

BASES OF FUNCTIONAL ANALYSIS

Jean-Yves CHEMIN

Laboratoire J.-L. Lions, Campus Pierre et Marie Curie, Case 187

Sorbonne Université, 4 Place Jussieu

75230 Paris Cedex 05, France

Fax: +33 1 44 27 72 00, e-mail: chemin@ann.jussieu.fr

September 11, 2018

Update: frederic.klopp@imj-prg.fr (25-11-21)

Contents

1	Metric spaces	7
1.1	Definition of metric spaces	7
1.2	Complete spaces	14
1.3	Compactness	18
2	Normed spaces, Banach spaces	27
2.1	Definitions of normed spaces and Banach spaces	27
2.2	Spaces of continuous linear maps	32
2.3	Banach spaces, compactness and finite dimension	38
2.4	Compactness in the space of continuous functions: Ascoli's theorem	40
2.5	About the Stone-Weierstrass theorem	41
2.6	Notions on separable metric spaces	46
3	Duality in Banach spaces	51
3.1	A presentation of the concept of duality	51
3.2	Identifying a normed space as a dual space	54
3.3	A weaker sense for convergence in E'	57
4	Hilbert spaces	63
4.1	Orthogonality	63
4.2	Properties of Hilbert spaces	65
4.3	Duality in Hilbert spaces	70
4.4	The adjoint of an operator, self-adjoint operators	73
5	L^p spaces	81
5.1	Measure theory and definition of the L^p spaces	82
5.2	The L^p spaces are Banach spaces	84
5.3	Density in L^p spaces	90
5.4	Convolution and smoothing	97
5.5	Duality between L^p and $L^{p'}$	103
6	The Dirichlet problem	109
6.1	A classical approach to the problem	110
6.2	The concept of quasi-derivatives	111
6.3	The space $H_0^1(\Omega)$ and the Dirichlet problem	113

7	The Fourier transform	117
7.1	The Fourier transform on $L^1(\mathbb{R}^d)$	117
7.2	The inversion formula and the Fourier-Plancherel theorem	121
7.3	Proof of Rellich's theorem	123
8	Tempered distributions in dimension 1	125
8.1	Definition of tempered distributions and examples	126
8.2	Operations on tempered distributions	131
8.3	Two applications	142

Introduction

These are the lecture notes for the “Bases of Functional Analysis” course, part of the 1st year of the Master’s degree in Mathematics at Pierre & Marie Curie University, Paris. The aim of the course is for the student to acquire an elementary and solid mastery of tools which are fundamental in mathematics, both as a core subject (geometry, probability theory, partial differential equations) and in its applications to physics, mechanics, the study of large-scale systems, imagery, statistics, etc.

First, let us make some general remarks about these notes. It is not a treatise, in that some classical results, like the Cauchy-Lipschitz theorem, are absent, having been taught in other courses, while others, like the Stone-Weierstrass theorem, are stated in a specific framework. Some easy proofs, which are merely simple applications of definitions, may be sketched or omitted - the detailed writing of these proofs make for excellent learning exercises. Generally speaking, readers who wish to have a good understanding of the concepts we introduce should rework the course’s proofs by themselves.

Proofs which are considered to be either non-essential to the course or too difficult, shall be written in small print. These are not studied during the lectures, but are included in the notes to satisfy the curiosity of motivated students.

The notes are structured as follows: In the first chapter, we show basic results on the topology of metric spaces, including the notions of complete metric spaces and compact metric spaces. A solid mastery of the contents of this chapter is absolutely imperative.

The second chapter deals with the study of normed vector spaces, fundamental examples of which are function spaces. A key point here is understanding the effects of working in infinite dimension (which is the case in function spaces) on topology. Ascoli’s theorem, a compactness criterion for parts of continuous function spaces, illustrates the difficulties which appear in infinite dimension frameworks.

The third chapter deals with the notion of duality. It may be short, but it is fundamental. Duality is the basis of the theory of distributions, a major breakthrough in analysis at the start of the XXth century - this will be studied in chapter 8. Beyond the concept of transposes of linear maps, this chapter explains the procedure which allows one to identify the dual of a Banach space - another Banach space with a weaker notion of convergence, induced by the fact that it is a dual space, that we call “weak-star convergence”.

The fourth chapter is a classic: Hilbert spaces, which extend the notion of Euclidean spaces to infinite dimension.

The fifth chapter studies the spaces of functions which have a finite integral relative to a measure when elevated to a power p . We start by recalling fundamental results of integration theory, without proof. Another important notion, the convolution of functions, is defined, studied, then applied to approximation.

The sixth chapter studies the Dirichlet problem on a bounded domain. The goal is to prove that, for any square-integrable function f defined on a bounded, connected open set Ω of \mathbb{R}^d , there exists a unique solution u in a function space we will define (the Sobolev space $H_0^1(\Omega)$), which solves, in a broader sense than usual, the Laplace equation

$$-\sum_{j=1}^d \frac{\partial^2 u}{\partial x_j^2} = f,$$

and such that the function u is zero on the boundary of the open set Ω . We find this solution by determining if the infimum of the function

$$u \mapsto \int_{\Omega} \sum_{j=1}^d \left| \frac{\partial u}{\partial x_j} \right|^2 dx - \int_{\Omega} f(x)u(x)dx$$

on the set of continuously differentiable functions with compact support in Ω , is reached. Many previously established concepts and results come into play here, and ideas from the theory of distributions, which will be developed in chapter 8, manifest themselves for the first time.

The seventh chapter deals with the Fourier transform on the space of integrable functions on \mathbb{R}^d and many applications. This chapter is fundamental and will be crucial in the following two.

The eighth chapter presents the theory of tempered distributions. We choose not to go into the general theory of distributions to keep things simple. The basic idea is that, when we know how to define an operation on very smooth and rapidly decreasing functions on \mathbb{R}^d (e.g. functions in the Schwartz space), we can use duality to extend it to a space of tempered distributions, which contain functions (hence “distributions” are also called “generalised functions”) as well as some very singular objects. Of course, this chapter contains examples which must be studied and understood in order to properly comprehend and apply this theory.

Translator’s note. The convention we use throughout these notes is that “positive” means “strictly positive” ($x > 0$), and that “increasing” means “strictly increasing” ($x < y$ implies $f(x) < f(y)$). For coherence, mathematical notations have been chosen strictly identical to the original version.

Chapter 1

Metric spaces

Introduction

This chapter is a summary of basic results on metric spaces. The first section presents the notion of metrics, or distances, which is very natural and intuitive, and shows how it allows one to generalise the convergence of sequences and the continuity of functions (both notions that one is already familiar with in the real- and complex-variable settings). Metrics also allow one to give abstract definitions of open and closed subsets, concepts which are constantly used in functional analysis.

In the second section, we introduce complete metric spaces, which are spaces in which all Cauchy sequences are convergent. This notion is fundamental: in these spaces, we can prove that certain sequences converge without needing any prior information about their limit. A crucial tool, it is very commonly used in analysis to prove existence theorems - the Cauchy-Lipschitz theorem is one such result.

In the third section, we discuss compact metric spaces. Functional analysis requires the user to go beyond the elementary description of compact subsets of \mathbb{R}^N .

On the whole, this chapter is rather abstract. Examples, illustrations and applications of the fundamental notions presented here will be plentiful in the remainder of the course.

1.1 Definition of metric spaces

Definition 1.1.1. *Let X be a set. We call a metric, or distance function, on X any function d from $X \times X$ to \mathbb{R}^+ such that*

$$\begin{aligned}d(x, y) &= 0 \iff x = y \\d(x, y) &= d(y, x) \\d(x, y) &\leq d(x, z) + d(z, y)\end{aligned}$$

The pair (X, d) is called a metric space.

Some examples

- Set $X = \mathbb{R}$ and $d(x, y) = |x - y|$. These define a metric space.

- Set $X = \mathbb{R}^N$, and choose among the following distance functions:

$$d_e(x, y) \stackrel{\text{def}}{=} \left(\sum_{j=1}^N (x_j - y_j)^2 \right)^{\frac{1}{2}}$$

$$d_\infty(x, y) \stackrel{\text{def}}{=} \max_{1 \leq j \leq N} |x_j - y_j|$$

$$d_1(x, y) \stackrel{\text{def}}{=} \sum_{j=1}^N |x_j - y_j|.$$

- More generally, let us consider $(X_j, d_j)_{1 \leq j \leq N}$, a finite family of metric spaces.

Set $X = \prod_{j=1}^N X_j$, and define

$$D_\infty \begin{cases} X \times X & \longrightarrow \mathbb{R}^+ \\ (x, x') & \longmapsto \max_{1 \leq j \leq N} d_j(x_j, y_j) \end{cases} \quad \text{and} \quad D_1 \begin{cases} X \times X & \longrightarrow \mathbb{R}^+ \\ (x, x') & \longmapsto \sum_{j=1}^N d_j(x_j, y_j) \end{cases} \quad (1.1)$$

Both of these functions are metrics on X .

- Let us return to $X = \mathbb{R}$, let f be an one to one function from \mathbb{R} to \mathbb{R} , and we define

$$d_f(x, y) = |f(x) - f(y)|.$$

The function d_f is a metric on X .

The following exercise shows how we can define a metric on the space of sequences of elements of a given metric space.

Exercise 1.1.1. Let (X, d) be a metric space, and let $(a_n)_{n \in \mathbb{N}}$ be a sequence of positive real numbers such that

$$\sum_{n \in \mathbb{N}} a_n < \infty.$$

We consider the set $X^{\mathbb{N}}$ made up of the sequences of elements of X . We define

$$D_a \begin{cases} X^{\mathbb{N}} \times X^{\mathbb{N}} & \longrightarrow \mathbb{R}^+ \\ (x, y) & \longmapsto \sum_{n \in \mathbb{N}} \min\{a_n, d(x(n), y(n))\}. \end{cases}$$

Prove that D_a is a metric on $X^{\mathbb{N}}$.

Definition 1.1.2. Let (X, d) be a metric space, x be a point in X and α a positive real number. The open (resp. closed) ball with centre x and radius α , denoted $B(x, \alpha)$ (resp. $B_f(x, \alpha)$) the set of points y in X such that $d(x, y) < \alpha$ (resp. $d(x, y) \leq \alpha$).

Metrics allow one to give very simple and general definitions for the concepts of the limit of a sequence and the continuity of a function.

Definition 1.1.3 (convergent sequence). Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a metric space (X, d) and let ℓ be a point of X . The sequence $(x_n)_{n \in \mathbb{N}}$ is said to converge to ℓ if and only if

$$\forall \varepsilon > 0, \exists n_0 / n \geq n_0 \implies d(x_n, \ell) < \varepsilon.$$

Exercise 1.1.2. We consider a metric space (X, d) and the metric D_a defined on $X^{\mathbb{N}}$ in exercise 1.1.1. Show that a sequence $(x_p)_{p \in \mathbb{N}}$ of elements of $X^{\mathbb{N}}$ converges to x in the sense of the metric D_a if and only if

$$\forall n \in \mathbb{N}, \lim_{p \rightarrow \infty} d(x_p(n), x(n)) = 0.$$

Definition 1.1.4 (continuous function). Let (X, d) and (Y, δ) be two metric spaces. We consider a function f from X to Y and a point x_0 in X . The function f is said to be continuous at the point x_0 if and only if

$$\forall \varepsilon > 0, \exists \alpha > 0 / d(x, x_0) < \alpha \implies \delta(f(x), f(x_0)) < \varepsilon.$$

Proposition 1.1.1 (composition of continuous functions). Let (X, d) , (Y, δ) and (Z, ρ) be three metric spaces, and f et g be two functions, from X to Y and from Y to Z respectively. Let x_0 be a point of X such that f is continuous at x_0 and g is continuous at $f(x_0)$. Then the function $g \circ f$ is continuous at the point x_0 .

Proof. Since the function g is continuous at $f(x_0)$, we have

$$\forall \varepsilon > 0, \exists \alpha > 0 / \delta(y, f(x_0)) < \alpha \implies \rho(g(y), (g \circ f)(x_0)) < \varepsilon.$$

Since the function f is continuous at x_0 , we have

$$\exists \beta > 0 / d(x, x_0) < \beta \implies \delta(f(x), f(x_0)) < \alpha.$$

As the real number β is chosen for each ε , we deduce that

$$\exists \beta > 0 / d(x, x_0) < \beta \implies \rho(g \circ f(x), (g \circ f)(x_0)) < \varepsilon.$$

This ends the proof. □

Proposition 1.1.2 (sequences and continuous functions). Let (X, d) and (Y, δ) be two metric spaces, f be a function from X to Y , and $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of X . We assume that $(x_n)_{n \in \mathbb{N}}$ converges to ℓ and that the function f is continuous at point ℓ . Then the sequence $(f(x_n))_{n \in \mathbb{N}}$ of elements of Y converges to $f(\ell)$.

The proof of this proposition is similar to the previous one, and is left as an exercise for the reader.

Definition 1.1.5 (Interior and closure, open and closed sets). Let (X, d) be a metric space, and A be a subset of X .

- We call the interior of A , denoted $\overset{\circ}{A}$, the set of points x in X such that there exists an open ball $B(x, \alpha)$, with $\alpha > 0$, which is a subset of A .
- We call the closure of A , denoted \overline{A} , the set of points x in X such that, for any open ball $B(x, \alpha)$, with $\alpha > 0$, the intersection of $B(x, \alpha)$ with A is non-empty.
- A set A is said to be dense in X if and only if $\overline{A} = X$.
- A set A is said to be open if and only if $\overset{\circ}{A} = A$.
- A set A is said to be closed if and only if $A = \overline{A}$.

Remark The definitions clearly imply that $\overset{\circ}{A} \subset A \subset \overline{A}$.

Exercise 1.1.3. Let A be a finite part of a metric space (X, d) . Prove that $\overline{A} = A$.

Proposition 1.1.3. The interior of an open ball B is B . The closure of a closed ball B' is B' .

Proof. Let y be a point of the open ball $B(x, \alpha)$, and let us consider the open ball with centre y and radius $\alpha - d(x, y)$ (which is a positive number). Using the triangular inequality, we have, for any z in $B(y, \alpha - d(x, y))$,

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + \alpha - d(x, y) = \alpha.$$

Now let us consider a point y in the closure of the closed ball with centre x and radius α . By definition, for any positive real number β , there exists a point z in $B(y, \beta) \cap B_f(x, \alpha)$. Once again using the triangular inequality, we get

$$d(x, y) \leq d(x, z) + d(z, y) < \beta + \alpha.$$

Hence, for any positive number β , $d(x, y) < \alpha + \beta$, so $d(x, y) \leq \alpha$, which ends the proof. \square

Remarks

- The above proposition states precisely that open balls are open sets, and that closed balls are closed sets (this consistency in terminology is reassuring news).
- The set X is both open and closed. We set the standard that the same occurs for the empty set.
- The proof above shows that the closure of the open ball $B(x_0, \alpha)$ is a subset of the closed ball $B_f(x_0, \alpha)$. However, it is not true in general that the closure of an open ball is the corresponding closed ball. For example, considering a generic set X and the metric d defined by $d(x, y) = 1$ if x is not equal to y , and $d(x, x) = 0$, we have $B(x_0, 1) = \{x_0\}$, and this set is also closed, therefore equal to its closure. Meanwhile, $B_f(x_0, 1) = X$.

Proposition 1.1.4. Let A be a subset of a metric space (X, d) . We have

$$(\overset{\circ}{A})^c = \overline{A^c} \quad \text{and} \quad \overline{A}^c = (\overset{\circ}{A^c}).$$

Proof. A point x in X belongs to $(\overset{\circ}{A})^c$ if and only if

$$\forall \alpha > 0, B(x, \alpha) \cap A^c \neq \emptyset,$$

which means precisely that x belongs to the closure of the complement of A . A point x belongs to $(\overline{A})^c$ if and only if

$$\exists \alpha > 0, B(x, \alpha) \cap A = \emptyset,$$

in other words $B(x, \alpha) \subset A^c$, which means that x belongs to $(\overset{\circ}{A^c})$. The proposition is proved. \square

Remark Proposition 1.1.4 states that the complement of an open set is closed, and vice-versa.

Proposition 1.1.5. Let A be a subset of a metric space (X, d) . A point x in X belongs to \overline{A} if and only if there exists a sequence $(a_n)_{n \in \mathbb{N}}$ of elements of A such that $\lim_{n \rightarrow \infty} a_n = x$.

Proof. Let us assume that x is the limit of a sequence $(a_n)_{n \in \mathbb{N}}$ of elements of A . For any positive real number α , there exists an integer n_0 such that x_{n_0} belongs to the ball $B(x, \alpha)$, which implies that $B(x, \alpha) \cap A \neq \emptyset$. So $x \in \overline{A}$.

Conversely, let us assume that $x \in \overline{A}$. Then, for any positive integer n , there exists an element a_n of X such that

$$a_n \in A \cap B(x, n^{-1})$$

Consider the sequence $(a_n)_{n \in \mathbb{N}}$ that we have defined: it satisfies

$$d(a_n, x) \leq \frac{1}{n}.$$

The sequence $(a_n)_{n \in \mathbb{N}}$ therefore converges to x , and the proposition is proved. \square

Proposition 1.1.6. *Any union of open sets is open. The intersection of a finite number of open sets is open. Any intersection of closed sets is closed. The union of a finite number of closed sets is closed.*

Proof. Let $(U_\lambda)_{\lambda \in \Lambda}$ be a family of open sets, and x be a point of $U \stackrel{\text{def}}{=} \bigcup_{\lambda \in \Lambda} U_\lambda$. Let $\lambda \in \Lambda$ be such that $x \in U_\lambda$. Since U_λ is open (therefore equal to its interior),

$$\exists \alpha > 0 / B(x, \alpha) \subset U_\lambda \subset U$$

so U is open. Now let $U = \bigcap_{j=1}^N U_j$ where each U_j is open. For every j in $\{1, \dots, N\}$, there exists

a positive real number α_j such that $B(x, \alpha_j) \subset U_j$. Set $\alpha \stackrel{\text{def}}{=} \min\{\alpha_j, j \in \{1, \dots, N\}\}$. For every j , the open ball $B(x, \alpha)$ is a subset of U_j , and is therefore in the intersection U . \square

Remark For a set X , let us consider Θ a part of $\mathcal{P}(X)$, which is the set of all subsets of X , which satisfies:

- the empty set and X belong to Θ ,
- if $(U_j)_{1 \leq j \leq N}$ is a finite family of elements of Θ , then $\bigcap_{j=1}^N U_j$ also belongs to Θ ,
- if $(U_\lambda)_{\lambda \in \Lambda}$ is a family of elements of Θ , then $\bigcup_{\lambda \in \Lambda} U_\lambda$ also belongs to Θ .

This defines what is called a topology on X , with the elements of Θ being defined as open sets of X , and closed sets are by definition complements of open sets. As the following proposition will show, the notions of convergent sequences and continuous functions can be defined in terms of open sets.

Proposition 1.1.7. *Let (X, d) be a metric space. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of X , and x be a point of X . The sequence $(x_n)_{n \in \mathbb{N}}$ converges to x if and only if for any open subset U of X containing x , there exists an integer n_0 such that*

$$\forall n \geq n_0, x_n \in U.$$

Let (Y, δ) be another metric space, f be a function from X to Y , and x_0 be an element of X . The function f is continuous at x_0 if and only if for any open subset V of Y containing $f(x_0)$, there exists an open subset U of X containing x_0 such that

$$f(U) \subset V.$$

The proof of this proposition makes for an educational exercise we highly recommend.

Theorem 1.1.1 (Characterisation of continuous functions). *Let f be a function between two metric space (X, d) and (Y, δ) . The following three statements are equivalent.*

- *The function f is continuous at every point of X .*
- *The preimage of an open set is open.*
- *The preimage of a closed set is closed.*

Proof. Assume that f is continuous at every point in X , and let us consider an open subset V of (Y, δ) and a point x in $f^{-1}(V)$. Since the set V is open, by definition there exists a positive number ε_0 such that $B(f(x), \varepsilon_0) \subset V$. As the function f is continuous at point x ,

$$\exists \alpha > 0, f(B(x, \alpha)) \subset B(f(x), \varepsilon_0).$$

Thus,

$$B(x, \alpha) \subset f^{-1}\left(f(B(x, \alpha))\right) \subset f^{-1}(B(f(x), \varepsilon_0)) \subset f^{-1}(V).$$

Conversely, let x_0 be a point of X . For any $\varepsilon > 0$, the ball $B(f(x_0), \varepsilon)$ is an open subset of (Y, δ) . By assumption, $f^{-1}(B(f(x_0), \varepsilon))$ is an open subset of (X, d) . So, there exists a positive α such that $B(x_0, \alpha)$ is a subset of $f^{-1}(B(f(x_0), \varepsilon))$, and therefore

$$f(B(x_0, \alpha)) \subset f\left(f^{-1}(B(f(x_0), \varepsilon))\right) \subset B(f(x_0), \varepsilon).$$

The equivalence between the second and third statements is a consequence of closed sets being complements of open sets (and vice-versa), and of

$$\begin{aligned} f^{-1}(V) &= \{x \in X / f(x) \in V\} \\ &= \{x \in X / f(x) \in V^c\}^c \\ &= (f^{-1}(V^c))^c. \end{aligned}$$

The theorem is proved. □

We can now ask questions about the effects of a change of metric on the properties of a metric space.

Definition 1.1.6. *Let X be a set, on which we consider two metrics d and δ . We say that these two metrics are topologically equivalent if and only if the identity map Id is continuous from (X, d) to (X, δ) , and from (X, δ) to (X, d) .*

Remark The open sets induced by two topologically equivalent metrics are the same; thus the same functions are continuous and the same sequences are convergent in both settings.

Proposition 1.1.8. *Let X be a set, and d and δ be two metrics on X . Then we have the following property.*

The metrics d and δ are topologically equivalent if and only if

$$\begin{aligned} \forall x \in X, \forall \varepsilon > 0, \exists \eta > 0 / \forall y \in Y, \\ d(x, y) < \eta \implies \delta(x, y) < \varepsilon \quad \text{and} \quad \delta(x, y) < \eta \implies d(x, y) < \varepsilon. \end{aligned}$$

Proving this proposition merely requires using the definitions; it is left to the reader.

Definition 1.1.7. Let A be a subset of a metric space (X, d) , and x be a point of X . We call distance between the point x and the set A the following quantity

$$d(x, A) \stackrel{\text{def}}{=} \inf_{a \in A} d(x, a).$$

Exercise 1.1.4. Prove that \overline{A} is the set of points x in X such that $d(x, A) = 0$.

Proposition 1.1.9. Let A be a subset of a metric space (X, d) . The function

$$d_A \begin{cases} X & \longrightarrow \mathbb{R}^+ \\ x & \longmapsto d(x, A) \end{cases}$$

is Lipschitz continuous, and 1 is a Lipschitz constant, that is

$$|d(x, A) - d(x', A)| \leq d(x, x').$$

Proof. Using the triangular inequality, we have, for any (x, y) in X^2 and any point a in A ,

$$d(x, a) \leq d(x, y) + d(y, a).$$

Since the infimum of a set is a lower bound of this set, we have, for any (x, y) in X^2 and any point a in A ,

$$d(x, A) \leq d(x, y) + d(y, A) \quad \text{and thus} \quad d(x, A) - d(x, y) \leq d(y, A).$$

As the infimum of a set is the largest lower bound of this set, we deduce that

$$d(x, A) - d(x, y) \leq d(y, A) \quad \text{and thus} \quad d(x, A) - d(y, A) \leq d(x, y).$$

The proof is finished by exchanging the positions of x and y . □

We will now introduce the notion of metric subspaces. Let (X, d) be a metric space, and A be a subset of X . It is natural to consider the metric space $(A, d|_{A \times A})$. We have the following property.

Proposition 1.1.10. Let (X, d) be a metric space, and A be a subset of X . Then B , a part of A , is an open (resp. closed) subset of the metric space $(A, d|_{A \times A})$ if and only if there exists an open (resp. closed) subset \tilde{B} of X such that $B = \tilde{B} \cap A$.

Proof. We will only deal with open subsets, and the case of closed subsets can be obtained by taking the complements¹ Let \tilde{B} be an open subset of X , and let us prove that $\tilde{B} \cap A$ is an open set in $(A, d|_{A \times A})$. Let a_0 be a point in $\tilde{B} \cap A$. Since \tilde{B} is open in X , there exists an open ball (for the metric space (X, d)) such that

$$B(a_0, \alpha) \subset \tilde{B}.$$

By intersecting with A , we know that $B(a_0, \alpha) \cap A$ is a subset of B . But $B(a_0, \alpha) \cap A$ is exactly the set of all a in A such that $d(a_0, a) < \alpha$. So B is an open subset of the metric space $(A, d|_{A \times A})$.

Conversely, let B be an open subset of the metric space $(A, d|_{A \times A})$. For every a in B , there exists a positive number α_a such that

$$B_A(a, \alpha_a) \subset A \quad \text{with} \quad B_A(a, \alpha_a) = \{a' \in A / d(a, a') < \alpha_a\}.$$

¹We strongly recommend writing the details of the closed subset case as an exercise.

Set $\tilde{B} \stackrel{\text{def}}{=} \bigcup_{a \in A} B_X(a, \alpha)$, which is open since it is a union of open sets. Since $B_A(a, \alpha) = B_X(a, \alpha) \cap A$, we have $\tilde{B} \cap A = B$, and the proposition is proved. \square

To finish off this introduction on metric spaces, we define the notion of diameter in a metric space.

Definition 1.1.8. *Let (X, d) a metric space. We will say that a subset A of X has finite diameter if and only if there exists a positive number C such that*

$$\forall (a, a') \in A^2, d(a, a') \leq C.$$

If A has finite diameter, the diameter of A is defined as the supremum of the set of quantities $d(a, a')$ with (a, a') in the set $A \times A$.

1.2 Complete spaces

Definition 1.2.1. *Let (X, d) be a metric space. We call a Cauchy sequence in X any sequence $(x_n)_{n \in \mathbb{N}}$ of elements of X that satisfies*

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} / \forall n \geq n_0, \forall m \geq n_0, d(x_n, x_m) < \varepsilon.$$

We note that if $(x_n)_{n \in \mathbb{N}}$ is a sequence of elements of a metric space (X, d) that is convergent to a point ℓ in X , then, since

$$d(x_n, x_m) \leq d(x_n, \ell) + d(\ell, x_m),$$

it is also a Cauchy sequence. Complete spaces are metric spaces in which the converse is true. This is told precisely in the following definition.

Definition 1.2.2. *Let (X, d) be a metric space. This space is said to be complete if and only if all Cauchy sequences in X are convergent.*

Let us provide a few examples of complete spaces. A fundamental example is the space \mathbb{R} endowed with the metric $d(x, y) = |x - y|$, which is complete due to the way it is constructed. We will now show how to make complete spaces based on some we already know. In other words, we will look for operations on metric spaces for which the completeness of these spaces is preserved.

Proposition 1.2.1. *Let $(X_1, d_1), \dots, (X_N, d_N)$ be a family of N complete metric spaces. If we set*

$$X = X_1 \times \dots \times X_N \quad \text{and} \quad d((x_1, \dots, x_N), (y_1, \dots, y_N)) = \max_{1 \leq j \leq N} d_j(x_j, y_j),$$

then the metric space (X, d) is complete.

The proof is left as an exercise. An immediate consequence is that the space \mathbb{R}^N , endowed with any one of the metrics d_e , d_1 or d_∞ , is complete.

The following exercise provides us with an interesting example.

Exercise 1.2.1. *Let (X, d) be a complete metric space. We consider on $X^{\mathbb{N}}$ the metric D_a introduced in exercise 1.1.1. Show that the metric space $(X^{\mathbb{N}}, D_a)$ is complete.*

Proposition 1.2.2. *Let (X, d) be a complete space. We consider a subset A of X . The metric space $(A, d|_{A \times A})$ is complete if and only if A is closed.*

Proof. Assume that $(A, d|_{A \times A})$ is complete. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of elements of A that converges to x in X . The sequence $(a_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in X made up of elements of A , so it is a Cauchy sequence in the space $(A, d|_{A \times A})$, which is complete. Therefore, there exists a in A such that the sequence $(a_n)_{n \in \mathbb{N}}$ converges to a in $(A, d|_{A \times A})$, which is a metric subspace of (X, d) . Since the limit of a sequence is unique, we have $a = x$ and x is in A .

Conversely, let us assume that A is closed, and consider a Cauchy sequence $(a_n)_{n \in \mathbb{N}}$ in the metric space $(A, d|_{A \times A})$. It is therefore a Cauchy sequence in (X, d) which is complete, so it is convergent to x in X . The fact that A is closed implies that x is in A , and thus, $(A, d|_{A \times A})$ is complete. \square

When a metric space (X, d) is complete, we can prove the existence of some objects. The following theorem is the most spectacular illustration of this.

Theorem 1.2.1 (contracting map fixed-point theorem, Picard). *Let f be a function from a complete metric space (X, d) to itself such that there exists a number k in $]0, 1[$ such that, for all $(x, y) \in X^2$,*

$$d(f(x), f(y)) \leq kd(x, y).$$

Then there exists a unique fixed point z in X , satisfying $f(z) = z$.

Proof. For a given an element x_0 in X , we consider the sequence $(x_n)_{n \in \mathbb{N}}$ defined by the iterative formula $x_{n+1} = f(x_n)$. We therefore write that

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq kd(x_n, x_{n-1}). \end{aligned}$$

We iterate these multiplications, and get that

$$d(x_{n+1}, x_n) \leq k^n d(x_1, x_0).$$

Thus, for any pair of integers (n, p) , we have

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{m=1}^p d(x_{n+m}, x_{n+m-1}) \\ &\leq d(x_1, x_0) \sum_{m=1}^p k^{n+m-1} \\ &\leq \frac{k^n}{1-k} d(x_1, x_0). \end{aligned}$$

Hence, the sequence $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. Let z be its limit. As the function f is continuous, seeing that it is Lipschitz continuous, we take the limit in the iterative formula defining $(x_n)_{n \in \mathbb{N}}$, and get that $z = f(z)$.

It remains to prove that the fixed point is unique. Let z_1 and z_2 be two solutions of $z = f(z)$. The hypothesis on f leads to

$$d(z_1, z_2) \leq kd(z_1, z_2).$$

The fact that k must be smaller than 1 means that $d(z_1, z_2) = 0$, hence $z_1 = z_2$. Thus, the theorem is proved. \square

Remark. This theorem is the key ingredient for numerous existence and uniqueness theorems. An important example is the Cauchy-Lipschitz theorem, guaranteeing existence and uniqueness for solutions of ordinary differential equations.

We shall now prove Baire's classic theorem, which has many applications in functional analysis. There will be some examples of these in chapter 2.

Theorem 1.2.2 (Baire category theorem). *Let (X, d) be a complete metric space. We consider a sequence $(U_n)_{n \in \mathbb{N}}$ of dense open subsets of X . Then, $\bigcap_{n \in \mathbb{N}} U_n$ is also dense.*

Proof. The proof relies largely on the following lemma, which is interesting in itself.

Lemma 1.2.1. *Let (X, d) be a complete metric space, and $(F_n)_{n \in \mathbb{N}}$ be a non-increasing sequence of non-empty closed subsets of X , such that the diameter of F_n tends to zero. Then there exists an element x in X such that*

$$\bigcap_{n \in \mathbb{N}} F_n = \{x\}.$$

Proof. Let us consider a sequence $(x_n)_{n \in \mathbb{N}}$ of elements of X such that, for every integer n , x_n belongs to F_n . Since the sequence of sets $(F_n)_{n \in \mathbb{N}}$ is non-increasing (in the sense that $F_{n+1} \subset F_n$), we have

$$\forall p \in \mathbb{N}, d(x_n, x_{n+p}) \leq \delta(F_n).$$

The fact that the diameter of the sets F_n tends to zero implies that the sequence $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, and therefore it converges to a point x in X . By making p go to infinity in the above statement, we get that every sequence $(x_n)_{n \in \mathbb{N}}$ such that x_n is in F_n satisfies $d(x_n, x) \leq \delta(F_n)$. This ends the proof of the lemma. \square

Back to the proof of theorem 1.2.2. Let V be an open subset of X . We are going to prove that $\bigcap_n U_n \cap V \neq \emptyset$, which is enough to prove the theorem.

The open set U_0 is dense in X , so $U_0 \cap V$ is a non-empty open set. Hence, there exist a positive number α_0 (which we can assume is less than 1) and a point x_0 in X such that

$$B_f(x_0, \alpha_0) \subset U_0 \cap V. \tag{1.2}$$

The open set U_1 is also dense, so the set $U_1 \cap B(x_0, \alpha_0)$ is a non-empty open set. So there exist a positive number α_1 (which we can assume is less than 1/2) and a point x_1 in X such that

$$B_f(x_1, \alpha_1) \subset U_1 \cap B(x_0, \alpha_0).$$

We proceed by induction and assume that we have built a sequence $(x_j)_{0 \leq j \leq n}$ of elements of X and a sequence $(\alpha_j)_{0 \leq j \leq n}$ such that, for each $j \leq n$, we have

$$\alpha_j \leq \frac{1}{j+1} \quad \text{and} \quad B_f(x_j, \alpha_j) \subset U_j \cap B(x_{j-1}, \alpha_{j-1}). \tag{1.3}$$

The open set U_{n+1} is dense, so the set $U_{n+1} \cap B(x_n, \alpha_n)$ is a non-empty open set. Therefore there exist a positive number α_{n+1} (which we can assume is less than $1/(n+2)$) and a point x_{n+1} in X such that

$$B_f(x_{n+1}, \alpha_{n+1}) \subset B(x_n, \alpha_n) \cap U_{n+1}.$$

Applying lemma 1.2.1 to the sequence $F_n = B_f(x_n, \alpha_n)$ implies that there exists a point x that belongs to the intersection of the closed balls $B_f(x_n, \alpha_n)$. Given the equations (1.2) and (1.3), we have

$$\forall n \in \mathbb{N}, x \in V \cap \bigcap_{j \leq n} U_j.$$

The Baire category theorem is proved. □

Often, the following statement is used, which is the Baire theorem in which we consider the complements.

Theorem 1.2.3. *Let (X, d) be a complete metric space, and consider a sequence $(F_n)_{n \in \mathbb{N}}$ of closed subsets of X which have empty interiors. Then the interior of the set $\bigcup_{n \in \mathbb{N}} F_n$ is empty.*

The Baire category theorem is also used as formulated in the following corollary, the proof of which is very easy, and is left as an exercise.

Corollary 1.2.1. *Let $(F_n)_{n \in \mathbb{N}}$ be a sequence of closed sets in a complete metric space (X, d) , such that the union of all the F_n is X . Then, there exists an integer n_0 such that $F_{n_0}^\circ \neq \emptyset$.*

In other words, a complete metric space cannot be the union of countably many closed sets with empty interiors. One noteworthy non-complete metric space is built this way: \mathbb{Q} .

Let us now introduce the notion of uniformly continuous functions.

Definition 1.2.3. *Let (X, d) and (Y, δ) be two metric spaces, and we consider a function f from X to Y . The function f is said to be uniformly continuous if and only if*

$$\forall \varepsilon > 0, \exists \alpha > 0 / d(x, x') < \alpha \implies \delta(f(x), f(x')) < \varepsilon.$$

Here are some examples and counter-examples. The function $x \mapsto x^2$ is uniformly continuous on any bounded interval $[a, b]$, but not on \mathbb{R} . The function $x \mapsto \sqrt{|x|}$ is uniformly continuous on \mathbb{R} .

Theorem 1.2.4 (extension of uniformly continuous functions). *Let (X, d) and (Y, δ) be two metric spaces, A be a dense subset of X , and f be a uniformly continuous function from (A, d) to (Y, δ) . If Y is complete, then there exists a unique uniformly continuous function \tilde{f} from (X, d) to (Y, δ) such that $\tilde{f}|_A = f$.*

Proof. Let us consider an element x in X , and we will try to give a value to $\tilde{f}(x)$. Since the set A is dense in X , proposition 1.1.5 ensures that there exists a sequence $(a_n)_{n \in \mathbb{N}}$ of elements of A which converges to x . Also, the function f is uniformly continuous on A , which means that

$$\forall \varepsilon > 0, \exists \alpha / \forall (a, b) \in A^2, d(a, b) < \alpha \implies \delta(f(a), f(b)) < \varepsilon.$$

As the sequence $(a_n)_{n \in \mathbb{N}}$ converges, it is a Cauchy sequence, so there exists an integer n_0 such that

$$\forall n \geq n_0, \forall p, d(a_n, a_{n+p}) < \alpha.$$

This means that the sequence $(f(a_n))_{n \in \mathbb{N}}$ is a Cauchy sequence in Y , so it converges to a limit y .

The first thing to check is that this limit does not depend on the sequence $(a_n)_{n \in \mathbb{N}}$ we chose to construct it. To see that this is true, take $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$, two sequences that

converge to x . Reasoning similarly to above, the uniform continuity of the function f on A gives us

$$\lim_{n \rightarrow \infty} \delta(f(a_n), f(b_n)) = 0.$$

So the limit y does indeed not depend on the choice of the sequence $(a_n)_{n \in \mathbb{N}}$. Therefore, we define \tilde{f} as follows:

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} f(a_n) \text{ for any sequence } (a_n)_{n \in \mathbb{N}} \in A^{\mathbb{N}} / \lim_{n \rightarrow \infty} a_n = x.$$

Since the function f is continuous on A , it is clear that $\tilde{f}|_A = f$. We must now prove that \tilde{f} is uniformly continuous on X . The uniform continuity of \tilde{f} can be written mathematically as

$$\forall \varepsilon > 0, \exists \alpha > 0 / \forall (a, b) \in A^2, d(a, b) < \alpha \Rightarrow \delta(f(a), f(b)) < \varepsilon. \quad (1.4)$$

So let us consider a pair (x, y) of elements of X such that $d(x, y) < \alpha$. There exist two sequences $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ of elements of A such that

$$\lim_{n \rightarrow \infty} a_n = x \quad \text{and} \quad \lim_{n \rightarrow \infty} b_n = y.$$

There must exist an integer n_0 such that

$$n \geq n_0 \implies d(a_n, b_n) < \alpha.$$

Therefore, using equation (1.4), we have

$$\delta(f(a_n), f(b_n)) < \varepsilon.$$

Taking the limit, we obtain

$$d(x, y) < \alpha \implies \delta(\tilde{f}(x), \tilde{f}(y)) \leq \varepsilon,$$

which ends the proof of the theorem. □

1.3 Compactness

An important concept we will use in this section is that of cluster points of a sequence.

Definition 1.3.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a metric space (X, d) . We call the set of cluster points of the sequence $(x_n)_{n \in \mathbb{N}}$, and denote it $\text{Adh}(x_n)$ (from the French “valeur d’adhérence”), the set -which may be empty- defined by

$$\text{Adh}(x_n) \stackrel{\text{def}}{=} \bigcap_{n \in \mathbb{N}} \overline{A_n} \quad \text{with} \quad A_n \stackrel{\text{def}}{=} \{x_m, m \geq n\}.$$

Examples In \mathbb{R} endowed with the metric $|x - y|$, the sequence $(x_n)_{n \in \mathbb{N}}$ defined by $x_n = n$ is such that $\text{Adh}(x_n) = \emptyset$, and the sequence $(y_n)_{n \in \mathbb{N}}$ defined by $y_n = (-1)^{n+1}$ satisfies $\text{Adh}(y_n) = \{-1, 1\}$.

Proposition 1.3.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in a metric space (X, d) that we assume converges to a point ℓ in X . Then $\text{Adh}(x_n) = \{\ell\}$.

Proof. By definition of the limit of a sequence, for every $\varepsilon > 0$, there exists n_0 such that the set A_{n_0} defined above is included in the open ball $B(\ell, \varepsilon)$. So, if x is a point of X not equal to ℓ , by taking ε smaller than $d(\ell, x)$, we see that $x \notin \overline{A_{n_0}}$, thus $x \notin \overline{A_n}$ for all $n \geq n_0$. Hence the proposition is proved. □

Remark It is possible that a sequence $(x_n)_{n \in \mathbb{N}}$ satisfies $\text{Adh}(x_n) = \{\ell\}$, but does not converge. Consider, for example, the sequence of real numbers defined by

$$x_{2n} = 2n \quad \text{and} \quad x_{2n+1} = \frac{1}{2n+1}.$$

Even though this sequence does not converge to 0, we still have $\text{Adh}(x_n) = \{0\}$.

Proposition 1.3.2. *In a metric space (X, d) , any Cauchy sequence that possesses a cluster point ℓ converges to ℓ .*

Proof. Let us consider a Cauchy sequence $(x_n)_{n \in \mathbb{N}}$ with a cluster point ℓ . As it is a Cauchy sequence, there exists an integer n_0 such that

$$\forall n \geq n_0, \forall m \geq n_0, d(x_n, x_m) \leq \frac{\varepsilon}{2}.$$

As the sequence has a cluster point ℓ , there exists an integer $n_1 \geq n_0$ such that

$$d(x_{n_1}, \ell) < \frac{\varepsilon}{2}.$$

We therefore see that, for every integer $n \geq n_0$, we have

$$d(x_n, \ell) \leq d(x_n, x_{n_1}) + d(x_{n_1}, \ell) < \varepsilon.$$

The proposition is proved. □

Proposition 1.3.3. *Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a metric space (X, d) . A point ℓ in X belongs to $\text{Adh}(x_n)$ if and only if there exists an increasing function ϕ from \mathbb{N} to \mathbb{N} such that*

$$\lim_{n \rightarrow \infty} x_{\phi(n)} = \ell.$$

Proof. Before we starting the proof of this proposition, we remark that an increasing function from \mathbb{N} to \mathbb{N} satisfies

$$\forall n \in \mathbb{N}, \phi(n) \geq n. \tag{1.5}$$

This property is proved by induction. The statement is naturally true for $n = 0$. Now, we assume that it is true for some n . Since f is increasing, we have $\phi(n+1) > \phi(n) \geq n$, which implies that $\phi(n+1) \geq n+1$. In particular, the sequence $(\phi(n))_{n \in \mathbb{N}}$ must tend to infinity when n does so.

We return proving the proposition. If ℓ belongs to $\text{Adh}(x_n)$, we define the function ϕ by induction as follows: we choose $\phi(0) = 0$, then get $\phi(n+1)$ using $\phi(n)$ like so

$$\phi(n+1) \stackrel{\text{def}}{=} \min \left\{ m > \phi(n) / d(x_m, \ell) < \frac{1}{n+1} \right\}.$$

The function ϕ is increasing by construction, and we have, for every $n \geq 1$,

$$d(x_{\phi(n)}, \ell) \leq \frac{1}{n}, \quad \text{so} \quad \lim_{n \rightarrow \infty} x_{\phi(n)} = \ell.$$

Conversely, if there exists an increasing function ϕ such that $\lim_{n \rightarrow \infty} x_{\phi(n)} = \ell$, then for every $\varepsilon > 0$, there exists n_0 such that

$$\forall n \geq n_0, d(x_{\phi(n)}, \ell) < \varepsilon.$$

Using (1.5), that is $\phi(n) \geq n$, we deduce that

$$\forall \varepsilon > 0, \forall n \in \mathbb{N}, \exists m \geq n / d(x_m, \ell) < \varepsilon,$$

which means that $\ell \in \text{Adh}(x_n)$, hence the theorem is proved. □

Definition 1.3.2. We call extraction function, any increasing function from \mathbb{N} to \mathbb{N} . If $(x_n)_{n \in \mathbb{N}}$ is a sequence of elements of a set X , a subsequence is a sequence which can be written as $(x_{\phi(n)})_{n \in \mathbb{N}}$, where ϕ is an extraction function.

