributions then correspond exactly to positive Radon measures: If a Radon measure μ satisfies $\int \varphi\,d\mu \ge 0$ for every positive $\varphi \in \mathscr{D}(\Omega)$, the remark following Corollary 1.3 implies that the same is true for every positive $f \in C_c(\Omega)$.

Distributions of nonzero finite order

Let m be a positive integer. A simple example of a distribution of order m on an arbitrary open set Ω is the distribution T defined by

$$\langle T, \varphi\rangle = (D^p\varphi)(a) \quad \text{for all } \varphi \in \mathscr{D}(\Omega),$$

where p is a multiindex of length m and a is any point of Ω. That T is a distribution of order at most m follows directly from the definitions. To prove that the order cannot be less than m, consider a function $\psi \in \mathscr{D}(\mathbb{R}^d)$ such that $\psi(0) = 1$ and $\operatorname{Supp}\psi \subset \bar{B}(0,1)$. For every $\alpha > 0$, put

$$\varphi_\alpha(x) = (x-a)^p\,\psi\big((x-a)/\alpha\big),$$

where, for $y \in \mathbb{R}^d$, we have set $y^p = y_1^{p_1}\dots y_d^{p_d}$. Since the support of φ_α is contained in $\bar{B}(a,\alpha)$, we see that, at least for $\alpha \le \alpha_0 < d(a, \mathbb{R}^d\backslash\Omega)$, we have $\varphi_\alpha \in \mathscr{D}(\Omega)$. Moreover, we deduce easily from Leibniz's formula that, first, $\langle T, \varphi_\alpha\rangle = p!$ for every $\alpha > 0$, and secondly, if q is a multiindex of length strictly less than m, then

$$(D^q\varphi_\alpha)(x) = \sum_{r\le q} C_{r,q}\,(x-a)^{p-q+r}\,\alpha^{-|r|}D^r\,\psi\big((x-a)/\alpha\big),$$

so that the uniform norm of $D^q\varphi_\alpha$, when $\alpha \le 1$, satisfies

$$\|D^q\varphi_\alpha\| \le C_q\alpha^{|p|-|q|} \le C_q\alpha,$$

where the constant C_q depends only on q and on the chosen function ψ. It follows that

$$\|\varphi_\alpha\|^{(m-1)} \le C\alpha,$$

where the constant C depends only on ψ. Since all the functions φ_α are supported in the compact $K = \bar{B}(a, \alpha_0)$, this makes it impossible for condition $(*)$ on page 268 to hold with m replaced by $m-1$. Therefore T has order exactly m.

A distribution of infinite order

Let T be the linear form on $\mathscr{D}(\mathbb{R})$ defined by

$$\langle T, \varphi \rangle = \sum_{n=0}^{+\infty} \varphi^{(n)}(n) \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}).$$

Since the intersection of any compact subset of $\mathbb{R}$ with $\mathbb{N}$ is finite, this sum has only finitely many nonzero terms. Moreover, it is clear that, if K is a compact subset of $\mathbb{R}$ and $N = \max(\mathbb{N} \cap K)$, we have

$$\bigl|\langle T, \varphi \rangle\bigr| \le \|\varphi\|^{(N)} \quad \text{for every } \varphi \in \mathscr{D}_K(\mathbb{R}),$$

which proves that T is a distribution.

Now take $m \in \mathbb{N}$ and set $K = \bigl[m - \frac{1}{2}, m + \frac{1}{2}\bigr]$. For every $\varphi \in \mathscr{D}_K(\mathbb{R})$, we have $\langle T, \varphi \rangle = \varphi^{(m)}(m)$. It follows form the preceding example that the smallest integer n for which there exists $C > 0$ with

$$\bigl|\langle T, \varphi \rangle\bigr| \le C\|\varphi\|^{(n)} \quad \text{for all } \varphi \in \mathscr{D}_K(\mathbb{R})$$

is m. Thus the distribution T cannot have order less than m, and this for every $m \in \mathbb{N}$. This means T has infinite order.

2C Restriction and Extension of a Distribution to an Open Set

Let T be a distribution on Ω and let Ω' be an open subset of Ω. We know that $\mathscr{D}(\Omega')$ can be identified with a subspace of $\mathscr{D}(\Omega)$ (by extending each function of $\mathscr{D}(\Omega')$ to Ω with the value 0 on $\Omega \setminus \Omega'$). Thus we can define the restriction T_0 of T to $\mathscr{D}(\Omega')$, which is certainly a distribution on Ω', called the **restriction of T to Ω'**. Conversely, T is called an **extension of T_0 to Ω**.

Remark. The expression "restriction of T to Ω'" is an abuse of language, since the domain of T is the set of test functions $\mathscr{D}(\Omega)$, and not Ω itself. A similar remark applies to "extension of T_0 to Ω".

2D Convergence of Sequences of Distributions

By definition, a sequence $(T_n)_{n\in\mathbb{N}}$ in $\mathscr{D}'(\Omega)$ converges to $T \in \mathscr{D}'(\Omega)$ if

$$\lim_{n\to+\infty} \langle T_n, \varphi\rangle = \langle T, \varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

Therefore this notion is a type of weak convergence.

This definition extends immediately to families (T_λ) in $\mathscr{D}'(\Omega)$, where λ ranges over a subset of $\mathbb{R}$ and tends to $\lambda_0 \in [-\infty, +\infty]$. For example, when we write $\lim_{\varepsilon\to 0} T_\varepsilon = T$ in $\mathscr{D}'(\Omega)$ we mean that $T_\varepsilon, T \in \mathscr{D}'(\Omega)$ and that

$$\lim_{\varepsilon\to 0} \langle T_\varepsilon, \varphi\rangle = \langle T, \varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

We now give an example of a distribution defined as a limit of distributions.

2E Principal Values

Consider the function $x \mapsto 1/x$ from $\mathbb{R}$ to $\mathbb{R}$. This function is clearly not locally integrable on $\mathbb{R}$, but it is on $\mathbb{R}^*$. We will see how we can extend to $\mathbb{R}$ the distribution defined by this function on $\mathbb{R}^*$.

Proposition 2.4 *For every $\varphi \in \mathscr{D}(\mathbb{R})$, the limit*

$$T(\varphi) = \lim_{\varepsilon\to 0^+} \int_{\{|x|>\varepsilon\}} \frac{\varphi(x)}{x}\,dx \tag{$*$}$$

exists. The linear form T thus defined is a distribution of order 1 *on $\mathbb{R}$, and is an extension to $\mathbb{R}$ of the distribution $[1/x] \in \mathscr{D}'(\mathbb{R}^*)$.*

We call T the **principal value** of $1/x$ and denote it by $\operatorname{pv}(1/x)$.

Proof. Take $\varphi \in \mathscr{D}(\mathbb{R})$ and $A > 0$ such that $\operatorname{Supp}\varphi \subset [-A, A]$. If $\varepsilon < A$,

$$\begin{aligned}\int_{\{|x|>\varepsilon\}} \frac{\varphi(x)}{x}\,dx &= \int_{-A}^{-\varepsilon} \frac{\varphi(x)}{x}\,dx + \int_{\varepsilon}^{A} \frac{\varphi(x)}{x}\,dx \\ &= \int_{-A}^{-\varepsilon} \frac{\varphi(x)-\varphi(0)}{x}\,dx + \int_{\varepsilon}^{A} \frac{\varphi(x)-\varphi(0)}{x}\,dx,\end{aligned}$$

because $1/x$ is an odd function. Since $\big(\varphi(x)-\varphi(0)\big)/x$ can be continuously extended to the point 0 with the value $\varphi'(0)$, we get

$$\lim_{\varepsilon\to 0^+} \int_{\{|x|>\varepsilon\}} \frac{\varphi(x)}{x}\,dx = \int_{-A}^{A} \frac{\varphi(x)-\varphi(0)}{x}\,dx.$$

At the same time, by the Mean Value Theorem,

$$\left|\int_{-A}^{A} \frac{\varphi(x)-\varphi(0)}{x}\,dx\right| \le 2A\,\|\varphi\|^{(1)}.$$

This shows that equation $(*)$ defines a distribution of order at most 1. On the other hand, if $\varphi \in \mathscr{D}(\mathbb{R}^*)$, there exists a real $\delta > 0$ such that $\varphi = 0$ on $[-\delta, \delta]$, so

$$\lim_{\varepsilon \to 0^+} \int_{\{|x|>\varepsilon\}} \frac{\varphi(x)}{x}\,dx = \int_{\{|x|>\delta\}} \frac{\varphi(x)}{x}\,dx = \int_{\mathbb{R}^*} \frac{\varphi(x)}{x}\,dx.$$

This shows that $\mathrm{pv}(1/x)$ coincides with $[1/x]$ on $\mathscr{D}(\mathbb{R}^*)$. It remains to prove that the distribution $\mathrm{pv}(1/x)$ has order 1, which will follow if we show that it does not have order 0. For each integer $n \geq 2$, take $\psi_n \in \mathscr{D}(\mathbb{R})$ such that $0 \leq \psi_n \leq 1$, $\operatorname{Supp}\psi_n \subset (0,1)$ and $\psi_n = 1$ on $[1/n, (n-1)/n]$. Let φ_n be the odd function that coincides with ψ_n on $\mathbb{R}^+$. If $K = [-1,1]$, we have $\varphi_n \in \mathscr{D}_K(\mathbb{R})$, $\|\varphi_n\| = 1$, and

$$\big\langle \mathrm{pv}(1/x), \varphi_n \big\rangle = 2\int_0^1 \frac{\psi_n(x)}{x}\,dx \geq 2\log(n-1).$$

Thus there is no constant $C \geq 0$ such that

$$\big|\langle \mathrm{pv}(1/x), \varphi \rangle\big| \leq C\,\|\varphi\| \quad \text{for all } \varphi \in \mathscr{D}_K(\mathbb{R}),$$

proving the desired result. □

Another calculation of a principal value is given in Exercise 7 on page 291.

2F Finite Parts

In the previous example we used the fact that the function $1/x$ is odd in order to define the distribution $\mathrm{pv}(1/x)$ as the limit, when ε tends to 0, of the distribution defined on $\mathbb{R}$ by the locally integrable function $1_{\{|x|>\varepsilon\}}(x)/x$. If we are dealing with a function that is not odd, this approximation procedure does not converge, and it is necessary to apply a correction, represented by a divergent term. This is called the method of finite parts, and we will illustrate it with two examples.

We first introduce some notation that will often be useful. We define the **Heaviside function**, denoted by Y, as the characteristic function of $\mathbb{R}^+$. Thus, for $x \in \mathbb{R}$, we have $Y(x) = 0$ if $x < 0$ and $Y(x) = 1$ if $x \geq 0$.

Proposition 2.5 *For every $\varphi \in \mathscr{D}(\mathbb{R})$, the limit*

$$\langle T, \varphi \rangle = \lim_{\varepsilon \to 0^+} \left(\int_\varepsilon^{+\infty} \frac{\varphi(x)}{x}\,dx + \varphi(0)\log\varepsilon \right)$$

exists. The linear form T thus defined is a distribution of order 1 *on* $\mathbb{R}$, *called the **finite part** of $Y(x)/x$ and denoted by* $\mathrm{fp}(Y(x)/x)$.

Proof. Take $\varphi \in \mathscr{D}(\mathbb{R})$ and $A > 0$ such that $\operatorname{Supp}\varphi \subset [-A, A]$. Then

$$\int_\varepsilon^{+\infty} \frac{\varphi(x)}{x}\,dx = \int_\varepsilon^A \frac{\varphi(x)-\varphi(0)}{x}\,dx + \varphi(0)\log A - \varphi(0)\log\varepsilon.$$

Thus

$$\lim_{\varepsilon\to 0+}\left(\int_\varepsilon^{+\infty} \frac{\varphi(x)}{x}\,dx + \varphi(0)\log\varepsilon\right) = \int_0^A \frac{\varphi(x)-\varphi(0)}{x}\,dx + \varphi(0)\log A,$$

and this expression is bounded in absolute value by $\|\varphi\|^{(1)}\max(A, |\log A|)$, by the Mean Value Theorem. It follows that $\operatorname{fp}(Y(x)/x)$ is indeed a distribution of order at most 1.

For each integer $n \geq 2$, take $\psi_n \in \mathscr{D}(\mathbb{R})$ such that $0 \leq \psi_n \leq 1$, $\operatorname{Supp}\psi_n \subset (0,1)$, and $\psi_n = 1$ on $[1/n,\ (n-1)/n]$. We see that

$$\psi_n \in \mathscr{D}_{[0,1]}(\mathbb{R}),\quad \|\psi_n\| = 1,\quad \text{and}\quad \big\langle \operatorname{fp}(Y(x)/x),\ \psi_n\big\rangle \geq \log(n-1),$$

which proves that $\operatorname{fp}(Y(x)/x)$ is not of order 0, and so is of order 1. □

Other examples of finite parts on $\mathbb{R}$ will be given in Exercises 3 and 19. Here is another example, this time on $\mathbb{R}^2$. Put $r = \sqrt{x^2+y^2}$ and

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}.$$

Proposition 2.6 *For every $\varphi \in \mathscr{D}(\mathbb{R}^2)$, the limit*

$$\langle T, \varphi\rangle = \lim_{\varepsilon\to 0+}\left(\iint_{\{r\geq\varepsilon\}} r^{-4}\varphi\,dx\,dy - \pi\varphi(0,0)\varepsilon^{-2} + \frac{\pi}{2}\Delta\varphi(0,0)\log\varepsilon\right)$$

exists. The linear form T thus defined is a distribution of order 3 on $\mathbb{R}^2$.

T is called the **finite part** of $1/r^4$ and is denoted by $\operatorname{fp}(1/r^4)$. (Note that the function $1/r^4$ is not locally integrable on $\mathbb{R}^2$.)

Summary of proof. Take $\varphi \in \mathscr{D}(\mathbb{R}^2)$ and $A > 0$ such that $\operatorname{Supp}\varphi \subset \bar{B}(0, A)$. A quick calculation shows that

$$\begin{aligned}
&\iint_{\{r\geq\varepsilon\}} \frac{\varphi}{r^4}\,dx\,dy \\
&\quad = \iint_{\{A\geq r\geq\varepsilon\}} r^{-4}\bigg(\varphi(x,y) - \varphi(0,0) - x\,\frac{\partial\varphi}{\partial x}(0,0) - y\,\frac{\partial\varphi}{\partial y}(0,0) \\
&\qquad\qquad - \Big(\frac{x^2}{2}\,\frac{\partial^2\varphi}{\partial x^2}(0,0) + xy\,\frac{\partial^2\varphi}{\partial x\partial y}(0,0) + \frac{y^2}{2}\,\frac{\partial^2\varphi}{\partial y^2}(0,0)\Big)\bigg)\,dx\,dy \\
&\qquad\qquad - \pi\varphi(0,0)(A^{-2} - \varepsilon^{-2}) + \frac{\pi}{2}\Delta\varphi(0,0)\log\frac{A}{\varepsilon}.
\end{aligned}$$

We then deduce from Taylor's formula that the limit given in the statement of the proposition exists and is bounded in absolute value by $C_A\|\varphi\|^{(3)}$, with $C_A > 0$. Therefore the distribution thus defined has order at most 3. It remains to show that it is not of order less than 3. $\square$

Exercises

1. Let $\mathscr{D}^{\mathbb{R}}(\Omega)$ be the set of real-valued elements of $\mathscr{D}(\Omega)$. A distribution T on Ω is called *real* if $\langle T, \varphi\rangle \in \mathbb{R}$ for every $\varphi \in \mathscr{D}^{\mathbb{R}}(\Omega)$. Show that every distribution T on Ω can be written in a unique way as $T = T_1 + iT_2$, where T_1 and T_2 are real distributions on Ω.

Show that real distributions can be identified with continuous linear forms on $\mathscr{D}^{\mathbb{R}}(\Omega)$.

2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of points in Ω having no cluster point in Ω. Show that the map defined by

$$\langle T, \varphi\rangle = \sum_{n=0}^{\infty} (D^{p_n}\varphi)(x_n),$$

where each p_n is a multiindex, is a distribution. Compute its order.

3. Show that, for every function $\varphi \in \mathscr{D}(\mathbb{R})$, the limit as ε tends to 0 of

$$\int_{\{|x|\geq\varepsilon\}} \frac{\varphi(x)}{x^2}\,dx - 2\,\frac{\varphi(0)}{\varepsilon}$$

exists, and that this defines a distribution (the finite part of $1/x^2$). Determine its order.

4. Take $f \in C((\mathbb{R}^d)^*)$.

a. Assume there exists a constant $C > 0$ and an integer $n > 0$ such that, for every $x \in B(0,1)\setminus\{0\}$,

$$|f(x)| \leq \frac{C}{|x|^n}\,.$$

Show that f extends to a distribution of order at most n on $\mathbb{R}^d$.
Hint. Consider, for $\varphi \in \mathscr{D}(\mathbb{R}^d)$,

$$\langle T, \varphi\rangle = \int_{\{|x|\geq 1\}} f(x)\varphi(x)\,dx + \int_{\{|x|\leq 1\}} f(x)\big(\varphi(x) - P_n(x)\big)\,dx,$$

where P_n is the sum of the terms of order at most $n-1$ in the Taylor series expansion of φ at 0.

b. Suppose that f is positive and that

$$\lim_{x\to 0} |x|^n f(x) = +\infty \quad \text{for all } n \in \mathbb{N}.$$

Show that there is no distribution on $\mathbb{R}^d$ whose restriction to $(\mathbb{R}^d)^*$ is f.
Hint. Take $\varphi \in \mathscr{D}(\mathbb{R}^d)$ supported in $B(0,4) \setminus \bar{B}(0,1)$ and such that $\varphi = 1$ on $\bar{B}(0,3) \setminus B(0,2)$. For $n \geq 1$, set

$$\lambda_n = \inf_{x \in \bar{B}(0,3/n) \setminus B(0,2/n)} f(x) \qquad \text{and} \qquad \varphi_n(x) = \frac{n^{d+1}}{\lambda_n} \varphi(nx).$$

Show that $(\varphi_n)_{n \in \mathbb{N}^*}$ tends to 0 in $\mathscr{D}(\mathbb{R}^d)$ and that

$$\lim_{n \to +\infty} \int f(x) \varphi_n(x) \, dx = +\infty.$$

5. Let T be a distribution on Ω such that every point in Ω has an open neighborhood on which the restriction of T vanishes. Show that $T = 0$.
Hint. Take $\varphi \in \mathscr{D}(\Omega)$. Cover the support of φ with finitely many sets on which T vanishes; then use a C^∞ partition of unity (or Exercise 11 on page 266).

6. *Piecing distributions together.* Let $\Omega_1, \ldots, \Omega_n$ be open sets in $\mathbb{R}^d$ whose union is Ω. For each $j \in \{1, \ldots, n\}$, let T_j be a distribution on Ω_j. Suppose that, for every pair of integers $(i, j) \in \{1, \ldots, n\}^2$, the distributions T_i and T_j coincide on the open set $\Omega_i \cap \Omega_j$. We wish to show that there is a unique distribution T on Ω whose restriction to each Ω_j is T_j.

a. Using Exercise 5, prove that such a distribution T must be unique.

b. For each $j \in \{1, \ldots, n\}$, take $\varphi_j \in \mathscr{D}(\Omega_j)$. Show that $\sum_{j=1}^n \varphi_j = 0$ implies $\sum_{j=1}^n \langle T_j, \varphi_j \rangle = 0$.
Hint. Use a C^∞ partition of unity associated with the open sets Ω_j, $1 \leq j \leq n$, and with the compact $K = \bigcup_{j=1}^n \operatorname{Supp} \varphi_j$.

c. Take $\varphi \in \mathscr{D}(\Omega)$. Show that the expression

$$\langle T, \varphi \rangle = \sum_{j=1}^n \langle T_j, \varphi_j \rangle$$

is independent of the choice of a family $\varphi_1, \ldots, \varphi_n$ such that

$$\varphi_j \in \mathscr{D}(\Omega_j) \ \text{ for all } j \in \{1, \ldots, n\} \qquad \text{and} \qquad \varphi = \sum_{j=1}^n \varphi_j.$$

(The existence of such a family follows from Proposition 1.4; see Exercise 11 on page 266.)

d. Show that the map T defined above is a distribution on Ω having the desired properties.

e. Show that, if each distribution T_j has order at most m, so does T.

7. Let Ω be open in $\mathbb{R}^d$ and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^1_{\text{loc}}(\Omega)$. Show that, if the sequence (f_n) converges in $L^1_{\text{loc}}(\Omega)$ to an element f in $L^1_{\text{loc}}(\Omega)$, then f_n tends to f in $\mathscr{D}'(\Omega)$. (Convergence in $L^1_{\text{loc}}(\Omega)$ is defined in Exercise 19 on page 159.)
8. Compute the limit of the sequence of distributions in $\mathbb{R}^d$ defined by the functions $T_n(x) = n^d\chi(nx)$, where $\chi \in L^1(\mathbb{R}^d)$.
9. Compute the limit of the sequences of distributions on $\mathbb{R}$ defined by the following functions:

   a. $T_n(x) = \sin nx$;
   b. $T_n(x) = (\sin nx)/x$;
   c. $T_n(x) = n \sin(nx)\, 1_{\{x\geq 0\}}$;
   d. $T_n(x) = |x|^{\frac{1}{n}-1}/(2n)$.

   Hint. For parts a, b, and c, you might use the Riemann–Lebesgue Lemma (Exercise 4a on page 179) and/or integration by parts.
10. Study the convergence in $\mathscr{D}'(\mathbb{R}^*)$, then in $\mathscr{D}'(\mathbb{R})$, of the sequence of distributions

$$T_n = \sum_{k=1}^{n} a_k\big(\delta_{1/k} - \delta_{-1/k}\big),$$

   where (a_n) is a sequence of complex numbers.
11. Show that the equation

$$\langle T, \varphi\rangle = \sum_{n=1}^{+\infty} \varphi^{(n)}(1/n) \quad \text{for all } \varphi \in \mathscr{D}((0, +\infty))$$

   defines a distribution T on $(0, +\infty)$ of infinite order, and that T cannot be extended to $\mathbb{R}$.
12. Find the limit in $\mathscr{D}'(\mathbb{R}^d)$ as ε tends to 0 of the family (T_ε) defined by

$$T_\varepsilon(x) = \frac{1}{\omega_d\varepsilon^d}\, 1_{\{|x|\leq\varepsilon\}}(x),$$

   where ω_d is the volume of the unit ball in $\mathbb{R}^d$.
13. For $x \in \mathbb{R}$ and $N \in \mathbb{N}$, write

$$S_N(x) = \sum_{n=-N}^{N} e^{inx} = \frac{\sin\big((N+\frac{1}{2})x\big)}{\sin(x/2)}.$$

   a. Take $\varphi \in \mathscr{D}(\mathbb{R})$. Show that, for every $p \in \mathbb{Z}$,

$$\lim_{N\to+\infty} \int_{(2p-1)\pi}^{(2p+1)\pi} S_N(x)\varphi(x)\, dx = 2\pi\varphi(2p\pi).$$

Hint. Show that

$$\int_{(2p-1)\pi}^{(2p+1)\pi} S_N(x)\varphi(x)\,dx$$
$$= \int_{-\pi}^{\pi} S_N(x)\big(\varphi(x+2p\pi)-\varphi(2p\pi)\big)\,dx \;+\; 2\pi\varphi(2p\pi);$$

then apply the Riemann–Lebesgue Lemma (Exercise 4a on page 179).

b. Deduce that the sequence of distributions $([S_N])_{N\in\mathbb{N}}$ converges in $\mathscr{D}'(\mathbb{R})$ to $2\pi\sum_{p\in\mathbb{Z}}\delta_{2p\pi}$, where $\delta_{2p\pi}$ is the Dirac measure at the point $2p\pi$.

Remark. One can show that the sequence $([S_N])_{N\in\mathbb{N}}$ (considered as a sequence of Radon measures on $\mathbb{R}$) does not converge vaguely (this concept is defined in Exercise 6 on page 91). Compare with Exercise 1 on page 284.

14. Let $(c_n)_{n\in\mathbb{Z}}$ be a family in $\mathbb{C}$ such that there exist $C\geq 0$ and $\gamma\geq 0$ satisfying

$$|c_n|\leq C\,|n|^{\gamma}\quad\text{for all } n\in\mathbb{Z}^*.$$

Show that the series $\sum_{n\in\mathbb{Z}} c_n\,[e^{inx}]$ converges in $\mathscr{D}'(\mathbb{R})$ and that the sum has finite order.

Hint. If $\varphi\in\mathscr{D}(\mathbb{R})$ with $\operatorname{Supp}\varphi\subset[-A,A]$ (where $A>0$), prove using integration by parts that, for every $r\in\mathbb{N}$ and $n\in\mathbb{Z}^*$,

$$\left|\int e^{inx}\,\varphi(x)\,dx\right|\leq 2A\,\|\varphi\|^{(r)}\,|n|^{-r}.$$

15. Let (f_n) be a sequence of functions in $L^2(\Omega)$ and suppose $f\in L^2(\Omega)$.

a. Show that, if the sequence $(f_n)_{n\in\mathbb{N}}$ converges weakly to f in the Hilbert space $L^2(\Omega)$, it converges to f in $\mathscr{D}'(\Omega)$.

b. Is the converse true? (You might consider, for instance, the open set $\Omega=(0,1)$ and the functions $f_n=n\,1_{[1/n,2/n]}$.)

c. Show that $(f_n)_{n\in\mathbb{N}}$ converges weakly to f in $L^2(\Omega)$ if and only if it is bounded in $L^2(\Omega)$ and converges to f in $\mathscr{D}'(\Omega)$.

Recall that every weakly convergent sequence is bounded in $L^2(\Omega)$; see Exercise 10a on page 120 (this follows from Baire's Theorem).

Hint. Show first that $\mathscr{D}(\Omega)$ is dense in $L^2(\Omega)$.

16. *Banach–Steinhaus Theorem in $\mathscr{D}'$.* Let Ω be an open set in $\mathbb{R}^d$. Let (T_n) be a sequence of distributions on Ω such that, for every $\varphi\in\mathscr{D}(\Omega)$, the sequence of numbers $(\langle T_n,\varphi\rangle)$ is bounded. We wish to show that, for every compact K contained in Ω, there exists an $m\in\mathbb{N}$ and a real constant $C>0$ such that

$$\big|\langle T_n,\varphi\rangle\big|\leq C\,\|\varphi\|^{(m)}\quad\text{for all }\varphi\in\mathscr{D}_K(\Omega)\ \text{ and }\ n\in\mathbb{N}.\qquad(*)$$

To do this we define, for every $k\in\mathbb{N}$, a set

$$F_k=\{\varphi\in\mathscr{D}_K(\Omega):|\langle T_n,\varphi\rangle|\leq k\text{ for all }n\in\mathbb{N}\}.$$

a. Show that each F_k is closed in the metric space $\mathscr{D}_K(\Omega)$ defined in Exercise 7 on page 265. Deduce that, for at least one $k_0 \in \mathbb{N}$, the set F_{k_0} has nonempty interior. (Use Baire's Theorem, Exercise 6 on page 22.)

b. Show that each F_k is convex and symmetric with respect to 0 and deduce that there exists $r > 0$ such that the ball $B(0, r)$ (in $\mathscr{D}_K(\Omega)$) is contained in F_{k_0}.

c. Let $m \in \mathbb{N}$ be such that $\sum_{n>m} 2^{-n} \le r/2$, and set $C = 4k_0/r$. Show that m and C satisfy condition $(*)$.

17. Let (T_n) be a sequence of distributions on Ω such that, for every $\varphi \in \mathscr{D}(\Omega)$, the sequence of numbers $(\langle T_n, \varphi\rangle)$ converges. Show that, for any $t \in (a, b)$, the linear form T on $\mathscr{D}(\Omega)$ defined by

$$\langle T, \varphi\rangle = \lim_{n\to+\infty} \langle T_n, \varphi\rangle$$

is a distribution on Ω.

Is it true that, if all the distributions T_n have order at most m, then so does T?

Hint. Use Exercise 16.

18. Let $(T_t)_{t\in(a,b)}$ be a family of distributions on Ω. Suppose that, for every $\varphi \in \mathscr{D}(\Omega)$, the function $t \mapsto \langle T_t, \varphi\rangle$ is differentiable on (a, b). Show that, for any $t \in (a, b)$, the linear form dT_t/dt defined by

$$\left\langle \frac{dT_t}{dt}, \varphi \right\rangle = \frac{d}{dt} \langle T_t, \varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega)$$

is a distribution on Ω.

Hint. Use Exercise 17.

19. *Finite part of $Y(x)/x^\alpha$, for $\alpha \in \mathbb{R}^+$*

a. Take $m \in \mathbb{N}^*$. Prove that, for every $\varphi \in \mathscr{D}(\mathbb{R})$, the limit

$$\langle T, \varphi\rangle = \lim_{\varepsilon\to 0+} \left(\int_\varepsilon^{+\infty} \frac{\varphi(x)}{x^m}\, dx - \sum_{k=0}^{m-2} \frac{\varphi^{(k)}(0)}{(m-k-1)k!} \frac{1}{\varepsilon^{m-k-1}} + \frac{\varphi^{(m-1)}(0)}{(m-1)!} \log \varepsilon \right)$$

exists and that the linear form T thus defined is a distribution of order (at most) m on $\mathbb{R}$. This distribution is called the *finite part of* $Y(x)/x^m$, and is denoted $\operatorname{fp}(Y(x)/x^m)$.

b. Take $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$. Let m be the integer such that $m < \alpha < m+1$. Show that, for every $\varphi \in \mathscr{D}(\mathbb{R})$, the limit

$$\langle T, \varphi\rangle = \lim_{\varepsilon\to 0+} \left(\int_\varepsilon^{+\infty} \frac{\varphi(x)}{x^\alpha}\, dx - \sum_{k=0}^{m-1} \frac{\varphi^{(k)}(0)}{(\alpha-k-1)k!} \frac{1}{\varepsilon^{\alpha-k-1}} \right)$$

exists and that the linear form T thus defined is a distribution of order (at most) m on $\mathbb{R}$. This distribution is called the *finite part of* $Y(x)/x^\alpha$, and is denoted $\mathrm{fp}(Y(x)/x^\alpha)$.

20. Complete the proof of Proposition 2.6.

3 Complements

In this section, we study under what conditions a distribution can be extended to test function spaces larger than $\mathscr{D}(\Omega)$, namely $\mathscr{D}^m(\Omega)$ or $\mathscr{E}(\Omega)$. We will introduce to this effect the important notion of the support of a distribution.

3A Distributions of Finite Order

The next proposition provides a characterization of distributions of finite order.

Proposition 3.1 *Let T be a distribution on Ω and suppose $m \in \mathbb{N}$. A necessary and sufficient condition for T to have order at most m is that T can be extended to a continuous linear form on $\mathscr{D}^m(\Omega)$. The extension is then unique.*

Proof. Suppose that T has order at most m. Property $(*)$ on page 268 then implies that T is continuous (and even uniformly continuous) on the space $\mathscr{D}(\Omega)$ regarded, topologically speaking, as a subspace of $\mathscr{D}^m(\Omega)$. Since $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}^m(\Omega)$ by Corollary 1.3, we can apply the theorem of extension of continuous linear forms. This theorem applies a priori to continuous linear forms on normed spaces, but we can reduce the problem to that situation by considering the normed spaces $\mathscr{D}^m_K(\Omega)$. Similarly, since $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}^m(\Omega)$, this extension is unique.

In the other direction, it is clear from the definitions that the restriction of a continuous linear form on $\mathscr{D}^m(\Omega)$ to $\mathscr{D}(\Omega)$ is a distribution of order at most m. □

Conversely, the restriction to $\mathscr{D}(\Omega)$ of a continuous linear form on $\mathscr{D}^m(\Omega)$ is a distribution (since a sequence in $\mathscr{D}(\Omega)$ that converges in $\mathscr{D}(\Omega)$ obviously converges in $\mathscr{D}^m(\Omega)$), and it has order at most m by the preceding reasoning. Thus we can identify the space of distributions of order at most m on Ω with the space of continuous linear forms on $\mathscr{D}^m(\Omega)$, which we denote by $\mathscr{D}'^m(\Omega)$. We will make this identification from now on, and for $T \in \mathscr{D}'^m(\Omega)$ and $\varphi \in \mathscr{D}^m(\Omega)$ we will still denote by $\langle T, \varphi\rangle$ the result of evaluating T at φ.

An important particular case, already discussed on page 270, is when $m = 0$. Then $\mathscr{D}^0(\Omega) = C_c(\Omega)$ and the space $\mathscr{D}'^0(\Omega)$ of distributions of order 0 can be identified with the space $\mathfrak{M}(\Omega)$ of complex Radon measures on Ω. Consequently, $\mathfrak{M}(\Omega) = \mathscr{D}'^0(\Omega) \subset \mathscr{D}'(\Omega)$ and, if $\varphi \in C_c(\Omega)$ and $\mu \in \mathfrak{M}(\Omega)$, we have

$$\langle \mu, \varphi \rangle = \mu(\varphi) = \int \varphi \, d\mu.$$

3B The Support of a Distribution

Let T be a distribution on Ω. By definition, a **domain of nullity** of T is an open set Ω' contained in Ω and such that the restriction of T to Ω' is the zero distribution on Ω'.

Proposition 3.2 *Any distribution T on Ω has a largest domain of nullity Ω_0.*

The complement of this set, $\Omega \setminus \Omega_0$, which is closed in Ω, is called the **support** of T and is denoted by $\operatorname{Supp} T$.

Proof. Let $\mathscr{U}$ be the set of domains of nullity of T, and let $\Omega_0 = \bigcup_{O \in \mathscr{U}} O$ be their union. It suffices to show that Ω_0 is itself a domain of nullity of T. Take $\varphi \in \mathscr{D}(\Omega_0) \subset \mathscr{D}(\Omega)$. By the compactness of $\operatorname{Supp} \varphi$, there exist finitely many elements $\omega_1, \dots, \omega_n$ of $\mathscr{U}$ whose union contains $\operatorname{Supp} \varphi$. By Proposition 1.4, there exists a C^∞ partition of unity associated with this open cover; that is, there exist functions $\varphi_1, \dots, \varphi_n \in \mathscr{D}$ such that $0 \le \varphi_j \le 1$ and $\operatorname{Supp} \varphi_j \subset \omega_j$ for every $j \in \{1, \dots, n\}$ and such that $\sum_{j=1}^n \varphi_j(x) = 1$ for every $x \in \operatorname{Supp} \varphi$. It follows that

$$\varphi = \sum_{j=1}^n \varphi \varphi_j .$$

Since each $\varphi \varphi_j$ is supported in the domain of nullity ω_j, this implies that

$$\langle T, \varphi \rangle = \sum_{j=1}^n \langle T, \varphi \varphi_j \rangle = 0.$$

This proves that Ω_0 is indeed a domain of nullity of T, and by the construction it is the largest such domain. □

The support of a complex Radon measure μ on Ω was defined in Exercise 2 on page 90. Since, for every open set O, the space $\mathscr{D}(O)$ is dense in $C_c(O)$ (Corollary 1.3), one can check easily that this definition coincides with the one just given for distributions.

3C Distributions with Compact Support

The next proposition characterizes distributions having compact support.

Proposition 3.3 *Let T be a distribution on Ω. A necessary and sufficient condition for the support of T to be compact is that T have an extension to a continuous linear form on $\mathscr{E}(\Omega)$. The extension is then unique.*

Proof. Suppose first that the support of T is compact. Then there exists a compact K in Ω whose interior contains the support of T. It follows from Proposition 1.4 that there exists $\rho \in \mathscr{D}(\Omega)$ such that $0 \leq \rho \leq 1$ and $\rho(x) = 1$ for all $x \in K$. We then set, for $f \in \mathscr{E}(\Omega)$,

$$\hat{T}(f) = \langle T, f\rho \rangle .$$

It is clear that this does define a linear form $\hat{T}$ on $\mathscr{E}(\Omega)$. On the other hand, if $\varphi \in \mathscr{D}(\Omega)$, we have

$$\operatorname{Supp}(\varphi - \varphi\rho) \subset \Omega \backslash \overset{\circ}{K} \subset \Omega \backslash \operatorname{Supp} T,$$

so that

$$\langle T, \varphi \rangle = \langle T, \varphi\rho \rangle .$$

It follows that $\hat{T}$ is an extension of T to $\mathscr{E}(\Omega)$.

Finally, if (f_n) is a sequence in $\mathscr{E}(\Omega)$ that tends to 0 in $\mathscr{E}(\Omega)$, it is easy to see from the definitions and from Leibniz's formula that the sequence $(f_n\rho)$ tends to 0 in $\mathscr{D}(\Omega)$, so that

$$\lim_{n \to +\infty} \langle T, f_n\rho \rangle = 0.$$

This proves that $\hat{T}$ is continuous on $\mathscr{E}(\Omega)$.

Thus T has an extension $\hat{T}$ that is a continuous linear form on $\mathscr{E}(\Omega)$. Since $\mathscr{D}(\Omega)$ is dense in $\mathscr{E}(\Omega)$, this extension is unique.

For the converse, assume that T can be extended to a continuous linear form $\hat{T}$ on $\mathscr{E}(\Omega)$. Let $(K_n)_{n \in \mathbb{N}}$ be an exhausting sequence of compact subsets of Ω. If the support of T is not compact, there exists, for every integer $n \in \mathbb{N}$, an element φ_n of $\mathscr{D}(\Omega)$ such that

$$\operatorname{Supp} \varphi_n \subset (\Omega \setminus K_n) \quad \text{and} \quad \langle T, \varphi_n \rangle \neq 0$$

(by the definition of the support of T). Dividing φ_n by $\langle T, \varphi_n \rangle$, if necessary, we can assume that

$$\langle T, \varphi_n \rangle = 1.$$

Now, we claim that the series $\sum_{n=0}^{+\infty} \varphi_n$ converges in $\mathscr{E}(\Omega)$. Indeed, if K is a compact subset of Ω, then K is contained in some K_{n_0}, for $n_0 \in \mathbb{N}$; but, for every $n > n_0$, we have $\varphi_n = 0$ on K_{n_0} and so on K, so the sum $\sum_{n=0}^{+\infty} \varphi_n$ reduces to a finite sum on K, and this for every compact subset K of Ω. So the sum converges in $\mathscr{E}(\Omega)$. By the continuity of T, it follows that the series $\sum_{n=0}^{+\infty} \langle T, \varphi_n \rangle$ converges, contradicting our assumption that $\langle T, \varphi_n \rangle = 1$. $\square$

Conversely, the restriction to $\mathscr{D}(\Omega)$ of a continuous linear form on $\mathscr{E}(\Omega)$ is a distribution on Ω (since a sequence in $\mathscr{D}(\Omega)$ that converges in $\mathscr{D}(\Omega)$ also converges in $\mathscr{E}(\Omega)$), and this distribution has compact support by the preceding results. Thus we can identify the space of distributions on Ω having compact support with the space of continuous linear forms on $\mathscr{E}(\Omega)$, denoted by $\mathscr{E}'(\Omega)$. We will make this identification from now on. In particular, for $T \in \mathscr{E}'(\Omega)$ and $\varphi \in \mathscr{E}(\Omega)$, we will still write $T(\varphi)$ as $\langle T, \varphi\rangle$.

We remark also that a distribution on Ω with compact support can be identified with an element of $\mathscr{E}'(\mathbb{R}^d)$ by setting

$$\langle T, \varphi\rangle = \left\langle T, \varphi_{|\Omega}\right\rangle \quad \text{for all } \varphi \in \mathscr{E}(\mathbb{R}^d). \tag{$*$}$$

Indeed, if φ is in $\mathscr{E}(\mathbb{R}^d)$ the restriction $\varphi_{|\Omega}$ of φ to Ω lies in $\mathscr{E}(\Omega)$.

Proposition 3.4 *Every distribution T with compact support in Ω has finite order. More precisely, there exists an integer $m \in \mathbb{N}$ and a constant $C \geq 0$ such that*

$$\left|\langle T, \varphi\rangle\right| \leq C \left\|\varphi\right\|^{(m)} \quad \textit{for all } \varphi \in \mathscr{D}(\Omega).$$

Proof. Let K be the support of T and let K', K'' be compact sets such that

$$K \subset \mathring{K}' \subset K' \subset \mathring{K}'' \subset K'' \subset \Omega.$$

By Proposition 2.1, there exists a constant $C \geq 0$ and an integer $m \in \mathbb{N}$ such that

$$\left|\langle T, \varphi\rangle\right| \leq C \left\|\varphi\right\|^{(m)} \quad \text{for all } \varphi \in \mathscr{D}_{K''}(\Omega).$$

By Proposition 1.4, there exists $\psi \in \mathscr{D}$ such that $0 \leq \psi \leq 1$, $\psi = 1$ on K' and $\operatorname{Supp}\psi \subset \mathring{K}''$. If $\varphi \in \mathscr{D}(\Omega)$, then $\varphi\psi \in \mathscr{D}_{K''}(\Omega)$ and

$$\operatorname{Supp}(\varphi - \varphi\psi) \subset \Omega \backslash \mathring{K}' \subset \Omega \backslash K.$$

Since the compact K is the support of T, it follows that there is a positive constant C' depending only on C, m and ψ, and such that

$$\left|\langle T, \varphi\rangle\right| = \left|\langle T, \varphi\psi\rangle\right| \leq C \left\|\varphi\psi\right\|^{(m)} \leq C' \left\|\varphi\right\|^{(m)},$$

the last inequality being a consequence of Leibniz's formula. □

Remark. One can easily deduce from the preceding results that, if T is a distribution with compact support, there exists an integer $m \in \mathbb{N}$ (any integer not less than the order of T will do) such that T extends to a continuous linear form on $\mathscr{E}^m(\Omega)$, and that this extension is unique.

Exercises

1. Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of positive Radon measures that converges in $\mathscr{D}'(\Omega)$ to a distribution T. Show that T is a positive Radon measure and that the sequence $(T_n)_{n\in\mathbb{N}}$ converges vaguely to T (this term is defined in Exercise 6 on page 91). Compare with Exercise 13 on page 277.
2. Let T be a distribution on $\mathbb{R}^n$ and suppose $\varphi \in \mathscr{D}(\mathbb{R}^n)$ vanishes at every point in the support of T. Does it follow that $\langle T, \varphi\rangle = 0$?
3. Let T be a distribution on Ω with compact support K and order m, and suppose $\varphi \in \mathscr{D}(\Omega)$ satisfies the following property: For every multiindex p of length at most m and every x in the support of T, we have $D^\alpha\varphi(x) = 0$. We wish to prove that $\langle T, \varphi\rangle = 0$. Put $K_\varepsilon = \{x \in \mathbb{R}^d : d(x, K) \le \varepsilon\}$, for $\varepsilon > 0$.
   a. Take $\eta > 0$. By assumption, there exists a real number $r_\eta > 0$ such that $|D^p\varphi(x)| \le \eta$, for every $x \in K_{r_\eta}$ and every multiindex p of length at most m. Show that, for every $x \in K_{r_\eta}$ and every $p \in \mathbb{N}^d$ such that $|p| \le m$,
   $$|D^p\varphi(x)| \le \eta d^{m-|p|} d(x, K)^{m-|p|}.$$
   Hint. You might use reverse induction on $n = |p|$, applying at each step the Mean Value Theorem on the segment $[x, y]$, where y is a point in K such that $d(x, K) = d(x, y)$.
   b. Suppose $\chi \in \mathscr{D}(\mathbb{R}^d)$ has its support contained in $B(0, 1)$ and satisfies $\int \chi(x)\, dx = 1$. Let χ_ε be the element of $\mathscr{D}(\mathbb{R}^d)$ defined by
   $$\chi_\varepsilon(x) = \varepsilon^{-d} \int_{K_{2\varepsilon}} \chi\Big(\frac{x-y}{\varepsilon}\Big)\, dy.$$
   Show the following facts:
      i. For every $\varepsilon > 0$, the support of χ_ε is contained in $K_{3\varepsilon}$; moreover, $\chi_\varepsilon = 1$ in K_ε.
      ii. For every multiindex p we have
      $$\|D^p\chi_\varepsilon\|^{(0)} \le \|\chi\|^{(|p|)}\omega_d\, \varepsilon^{-|p|},$$
      where ω_d is the volume of the unit ball in $\mathbb{R}^d$.
   c. Show that there exists a constant $C > 0$ (depending only on d and m) such that, for every $\varepsilon < r_\eta/3$,
   $$\|\chi_\varepsilon\varphi\|^{(m)} \le \eta C \|\chi\|^{(m)}.$$
   (Use Leibniz's formula.)
   d. Show that there exists a constant $C' > 0$ such that
   $$|\langle T, \varphi\rangle| \le \eta C' \|\chi\|^{(m)}.$$
   Finish the proof.

4. *Structure of distributions with finite support*

   a. Let $f_1, f_2, \ldots, f_n$ and f be linear forms on a vector space E with the property that, for every $x \in E$,

   $$f_1(x) = f_2(x) = \cdots = f_n(x) = 0 \quad \Longrightarrow \quad f(x) = 0.$$

   Show that there exist scalars $c_1, \ldots, c_n$ such that $f = c_1 f_1 + \cdots + c_n f_n$.
   Hint. Let E, F, and G be vector spaces, f a linear map from E to G, and g a linear map from E to F such that $\ker g \subseteq \ker f$. Suppose F is finite-dimensional. Then there exists a linear map h from F to G such that $f = h \circ g$. (Why?) Apply this result to $F = \mathbb{C}^n$ and $g = (f_1, \ldots, f_n)$.

   b. Let T be a distribution on an open subset Ω of $\mathbb{R}^d$, and suppose $\operatorname{Supp} T = \{0\}$. Show that T is given by

   $$\langle T, \varphi \rangle = \sum_{|p| \le m} c_p D^p \varphi(0) \quad \text{for all } \varphi \in \mathscr{D}(\Omega),$$

   for appropriate constants c_p.
   Hint. Use Exercise 3.

   c. Determine likewise the general form of a distribution whose support is finite.

5. Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of distributions on an open subset Ω of $\mathbb{R}^d$. We assume that the sequence (T_n) converges in $\mathscr{D}'(\Omega)$ and that the supports of the distributions T_n are all contained in the same compact K. Show that the orders of distributions T_n have a uniform upper bound $m \in \mathbb{N}$.
   Hint. Use the Banach–Steinhaus Theorem in $\mathscr{D}'(\Omega)$, stated in Exercise 16 on page 278.

8

Multiplication and Differentiation

We define in this chapter two important operations involving distributions. Again we will be working with an open subset Ω of $\mathbb{R}^d$.

1 Multiplication

In this section we define the product of a distribution by a smooth function. This definition arises from the following lemma.

Lemma 1.1 *Suppose $\alpha \in \mathscr{E}(\Omega)$. The map $\varphi \mapsto \alpha\varphi$ from $\mathscr{D}(\Omega)$ to $\mathscr{D}(\Omega)$ is continuous. Likewise, if $\alpha \in \mathscr{E}^m(\Omega)$, with $m \in \mathbb{N}$, the map $\varphi \mapsto \alpha\varphi$ from $\mathscr{D}^m(\Omega)$ to $\mathscr{D}^m(\Omega)$ is continuous.*

In other words, if $(\varphi_n)_{n\in\mathbb{N}}$ is a sequence in $\mathscr{D}(\Omega)$ or $\mathscr{D}^m(\Omega)$ converging to 0 in $\mathscr{D}(\Omega)$ or $\mathscr{D}^m(\Omega)$, respectively, the same is true about the sequence $(\alpha\varphi_n)_{n\in\mathbb{N}}$.

Proof. The lemma follows immediately from Leibniz's formula (page 258) and from the fact that, if $\varphi_n \in \mathscr{D}(\Omega)$, the support of $\alpha\varphi_n$ is contained in the support of φ_n. $\square$

Thus we can define the product of a function and a distribution as follows:

Definition 1.2 If $T \in \mathscr{D}'(\Omega)$ and $\alpha \in \mathscr{E}(\Omega)$, the product distribution αT on Ω is defined by setting

$$\langle \alpha T, \varphi\rangle = \langle T, \alpha\varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

If $T \in \mathscr{D}'^m(\Omega)$ and $\alpha \in \mathscr{E}^m(\Omega)$, the product $\alpha T \in \mathscr{D}'^m(\Omega)$ is defined by

$$\langle \alpha T, \varphi \rangle = \langle T, \alpha\varphi \rangle \quad \text{for all } \varphi \in \mathscr{D}^m(\Omega).$$

(Recall that $\mathscr{D}'^m(\Omega)$ is the set of continuous linear forms on the space $\mathscr{D}^m(\Omega)$, which by Proposition 3.3 on page 282 can be identified with the space of distributions of order at most m).

That αT really is a distribution follows from the preceding lemma: If $(\varphi_n)_{n\in\mathbb{N}}$ is a sequence in $\mathscr{D}(\Omega)$ or $\mathscr{D}^m(\Omega)$ that converges to 0, Lemma 1.1 implies that the sequence $(\langle T, \alpha\varphi_n\rangle)_{n\in\mathbb{N}}$ tends to 0 since T is a distribution. Thus αT really is a continuous linear form on $\mathscr{D}(\Omega)$ or $\mathscr{D}^m(\Omega)$, as the case may be.

Obviously, if $f \in L^1_{\text{loc}}(\Omega)$ and $\alpha \in C(\Omega)$, we have

$$\alpha[f] = [\alpha f].$$

In this sense, this multiplication extends the usual product of functions. We will see in Exercise 1 below that this extension cannot be pushed further to the case of the product of two arbitrary distributions without the loss of the elementary algebraic properties of multiplication, such as associativity and commutativity.

Remark. The definition immediately implies that if $\alpha \in \mathscr{E}(\Omega)$ the linear map $T \mapsto \alpha T$ from $\mathscr{D}'(\Omega)$ to $\mathscr{D}'(\Omega)$ is continuous, in the sense that, if $(T_n)_{n\in\mathbb{N}}$ converges to T in $\mathscr{D}'(\Omega)$, then $(\alpha T_n)_{n\in\mathbb{N}}$ converges to αT in $\mathscr{D}'(\Omega)$.

Proposition 1.3 *With the notation introduced in Definition* 1.2, *we have*

$$\operatorname{Supp}(\alpha T) \subset \operatorname{Supp}\alpha \cap \operatorname{Supp} T$$

and, if $\beta \in \mathscr{E}(\Omega)$ (or $\beta \in \mathscr{E}^m(\Omega)$), we have

$$\alpha(\beta T) = (\alpha\beta)T.$$

Proof. The second claim is obvious. To show the first, take $\varphi \in \mathscr{D}(\Omega)$. If $\operatorname{Supp}\varphi \subset \Omega \setminus \operatorname{Supp}\alpha$, then $\alpha\varphi = 0$, so $\langle \alpha T, \varphi \rangle = 0$. It follows that $\Omega \setminus \operatorname{Supp}\alpha$ is contained in $\Omega \setminus \operatorname{Supp}(\alpha T)$, so $\operatorname{Supp}(\alpha T) \subset \operatorname{Supp}\alpha$.

Now if $\operatorname{Supp}\varphi \subset \Omega \setminus \operatorname{Supp} T$, then

$$\operatorname{Supp}\alpha\varphi \subset \operatorname{Supp}\varphi \subset \Omega \setminus \operatorname{Supp} T,$$

which implies that $\langle \alpha T, \varphi \rangle = 0$. Therefore $\Omega \setminus \operatorname{Supp} T$ is contained in $\Omega \setminus \operatorname{Supp}(\alpha T)$, so $\operatorname{Supp}(\alpha T) \subset \operatorname{Supp} T$. The result follows. □

The inclusion in the proposition may be strict. For example, if $T = \delta$ is the Dirac measure at 0 in $\mathbb{R}^d$, and if $\alpha \in C(\mathbb{R}^d)$ is such that $\alpha(0) = 0$ and $0 \in \operatorname{Supp}\alpha$ (say $\alpha(x) = x$), then $\alpha T = \alpha(0)\delta = 0$ and the support of αT is empty, whereas $\operatorname{Supp}\alpha \cap \operatorname{Supp} T = \{0\}$.

Division of distributions is an important problem: If $S \in \mathscr{D}'(\Omega)$ and $\alpha \in \mathscr{E}(\Omega)$, is there a $T \in \mathscr{D}'(\Omega)$ such that $\alpha T = S$? Clearly, if α vanishes nowhere in Ω, the product $T = (1/\alpha)S$ is the unique solution to the problem, by the second part of Proposition 1.3. In the general case, the restriction of T to the open set $\{x \in \mathbb{R}^d : \alpha(x) \neq 0\}$ is uniquely defined by the same equality, but the global problem may have infinitely many solutions. Here is an example in dimension 1.

Proposition 1.4 *For every $S \in \mathscr{D}'(\mathbb{R})$, there exists $T \in \mathscr{D}'(\mathbb{R})$ such that $xT = S$. If T_0 is such that $xT_0 = S$, the set of solutions of the equation $xT = S$ equals $\{T_0 + C\delta : C \in \mathbb{C}\}$.*

Proof. Take $\chi \in \mathscr{D}(\mathbb{R})$ such that $\chi(0) = 1$. To each $\varphi \in \mathscr{D}(\mathbb{R})$ we associate $\tilde{\varphi}$, defined by

$$\tilde{\varphi}(x) = \int_0^1 \big(\varphi'(tx) - \varphi(0)\chi'(tx)\big)\,dt.$$

One easily checks that $\tilde{\varphi} \in \mathscr{D}(\mathbb{R})$ and that the map $\varphi \mapsto \tilde{\varphi}$ from $\mathscr{D}(\mathbb{R})$ to $\mathscr{D}(\mathbb{R})$ is continuous. Moreover, if $x \in \mathbb{R}^*$, $\tilde{\varphi}(x) = \big(\varphi(x) - \varphi(0)\chi(x)\big)/x$. Now put

$$\langle T, \varphi\rangle = \langle S, \tilde{\varphi}\rangle \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}).$$

Since $\varphi \mapsto \tilde{\varphi}$ is continuous, T belongs to $\mathscr{D}'(\mathbb{R})$; since $\widetilde{x\varphi} = \varphi$, we get $xT = S$.

Now take $T \in \mathscr{D}'(\mathbb{R})$ with $xT = 0$. If $\varphi \in \mathscr{D}(\mathbb{R})$, we have

$$0 = \langle xT, \tilde{\varphi}\rangle = \big\langle T, \varphi - \varphi(0)\chi\big\rangle = \langle T, \varphi\rangle - \langle T, \chi\rangle\,\langle \delta, \varphi\rangle\,.$$

It follows that $T = \langle T, \chi\rangle\,\delta$. □

Here is a particular case.

Proposition 1.5 *Suppose $T \in \mathscr{D}'(\mathbb{R})$. Then $xT = 1$ if and only if there exists $C \in \mathbb{C}$ such that $T = \operatorname{pv}(1/x) + C\delta$.*

Note that, in the equality $xT = 1$, the symbol 1 represents the constant function equal to 1, identified with the distribution $[1]$, which is none other than Lebesgue measure λ.

Proof. By Proposition 1.4, it suffices to show that $x\operatorname{pv}(1/x) = 1$. To do this, take $\varphi \in \mathscr{D}(\mathbb{R})$. By definition,

$$\begin{aligned}\langle x\operatorname{pv}(1/x), \varphi\rangle &= \big\langle \operatorname{pv}(1/x), x\varphi\big\rangle = \lim_{\varepsilon\to 0^+} \int_{\{|x|>\varepsilon\}} (1/x)\,x\,\varphi(x)\,dx \\ &= \int \varphi(x)\,dx = \big\langle [1], \varphi\big\rangle,\end{aligned}$$

as we wished to show. □

Exercises

1. Show that it is impossible to define a multiplication operation on the set $\mathcal{D}'(\mathbb{R})$ that is at once associative, commutative and an extension of the multiplication defined in the text.

Hint. Suppose there is such a multiplication and compute in two ways the product $x\delta\,\mathrm{pv}(1/x)$, where δ is the Dirac measure at 0.

2. Consider an open set Ω in $\mathbb{R}^d$ and elements $\alpha \in \mathcal{E}(\Omega)$ and $T \in \mathcal{D}'(\Omega)$. Assume that $\alpha = 1$ on an open set that contains the support of T. Show that $\alpha T = T$.

3. Suppose $T \in \mathcal{D}'(\mathbb{R}^d)$, $a \in \mathbb{R}^d$, and $m \in \mathbb{N}$. Show that $(x-a)^p T = 0$ for every multiindex p of length $m+1$ if and only if T can be written as

$$\langle T, \varphi\rangle = \sum_{|q|\le m} c_q D^q\varphi(a) \quad \text{for all } \varphi \in \mathcal{D}(\mathbb{R}^d),$$

with $c_q \in \mathbb{C}$ for $|q| \le m$. (As might be expected, by $(x-a)^p$ we mean the product $(x_1-a_1)^{p_1}\dots(x_d-a_d)^{p_d}$.)

Hint. Show first that, if $(x-a)^p T = 0$ for every multiindex p of length $m+1$, the support of T is contained in $\{a\}$ and so is compact. Using Taylor's formula (Exercise 1 on page 264), prove then that, for every $\varphi \in \mathcal{D}(\mathbb{R}^d)$,

$$\langle T, \varphi\rangle = \sum_{k=0}^{m} \frac{1}{k!} \sum_{|q|=k} D^q\varphi(a)\langle T, (x-a)^q\rangle.$$

4. Suppose $S \in \mathcal{D}'(\mathbb{R})$, $a \in \mathbb{R}$, and $m \in \mathbb{N}$.

a. Choose $\chi \in \mathcal{D}(\mathbb{R})$ such that $\chi(a) = 1$ and $\chi^{(k)}(a) = 0$ for $k \in \{1,\dots,m\}$. Given $\varphi \in \mathcal{D}(\mathbb{R})$, define a function $\tilde\varphi$ by

$$\tilde\varphi(x) = \frac{\varphi(x) - \big(\varphi(a) + \varphi'(a)(x-a) + \cdots + (1/m!)\varphi^{(m)}(a)(x-a)^m\big)\chi(x)}{(x-a)^{m+1}}$$

if $x \neq a$, and extend it to $x = a$ by continuity. Show that the map $\varphi \mapsto \tilde\varphi$ from $\mathcal{D}(\mathbb{R})$ to $\mathcal{D}(\mathbb{R})$ is continuous (in the sense of sequences).

b. Show that the equality

$$\langle T, \varphi\rangle = \langle S, \tilde\varphi\rangle$$

defines a distribution on $\mathbb{R}$ that is a solution of the equation

$$(x-a)^{m+1}T = S.$$

c. Determine all solutions of the equation $(x-a)^{m+1}T = S$.

Hint. Use Exercise 3.

5. In each of the following cases, the question is to show the existence in $\mathscr{D}'(\mathbb{R})$ of solutions T for the equation $fT = S$, with $S \in \mathscr{D}'(\mathbb{R})$ arbitrary, and to find the general form of the solutions in terms of a particular solution T_0.

a. Suppose that $f \in \mathscr{E}(\mathbb{R})$ and that f has a unique zero $a \in \mathbb{R}$, which furthermore is of finite order; that is, there exists an integer $m \in \mathbb{N}^*$ such that $f^{(m)}(a) \neq 0$.

Hint. Let m be the smallest integer such that $f^{(m)}(a) \neq 0$. The function g defined by $g(x) = (x-a)^{-m} f(x)$ and extended by continuity to $x = a$ belongs to $\mathscr{E}(\mathbb{R})$ and vanishes nowhere. Then $fT = S$ if and only if $(x-a)^m T = g^{-1} S$. Now apply Exercise 4.

b. Suppose that $f \in \mathscr{E}(\mathbb{R})$, that the set of zeros of f has no cluster point, and that each zero has finite order.

Hint. Let $(O_k)_{k\in\mathbb{N}}$ be a locally finite cover of $\mathbb{R}$ by bounded open sets, each containing at most one zero of f. Write S in the form $S = \sum_{k\in\mathbb{N}} S_k$, where $\operatorname{Supp} S_k \subset O_k$ for each $k \in \mathbb{N}$ (see Exercise 14 on page 267). Solve the equation $fT_k = S_k$ for each k, using the preceding case as inspiration.

6. **a.** Show that the distributions T on $\mathbb{R}$ such that $xT = Y$ are exactly those of the form $T = \operatorname{fp}(Y(x)/x) + C\delta$, for $C \in \mathbb{C}$.

b. More generally, prove that, for every $m \in \mathbb{N}^*$, the distributions T on $\mathbb{R}$ such that $x^m T = Y$ are exactly those of the form $T = \operatorname{fp}(Y(x)/x^m) + \sum_{k=0}^{m-1} c_k \delta^{(k)}$, for $c_k \in \mathbb{C}$ (see Exercise 19 on page 279).

7. **a.** Prove that the equality

$$\langle T_0, \varphi \rangle = \lim_{\varepsilon \to 0^+} \sum_{n\in\mathbb{Z}} \left(\int_{n\pi-\pi/2}^{n\pi-\varepsilon} \frac{\varphi(x)}{\sin x}\, dx + \int_{n\pi+\varepsilon}^{n\pi+\pi/2} \frac{\varphi(x)}{\sin x}\, dx \right)$$

defines a distribution T_0 of order 1 on $\mathbb{R}$. (T_0 is the *principal value* of $1/\sin x$.)

b. Show that $\sin x \; T_0 = 1$ and deduce the general form of the solutions of the equation $\sin x \; T = 1$.

8. Suppose $T \in \mathscr{D}'(\Omega)$ and $\varphi \in \mathscr{D}(\Omega)$ are such that, for every multiindex p (of length equal to at most the order of T if T has finite order) and for every x in the support of T, we have $D^p\varphi(x) = 0$. Show that $\langle T, \varphi \rangle = 0$.
Hint. Apply Exercise 3 on page 284 to the distribution $S = \chi T$, where χ is a test function that has the value 1 on an open set containing the support of φ.

9. Take $T \in \mathscr{D}'^m(\Omega)$ and $\alpha \in \mathscr{E}^m(\Omega)$ (or $T \in \mathscr{D}'(\Omega)$ and $\alpha \in \mathscr{E}(\Omega)$). Suppose that, for any $p \in \mathbb{N}^d$ such that $|p| \leq m$ (or any $p \in \mathbb{N}^d$, respectively), $D^p\alpha$ vanishes on the support of T. Show that $\alpha T = 0$. (Use Exercise 8).

10. Let $(K_n)_{n\in\mathbb{N}}$ be an exhausting sequence of compact subsets of Ω. For each $n \in \mathbb{N}$, let $\varphi_n \in \mathscr{D}(\Omega)$ be such that $0 \leq \varphi_n \leq 1$, $\varphi_n = 1$ on K_n, and

$\operatorname{Supp}\varphi_n \subset \mathring{K}_{n+1}$. Show that $\lim_{n\to+\infty}\varphi_n T = T$ in $\mathscr{D}'(\Omega)$, for every $T \in \mathscr{D}'(\Omega)$. Deduce that $\mathscr{E}'(\Omega)$ is dense in $\mathscr{D}'(\Omega)$ (in the sense of convergence of sequences).

2 Differentiation

For $p \in \mathbb{N}^d$, the differentiation operator of order p on $\mathscr{D}'(\Omega)$ is defined as follows: If $T \in \mathscr{D}(\Omega)$, set

$$\langle D^p T, \varphi\rangle = (-1)^{|p|}\langle T, D^p\varphi\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

Since the map $D^p : \varphi \mapsto D^p\varphi$ from $\mathscr{D}(\Omega)$ to $\mathscr{D}(\Omega)$ is continuous, the linear form D^pT thus defined on $\mathscr{D}(\Omega)$ is indeed a distribution. This map D^p is also continuous as a map from $\mathscr{D}^{m+|p|}(\Omega)$ to $\mathscr{D}^m(\Omega)$, which leads to the following property:

Proposition 2.1 *Suppose $m \in \mathbb{N}$. For every $T \in \mathscr{D}'^m(\Omega)$, we have $D^pT \in \mathscr{D}'^{m+|p|}(\Omega)$ and*

$$\langle D^p T, \varphi\rangle = (-1)^{|p|}\langle T, D^p\varphi\rangle \quad \textit{for all } \varphi \in \mathscr{D}^{m+|p|}(\Omega).$$

We also use the notation

$$D^pT = \frac{\partial^{|p|}T}{\partial x^p} = \partial_x^p T,$$

or, if $d = 1$,

$$DT = T' = \frac{dT}{dx}, \qquad D^mT = T^{(m)} = \frac{d^mT}{dx^m} \quad \text{for } m \in \mathbb{N},$$

as for functions. Indeed, the differentiation operator defined above on $\mathscr{D}(\Omega)$ extends ordinary differentiation of functions of class C^1:

Proposition 2.2 *Let $m \in \mathbb{N}$ and $p \in \mathbb{N}^d$ satisfy $|p| \le m$. If $f \in \mathscr{E}^m(\Omega)$, then*

$$D^p([f]) = [D^pf].$$

In this equality, the first D^p denotes **differentiation in the sense of distributions** as defined above, and the second denotes **ordinary differentiation** in the sense of functions.

The proposition is easily obtained by induction on $|p|$ starting from the case $|p| = 1$, which is a consequence of the following lemma.

Lemma 2.3 (Integration by parts) *If $f \in \mathscr{E}^1(\Omega)$ and $\varphi \in \mathscr{D}^1(\Omega)$, then, for every $j \in \{1, \dots, d\}$,*

$$\int_\Omega D_jf\,\varphi\,dx = -\int_\Omega f\,D_j\varphi\,dx.$$

Proof. By Fubini's Theorem, we can reduce to the case $d = 1$, to which we apply the classical theorem of integration by parts, taking into account that the support of φ is a compact subset of Ω, so that the "boundary" terms vanish. □

Examples

1. Take $a \in \Omega$. The derivatives of the Dirac measure at a (denoted by δ_a and defined by $\langle \delta_a, \varphi \rangle = \varphi(a)$) are given by

$$\langle D^p \delta_a, \varphi \rangle = (-1)^{|p|} D^p \varphi(a) \quad \text{for all } p \in \mathbb{N}^d;$$

these distributions were studied on page 270. Thus, for every $p \in \mathbb{N}$, the distribution $D^p \delta_a$ has order $|p|$.
In particular, if $a = 0$ (in which case we write $\delta = \delta_0$) and $d = 1$, we have

$$\langle \delta', \varphi \rangle = -\varphi'(0) = -\lim_{h \to 0} \frac{\langle \delta_h, \varphi \rangle - \langle \delta_0, \varphi \rangle}{h} \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

It follows that

$$\delta' = \lim_{h \to 0} -\frac{\delta_h - \delta}{h} \quad \text{in } \mathscr{D}'(\Omega).$$

2. The derivative in the sense of distributions of the Heaviside function Y is the Dirac measure at 0: indeed, if $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\langle Y', \varphi \rangle = -\int_0^{+\infty} \varphi'(t)\, dt = \varphi(0),$$

where we have used, in calculating the integral, the fact that φ has compact support. Therefore $Y' = \delta$.

3. The function $x \mapsto \log(|x|)$ is locally integrable on $\mathbb{R}$ and as such defines a distribution. We compute its derivative in the sense of distributions. If $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\begin{aligned}\left\langle \frac{d}{dx}\left[\log(|x|)\right], \varphi \right\rangle &= -\int \varphi'(x) \log(|x|)\, dx \\ &= -\lim_{\varepsilon \to 0+} \int_{\{|x| \geq \varepsilon\}} \varphi'(x) \log(|x|)\, dx.\end{aligned}$$

Integrating by parts, we deduce that

$$\left\langle \frac{d}{dx}\left[\log(|x|)\right], \varphi \right\rangle = -\lim_{\varepsilon \to 0+} \left(-\varphi(\varepsilon) \log \varepsilon + \varphi(-\varepsilon) \log \varepsilon - \int_{\{|x| \geq \varepsilon\}} \frac{\varphi(x)}{x}\, dx \right).$$

Now, $\log \varepsilon (\varphi(-\varepsilon) - \varphi(\varepsilon))$ tends to 0 as ε tends to 0. Therefore

$$\frac{d}{dx}\left[\log(|x|)\right] = \mathrm{pv}\left(\frac{1}{x}\right).$$

One shows likewise that

$$\frac{d}{dx}[Y\log x] = \operatorname{fp}\Big(\frac{Y(x)}{x}\Big).$$

The next proposition follows easily from the definitions.

Proposition 2.4 *Suppose $p \in \mathbb{N}^d$.*

1. *The application $T \mapsto D^pT$ from $\mathscr{D}'(\Omega)$ to $\mathscr{D}'(\Omega)$ is continuous in the following sense: For every sequence $(T_n)_{n\in\mathbb{N}}$ in $\mathscr{D}'(\Omega)$ that converges to T in $\mathscr{D}'(\Omega)$, the sequence $(D^pT_n)_{n\in\mathbb{N}}$ in $\mathscr{D}'(\Omega)$ converges to D^pT in $\mathscr{D}'(\Omega)$.*
2. *For every $T \in \mathscr{D}'(\Omega)$,*

$$\operatorname{Supp}(D^pT) \subset \operatorname{Supp} T.$$

We remark that the property of continuity extends immediately to families somewhat more general than sequences. For example, we deduce from Example 1 above that

$$\delta'' = \lim_{h\to 0} \frac{\delta' - \delta_h{}'}{h}$$

in $\mathscr{D}'(\mathbb{R})$.

Leibniz's formula also generalizes without change:

Proposition 2.5 (Leibniz's formula) *Consider $T \in \mathscr{D}'(\Omega)$, $\alpha \in \mathscr{E}(\Omega)$, and $p \in \mathbb{N}^d$. Then*

$$D^p(\alpha T) = \sum_{q\le p} \binom{p}{q} D^{p-q}\alpha\, D^qT.$$

This formula remains true for $T \in \mathscr{D}'^m(\Omega)$ and $\alpha \in \mathscr{E}^{m+|p|}(\Omega)$.

Proof. This is obvious if $|p| = 0$. Consider the case $|p| = 1$. If $j \in \{1,\dots,d\}$, we have

$$\langle D_j(\alpha T),\varphi\rangle = -\langle \alpha T, D_j\varphi\rangle = -\langle T, \alpha D_j\varphi\rangle = -\langle T, D_j(\alpha\varphi)\rangle + \langle T, (D_j\alpha)\varphi\rangle,$$

so that

$$\langle D_j(\alpha T),\varphi\rangle = \langle (D_j\alpha)T + \alpha D_jT,\ \varphi\rangle.$$

Thus $D_j(\alpha T) = \alpha D_jT + (D_j\alpha)T$. From here the formula can be extended by induction on $|p|$ as in the case of functions. □

Remark. We will show in Chapter 9 (proposition 2.14 on page 334) that $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}'(\Omega)$ (in the sense of sequences). The preceding proposition then becomes a consequence of Leibniz's formula for functions, together with the denseness result just mentioned and the continuity of the operators of differentiation and of multiplication by a function.

By Proposition 2.2, the first derivatives $D_j([f])$, for $j \in \{1, \dots, d\}$, of a distribution defined by a function $f \in \mathscr{E}^1(\Omega)$ correspond to continuous functions on Ω. We will show that, conversely, any distribution whose first derivatives are defined by continuous functions corresponds to a function in $\mathscr{E}^1(\Omega)$.

Theorem 2.6 *Let $T \in \mathscr{D}'(\Omega)$. Suppose that there exists, for every $j \in \{1, \dots, d\}$, a function $g_j \in C(\Omega)$ such that $D_j T = [g_j]$. Then there exists $f \in \mathscr{E}^1(\Omega)$ such that $T = [f]$.*

Proof

- Suppose first that the result has been proved for the case where Ω is an open parallelepiped in $\mathbb{R}^d$:

$$\Omega = (a, b) = \prod_{j=1}^{d} (a_j, b_j).$$

We derive the general case. Let Ω be any open set in $\mathbb{R}^d$ and let T be a distribution on Ω for which there exists, for every $j \in \{1, \dots, d\}$, a function $g_j \in C(\Omega)$ such that $D_j T = [g_j]$. Let $\mathscr{U}$ be the set of open parallelepipeds contained in Ω. For every $\omega \in \mathscr{U}$, there exists $f_\omega \in \mathscr{E}^1(\omega)$ such that the restriction of T to ω is $[f_\omega]$. It is clear that, for $\omega_1, \omega_2 \in \mathscr{U}$ with $\omega_1 \cap \omega_2 \neq \varnothing$, we have $f_{\omega_1} = f_{\omega_2}$ on $\omega_1 \cap \omega_2$. Thus there exists $f \in \mathscr{E}^1(\Omega)$ such that, for every $\omega \in \mathscr{U}$, the restriction of f to ω is f_ω. It follows that every $\omega \in \mathscr{U}$ is a domain of nullity for $T - [f]$, in the sense of Proposition 3.2 on page 281. By this same proposition, this implies that the support of $T - [f]$ is empty and so that $T = [f]$.
Thus we can assume that we are in the case $\Omega = (a, b)$. We argue by induction on the dimension d.
- Case $d = 1$. Suppose $T \in \mathscr{D}'(\Omega)$ satisfies $T' = [g]$ with $g \in C((a, b))$. Let $\alpha \in (a, b)$. The function G defined by $G(x) = \int_\alpha^x g(t)\, dt$ belongs to $\mathscr{E}^1((a, b))$ and satisfies $[G]' = [g]$. Therefore the distribution $S = T - [G]$ satisfies $S' = 0$. Now let $\chi \in \mathscr{D}((a, b))$ be such that $\int_a^b \chi(x)\, dx = 1$. We define, for each $\varphi \in \mathscr{D}((a, b))$, a function $\tilde{\varphi}$ by setting

$$\tilde{\varphi}(x) = \varphi(x) - \left(\int_a^b \varphi(t)\, dt \right) \chi(x) \quad \text{for all } x \in (a, b).$$

Then $\tilde{\varphi} \in \mathscr{D}((a, b))$ and $\int_a^b \tilde{\varphi}(x)\, dx = 0$. Therefore the function Φ defined on (a, b) by

$$\Phi(x) = \int_a^x \tilde{\varphi}(t)\, dt$$

satisfies $\Phi(x) = 0$ if $x \notin [\min \operatorname{Supp} \tilde{\varphi}, \max \operatorname{Supp} \tilde{\varphi}]$. Thus $\Phi \in \mathscr{D}((a, b))$. Then

$$0 = \langle S', \Phi \rangle = -\langle S, \Phi' \rangle = -\langle S, \tilde{\varphi} \rangle,$$

so that

$$\langle S, \varphi \rangle = \int_a^b \varphi(t)\,dt\ \langle S, \chi \rangle = \big\langle \langle S, \chi \rangle,\ \varphi \big\rangle \quad \text{for all } \varphi \in \mathscr{D}((a,b)).$$

Thus, if we set $f = G + \langle S, \chi \rangle$, we have $f \in \mathscr{E}^1((a,b))$ and $T = [f]$.

- Suppose the result has been proved for $d \geq 1$. For $(a,b) = \prod_{j=1}^{d+1}(a_j, b_j)$ take $T \in \mathscr{D}'((a,b))$ such that, for every $j \in \{1, \ldots, d+1\}$, there exists $g_j \in C((a,b))$ satisfying $D_j T = [g_j]$. Put

$$G(x_1, \ldots, x_{d+1}) = \int_\alpha^{x_{d+1}} g_{d+1}(x_1, \ldots, x_d, t)\,dt,$$

where $\alpha \in (a_{d+1}, b_{d+1})$. Using Fubini's Theorem and integration by parts, one sees that $D_{d+1}[G] = [g_{d+1}]$. The distribution $S = T - [G]$ then satisfies $D_{d+1}S = 0$.
Take $\chi \in \mathscr{D}((a_{d+1}, b_{d+1}))$ such that $\int_{a_{d+1}}^{b_{d+1}} \chi(x)\,dx = 1$. If $\varphi \in \mathscr{D}((a,b))$, define $\tilde{\varphi} \in \mathscr{D}((a,b))$ by

$$\tilde{\varphi}(x_1, \ldots, x_{d+1}) = \varphi(x_1, \ldots, x_{d+1}) - \chi(x_{d+1}) \int_{a_{d+1}}^{b_{d+1}} \varphi(x_1, \ldots, x_d, t)\,dt.$$

Then, for every $(x_1, \ldots, x_d) \in \prod_{j=1}^{d}(a_j, b_j)$,

$$\int_{a_{d+1}}^{b_{d+1}} \tilde{\varphi}(x_1, \ldots, x_d, t)\,dt = 0.$$

Now set

$$\Phi(x_1, \ldots, x_{d+1}) = \int_{a_{d+1}}^{x_{d+1}} \tilde{\varphi}(x_1, \ldots, x_d, t)\,dt.$$

As in the case $d = 1$, we have $\Phi \in \mathscr{D}((a,b))$ and

$$0 = \langle D_{d+1}S, \Phi \rangle = -\langle S, D_{d+1}\Phi \rangle = -\langle S, \tilde{\varphi} \rangle,$$

so that

$$\langle S, \varphi \rangle = \langle S,\ \hat{\varphi} \otimes \chi \rangle,$$

where we have used the notation

$$\hat{\varphi}(x_1, \ldots, x_d) = \int_{a_{d+1}}^{b_{d+1}} \varphi(x_1, \ldots, x_d, t)\,dt$$

and

$$\hat{\varphi} \otimes \chi(x_1, \ldots, x_{d+1}) = \hat{\varphi}(x_1, \ldots, x_d)\chi(x_{d+1}).$$

Consider the distribution $U \in \mathscr{D}'\big(\prod_{j=1}^{d}(a_j, b_j)\big)$ defined by

$$\langle U, \psi \rangle = \langle T,\ \psi \otimes \chi \rangle \quad \text{for all } \psi \in \mathscr{D}\big(\textstyle\prod_{j=1}^{d}(a_j, b_j)\big).$$

It is clear that the linear form U is indeed a distribution and that, if $j \in \{1,\dots,d\}$,

$$\begin{aligned}\langle D_jU, \psi\rangle &= -\big\langle T, (D_j\psi)\otimes\chi\big\rangle = -\big\langle T, D_j(\psi\otimes\chi)\big\rangle \\ &= \langle D_jT, \psi\otimes\chi\rangle = \big\langle [g_j], \psi\otimes\chi\big\rangle.\end{aligned}$$

Consequently, if $j \in \{1,\dots,d\}$, we have $D_jU = [\hat g_j]$, where

$$\hat g_j(x_1,\dots,x_d) = \int_{a_{d+1}}^{b_{d+1}} g_j(x_1,\dots,x_d,t)\chi(t)\,dt.$$

Thus U satisfies the induction hypothesis and there exists an element $u \in \mathscr{E}^1(\prod_{j=1}^d (a_j,b_j))$ such that $U = [u]$. Now, for $\varphi \in \mathscr{D}((a,b))$, we have

$$\langle T,\varphi\rangle = \big\langle [G],\varphi\big\rangle - \big\langle [G], \hat\varphi\otimes\chi\big\rangle + \langle U,\hat\varphi\rangle.$$

It follows that $T = [f]$ with

$$\begin{aligned}&f(x_1,\dots,x_{d+1})\\ &\quad= G(x_1,\dots,x_{d+1}) - \int_{a_{d+1}}^{b_{d+1}} G(x_1,\dots,x_d,t)\chi(t)\,dt + u(x_1,\dots,x_d).\end{aligned}$$

Thus $f \in C((a,b))$ and the derivative in the ordinary sense, $\partial f/\partial x_{d+1}$, exists on (a,b) and equals g_{d+1}. One shows similarly that the other partial first derivatives of f in the ordinary sense exist and are continuous, which implies that $f \in \mathscr{E}^1((a,b))$. □

We deduce from this theorem an important uniqueness result.

Theorem 2.7 *Let Ω be a connected open subset of $\mathbb{R}^d$ and suppose that T is a distribution on Ω such that $D_jT = 0$ for every $j \in \{1,\dots,d\}$. Then $T = C$ for some $C \in \mathbb{C}$.*

Proof. By the preceding theorem, there exists $f \in \mathscr{E}^1(\Omega)$ such that $T = [f]$ and $D_jf = 0$ in the ordinary sense, for all $j \in \{1,\dots,d\}$. The result follows. □

Working by induction starting from Theorem 2.6, we see also that, for $r \in \mathbb{N}$ and $T \in \mathscr{D}'(\Omega)$, if $D^p(T) \in C(\Omega)$ for every multiindex p of length r, then $T \in \mathscr{E}^r(\Omega)$.

We will now study in more detail the case of dimension $d = 1$, starting with a characterization of distributions whose derivative is locally integrable.

Theorem 2.8 *Suppose that Ω is an open interval in $\mathbb{R}$ and that $\alpha \in \Omega$. Let $T \in \mathscr{D}'(\Omega)$ and $f \in L^1_{\mathrm{loc}}(\Omega)$. The following properties are equivalent:*

i. $T' = [f]$.

ii. *There exists $C \in \mathbb{C}$ such that $T = [F]$, with $F(x) = C + \int_\alpha^x f(t)\,dt$.*

Functions of the form $F(x) = C + \int_\alpha^x f(t)\,dt$ as above are called **absolutely continuous** on Ω. Thus, a distribution has for derivative a locally integrable function if and only if it "is" an absolutely continuous function. Another way to say this is that, if $f \in L^1_{\mathrm{loc}}(\Omega)$, the function F defined by $F(x) = C + \int_\alpha^x f(t)\,dt$ is a **primitive** of f in the sense of distributions.

Proof. Suppose $\Omega = (a, b)$. Take $f \in L^1_{\mathrm{loc}}((a, b))$ and let $F(x) = \int_\alpha^x f(t)\,dt$. Then, for every $\varphi \in \mathscr{D}((a, b))$,

$$\langle [F]', \varphi \rangle = -\int_a^b \varphi'(x) \left(\int_\alpha^x f(t)\,dt \right) dx.$$

Therefore, by Fubini's Theorem,

$$\begin{aligned}\langle [F]', \varphi \rangle &= \iint_{\{a \le x \le t \le \alpha\}} \varphi'(x) f(t)\,dt\,dx - \iint_{\{\alpha \le t \le x \le b\}} \varphi'(x) f(t)\,dt\,dx \\ &= \int_a^\alpha \varphi(t) f(t)\,dt + \int_\alpha^b \varphi(t) f(t)\,dt = \langle [f], \varphi \rangle.\end{aligned}$$

Thus $[F]' = [f]$ and the desired result follows from the uniqueness theorem proved earlier (Theorem 2.7). □

Still in the case of an open interval $\Omega = (a, b)$ in $\mathbb{R}$, one can characterize distributions whose derivative is positive—which is to say, by Proposition 2.3 on page 270, those whose derivative belongs to the space $\mathfrak{M}^+((a, b))$ of positive Radon measures on (a, b). Recall that, if α is an increasing function on (a, b), we can associate to α a positive Radon measure on (a, b), namely the Stieltjes measure $d\alpha$, and that we obtain in this way all elements of $\mathfrak{M}^+((a, b))$. (We saw this in Section 3A of Chapter 2 (page 71) for the case $(a, b) = \mathbb{R}$, and it extends immediately to the case of an arbitrary open interval (a, b).)

Theorem 2.9 *Suppose that Ω is an open interval in $\mathbb{R}$, and that $T \in \mathscr{D}'(\Omega)$. If there exists an increasing function α on Ω such that $T = [\alpha]$, then $T' = d\alpha$ and therefore T' is positive.*

Conversely, if T' is positive, there exists an increasing function α on Ω and a constant $C \in \mathbb{R}$ such that $T = [\alpha + iC]$.

Proof. Set $\Omega = (a, b)$. Let α be an increasing function on (a, b). Take $\varphi \in \mathscr{D}(\Omega)$ and let c, d be such that $a < c < d < b$ and the support of φ is contained is $[c, d]$. For $n \in \mathbb{N}^*$ and $k \in \{0, \dots, n\}$, set

$$x_k = c + k\,\frac{d - c}{n}.$$

Then, by definition,

$$\int \varphi\, d\alpha = \int_c^d \varphi(x)\, d\alpha(x) = \lim_{n\to+\infty} \sum_{k=0}^{n-1} \varphi(x_k)\big(\alpha(x_{k+1}) - \alpha(x_k)\big).$$

We perform a summation by parts. Since $\varphi(c) = \varphi(d) = 0$, we have

$$\sum_{k=0}^{n-1} \varphi(x_k)\big(\alpha(x_{k+1}) - \alpha(x_k)\big) = \sum_{k=1}^{n} \alpha(x_k)\big(\varphi(x_{k-1}) - \varphi(x_k)\big).$$

Consequently,

$$\int \varphi\, d\alpha = -\lim_{n\to+\infty} \int_c^d \varphi'(x)\bigg(\sum_{k=1}^{n} 1_{[x_{k-1},x_k)}(x)\alpha(x_k)\bigg)\, dx.$$

Using the Dominated Convergence Theorem, we obtain

$$\int \varphi\, d\alpha = -\int \varphi'(x)\alpha(x_+)\, dx.$$

(Recall that $\alpha(x_+)$ denotes the limit from the right of the function α at x.) Now, $\alpha(x_+) = \alpha(x)$ except at a set of points x that is countable, and so of Lebesgue measure zero (see Exercise 6 on page 5). Therefore

$$\int \varphi\, d\alpha = -\int \varphi'(x)\alpha(x)\, dx,$$

so $d\alpha = [\alpha]'$. This proves the first part of the theorem.

Now suppose that T' is positive. By Proposition 2.3 on page 270, T' is a positive Radon measure on Ω. By Theorem 3.8 on page 73 (applied to Ω rather than $\mathbb{R}$), there exists an increasing function α such that $T' = d\alpha$ (we may assume α is right continuous). Then, by the first part of this proof, $T' = [\alpha]'$. Now it suffices to apply the uniqueness theorem (Theorem 2.7) to obtain $T = [\alpha + C]$, for $C \in \mathbb{C}$. The desired result follows by replacing α with $\alpha + \operatorname{Re} C$ and C with $\operatorname{Im} C$. □

Obviously, in the preceding theorem, we can assume that $C = 0$ if T is real—that is, if $\langle T, \varphi\rangle \in \mathbb{R}$ for every real-valued $\varphi \in \mathscr{D}(\Omega)$.

We also see from Theorem 2.9 that every positive Radon measure of finite mass μ on $\mathbb{R}$ is the derivative in the sense of distributions of its distribution function F, defined by $F(x) = \mu((-\infty, x])$. Indeed, by Remark 1 on page 74, we have $\mu = dF$. In particular, we recover the result that $\mu = F'(x)\, dx$ if F is of class C^1.

The next theorem, applicable to a large class of functions of one variable, links the derivative in the sense of distributions with the ordinary derivative.

Theorem 2.10 *Suppose that Ω is an open subset of $\mathbb{R}$ and that f is a function on Ω for which there exist points $x_1 < \cdots < x_n$ in Ω satisfying these conditions:*

- *f is of class C^1 on $\Omega \setminus \{x_1, \ldots, x_n\}$.*
- *For every $j \in \{1, \ldots, n\}$, f has right and left limits at x_j, which we denote by $f(x_{j+})$ and $f(x_{j-})$, respectively.*
- *The ordinary derivative f' of f, defined on $\Omega \setminus \{x_1, \ldots, x_n\}$, belongs to $L^1_{\mathrm{loc}}(\Omega)$.*

Then

$$[f]' = [f'] + \sum_{j=1}^{n} \big(f(x_{j+}) - f(x_{j-})\big)\delta_{x_j}.$$

Proof. Considering separately each of the connected components of Ω, we can assume that Ω is an open interval (a, b). Put $x_0 = a$ and $x_{n+1} = b$. Then, if $\varphi \in \mathscr{D}(\Omega)$, we have

$$\langle [f]', \varphi \rangle = -\langle [f], \varphi' \rangle = -\sum_{j=0}^{n} \int_{x_j}^{x_{j+1}} f(t)\varphi'(t)\, dt,$$

or yet, integrating by parts (and setting $\varphi(a) = \varphi(b) = 0$),

$$\begin{aligned}\langle [f]', \varphi \rangle &= \sum_{j=0}^{n} \Big(\int_{x_j}^{x_{j+1}} \varphi(t) f'(t)\, dt + f(x_{j+})\varphi(x_j) - f((x_{j+1})_-)\varphi(x_{j+1}) \Big) \\ &= \int \varphi(t) f'(t)\, dt + \sum_{j=1}^{n} \varphi(x_j)\big(f(x_{j+}) - f(x_{j-})\big),\end{aligned}$$

which concludes the proof. □

By induction on $p \in \mathbb{N}^*$, we deduce the following corollary.