The notion of compact space can be seen from two angles: one using sequences, the other with balls that cover the space. The equivalence of these two viewpoints is given by the following theorem.

Theorem 1.3.1. Let (X, d) be a metric space. The following two statements are equivalent:
i) every sequence of elements of (X, d) has a cluster point;
ii) the metric space (X, d) is complete, and we have

$$\forall \varepsilon > 0, \exists (x_j)_{1 \leq j \leq N} / X = \bigcup_{j=1}^N B(x_j, \varepsilon). \quad (1.6)$$

Definition 1.3.3. Let (X, d) be a metric space. We say that X is a compact space if and only if one of the two conditions of the above theorem is satisfied.

Let us start a list of examples of compact metric spaces by noting that any finite metric space is compact.

Theorem 1.3.2. The metric spaces $([a, b], |x - y|)$ are compact.

Proof. If we use condition *ii)*, which is reasonable because the space $([a, b], |x - y|)$ is complete, it suffices to see that, for every positive ε ,

$$[a, b] = [a, a + \varepsilon] \cup \left(\bigcup_{k=1}^{N_\varepsilon - 1} [a + (k-1)\varepsilon, a + (k+1)\varepsilon] \right) \cup [b - N_\varepsilon \varepsilon, b] \quad \text{with} \quad N_\varepsilon \stackrel{\text{def}}{=} \left\lceil \frac{b-a}{\varepsilon} \right\rceil.$$

Using criterion *i)* requires a bit more work. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in the interval $[a, b]$, and we consider the set $A_n \stackrel{\text{def}}{=} \{x_m, m \geq n\}$, which is a set in \mathbb{R} which has an upper bound. It therefore has a supremum, which we denote M_n . The sequence $(M_n)_{n \in \mathbb{N}}$ is non-increasing and bounded from below, so it converges. The reader will notice that the limit of the sequence $(M_n)_{n \in \mathbb{N}}$ is a cluster point of $(x_n)_{n \in \mathbb{N}}$ (it is, in fact, the largest cluster point). \square

Proof of theorem 1.3.1. By using proposition 1.3.2, in order to prove that *i)* implies *ii)*, it suffices to show that *i)* implies statement (1.6). We will prove the contrapositive. Let us assume that the statement (1.6) is not satisfied. This means that there exists a positive number α such that we cannot cover X with a finite number of balls with radius α . Let x_0 be any element of X . There exists an element x_1 in X which does not belong to the ball $B(x_0, \alpha)$. Now let (x_0, \dots, x_p) be a p -tuple of elements of X such that, for every m not equal to n , we have $d(x_m, x_n) \geq \alpha$. We assume that

$$\bigcup_{n=1}^p B(x_n, \alpha) \neq X.$$

So, there exists a point x_{p+1} in X that does not belong to the above union of balls. By induction, we construct a sequence $(x_n)_{n \in \mathbb{N}}$ such that

$$m \neq n \Rightarrow d(x_m, x_n) \geq \alpha.$$

Such a sequence cannot have a cluster point, so we have proved that *i*) implies *ii*).

Let us now assume *ii*), and consider an infinite sequence $(x_n)_{n \in \mathbb{N}}$ of elements of X . The covering hypothesis implies that there exists an element α_0 in X such that the set \mathcal{X}_0 defined by

$$\mathcal{X}_0 \stackrel{\text{def}}{=} \left\{ m / x_m \in F_0 \stackrel{\text{def}}{=} B_f \left(\alpha_0, \frac{1}{2} \right) \right\}$$

is infinite. We set

$$\phi(0) \stackrel{\text{def}}{=} \min \mathcal{X}_0.$$

Likewise, there exists an element α_1 in X such that the set \mathcal{X}_1 defined by

$$\mathcal{X}_1 \stackrel{\text{def}}{=} \left\{ m \in \mathcal{X}_0 / x_m \in F_1 \stackrel{\text{def}}{=} F_0 \cap B_f \left(\alpha_1, \frac{1}{4} \right) \right\}$$

is infinite. We then set

$$\phi(1) \stackrel{\text{def}}{=} \min \{ m \in \mathcal{X}_1 / m > \phi(0) \}.$$

Let us assume that we have built a sequence $(F_m)_{1 \leq m \leq n}$ of closed subsets of X , an increasing sequence of integers $(\phi_m)_{1 \leq m \leq n}$ such that $\delta(F_m) \leq 2^{-m}$, and a sequence $(\mathcal{X}_m)_{1 \leq m \leq n}$ of subsets of \mathbb{N} defined such that

$$\mathcal{X}_m \stackrel{\text{def}}{=} \{ m' \in \mathcal{X}_{m-1} / x_{m'} \in F_{m-1} \}$$

is infinite. Then there exists an element α_{n+1} in X such that

$$\mathcal{X}_{n+1} \stackrel{\text{def}}{=} \left\{ m / x_m \in F_{n+1} \stackrel{\text{def}}{=} F_n \cap B_f \left(\alpha_{n+1}, \frac{1}{2^{n+1}} \right) \right\}$$

is infinite. We set

$$\phi(n+1) \stackrel{\text{def}}{=} \min \{ m \in \mathcal{X}_{n+1} / m > \phi(n) \}.$$

We therefore have, for every positive integer n ,

$$x_{\phi(n)} \in F_n \quad \text{and} \quad \delta(F_n) \leq \frac{1}{2^n}.$$

By lemma 1.2.1, there exists a point ℓ which belongs to the intersection of all the F_n . By definition of the diameter of a set, we have

$$\forall n \in \mathbb{N}, \quad d(x, x_{\phi(n)}) \leq \frac{1}{2^n}.$$

This ends the proof of the theorem. □

The following proposition yields many examples of compact spaces.

Proposition 1.3.4. *Let $(X_1, d_1), \dots, (X_N, d_N)$ be a family of N compact metric spaces. Set*

$$X = X_1 \times \dots \times X_N \quad \text{and} \quad d((x_1, \dots, x_N), (y_1, \dots, y_N)) = \max_{1 \leq j \leq N} d_j(x_j, y_j).$$

The metric space (X, d) is compact.

Proof. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements in X . By definition, for each integer j , there exists a sequence $(x_n^j)_{n \in \mathbb{N}}$ in X_j such that $x_n = (x_n^1, \dots, x_n^N)$. By definition of a compact space, there exists an extraction function ϕ_1 and a point x^1 in X^1 such that

$$\lim x_{\phi_1(n)}^1 = x^1.$$

Likewise, there exists an extraction function ϕ_2 and a point x_2 in X_2 such that

$$\lim x_{\phi_1 \circ \phi_2(n)}^2 = x^2.$$

We reiterate this process N times, and we construct an extraction function ϕ by setting

$$\phi \stackrel{\text{def}}{=} \phi_1 \circ \cdots \circ \phi_N,$$

so that, for every j in $\{1, \dots, N\}$, we have

$$\lim_{n \rightarrow \infty} x_{\phi(n)}^j = x^j.$$

Using the definition on the distance we defined on X , we see that the sequence $x_{\phi(n)}$ converges to (x^1, \dots, x^N) . The result is proved. \square

Definition 1.3.4. Let (X, d) be a metric space. A subset A is said to be a compact subset of X if and only if the metric space $(A, d|_{A \times A})$ is compact.

Proposition 1.3.5. Let A be a compact subset of a metric space (X, d) . If A is compact, then A is closed and has finite diameter, which means that

$$\delta(A) \stackrel{\text{def}}{=} \sup_{(a, a') \in A^2} d(a, a') < \infty.$$

Proof. Let x be a point in \overline{A} . By proposition 1.1.5, there exists a sequence $(a_n)_{n \in \mathbb{N}}$ of elements in A which converges to x . As we assume that A is compact, there exists an extraction function ϕ and a point ℓ in A such that $\lim_{n \rightarrow \infty} a_{\phi(n)} = \ell$. Since the sequence $(a_{\phi(n)})_{n \in \mathbb{N}}$ also converges to x , the uniqueness of the limit of a sequence implies that $\ell = x$, and therefore that $x \in A$. So A is a closed subset of X .

Now we prove that a compact subset must have finite diameter. Let us consider two sequences $(a_n)_{n \in \mathbb{N}}$ and $(b_n)_{n \in \mathbb{N}}$ of elements in A such that

$$\lim_{n \rightarrow \infty} d(a_n, b_n) = \delta(A).$$

Since A is compact, there exists an extraction function ϕ_0 and a point a in A such that the sequence $(a_{\phi_0(n)})_{n \in \mathbb{N}}$ converges to a . Likewise, there exists an extraction function ϕ_1 and a point b in A such that the sequence $(b_{\phi_1(n)})_{n \in \mathbb{N}}$ converges to b . Taking the limit in the above equation, we get

$$\delta(A) = \lim_{n \rightarrow \infty} d(a_{\phi_0 \circ \phi_1(n)}, b_{\phi_0 \circ \phi_1(n)}) = d(a, b).$$

The diameter of A is therefore finite. \square

Proposition 1.3.6. Let A be a subset of a complete metric space (X, d) . Its closure \overline{A} is compact if and only if for every positive real number ε , there exists a finite number of points $(x_j)_{1 \leq j \leq N}$ in X such that

$$A \subset \bigcup_{j=1}^N B(x_j, \varepsilon).$$

Proof. If \overline{A} is compact, then we are just rewriting definition 1.3.3. Conversely, let us assume the covering property for a subset A of X . First, we prove that the property is also true for its closure \overline{A} .

We are assuming that, for any positive number ε , there exists a finite sequence $(x_j)_{1 \leq j \leq N}$ of elements in X such that

$$A \subset \bigcup_{j=1}^N B\left(x_j, \frac{\varepsilon}{2}\right).$$

Let x be a point in \overline{A} . By definition, there exists a point a in A such that $d(x, a)$ is smaller than $\varepsilon/2$. So there exists an integer j in $\{1, \dots, N\}$ such that a belongs to $B(x_j, \varepsilon/2)$. Thus,

$$d(x, x_j) \leq d(x, a) + d(a, x_j) < \varepsilon,$$

so x belongs to $B(x_j, \varepsilon)$, which shows that the covering property is satisfied by \overline{A} . As the metric space (X, d) is complete, in order to prove that \overline{A} is compact, it suffices to show that we have the covering property with the points x_j being in the set \overline{A} . We know that, for every positive number ε , there exists a finite sequence $(x_j)_{1 \leq j \leq N}$ of elements of X such that

$$\overline{A} \subset \bigcup_{j=1}^N B\left(x_j, \frac{\varepsilon}{2}\right).$$

Of course, we only consider points x_j such that the open ball with centre x_j and radius $\varepsilon/2$ has a non-empty intersection with \overline{A} . Let a_j be a point in said intersection. Then, for any a in \overline{A} , there exists j such that

$$d(a, a_j) \leq d(a, x_j) + d(x_j, a_j) < \varepsilon,$$

which ends the proof. □

From theorem 1.3.2 and propositions 1.3.4 and 1.3.5, we deduce the following corollary.

Corollary 1.3.1. *A closed subset A in \mathbb{R}^d is compact if and only if*

$$\exists r > 0 / A \subset [-r, r]^d.$$

Exercise 1.3.1. *Let A be a subset of a metric space (X, d) . Prove that the closed set \overline{A} is compact if and only if any sequence of elements of A has a cluster point in A .*

Proposition 1.3.7. *Let A be a compact subset of a metric space (X, d) . If B is a closed subset of X that is contained in A , then B is also compact.*

Proof. Let $(b_n)_{n \in \mathbb{N}}$ be a sequence of elements of B . Since B is a subset of A , which is compact, there exists an extraction function ϕ such that $(b_{\phi(n)})_{n \in \mathbb{N}}$ converges to ℓ , which is an element of A . Since B is closed in X , $\ell \in \overline{B} = B$. Hence, B is compact. □

Theorem 1.3.3 (Heine). *Let (X, d) and (Y, δ) be two metric spaces, and let f be a continuous map from X to Y . Then, for any compact subset A of X , $f(A)$ is a compact subset of Y . Moreover, if X is compact, then f is uniformly continuous from X to Y .*

Proof. Let $(y_n)_{n \in \mathbb{N}}$ be a sequence of elements in $f(A)$, that is, there exists a sequence $(a_n)_{n \in \mathbb{N}}$ of elements in A such that $y_n = f(a_n)$. Since A is a compact subset of X , there exists a point a' in A and an extraction function ϕ such that

$$\lim_{n \rightarrow \infty} a_{\phi(n)} = a'$$

As the function f is continuous, we have

$$\lim_{n \rightarrow \infty} y_{\phi(n)} = f(a') \in f(A).$$

This proves the first part of the theorem.

To prove the second part, we consider the contrapositive. We assume that the metric space (X, d) is compact, and consider a function f on X which is not uniformly continuous. This means that there exists a positive real number ε and two sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ of elements in X such that

$$\forall n \in \mathbb{N}, d(x_n, y_n) \leq \frac{1}{n+1} \quad \text{and} \quad \delta(f(x_n), f(y_n)) \geq \varepsilon_0.$$

As the metric space (X, d) is compact, there also exists an extraction function ϕ and a point x in X such that $\lim_{n \rightarrow \infty} x_{\phi(n)}$ converges to x . As the distance between x_n and y_n converges to zero, the sequence $(y_{\phi(n)})_{n \in \mathbb{N}}$ also converges to x . The triangular inequality leads to

$$\delta(f(x_{\phi(n)}), f(x)) + \delta(f(x), f(y_{\phi(n)})) \geq \delta(f(x_{\phi(n)}), f(y_{\phi(n)})) \geq \varepsilon_0,$$

which implies that the function f is not continuous at the point x , by proposition 1.1.2. The theorem is proved. \square

Corollary 1.3.2. *Let f be a continuous function from (X, d) to $(\mathbb{R}, |x - y|)$. If (X, d) is compact, then the function f has a minimum and a maximum, i. e. there exist x_m and x_M in X such that*

$$\forall x \in X, f(x_m) \leq f(x) \leq f(x_M).$$

Proof. By Heine's theorem, theorem 1.3.3, the set $f(X)$ is compact in \mathbb{R} , so it is closed and a subset of some bounded interval $[a, b]$. This set therefore has an infimum and a supremum, which are cluster points of the closed set $f(X)$, and therefore, they belong to $f(X)$. \square

We will now give a characterisation of compactness in terms of covers by open sets.

Theorem 1.3.4. *Let (X, d) be a metric space. The following three statements are equivalent.*

i) *For any family $(U_\lambda)_{\lambda \in \Lambda}$ of open subsets of X that cover X , which means*

$$X = \bigcup_{\lambda \in \Lambda} U_\lambda,$$

we can extract a finite cover of X , i. e. there exists a finite sequence $(\lambda_j)_{1 \leq j \leq N}$ such that

$$X = \bigcup_{j=1}^N U_{\lambda_j}.$$

ii) *For any family $(F_\lambda)_{\lambda \in \Lambda}$ of closed subsets of X , we have*

$$\forall N, \forall (\lambda_1, \dots, \lambda_N) \in \Lambda^N / \bigcap_{j=1}^N F_{\lambda_j} \neq \emptyset \implies \bigcap_{\lambda \in \Lambda} F_\lambda \neq \emptyset.$$

iii) *Every sequence of elements of X has a cluster point.*

Proof. The equivalence of parts *i*) and *ii*) can be proved by considering the contrapositives. Indeed, the contrapositive of *ii*) is equivalent to:

“For any family $(F_\lambda)_{\lambda \in \Lambda}$ of closed subsets of X ,

$$\bigcap_{\lambda \in \Lambda} F_\lambda = \emptyset \implies \exists (\lambda_1, \dots, \lambda_N) \in \Lambda^N / \bigcap_{j=1}^N F_{\lambda_j} = \emptyset.”$$

This is seen to be equivalent to *i*) by taking the complements of the closed sets F_λ .

Let us now prove that *ii*) implies *iii*). By definition of the set of cluster points, we have

$$\text{Adh}(x_n) = \bigcap_n \bar{A}_n \quad \text{with} \quad A_n \stackrel{\text{def}}{=} \{x_m, m \geq n\}.$$

The sequence $(\bar{A}_n)_{n \in \mathbb{N}}$ is a non-increasing sequence of non-empty closed sets. So, by property *ii*), the intersection of all the \bar{A}_n (which is the set of cluster points) is non-empty, which proves point *iii*).

Let us now prove that *iii*) implies *i*), which is the trickiest part of the proof. We have to deduce a property about covers by open sets from a property on sequences. The following lemma is a crucial element, and it can be useful to prove other properties on compact spaces.

Lemma 1.3.1 (Lebesgue). *Let (X, d) be a compact metric space. For any family $(U_\lambda)_{\lambda \in \Lambda}$ of open sets which cover X , there exists a positive real number α such that*

$$\forall x \in X, \exists \lambda \in \Lambda / B(x, \alpha) \subset U_\lambda.$$

Proof. To prove this lemma, let us consider the function δ defined by

$$\delta \begin{cases} X & \longrightarrow \mathbb{R}_+^* \\ x & \longmapsto \sup \{ \beta / \exists \lambda / B(x, \beta) \subset U_\lambda \}. \end{cases}$$

Proving that the function δ has a minimum suffices to get the result. To prove this, let us establish the following.

$$\forall \varepsilon > 0, \forall x \in X, \exists \beta / d(x, y) < \beta \implies \delta(y) > \delta(x) - \varepsilon. \quad (1.7)$$

Let x be a point in X , and ε be a positive number smaller than $\delta(x)$. Assume that y is a point in X such that $d(x, y) < \varepsilon/2$. The triangular inequality implies that

$$B(y, \delta(x) - \varepsilon) \subset B(x, \delta(x) - \varepsilon/2) \subset U_\lambda,$$

so the statement (1.7) is proved.

Now, let us consider a sequence $(x_n)_{n \in \mathbb{N}}$ of elements in X such that $\lim_{n \rightarrow \infty} \delta(x_n) = \inf_{x \in X} \delta(x)$. By assumption, we can extract from the sequence $(x_n)_{n \in \mathbb{N}}$ a subsequence that converges to a point x_∞ in X . We will denote this subsequence $(x_n)_{n \in \mathbb{N}}$. By statement (1.7), for any $\varepsilon > 0$, there exists an integer n_0 such that

$$n \geq n_0 \implies \delta(x_n) > \delta(x_\infty) - \varepsilon.$$

Taking the limit, we deduce that

$$\inf_{x \in X} \delta(x) \geq \delta(x_\infty) - \varepsilon,$$

which is valid for any positive real number ε . By definition of the infimum, we get that $\delta(x_\infty) = \inf_{x \in X} \delta(x)$, and therefore $\inf_{x \in X} \delta(x)$ is positive. This ends the proof of Lebesgue’s lemma. \square

Conclusion of the proof of theorem 1.3.4

To prove that *iii*) implies *i*), we assume that X is compact (that is, any sequence has a converging subsequence), and we consider a family $(U_\lambda)_{\lambda \in \Lambda}$ of open subsets of X that cover X . The aim is to extract a finite ‘subcover’. Let $\alpha > 0$ be the radius given by Lebesgue’s lemma. By definition 1.3.3, there exists a finite number of balls B_1, \dots, B_N with radius α that cover X . Given the property that α

has, owing to Lebesgue's lemma, each of these balls is included in some open set U_{λ_j} in the family, and therefore

$$X \subset \bigcup_{j=1}^N B_j \subset \bigcup_{j=1}^N U_{\lambda_j}.$$

This ends the proof of the theorem. □

Exercise 1.3.2. Let (X, d) be a metric space. We consider the set $X^{\mathbb{N}}$ of sequences of elements in X , and the metric D_a on $X^{\mathbb{N}}$ given in exercise 1.1.1.

1) Prove that the sequence $(x_p)_{p \in \mathbb{N}}$ is a Cauchy sequence in $X^{\mathbb{N}}$ if and only if, for every integer n , the sequence $(x_p(n))_{p \in \mathbb{N}}$ is a Cauchy sequence in X .

2) Deduce that if the metric space (X, d) is complete, then $(X^{\mathbb{N}}, D_a)$ is also complete.

3) Prove that if the metric space (X, d) is compact, then $(X^{\mathbb{N}}, D_a)$ is also compact.

4) Prove that U is an open subset of $(X^{\mathbb{N}}, D_a)$ if and only if for every $x \in U$, there exists a finite subset J of \mathbb{N} and a positive number α such that

$$\forall y \in X^{\mathbb{N}}, \forall j \in J, d(y(j), x(j)) < \alpha \implies y \in U.$$

Chapter 2

Normed spaces, Banach spaces

This chapter is devoted to the study of Banach spaces, which are normed vector spaces that are complete with respect to the metric associated with the norm. These are the basic objects of functional analysis. This chapter provides our first opportunity to use the fundamental concepts of metric topology we saw in the previous chapter on concrete examples.

In the first section, the notion of norm and metric associated with a norm is introduced, and, beyond \mathbb{R}^N , we present and study spaces of bounded continuous functions and spaces of sequences which are summable when elevated to a power p . These are elementary and important examples that the reader should keep in mind as basic models of Banach spaces.

In the second section, we study the space of linear continuous maps between Banach spaces. It is important for the rest of the course to fully assimilate the continuity criterion for linear maps, as well as the notion of norm of a linear map. This part ends with a continuity criterion for multilinear maps.

In the third section, we study the case of vector spaces with finite dimension N , and show that all of these can be identified with \mathbb{R}^N .

The fourth section deals with Ascoli's theorem, which provides us with a compactness criterion for subsets of the space of continuous functions from a compact space to a Banach space. This theorem allows one to witness how far from true the equivalence between being compact and being closed and bounded is in general, unlike in the finite dimension case.

The fifth part contains results on dense sets in the space of continuous functions from a compact set to \mathbb{R} or \mathbb{C} . We take the opportunity to introduce the concept of separability, which will be put to use in the next chapter.

Finally, we point out that, for the rest of the course, \mathbb{K} will denote \mathbb{R} or \mathbb{C} .

2.1 Definitions of normed spaces and Banach spaces

Definition 2.1.1. *Let E be a vector space on \mathbb{K} . A map N from E to \mathbb{R}^+ is called a semi-norm on E if and only if the two following conditions are verified.*

- $N(x + y) \leq N(x) + N(y)$,
- For any λ in \mathbb{K} , we have $N(\lambda x) = |\lambda|N(x)$.

A map N from E to \mathbb{R}^+ is a norm on E if and only if it is a semi-norm with the property: $N(x) = 0$ implies $x = 0$.

Definition 2.1.2. Let E be a vector space and N be a norm on E . The pair (E, N) is called a normed space.

Definition 2.1.3. Let N_1 and N_2 be two norms on a vector space E . These two norms are said to be equivalent if and only if there exists a constant C such that

$$\forall x \in E, C^{-1}N_1(x) \leq N_2(x) \leq CN_1(x).$$

Notation. Most often, a norm is denoted $\|\cdot\|_E$, or $\|\cdot\|$.

Proposition 2.1.1. Let $(E, \|\cdot\|)$ be a normed space. Then, the map defined as follows,

$$\begin{cases} E \times E & \longrightarrow & \mathbb{R}^+ \\ (x, y) & \longmapsto & \|x - y\| \end{cases}$$

is a metric on E ; it is called the metric induced by the norm $\|\cdot\|$.

The easy proof of this is left as an exercise.

Convention. Unless it is expressly mentioned otherwise, we will always consider the space E to be endowed with the metric structure we have just described.

Definition 2.1.4. Let $(E, \|\cdot\|)$ be a normed space. We say that $(E, \|\cdot\|)$ is a Banach space if and only if the metric space (E, d) , in which d is the metric induced by the norm $\|\cdot\|$ (i.e. $d(x, y) = \|x - y\|$), is a complete space.

We will now give a sequence of examples of Banach spaces.

Proposition 2.1.2. For p in the interval $[1, \infty[$, we consider the map defined by the following:

$$\|\cdot\|_{\ell^p} \begin{cases} \mathbb{K}^N & \longrightarrow & \mathbb{R}^+ \\ (x_j)_{1 \leq j \leq N} & \longmapsto & \left(\sum_{j=1}^N |x_j|^p \right)^{\frac{1}{p}}. \end{cases}$$

This is a norm on \mathbb{K}^N , and $(\mathbb{K}^N, \|\cdot\|_{\ell^p})$ is a Banach space. Moreover, for any pair (p, q) in $[1, \infty[$ such that $p \leq q$, we have

$$\|x\|_{\ell^\infty} \stackrel{\text{def}}{=} \sup_{1 \leq j \leq N} |x_j| \leq \|x\|_{\ell^q} \leq \|x\|_{\ell^p} \leq N^{\frac{1}{p} - \frac{1}{q}} \|x\|_{\ell^q} \leq N \sup_{1 \leq j \leq N} \|x\|_{\ell^\infty}. \quad (2.1)$$

Proof. We begin by proving what is known as Hölder's inequality. For any p in the interval $[1, \infty[$, we have, for any (a, b) in $\mathbb{K}^N \times \mathbb{K}^N$,

$$\left| \sum_{j=1}^N a_j b_j \right| \leq \|a\|_{\ell^p} \|b\|_{\ell^{p'}} \quad \text{with} \quad \frac{1}{p} + \frac{1}{p'} = 1. \quad (2.2)$$

To prove this inequality (which is obvious when $p = 1$ or $p = \infty$), we note that the concave nature of the logarithm function implies that

$$\forall (a, b) \in]0, \infty[^2, \forall \theta \in [0, 1], \theta \log a + (1 - \theta) \log b \leq \log(\theta a + (1 - \theta)b).$$

As the exponential function is increasing, we deduce that

$$a^\theta b^{1-\theta} \leq \theta a + (1 - \theta)b,$$

which is often stated, equivalently, as: for any θ in the interval $]0, 1[$,

$$ab \leq \theta a^{\frac{1}{\theta}} + (1 - \theta)b^{\frac{1}{1-\theta}}. \quad (2.3)$$

Dividing the vectors by their norms, we can assume that $\|a\|_{\ell^p} = \|b\|_{\ell^{p'}} = 1$. Then,

$$\begin{aligned} \left| \sum_{j=1}^N a_j b_j \right| &\leq \sum_{j=1}^N |a_j| |b_j| \\ &\leq \frac{1}{p} \sum_{j=1}^N |a_j|^p + \left(1 - \frac{1}{p}\right) \sum_{j=1}^N |b_j|^{p'} \\ &\leq 1. \end{aligned}$$

Now that inequality (2.2) is proved, we can write that

$$\begin{aligned} \sum_{j=1}^N |x_j + y_j|^p &\leq \sum_{j=1}^N |x_j| |x_j + y_j|^{p-1} + \sum_{j=1}^N |y_j| |x_j + y_j|^{p-1} \\ &\leq \left(\left(\sum_{j=1}^N |x_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^N |y_j|^p \right)^{\frac{1}{p}} \right) \left(\sum_{j=1}^N |x_j + y_j|^{(p-1)\frac{p}{p-1}} \right)^{1-\frac{1}{p}} \\ &\leq \left(\left(\sum_{j=1}^N |x_j|^p \right)^{\frac{1}{p}} + \left(\sum_{j=1}^N |y_j|^p \right)^{\frac{1}{p}} \right) \left(\sum_{j=1}^N |x_j + y_j|^p \right)^{1-\frac{1}{p}}. \end{aligned}$$

Simplifying this, we get that, for p in the interval $]1, \infty[$, the map $\|\cdot\|_{\ell^p}$ is a norm (the cases $p = 1$ and $p = \infty$ are again obvious).

Now, we prove (2.1). For any (p, q) in $[1, \infty]^2$ such that $p < q$, we have

$$\frac{x_j}{\|x\|_{\ell^q}} \leq 1.$$

This implies that

$$\left| \frac{x_j}{\|x\|_{\ell^q}} \right|^q \leq \left| \frac{x_j}{\|x\|_{\ell^q}} \right|^p.$$

Summing all the terms, we get

$$1 \leq \left(\frac{\|x\|_{\ell^p}}{\|x\|_{\ell^q}} \right)^p$$

which proves the result. \square

Proposition 2.1.3. *Let X be a set, and $(E, \|\cdot\|)$ be a Banach space. We consider $\mathcal{B}(X, E)$, the set of bounded functions from X to E , i.e. functions f that satisfy*

$$\|f\|_{\mathcal{B}(X, E)} \stackrel{\text{def}}{=} \sup_{x \in X} \|f(x)\| < \infty.$$

Then $(\mathcal{B}(X, E), \|\cdot\|_{\mathcal{B}(X, E)})$ is a Banach space. Moreover, if X is endowed with a metric d , we can define $\mathcal{C}_b(X, E)$, the set of continuous functions in $\mathcal{B}(X, E)$. Then $(\mathcal{C}_b(X, E), \|\cdot\|_{\mathcal{B}(X, E)})$ is a Banach space.

Proof. Let $(f_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in $\mathcal{B}(X, E)$. By definition, we have

$$\forall \varepsilon, \exists n_0 / \forall n \geq n_0, \forall p, \forall x \in X, \|f_n(x) - f_{n+p}(x)\| < \varepsilon. \quad (2.4)$$

In particular, for every x , the sequence $(f_n(x))_{n \in \mathbb{N}}$ is a Cauchy sequence in E . As this space is complete, the sequence converges. So, for every x in X , there exists an element in E , which we denote $f(x)$, such that

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

We must check that f belongs to $\mathcal{B}(X, E)$. By inequality (2.4), in which we take $\varepsilon = 1$, we have

$$\forall p, \forall x \in X, \|f_{n_0}(x) - f_{n_0+p}(x)\| < 1.$$

Taking the limit as p tends to infinity, we get

$$\forall x \in X, \|f_{n_0}(x) - f(x)\| \leq 1.$$

As the function f_{n_0} is bounded, we see that f is also bounded. Indeed, we have

$$\begin{aligned} \|f(x) - f(x')\| &\leq \|f(x) - f_{n_0}(x)\| + \|f_{n_0}(x) - f_{n_0}(x')\| + \|f_{n_0}(x') - f(x')\| \\ &\leq 2 + \|f_{n_0}(x) - f_{n_0}(x')\|. \end{aligned}$$

Now, we must check that the sequence $(f_n)_{n \in \mathbb{N}}$ converges to f in the space $\mathcal{B}(X, E)$. In order to do so, we take the limit as p tends to infinity in inequality (2.4), which yields

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} / \forall n \geq n_0, \forall x \in X, \|f_n(x) - f(x)\| \leq \varepsilon,$$

which ends the proof that $\mathcal{B}(X, E)$ is complete.

To prove that $\mathcal{C}_b(X, E)$ is complete, it suffices, by proposition 1.2.2, to prove that $\mathcal{C}_b(X, E)$ is closed in $\mathcal{B}(X, E)$. In other words, we must prove that a uniform limit of continuous functions is continuous. For this, we consider a sequence $(f_n)_{n \in \mathbb{N}}$ of elements of $\mathcal{C}_b(X, E)$ which converges to f in $\mathcal{B}(X, E)$, and we must show that f is continuous from X to E . By definition, for any $\varepsilon > 0$, there exists an integer n_0 such that

$$\forall x \in X, \|f_{n_0}(x) - f(x)\| < \frac{\varepsilon}{4}. \quad (2.5)$$

Repeatedly using the triangular inequality and (2.5) above, we can write, for any pair (x, x_0) in $X \times X$,

$$\begin{aligned} \|f(x) - f(x_0)\| &\leq \|f(x) - f_{n_0}(x)\| + \|f_{n_0}(x) - f_{n_0}(x_0)\| + \|f_{n_0}(x_0) - f(x_0)\| \\ &\leq \frac{\varepsilon}{2} + \|f_{n_0}(x) - f_{n_0}(x_0)\|. \end{aligned}$$

As the function f_{n_0} is continuous, for any x_0 in X , there exists $\alpha > 0$ such that

$$\forall x \in X, d(x, x_0) < \alpha \implies \|f_{n_0}(x) - f_{n_0}(x_0)\| < \frac{\varepsilon}{2}.$$

We deduce that

$$\forall x \in X, d(x, x_0) < \alpha \implies \|f(x) - f(x_0)\| < \varepsilon,$$

which ends the proof of the proposition. \square

Remark. When $X = \mathbb{N}$ and $E = \mathbb{K}$, which is \mathbb{R} or \mathbb{C} , we denote $\mathcal{B}(\mathbb{N}, \mathbb{K})$ as $\ell^\infty(\mathbb{N}, \mathbb{K})$. We may omit \mathbb{K} in this notation when it does not matter whether the sequences have real or complex values.

Exercise 2.1.1. We denote $\mathcal{C}_u(X, E)$ the set of uniformly continuous functions f in $\mathcal{B}(X, E)$. Prove that $(\mathcal{C}_u(X, E), \|\cdot\|_{\mathcal{B}(X, E)})$ is a Banach space.

Theorem 2.1.1. For p in the interval $[1, \infty[$, we consider $\ell^p(\mathbb{N})$, the set of all sequences x taking values in \mathbb{K} such that

$$\sum_{n \in \mathbb{N}} |x(n)|^p < \infty.$$

We then set

$$\|x\|_{\ell^p} \stackrel{\text{def}}{=} \left(\sum_{n \in \mathbb{N}} |x(n)|^p \right)^{\frac{1}{p}}.$$

The space $(\ell^p(\mathbb{N}), \|\cdot\|_{\ell^p})$ is a Banach space. Moreover, if $p \leq q$, then $\ell^p(\mathbb{N})$ is a subspace of $\ell^q(\mathbb{N})$, and the embedding

$$\begin{cases} \ell^p(\mathbb{N}) & \longrightarrow & \ell^q(\mathbb{N}) \\ x & \longmapsto & x \end{cases}$$

is a continuous linear map.

Proof. First, we prove that $\ell^p(\mathbb{N})$ is a vector space. Let x and y be two elements of $\ell^p(\mathbb{N})$. According to proposition 2.1.2, for any integer N , we have

$$\begin{aligned} \sum_{j=0}^N |x(j) + y(j)|^p &\leq \left(\left(\sum_{j=0}^N |x(j)|^p \right)^{\frac{1}{p}} + \left(\sum_{j=0}^N |y(j)|^p \right)^{\frac{1}{p}} \right)^p \\ &\leq (\|x\|_{\ell^p} + \|y\|_{\ell^p})^p. \end{aligned}$$

As a result, the series with general term $|x(j) + y(j)|^p$ converges, and

$$\sum_{j \in \mathbb{N}} |x(j) + y(j)|^p \leq (\|x\|_{\ell^p} + \|y\|_{\ell^p})^p,$$

which ensures that $(\ell^p(\mathbb{N}), \|\cdot\|_{\ell^p})$ is a normed space. Now we prove that it is a Banach space. Let $(x_q)_{q \in \mathbb{N}}$ be a Cauchy sequence in $\ell^p(\mathbb{N})$. By definition,

$$\forall \varepsilon > 0, \exists q_\varepsilon / \forall q' \geq q_\varepsilon, \forall q'' \geq q_\varepsilon, \|x_{q'} - x_{q''}\|_{\ell^p(\mathbb{N})} < \varepsilon. \quad (2.6)$$

As, for any p , $|x(n)| \leq \|x\|_{\ell^p}$, then, for any n , the sequence $(x_q(n))_{q \in \mathbb{N}}$ is a Cauchy sequence in \mathbb{K} , and thus it converges to an element $x(n)$ in \mathbb{K} . It remains to prove that $x = (x(n))_{n \in \mathbb{N}}$ is in $\ell^p(\mathbb{N})$, and that the sequence $(x_q)_{q \in \mathbb{N}}$ converges to x in $\ell^p(\mathbb{N})$.

Using (2.6) with $\varepsilon = 1$, we get

$$\forall q \geq q_1, \|x_q\|_{\ell^p} \leq 1 + \|x_{q_1}\|_{\ell^p}.$$

By definition of the ℓ^p norm, this implies that

$$\forall q \geq q_1, \forall N \in \mathbb{N}, \sum_{n=0}^N |x_q(n)|^p \leq (1 + \|x_{q_1}\|_{\ell^p})^p.$$

Taking the limit as q tends to infinity, we get

$$\forall N \in \mathbb{N}, \sum_{n=0}^N |x(n)|^p \leq (1 + \|x_{q_1}\|_{\ell^p})^p,$$

which ensures that x is in $\ell^p(\mathbb{N})$. To prove convergence, we note that (2.6) implies that

$$\forall \varepsilon > 0, \exists q_\varepsilon / \forall q \geq q_\varepsilon, \forall q' \geq q_\varepsilon, \forall N \in \mathbb{N}, \sum_{n=0}^N |x_q(n) - x_{q'}(n)|^p < \varepsilon^p.$$

Taking the limit as q' tends to infinity, we deduce that

$$\forall \varepsilon > 0, \exists q_\varepsilon / \forall q \geq q_\varepsilon, \forall q' \geq q_\varepsilon, \forall N \in \mathbb{N}, \sum_{n=0}^N |x_q(n) - x(n)|^p \leq \varepsilon^p.$$

Now, we make N tend to infinity, and we get

$$\forall \varepsilon > 0, \exists q_\varepsilon / \forall q \geq q_\varepsilon, \forall q' \geq q_\varepsilon, \sum_{n \in \mathbb{N}} |x_q(n) - x(n)|^p \leq \varepsilon^p.$$

Convergence is proved.

To conclude the theorem, observe that, by equation (2.1), for any x in $\ell^{p_1}(\mathbb{N})$, any $p_2 \geq p_1$ and any positive integer N ,

$$\sum_{n=0}^N \left(\frac{|x(n)|}{\|x\|_{\ell^{p_1}}} \right)^{p_2} \leq \sum_{n=0}^N \left(\frac{|x(n)|}{\|x\|_{\ell^{p_1}}} \right)^{p_1} \leq 1.$$

We deduce that x belongs to $\ell^{p_2}(\mathbb{N})$, and, by taking the limit in N , that

$$\|x\|_{\ell^{p_2}} \leq \|x\|_{\ell^{p_1}}.$$

Using this on $x - y$, we see that the embedding is 1-Lipschitz continuous. The theorem is proved. \square

Exercise 2.1.2. Prove that the set of sequences that have only a finite number of non-zero terms is dense in $\ell^p(\mathbb{N})$, for any p in the interval $[1, \infty[$.

Exercise 2.1.3. Show that, for any p in $]1, \infty[$, the map $x \mapsto \|x\|_{\ell^p}^p$ is differentiable at any point x in $\ell^p(\mathbb{N})$, and that, for any p in $[2, \infty[$, the map is twice differentiable.

2.2 Spaces of continuous linear maps

As we have endowed a normed space with a metric, we can define continuous functions on them. Being vector spaces, linear maps obviously play an important role. The continuity of such maps can be checked in a very simple way.

Theorem 2.2.1. Let E and F be two normed vector spaces, and ℓ be a linear map from E to F . The following three statements are equivalent:

- the map ℓ is Lipschitz continuous,

- the map ℓ is continuous,
- there exists an open subset U of E such that ℓ is bounded on U , that is

$$\sup_{x \in U} \|\ell(x)\|_F < \infty.$$

Proof. It is clear that the only point we need to prove is that the third condition implies the first one. Let U be an open set on which the linear map ℓ is bounded, and let x_0 be a point in U . The set $U - x_0$ is an open set (exercise: prove it!) containing the origin. As it is open, there exists an open ball with centre 0 and radius α which is included in $U - x_0$. So, we can write that

$$\begin{aligned} M \stackrel{\text{def}}{=} \sup_{x \in B(0, \alpha)} \|\ell(x)\| &\leq \sup_{x \in U - x_0} \|\ell(x)\| \\ &\leq \sup_{y \in U} \|\ell(y - x_0)\| \\ &\leq \sup_{y \in U} \|\ell(y)\| + \|\ell(x_0)\|. \end{aligned}$$

Thus, the map ℓ is bounded on a ball with centre 0 and radius α . For any y in $E \setminus \{0\}$, we have

$$\frac{\alpha}{2\|y\|_E} y \in B(0, \alpha).$$

We deduce that

$$\left\| \ell \left(\frac{\alpha}{2\|y\|_E} y \right) \right\| \leq M.$$

Using the linearity of ℓ and the homogeneity of the norm, we see that

$$\forall y \in E, \|\ell(y)\|_F \leq \frac{2M}{\alpha} \|y\|_E.$$

This inequality also holds for $y = x - x'$, so the theorem is proved. \square

Proposition 2.2.1. *Let E and F be two normed vector spaces; we denote $\mathcal{L}(E, F)$ the set of continuous linear maps from E to F . The mapping defined by*

$$\|L\|_{\mathcal{L}(E, F)} \stackrel{\text{def}}{=} \sup_{\|x\|_E \leq 1} \|L(x)\|_F$$

is a norm on $\mathcal{L}(E, F)$. If $(F, \|\cdot\|_F)$ is a Banach space, then $(\mathcal{L}(E, F), \|\cdot\|_{\mathcal{L}(E, F)})$ is one too.

Proof. It is very easy to see that $\mathcal{L}(E, F)$ is a vector space. Let us now check the three properties that define a norm. Assume that $\|L\|_{\mathcal{L}(E, F)} = 0$. This implies that $L(x)$ is equal to 0 for any x in the unit ball (which is what we usually call the ball with centre 0 and radius 1). Then, for any $x \neq 0$,

$$L \left(\frac{x}{2\|x\|_E} \right) = \frac{1}{2\|x\|_E} L(x) = 0.$$

We deduce that $L(x) = 0$, and therefore $L = 0$. Now, let L_1 and L_2 be two elements of $\mathcal{L}(E, F)$. We have

$$\|L_1(x) + L_2(x)\|_F \leq \|L_1(x)\|_F + \|L_2(x)\|_F.$$

Thus, for any x in E with a norm less than 1, we have

$$\|L_1(x) + L_2(x)\|_F \leq \|L_1\|_{\mathcal{L}(E,F)} + \|L_2\|_{\mathcal{L}(E,F)}.$$

The supremum being the smallest upper bound, we deduce that

$$\|L_1 + L_2\|_{\mathcal{L}(E,F)} \leq \|L_1\|_{\mathcal{L}(E,F)} + \|L_2\|_{\mathcal{L}(E,F)}.$$

Finally, let $L \in \mathcal{L}(E, F)$ and $\lambda \in \mathbb{K}$. We have

$$\|\lambda L(x)\|_F = |\lambda| \|L(x)\|_F,$$

hence,

$$\sup_{\|x\|_E \leq 1} \|\lambda L(x)\|_F = |\lambda| \sup_{\|x\|_E \leq 1} \|L(x)\|_F.$$

Thus, $\|\cdot\|_{\mathcal{L}(E;F)}$ is a norm on $\mathcal{L}(E; F)$.