Corollary 2.11 *Suppose that Ω is an open subset of $\mathbb{R}$ and that f is a function on Ω for which there exist points $x_1 < \cdots < x_n$ in Ω and an integer $p \in \mathbb{N}^*$ satisfying these conditions:*

- *f is of class C^p on $\Omega \setminus \{x_1, \ldots, x_n\}$.*
- *For every $j \in \{1, \ldots, n\}$ and every integer $k \in \{0, \ldots, p-1\}$, the right and left limits of $f^{(k)}$ at x_j exist.*
- *The ordinary p-th derivative $f^{(p)}$ of f, defined on $\Omega \setminus \{x_1, \ldots, x_n\}$, belongs to $L^1_{\mathrm{loc}}(\Omega)$.*

Then

$$[f]^{(p)} = [f^{(p)}] + \sum_{j=1}^{n} \sum_{k=0}^{p-1} \big(f^{(k)}(x_{j+}) - f^{(k)}(x_{j-})\big)\delta_{x_j}^{(p-1-k)}.$$

Examples

The following examples are immediate applications of the two preceding results.

1. Recall our notation $x^+ = \max(0, x)$. Then $[x^+]' = Y$.
2. We also find that $Y' = \delta$. More generally, if $Y_a = 1_{[a,+\infty)}$, then $Y_a' = \delta_a$.
3. $\big[|x|/2\big]'' = \delta$.
4. Let f be a function of class C^p on $\mathbb{R}$. Then

$$[Yf]^{(p)} = [Yf^{(p)}] + \sum_{k=0}^{p-1} f^{(k)}(0)\delta^{(p-1-k)}.$$

In dimension $d \geq 2$, Theorem 2.10 has the following partial generalization (see also Exercise 15).

Theorem 2.12 *Suppose that $d \geq 2$ and, if $(x_2, \ldots, x_d) \in \mathbb{R}^{d-1}$, write*

$$\Omega_{x_2,\ldots,x_d} = \{x_1 \in \mathbb{R} : (x_1, x_2, \ldots, x_d) \in \Omega\}.$$

Let $f \in L^1_{\text{loc}}(\Omega)$ satisfy the following conditions:

- *For almost every $(x_2, \ldots, x_d) \in \mathbb{R}^{d-1}$, the map on $\Omega_{x_2,\ldots,x_d}$ defined by*

$$x_1 \mapsto f(x_1, x_2, \ldots, x_d)$$

  is continuous on $\Omega_{x_2,\ldots,x_d}$ and of class C^1 except at finitely many points of $\Omega_{x_2,\ldots,x_d}$.
- *The ordinary partial derivative $\partial f/\partial x_1$, defined almost everywhere on Ω, is an element of $L^1_{\text{loc}}(\Omega)$.*

Then

$$D_1[f] = \left[\frac{\partial f}{\partial x_1}\right].$$

Of course, an analogous result holds if we replace the subscript 1 by any $j \in \{2, \ldots, d\}$.

Proof. Argue as in the proof of Theorem 2.10 and apply Fubini's Theorem. □

Examples

1. $D_j(x_1^+) = \begin{cases} 0 & \text{if } j \neq 1, \\ Y(x_1) & \text{if } j = 1. \end{cases}$
2. Set $r = |x| = \sqrt{x_1^2 + \cdots + x_d^2}$ and $B_1 = \{x \in \mathbb{R}^d : |x| < 1\}$. By Theorem 3.9 on page 74 and Remark 2 on page 76,

$$\int_{B_1} r^{-\alpha}\, dx = s_d \int_0^1 r^{d-1-\alpha}\, dr \leq +\infty,$$

   where s_d is the area of the unit sphere in $\mathbb{R}^d$ ($s_d = d\omega_d$, where $\omega_d = \lambda(B_1)$ is the volume of B_1). As a consequence:

Proposition 2.13 *The function $x \mapsto r^{-\alpha}$ is locally integrable on $\mathbb{R}^d$ if and only if $\alpha < d$.*

Therefore we obtain, as a consequence of Theorem 2.12:

- If $\alpha < d-1$ and $1 \le j \le d$,

$$D_j\left(\frac{1}{r^\alpha}\right) = -\alpha\,\frac{x_j}{r^{\alpha+2}}\,.$$

- If $d \ge 2$ and $1 \le j \le d$,

$$D_j(\log r) = \frac{x_j}{r^2}\,.$$

(The local integrability of the derivatives follows from the preceding criterion and the fact that $|x_j| \le r$.)

Exercises

1. Show that, for every distribution T on an open subset Ω of $\mathbb{R}^d$ and for every $i, j \in \{1, \dots, d\}$,

$$D_i D_j T = D_j D_i T.$$

2. **a.** For $h \in \mathbb{R}^d$, let τ_h be the operator on $\mathscr{D}'(\mathbb{R}^d)$ defined by

$$\langle \tau_h T, \varphi\rangle = \big\langle T,\, \varphi(\,\cdot + h)\big\rangle.$$

If the distribution T is defined by a locally integrable function f, what does $\tau_h T$ correspond to?

Show that

$$\lim_{h_1 \to 0} \frac{\tau_{(h_1,0,\dots,0)}T - T}{h_1} = -D_1 T$$

in $\mathscr{D}'(\mathbb{R}^d)$.

b. We say of a distribution T on $\mathbb{R}^d$ that it does not depend on the first variable (say) if $\tau_{(h_1,0,\dots,0)}T = T$ for every $h_1 \in \mathbb{R}$. Show that T does not depend on the first variable if and only if $D_1 T = 0$. (See also Exercise 6 on page 324.)

Hint. For any function $\varphi \in \mathscr{D}(\mathbb{R}^d)$, find the derivative of the function f defined on $\mathbb{R}$ by

$$f(h_1) = \langle \tau_{(h_1,0,\dots,0)}T,\, \varphi\rangle.$$

3. Let T be a distribution on $\mathbb{R}$. Show that T is defined by a Lipschitz function if and only if $T' \in L^\infty$. (In particular, Lipschitz functions are absolutely continuous.)

Hint. The "if" part follows from Theorem 2.8. To prove the "only if" part, use the first part of Exercise 2, the duality $L^\infty = (L^1)'$, and the fact that $\mathscr{D}$ is dense in L^1.

4. Let T be a distribution on $\mathbb{R}^d$ such that $D^p T = 0$ for every multiindex p of length $m+1$. Show that T is defined by a polynomial function of degree at most m.

Hint. Work by induction on m.

5. Let Ω be an open interval in $\mathbb{R}$.

a. Show that, if f is an absolutely continuous function on Ω and $g \in \mathscr{E}^1(\Omega)$, then gf is absolutely continuous on Ω and $[gf]' = [g'f + gf_1]$, where f_1 is the element of $L^1_{\mathrm{loc}}(\Omega)$ defined by $[f]' = [f_1]$.

Hint. Write $[gf] = g[f]$ and apply Leibniz's formula.

b. Let g be an absolutely continuous function on Ω and suppose $g_1 \in L^1_{\mathrm{loc}}(\Omega)$ satisfies $[g]' = [g_1]$. Show that there is a sequence $(g_n)_{n\in\mathbb{N}}$ in $\mathscr{E}^1(\Omega)$ such that (g_n') converges to g_1 in $L^1_{\mathrm{loc}}(\Omega)$ and (g_n) converges to g in $C(\Omega)$.

c. Deduce from this that, if f and g are absolutely continuous on Ω, so is fg. Write down $[fg]'$ in this case.

6. Show that the map defined on $\mathscr{D}(\mathbb{R}^2)$ by

$$\langle T, \varphi \rangle = \int_{\mathbb{R}} \varphi(x, x)\, dx$$

is a distribution. Find its support and its order, and compute $\dfrac{\partial T}{\partial x_1} + \dfrac{\partial T}{\partial x_2}$.

7. a. Let $\alpha \in \mathbb{R}^+ \setminus \mathbb{N}$. Show that

$$\frac{d}{dx}\,\mathrm{fp}\left(\frac{Y(x)}{x^\alpha}\right) = -\alpha\,\mathrm{fp}\left(\frac{Y(x)}{x^{\alpha+1}}\right).$$

b. Let $m \in \mathbb{N}^*$. Show that

$$\frac{d}{dx}\,\mathrm{fp}\left(\frac{Y(x)}{x^m}\right) = -m\,\mathrm{fp}\left(\frac{Y(x)}{x^{m+1}}\right) + \frac{(-1)^m}{m!}\,\delta^{(m)}.$$

(The finite part of a function $x \mapsto Y(x)/x^\alpha$, where $\alpha > 0$, was defined in Exercise 19 on page 279.)

c. Use this to find the successive derivatives of $\mathrm{fp}(Y(x)/x)$.

8. Compute the second derivative, in the sense of distributions on $\mathbb{R}$, of the function f defined by $f(x) = \max(1-|x|, 0)$.

9. We denote by σ the surface measure of the unit sphere in $\mathbb{R}^2$. Recall from Exercise 16 on page 83 that

$$\int \varphi\, d\sigma = \int_0^{2\pi} \varphi(\cos\theta, \sin\theta)\, d\theta \quad \text{for all } \varphi \in C_c(\mathbb{R}^2).$$

We set $f(x,y) = \max\bigl(1 - \sqrt{x^2+y^2}, 0\bigr)$ and

$$\chi(x,y) = \begin{cases} (x^2+y^2)^{-1/2} & \text{if } 0 < x^2+y^2 < 1, \\ 0 & \text{otherwise.} \end{cases}$$

a. Calculate $\partial f/\partial x$ and $\partial f/\partial y$ in the sense of distributions on $\mathbb{R}^2$ (show in particular that these derivatives are functions).

b. Take $\varphi \in \mathscr{D}(\mathbb{R}^2)$ and set $\psi(\rho,\theta) = \varphi(\rho\cos\theta,\, \rho\sin\theta)$. Show that

$$\begin{aligned}\iint_{(0,1)\times(0,2\pi)} \frac{\partial\psi}{\partial\rho}\,\rho\,d\rho\,d\theta \\ &= \iint_{\{x^2+y^2<1\}} \frac{1}{\sqrt{x^2+y^2}}\left(x\frac{\partial\varphi}{\partial x} + y\frac{\partial\varphi}{\partial y}\right)dx\,dy \\ &= \langle \Delta f, \varphi\rangle.\end{aligned}$$

c. Deduce that $\Delta f = \sigma - \chi$ in the sense of distributions.

10. For $r > 0$, let σ_r be the surface measure of the sphere of center 0 and radius r in $\mathbb{R}^d$. Show that

$$\lim_{r\to 0^+} \frac{2d}{r^2}\left(\frac{1}{s_d r^{d-1}}\sigma_r - \delta\right) = \Delta\delta$$

in $\mathscr{D}'(\mathbb{R}^d)$.

Hint. Use the Taylor–Young formula and (after having proved them) the equalities

$$\int x_j\,d\sigma_r = 0, \qquad \int x_i x_j\,d\sigma_r = 0 \ \text{ if } i \neq j, \qquad \int x_j^2\,d\sigma_r = \frac{s_d}{d} r^{d+1}.$$

11. Let f be a real-valued function of class C^2 on $\mathbb{R}^d$, satisfying $\Delta f = 0$.

a. For $\varepsilon > 0$, set $g_\varepsilon = (\varepsilon^2 + f^2)^{1/2}$. Compute Δg_ε and show that it is a positive function.

b. Show that g_ε tends to $|f|$ in $\mathscr{D}'(\mathbb{R}^d)$ when ε tends to 0.

c. Show that there exists a positive Radon measure μ on $\mathbb{R}^d$ such that $\Delta|f| = \mu$ in $\mathscr{D}'(\mathbb{R}^d)$. Show that the support of μ is contained in $f^{-1}(0)$.

d. Determine μ by direct calculation when $d = 2$ and $f(x,y) = xy$.

12. Let $\Omega = (a,b)$ be an open interval in $\mathbb{R}$.

a. Let f be a convex function on Ω.

i. Show that, if $\varphi \in \mathscr{D}(\mathbb{R})$,

$$\varphi'' = \lim_{h\to 0^+} h^{-2}(\tau_h\varphi + \tau_{-h}\varphi - 2\varphi)$$

in $\mathscr{D}(\mathbb{R})$, where, if $k \in \mathbb{R}$, $\tau_k\varphi(x) = \varphi(x-k)$.

ii. Deduce that $[f]''$ is a positive Radon measure on Ω.

b. Conversely, suppose that T is a distribution on Ω and that T'' is a positive Radon measure on Ω. Show that there exists a convex function f on Ω such that $T - [f]$ is a first-degree polynomial with coefficients in $\mathbb{C}$, and that we can assume this polynomial to be zero if T is real in the sense of Exercise 1 on page 275.

Hint. Check that, if α is an increasing function on Ω and if $c \in \Omega$, the function f defined by $f(x) = \int_c^x \alpha(t)\,dt$ is convex.

c. Deduce that a function f is a difference of two convex functions if and only if it is continuous, real-valued, and $[f]''$ is a Radon measure on (a, b).

13. Let Ω be an open interval in $\mathbb{R}$. Show that a distribution T on Ω has as its first derivative a Radon measure on Ω if and only if there exists a function α of bounded variation on every compact interval contained in Ω (see Exercise 13 on page 93) such that $T = [\alpha]$. (You might also recall Exercise 15 on page 94.)

14. Let $r \in \mathbb{N}$. Show that

$$\lim_{N \to +\infty} \sum_{n=-N}^{N} n^r e^{inx} = 2\pi(-i)^r \sum_{p \in \mathbb{Z}} \delta_{2p\pi}^{(r)}.$$

Hint. Use Exercise 13 on page 277.

15. Let S_1 be the unit sphere in $\mathbb{R}^d$ and let σ_1 be its surface measure (page 74). For $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, write $\tilde{x} = (x_2, \ldots, x_d) \in \mathbb{R}^{d-1}$ and $\tilde{r} = \sqrt{1 - |\tilde{x}|^2}$.

a. Take $\varphi \in C(S_1)$. Extend φ to the ball $B_1 = \{x \in \mathbb{R}^d : |x| < 1\}$ by setting

$$\varphi(x) = \frac{1}{2}\Big(\Big(1 - \frac{x_1}{\tilde{r}}\Big)\varphi(-\tilde{r}, \tilde{x}) + \Big(1 + \frac{x_1}{\tilde{r}}\Big)\varphi(\tilde{r}, \tilde{x})\Big).$$

i. Show that the extended φ is continuous on $\bar{B}_1$.

ii. Show that, for $r \le 1$,

$$\int_{B_r} \varphi(x)\, dx = \int_{\{|\tilde{x}| < r\}} (r^2 - |\tilde{x}|^2)^{1/2} \big(\varphi(-\tilde{r}, \tilde{x}) + \varphi(\tilde{r}, \tilde{x})\big)\, d\tilde{x},$$

where $B_r = \{x \in \mathbb{R}^d : |x| < r\}$.

iii. Show that the map

$$r \mapsto \int_{B_r} \varphi(x)\, dx$$

is left differentiable at the point 1, and find its left derivative. Deduce that

$$\int \varphi(x)\, d\sigma_1(x) = \int_{\{|\tilde{x}| < 1\}} \frac{\varphi(-\tilde{r}, \tilde{x}) + \varphi(\tilde{r}, \tilde{x})}{\tilde{r}}\, d\tilde{x}.$$

iv. Show the same result with $\tilde{x} = (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_d)$.

b. For $\rho > 0$, let S_ρ be the sphere of center 0 and radius ρ in $\mathbb{R}^d$ and let σ_ρ be its surface measure. Let f be an element of $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ whose restriction to $\mathbb{R}^d \setminus S_\rho$ is of class C^1. Take $j \in \{1, \dots, d\}$, and assume that $(\partial f/\partial x_j) \in L^1_{\text{loc}}(\mathbb{R}^d)$ and that, for every $x \in S_\rho$, the limits

$$f_+^\rho(x) = \lim_{\substack{y \to x \\ |y|>\rho}} f(y), \qquad f_-^\rho(x) = \lim_{\substack{y \to x \\ |y|<\rho}} f(y)$$

exist.

i. Show that the functions f_+^ρ and f_-^ρ are continuous on S_ρ.

ii. Show that

$$D_j[f] = \left[\frac{\partial f}{\partial x_j}\right] + \frac{x_j}{\rho}(f_+^\rho - f_-^\rho)\sigma_\rho.$$

Hint. Reduce to the case $\rho = 1$ by setting $f_\rho(x) = f(\rho x)$; then use the representation of the measure σ_1 given in part a.

iii. Use this result to compute Δf in Exercise 9.

iv. State and prove a similar result when S_ρ is replaced by a hyperplane in $\mathbb{R}^d$.

16. Consider in $\mathscr{D}'(\mathbb{R})$ the equation

$$2xT' - T = \delta, \qquad (*)$$

where δ is the Dirac measure at 0.

a. For an arbitrary integer $j \geq 1$, express the distribution $x\delta^{(j)}$ in terms of $\delta^{(j-1)}$.

b. Determine the solutions of $(*)$ whose support is $\{0\}$. (You might use the result from Exercise 4 on page 285.)

c. Let T be a solution of $(*)$. Denote by U and V the restrictions of T to $(0, +\infty)$ and $(-\infty, 0)$, respectively. Thus $U \in \mathscr{D}'((0, +\infty))$ and $V \in \mathscr{D}'((-\infty, 0))$. By computing $(x^{-1/2}U)'$ in $\mathscr{D}'((0, +\infty))$ and $((-x)^{-1/2}V)'$ in $\mathscr{D}'((-\infty, 0))$, determine U and V.

d. Show that, for every $(\lambda, \mu) \in \mathbb{R}^2$, the distribution S defined by

$$S(x) = \lambda\sqrt{xY(x)} + \mu\sqrt{-xY(-x)}$$

satisfies $2xS' - S = 0$.

e. Deduce from this the general form of the solutions of $(*)$.

3 Fundamental Solutions of a Differential Operator

Let P be a complex polynomial of degree m in d indeterminates:

$$P(X) = \sum_{|p| \leq m} a_p X_1^{p_1} \dots X_d^{p_d}.$$

The linear map

$$P(D) = \sum_{|p| \le m} a_p D^p,$$

which is a linear combination of differentiation operators, is called a **linear differential operator of order m with constant coefficients** on $\mathbb{R}^d$.

For example, if $P(X) = X_1^2 + \cdots + X_d^2$, the operator $P(D)$ is exactly the Laplacian on $\mathbb{R}^d$:

$$P(D) = \Delta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}\,.$$

If $P(D)$ is such an operator, we define a **fundamental solution** of $P(D)$ as any distribution $E \in \mathscr{D}'(\mathbb{R}^d)$ such that $P(D)E = \delta$. This notion will play an important role in the next chapter. For example, if $d = 1$, the Heaviside function is a fundamental solution of the differential operator $P(D) = D$, since $Y' = \delta$. The next theorem shows that, if $d = 1$, every linear differential operator with constant coefficients has a fundamental solution.

Theorem 3.1 *Let $P(X) = \sum_{j=0}^{m} a_j X^j$, where $m \in \mathbb{N}^*$, $a_1, \dots, a_m \in \mathbb{C}$, and $a_m \ne 0$. Let φ be the solution on $\mathbb{R}$ of the differential equation*

$$\sum_{j=0}^{m} a_j \varphi^{(j)} = 0$$

such that $\varphi^{(m-1)}(0) = 1$ and $\varphi^{(j)}(0) = 0$ for every $j \le m-2$. Then $E = (1/a_m) Y\varphi$ is a fundamental solution of $P(D)$.

Proof. As a particular case of Example 4 on page 301, we have

$$\begin{aligned} [Y\varphi]^{(m)} &= [Y\varphi^{(m)}] + \delta, \\ [Y\varphi]^{(k)} &= [Y\varphi^{(k)}] \quad \text{for all } k \le m-1, \end{aligned}$$

so that

$$P(D)E = a_m^{-1} \sum_{j=0}^{m} a_j [Y\varphi]^{(j)} = a_m^{-1} \left[\sum_{j=0}^{m} a_j Y \varphi^{(j)} \right] + \delta = \delta. \qquad \square$$

Obviously, there is no uniqueness for fundamental solutions: two fundamental solutions differ by a solution of the associated differential equation.

We will now exhibit fundamental solutions of certain classical linear differential operators.

3A The Laplacian

Consider the Laplace operator, or Laplacian, in dimension d:

$$\Delta = \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} = \sum_{j=1}^{d} D_j^2.$$

As before, we set $r = |x|$.

Theorem 3.2 *Let E be the distribution on $\mathbb{R}^d$ defined by*

- $E = r/2$ *if* $d = 1$,
- $E = \dfrac{1}{2\pi}\log r$ *if* $d = 2$,
- $E = -\dfrac{1}{s_d(d-2)}\dfrac{1}{r^{d-2}}$ *if* $d \geq 3$.

Then $\Delta E = \delta$.

Proof. The case $d = 1$ was dealt with in Example 3 on page 301.

Case $d = 2$. Suppose $d = 2$ and let $f(x) = \log r$. Since the first derivatives of the function f do not satisfy the hypotheses of Theorem 2.12, we cannot use that theorem directly to compute the Laplacian in the sense of distributions. For this reason we approximate in $\mathscr{D}'(\mathbb{R}^2)$ the distribution $[f]$ by a family $([f_\varepsilon])$ of distributions defined by functions whose Laplacian we *can* compute by applying Theorem 2.12 to the functions f_ε and to their first-order derivatives. We then obtain the Laplacian $\Delta[f]$ by passing to the limit.

Thus we define, for $\varepsilon \in (0,1)$, a function f_ε by

$$f_\varepsilon(x) = \begin{cases} \log r & \text{if } r \geq \varepsilon, \\ a_\varepsilon r^2 + b_\varepsilon & \text{if } r \leq \varepsilon, \end{cases}$$

where a_ε and b_ε are real numbers chosen so that the function f_ε is of class C^1 on $\mathbb{R}^2$; that is, so that

$$a_\varepsilon \varepsilon^2 + b_\varepsilon = \log \varepsilon \qquad \text{and} \qquad 2a_\varepsilon \varepsilon = \frac{1}{\varepsilon}.$$

Thus

$$a_\varepsilon = \frac{1}{2\varepsilon^2}, \qquad b_\varepsilon = \log \varepsilon - \frac{1}{2}.$$

Now, if $r \leq \varepsilon$,

$$|a_\varepsilon r^2 + b_\varepsilon| = \log \frac{1}{\varepsilon} + \frac{1}{2}\left(1 - \left(\frac{r}{\varepsilon}\right)^2\right) \leq \log \frac{1}{r} + \frac{1}{2}.$$

We deduce that, for every $x \in \mathbb{R}^2$,

$$|f_\varepsilon(x)| \leq |\log r| + \tfrac{1}{2};$$

thus, by the Dominated Convergence Theorem,

$$\lim_{\varepsilon \to 0} [f_\varepsilon] = [f]$$

in $\mathscr{D}'(\mathbb{R}^2)$. At the same time, for $j = 1, 2$,

$$D_j f_\varepsilon(x) = \begin{cases} x_j/r^2 & \text{if } r \geq \varepsilon, \\ x_j/\varepsilon^2 & \text{if } r \leq \varepsilon. \end{cases}$$

This function $D_j f_\varepsilon$ satisfies, by construction, the hypotheses of Theorem 2.12. We deduce that $D_j^2[f_\varepsilon] = [g_j^\varepsilon]$, with

$$g_j^\varepsilon(x) = \begin{cases} \dfrac{1}{r^2} - 2\dfrac{x_j^2}{r^4} & \text{if } r \geq \varepsilon, \\ 1/\varepsilon^2 & \text{if } r \leq \varepsilon. \end{cases}$$

Therefore

$$\Delta[f_\varepsilon] = \frac{2}{\varepsilon^2} 1_{\bar{B}(0,\varepsilon)}.$$

An elementary calculation shows that

$$\lim_{\varepsilon \to 0} \frac{1}{\pi\varepsilon^2} 1_{\bar{B}(0,\varepsilon)} = \delta$$

in $\mathscr{D}'(\mathbb{R}^2)$; thus, by continuity in $\mathscr{D}'(\mathbb{R}^2)$ of the operator Δ,

$$\Delta[f] = 2\pi\delta,$$

which proves our result since $E = f/(2\pi)$.

Case $d = 3$. We work as in the previous case. For $\varepsilon > 0$, set

$$f_\varepsilon(x) = \begin{cases} r^{2-d} & \text{if } r \geq \varepsilon, \\ a_\varepsilon r^2 + b_\varepsilon & \text{if } r \leq \varepsilon, \end{cases}$$

where a_ε and b_ε are real numbers chosen so that the function f_ε is of class C^1 on $\mathbb{R}^d$; that is, so that (one concludes after some calculations),

$$a_\varepsilon = \frac{2-d}{2}\varepsilon^{-d}, \qquad b_\varepsilon = \frac{d}{2}\varepsilon^{2-d}.$$

Thus

$$a_\varepsilon r^2 + b_\varepsilon = \varepsilon^{2-d}\left(\frac{d}{2} - \frac{d-2}{2}\left(\frac{r}{\varepsilon}\right)^2\right),$$

which implies that

$$0 \leq f_\varepsilon(x) \leq \frac{d}{2} r^{2-d}.$$

Thus, by the Dominated Convergence Theorem, $\lim_{\varepsilon \to 0}[f_\varepsilon] = [f]$ in $\mathscr{D}'(\mathbb{R}^d)$, with $f(x) = r^{2-d}$. A calculation similar to the one carried out in the case $d = 2$ yields

$$\Delta[f_\varepsilon] = d(2-d)\varepsilon^{-d} 1_{\bar{B}(0,\varepsilon)}.$$

Now,

$$\lim_{\varepsilon \to 0} \frac{1}{\omega_d \varepsilon^d} 1_{\bar{B}(0,\varepsilon)} = \delta$$

in $\mathscr{D}'(\mathbb{R}^d)$; therefore, by the continuity of Δ,

$$\Delta[f] = d\omega_d(2-d)\delta.$$

Since $d\omega_d = s_d$ and $E = -f/(s_d(d-2))$, the result follows. $\square$

3B The Heat Operator

We now place ourselves in the space $\mathbb{R}^{d+1} = \mathbb{R} \times \mathbb{R}^d$, a generic point of which will be denoted by (t, x). For $c > 0$, we define the **heat operator** $\mathscr{C}$ by

$$\mathscr{C} = \frac{\partial}{\partial t} - c \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2} = \frac{\partial}{\partial t} - c\Delta.$$

Theorem 3.3 *For $(t, x) \in \mathbb{R} \times \mathbb{R}^d$, let*

$$\Gamma(t, x) = 1_{(0,+\infty)}(t) \frac{1}{(4c\pi t)^{d/2}} e^{-|x|^2/(4ct)}.$$

Then $\Gamma \in L^1_{\mathrm{loc}}(\mathbb{R}^{d+1})$ and $\mathscr{C}\Gamma = \delta$ in $\mathscr{D}'(\mathbb{R}^{d+1})$.

Proof. For $t > 0$, we obtain, by applying the change of variables $u = x/\sqrt{2ct}$ and then Fubini's Theorem,

$$\int_{\mathbb{R}^d} \Gamma(t, x)\, dx = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-|u|^2/2}\, du = \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2}\, dx \right)^d.$$

Since $\displaystyle\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-x^2/2} dx = 1$ (a classical result), we get

$$\int_{\mathbb{R}^d} \Gamma(t, x)\, dx = 1 \quad \text{for all } t > 0,$$

which in particular proves, by Fubini's Theorem, that $\Gamma \in L^1_{\mathrm{loc}}(\mathbb{R}^{d+1})$. Now take $\varphi \in \mathscr{D}(\mathbb{R}^{d+1})$. Then

$$\begin{aligned}\left\langle \frac{\partial}{\partial t}\Gamma, \varphi \right\rangle &= -\int_0^{+\infty} \int_{\mathbb{R}^d} \frac{\partial \varphi}{\partial t}(t, x)\Gamma(t, x)\, dt\, dx \\ &= -\lim_{\varepsilon \to 0} \int_{\varepsilon}^{+\infty} \int_{\mathbb{R}^d} \frac{\partial \varphi}{\partial t}(t, x)\Gamma(t, x)\, dt\, dx.\end{aligned}$$

Integrating by parts, we get

$$\left\langle \frac{\partial}{\partial t}\Gamma, \varphi \right\rangle = \lim_{\varepsilon \to 0} \left(\int_{\varepsilon}^{+\infty} \int_{\mathbb{R}^d} \varphi(t,x) \frac{\partial \Gamma}{\partial t}(t,x)\,dt\,dx + \int_{\mathbb{R}^d} \varphi(\varepsilon,x)\Gamma(\varepsilon,x)\,dx \right). \quad (*)$$

But, if $t > 0$, we have $\Gamma(t,x) = t^{-d/2}\Gamma(1, x/\sqrt{t})$. Therefore, applying the change of variables $x = \sqrt{\varepsilon}\,u$, we get

$$\int_{\mathbb{R}^d} \varphi(\varepsilon,x)\Gamma(\varepsilon,x)\,dx = \int_{\mathbb{R}^d} \varphi(\varepsilon, \sqrt{\varepsilon}\,u)\Gamma(1,u)\,du.$$

This expression tends to $\varphi(0)\int \Gamma(1,u)\,du = \varphi(0)$ as ε tends to 0, by the Dominated Convergence Theorem. Moreover, Γ is of class C^∞ on the complement of the set $\{t=0\} \times \mathbb{R}^d$, and an elementary calculation shows that $\partial\Gamma/\partial t = c\,\Delta\Gamma$ (in the classical sense) on the set $\{t > 0\} \times \mathbb{R}^d$; therefore, if $\varepsilon > 0$,

$$\begin{aligned} \int_{\{t \geq \varepsilon\} \times \mathbb{R}^d} \varphi(t,x) \frac{\partial \Gamma}{\partial t}(t,x)\,dt\,dx &= \int_{\{t \geq \varepsilon\} \times \mathbb{R}^d} c\varphi(t,x)\Delta\Gamma(t,x)\,dt\,dx \\ &= \int_{\{t \geq \varepsilon\} \times \mathbb{R}^d} c\Delta\varphi(t,x)\Gamma(t,x)\,dt\,dx \end{aligned}$$

(again integrating by parts). Taking the limit, we deduce then from equality $(*)$ that

$$\left\langle \frac{\partial}{\partial t}\Gamma, \varphi \right\rangle = c\,\langle \Delta\Gamma, \varphi \rangle + \varphi(0),$$

and so that $\mathscr{C}\Gamma = \delta$. □

3C The Cauchy–Riemann Operator

The **Cauchy–Riemann operator** is important in the theory of holomorphic functions. It is denoted by $\partial/\partial\bar{z}$ and is defined, for $d = 2$, by

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y} \right).$$

In the sequel, we use the notation $z = x + iy$.

Theorem 3.4 *In* $\mathscr{D}'(\mathbb{R}^2)$,

$$\frac{\partial}{\partial \bar{z}}\left(\frac{1}{\pi z} \right) = \delta.$$

Proof. We follow a method analogous to the one used in the proof of Theorem 3.2. For $\varepsilon > 0$, put

$$f_\varepsilon(x,y) = \begin{cases} 1/z & \text{if } |z| > \varepsilon, \\ \bar{z}/\varepsilon^2 & \text{if } |z| \leq \varepsilon. \end{cases}$$

Then f_ε is continuous on $\mathbb{R}^2$ and, by Theorem 2.12,

$$\frac{\partial}{\partial x}[f_\varepsilon] = [g_{1,\varepsilon}], \qquad \frac{\partial}{\partial y}[f_\varepsilon] = [g_{2,\varepsilon}],$$

with

$$g_{1,\varepsilon}(x,y) = \begin{cases} -\dfrac{1}{z^2} & \text{if } |z| > \varepsilon, \\ \dfrac{1}{\varepsilon^2} & \text{if } |z| < \varepsilon; \end{cases} \qquad g_{2,\varepsilon}(x,y) = \begin{cases} -\dfrac{i}{z^2} & \text{if } |z| > \varepsilon, \\ -\dfrac{i}{\varepsilon^2} & \text{if } |z| < \varepsilon. \end{cases}$$

Thus

$$\frac{\partial}{\partial \bar z}[f_\varepsilon] = \frac{1}{\varepsilon^2} 1_{\bar B(0,\varepsilon)},$$

which tends to $\pi\delta$ in $\mathscr{D}'(\mathbb{R}^2)$ when ε tends to 0. We have $|f_\varepsilon(x,y)| \le 1/|z|$, so the Dominated Convergence Theorem implies that $[f_\varepsilon]$ tends to $[f]$ in $\mathscr{D}'(\mathbb{R}^2)$, with $f(x,y) = 1/z$. Therefore

$$\frac{\partial}{\partial \bar z}[f] = \lim_{\varepsilon \to 0} \frac{\partial}{\partial \bar z}[f_\varepsilon] = \pi\delta,$$

whence the result. □

Exercises

1. Determine a fundamental solution of the differential operator defined on $\mathbb{R}$ by $P(D) = D^2 - 2D - 3$.

2. Let T be the distribution on $\mathbb{R}^2$ defined by the characteristic function of the set $\{(x,y) \in \mathbb{R}^2 : 0 \le y \le x\}$. Show that

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial x\, \partial y}\right) T = \delta.$$

3. Let E be the fundamental solution of the Laplacian given in Theorem 3.2. Define a function Φ on $(0, +\infty)$ by $\Phi(r) = E(x)$ and put

$$E^\rho(x) = \begin{cases} \Phi(r) & \text{if } r \ge \rho, \\ \Phi(\rho) & \text{if } r \le \rho. \end{cases}$$

Show that

$$\Delta E^\rho = \frac{1}{s_d}\rho^{1-d}\sigma_\rho,$$

where σ_ρ is the surface measure on the sphere of center 0 and radius ρ. Derive another proof of Theorem 3.2.

Hint. Use Exercise 15b on page 305. In the case $d = 3$ (for example), you might also use the following more elementary reasoning:

a. Reduce to the case $\rho = 1$.

b. Take $\varepsilon \in (0,1)$. Determine real numbers a_ε, b_ε, and c_ε such that the function Φ_ε defined by

$$\Phi_\varepsilon(t) = \begin{cases} 1/t & \text{if } t \geq 1, \\ a_\varepsilon t^2 + b_\varepsilon t + c_\varepsilon & \text{if } 1-\varepsilon \leq t < 1, \\ \Phi_\varepsilon(1-\varepsilon) & \text{if } 0 \leq t < 1-\varepsilon \end{cases}$$

is of class C^1 on $[0, +\infty)$.
Then show that the function Φ_ε is decreasing and that $\Phi_\varepsilon = 1+(\varepsilon/2)$ on $[0,\, 1-\varepsilon]$.

c. Put $E^1_\varepsilon(x) = \Phi_\varepsilon(r)$. Show that the function E^1_ε is of class C^1 on $\mathbb{R}^3$ and that the family of distributions $([E^1_\varepsilon])_{\varepsilon>0}$ tends to $[-4\pi E^1]$ in $\mathscr{D}'(\mathbb{R}^3)$.

d. Show that, for every $\varepsilon > 0$, $\Delta[E^1_\varepsilon]$ is a nonpositive-valued locally integrable function that vanishes on the complement of the set $\{x \in \mathbb{R}^3 : 1-\varepsilon \leq r \leq 1\}$. Deduce that there exists a positive Radon measure σ such that $\Delta[E^1] = \sigma$.

e. Show that σ is invariant under orthogonal transformations, that the support of σ is contained in the unit sphere S_1 in $\mathbb{R}^3$, and that $\int d\sigma = 1$. Deduce that $\sigma = \sigma_1/(4\pi)$.
Hint. Use Exercise 17 on page 83.

4. *Fundamental solution of Δ^k, for $k \in \mathbb{N}^*$.* We work in $\mathbb{R}^d$.

a. Show that, if $m \in \mathbb{N}^*$, $\alpha \in \mathbb{R}$, and $2m < \alpha + d$,

$$\Delta^m r^\alpha = \left(\prod_{j=0}^{m-1}(\alpha - 2j)\right)\left(\prod_{j=1}^{m}(\alpha + d - 2j)\right) r^{\alpha-2m}.$$

Deduce in particular that, if $k \geq 2$,

$$\Delta^{k-1} r^{2k-d} = \left(\prod_{j=2}^{k}(2j-d)\right) 2^{k-1}(k-1)!\, r^{2-d}.$$

b. Show that, if d is odd or $d > 2k$, there exists a constant C^k_d (which you should determine) such that

$$\Delta^k(C^k_d r^{2k-d}) = \delta.$$

c. Similarly, show that, if d is even and $d \leq 2k$, there exists a constant B^k_d such that

$$\Delta^k(B^k_d r^{2k-d} \log r) = \delta.$$

Hint. In the case $d = 2$ and $k \geq 1$,

$$\Delta^{k-1}(r^{2k-2}\log r) = 2^{2k-2}\big((k-1)!\big)^2 \log r + C_k.$$

In the case $d = 2d'$ with $d' \geq 2$ and $k \geq d'$,

$$\Delta^{k-(d'-1)}(r^{2k-d} \log r) = 2^{2k-d+1} \frac{(k-1)!}{(d'-2)!}(k-d')!\, r^{-2}$$

and

$$\Delta^{d'-1} r^{-2} = \frac{1}{C_d^{d'-1}}\, \delta.$$

It follows in each case that, if we put $d = 2d'$,

$$B_d^k = (s_d)^{-1} 2^{2-2k} \frac{(-1)^{d'-1}}{(k-1)!\,(k-d')!\,(d'-1)!}.$$

d. Deduce from the preceding calculations that, if $2k \geq d+1$, then Δ^k has a fundamental solution of class C^{2k-d-1}.

5. *Fundamental solution of* $\lambda + \Delta$ *in* $\mathbb{R}^3$, *for* $\lambda \in \mathbb{R}$. Denote by $x = (x_1, x_2, x_3)$ a generic point in $\mathbb{R}^3$ and, as usual, write $r = |x|$.

a. Take $\varphi \in C^2([0, +\infty))$ and set $f(x) = \varphi(r)/r$.

i. Show that, if $\varphi(0) = 0$, the derivatives $D_j f$ and $D_j^2 f$ in $\mathscr{D}'(\mathbb{R}^3)$, for $j \in \{1, 2, 3\}$, are locally integrable functions. Write them down in terms of φ, φ', and φ''. Write down Δf as well.

ii. Deduce an expression for Δf in the general case.

Hint. Write $f(x) = \dfrac{\varphi(r) - \varphi(0)}{r} + \varphi(0)\, \dfrac{1}{r}$.

b. Take $\lambda \in \mathbb{R}$. Determine the fundamental solutions of the operator $\Delta + \lambda$ having the form $E_\lambda(x) = \varphi(r)/r$. (Distinguish cases according to the sign of λ.)

c. Show that if $\lambda \leq 0$ there exists a unique fundamental solution E_λ such that $\lim_{|x| \to +\infty} E_\lambda(x) = 0$. Determine it. Show that this fundamental solution satisfies $E_\lambda(x) < 0$ for all $x \in (\mathbb{R}^d)^*$.

d. Show that E_λ does not have constant sign if $\lambda > 0$.

6. *Fundamental solution of the wave operator on* $\mathbb{R}^2$. Let E_1 be the distribution on $\mathbb{R}^2$ defined by the function

$$f(t, x) = \tfrac{1}{2}\, 1_{\{t > |x|\}}.$$

Show that

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right) E_1 = \delta.$$

7. *Fundamental solution of the wave operator on* $\mathbb{R}^4$. Denote by (t, x, y, z) a generic point in $\mathbb{R}^4$. If $r > 0$, denote by S_r the sphere in $\mathbb{R}^3$ of center 0 and radius r, and by σ_r its surface measure. For $\varphi \in \mathscr{D}(\mathbb{R}^4)$, write

$$\begin{aligned} \tilde{\varphi}(s, t) &= \frac{1}{4\pi s^2} \int_{S_s} \varphi(t, x, y, z)\, d\sigma_s(x, y, z) \\ &= \frac{1}{4\pi} \int_{S_1} \varphi(t, sx, sy, sz)\, d\sigma_1(x, y, z). \end{aligned}$$

Write $\Delta = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$.

a. Show that, for every $\varphi \in \mathscr{D}(\mathbb{R}^4)$,

$$\widetilde{\frac{\partial^2 \varphi}{\partial t^2}} = \frac{\partial^2 \tilde{\varphi}}{\partial t^2} \qquad \text{and} \qquad \widetilde{\Delta\varphi} = \frac{\partial^2 \tilde{\varphi}}{\partial s^2} + \frac{2}{s}\frac{\partial \tilde{\varphi}}{\partial s}.$$

Hint. For the second equality, you might use the expression of the Laplacian in spherical coordinates and Exercise 16b on page 83. Recall that, if we write

$$x = r\cos\theta\cos\varphi, \quad y = r\sin\theta\cos\varphi, \quad z = r\sin\varphi,$$

with $\theta \in (0, 2\pi)$ and $\varphi \in (-\pi/2, \pi/2)$, the Laplacian of a function $f(x, y, z) = F(r, \theta, \varphi)$ is given by

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial F}{\partial r}\right) + \frac{1}{r^2\cos\varphi}\frac{\partial}{\partial\varphi}\left(\cos\varphi\frac{\partial F}{\partial\varphi}\right) + \frac{1}{r^2\cos^2\varphi}\frac{\partial^2 F}{\partial\theta^2}.$$

b. Show that the relation

$$\langle E_3, \varphi\rangle = \int_0^{+\infty} t\tilde{\varphi}(t,t)\,dt = \int_{\mathbb{R}^3} \frac{\varphi(|u|, u)}{4\pi\,|u|}\,du \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^4)$$

defines a distribution E_3 on $\mathbb{R}^4$ (in fact, a positive Radon measure) and that

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right)E_3 = \delta.$$

Hint. If $v : (s,t) \mapsto v(s,t)$ is a function of class C^2 on $\mathbb{R}^2$, compute the derivative of the univariate function h defined by

$$h(t) = t\frac{\partial v}{\partial t}(t,t) - t\frac{\partial v}{\partial s}(t,t) - v(t,t).$$

c. Show that the support of E_3 equals the set

$$\{(t, x, y, z) \in \mathbb{R}^4 : t^2 = x^2 + y^2 + z^2 \text{ and } t \geq 0\}.$$

8. *Fundamental solution of the wave operator on* $\mathbb{R}^3$. Denote by (t, x, y) a generic point in $\mathbb{R}^3$. If $r > 0$, denote by S_r the sphere in $\mathbb{R}^3$ of center 0 and radius r, and by σ_r its surface measure.

a. Show that the relation

$$\langle E_2, \varphi\rangle = \int_0^{+\infty} \frac{1}{4\pi t}\left(\int_{S_t} \varphi(t, x, y)\,d\sigma_t(x, y, z)\right)dt \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^3)$$

defines a distribution on $\mathbb{R}^3$ and that

$$\left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}\right)E_2 = \delta.$$

Hint. Start by verifying that, for every compact subset K of $\mathbb{R}^3$, the set $\operatorname{Supp} E_3 \cap (K \times \mathbb{R})$ is compact. Then use Exercise 7.

b. Show that E_2 is given by the function

$$E_2(t,x,y) = \begin{cases} \dfrac{1}{2\pi}\dfrac{1}{\sqrt{t^2-x^2-y^2}} & \text{if } t > \sqrt{x^2+y^2}, \\ 0 & \text{otherwise.} \end{cases}$$

Hint. Show that

$$\langle E_2, \varphi\rangle = \frac{1}{2\pi}\int_{\{z\geq 0\}} \frac{\varphi\big(\sqrt{x^2+y^2+z^2},\, x,\, y\big)}{\sqrt{x^2+y^2+z^2}}\, dx\, dy\, dz$$

and set $z = \sqrt{t^2-x^2-y^2}$.

9

Convolution of Distributions

1 Tensor Product of Distributions

We start by proving two preliminary results, which are interesting in their own right. In the sequel, d and d' will denote integers greater than or equal to 1, while Ω and Ω' will denote open sets in $\mathbb{R}^d$ and $\mathbb{R}^{d'}$.

Theorem 1.1 (Differentiation inside the brackets) *Let $m \in \mathbb{N}$ and $r \in \mathbb{N}$. If $T \in \mathscr{D}'^m(\Omega)$ and $\varphi \in \mathscr{D}^{m+r}(\Omega \times \Omega')$, the map on Ω' defined by*

$$y \mapsto \langle T, \varphi(\cdot, y)\rangle \tag{$*$}$$

belongs to $\mathscr{D}^r(\Omega')$ and, for every multiindex $p \in \mathbb{N}^{d'}$ of length at most r,

$$\frac{\partial^{|p|}}{\partial y^p}\langle T, \varphi(\cdot, y)\rangle = \left\langle T, \frac{\partial^{|p|}}{\partial y^p}\varphi(\cdot, y)\right\rangle \tag{$**$}$$

for every $y \in \Omega'$.

If $T \in \mathscr{D}'(\Omega)$ and $\varphi \in \mathscr{D}(\Omega \times \Omega')$, the map defined in $(*)$ belongs to $\mathscr{D}(\Omega')$ and the relation $(**)$ is valid for all $p \in \mathbb{N}^{d'}$.

Proof. We carry out the proof in the case $T \in \mathscr{D}'^m(\Omega)$, $\varphi \in \mathscr{D}^{m+r}(\Omega \times \Omega')$. The other case is very similar.

Case $r = 0$. Take $T \in \mathscr{D}'^m(\Omega)$ and $\varphi \in \mathscr{D}^m(\Omega \times \Omega')$, and let K and K' be compact subsets of Ω and Ω', respectively, satisfying $\operatorname{Supp}\varphi \subset K \times K'$. Since, for every multiindex p of length at most m, the function

$(\partial^{|p|}\varphi)/(\partial x^p)$ is uniformly continuous (being continuous and having compact support), and since all the functions $\varphi(\,\cdot\,,y)$, with $y \in \Omega'$, are supported within the same compact K, we see that, if $(y_n)_{n\in\mathbb{N}}$ is a sequence in Ω' converging to $y \in \Omega'$, the sequence of functions $(\varphi(\,\cdot\,,y_n))_{n\in\mathbb{N}}$ converges to $\varphi(\,\cdot\,,y)$ in $\mathscr{D}^m(\Omega)$, so the sequence $\big(\langle T, \varphi(\,\cdot\,,y_n)\rangle\big)_{n\in\mathbb{N}}$ converges to $\langle T, \varphi(\,\cdot\,,y)\rangle$. We deduce that the map $y \mapsto \langle T, \varphi(\,\cdot\,,y)\rangle$ is continuous on Ω'. Since its support is compact (being contained in K'), this map does belong to $C_c(\Omega') = \mathscr{D}^0(\Omega')$.

Case $r = 1$. Take $T \in \mathscr{D}'^m(\Omega)$ and $\varphi \in \mathscr{D}^{m+1}(\Omega \times \Omega')$, and again let K and K' be compact subsets of Ω and Ω', respectively, satisfying $\operatorname{Supp}\varphi \subset K \times K'$. For $1 \le j \le d'$, let e_j be the j-th vector of the canonical basis of $\mathbb{R}^{d'}$. Take $y \in \Omega'$. If $x \in \Omega$ and $t \neq 0$, we have

$$\left|\frac{\varphi(x,\, y{+}te_j) - \varphi(x,y)}{t} - \frac{\partial\varphi}{\partial y_j}(x,y)\right| \le \sup_{t'\in[0,t]}\left|\frac{\partial\varphi}{\partial y_j}(x,\, y{+}t'e_j) - \frac{\partial\varphi}{\partial y_j}(x,y)\right|.$$

Using the fact that $\partial\varphi/\partial y_j$ is uniformly continuous, we easily deduce that the family of functions

$$\frac{\varphi(\,\cdot\,,\, y{+}te_j) - \varphi(\,\cdot\,,y)}{t}$$

converges in $\mathscr{D}^0(\Omega)$, as t tends to 0, to $(\partial\varphi(\,\cdot\,,y))/(\partial y_j)$. The reasoning we have used here for φ can be repeated without change for the partial derivatives $(\partial^{|p|}\varphi)/(\partial x^p)$, for $|p| \le m$; therefore

$$\frac{\varphi(\,\cdot\,,y + te_j) - \varphi(\,\cdot\,,y)}{t}$$

converges to $(\partial\varphi(\,\cdot\,,y))/(\partial y_j)$ in $\mathscr{D}^m(\Omega)$, as t tends to 0. It follows that

$$\frac{\big\langle T,\, \varphi(\,\cdot\,,\, y{+}te_j)\big\rangle - \big\langle T,\, \varphi(\,\cdot\,,y)\big\rangle}{t}$$

has the limit $\big\langle T,\, (\partial\varphi(\,\cdot\,,y))/(\partial y_j)\big\rangle$ as t tends to 0; that is, the partial derivative

$$\frac{\partial}{\partial y_j}\big\langle T,\, \varphi(\,\cdot\,,y)\big\rangle$$

exists and satisfies

$$\frac{\partial}{\partial y_j}\big\langle T,\, \varphi(\,\cdot\,,y)\big\rangle = \left\langle T,\, \frac{\partial}{\partial y_j}\varphi(\,\cdot\,,y)\right\rangle;$$

moreover this is the case for every $y \in \Omega'$ and every $j \in \{1,\dots,d'\}$. Since the maps $y \mapsto \big\langle T,\, (\partial\varphi(\,\cdot\,,y))/(\partial y_j)\big\rangle$ are continuous on Ω' (by the case $r = 0$), this shows also that $y \mapsto \big\langle T,\, \varphi(\,\cdot\,,y)\big\rangle$ belongs to $\mathscr{D}^1(\Omega')$, which concludes the proof in the case $r = 1$.

The general case follows from the two preceding ones by induction. □

Theorem 1.2 *The vector space* $\mathscr{D}(\Omega) \otimes \mathscr{D}(\Omega')$ *spanned by the functions*

$$f \otimes g : (x, y) \mapsto f(x)g(y),$$

with $f \in \mathscr{D}(\Omega)$ *and* $g \in \mathscr{D}(\Omega')$, *is dense in* $\mathscr{D}(\Omega \times \Omega')$.

Proof. We use a lemma that allows us to approximate the convolution by means of a "discrete convolution":

Lemma 1.3 *Suppose* $\varphi, \psi \in \mathscr{D}(\mathbb{R}^n)$. *For* $\varepsilon > 0$ *and* $x \in \mathbb{R}^n$, *set*

$$g_\varepsilon(x) = \varepsilon^n \sum_{\nu \in \mathbb{Z}^n} \varphi(x - \varepsilon\nu)\psi(\varepsilon\nu).$$

Then $g_\varepsilon \in \mathscr{D}(\mathbb{R}^n)$, $\operatorname{Supp} g_\varepsilon \subset \operatorname{Supp} \varphi + \operatorname{Supp} \psi$, *and*

$$\lim_{\varepsilon \to 0} g_\varepsilon = \varphi * \psi$$

in $\mathscr{D}(\mathbb{R}^n)$.

Proof. The function g_ε is defined by a finite sum whose number of terms depends only on ε (since ψ has compact support). Since each of these terms is an element of $\mathscr{D}(\mathbb{R}^n)$ and is supported within $\operatorname{Supp} \varphi + \operatorname{Supp} \psi$, the same holds for g_ε. At the same time, for every $p \in \mathbb{N}^d$,

$$D^p g_\varepsilon(x) = \varepsilon^n \sum_{\nu \in \mathbb{Z}^n} D^p \varphi(x - \varepsilon\nu)\psi(\varepsilon\nu).$$

Thus the result will be proved if we show that g_ε converges uniformly to $\varphi * \psi$ (for then we will be able to apply the same result to $D^p\varphi$ and ψ instead of φ and ψ).

Denote by $\|\cdot\|$ the uniform norm on $\mathbb{R}^n$ and set $N = \max_{x \in \operatorname{Supp} \psi} \|x\|$. By the Mean Value Theorem, there exists $C > 0$ such that, for every $x, y, y' \in \mathbb{R}^n$,

$$\left|\varphi(x - y)\psi(y) - \varphi(x - y')\psi(y')\right| \le C\|y - y'\|.$$

For $\nu \in \mathbb{Z}^n$, set

$$Q_\nu^\varepsilon = \prod_{j=1}^{n} \left[\nu_j\varepsilon,\ (\nu_j + 1)\varepsilon\right).$$

Then

$$\varphi * \psi(x) = \sum_{\|\nu\| \le (N/\varepsilon)+1} \int_{Q_\nu^\varepsilon} \varphi(x - y)\psi(y)\, dy,$$

so that

$$\begin{aligned}\left|\varphi*\psi(x)-g_\varepsilon(x)\right| &\le \sum_{\|\nu\|\le(N/\varepsilon)+1}\int_{Q_\nu^\varepsilon}\left|\varphi(x-y)\psi(y)-\varphi(x-\nu\varepsilon)\psi(\nu\varepsilon)\right|dy\\ &\le C\varepsilon^{n+1}\left(2\left(\frac{N}{\varepsilon}+1\right)+1\right)^n \le C'\varepsilon \qquad (\text{for } \varepsilon<1),\end{aligned}$$

proving the result. □

Now consider two smoothing sequences, $(\chi_n)_{n\in\mathbb{N}}$ and $(\tilde{\chi}_n)_{n\in\mathbb{N}}$, on $\mathbb{R}^d$ and $\mathbb{R}^{d'}$, respectively. Clearly, $(\chi_n\otimes\tilde{\chi}_n)_{n\in\mathbb{N}}$ is a smoothing sequence on $\mathbb{R}^d\times\mathbb{R}^{d'}$. Take $\varphi\in\mathscr{D}(\Omega\times\Omega')$. Then there exist compact sets K and K_1 in Ω and compact sets K' and K_1' in Ω' such that $\operatorname{Supp}\varphi\subset K\times K'$ and $K\subset\mathring{K}_1$, $K'\subset\mathring{K}_1'$. By Proposition 1.2 on page 261, φ can be approximated arbitrarily close, in the metric space $\mathscr{D}_{K_1\times K_1'}(\Omega\times\Omega')$, by some function $\varphi*(\chi_n\otimes\tilde{\chi}_n)$, with n so large that $K+\operatorname{Supp}\chi_n\subset K_1$ and $K'+\operatorname{Supp}\tilde{\chi}_n\subset K_1'$ (where, as usual, we identify φ with an element of $\mathscr{D}(\mathbb{R}^d\times\mathbb{R}^{d'})$ by giving it the value 0 outside $\Omega\times\Omega'$). By the lemma, $\varphi*(\chi_n\otimes\tilde{\chi}_n)$ can in turn be approximated arbitrarily close, in the space $\mathscr{D}_{K_1\times K_1'}(\Omega\times\Omega')$, by a function of the form

$$\varepsilon^{d+d'}\sum_{\nu\in\mathbb{Z}^d,\,\tilde{\nu}\in\mathbb{Z}^{d'}}\chi_n(x-\varepsilon\nu)\tilde{\chi}_n(y-\varepsilon\tilde{\nu})\varphi(\varepsilon\nu,\varepsilon\tilde{\nu}),$$

which lies in $\mathscr{D}_{K_1}(\Omega)\otimes\mathscr{D}_{K_1'}(\Omega')$. The result follows. □

By the same method or by induction, one shows that, if Ω_j is open in $\mathbb{R}^{d_j}$ for each $j\in\{1,\dots,r\}$, then $\mathscr{D}(\Omega_1)\otimes\cdots\otimes\mathscr{D}(\Omega_r)$ is dense in $\mathscr{D}(\Omega_1\times\cdots\times\Omega_r)$.

In what follows x will denote a generic point of $\mathbb{R}^d$ and y a generic point of $\mathbb{R}^{d'}$. If T is a distribution on Ω and if $\varphi\in\mathscr{D}(\Omega)$, we write, if there is a risk of confusion in the space under consideration (Ω or Ω'),

$$\langle T,\varphi\rangle=\big\langle T_x,\,\varphi(x)\big\rangle.$$

Likewise, if S is a distribution on Ω' and if $\psi\in\mathscr{D}(\Omega')$, we write $\langle S,\psi\rangle=\big\langle S_y,\,\psi(y)\big\rangle$.

Proposition 1.4 *Suppose $T\in\mathscr{D}'(\Omega)$ and $S\in\mathscr{D}'(\Omega')$. There exists a unique distribution on $\Omega\times\Omega'$, denoted $T\otimes S$ and called the **tensor product** of T and S, such that*

$$\big\langle T\otimes S,\,\varphi\otimes\psi\big\rangle=\langle T,\varphi\rangle\,\langle S,\psi\rangle$$

for all $\varphi\in\mathscr{D}(\Omega)$ and $\psi\in\mathscr{D}(\Omega')$. Moreover, for every $\varphi\in\mathscr{D}(\Omega\times\Omega')$,

$$\big\langle T\otimes S,\,\varphi\big\rangle=\big\langle T_x,\,\langle S_y,\,\varphi(x,y)\rangle\big\rangle=\big\langle S_y,\,\langle T_x,\,\varphi(x,y)\rangle\big\rangle.$$

Proof. Uniqueness follows immediately from Theorem 1.2. For existence, consider the linear map on $\mathscr{D}(\Omega \times \Omega')$ defined by

$$\varphi \mapsto \big\langle S_y,\ \langle T_x,\ \varphi(x,y)\rangle\big\rangle. \tag{$*$}$$

This map is well defined, by Theorem 1.1. Let K_1 be a compact subset of $\Omega \times \Omega'$, and let K and K' be compact subsets in Ω and Ω', respectively, such that $K_1 \subset K \times K'$. Take $m, m' \in \mathbb{N}$ and $C, C' > 0$ such that

$$\big|\langle T, \varphi\rangle\big| \le C\|\varphi\|^{(m)} \quad \text{for all } \varphi \in \mathscr{D}_K(\Omega)$$

and

$$\big|\langle S, \varphi\rangle\big| \le C'\|\varphi\|^{(m')} \quad \text{for all } \varphi \in \mathscr{D}_{K'}(\Omega')$$

(see Proposition 2.1 on page 268). Then, again by Theorem 1.1, there exists a constant $C'' \ge 0$ such that

$$\big|\big\langle S_y,\ \langle T_x, \varphi(x,y)\rangle\big\rangle\big| \le C''\|\varphi\|^{(m+m')} \quad \text{for all } \varphi \in \mathscr{D}_{K_1}(\Omega \times \Omega').$$

Thus, the linear map defined in $(*)$ is indeed a distribution on $\Omega \times \Omega'$ satisfying the indicated condition, namely

$$\big\langle S_y,\ \langle T_x, \varphi(x)\psi(y)\rangle\big\rangle = \langle T,\ \varphi\rangle\,\langle S, \psi\rangle$$

for all $\varphi \in \mathscr{D}(\Omega)$ and $\psi \in \mathscr{D}(\Omega')$. One argues likewise for the expression $\big\langle T_x,\ \langle S_y, \varphi(x,y)\rangle\big\rangle$, interchanging the roles of x and y. □

We see simply that, if f and g are locally integrable functions on Ω and Ω', respectively, then $[f] \otimes [g] = [f \otimes g]$. Similarly, the tensor product in the sense of distributions of two complex Radon measures equals their tensor product in the sense of measures. All of this follows from Fubini's Theorem.

From the definition we see also that, if T and S are distributions on Ω and if $\varphi \in \mathscr{D}(\Omega \times \Omega)$ is such that $\varphi(x,y) = \varphi(y,x)$ for every $(x,y) \in \Omega \times \Omega$, then $\langle T \otimes S,\ \varphi\rangle = \langle S \otimes T,\ \varphi\rangle$.

Proposition 1.5 *Suppose $T \in \mathscr{D}'(\Omega)$ and $S \in \mathscr{D}'(\Omega')$. Then:*

i. $\operatorname{Supp}(T \otimes S) = (\operatorname{Supp} T) \times (\operatorname{Supp} S)$.
ii. *For any $p \in \mathbb{N}^d$ and $q \in \mathbb{N}^{d'}$,*

$$\partial_x^p \partial_y^q (T \otimes S) = (\partial_x^p T) \otimes (\partial_y^q S).$$

Proof. If φ is supported within $(\Omega \setminus \operatorname{Supp} T) \times \Omega'$, the support of $\varphi(\,\cdot\,, y)$, for every $y \in \Omega'$, is contained in $\Omega \setminus \operatorname{Supp} T$. Therefore

$$\langle T \otimes S,\ \varphi\rangle = \big\langle S_y,\ \langle T_x, \varphi(x,y)\rangle\big\rangle = 0.$$

It follows that the support of $T \otimes S$ is contained in $\operatorname{Supp} T \times \Omega'$; similarly, it is contained in $\Omega \times \operatorname{Supp} S$, and so also in the intersection of these two sets, which is $\operatorname{Supp} T \times \operatorname{Supp} S$.

Conversely, if $(x, y) \in \operatorname{Supp} T \times \operatorname{Supp} S$ and if $(x, y) \notin \operatorname{Supp}(T \otimes S)$, let O denote the complement of the support of $T \otimes S$ in $\Omega \times \Omega'$. Then there exist open sets O_1 and O_2 containing x and y, respectively, and such that $O \supset O_1 \times O_2$. By the definition of x and y, there exist $\varphi \in \mathscr{D}(O_1)$ and $\psi \in \mathscr{D}(O_2)$ such that $\langle T, \varphi \rangle \neq 0$ and $\langle S, \psi \rangle \neq 0$. But then $\varphi \otimes \psi \in \mathscr{D}(O)$ and $\langle T \otimes S,\, \varphi \otimes \psi \rangle \neq 0$, which contradicts the definition of O. Therefore $\operatorname{Supp} T \times \operatorname{Supp} S \subset \operatorname{Supp}(T \otimes S)$, and the first assertion of the theorem is proved.

Next, if $\varphi \in \mathscr{D}(\Omega)$ and $\psi \in \mathscr{D}(\Omega')$,

$$\begin{aligned}\big\langle \partial_x^p \partial_y^q (T \otimes S),\, \varphi \otimes \psi \big\rangle &= (-1)^{|p|+|q|} \big\langle T \otimes S,\, (\partial_x^p \varphi) \otimes (\partial_y^q \psi) \big\rangle \\ &= (-1)^{|p|+|q|} \langle T,\, \partial_x^p \varphi \rangle \langle S, \partial_y^q \psi \rangle \\ &= \big\langle (\partial_x^p T) \otimes (\partial_y^q S),\, \varphi \otimes \psi \big\rangle.\end{aligned}$$

Now just apply the denseness theorem (Theorem 1.2) to obtain the second part of the theorem. □

One can, in a completely analogous way, define the tensor product of finitely many distributions. The tensor product thus constructed is associative.

Exercises

1. Suppose $T \in \mathscr{D}'^m(\Omega)$ and $S \in \mathscr{D}'^n(\Omega')$. Show that

$$T \otimes S \in \mathscr{D}'^{m+n}(\Omega \times \Omega'),$$

and that in this situation the formulas in Proposition 1.4 are valid for every $\varphi \in \mathscr{D}^{m+n}(\Omega \times \Omega')$.

2. Show that, if T is a distribution on Ω, the map $S \mapsto T \otimes S$ from $\mathscr{D}'(\Omega')$ to $\mathscr{D}'(\Omega \times \Omega')$ is continuous (in the sense of sequences). Show also that, if S is a distribution on Ω', the map $T \mapsto T \otimes S$ from $\mathscr{D}'(\Omega)$ to $\mathscr{D}'(\Omega \times \Omega')$ is continuous (in the sense of sequences).

3. *Homogeneous distributions.* Let Ω be an open set in $\mathbb{R}^d$ such that

$$\lambda \Omega \subset \Omega \quad \text{for all } \lambda > 0.$$

If $T \in \mathscr{D}'(\Omega)$ and $\lambda > 0$, define a distribution T_λ on Ω by

$$\langle T_\lambda, \varphi \rangle = \lambda^{-d} \big\langle T,\, \varphi(\cdot / \lambda) \big\rangle \quad \text{for all } \varphi \in \mathscr{D}(\Omega).$$

a. Determine T_λ if $T \in L^1_{\text{loc}}(\Omega)$.

b. A distribution T on Ω is said to be *homogeneous* of degree $\alpha \in \mathbb{R}$ if

$$T_\lambda = \lambda^\alpha T \quad \text{for all } \lambda > 0.$$

Show that a distribution T on Ω is homogeneous of degree α if and only if it satisfies *Euler's equation* in $\mathscr{D}'(\Omega)$:

$$\sum_{j=1}^{d} x_j D_j T = \alpha T.$$

Hint. You might use Theorem 1.1 to compute the derivative of the function $\lambda \mapsto \langle \lambda^{-\alpha} T_\lambda, \varphi \rangle$.

c. Show that the only homogeneous distributions on $\mathbb{R}$ having support $\{0\}$ are those of the form $\lambda\delta^{(k)}$, with $\lambda \in \mathbb{C}^*$ and $k \in \mathbb{N}$. Determine their degrees.
Hint. Use Exercise 4 on page 285.

d. Show that $\mathrm{pv}(1/x)$ is a homogeneous distribution on $\mathbb{R}$ and find its degree. What about $\mathrm{fp}(Y(x)/x)$?

e. Determine all homogeneous distributions of degree 0 on $\mathbb{R}$.

f. Let T be a homogeneous distribution of degree α on Ω and S a homogeneous distribution of degree β on Ω'. Show that $T \otimes S$ is a homogeneous distribution of degree $\alpha + \beta$ on $\Omega \times \Omega'$.

g. Show that the distribution $(x^2 Y(x)) \otimes \delta'$ on $\mathbb{R}^2$ is homogeneous; find its degree and order.

4. a. If $J \subseteq \{1, 2, \dots, d\}$, denote by Y^J the distribution defined by

$$Y^J = Y_1^J \otimes Y_2^J \otimes \cdots \otimes Y_d^J,$$

where $Y_i^J = Y$ (the Heaviside function) if $i \in J$ and $Y_i^J = \delta$ otherwise. What differential operator is Y^J a fundamental solution of?

b. Compute the p-th derivative of the function f defined by $f(x) = x^p Y(x)$. Deduce a fundamental solution of the one-variable differential operator D^p, for $p \in \mathbb{N}$.

c. Determine a fundamental solution E of the d-variable differential operator D^p, where $p = (p_1, \dots, p_d) \in \mathbb{N}^d$.
If $p_1 = p_2 = \cdots = p_d = k$ with $k \geq 2$, prove that D^p has a fundamental solution of class C^{k-2} in $\mathbb{R}^d$.

5. Show that the following relation defines a distribution T on $\mathbb{R}^2$:

$$\langle T, \varphi \rangle = \lim_{\varepsilon \to 0} \iint_{\{|x|, |y| > \varepsilon\}} \frac{\varphi(x,y)}{xy}\, dx\, dy \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^2).$$

Show that $T = \mathrm{pv}(1/x) \otimes \mathrm{pv}(1/y)$. What is the order of T?
Hint. Introduce $\varphi(x,y) - \varphi(x,0) - \varphi(0,y) + \varphi(0,0)$.

6. *Distributions that do not depend on a certain variable.* (See also Exercise 2 on page 302.) Let T be a distribution on $\mathbb{R}^d$, where $d \geq 2$.

a. Suppose that $(\partial T)/(\partial x_1) = 0$.

i. Show that, for every $\psi \in \mathscr{D}(\mathbb{R}^{d-1})$, there exists a constant $S(\psi)$ such that

$$\langle T,\, \varphi \otimes \psi \rangle = S(\psi) \int_{\mathbb{R}} \varphi(x)\, dx \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}).$$

Hint. Fix ψ and prove that the linear form U defined on $\mathscr{D}(\mathbb{R})$ by

$$\langle U, \varphi \rangle = \langle T,\, \varphi \otimes \psi \rangle \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R})$$

is a distribution and that $U' = 0$.

ii. Show that the map $\psi \mapsto S(\psi)$ is a distribution on $\mathbb{R}^{d-1}$ and that

$$T = 1 \otimes S.$$

Hint. Take $\chi \in \mathscr{D}(\mathbb{R})$ such that $\int \chi\, dx = 1$. Then

$$S(\psi) = \langle T,\, \chi \otimes \psi \rangle.$$

b. Show that, conversely, if there exists $S \in \mathscr{D}'(\mathbb{R}^{d-1})$ such that $T = 1 \otimes S$, then $(\partial T)/(\partial x_1) = 0$.

7. Let T be a distribution on $\mathbb{R}^d$, where $d \geq 2$. Show that $x_1 T = 0$ if and only if there exists $S \in \mathscr{D}'(\mathbb{R}^{d-1})$ such that $T = \delta \otimes S$.
Hint. Argue as in Exercise 6 and use Proposition 1.4 on page 289.

2 Convolution of Distributions

2A Convolution in $\mathscr{E}'$

We define first the convolution product of distributions with compact support on $\mathbb{R}^d$.

Let T and S be elements of $\mathscr{E}'(\mathbb{R}^d)$. We know from Proposition 1.5 that $T \otimes S$ is a distribution on $\mathbb{R}^d \times \mathbb{R}^d$ with a compact support that coincides with $\operatorname{Supp} T \times \operatorname{Supp} S$. On the other hand, if $\varphi \in \mathscr{D}(\mathbb{R}^d)$, the function defined on $\mathbb{R}^d \times \mathbb{R}^d$ by $(x, y) \mapsto \varphi(x+y)$ belongs to $\mathscr{E}(\mathbb{R}^d \times \mathbb{R}^d)$. Proposition 3.3 on page 282 then says that the bracket $\langle T_x \otimes S_y,\, \varphi(x+y) \rangle$ is well-defined. Moreover, the map from $\mathscr{D}(\mathbb{R}^d)$ to $\mathscr{E}(\mathbb{R}^d \times \mathbb{R}^d)$ that takes $\varphi \in \mathscr{D}(\mathbb{R}^d)$ to $(x, y) \mapsto \varphi(x+y)$ is clearly continuous. This leads to the following definition:

Definition 2.1 If $T, S \in \mathscr{E}'(\mathbb{R}^d)$, the **convolution** of T and S is the distribution $T * S$ defined by

$$\langle T * S,\, \varphi \rangle = \big\langle T_x \otimes S_y,\, \varphi(x+y) \big\rangle \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^d).$$

Proposition 2.2 *If $T, S \in \mathscr{E}'(\mathbb{R}^d)$, then $T * S \in \mathscr{E}'(\mathbb{R}^d)$ and*

$$\operatorname{Supp}(T * S) \subset \operatorname{Supp} T + \operatorname{Supp} S.$$

The convolution product is a commutative and associative binary operation in $\mathscr{E}'(\mathbb{R}^d)$, having δ (the Dirac measure at 0) as a unity element. In other words, the convolution product makes the space $\mathscr{E}'(\mathbb{R}^d)$ into a commutative algebra with unity.

Proof. Let $\varphi \in \mathscr{D}(\mathbb{R}^d)$. If $x \notin \operatorname{Supp} \varphi - \operatorname{Supp} S$, then

$$\operatorname{Supp} \varphi(x + \cdot\,) \cap \operatorname{Supp} S = (\operatorname{Supp} \varphi - x) \cap \operatorname{Supp} S = \varnothing,$$

so $\langle S_y, \varphi(x + y)\rangle = 0$. Thus, $\operatorname{Supp} \langle S_y, \varphi(\,\cdot + y)\rangle \subset \operatorname{Supp} \varphi - \operatorname{Supp} S$. It follows that, if the support of T does not intersect $\operatorname{Supp} \varphi - \operatorname{Supp} S$, we have $\langle T * S, \varphi\rangle = 0$ and therefore $\operatorname{Supp}(T * S) \subset \operatorname{Supp} T + \operatorname{Supp} S$. The rest of the proposition follows immediately from the results proved in Section 1. $\square$

As a consequence of Proposition 1.5, we have the following fundamental property:

Proposition 2.3 *If $T, S \in \mathscr{E}'(\mathbb{R}^d)$ and $j \in \{1, \ldots, d\}$, then*

$$D_j(T * S) = (D_j T) * S = T * (D_j S).$$

Obviously, this result extends to every differential operator $P(D)$, of any order: if $T, S \in \mathscr{E}'(\mathbb{R}^d)$, then

$$P(D)(T * S) = (P(D)T) * S = T * (P(D)S)$$

for every polynomial P with complex coefficients.

2B Convolution in $\mathscr{D}'$

One cannot hope to define a convolution product on all of $\mathscr{D}'$ that extends the convolution product of functions, because, in general, two locally integrable functions are not convolvable: for example, $1 * 1$ has no meaning. We will define the convolution product in $\mathscr{D}'$ in case the supports satisfy a condition that we now introduce.

Definition 2.4 We say that a family of closed subsets $F_1, \ldots, F_n$ of $\mathbb{R}^d$ satisfies condition (C) if, for every compact subset K of $\mathbb{R}^d$, the set

$$\{(x^1, \ldots, x^n) \in F_1 \times \cdots \times F_n \,:\, x^1 + \cdots + x^n \in K\}$$

is a compact subset of $(\mathbb{R}^d)^n$.

Obviously, we could have written this condition with "bounded" instead of "compact".

Let's first give some examples and simple properties. Most of the proofs are left as exercises.

1. Suppose $(F_1, \ldots, F_n)$ is a family of nonempty closed sets that satisfies condition (C). Then every family of closed sets $(\tilde{F}_1, \ldots, \tilde{F}_p)$, where $1 \le p \le n$ and $\tilde{F}_j \subset F_j$ for all $j \in \{1, \ldots, p\}$, also satisfies (C).
2. Clearly, every family of compact subsets satisfies condition (C).
3. If $(F_1, \ldots, F_n)$ satisfies condition (C), so does the family $(F_1, \ldots, F_n, L)$, for every compact L in $\mathbb{R}^d$. Indeed, if K is a compact subset of $\mathbb{R}^d$,

$$\{(x^1, \ldots, x^n, x^{n+1}) \in F_1 \times \cdots \times F_n \times L \,:\, x^1 + \cdots + x^n + x^{n+1} \in K\}$$
$$\subset \{(x^1, \ldots, x^n) \in F_1 \times \cdots \times F_n \,:\, x^1 + \cdots + x^n \in K - L\} \times L,$$

   and the set $K - L$ is compact.
   It follows by induction that a family of closed sets all or all but one of which are compact satisfies property (C).
4. Let F be a closed subset of $\mathbb{R}^d$ containing a one-dimensional subspace $\mathbb{R}u$ of $\mathbb{R}^d$, where $u \neq 0$. Then the family (F, F) does not satisfy condition (C). Indeed, the set

$$\{(x^1, x^2) \in \mathbb{R}u \times \mathbb{R}u \,:\, x^1 + x^2 = 0\} = \{(tu, -tu) : t \in \mathbb{R}\}$$

   is unbounded.
5. If $a, b \in \mathbb{R}$, the family $((-\infty, a],\ [b, +\infty))$ does not satisfy condition (C). By Example 1 and because $\mathbb{R} \supset (-\infty, 0]$, neither does the pair $(\mathbb{R}, \mathbb{R}^+)$. By contrast, for every $a_1, \ldots, a_n \in \mathbb{R}$, the family

$$\big([a_1, +\infty),\ \ldots,\ [a_n, +\infty)\big)$$

   satisfies (C). In particular, $(\mathbb{R}^+, \ldots, \mathbb{R}^+)$ satisfies (C). For a generalization to dimension d, see Exercise 4 on page 335.
6. If $(F_1, \ldots, F_n)$ satisfies condition (C), the set $F_1 + \cdots + F_n$ is closed. (Recall that, in general, the sum $F_1 + F_2$ of closed sets F_1 and F_2 need not be closed.)
7. If $(F_1, \ldots, F_n)$ satisfies condition (C) and if (I, J) is a partition of the set $\{1, \ldots, n\}$ (that is, $I \cap J = \varnothing$ and $I \cup J = \{1, \ldots, n\}$), then the family (F_I, F_J) satisfies (C), with $F_I = \sum_{k \in I} F_k$ and $F_J = \sum_{k \in J} F_k$.

The next step in the construction consists in extending the bracket. If $\varphi \in \mathscr{E}(\mathbb{R}^d)$, the expression $\langle T, \varphi \rangle$ has so far been defined only when T is a distribution with compact support on $\mathbb{R}^d$ (see Proposition 3.3 on page 282). The next proposition allows us to extend this definition to the case where $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$ is compact.

Proposition 2.5 *Let Ω be open in $\mathbb{R}^d$. Let $T \in \mathscr{D}'(\Omega)$ and $\varphi \in \mathscr{E}(\Omega)$ be such that* $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$ *is compact. Then, if $\rho \in \mathscr{D}(\Omega)$ is a function taking the value* 1 *on an open set containing* $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$*, the value of $\langle T, \rho\varphi \rangle$ does not depend on ρ.*

This value is denoted by $\langle T, \varphi \rangle$.

Proof. Take $\rho \in \mathscr{D}(\Omega)$ such that $\rho = 0$ on an open set containing $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$. Then the support of ρ is contained in the complement of $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$, and therefore

$$\operatorname{Supp} \rho\varphi \subset \operatorname{Supp} \varphi \cap (\mathbb{R}^d \setminus (\operatorname{Supp} T \cap \operatorname{Supp} \varphi)) = \operatorname{Supp} \varphi \cap (\mathbb{R}^d \setminus \operatorname{Supp} T),$$

which implies that $\langle T, \rho\varphi \rangle = 0$.

Consequently, if ρ and $\tilde{\rho}$ are functions in $\mathscr{D}(\Omega)$ that coincide on an open set containing $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$, we have $\langle T, \rho\varphi \rangle = \langle T, \tilde{\rho}\varphi \rangle$. □

Naturally, if $T \in \mathscr{E}'(\Omega)$ and $\varphi \in \mathscr{E}(\Omega)$, we recover the meaning of the brackets defined in Proposition 3.3 on page 282. If $T \in \mathscr{D}'(\Omega)$ and $\varphi \in \mathscr{D}(\Omega)$, we recover the usual meaning of the brackets.

Note that we can define similarly the value of $\langle T, \varphi \rangle$ for $T \in \mathscr{D}'^m(\Omega)$ and $\varphi \in \mathscr{E}^m(\Omega)$ if $\operatorname{Supp} T \cap \operatorname{Supp} \varphi$ is compact.

We can now define the convolution product of a family of distributions whose supports satisfy condition (C) of Definition 2.4. *We will say from now on that such a family of distributions itself satisfies condition* (C).

Proposition 2.6 *Let $(T_1, \dots, T_n)$ be a family of distributions on $\mathbb{R}^d$ satisfying condition* (C).

1. *If $\varphi \in \mathscr{D}(\mathbb{R}^d)$, we define a function $\hat{\varphi}$ on $(\mathbb{R}^d)^n$ by*

$$\hat{\varphi}(x^1, \dots, x^n) = \varphi(x^1 + \cdots + x^n).$$

   Then $\hat{\varphi} \in \mathscr{E}((\mathbb{R}^d)^n)$ and $\operatorname{Supp}(T_1 \otimes \cdots \otimes T_n) \cap \operatorname{Supp} \hat{\varphi}$ *is compact. The map defined on $\mathscr{D}(\mathbb{R}^d)$ by*

$$\varphi \mapsto \langle T_1 \otimes \cdots \otimes T_n, \hat{\varphi} \rangle$$

   is a distribution on $\mathbb{R}^d$, denoted $T_1 * \cdots * T_n$ and called the ***convolution*** *of $T_1, \dots, T_n$.*
2. *For each $l > 0$, let $\rho_l \in \mathscr{D}(\mathbb{R}^d)$ be such that $\rho_l = 1$ on $\bar{B}(0, l)$. For every open bounded set Ω in $\mathbb{R}^d$, there exists a real number $l > 0$ such that the restrictions of $T_1 * \cdots * T_n$ and of $(\rho_{l'} T_1) * \cdots * (\rho_{l'} T_n)$ to Ω coincide for every $l' \geq l$. In particular,*

$$T_1 * \cdots * T_n = \lim_{l \to +\infty} (\rho_l T_1) * \cdots * (\rho_l T_n)$$

   in $\mathscr{D}'(\mathbb{R}^d)$.

In the preceding statement we have $\rho_l T_j \in \mathscr{E}'(\mathbb{R}^d)$, so the convolution $(\rho_l T_1) * \cdots * (\rho_l T_n)$ is defined in the sense of Section 2A. Indeed, the preceding definition coincides with Definition 2.1 when all distributions have compact support.

Proof. Let Ω be a bounded open set in $\mathbb{R}^d$ and φ an element of $\mathscr{D}(\Omega)$. We know that $\operatorname{Supp}(T_1 \otimes \cdots \otimes T_n) = \operatorname{Supp} T_1 \times \cdots \times \operatorname{Supp} T_n$, so

$$\operatorname{Supp}(T_1 \otimes \cdots \otimes T_n) \cap \operatorname{Supp} \hat{\varphi} \\ \subset \{(x^1, \ldots, x^n) \in \operatorname{Supp} T_1 \times \cdots \times \operatorname{Supp} T_n \,:\, x^1 + \cdots + x^n \in \bar{\Omega}\}.$$

By condition (C), we deduce that $\operatorname{Supp}(T_1 \otimes \cdots \otimes T_n) \cap \operatorname{Supp} \hat{\varphi}$ is a compact subset of $(\mathbb{R}^d)^n$ contained in a compact K_Ω that depends only on Ω, not on φ. Thus, $\langle T_1 \otimes \cdots \otimes T_n, \hat{\varphi}\rangle$ is well defined and coincides with $\big\langle T_1 \otimes \cdots \otimes T_n, (\rho_l \otimes \cdots \otimes \rho_l)\hat{\varphi}\big\rangle$ if $K_\Omega \subset (B(0,l))^n$. Now,

$$\begin{aligned} \langle T_1 \otimes \cdots \otimes T_n,\, (\rho_l \otimes \cdots \otimes \rho_l)\hat{\varphi}\rangle &= \langle \rho_l T_1 \otimes \cdots \otimes \rho_l T_n,\, \hat{\varphi}\rangle \\ &= \langle (\rho_l T_1) * \cdots * (\rho_l T_n),\, \varphi\rangle. \end{aligned}$$

This shows that $T_1 * \cdots * T_n$ is a distribution, and proves the second part of the proposition as well. □

We now state the essential properties of the convolution product in $\mathscr{D}'(\mathbb{R}^d)$.

Proposition 2.7 **1.** *If* (T, S) *satisfies condition* (C), *then* $T * S = S * T$.
2. *If* $(T_1, \ldots, T_n)$ *satisfies* (C), *then*

$$\operatorname{Supp}(T_1 * \cdots * T_n) \subset \operatorname{Supp} T_1 + \cdots + \operatorname{Supp} T_n.$$

3. $\delta * T = T * \delta$ *for all* $T \in \mathscr{D}'(\mathbb{R}^d)$.

Proof. The second part of Proposition 2.6 allows us, by passing to the limit, to reduce the problem to the case of distributions with compact support, for which these properties were stated in Proposition 2.2. The reasoning is straightforward for the proof of parts 1 and 3. We spell it out for part 2.

If $(T_1, \ldots, T_n)$ satisfies (C), then, by property 6 on page 326, the set $F = \operatorname{Supp} T_1 + \cdots + \operatorname{Supp} T_n$ is closed. On the other hand, if $l > 0$, we have $\operatorname{Supp}(\rho_l T_j) \subset \operatorname{Supp} T_j$ for every $j \in \{1, \ldots, d\}$ (in the notation of Proposition 2.6); thus, by Proposition 2.2, $\operatorname{Supp}((\rho_l T_1) * \cdots * (\rho_l T_n)) \subset F$. We deduce that, for every $\varphi \in \mathscr{D}(\mathbb{R}^d)$ satisfying $\operatorname{Supp}\varphi \subset \mathbb{R}^d \setminus F$, Proposition 2.6 yields

$$\langle T_1 * \cdots * T_n,\, \varphi\rangle = \lim_{l \to +\infty} \big\langle (\rho_l T_1) * \cdots * (\rho_l T_n),\, \varphi\big\rangle = 0.$$

Therefore $\mathbb{R}^d \setminus F$ is a domain of nullity of $T_1 * \cdots * T_n$, which proves part 2 of the proposition. □

Proposition 2.8 (Continuity) *Let* $(T_n)_{n \in \mathbb{N}}$ *be a sequence in* $\mathscr{D}'(\mathbb{R}^d)$, *and let* T, S *belong to* $\mathscr{D}'(\mathbb{R}^d)$. *Suppose that the sequence* $(T_n)_{n \in \mathbb{N}}$ *converges* *to* T *in* $\mathscr{D}'(\mathbb{R}^d)$*, that there exists a closed set* F *in* $\mathbb{R}^d$ *such that* $\operatorname{Supp} T_n \subset F$ *for all* $n \in \mathbb{N}$*, and that* $(F, \operatorname{Supp} S)$ *satisfies* (C)*. Then*

$$\lim_{n \to +\infty} T_n * S = T * S$$

in $\mathscr{D}'(\mathbb{R}^d)$.

Proof. Take $\varphi \in \mathscr{D}(\mathbb{R}^d)$. As above, write $\hat{\varphi}(x, y) = \varphi(x + y)$. Since the family $(F, \operatorname{Supp} S)$ satisfies (C), the intersection $\operatorname{Supp} \hat{\varphi} \cap (F \times \operatorname{Supp} S)$ is compact. Let $\rho \in \mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$ satisfy $\rho = 1$ on an open set that contains this compact. Then, by definition,

$$\langle T_n * S,\, \varphi \rangle = \big\langle (T_n)_x,\, \langle S_y, \rho(x, y) \hat{\varphi}(x, y) \rangle \big\rangle.$$

Since the map $x \mapsto \big\langle S_y,\, \rho(x, y) \hat{\varphi}(x, y) \big\rangle$ belongs to $\mathscr{D}(\mathbb{R}^d)$, we deduce that

$$\lim_{n \to +\infty} \langle T_n * S, \varphi \rangle = \big\langle T_x,\, \langle S_y, \rho(x, y) \hat{\varphi}(x, y) \rangle \big\rangle = \langle T * S, \varphi \rangle,$$

which is the desired result. □

Obviously, this result extends to families (T_λ), with $\lambda \to \lambda_0$ (where λ runs over a subset of $\mathbb{R}$ and $\lambda_0 \in [-\infty, \infty]$).

The next proposition explicitly defines the convolution product.