Let us assume that $(F, \|\cdot\|_F)$ is complete, and let us consider a Cauchy sequence $(L_n)_{n \in \mathbb{N}}$ in $\mathcal{L}(E, F)$. By definition of the norm on $\mathcal{L}(E, F)$, we deduce that, for any x in E , the sequence $(L_n(x))_{n \in \mathbb{N}}$ is a Cauchy sequence in F . As the space F is assumed to be complete, for every x in E , there exists an element in F , which we denote $L(x)$ such that

$$\lim_{n \rightarrow \infty} L_n(x) = L(x).$$

Now it suffices to prove that L is an element of $\mathcal{L}(E, F)$, and that

$$\lim_{n \rightarrow \infty} L_n = L \quad \text{in} \quad \mathcal{L}(E, F).$$

The uniqueness property for limits allows us to easily see that L is a linear map. As the sequence $(L_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, it is bounded. So there exists a constant C such that

$$\sup_{\substack{n \in \mathbb{N} \\ \|x\|_E \leq 1}} \|L_n(x)\|_F \leq C.$$

Taking the limit, we deduce that

$$\sup_{\|x\|_E \leq 1} \|L(x)\|_F \leq C.$$

So the linear map L is continuous. We now prove the convergence of the sequence $(L_n)_{n \in \mathbb{N}}$ to L in $\mathcal{L}(E, F)$. As the sequence $(L_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, for any positive ε , there exists an integer n_0 such that, for any integer $n \geq n_0$, we have

$$\sup_{\substack{\|x\|_E \leq 1 \\ p \in \mathbb{N}}} \|L_n(x) - L_{n+p}(x)\|_F < \varepsilon.$$

Taking the limit as p tends to infinity, we get

$$\sup_{\|x\|_E \leq 1} \|L_n(x) - L(x)\|_F \leq \varepsilon,$$

which ends the proof of the proposition. □

Proposition 2.2.2. Let E , F and G be three normed vector spaces, and (L_1, L_2) be an element of $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$. Then the composition $L_2 \circ L_1$ belongs to $\mathcal{L}(E, G)$, and

$$\|L_2 \circ L_1\|_{\mathcal{L}(E, G)} \leq \|L_1\|_{\mathcal{L}(E, F)} \|L_2\|_{\mathcal{L}(F, G)}.$$

Proof. As the supremum is an upper bound, we have, for any x in the unit ball E ,

$$\begin{aligned} \|L_2 \circ L_1(x)\|_G &\leq \|L_2\|_{\mathcal{L}(F, G)} \|L_1(x)\|_F \\ &\leq \|L_2\|_{\mathcal{L}(F, G)} \|L_1\|_{\mathcal{L}(E, F)}. \end{aligned}$$

Noting that the supremum is the smallest upper bound yields the result. \square

Remark. Even if $F = \mathbb{K}$ (the case of linear functionals), the supremum on the unit ball may not be attained, as the following exercise, which we highly recommend, shows.

Exercise 2.2.1. Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of elements in the interval $]0, 1[$, which we assume converges to 0. We define the linear form ℓ_a on $\ell^1(\mathbb{N})$ by

$$\langle \ell_a, x \rangle \stackrel{\text{def}}{=} \sum_{n \in \mathbb{N}} (1 - a_n)x(n).$$

Prove that $\|\ell_a\|_{(\ell^1(\mathbb{N}))'} = 1$, and that, for every x in the unit ball of $\ell^1(\mathbb{N})$, $|\langle \ell_a, x \rangle|$ is strictly less than 1.

When $E = F$, the space $\mathcal{L}(E, F)$ is denoted $\mathcal{L}(E)$. We are going to study this set, and, in particular, its invertible elements. One of the basic results on invertible elements of $\mathcal{L}(E)$ is the following.

Theorem 2.2.2. Let $(E, \|\cdot\|)$ be a Banach space. The set of elements in $\mathcal{L}(E)$ which have a distance to Id strictly less than 1 are invertible in $\mathcal{L}(E)$. In other words,

$$\forall A \in B_{\mathcal{L}(E)}(\text{Id}, 1), \exists! A^{-1} \in \mathcal{L}(E) / A \circ A^{-1} = A^{-1} \circ A = \text{Id}.$$

Proof. We will use the following property.

Proposition 2.2.3. Let E be a Banach space, and $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of E such that

$$\sum_{n \in \mathbb{N}} \|x_n\|_E < \infty.$$

Then the sequence $S_N \stackrel{\text{def}}{=} \sum_{n=0}^N x_n$ converges.

Proof. The sequence $(S_N)_{N \in \mathbb{N}}$ is a Cauchy sequence, due to

$$\|S_{N+P} - S_N\| \leq \sum_{n=N+1}^{N+P} \|x_n\|.$$

This proves the proposition. \square

Back to the proof of theorem 2.2.2. An element A in $B_{\mathcal{L}(E)}(\text{Id}, 1)$ can be written as $A = \text{Id} - L$, with $\|L\|_{\mathcal{L}(E)}$ strictly less than 1. Set

$$S_N = \sum_{n=0}^N L^n.$$

Proposition 2.2.3 states that the sequence $(S_N)_{N \in \mathbb{N}}$ converges to an element A^{-1} in $\mathcal{L}(E)$ as soon as $\|L\|_{\mathcal{L}(E)}$ is strictly less than 1. Also, we have

$$S_N A = S_N A = \text{Id} - L^{N+1}.$$

By taking the limit in the above equality, we conclude the proof of the theorem. \square

Corollary 2.2.1. *The set $U(E)$ of invertible elements in $\mathcal{L}(E)$ is open, and the mapping Inv , defined by*

$$\begin{cases} U(E) & \longrightarrow & U(E) \\ L & \longmapsto & L^{-1} \end{cases}$$

is continuous.

Proof. If L_0 is a fixed element in $U(E)$, we can write that

$$L = L_0 - (L_0 - L) = (\text{Id} - (L_0 - L)L_0^{-1})L_0.$$

If we assume that

$$\|L - L_0\|_{\mathcal{L}(E)} < \frac{1}{\|L_0^{-1}\|_{\mathcal{L}(E)}},$$

we deduce from theorem 2.2.2 that $\text{Id} - (L_0 - L)L_0^{-1}$ is invertible, and therefore L is too. The set of invertible elements is therefore open, because if L_0 belongs to $U(E)$, then the ball \mathcal{B}_{L_0} with centre L_0 and radius $\|L_0^{-1}\|_{\mathcal{L}(E)}^{-1}$ is included in $U(E)$. Theorem 2.2.2 then implies that

$$L^{-1} = L_0^{-1} \sum_{n=0}^{\infty} ((L_0 - L)L_0^{-1})^n.$$

We deduce that, if L belongs to $B(L_0, \rho_0)$, with ρ_0 strictly less than $\|L_0^{-1}\|_{\mathcal{L}(E)}^{-1}$, we have

$$\begin{aligned} \|L^{-1} - L_0^{-1}\|_{\mathcal{L}(E)} &\leq \|L_0^{-1}\|_{\mathcal{L}(E)} \sum_{n=1}^{\infty} \|L - L_0\|_{\mathcal{L}(E)}^n \|L_0^{-1}\|_{\mathcal{L}(E)}^n \\ &\leq \|L - L_0\|_{\mathcal{L}(E)} \|L_0^{-1}\|_{\mathcal{L}(E)}^2 \sum_{n=0}^{\infty} \|L - L_0\|_{\mathcal{L}(E)}^n \|L_0^{-1}\|_{\mathcal{L}(E)}^n \\ &\leq \|L - L_0\|_{\mathcal{L}(E)} \|L_0^{-1}\|_{\mathcal{L}(E)}^2 \frac{1}{1 - \rho \|L_0^{-1}\|_{\mathcal{L}(E)}}. \end{aligned}$$

This ends the proof of the corollary. \square

Remark. In fact, the proof shows that, when ρ_0 is strictly less than $\|L_0^{-1}\|_{\mathcal{L}(E)}^{-1}$, the map $L \mapsto L^{-1}$ is k -Lipschitz continuous on $B_{\mathcal{L}(E)}(L_0, \rho_0)$, with

$$k = \frac{\|L_0^{-1}\|_{\mathcal{L}(E)}^2}{1 - \rho_0 \|L_0^{-1}\|_{\mathcal{L}(E)}}.$$

Exercise 2.2.2. Prove that the map Inv is in the class $C^1(U(E))$.

To conclude this section, we will characterise continuous multilinear maps. The following theorem is to multilinear maps what theorem 2.2.1 is to linear maps, a condition for continuity.

Theorem 2.2.3. Let $((E_j, \|\cdot\|_j))_{1 \leq j \leq N}$ be a family of normed vector spaces, and $(F, \|\cdot\|)$ be a normed vector space. We consider an N -linear map from $E_1 \times \cdots \times E_N$ to F . This map f is continuous from $E = E_1 \times \cdots \times E_N$ endowed with the norm

$$\|(x_1, \cdots, x_N)\|_E = \sum_{1 \leq j \leq N} \|x_j\|_{E_j}$$

to F if and only if

$$\sup_{\|(x_1, \cdots, x_N)\|_E \leq 1} \|f(x_1, \cdots, x_N)\|_F < \infty. \quad (2.7)$$

Moreover, if f is continuous from $E_1 \times \cdots \times E_N$ to F , then it is Lipschitz continuous on all bounded subsets of $E_1 \times \cdots \times E_N$.

Proof. If f is continuous at the point 0, then there exists a positive real number α such that

$$\sum_{j=1}^N \|x_j\|_{E_j} < \alpha \implies \|f(x_1, \cdots, x_N)\|_F < 1.$$

We then use a homogeneity argument. Let $x = (x_1, \cdots, x_N)$ be a non-zero element of E with a norm in E less than 1. The norm of $\frac{\alpha}{2}x$ is therefore less than α . Thus,

$$\|f(x_1, \cdots, x_N)\|_F < \left(\frac{2}{\alpha}\right)^N.$$

Conversely, assume that inequality (2.7) holds. Set $h = (h_1, \cdots, h_N)$, and we denote $h_0 \stackrel{\text{def}}{=} 0$, and, for each j in $\{1, \cdots, N\}$, $\tilde{h}_j \stackrel{\text{def}}{=} (h_1, \cdots, h_j, 0, \cdots, 0)$. Then, for any fixed x in E , we have

$$f(x+h) - f(x) = \sum_{j=1}^N f(x+h_j) - f(x+h_{j-1}).$$

Inequality (2.7) implies that, for any integer j between 1 and N , we have

$$\|f(x+h_j) - f(x+h_{j-1})\|_F \leq C \left(\prod_{k=1}^{j-1} \|x_k + h_k\|_{E_k} \right) \|h_j\|_{E_j} \left(\prod_{k=j+1}^N \|x_k\|_{E_k} \right).$$

We can assume that $\|h\|_E \leq 1$. Thus, for any integer j between 1 and N , we have

$$\|f(x+h_j) - f(x+h_{j-1})\|_F \leq C_x \|h_j\|_{E_j}.$$

Taking the sum for j , we deduce that

$$\|f(x+h) - f(x)\|_F \leq C_x \|h\|_E,$$

and the theorem is proved. \square

2.3 Banach spaces, compactness and finite dimension

Whether the dimension of a space is finite or not has large implications on topology, as shown by the following statement.

Theorem 2.3.1. *Let E be a normed vector space. If the dimension of E is finite, then the closed unit ball is compact, and all norms on E are equivalent. Conversely, if the closed unit ball of E is compact, then the dimension of the vector space E is finite.*

Proof. Let E be a normed vector space with finite dimension N . We will prove that there is a continuous linear bijection from \mathbb{R}^N endowed with the norm $\|\cdot\|_\infty$ to E , such that the inverse is also continuous. Consider $(\vec{e}_j)_{1 \leq j \leq N}$ a basis of the space E , and the linear bijection I defined by

$$I \begin{cases} \mathbb{R}^N & \longrightarrow E \\ x = (x_j)_{1 \leq j \leq N} & \longmapsto \sum_{j=1}^N x_j \vec{e}_j. \end{cases}$$

The map $x \mapsto I(x)$ is a linear bijection. Let us prove that it is continuous. We have

$$\begin{aligned} \|I(x)\|_E &\leq \sum_{j=1}^N |x_j| \|\vec{e}_j\|_E \\ &\leq M \|x\|_\infty \quad \text{with} \quad M \stackrel{\text{def}}{=} \sum_{j=1}^N \|\vec{e}_j\|_E. \end{aligned}$$

As the map I is linear, we deduce that

$$\|I(x) - I(y)\|_E \leq M \|x - y\|_\infty.$$

The function

$$x \mapsto \|I(x)\|_E$$

is therefore continuous from \mathbb{R}^N to \mathbb{R}^+ . By corollary 1.3.1 on page 23, the sphere \mathbb{S}^{N-1} , that is the set of points x in \mathbb{R}^N such that $\|x\|_\infty = 1$, is compact. Moreover, as I is bijective, it does not vanish on \mathbb{S}^{N-1} . By corollary 1.3.2 on page 24, there exists a positive real number m such that

$$\forall x \in \mathbb{R}^N \setminus \{0\}, \quad m \leq \left\| I\left(\frac{x}{\|x\|_\infty}\right) \right\|_E \leq M.$$

As the map I is linear, we have

$$\forall x \in \mathbb{R}^N, \quad m \|x\|_\infty \leq \|I(x)\|_E \leq M \|x\|_\infty. \quad (2.8)$$

So I and I^{-1} are continuous linear bijections. This implies that all norms on E are equivalent. Therefore the open sets (resp. closed sets, compact sets) of E are precisely the images of open (resp. closed, compact) sets of \mathbb{R}^N through I . Thus, the closed unit ball of E is compact, as it is a closed subset of the image through I of the closed ball of \mathbb{R}^N with centre 0 and radius m^{-1} .

The converse is based on the following geometric lemma.

Lemma 2.3.1. *Let $(E, \|\cdot\|)$ be a normed vector space. We assume that there exists a real number d in the interval $]0, 1[$ and a finite family $(x_j)_{1 \leq j \leq N}$ such that*

$$\mathbb{S}_E \stackrel{\text{def}}{=} \{x \in E, \|x\| = 1\} \subset \bigcup_{j=1}^N B(x_j, \delta).$$

Then the family $(x_j)_{1 \leq j \leq N}$ generates the space E .

If we accept the lemma for a moment, we only need to use the characterisation of compactness provided by theorem 1.3.4 on page 24 to conclude the proof. \square

Proof of lemma 2.3.1 Let F be the subspace of E generated by the family $(x_j)_{1 \leq j \leq N}$. It is a vector space with finite dimension, so, by the first part of our theorem, the space $(F, \|\cdot\|)$ is complete. It is therefore a closed subspace of E .

We now consider a vector y in E , and we will prove that it is a cluster point of F , which is enough to prove the lemma. The key point we need to establish is the following:

$$\forall y \in E, \exists y_1 \in F / \|y - y_1\| \leq \delta \|y\|. \quad (2.9)$$

If $y = 0$, obviously $y_1 = 0$. Let us assume that y is non-zero. This means that there exists j in $\{1, \dots, N\}$ such that

$$\left\| \frac{y}{\|y\|} - x_j \right\| \leq \delta$$

which we can rewrite $\|y - \|y\|x_j\| \leq \delta \|y\|$, and this proves statement (2.9).

We iterate this process. Assumption (2.9) implies that

$$\exists y_2 \in F / \|y - y_1 - y_2\| \leq \delta \|y - y_1\| \leq \delta^2 \|y\|.$$

By induction, we construct a sequence $(y_n)_{n \in \mathbb{N}}$ of elements of F such that

$$\forall n \in \mathbb{N}, \left\| y - \sum_{j=1}^n y_j \right\| \leq \delta^n \|y\|.$$

It is clear that

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n y_j = y$$

which proves that y is in F , because F is closed in E . The lemma is therefore proved. \square

Remark. In a vector space with finite dimension, and **only in these**, the compact subsets are exactly the closed and bounded subsets. As the above theorem states, this is always false in infinite-dimension spaces. To illustrate, let us consider the space E of continuous functions from $[0, 1]$ to \mathbb{R} , endowed with the norm inducing uniform convergence. For every $n \in \mathbb{N}$, the function defined by $f_n(x) \stackrel{\text{def}}{=} x^n$ has norm 1, but the sequence $(f_n)_{n \in \mathbb{N}}$ has no cluster point **in the space E** , because

$$\forall x \in [0, 1[, \lim_{n \rightarrow \infty} x^n = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} 1^n = 1.$$

2.4 Compactness in the space of continuous functions: Ascoli's theorem

In finite dimension, and **only in this setting**, the compact subsets of a vector space are exactly the closed and bounded subsets. Theorem 2.3.1 below states that this is always false in normed vector spaces with infinite dimension. To illustrate this, let us consider, in the space E of continuous functions from $[0, 1]$ to \mathbb{R} , endowed with the norm inducing uniform convergence, the sequence $(f_n)_{n \in \mathbb{N}}$ defined by $f_n(x) \stackrel{\text{def}}{=} x^n$. The norm of each element of this sequence is equal to 1, and yet, the sequence does not have a cluster point **in the space E** because

$$\forall x \in [0, 1[, \lim_{n \rightarrow \infty} x^n = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} 1^n = 1.$$

The aim of this section is to establish a criterion for compactness in the space of continuous functions from a compact metric space to a Banach space.

Theorem 2.4.1 (Ascoli). *Let (X, d) be a compact metric space, and $(E, \|\cdot\|_E)$ be a Banach space. We consider a subset A of $\mathcal{C}(X, E)$ which satisfies the following two properties:*

i) *the subset A is uniformly equicontinuous, i.e.*

$$\forall \varepsilon > 0, \exists \alpha > 0 / \forall (x, x') \in X^2, d(x, x') < \alpha \implies \forall f \in A, \|f(x) - f(x')\|_E < \varepsilon; \quad (2.10)$$

ii) *for any $x \in X$, the closure of the set $\{f(x), f \in A\}$ is compact.*

Then, the closure of A is compact.

Proof. To prove this theorem, we will use the criterion stated in proposition 1.3.6 on page 22. Let ε be a positive number. We will prove that we can cover A with a finite number of open balls with radius ε , which, according to proposition 1.3.6 on p. 22, yields the result.

By assumption (2.10), there exists a positive real number α such that

$$\forall (x, x') \in X^2, d(x, x') < \alpha \implies \forall f \in A, \|f(x) - f(x')\|_E < \frac{\varepsilon}{3}. \quad (2.11)$$

The compactness of X ensures the existence of a finite sequence $(x_j)_{1 \leq j \leq N}$ of elements of X such that

$$X = \bigcup_{j=1}^N B(x_j, \alpha).$$

Set $A_j \stackrel{\text{def}}{=} \{f(x_j), f \in A\}$. By assumption, the closure of A_j , $\overline{A_j}$, is compact in E . By proposition 1.3.4 on page 21, the product $\overline{A_1} \times \cdots \times \overline{A_N}$ is compact in the Banach space E^N , endowed with the norm

$$\|(y_j)_{1 \leq j \leq N}\|_\infty \stackrel{\text{def}}{=} \max_{1 \leq j \leq N} \|y_j\|_E.$$

So the closure of the product $A_1 \times \cdots \times A_N$ is compact in the Banach space E^N . As the subset \mathcal{A}_N of E^N defined by

$$\mathcal{A} \stackrel{\text{def}}{=} \{(f(x_1), \dots, f(x_N)), f \in A\}$$

is included in the product $A_1 \times \cdots \times A_N$, so its closure is also compact. By theorem 1.3.1 on page 20, there exists a finite sequence $(f_k)_{1 \leq k \leq M}$ of elements of A such that the balls (for

the norm $\|\cdot\|_\infty$ we defined above) with centre $(f_k(x_j))_{1 \leq j \leq N}$ and radius $\varepsilon/3$ cover \mathcal{A} , which means that

$$\forall f \in A, \exists k \in \{1, \dots, M\} / \forall j \in \{1, \dots, N\}, \|f(x_j) - f_k(x_j)\|_E < \frac{\varepsilon}{3}.$$

We now prove that the balls in $\mathcal{C}(X, E)$ with centres f_k and radius ε cover A . To do so, let us consider f , an element of A . There exists an index k in $\{1, \dots, M\}$ such that

$$\forall j \in \{1, \dots, N\}, \|f(x_j) - f_k(x_j)\|_E < \frac{\varepsilon}{3}.$$

Let x be a point in X . There exists an index J in $\{1, \dots, N\}$ such that $d(x, x_j) < \alpha$. We can therefore write

$$\begin{aligned} \|f(x) - f_k(x)\|_E &\leq \|f(x) - f(x_j)\|_E + \|f(x_j) - f_k(x_j)\|_E + \|f_k(x_j) - f_k(x)\|_E \\ &< \varepsilon. \end{aligned}$$

This proves the theorem. □

Exercise 2.4.1. Prove the converse of Ascoli's theorem: if A is a compact subset of $\mathcal{C}(X, E)$, then it satisfies conditions *i)* and *ii)* of Ascoli's theorem. The equivalence is known as the Arzelà-Ascoli theorem.

Exercise 2.4.2. Let (X, d) be a compact metric space, and Y be a compact subset of a Banach space $(E, \|\cdot\|_E)$. Prove that the set of Lipschitz continuous functions with a given Lipschitz constant k from X to Y is compact in $\mathcal{C}(X, E)$.

Exercise 2.4.3. Give an example of a sequence of functions $(f_n)_{n \in \mathbb{N}}$ from $[0, 1]$ to \mathbb{R} which is bounded in the space of Lipschitz continuous functions, and which converges to 0 in the space $\mathcal{C}([0, 1], \mathbb{R})$ but not in the space of Lipschitz continuous functions from $[0, 1]$ to \mathbb{R} .

2.5 About the Stone-Weierstrass theorem

The aim of this section is to find dense subspaces of the space of continuous functions on a compact metric space (X, d) . We begin with the case where X is the interval $[0, 1]$.

Theorem 2.5.1 (Bernstein). Let f be a continuous function from the interval $[0, 1]$ to a Banach space E . The sequence of functions $(S_n(f))_{n \in \mathbb{N}}$ defined by

$$S_n(f)(x) \stackrel{\text{def}}{=} \sum_{k=0}^n f\left(\frac{k}{n}\right) C_n^k x^k (1-x)^{n-k},$$

in which the C_n^k are the binomial coefficients, is uniformly convergent to f on the interval $[0, 1]$.

Proof. For any x in the interval $[0, 1]$ and any positive integer n , we have $1 = (x + 1 - x)^n$, so the binomial formula implies that

$$\sum_{k=0}^n C_n^k x^k (1-x)^{n-k} = 1 \tag{2.12}$$

By definition of the sequence $S_n(f)$, we deduce that

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

By Heine's theorem, theorem 1.3.3 on page 23, the function f is uniformly continuous on the compact interval $[0, 1]$, so, if we consider an arbitrary positive number ε , there exists a positive real number α_ε such that

$$|x - y| < \alpha_\varepsilon \implies \|f(x) - f(y)\|_E < \frac{\varepsilon}{2}.$$

We can then write

$$\begin{aligned} \|f(x) - S_n(f)(x)\|_E &\leq \sum_{k=0}^n \left\| f(x) - f\left(\frac{k}{n}\right) \right\|_E x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + 2 \sup_{x \in [0,1]} \|f(x)\|_E \sum_{|k-nx| \geq n\alpha_\varepsilon} C_n^k x^k (1-x)^{n-k}. \end{aligned}$$

By multiplying and dividing by $(k - nx)^2$, we get

$$\begin{aligned} \|f(x) - S_n(f)(x)\|_E &\leq \frac{\varepsilon}{2} + \sup_{x \in [0,1]} \|f(x)\|_E \sum_{|k-nx| \geq n\alpha} \frac{(k - nx)^2}{(k - nx)^2} C_n^k x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + \frac{2}{n^2 \alpha^2} \sup_{x \in [0,1]} \|f(x)\|_E \sum_{|k-nx| \geq n\alpha_\varepsilon} (k - nx)^2 C_n^k x^k (1-x)^{n-k} \\ &\leq \frac{\varepsilon}{2} + \frac{2}{n^2 \alpha^2} \sup_{x \in [0,1]} \|f(x)\|_E \sum_{k=0}^n (k - nx)^2 C_n^k x^k (1-x)^{n-k}. \end{aligned} \quad (2.13)$$

We can compute the sum on the right. By differentiating equality (2.12), then multiplying by x , we get

$$\sum_{k=0}^n k C_n^k x^k (1-x)^{n-k} - nx \sum_{k=0}^{n-1} C_{n-1}^k x^k (1-x)^{n-1-k} = 0,$$

which, by using (2.12) again, yields that

$$\sum_{k=0}^n k C_n^k x^k (1-x)^{n-k} = nx. \quad (2.14)$$

Then, by differentiating and multiplying by x the above equality, we get

$$\sum_{k=0}^n k^2 C_n^k x^k (1-x)^{n-k} - nx \sum_{k=0}^{n-1} k C_{n-1}^k x^k (1-x)^{n-1-k} = nx$$

Using equality (2.14) in this leads to

$$\sum_{k=0}^n k^2 C_n^k x^k (1-x)^{n-k} = nx(1 + (n-1)x)$$

which implies that

$$\begin{aligned} \sum_{k=0}^n (k - nx)^2 C_n^k x^k (1-x)^{n-k} &= \sum_{k=0}^n k^2 C_n^k x^k (1-x)^{n-k} - 2nx \sum_{k=0}^n k C_n^k x^k (1-x)^{n-k} + n^2 x^2 \\ &= nx(1 + (n-1)x) - 2n^2 x^2 + n^2 x^2 \\ &= nx(1 - x). \end{aligned}$$

Equality (2.13) then becomes

$$\|f(x) - S_n(f)(x)\|_E \leq \frac{\varepsilon}{2} + \frac{2}{n\alpha^2} \sup_{x \in [0,1]} \|f(x)\|_E.$$

All that remains to do is to choose n such that $n \geq n_\varepsilon \stackrel{\text{def}}{=} \left[\sup_{x \in [0,1]} \|f(x)\|_E \frac{4}{\varepsilon\alpha^2} \right] + 1$, and the theorem is proved. \square

We are going to state a criterion for density for sub-algebras of $\mathcal{C}(X, \mathbb{K})$, where (X, d) is a compact metric space. First, here is a definition.

Definition 2.5.1. *Let (X, d) be a metric space and A a subset of $\mathcal{C}(X, \mathbb{K})$. We say that the subset A separates the points of X if and only if for any pair (x, y) in X^2 such that $x \neq y$, there exists a function f in A such that $f(x) \neq f(y)$.*

We can now state the following theorem on density.

Theorem 2.5.2 (Stone-Weierstrass). *Let (X, d) be a compact metric space, and A be a sub-algebra of the vector space $\mathcal{C}(X, \mathbb{R})$. If A contains all the constant functions from X to \mathbb{R} , and separates the points of X , then A is a dense subset of $\mathcal{C}(X, \mathbb{R})$.*

The complex-valued case differs slightly.

Theorem 2.5.3. *Let (X, d) be a compact metric space, and A be a sub-algebra of the vector space $\mathcal{C}(X, \mathbb{C})$. If A contains all the constant functions from X to \mathbb{C} , separates the points of X , and is stable by conjugation (i.e. $f \in A \Rightarrow \bar{f} \in A$), then A is a dense subset of $\mathcal{C}(X, \mathbb{C})$.*

We can deduce the following result.

Corollary 2.5.1. *If X is a compact space, then the Banach space $\mathcal{C}(X, \mathbb{K})$ is separable, which means that there exists a dense sequence of points of $\mathcal{C}(X, \mathbb{K})$.*

The proof of the general Stone-Weierstrass theorem is given for the reader's culture.

Proof of theorem 2.5.2.

One of the main steps of the proof is the following lemma, which we will accept without proof for the moment.

Lemma 2.5.1. *Let A be a sub-algebra of $\mathcal{C}(X, \mathbb{R})$. For any function f in \bar{A} , the function $|f|$ belongs to \bar{A} .*

Remark. Since $\max\{f, g\} = \frac{1}{2}(f + g + |f - g|)$ and $\min\{f, g\} = \frac{1}{2}(f + g - |f - g|)$, if f and g are two functions that belong to A , then their minimum and maximum are in \bar{A} .

Back to the proof of theorem 2.5.2. The fact that A separates the points of X implies that

$$(S^+) \quad \forall (x_1, x_2) \in X^2 / x_1 \neq x_2, \forall (\alpha_1, \alpha_2) \in \mathbb{R}^2, \exists h \in A / h(x_j) = \alpha_j.$$

For the same reason, there exists g in A such that $g(x_1) \neq g(x_2)$. We then set

$$h(z) \stackrel{\text{def}}{=} \alpha_1 + (\alpha_2 - \alpha_1) \frac{g(z) - g(x_1)}{g(x_2) - g(x_1)}.$$

Let us consider f a function in $\mathcal{C}(X, \mathbb{R})$, x_0 a point of X and ε a positive real number. We can prove that

$$\exists h_{x_0} \in \bar{A} / h_{x_0}(x_0) = f(x_0) \quad \text{and} \quad \forall z \in X, h_{x_0}(z) \leq f(z) + \varepsilon. \quad (2.15)$$

Indeed, by statement (S^+) , for any y in X , there exists a function h_{x_0} in A such that h coincides with f at the points x_0 and y . Both functions f and $h_{x_0,y}$ are continuous, so there exists a positive number $\alpha_{x_0,y}$ such that

$$d(z, y) < \alpha_{x_0,y} \implies h_{x_0,y}(z) < f(y) + \varepsilon.$$

As the space (X, d) is assumed to be compact, it can be covered with a finite number of balls like the one described above, which means that there exists a finite sequence $(y_j)_{1 \leq j \leq N}$ such that

$$X = \bigcup_{j=1}^N B(y_j, \alpha_{x_0,y_j}).$$

Set $h_{x_0} \stackrel{\text{def}}{=} \min_{1 \leq j \leq N} h_{x_0,y_j}$. By lemma 2.5.1, the function h_{x_0} belongs to \overline{A} and satisfies, for every z in X , $h_{x_0}(z) < f(z) + \varepsilon$. Thus, (2.15) is proved.

Therefore, for any $x \in X$, there exists a function $h_x \in \overline{A}$ such that

$$h_x(x) = f(x) \quad \text{and} \quad \forall y \in X, h_x(y) < f(y) + \varepsilon.$$

The functions h_x and f are continuous, so, for any $y \in X$, there exists a positive number α_x such that

$$\forall z \in B(x, \alpha_x), h_x(z) > f(z) - \varepsilon.$$

Again, X can be covered by a finite number of such balls, so there is a finite family $(x_j)_{1 \leq j \leq N}$ of points of X such that

$$X = \bigcup_{j=1}^N B(x_j, \alpha_{x_j}).$$

Let us set $h \stackrel{\text{def}}{=} \max_{1 \leq j \leq N} h_{x_j}$. By lemma 2.5.1, h belongs to \overline{A} , and, moreover, it satisfies

$$\forall y \in X, f(y) - \varepsilon < h(y) < f(y) + \varepsilon,$$

which means that $\|f - h\|_{\mathcal{C}(X, \mathbb{R})} < \varepsilon$. The theorem is therefore proved, providing we prove lemma 2.5.1. \square

Proof of lemma 2.5.1. Let us first note that we can assume that A is closed, because if A is a sub-algebra of $\mathcal{C}(X, \mathbb{K})$, then so is \overline{A} (exercise: prove it!). The proof then relies on the following lemma, which we accept for the time being.

Lemma 2.5.2. *There exists a sequence of polynomial functions $(P_n)_{n \in \mathbb{N}}$ such that*

$$\lim_{n \rightarrow \infty} P_n = \sqrt{\cdot} \quad \text{in the space } C([0, 1]; \mathbb{R}).$$

Let f be an element of A , and let us set

$$f_n \stackrel{\text{def}}{=} P_n \left(\frac{f^2}{\|f\|} \right) \quad \text{where, of course,} \quad \|f\| = \sup_{x \in X} |f(x)|.$$

The sequence $(f_n)_{n \in \mathbb{N}}$ is a sequence of elements of A because A is an algebra. By lemma 2.5.2, the sequence of functions $(f_n)_{n \in \mathbb{N}}$ converges to $\|f\|^{-1} \sqrt{f^2} = \|f\|^{-1} |f|$. Thus, $|f|$ is in \overline{A} . This means that theorem 2.5.2 is now proved providing we show lemma 2.5.2. \square

Proof of lemma 2.5.2. We define by induction the sequence of polynomial functions

$$P_{n+1}(t) = P_n(t) + \frac{1}{2}(t - P_n^2(t)) \quad \text{and} \quad P_0 = 0.$$

We will show by induction that, for every $t \in [0, 1]$, $0 \leq P_n(t) \leq \sqrt{t}$. This is true for P_0 . Let us assume that it is true for P_n . The polynomial function P_{n+1} is non-negative on $[0, 1]$ as it is the sum of two non-negative functions. Moreover, as P_n is non-negative on $[0, 1]$,

$$\begin{aligned} t - P_{n+1}^2(t) &= t - \left(P_n(t) + \frac{1}{2}(t - P_n(t))^2 \right)^2 \\ &\geq t - P_n^2(t) \\ &\geq 0 \end{aligned}$$

As $P_{n+1}(t) - P_n(t) = 1/2(t - P_n^2(t))$ for any $t \in [0, 1]$, we have $P_{n+1}(t) \geq P_n(t)$. Thus, for any $t \in [0, 1]$, the sequence $(P_n(t))_{n \in \mathbb{N}}$ is non-decreasing and bounded from above by \sqrt{t} , so it converges. Let $\ell(t)$ denote the limit. Taking the limit in the relation between P_n and P_{n+1} , we see that $t - \ell^2(t) = 0$, and the fact that this convergence is in fact uniform is given by the following theorem.

Theorem 2.5.4 (Dini). *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of elements of $\mathcal{C}(X, \mathbb{R})$ such that, for any x in X , the sequence $(f_n(x))_{n \in \mathbb{N}}$ is a non-decreasing sequence with an upper bound. If $g(x) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} f_n(x)$ is a continuous function on X , then the sequence $(f_n)_{n \in \mathbb{N}}$ converges uniformly to g , that is*

$$\lim_{n \rightarrow \infty} \sup_{x \in X} |f_n(x) - g(x)| = 0.$$

Proof. For any $x \in X$, the sequence $(f_n(x))_{n \in \mathbb{N}}$ is non-decreasing and bounded from above, so, for any $\varepsilon > 0$, there exists an integer n_x such that

$$g(x) - \frac{\varepsilon}{2} < f_{n_x}(x).$$

As both functions g and f_{n_x} are continuous on the compact space (X, d) , they are uniformly continuous on (X, d) . Thus, there exists a positive real number α_x such that

$$d(y, y') < \alpha_x \implies |g(y) - g(y')| + |f_{n_x}(y) - f_{n_x}(y')| < \frac{\varepsilon}{2}.$$

The family $(B(x, \alpha_x))_{x \in X}$ covers the compact space (X, d) , so we can extract a finite sub-family that also covers X , which means that there exists a sequence $(x_j)_{1 \leq j \leq N}$ such that

$$X = \bigcup_{j=1}^N B(x_j, \alpha_{x_j}).$$

Set $n_0 \stackrel{\text{def}}{=} \max_{1 \leq j \leq N} n_{x_j}$. The sequence $(f_n(y))_{n \in \mathbb{N}}$ is non-decreasing, so, for any j in $\{1, \dots, N\}$, we have

$$\begin{aligned} f_{n_0}(y) &\geq f_{n_{x_j}}(y) \\ &\geq f_{n_{x_j}}(x_j) - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)| \\ &\geq g(x_j) - \frac{\varepsilon}{2} - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)| \\ &\geq g(y) - \frac{\varepsilon}{2} - |g(y) - g(x_j)| - |f_{n_{x_j}}(y) - f_{n_{x_j}}(x_j)|. \end{aligned}$$

Let j be such that $y \in B(x_j, \alpha_{x_j})$. We get

$$f_{n_0}(y) \geq g(y) - \varepsilon.$$

As the sequence $(f_n(y))_{n \in \mathbb{N}}$ is non-decreasing, we obtain

$$n \geq n_0 \implies \forall y \in X, g(y) \geq f_n(y) > g(y) - \varepsilon.$$

The theorem is proved. □

2.6 Notions on separable metric spaces

The notion of separable spaces is based on countability.

Definition 2.6.1. *Let X be a set. The set X is said to be countable if and only if there exists an one to one map from X to \mathbb{N} .*

Obviously, any set with finite cardinality is countable. Moreover, we have the following proposition.

Proposition 2.6.1. *Let X be a countable set with infinite cardinality (also called countably infinite). Then there exists a bijective map from X to \mathbb{N} .*

Proof. By assumption, there exists an one to one function f from X to \mathbb{N} . Therefore it is a bijective function from X to $Y = f(X)$. It suffices then to show that an infinite part Y of \mathbb{N} is in one-to-one correspondance with \mathbb{N} . We define by induction the sequence

$$b(n) \stackrel{\text{def}}{=} \min Y \setminus \{b(0), \dots, b(n-1)\}$$

The function b is increasing, and therefore one to one. It is bijective because the elements of Y are taken in ascending order. \square

Proposition 2.6.2. *Any finite product of countable sets is countable.*

Proof. It suffices to show the result for the product of two sets, which, in the same way as in the previous proposition, means we must construct an one to one function from $\mathbb{N} \times \mathbb{N}$ to \mathbb{N} . We define

$$\Phi \begin{cases} \mathbb{N} \times \mathbb{N} & \implies \mathbb{N} \\ (p, q) & \mapsto p + \frac{(p+q+1)(p+q+2)}{2}. \end{cases}$$

Let us prove that Φ is one to one. To do so, we start by observing that if (p, q) and (p', q') are two pairs of \mathbb{N}^2 such that $p' + q'$ is strictly greater than $p + q$, then $\Phi(p', q')$ is strictly greater than $\Phi(p, q)$. Indeed, we have $p' + q'$ greater than $p + q + 1$, so

$$\begin{aligned} (p' + q' + 1)(p' + q' + 2) &\geq (p + q + 3)(p + q + 2) \\ &\geq 2(p + q + 2) + (p + q + 1)(p + q + 2). \end{aligned}$$

We deduce that

$$\Phi(p', q') \geq p' + q' + 2 + \Phi(p, q).$$

Thus, if $\Phi(p, q) = \Phi(p', q')$, then $p + q = p' + q'$, which implies immediately that $p = p'$ and $q = q'$, proving that Φ is one to one. \square

Corollary 2.6.1. *The sets \mathbb{Z} and \mathbb{Q} are countable.*

Corollary 2.6.2. *A countable disjoint union of countable sets is countable. To be precise, let A_n be a sequence of countable sets, and we set*

$$\mathcal{A} = \{(n, a) \text{ with } a \in A_n\}.$$

The set \mathcal{A} is countable.

Proof. Each set A_n is countable. So there exists, for each n , an one to one map f_n from A_n to \mathbb{N} . We then define

$$\Phi \begin{cases} \mathcal{A} & \longrightarrow & \mathbb{N}^2 \\ a = (n, a_n) & \longmapsto & (n, \phi_n(a_n)). \end{cases}$$

This map is one to one. Indeed, if $\Phi(a) = \Phi(b)$, then this means that $(n, a_n) = (m, b_m)$, so $n = m$ and $a_n = b_m$. Now, let ϕ be the map from proposition 2.6.1: the map $\phi \circ \Phi$ is one to one from \mathcal{A} to \mathbb{N} as it is the composition of two one to one maps. \square

We have the following negative result.

Theorem 2.6.1. *Let (X, d) be a complete metric space with infinite cardinality, such that the interior of each singleton is empty. Then the set X is not countable.*

Proof. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of X (since all countable set can be represented this way). Baire's theorem, as stated in theorem 1.2.3, yields that the interior of the set

$$\bigcup_{n \in \mathbb{N}} \{x_n\}$$

is empty, so it cannot be equal to X . \square

We can now define the notion of separable metric spaces.

Definition 2.6.2. *Let (X, d) be a metric space. The space (X, d) is said to be separable if and only if there exists a dense sequence of elements of X .*

In other words, this definition means that, in a separable metric space (X, d) , there exists a sequence $(x_n)_{n \in \mathbb{N}}$ of elements of X such that

$$\forall x \in X, \forall \varepsilon > 0, \exists m / d(x, x_m) < \varepsilon.$$

Example. The space \mathbb{R} endowed with the metric $d(x, y) = |x - y|$ is separable, because the sequence of all rationals is dense in \mathbb{R} .

Proposition 2.6.3. *Let $(X_1, d_1), \dots, (X_N, d_N)$ be a family of N separable metric spaces. If we set*

$$X = X_1 \times \dots \times X_N \quad \text{and} \quad d((x_1, \dots, x_N), (y_1, \dots, y_N)) = \max_{1 \leq j \leq N} d_j(x_j, y_j),$$

then the metric space (X, d) is separable.

Proof. In each space X_j , there exists a countable subset D_j which is dense in X_j . Consider $x = (x_1, \dots, x_N)$ a point in X , and set $D = D_1 \times \dots \times D_N$. As it is a finite product of countable sets, D is a countable subset of X , and, as D_j is dense in X_j , for every positive ε , and for every integer j between 1 and N , there exists y_j in D_j such that

$$d_j(x_j, y_j) < \varepsilon.$$

By definition of the metric d , we have $d(x, y) < \varepsilon$. The proposition is proved. \square

Proposition 2.6.4. *Let (X, d) be a separable metric space, and A be a subset of X . Then the metric space (A, d) is separable.*

Proof. To prove this proposition, we denote \mathbb{D} a countable dense subset of X . For any integer n and any x in \mathbb{D} , we choose an element $a_{x,n}$ in $B(x, n^{-1}) \cap A$, if this set is not empty. If it is empty, we select any element of A . We have thus constructed a countable subset of A , $\tilde{\mathbb{D}} \stackrel{\text{def}}{=} (a_{x,n})_{(x,n) \in \mathbb{D} \times \mathbb{N}}$.

Let us prove that $\tilde{\mathbb{D}}$ is dense in A . Let a be an element in A . There exists an element x in \mathbb{D} such that $d(x, a) < n^{-1}$. By definition of $\tilde{\mathbb{D}}$, the element $a_{x,n}$ is in $\tilde{\mathbb{D}} \cap B(x, n^{-1})$, so we have that

$$d(a, a_{x,n}) \leq d(a, x) + d(x, a_{x,n}) < \frac{2}{n},$$

hence the proposition. \square

Theorem 2.6.2. *All compact metric spaces (X, d) are separable.*

Proof. Let \mathcal{B}_n be the set of balls with radius n^{-1} . Of course, this set covers X with open balls, and, given that X is compact, for every integer n , there exists a finite family of points $(x_{j,n})_{1 \leq j \leq J(n)}$ such that

$$X = \bigcup_{j=1}^{J(n)} B(x_{j,n}, n^{-1}).$$

Let D be the set defined by

$$D = \{x_{j,n}, 1 \leq j \leq J(n), n \in \mathbb{N}\}.$$

Corollary 2.6.2 implies that D is countable. Let us prove that it is dense. To do so, consider an arbitrary positive real number ε and any point x in X . We then choose an integer n such that n is greater than the inverse of ε . By definition of the sequence $(x_{j,n})$, there exists j such that

$$d(x, x_{j,n}) < \frac{1}{n}.$$

Given that n is arbitrary, the theorem is proved. \square

We end the chapter with an example of a non-separable metric space.

Proposition 2.6.5. *Let X be the set of functions from $[0, 1]$ to $[0, 1]$, endowed with the metric*

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|.$$

The metric space (X, d) is not separable.

Proof. Consider the set A defined by

$$A = \{\mathbf{1}_P, P \in \mathcal{P}([0, 1])\},$$

where $\mathcal{P}(X)$ is the power set of X . Let us make two easy points. First, if $P \neq Q$ are two subsets of $[0, 1]$, then $d(\mathbf{1}_P, \mathbf{1}_Q) = 1$. Second, the set $\mathcal{P}([0, 1])$, as well as A , is not countable. The proof of the proposition is then finished off by the following lemma.

Lemma 2.6.1. *Let (X, d) be a metric space. If there exists an uncountable subset A of X , and a positive number α such that*

$$\forall (x_1, x_2) \in A^2, x_1 \neq x_2 \Rightarrow d(x_1, x_2) \geq \alpha,$$

then the metric space (X, d) is not separable.

Proof. Let D be a dense subset of X . for any a in A , there exists z in D such that

$$d(a, z) \leq \frac{\alpha}{3}.$$

Let a_1 and a_2 be two different elements of A . They correspond to two elements z_1 and z_2 in D which satisfy the above inequality. As a result, we have

$$\begin{aligned} \alpha &\leq d(a_1, a_2) \\ &\leq d(a_1, z_1) + d(a_2, z_2) + d(z_1, z_2) \\ &\leq \frac{\alpha}{3} + \frac{\alpha}{3} + d(z_1, z_2) \\ &\leq \frac{2\alpha}{3} + d(z_1, z_2). \end{aligned}$$

Hence z_1 cannot be equal to z_2 . So a dense subset of X cannot be countable, which proves the lemma. \square

Chapter 3

Duality in Banach spaces

Introduction

This short chapter is very important. It deals with a particular type of Banach space: the dual space (or topological dual space) of a Banach space, which is the set of continuous linear functionals on that space. First, we present the general concept of transposing a linear map. We point out that, in infinite dimension, this is far from simply transposing a matrix. A striking new fact we will discover in this chapter is that the dual of a Banach space may be very different from that space. This is of course not the case in finite dimension, where the dimension of the dual space is equal to that of the original vector space.

In the second section, we will give a precise mathematical meaning to the sentence “the dual space of E is F ”. This definition is fundamental. It allows one to construct isomorphisms between the dual E' of a Banach space and another Banach space F , which gives us a description of E' . The duals of the spaces of sequences which are summable when elevated to a power p will be studied in detail as an essential example.

The third section introduces a crucial concept, which is that of a weaker notion of convergence for a sequence of elements in the dual of a Banach space. It is the concept of weak- \star convergence, which is essentially pointwise convergence. This allows one to extract subsequences from any bounded sequence in the dual space which converge in this sense. Hence the importance of the identification theorems proved beforehand: if a space can be identified as the dual space of another, then it will inherit this property. We will see an application of this result in the next chapter, when we will prove theorem 4.4.2 on page 75, in which we study the structure of self-adjoint compact operators.

3.1 A presentation of the concept of duality

The reader should already be well informed on the notion of dual vector spaces. Nonetheless, we briefly recall some fundamental points. If E is any vector space, we denote E' the space of linear forms defined on E . Moreover, if ℓ is a linear form on E and x is a vector in E , we denote

$$\langle \ell, x \rangle \stackrel{\text{def}}{=} \ell(x).$$

The first key notion is that of the transpose of a linear map. Consider a linear map L from E to E . Its transpose is the map from E' (the set of linear functionals on E) to E' defined by

$$\langle {}^tL(\ell), \vec{h} \rangle = \langle \ell, L\vec{h} \rangle.$$

Let us note that if L_1 and L_2 are two linear maps from E to E , then

$${}^t(L_1 \circ L_2) = {}^tL_2 \circ {}^tL_1.$$

Indeed, by definition of the transpose of a map, we have

$$\begin{aligned} \langle {}^t(L_1 \circ L_2)(\ell), \vec{h} \rangle &= \langle \ell, L_1 \circ L_2(\vec{h}) \rangle \\ &= \langle \ell, L_1(L_2(\vec{h})) \rangle \\ &= \langle {}^tL_1(\ell), L_2(\vec{h}) \rangle \\ &= \langle {}^tL_2({}^tL_1(\ell)), \vec{h} \rangle. \end{aligned}$$

Hence the result.

In the case where the vector space E has finite dimension, we can choose $(\vec{e}_j)_{1 \leq j \leq N}$, a basis of E . We can easily prove that the family of linear forms $(e_j^*)_{1 \leq j \leq N}$, defined by

$$\langle e_j^*, \vec{e}_k \rangle = \delta_{j,k}$$

forms a basis of the vector space of linear forms E' , and is called a dual basis. Moreover, it is easy to see that if the linear map L is represented by the matrix $(L_j^i)_{1 \leq i, j \leq N}$ in a basis $(\vec{e}_j)_{1 \leq j \leq N}$ of E , then its transpose tL is represented by the matrix $(\tilde{L}_j^i)_{1 \leq i, j \leq N}$, with $\tilde{L}_j^i = L_j^i$, in the dual basis $(e_j^*)_{1 \leq j \leq N}$. Thus we see that, if E has finite dimension, then the set of continuous linear forms (note that linear forms in this case are always continuous) is a vector space with the same dimension.

In this chapter, we will study the case where E is a general Banach space.

Definition 3.1.1. We call the topological dual space (dual for short) of a normed \mathbb{K} vector space $(E, \|\cdot\|)$, denoted E' , the vector space of continuous linear functionals on E . It is endowed with the norm

$$\|\ell\|_{E'} \stackrel{\text{def}}{=} \sup_{\|x\| \leq 1} |\langle \ell, x \rangle|.$$

Compared with the finite dimension case, a radically new phenomenon appears. As we will see after the proof of theorem 3.2.1, the Banach space $\ell^1(\mathbb{N})$ of absolutely convergent series is separable, while its dual is not. At this stage, this sounds like it will create some difficulties, but transposing in this setting will in fact prove to be extremely useful. This idea is at the heart of the notion of quasi-derivatives that we will introduce in chapter 6 and, more generally, in the theory of distributions, which we will study in chapter 8.

There exists a natural bilinear form defined on the space $E' \times E$.

Proposition 3.1.1. Let E be a normed vector space. Then the mapping $\langle \cdot, \cdot \rangle$ defined by

$$\langle \cdot, \cdot \rangle \begin{cases} E' \times E & \longrightarrow \mathbb{K} \\ (\ell, x) & \longmapsto \langle \ell, x \rangle. \end{cases}$$

is a continuous bilinear form.

Proof. By definition of the norm on E' , we have $|\langle \ell, x \rangle| \leq \|\ell\|_{E'} \|x\|$. Theorem 2.2.3 on page 37, characterising continuous multilinear maps, then ends the proof. \square

We are now going to present the transpose of a linear map. This notion will play a central role in chapter 8.

Proposition 3.1.2. *Let E and F be two normed vector spaces, and let A be an element of the space $\mathcal{L}(E, F)$. The map tA , defined by*

$$\begin{cases} F' & \longrightarrow E' \\ \ell & \longmapsto {}^tA(\ell) : x \mapsto \langle \ell, Ax \rangle \end{cases}$$

is a continuous linear map.

Proof. We can write

$$\begin{aligned} |\langle \ell, Ax \rangle| &\leq \|\ell\|_{F'} \|Ax\|_E \\ &\leq \|\ell\|_{F'} \|A\|_{\mathcal{L}(E, F)} \|x\|_E. \end{aligned}$$

Hence, we have

$$\begin{aligned} \|{}^tA(\ell)\|_{E'} &= \sup_{\|x\|_E \leq 1} |\langle \ell, Ax \rangle| \\ &\leq \|\ell\|_{F'} \|A\|_{\mathcal{L}(E, F)}. \end{aligned}$$

The proposition is proved. □

Definition 3.1.2. *The linear map defined above is called the transpose of the linear map A .*

Proposition 3.1.3. *Let E , F and G be three normed vector spaces, and let (A, B) be an element of $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$. Then,*

$${}^t(B \circ A) = {}^tA \circ {}^tB$$

Moreover, if $E = F$, then

$$A \in U(E) \iff {}^tA \in U(E') \quad \text{and} \quad ({}^tA)^{-1} = {}^t(A^{-1}).$$

Proof. The first equality is purely algebraic. Indeed, for any linear form ℓ on G , and for any x in E , we have

$$\begin{aligned} \langle {}^tA \circ {}^tB(\ell), x \rangle &= \langle {}^tB(\ell), Ax \rangle \\ &= \langle \ell, B \circ Ax \rangle \\ &= \langle {}^t(B \circ A)(\ell), x \rangle. \end{aligned}$$

Then, if A is in $U(E)$, we have $A \circ A^{-1} = A^{-1} \circ A = \text{Id}_E$. By transposing this equality, we get

$${}^tA \circ {}^t(A^{-1}) = {}^t(A^{-1}) \circ {}^tA = \text{Id}_{E'}.$$

Hence, tA is an invertible element of $\mathcal{L}(E')$, and $({}^tA)^{-1} = {}^t(A^{-1})$. The proposition is proved. □

3.2 Identifying a normed space as a dual space

This section is very important for the remainder of the course, in particular the chapter on distribution theory. It is common to say that a given space “is the dual” of another. We need to give a precise meaning to this sentence.

Proposition 3.2.1. *Let E and F be two Banach spaces, and let B be a bilinear continuous functional on $F \times E$. The mapping δ_B defined by*

$$\delta_B \begin{cases} F & \longrightarrow & E' \\ y & \longmapsto & \delta_B(y) : \langle \delta_B(y), x \rangle = B(y, x) \end{cases} \quad (3.1)$$

is in $\mathcal{L}(F, E')$.

Proof. As the bilinear form B is continuous, we can write that

$$|\langle \delta_B(y), x \rangle| \leq C \|y\|_F \|x\|_E.$$

By definition of the norm on E' , we have

$$\|\delta_B(y)\|_{E'} \leq C \|y\|_F ;$$

hence the proposition. □

Definition 3.2.1. *Let E and F be two Banach spaces, and let B be a bilinear continuous functional on $F \times E$. We say that B identifies E' with F if and only if the map δ_B defined by (3.1) is an isomorphism from F to E' .*

In other words, this means that there exists a constant C such that, for any continuous linear functional ℓ on E , there exists a unique element y in F such that

$$\forall x \in E, \langle \ell, x \rangle = B(x, y) \quad \text{and} \quad C^{-1} \|y\|_F \leq \|\ell\|_{E'} \leq C \|y\|_F.$$

Theorem 3.2.1. *Let p be a number in the interval $[1, +\infty[$, and set $E = \ell^p(\mathbb{N})$ and $F = \ell^{p'}(\mathbb{N})$, where $p' = p/(p-1)$. We consider the following bilinear map B , defined by*

$$B \begin{cases} \ell^{p'}(\mathbb{N}) \times \ell^p(\mathbb{N}) & \longrightarrow & \mathbb{K} \\ (y, x) & \longmapsto & B(y, x) \stackrel{\text{def}}{=} \sum_{n \in \mathbb{N}} x(n)y(n). \end{cases}$$

The bilinear form B identifies $(\ell^p(\mathbb{N}))'$ with $\ell^{p'}(\mathbb{N})$. Moreover, the map δ_B induced by proposition 3.2.1 is an isometry.

Proof. Hölder’s inequality states that

$$|B(y, x)| \leq \|x\|_{\ell^p(\mathbb{N})} \|y\|_{\ell^{p'}(\mathbb{N})}. \quad (3.2)$$

So the bilinear form B is continuous. Let us prove that the map δ_B is a bijection. Consider T a linear functional on $\ell^p(\mathbb{N})$, and let us look for an element y in $\ell^{p'}(\mathbb{N})$ such that

$$\forall x \in \ell^p(\mathbb{N}), \langle T, x \rangle = B(y, x). \quad (3.3)$$

If the above is satisfied by y , then it must be satisfied for $x = e_n$, where e_n denotes the sequence with all terms equal to zero, except the one with index n , which is equal to 1. Therefore, y must verify

$$y(n) = \langle T, e_n \rangle. \quad (3.4)$$

This proves that δ_B is one to one. Moreover, by linearity, if V denotes the set of finite linear combinations of elements of the sequence $(e_n)_{n \in \mathbb{N}}$, then, for any x in V ,

$$B(y, x) = \langle T, x \rangle. \quad (3.5)$$

We must now prove that the sequence y defined by (3.4) is an element of $\ell^{p'}(\mathbb{N})$. Let us first assume that p is equal to 1. For any integer n , we have

$$\begin{aligned} |y(n)| &\leq \|T\|_{(\ell^1(\mathbb{N}))'} \|e_n\|_{\ell^1(\mathbb{N})} \\ &\leq \|T\|_{(\ell^1(\mathbb{N}))'}. \end{aligned}$$

So the sequence $y = (y(n))_{n \in \mathbb{N}}$ is in $\ell^\infty(\mathbb{N})$, and we effectively have $\|y\|_{\ell^\infty(\mathbb{N})} \leq \|T\|_{(\ell^1(\mathbb{N}))'}$. Both terms in equality (3.5) are continuous and coincide on a dense subspace of $\ell^1(\mathbb{N})$. This ends the proof of the theorem for $p = 1$.

Now, let us assume that p is strictly greater than 1. We consider the sequence $(x_r)_{r \in \mathbb{N}}$ of elements of $\ell^p(\mathbb{N})$ defined by

$$x_r \stackrel{\text{def}}{=} \sum_{n \leq r} \frac{\bar{y}(n)}{|y(n)|} |y(n)|^{\frac{1}{p-1}} e_n.$$

This implies that

$$x_r(n) = \frac{\bar{y}(n)}{|y(n)|} |y(n)|^{\frac{1}{p-1}} \quad \text{if } n \leq r, \quad \text{and } 0 \quad \text{otherwise.}$$

We therefore have

$$\sum_{n \in \mathbb{N}} |x_r(n)|^p = \sum_{n \leq r} |y(n)|^{p'}.$$

$$\begin{aligned} B(y, x_r) &= \sum_{n=0}^r \frac{\bar{y}(n)}{|y(n)|} |y(n)|^{\frac{1}{p-1}} B(e_n, y) \\ &= \sum_{n=0}^r |y(n)|^{\frac{1}{p-1} + 1} \\ &= \sum_{n=0}^r |y(n)|^{p'}. \end{aligned}$$

Moreover, we know that

$$B(y, x_r) = \sum_{n=0}^r \frac{\bar{y}(n)}{|y(n)|} |y(n)|^{\frac{1}{p-1}} T(e_n) = \langle T, x_r \rangle.$$

Since the linear form T is continuous, we have

$$\begin{aligned} \sum_{n \leq r} |y(n)|^{p'} &\leq \|T\|_{(\ell^p(\mathbb{N}))'} \|x_r\|_{\ell^p(\mathbb{N})} \\ &\leq \|T\|_{(\ell^p(\mathbb{N}))'} \left(\sum_{n \leq r} |y(n)|^{p'} \right)^{\frac{1}{p}}. \end{aligned}$$

Hence, we get

$$\left(\sum_{n \leq r} |y(n)|^{p'} \right)^{1 - \frac{1}{p'}} \leq \|T\|_{(\ell^p(\mathbb{N}))'}.$$

As this inequality is true for any integer r , the sequence y is indeed in $\ell^{p'}(\mathbb{N})$, and we have

$$\|y\|_{\ell^{p'}(\mathbb{N})} \leq \|T\|_{(\ell^p(\mathbb{N}))'}. \quad (3.6)$$

The subspace V is dense in $\ell^p(\mathbb{N})$, so we observe again that equality (3.5) states that two continuous linear functionals coincide on a dense subspace. Hence, they are equal on the entire space, and this ends the proof of the theorem. \square

Exercise 3.2.1. Let $c_0(\mathbb{N})$ denote the Banach space defined by

$$c_0(\mathbb{N}) \stackrel{\text{def}}{=} \{x = (x(n))_{n \in \mathbb{N}} / \lim_{n \rightarrow \infty} x(n) = 0\} \quad \text{endowed with the norm} \quad \|x\|_{\ell^\infty(\mathbb{N})} \stackrel{\text{def}}{=} \sup_{n \in \mathbb{N}} |x(n)|.$$

Prove that $c_0(\mathbb{N})$ is a closed subspace of $\ell^\infty(\mathbb{N})$. Then, by following the ideas of the proof we have just completed, prove that the bilinear form B identifies $c_0(\mathbb{N})'$ with $\ell^1(\mathbb{N})$.

Remark. At the start of this chapter, we noted that, in the case of a vector space E with finite dimension, the dual space is a vector space with the same dimension, and therefore a space that is isomorphic to E . No such luck in infinite dimension: the previous theorem showed that the dual space of $\ell^1(\mathbb{N})$ is isometric to $\ell^\infty(\mathbb{N})$, and thus the space $(\ell^1(\mathbb{N}))'$ has the same topological properties as $\ell^\infty(\mathbb{N})$. In particular, it is not separable. However, we can show that $\ell^1(\mathbb{N})$ is separable. As an exercise, prove that the space of sequences which have all terms, apart from a finite number of them, equal to zero, is dense in $\ell^1(\mathbb{N})$ (and, in fact, dense in $\ell^p(\mathbb{N})$ for any p in $[1, +\infty[$).

Exercise 3.2.2. Let s be a real number. We define the space $\ell^{2,s}(\mathbb{N})$ by

$$\ell^{2,s}(\mathbb{N}) \stackrel{\text{def}}{=} \{u = (u(n))_{n \in \mathbb{N}} / ((1 + n^2)^{\frac{s}{2}} u(n))_{n \in \mathbb{N}} \in \ell^2(\mathbb{N})\},$$

endowed with the norm

$$\|u\|_{\ell^{2,s}(\mathbb{N})} \stackrel{\text{def}}{=} \left(\sum_{n \in \mathbb{N}} (1 + n^2)^s |u(n)|^2 \right)^{\frac{1}{2}}.$$

Let B_1 and B_2 be two bilinear functionals defined as follows:

$$B_1 \begin{cases} \ell^{2,s}(\mathbb{N}) \times \ell^{2,s}(\mathbb{N}) & \longrightarrow \mathbb{K} \\ (y, x) & \longmapsto B_1(y, x) \stackrel{\text{def}}{=} \sum_{n \in \mathbb{N}} (1 + n^2)^s x(n)y(n). \end{cases}$$

$$B_2 \begin{cases} \ell^{2,-s}(\mathbb{N}) \times \ell^{2,s}(\mathbb{N}) & \longrightarrow \mathbb{K} \\ (y, x) & \longmapsto B_2(y, x) \stackrel{\text{def}}{=} \sum_{n \in \mathbb{N}} x(n)y(n). \end{cases}$$

1) Prove that the bilinear form B_1 identifies $(\ell^{2,s}(\mathbb{N}))'$ with $\ell^{2,s}(\mathbb{N})$, and that B_2 identifies $(\ell^{2,s}(\mathbb{N}))'$ with $\ell^{2,-s}(\mathbb{N})$.

2) Propose a onto linear isometry from $\ell^{2,s}(\mathbb{N})$ to $\ell^{2,-s}(\mathbb{N})$.

3.3 A weaker sense for convergence in E'

Definition 3.3.1. Let E be a normed vector space. We consider a sequence $(\ell_n)_{n \in \mathbb{N}}$ of elements of E' , and ℓ an element of E' . We say that the sequence $(\ell_n)_{n \in \mathbb{N}}$ is weak- \star convergent to ℓ , denoted $\lim_{n \rightarrow \infty} \ell_n = \ell$, if and only if

$$\forall x \in E, \quad \lim_{n \rightarrow \infty} \langle \ell_n, x \rangle = \langle \ell, x \rangle.$$

First, we must note that weak- \star convergence is implied by convergence in norm (the usual sense). Indeed, for any x in E , we have

$$\begin{aligned} |\langle \ell_n, x \rangle - \langle \ell, x \rangle| &= |\langle \ell_n - \ell, x \rangle| \\ &\leq \|\ell_n - \ell\|_{E'} \|x\|_E. \end{aligned}$$

Let us now show a sequence $(\ell_n)_{n \in \mathbb{N}}$ of E' which is weak- \star convergent, but not convergent in norm. Consider the space $\ell^1(\mathbb{N})$ and the sequence $(\ell_n)_{n \in \mathbb{N}}$ of linear functionals defined by

$$\langle \ell_n, x \rangle \stackrel{\text{def}}{=} x(n).$$

For any x in $\ell^1(\mathbb{N})$, $\lim_{n \rightarrow \infty} x(n) = 0$. Hence, the sequence $(\ell_n)_{n \in \mathbb{N}}$ is weak- \star convergent to 0. However, $\|\ell_n\|_{(\ell^1(\mathbb{N}))'} = 1$, so the sequence does not converge to 0 in norm.

We will now show some properties of weak- \star convergence.

Theorem 3.3.1. Let E be a Banach space, and $(\ell_n)_{n \in \mathbb{N}}$ be a sequence of elements of E' . Assume that $\lim_{n \rightarrow \infty} \ell_n = \ell$. Then the sequence $(\ell_n)_{n \in \mathbb{N}}$ is bounded in E' , which implies that ℓ is an element of E' . Moreover, we have

$$\|\ell\|_{\mathcal{L}(E, \mathbb{K})} \leq \liminf_{n \rightarrow \infty} \|\ell_n\|_{\mathcal{L}(E, \mathbb{K})}.$$

Proof. This is merely a statement of the classic Banach-Steinhaus theorem in the setting of linear functionals. We will prove the general theorem, which is the following.

Theorem 3.3.2 (Banach-Steinhaus). Let E and F be two normed spaces, and let $(L_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{L}(E, F)$. Assume that E is complete, and that, for any x in E , the sequence $(L_n(x))_{n \in \mathbb{N}}$ has a limit, denoted $L(x)$.

Then the sequence $(L_n)_{n \in \mathbb{N}}$ is bounded in $\mathcal{L}(E, F)$, so the map L is also in $\mathcal{L}(E, F)$. Also, we have

$$\|L\|_{\mathcal{L}(E, F)} \leq \liminf_{n \rightarrow \infty} \|L_n\|_{\mathcal{L}(E, F)}.$$

Proof. The key ingredient is the following lemma.

Lemma 3.3.1. Let E and F be two normed spaces, and let $(L_n)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{L}(E, F)$. Assume that E is complete, and that, for any x in E , the sequence $(L_n(x))_{n \in \mathbb{N}}$ is bounded in F .

Then, the sequence $(L_n)_{n \in \mathbb{N}}$ is bounded in $\mathcal{L}(E, F)$.

Proof. Let us consider the sets $\mathcal{F}_{n,p}$ defined by

$$\mathcal{F}_{n,p} = \{x \in E / \|L_n(x)\|_F \leq p\}.$$

These sets are closed, as they are preimages of closed sets by a continuous map. So, the sets \mathcal{F}_p defined by

$$\mathcal{F}_p \stackrel{\text{def}}{=} \bigcap_{n \in \mathbb{N}} \mathcal{F}_{n,p}$$

are also closed, because they are intersections of closed sets. Moreover, let x be an element of E . As the sequence $(L_n(x))_{n \in \mathbb{N}}$ is assumed to be bounded, there exists an integer p such that x belongs to \mathcal{F}_p . This means that the union of the closed sets \mathcal{F}_p is equal to the entire space E . By corollary 1.2.1 on page 17, deduced from Baire's theorem (theorem 1.2.2), this implies that there exists an integer p_0 such that the set $\overset{\circ}{\mathcal{F}}_{p_0} \neq \emptyset$. Thus, there exists a point x_0 and a positive real number α such that $B(x_0, \alpha)$ is included in \mathcal{F}_{p_0} . Hence,

$$\begin{aligned} \sup_{\substack{n \in \mathbb{N} \\ \|x\| \leq \alpha}} \|L_n(x)\|_F &\leq \sup_{\substack{n \in \mathbb{N} \\ \|x\| \leq \alpha}} \|L_n(x + x_0)\|_F + \|L_n(x_0)\|_F \\ &\leq \sup_{\substack{n \in \mathbb{N} \\ x \in B(x_0, \alpha)}} \|L_n(x)\|_F + \|L_n(x_0)\|_F \\ &\leq 2p_0. \end{aligned}$$

Thus, for any x in E with a norm smaller than or equal to 1, we can write

$$\begin{aligned} \|L_n(x)\|_F &\leq \alpha^{-1} \|L_n(\alpha x)\|_F \\ &\leq \frac{2p_0}{\alpha}. \end{aligned}$$

Therefore, the sequence $(L_n)_{n \in \mathbb{N}}$ is bounded in $\mathcal{L}(E, F)$. □

Back to the proof of theorem 3.3.2. Taking the limit, we get that

$$\forall x \in B(x_0, \alpha), \|L(x)\|_F \leq 2p_0.$$

We must now prove the bound on the norm of L . Let x be an element of E with a norm equal to 1. We can write that

$$\begin{aligned} \|L(x)\|_F &= \lim_{n \rightarrow \infty} \|L_n(x)\|_F \\ &= \liminf_{n \rightarrow \infty} \|L_n(x)\|_F \\ &\leq \liminf_{n \rightarrow \infty} \|L_n\|_{\mathcal{L}(E, F)} \quad (\text{because } \|x\|_E = 1). \end{aligned}$$

All statements in the theorem are now proved. □

Theorem 3.3.1 is obtained by applying this result to the case $F = \mathbb{K}$. □

Although the following theorem has a very simple proof, it will be very useful.

Theorem 3.3.3. *Let E and F be normed vector spaces. We consider an element A in $\mathcal{L}(E, F)$, a sequence $(\ell_n)_{n \in \mathbb{N}}$ in F' , and an element ℓ in F' . We then have*

$$\lim_{n \rightarrow \infty} \star \ell_n = \ell \quad \text{in } F' \implies \lim_{n \rightarrow \infty} \star {}^t A \ell_n = {}^t A \ell \quad \text{in } E'.$$

Proof. Indeed, for any x in E , we have

$$\langle {}^t A \ell_n, x \rangle = \langle \ell_n, Ax \rangle.$$

We assume that

$$\lim_{n \rightarrow \infty} \langle \ell_n, Ax \rangle = \langle \ell, Ax \rangle.$$

However, the definition of the transpose of a linear map yields $\langle \ell, Ax \rangle = \langle {}^t A \ell, x \rangle$. So

$$\lim_{n \rightarrow \infty} \langle {}^t A \ell_n, x \rangle = \langle {}^t A \ell, x \rangle.$$

The theorem is proved. \square

Proposition 3.3.1. *Let $(x_n)_{n \in \mathbb{N}}$ and $(\ell_n)_{n \in \mathbb{N}}$ be two sequences in E and E' respectively, where E is a Banach space. Then,*

$$\left(\liminf_{n \rightarrow \infty} \star \ell_n = \ell \quad \text{and} \quad \lim_{n \rightarrow \infty} \|x_n - x\| = 0 \right) \implies \lim_{n \rightarrow \infty} \langle \ell_n, x_n \rangle = \langle \ell, x \rangle.$$

Proof. We begin by writing

$$\langle \ell_n, x_n \rangle = \langle \ell_n, x_n - x \rangle + \langle \ell_n - \ell, x \rangle + \langle \ell, x \rangle.$$

By the Banach-Steinhaus theorem, theorem 3.3.1, the sequence $(\ell_n)_{n \in \mathbb{N}}$ is bounded, so there exists a positive number M such that

$$|\langle \ell_n, x_n \rangle - \langle \ell, x \rangle| \leq M \|x_n - x\| + |\langle \ell_n - \ell, x \rangle|.$$

Let ε be a positive real number. There exists an integer n_0 such that

$$n \geq n_0 \implies \|x_n - x\| < \frac{\varepsilon}{2M + 1}.$$

The point x is fixed, so there exists an integer n_1 such that

$$n \geq n_1 \implies |\langle \ell_n - \ell, x \rangle| < \frac{\varepsilon}{2}.$$

So, we have

$$n \geq \max\{n_0, n_1\} \implies |\langle \ell_n, x_n \rangle - \langle \ell, x \rangle| < \varepsilon.$$

This proves the proposition. \square

Once we have identified a dual space with another, we can define a notion of weak- \star convergence on the latter as follows.

Definition 3.3.2. *Let E and F be two Banach spaces, and let B be a bilinear form on $F \times E$ which identifies E' with F . By proxy, we say that a sequence $(y_n)_{n \in \mathbb{N}}$ is weak- \star convergent to an element y in F if and only if the sequence $(\delta_B(y_n))_{n \in \mathbb{N}}$ is weak- \star convergent to $\delta_B(y)$, which is stated mathematically like so:*

$$\forall x \in E, \quad \lim_{n \rightarrow \infty} B(y_n, x) = B(y, x).$$

Let us illustrate the extension of this notion. For a given q in $]1, +\infty[$, we consider a sequence $(y_p)_{p \in \mathbb{N}}$ of elements of $\ell^q(\mathbb{N})$. By theorem 3.2.1, we identify the dual of $\ell^q(\mathbb{N})$ with $\ell^{q'}(\mathbb{N})$, and “the sequence $(y_p)_{p \in \mathbb{N}}$ is weak- \star convergent to an element y in $\ell^q(\mathbb{N})$ ” means

$$\forall x \in \ell^{q'}(\mathbb{N}), \quad \lim_{p \rightarrow \infty} \sum_{n=0}^{\infty} (y_p(n) - y(n))x(n) = 0.$$

The following theorem is a fundamental result, which can be interpreted as a weak- \star compactness property for the unit ball of a dual space.

Theorem 3.3.4 (weak- \star compactness). *Let E be a separable normed space. We consider a bounded sequence $(\ell_n)_{n \in \mathbb{N}}$ in E' . There exists an element ℓ in E' and an extraction function ψ from \mathbb{N} to \mathbb{N} such that*

$$\lim_{n \rightarrow \infty} \star \ell_{\psi(n)} = \ell.$$

Proof. Let $(a_p)_{p \in \mathbb{N}}$ be a dense sequence of E . We start by extracting from $(a_p)_{p \in \mathbb{N}}$ a linearly independent family of vectors, and we will work on the subspace generated by this. More precisely, we define $\varphi(0)$ as the smallest integer p such that a_p is non-zero. Then, set $\varphi(1)$ the smallest integer p such that a_p does not belong to $\mathbb{K}a_{\varphi(0)}$. Let us assume that we have built $(\varphi(0), \dots, \varphi(p))$ such that

$$\forall p' < p, a_{\varphi(p')} \notin V_{p'} \stackrel{\text{def}}{=} \text{Span}\{a_{\varphi(0)}, \dots, a_{\varphi(p')}\}.$$

Then, we set $\varphi(p+1)$ to be the smallest integer q such that a_q is not in V_p . We define

$$V = \bigcup_{p \in \mathbb{N}} V_p = \text{Span}\{a_{\varphi(p)}, p \in \mathbb{N}\}.$$

The family $(a_{\varphi(p)})_{p \in \mathbb{N}}$ is an algebraic basis of the vector space V . Henceforth, we omit the extraction function φ in our notation, and the sequence $(a_p)_{p \in \mathbb{N}}$ is assumed to be a family of linearly independent vectors of E such that $\text{Span}\{a_p, p \in \mathbb{N}\}$ is dense.

We are now going to use Cantor's diagonal argument to construct an extraction function ψ such that

$$\forall j \in \mathbb{N}, \lim_{n \rightarrow \infty} \langle \ell_{\psi(n)}, a_j \rangle = \lambda_j. \quad (3.7)$$

The sequence $(\langle \ell_n, a_0 \rangle)_{n \in \mathbb{N}}$ is a bounded sequence in \mathbb{K} , so there exists an element λ_0 in \mathbb{K} and an extraction function ϕ_0 from \mathbb{N} to \mathbb{N} such that

$$\lim_{n \rightarrow \infty} \langle \ell_{\phi_0(n)}, a_0 \rangle = \lambda_0.$$

Assume that we have constructed a finite family $(\phi_j)_{0 \leq j \leq m}$ of increasing functions from \mathbb{N} to \mathbb{N} and a finite sequence $(\lambda_j)_{0 \leq j \leq m}$ such that, for every $j \leq m$,

$$\lim_{n \rightarrow \infty} \langle \ell_{\phi_0 \circ \dots \circ \phi_j(n)}, a_j \rangle = \lambda_j.$$

The sequence $(\langle \ell_{\phi_0 \circ \dots \circ \phi_m(n)}, a_{m+1} \rangle)_{n \in \mathbb{N}}$ is bounded in \mathbb{K} , so there exists an increasing function ϕ_{m+1} from \mathbb{N} to \mathbb{N} and an element λ_{m+1} in \mathbb{K} such that

$$\forall j \leq m+1, \lim_{n \rightarrow \infty} \langle \ell_{\phi_0 \circ \dots \circ \phi_{m+1}(n)}, a_j \rangle = \lambda_j.$$

Set $\psi(n) = \phi_0 \circ \dots \circ \phi_n(n)$. This is an increasing function from \mathbb{N} to \mathbb{N} . Indeed, observe that any increasing function μ from \mathbb{N} to \mathbb{N} satisfies $\mu(n) \geq n$. As a result, if $n < m$, we have

$$\psi(n) = \phi_0 \circ \dots \circ \phi_n(n) < \phi_0 \circ \dots \circ \phi_n \circ \dots \circ \phi_m(m) = \psi(m).$$

Moreover, for any integer j , we have, for any n strictly greater than j ,

$$\langle \ell_{\psi(n)}, a_j \rangle = \langle \ell_{\phi_0 \circ \phi_1 \circ \dots \circ \phi_j(\phi_{j+1} \circ \dots \circ \phi_n(n))}, a_j \rangle.$$

Thus, as any subsequence of a convergent sequence is convergent, we have

$$\forall j \in \mathbb{N}, \lim_{n \rightarrow \infty} \langle \ell_{\psi(n)}, a_j \rangle = \lambda_j.$$

We will now prove that, for any x in E , the sequence $(\langle \ell_{\psi(n)}, x \rangle)_{n \in \mathbb{N}}$ is a Cauchy sequence. Set x in E . As the sequence $(a_j)_{j \in \mathbb{N}}$ is dense in E , we have

$$\forall \varepsilon > 0, \exists j_\varepsilon \in \mathbb{N} / \|x - a_{j_\varepsilon}\|_E < \frac{\varepsilon}{4M} \quad \text{with} \quad M \stackrel{\text{def}}{=} \sup_{n \in \mathbb{N}} \|\ell_n\|_{E'}.$$

Now, we can write that

$$\begin{aligned} |\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n)}, x \rangle| &\leq |\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n+p)}, a_{j_\varepsilon} \rangle| + |\langle \ell_{\psi(n+p)}, a_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle| \\ &\quad + |\langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, x \rangle| \\ &\leq \|\ell_{\psi(n+p)}\|_{E'} \|x - a_{j_\varepsilon}\|_E + |\langle \ell_{\psi(n+p)}, a_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle| \\ &\quad + \|\ell_{\psi(n)}\|_{E'} \|a_{j_\varepsilon} - x\|_E \\ &\leq \frac{\varepsilon}{2} + |\langle \ell_{\psi(n+p)}, a_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle|. \end{aligned}$$

As the sequence $(\langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle)_{n \in \mathbb{N}}$ converges, it is also a Cauchy sequence. Thus,

$$\exists n_\varepsilon / \forall n \geq n_\varepsilon, \forall p \in \mathbb{N}, |\langle \ell_{\psi(n+p)}, a_{j_\varepsilon} \rangle - \langle \ell_{\psi(n)}, a_{j_\varepsilon} \rangle| < \frac{\varepsilon}{2}.$$

We deduce that

$$\forall \varepsilon > 0, \exists n_\varepsilon / \forall n \geq n_\varepsilon, \forall p \in \mathbb{N}, |\langle \ell_{\psi(n+p)}, x \rangle - \langle \ell_{\psi(n)}, x \rangle| < \varepsilon$$

The sequence $(\langle \ell_{\psi(n)}, x \rangle)_{n \in \mathbb{N}}$ is a Cauchy sequence, so it converges to a limit which we will denote $\ell(x)$. By uniqueness of the limit, the map ℓ defined here is linear (check this as an exercise). Moreover, we have

$$|\ell(x)| \leq M \|x\|_E.$$

This ends the proof of the theorem. □

In the setting of definition 3.3.2, theorem 3.3.4 can be stated as follows.

Theorem 3.3.5. *Let E be a separable Banach space, and let F be a Banach space. Let B be a bilinear form which identifies E' with F . Then, for any bounded sequence $(y_n)_{n \in \mathbb{N}}$ of elements of F , there exists an element y in F and an extraction function ϕ such that*

$$\forall x \in E, \lim_{n \rightarrow \infty} B(y_{\phi(n)}, x) = B(y, x).$$

This is a very important theorem. It means that when a space F can be identified as a dual space, then any bounded sequence has a subsequence that is convergent in the weaker sense given by definition 3.3.2. This is a widely used property, and we will see an important application of this in the next chapter, where we will prove a diagonalisation theorem in Hilbert spaces (theorem 4.4.2 on page 75).

Chapter 4

Hilbert spaces

Introduction

This chapter is dedicated to the study of what are named Hilbert spaces. These are Banach spaces which have Euclidean-type norms, in other words norms that are induced by an inner product. These extend to infinite dimension the notion of Euclidean spaces, which have finite dimension. The key advantage is that we can talk about geometry in these spaces. For example, we can explain what two orthogonal vectors (which may be functions) are. This is a fundamental point of functional analysis worth insisting on: an element of a Banach space (often a function) is truly viewed as a point, or vector, of a geometric space.

The first section presents the concept of Hilbert spaces. In particular, we explain precisely what it means that a norm is associated with an inner product.

The second section shows how this structure generalises the notion of Euclidean space to the infinite dimension setting. For example,

- it is possible to define an orthogonal projection onto a closed convex subset;
- it is possible to extend the familiar concept of orthonormal bases in finite-dimension Euclidean spaces to the case of infinite dimension spaces.

In the third section, we show that the dual space of a Hilbert space \mathcal{H} can be identified with \mathcal{H} itself. We will see in chapter 6 how this provides a spectacularly simple proof of a theorem guaranteeing existence and uniqueness of solutions of a partial differential equation. We finally note that the phenomenon seen in section 3.2 does not take place in Hilbert spaces.

In the fourth and final section of this chapter, we introduce the adjoint of a continuous operator from a Hilbert space to itself, as well as the trickier, new notion of compact operators. The main result of this section is the generalisation of the classic diagonalisation theorem for symmetric operators in finite dimension. This theorem will be used in chapter 6.

4.1 Orthogonality

In the rest of the course, it is understood that, if $\mathbb{K} = \mathbb{R}$, $\bar{\lambda} = \lambda$, while, if $\mathbb{K} = \mathbb{C}$, this is the usual conjugation operation. We first recall the definition of an inner product.

Definition 4.1.1. *Let E and F be two \mathbb{K} -vector spaces. A map ℓ from E to F is said to be antilinear if, for any pair (x, y) in $E \times F$, and any scalar λ in \mathbb{K} , we have*

$$\ell(\lambda x + y) = \bar{\lambda}\ell(x) + \ell(y).$$

A map f from $E \times E$ to \mathbb{K} is said to be sesquilinear¹ if it is linear with respect to the first variable, and antilinear with respect to the second one. In other words,

$$\begin{aligned}\forall \lambda \in \mathbb{K}, \forall (x, y, z) \in E^3 \quad f(\lambda x + y, z) &= \lambda f(x, y) + f(y, z) \\ f(z, \lambda x + y) &= \bar{\lambda} f(z, x) + f(z, y).\end{aligned}$$

A map f from $E \times E$ to \mathbb{K} is an inner product if and only if it is sesquilinear, Hermitian and positive definite, i.e. it satisfies

$$\begin{aligned}\forall (x, y) \in E^2, \quad f(y, x) &= \overline{f(x, y)} \quad (\text{Hermitian}) \\ \forall x \in E, \quad f(x, x) &\in \mathbb{R}^+, \quad (\text{positive}) \\ f(x, x) = 0 &\iff x = 0 \quad (\text{definite}).\end{aligned}$$

Remark. An inner product is often denoted $(\cdot | \cdot)$, and we write $\| \cdot \|^2 = (\cdot | \cdot)$.

Proposition 4.1.1. Let E be a \mathbb{K} -vector space, and let $(\cdot | \cdot)$ be an inner product on E . The map $x \mapsto (x|x)^{\frac{1}{2}}$ is a norm on E , and we have the following properties.

$$\forall (x, y) \in E^2, \quad |(x|y)| \leq \|x\| \|y\| \quad (\text{Cauchy-Schwarz inequality}).$$

$$\forall (x, y) \in E^2, \quad \|x\|^2 + \|y\|^2 = 2 \left\| \frac{x+y}{2} \right\|^2 + 2 \left\| \frac{x-y}{2} \right\|^2 \quad (\text{Appolonius's theorem}).$$

Proof. Obviously, we have

$$0 \leq \left\| \frac{\overline{(x|y)}}{\|x\|} x - \|x\| y \right\|^2.$$

By expanding the right-hand side, we get

$$0 \leq |\overline{(x|y)}|^2 - 2 \Re(\overline{(x|y)}(x|y)) + \|x\|^2 \|y\|^2.$$

To prove that $x \mapsto (x|x)^{\frac{1}{2}}$ is a norm, it suffices to note that

$$\begin{aligned}(x+y|x+y) &\leq (x|x) + 2 \Re(x|y) + (y|y) \\ &\leq (x|x) + 2(x|x)^{\frac{1}{2}}(y|y)^{\frac{1}{2}} + (y|y) \\ &\leq \left((x|x)^{\frac{1}{2}} + (y|y)^{\frac{1}{2}} \right)^2, \quad \text{and that} \\ (\lambda x|\lambda x) &\leq \lambda \bar{\lambda} \|x\|^2.\end{aligned}$$

To prove Appolonius's theorem, we simply observe that

$$\begin{aligned}2 \left\| \frac{x+y}{2} \right\|^2 + 2 \left\| \frac{x-y}{2} \right\|^2 &= \frac{1}{2} \|x+y\|^2 + \frac{1}{2} \|x-y\|^2 \\ &= \frac{1}{2} (\|x\|^2 + 2 \Re(x|y) + \|y\|^2 + \|x\|^2 - 2 \Re(x|y) + \|y\|^2) \\ &= \|x\|^2 + \|y\|^2.\end{aligned}$$

The proposition is proved. □

¹from the Latin "sesqui", meaning "one and a half"

Exercise 4.1.1. Let $(E, \|\cdot\|)$ be a \mathbb{K} -vector space. Assume that

$$\forall (x, y) \in E^2, \quad \|x\|^2 + \|y\|^2 = 2 \left\| \frac{x+y}{2} \right\|^2 + 2 \left\| \frac{x-y}{2} \right\|^2.$$

Prove that there exists a sesquilinear map such that $\|x\|^2 = (x|x)$. We recommend starting with $\mathbb{K} = \mathbb{R}$, which is easier, before dealing with the case $\mathbb{K} = \mathbb{C}$.

An inner product allows one to define the notion of orthogonal vectors. This concept, which we are familiar with in the plane and in 3-dimensional space, will be extremely rich with applications in infinite-dimension spaces, in particular function spaces.

Definition 4.1.2. Let E be a vector space endowed with an inner product, and let x and y be two elements of E . The vectors x and y are said to be orthogonal (denoted $x \perp y$) if and only if $(x|y) = 0$.

We recall the highly famous Pythagoras equality:

$$x \perp y \implies \|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

Proposition 4.1.2. Let E be a \mathbb{K} -vector space endowed with an inner product, and let A be a subset of E . Define

$$A^\perp \stackrel{\text{def}}{=} \{x \in E / \forall a \in A, x \perp a\}.$$

The set A^\perp is a closed vector subspace of E .

Proof. The linear map \mathcal{L}_a from E to \mathbb{K} defined by $x \mapsto (x|a)$ is continuous, so its kernel is a closed vector subspace of E . It is clear that

$$A^\perp = \bigcap_{a \in A} \ker \mathcal{L}_a,$$

so A^\perp is a closed vector subspace of E , as it is an intersection of the same. This proves the proposition. \square

Exercise 4.1.2. Let E be a \mathbb{K} vector space endowed with a Hermitian inner product, and let A be a subset of E . Prove that $\overline{A^\perp} = A^\perp = \text{Span}(A)^\perp$.