Proposition 2.9 *Suppose* (T, S) *satisfies property* (C)*. Then, for every* $\varphi \in \mathscr{D}(\mathbb{R}^d)$*, the function* $\tilde{\varphi}$ *on* $\mathbb{R}^d$ *defined by*

$$\tilde{\varphi}(x) = \big\langle S_y,\, \varphi(x + y) \big\rangle$$

belongs to $\mathscr{E}(\mathbb{R}^d)$*, the intersection* $\operatorname{Supp} \tilde{\varphi} \cap \operatorname{Supp} T$ *is compact, and*

$$\langle T * S,\, \varphi \rangle = \langle T, \tilde{\varphi} \rangle = \big\langle T_x,\, \langle S_y, \varphi(x + y) \rangle \big\rangle.$$

Proof. Put $K = \{(x, y) \in \operatorname{Supp} T \times \operatorname{Supp} S : x + y \in \operatorname{Supp} \varphi\}$. Then the support of $\tilde{\varphi}$ is contained in $\operatorname{Supp} \varphi - \operatorname{Supp} S$ and $(\operatorname{Supp} \varphi - \operatorname{Supp} S) \cap \operatorname{Supp} T$ is the projection of K on the first factor. Therefore $\operatorname{Supp} \tilde{\varphi} \cap \operatorname{Supp} T$ is compact. At the same time, if $\rho_l \in \mathscr{D}(\mathbb{R}^d)$ satisfies $\rho_l = 1$ on $B(0, l)$, the function

$$\rho_l \tilde{\varphi} : x \mapsto \big\langle S_y,\, \rho_l(x) \varphi(x + y) \big\rangle$$

belongs to $\mathscr{D}(\mathbb{R}^d)$, by Theorem 1.1. Therefore $\tilde{\varphi}$ is of class C^∞ on $B(0, l)$ for every $l > 0$, which is to say that $\tilde{\varphi} \in \mathscr{E}(\mathbb{R}^d)$.

At the same time, by Proposition 2.8,

$$\begin{aligned} \langle T * S,\, \varphi \rangle &= \lim_{l \to +\infty} \lim_{l' \to +\infty} \langle \rho_l T * \rho_{l'} S,\, \varphi \rangle \\ &= \lim_{l \to +\infty} \lim_{l' \to +\infty} \big\langle T_x,\, \rho_l(x) \langle S_y, \rho_{l'}(y) \varphi(x + y) \rangle \big\rangle. \end{aligned}$$

Now, if $B(0,l') \supset \operatorname{Supp}\varphi - \operatorname{Supp}\rho_l$, we have

$$\operatorname{Supp}(\varphi(x+\cdot)) \subset B(0,l') \quad \text{for every } x \in \operatorname{Supp}\rho_l.$$

Therefore $\rho_{l'}(y)\varphi(x+y) = \varphi(x+y)$. We deduce that

$$\langle T * S,\, \varphi\rangle = \lim_{l\to+\infty} \bigl\langle T_x,\, \rho_l(x)\langle S_y, \varphi(x+y)\rangle\bigr\rangle.$$

By definition, if $B(0,l) \supset \operatorname{Supp}\tilde{\varphi} \cap \operatorname{Supp} T$, then

$$\bigl\langle T_x,\, \rho_l(x)\langle S_y, \varphi(x+y)\rangle\bigr\rangle = \bigl\langle T_x,\, \langle S_y, \varphi(x+y)\rangle\bigr\rangle,$$

which proves the result. □

This result can be extended to the case where $T \in \mathscr{D}'^m(\mathbb{R}^d)$, $S \in \mathscr{D}'^n(\mathbb{R}^d)$, and $\varphi \in \mathscr{D}^{m+n}(\mathbb{R}^d)$; see Exercise 7 below.

Corollary 2.10 *Let f and g be elements of $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ whose supports satisfy condition* (C). *Then f and g are convolvable in the sense of the definition on page* 171; *moreover $f * g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and*

$$[f] * [g] = [f * g].$$

Proof. For every $\varphi \in \mathscr{D}(\mathbb{R}^d)$,

$$\iint |f(x-y)|\,|g(y)|\,|\varphi(x)|\,dx\,dy = \iint |f(x)|\,|g(y)|\,|\varphi(x+y)|\,dx\,dy$$

(because Lebesgue measure is invariant under translations); the term on the right is finite because the supports of f and g satisfy condition (C). This proves that f and g are convolvable and that $f * g \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Moreover, if $\varphi \in \mathscr{D}(\mathbb{R}^d)$, we have

$$\bigl\langle [f*g],\, \varphi\bigr\rangle = \int f(x)\Bigl(\int g(y)\varphi(x+y)\,dy\Bigr)\,dx$$

by Fubini's Theorem, and this quantity equals $\bigl\langle [f]*[g],\, \varphi\bigr\rangle$ by Proposition 2.9. □

Proposition 2.11 (Associativity) *Let (T_1, T_2, T_3) be a family of distributions on $\mathbb{R}^d$ satisfying* (C). *The distributions $(T_1*T_2)*T_3$ and $T_1*(T_2*T_3)$ are well-defined and coincide.*

Proof. By property 1 on page 326, the distributions $T_1 * T_2$ and $T_2 * T_3$ are well defined and, by Proposition 2.7,

$$\operatorname{Supp}(T_1*T_2) \subset \operatorname{Supp}T_1 + \operatorname{Supp}T_2, \qquad \operatorname{Supp}(T_2*T_3) \subset \operatorname{Supp}T_2 + \operatorname{Supp}T_3.$$

It follows then from properties 1 and 7 on page 326 that the distributions $(T_1 * T_2) * T_3$ and $T_1 * (T_2 * T_3)$ are well defined.

In view of this we obtain, by applying Proposition 2.8 several times,

$$(T_1 * T_2) * T_3 = \lim_{l_1 \to +\infty} \lim_{l_2 \to +\infty} \lim_{l_3 \to +\infty} (\rho_{l_1} T_1 * \rho_{l_2} T_2) * \rho_{l_3} T_3,$$

where, for $l > 0$, we have $\rho_l \in \mathscr{D}(\mathbb{R}^d)$ and $\rho_l = 1$ on $B(0, l)$. Because the convolution product is associative in $\mathscr{E}'(\mathbb{R}^d)$ (Proposition 2.2), we get

$$(\rho_{l_1} T_1 * \rho_{l_2} T_2) * \rho_{l_3} T_3 = \rho_{l_1} T_1 * (\rho_{l_2} T_2 * \rho_{l_3} T_3).$$

Then it suffices to use Proposition 2.8 several times again to obtain

$$(T_1 * T_2) * T_3 = T_1 * (T_2 * T_3).$$

□

The same reasoning shows that, if $(T_1, \dots, T_n)$ satisfies (C), one can compute the product $T_1 * \cdots * T_n$ by grouping the terms in any desired way. On the contrary, if (T_1, T_2, T_3) does not satisfy (C), the distributions $(T_1 * T_2) * T_3$ and $T_1 * (T_2 * T_3)$ may both be defined but not be equal; see Example 4 below.

Proposition 2.12 *If $(T_1, \dots, T_n)$ satisfies condition (C), we have*

$$D_j(T_1 * \cdots * T_n) = T_1 * \cdots * T_{k-1} * D_j T_k * T_{k+1} * \cdots * T_n$$

for all $j \in \{1, \dots, d\}$ and $k \in \{1, \dots, n\}$. This remains so if we replace D_j by an arbitrary differential operator of the form $P(D)$.

Proof. Note first that $\operatorname{Supp} D_j T_k \subset \operatorname{Supp} T_k$, so the two sides in the equality above are well defined (see property 1 on page 326). By associativity and commutativity, it suffices to show that, if (T, S) satisfies (C), then $D_j(T * S) = (D_j T) * S$. We already know this is so when T and S have compact support (Proposition 2.3 on page 325). The general case follows by passing to the limit, using Proposition 2.8 and the continuity in $\mathscr{D}'(\mathbb{R}^d)$ of the map $T \mapsto D_j T$ (as well as the formula for the derivative of a product):

$$\begin{aligned} D_j(T * S) &= \lim_{l' \to +\infty} \lim_{l \to +\infty} D_j(\rho_l T * \rho_{l'} S) \\ &= \lim_{l' \to +\infty} \lim_{l \to +\infty} (\rho_l D_j T * \rho_{l'} S) + \lim_{l' \to +\infty} \lim_{l \to +\infty} ((D_j \rho_l) T * \rho_{l'} S) \\ &= D_j T * S, \end{aligned}$$

where the latter equality comes from the fact that $\lim_{l \to +\infty} (D_j \rho_l) T = 0$.

□

Examples

1. Let $P(D)$ be a linear differential operator with constant coefficients. Then, for every $T \in \mathscr{D}'(\mathbb{R}^d)$,

$$P(D)T = (P(D)\delta) * T.$$

2. Suppose $T \in \mathcal{D}'(\mathbb{R}^d)$ and $a \in \mathbb{R}^d$. The **translate** of T by a, denoted by $\tau_a T$, is the distribution defined by

$$\langle \tau_a T, \varphi \rangle = \langle T, \tau_{-a}\varphi \rangle \quad \text{for all } \varphi \in \mathcal{D}'(\mathbb{R}^d),$$

where, for every function f, $\tau_a f$ is the translate of f by a: that is, $\tau_a f(x) = f(x-a)$ (see page 169). One easily checks, using the invariance of Lebesgue measure under translations, that $\tau_a[f] = [\tau_a f]$ if $f \in L^1_{\text{loc}}(\mathbb{R}^d)$.

One deduces immediately from the definitions that

$$\tau_a T = \delta_a * T.$$

In particular, if $d = 1$,

$$T' = \delta' * T = \lim_{h \to 0} \frac{\delta - \delta_h}{h} * T$$

(see equation $(*)$ on page 293); equivalently,

$$T' = \lim_{h \to 0} \frac{T - \tau_h T}{h}.$$

3. Let σ_r be the surface measure on the sphere in $\mathbb{R}^d$ having center 0 and radius r. In view of Example 1 above, we deduce from Exercise 10 on page 304 that, for every distribution T on $\mathbb{R}^d$,

$$\Delta T = \lim_{r \to 0^+} \frac{2d}{r^2} \left(T * \frac{\sigma_r}{s_d r^{d-1}} - T \right).$$

Thus, if

$$T = T * \frac{\sigma_r}{s_d r^{d-1}}$$

for every $r > 0$ (or at least for r sufficiently small), we have $\Delta T = 0$. (In this case we say that T is a **harmonic distribution**.) The converse also holds; see Exercise 1 on page 344.

4. One easily checks that $(1 * \delta') * Y = 0$ and $1 * (\delta' * Y) = 1$, which shows that the convolution product is in general not associative.

2C Convolution of a Distribution with a Function

Proposition 2.13 *Consider $T \in \mathcal{D}'(\mathbb{R}^d)$ and $f \in \mathcal{E}(\mathbb{R}^d)$, and suppose (T, f) satisfies condition* (C). *Then $T * f \in \mathcal{E}(\mathbb{R}^d)$ and, for all $x \in \mathbb{R}^d$, the intersection* $\operatorname{Supp} f(x - \cdot\,) \cap \operatorname{Supp} T$ *is compact and*

$$T * f(x) = \langle T_y, f(x-y) \rangle .$$

This remains true if $T \in \mathcal{D}'^m(\mathbb{R}^d)$ and $f \in \mathcal{E}^{m+r}(\mathbb{R}^d)$ (with $m, r \in \mathbb{N}$), except that in this case $T * f \in \mathcal{E}^r(\mathbb{R}^d)$.

Proof. For each $l > 0$, we again fix an element ρ_l of $\mathscr{D}(\mathbb{R}^d)$ equal to 1 on $B(0,l)$. Take $T \in \mathscr{D}'^m(\mathbb{R}^d)$ and $f \in \mathscr{E}^{m+r}(\mathbb{R}^d)$ (or $T \in \mathscr{D}'(\mathbb{R}^d)$ and $f \in \mathscr{E}(\mathbb{R}^d)$), and suppose that (T, f) satisfies (C). For every compact subset K of $\mathbb{R}^d$, the set

$$\tilde{K} = \{(x,y) \in \operatorname{Supp} f \times \operatorname{Supp} T \,:\, x + y \in K\}$$

is a compact subset of $\mathbb{R}^d \times \mathbb{R}^d$. Denote by K' its projection on the second factor; then K' is compact in $\mathbb{R}^d$. For every $x \in K$,

$$\begin{aligned}\operatorname{Supp}(f(x-\cdot\,)) \cap \operatorname{Supp} T &= (x - \operatorname{Supp} f) \cap \operatorname{Supp} T \\ &= \{y \in \operatorname{Supp} T : \exists z \in \operatorname{Supp} f \text{ such that } y + z = x\} \\ &\subset K'.\end{aligned}$$

Now take $l > \max_{x\in K} |x| + \max_{y \in K'} |y|$. For every $x \in K$, the function $y \mapsto \rho_l(y)\rho_l(x-y)$ equals 1 on an open that contains K', so

$$\big\langle T_y,\, f(x-y)\big\rangle = \big\langle T_y,\, \rho_l(y)\rho_l(x-y) f(x-y)\big\rangle \quad \text{for all } x \in K. \qquad (*)$$

Since the function $(x,y) \mapsto \rho_l(y)\rho_l(x-y)f(x-y)$ lies in $\mathscr{D}^{m+r}(\mathbb{R}^d \times \mathbb{R}^d)$ (or $\mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$, as the case may be), we deduce from Theorem 1.1 that the function $x \mapsto \langle T_y, f(x-y)\rangle$ is of class C^r (or C^∞) in $\mathring{K}$. This reasoning is valid for every compact subset K of $\mathbb{R}^d$, so the function belongs to $\mathscr{E}^r(\mathbb{R}^d)$ (or $\mathscr{E}(\mathbb{R}^d)$).

Now consider $\varphi \in \mathscr{D}(\mathbb{R}^d)$. By the definition of the convolution product in $\mathscr{E}'(\mathbb{R}^d)$,

$$\begin{aligned}\big\langle (\rho_l T) * (\rho_l f),\, \varphi\big\rangle &= \big\langle (\rho_l T)_y,\, \langle (\rho_l f)_x, \varphi(x+y)\rangle\big\rangle \\ &= \Big\langle T_y,\, \rho_l(y) \int \rho_l(x) f(x) \varphi(x+y)\, dx\Big\rangle \\ &= \Big\langle T_y,\, \rho_l(y) \int \rho_l(x-y) f(x-y) \varphi(x)\, dx\Big\rangle \\ &= \big\langle \varphi(x) \otimes T_y,\, \rho_l(y)\rho_l(x-y) f(x-y)\big\rangle \\ &= \int \varphi(x) \big\langle T_y,\, \rho_l(y)\rho_l(x-y) f(x-y)\big\rangle\, dx.\end{aligned}$$

Now, applying equality $(*)$ to the compact $K = \operatorname{Supp} \varphi$, we see that, for l large enough,

$$\big\langle T_y,\, \rho_l(y)\rho_l(x-y) f(x-y)\big\rangle = \big\langle T, f(x-\cdot\,)\big\rangle \quad \text{for all } x \in \operatorname{Supp} \varphi.$$

Therefore, making l go to infinity, we obtain, by virtue of Proposition 2.6,

$$\langle T * f,\, \varphi\rangle = \int \varphi(x) \big\langle T,\, f(x-\cdot\,)\big\rangle\, dx,$$

which concludes the proof. $\square$

Remarks

1. In particular, consider a complex Radon measure μ on $\mathbb{R}^d$ and a map $f \in C(\mathbb{R}^d)$ such that $(\operatorname{Supp}\mu, \operatorname{Supp} f)$ satisfies (C). Then $\mu * f \in C(\mathbb{R}^d)$ and
$$\mu * f(x) = \int f(x-y)\, d\mu(y) \quad \text{for all } x \in \mathbb{R}^d.$$
For an extension to the case $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, see Exercise 10 below.
2. Take $T \in \mathscr{D}'(\mathbb{R}^d)$ and $f \in \mathscr{E}(\mathbb{R}^d)$ such that (T, f) satisfies (C), and recall from page 169 the notation $\check{f}$, defined by $\check{f}(x) = f(-x)$. By Proposition 2.13, $\langle T, \check{f}\rangle$ is well defined and
$$\langle T, \check{f}\rangle = T * f(0).$$
More generally, if $T \in \mathscr{D}'(\mathbb{R}^d)$, we define a distribution $\check{T}$ by
$$\langle \check{T}, \varphi\rangle = \langle T, \check{\varphi}\rangle \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^d).$$
Clearly, $\operatorname{Supp}\check{T} = -\operatorname{Supp} T$. Therefore, if (T, S) satisfies (C), so does $(\check{T}, \check{S})$. Moreover,
$$(T * S)\check{} = \check{T} * \check{S}.$$
This follows immediately from the definition of the convolution product (Proposition 2.6) and from the obvious fact that $(T \otimes S)\check{} = \check{T} \otimes \check{S}$.
As a consequence, by the associativity of the convolution product, we conclude that, if (T, S) satisfies (C), we have, for every $\varphi \in \mathscr{D}(\mathbb{R}^d)$,
$$\begin{aligned}\langle T * S, \varphi\rangle &= T * S * \check{\varphi}\,(0) = T * (S * \check{\varphi})\,(0) = T * (\check{S} * \varphi)\check{}\,(0)\\ &= \langle T, \check{S} * \varphi\rangle = \big\langle T_x, \langle S_y, \varphi(x+y)\rangle\big\rangle.\end{aligned}$$
We thus recover Proposition 2.9.

We now give an application of Proposition 2.13 to the smoothing of distributions.

Proposition 2.14 *For every open Ω in $\mathbb{R}^d$, the set $\mathscr{D}(\Omega)$ is dense in $\mathscr{D}'(\Omega)$. In other words, every distribution on Ω is the limit in $\mathscr{D}'(\Omega)$ of a sequence of elements of $\mathscr{D}(\Omega)$.*

Proof. Let Ω be open in $\mathbb{R}^d$, and let $(K_n)_{n\in\mathbb{N}}$ be a sequence of compact sets exhausting Ω. For every $n \in \mathbb{N}$, take $\varphi_n \in \mathscr{D}(\Omega)$ such that $\varphi_n = 1$ on K_n. Also let $(\chi_n)_{n\in\mathbb{N}}$ be a smoothing sequence in $\mathbb{R}^d$ and $(\chi_{p_n})_{n\in\mathbb{N}}$ a subsequence such that $\operatorname{Supp}\varphi_n + \operatorname{Supp}\chi_{p_n} \subset \Omega$ for every $n \in \mathbb{N}$.

Take $T \in \mathscr{D}'(\Omega)$ and write $\psi_n = (\varphi_n T) * \chi_{p_n}$ for every integer $n \in \mathbb{N}$. (The distribution $\varphi_n T$ has compact support in Ω and so can be identified with a distribution on $\mathbb{R}^d$ with compact support, as explained on page 283; see particularly Equation $(*)$ on that page. Thus the convolution product

$(\varphi_n T) * \chi_{p_n}$ does make sense.) By Proposition 2.13 and our assumption on the supports of φ_n and χ_{p_n}, we have $\psi_n \in \mathscr{D}(\Omega)$. We will show that the sequence $(\psi_n)_{n\in\mathbb{N}}$ converges in $\mathscr{D}'(\Omega)$ to T, and this will prove the proposition.

To do this, take $\varphi \in \mathscr{D}(\Omega)$. By definition,

$$\langle \psi_n, \varphi \rangle = \left\langle T_x, \varphi_n(x) \int \chi_{p_n}(y)\varphi(x+y)\,dy \right\rangle = \langle T, \varphi_n(\varphi * \check{\chi}_{p_n}) \rangle.$$

Now, for n large enough, $\operatorname{Supp}(\varphi * \check{\chi}_{p_n}) \subset \operatorname{Supp}\varphi - \operatorname{Supp}\chi_{p_n} \subset K_n$, so $\varphi_n(\varphi * \check{\chi}_{p_n}) = \varphi * \check{\chi}_{p_n}$, whence

$$\langle \psi_n, \varphi \rangle = \langle T, \varphi * \check{\chi}_{p_n} \rangle.$$

The sequence $(\varphi * \check{\chi}_{p_n})_{n\in\mathbb{N}}$ converges to φ in $\mathscr{D}(\Omega)$, by Proposition 1.2 of page 261 applied to every $m \in \mathbb{N}$, since $(\check{\chi}_n)_{n\in\mathbb{N}}$ is also a smoothing sequence. Therefore the sequence $(\psi_n)_{n\in\mathbb{N}}$ converges to T in $\mathscr{D}'(\Omega)$. □

Remark. With the notation used in the preceding proof, we see that, for a distribution T of order m and any $\varphi \in \mathscr{D}^m(\Omega)$, we have

$$\lim_{n\to+\infty} \int_\Omega \varphi(x)\psi_n(x)\,dx = \langle T, \varphi \rangle.$$

Exercises

1. Compute $\delta_x * \delta_y$, for $x, y \in \mathbb{R}^d$.

2. Let P and Q be polynomials in d variables:

$$P(x) = \sum_{|\alpha|\le p} a_\alpha x^\alpha, \quad Q(x) = \sum_{|\alpha|\le q} b_\alpha x^\alpha, \qquad \text{where } x = (x_1, \dots, x_d).$$

Compute $P(D)\delta * Q(D)\delta$.

3. Prove assertions 1, 5, 6, and 7 on page 326.

4. Let F be a closed subset of $\mathbb{R}^d$ containing 0 and such that $\lambda F \subset F$ for all $\lambda \in \mathbb{R}^+$.

- **a.** Show that (F, F) satisfies (C) if and only if $F \cap (-F) = \{0\}$.
- **b.** Suppose that $F \cap (-F) = \{0\}$ and that $F + F \subset F$. (For example, $F = (\mathbb{R}^+)^d$.) Show that, for every $r \ge 1$, the family $(F, \dots, F)$, where F is repeated r times, satisfies (C).

5. Let L be the function defined on $\mathbb{R}^d$ by $L(x) = a \cdot x$ with $a \in \mathbb{C}^d$ (where the dot represents the canonical scalar product in $\mathbb{C}^d$).

- **a.** If T and S are distributions satisfying (C), prove that
  - **i.** $L(S * T) = (LS) * T + S * (LT)$, and
  - **ii.** $e^L(S * T) = (e^L S) * (e^L T)$.

b. Let P be a polynomial in d variables.

i. Find a polynomial Q such that, for every $T \in \mathscr{D}'$,

$$e^L P(D)T = Q(D)(e^L T).$$

ii. Let E be a fundamental solution of $P(D)$. Determine a fundamental solution of $Q(D)$.

c. Derive from this and from Exercise 4 on page 323 a fundamental solution of $\prod_{j=1}^{d}(D_j - a_j)$, where $a_1, \dots, a_d \in \mathbb{C}$.

6. a. Let P be a polynomial in d variables and T a distribution with compact support. Show that $T * P$ is a polynomial.

Hint. Use Exercise 4 on page 303 (or Proposition 2.13).

b. Find the limit in $\mathscr{D}'(\mathbb{R}^d)$ of the sequence of polynomials (P_n) on $\mathbb{R}^d$ defined by

$$P_n(x) = \frac{n^d}{\pi^{d/2}}\left(1 - \frac{|x|^2}{n}\right)^{n^3}.$$

c. Deduce that every distribution with compact support is the limit in $\mathscr{D}'(\mathbb{R}^d)$ of a sequence of polynomials.

7. Let $m, n \in \mathbb{N}$, and consider $T \in \mathscr{D}'^m(\mathbb{R}^d)$ and $S \in \mathscr{D}'^n(\mathbb{R}^d)$ such that (T, S) satisfies (C). Show that $T * S \in \mathscr{D}'^{m+n}(\mathbb{R}^d)$ and that

$$\langle T * S, \varphi\rangle = \big\langle T_x, \langle S_y, \varphi(x+y)\rangle\big\rangle \quad \text{for all } \varphi \in \mathscr{D}^{m+n}(\mathbb{R}^d).$$

8. *Convolution of measures*

a. Show that, if μ and ν are complex Radon measures on $\mathbb{R}^d$ whose supports satisfy (C), the convolution $\mu * \nu$ is a Radon measure on $\mathbb{R}^d$ and

$$\langle \mu * \nu, \varphi\rangle = \iint \varphi(x+y)\, d\mu(x)\, d\nu(y) \quad \text{for all } \varphi \in C_c(\mathbb{R}^d).$$

(The double integral is defined by decomposing μ and ν into positive measures: see page 89.)

Hint. See Exercise 7 with $m = n = 0$.

b. Let μ and ν be bounded complex Radon measures on $\mathbb{R}^d$. Show that one can define $\mu * \nu$ by the formula of the previous question and that $\mu * \nu$ is a bounded Radon measure.

c. Show that the space $\mathfrak{M}_f(\mathbb{R}^d)$, with the convolution product $*$ and the norm of $(C_0(\mathbb{R}^d))'$, is a commutative Banach algebra with unity and that $L^1(\mathbb{R}^d)$ is a closed subalgebra of it (without unity).

9. a. Show that, if μ is a Radon measure on $\mathbb{R}^d$ and if $\varphi \in C_c(\mathbb{R}^d)$, then $\mu * \varphi \in C(\mathbb{R}^d)$.

Hint. See the first remark following Proposition 2.13.

b. Conversely, let T be a distribution on $\mathbb{R}^d$ such that $T * \varphi \in C(\mathbb{R}^d)$ for every $\varphi \in C_c(\mathbb{R}^d)$.

i. Let $(\chi_n)_{n\in\mathbb{N}}$ be a smoothing sequence. Show that the sequence $(T * \chi_n)_{n\in\mathbb{N}}$ converges vaguely in the sense of Exercise 6 on page 91.
Hint. For every $\varphi \in C_c(\mathbb{R}^d)$,

$$\lim_{n\to+\infty} \int (T * \chi_n)(x)\varphi(x)\,dx = T * \check{\varphi}(0).$$

ii. Deduce that T is a Radon measure.

10. *Convolution of a measure with a locally integrable function.* Suppose μ is a complex Radon measure on $\mathbb{R}^d$, that $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, and that the supports of μ and f satisfy (C). Show that $\mu * f \in L^1_{\text{loc}}(\mathbb{R}^d)$ and that

$$\mu * f(x) = \int f(x-y)\,d\mu(y) \quad \text{for almost every } x$$

(where the integral is defined by considering a particular Borel function representing f).

11. **a.** Let L be a continuous linear map from $\mathscr{D}(\mathbb{R}^d)$ to $\mathscr{E}(\mathbb{R}^d)$, commuting with translations. Show that there exists a distribution $T \in \mathscr{D}'(\mathbb{R}^d)$ such that

$$L(\varphi) = T * \varphi \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^d).$$

(You might note that the equality $\langle T, \varphi\rangle = (L(\check{\varphi}))(0)$ must hold.)

b. Let L be a continuous linear map from $\mathscr{D}(\mathbb{R}^d)$ to $\mathscr{E}(\mathbb{R}^d)$ commuting with each differentiation D_j, for $1 \le j \le d$. Show that there exists a distribution $T \in \mathscr{D}'(\mathbb{R}^d)$ such that

$$L(\varphi) = T * \varphi \quad \text{for all } \varphi \in \mathscr{D}(\mathbb{R}^d).$$

Hint. You might show that L commutes with translations, as follows: Take $\varphi \in \mathscr{D}$ and $u \in \mathbb{R}^d$, and let h be the function defined by

$$h(x) = (\tau_{-x} L \tau_x \varphi)(u) = (L\tau_x\varphi)(u+x),$$

where $\tau_x\psi(y) = \psi(y-x)$. Show that all partial derivatives $D_j h$ are zero. Deduce that h is constant and finish the proof.

3 Applications

3A Primitives and Sobolev's Theorem

The next proposition allows one to recover a distribution with compact support from its first derivatives. Thus it is a formula for finding a primitive.

Proposition 3.1 *If $T \in \mathscr{E}'(\mathbb{R}^d)$, then*

$$T = \frac{1}{s_d} \sum_{j=1}^{d} \left(\frac{x_j}{r^d} \right) * D_j T,$$

where $r = |x|$ and s_d is the area of the unit sphere in $\mathbb{R}^d$.

Proof. Let E be the fundamental solution of the Laplacian given in Theorem 3.2 on page 308. A simple calculation using Theorem 2.12 on page 301 shows that

$$D_j E = \frac{1}{s_d} \frac{x_j}{r^d} \quad \text{for all } j \in \{1, \dots, d\},$$

and this in any dimension d. At the same time, $\Delta E * T = T$, since $\Delta E = \delta$. Since T has compact support (so that (E, T) satisfies (C)), we deduce from Proposition 2.12 that

$$\Delta E * T = \Delta(E * T) = \sum_{j=1}^{d} D_j^2 (E * T) = \sum_{j=1}^{d} (D_j E) * (D_j T).$$

Therefore,

$$T = \sum_{j=1}^{d} (D_j E) * (D_j T),$$

which yields the result. □

We now introduce the *Sobolev spaces* $W^{1,p}$ over $\mathbb{R}^d$, where $1 \le p \le \infty$. By definition, the Sobolev space $W^{1,p}(\mathbb{R}^d)$ is the set of elements $f \in L^p(\mathbb{R}^d)$ for which, for every $j \in \{1, \dots, d\}$, there exists $g_j \in L^p(\mathbb{R}^d)$ such that $D_j[f] = [g_j]$. In the sequel we will omit the brackets, writing simply $D_j f = g_j$.

We define on the space $W^{1,p}(\mathbb{R}^d)$ a norm $\|\cdot\|_{1,p}$, as follows:

$$\|f\|_{1,p} = \|f\|_p + \sum_{j=1}^{d} \|D_j f\|_p \quad \text{for all } f \in W^{1,p}(\mathbb{R}^d).$$

Here $\|\cdot\|_p$ is the norm on $L^p(\mathbb{R}^d)$.

Proposition 3.2 *The norm $\|\cdot\|_{1,p}$ makes $W^{1,p}(\mathbb{R}^d)$ into a Banach space.*

Proof. Let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in $W^{1,p}(\mathbb{R}^d)$. Since the space $L^p(\mathbb{R}^d)$ is complete, the sequences (f_n), $(D_1 f_n)$, ..., $(D_d f_n)$, which are clearly Cauchy sequences in $L^p(\mathbb{R}^d)$, converge in $L^p(\mathbb{R}^d)$. Let f, g_1, ..., g_d be their limits in $L^p(\mathbb{R}^d)$. Since $\mathscr{D}(\mathbb{R}^d)$ is contained in $L^{p'}(\mathbb{R}^d)$ (where p' is the exponent conjugate to p), we deduce easily from Hölder's inequality that the same sequences also converge in $\mathscr{D}'(\mathbb{R}^d)$. Since the operators D_j are continuous in $\mathscr{D}'(\mathbb{R}^d)$, we deduce that

$$D_j f = \lim_{n\to+\infty} D_j f_n = g_j \quad \text{for } 1 \le j \le d,$$

by the uniqueness of the limit in $\mathscr{D}'(\mathbb{R}^d)$. This shows that $f \in W^{1,p}(\mathbb{R}^d)$ and that the sequence $(f_n)_{n\in\mathbb{N}}$ converges to f in $W^{1,p}(\mathbb{R}^d)$. □

Remark. The space $W^{1,2}(\mathbb{R}^d)$ is often denoted by $H^1(\mathbb{R}^d)$ and given the equivalent norm $\|\cdot\|_{H^1}$ defined by

$$\|f\|_{H^1} = \Big(\|f\|_2^2 + \sum_{j=1}^d \|D_j f\|_2^2\Big)^{1/2},$$

which comes from the scalar product defined by

$$(f \mid g) = \int f(x)\overline{g(x)}\,dx + \sum_{j=1}^d \int D_j f(x)\overline{D_j g(x)}\,dx.$$

Thus $H^1(\mathbb{R}^d)$ is a Hilbert space.

The next theorem says that, if p is finite, $W^{1,p}(\mathbb{R}^d)$ is continuously embeddable in some spaces $L^r(\mathbb{R}^d)$ with $r > p$, and that, if $d < p < \infty$, it is continuously embeddable in $C_0(\mathbb{R}^d)$ ("continuously embeddable" means that $W^{1,p}$ is contained in each of the spaces considered and that the corresponding canonical injections are continuous).

Theorem 3.3 (Sobolev Injection Theorem) *Suppose that $p \in [1,\infty]$ and that r satisfies*

- *$r \in [p, pd/(d-p))$ if $p < d$,*
- *$r \in [p,\infty)$ if $p = d$,*
- *$r \in [p,\infty]$ if $p > d$.*

Then $W^{1,p}(\mathbb{R}^d) \subset L^r(\mathbb{R}^d)$ and there exists $C_{r,p} \ge 0$ such that

$$\|f\|_r \le C_{r,p}\|f\|_{1,p} \quad \text{for all } f \in W^{1,p}(\mathbb{R}^d).$$

Moreover, if $d < p < \infty$, every element of $W^{1,p}(\mathbb{R}^d)$ has a representative in $C_0(\mathbb{R}^d)$. Finally, every element in $W^{1,\infty}(\mathbb{R}^d)$ has a uniformly continuous representative.

Proof. Let $\gamma \in \mathscr{D}(\mathbb{R}^d)$ be such that $\gamma = 1$ in a neighborhood of 0. Since $|x_j r^{-d}| \le r^{1-d}$, we have $\gamma x_j r^{-d} \in L^\alpha(\mathbb{R}^d)$ for every $\alpha \ge 1$ such that $\alpha(d-1) < d$ (see Proposition 2.13 on page 302), and so also for every $\alpha \in [1, d/(d-1))$.

Let E be the fundamental solution of the Laplacian given by Theorem 3.2 on page 308. Since γE has compact support (so that $(T, \gamma E)$ satisfies (C)), we deduce from Proposition 2.12 that

$$T * \Delta(\gamma E) = \Delta(T * \gamma E) = \sum_{j=1}^{d} D_j^2(T * \gamma E) = \sum_{j=1}^{d} (D_j T) * D_j(\gamma E).$$

Now,

$$\Delta(\gamma E) = (\Delta\gamma)E + 2\sum_{j=1}^{d} D_j\gamma D_j E + \gamma \Delta E \tag{$*$}$$

(Leibniz's formula). Since $\Delta E = \delta$, we get $\gamma\Delta E = \gamma(0)\delta = \delta$. Since E is of class C^∞ on $\mathbb{R}^d \setminus \{0\}$ and since $\Delta\gamma$ and $D_j\gamma$ vanish near 0, we deduce from $(*)$ that

$$\eta = \Delta(\gamma E) - \delta \in \mathscr{D}(\mathbb{R}^d).$$

Similarly, we can show that, for each $j \in \{1, \dots, d\}$, there exists an $\eta_j \in \mathscr{D}(\mathbb{R}^d)$ such that

$$D_j(\gamma E) = \eta_j + \frac{1}{s_d}\gamma\frac{x_j}{r^d}.$$

We then get

$$T = -T * \eta + \sum_{j=1}^{d} D_j T * \eta_j + \frac{1}{s_d}\sum_{j=1}^{d} (D_j T) * \left(\gamma\frac{x_j}{r^d}\right). \tag{$**$}$$

Suppose $T \in W^{1,p}(\mathbb{R}^d)$. Then T, D_1T, ..., $D_dT \in L^p$ and we can apply Young's inequality (Theorem 3.4 on page 172) to equation $(**)$. We conclude that $T \in L^r(\mathbb{R}^d)$ for every r such that $1/p + 1/\alpha - 1 = 1/r$, where $1 \le \alpha \le p/(p-1)$ and $\alpha < d/(d-1)$. If $p \le d$, we must have $1 \le \alpha < d/(d-1)$, so that $r \in [p, pd/(d-p))$ (with $pd/(d-p) = \infty$ if $p = d$). If $p > d$, we must have $1 \le \alpha \le p/(p-1)$, so that $r \in [p, \infty]$. In particular, we can take $\alpha = p/(p-1) = p'$, the conjugate exponent of p. The last part of the theorem then follows from Proposition 3.2 on page 171. Finally, the existence of constants $C_{r,p}$ also follows from equation $(**)$ and Young's inequality. □

3B *Regularity*

Let Ω be open in $\mathbb{R}^d$. If $p \in [1, \infty]$, denote by $L^p_{\mathrm{loc}}(\Omega)$ the set of equivalence classes (with respect to Lebesgue measure) of functions on Ω such that $1_K f \in L^p(\Omega)$ for every compact K in Ω.

Theorem 3.4 *Let T be a distribution on an open set Ω in $\mathbb{R}^d$. Suppose that $p \in [1, \infty]$ and that $D_jT \in L^p_{\mathrm{loc}}(\Omega)$ for every $j \in \{1, \dots, d\}$.*

– *If $p \le d$, then $T \in L^r_{\mathrm{loc}}(\Omega)$ for every $r \in [p, pd/(d-p))$.*
– *If $p > d$, then $T \in C(\Omega)$.*

If $p = d$, we interpret $pd/(d-p)$ as ∞.

Proof. Let K be a compact subset of Ω and K' a compact subset of Ω whose interior contains K. Let $\varphi \in \mathscr{D}(\Omega)$ be such that $\varphi = 1$ on K'. Put $\eth = d(K, \Omega\backslash \mathring{K}') > 0$ and let $\gamma \in \mathscr{D}(\mathbb{R}^d)$ be such that $\gamma = 1$ in a neighborhood of 0 and $\operatorname{Supp}\gamma \subset B(0, \eth/2)$. If E is the fundamental solution of the Laplacian provided by Theorem 3.2 on page 308, we saw in the proof of Theorem 3.3 that there exist $\eta, \eta_1, \dots, \eta_d \in \mathscr{D}(\mathbb{R}^d)$ such that

$$\Delta(\gamma E) = \eta + \delta, \qquad D_j(\gamma E) = \eta_j + \frac{1}{s_d}\gamma\frac{x_j}{r^d} \quad \text{for all } j \in \{1, \dots, d\}.$$

Using formula (**) from the previous page and replacing T by φT (considered as a distribution on $\mathbb{R}^d$: see page 283), we obtain

$$\begin{aligned}\varphi T = -(\varphi T) * \eta + \sum_{j=1}^{d} D_j(\varphi T) * \eta_j \\ + \frac{1}{s_d}\sum_{j=1}^{d}((D_j\varphi)T) * \left(\gamma\frac{x_j}{r^d}\right) + \frac{1}{s_d}\sum_{j=1}^{d}(\varphi D_j T) * \left(\gamma\frac{x_j}{r^d}\right).\end{aligned}$$

By Proposition 2.13, $(\varphi T) * \eta$ and the $D_j(\varphi T) * \eta_j$, for every j, belong to $\mathscr{D}(\mathbb{R}^d)$. At the same time,

$$\operatorname{Supp}\left((D_j\varphi)T * \left(\gamma\frac{x_j}{r^d}\right)\right) \subset (\mathbb{R}^d\backslash \mathring{K}') + \bar{B}(0, \eth/2) \subset (\mathbb{R}^d \setminus K).$$

Finally, $\varphi D_j T \in L^p(\mathbb{R}^d)$. We then apply Young's inequality (Theorem 3.4 on page 172) as in the proof of Theorem 3.3. We conclude that

$$\Omega_K = \mathbb{R}^d \setminus \left((\mathbb{R}^d\backslash \mathring{K}') + \bar{B}(0, \eth/2)\right)$$

is an open set satisfying $K \subset \Omega_K \subset K'$ and that the restriction of T to Ω_K belongs to $L^r(\Omega_K)$ if $p \le d$ and $r \in [p, pd/(p-d))$ and to $C(\Omega_K)$ if $p > d$. Since this happens for every compact K, the theorem is proved. □

Hypoelliptic Differential Operators

We now state another fairly general regularity criterion. We start with a definition: If P is a polynomial over $\mathbb{C}$, the linear differential operator $P(D)$ is said to be **hypoelliptic** if, for every open subset Ω of $\mathbb{R}^d$ and every $T \in \mathscr{D}'(\Omega)$,

$$P(D)T \in \mathscr{E}(\Omega) \implies T \in \mathscr{E}(\Omega).$$

In particular, if $P(D)$ is hypoelliptic, every solution in $\mathscr{D}'(\Omega)$ of the partial differential equation $P(D)T = 0$ is a function of class C^∞, and so also a solution in the ordinary sense.

Theorem 3.5 *Any differential operator with constant coefficients having a fundamental solution whose restriction to $\mathbb{R}^d \setminus \{0\}$ is a function of class C^∞ is hypoelliptic.*

Proof. The proof is analogous to that of Theorem 3.4. Let Ω be an open subset of $\mathbb{R}^d$ and let K, K' be compact subsets of Ω such that $K \subset \mathring{K}'$. Write $\eth = d(K, \mathbb{R}^d \setminus \mathring{K}')$. Let $\varphi \in \mathscr{D}(\Omega)$ have the value 1 on K' and let $\gamma \in \mathscr{D}(\mathbb{R}^d)$ have the value 1 on a neighborhood of 0; assume also that $\operatorname{Supp} \gamma \subset \bar{B}(0, \eth/2)$. Finally, set

$$\Omega_K = \mathbb{R}^d \setminus \big((\mathbb{R}^d \setminus \mathring{K}') + \bar{B}(0, \eth/2)\big).$$

Then

$$K \subset \Omega_K \subset K'.$$

Consider a differential operator $P(D)$ having a fundamental solution E of class C^∞ on $\mathbb{R}^d \setminus \{0\}$. Let $T \in \mathscr{D}'(\Omega)$ be such that $f = P(D)T \in \mathscr{E}(\Omega)$. By Leibniz's formula,

$$P(D)(\gamma E) = \delta + \eta \quad \text{with } \eta \in \mathscr{D}(\mathbb{R}^d)$$

and

$$P(D)(\varphi T) = \varphi f + S \quad \text{with } \operatorname{Supp} S \subset (\mathbb{R}^d \setminus \mathring{K}').$$

Then

$$P(D)(\gamma E * \varphi T) = \varphi T + \varphi T * \eta = \gamma E * \varphi f + \gamma E * S,$$

that is,

$$\varphi T = -(\varphi T) * \eta + \gamma E * \varphi f + \gamma E * S.$$

Since η and φf belong to $\mathscr{D}(\mathbb{R}^d)$, we deduce from Proposition 2.13 that $-(\varphi T) * \eta + \gamma E * \varphi f \in \mathscr{D}(\mathbb{R}^d)$. On the other hand, $\operatorname{Supp}(\gamma E * S) \subset \mathbb{R}^d \setminus \Omega_K$. We deduce that the restriction of T to Ω_K is of class C^∞. Since K is arbitrary and $\Omega_K \supset K$, this implies that $T \in \mathscr{E}(\Omega)$. □

Examples

The operators Δ, $\mathscr{C}$, $\partial/\partial\bar{z}$ (see Chapter 8), as well as Δ^k for $k \geq 2$ (see Exercise 4 on page 313), are hypoelliptic. In particular, a harmonic distribution T on Ω is a harmonic function in the classical sense; a distribution T on an open subset Ω of $\mathbb{R}^2$ such that $\partial T/\partial\bar{z} = 0$ is a holomorphic function on Ω.

If $d = 1$, every operator is hypoelliptic.

Conversely, note that, if E is a fundamental solution of a hypoelliptic operator $P(D)$, the restriction $\tilde{E}$ of E to $\mathbb{R}^d \setminus \{0\}$ is of class C^∞ (since $P(D)\tilde{E} = 0 \in \mathscr{E}(\mathbb{R}^d \setminus \{0\})$). This allows one to show that, for example, the operator

$$\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2}$$

on $\mathbb{R}^2$ is not hypoelliptic (see Exercise 6 on page 314).

3C Fundamental solutions and Partial Differential Equations

The existence of a fundamental solution for a differential operator allows one to find solutions of the corresponding partial differential equation if the right-hand side is an operator with compact support.

Theorem 3.6 *Consider a linear differential operator $P(D)$ with constant coefficients, a fundamental solution E of $P(D)$, and $S \in \mathscr{E}'(\mathbb{R}^d)$. The distribution $T_0 = E * S$ satisfies $P(D)T_0 = S$. Moreover, the set of solutions $T \in \mathscr{D}'(\mathbb{R}^d)$ of the equation*

$$P(D)T = S$$

equals $\{T = T_0 + U : U \in \mathscr{D}'(\mathbb{R}^d)$ such that $P(D)U = 0\}$.