Definition 4.1.3. Let \mathcal{H} be a \mathbb{K} -vector space endowed with an inner product $(\cdot|\cdot)$. The space \mathcal{H} is called a Hilbert space if and only if the space \mathcal{H} is complete for the norm associated with the inner product.

The Hilbert space structure is absolutely fundamental. Many simple methods and concepts from geometry in Euclidean spaces with finite dimension will reappear here.

4.2 Properties of Hilbert spaces

Pythagoras's equality is extended to Hilbert spaces as follows.

Theorem 4.2.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a Hilbert space \mathcal{H} such that, for $n \neq m$, $x_n \perp x_m$. Then the series $\sum_n x_n$ is convergent if and only if the series $\sum_n \|x_n\|^2$ is convergent, and in that case,

$$\left\| \sum_{n \in \mathbb{N}} x_n \right\|^2 = \sum_{n \in \mathbb{N}} \|x_n\|^2$$

Proof. Let us set $S_q = \sum_{p \leq q} x_p$. Pythagoras's theorem yields

$$\sum_{p \leq q} \|x_p\|^2 = \|S_q\|^2. \quad (4.1)$$

If the right-hand side of the above equality is assumed to be convergent, then it is bounded independently of q , and the series $\sum_n \|x_n\|^2$ is convergent. Conversely, if the series $\sum_n \|x_n\|^2$ is convergent, then

$$\begin{aligned} \|S_{q+q'} - S_q\|^2 &= \sum_{p=q+1}^{q+q'} \|x_p\|^2 \\ &\leq \sum_{p=q+1}^{\infty} \|x_p\|^2. \end{aligned}$$

By taking the limit in equality (4.1), we finish the proof of the theorem. \square

Exercise 4.2.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a Hilbert space \mathcal{H} such that

$$\sum_0^{\infty} \|x_n\|^2 < \infty.$$

Assume that there exists an integer N_0 such that, if $|n - m| \geq N_0$, the vectors x_n and x_m are orthogonal. Prove that the series $\sum_n x_n$ is convergent, and that there exists a constant C , which only depends on N_0 , such that

$$\left\| \sum_{n \in \mathbb{N}} x_n \right\|^2 \leq C \sum_{n \in \mathbb{N}} \|x_n\|^2.$$

Many remarkable properties of Hilbert spaces are based on the following projection theorem.

Theorem 4.2.2 (projection onto a closed convex subset). *Let Γ be a closed convex subset of a Hilbert space \mathcal{H} . Then, for any point x in \mathcal{H} , there exists a unique point in Γ , which we denote $p_{\Gamma}(x)$ and call the projection of x on Γ , such that*

$$\|x - p_{\Gamma}(x)\| = \inf_{g \in \Gamma} \|x - g\|.$$

Proof. Recall Apollonius's equality: if γ and γ' are two points in Γ , then we have

$$\frac{1}{2} \|\gamma - \gamma'\|^2 = \|x - \gamma\|^2 + \|x - \gamma'\|^2 - 2 \left\| x - \frac{\gamma + \gamma'}{2} \right\|^2. \quad (4.2)$$

Note that, since the set Γ is convex, the point $\frac{\gamma + \gamma'}{2}$ is also in Γ . We first prove uniqueness. Assume that there are two points γ_1 and γ_2 in Γ which both minimise $\|x - g\|$ for $g \in \Gamma$. Then, Apollonius's equality yields

$$\begin{aligned} \frac{1}{2} \|\gamma_1 - \gamma_2\|^2 &= \|x - \gamma_1\|^2 + \|x - \gamma_2\|^2 - 2 \left\| x - \frac{\gamma_1 + \gamma_2}{2} \right\|^2 \\ &= 2d^2 - 2 \left\| x - \frac{\gamma_1 + \gamma_2}{2} \right\|^2 \\ &\leq 2d^2 - 2d^2 = 0. \end{aligned}$$

This provides uniqueness. Now we prove existence. By definition of the infimum, there exists a minimising sequence, that is a sequence $(\gamma_n)_{n \in \mathbb{N}}$ of elements of Γ such that

$$\lim_{n \rightarrow \infty} \|x - \gamma_n\| = d \stackrel{\text{def}}{=} \inf_{\gamma \in \Gamma} \|x - \gamma\|.$$

We use relation (4.2) again to get that, for any pair of integers (n, m) ,

$$\frac{1}{2} \|\gamma_n - \gamma_m\|^2 = \|x - \gamma_n\|^2 + \|x - \gamma_m\|^2 - 2 \left\| x - \frac{\gamma_n + \gamma_m}{2} \right\|^2.$$

As Γ is convex, the midpoint between γ_n and γ_m is also in Γ , thus,

$$\frac{1}{2} \|\gamma_n - \gamma_m\|^2 \leq \|x - \gamma_n\|^2 + \|x - \gamma_m\|^2 - 2d^2.$$

Given the definition of the sequence $(\gamma_n)_{n \in \mathbb{N}}$, the above yields that it is a Cauchy sequence. As \mathcal{H} is a Hilbert space, it is complete, and, Γ being a closed subset of \mathcal{H} , it is also complete, so the sequence $(\gamma_n)_{n \in \mathbb{N}}$ converges to a point in Γ . This ends the proof of the theorem. \square

Proposition 4.2.1. *Let Γ be a closed convex subset of a Hilbert space \mathcal{H} , and let x and x' be two points in \mathcal{H} . We denote $p_\Gamma(x)$ the unique point in Γ such that $d(x, \Gamma) = \|x - p_\Gamma(x)\|$. Then, for any γ in Γ ,*

$$\Re(x - p_\Gamma(x) | p_\Gamma(x) - \gamma) \geq 0 \quad \text{and} \quad \|p_\Gamma(x) - p_\Gamma(x')\| \leq \|x - x'\|.$$

Proof. To get the first inequality, we observe that, as Γ is convex, for any λ in $[0, 1]$ and any γ in Γ ,

$$\|x - ((1 - \lambda)p_\Gamma(x) + \lambda\gamma)\|^2 = \|x - p_\Gamma(x) + \lambda(p_\Gamma(x) - \gamma)\|^2 \geq \|x - p_\Gamma(x)\|^2.$$

A quick computation yields that

$$\forall \lambda \in [0, 1], \forall \gamma \in \Gamma \quad \lambda(2 \Re(x - p_\Gamma(x) | p_\Gamma(x) - \gamma) + \lambda \|p_\Gamma(x) - \gamma\|^2) \geq 0$$

which proves the first inequality. To prove the second one, we use the first result on the points x and $\gamma = p_\Gamma(x')$, then on the points x' and $\gamma = p_\Gamma(x)$. This yields

$$\begin{aligned} \Re(x - p_\Gamma(x) | p_\Gamma(x) - p_\Gamma(x')) &\geq 0 \\ \text{and} \quad \Re(x' - p_\Gamma(x') | p_\Gamma(x') - p_\Gamma(x)) &\geq 0. \end{aligned}$$

Adding the two together, we get

$$\Re(x - x' - (p_\Gamma(x) - p_\Gamma(x')) | p_\Gamma(x) - p_\Gamma(x')) \geq 0.$$

Using the Cauchy-Schwarz inequality, this implies

$$\|p_\Gamma(x) - p_\Gamma(x')\|^2 \leq \Re(x - x' | p_\Gamma(x) - p_\Gamma(x')) \leq \|x - x'\| \|p_\Gamma(x) - p_\Gamma(x')\|,$$

which ends the proof of the theorem. \square

Nearly all properties in Hilbert spaces can be seen as corollaries of the above theorem 4.2.2.

Corollary 4.2.1. *Let F be a closed vector subspace of a Hilbert space \mathcal{H} . Then,*

$$\mathcal{H} = F \oplus F^\perp \quad \text{and} \quad x = p_F(x) + p_{F^\perp}(x).$$

Proof. Let x be an element of \mathcal{H} . For any point f in F , and for any real number λ , we have

$$\|x - p_F(x) + \lambda f\|^2 = \lambda^2 \|f\|^2 + 2\lambda \Re e(x - p_F(x)|f) + \|x - p_F(x)\|^2.$$

By definition of the projection on F , we must have

$$\|x - p_F(x) + \lambda f\|^2 \geq \|x - p_F(x)\|^2.$$

This implies that, for any f in F , we must have

$$\Re e(x - p_F(x)|f) = 0.$$

If $\mathbb{K} = \mathbb{R}$, this is all we need in order to state that

$$x - p_F(x) \in F^\perp.$$

If $\mathbb{K} = \mathbb{C}$ however, we must repeat the above argument with if instead of f to reach that conclusion. Hence, we have proved that $\mathcal{H} = F + F^\perp$. But, if $x \in F \cap F^\perp$, then $(x|x) = 0$, thus $x = 0$. The corollary is proved. \square

Corollary 4.2.2. *Let A be a subset of a Hilbert space \mathcal{H} . This subset A is total (i.e. the vector subspace it spans is dense) if and only if $A^\perp = \{0\}$.*

Proof. We use the result of exercise 4.1.2, which states that $A^\perp = \overline{(\text{Span}(A))}^\perp$. If $A^\perp = \{0\}$, then $\overline{\text{Span}(A)}$ is equal to \mathcal{H} , by corollary 4.2.1, which precisely means that the subset A is total.

Conversely, if the subset A is total, then $\overline{\text{Span}(A)}$ is equal to \mathcal{H} by definition, hence $A^\perp = \{0\}$. The property is proved. \square

Exercise 4.2.2. *Let A be a subset of a Hilbert space \mathcal{H} . Prove that*

$$(A^\perp)^\perp = \overline{\text{Span}(A)}.$$

Definition 4.2.1. *Let \mathcal{H} be a separable Hilbert space with infinite dimension. We call an orthonormal basis or Hilbertian basis (French term) of \mathcal{H} , any sequence $(e_n)_{n \in \mathbb{N}}$ of elements of \mathcal{H} which is total, and such that*

$$(e_n | e_m) = \delta_{n,m} \quad \text{with} \quad \delta_{n,m} = 1 \quad \text{if} \quad n = m \quad \text{and} \quad 0 \quad \text{otherwise.} \quad (4.3)$$

Remark. In this setting, an orthonormal basis is never a basis in the algebraic sense.

Exercise 4.2.3. *Consider the space $\ell^2(\mathbb{N})$, and the sequence $(e_n)_{n \in \mathbb{N}}$ of elements of $\ell^2(\mathbb{N})$ defined by*

$$e_n(k) = \delta_{n,k}.$$

Prove that $(e_n)_{n \in \mathbb{N}}$ is an orthonormal basis of $\ell^2(\mathbb{N})$. What is the vector space spanned by the family $(e_n)_{n \in \mathbb{N}}$?

Theorem 4.2.3. *In every separable Hilbert space \mathcal{H} , there exists an orthonormal basis.*

Proof. Let $(a_n)_{n \in \mathbb{N}}$ be a total countable subset of \mathcal{H} . We can assume that the family $(a_n)_{n \in \mathbb{N}}$ consists of linearly independent vectors. Indeed, we define the following extraction function

$$\phi(0) = \min\{n / a_n \neq 0\} \quad \text{and} \quad \phi(n) = \min\{m / a_m \notin \text{Span}\{a_0, \dots, a_{\phi(n)}\}\}.$$

We let the reader check, as an exercise, that the vector subspace generated by the $a_{\phi(n)}$ is the same as the one spanned by the a_n . We will now apply the Gram-Schmidt orthonormalisation method to this family. Let us recall the principle. First, we set

$$e_1 = \frac{a_1}{\|a_1\|}.$$

We assume that the terms $(e_j)_{1 \leq j \leq n}$ satisfy property (4.3), so that

$$\text{Span}\{a_1, \dots, a_n\} = \text{Span}\{e_1, \dots, e_n\}.$$

Set

$$e_{n+1} = \frac{f_{n+1}}{\|f_{n+1}\|} \quad \text{with} \quad f_{n+1} = a_{n+1} - \sum_{j=1}^n (a_{n+1}|e_j)e_j.$$

It is quick to check the equalities in (4.3) are satisfied by e_{n+1} . The theorem is proved. \square

Theorem 4.2.4. *Let \mathcal{H} be a separable Hilbert space, and let $(e_n)_{n \in \mathbb{N}}$ be an orthonormal basis of \mathcal{H} . The map \mathcal{I} defined by*

$$\begin{cases} \mathcal{H} & \longrightarrow & \ell^2(\mathbb{N}) \\ x & \longmapsto & ((x|e_n))_{n \in \mathbb{N}} \end{cases}$$

is a linear isometric bijection. In particular, the Bessel and Parseval equalities hold in \mathcal{H} :

$$x = \sum_{n \in \mathbb{N}} (x|e_n)e_n \quad \text{and} \quad \|x\|^2 = \sum_{n \in \mathbb{N}} |(x|e_n)|^2.$$

Proof. The main point to prove is that \mathcal{I} is indeed a map from \mathcal{H} to $\ell^2(\mathbb{N})$. We set

$$x_q \stackrel{\text{def}}{=} \sum_{p \leq q} (x|e_p)e_p.$$

It is clear that

$$\begin{aligned} (x|x_q) &= \sum_{p \leq q} \overline{(x|e_p)}(x|e_p) \\ &= \sum_{p \leq q} |(x|e_p)|^2 \\ &= (x_q|x_q). \end{aligned}$$

The Cauchy-Schwarz inequality states that $\|x_q\|^2 \leq \|x\| \times \|x_q\|$. We deduce that, for any integer q , we have

$$\sum_{p=0}^q |(x|e_p)|^2 \leq \|x\|^2.$$

So $\mathcal{I}(x)$ does belong to $\ell^2(\mathbb{N})$.

The fact that the map \mathcal{I} is one to one is deduced from corollary 4.2.2, which states that the orthogonal set of a total subset is the singleton $\{0\}$. That the map \mathcal{I} is onto and an isometry, is deduced from theorem 4.2.1. Thus, theorem 4.2.3 is proved. \square

4.3 Duality in Hilbert spaces

Let \mathcal{H} be a Hilbert space with finite dimension d ; in other words, \mathcal{H} is a Euclidean or Hermitian space. We know that the map δ defined by

$$\delta \begin{cases} \mathcal{H} & \longrightarrow \mathcal{H}' \\ x & \longmapsto \delta(x) : h \mapsto (h|x) \end{cases}$$

is an antilinear isometric bijection from \mathcal{H} to \mathcal{H}' . When \mathcal{H} has infinite dimension, we will see few changes.

Theorem 4.3.1 (Riesz representation theorem). *Let \mathcal{H} be a Hilbert space. We consider the map δ defined by*

$$\delta \begin{cases} \mathcal{H} & \longrightarrow \mathcal{H}' \\ x & \longmapsto \delta(x) : h \mapsto (h|x) \end{cases}$$

This is an antilinear isometric bijection.

Proof. The fact that $\|\delta(x)\|_{\mathcal{H}'} \leq \|x\|$ (and therefore $\delta(x)$ is a continuous linear form) is obtained by using the Cauchy-Schwarz inequality. It is obvious that δ is antilinear. Moreover,

$$\langle \delta(x), \frac{x}{\|x\|} \rangle = \|x\|;$$

so δ is an isometry. An isometry is always one to one, so it remains to prove that δ is onto. To do this, let us consider an element ℓ in $\mathcal{H}' \setminus \{0\}$. Its kernel is a closed subspace of \mathcal{H} , not equal to \mathcal{H} . Let x be a non-zero vector in $(\ker \ell)^\perp$, such that $\langle \ell, x \rangle = \|x\|^2$. Both ℓ and $\delta(x)$ have the same kernel, and there exists a non-zero h_0 such that $\langle \ell, h_0 \rangle = \langle \delta(x), h_0 \rangle$, so $\delta(x) = \ell$. Thus, the theorem is proved. \square

The above theorem can be seen from a different angle when $\mathbb{K} = \mathbb{R}$, which will be useful in chapter 5 to describe the duals of the L^p spaces.

Theorem 4.3.2. *Let \mathcal{H} be a real Hilbert space. For ℓ a continuous linear functional on \mathcal{H} , we define the function*

$$F \begin{cases} \mathcal{H} & \longrightarrow \mathbb{R} \\ u & \longmapsto \frac{1}{2}\|u\|_{\mathcal{H}}^2 - \langle \ell, u \rangle. \end{cases}$$

The function F is bounded from below, and there exists a unique v in \mathcal{H} such that $F(v) = \inf_{u \in \mathcal{H}} F(u)$ which satisfies

$$\forall h \in \mathcal{H}, (v|h)_{\mathcal{H}} = \langle \ell, h \rangle.$$

Proof. As the linear functional ℓ is continuous, we can write that

$$\begin{aligned} F(u) &\geq \frac{1}{2}\|u\|_{\mathcal{H}}^2 - \|\ell\|_{\mathcal{H}'}\|u\| \\ &\geq \frac{1}{2}(\|u\|_{\mathcal{H}} - \|\ell\|_{\mathcal{H}'})^2 - \frac{1}{2}\|\ell\|_{\mathcal{H}'}^2. \end{aligned}$$

Let $\mu \stackrel{\text{def}}{=} \inf_{u \in \mathcal{H}} F(u)$, and let $(u_n)_{n \in \mathbb{N}}$ be a minimising sequence, that is a sequence of elements of \mathcal{H} such that

$$F(u_n) = \mu + \varepsilon_n \quad \text{with} \quad \lim_{n \rightarrow \infty} \varepsilon_n = 0. \quad (4.4)$$

By using Appolonius's relation, we get

$$\left\| \frac{u_n - u_m}{2} \right\|_{\mathcal{H}}^2 = \frac{1}{2} \|u_n\|_{\mathcal{H}}^2 + \frac{1}{2} \|u_m\|_{\mathcal{H}}^2 - \left\| \frac{u_n + u_m}{2} \right\|_{\mathcal{H}}^2.$$

As ℓ is a linear map, we have

$$\begin{aligned} \left\| \frac{u_n - u_m}{2} \right\|_{\mathcal{H}}^2 &= \frac{1}{2} \|u_n\|_{\mathcal{H}}^2 - \langle \ell, u_n \rangle + \frac{1}{2} \|u_m\|_{\mathcal{H}}^2 - \langle \ell, u_m \rangle \\ &\quad - 2 \left(\frac{1}{2} \left\| \frac{u_n + u_m}{2} \right\|_{\mathcal{H}}^2 - \left\langle \ell, \frac{u_n + u_m}{2} \right\rangle \right). \end{aligned}$$

Using (4.4), and by definition of μ , we obtain

$$\left\| \frac{u_n - u_m}{2} \right\|_{\mathcal{H}}^2 \leq \varepsilon_n + \varepsilon_m.$$

This proves that the sequence $(u_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, so it converges to an element v in \mathcal{H} . As the function F is continuous, we have $F(v) = \mu$. Since, for any h in \mathcal{H} and any real number t , we have

$$\mu \leq F(v + th) = \frac{1}{2} \|v\|_{\mathcal{H}}^2 + t(v|h) + \frac{1}{2} t^2 \|h\|_{\mathcal{H}}^2 - t\langle \ell, h \rangle,$$

from which we deduce

$$\forall h \in \mathcal{H}, \forall t \in \mathbb{R}, t(v|h) + \frac{1}{2} t^2 \|h\|_{\mathcal{H}}^2 - t\langle \ell, h \rangle \geq 0,$$

which implies that

$$\forall h \in \mathcal{H}, (v|h) = \langle \ell, h \rangle.$$

Such an element v is of course unique, because if $(v_1 - v_2|h) = 0$ for every h , we can choose $h = v_1 - v_2$, and this proves uniqueness. \square

We can deduce the following two consequences.

Definition 4.3.1. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a Hilbert space \mathcal{H} , and let x be an element of \mathcal{H} . The sequence $(x_n)_{n \in \mathbb{N}}$ is said to be weakly convergent to x , denoted $\limf_{n \rightarrow \infty} x_n = x$ or $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$, if and only if

$$\forall h \in \mathcal{H}, \lim_{n \rightarrow \infty} (h|x_n) = (h|x).$$

Observing this definition, we note that

$$\limf_{n \rightarrow \infty} x_n = x \iff \limf_{n \rightarrow \infty} \star \delta(x_n) = \delta(x).$$

Among other things, the following theorem explains the terminology, why the convergence is called weak.

Theorem 4.3.3. Let $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of a Hilbert space \mathcal{H} , and let x be an element of \mathcal{H} . Then, we have:

$$\text{if } \lim_{n \rightarrow \infty} \|x_n - x\| = 0, \text{ then } \limf_{n \rightarrow \infty} x_n = x; \quad (4.5)$$

$$\text{if } \limf_{n \rightarrow \infty} x_n = x \text{ and } \lim_{n \rightarrow \infty} \|x_n\| = \|x\|, \text{ then } \lim_{n \rightarrow \infty} \|x_n - x\| = 0. \quad (4.6)$$

Moreover, if $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$, then the sequence $(x_n)_{n \in \mathbb{N}}$ is bounded.

Proof. The first point of the theorem is a direct consequence of the fact that

$$|(h|x_n) - (h|x)| \leq \|h\| \times \|x_n - x\|.$$

For the second point, it suffices to write that

$$\|x_n - x\|^2 = \|x_n\|^2 - 2\Re(x|x_n) + \|x\|^2;$$

As the sequence $(x_n)_{n \in \mathbb{N}}$ is weakly convergent to x , we have

$$-2 \lim_{n \rightarrow \infty} \Re(x|x_n) = -2\|x\|^2.$$

This proves the second point. The third is a corollary of the Banach-Steinhaus theorem. Indeed, the weak convergence hypothesis means that, for any y in \mathcal{H} ,

$$\lim_{n \rightarrow \infty} \delta(x_n)(y) = \lim_{n \rightarrow \infty} (y|x_n) = (y|x).$$

Theorem 3.3.2 (Banach-Steinhaus) states that the sequence $(\delta(x_n))_{n \in \mathbb{N}}$ is bounded in \mathcal{H} . By the previous Riesz representation theorem, we get that $(x_n)_{n \in \mathbb{N}}$ is a bounded sequence. The theorem is proved.

Example. Let \mathcal{H} be a separable Hilbert space with infinite dimension. Consider an orthonormal basis $(e_n)_{n \in \mathbb{N}}$ of \mathcal{H} . By the Riesz representation theorem 4.3.1, for any x in \mathcal{H} , the sequence $((x|e_n))_{n \in \mathbb{N}}$ belongs to $\ell^2(\mathbb{N})$, so it converges to 0. However, the sequence $(e_n)_{n \in \mathbb{N}}$ does not converge, as the norm of each e_n is equal to 1. This is therefore a sequence which is weakly convergent, but not convergent in norm.

Proposition 4.3.1. *Let $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ be two sequences of elements in \mathcal{H} such that*

$$\lim_{n \rightarrow \infty} x_n = x \quad \text{and} \quad \liminf_{n \rightarrow \infty} y_n = y.$$

Then, $\lim_{n \rightarrow \infty} (x_n|y_n) = (x|y)$.

Proof. We simply write that

$$\begin{aligned} |(x_n|y_n) - (x|y)| &\leq |(x_n - x|y_n)| + |(x|y_n - y)| \\ &\leq \|x_n - x\| \times \|y_n\| + |(x|y_n - y)|. \end{aligned}$$

Theorem 4.3.3 yields that the sequence $(y_n)_{n \in \mathbb{N}}$ is bounded, so

$$|(x_n|y_n) - (x|y)| \leq C\|x_n - x\| + |(x|y_n - y)|,$$

The proposition is proved.

Exercise 4.3.1. *Find two sequences $(x_n)_{n \in \mathbb{N}}$ and $(y_n)_{n \in \mathbb{N}}$ in \mathcal{H} such that*

$$\liminf_{n \rightarrow \infty} x_n = x, \quad \liminf_{n \rightarrow \infty} y_n = y \quad \text{and} \quad \lim_{n \rightarrow \infty} (x_n|y_n) \neq (x|y).$$

Exercise 4.3.2. *Prove that, in a Hilbert space \mathcal{H} , if a sequence $(x_n)_{n \in \mathbb{N}}$ is weakly convergent to x , then*

$$\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$$

The following theorem is a weak compactness result for the unit ball of a Hilbert space. When the space does not have finite dimension, we know, by theorem 2.3.1, that the unit ball is not compact in the topology defined by the norm. However, we have the next theorem, which is merely theorem 3.3.4 on page 60 in this setting.

Theorem 4.3.4 (weak compactness). *Let $(x_n)_{n \in \mathbb{N}}$ be a bounded sequence in a separable Hilbert space \mathcal{H} . We can extract from $(x_n)_{n \in \mathbb{N}}$ a weakly convergent subsequence.*

4.4 The adjoint of an operator, self-adjoint operators

The adjoint of a linear map is a well-known notion in finite dimension. Given an inner product on a vector space \mathcal{H} with finite dimension d , the adjoint of A is denoted A^* , and defined by the relation

$$(Ax|y) = (x|A^*y).$$

It is common knowledge that, if $(A_{i,j})_{1 \leq i,j \leq d}$ is the matrix of the map A in an orthonormal basis, then the matrix representing A^* in the same basis is such that $A_{i,j}^* = \overline{A_{j,i}}$. Moreover, there is a classic linear algebra theorem which states that if A is self-adjoint (i.e. $A = A^*$), then it can be diagonalised in an orthonormal basis.

In this section, we generalise these concepts and results in Hilbert spaces with infinite dimension.

Theorem 4.4.1. *Let A be a continuous linear operator on a Hilbert space \mathcal{H} . There exists a unique linear operator A^* , which is continuous on \mathcal{H} , such that*

$$\forall (x, y) \in \mathcal{H} \times \mathcal{H}, (Ax|y) = (x|A^*y).$$

Moreover, the map from $\mathcal{L}(\mathcal{H})$ to itself, defined by $A \mapsto A^*$, is an antilinear isometry.

We extend from this theorem the following definition.

Definition 4.4.1. *The map A^* given above is called the adjoint of the operator A . An element A of $\mathcal{L}(\mathcal{H})$ is self-adjoint if and only if $A = A^*$.*

Proof of theorem 4.4.1. Uniqueness of the operator A^* is a consequence of the fact that $\mathcal{H}^\perp = \{0\}$. Let \mathcal{L}_A be the map defined by

$$\mathcal{L}_A \begin{cases} \mathcal{H} & \longrightarrow \mathcal{H}' \\ y & \longmapsto \mathcal{L}_A(y) : x \longmapsto (Ax|y). \end{cases}$$

It is clear that the map \mathcal{L}_A is a continuous antilinear map from \mathcal{H} to \mathcal{H}' . Set $A^* \stackrel{\text{def}}{=} \delta^{-1} \circ \mathcal{L}_A$. By definition of A^* , we have

$$(x|A^*y) = (\mathcal{L}_A y|x) = (Ax|y).$$

So, the operator A^* satisfies the desired relation with respect to A . We also have

$$\begin{aligned} \|A\| &= \sup_{\|x\| \leq 1, \|y\| \leq 1} |(Ax|y)| \\ &= \sup_{\|x\| \leq 1, \|y\| \leq 1} |(x|A^*y)| \\ &= \|A^*\|. \end{aligned}$$

So, theorem 4.4.1 is proved.

Exercise 4.4.1. *Let \mathcal{H} be a real Hilbert space, and let Ω be an open subset of \mathcal{H} . We define the “gradient” operator by*

$$\text{grad} \begin{cases} C^1(\Omega; \mathbb{R}) & \longrightarrow \mathcal{C}(\Omega, \mathcal{H}) \\ g & \longmapsto \delta^{-1} Dg \end{cases}$$

Let g be a C^1 function from Ω to \mathbb{R} , another open set of \mathcal{H} , and let f be a C^1 function from Ω' to \mathbb{R} . Compute $\text{grad}(f \circ g)$.

Proposition 4.4.1. *Let A be a continuous linear map on a Hilbert space \mathcal{H} . The following properties hold.*

The \star operation is involutory, i.e. $A = A^{\star\star}$. We also have

$$(\ker A)^\perp = \overline{\operatorname{Im} A^\star} \quad \text{and} \quad \ker A = (\operatorname{Im} A^\star)^\perp. \quad (4.7)$$

Finally, if a sequence $(x_n)_{n \in \mathbb{N}}$ is weakly convergent to x , then the sequence $(Ax_n)_{n \in \mathbb{N}}$ is weakly convergent to Ax .

Proof. The first point is simple: since $(Ax|y) = (x|A^\star y)$, we have $(A^\star y|x) = (y|Ax)$. To prove the equalities in (4.7), write that

$$\forall y \in \mathcal{H}, (Ax|y) = 0 \iff \forall y \in \mathcal{H}, (x|A^\star y) = 0.$$

Restating the above equivalence in terms of sets yields that $\ker A$ and $(\operatorname{Im} A^\star)^\perp$ are equal. The other relation is deduced from this by taking the orthogonal.

Let $(x_n)_{n \in \mathbb{N}}$ be a sequence that converges weakly to x . For any y in \mathcal{H} , we have

$$(y|Ax_n) = (A^\star y|x_n).$$

So, the weak convergence of the sequence $(x_n)_{n \in \mathbb{N}}$ implies the weak convergence of the sequence $(Ax_n)_{n \in \mathbb{N}}$. \square

Proposition 4.4.2. *Let A be a self-adjoint operator in $\mathcal{L}(\mathcal{H})$. Then,*

$$\|A\|_{\mathcal{L}(\mathcal{H})} = \sup_{\|u\|=1} |(Au|u)|.$$

Proof. First, we note that

$$\|A\|_{\mathcal{L}(\mathcal{H})} = \sup_{\substack{\|u\|=1 \\ \|v\|=1}} \Re e(Au|v).$$

It is clear that $M \stackrel{\text{def}}{=} \sup_{\|u\|=1} |(Au|u)|$ is less than or equal to $\|A\|_{\mathcal{L}(\mathcal{H})}$. Let u and v be two vectors in \mathcal{H} with norm 1. We have

$$\begin{aligned} \Re e(Au|v) &= \frac{1}{4}((A(u+v)|u+v) - (A(u-v)|u-v)) \\ &\leq \frac{M}{4}(\|u+v\|^2 + \|u-v\|^2). \end{aligned}$$

As $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$, and the norms of u and v are equal to 1, we get

$$\Re e(Au|v) \leq M,$$

which proves the proposition. \square

We are now going to study a generalisation of the diagonalisation theorem for self-adjoint operators on finite-dimension spaces. For this, we require an extra hypothesis on the operator: a compactness hypothesis. So, we first introduce the notion of compact operators. This notion is defined more generally for linear maps between Banach spaces.

Definition 4.4.2. *Let E and F be two Banach spaces. An element u in $\mathcal{L}(E, F)$ is called compact if and only if, for any bounded subset A of E , the closure of the set $u(A)$ is compact in F .*

Example. If $u(E)$ is a vector space with finite dimension, then the operator u is compact. Indeed, if A is a bounded subset of E , the image of A is a bounded subset of the space $u(E)$, which has finite dimension, and the closure of such a set is compact.

Exercise 4.4.2. Let E be the vector space of functions f from $[0, 1]$ to \mathbb{R} such that

$$\sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^{\frac{1}{2}}} < \infty$$

1) Prove that

$$\|f\|_E \stackrel{\text{def}}{=} |f(0)| + \sup_{x \neq x'} \frac{|f(x) - f(x')|}{|x - x'|^{\frac{1}{2}}}$$

is a norm on E , and that $(E, \|\cdot\|_E)$ is a Banach space.

2) Prove that the map

$$\iota \begin{cases} E & \longrightarrow \mathcal{C}([0, 1], \mathbb{R}) \\ f & \longmapsto f \end{cases}$$

is compact.

Exercise 4.4.3. Let $(A(n, m))_{(n, m) \in \mathbb{N}^2}$ be a sequence of real numbers such that $\sum_{(n, m) \in \mathbb{N}^2} A_{n, m}^2$ is finite.

1) Prove that the relation

$$(\mathcal{A}x)(n) \stackrel{\text{def}}{=} \sum_{m \in \mathbb{N}} A(n, m)x(m)$$

defines a continuous operator \mathcal{A} from $\ell^2(\mathbb{N})$ to itself, and that $\|\mathcal{A}\|_{\mathcal{L}(\ell^2)} \leq \left(\sum_{(n, m) \in \mathbb{N}^2} A_{n, m}^2 \right)^{\frac{1}{2}}$.

2) Let \mathcal{A}_N be an operator defined by $(\mathcal{A}_N x)(n) = \mathbf{1}_{\{0, \dots, N\}}(n)(\mathcal{A}x)(n)$. Prove that

$$\lim_{N \rightarrow \infty} \|\mathcal{A}_N - \mathcal{A}\|_{\mathcal{L}(\ell^2)} = 0.$$

3) Deduce that the operator \mathcal{A} is compact.

We now state the diagonalisation theorem for compact, self-adjoint operators on a Hilbert space.

Theorem 4.4.2. Let A be a compact, self-adjoint operator on a separable Hilbert space \mathcal{H} with infinite dimension. There exists a sequence of real numbers $(\lambda_j)_{j \in \mathbb{N}}$, such that $(|\lambda_j|)_{j \in \mathbb{N}}$ is nonincreasing and converges to 0, such that:

- for any j such that λ_j is non-zero, λ_j is an eigenvalue of A , and $E_j \stackrel{\text{def}}{=} \ker A - \lambda_j \text{Id}$ is a subspace of \mathcal{H} with finite dimension; moreover, if λ_j and $\lambda_{j'}$ are distinct, their corresponding eigenspaces are orthogonal;
- if $E \stackrel{\text{def}}{=} \text{Span} \bigcup_{\substack{j \in \mathbb{N} \\ \lambda_j \neq 0}} E_j$, then $\ker A = E^\perp$;
- we have $\|A\|_{\mathcal{L}(\mathcal{H})} = \max_{j \in \mathbb{N}} |\lambda_j|$.

Before we prove this theorem, we recall and re-prove the diagonalisation theorem for self-adjoint operators on Euclidean (or Hermitian) spaces with finite dimension.

Theorem 4.4.3. *Let A be a self-adjoint operator on a Hilbert space with finite dimension N . There exists a finite sequence $(\lambda_j)_{1 \leq j \leq N}$ and an orthonormal basis $(e_j)_{1 \leq j \leq N}$ such that, for any j , e_j is an eigenvector corresponding to the eigenvalue λ_j .*

Proof. Consider the quadratic (or Hermitian) form associated with A , that is

$$Q_A \quad \begin{cases} \mathcal{H} & \longrightarrow \mathbb{R} \\ x & \longmapsto (Ax|x) \end{cases}$$

This is a continuous function on \mathcal{H} . We denote $\mathbb{S}_{\mathcal{H}}$ the unit sphere of \mathcal{H} . As the space has finite dimension, corollary 1.3.2 on page 24, a consequence of Heine's theorem 1.3.3 on page 23, yields that there exists a point x_M in $\mathbb{S}_{\mathcal{H}}$ such that

$$|(Ax_M|x_M)| = M_0 \stackrel{\text{def}}{=} \sup_{x \in \mathbb{S}_{\mathcal{H}}} |(Ax|x)| \quad (4.8)$$

We aim to show that x_M is an eigenvector of A , corresponding to the eigenvalue $\pm M_0$. This proof is general, it does not depend on whether the dimension of the space is finite or not. Let us state a lemma.

Lemma 4.4.1. *Let \mathcal{H} be a Hilbert space (with finite or infinite dimension). Let x_0 be a vector in $\mathbb{S}_{\mathcal{H}}$ such that*

$$|(Ax_0|x_0)| = M_0 \stackrel{\text{def}}{=} \sup_{x \in \mathbb{S}_{\mathcal{H}}} |(Ax|x)|$$

Then x_0 is an eigenvector of A , corresponding to the eigenvalue M_0 or $-M_0$.

Proof. By switching A for $-A$ if needed, we can assume that $(Ax_0|x_0) = M_0$. Observe that, for any non-zero vector y in \mathcal{H} ,

$$\left(A \left(\frac{y}{\|y\|} \right) \middle| \frac{y}{\|y\|} \right) \leq M_0,$$

which implies that

$$\forall y \in \mathcal{H}, \quad F(y) \stackrel{\text{def}}{=} M_0 \|y\|^2 - (Ay|y) \geq 0. \quad (4.9)$$

We can observe that the point x_0 satisfies $F(x_0) = 0$, and is a minimum of the function F . Therefore, the differential of F at the point x_0 is equal to zero. Alternatively, we can prove (4.9) directly, seeing that

$$\forall \lambda \in \mathbb{R}, \quad \forall h \in \mathcal{H}, \quad M_0 \|x_0 + \lambda h\|^2 - (A(x_0 + \lambda h)|x_0 + \lambda h) \geq 0$$

As $F(x_0) = 0$, by expanding the above, we get

$$\forall \lambda \in \mathbb{R}, \quad \forall h \in \mathcal{H}, \quad 2\lambda \Re(M_0 x_0 - Ax_0|h) + \lambda^2 (M_0 \|h\|^2 - A(h|h)) \geq 0.$$

As the discriminant of a non-negative polynomial is non-positive, we get that

$$\forall h \in \mathcal{H}, \quad \Re(M_0 x_0 - Ax_0|h) = 0$$

In the Hermitian setting, we need to use the above relation with ih instead of h to find that, for any h in \mathcal{H} , $(M_0 x_0 - Ax_0|ih) = 0$, which ends the proof of the lemma. \square

Back to the proof of theorem 4.4.3. We are now going to prove a very simple linear algebra result, which also holds regardless of whether the dimension of the space is finite or not.

Lemma 4.4.2. *Let A be a self-adjoint operator on a Hilbert space \mathcal{H} . When two eigenvectors of A are associated with distinct eigenvalues, they are orthogonal.*

Proof. Let u and v be two eigenvectors corresponding to distinct eigenvalues λ and λ' . We have

$$\lambda(u|v) = (Au|v) = (Av|u) = \lambda'(u|v).$$

This yields $(\lambda - \lambda')(u|v) = 0$, so the lemma is proved. \square

Back to the proof of theorem 4.4.3. We define

$$E_0^\pm \stackrel{\text{def}}{=} \ker(A \pm M_0 \text{Id}) \quad \text{and} \quad \mathcal{H}_1 \stackrel{\text{def}}{=} (E_0^+ \oplus E_0^-)^\perp = (E_0^+)^\perp \cap (E_0^-)^\perp. \quad (4.10)$$

Lemma 4.7 states, in particular, that

$$(A \pm M_0 \text{Id})(\mathcal{H}_1) \subset (E_0^\pm)^\perp \quad \text{and therefore} \quad A(\mathcal{H}_1) \subset \mathcal{H}_1 \quad (4.11)$$

We reiterate the process on \mathcal{H}_1 , and find either that there exist one or two non-zero eigenvalues $\pm M_1$, or that 0 is an eigenvalue. As the dimension of \mathcal{H} is finite, the decreasing sequence of subspaces \mathcal{H}_j is finite, and theorem 4.4.3 is proved. \square

Proof of theorem 4.4.2. We follow the same lines as above, but some complications arise due to the space having infinite dimension. Here are the three main tricky points to prove, ordered by increasing difficulty:

- the fact that the eigenspaces corresponding to non-zero eigenvalues have finite dimension;
- the fact that the sequence of eigenvalues converges to 0;
- the existence of a point x_M that attains the supremum M_0 .

For the first point, we need to prove that the unit ball of the eigenspace E_j is compact, which, by theorem 2.3.1 on page 38, ensures that the dimension of E_j is finite.

Let $(x_n)_{n \in \mathbb{N}}$ be a bounded sequence of E_j . As the operator A is compact, we can extract a subsequence $(x_{\varphi(n)})_{n \in \mathbb{N}}$ such that the sequence $(Ax_{\varphi(n)})_{n \in \mathbb{N}}$ converges. But, we have $Ax_{\varphi(n)} = \lambda_j x_{\varphi(n)}$, and λ_j is non-zero, hence the unit ball of E_j is compact.

In order to prove the second point, let us consider an infinite sequence $(\lambda_j)_{j \in \mathbb{N}}$ of distinct eigenvalues, and let $(e_j)_{j \in \mathbb{N}}$ be a sequence of vectors with norm 1, such that e_j is an eigenvector corresponding to the eigenvalue λ_j . The sequence $(e_j)_{j \in \mathbb{N}}$ is weakly convergent to 0 (the proof is left as an exercise), and we can extract a strongly convergent subsequence from this. As $Ae_j = \lambda_j e_j$, we have

$$\lim_{j \rightarrow \infty} |\lambda_j| \|e_j\| = 0.$$

As the norm of e_j is equal to 1 for every j , it is the sequence $(\lambda_j)_{j \in \mathbb{N}}$ that converges to 0.

Now we need to construct the sequence $(\lambda_j)_{j \in \mathbb{N}}$. The crucial element is the following lemma.

Lemma 4.4.3. *Let \mathcal{H} be a separable Hilbert space with infinite dimension, and let A be a compact, self-adjoint operator on \mathcal{H} . There exists a vector x_M on the unit sphere of \mathcal{H} such that*

$$|(Ax_M|x_M)| = M_0 \stackrel{\text{def}}{=} \sup_{\|x\|=1} |(Ax|x)|.$$

Moreover, $Ax_M = \pm M_0 x_M$.

Proof. If $M_0 = 0$, then by proposition 4.4.2, we have $A = 0$, and there is nothing to prove. So now let M_0 be positive. As $(Ax|x)$ is real, and switching A for $-A$ if needed, we assume that

$$M_0 = \sup_{\|x\|=1} (Ax|x).$$

By definition, there exists a sequence $(x_n)_{n \in \mathbb{N}}$ of elements of the unit ball B such that

$$\lim_{n \rightarrow \infty} (Ax_n|x_n) = M_0$$

As the operator A is compact, there exist a subsequence, still denoted $(x_n)_{n \in \mathbb{N}}$, and an element y in \mathcal{H} such that

$$\lim_{n \rightarrow \infty} Ax_n = y.$$

By using theorem 4.3.4, we can extract another subsequence and assume that there exists x in \mathcal{H} such that

$$\liminf_{n \rightarrow \infty} x_n = x.$$

We aim to prove that $Ax = y$. As the operator A is self-adjoint, we have, for any z in \mathcal{H} ,

$$\lim_{n \rightarrow \infty} (x_n|Az) = (x|Az) = (Ax|z).$$

It is also clear that

$$\lim_{n \rightarrow \infty} (Ax_n|z) = (y|z).$$

As a result, for any z in \mathcal{H} , we have $(y|z) = (Ax|z)$, which implies that $y = Ax$. So, by proposition 4.3.1,

$$\lim_{n \rightarrow \infty} (Ax_n|x_n) = (Ax|x).$$

The supremum M_0 is therefore attained at a point x which is, of course, non-zero, since M_0 is strictly positive. Moreover, if we had $\|x\| < 1$, we would have

$$\left(A \frac{x}{\|x\|} \middle| \frac{x}{\|x\|} \right) = \frac{M_0}{\|x\|^2} > M_0,$$

which contradicts the fact that M_0 is the maximum value. So $\|x\| = 1$, which means that the sequence $(x_n)_{n \in \mathbb{N}}$ converges to x in norm. \square

We can apply lemma 4.4.3 and work in the space

$$\mathcal{H}_1 \stackrel{\text{def}}{=} \left(\ker(A - M_0 \text{Id}) + \ker(A + M_0 \text{Id}) \right)^\perp = \ker(A - M_0 \text{Id})^\perp \cap \ker(A + M_0 \text{Id})^\perp.$$

We prove that \mathcal{H}_1 is invariant under A . It is a simple lemma.

Lemma 4.4.4. *Let B be an element of $\mathcal{L}(\mathcal{H})$. Then $(\ker B)^\perp$ is invariant under B^* .*

Proof. Let $x \in (\ker B)^\perp$ and $y \in \ker B$. We have

$$(y|B^*x) = (By|x) = 0,$$

hence B^*x is in the orthogonal space of $\ker B$. The lemma is proved. \square

Back to the proof of theorem 4.4.2. If we apply the above lemma with $B = A \pm M_0 \text{Id}$, we get that

$$(A \pm M_0 \text{Id})(\ker(A \pm M_0 \text{Id}))^\perp \subset (\ker(A \pm M_0 \text{Id}))^\perp,$$

because A is self-adjoint. This immediately implies that

$$A(\ker(A \pm M_0 \text{Id}))^\perp \subset (\ker(A \pm M_0 \text{Id}))^\perp,$$

hence $A\mathcal{H}_1 \subset \mathcal{H}_1$. We now study the operator A restricted to the Hilbert space \mathcal{H}_1 . The operator $A|_{\mathcal{H}_1}$ is a compact, self-adjoint operator on \mathcal{H}_1 (exercise). Set

$$M_1 \stackrel{\text{def}}{=} \sup_{\substack{\|x\|=1 \\ x \in \mathcal{H}_1}} (Ax|x).$$

By the preceding study, if $M_1 \neq 0$, this supremum is a maximum attained for at least one vector with norm 1 in $(\ker(A - M_1 \text{Id}) + \ker(A + M_1 \text{Id}))^\perp$. Thus $M_1 < M_0$. This process is reiterated as follows. Set

$$\mathcal{H}_j \stackrel{\text{def}}{=} (E_0 + \cdots + E_{j-1})^\perp \quad \text{and} \quad \sup_{\substack{\|u\|=1 \\ u \in \mathcal{H}_{j+1}}} |(Au|u)| = M_{j+1}.$$

If $M_{j+1} = 0$, proposition 4.4.2 implies that $\mathcal{H}_{j+1} = \ker A$, and the procedure ends. If M_{j+1} is positive, the procedure continues.

If the sequence $(M_j)_{j \in \mathbb{N}}$ is a sequence of positive real numbers, set

$$E^\perp = \bigcap_{j \in \mathbb{N}} \mathcal{H}_j.$$

This means that

$$\forall j \in \mathbb{N}, \quad \sup_{\substack{\|u\|=1 \\ u \in E^\perp}} |(Au|u)| \leq |\lambda_j|,$$

hence

$$\sup_{\substack{\|u\|=1 \\ u \in E^\perp}} |(Au|u)| = 0,$$

which implies, using proposition 4.4.2, that $A|_{E^\perp} = 0$. The proof of theorem 4.4.2 is complete.

□

Chapter 5

L^p spaces

Introduction

This chapter is dedicated to the study of functions from X to \mathbb{K} which are integrable with respect to a positive measure μ , defined on a σ -algebra \mathcal{B} of X , when elevated to a power p . Two fundamental examples will be looked at closely.

First, we will study the case where $X = \mathbb{N}$, the σ -algebra \mathcal{B} is the power set of \mathbb{N} , and μ is the counting measure, i.e. the measure of a subset of \mathbb{N} is the number of elements (a finite number or infinity) the subset contains. Then the space of functions which are p -th power integrable is the set $\ell^p(\mathbb{N})$, which we have already worked with in chapters 2 and 3.

The other important example is where $X = \mathbb{R}^d$ or a subset of \mathbb{R}^d , endowed with the Borel σ -algebra and the Lebesgue measure.

The first section will recall, without proof, a certain number of fundamental results from integration theory that we will constantly use in this chapter. They must imperatively be remembered.

In the second section, we prove that the spaces of (equivalence classes of) p -th power integrable functions are Banach spaces. The fact that such functions are defined modulo negligible sets adds some difficulty compared to the $\ell^p(\mathbb{N})$ spaces studied in chapter 2 (see theorem 2.1.1 on page 31 and its proof). Hölder's inequality plays an important role in the proof. In this section, we will also show a criterion for belonging to a space of p -th power integrable functions based on an inequality on a collection of weighted means: this is the important lemma 5.2.2. The notion that such a collection of weighted means is characteristic to each function is at the heart of the notion of weak derivative that we will introduce in chapter 6 and, more generally, of the concept of distributions (chapter 8).

The third section consists of the proof of the theorem stating that the space of continuous functions with compact support is dense in the space of p -th power integrable functions, when X is an open subset of \mathbb{R}^d .

The fourth section is dedicated to defining the convolution of two functions defined on \mathbb{R}^d . This is a massively used operation in analysis, and provides "explicit" approximations of p -th power integrable functions using smooth functions (i.e. infinitely differentiable) with compact support.

The fifth and final section in this chapter aims to identify the dual of the space of p -th power integrable functions as the space of p' -th power integrable functions, where p' is the Hölder conjugate of p . This has already been observed in chapter 3 (theorem 3.2.1 on page 54). We will only show the case where p is in the interval $[1, 2]$ in this chapter.

5.1 Measure theory and definition of the L^p spaces

In this section, we are going to define the L^p spaces and recall the main, basic theorems of measure theory. Let X be a set, endowed with a σ -algebra \mathcal{B} and a positive measure μ on \mathcal{B} . As is standard, the measure is assumed to be σ -finite, which means that X is the union of countably many sets with finite measure.

Definition 5.1.1. *Let p be in the interval $[1, +\infty[$. We denote $L^p(X, d\mu)$ the space of (equivalence classes modulo equality almost everywhere of) measurable functions f such that*

$$\int_X |f(x)|^p d\mu < \infty.$$

If $p = \infty$, we denote $L^p(X, d\mu)$ the space of (equivalence classes modulo equality almost everywhere of) measurable functions such that

$$\|f\|_{L^\infty} \stackrel{\text{def}}{=} \sup\{\lambda / \mu\{x / |f(x)| > \lambda\} > 0\} = \inf\{M / \mu\{x / |f(x)| > M\}\} = 0.$$

Let us prove that the two quantities in the definition of $\|f\|_{L^\infty}$ are equal. First, we note that for any pair (λ, M) such that $\mu\{x / |f(x)| > M\} = 0$ and $\mu\{x / |f(x)| > \lambda\} > 0$, we have that λ is strictly less than M . Hence,

$$\sup\{\lambda / \mu\{x / |f(x)| > \lambda\} > 0\} \leq \inf\{M / \mu\{x / |f(x)| > M\}\}.$$

Now, let M_1 be a real number strictly greater than $\sup\{\lambda / \mu\{x / |f(x)| > \lambda\} > 0\}$. By definition of the upper bound, we have $\mu\{x / |f(x)| > M_1\} = 0$. Thus,

$$\inf\{M / \mu\{x / |f(x)| > M\}\} \leq \sup\{\lambda / \mu\{x / |f(x)| > \lambda\} > 0\}.$$

We now recall the fundamental theorems of measure theory.

Theorem 5.1.1 (monotone convergence, Beppo Levi). *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of positive, measurable functions on X . We assume that, for every x , the sequence $(f_n(x))_{n \in \mathbb{N}}$ is non-decreasing. Let f be the function defined by*

$$f \begin{cases} X & \longrightarrow \mathbb{R}^+ \cup \{+\infty\} \\ x & \longmapsto \lim_{n \rightarrow \infty} f_n(x). \end{cases}$$

Then, in the set $\mathbb{R}^+ \cup \{+\infty\}$ with its usual topology, we have

$$\lim_{n \rightarrow \infty} \int_X f_n d\mu = \int_X f d\mu$$

Lemma 5.1.1 (Fatou). *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of measurable functions from X to $[0, +\infty]$. Then, we have*

$$\int_X (\liminf_{n \rightarrow \infty} f_n) d\mu \leq \liminf_{n \rightarrow \infty} \int_X f_n d\mu.$$

The link between almost-everywhere dominated convergence and convergence in the sense of the L^p norm is described in the next two theorems.

Theorem 5.1.2. Let p be a number greater than or equal to 1. Consider a sequence $(f_n)_{n \in \mathbb{N}}$ of elements of L^p , and f a function such that

$$\lim_{n \rightarrow \infty} \int_X |f(x) - f_n(x)|^p d\mu(x) = 0.$$

Then, there exists an extraction function ψ such that

$$\forall a.e d\mu(x), \lim_{n \rightarrow \infty} f_{\psi(n)}(x) = f(x).$$

Theorem 5.1.3 (dominated convergence, Lebesgue). Let p be a real number greater than or equal to 1, and let $(f_n)_{n \in \mathbb{N}}$ be a sequence of functions in $L^p(X, d\mu)$. If, for almost every x in X , we have

$$\lim_{n \rightarrow \infty} f_n(x) = f(x),$$

and if there exists a function g in $L^p(X, d\mu)$ such that, for almost every x in X ,

$$|f_n(x)| \leq g(x),$$

then

$$f \in L^p \quad \text{and} \quad \lim_{n \rightarrow \infty} \|f_n - f\|_{L^p} = 0.$$

Theorem 5.1.4 (differentiability of functions defined by Lebesgue integrals). Let (X, μ) be a measured space, Ω be an open subset of \mathbb{R}^d , and f be a measurable function from $\Omega \times X$ to \mathbb{K} . If, for almost every x in X , the function

$$z \mapsto f(z, x)$$

is differentiable on Ω , and if, for every x in X , the functions

$$x \mapsto f(z, x) \quad \text{and} \quad x \mapsto Df(z, x)$$

are integrable on X , and, finally, if for every x in X and every z in Ω , we have

$$|Df(z, x)| \leq g(x) \quad \text{with} \quad g \in L^1(X, d\mu),$$

then the function

$$z \mapsto \int_X f(z, x) d\mu(x)$$

is differentiable on Ω , and we have the formula

$$D \int_X f(z, x) d\mu(x) = \int_X Df(z, x) d\mu(x).$$

Theorem 5.1.5 (Fubini). Let (X_1, μ_1) and (X_2, μ_2) be two measured spaces, and let F be a positive, measurable function from $X_1 \times X_2$ to $[0, +\infty]$. Then,

$$\begin{aligned} \int_{X_1 \times X_2} F(x_1, x_2) d\mu_1(x_1) \otimes d\mu_2(x_2) &= \int_{X_1} d\mu_1(x_1) \int_{X_2} F(x_1, x_2) d\mu_2(x_2) \\ &= \int_{X_2} d\mu_2(x_2) \int_{X_1} F(x_1, x_2) d\mu_1(x_1). \end{aligned}$$

By applying this theorem, one can solve the following exercise.

Exercise 5.1.1. Let f be a measurable function on a measured space (X, μ) . Show that

$$\int_X |f(x)|^p d\mu(x) = p \int_0^{+\infty} \lambda^{p-1} \mu(|f| > \lambda) d\lambda.$$

Theorem 5.1.6 (Fubini-Tonelli). Let (X_1, μ_1) and (X_2, μ_2) be two measured spaces, and let F be a measurable function on $X_1 \times X_2$. The following two statements are equivalent:

- i) the function F belongs to $L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)$;
- ii) for almost every point x_1 in X_1 , the function $F(x_1, \cdot)$ belongs to $L^1(X_2, d\mu_2)$,

$$\int_{X_2} F(x_1, x_2) d\mu_2(x_2) \in L^1(X_1, d\mu_1),$$

$$\text{and } \left\| \int_{X_2} F(x_1, x_2) d\mu_2(x_2) \right\|_{L^1(X_1, d\mu_1)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

Identically, i) is equivalent to: for almost any point x_2 in X_2 , the function $F(\cdot, x_2)$ belongs to $L^1(X_1, d\mu_1)$,

$$\int_{X_1} F(x_1, x_2) d\mu_1(x_1) \in L^1(X_2, d\mu_2),$$

$$\text{and } \left\| \int_{X_1} F(x_1, x_2) d\mu_1(x_1) \right\|_{L^1(X_2, d\mu_2)} \leq \|F\|_{L^1(X_1 \times X_2, d\mu_1 \otimes d\mu_2)}.$$

Moreover, if i) is satisfied, we have the formula

$$\begin{aligned} \int_{X_2} \left(\int_{X_1} F(x_1, x_2) d\mu_1(x_1) \right) d\mu(x_2) &= \int_{X_1} \left(\int_{X_2} F(x_1, x_2) d\mu_2(x_2) \right) d\mu(x_1) \\ &= \int_{X_1 \times X_2} F(x_1, x_2) d\mu_1 \otimes d\mu_2(x_1, x_2). \end{aligned}$$

5.2 The L^p spaces are Banach spaces

The study of L^p spaces requires the introduction of the Hölder conjugate.

$$\text{If } p \in]1, \infty[, p' \stackrel{\text{def}}{=} \frac{p}{p-1}, \text{ if } p = 1, p' \stackrel{\text{def}}{=} +\infty, \text{ and if } p = +\infty, p' \stackrel{\text{def}}{=} 1.$$

The numbers p and p' are called conjugate exponents, and, accepting the convention that $\frac{1}{\infty} = 0$, we have

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

The first important new theorem of this chapter is the following.

Theorem 5.2.1. For any p in $[1, \infty]$, $(L^p(X, d\mu), \|\cdot\|_{L^p})$ is a Banach space.

Proof. We start with the case where p is infinity. First of all, we note that if $\|f\|_{L^\infty} = 0$, there exists a sequence $(\lambda_n)_{n \in \mathbb{N}}$ which converges to 0 such that, for any n , the set $\{x / |f(x)| > \lambda_n\}$ is negligible. As the union of countably many negligible sets is also negligible, the set of points x such that $f(x)$ is non-zero is negligible. Thus f is equal to 0 as an element of L^∞ . The proof of the homogeneity of the L^∞ norm is left as an exercise, and uses the fact that

$$\{x / |\lambda f(x)| > M\} = \left\{x / |f(x)| > \frac{M}{|\lambda|}\right\}.$$

Now, let us consider f and g two (equivalence classes of) functions in L^∞ , and set two positive numbers M_f and M_g such that the measures of the sets $\{x / |f(x)| > M_f\}$ and $\{x / |g(x)| > M_g\}$ are equal to zero. Note that

$$\{x / |f(x) + g(x)| > M_f + M_g\} \subset \{x / |f(x)| > M_f\} \cup \{x / |g(x)| > M_g\},$$

so the space L^∞ is stable under addition. Moreover, as the infimum of a set is a lower bound of the set, we have $\|f + g\|_{L^\infty} \leq M_f + M_g$. As the infimum is the greatest lower bound, we have $\|f + g\|_{L^\infty} \leq \|f\|_{L^\infty} + \|g\|_{L^\infty}$, thus we conclude that $(L^\infty(X, d\mu), \|\cdot\|_{L^\infty})$ is a normed vector space.

In order to prove that $(L^\infty, \|\cdot\|_{L^\infty})$ is a Banach space, we are going to apply proposition 2.1.3 on page 29. Let us consider a Cauchy sequence $(f_n)_{n \in \mathbb{N}}$ in $(L^\infty, \|\cdot\|_{L^\infty})$, and define the set $E_{m,p,n}$ as follows:

$$E_{m,p,n} \stackrel{\text{def}}{=} \left\{ x \in X / |f_m(x) - f_{m+p}(x)| > \frac{1}{n+1} \right\}.$$

Here, the equivalence classes of the functions f_n and representatives of these classes are denoted the same way. As the sequence $(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, we can define a function ϕ from \mathbb{N} to \mathbb{N} by

$$\phi(n) = \min \left\{ m \geq \phi(n-1) + 1 / \forall m' \geq m, \forall p \geq 0, \|f_{m'} - f_{m'+p}\|_{L^\infty} < \frac{1}{n+1} \right\}.$$

By definition of the essential supremum, for any $m \geq \phi(n)$, the measure of the set $E_{m,p,n}$ is zero. Thus, the set E defined as

$$E \stackrel{\text{def}}{=} \bigcup_{\substack{(n,p) \in \mathbb{N}^2 \\ m \geq \phi(n)}} E_{m,p,n}$$

has zero measure. We can therefore choose representatives of the (classes of) functions f_n that are equal to zero on the negligible set E (these representatives are still denoted f_n). Thus,

$$\forall n, \forall m \geq \phi(n), \forall p, \forall x \in X, |f_m(x) - f_{m+p}(x)| < \frac{1}{n+1}$$

and the sequence $(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{B}(X, \mathbb{K})$. We can now use proposition 2.1.3 on page 29 to conclude that $L^\infty(X, d\mu)$ is complete.

Now, we study the case where p is finite. The first step is to prove that $(L^p(X, d\mu), \|\cdot\|_{L^p})$ is a normed vector space, where

$$\|f\|_{L^p} \stackrel{\text{def}}{=} \left(\int_X |f(x)|^p d\mu(x) \right)^{\frac{1}{p}}.$$

This is clear for $p = 1$. When p belongs to $]1, \infty[$, we first note that, since

$$|f(x) + g(x)|^p \leq 2^p (|f(x)|^p + |g(x)|^p),$$

$L^p(X, d\mu)$ is a vector space. Let us show that it is a normed space. This relies on Hölder's inequality, which we state and prove.

Proposition 5.2.1 (Hölder's inequality). *Let (X, μ) be a measured space, f be a function in $L^p(X, d\mu)$ and g be a function in $L^{p'}(X, d\mu)$. Then, the function fg is in $L^1(X, d\mu)$, and we have*

$$\int_X |f(x)g(x)| d\mu(x) \leq \|f\|_{L^p} \|g\|_{L^{p'}}.$$

Proof. First, note that if $p = 1$ or $p = \infty$, there is nothing to prove. Moreover, we can assume that $\|f\|_{L^p} = \|g\|_{L^{p'}} = 1$, and that the functions f and g are non-negative. Using the convexity inequality (2.3) on page 29, we have

$$f(x)g(x) = (f(x)^p)^{\frac{1}{p}}(g(x)^{p'})^{\frac{1}{p'}} \leq \frac{1}{p}f(x)^p + \frac{1}{p'}g(x)^{p'},$$

which proves Hölder's inequality. Indeed, we integrate the above, and we get

$$\begin{aligned} \int_X |fg| d\mu &\leq \frac{1}{p} \int_X |f|^p d\mu + \frac{1}{p'} \int_X |g|^{p'} d\mu \\ &\leq \frac{1}{p} + \frac{1}{p'} = 1. \end{aligned}$$

This ends the proof of the proposition. \square

Back to the proof of theorem 5.2.1. As $f + g$ belongs to L^p , Hölder's inequality implies that

$$\begin{aligned} \int_X |f + g|^p d\mu &\leq \int_X |f||f + g|^{p-1} d\mu + \int_X |g||f + g|^{p-1} d\mu \\ &\leq \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} \left(\int_X |f + g|^{(p-1)\frac{p}{p-1}} d\mu \right)^{\frac{1}{p'}} \\ &\quad + \left(\int_X |g|^p d\mu \right)^{\frac{1}{p}} \left(\int_X |f + g|^{(p-1)\frac{p}{p-1}} d\mu \right)^{\frac{1}{p'}} \\ &\leq \left(\left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} + \left(\int_X |g|^p d\mu \right)^{\frac{1}{p}} \right) \left(\int_X |f + g|^{(p-1)\frac{p}{p-1}} d\mu \right)^{1-\frac{1}{p}}. \end{aligned}$$

This implies that $\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$, and therefore that $L^p(X, d\mu)$ is a normed space.

Now we will show that these spaces are complete. The proof relies on the following lemma.

Lemma 5.2.1. *If E is a normed space such that, for any sequence $(x_n)_{n \in \mathbb{N}}$ of elements of E , we have*

$$\sum_{n \in \mathbb{N}} \|x_n\|_E < \infty \implies S_N \stackrel{\text{def}}{=} \sum_{n=0}^N x_n \text{ is convergent,}$$

then E is a Banach space.

Proof. Recall proposition 1.3.2 on page 19, which states that a Cauchy sequence that has a cluster point is convergent. Let $(a_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in E : we will extract a convergent subsequence, and this will yield the result. Define a map ϕ from \mathbb{N} to \mathbb{N} by $\phi(0) = 0$ and

$$\phi(n+1) \stackrel{\text{def}}{=} \min \left\{ m \geq \phi(n) + 1, \sup_{p \geq 0} \|a_m - a_{m+p}\| \leq \frac{1}{(2+n)^2} \right\}.$$

Thus, for any integer n , we have

$$\|a_{\phi(n+1)} - a_{\phi(n)}\| \leq \frac{1}{(1+n)^2}.$$

Set $x_n \stackrel{\text{def}}{=} a_{\phi(n+1)} - a_{\phi(n)}$. We have

$$\sum_{n \in \mathbb{N}} \|x_n\| \leq \sum_{n \in \mathbb{N}} \frac{1}{(1+n)^2} < \infty.$$

So, by hypothesis, the sequence $\sum_{n=0}^N x_n = a_{\phi(N)} - a_0$ is convergent. Proposition 1.3.2 on page 19 states that any Cauchy sequence that has a cluster point is convergent, so the lemma is proved. \square

Back to the proof of theorem 5.2.1. Let $(f_n)_{n \in \mathbb{N}}$ be a sequence of elements of L^p such that

$$\sum_{n \in \mathbb{N}} \|f_n\|_{L^p} < \infty.$$

Define the functions

$$S_N(x) = \sum_{n=0}^N f_n(x) \quad \text{and} \quad S_N^+(x) = \sum_{n=0}^N |f_n(x)|.$$

For any integer N , we have

$$\|S_N^+\|_{L^p} \leq \sum_{n \in \mathbb{N}} \|f_n\|_{L^p},$$

which means that

$$\forall n \in \mathbb{N}, \quad \int_X S_N^+(x)^p d\mu(x) \leq \left(\sum_{n \in \mathbb{N}} \|f_n\|_{L^p} \right)^p.$$

So the sequence $(S_N^+(x))_{N \in \mathbb{N}}$ is non-decreasing. The monotone convergence theorem 5.1.1 yields that

$$\int_X S^+(x)^p d\mu(x) \leq \left(\sum_{n \in \mathbb{N}} \|f_n\|_{L^p} \right)^p \quad \text{with} \quad S^+(x) = \sum_{n \in \mathbb{N}} |f_n(x)|.$$

The function S^+ takes finite values almost everywhere, and belongs to L^p . Thus, for almost any x , the series with general term $f_n(x)$ converges in \mathbb{C} . We denote S its sum; S is an element of L^p , and we have

$$\left\| S - \sum_{n=0}^N f_n \right\|_{L^p}^p = \int_X \left| S(x) - \sum_{n=0}^N f_n(x) \right|^p d\mu(x).$$

We know that

$$\left| S(x) - \sum_{n=0}^N f_n(x) \right| \leq 2S^+(x).$$

As S^+ is in L^p , the dominated convergence theorem yields that

$$\lim_{N \rightarrow \infty} \left\| S - \sum_{n=0}^N f_n \right\|_{L^p}^p = 0.$$

The theorem is proved. \square

Corollary 5.2.1. *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence that converges to f in $L^1(X, d\mu)$. Then there exists an extraction function ϕ such that*

$$\forall \text{a.e. } x \in X, \quad \lim_{n \rightarrow \infty} f_{\phi(n)}(x) = f(x).$$

We are now going to make a few remarks on the Hölder inequality. We begin with the following corollary.

Corollary 5.2.2. *Let p and q be two real numbers strictly greater than 1, such that*

$$\frac{1}{p} + \frac{1}{q} \leq 1.$$

Then the mapping

$$\begin{cases} L^p \times L^q & \longrightarrow L^r \\ (f, g) & \longmapsto fg \end{cases}$$

is bilinear and continuous if

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q}.$$

Proof. It suffices to apply Hölder's inequality with $s = r/p$. This yields

$$\int_X |f(x)g(x)|^r dx \leq \|f\|_{L^p}^r \|g\|_{L^q}^r,$$

which means that

$$\|fg\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q},$$

and the corollary is proved. □

Corollary 5.2.3. *If X has finite measure, then $L^p(X, d\mu) \subset L^q(X, d\mu)$ if $p \geq q$.*

Proof. As the function $\mathbf{1}$ is clearly in the space $L^\infty(X, d\mu) \cap L^1(X, d\mu)$, it belongs to $L^p(X, d\mu)$ for any p . We even have

$$\forall p \in [1, \infty], \quad \|\mathbf{1}\|_{L^p} = \mu(X)^{\frac{1}{p}}.$$

Corollary 5.2.2 implies that

$$\|f\|_{L^q} \leq \mu(X)^{\frac{1}{s}} \|f\|_{L^p} \quad \text{with} \quad \frac{1}{s} = \frac{1}{q} - \frac{1}{p},$$

so the corollary is proved. □

We stress again that Hölder's inequality is fundamental. It is also optimal in the following sense.

Lemma 5.2.2. *Let (X, μ) be a measured space. Let f be a measurable function and p be an element of $[1, \infty]$. Assume that*

$$\sup_{\substack{\|g\|_{L^{p'}} \leq 1 \\ g \geq 0}} \int_X |f(x)g(x)| d\mu(x) < +\infty. \tag{5.1}$$

Then f belongs to L^p , and

$$\|f\|_{L^p} = \sup_{\|g\|_{L^{p'}} \leq 1} \left| \int_X f(x)g(x) d\mu(x) \right|.$$

Proof. We begin with the case $p = 1$. By choosing the constant function equal to 1 for g , we have g bounded with an L^∞ norm equal to 1, and, using inequality (5.1), we get that

$$\int_X |f(x)|g(x)d\mu(x) < \infty.$$

Now, consider g the function defined by

$$g(x) = \frac{\overline{f}(x)}{|f(x)|} \quad \text{if } f(x) \neq 0 \quad \text{and} \quad 0 \text{ otherwise.}$$

The function g is bounded and its norm is equal to 1. Moreover,

$$\int_X f(x)g(x)d\mu(x) = \int_X |f(x)|d\mu(x).$$

The lemma is proved in this case.

Now, let p be a real number strictly greater than 1, and consider an increasing sequence of sets with finite measure $(K_n)_{n \in \mathbb{N}}$, such that the union of these sets is equal to X . Set

$$f_n^+(x) = \mathbf{1}_{K_n \cap (|f| \leq n)} |f| \quad \text{and} \quad g_n(x) = \frac{f_n^+(x)^{p-1}}{\|f_n^+\|_{L^p}^{\frac{p}{p'}}}.$$

It is clear that the function f_n is non-negative, and belongs to $L^1 \cap L^\infty$, so it belongs to L^q for any q , and we have

$$\|g_n\|_{L^{p'}}^{p'} = \frac{1}{\|f_n^+\|_{L^p}^p} \int_X f_n^+(x)^{(p-1)\frac{p}{p-1}} d\mu(x) = 1.$$

By definition of functions f_n and g_n , we have

$$\begin{aligned} \int_X |f(x)| \mathbf{1}_{K_n \cap (|f| \leq n)} g_n(x) d\mu(x) &= \int_X f_n^+(x) g_n(x) d\mu(x) \\ &= \left(\int_X f_n^+(x)^p d\mu(x) \right) \|f_n^+\|_{L^p}^{-\frac{p}{p'}} \\ &= \|f_n^+\|_{L^p}^{p-\frac{p}{p'}} \\ &= \|f_n^+\|_{L^p}. \end{aligned}$$

Therefore,

$$\int_X f_n^+(x)^p d\mu(x) \leq \left(\sup_{\substack{\|g\|_{L^{p'}} \leq 1 \\ g \geq 0}} \int_X |f(x)|g(x)d\mu(x) \right)^p.$$

The monotone convergence theorem applied to the non-decreasing sequence $((f_n^+)^p)_{n \in \mathbb{N}}$ immediately yields that

$$\|f\|_{L^p} \leq \sup_{\substack{\|g\|_{L^{p'}} \leq 1 \\ g \geq 0}} \int_X f(x)g(x)d\mu(x).$$

Thus, if hypothesis (5.1) holds, the function f belongs to L^p . Now, we assume that f belongs to L^p . Then, if we set

$$g(x) = \frac{\overline{f}(x)|f(x)|^{p-1}}{|f(x)| \times \|f\|_{L^p}^{\frac{p}{p'}}},$$

we have

$$\|g\|_{L^{p'}}^{p'} = \frac{1}{\|f\|_{L^p}^p} \int_X |f(x)|^{(p-1)\frac{p}{p-1}} d\mu(x) = 1 \quad \text{and} \quad \|f\|_{L^p} = \int_X f(x)g(x)d\mu(x).$$

So the lemma is proved for every p in $[1, \infty[$. In the case where $p = +\infty$, consider a positive number λ such that $\mu(|f| \geq \lambda)$ is positive. Set $E_\lambda \stackrel{\text{def}}{=} (|f| \geq \lambda)$, and let g_0 be a non-negative function in L^1 , with support in E_λ , and with integral equal to 1. Then, if

$$g(x) = \frac{\bar{f}(x)}{|f(x)|} g_0,$$

we get

$$\int_X |f(x)g(x)|d\mu(x) = \int_X |f(x)|g_0(x)d\mu(x) \geq \lambda \int_X g_0d\mu(x) \geq \lambda.$$

The lemma is proved. □

5.3 Density in L^p spaces

Throughout this chapter, Ω will denote an open set of \mathbb{R}^d , endowed with the usual Euclidean metric, and μ will be a non-negative measure such that the measure of every compact set is finite. Let us present the main result of this section.

We begin with a digression on the topology of open sets of \mathbb{R}^d .

Theorem 5.3.1 (characterisation of compact subsets of an open set of \mathbb{R}^d). *Let A be a closed subset of an open set Ω of \mathbb{R}^d , such that there exists a positive number r such that $A \subset [-r, r]^d$. Then*

$$A \text{ is compact} \iff \inf_{x \in A} d(x, \mathbb{R}^d \setminus \Omega) > 0.$$

Proof. We start by assuming that A is compact. As Ω is open, then, for any x in A , $d(x, \mathbb{R}^d \setminus \Omega)$ is strictly positive. By exercise 1.1.4 on page 13, we know that the function that maps a point to its distance to a given set is continuous. Corollary 1.3.2 states that its infimum is attained, so it is strictly positive.

Conversely, let A be a closed set in Ω such that there exists a positive number r such that $A \subset [-r, r]^d$, and that satisfies

$$\delta \stackrel{\text{def}}{=} \inf_{x \in A} d(x, X \setminus \Omega) > 0.$$

In order to prove that A is compact, it suffices to prove that A is closed in \mathbb{R}^d . As A is closed in Ω , this means that there exists a set B that is closed in \mathbb{R}^d such that $A = B \cap \Omega$. Let $\Omega_{\delta/2}$ be the closed set defined by

$$\Omega_{\delta/2} \stackrel{\text{def}}{=} \{x / d(x, X \setminus \Omega) \geq \delta/2\}.$$

We have the following relations between sets:

$$A = A \cap \Omega_{\delta/2} = B \cap \Omega \cap \Omega_{\delta/2} = B \cap \Omega_{\delta/2}.$$

The set $B \cap \Omega_{\delta/2}$ is closed in \mathbb{R}^d as the intersection of two closed sets, so A is closed in \mathbb{R}^d . The theorem is proved. □

The following result will also be useful to us.

Theorem 5.3.2. *Let Ω be an open set of \mathbb{R}^d . There exists a sequence of compact subsets $(K_n)_{n \in \mathbb{N}}$ such that*

$$\bigcup_{n \in \mathbb{N}} K_n = \Omega, \quad K_n \subset K_{n+1}^\circ \quad \text{and} \quad \forall K \text{ compact subset of } \Omega, \exists n / K \subset K_n^\circ.$$

Proof. Set

$$K_n \stackrel{\text{def}}{=} B(0, n) \cap \left\{ x \in \mathbb{R}^d / d(x, \mathbb{R}^d \setminus \Omega) \geq \frac{1}{n} \right\}.$$

As Ω is open, for each x in Ω , $d(x, \mathbb{R}^d \setminus \Omega)$ is positive. So, there exists an integer n such that $x \in K_n$. Moreover,

$$K_n \subset \overset{\circ}{B}(0, n+1) \cap \left\{ x \in \mathbb{R}^d / d(x, \mathbb{R}^d \setminus \Omega) > \frac{1}{n} \right\} \subset K_{n+1}^\circ.$$

Hence the first point of the theorem. To prove the second one, we notice that $\bigcup_{n \in \mathbb{N}} K_{n+1}^\circ = \Omega$,

since $K_n \subset K_{n+1}^\circ$. This proves the theorem. \square

Definition 5.3.1. *Let Ω be an open set of \mathbb{R}^d . We call an exhaustion of Ω by compact sets a sequence of compact subsets of Ω which satisfy the results of theorem 5.3.2 above.*

We now return to the L^p spaces.

Theorem 5.3.3. *Let p be a number in $[1, +\infty[$. The space $C_c(\Omega)$, containing all continuous functions with compact support in Ω , is dense in $L^p(\Omega, d\mu)$.*

The proof of this theorem is long and intricate. We show it here for the reader's culture.

To prove this theorem, we will first show two density results which hold for functions defined on a generic measured space. We start with the following proposition.

Proposition 5.3.1. *If p belongs to $[1, \infty[$, then $L^1(X, d\mu) \cap L^\infty(X, d\mu)$ is dense in $L^p(X, d\mu)$.*

Proof. Consider a non-decreasing sequence $(K_n)_{n \in \mathbb{N}}$ of sets with finite measure, the union of which is equal to X . Set

$$f_n = \mathbf{1}_{K_n \cap \{|f| \leq n\}} f.$$

It is clear that, for almost every x , we have

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Moreover, we have the inequality

$$|f_n(x) - f(x)|^p \leq 2^p |f(x)|^p.$$

The dominated convergence theorem then ensures that the result holds. \square

Now, we prove a second density result.

Proposition 5.3.2. *If p belongs to $[1, \infty[$, then the set of integrable simple functions (i.e. which only take a finite number of non-zero values on sets with finite measure) is dense in $L^p(X, d\mu)$.*

Proof. We will use the following fact, which the reader can prove as an exercise: if (Y, d) is a metric space, if A is a dense subset of Y , and if every element of A is the limit of a sequence of elements of B , then B is also dense in Y . Courtesy of this, it suffices to show that every function in $L^1(X, d\mu) \cap L^\infty(X, d\mu)$ is the limit of a sequence of integrable simple functions. To do so, for a given function f in $L^1(X, d\mu) \cap L^\infty(X, d\mu)$, consider the sequence of functions $(f_n)_{n \in \mathbb{N}}$ defined by

$$f_n \stackrel{\text{def}}{=} \sum_{j=0}^{+\infty} \frac{j}{n+1} \mathbf{1}_{f^{-1}([\frac{j}{n+1}, \frac{j+1}{n+1}[})} + \sum_{j=-\infty}^{-1} \frac{j+1}{n+1} \mathbf{1}_{f^{-1}([\frac{j}{n+1}, \frac{j+1}{n+1}[}). \quad (5.2)$$

It is obvious that

$$\forall x \in X, |f_n(x) - f(x)| \leq \frac{1}{n+1} \quad \text{and} \quad |f_n(x)| \leq |f(x)|.$$

The dominated convergence theorem ensures that $f_n \rightarrow f$ in L^p , which means that there exists n_0 such that, for n greater than n_0 ,

$$\|f_n - f\|_{L^p} < \frac{\varepsilon}{2}. \quad (5.3)$$

Moreover, as the function f is essentially bounded, the sums that appear in (5.2) are finite, since there exists a real number λ that satisfies

$$\frac{|j|}{n+1} \geq \lambda \implies \mu\left(f^{-1}\left(\left[\frac{j}{n+1}, \frac{j+1}{n+1}\right]\right)\right) = 0.$$

Hence each function f_n only takes a finite number of values. Also, as the function f is integrable, the sets $f^{-1}\left(\left[\frac{j}{n+1}, \frac{j+1}{n+1}\right]\right)$ have finite measure, which ends the proof of the proposition. \square

Proof of theorem 5.3.3. Let f be a function in $L^p(\Omega)$. We must prove that, for every positive ε , there exists a continuous function g with compact support such that

$$\|f - g\|_{L^p} < \varepsilon. \quad (5.4)$$

By proposition 5.3.2, there exists an integrable simple function \tilde{f} such that

$$\|f - \tilde{f}\|_{L^p} < \frac{\varepsilon}{2} \quad \text{with} \quad \tilde{f} = \sum_{j=1}^N \alpha_j \mathbf{1}_{A_j}, \quad (5.5)$$

where the α_j are real numbers, and the A_j are Borel sets with finite measure. We will accept the following result at first.

Lemma 5.3.1. *For any p in $[1, \infty[$, for any Borel set A with finite measure, and for any positive number ε , there exists a continuous function h with compact support such that*

$$\|h - \mathbf{1}_A\|_{L^p} < \varepsilon.$$

This allows us to state that, for each j in $\{1, \dots, N\}$, there exists a continuous function g_j with compact support in Ω such that

$$\|\mathbf{1}_{A_j} - g_j\|_{L^p} \leq \frac{\varepsilon}{2N|\alpha_j|}.$$

Now define the function g as follows: $g \stackrel{\text{def}}{=} \sum_{j=1}^N \alpha_j g_j$. We have

$$\begin{aligned} \|\tilde{f} - g\|_{L^p} &\leq \sum_{j=1}^N |\alpha_j| \|\mathbf{1}_{A_j} - g_j\|_{L^p} \\ &\leq \frac{\varepsilon}{2}. \end{aligned}$$

Theorem 5.3.3 is therefore proved, providing we prove lemma 5.3.1. \square

Proof of lemma 5.3.1. Let $(K_n)_{n \in \mathbb{N}}$ be an exhaustion of Ω by compact sets (see theorem 5.3.2 and definition 5.3.1 on page 91). As A has finite measure, the characteristic function of A belongs to L^p for every p . Moreover, for every x in Ω , we have

$$\lim_{n \rightarrow \infty} \mathbf{1}_{K_n \cap A}(x) = \mathbf{1}_A(x) \quad \text{and} \quad 0 \leq \mathbf{1}_{K_n \cap A}(x) \leq \mathbf{1}_A(x).$$

So, for every positive ε , there exists n_0 such that

$$\|\mathbf{1}_{K_n \cap A} - \mathbf{1}_A\|_{L^p} < \frac{\varepsilon}{2}. \quad (5.6)$$

We now need a fundamental result from measure theory, called the regularity theorem for the Borel measure on \mathbb{R}^d . We will use it without proof for now.

Theorem 5.3.4. *For every set A belonging to the complete Borel σ -algebra \mathcal{B} with finite measure,*

$$\forall \varepsilon > 0, \exists \text{ a compact set } K, \exists \text{ an open set } U / K \subset A \subset U \quad \text{and} \quad \mu(U \setminus K) < \varepsilon. \quad (5.7)$$

End of the proof of lemma 5.3.1. Thus there exists an open set U and a compact set K such that

$$K \subset K_n \cap A \subset U \quad \text{and} \quad \mu(U \setminus K) < \left(\frac{\varepsilon}{2}\right)^p.$$

We can assume that the closure of U is compact, for example by replacing the open set U by $U \cap K_n + B(0, \varepsilon)$, which will still be denoted U . Set

$$h(x) \stackrel{\text{def}}{=} \frac{d(x, U^c)}{d(x, U^c) + d(x, K)}.$$

The function h is continuous, takes its values in $[0, 1]$, and satisfies

$$h|_{X \setminus U} = 0 \quad \text{and} \quad h|_K = 1.$$

We easily deduce that

$$|h - \mathbf{1}_A| \leq \mathbf{1}_{U \setminus K} \quad \text{and therefore} \quad |h - \mathbf{1}_A|^p \leq \mathbf{1}_{U \setminus K}.$$

Hence,

$$\|h - \mathbf{1}_A\|_{L^p} < \frac{\varepsilon}{2}.$$

The lemma is now proved, because the function g is zero outside of the closure of U , which we know to be compact. \square

Proof of theorem 5.3.4. We present this proof for completeness and for the reader's culture. The first set consists of reducing the problem to the case where the open set Ω has finite measure. To do so, let us assume the theorem proved in that setting. Consider a subset A of the complete σ -algebra of Ω with finite measure. We then define the sequences

$$\Omega_p \stackrel{\text{def}}{=} \Omega \cap B(0, p), \quad A_p \stackrel{\text{def}}{=} A \cap B(0, p), \quad B_0 \stackrel{\text{def}}{=} A_0 \quad \text{and} \quad B_p = A_p \setminus A_{p-1} \quad \text{for } p \geq 1.$$

The sets B_p satisfy

$$\bigcup_{p'=0}^p B_{p'} = \bigcup_{p'=0}^p A_{p'}.$$

As the sets B_p are pairwise distinct, we have

$$\sum_{p=0}^{\infty} \mu(B_p) = \mu(A) < \infty. \quad (5.8)$$

Assertion (5.7) implies that

$$\forall \varepsilon > 0, \forall p, \exists K_{p,\varepsilon} \text{ compact}, \exists U_{p,\varepsilon} \text{ open in } \Omega / K_{p,\varepsilon} \subset A_p \subset U_{p,\varepsilon} \text{ and } \mu(U_{p,\varepsilon} \setminus K_{p,\varepsilon}) < \varepsilon 2^{-p-2}.$$

Assertion (5.8), meanwhile, implies that the series with general term $(\mu(K_{p,\varepsilon}))_{p \in \mathbb{N}}$ is convergent. So there exists an integer p_ε such that

$$\sum_{p=p_\varepsilon+1}^{\infty} \mu(B_p) < \frac{\varepsilon}{2}. \quad (5.9)$$

We then set

$$U_\varepsilon \stackrel{\text{def}}{=} \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \quad \text{and} \quad K_\varepsilon \stackrel{\text{def}}{=} \bigcup_{p=0}^{p_\varepsilon} K_{p,\varepsilon}$$

We have

$$A = \bigcup_{p=0}^{\infty} B_p \subset U_\varepsilon \quad \text{and} \quad K_\varepsilon \subset \bigcup_{p=0}^{p_\varepsilon} B_p \subset A.$$

Moreover, U_ε is an open subset of Ω as it is a union of open sets, and K_ε is compact as it is a finite union of compact sets. Finally, we have

$$\begin{aligned} U_\varepsilon \setminus K_\varepsilon &= \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \left(\bigcup_{p=0}^{p_\varepsilon} K_{p,\varepsilon} \right)^c \\ &\subset \bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{p_\varepsilon} K_{p,\varepsilon}^c \\ &\subset \left(\bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{\infty} K_{p,\varepsilon}^c \right) \cup \bigcup_{p=p_\varepsilon+1}^{\infty} K_{p,\varepsilon}. \end{aligned}$$

Since

$$\bigcup_{p=0}^{\infty} U_{p,\varepsilon} \cap \bigcap_{p=0}^{\infty} K_{p,\varepsilon}^c \subset \bigcup_{p=0}^{\infty} (U_{p,\varepsilon} \setminus K_{p,\varepsilon}),$$

we deduce that

$$U_\varepsilon \setminus K_\varepsilon \subset \left(\bigcup_{p=0}^{\infty} (U_{p,\varepsilon} \setminus K_{p,\varepsilon}) \right) \cup \bigcup_{p=p_\varepsilon+1}^{\infty} K_{p,\varepsilon}.$$

As $K_{p,\varepsilon}$ is a subset of B_p , inequality (5.12) ensures that

$$\begin{aligned} \mu(U_\varepsilon \setminus K_\varepsilon) &\leq \sum_{p=0}^{\infty} \mu(U_{p,\varepsilon} \setminus K_{p,\varepsilon}) + \sum_{p=p_\varepsilon+1}^{\infty} \mu(K_{p,\varepsilon}) \\ &\leq \sum_{p=0}^{\infty} \varepsilon 2^{-p-2} + \frac{\varepsilon}{2} \\ &\leq \varepsilon. \end{aligned}$$

We must now prove the theorem in the case where Ω has finite measure. Let \mathcal{A} be the set of subsets of Ω which satisfy (5.7). We will prove that all open subsets belong to \mathcal{A} , then that \mathcal{A} is a σ -algebra, which is enough to yield the result.