Proof. If $S \in \mathscr{E}'(\mathbb{R}^d)$, Proposition 2.12 yields

$$P(D)(E * S) = P(D)E * S = \delta * S = S.$$

Set $U = T - E * S$. Clearly, $P(D)T = S$ if and only if $P(D)U = 0$. □

3D The Algebra $\mathscr{D}'_+$

We now consider the case $d = 1$ and write

$$\mathscr{D}'_+ = \{T \in \mathscr{D}'(\mathbb{R}) : \operatorname{Supp} T \subset \mathbb{R}^+\}.$$

Because $(\mathbb{R}^+, \dots, \mathbb{R}^+)$ satisfies condition (C), the convolution of two elements of $\mathscr{D}'_+$ is always defined and this operation makes $\mathscr{D}'_+$ into a commutative algebra with unity, by Propositions 2.7 and 2.11. We will apply this fact to the resolution of linear differential equations with constant coefficients and continuous right-hand side.

Let $P(D) = a_0 + a_1 D + \cdots + a_m D^m$ be a linear differential operator with constant coefficients such that $m \geq 1$ and $a_m \neq 0$. We know from Theorem 3.1 on page 307 that $P(D)$ has a fundamental solution $E = (1/a_m)Yf$, where f is the solution on $\mathbb{R}$ of the differential equation $P(D)f = 0$ satisfying the conditions $f(0) = f'(0) = \cdots = f^{(m-2)}(0) = 0$ and $f^{(m-1)}(0) = 1$. In particular, $E \in \mathscr{D}'_+$. It follows that $P(D)\delta$ is invertible in the algebra $\mathscr{D}'_+$ and that its (necessarily unique) inverse is E. Thus, for every $S \in \mathscr{D}'_+$, there exists a unique distribution $T \in \mathscr{D}'_+$ such that $P(D)T = S$: namely, $T = E * S$. If we take, for example, $\psi \in C(\mathbb{R}^+)$ (and extend it to $\mathbb{R}^-$ with the value 0), and if we put

$$\varphi(x) = \begin{cases} \dfrac{1}{a_m} \displaystyle\int_0^x f(x-y)\psi(y)\,dy & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases}$$

then, in the sense of distributions,

$$P(D)\varphi = \psi.$$

On the other hand, we know from the theory of differential equations that the equation $P(D)g = \psi$ on $\mathbb{R}^+$ has a unique solution g that satisfies the conditions $g(0) = g'(0) = \cdots = g^{(m-1)}(0) = 0$. Extending g to $(-\infty, 0]$ with the value 0, we get as well $P(D)g = \psi$ in the sense of distributions. We deduce, by identification, that

$$g(x) = \frac{1}{a_m}\int_0^x f(x-y)\psi(y)\,dy \quad \text{for all } x \geq 0.$$

Applying a similar reasoning to the case $x \leq 0$, we finally see that, for every $\psi \in C(\mathbb{R})$, the unique solution g of the equation $P(D)g = \psi$ satisfying $g(0) = g'(0) = \cdots = g^{(m-1)}(0) = 0$ is given by

$$g(x) = \frac{1}{a_m}\int_0^x f(x-y)\psi(y)\,dy \quad \text{for all } x \in \mathbb{R}.$$

Exercises

1. *Harmonic functions and the mean value property.* This exercise is a continuation of Exercise 3 on page 312, whose notation we keep.
   a. Show that $E - E^\rho \in \mathscr{E}'(\mathbb{R}^d)$ for every $\rho > 0$.
   b. Deduce that, for every $T \in \mathscr{D}'(\mathbb{R}^d)$,

   $$(E - E^\rho) * \Delta T = T - T * \frac{1}{s_d\rho^{d-1}}\sigma_\rho. \qquad (*)$$

   Show that, in particular, any harmonic function f on $\mathbb{R}^d$ (that is, any $f \in \mathscr{E}^2(\mathbb{R}^2)$ satisfying $\Delta f = 0$ in the ordinary sense) satisfies the following *mean value property*:

   $$f = f * \frac{1}{s_d\rho^{d-1}}\sigma_\rho \quad \text{for all } \rho > 0.$$

   (This is converse to the property of Example 3 on page 332.)
   c. Applying $(*)$ to the distribution $T = E$, prove that

   $$E^\rho = E * \frac{1}{s_d\rho^{d-1}}\sigma_\rho.$$

2. *Subharmonic distributions.* A distribution is said to be *subharmonic* if its Laplacian is a positive distribution. (For example, if f is a harmonic real-valued function, $|f|$ defines a subharmonic distribution: see Exercise 11 on page 304.)

a. Characterize subharmonic distributions on $\mathbb{R}$.
Hint. See Exercise 12 on page 304.

b. Show that a distribution T on $\mathbb{R}^d$ is subharmonic if and only if, for every $\rho > 0$,

$$T * \frac{1}{s_d \rho^{d-1}} \sigma_\rho - T$$

is a positive distribution.
Hint. Sufficiency follows from Example 3 on page 332 and necessity from equation $(*)$ in Exercise 1.

c. Show that every subharmonic distribution can be represented by a locally integrable function.
Hint. Revisit the proof of Theorem 3.5 and use Exercise 10 on page 337.

d. Let f and g be locally integrable real functions on $\mathbb{R}^d$, and assume f and g are subharmonic (this means that the distributions $[f]$ and $[g]$ are subharmonic). Show that $\sup(f, g)$ is subharmonic.

e. Recall that a function f from $\mathbb{R}^d$ to $[-\infty, \infty]$ is said to be *upper semicontinuous* if, for every $a \in \mathbb{R}$, the set $\{f < a\}$ is open. Recall also that the pointwise limit of a decreasing sequence of continuous real functions is an upper semicontinuous function with values in $[-\infty, +\infty)$. Show that, if f is a subharmonic real-valued function, there exists an upper semicontinuous function $\tilde{f}$ with values in $[-\infty, +\infty)$ such that $f = \tilde{f}$ almost everywhere.
Hint. Take again the proof of Theorem 3.5 and note that there exists a decreasing sequence in $C_c(\mathbb{R}^d)$ that converges pointwise to γE.

3. *Harmonic functions and the mean value property, continued.* We wish to characterize harmonicity *on an open set* by the mean value property. We fix an open subset Ω of $\mathbb{R}^d$ and $f \in C(\Omega)$, and keep the notation of Exercise 1.

a. Suppose that f is a harmonic function on Ω. Take $x \in \Omega$ and $\rho \in (0, d(x, \mathbb{R}^d \backslash \Omega))$. Take also $\varepsilon > 0$ such that $\varepsilon < d(x, \mathbb{R}^d \backslash \Omega) - \rho$. Denote by φ an element of $C_c(\Omega)$ such that $\varphi = 1$ on $\bar{B}(x, \rho + \varepsilon)$, and identify φf with an element of $C_c(\mathbb{R}^d)$.

i. Show that $(E - E^\rho) * \Delta(\varphi f) = 0$ on $B(x, \varepsilon)$.

ii. Deduce that

$$\varphi f = \varphi f * \frac{1}{s_d \rho^{d-1}} \sigma_\rho$$

on $B(x, \varepsilon)$. (Use part b of Exercise 1.)

iii. Show that f satisfies the following *mean value property*: For all $x \in \Omega$ and all $\rho \in (0, d(x, \mathbb{R}^d \backslash \Omega))$, we have

$$f(x) = \frac{1}{s_d \rho^{d-1}} \int f(x - y)\, d\sigma_\rho(y).$$

Hint. Use the first remark after Proposition 2.13 on page 332.

b. Conversely, suppose that, for all $x \in \Omega$ and all $\rho \in (0, d(x, \mathbb{R}^d \backslash \Omega))$, we have

$$f(x) = \frac{1}{s_d \rho^{d-1}} \int f(x-y)\, d\sigma_\rho(y).$$

Let K and K' be compact subsets of Ω such that $K \subset \mathring{K}'$, and take $\varphi \in C_c(\Omega)$ such that $\varphi = 1$ on K'.

i. Show that

$$\varphi f = \varphi f * \frac{1}{s_d \rho^{d-1}} \sigma_\rho$$

on K if $0 < \rho < d(K, (\mathbb{R}^d \backslash \mathring{K}'))$. Deduce that $\Delta f = 0$ in the interior of K.

Hint. See Example 3 on page 332.

ii. Show that in this case f is a harmonic function on Ω.

c. Show likewise that f is subharmonic (see Exercise 2) if and only if, for all $x \in \Omega$ and all $\rho \in (0, d(x, \mathbb{R}^d \backslash \Omega))$,

$$f(x) \leq \frac{1}{s_d \rho^{d-1}} \int f(x-y)\, d\sigma_\rho(y).$$

4. Show that, for every $p \in [1, +\infty)$, the space $W^{1,p}(\mathbb{R}^d)$ is separable and $\mathscr{D}(\mathbb{R}^d)$ is dense in $W^{1,p}(\mathbb{R}^d)$.

Hint. For separability, note that $W^{1,p}(\mathbb{R}^d)$ is isometric to a subspace of $(L^p(\mathbb{R}^d))^{d+1}$.

5. a. Show that, if $f \in W^{1,1}(\mathbb{R})$, there exists an element $g \in L^1(\mathbb{R})$ such that $\int_{-\infty}^{+\infty} g(x)\, dx = 0$ and $\int_{-\infty}^{x} g(t)\, dt = f(x)$ almost everywhere. Deduce that f has a representative in $C_0(\mathbb{R})$. (By Theorem 3.3, this is still true if $f \in W^{1,p}(\mathbb{R})$ for $1 < p < \infty$.)

b. Show that there exists $f \in W^{1,2}(\mathbb{R}^2)$ such that $f \notin L^\infty(\mathbb{R}^2)$.

Hint. Take f with compact support and equal to $\log(\log(1/r))$ in a neighborhood of 0.

6. Let T be a distribution on an open Ω of $\mathbb{R}^d$. Suppose that $D^p T \in L^1_{\mathrm{loc}}(\Omega)$ for every multiindex p of length $d+1$. Show that $T \in C(\Omega)$. (Apply Theorem 3.4 $d+1$ times.) Deduce that if $D^p T \in L^1_{\mathrm{loc}}(\Omega)$ for every multiindex $p \in \mathbb{N}^d$, then $T \in \mathscr{E}(\Omega)$.

7. Let F be a closed subset of $\mathbb{R}^d$ such that

$$\lambda F \subset F \quad \text{for all } \lambda \in \mathbb{R}^+, \qquad F \cap (-F) = \{0\}, \qquad F + F \subset F.$$

Write $\mathscr{D}'_F = \{T \in \mathscr{D}'(\mathbb{R}^d) : \operatorname{Supp} T \subset F\}$. Show that the convolution product makes $\mathscr{D}'_F$ into a commutative algebra with unity. (See Exercise 4 on page 335.)

8. Denote by $Y^{(d)}$ the function on $\mathbb{R}^d$ defined by

$$Y^{(d)}(x) = Y(x_1) \ldots Y(x_d).$$

Also set $F = (\mathbb{R}^+)^d$, a closed set.

a. **i.** Show that $Y^{(d)}$ is a fundamental solution of the operator $D_1 \dots D_d$ and that the support of $Y^{(d)}$ is contained in F.

ii. Let $\mathscr{D}'_F$ be the space defined in Exercise 7. Show that, if $S \in \mathscr{D}'_F$, there exists a unique $T \in \mathscr{D}'_F$ such that $D_1 \dots D_d T = S$.

b. **i.** Let T be a distribution on $\mathbb{R}^d$ supported within $a + F$, where $a \in \mathbb{R}^d$. Show that the convolution $Y^{(d)} * T$ is well defined, that $\operatorname{Supp}(Y^{(d)} * T) \subset a + F$ and that

A. if $T \in L^1_{\text{loc}}(\mathbb{R}^d)$, then $Y^{(d)} * T \in C(\mathbb{R}^d)$;

B. if T is a Radon measure, then $Y^{(d)} * T \in L^\infty_{\text{loc}}(\mathbb{R}^d)$;

C. if T is of order at most m with $m \geq 1$, then $Y^{(d)} * T$ is a distribution of order at most $m - 1$.

ii. If $r \in \mathbb{N}^*$, set $Y_r^{(d)} = Y^{(d)} * \dots * Y^{(d)}$ ($Y^{(d)}$ appears r times). Show that, if T is a distribution with compact support of order at most m (with $m \geq 0$), then $Y_{m+2}^{(d)} * T \in C(\mathbb{R}^d)$.

c. Let T be a distribution on an open Ω of $\mathbb{R}^d$ such that, for every $p \in \mathbb{N}^d$,

$$\max_{1 \leq j \leq d} p_j \leq 1 \implies D^p T \in L^1_{\text{loc}}(\Omega).$$

Show that $T \in C(\Omega)$.

Hint. Take $\varphi \in \mathscr{D}(\Omega)$. Show that $D_1 \dots D_d(\varphi T) \in L^1(\mathbb{R}^d)$ and use parts a-i and b-i.

d. Let T be a distribution on an open Ω in $\mathbb{R}^d$. Suppose that T and its derivatives of all orders have order at most m. Show that $T \in \mathscr{E}(\Omega)$.

Hint. Start by showing that $T \in C(\Omega)$ using parts a-i and b-ii; then consider the derivatives of T.

9. Let J be a nonempty subset of $\{1, \dots, d\}$ and set $D_J = \prod_{j \in J} D_j$. Show that D_J is not hypoelliptic if $d \geq 2$.

Hint. Use Exercise 4 on page 323.

10. *Local and global structures of a distribution.*

a. Let T be a distribution with compact support of order k on $\mathbb{R}^d$. Show that there exists a continuous function f such that

$$D_1^{k+2} D_2^{k+2} \dots D_d^{k+2} [f] = T.$$

Hint. Use Exercise 4 on page 323 or Exercise 8b-ii above.

b. Let T be a distribution on $\mathbb{R}^d$. Show that, for every compact subset K of $\mathbb{R}^d$, there exists a continuous function f and a multiindex $p \in \mathbb{N}^d$ such that T coincides with $D^p[f]$ on $\mathscr{D}_K$.

c. Let T be a distribution on $\mathbb{R}^d$. Show that there exists a sequence $(p_n)_{n \in \mathbb{N}}$ of multiindices and a sequence $(f_n)_{n \in \mathbb{N}}$ of continuous functions on $\mathbb{R}^d$ such that

$$T = \sum_{n=0}^{\infty} D^{p_n} f_n.$$

Hint. Let $(\Omega_n)_{n\in\mathbb{N}}$ be a locally finite covering of $\mathbb{R}^d$ by bounded open sets and take a C^∞ partition of unity (φ_n) subordinate to this covering (see Exercise 14 on page 267). Apply the result of the first part of this exercise to the distributions $\varphi_n T$.

11. We work in $\mathbb{R}^d$, where $d \geq 3$.

a. Show that the only harmonic function that tends to 0 at infinity is the zero function.

Hint. Use the mean value property from Exercise 1.

b. Take $\varphi \in L_c^\infty(\mathbb{R}^d)$ (this space was defined in Exercise 19 on page 159) and let E be the fundamental solution of Δ. Show that the *Poisson equation*

$$\Delta f = \varphi$$

has a unique solution in $C_0(\mathbb{R}^d)$, namely $f = E * \varphi$. Show that f is of class C^1 on $\mathbb{R}^d$ and harmonic on $\mathbb{R}^d \setminus \operatorname{Supp}\varphi$. Show that if, in addition, $D_1\varphi, \dots, D_d\varphi \in L^\infty(\mathbb{R}^d)$, then $\varphi \in C_c(\mathbb{R}^d)$, $f \in \mathscr{E}^2(\mathbb{R}^d)$, and $\Delta f = \varphi$ in the ordinary sense.

12. Solve in $\mathscr{D}'_+$ the following equation in T:

$$(Y(x)\sin x) * T = S,$$

where $S \in \mathscr{D}'_+$. Under what condition on S is the distribution T defined by a locally integrable function?

Hint. Find a differential operator of which $Y(x)\sin x$ is a fundamental solution.

10

The Laplacian on an Open Set

Conventions. In this whole chapter, Ω will denote an open subset of $\mathbb{R}^d$. The elements of $L^1_{\text{loc}}(\Omega)$ will always be identified with the distributions they define on Ω. Recall that, for every $p \in [1, \infty]$, $L^p(\Omega) \subset L^1_{\text{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$. Differentiation operators should always be understood in the sense of distributions, unless otherwise stated.

If $f \in L^1_{\text{loc}}(\Omega)$, we denote by ∇f the gradient of f:

$$\nabla f = (D_1 f, \dots, D_d f).$$

If all derivatives $D_1 f, \dots, D_d f$ belong to $L^1_{\text{loc}}(\Omega)$, we also write

$$|\nabla f| = \Big(\sum_{j=1}^{d} |D_j f|^2 \Big)^{1/2}.$$

If $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ are elements of $\mathbb{C}^d$, we write

$$x \cdot y = \sum_{j=1}^{d} x_j y_j.$$

1 The spaces $H^1(\Omega)$ and $H^1_0(\Omega)$

The Sobolev spaces $W^{1,p}(\mathbb{R}^d)$ over $\mathbb{R}^d$, and in particular the space $H^1(\mathbb{R}^d) = W^{1,2}(\mathbb{R}^d)$, were defined in Chapter 9, on page 338. One can define in an analogous way the same spaces over an arbitrary open set Ω. In particular, $H^1(\Omega)$ is the space consisting of elements $f \in L^2(\Omega)$ all of whose first derivatives $D_1 f, \ldots, D_d f$ belong to $L^2(\Omega)$. This space is given the scalar product $(\cdot\,,\cdot)_{H^1(\Omega)}$ defined by

$$(f \mid g)_{H^1(\Omega)} = (f \mid g)_{L^2(\Omega)} + \sum_{j=1}^{d} (D_j f \mid D_j g)_{L^2(\Omega)} = \int_\Omega f\bar{g}\,dx + \int_\Omega \nabla f \cdot \overline{\nabla g}\,dx.$$

The norms on the spaces $L^2(\Omega)$ and $H^1(\Omega)$ will be denoted by $\|\cdot\|_{L^2(\Omega)}$ and $\|\cdot\|_{H^1(\Omega)}$, or simply by $\|\cdot\|_{L^2}$ and $\|\cdot\|_{H^1}$ if there is no danger of confusion.

Imitating the proof of Proposition 3.2 on page 338 we obtain this result:

Proposition 1.1 *The scalar product $(\cdot\,,\cdot)_{H^1(\Omega)}$ makes $H^1(\Omega)$ into a Hilbert space.*

In dimension $d = 1$, the Sobolev space $H^1(\Omega)$ has certain particular properties.

Proposition 1.2 *Suppose that $\Omega = (a, b)$, with $-\infty \le a < b \le +\infty$. Every element f of $H^1(\Omega)$ has a continuous representative on Ω (still denoted by f) that has finite limits at a and b. Moreover, if $a = -\infty$, we have $\lim_{x\to a} f(x) = 0$; similarly, if $b = +\infty$, we have $\lim_{x\to b} f(x) = 0$.*

Proof. By Theorem 2.8 on page 297, every element of $H^1(\Omega)$ has a continuous representative f satisfying, for $\alpha \in \Omega$,

$$f(t) = f(\alpha) + \int_\alpha^t f'(u)\,du \quad \text{for all } t \in \Omega. \tag{$*$}$$

If, for example, $b < +\infty$, then $L^2((\alpha, b)) \subset L^1((\alpha, b))$, so $f' \in L^1((\alpha, b))$. Therefore f does have a finite limit at b. Similarly, f certainly has a finite limit at a if $a > -\infty$.

Now suppose that $b = +\infty$. Multiplying equality $(*)$ by $f'(t)$ and integrating the resulting equality between α and x, we conclude that, if $x > \alpha$,

$$\int_\alpha^x f(t) f'(t)\,dt = f(\alpha)\bigl(f(x) - f(\alpha)\bigr) + \int_\alpha^x \Bigl(\int_\alpha^t f'(u)\,du\Bigr) f'(t)\,dt.$$

By Fubini's Theorem,

$$\begin{aligned}\int_\alpha^x \Bigl(\int_\alpha^t f'(u)\,du\Bigr) f'(t)\,dt &= \iint_{[\alpha,x]^2} 1_{\{u\le t\}} f'(u) f'(t)\,du\,dt \\ &= \frac{1}{2}\iint_{[\alpha,x]^2} f'(u) f'(t)\,du\,dt = \frac{1}{2}\bigl(f(x) - f(\alpha)\bigr)^2.\end{aligned}$$

We deduce that

$$f^2(x) = f^2(\alpha) + 2\int_\alpha^x f(t) f'(t)\,dt. \tag{$**$}$$

Since the two functions f and f' belong to $L^2(\Omega)$, their product ff' belongs to $L^1(\Omega)$ by the Schwarz inequality; therefore, by (**), f^2 has a finite limit at $+\infty$. Since $f^2 \in L^1(\Omega)$, this limit can only be 0. The reasoning is similar if $a = -\infty$. □

Remarks

1. If $\Omega = \mathbb{R}$, we recover the inclusion $H^1(\mathbb{R}) \subset C_0(\mathbb{R})$, which is a particular case of the Sobolev Injection Theorem (Theorem 3.3 on page 339).
2. This result does not generalize to the case $d \geq 2$: if $d \geq 2$, there exist elements of $H^1(\mathbb{R}^d)$ having no continuous representative (see Exercise 5 on page 346).

When Ω is a bounded interval in $\mathbb{R}$, an interesting denseness result holds.

Proposition 1.3 *Suppose that $\Omega = (a, b)$, with $-\infty < a < b < +\infty$. Then $C^1(\overline{\Omega})$ is a dense subspace of $H^1(\Omega)$.*

Proof. Clearly $C^1([a, b])$ is a subspace of $H^1(\Omega)$. Consider an element of $H^1(\Omega)$, having a continuous representative f. By the preceding proposition (and Theorem 2.8 on page 297), f has a continuous extension to $[a, b]$ and

$$f(x) = f(a) + \int_a^x f'(t)\,dt \quad \text{for all } x \in [a, b].$$

Since $C_c(\Omega)$ is dense in $L^2(\Omega)$, the derivative f' is the limit in $L^2(\Omega)$ of a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of elements of $C_c(\Omega)$. For each $n \in \mathbb{N}$, set

$$f_n(x) = f(a) + \int_a^x \varphi_n(t)\,dt.$$

Clearly $f_n \in C^1([a, b])$ and, for every $x \in [a, b]$,

$$\left|f(x) - f_n(x)\right| \leq \int_a^b \left|f'(t) - \varphi_n(t)\right| dt \leq \sqrt{b-a}\, \left\|f' - \varphi_n\right\|_{L^2},$$

by the Schwarz inequality. Thus $(f_n)_{n\in\mathbb{N}}$ is a sequence in $C^1([a, b])$ that converges uniformly, and so in $L^2(\Omega)$, to the element f. Since, in addition, $f_n' = \varphi_n$ for every n, which implies that the sequence $(f_n')_{n\in\mathbb{N}}$ converges to f' in $L^2(\Omega)$, we deduce that $(f_n)_{n\in\mathbb{N}}$ converges to f in $H^1(\Omega)$. □

The Space $H_0^1(\Omega)$

This denseness theorem of $C^1(\overline{\Omega})$ in $H^1(\Omega)$ remains valid if Ω is a bounded open subset of $\mathbb{R}^d$ under an additional regularity assumption (for example, that Ω be "of class C^1"). We will not use this result; instead we will introduce a subspace of $H^1(\Omega)$ in which $\mathscr{D}(\Omega)$ is dense, namely the space $H_0^1(\Omega)$, which is by definition the closure of $\mathscr{D}(\Omega)$ in $H^1(\Omega)$. The space

$H_0^1(\Omega)$, with the scalar product $(\cdot,\cdot)_{H^1}$, is a Hilbert space by construction, being a closed subspace of the Hilbert space $H^1(\Omega)$. In the case $d=1$, the relationship between $H_0^1(\Omega)$ and $H^1(\Omega)$ is simple:

Proposition 1.4 *If $\Omega=(a,b)$, with $-\infty \le a < b \le +\infty$, then*

$$H_0^1(\Omega) = H^1(\Omega)\cap C_0(\Omega).$$

In other words, $H_0^1(\Omega)$ consists of those elements of $H^1(\Omega)$ whose continuous representative tends to 0 at the boundary of Ω in $[-\infty,+\infty]$.

Proof. Consider an element of $H_0^1(\Omega)$ having f as its continuous representative and let $(\varphi_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{D}(\Omega)$ that converges to f in $H^1(\Omega)$. Since the sequence $(\varphi_n)_{n\in\mathbb{N}}$ converges to f in $L^2(\Omega)$, it has a subsequence that converges to f almost everywhere. Replacing it by a subsequence if necessary, we can suppose that there is a point $\alpha\in\Omega$ such that the sequence $(\varphi_n(\alpha))_{n\in\mathbb{N}}$ converges to $f(\alpha)$. Then, by the proof of Proposition 1.2 (and particularly by Equation (**), with f replaced by $f-\varphi_n$), we have, for every $x\in\Omega$,

$$\left|f(x)-\varphi_n(x)\right|^2 \le \left|f(\alpha)-\varphi_n(\alpha)\right|^2+2\left|\int_\alpha^x \left|f(t)-\varphi_n(t)\right|\left|f'(t)-\varphi_n'(t)\right|dt\right|,$$

which proves, by the Schwarz inequality, that the sequence $(\varphi_n)_{n\in\mathbb{N}}$ converges uniformly to f on Ω, and so that $f\in C_0(\Omega)$.

At the same time, since $C_c(\Omega)$ is dense in $L^2(\Omega)$ by Proposition 2.6 on page 107, since $\mathscr{D}(\Omega)$ is dense in $C_c(\Omega)=\mathscr{D}^0(\Omega)$ by Corollary 1.3 on page 262, and since there is a continuous injection from $C_c(\Omega)$ into $L^2(\Omega)$ (so that convergence in $C_c(\Omega)$ implies convergence in $L^2(\Omega)$), we deduce that the space $\mathscr{D}(\Omega)$ is dense in $L^2(\Omega)$.

Now consider $f\in H^1(\Omega)\cap C_0(\Omega)$. First suppose that $-\infty<a<b<+\infty$. Then, for every $x\in[a,b]$, we have $f(x)=\int_a^x f'(t)\,dt$ and, in particular, $\int_a^b f'(t)\,dt=0$. Let $(\psi_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{D}(\Omega)$ that tends to f' in $L^2(\Omega)$, and set $\lambda_n=\int_a^b \psi_n(t)\,dt$ for $n\in\mathbb{N}$. Then

$$\lim_{n\to+\infty}\lambda_n=\int_a^b f'(t)\,dt=0.$$

Take $\chi\in\mathscr{D}(\Omega)$ such that $\int_a^b \chi(t)\,dt=1$ and define φ_n by

$$\varphi_n(x)=\int_a^x \bigl(\psi_n(t)-\lambda_n\chi(t)\bigr)\,dt.$$

We then check that $\varphi_n\in\mathscr{D}(\Omega)$, that the sequence $(\varphi_n')_{n\in\mathbb{N}}$ converges to f' in $L^2(\Omega)$, and that $(\varphi_n)_{n\in\mathbb{N}}$ converges to f uniformly and so in $L^2(\Omega)$. Thus $f\in H_0^1(\Omega)$, which proves the desired result if Ω is bounded.

Finally, suppose that, for example, $a=-\infty$ and $b<+\infty$. Let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{E}((-\infty,b))$ such that, for every $n\in\mathbb{N}$, $0\le\rho_n\le1$,

$\rho_n = 1$ on $[-n, b)$, $\rho_n = 0$ on $(-\infty, -n-2]$, and $|\rho_n'| \leq 1$. By the Dominated Convergence Theorem, we see that, on the one hand, the sequence $(\rho_n f)_{n\in\mathbb{N}}$ converges to f in $L^2(\Omega)$ and that, on the other, since $(\rho_n f)' = \rho_n' f + \rho_n f'$, the sequence $((\rho_n f)')_{n\in\mathbb{N}}$ tends to f' in $L^2(\Omega)$. Therefore $(\rho_n f)_{n\in\mathbb{N}}$ converges to f in $H^1(\Omega)$. Moreover, since $(\rho_n f)(b) = 0$ and $(\rho_n f)(x) = 0$ for $x \leq -n-2$, we can, by the result of the preceding paragraph, approximate each $\rho_n f$ by elements of $\mathscr{D}((-n-2, b))$ in $H^1((-n-2, b))$, and so also in $H^1(\Omega)$. Thus the result is proved in this case. The cases $a > -\infty$, $b = +\infty$, and $a = -\infty$, $b = +\infty$ are analogous. □

Proposition 1.2 says that $H^1(\mathbb{R}) \subset C_0(\mathbb{R})$, so Proposition 1.4 implies that $H^1(\mathbb{R}) = H_0^1(\mathbb{R})$. The next proposition shows that this remains true in all dimensions. It is also clear from Proposition 1.4 that $H^1(\Omega) \neq H_0^1(\Omega)$ if Ω is an interval distinct from $\mathbb{R}$. Intuitively, if Ω is a bounded open set in $\mathbb{R}^d$, the elements of $H_0^1(\Omega)$ are, as in the case $d = 1$, those elements of $H^1(\Omega)$ that "vanish on the boundary of Ω". In dimension 1, this expression makes sense since the elements of $H^1(\Omega)$ have a continuous representative. In higher dimensions, the elements of $H^1(\Omega)$ are only defined almost everywhere, so talking about their value on the boundary of Ω, which generally has measure zero, makes no sense a priori. Nonetheless, it is possible, if Ω is sufficiently regular, to define the value of an element of $H^1(\Omega)$ at the boundary of Ω. We will not do this; in this regard see Exercises 16, 17, and 18 below.

Proposition 1.5 *The spaces $H^1(\mathbb{R}^d)$ and $H_0^1(\mathbb{R}^d)$ coincide.*

Proof. We must show that $\mathscr{D}(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$. Take $\xi \in \mathscr{D}(\mathbb{R}^d)$ such that $\xi(0) = 1$. For each $n \in \mathbb{N}^*$, put $\xi_n(x) = \xi(x/n)$. If $f \in H^1(\mathbb{R}^d)$, then $\xi_n f \in H^1(\mathbb{R}^d)$ and, by the Dominated Convergence Theorem, the sequence $(\xi_n f)_{n\in\mathbb{N}^*}$ converges to f in $L^2(\mathbb{R}^d)$. Moreover, for each $j \in \{1, \ldots, d\}$,

$$D_j(\xi_n f)(x) = \xi_n(x) D_j f(x) + \frac{1}{n}(D_j\xi)\Big(\frac{x}{n}\Big) f(x),$$

so, again by Dominated Convergence, the sequence $(D_j(\xi_n f))_{n\in\mathbb{N}}$ converges to $D_j f$ in $L^2(\mathbb{R}^d)$. Therefore the sequence $(\xi_n f)_{n\in\mathbb{N}^*}$ converges to f in $H^1(\mathbb{R}^d)$. Consequently, the space $H_c^1(\mathbb{R}^d)$ consisting of elements of $H^1(\mathbb{R}^d)$ with compact support is dense in $H^1(\mathbb{R}^d)$.

Now take $f \in H_c^1(\mathbb{R}^d)$ and let $(\chi_n)_{n\in\mathbb{N}}$ be a smoothing sequence. Then, for every $n \in \mathbb{N}$, the convolution $f * \chi_n$ is a function of class C^∞ on $\mathbb{R}^d$ (this follows from the theorems on differentiation under the summation sign) whose support is contained in $\operatorname{Supp} f + \operatorname{Supp} \chi_n$. Therefore $f * \chi_n \in \mathscr{D}(\mathbb{R}^d)$. On the other hand, by Proposition 3.7 on page 174, the sequence $(\chi_n * f)_{n\in\mathbb{N}}$ converges to f in $L^2(\mathbb{R}^d)$. Since, in addition, $D_j(f * \chi_n) = (D_j f) * \chi_n$ (by Corollary 2.10 on page 330 and Proposition 2.12 on page 331), we again deduce from Proposition 3.7 on page 174 that the sequence $(D_j(f * \chi_n))_{n\in\mathbb{N}}$

converges to $D_j f$ in $L^2(\mathbb{R}^d)$. It follows that $(f * \chi_n)_{n\in\mathbb{N}}$ converges to f in $H^1(\mathbb{R}^d)$, which concludes the proof. $\square$

The next inequality, applicable when Ω is bounded, will play an important role later.

Proposition 1.6 (Poincaré inequality) *If Ω is a bounded open set in $\mathbb{R}^d$ (more generally, if one of the projections of Ω on the coordinate axes is bounded), there exists a constant $C \geq 0$ depending only on Ω and such that*

$$\|u\|_{L^2(\Omega)} \leq C \,\big\||\nabla u|\big\|_{L^2(\Omega)} \quad \textit{for all } u \in H_0^1(\Omega).$$

If Ω is bounded, we can take $C = d(\Omega)$.

Proof. By denseness, we just have to show the inequality for every $u \in \mathscr{D}(\Omega)$, that is, for every $u \in \mathscr{D}(\mathbb{R}^d)$ such that $\operatorname{Supp} u \subset \Omega$. Suppose for example that the projection on Ω onto the first factor is bounded, so there exist real numbers $A < B$ such that $\Omega \subset [A, B] \times \mathbb{R}^{d-1}$. Since u is of class C^1,

$$u(x) = \int_A^{x_1} \frac{\partial u}{\partial x_1}(t, x_2, \ldots, x_d)\, dt \quad \text{for all } x \in \Omega.$$

It follows, by the Schwarz inequality, that

$$\big|u(x)\big|^2 \leq (B - A) \int_A^B \left|\frac{\partial u}{\partial x_1}\right|^2 (t, x_2, \ldots, x_d)\, dt \quad \text{for all } x \in \Omega.$$

Integrating this inequality over $[A, B] \times \mathbb{R}^{d-1}$ gives

$$\|u\|_{L^2}^2 \leq (B - A)^2 \,\|D_1 u\|_{L^2}^2.$$

Since $|D_1 u| \leq |\nabla u|$, the result is proved. $\square$

It follows in particular that, if Ω is a bounded open set, nonzero constant functions belong to $H^1(\Omega)$ but not to $H_0^1(\Omega)$; thus $H_0^1(\Omega) \neq H^1(\Omega)$.

The Poincaré inequality can be interpreted in the following way:

Corollary 1.7 *Suppose that Ω is a bounded open set (more generally, that one of the coordinate projections of Ω is bounded). The map*

$$u \mapsto \|u\|_{H_0^1(\Omega)} = \big\||\nabla u|\big\|_{L^2(\Omega)}$$

is a Hilbert norm on $H_0^1(\Omega)$ equivalent to the norm $\|\cdot\|_{H^1(\Omega)}$.

Proof. If C is the constant that appears in the Poincaré inequality, we have, for every $u \in H_0^1(\Omega)$,

$$\big\||\nabla u|\big\|_{L^2(\Omega)}^2 \leq \|u\|_{L^2(\Omega)}^2 + \big\||\nabla u|\big\|_{L^2(\Omega)}^2 = \|u\|_{H^1(\Omega)}^2 \leq (1 + C^2)\big\||\nabla u|\big\|_{L^2(\Omega)}^2,$$

which proves the equivalence between the two norms. At the same time, we see that the norm $\|u\|_{H_0^1(\Omega)}$ is defined by the scalar product

$$(u \mid v)_{H_0^1(\Omega)} = \sum_{j=1}^{d} (D_j u \mid D_j v)_{L^2(\Omega)}. \qquad \square$$

Proposition 1.8 *For $f \in H_0^1(\Omega)$, set*

$$\tilde{f} = \begin{cases} f & \text{on } \Omega, \\ 0 & \text{on } \mathbb{R}^d \setminus \Omega. \end{cases}$$

Then $\tilde{f} \in H^1(\mathbb{R}^d)$ and the map that takes f to $\tilde{f}$ is an isometry between $\big(H_0^1(\Omega), \|\cdot\|_{H^1(\Omega)}\big)$ and $\big(H^1(\mathbb{R}^d), \|\cdot\|_{H^1(\mathbb{R}^d)}\big)$.

Proof. If $f \in \mathscr{D}(\Omega)$, we clearly have $\tilde{f} \in \mathscr{D}(\mathbb{R}^d)$ and $\widetilde{D_j f} = D_j \tilde{f}$ for $j \in \{1, \dots, d\}$. Consequently, the map $f \mapsto \tilde{f}$ is an isometry from $\mathscr{D}(\Omega)$, with the norm $\|\cdot\|_{H^1(\Omega)}$, into $H^1(\mathbb{R}^d)$. Since $H^1(\mathbb{R}^d)$ is complete, the extension theorem says that this isometry extends to an isometry $f \mapsto \hat{f}$ from $\big(H_0^1(\Omega), \|\cdot\|_{H^1(\Omega)}\big)$ to $\big(H^1(\mathbb{R}^d), \|\cdot\|_{H^1(\mathbb{R}^d)}\big)$. Now, convergence in H^1 implies convergence in L^2, so, if $(\varphi_n)_{n\in\mathbb{N}}$ is a sequence in $\mathscr{D}(\Omega)$ converging to f in $H_0^1(\Omega)$, the sequence $(\tilde{\varphi}_n)_{n\in\mathbb{N}}$ converges to $\tilde{f}$ in $L^2(\mathbb{R}^d)$. Since it also converges to $\hat{f}$ in $H^1(\mathbb{R}^d)$ (by the definition of $\hat{f}$) and so also in $L^2(\mathbb{R}^d)$, it follows that $\tilde{f} = \hat{f}$, which concludes the proof. $\square$

Lemma 1.9 *For every $u \in H^1(\mathbb{R}^d)$ and every $h \in \mathbb{R}^d$,*

$$\|\tau_h u - u\|_{L^2} \leq \big\||\nabla u|\big\|_{L^2} \, |h|.$$

Proof. By Proposition 1.5, the space $\mathscr{D}(\mathbb{R}^d)$ is dense in $H^1(\mathbb{R}^d)$. Thus it suffices to prove the property for $u \in \mathscr{D}(\mathbb{R}^d)$. If $u \in \mathscr{D}(\mathbb{R}^d)$, we have

$$u(x-h) - u(x) = -\int_0^1 \nabla u(x - th) \cdot h \, dt;$$

thus, by the Schwarz inequality,

$$\big|\tau_h u(x) - u(x)\big|^2 \leq |h|^2 \int_0^1 |\nabla u|^2 (x - th) \, dt.$$

Now it suffices to integrate this inequality over $\mathbb{R}^d$ using the fact that Lebesgue measure is invariant under translations. $\square$

We now derive from the preceding results an important compactness theorem.

Theorem 1.10 (Rellich) *If Ω is a bounded open set, the canonical injection $u \mapsto u$ from $H_0^1(\Omega)$ into $L^2(\Omega)$ is a compact operator.*

In other words, every bounded sequence in $H_0^1(\Omega)$ has a convergent subsequence in $L^2(\Omega)$.

Proof. Since the map $u \mapsto u_{|\Omega}$ from $L^2(\mathbb{R}^d)$ to $L^2(\Omega)$ is clearly continuous, it suffices to prove that the map $f \mapsto \tilde{f}$ from $H_0^1(\Omega)$ to $L^2(\mathbb{R}^d)$ is a compact operator (where $\tilde{f}$ is as in Proposition 1.8). Let B be the closed unit ball in $H_0^1(\Omega)$, and put $\tilde{B} = \{\tilde{f} : f \in B\}$. By Proposition 1.8, $\tilde{B}$ is contained in the closed unit ball of $H^1(\mathbb{R}^d)$. We must show that $\tilde{B}$ is relatively compact in $L^2(\mathbb{R}^d)$. To do this, we use the criterion provided by Theorem 3.8 on page 175 in the case $p = 2$.

Properties i and ii in the statement of that theorem are clearly satisfied since, for every $f \in B$, we have $\|\tilde{f}\|_{L^2} \le 1$ and

$$\int_{\{|x|>R\}} |\tilde{f}(x)|^2 \, dx = 0$$

for every $R > 0$ such that $\Omega \subset B(0, R)$. On the other hand, $\tilde{B}$ is contained in the closed unit ball of $H^1(\mathbb{R}^d)$; thus, by Lemma 1.9,

$$\left\|\tau_h \tilde{f} - \tilde{f}\right\|_{L^2} \le |h| \quad \text{for all } f \in B \text{ and } h \in \mathbb{R}^d,$$

which proves property iii. □

Exercises

Here Ω is still an open subset of $\mathbb{R}^d$.

1. Show that, if $u \in H^1(\Omega)$ and $v \in H_0^1(\Omega)$, then

$$(D_j u \,|\, v)_{L^2} = -(u \,|\, D_j v)_{L^2} \quad \text{for } j \in \{1, \dots, d\}.$$

2. Let $(u_n)_{n \in \mathbb{N}}$ be a sequence in $H^1(\Omega)$ that converges in $L^2(\Omega)$ to an element $u \in L^2(\Omega)$ and such that, for each $j \in \{1, \dots, d\}$, the sequence $(D_j u_n)_{n \in \mathbb{N}}$ converges in $L^2(\Omega)$ to an element $v_j \in L^2(\Omega)$. Show that $u \in H^1(\Omega)$, that $(u_n)_{n \in \mathbb{N}}$ converges to u in $H^1(\Omega)$, and that $v_j = D_j u$ for each $j \in \{1, \dots, d\}$.
Hint. Show that $(u_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $H^1(\Omega)$.

3. Let $(u_n)_{n \in \mathbb{N}}$ be a bounded sequence in $H^1(\Omega)$ that converges in $L^2(\Omega)$ to $u \in L^2(\Omega)$. Show that $u \in H^1(\Omega)$ and that there exists a subsequence $(u_{n_k})_{k \in \mathbb{N}}$ such that

$$\lim_{k \to +\infty} \frac{u_{n_0} + \cdots + u_{n_k}}{k + 1} = u$$

in $H^1(\Omega)$.
Hint. Use the Banach–Saks Theorem (Exercise 16 on page 121).

4. Let $H_c^1(\Omega)$ be the set of elements of $H^1(\Omega)$ having compact support. Prove that $H_c^1(\Omega)$ is a dense subspace of $H_0^1(\Omega)$.

5. Show that the canonical injection from $H^1(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ is not compact.

6. Suppose $u \in H^1(\Omega)$. Show that there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathscr{D}(\Omega)$ that converges to u in $L^2(\Omega)$ and is such that, for every $j \in \{1,\dots,d\}$, the sequence $(D_j u_n)_{n\in\mathbb{N}}$ converges to $D_j u$ in $L^2_{\text{loc}}(\Omega)$. (Convergence in $L^2_{\text{loc}}(\Omega)$ was defined in Exercise 19 on page 159.)

7. Suppose that $\Omega = (a,b)$ is a bounded interval in $\mathbb{R}$. Show that the best constant in the Poincaré inequality is $(b-a)/\pi$.
Hint. Use Wirtinger's inequality, Exercise 16d on page 137.

8. *The Meyers–Serrin Theorem.* Let Ω be an open subset of $\mathbb{R}^d$. Show that $\mathscr{E}(\Omega) \cap H^1(\Omega)$ is dense in $H^1(\Omega)$.
Hint. Let $(\Omega_j)_{j\in\mathbb{N}}$ be a family of relatively compact open subsets of Ω covering Ω and such that $\Omega_0 = \varnothing$ and $\overline{\Omega}_j \subset \Omega_{j+1}$ for every $j \in \mathbb{N}$. Let $(\varphi_j)_{j\in\mathbb{N}}$ be a partition of unity relative to the family of open sets $(\Omega_{j+2} \setminus \overline{\Omega}_j)_{j\in\mathbb{N}}$ (see Exercise 14 on page 267). Finally, take a smoothing sequence $(\chi_n)_{n\in\mathbb{N}}$ on $\mathbb{R}^d$, an element $u \in H^1(\Omega)$, and $\varepsilon > 0$. Show that, for every $j \in \mathbb{N}$, there exists an integer $n_j \in \mathbb{N}$ such that

$$\operatorname{Supp}(\chi_{n_j} * (\varphi_j u)) \subset \Omega_{j+2} \setminus \overline{\Omega}_j \quad \text{and} \quad \big\|(\chi_{n_j} * (\varphi_j u)) - \varphi_j u\big\|_{H^1(\Omega)} \le \varepsilon 2^{-j-1}.$$

Then consider $v = \sum_{j\in\mathbb{N}} \chi_{n_j} * (\varphi_j u)$.