If $A = U$ is open, then the (non-decreasing) sequence of compact sets $(K_n)_{n \in \mathbb{N}}$ defined by

$$K_N \stackrel{\text{def}}{=} \{x \in U / \|x\| \leq n \text{ and } d(x, U^c) \geq 2^{-n}\}$$

satisfies

$$\forall x \in U, \lim_{n \rightarrow \infty} \mathbf{1}_{K_n}(x) = \mathbf{1}_U(x).$$

The monotone convergence theorem 5.1.1 implies that

$$\lim_{n \rightarrow \infty} \mu(U \setminus K_n) = 0,$$

which ensures that U is an element of \mathcal{A} . Now, let us consider an element A in \mathcal{A} , ε a positive number, U an open set, and K a compact set such that

$$\mu(U \setminus K) < \frac{\varepsilon}{2}.$$

If C is a subset of B , we have the relation $B \setminus C = B \cap C^c$, and therefore $C^c \setminus B^c = B \setminus C$. Hence,

$$U^c \subset A^c \subset K^c \quad \text{and} \quad \mu(K^c \setminus U^c) = \mu(U \setminus K) < \frac{\varepsilon}{2}.$$

However, U^c may not be compact. To circumvent this difficulty, we introduce an exhaustion of U^c by compact sets, that is a sequence K_n satisfying theorem 5.3.2 on page 91. It is clear that

$$\forall x \in \Omega, \mathbf{1}_{K^c \setminus (U^c \cap K_n)}(x) = \mathbf{1}_{K^c}(x) - \mathbf{1}_{U^c \cap K_n}(x) = \mathbf{1}_{K^c}(x) - \mathbf{1}_{U^c}(x)$$

The monotone convergence theorem yields that A^c belongs to \mathcal{A} .

It remains to prove that, if $(A_n)_{n \in \mathbb{N}}$ is a sequence of elements of \mathcal{A} , then the union of all the A_n belongs to \mathcal{A} . We start by proving the result for the union of a finite number of these. Let A_1 and A_2 be two elements of \mathcal{A} . For any positive real number ε , there exist two open sets U_1 and U_2 , and two compact sets K_1 and K_2 , such that

$$K_j \subset A_j \subset U_j \quad \text{and} \quad \mu(U_j \setminus K_j) < \frac{\varepsilon}{2}.$$

$$K_1 \cup K_2 \subset A_1 \cup A_2 \subset U_1 \cup U_2. \tag{5.10}$$

The set $U_1 \cup U_2$ is open, and the set $K_1 \cup K_2$ is compact. Moreover, we have

$$\begin{aligned} (U_1 \cup U_2) \setminus (K_1 \cup K_2) &= (U_1 \cup U_2) \cap (K_1 \cup K_2)^c \\ &= (U_1 \cap (K_1 \cup K_2)^c) \cup (U_2 \cap (K_1 \cup K_2)^c) \\ &\subset (U_1 \cap K_1^c) \cup (U_2 \cap K_2^c). \end{aligned}$$

Hence

$$\begin{aligned} \mu((U_1 \cup U_2) \setminus (K_1 \cup K_2)) &\leq \mu(U_1 \setminus K_1) + \mu(U_2 \setminus K_2) \\ &\leq \varepsilon. \end{aligned}$$

Using (5.8), we see that $U_1 \cup U_2$ belongs to \mathcal{A} . Also, as \mathcal{A} is stable under taking the complements, if A_1 and A_2 belong to \mathcal{A} , then

$$A_1 \cap A_2 = (A_1^c \cup A_2^c)^c \in \mathcal{A}.$$

So \mathcal{A} is stable under finite union.

Now, let $(A_n)_{n \in \mathbb{N}}$ be a sequence of elements of \mathcal{A} . We follow a similar technique to the one we used to reduce the proof to the case where Ω has finite measure. We define

$$B_n \stackrel{\text{def}}{=} \bigcup_{k=0}^n A_k \setminus \bigcup_{j=0}^{k-1} A_j.$$

The sets B_n are pairwise distinct, their union is equal to that of the A_n , and we have

$$\sum_{n \in \mathbb{N}} \mu(B_n) = \mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) < \infty.$$

As each set B_n belongs to \mathcal{A} , there exists a sequence of open sets $(U_n)_{n \in \mathbb{N}}$ and a sequence of compact sets $(K_n)_{n \in \mathbb{N}}$ such that

$$\forall n \in \mathbb{N}, K_{n,\varepsilon} \subset B_n \subset U_{n,\varepsilon} \quad \text{and} \quad \mu(U_{n,\varepsilon} \setminus K_{n,\varepsilon}) < \varepsilon 2^{-n-2}. \quad (5.11)$$

There exists an integer n_ε such that

$$\sum_{n=n_\varepsilon+1}^{\infty} \mu(B_n) \leq \frac{\varepsilon}{2}. \quad (5.12)$$

We then set

$$U_\varepsilon \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} U_{n,\varepsilon} \quad \text{and} \quad K_\varepsilon \stackrel{\text{def}}{=} \bigcup_{n=0}^{n_\varepsilon} K_{n,\varepsilon}$$

The set U_ε is open, as it is the union of a sequence of open sets, and K_ε is compact as it is the union of a finite sequence of compact sets. It is then obvious that

$$K_\varepsilon \subset \bigcup_{n \in \mathbb{N}} A_n \subset U_\varepsilon. \quad (5.13)$$

Moreover,

$$K_\varepsilon^c = \left(\bigcup_{n=0}^{n_\varepsilon} K_{n,\varepsilon} \right)^c = \bigcap_{n=0}^{n_\varepsilon} K_{n,\varepsilon}^c.$$

Thus, we can deduce that

$$\begin{aligned} U_\varepsilon \setminus K_\varepsilon &= \left(\bigcup_{n \in \mathbb{N}} U_{n,\varepsilon} \right) \cap \left(\bigcap_{n=0}^{n_\varepsilon} K_{n,\varepsilon}^c \right) \\ &\subset \left(\left(\bigcup_{n \in \mathbb{N}} U_{n,\varepsilon} \right) \cap \left(\bigcap_{n \in \mathbb{N}} K_{n,\varepsilon}^c \right) \right) \cup \bigcup_{n=n_\varepsilon+1}^{\infty} K_{n,\varepsilon} \\ &\subset \left(\bigcup_{n \in \mathbb{N}} U_{n,\varepsilon} \cap K_{n,\varepsilon}^c \right) \cup \bigcup_{n=n_\varepsilon+1}^{\infty} K_{n,\varepsilon} \\ &\subset \left(\bigcup_{n \in \mathbb{N}} U_{n,\varepsilon} \setminus K_{n,\varepsilon} \right) \cup \bigcup_{n=n_\varepsilon+1}^{\infty} K_{n,\varepsilon}. \end{aligned}$$

By (5.11) and (5.12), we deduce that

$$\begin{aligned} \mu(U_\varepsilon \setminus K_\varepsilon) &\leq \sum_{n \in \mathbb{N}} \mu(U_{n,\varepsilon} \setminus K_{n,\varepsilon}) + \sum_{n=n_\varepsilon+1}^{\infty} \mu(K_{n,\varepsilon}) \\ &\leq \varepsilon \sum_{n \in \mathbb{N}} 2^{-n-2} + \sum_{n=n_\varepsilon+1}^{\infty} \mu(B_n) \\ &\leq \varepsilon. \end{aligned}$$

Thus, the theorem is proved. \square

From the above theorem, which can be admitted, we deduce the following fundamental result.

Corollary 5.3.1. *Let p be a real number in the interval $[1, \infty[$. The space $L^p(\Omega)$ is separable.*

Proof. Let $(K_n)_{n \in \mathbb{N}}$ be an exhaustion by compact sets of Ω (see definition 5.3.1 on page 91). We only need to find a countable subset A of $L^p(\Omega)$ such that, for any continuous function with compact support in K_n , we have

$$\forall \varepsilon, \exists g \in A / \|f - g\|_{L^\infty(K_n)} \leq \varepsilon. \quad (5.14)$$

Let $(\psi_n)_{n \in \mathbb{N}}$ be a sequence of continuous functions which take the value 1 near K_n , and which are supported in $\overset{\circ}{K}_{n+1}$. The set A will be chosen as the set of functions written as $\psi_n P$, with P a polynomial function with d variables and rational coefficients. Proving that this set satisfies property (5.14) is a great exercise for the reader, which we highly recommend. \square

Exercise 5.3.1. Set $x \in \mathbb{R}^d$ and τ_x the mapping defined by

$$\tau_x \begin{cases} L^p & \longrightarrow L^p \\ f & \longmapsto \tau_x(f) : y \mapsto f(x-y). \end{cases}$$

Prove that the map τ_x is an isometry from L^p to L^p , and, if p is a real number, then

$$\forall f \in L^p, \lim_{x \rightarrow 0} \|\tau_x f - \check{f}\|_{L^p} = 0 \quad \text{where} \quad \check{f}(y) \stackrel{\text{def}}{=} f(-y).$$

The following exercise shows how one can linearly “smooth” functions in L^p .

Exercise 5.3.2. 1) Let K be a compact subset of \mathbb{R}^d . Prove that, for any positive ε , there exists a finite sequence $(A_{k,\varepsilon})_{k \in \mathbb{N}}$ of distinct sets with compact closure and diameter less than ε , such that

$$\bigcup_{k=1}^{\infty} A_{k,\varepsilon} = K.$$

We consider the map P_n defined by

$$P_n \begin{cases} L^1 & \longrightarrow L^\infty \\ f & \longmapsto P_n(f) \stackrel{\text{def}}{=} \sum_{j=1}^{\infty} \frac{1}{\mu(A_{j, \frac{1}{n+1}})} \left(\int_{A_{j, \frac{1}{n+1}}} f(x) d\mu(x) \right) \mathbf{1}_{A_{j, \frac{1}{n+1}}}. \end{cases}$$

- 2) Prove that P_n is a continuous linear map from $L^1(K)$ to $L^\infty(K)$.
- 3) Prove that, for any p , we have $\|P_n\|_{\mathcal{L}(L^p; L^p)} \leq 1$.
- 4) Prove that, if f belongs to L^p and g belongs to $L^{p'}$, then

$$\int_K f(x) P_n g(x) d\mu(x) = \int_K P_n f(x) g(x) d\mu(x).$$

- 5) Prove that

$$\forall f \in L^p, \lim_{n \rightarrow \infty} \|P_n f - f\|_{L^p} = 0.$$

5.4 Convolution and smoothing

In this section, we work in the whole space \mathbb{R}^d . The convolution operation is a critical one in the study of functions on \mathbb{R}^d .

Theorem 5.4.1. Let f and g be two functions in $L^1(\mathbb{R}^d)$. Then, for almost every x in \mathbb{R}^d , the function

$$y \mapsto f(x-y)g(y)$$

is integrable, and the function F defined by

$$F(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

belongs to L^1 . It is called the convolution of f and g , and is denoted $f \star g$. The thus defined \star operation is a continuous bilinear map from $L^1 \times L^1$ to L^1 . The following properties hold:

$$\|f \star g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1} \quad \text{and} \quad \int_{\mathbb{R}^d} (f \star g)(x) dx = \left(\int_{\mathbb{R}^d} f(x) dx \right) \left(\int_{\mathbb{R}^d} g(x) dx \right).$$

Proof. As the functions f and g are assumed to be in L^1 , we have

$$|f(x-y)| \times |g(y)| \in L^1(\mathbb{R}^d \times \mathbb{R}^d) \quad \text{and} \quad \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)| \times |g(y)| dx dy = \|f\|_{L^1} \|g\|_{L^1}.$$

The conclusions of Fubini’s theorem 5.1.5 immediately yield the first part of the theorem. The second part is a consequence of the Fubini-Tonelli theorem. \square

We can also define the convolution of a function in L^p and a function in $L^{p'}$.

Theorem 5.4.2. *Let f be a function in L^p , and g be a function in $L^{p'}$. The formula*

$$(f \star g)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

defines a continuous bilinear map from $L^p \times L^{p'}$ to L^∞ .

Proof. By Hölder's inequality, for almost every x in \mathbb{R}^d , the map

$$y \mapsto f(x-y)g(y)$$

is integrable, and

$$\left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \leq \|f\|_{L^p} \|g\|_{L^{p'}}.$$

The theorem is proved. □

Both of the above theorems extend to allow one to define the convolution of any two functions, providing these are in the appropriate L^p spaces. More precisely, the following holds.

Theorem 5.4.3 (Young's inequality). *Let (p, q, r) be a triple of real numbers such that*

$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}. \tag{5.15}$$

Consider a couple of functions (f, g) in $L^p \times L^q$. Then, the formula

$$(f \star g)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(x-y)g(y)dy$$

defines a function in L^r , and we have

$$\|f \star g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}.$$

Proof. This is an application of lemma 5.2.2. First, we assume that f and g are non-negative. Let φ be a positive function in $L^{r'}$. We are going to bound

$$I(f, g, \varphi) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x-y)g(y)\varphi(x)dx dy.$$

Let us assume that the inequality

$$I(f, g, \varphi) \leq \|f\|_{L^p} \|g\|_{L^q} \|\varphi\|_{L^{r'}} \tag{5.16}$$

is proved. This means that

$$(x, y) \mapsto f(x-y)g(y)\varphi(x)$$

is in $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Fubini's theorem 5.1.5 states that, for almost every x ,

$$y \mapsto f(x-y)g(y)$$

is in $L^1(\mathbb{R}^d, dy)$. Thus, $f \star g$ is a function defined for almost every x in \mathbb{R}^d . Moreover, Fubini's theorem ensures that

$$\begin{aligned} \int_{\mathbb{R}^d} f \star g(x) \varphi(x) dx &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x-y)g(y)dy\varphi(x)dx \\ &\leq I(f, g, \varphi). \end{aligned}$$

Inequality (5.16) and lemma 5.2.2 then imply that $f \star g$ is in L^r , and that

$$\|f \star g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^p}.$$

If f and g do not have a specific sign (or are complex-valued), simply apply the above to $|f|$ and $|g|$, and note that

$$\left| \int_{\mathbb{R}^d} f(x-y)g(y)dy \right| \leq \int_{\mathbb{R}^d} |f(x-y)| |g(y)| dy.$$

Now we prove inequality (5.16). Without loss of generality, we can assume that $\|f\|_{L^p} = \|g\|_{L^q} = 1$. Let α and β be two numbers in the interval $]0, 1[$. We write that

$$I(f, g, \varphi) = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x-y)^{1-\alpha} g(y)^{1-\beta} \varphi(x) \times f(x-y)^\alpha g(y)^\beta dx dy.$$

We apply Hölder's inequality with the measure

$$d\mu(x, y) = f(x-y)^\alpha g(y)^\beta dx dy$$

and the pair (r, r') . As a result, we have

$$\begin{aligned} I(f, g, \varphi) &\leq I^{(1)}(f, g, \varphi)^{\frac{1}{r}} I^{(2)}(f, g, \varphi)^{1-\frac{1}{r}} \\ \text{with } I^{(1)}(f, g, \varphi) &\stackrel{\text{def}}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} |f(x-y)|^{(1-\alpha)r+\alpha} |g(y)|^{(1-\beta)r+\beta} dx dy \\ \text{and } I^{(2)}(f, g, \varphi) &\stackrel{\text{def}}{=} \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x)^{r'} f(x-y)^\alpha g(y)^\beta dx dy. \end{aligned}$$

We start by bounding $I^{(1)}(f, g, \varphi)$. Fubini's theorem for non-negative functions and the invariance of the Lebesgue measure under translations mean that

$$I^{(1)}(f, g, \varphi) = \left(\int_{\mathbb{R}^d} f(x)^{(1-\alpha)r+\alpha} dx \right) \left(\int_{\mathbb{R}^d} g(y)^{(1-\beta)r+\beta} dy \right)$$

It is then natural to choose α such that $(1-\alpha)r+\alpha = p$ and β such that $(1-\beta)r+\beta = q$, which yields

$$\alpha = \frac{r-p}{r-1} \quad \text{and} \quad \beta = \frac{r-q}{r-1}.$$

As $\|f\|_{L^p} = \|g\|_{L^q} = 1$, we get that $I^{(1)}(f, g, \varphi) = 1$. Now we bound $I^{(2)}(f, g)$. To do so, observe that, by the definitions of α and β ,

$$\frac{\alpha}{p} + \frac{\beta}{q} = 1.$$

Hölder's inequality then implies that

$$\begin{aligned} \int_{\mathbb{R}^d} f(x-y)^\alpha |g(y)|^\beta dy &\leq \left(\int_{\mathbb{R}^d} f(x-y)^{\alpha \times \frac{p}{\alpha}} dy \right)^{\frac{\alpha}{p}} \left(\int_{\mathbb{R}^d} g(y)^{\beta \times \frac{q}{\beta}} dy \right)^{\frac{\beta}{q}} \\ &\leq \|f\|_{L^p}^\alpha \|g\|_{L^q}^\beta \\ &\leq 1. \end{aligned}$$

Hence,

$$I^{(2)}(f, g, \varphi) \leq \|\varphi\|_{L^{r'}}.$$

Inequality (5.16), and therefore theorem 5.4.3, are proved. \square

Theorem 5.4.4. *Let f and g be two functions in L^p and L^q respectively, such that*

$$\frac{1}{p} + \frac{1}{q} \geq 1.$$

Then

$$\text{Supp } (f \star g) \subset \text{Adh } (\text{Supp } f + \text{Supp } g).$$

Proof. Let x be a point in \mathbb{R}^d , and ρ be a positive number such that

$$B(x, \rho) \cap (\text{Supp } f + \text{Supp } g) = \emptyset.$$

For any bounded function φ which is zero outside $B(x, \rho)$, we have

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x) f(x-y) g(y) dx dy &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(x+y) f(x) g(y) dx dy \\ &= 0. \end{aligned}$$

The theorem is then proved by applying lemma 5.2.2 on $X = B(x, \rho)$. \square

Convolution is a crucial operation, because it is used to give an explicit means to smooth and approximate functions in L^p .

Theorem 5.4.5. *Let φ be a function in $L^1(\mathbb{R}^d)$ with integral equal to 1, and let p be a real number greater or equal to 1. Set*

$$\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi\left(\frac{x}{\varepsilon}\right).$$

Then, for every function f in L^p ,

$$\lim_{\varepsilon \rightarrow 0} \|\varphi_\varepsilon \star f - f\|_{L^p} = 0.$$

Proof. Let f be a function in L^p , and η be a positive real number. There exist two continuous functions g and ψ with compact support such that

$$\|f - g\|_{L^p} < \frac{\eta}{8\|\varphi\|_{L^1}} \quad \text{and} \quad \|\varphi - \psi\|_{L^1} < \frac{\eta}{8\|g\|_{L^p} + 1}.$$

As the integral of the function φ , and therefore also that of φ_ε , is equal to 1, we can write that

$$\begin{aligned}\varphi_\varepsilon \star f - f &= \varphi_\varepsilon \star f - \left(\int_{\mathbb{R}^d} \varphi_\varepsilon(x) dx \right) f \\ &= \varphi_\varepsilon \star (f - g) - \left(\int_{\mathbb{R}^d} \varphi_\varepsilon(x) dx \right) (f - g) \\ &\quad + (\varphi_\varepsilon - \psi_\varepsilon) \star g - \left(\int_{\mathbb{R}^d} (\varphi_\varepsilon - \psi_\varepsilon)(x) dx \right) g \\ &\quad + \psi_\varepsilon \star g - \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x) dx \right) g.\end{aligned}$$

Seeing the choice of the function g we made, we have

$$\|\varphi_\varepsilon \star (f - g)\|_{L^p} < \frac{\eta}{8} \quad \text{and} \quad \left\| \left(\int_{\mathbb{R}^d} \varphi_\varepsilon(x) dx \right) (f - g) \right\|_{L^p} < \frac{\eta}{8}.$$

Also, we chose ψ so that

$$\|(\varphi_\varepsilon - \psi_\varepsilon) \star g\|_{L^p} < \frac{\eta}{8} \quad \text{and} \quad \left\| \left(\int_{\mathbb{R}^d} (\varphi_\varepsilon - \psi_\varepsilon)(x) dx \right) g \right\|_{L^p} < \frac{\eta}{8}.$$

Ultimately, we get

$$\|\varphi_\varepsilon \star f - f\|_{L^p} < \frac{\eta}{2} + \left\| \psi_\varepsilon \star g - \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x) dx \right) g \right\|_{L^p}.$$

But, by the definition of convolution, we gave, by a change of variables,

$$(\psi_\varepsilon \star g)(x) - \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x) dx \right) g(x) = \int_{\mathbb{R}^d} \psi(z) (g(x - \varepsilon z) - g(x)) dz.$$

We deduce immediately that

$$\left| (\psi_\varepsilon \star g)(x) - \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x) dx \right) g(x) \right| \leq \int_{\mathbb{R}^d} |\psi(z)| |g(x - \varepsilon z) - g(x)| dz.$$

As the function g is uniformly continuous, for any η , there exists a positive α such that

$$|y - y'| \leq \alpha \Rightarrow |g(y) - g(y')| < \frac{\eta}{2\|\psi\|_{L^1} |\text{Supp } g + B(0, \varepsilon R)|^{\frac{1}{p}} + 1}.$$

Set

$$R \stackrel{\text{def}}{=} \sup_{z \in \text{Supp } \psi} |z|.$$

Then we have

$$\varepsilon < \frac{\alpha}{R} \Rightarrow \|\psi_\varepsilon \star g - g\|_{L^\infty} \leq \frac{\eta}{2|\text{Supp } g + B(0, \varepsilon R)|^{\frac{1}{p}}}.$$

Hence, we get

$$\|\varphi_\varepsilon \star f - f\|_{L^p} < \eta.$$

Thus, the theorem is proved. \square

Important remark. As the reader will prove in the next exercise, this result does not hold when $p = \infty$.

Exercise 5.4.1. Consider for f the Heaviside function H (the characteristic function of the set of non-negative real numbers), and for φ any even continuous function with compact support and integral equal to 1. Prove that $\|\varphi_\varepsilon \star f - f\|_{L^\infty} \geq 1/2$.

As we will see, the previous theorem is extremely important when it comes to smoothing functions. Indeed, theorem 5.1.4 on page 83 ensures that if φ is smooth with compact support, then the function $\varphi_\varepsilon \star f$ is also smooth with compact support. Thus, every function in L^p can be approximated by a compactly supported smooth function. First, the existence of such functions must be assured.

Proposition 5.4.1. Let f be the function from \mathbb{R} to \mathbb{R} defined by

$$f(x) = e^{\frac{1}{x(x-1)}} \quad \text{if } x \in [0, 1] \quad \text{and} \quad f(x) = 0 \quad \text{otherwise.}$$

This function is smooth with compact support.

Proof. It suffices to notice that

$$f^{(k)}(x) = \frac{P_k(x)}{x^{k+1}(1-x)^{k+1}} e^{\frac{1}{x(x-1)}},$$

where P_k is a polynomial function with degree δ_k . We let the reader fill in the details.

Remark. The functions φ_ε are known as mollifiers, or approximations to the identity.

Corollary 5.4.1. Let Ω be an open set of \mathbb{R}^d . We denote $\mathcal{D}(\Omega)$ the set of smooth functions with compact support in Ω . The space $\mathcal{D}(\Omega)$ is not equal to $\{0\}$.

Proof. Consider the function $\varphi(x) = f(|x|)$, where f is given by proposition 5.4.1 above. Simply apply a homothety-translation to this function to fit its support in Ω . \square

Corollary 5.4.2. For every real number $p \geq 1$, the space $\mathcal{D}(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$.

Proof. By corollary 5.4.1, there exists a function φ that is smooth with compact support that has its integral equal to 1. Consider the family $(\varphi_\varepsilon)_{\varepsilon>0}$ defined by

$$\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi\left(\frac{x}{\varepsilon}\right).$$

For any function f in L^p , we can find a continuous function g with compact support as close to f as we like in L^p . Set

$$g_\varepsilon = \varphi_\varepsilon \star g.$$

By proposition 5.4.4, the function g_ε has compact support. Also, as φ_ε is smooth, so is g_ε , by theorem 5.1.4. The corollary is proved. \square

Corollary 5.4.3. Let Ω be an open set of \mathbb{R}^d , and let p be a number in the interval $[1, \infty[$. The space $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$.

Proof. Let $(K_n)_{n \in \mathbb{N}}$ be an exhaustion of Ω by compact sets (recall definition 5.3.1 on page 91). For f in L^p , set $f_n \stackrel{\text{def}}{=} \mathbf{1}_{K_n} f$. By theorem 5.1.3, we have

$$\lim_{n \rightarrow \infty} \|f_n - f\|_{L^p(\Omega)} = 0.$$

Now set $f_{n,\varepsilon} \stackrel{\text{def}}{=} \varphi_\varepsilon \star f_n$. This function is smooth, and its support is the compact set $K_n + B_f(0, \varepsilon)$, which is the set of points that are at a distance at most ε from K_n . By proposition 5.3.1 on page 90, for $\varepsilon \leq \varepsilon_n$, the compact set $K_n + B_f(0, \varepsilon)$ is a subset of Ω , hence the corollary. \square

Corollary 5.4.4. *Let K be a compact subset of Ω , an open set in \mathbb{R}^d . There exists a function ψ in $\mathcal{D}(\Omega)$ such that $\psi = 1$ on a neighbourhood of K .*

Proof. Let δ be a positive number such that, if

$$K_r \stackrel{\text{def}}{=} \{x \in \Omega / d(x, K) \leq 2r\},$$

then $K_{2\delta}$ is a subset of Ω (such a number exists by theorem 5.3.1 on page 90). We consider a sequence of mollifiers $(\varphi_\varepsilon)_\varepsilon$ and set

$$\psi_\varepsilon(x) \stackrel{\text{def}}{=} \varphi_\varepsilon \star \mathbf{1}_{K_\delta}(x) = \int_{K_\delta} \varphi_\varepsilon(x - y) dy.$$

There exists a constant C such that

$$x \notin K_\delta + B(0, C\varepsilon) \implies \varphi_\varepsilon \star \mathbf{1}_{K_\delta}(x) = 0.$$

So, if $C\varepsilon < \delta$, then $\psi_\varepsilon \in \mathcal{D}(\Omega)$, as $K_{2\delta} \subset \Omega$. Finally, if x is in $K_{\delta - C\varepsilon}$ (which is non-empty since $\delta > C\varepsilon$), then, for any $y \in B(x, C\varepsilon)$, y belongs to K_δ . Thus,

$$\begin{aligned} \varphi_\varepsilon \star \mathbf{1}_{K_\delta}(x) &= \int_{\mathbb{K}_\delta} \varphi_\varepsilon(x - y) dy \\ &= \int_{\mathbb{R}^d} \varphi_\varepsilon(x - y) dy \\ &= 1. \end{aligned}$$

The corollary is proved. \square

5.5 Duality between L^p and $L^{p'}$

The main result of this section is the following theorem, which explains how continuous linear functionals on L^p spaces, when p is real, can be represented.

Theorem 5.5.1. *Let p be a real number (strictly) greater than 1, and let B be the bilinear form defined by*

$$B \begin{cases} L^{p'} \times L^p & \longrightarrow \mathbb{K} \\ (g, f) & \longmapsto \int_X f(x)g(x)d\mu(x). \end{cases}$$

Then the bilinear form B identifies the dual of $L^p(X, d\mu)$ with $L^{p'}(X, d\mu)$.

Before proving this theorem, we will show an important consequence of it, and make a few comments about the result.

Corollary 5.5.1. *Let p be a number in $[1, \infty[$. Consider a bounded sequence $(g_n)_{n \in \mathbb{N}}$ in $L^{p'}$. There exists an extraction function ϕ and a function g in $L^{p'}$ such that*

$$\forall f \in L^p, \lim_{n \rightarrow \infty} \int_X g_{\phi(n)}(x) f(x) d\mu(x) = \int_X g(x) f(x) d\mu(x). \quad (5.17)$$

Proof. We apply theorem 3.3.4 on page 60 to the sequence $(L_n)_{n \in \mathbb{N}}$ of linear functionals on L^p defined by

$$\langle L_n, f \rangle \stackrel{\text{def}}{=} \int_X g_n(x) f(x) d\mu(x).$$

This yields the existence of an extraction function ϕ and a linear form L such that

$$\forall f \in L^p, \lim_{n \rightarrow \infty} \int_X g_{\phi(n)}(x) f(x) d\mu(x) = \langle L, f \rangle.$$

Theorem 5.5.1 above then ensures the existence of a function g in $L^{p'}$ such that

$$\forall f \in L^p, \lim_{n \rightarrow \infty} \int_X g(x) f(x) d\mu(x) = \langle L, f \rangle.$$

The corollary is proved. □

Let us consider the case where X is an open set in \mathbb{R}^d , and μ is the Lebesgue measure. Note that, as p is finite, the space $C_c(\Omega)$ is dense in $L^p(\Omega)$. We let the reader establish, as an exercise, that assertion (5.17) is equivalent to the following:

$$\forall \varphi \in C_c(\Omega), \lim_{n \rightarrow \infty} \int_X g_{\phi(n)}(x) \varphi(x) d\mu(x) = \int_X g(x) \varphi(x) d\mu(x). \quad (5.18)$$

We are going to observe that this property is false when p' is equal to 1, in other words when p is infinity. Indeed, let χ be a function in $\mathcal{D}(\mathbb{R}^d)$ with integral equal to 1. As we showed when proving theorem 5.4.5, if $(\varepsilon_n)_{n \in \mathbb{N}}$ is a sequence of positive numbers converging to 0, we have

$$\forall \varphi \in C_c(\Omega), \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} \frac{1}{\varepsilon_n^d} \chi\left(\frac{x}{\varepsilon_n}\right) \varphi(x) dx = \varphi(0).$$

But the linear functional δ_0 defined on $C_c(\Omega)$ by $\langle \delta_0, \varphi \rangle = \varphi(0)$ is the Dirac measure at 0. No function g in L^1 can represent this via the bilinear form B because, if such was the case,

$$\int_{\Omega} g(x) \varphi(x) dx = \varphi(0)$$

implies that g is zero except at $x = 0$, and therefore g is zero almost everywhere.

Proof of theorem 5.5.1 when p belongs to $[1, 2]$. The theorem was proved in the case $p = 2$ in chapter 4 on Hilbert spaces. The set X is assumed to be the union of countably many subsets K_n , each with finite measure. Without loss of generality, we can assume that the sequence $(K_n)_{n \in \mathbb{N}}$ is non-decreasing, in the sense that $K_n \subset K_{n+1}$. Let Φ be a linear functional on $L^p(X, d\mu)$, and let K be any subset of X with finite measure. We denote Φ_K the restriction of Φ to K , that is the linear form defined by

$$\langle \Phi_K, f \rangle \stackrel{\text{def}}{=} \langle \Phi, \mathbf{1}_K f \rangle.$$

We are now going to work in the space $L^p(K, d\mu)$. Let Φ be a linear form on $L^p(K, d\mu)$. Assuming for now that there exists a unique function g_K on K such that

$$\forall f \in L^p(K, d\mu), \quad \langle \Phi, f \rangle = B(g_K, f) = \int_K g_K(x) f(x) d\mu(x), \quad (5.19)$$

and such that

$$\|g_K\|_{L^{p'}} = \|\Phi_K\|_{(L^p)'}, \quad (5.20)$$

we consider the non-decreasing sequence $(K_n)_{n \in \mathbb{N}}$ of subsets of X with finite measure such that the union of the K_n is the whole space X . Then there exists a sequence of functions $(g_n)_{n \in \mathbb{N}}$ satisfying

$$\forall f \in L^p(K_n, d\mu), \quad \langle \Phi_{K_n}, f \rangle = B(g_n, f).$$

By lemma 5.2.2, we know that

$$g_n \in L^{p'} \quad \text{and that} \quad \|g_n\|_{L^{p'}} = \|\Phi_{K_n}\|_{(L^p)' } \leq \|\Phi\|_{(L^p)' }.$$

Moreover, the uniqueness of the function g_n satisfying (5.19) and (5.20) ensures that, if $m \geq n$,

$$\mathbf{1}_{K_n} g_m = g_n,$$

because the function $\mathbf{1}_{K_n} g_m$ also satisfies relations (5.19) and (5.20) for the linear form Φ_{K_n} . The sequence $(|g_n|^{p'})_{n \in \mathbb{N}}$ is therefore an non-decreasing sequence of functions such that

$$\sup_n \int_{\Omega} |g_n(x)|^{p'} d\mu(x) \leq \|\Phi\|_{(L^p)'}^{p'}.$$

The monotone convergence theorem ensures that the function

$$g(x) = \lim_{n \rightarrow \infty} g_n(x) \in L^{p'} \quad \text{and that} \quad \|g\|_{L^{p'}} \leq \|\Phi\|_{(L^p)' }.$$

Lemma 5.2.2 then ensures that $\|g\|_{L^{p'}} = \|\Phi\|_{(L^p)'}$. Moreover, for any function f in L^p and for any integer n , we have

$$\langle \Phi, \mathbf{1}_{K_n} f \rangle = \int_X g(x) \mathbf{1}_{K_n} f(x) d\mu(x).$$

The dominated convergence theorem then finishes the proof off. But we are yet to prove relations (5.19) and (5.20). By Hölder's inequality, we can say that

$$\begin{aligned} \|f\|_{L^p} &= \left(\int_K |f(x)|^p d\mu(x) \right)^{\frac{1}{p}} \\ &\leq \left(\int_K |f(x)|^{p \frac{2}{p}} d\mu(x) \right)^{\frac{p}{2} \times \frac{1}{p}} \mu(K)^{\frac{1}{p} (1 - \frac{p}{2})} \\ &\leq \|f\|_{L^2} \mu(K)^{\frac{1}{p} - \frac{1}{2}}. \end{aligned}$$

Thus, the linear form Φ appears as a continuous linear form on the space L^2 , so there exists a function g in L^2 such that

$$\forall f \in L^2, \quad \langle \Phi, f \rangle = \int_X f(x) g(x) d\mu(x).$$

We deduce that, for any function f in L^2 , we have

$$\int_K |f(x) g(x)| d\mu(x) \leq \|\Phi\|_{(L^p)' } \|f\|_{L^p}.$$

But, by corollary 5.3.1, the space $L^\infty \cap L^p$ is dense in L^p . As p is between 1 and 2 here, and since K is compact, L^2 is dense in L^p . By lemma 5.2.2, we deduce that g belongs to $L^{p'}$, and $\|g\|_{L^{p'}} = \|\Phi\|_{(L^p)'}$. Thus, (5.19) and (5.20) are proved.

Proof of theorem 5.5.1 when p belongs to $]2, \infty[$. We start by proving that, for p in the interval $]2, \infty[$, the following inequality holds:

$$\forall (a, b) \in \mathbb{R}^2, \quad \frac{|a|^p + |b|^p}{2} \geq \left| \frac{a+b}{2} \right|^p + \frac{p(p-1)}{2^{p-2}} \left| \frac{a-b}{2} \right|^p. \quad (5.21)$$

As the function $t \mapsto |t|^p$ is twice continuously differentiable on \mathbb{R} , the second-order Taylor formula with integral remainder between a and $(a+b)/2$, then between b and $(a+b)/2$, yields

$$\begin{aligned} |a|^p &= \left| \frac{a+b}{2} \right|^p + p \frac{a-b}{2} \frac{a+b}{2} \left| \frac{a+b}{2} \right|^{p-2} \\ &\quad + p(p-1) \left(\frac{a-b}{2} \right)^2 \int_0^1 (1-\theta) \left| \frac{a+b}{2} + \theta \frac{a-b}{2} \right|^{p-2} d\theta \\ \text{and } |b|^p &= \left| \frac{a+b}{2} \right|^p + p \frac{b-a}{2} \frac{a+b}{2} \left| \frac{a+b}{2} \right|^{p-2} \\ &\quad + p(p-1) \left(\frac{a-b}{2} \right)^2 \int_0^1 (1-\theta) \left| \frac{a+b}{2} + \theta \frac{b-a}{2} \right|^{p-2} d\theta. \end{aligned}$$

The mean of these two equalities gives

$$\begin{aligned} \frac{|a|^p + |b|^p}{2} &= \left| \frac{a+b}{2} \right|^p \\ &\quad + \frac{p(p-1)}{2} \left(\frac{a-b}{2} \right)^2 \int_0^1 (1-\theta) \left(\left| \frac{a+b}{2} + \theta \frac{a-b}{2} \right|^{p-2} + \left| \frac{a+b}{2} + \theta \frac{b-a}{2} \right|^{p-2} \right) d\theta. \end{aligned}$$

If we have

$$\left| \frac{a+b}{2} + \theta \frac{a-b}{2} \right| \leq \frac{1}{4} |b-a| \quad \text{and} \quad \left| \frac{a+b}{2} + \theta \frac{b-a}{2} \right| \leq \frac{1}{4} |b-a|, \quad (5.22)$$

then

$$\begin{aligned} \theta |b-a| &= \left| \theta \frac{a-b}{2} - \theta \frac{b-a}{2} \right| \\ &= \left| \frac{a+b}{2} + \theta \frac{a-b}{2} - \frac{a+b}{2} - \theta \frac{b-a}{2} \right| \\ &\leq \left| \frac{a+b}{2} + \theta \frac{a-b}{2} \right| + \left| \frac{a+b}{2} - \theta \frac{b-a}{2} \right| \\ &\leq \frac{1}{2} |b-a|. \end{aligned}$$

Assuming that a and b are distinct (there is nothing worth seeing otherwise), we get that assertion (5.22) can only be satisfied if θ is less than or equal to $1/2$. Hence,

$$\begin{aligned} \frac{|a|^p + |b|^p}{2} &\geq \left| \frac{a+b}{2} \right|^p + \frac{p(p-1)}{2} \left(\frac{b-a}{2} \right)^2 \left| \frac{b-a}{4} \right|^{p-2} \int_{\frac{1}{2}}^1 (1-\theta) d\theta \\ &\geq \left| \frac{a+b}{2} \right|^p + \frac{p(p-1)}{2^{p-2}} \left| \frac{b-a}{2} \right|^p, \end{aligned}$$

which proves inequality (5.21). The proof of the theorem will now continue, following the lines of that of theorem 4.3.2 on page 70. Let ℓ be a continuous linear form on $L^p(X, d\mu)$. Consider the function

$$F_p \begin{cases} L^p(X, d\mu) & \longrightarrow \mathbb{R} \\ f & \longmapsto \frac{1}{p} \|f\|_{L^p}^p - \langle \ell, f \rangle \end{cases}$$

The function F_p is bounded from below, because

$$F_p(f) \geq \frac{1}{p} \|f\|_{L^p}^p - \|\ell\|_{(L^p)'} \|f\|_{L^p}.$$

Set $m_p \stackrel{\text{def}}{=} \inf_{f \in L^p} F_p(f)$. We consider a minimising sequence $(f_n)_{n \in \mathbb{N}}$, that is a sequence such that

$$F_p(f_n) = m_p + \varepsilon_n \quad \text{with} \quad \lim_{n \rightarrow \infty} \varepsilon_n = 0.$$

We are going to prove that this is a Cauchy sequence. First, we apply inequality (5.21) with $a = f(x)$ and $b = g(x)$, then we integrate with respect to the measure $d\mu$; we get that

$$\forall (f, g) \in (L^p(X, d\mu))^2, \quad \|f\|_{L^p}^p + \|g\|_{L^p}^p \geq 2 \left\| \frac{f+g}{2} \right\|_{L^p}^p + \frac{p(p-1)}{2^{p-1}} \left\| \frac{f-g}{2} \right\|_{L^p}^p. \quad (5.23)$$

Applying this inequality with $f = f_n$ and $g = f_m$, we can write

$$\frac{1}{p} \|f_n\|_{L^p}^p + \frac{1}{p} \|f_m\|_{L^p}^p \geq \frac{2}{p} \left\| \frac{f_n + f_m}{2} \right\|_{L^p}^p + \frac{(p-1)}{2^{p-1}} \left\| \frac{f_n - f_m}{2} \right\|_{L^p}^p.$$

After subtracting $\langle \ell, f_n + f_m \rangle$, the inequality becomes

$$\begin{aligned} & \frac{1}{p} \|f_n\|_{L^p}^p - \langle \ell, f_n \rangle + \|f_m\|_{L^p}^p - \langle \ell, f_m \rangle \\ & \geq 2 \left(\left\| \frac{f_n + f_m}{2} \right\|_{L^p}^p - \langle \ell, \frac{f_n + f_m}{2} \rangle \right) + \frac{(p-1)}{2^{p-1}} \left\| \frac{f_n - f_m}{2} \right\|_{L^p}^p. \end{aligned}$$

By definition of F_p , this can be written as

$$\frac{(p-1)}{2^{p-1}} \left\| \frac{f_n - f_m}{2} \right\|_{L^p}^p \leq F_p(f_n) + F_p(f_m) - 2F_p\left(\frac{f_n + f_m}{2}\right)$$

As the sequence $(f_n)_{n \in \mathbb{N}}$ minimises F_p , we deduce that

$$\frac{(p-1)}{2^{p-1}} \left\| \frac{f_n - f_m}{2} \right\|_{L^p}^p \leq \varepsilon_n + \varepsilon_m,$$

which ensures that $(f_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. Let f be its limit. As the function F_p is continuous, we deduce that $F_p(f) = m_p$; in other words, the infimum of F_p is a minimum.

Now, for a given function h in L^p , consider the function g_h from \mathbb{R} to \mathbb{R} defined by

$$g_h(t) \stackrel{\text{def}}{=} F_p(f + th).$$

Using the differentiability theorem for functions defined by integrals, we see that the function g_h is differentiable, and that

$$g_h'(t) = \int_X (f(x) + th(x)) |f(x) + th(x)|^{p-2} h(x) d\mu(x) - \langle \ell, h \rangle.$$

The point $t = 0$ is a minimum of the function g_h , so its derivative vanishes at this point, which means that

$$\forall h \in L^p(X, d\mu), \quad \int_X f(x) |f(x)|^{p-2} h(x) d\mu(x) = \langle \ell, h \rangle,$$

which ends the proof of theorem 5.5.1. □

The result of theorem 5.4.5 on families of mollifiers can be interpreted in terms of duality: using the bilinear form

$$B(f, g) = \int_{\Omega} f(x)g(x)dx,$$

the space $L^1(\Omega, dx)$ can be identified as the dual of a normed space containing only continuous functions with compact support in Ω . Indeed, let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers converging to 0. We consider a sequence $(\varphi_{\varepsilon_n})_{n \in \mathbb{N}}$ of approximations of unity as defined in theorem 5.4.5. Then, for any continuous function g with compact support, we have

$$\lim_{n \rightarrow \infty} \int_{\Omega} \varphi_{\varepsilon_n}(x)g(x)dx = g(0).$$

Chapter 6

The Dirichlet problem

Introduction

The aim of this chapter is to solve the Dirichlet problem on a bounded domain of \mathbb{R}^d . We shall see how this problem, when formulated in terms of functions of class C^1 , will naturally lead to the introduction of the weak derivative, or quasi-derivative, concept. This concept will then lead us to the more general sense of derivation for distributions.