9. *The Poincaré inequality in $H^1(\Omega)$.* (This result generalizes Exercise 16b on page 137.) Suppose Ω is a bounded and convex open set in $\mathbb{R}^d$.

a. Show that the relation

$$Tf(x) = \int_\Omega |x-y|^{1-d} f(y)\, dy$$

defines a continuous linear operator T from $L^2(\Omega)$ to $L^2(\Omega)$.
Hint. Use the Young inequality in $\mathbb{R}^d$ (see page 172 and also Exercise 9b on page 182).

b. Take $u \in C^1(\Omega) \cap L^1(\Omega)$, and put $m(u) = \dfrac{1}{\operatorname{vol}\Omega} \displaystyle\int_\Omega u(x)\, dx$, $\delta = d(\Omega)$. Show that, for every $x \in \Omega$,

$$\big|u(x) - m(u)\big| \le \frac{1}{\operatorname{vol}\Omega} \frac{\delta^d}{d} \int_\Omega |x-y|^{1-d} \big|\nabla u(y)\big|\, dy.$$

Hint. Prove, then integrate with respect to y, the equality

$$u(x) - u(y) = \int_0^{|x-y|} \frac{x-y}{|x-y|} \cdot \nabla u\left(x + t\frac{y-x}{|y-x|}\right) dt.$$

Deduce that

$$|u(x) - m(u)| \le \frac{1}{\operatorname{vol}\Omega} \int_{\{z\in\mathbb{R}^d : |z|\le\delta\}} \left(\int_0^{+\infty} \left|\nabla u\left(x + t\frac{z}{|z|}\right)\right| dt \right) dz,$$

where ∇u is extended with the value 0 outside Ω. Then use Theorem 3.9 on page 74 twice.

c. Deduce the existence of a constant $C > 0$ such that

$$\left\|u - m(u)\right\|_{L^2(\Omega)} \leq C\left\||\nabla u|\right\|_{L^2(\Omega)} \quad \text{for all } u \in H^1(\Omega). \qquad (*)$$

Hint. Show this inequality for every $u \in C^1(\Omega) \cap H^1(\Omega)$, then argue by denseness, using Exercise 8.

d. Show that the norm $|\cdot|$ defined on $H^1(\Omega)$ by

$$|u| = \left|\int_\Omega u(x)\,dx\right| + \left\||\nabla u|\right\|_{L^2(\Omega)}$$

is equivalent to the norm $\|\cdot\|_{H^1(\Omega)}$.

10. All functions considered here are real-valued.

a. Suppose $u \in H^1(\Omega)$. Show that there exists a sequence $(u_n)_{n\in\mathbb{N}}$ in $\mathscr{D}(\Omega)$ that converges to u almost everywhere and in $L^2(\Omega)$ and is such that, for every $j \in \{1, \dots, d\}$, the sequence $(D_j u_n)_{n\in\mathbb{N}}$ converges to $D_j u$ in $L^2_{\text{loc}}(\Omega)$ (see Exercise 6).

b. Let $G \in C^1(\mathbb{R})$ satisfy

$$G(0) = 0, \qquad \left|G'(t)\right| \leq M \quad \text{for all } t \in \mathbb{R}.$$

Show that, if $u \in H^1(\Omega)$, then $G \circ u \in H^1(\Omega)$ and

$$D_j(G \circ u) = (G' \circ u) D_j u$$

for every $j \in \{1, \dots, d\}$. In particular, $\|G \circ u\|_{H^1} \leq M\|u\|_{H^1}$.

Hint. Take an approximating sequence $(u_n)_{n\in\mathbb{N}}$ of u as in the first part of this exercise. Show that $(G \circ u_n)_{n\in\mathbb{N}}$ converges to $G \circ u$ in $L^2(\Omega)$ and, for every $j \in \{1, \dots, d\}$, the sequence whose general term is $D_j(G \circ u_n) = (G' \circ u_n) D_j u_n$ converges to $(G' \circ u) D_j u$ in $L^2_{\text{loc}}(\Omega)$.

c. Show that, if G is as above and $u \in H_0^1(\Omega)$, then $G \circ u \in H_0^1(\Omega)$.

Hint. Consider again the preceding proof and notice that, if $v \in \mathscr{D}(\Omega)$, then $G \circ v \in H_c^1(\Omega)$ (see exercise 4).

11. All functions considered here are real-valued.

a. Show that, for every $n \in \mathbb{N}$, there exists a function $G_n \in \mathscr{E}(\mathbb{R})$ such that $|G_n'(t)| \leq 1$ for all $t \in \mathbb{R}$ and

$$G_n(t) = \begin{cases} -1/2n & \text{if } t \leq -1/n, \\ t & \text{if } t \geq 0. \end{cases}$$

b. Suppose $u \in H^1(\Omega)$. Show that $u^+ \in H^1(\Omega)$ and that $D_j(u^+) = 1_{\{u\geq 0\}} D_j u$ for every $j \in \{1, \dots, d\}$.

Hint. Compute the limits in $L^2(\Omega)$ of the sequence $(G_n \circ u)_{n\in\mathbb{N}}$ and of the sequences $(D_j(G_n \circ u))_{n\in\mathbb{N}}$, for $j \in \{1, \dots, d\}$, using Exercise 10 for the latter. Then use Exercise 2.

c. Show that, if $u \in H^1(\Omega)$,

$$1_{\{u=0\}} D_j u = 0 \quad \text{for } j \in \{1, \ldots, d\}.$$

Hint. Compute $D_j(u^-)$.

d. Let $u \in H_0^1(\Omega)$. Show that $u^+ \in H_0^1(\Omega)$.
Hint. Let $(\varphi_n)_{n \in \mathbb{N}}$ be a sequence in $\mathscr{D}(\Omega)$ converging to u in $H_0^1(\Omega)$ and almost everywhere. Then consider the sequence $(G_n \circ \varphi_n)_{n \in \mathbb{N}}$.

e. Assume either that $u \in H^1(\Omega)$ or that $u \in H_0^1(\Omega)$. Show that $|u| \in H^1(\Omega)$ or $|u| \in H_0^1(\Omega)$, respectively, and that

$$D_j |u| = 1_{\{u \geq 0\}} D_j u - 1_{\{u<0\}} D_j u \quad \text{for } j \in \{1, \ldots, d\}.$$

Deduce that $\big\| |u| \big\|_{H^1} = \|u\|_{H^1}$.

f. Show that $H^1(\Omega)$ and $H_0^1(\Omega)$ are lattices.

12. All functions considered here are real-valued.

a. Let K be a compact subset of $\mathbb{R}$ of Lebesgue measure zero.

i. Show that there exists a sequence $(\varphi_n)_{n \in \mathbb{N}}$ in $C_c(\mathbb{R})$ such that, for every $n \in \mathbb{N}$, we have $0 \leq \varphi_n \leq 1$, $\varphi_n \geq \varphi_{n+1}$ and

$$\lim_{n \to +\infty} \varphi_n(t) = 1_K(t) \quad \text{for } t \in \mathbb{R}.$$

ii. Take $u \in H^1(\Omega)$ and suppose $\Phi_n(x) = \int_0^x \varphi_n(t)\, dt$. Show that, for every $n \in \mathbb{N}$, $\Phi_n \circ u \in H^1(\Omega)$, $(\Phi_n \circ u)_{n \in \mathbb{N}}$ converges to 0 in $H^1(\Omega)$, and

$$1_{\{u \in K\}} D_j u = 0 \quad \text{for } j \in \{1, \ldots, d\}.$$

(This generalizes Exercise 11c.)
Hint. Use Exercises 2 and 10.

b. Show that, if A is a Borel set in $\mathbb{R}$ with measure zero and $u \in H^1(\Omega)$, we have $\nabla u = 0$ almost everywhere on $u^{-1}(A)$.
Hint. Use the fact that the Radon measure μ defined by

$$\int \varphi\, d\mu = \int (\varphi \circ u)\, |\nabla u|^2\, dx \quad \text{for all } \varphi \in C_c(\mathbb{R})$$

is regular (see Exercise 5 on page 77).

13. All functions considered here are real-valued.

a. Let $G : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant M and satisfying $G(0) = 0$. Show that, if $u \in H^1(\Omega)$, then $G \circ u \in H^1(\Omega)$ and $\|G \circ u\|_{H^1} \leq M \|u\|_{H^1}$.
Hint. Approximate G by functions

$$G_n(x) = n\left(\int_x^{x+(1/n)} G(t)\, dt - \int_0^{1/n} G(t)\, dt\right)$$

and use Exercises 3 and 10.

b. Under the same assumptions on G, prove that $u \in H_0^1(\Omega)$ implies $G \circ u \in H_0^1(\Omega)$.
Hint. Show first that $u \in \mathscr{D}(\Omega)$ implies $G \circ u \in H_0^1(\Omega)$ (see Exercise 4). Then use Exercise 3.

14. Show that, if f and g belong to $H^1(\Omega) \cap L^\infty(\Omega)$, so does the product fg, and that

$$D_j(fg) = gD_jf + fD_jg \quad \text{for } j \in \{1, \dots, d\}.$$

Show that if, in addition, f and g belong to $H_0^1(\Omega)$, so does fg.
Hint. Using Exercise 10, prove first that, if $h \in H^1(\Omega) \cap L^\infty(\Omega)$, then $h^2 \in H^1(\Omega)$ and $D_j(h^2) = 2hD_jh$ (and similarly with $H_0^1(\Omega)$ instead of $H^1(\Omega)$).

15. Show that every positive element of $H_0^1(\Omega)$ is the limit in $H^1(\Omega)$ of a sequence of positive elements of $\mathscr{D}(\Omega)$.
Hint. If u is a positive element of $H_0^1(\Omega)$, there exists a sequence $(\psi_n)_{n\in\mathbb{N}}$ of real-valued elements of $\mathscr{D}(\Omega)$ that converges to u in $H^1(\Omega)$ and almost everywhere. Show that there exists a sequence $(G_n)_{n\in\mathbb{N}}$ in $\mathscr{E}(\mathbb{R})$ such that, for every $n \in \mathbb{N}$, we have $0 \le G_n' \le 1$ and

$$G_n(t) = \begin{cases} 0 & \text{if } t \le 0, \\ t - 1/n & \text{if } t \ge 2/n. \end{cases}$$

Show that the sequence $(G_n \circ \psi_n)_{n\in\mathbb{N}}$ converges to u in $H^1(\Omega)$ (use Exercise 11c).

16. We wish to show that $C_0(\Omega) \cap H^1(\Omega) \subset H_0^1(\Omega)$. Take $u \in C_0^{\mathbb{R}}(\Omega) \cap H^1(\Omega)$ and let φ be a function of class C^1 on $\mathbb{R}$ such that $\varphi = 0$ on $[-1, 1]$ and $\varphi = 1$ on $(-\infty, -2) \cup (2, +\infty)$.

a. For $n \in \mathbb{N}^*$, set $u_n = \varphi(nu)u$. Show that $u_n \in H^1(\Omega) \cap C_c(\Omega)$.
Hint. Use Exercise 10.

b. Show that the sequence $(u_n)_{n\in\mathbb{N}}$ converges to u in $H^1(\Omega)$.
Hint. Use Exercise 2.

c. Complete the proof.
Hint. Use Exercise 4.

(Note that this exercise yields another proof of the result in the one-dimensional case.)

17. Assume that $d = 2$ and that

$$\Omega = \{x \in \mathbb{R}^2 : 0 < |x| < 1\}.$$

Let φ be an element of $\mathscr{D}(\mathbb{R})$ such that $\varphi(t) = 1$ if $|t| \le 1/2$ and $\varphi(t) = 0$ if $|t| \ge 1$. For $n \in \mathbb{N}^*$ and $x \in \Omega$, put

$$u_n(x) = \varphi(2|x|)\varphi\left(\frac{1}{n}\log\frac{1}{|x|}\right)$$

and $u(x) = \varphi(2|x|)$. Show that $u_n \in \mathscr{D}(\Omega)$ for every $n \in \mathbb{N}^*$ and that the sequence $(u_n)_{n\in\mathbb{N}}$ converges to u in $H^1(\Omega)$. Deduce that

$$H_0^1(\Omega) \cap C(\Omega) \not\subset C_0(\Omega).$$

18. *The trace theorem in a half-space.* Assume that $d \geq 2$ and that $\Omega = (0, +\infty) \times \mathbb{R}^{d-1}$.

a. Show that the space $\mathscr{D}(\bar{\Omega})$ consisting of functions of class C^∞ on $\bar{\Omega}$ with compact support, is dense in $H^1(\Omega)$.
Hint. Argue as in the proof of Proposition 1.5, choosing a smoothing sequence $(\chi_n)_{n\in\mathbb{N}}$ consisting of elements of $\mathscr{D}(-\Omega)$.

b. Take $u \in \mathscr{D}(\bar{\Omega})$. Show that, for every $x \in \mathbb{R}^{d-1}$,

$$\left|u(0,x)\right|^2 \leq \int_0^{+\infty} \left(\left|u(x_1,x)\right|^2 + \left|\frac{\partial u}{\partial x_1}(x_1,x)\right|^2 \right) dx_1.$$

Deduce that

$$\left\|u(0,\,\cdot\,)\right\|_{L^2(\mathbb{R}^{d-1})} \leq \|u\|_{H^1(\Omega)}.$$

c. Show that there exists a unique continuous linear map γ_0 from $H^1(\Omega)$ to $L^2(\mathbb{R}^{d-1})$ such that

$$\gamma_0 u = u(0,\cdot\,) \quad \text{for all } u \in \mathscr{D}(\bar{\Omega}).$$

d. Show that

$$\gamma_0 u = u(0,\cdot\,) \quad \text{for all } u \in C(\bar{\Omega}) \cap H^1(\Omega).$$

Hint. Show that every element of $C(\bar{\Omega}) \cap H^1(\Omega)$ can be approximated in $C(\bar{\Omega})$ and in $H^1(\Omega)$ by a sequence of elements of $\mathscr{D}(\bar{\Omega})$.

e. *Green's formula in $H^1(\Omega)$.* Suppose $u, v \in H^1(\Omega)$. Show that

$$\langle v, D_j u\rangle_{L^2(\Omega)} = -\langle D_j v, u\rangle_{L^2(\Omega)} \quad \text{for all } j \in \{2, \dots, d\}$$

and that

$$\langle v, D_1 u\rangle_{L^2(\Omega)} = -\langle D_1 v, u\rangle_{L^2(\Omega)} - \langle \gamma_0 u, \gamma_0 v\rangle_{L^2(\mathbb{R}^{d-1})}.$$

Hint. Use the first part of the exercise.

f. **i.** If $u \in H^1(\Omega)$, denote by $\tilde{u}$ the extension of u to $\mathbb{R}^d$ having the value 0 outside Ω. Show that, for every $u \in \ker \gamma_0$,

$$\frac{\partial \tilde{u}}{\partial x_j} = \widetilde{\frac{\partial u}{\partial x_j}} \quad \text{for } j \in \{1, \dots, d\},$$

and so that $\tilde{u} \in H^1(\mathbb{R}^d)$.
Hint. Use the preceding question.

ii. Show that the space $\mathscr{D}(\Omega)$ is dense in $\ker\gamma_0$ (with respect to the norm $\|\cdot\|_{H^1(\Omega)}$).

g. Deduce from these results that $H_0^1(\Omega) = \ker\gamma_0$ and that

$$H_0^1(\Omega) \cap C(\overline{\Omega}) = \{u \in C(\overline{\Omega}) \cap H^1(\Omega) : u(0,\cdot\,) = 0\}.$$

(In this sense, we can say that $H_0^1(\Omega)$ consists of those elements of $H^1(\Omega)$ that vanish on the boundary of Ω.)

19. *The maximum principle for the Laplacian in* $H^1(\Omega)$. Suppose that Ω is bounded. All functions considered will be real-valued.

If $u \in H^1(\Omega)$, we say that $u \le 0$ on the boundary $\partial\Omega$ of Ω if $u^+ \in H_0^1(\Omega)$. (It was proved in Exercise 10 that $u^+ \in H^1(\Omega)$ for every $u \in H^1(\Omega)$.)

a. Show that, if $u \in C(\overline{\Omega}) \cap H^1(\Omega)$ and $u(x) \le 0$ for every $x \in \partial\Omega$, then $u \le 0$ on $\partial\Omega$ in the sense defined above.

Hint. Use Exercise 16.

b. Take $u \in H^1(\Omega)$. Show that, if $\Delta u \ge 0$ (that is, if Δu is a positive distribution), we have

$$u(x) \le \sup_{\partial\Omega} u \quad \text{for almost every } x \in \Omega,$$

where $\sup_{\partial\Omega} u = \inf\{l \in \mathbb{R} : u - l \le 0 \text{ on } \partial\Omega\}$.

Hint. Take $l \in \mathbb{R}$ such that $v = (u-l)^+ \in H_0^1(\Omega)$. Using Exercise 10 show that $\nabla v = 0$ on $\{u - l < 0\}$ and $\nabla v = \nabla u$ on $\{u - l \ge 0\}$. Deduce that

$$\big\||\nabla v|\big\|_{L^2(\Omega)}^2 = \int_\Omega \nabla u(x) \cdot \nabla v(x)\, dx \le 0$$

(using Exercise 15) and conclude the proof.

c. Take $u \in H^1(\Omega)$. Show that $\Delta u = 0$ implies

$$\inf_{\partial\Omega} u \le u(x) \le \sup_{\partial\Omega} u \quad \text{for almost every } x \in \Omega,$$

where $\inf_{\partial\Omega} u = -\sup_{\partial\Omega}(-u)$.

d. Show that these results remain true if we replace the assumptions $\Delta u \ge 0$ and $\Delta u = 0$ by, respectively, $\mathscr{L}u \ge 0$ and $\mathscr{L}u = 0$, where $\mathscr{L}$ is an *elliptic* homogeneous operator of order 2, that is, a linear operator on $H^1(\Omega)$ of the form

$$\mathscr{L}u = \sum_{1 \le i,j \le d} \frac{\partial}{\partial x_i}\left(a_{i,j}\frac{\partial}{\partial x_j}u\right),$$

with $a_{i,j} \in L^\infty_{\mathbb{R}}(\Omega)$ for $1 \le i, j \le d$, satisfying the condition that there exists $\alpha > 0$ such that, for almost every $x \in \Omega$,

$$\sum_{1 \le i,j \le d} a_{i,j}(x)\xi_i\xi_j \ge \alpha|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d.$$

20. Suppose that Ω is a bounded interval in $\mathbb{R}$. Show that the canonical injection of $H^1(\Omega)$ in $C(\overline{\Omega})$ is a compact operator. (This is a stronger result than Rellich's Theorem in this case.)
Hint. Show that, if $u \in H^1(\Omega)$ and $x, y \in \overline{\Omega}$,

$$\left|u(x) - u(y)\right| \leq |x - y|^{1/2} \, \|u\|_{H^1}.$$

(In other words, $H^1(\Omega)$ injects continuously in $C^{1/2}(\overline{\Omega})$, the space of Hölder functions of order $1/2$ on $\overline{\Omega}$; see Exercise 5 on page 45.) Then use Ascoli's Theorem or Exercise 5 on page 45.

21. Suppose that Ω is a bounded open subset of $\mathbb{R}^d$. Let $(u_n)_{n\in\mathbb{N}}$ be a sequence in $H_0^1(\Omega)$. Suppose that the sequence $(u_n)_{n\in\mathbb{N}}$ converges weakly in $L^2(\Omega)$ and that, for every $j \in \{1, \ldots, d\}$, the sequence $(\partial u_n/\partial x_j)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$. Show that the sequence $(u_n)_{n\in\mathbb{N}}$ converges strongly in $L^2(\Omega)$.
Hint. By Exercise 10a on page 120, the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in $H_0^1(\Omega)$. Then use Rellich's Theorem and Exercise 12 on page 121.

2 The Dirichlet Problem

We consider a bounded open subset Ω of $\mathbb{R}^d$. The space $H_0^1(\Omega)$ is from now on given the Hilbert space structure defined by the scalar product

$$(u \mid v)_{H_0^1} = \sum_{j=1}^{d} (D_j u \mid D_j v)_{L^2(\Omega)} = \int_\Omega \nabla u \cdot \overline{\nabla v} \, dx.$$

We denote by $\|\cdot\|_{H_0^1}$ the norm associated with this scalar product.

If $f \in L^2(\Omega)$, a **solution of the Dirichlet problem on Ω with right-hand side f** is, by definition, an element u of $H_0^1(\Omega)$ such that

$$\Delta u = f.$$

Classically, to impose a **Dirichlet condition** on the solution (in the ordinary sense) of a partial differential equation over Ω means stipulating the value of the solution on the boundary of Ω. In the present context, the condition $u \in H_0^1(\Omega)$ is of this type, since it amounts, in a sense already discussed, to requiring that u "vanish on the boundary of Ω".

The next proposition gives the so-called **variational formulation** of the Dirichlet problem, which underlies the Galerkin-type algorithms for numerical solution of the Dirichlet problem (see Exercise 1 on page 116).

Proposition 2.1 *If $f \in L^2(\Omega)$, these statements are equivalent:*

- *$u \in H_0^1(\Omega)$ and $\Delta u = f$.*

– *$u \in H_0^1(\Omega)$ and $(u \mid v)_{H_0^1} = -(f \mid v)_{L^2}$ for all $v \in H_0^1(\Omega)$.*

Proof. If $f \in L^2(\Omega)$ and $u \in H_0^1(\Omega)$, we have the following chain of equivalences:

$$\begin{aligned}\Delta u = f &\iff \langle \Delta u, \bar\varphi\rangle = (f \mid \varphi)_{L^2} \quad \text{for all } \varphi \in \mathscr{D}(\Omega)\\ &\iff \sum_{j=1}^{d}(D_j u \mid D_j\varphi)_{L^2} = -(f \mid \varphi)_{L^2} \quad \text{for all } \varphi \in \mathscr{D}(\Omega)\\ &\iff (u \mid \varphi)_{H_0^1} = -(f \mid \varphi)_{L^2} \quad \text{for all } \varphi \in \mathscr{D}(\Omega).\end{aligned}$$

The result follows because $\mathscr{D}(\Omega)$ is dense in $H_0^1(\Omega)$. □

This will allow us, in particular, to prove the existence and uniqueness of the solution of the Dirichlet problem.

Theorem 2.2 *For every $f \in L^2(\Omega)$, the Dirichlet problem on Ω with right-hand side f has a unique solution $u \in H_0^1(\Omega)$. The operator*

$$\begin{aligned}\Delta^{-1} \ :\ L^2(\Omega) &\to H_0^1(\Omega)\\ f &\mapsto u\end{aligned}$$

thus defined is continuous and has norm at most C, the constant that appears in the Poincaré inequality (Proposition 1.6).

Proof. If $f \in L^2(\Omega)$,

$$\bigl|(f \mid v)_{L^2}\bigr| \le C\|f\|_{L^2}\,\|v\|_{H_0^1} \quad \text{for all } v \in H_0^1(\Omega),$$

where C is the constant in the Poincaré inequality. Thus, the map $L : v \mapsto (v \mid f)_{L^2}$ is a continuous linear form on $H_0^1(\Omega)$ of norm at most $C\|f\|_{L^2}$. Therefore the existence and uniqueness of the solution u, together with the inequality $\|u\|_{H_0^1} \le C\|f\|_{L^2}$, follow immediately (in view of the preceding proposition) from an application of Riesz's Theorem (page 111) to the Hilbert space $H_0^1(\Omega)$. □

The Dirichlet problem can also be interpreted as a minimization problem:

Proposition 2.3 *Let $f \in L^2(\Omega)$. For every $v \in H_0^1(\Omega)$, put*

$$J_f(v) = \tfrac{1}{2}\bigl(\|v\|_{H_0^1}\bigr)^2 + \operatorname{Re}(f \mid v)_{L^2}.$$

These statements are equivalent:

– *$u \in H_0^1(\Omega)$ and $\Delta u = f$.*
– *$u \in H_0^1(\Omega)$ and $J_f(u) = \min_{v \in H_0^1(\Omega)} J_f(v)$.*

Proof. Suppose $h \in H_0^1(\Omega)$. Then

$$J_f(u+h) = J_f(u) + \operatorname{Re}\big((f \mid h)_{L^2} + (u \mid h)_{H_0^1}\big) + \tfrac{1}{2}\big(\|h\|_{H_0^1}\big)^2.$$

Therefore, by Proposition 2.1, $\Delta u = f$ implies that

$$J_f(u+h) = J_f(u) + \tfrac{1}{2}\big(\|h\|_{H_0^1}\big)^2 \quad \text{for all } h \in H_0^1(\Omega).$$

Thus J_f attains its minimum on $H_0^1(\Omega)$ at u and only at u. □

We now study the spectral properties of the Laplacian on Ω with "Dirichlet conditions". More precisely, we will say that a complex number λ is an **eigenvalue of the Dirichlet Laplacian** if there exists a nonzero $u \in H_0^1(\Omega)$ such that $\Delta u = \lambda u$. Such functions u are the **eigenfunctions** associated with the eigenvalue λ. The **eigenspace** associated with λ is the space of $u \in H_0^1(\Omega)$ such that $\Delta u = \lambda u$.

Proposition 2.4 *The operator*

$$\begin{aligned} T : H_0^1(\Omega) &\to H_0^1(\Omega) \\ v &\mapsto u \ \text{ such that } \Delta u = -v \end{aligned}$$

is an injective, compact, positive selfadjoint operator on $H_0^1(\Omega)$.

Proof. Let $J : u \mapsto u$ be the canonical injection from $H_0^1(\Omega)$ into $L^2(\Omega)$. Then $T = -\Delta^{-1} \circ J$, so, by Proposition 1.2 on page 215, T is compact, because Δ^{-1} is continuous (Theorem 2.2) and J is compact (Rellich's Theorem, page 355).

On the other hand, if $u \in H_0^1(\Omega)$ and $Tu = w$, we have $\Delta w = -u$. Therefore, by Proposition 2.1,

$$(Tu \mid v)_{H_0^1} = (u \mid v)_{L^2} \quad \text{for all } u, v \in H_0^1(\Omega),$$

which easily implies that T is selfadjoint, positive, and injective. □

It follows that we can apply to the operator T the results established in Chapter 6 concerning the spectrum of compact selfadjoint operators. Now, if $\lambda \in \mathbb{C}$ and $u \in H_0^1(\Omega)$ is nonzero, we have

$$\Delta u = \lambda u \iff T(\lambda u) = -u \iff \Big(\lambda \neq 0 \text{ and } Tu = -\frac{1}{\lambda} u\Big).$$

It follows that λ is an eigenvalue of the Dirichlet Laplacian if and only if $\lambda \neq 0$ and $-1/\lambda$ is an eigenvalue of T, and that in this case the associated eigenfunctions are the same. Since T is not of finite rank (its image clearly contains $\mathscr{D}(\Omega)$), we deduce from Theorem 2.2 on page 235 and Corollary 2.7 on page 238 the following properties:

Theorem 2.5 **1.** *The set of eigenvalues of Dirichlet Laplacian on Ω forms a sequence $0 > \lambda_0 > \lambda_1 > \cdots > \lambda_n > \cdots$ tending to $-\infty$.*

2. *The eigenspace associated with each eigenvalue λ has finite dimension d_λ.*

3. *Let $(\mu_n)_{n\in\mathbb{N}}$ be the decreasing sequence of eigenvalues of the Dirichlet Laplacian, where each eigenvalue λ is repeated d_λ times. There exists a Hilbert basis $(u_n)_{n\in\mathbb{N}}$ of $H_0^1(\Omega)$ such that*

$$\Delta u_n = \mu_n u_n \quad \textit{for all } n \in \mathbb{N}.$$

Remarks

1. By Proposition 2.1,

$$(u_n \,|\, v)_{H_0^1} = -\mu_n (u_n \,|\, v)_{L^2} \quad \text{for all } n \in \mathbb{N} \text{ and } v \in H_0^1(\Omega).$$

In particular, $\big(\|u_n\|_{L^2}\big)^2 = -1/\mu_n$ and $(u_n \,|\, u_m)_{L^2} = 0$ if $n \neq m$. At the same time, the space $\mathscr{D}(\Omega)$ is dense in $L^2(\Omega)$ (see the proof of Proposition 1.4), and a fortiori $H_0^1(\Omega)$ is dense in $L^2(\Omega)$. Since convergence in $H_0^1(\Omega)$ implies convergence in $L^2(\Omega)$, the family (u_n), which is fundamental in $H_0^1(\Omega)$, is also fundamental in the closure of $H_0^1(\Omega)$ in $L^2(\Omega)$, namely in $L^2(\Omega)$. It follows that the sequence $\big(\sqrt{-\mu_n}\, u_n\big)_{n\in\mathbb{N}}$ is a Hilbert basis for $L^2(\Omega)$.

2. For every $n \in \mathbb{N}$, we have $u_n \in \mathscr{E}(\Omega)$ (see Exercise 8). Therefore u_n satisfies the equation $\Delta u_n = \mu_n u_n$ in the ordinary sense.

Using the sequence $(u_n)_{n\in\mathbb{N}}$ and the eigenvalues $(\mu_n)_{n\in\mathbb{N}}$, we will now describe the solutions of various partial differential problems with Dirichlet conditions.

2A The Dirichlet Problem

Proposition 2.6 *Suppose $f \in L^2(\Omega)$. The solution u of the Dirichlet problem on Ω with right-hand side f is given by*

$$u = -\sum_{n=0}^{+\infty} (f \,|\, u_n)_{L^2}\, u_n,$$

the series being convergent in $H_0^1(\Omega)$.

Proof. By remark 1 above, $\big(\sqrt{-\mu_n}\, u_n\big)_{n\in\mathbb{N}}$ is a Hilbert basis of $L^2(\Omega)$, so

$$f = -\sum_{n=0}^{+\infty} \mu_n (f \,|\, u_n)_{L^2}\, u_n,$$

with convergence in $L^2(\Omega)$. In particular,

$$\left(\|f\|_{L^2}\right)^2 = \sum_{n=0}^{+\infty} -\mu_n \left|(f \mid u_n)_{L^2}\right|^2 < +\infty$$

and, since the sequence $(-\mu_n)_{n\in\mathbb{N}}$ is increasing and thus bounded below by $-\mu_0 > 0$,

$$\sum_{n=0}^{+\infty} \left|(f \mid u_n)_{L^2}\right|^2 < +\infty.$$

Since the sequence $(u_n)_{n\in\mathbb{N}}$ is a Hilbert basis for $H_0^1(\Omega)$, it follows that the series

$$v = -\sum_{n=0}^{+\infty} (f \mid u_n)_{L^2}\, u_n$$

converges in $H_0^1(\Omega)$. Since convergence in $H_0^1(\Omega)$ implies convergence in $L^2(\Omega)$, which in turn implies convergence in $\mathscr{D}'(\Omega)$, we deduce that, in $\mathscr{D}'(\Omega)$,

$$\Delta v = -\sum_{n=0}^{+\infty} \mu_n (f \mid u_n)_{L^2}\, u_n$$

(by the definition of the sequence μ_n). Likewise

$$f = -\sum_{n=0}^{+\infty} \mu_n (f \mid u_n)_{L^2}\, u_n,$$

with convergence in $\mathscr{D}'(\Omega)$. It follows that $\Delta v = f$ and so that $v = u$, the solution of the Dirichlet problem on Ω with right-hand side f. □

2B The Heat Problem

Proposition 2.7 *Suppose $f \in H_0^1(\Omega)$. There exists a unique function u from $(0, +\infty)$ to $H_0^1(\Omega)$, differentiable in $(0, +\infty)$ and satisfying the following conditions:*

- *$u'(t) = \Delta u(t)$ for all $t > 0$.*
- *$\lim_{t\to 0} u(t) = f$ in $H_0^1(\Omega)$.*

This function u is given by

$$u(t) = \sum_{n=0}^{+\infty} (f \mid u_n)_{H_0^1}\, e^{t\mu_n} u_n \quad \textit{for all } t > 0,$$

the series being convergent in $H_0^1(\Omega)$. If we write $u(t)(x) = u(t,x)$, we have

$$\left(\frac{\partial}{\partial t} - \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}\right) u = 0 \quad \textit{in } \mathscr{D}'((0, +\infty) \times \Omega).$$

The function u is called the **solution of the heat problem on Ω with initial data f and Dirichlet conditions.**

Proof. Suppose u satisfies the conditions of the statement. Then

$$\frac{d}{dt}\big(\|u(t)\|_{L^2}\big)^2 = 2\,\mathrm{Re}\big(u'(t)\,|\,u(t)\big)_{L^2} = -2\big(\|u(t)\|_{H_0^1}\big)^2,$$

by Proposition 2.1. It follows that the function $t \mapsto (\|u(t)\|_{L^2})^2$ is decreasing and, in particular, that

$$\|u(t)\|_{L^2} \le \|f\|_{L^2} \quad \text{for all } t > 0.$$

Consequently, if $f = 0$, we have $u(t) = 0$ for all $t > 0$, which proves uniqueness.

Regarding existence, it suffices to check that the given formula is good. This is easy if we take into account that $\mu_n \le \mu_0 < 0$ for every n. □

2C The Wave Problem

Proposition 2.8 *Suppose $f, g \in H_0^1(\Omega)$. There exists at most one function u from $\mathbb{R}$ to $H_0^1(\Omega)$, twice differentiable on $\mathbb{R}$ and satisfying these conditions:*

- *$u''(t) = \Delta u(t)$ for all $t \in \mathbb{R}$.*
- *$u(0) = f$ and $u'(0) = g$.*

If the sequences $\big(\mu_n (f\,|\,u_n)_{H_0^1}\big)_{n\in\mathbb{N}}$ and $\big(\sqrt{-\mu_n}\,(g\,|\,u_n)_{H_0^1}\big)_{n\in\mathbb{N}}$ lie in ℓ^2, such a function u exists and is given by

$$u(t) = \sum_{n=0}^{+\infty}\left(\cos\big(\sqrt{-\mu_n}\,t\big)(f\,|\,u_n)_{H_0^1} + \frac{1}{\sqrt{-\mu_n}}\sin\big(\sqrt{-\mu_n}\,t\big)(g\,|\,u_n)_{H_0^1}\right)u_n$$

for all $t \in \mathbb{R}$, the series being convergent in $H_0^1(\Omega)$.

Proof. Let u satisfy the conditions of the statement. By Proposition 2.1,

$$\begin{aligned}\frac{d}{dt}\big(\|u'(t)\|_{L^2}\big)^2 &= 2\,\mathrm{Re}\big(u''(t)\,|\,u'(t)\big)_{L^2}\\ &= -2\,\mathrm{Re}\big(u(t)\,|\,u'(t)\big)_{H_0^1} = -\frac{d}{dt}\big(\|u(t)\|_{H_0^1}\big)^2.\end{aligned}$$

It follows that the expression $\big(\|u'(t)\|_{L^2}\big)^2 + \big(\|u(t)\|_{H_0^1}\big)^2$ does not depend on t. In particular, if $f = g = 0$, we have $u(t) = 0$ for $t \in \mathbb{R}$, which proves uniqueness.

The proof of existence, as in the previous example, is straightforward. □

Here again, if we write $u(t)(x) = u(t,x)$, we conclude that, under the assumptions made in the existence part of Proposition 2.8, the given function u is a solution in $\mathscr{D}'(\mathbb{R} \times \mathbb{R}^d)$ of the equation

$$\left(\frac{\partial^2}{\partial t^2} - \sum_{j=1}^{d} \frac{\partial^2}{\partial x_j^2}\right) u = 0.$$

Note that $\left(\dfrac{\sqrt{-\mu_n}}{2\pi}\right)_{n\in\mathbb{N}}$ is the sequence of fundamental frequencies of the wave u.

Exercises

Unless otherwise stated, Ω is a bounded open subset of $\mathbb{R}^d$.

1. *A generalized Dirichlet problem*

a. Suppose $f, g_1, \ldots, g_d$ are elements of $L^2(\Omega)$. Show that there exists a unique element u in $H_0^1(\Omega)$ such that

$$\Delta u = f + \sum_{j=1}^{d} D_j g_j.$$

Show that, in addition,

$$\big\| |\nabla u| \big\|_{L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)} + \sum_{j=1}^{d} \|g_j\|_{L^2(\Omega)},$$

where C is the constant that appears in the Poincaré inequality for the open set Ω.

b. Suppose $f \in L^2(\Omega)$ and $g \in H^1(\Omega)$. Show that there exists a unique element u of $H^1(\Omega)$ such that

$$\Delta u = f \quad\text{and}\quad u - g \in H_0^1(\Omega).$$

Show that, in addition,

$$\big\| |\nabla u| \big\|_{L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)} + 2 \big\| |\nabla g| \big\|_{L^2(\Omega)}.$$

2. *The Sturm–Liouville problem.* All functions considered here will be real-valued. Let $a, b \in \mathbb{R}$ be such that $a < b$ and let p and q be elements of $L^\infty((a,b))$. Suppose that $q \geq 0$ and that there exists a real number $\alpha > 0$ such that $p \geq \alpha$.

a. Show that, for every $f \in L^2((a,b))$, the equation

$$-(pu')' + qu = f \tag{$*$}$$

has a unique solution u in $H_0^1((a,b))$, and that this solution minimizes over $H_0^1((a,b))$ the functional

$$J_f(v) = \frac{1}{2}\int_{(a,b)} (pv'^2 + qv^2)(x)\,dx - \int_{(a,b)} (fv)(x)\,dx.$$

Hint. Apply the Lax–Milgram Theorem (Exercise 1 on page 116) to the space $H_0^1((a,b))$ and to the bilinear form

$$\mathfrak{a}(u,v) = (pu' \mid v')_{L^2} + (qu \mid v)_{L^2}. \qquad (**)$$

b. Show that the linear operator T from $H_0^1((a,b))$ to $H_0^1((a,b))$ that maps each $f \in H_0^1((a,b))$ to the corresponding solution of $(*)$ is a compact, injective, positive selfadjoint operator on the Hilbert space $(H_0^1((a,b)),\ \mathfrak{a})$, where $\mathfrak{a}$ is defined by $(**)$.

c. Show that there exists a Hilbert basis $(e_n)_{n\in\mathbb{N}}$ of $(H_0^1((a,b)),\ \mathfrak{a})$ and an increasing sequence $(\lambda_n)_{n\in\mathbb{N}}$ of positive real numbers with limit $+\infty$ such that

$$-(pe_n')' + qe_n = \lambda_n e_n \quad \text{for all } n \in \mathbb{N}.$$

Show that the family $\bigl(\sqrt{\lambda_n}\,e_n\bigr)_{n\in\mathbb{N}}$ is a Hilbert basis for $L^2((a,b))$.

d. Show that, if $p \in C^1([a,b])$ and $q \in C([a,b])$, we have $e_n \in C^2([a,b])$ for every $n \in \mathbb{N}$. Compare Exercise 13 on page 224.

e. Let $\mathscr{V}_p$ be the set of p-dimensional subspaces of $H_0^1((a,b))$. Show that

$$\lambda_n = \min_{W\in\mathscr{V}_{n+1}} \max_{f\in W\setminus\{0\}} \frac{\displaystyle\int_a^b (pf'^2 + qf^2)\,dx}{\displaystyle\int_a^b f^2\,dx}$$

(see Exercise 11 on page 247). Deduce that, if α, β, γ, and δ are real numbers satisfying $0 < \alpha \le p \le \beta$ and $0 \le \gamma \le q \le \delta$, we have, for every $n \in \mathbb{N}$,

$$\gamma + \alpha\left(\frac{\pi(n+1)}{b-a}\right)^2 \le \lambda_n \le \delta + \beta\left(\frac{\pi(n+1)}{b-a}\right)^2.$$

Show that, in particular,

$$\lambda_0 = \min_{f\in H_0^1((a,b))\setminus\{0\}} \frac{\displaystyle\int_a^b (pf'^2 + qf^2)\,dx}{\displaystyle\int_a^b f^2\,dx}$$

and that this generalizes Wirtinger's inequality (Exercise 16d on page 137).

3. *A more general elliptic problem of order* 2. Take elements $a_{i,j}$ (for $1 \le i, j \le d$) and c in $L^\infty_{\mathbb{R}}(\Omega)$, with $c \le 0$. Let $\mathscr{L}$ be the differential operator defined by

$$\mathscr{L}u = \sum_{1\le i,j\le d} \frac{\partial}{\partial x_i}\left(a_{i,j}\frac{\partial}{\partial x_j}u\right) + cu.$$

Suppose that there exists $\alpha > 0$ such that, for almost every $x \in \Omega$,

$$\sum_{1\le i,j\le d} a_{i,j}(x)\xi_i\xi_j \ge \alpha|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d.$$

(In this case the operator $\mathscr{L}$ is said to be **strongly elliptic** or **uniformly elliptic**.) We now restrict ourselves to real-valued functions.

a. Show that $\mathscr{L}$ is well defined as a linear operator from $H^1_0(\Omega)$ to $\mathscr{D}'(\Omega)$.

b. Show that, for every $f \in L^2(\Omega)$, there exists a unique $u \in H^1_0(\Omega)$ such that $\mathscr{L}u = f$.

Hint. Apply the Lax–Milgram Theorem (Exercise 1 on page 116) to the space $H^1_0(\Omega)$ and to the bilinear form

$$\mathfrak{a}(v,w) = \sum_{1\le i,j\le d} (a_{i,j}D_jv \mid D_iw)_{L^2} - (cv \mid w)_{L^2}. \qquad (*)$$

c. Show that, if the matrix $(a_{i,j}(x))_{i,j}$ is symmetric, the solution $u \in H^1_0(\Omega)$ of the equation $\mathscr{L}u = f$ is characterized by a certain minimization property.

d. Suppose that the matrix $(a_{i,j}(x))_{i,j}$ is symmetric. Show that the operator T on $H^1_0(\Omega)$ that maps $f \in H^1_0(\Omega)$ to the element $Tf = u \in H^1_0(\Omega)$ such that $-\mathscr{L}u = f$ is a compact, injective, positive selfadjoint operator on the Hilbert space $(H^1_0(\Omega), \mathfrak{a})$, where $\mathfrak{a}$ is defined by $(*)$. Derive the existence of a Hilbert basis $(u_n)_{n\in\mathbb{N}}$ of $(H^1_0(\Omega), \mathfrak{a})$ and of a decreasing sequence $(\mu_n)_{n\in\mathbb{N}}$ of negative real numbers with limit $-\infty$ such that $\mathscr{L}u_n = \mu_n u_n$ for every $n \in \mathbb{N}$. Deduce, in particular, that Propositions 2.6, 2.7, and 2.8 extend immediately to the case where the Laplacian is replaced by the operator $\mathscr{L}$ and $(\cdot\mid\cdot)_{H^1_0}$ is replaced by $\mathfrak{a}$.

e. Let $\mathscr{V}_p$ be the set of p-dimensional subspaces of $H^1_0(\Omega)$. Show that, for every $n \in \mathbb{N}$,

$$-\mu_n = \min_{W\in\mathscr{V}_{n+1}} \max_{f\in W\setminus\{0\}} \frac{\mathfrak{a}(f,f)}{\|f\|^2_{L^2}}$$

(see Exercise 11 on page 247). Deduce that

$$(-\mu_n) \ge \alpha(-\mu_n(\Omega)) \quad \text{for all } n \in \mathbb{N},$$

where $(\mu_n(\Omega))_{n\in\mathbb{N}}$ is the sequence of eigenvalues of the Dirichlet Laplacian on Ω.

4. *A mixed problem.* We maintain the hypotheses and notation of Exercise 3. Take elements $b_i \in L^\infty_{\mathbb{R}}(\Omega)$ and let $\mathfrak{b}$ be the bilinear form on $H^1(\Omega)$ defined by

$$\mathfrak{b}(u,v) = \mathfrak{a}(u,v) + \int_\Omega \sum_{i=1}^d b_i(x)\frac{\partial u}{\partial x_i}(x)v(x)\,dx.$$

If $f \in L^2(\Omega)$, we consider the following problem $(\mathrm{P})_f$: Determine an element $u \in H^1_0(\Omega)$ such that

$$\mathfrak{b}(u,v) = \int_\Omega f(x)v(x)\,dx \quad \text{for all } v \in H^1_0(\Omega).$$

a. Interpret $(\mathrm{P})_f$ as a partial differential problem in $H^1_0(\Omega)$.

b. **i.** Show that, for any $\varepsilon > 0$ and any real numbers α and β,

$$\alpha\beta \le \frac{1}{2\varepsilon}\alpha^2 + \frac{\varepsilon}{2}\beta^2.$$

ii. Derive the existence of a real number $B > 0$ such that, for every $\varepsilon > 0$ and $u \in H^1_0(\Omega)$,

$$\int_\Omega \sum_{i=1}^d b_i(x)\,\frac{\partial u}{\partial x_i}(x)u(x)\,dx \ge -\frac{B}{2\varepsilon}\big\|\,|\nabla u|\,\big\|^2_{L^2(\Omega)} - \frac{B\varepsilon}{2}\|u\|^2_{L^2(\Omega)}.$$

c. Deduce from this that, if the diameter of Ω is small enough, the form $\mathfrak{b}$ is coercive in $H^1_0(\Omega)$, so the problem $(\mathrm{P})_f$ has a unique solution for every $f \in L^2(\Omega)$.

(We make no assumptions on the diameter of Ω in the remainder of the exercise.)

d. Show that there exists a constant $\lambda_0 > 0$ such that the bilinear form defined by

$$(u,v) \longmapsto \mathfrak{b}(u,v) + \lambda_0 \int_\Omega u(x)v(x)\,dx$$

is coercive on $H^1_0(\Omega)$.

e. Take $f \in L^2(\Omega)$. Explain why there is a unique $u \in H^1_0(\Omega)$ such that for every $v \in H^1_0(\Omega)$,

$$\mathfrak{b}(u,v) + \lambda_0 \int_\Omega u(x)v(x)\,dx = \int_\Omega f(x)v(x)\,dx.$$

Then prove that the operator T from $L^2(\Omega)$ to $L^2(\Omega)$ defined by $Tf = u$ is compact.

f. Let E be the vector space of solutions of $(\mathrm{P})_0$.

i. Show that E is finite-dimensional.

ii. Show that $(\mathrm{P})_f$ has a unique solution for every element $f \in L^2(\Omega)$ if and only if $E = \{0\}$.

5. *Neumann boundary conditions in dimension* 1. We restrict ourselves to real-valued functions. Suppose that $\Omega = (a, b)$.

a. Show that, for every $f \in L^2(\Omega)$, there exists a unique $u \in H^1(\Omega)$ such that

$$\int_a^b (u'v')(x)\,dx + \int_a^b (uv)(x)\,dx = \int_a^b f(x)v(x)\,dx \quad \text{for all } v \in H^1(\Omega).$$

Hint. Use the Riesz Theorem in the Hilbert space $H^1(\Omega)$.

b. Show that u is the element of $H^1(\Omega)$ that minimizes the functional

$$J_f(v) = \frac{1}{2}\int_a^b (v'^2 + v^2)(x)\,dx - \int_a^b (fv)(x)\,dx.$$

c. Check that, in this case, $u' \in H^1(\Omega)$ and

$$-u'' + u = f.$$

d. Deduce that, for every $v \in H^1(\Omega)$,

$$\int_a^b u'(x)v'(x)\,dx = u'(b)v(b) - u'(a)v(a) - \int_a^b u''(x)v(x)\,dx$$

(you might use Exercise 14 on page 360); then deduce that

$$u'(a) = u'(b) = 0.$$

e. Show that, for every function $f \in L^2(\Omega)$, there exists a unique function $u \in C^1([a, b])$ such that

$$-u'' + u = f, \quad u'(a) = u'(b) = 0,$$

and that if, in addition, $f \in C((a, b))$, then $u \in C^2((a, b))$.

6. *A variational problem with obstacles.* We restrict ourselves to real-valued functions. Fix $f \in L^2(\Omega)$ and let χ be a function on Ω such that the set

$$C = \{u \in H_0^1(\Omega) : u \geq \chi \text{ almost everywhere}\}$$

is nonempty.

a. Show that there exists a unique $u \in C$ minimizing over C the functional J defined by

$$J(v) = \tfrac{1}{2}\|v\|_{H_0^1}^2 + (f \,|\, v)_{L^2} \quad \text{for all } v \in C.$$

Show that u is also characterized by the following conditions:

$$\begin{aligned} &u \in C, \\ &\int_\Omega \nabla u \cdot \nabla(v - u)\,dx \geq -\int_\Omega f\,(v - u)\,dx \quad \text{for all } v \in C. \end{aligned} \tag{$*$}$$

Hint. Use the Lions–Stampacchia Theorem, Exercise 2 on page 117.

b. Now suppose that $\chi \in C(\Omega)$, $u \in C(\Omega)$, and u satisfies $(*)$. Show that u satisfies

$$\begin{aligned} &u \in C, \\ &-\Delta u + f \text{ is a positive distribution on } \Omega, \\ &-\Delta u + f = 0 \text{ on } \omega = \{u > \chi\}. \end{aligned} \qquad (**)$$

Hint. Prove first the following facts:

i. If $\varphi \in \mathscr{D}(\Omega)$ and $\varphi \geq 0$, then $u + \varphi \in C$.

ii. If $\varphi \in \mathscr{D}(\omega)$, there exists $\eta > 0$ such that $u + t\varphi \in C$ for every $t \in (-\eta, \eta)$.

c. Suppose that $u, \chi \in C(\Omega)$ and that, $D_j^2 u \in L^2(\Omega)$ for every $j \in \{1, \dots, d\}$. Show that, if u satisfies $(**)$, then u satisfies $(*)$.

d. Suppose in this part that u and χ belong to $C_0(\Omega)$ and that u satisfies $(**)$. Let $v \in C$.

i. Show that

$$\int_\Omega \nabla u(x) \cdot \nabla (v-u)^+(x)\, dx \geq -\int_\Omega f(x)(v-u)^+(x)\, dx$$

(see Exercise 15 on page 360).

ii. Show that

$$\int_\Omega \nabla u \cdot \nabla (u-v)^+ dx = \int_\omega \nabla u \cdot \nabla (u-v)^+ dx,$$
$$\int_\Omega f(u-v)^+ dx = \int_\omega f(u-v)^+ dx.$$

Hint. For the first equality, use Exercise 11 on page 358.

iii. For $n \geq 1$, let $w_n = (u - v - (1/n))^+$. Show that w_n belongs to $H_0^1(\Omega)$ and that the support of w_n is contained in the set $\{u - \chi \geq 1/n\}$. Deduce that $(w_n)_{|\omega} \in H_c^1(\omega)$, then that $(u-v)^+_{|\omega} \in H_0^1(\omega)$ (see Exercise 4 on page 357).

iv. Deduce that

$$\int_\omega \nabla u \cdot \nabla (u-v)^+ dx = -\int_\omega f(u-v)^+ dx.$$

v. Conclude from the preceding results that u satisfies $(*)$.

e. Suppose again that u and χ belong to $C_0(\Omega)$ and that u satisfies $(**)$. Let $v \in C$ be such that $-\Delta v + f$ is a positive distribution on Ω. Show that $v \geq u$. (Therefore, under the assumptions stated, u is also characterized as the smallest element $v \in C$ such that $-\Delta v + f \geq 0$.)

Hint. Show that $\Delta(u - v)$ is a positive distribution on ω and that $u - v \leq 0$ on $\partial\omega$ in the sense of Exercise 19 on page 362. Then deduce from this exercise that $v \geq u$ on ω. Conclude the proof.

7. Let Ω_1 and Ω_2 be disjoint open sets in $\mathbb{R}^d$ and put $\Omega = \Omega_1 \cup \Omega_2$. Determine the eigenvalues $(\lambda_n(\Omega))$ and the eigenfunctions of the Dirichlet Laplacian on Ω as a function of the eigenvalues $(\lambda_n(\Omega_1))$, $(\lambda_n(\Omega_2))$ and of the eigenfunctions of the Dirichlet Laplacian on Ω_1 and Ω_2.

8. Let $(u_n)_{n\in\mathbb{N}}$ be the Hilbert basis of $H_0^1(\Omega)$ defined in Theorem 2.5, consisting of eigenfunctions of the Dirichlet Laplacian. Let $(\mu_n)_{n\in\mathbb{N}}$ be the corresponding sequence of eigenvalues. Denote by E_k the fundamental solution of the operator Δ^k defined in Exercise 4 on page 313. We recall that, if $2k \geq d+1$, E_k belongs to $\mathscr{E}^{2k-d-1}(\mathbb{R}^d)$. Moreover, we denote by $\tilde{u}_n$ the function u_n extended to $\mathbb{R}^d$ with the value 0 outside Ω.

a. Show that, if T_k^n denotes the restriction of $(\mu_n)^k E_k * \tilde{u}_n$ to Ω, then

$$\Delta^k(u_n - T_k^n) = 0$$

in $\mathscr{D}'(\Omega)$.

b. Using the fact that Δ^k is hypoelliptic, prove that $u_n \in \mathscr{E}^{2k-d-1}(\Omega)$ if $2k \geq d+1$. Deduce that $u_n \in \mathscr{E}(\Omega)$.

9. Let λ_0 be the first eigenvalue of the Dirichlet Laplacian (that is, the smallest eigenvalue in absolute value). Show that $\sqrt{-1/\lambda_0}$ is the best possible constant C in the Poincaré inequality.

10. We retain the notation of Theorem 2.5.

a. Show that

$$\|\varphi\|^2_{L^2(\Omega)} \leq -\frac{1}{\lambda_0}\|\varphi\|^2_{H_0^1(\Omega)} \quad \text{for all } \varphi \in H_0^1(\Omega)$$

and that equality takes place if and only if

$$\Delta\varphi = \lambda_0\varphi.$$

Hint. Use the Bessel equality in the space $H_0^1(\Omega)$ with the basis $(u_n)_{n\in\mathbb{N}}$ and in $L^2(\Omega)$ with the basis $(\sqrt{-\mu_n}\, u_n)_{n\in\mathbb{N}}$.

b. Let E be the λ_0-eigenspace of the Dirichlet Laplacian on Ω:

$$E = \{\varphi \in H_0^1(\Omega) : \Delta\varphi = \lambda_0\varphi\}.$$

Let $E_{\mathbb{R}}$ be the set of real-valued elements of E. Show that $|\varphi| \in E_{\mathbb{R}}$ for every $\varphi \in E_{\mathbb{R}}$.

Hint. Use Exercise 11e on page 359.

c. Recall from Exercise 8 that the elements of E belong to $\mathscr{E}(\Omega)$. Take $\varphi \in E_{\mathbb{R}}$ and let $x \in \Omega$ be such that $\varphi(x) = 0$. Show that, if $\rho < d(x, \mathbb{R}^d \backslash \Omega)$, then

$$\int \bigl|\varphi(x-y)\bigr|\, d\sigma_\rho(y) = 0,$$

where σ_ρ is the surface measure on the sphere of center 0 and radius ρ in $\mathbb{R}^d$. Deduce that $\varphi = 0$ on $B\bigl(x, d(x, \mathbb{R}^d \backslash \Omega)\bigr)$.

Hint. Show that $\Delta(|\varphi|) \leq 0$ and use Exercise 3c on page 346.

d. Suppose that Ω is connected. Show that a nonzero element of E cannot vanish anywhere on Ω. Deduce that E has dimension 1 and that E is generated by a strictly positive function on Ω.
Show that no eigenfunction of the Dirichlet Laplacian associated with an eigenvalue other than λ_0 can have positive values everywhere.

e. What happens to the results of the previous question when Ω is not connected? (You might use Exercise 7 for inspiration.)

11. *An asymptotic estimation of the eigenvalues of the Dirichlet Laplacian.* If Ω is a bounded open set in $\mathbb{R}^d$, denote by $(\mu_n(\Omega))_{n\in\mathbb{N}}$ the decreasing sequence of eigenvalues of the Dirichlet Laplacian on Ω, each eigenvalue appearing as many times as the dimension of the corresponding eigenspace. Denote by $\Lambda(\Omega) = \{\mu_n(\Omega)\}_{n\in\mathbb{N}}$ the set of eigenvalues of the Dirichlet Laplacian on Ω.

a. Suppose first that $\Omega = (0,l)^d$, where $0 < l < +\infty$.

i. Show that, for every $p \in (\mathbb{N}^*)^d$, the real number

$$\lambda_p = -\frac{\pi^2}{l^2}(p_1^2 + \cdots + p_d^2)$$

lies in $\Lambda(\Omega)$.
Hint. Show that the function

$$u_p(x) = \left(\frac{2}{l}\right)^{d/2} \prod_{j=1}^{d} \sin\left(\frac{\pi p_j x_j}{l}\right)$$

is an eigenfunction corresponding to λ_p.

ii. Show that $\Lambda(\Omega) = \{\lambda_p\}_{p\in(\mathbb{N}^*)^d}$.
Hint. Show that the family $\{u_p\}_{p\in(\mathbb{N}^*)^d}$ is a fundamental orthonormal family in $L^2(\Omega)$.

iii. Let $r > 0$. Show that the number of points in the ball $\bar{B}(0,r)$ in $\mathbb{R}^d$ whose coordinates belong to $\mathbb{N}^*$ is at least $\omega_d((r-\sqrt{d})^+)^d 2^{-d}$ and at most $\omega_d r^d 2^{-d}$, where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

iv. Deduce that, for every $n \in \mathbb{N}$,

$$\frac{\omega_d}{2^d}\left(\frac{l}{\pi}\sqrt{|\mu_n|} - \sqrt{d}\right)^d \le n+1 \le \omega_d\left(\frac{l}{2\pi}\sqrt{|\mu_n|}\right)^d.$$

v. Deduce that there exist constants α, $\beta > 0$ such that, for every $n \in \mathbb{N}$,

$$-\alpha n^{2/d} \le \mu_n(\Omega) \le -\beta n^{2/d}.$$

(More precisely, $\mu_n(\Omega) \sim -4\pi^2 l^{-2}\omega_d^{-2/d} n^{2/d}$.)

b. Let Ω and Ω' be bounded open sets in $\mathbb{R}^d$. Show that, if $\Omega \subset \Omega'$, we have

$$|\mu_n(\Omega')| \le |\mu_n(\Omega)| \quad \text{for all } n \in \mathbb{N}.$$

Hint. Use Exercise 11 on page 247 and the fact that, using the same proof of Proposition 1.8, we can isometrically inject $H_0^1(\Omega)$ into $H_0^1(\Omega')$.

c. Show that, for every bounded open set Ω in $\mathbb{R}^d$, there exist two constants α_Ω, $\beta_\Omega > 0$ such that

$$-\alpha_\Omega n^{2/d} \le \mu_n(\Omega) \le -\beta_\Omega n^{2/d} \quad \text{for all } n \in \mathbb{N}.$$

12. Let $u : (0, +\infty) \mapsto H_0^1(\Omega)$ be the solution of the heat problem

$$u'(t) = \Delta u(t) \quad \text{for all } t > 0$$

with initial condition $\lim_{t\to 0} u(t) = f$ in $H_0^1(\Omega)$ (where $f \in H_0^1(\Omega)$). Show that, for every $t > 0$,

$$\|u(t)\|_{H_0^1} \le \|f\|_{H_0^1} e^{-t|\lambda_0|},$$

where λ_0 is the first eigenvalue of the Dirichlet Laplacian (the smallest eigenvalue in absolute value).

13. *Heat semigroup.*

a. Suppose $f \in H_0^1(\Omega)$ and $t > 0$. Denote by $P_t f$ the value at t of the solution of the heat problem on Ω with initial data f and Dirichlet conditions. Show that

$$P_t f = \sum_{n=0}^{+\infty} e^{t\mu_n} (f \,|\, v_n)_{L^2} v_n$$

with $v_n = \sqrt{-\mu_n}\, u_n$ (in the notation of Theorem 2.5), the series being convergent in $L^2(\Omega)$.

b. Deduce that, for every $t > 0$, P_t extends to a continuous linear operator from $L^2(\Omega)$ to $L^2(\Omega)$ of norm at most $e^{t\mu_0}$. We denote this operator again by P_t.

c. Show the following facts:

- $P_{t+s} = P_t P_s$ for all $t, s > 0$.
- $\lim_{t\to 0^+} P_t f = f$ in $L^2(\Omega)$ for all $f \in L^2(\Omega)$.

d. Show that, if $f \in L^2(\Omega)$, the limit $\lim_{t\to 0^+} (P_t f - f)/t$ exists in $L^2(\Omega)$ if and only if the series $\sum_{n=0}^{+\infty} \mu_n^2 |(f|v_n)_{L^2}|^2$ converges. Show that, if this is the case,

$$\lim_{t\to 0^+} \frac{P_t f - f}{t} = \sum_{n=0}^{+\infty} \mu_n (f \,|\, v_n)_{L^2} v_n = \Delta f.$$

Answers to the Exercises

Prologue

- *Page 4, Ex. 1.* ***a.*** No. ***b.*** Yes. ***c.*** Yes.
- *Page 5, Ex. 3a.* D is the set of dyadic numbers:

$$D = \{k2^{-n} : n \in \mathbb{N}, k \in \mathbb{N}, 0 \le k < 2^n\}.$$

- *Page 7, Ex. 9a-i.* If $(n, m) \in I$, then $(n, m) + 1 = (n, m+1)$. The limit points of I are the elements $(n, 0)$, for $n \in \mathbb{N}^*$.
- *Page 11, Ex. 7.* ***b.*** c_0. ***c.*** c_0.

Chapter 1

- *Page 30, Ex. 1.* With the notation of the proof of Proposition 1.1,

$$P_n(f) = \sum_{j=1}^{N_n} f(x_j^n)\varphi_{n,j}.$$

- *Page 37, Ex. 2c-iii.* $D_n * f = S_n$, $K_n * f = (\sum_{j=0}^{n-1} S_j)/n$.
- *Page 38, Ex. 3b-ii.* $B_n(1) = 1$, $B_n(X) = X$, and $B_n(X^2) = X^2 + X(1-X)/n$.
- *Page 45, Ex. 1.* $[0, 1)$.
- *Page 46, Ex. 5b-ii.* If $m > n$, then $\bar{B}(C^m([0,1]))$ is not closed in $C^n([0,1])$.
- *Page 47, Ex. 7b-iii.* Uniqueness does not hold in general, as can be seen from the example $f(x,t) = \sqrt{|x|}$, for which $\varphi(t) = 0$ and $\varphi(t) = t^2/4$ are solutions.

Chapter 2

- *Page 54, Ex. 1.* ***b*** and ***c.*** $(0,1)$. ***d.*** Any infinite-dimensional Banach space: for example, $C([0,1])$.
- *Page 54, Ex. 5b.* The space $\mathbb{R}[X]$ with $\|P\| = \max_{x\in[0,1]} |P(x)|$. Another example: $\mathbb{Q}$.
- *Page 79, Ex. 8c.* No. $\operatorname{Supp}\nu = \{(x,x) : x \in \mathbb{R}\}$.
- *Page 91, Ex. 7d.* $X = \mathbb{R}$, $\mu_n = \delta_n$, the Dirac measure at n.

Chapter 3

- *Page 109, Ex. 2b.* $F^\perp$ consists of the constant functions on $[0,1]$. For $f(x) = e^x$, we have $d(f,F) = e - 1$.
- *Page 109, Ex. 3.* The orthogonal projection onto E_n is $f \mapsto 1_{A_n} f$.
- *Page 120, Ex. 10b.* Set $E = \ell^2$ and let $(x_n)_{n\in\mathbb{N}}$ be the sequence defined in Section 3B on page 115. It converges to 0 weakly and $(x_n \,|\, x_n) = 1$ for all n.
- *Page 130, Ex. 3.* ***a.*** A Hilbert basis for E_A is given by $\{e_p\}_{p\notin A}$, with $e_p(x) = e^{ipx}$.
  b. $E_A^\perp = E_{\mathbb{Z}\setminus A}$.
  c. The orthogonal projection onto E_A is given by $f \mapsto \sum_{p\notin A} (f \,|\, e_p) e_p$.
- *Page 131, Ex. 4b.* $f \mapsto \sum_{j=0}^{n} (j + \frac{1}{2})(f \,|\, P_j) P_j$.
- *Page 132, Ex. 7b-i.* $(X^k, L_n) = 0$ if $k < n$; $(X^n, L_n) = (-1)^n n!$.

Chapter 4

- *Page 148, Ex. 2c.* If $p = \infty$, $f_\infty(x) = 1$. If $p \in [1,\infty)$,

$$f_p(x) = \Bigl(|x|(\log^2 |x| + 1)\Bigr)^{-1/p}.$$

- *Page 149, Ex. 3d.* If $K(x,y) = 1/(x+y)$, then $\|T\| = \pi/\sin(\pi/p')$. If $K(x,y) = 1/\max(x,y)$, then $\|T\| = p + p'$.
- *Page 152, Ex. 7d.* No — for example, if X has a non $\mathscr{F}$-measurable subset and the singletons of X are $\mathscr{F}$-measurable.
- *Page 155, Ex. 14c.* Example: $\mathscr{A}_n = \{[k2^{-n}, (k+1)2^{-n}] : k \in \{0, \dots, 2^n - 1\}\}$.
- *Page 159, Ex. 19b.* If $(K_n)_{n\in\mathbb{N}}$ is a sequence of compact sets exhausting X, one can give L^p_{loc} the metric $d(f,g) = \sum_{n=0}^{+\infty} 2^{-n} \inf(\|1_{K_n}(f-g)\|_p, 1)$.
- *Page 164, Ex. 4c.* $(\ell^\infty(I))'$ can be identified with the vector space spanned by the family $\{L_\mu\}_{\mu\in\Lambda(I)}$.

Chapter 5

- *Page 195, Ex. 4b.* $\operatorname{ev}(T) = \{\lambda_i : i \in \mathbb{N}\}$ and $\sigma(T) = \overline{\operatorname{ev}(T)}$.
- *Page 196, Ex. 9.* If T is considered as an operator on $L^p(m)$, then $\operatorname{ev}(T) = \{\lambda \in \mathbb{K} : m(\{\varphi = \lambda\}) > 0\}$ and $\sigma(T) = \varphi(\operatorname{Supp} m)$.
- *Page 197, Ex. 10.* $r(T) = \pi/2$.
- *Page 197, Ex. 11.* $\operatorname{ev}(S) = \sigma(S) = \{0, 1/5\}$.
- *Page 197, Ex. 12.* $\operatorname{ev}(T) = \sigma(T) = \{0, (4+\sqrt{31})/60, (4-\sqrt{31})/60\}$.
- *Page 198, Ex. 15b.* $r(T) = 0$, $\sigma(T) = \{0\}$.

- *Page 200, Ex. **20f**.* Let $E = \ell^p$, $p \in [1, \infty]$ and T defined by $(Tu)(0) = 0$ and, for $n \in \mathbb{N}^*$, $(Tu)(n) = u(n-1)$. Then $\operatorname{ev}(T) = \overline{\operatorname{ev}(T)} = \varnothing$,
$$\operatorname{aev}(T) = \{\lambda \in \mathbb{K} : |\lambda| = 1\}, \qquad \sigma(T) = \{\lambda \in \mathbb{K} : |\lambda| \le 1\}.$$
- *Page 208, Ex. **2**.* In the example,
$$\sigma(T) = \{\lambda \in \mathbb{K} : |\lambda| \le 1\} \cup \{\lambda \in \mathbb{K} : |\lambda - 1| \le 1\}.$$
- *Page 211, Ex. **11e**.* $Pf = |a|f$, $Uf = 1_{\{a \ne 0\}}(a/|a|)f$.
- *Page 212, Ex. **14**.* $f(T)u = (f \circ \varphi)u$.

Chapter 6

- *Page 221, Ex. **2**. **c.*** No. ***d.*** $\operatorname{ev}(TS) = \{0\}$ if there exists $n \ge 1$ such that $\lambda_n = 0$; otherwise $\operatorname{ev}(TS) = \varnothing$. $\sigma(TS) = \{0\}$.
- *Page 223, Ex. **11**.* $\sigma(T) = \{0\} \cup \{((\pi/2) + k\pi)^{-2} : k \in \mathbb{N}\}$.
- *Page 224, Ex. **12d**.* Yes.
- *Page 227, Ex. **14c**.* One can take $\mu = \sum_{n \in \mathbb{N}} 2^{-n} \delta_{x_n}$, where $(x_n)_{n \in \mathbb{N}}$ is a dense sequence in X.
- *Page 243, Ex. **4b**.* $r(TT^*) = 4/\pi^2$, $\|T\| = 2/\pi$.
- *Page 244, Ex. **6b**.* The family $(f_n)_{n \in \mathbb{Z}}$, with $f_n(x) = \sqrt{2}\cos((\pi/2 + 2n\pi)x)$.
- *Page 249, Ex. **12d**.*
  - ***i.*** $\lambda_n = -n^2\pi^2$, $\varphi_0 = 1$, $\varphi_n = \sqrt{2}\cos(n\pi x)$ if $n \ge 1$;
  - ***ii.*** $\lambda_n = -((\pi/2) + n\pi)^2$, $\varphi_n = \sqrt{2}\cos(((\pi/2) + n\pi)x)$;
  - ***iii.*** $\lambda_n = -(n+1)^2\pi^2$, $\varphi_n = \sqrt{2}\sin((n+1)\pi x)$.

Chapter 7

- *Page 264, Ex. **2**.* If $h \in \mathscr{E}^n(\mathbb{R})$, then $f \in \mathscr{E}^{n-1}(\mathbb{R}^2)$.
- *Page 264, Ex. **3**.* The value of f at 0 is $h^{(n+1)}(0)/(n+1)!$.
- *Page 266, Ex. **9**.* $\lim_{n \to +\infty} \varphi_{1/n} = \sum_{j=1}^d u_j D_j \varphi$.
- *Page 275, Ex. **2**.* The order of T is $\sup_{n \in \mathbb{N}} |p_n|$.
- *Page 275, Ex. **3**.* Distribution of order 2.
- *Page 277, Ex. **8**.* $(\int \chi(x)\,dx)\delta$.
- *Page 277, Ex. **9**. **a.*** 0. ***b.*** $(2\int_0^{+\infty}(\sin x/x)\,dx)\delta = \pi\delta$. ***c*** and ***d.*** δ.
- *Page 277, Ex. **10**.* In $\mathscr{D}'(\mathbb{R}^*)$ there is convergence for any sequence $(a_n)_{n \in \mathbb{N}}$. In $\mathscr{D}'(\mathbb{R})$ there is convergence if and only if the series $\sum_{k=1}^{+\infty} a_k/k$ converges.
- *Page 277, Ex. **12**.* δ.
- *Page 278, Ex. **15b**.* No.
- *Page 279, Ex. **17**.* If $\Omega = \mathbb{R}$ and $T_n = n(\delta_{1/n} - \delta)$, then, for every $n \in \mathbb{N}^*$, T_n is of order 0 and the limit δ' is of order 1.
- *Page 284, Ex. **2**.* No. Example: $\varphi(x) = x$ in a neighborhood of 0 and $T = \delta'$.
- *Page 285, Ex. **4c**.* If $\operatorname{Supp} T = \{x_1, \dots, x_r\}$, then
$$\langle T, \varphi \rangle = \sum_{j=1}^{r} \sum_{|p| \le m} c_{p,j} D^p \varphi(x_j),$$
where m is an upper bound for the order of T.

Chapter 8

- *Page 291, Ex. 7b.* $T = T_0 + \sum_{k \in \mathbf{Z}} C_k \delta_{k\pi}$.
- *Page 303, Ex. 6.* $\operatorname{Supp} T = \{(x,x),\ x \in \mathbb{R}\}$; T has order 0 and $(\partial T)/(\partial x_1) + (\partial T)/(\partial x_2) = 0$.
- *Page 303, Ex. 7c.*

$$\frac{d^m}{dx^m}\operatorname{fp}(Y(x)/x) = (-1)^m m! \operatorname{fp}\left(\frac{Y(x)}{x^{m+1}}\right) - \left(\sum_{j=1}^{m} \frac{1}{j}\right) \delta^{(m)}.$$

- *Page 303, Ex. 8.* $\delta_1 + \delta_{-1} - 2\delta$.
- *Page 304, Ex. 9a.* $(\partial f/\partial x) = -1_{\{x^2+y^2<1\}} x(x^2+y^2)^{-1/2}$.
  $(\partial f/\partial y) = -1_{\{x^2+y^2<1\}} y(x^2+y^2)^{-1/2}$.
- *Page 304, Ex. 11.* ***a.*** $\Delta g_\varepsilon = \varepsilon^2((\partial f/\partial x)^2 + (\partial f/\partial y)^2)(\varepsilon^2 + f^2)^{-3/2}$.
  d. $\int \varphi \, d\mu = 2\int \varphi(0,y)|y|\,dy + 2\int \varphi(x,0)|x|\,dx$.
- *Page 306, Ex. 15b-iv.* Let H be a hyperplane and $\vec{n}$ a normal unit vector. We set $H^+ = \{x + t\vec{n} : x \in H \text{ and } t > 0\}$ and $H^- = \{x + t\vec{n} : x \in H \text{ and } t < 0\}$. Take $j \in \{1, \dots, d\}$ and let n_j be the j-th component of $\vec{n}$. Finally, denote by σ Lebesgue measure on H (defined by introducing an orthonormal basis for H). Let $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ be such that the restriction to $\mathbb{R}^d \setminus H$ is of class C^1 and such that $(\partial f/\partial x_j) \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Suppose also that

$$f_+(x) = \lim_{h \in H^+,\, y \to x} f(y) \quad \text{and} \quad f_-(x) = \lim_{h \in H^-,\, y \to x} f(y)$$

  exist at every point x of H. Then

$$D_j[f] = \left[\frac{\partial f}{\partial x_j}\right] + n_j(f_+ - f_-)\sigma.$$

- *Page 306, Ex. 16.* ***a.*** $x\delta^{(j)} = -j\delta^{(j-1)}$. ***b.*** $-\delta/3$. ***c.*** $U = \lambda x^{1/2}$, $V = \mu(-x)^{1/2}$.
  e. $T = -(\delta/3) + \lambda\sqrt{x}\,Y(x) + \mu\sqrt{-x}\,Y(-x)$.
- *Page 312, Ex. 1.* $E(x) = (e^{3x} - e^{-x})Y(x)/4$.
- *Page 313, Ex. 4b.* $C_d^k = \left(s_d 2^{k-1}(k-1)! \prod_{j=1}^{k}(2j-d)\right)^{-1}$.
- *Page 314, Ex. 5.* ***a-i.*** $D_j f = ((\varphi'(r)/r) - (\varphi(r)/r^2))(x_j/r)$,

$$D_j^2 f = \frac{\varphi''(r)}{r}\frac{x_j^2}{r^2} + \left(\frac{\varphi'(r)}{r^2} - \frac{\varphi(r)}{r^3}\right)\left(1 - 3\frac{x_j^2}{r^2}\right), \quad \Delta f = \frac{\varphi''(r)}{r}.$$

  a-ii. $\Delta f = (\varphi''(r))/r) - 4\pi\varphi(0)\delta$.
  b. If $\lambda > 0$, then $\varphi(r) = -(1/4\pi)\cos\sqrt{\lambda}\,r + C\sin\sqrt{\lambda}\,r$.
  If $\lambda = 0$, then $\varphi(r) = -(1/4\pi) + Cr$.
  If $\lambda < 0$, then $\varphi(r) = -(1/4\pi)\cosh\sqrt{-\lambda}\,r + C\sinh\sqrt{-\lambda}\,r$.
  c. $E_\lambda(x) = -e^{-\sqrt{-\lambda}r}/(4\pi r)$.

Chapter 9

- *Page 322, Ex. 3.* ***a.*** $T_\lambda(x) = T(\lambda x)$. ***c.*** The degree of $\delta^{(k)}$ is $-1-k$. ***d.*** The degree of $\operatorname{pv}(1/x)$ is -1. The distribution $\operatorname{fp}(Y(x)/x)$ is not homogeneous. ***e.*** $T = \lambda Y(x) + \mu Y(-x)$. ***g.*** $(x^2 Y(x)) \otimes \delta'$ has order 1 and degree 0.

- *Page 323, Ex. 4.* ***a.*** Y^J is a fundamental solution of $D_J = \prod_{j\in J} D_j$.
  b. $D^p(x^pY(x)) = p!Y(x)$. If $p \geq 1$, then $x^{p-1}Y(x)/(p-1)!$ is a fundamental solution of D^p. ***c.*** $E = E_1 \otimes \cdots \otimes E_d$ with $E_j(x_j) = x_j^{p_j-1}Y(x_j)/(p_j-1)!$ if $p_j \geq 1$ and $E_j = \delta$ if $p_j = 0$.
- *Page 323, Ex. 5.* T has order 2.
- *Page 335, Ex. 1.* $\delta_x * \delta_y = \delta_{x+y}$.
- *Page 335, Ex. 2.* $P(D)\delta * Q(D)\delta = (PQ)(D)\delta$.
- *Page 335, Ex. 5.* ***b-i.*** $Q(X) = P(X_1-a_1, \dots, X_d-a_d)$. ***b-ii.*** $e^L E$.
  c. $\prod_{j=1}^d e^{a_j x_j} Y(x_j)$.
- *Page 336, Ex. 6b.* The sequence $(P_n)_{n\in\mathbb{N}}$ converges to δ.
- *Page 345, Ex. 2a.* They are the distributions associated with functions of the form $f+g$, where f is a convex function on $\mathbb{R}$ and g an affine function with values in $\mathbb{C}$.
- *Page 348, Ex. 12.* $T = S''+S$. T is locally integrable if and only if $S \in C^1(\mathbb{R})$ and S' is absolutely continuous on $\mathbb{R}$.

Chapter 10

- *Page 371, Ex. 3c.* $J(u) = \min_{v\in H_0^1(\Omega)} J(v)$, with

$$J(v) = \tfrac{1}{2} \sum_{1\leq i,j\leq d} (a_{i,j}D_jv \,|\, D_jv)_{L^2} - \tfrac{1}{2}(cv \,|\, v)_{L^2} + (f \,|\, v)_{L^2}.$$

- *Page 372, Ex. 4a.* $u \in H_0^1(\Omega)$ and

$$\sum_{1\leq i,j\leq d} \frac{\partial}{\partial x_i}\Big(a_{i,j}\frac{\partial}{\partial x_j}\Big) u + cu - \sum_{i=1}^d b_i \frac{\partial}{\partial x_i} u = -f.$$

- *Page 375, Ex. 7.* If $u \in H_0^1(\Omega)$, then $\Delta u = \lambda u$ on Ω if and only if, for $i=1$ and $i=2$, $u_i = u_{|\Omega_i} \in H_0^1(\Omega_i)$ satisfies $\Delta u_i = \lambda u_i$ on Ω_i. Therefore $\{\lambda_n(\Omega)\}_{n\in\mathbb{N}} = \{\lambda_n(\Omega_1)\}_{n\in\mathbb{N}} \cup \{\lambda_n(\Omega_2)\}_{n\in\mathbb{N}}$.
- *Page 376, Ex. 10e.* Suppose for example that Ω is the disjoint union of two open sets Ω_1 and Ω_2, and let $\lambda_0(\Omega_1)$ and $\lambda_0(\Omega_2)$ be the first eigenvalue of the Dirichlet Laplacian on Ω_1 and Ω_2, respectively. If $\lambda_0(\Omega_1) = \lambda_0(\Omega_2)$ (for example, if Ω_2 is a translate of Ω_1), then $\lambda_0 = \lambda_0(\Omega_1)$ and E has dimension 2. If $|\lambda_0(\Omega_2)| > |\lambda_0(\Omega_1)|$ (for example, if $d=1$ and $\Omega = (0,2)\cup(3,4)$), then $\lambda_0 = \lambda_0(\Omega_1)$, E has dimension 1 and every element of E vanishes on Ω_2.

References

For certain topics, one can consult, among the books whose level is comparable to that of ours:

- Haim Brézis, *Analyse fonctionnelle: théorie et applications*, Masson, 1983.
- Gustave Choquet, *Topology*, Academic Press, 1966.
- Jean Dieudonné, *Foundations of modern analysis*, Academic Press, 1960; enlarged and corrected edition 1969 (vol. 1 of his *Treatise on analysis*).
- Serge Lang, *Real and functional analysis*, Springer, 1983 (2nd edition), 1993 (3rd edition).
- Pierre Arnaud Raviart and Jean-Marie Thomas, *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, 1983.
- Daniel Revuz, *Mesure et intégration*, Hermann, 1994.
- Walter Rudin, *Real and complex analysis*, McGraw-Hill, 1966, 1974 (2nd edition), 1987 (3rd edition).
- Walter Rudin, *Functional analysis*, McGraw-Hill, 1973, 1991 (2nd edition).
- Laurent Schwartz, *Analyse hilbertienne*, Hermann, 1979.
- Laurent Schwartz, *Mathematics for the physical sciences*, Hermann and Addison-Wesley, 1966.
- Claude Zuily, *Distributions et équations aux dérivées partielles: exercices corrigés*, Hermann, 1986.

It may be interesting to consult the "foundational texts", among which we cite:

- Stefan Banach, *Theory of linear operations*, North-Holland and Elsevier, 1987. Original French: *Théorie des opérations linéaires*, 1932, reprinted by Chelsea (1955), Jacques Gabay.
- Frigyes Riesz and Béla Sz.-Nagy, *Functional analysis*, Ungar, 1955; reprinted by Dover, 1960.
- Richard Courant and David Hilbert, *Methods of mathematical physics*, Interscience, 1953.
- Laurent Schwartz, *Théorie des distributions*, Hermann, 1966 (revised edition).

Among the great treatises in functional analysis, we mention:

- Nelson Dunford and Jacob T. Schwartz, *Linear operators* (3 vols.), Interscience, 1958–1971; reprinted 1988.
- Robert E. Edwards, *Functional analysis: theory and applications*, Holt, Rinehart and Winston, 1965; reprinted by Dover.
- I. M. Gel'fand and G. E. Shilov, *Generalized functions*, vols. 1, 2, 3, and I. M. Gel'fand and N. Y. Vilenkin, *Generalized functions*, vol. 4, Academic Press, 1964–1968.
- Michael Reed and Barry Simon, *Methods of modern mathematical physics* (4 vols.), Academic Press, 1972–1980.
- Kosaku Yosida, *Functional analysis*, Springer, 1955, 1980 (5th edition).
- Robert Dautray and Jacques-Louis Lions,*Analyse mathématique et calcul numérique* (9 vols.), Masson, 1984–1988. This work is more applied that the preceding ones.

This is but a very partial bibliography on the subject. For more references, the reader is encouraged to consult the bibliographies of the books listed above.

Index

Graduate Texts in Mathematics

(continued from page ii)

120 ZIEMER. Weakly Differentiable Functions: Sobolev Spaces and Functions of Bounded Variation.
121 LANG. Cyclotomic Fields I and II. Combined 2nd ed.
122 REMMERT. Theory of Complex Functions. *Readings in Mathematics*
123 EBBINGHAUS/HERMES et al. Numbers. *Readings in Mathematics*
124 DUBROVIN/FOMENKO/NOVIKOV. Modern Geometry—Methods and Applications. Part III.
125 BERENSTEIN/GAY. Complex Variables: An Introduction.
126 BOREL. Linear Algebraic Groups. 2nd ed.
127 MASSEY. A Basic Course in Algebraic Topology.
128 RAUCH. Partial Differential Equations.
129 FULTON/HARRIS. Representation Theory: A First Course. *Readings in Mathematics*
130 DODSON/POSTON. Tensor Geometry.
131 LAM. A First Course in Noncommutative Rings.
132 BEARDON. Iteration of Rational Functions.
133 HARRIS. Algebraic Geometry: A First Course.
134 ROMAN. Coding and Information Theory.
135 ROMAN. Advanced Linear Algebra.
136 ADKINS/WEINTRAUB. Algebra: An Approach via Module Theory.
137 AXLER/BOURDON/RAMEY. Harmonic Function Theory.
138 COHEN. A Course in Computational Algebraic Number Theory.
139 BREDON. Topology and Geometry.
140 AUBIN. Optima and Equilibria. An Introduction to Nonlinear Analysis.
141 BECKER/WEISPFENNING/KREDEL. Gröbner Bases. A Computational Approach to Commutative Algebra.
142 LANG. Real and Functional Analysis. 3rd ed.
143 DOOB. Measure Theory.
144 DENNIS/FARB. Noncommutative Algebra.
145 VICK. Homology Theory. An Introduction to Algebraic Topology. 2nd ed.
146 BRIDGES. Computability: A Mathematical Sketchbook.
147 ROSENBERG. Algebraic K-Theory and Its Applications.
148 ROTMAN. An Introduction to the Theory of Groups. 4th ed.
149 RATCLIFFE. Foundations of Hyperbolic Manifolds.
150 EISENBUD. Commutative Algebra with a View Toward Algebraic Geometry.
151 SILVERMAN. Advanced Topics in the Arithmetic of Elliptic Curves.
152 ZIEGLER. Lectures on Polytopes.
153 FULTON. Algebraic Topology: A First Course.
154 BROWN/PEARCY. An Introduction to Analysis.
155 KASSEL. Quantum Groups.
156 KECHRIS. Classical Descriptive Set Theory.
157 MALLIAVIN. Integration and Probability.
158 ROMAN. Field Theory.
159 CONWAY. Functions of One Complex Variable II.
160 LANG. Differential and Riemannian Manifolds.
161 BORWEIN/ERDÉLYI. Polynomials and Polynomial Inequalities.
162 ALPERIN/BELL. Groups and Representations.
163 DIXON/MORTIMER. Permutation Groups.
164 NATHANSON. Additive Number Theory: The Classical Bases.
165 NATHANSON. Additive Number Theory: Inverse Problems and the Geometry of Sumsets.
166 SHARPE. Differential Geometry: Cartan's Generalization of Klein's Erlangen Program.
167 MORANDI. Field and Galois Theory.
168 EWALD. Combinatorial Convexity and Algebraic Geometry.
169 BHATIA. Matrix Analysis.
170 BREDON. Sheaf Theory. 2nd ed.
171 PETERSEN. Riemannian Geometry.
172 REMMERT. Classical Topics in Complex Function Theory.
173 DIESTEL. Graph Theory.
174 BRIDGES. Foundations of Real and Abstract Analysis.
175 LICKORISH. An Introduction to Knot Theory.
176 LEE. Riemannian Manifolds.
177 NEWMAN. Analytic Number Theory.
178 CLARKE/LEDYAEV/STERN/WOLENSKI. Nonsmooth Analysis and Control Theory.
179 DOUGLAS. Banach Algebra Techniques in Operator Theory. 2nd ed.

180 SRIVASTAVA. A Course on Borel Sets.
181 KRESS. Numerical Analysis.
182 WALTER. Ordinary Differential Equations.
183 MEGGINSON. An Introduction to Banach Space Theory.
184 BOLLOBAS. Modern Graph Theory.
185 COX/LITTLE/O'SHEA. Using Algebraic Geometry.
186 RAMAKRISHNAN/VALENZA. Fourier Analysis on Number Fields.
187 HARRIS/MORRISON. Moduli of Curves.
188 GOLDBLATT. Lectures on the Hyperreals: An Introduction to Nonstandard Analysis.
189 LAM. Lectures on Modules and Rings.
190 ESMONDE/MURTY. Problems in Algebraic Number Theory.
191 LANG. Fundamentals of Differential Geometry.
192 HIRSCH/LACOMBE. Elements of Fuctional Analysis.