In order to formulate the Dirichlet problem, we will need the following notations. Let Ω be a bounded open set of \mathbb{R}^d . The set of continuously differentiable functions on Ω with support in a compact subset of $\bar{\Omega}$ is denoted $C_0^1(\Omega)$. For any x in Ω and for any function u of class C^1 on Ω , we denote

$$\nabla u(x) = \text{grad } u(x) = (\partial_1 u(x), \dots, \partial_d u(x)) \in \mathbb{R}^d.$$

Moreover, we have

$$|\nabla u(x)|^2 = \sum_{j=1}^d |\partial_j u(x)|^2 \quad \text{and} \quad \|\nabla u\|_{L^2(\Omega)}^2 \stackrel{\text{def}}{=} \int_{\Omega} |\nabla u(x)|^2 dx = \sum_{j=1}^d \int_{\Omega} |\partial_j u(x)|^2 dx.$$

If there is no risk of confusion, we will omit to write the dependence of these norms on Ω . Let f be a function in $L^2(\Omega)$. We consider the functional

$$F \begin{cases} C_0^1(\Omega) & \longrightarrow \mathbb{R} \\ u & \longmapsto \frac{1}{2} \|\nabla u(x)\|_{L^2}^2 - \int_{\Omega} f(x)u(x)dx = \frac{1}{2} \sum_{j=1}^d \int_{\Omega} |\partial_j u(x)|^2 dx - \int_{\Omega} f(x)u(x)dx \end{cases} ,$$

and we want to determine if F has a minimum, that is if there exists a function u in $C_0^1(\Omega)$ such that

$$F(u) = \inf_{v \in C_0^1(\Omega)} F(v).$$

Note that F is a continuous function (exercise: prove it!).

In the first section of this chapter, we will present a classical approach to this problem, by studying a minimising sequence from which we extract a subsequence that converges weakly to a function u , and such that the sequences of partial derivatives is also weakly convergent.

This will lead us to introduce the concept of quasi-derivative of a function in L^2 in the second section. Originally used by the French mathematician Jean Leray in 1934 in order to solve a global existence problem for solutions of the equation that governs the flow of viscous

incompressible fluids, it constitutes the first soundly defined departure from the classical sense of the derivative.

In the third section, this concept will allow us to define the Sobolev space, which is a Hilbert space adapted for our problem. Once we have defined this space, solving the Dirichlet problem $-\Delta u = f$, with u “vanishing on the boundary”, will be a simple application of the Riesz representation theorem (theorem 4.3.1 on page 70). This example illustrates the power of functional analysis, by taking a seemingly complicated problem and solving it with elegant simplicity, providing the conceptual setting is chosen appropriately. As our final result in this chapter, we will use the diagonalisation theorem for compact self-adjoint operators (theorem 4.4.2 on page 75) to “diagonalise” the Laplacian operator Δ .

6.1 A classical approach to the problem

We must first prove that the function F is bounded from below. This relies on the Poincaré inequality.

Proposition 6.1.1 (Poincaré inequality). *Let Ω be a bounded open set of \mathbb{R}^d . There exists a constant C (which depends on Ω) such that*

$$\forall u \in C_0^1(\Omega), \|u\|_{L^2} \leq C \|\nabla u\|_{L^2}.$$

Proof. Let u be a function in $C_0^1(\Omega)$. We first observe that the function equal to zero outside Ω and that coincides with u in Ω is of class C^1 and has compact support in \mathbb{R}^d . To simplify the notations, this function will also be denoted u . We have

$$\forall (x_1, x') \in \Omega, u(x_1, x') = \int_{-\infty}^{x_1} \partial_1 u(y_1, x') dy_1.$$

By the Cauchy-Schwarz inequality, this implies

$$\forall (x_1, x') \in \Omega, |u(x_1, x')|^2 \leq \delta(\Omega) \int_{-\infty}^{+\infty} |\partial_1 u(y_1, x')|^2 dy_1.$$

Integrating this inequality with respect to $x = (x_1, x')$, we get

$$\int_{\Omega} |u(x_1, x')|^2 dx \leq \delta(\Omega)^2 \int_{\Omega} |\partial_1 u(y_1, x')|^2 dy_1 dx'.$$

Thus, we have proved the Poincaré inequality. □

We deduce from this the following corollary.

Corollary 6.1.1. *The functional F is bounded from below. Moreover, we have*

$$\forall A > 0, \exists B / \|\nabla u\|_{L^2} \geq B \implies F(u) \geq A. \tag{6.1}$$

Proof. It suffices to say that, by the Poincaré inequality, we have

$$\begin{aligned} F(u) &\geq \frac{1}{2} \|\nabla u\|_{L^2}^2 - \|f\|_{L^2} \|u\|_{L^2} \\ &\geq \frac{1}{2} \|\nabla u\|_{L^2}^2 - C \|f\|_{L^2} \|\nabla u\|_{L^2} \\ &\geq \frac{1}{2} (\|\nabla u\|_{L^2} - C \|f\|_{L^2})^2 - \frac{1}{2} C^2 \|f\|_{L^2}^2. \end{aligned}$$

So F is bounded from below. Moreover, if A is a positive number, by setting

$$B = (2A + C^2 \|f\|_{L^2}^2)^{\frac{1}{2}} + C \|f\|_{L^2},$$

we end the proof of the corollary. \square

Let us now consider a minimising sequence $(u_n)_{n \in \mathbb{N}}$ of functions in $C_0^1(\Omega)$, i.e. that satisfies

$$\lim_{n \rightarrow \infty} F(u_n) = m = \inf_{v \in C_0^1(\Omega)} F(v).$$

As we want to know if this infimum is reached, we are looking for a limit for this sequence. Inequality (6.1) in corollary 6.1.1 implies that the sequence $(\|\nabla u_n\|_{L^2})_{n \in \mathbb{N}}$ is bounded. The Poincaré inequality yields that the sequence $(u_n)_{n \in \mathbb{N}}$ is bounded in L^2 . Thus, the sequences $(u_n)_{n \in \mathbb{N}}$ and $(\partial_j u_n)_{n \in \mathbb{N}}$ are all bounded in L^2 . The weak compactness theorem 4.3.4 on page 72 then states that there exist a function u and functions $u^{(j)}$ such that, after extracting a subsequence, we have

$$\liminf_{n \rightarrow \infty} u_n = u \quad \text{and} \quad \liminf_{n \rightarrow \infty} \partial_j u_n = u^{(j)}.$$

Let us assume for a moment that u is in $C_0^1(\Omega)$, and let us consider φ in $C_0^1(\Omega)$. By definition of the weak convergence in a Hilbert space, we have

$$\int_{\Omega} u^{(j)}(x) \varphi(x) dx = \lim_{n \rightarrow \infty} \int_{\Omega} \partial_j u_n(x) \varphi(x) dx.$$

The functions u_n and φ are functions in $C_0^1(\Omega)$, so we have, by integrating by parts,

$$\int_{\Omega} u^{(j)}(x) \varphi(x) dx = - \lim_{n \rightarrow \infty} \int_{\Omega} u_n(x) \partial_j \varphi(x) dx.$$

Once again using the definition of weak convergence in L^2 , we deduce that

$$\int_{\Omega} u^{(j)}(x) \varphi(x) dx = - \int_{\Omega} u(x) \partial_j \varphi(x) dx.$$

By corollary 5.4.3 on page 102, we know that $C_0^1(\Omega)$ is dense in $L^2(\Omega)$, thus we have $u^{(j)} = \partial_j u$.

But there is no reason to believe that u is always of class C^1 ; this is in fact generally false. However, the above computation shows us a link between the partial derivatives of u and the limit functions $u^{(j)}$. We will give mathematical foundations to this link in the next section.

6.2 The concept of quasi-derivatives

Here is what we are talking about.

Definition 6.2.1. *Let u be a function in $L^2(\Omega)$. We say that u has a partial quasi-derivative in the direction j if and only if there exists a constant C such that*

$$\forall \varphi \in C_0^1(\Omega), \quad \left| \int_{\Omega} u(x) \partial_j \varphi(x) dx \right| \leq C \|\varphi\|_{L^2}$$

Fundamental remark. By theorem 5.3.1 on page 91, the space $C_0^1(\Omega)$ is dense in $L^2(\Omega)$. Theorem 1.2.4 on page 17, on prolonging continuous functions on a dense subspace, states that we can extend the linear form

$$\ell_j \begin{cases} C_0^1(\Omega) & \longrightarrow \mathbb{R} \\ \varphi & \longmapsto - \int_{\Omega} u(x) \partial_j \varphi(x) dx. \end{cases}$$

to the whole L^2 space. We denote $\tilde{\ell}_j$ this linear form. The Riesz representation theorem 4.3.1 on page 70 ensures that there exists a function $u^{(j)}$ in L^2 such that

$$\forall f \in L^2(\Omega), \int_{\Omega} u^{(j)}(x) f(x) dx = \langle \tilde{\ell}_j, f \rangle.$$

In particular, for $f = \varphi \in C_0^1(\Omega)$, this yields

$$\forall \varphi \in C_0^1(\Omega), \int_{\Omega} u^{(j)}(x) \varphi(x) dx = - \int_{\Omega} u(x) \partial_j \varphi(x) dx.$$

We can thus give a more complete definition of the quasi-derivative.

Definition 6.2.2. Let u be a function in $L^2(\Omega)$. We say that u has a partial quasi-derivative in the direction j if and only if there exists a constant C such that

$$\forall \varphi \in C_0^1(\Omega), \left| \int_{\Omega} u(x) \partial_j \varphi(x) dx \right| \leq C \|\varphi\|_{L^2}.$$

Moreover, the function $u^{(j)}$ in $L^2(\Omega)$ that satisfies

$$\forall \varphi \in C_0^1(\Omega), \int_{\Omega} u^{(j)}(x) \varphi(x) dx = - \int_{\Omega} u(x) \partial_j \varphi(x) dx.$$

is called the quasi-derivative of u in the direction j .

To illustrate this notion, let us examine the following example.

Proposition 6.2.1. Let $\Omega =]-1, 1[$. We consider the function $u(x) = |x|^\alpha$ with α in $]1/2, 1[$. This function has a quasi-derivative, which is the function

$$x \longmapsto \alpha \operatorname{sg}(x) |x|^{\alpha-1}.$$

Proof. We write that, for any function φ of class C^1 and with compact support in the interval $] -1, 1[$,

$$\int_{-1}^1 |x|^\alpha \varphi(x) dx = \lim_{\varepsilon \rightarrow 0} I_\varepsilon(\varphi) \quad \text{where} \quad I_\varepsilon(\varphi) \stackrel{\text{def}}{=} \int_{-1}^{-\varepsilon} (-x)^\alpha \varphi'(x) dx + \int_{\varepsilon}^1 x^\alpha \varphi'(x) dx.$$

Integrating by parts, we get that

$$I_\varepsilon(\varphi) = \varepsilon^\alpha (\varphi(-\varepsilon) - \varphi(\varepsilon)) + \int_{-1}^{-\varepsilon} \alpha (-x)^{\alpha-1} \varphi(x) dx - \int_{\varepsilon}^1 \alpha x^{\alpha-1} \varphi(x) dx$$

As α is strictly greater than $1/2$, the function $|x|^{\alpha-1}$ is square-integrable. We can thus take the limit in the above integrals, which yields

$$\int_{-1}^1 |x|^\alpha \varphi'(x) dx = - \int_{-1}^1 \alpha \operatorname{sg}(x) |x|^{\alpha-1} \varphi(x) dx.$$

The proposition is proved. \square

The following proposition shows that we have effectively generalised the concept of derivatives, at least when it comes to functions that are C^1 , continuously differentiable.

Proposition 6.2.2. *If u belongs to $C_0^1(\Omega)$, then u has quasi-derivatives in all directions, and these quasi-derivatives coincide with the partial derivatives.*

Proof. By integrating by parts and using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} \left| \int_{\Omega} u(x) \partial_j \varphi(x) dx \right| &= \left| \int_{\Omega} \partial_j u(x) \varphi(x) dx \right| \\ &\leq \|\partial_j u\|_{L^2} \|\varphi\|_{L^2} \\ &\leq C \|\partial_j u\|_{L^\infty} \|\varphi\|_{L^2}. \end{aligned}$$

Let $u^{(j)}$ be the quasi-derivative of u in the direction j . By definition, this function in $L^2(\Omega)$ satisfies

$$\forall \varphi \in C_0^1(\Omega), \quad \int_{\Omega} u^{(j)}(x) \varphi(x) dx = - \int_{\Omega} u(x) \partial_j \varphi(x) dx.$$

Moreover, by integrating by parts, we have

$$\forall \varphi \in C_0^1(\Omega), \quad \int_{\Omega} \partial_j u(x) \varphi(x) dx = - \int_{\Omega} u(x) \partial_j \varphi(x) dx.$$

This then implies that

$$\int_{\Omega} (\partial_j u - u^{(j)})(x) \varphi(x) dx = 0$$

As the space $C_0^1(\Omega)$ is dense in $L^2(\Omega)$, we conclude that $u^{(j)} = \partial_j u$, hence the proposition. \square

Notation. Henceforth, the quasi-derivative of u in the direction j will be denoted $\partial_j u$.

6.3 The space $H_0^1(\Omega)$ and the Dirichlet problem

In this section, we will define the appropriate function space for the minimisation problem introduced at the beginning of the chapter.

Definition 6.3.1. *Let Ω be an open set of \mathbb{R}^d . We denote $H_0^1(\Omega)$ the space of functions u in $L^2(\Omega)$ such that there exists a sequence $(u_n)_{n \in \mathbb{N}}$ of functions in $C_0^1(\Omega)$ such that*

$$\lim_{n \rightarrow \infty} \|u_n - u\|_{L^2(\Omega)} = 0$$

and such that the sequences $(\partial_j u_n)_{n \in \mathbb{N}}$ are Cauchy sequences in $L^2(\Omega)$.

The properties of this space are given in the following propositions.

Proposition 6.3.1. *If $u \in H_0^1(\Omega)$, then u has quasi-derivatives in all directions.*

Proof. Let $(u_n)_{n \in \mathbb{N}}$ be a sequence of functions in $C_0^1(\Omega)$ that converges in $L^2(\Omega)$ to u , and such that the sequences $(\partial_j u_n)_{n \in \mathbb{N}}$ are Cauchy sequences in $L^2(\Omega)$. As the space $L^2(\Omega)$, there exist functions $u^{(j)}$ in $L^2(\Omega)$ such that

$$\lim_{n \rightarrow \infty} \|\partial_j u_n - u^{(j)}\|_{L^2} = 0.$$

Integrating by parts, we can then write

$$\begin{aligned} - \int_{\Omega} u(x) \partial_j \varphi(x) dx &= - \lim_{n \rightarrow \infty} \int_{\Omega} u_n(x) \partial_j \varphi(x) dx \\ &= \lim_{n \rightarrow \infty} \int_{\Omega} \partial_j u_n(x) \varphi(x) dx \\ &= \int_{\Omega} u^{(j)}(x) \varphi(x) dx. \end{aligned}$$

This proves the proposition. \square

Proposition 6.3.2. *Endowed with the following norm,*

$$N(u) \stackrel{\text{def}}{=} \left(\|u\|_{L^2(\Omega)}^2 + \sum_{j=1}^d \|\partial_j u\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}},$$

the space $H_0^1(\Omega)$ is a Hilbert space.

Proof. The fact that N is associated with the inner product

$$(u|v) \stackrel{\text{def}}{=} \int_{\Omega} u(x)v(x)dx + \sum_{j=1}^d \int_{\Omega} \partial_j u(x)\partial_j v(x)dx$$

is an easy exercise we leave for the reader. Let us prove that the space $H_0^1(\Omega)$ is complete. Consider a Cauchy sequence $(u_n)_{n \in \mathbb{N}}$ of elements of $H_0^1(\Omega)$. As the space $L^2(\Omega)$ is complete, there exist functions u and $(u^{(j)})_{1 \leq j \leq d}$ such that

$$\lim_{n \rightarrow \infty} \|u_n - u\|_{L^2(\Omega)} = \lim_{n \rightarrow \infty} \|\partial_j u_n - u^{(j)}\|_{L^2(\Omega)} = 0. \quad (6.2)$$

By definition of the space $H_0^1(\Omega)$, for each n , there exists a function v_n in $C_0^1(\Omega)$ such that

$$\|u_n - v_n\|_{L^2(\Omega)} \leq \frac{1}{n} \quad \text{and} \quad \|\partial_j u_n - \partial_j v_n\|_{L^2(\Omega)} \leq \frac{1}{n}.$$

Thus, there exists a sequence $(v_n)_{n \in \mathbb{N}}$ in $C_0^1(\Omega)$ such that

$$\lim_{n \rightarrow \infty} \|v_n - u\|_{L^2(\Omega)} = \lim_{n \rightarrow \infty} \|\partial_j v_n - u^{(j)}\|_{L^2(\Omega)} = 0.$$

So $u \in H_0^1(\Omega)$, and $u^{(j)} = \partial_j u$. By (6.2), we have

$$\lim_{n \rightarrow \infty} N(u_n - u) = 0.$$

The proposition is proved. □

We are now going to examine the properties of the space $H_0^1(\Omega)$ specific to the case where the open set Ω is bounded.

Proposition 6.3.3. *Endowed with the norm*

$$\|u\|_{H_0^1(\Omega)}^2 \stackrel{\text{def}}{=} \|\nabla u\|_{L^2(\Omega)}^2 \stackrel{\text{def}}{=} \left(\sum_{j=1}^d \|\partial_j u\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}},$$

the space $H_0^1(\Omega)$ is a Hilbert space.

Proof. Compared with proposition 6.3.2, the only thing we need to show is that $\|\cdot\|_{H_0^1(\Omega)}$ is a norm that is equivalent to N . To do so, we only need to prove that, for any u in $H_0^1(\Omega)$,

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}.$$

Let $(u_n)_{n \in \mathbb{N}}$ be a sequence of $C_0^1(\Omega)$ such that

$$\lim_{n \rightarrow \infty} \|u_n - u\|_{L^2(\Omega)} = \lim_{n \rightarrow \infty} \|\partial_j u_n - \partial_j u\|_{L^2(\Omega)} = 0$$

for every j in $\{1, \dots, d\}$. The Poincaré inequality implies that, for any n , we have

$$\|u_n\|_{L^2(\Omega)} \leq C \|\nabla u_n\|_{L^2(\Omega)}.$$

Taking the limit, we get that, for any function u in $H_0^1(\Omega)$,

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}.$$

□

We have the following proposition, that the reader can prove as an exercise.

Proposition 6.3.4. *The functional F extends continuously to the space $H_0^1(\Omega)$, and we have*

$$m = \inf_{v \in C_0^1(\Omega)} F(v) = \inf_{v \in H_0^1(\Omega)} F(v)$$

This minimisation problem fits into the general, abstract setting: let \mathcal{H} be a Hilbert space, and ℓ be a continuous linear form on \mathcal{H} ; can we find the minimum of the function F defined by

$$F \begin{cases} \mathcal{H} & \longrightarrow \mathbb{R} \\ v & \longmapsto \frac{1}{2} \|v\|_{\mathcal{H}}^2 - \langle \ell, v \rangle \quad ? \end{cases}$$

Formulated this way, the problem is immediately solved by applying theorem 4.3.2.

Remark. We return to the original problem. The unique minimum of the function F defined on $H_0^1(\Omega)$ is the function u in $H_0^1(\Omega)$ that satisfies

$$\forall v \in H_0^1(\Omega), \quad \sum_{j=1}^d \int_{\Omega} \partial_j u(x) \partial_j v(x) dx = (u|v)_{H_0^1(\Omega)} = \int_{\Omega} f(x) v(x) dx = (f|v)_{L^2}. \quad (6.3)$$

Let us assume that the solution u happens to be of class C^2 on Ω . Then, for any function φ of class C^2 with compact support in Ω , we have

$$\sum_{j=1}^d \int_{\Omega} \partial_j u(x) \partial_j \varphi(x) dx = \int_{\Omega} f(x) \varphi(x) dx = (f|v)_{L^2}.$$

Integrating by parts, we get

$$\sum_{j=1}^d \int_{\Omega} \partial_j u(x) \partial_j \varphi(x) dx = \int_{\Omega} f(x) \varphi(x) dx = - \int_{\Omega} \Delta u(x) \varphi(x) dx.$$

When u is only in $H_0^1(\Omega)$, we will say that it solves $-\Delta u = f$ “in the weak sense”, or “in the distribution sense”.

The following compactness property, which will be proved in the next chapter, is crucial in proving the existence of a sequence of eigenvalues for the operator Δ .

Theorem 6.3.1 (Rellich). *Any bounded subset of $H_0^1(\Omega)$ has compact closure in $L^2(\Omega)$.*

We will now state a result that describes the structure of the Laplacian operator and its action on the space $H_0^1(\Omega)$.

Theorem 6.3.2. *There exist a non-decreasing sequence $(\lambda_j)_{j \in \mathbb{N}}$ of positive real numbers that tends to infinity, and an orthonormal (Hilbertian) basis of the space $L^2(\Omega)$, that we denote $(e_j)_{j \in \mathbb{N}}$, such that the sequence $(\lambda_j^{-1}e_j)_{j \in \mathbb{N}}$ is an orthonormal basis of $H_0^1(\Omega)$, and such that*

$$-\Delta e_j = \lambda_j^2 e_j \quad \text{"in the weak sense",}$$

which means that, for any function of class C^1 with compact support in Ω , we have

$$\sum_{k=1}^d \int_{\Omega} \partial_k e_j(x) \partial_k \varphi(x) dx = \lambda_j^2 \int_{\Omega} e_j(x) \varphi(x) dx.$$

Proof. Let us define an operator B by

$$B \begin{cases} L^2 & \longrightarrow H_0^1(\Omega) \subset L^2(\Omega) \\ f & \longmapsto u \end{cases}$$

such that u is the solution of the Dirichlet problem in $H_0^1(\Omega)$. The operator B is continuous from $L^2(\Omega)$ to $H_0^1(\Omega)$. By Rellich's theorem, theorem 6.3.1 above, the operator B is compact from $L^2(\Omega)$ to $L^2(\Omega)$. Let us prove that the operator B is self-adjoint. By (6.3), we have, for any f in L^2 ,

$$\forall v \in H_0^1(\Omega), (Bf|v)_{H_0^1(\Omega)} = (f|v)_{L^2}.$$

By applying this relation to $v = Bg$, for some g in L^2 , we deduce that

$$(f|Bg)_{L^2} = (Bf|Bg)_{H_0^1(\Omega)} \quad \text{and therefore} \quad (f|Bg)_{L^2} = (f|Bg)_{L^2}.$$

Thus, the operator B is self-adjoint. Let us prove that it is one to one. Consider a function f in L^2 such that $Bf = 0$. By (6.3) again, this means that

$$\forall v \in H_0^1(\Omega), \int_{\Omega} f(x)v(x) dx = 0.$$

Corollary 5.4.2 states that $\mathcal{D}(\Omega)$, and therefore $H_0^1(\Omega)$ which contains $\mathcal{D}(\Omega)$, is dense in $L^2(\Omega)$. We then deduce that if $Bf = 0$, then

$$\forall g \in L^2, \int_{\Omega} f(x)v(x) dx = 0,$$

which implies that $f = 0$. Finally, we observe that if λ is a non-zero eigenvalue, then it is positive. Indeed, there exists in that case a non-zero function f_{λ} in L^2 such that

$$\lambda \|f\|_{L^2}^2 = (Bf|f)_{L^2} = \|f\|_{H_0^1}^2.$$

Theorem 4.4.2 on page 75 ensure the existence of a non-increasing sequence $(\mu_j)_{j \in \mathbb{N}}$ of numbers that tends to 0 and an orthonormal basis $(e_j)_{j \in \mathbb{N}}$ of $L^2(\Omega)$ such that

$$Be_j = \mu_j^2 e_j.$$

By applying the Laplacian, we get

$$e_j = -\Delta Be_j = \mu_j^2 (-\Delta) e_j,$$

which means that

$$-\Delta e_j = \frac{1}{\mu_j^2} e_j.$$

To prove that the sequence $(\lambda_j^{-1}e_j)_{j \in \mathbb{N}}$ is an orthonormal basis of $H_0^1(\Omega)$, it suffices to observe that

$$\lambda_j^2 (e_j|e_k)_{L^2} = (e_j|e_k)_{H_0^1}.$$

Theorem 6.3.2 is proved. □

Chapter 7

The Fourier transform

Introduction

This chapter is fundamental, albeit brief. The Fourier transform is a very general operation on functions that are integrable with respect to the Haar measure (invariant by translation) on a locally compact commutative group G . Here, we will stick to the case of the space \mathbb{R}^d endowed with the Lebesgue measure.

In the first section of this chapter, we define the Fourier transform of an integrable function, and establish the main properties of this transform, which are:

- it turns differentiation into multiplication and vice-versa,
- it turns a convolution into a product.

Beyond these basic properties, we present a few examples in which we compute Fourier transforms. The fundamental example of the Gaussian function is of particular note.

In the second section, we prove the Fourier inversion theorem which states that, when the Fourier transform of an integrable function is itself integrable, we can recover the function from its Fourier transform. We deduce from this the Fourier-Plancherel theorem, which extends the Fourier transform to the space of square-integrable functions; this extension is, up to a constant, an isometry of this space.

The third section applies these results in order to prove Rellich's theorem used in the previous chapter (see theorem 6.3.1 on page 115).

7.1 The Fourier transform on $L^1(\mathbb{R}^d)$

Definition 7.1.1. Let f be a function in $L^1(\mathbb{R}^d)$. We call the Fourier transform of f , denoted \widehat{f} or $\mathcal{F}(f)$, the function defined by

$$\mathcal{F}(f) \begin{cases} \mathbb{R}^d & \longrightarrow \mathbb{C} \\ \xi & \longmapsto \int_{\mathbb{R}^d} e^{-i(\xi|x)} f(x) dx \end{cases} \quad \text{where} \quad (\xi|x) \stackrel{\text{def}}{=} \sum_{j=1}^d \xi_j x_j$$

Let us compute a few examples of Fourier transforms.

Proposition 7.1.1. *Let a be a positive real number. We have the following results:*

$$\begin{aligned}\mathcal{F}(\mathbf{1}_{[-a,a]})(\xi) &= \frac{2 \sin a\xi}{\xi}, \\ \mathcal{F}(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_{\pm}})(\xi) &= \frac{1}{a \pm i\xi}, \quad \text{so } \mathcal{F}(e^{-a|\cdot|})(\xi) = \frac{2a}{a^2 + \xi^2}, \\ \mathcal{F}\left(\frac{a}{a^2 + |\cdot|^2}\right)(\xi) &= \pi e^{-a|\xi|}, \\ \text{and } \mathcal{F}(e^{-a|\cdot|^2})(\xi) &= \left(\frac{\pi}{a}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4a}}.\end{aligned}$$

Proof. By definition of the Fourier transform, we have

$$\begin{aligned}\mathcal{F}(\mathbf{1}_{[-a,a]})(\xi) &= \int_{-a}^a e^{-i\xi x} dx \\ &= -\frac{1}{i\xi} (e^{-i\xi a} - e^{i\xi a}) \\ &= \frac{2 \sin a\xi}{\xi}.\end{aligned}$$

Again using the definition of the Fourier transform, we have

$$\begin{aligned}\mathcal{F}(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_+})(\xi) &= \int_0^{\infty} e^{-(i\xi+a)x} dx \\ &= \frac{1}{a + i\xi}.\end{aligned}$$

Likewise, we have $\mathcal{F}(e^{-a|\cdot|}\mathbf{1}_{\mathbb{R}_-})(\xi) = \frac{1}{a - i\xi}$. Adding these two results, we get

$$\mathcal{F}(e^{-a|\cdot|})(\xi) = \frac{2a}{a^2 + \xi^2}.$$

Still using the definition of the Fourier transform, we have

$$\begin{aligned}\mathcal{F}(e^{-a|\cdot|^2})(\xi) &= \lim_{R \rightarrow \infty} \int_{-R}^R e^{-i\xi x - ax^2} dx \\ &= e^{-\frac{\xi^2}{4a}} \lim_{R \rightarrow \infty} \int_{-R}^R e^{-a(x+i\frac{\xi}{2a})^2} dx.\end{aligned}$$

The function $z \mapsto e^{-az^2}$ is holomorphic on \mathbb{C} , hence

$$-\int_{-R}^R e^{-a(x+i\frac{\xi}{2a})^2} dx - \int_0^{\frac{\xi}{2a}} e^{-a(R-iy)^2} dy + \int_{-R}^R e^{-ax^2} dx + \int_0^{\frac{\xi}{2a}} e^{-a(R+iy)^2} dy = 0$$

By taking the limit as R tends to infinity, we find that

$$\int_{-\infty}^{\infty} e^{-a(x+i\frac{\xi}{2a})^2} dx = \int_{-\infty}^{\infty} e^{-ax^2} dx$$

It is well known that

$$\int_{\mathbb{R}} e^{-ax^2} dx = \left(\frac{\pi}{a}\right)^{\frac{1}{2}}.$$

This proves the result in dimension 1. For the case in dimension d , it suffices to observe that

$$e^{-|x|^2} = \prod_{j=1}^d e^{-x_j^2}$$

and that, if f_j are functions of $L^1(\mathbb{R})$, we have

$$\mathcal{F}(f_1 \otimes \cdots \otimes f_d) = \widehat{f}_1 \otimes \cdots \otimes \widehat{f}_d,$$

which ends the proof of this proposition. \square

Note that the continuity theorem for functions defined by integrals yields that the Fourier transform of an integrable function is continuous (and bounded) on \mathbb{R}^d .

One of the main properties of the Fourier transform is that it turns differentiation into multiplication and vice-versa, as shown by the following theorem.

Theorem 7.1.1. *Let f be a function on \mathbb{R}^d such that $(1 + |x|)f(x)$ defines an integrable function on \mathbb{R}^d . Then the Fourier transform of f is of class C^1 , and*

$$\partial_{\xi_j} \mathcal{F}(f) = -i \mathcal{F}(M_j f) \quad \text{where} \quad (M_j f)(x) \stackrel{\text{def}}{=} x_j f(x). \quad (7.1)$$

Let f be a function of class C^1 such that f and its partial derivatives are integrable on \mathbb{R}^d . Then we have

$$\mathcal{F}(\partial_{x_j} f) = i M_j \mathcal{F}(f). \quad (7.2)$$

Proof. The two results rely on the fact that

$$\partial_{\xi_j} e^{-i(\xi|x)} = -i x_j e^{-i(\xi|x)} \quad \text{and} \quad \partial_{x_j} e^{-i(\xi|x)} = -i \xi_j e^{-i(\xi|x)}. \quad (7.3)$$

In the first case, the hypothesis and the differentiability theorem for functions defined by integrals ensure that f has partial derivatives, as well as relation (7.1). The second case requires approximation. Let χ be a C^1 function with compact support such that $\chi(0) = 1$, and let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a sequence of positive numbers that converges to 0. Using (7.3), and integrating by parts, we get

$$\begin{aligned} \xi_j \mathcal{F}(\chi(\varepsilon_n \cdot) f)(\xi) &= -i \int_{\mathbb{R}^d} \partial_{x_j} (e^{-i(\xi|x)}) \chi(\varepsilon_n x) f(x) dx \\ &= -i \int_{\mathbb{R}^d} e^{-i(\xi|x)} \partial_{x_j} (\chi(\varepsilon_n x) f(x)) dx. \end{aligned}$$

Leibniz's law and the dominated convergence theorem yield the second formula. \square

This theorem gives us another way to prove the formula for Gaussian functions in proposition 7.1.1.

Proof. By proposition 7.1.2 below, applied to the case where the linear map A is a homothety with ration \sqrt{a} , it suffices to show the result when $a = 1$. Let us assume first that the dimension d is equal to 1, and set

$$g(\xi) \stackrel{\text{def}}{=} \int_{\mathbb{R}} e^{-i\xi x} e^{-x^2} dx.$$

Relation (7.1) in theorem 7.1.1 implies that the function g is differentiable, and we have

$$\begin{aligned} g'(\xi) &= \int_{\mathbb{R}} -ix e^{-i\xi x} e^{-x^2} dx \\ &= \int_{\mathbb{R}} i e^{-i\xi x} \frac{1}{2} (e^{-x^2})' dx. \end{aligned}$$

Applying relation (7.2) from theorem 7.1.1 yields

$$g'(\xi) = -\frac{\xi}{2} g(\xi).$$

This differential equation is easy to solve: we get $g(\xi) = g(0) e^{-\frac{\xi^2}{4}}$. A classical computation gives us

$$g(0) = \int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}.$$

So the result is proved in dimension 1. In general dimension d , we observe that

$$e^{-|x|^2} = \prod_{j=1}^d e^{-x_j^2}$$

and that, if the functions f_j are in $L^1(\mathbb{R})$, we have

$$\mathcal{F}(f_1 \otimes \cdots \otimes f_d) = \widehat{f}_1 \otimes \cdots \otimes \widehat{f}_d,$$

which ends the proof of the proposition. \square

We will now provide a corollary which describes the properties of the Fourier transforms of smooth functions with compact support. Its proof (which we omit) merely consists of repeated applications of the above theorem.

Corollary 7.1.1. *Let f be a function in $\mathcal{D}(\mathbb{R}^d)$. Its Fourier transform \widehat{f} is smooth, and satisfies the following:*

$$\forall N \in \mathbb{N}, \forall \alpha \in \mathbb{N}^d, \sup_{\xi \in \mathbb{R}^d} (1 + |\xi|)^N |\partial^\alpha \widehat{f}(\xi)| < \infty. \quad \text{where } \partial^\alpha \stackrel{\text{def}}{=} \partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d}.$$

We will now study the influence of linear transformations on the Fourier transform.

Proposition 7.1.2. *Let f be a function in $L^1(\mathbb{R}^d)$, and A be an invertible linear map from \mathbb{R}^d to \mathbb{R}^d . We have*

$$\mathcal{F}(f \circ A)(\xi) = |\det A|^{-1} \widehat{f} \circ {}^t A^{-1}.$$

Proof. We perform the change of variable $y = Ax$ in the integral that defines the Fourier transform; this yields

$$\mathcal{F}(f \circ A)(\xi) = |\det A|^{-1} \int_{\mathbb{R}^d} e^{-i(\xi | A^{-1} x)} f(x) dx.$$

By definition of the transpose, we have $(\xi | A^{-1} x) = ({}^t A^{-1} \xi | x)$, hence the proposition. \square

From theorem 7.1.1 and proposition 7.1.2 above, we will deduce the next proposition, which shows how to compute the Fourier transform of a Gaussian function. This result will play a crucial role proving the inversion formula.

Proposition 7.1.3. *Let a be a positive real number. We have*

$$\mathcal{F}\left(e^{-a|\cdot|^2}\right)(\xi) = \left(\frac{\pi}{a}\right)^{\frac{d}{2}} e^{-\frac{|\xi|^2}{4a}}.$$

The following proposition gives some ideas on how the Fourier transform works on convolutions.

Proposition 7.1.4. *Let f and g be two functions in $L^1(\mathbb{R}^d)$. We have*

$$\mathcal{F}(f \star g) = \widehat{f} \widehat{g}.$$

Proof. For every ξ in \mathbb{R}^d , the function F_ξ from $\mathbb{R}^d \times \mathbb{R}^d$ to \mathbb{C} defined by

$$F_\xi(x, y) \stackrel{\text{def}}{=} e^{-i(\xi|x)} f(x-y)g(y) = e^{-i(\xi|x-y)} f(x-y)e^{-i(\xi|y)} g(y)$$

is an element of $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. Fubini's theorem ensures that

$$\int_{\mathbb{R}^d} e^{-i(\xi|x)} \left(\int_{\mathbb{R}^d} f(x-y)g(y)dy \right) dx = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} e^{-i(\xi|x-y)} f(x-y)dx \right) e^{-i(\xi|y)} g(y)dy,$$

which shows precisely that the Fourier transform of a convolution is the product of the Fourier transforms. \square

The Fourier transform is symmetric in the sense given by the following proposition.

Proposition 7.1.5. *For any pair (f, g) of functions in $L^1(\mathbb{R}^d)$, we have*

$$\int_{\mathbb{R}^d} \widehat{f}(x)g(x)dx = \int_{\mathbb{R}^d} f(\xi)\widehat{g}(\xi)d\xi.$$

Proof. This is an immediate consequence of Fubini's theorem applied to the integrable function on $\mathbb{R}^d \times \mathbb{R}^d$ defined by $(x, \xi) \mapsto f(\xi)e^{-i(\xi|x)}g(x)$. \square

7.2 The inversion formula and the Fourier-Plancherel theorem

We are now going to prove the fundamental theorem of Fourier analysis.

Theorem 7.2.1. *Let f be a function in $L^1(\mathbb{R}^d)$ such that its Fourier transform is also a function in $L^1(\mathbb{R}^d)$. Then, the function f is continuous, and we have*

$$\forall x \in \mathbb{R}, f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|x)} \widehat{f}(\xi) d\xi, \quad (7.4)$$

as well as the Fourier-Plancherel relation

$$\|\widehat{f}\|_{L^2}^2 = (2\pi)^d \|f\|_{L^2}^2. \quad (7.5)$$

Remarks.

- The result can be rewritten as

$$\mathcal{F}^2 f = (2\pi)^d \check{f} \quad \text{where} \quad \check{f}(x) \stackrel{\text{def}}{=} f(-x). \quad (7.6)$$

- This theorem means that a function f in \mathcal{S} (defined in the next chapter) can be written as a superposition of oscillatory functions, the $e^{i(\xi|x)}$, and the Fourier transform appears as the density of these oscillations.

Proof of theorem 7.2.1. By proposition 7.1.3, the theorem is already proved for Gaussian functions. Of course, we are going to use this result. Set

$$G(x) \stackrel{\text{def}}{=} \pi^{-\frac{d}{2}} e^{-|x|^2} \quad \text{and} \quad G_\varepsilon(x) = \frac{1}{\varepsilon^d} G\left(\frac{x}{\varepsilon}\right).$$

Proposition 7.1.3 states that $\widehat{G}_\varepsilon(\xi) = e^{-\frac{\varepsilon^2|\xi|^2}{4}}$. By the dominated convergence theorem, we deduce from the fact that \widehat{f} belongs to L^1 that

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|x)} \widehat{f}(\xi) d\xi = \lim_{\varepsilon \rightarrow 0} (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|x)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi. \quad (7.7)$$

As the function f belongs to L^1 , we know that $\widehat{G}_\varepsilon \otimes f$ is a function in $L^1(\mathbb{R}^d \times \mathbb{R}^d)$. By Fubini's theorem, we have

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|x)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi = (2\pi)^{-d} \int_{\mathbb{R}^d \times \mathbb{R}^d} e^{i(\xi|x)} \widehat{G}_\varepsilon(\xi) e^{-i(\xi|y)} f(y) d\xi dy.$$

Proposition 7.1.3 applied with $a = \varepsilon^2/4$ states, in particular, that

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|y-x)} \widehat{G}_\varepsilon(\xi) d\xi = (2\pi)^{-d} \left(\frac{4\pi}{\varepsilon^2}\right)^{\frac{d}{2}} e^{-\frac{|x-y|^2}{\varepsilon^2}} = G_\varepsilon(x-y).$$

Thus, we get that

$$\forall \varepsilon > 0, (G_\varepsilon \star f)(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(x|\xi)} \widehat{G}_\varepsilon(\xi) \widehat{f}(\xi) d\xi.$$

Theorem 5.4.5 on page 100 states that if $(\varepsilon_n)_{n \in \mathbb{N}}$ is a sequence of positive real numbers converging to 0, then $\lim_{n \rightarrow \infty} \|G_{\varepsilon_n} \star f - f\|_{L^1(\mathbb{R}^d)} = 0$. Theorem 5.1.2 on page 83 adds that, up to the extraction of a subsequence that we do not denote differently, for almost every x in \mathbb{R}^d , we have

$$\lim_{n \rightarrow \infty} (G_{\varepsilon_n} \star f)(x) = f(x).$$

Statement (7.7) then ensures that, for almost every x in \mathbb{R}^d ,

$$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{i(\xi|x)} \widehat{f}(\xi) d\xi.$$

The continuity theorem for functions defined by integrals ends the proof of (7.4), as it yields that the right-hand side of the above equality, and therefore f , is continuous.

To prove (7.5), first observe that the fact that f and \widehat{f} are in $L^1(\mathbb{R}^d)$ implies that f is bounded, and thus both f and \widehat{f} are square-integrable. Noting that $\overline{\mathcal{F}(f)} = \mathcal{F}(\overline{f})$ and by using proposition 7.1.5, we can write that

$$\begin{aligned} (2\pi)^{-d} \|\widehat{f}\|_{L^2}^2 &= \int_{\mathbb{R}^d} (2\pi)^{-d} (\mathcal{F}f)(\xi) (\mathcal{F}(\overline{f}))(\xi) d\xi \\ &= \int_{\mathbb{R}^d} (2\pi)^{-d} (\mathcal{F}^2 f)(x) \overline{f}(-x) dx. \end{aligned}$$

The Fourier inversion formula as stated in (7.6) then ensures that

$$(2\pi)^{-d} \|\widehat{f}\|_{L^2}^2 = \int_{\mathbb{R}^d} f(-x) \overline{f}(-x) dx = \|f\|_{L^2}^2,$$

which ends the proof of the fundamental theorem of Fourier analysis. \square

This theorem, at the base of Fourier analysis, has countless applications. Let us provide two corollaries.

Corollary 7.2.1. *The Fourier transform extends into a continuous, invertible linear map from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$, which satisfies*

$$\forall f \in L^2(\mathbb{R}^d), \|\widehat{f}\|_{L^2(\mathbb{R}^d)} = (2\pi)^{\frac{d}{2}} \|f\|_{L^2(\mathbb{R}^d)}.$$

Proof. The Fourier transform is well defined on $\mathcal{D}(\mathbb{R}^d)$, which, by corollary 5.4.2 on page 102, is a dense subspace of $L^2(\mathbb{R}^d)$. Corollary 7.1.1 allows us to apply the Fourier-Plancherel relation (7.5) to functions in $\mathcal{D}(\mathbb{R}^d)$. The extension theorem 1.2.4 on page 17 then allows us to extend \mathcal{F} into a continuous linear map from $L^2(\mathbb{R}^d)$ to itself, and the Fourier-Plancherel relation naturally extends to the whole L^2 space. The only point left to prove is that \mathcal{F} is onto. Proposition 7.1.5 implies, in particular, that

$$\forall (f, g) \in \mathcal{D}(\mathbb{R}^d), (\mathcal{F}f|g)_{L^2(\mathbb{R}^d)} = (f|\check{\mathcal{F}}g)_{L^2(\mathbb{R}^d)}. \quad (7.8)$$

To prove that \mathcal{F} is onto, we first observe that the Fourier-Plancherel equality implies that $\mathcal{F}(L^2)$ is closed in L^2 , then we prove that $\mathcal{F}(L^2)$ is dense by using criterion 4.2.2 on page 68, which states that we only need to show that $(\mathcal{F}(L^2))^\perp$ consists only of the zero function. Indeed, if g is a function in $L^2(\mathbb{R}^d)$ such that

$$\forall f \in L^2(\mathbb{R}^d), (\mathcal{F}(f)|g)_{L^2(\mathbb{R}^d)} = 0,$$

then relation (7.8), extended to the whole L^2 space, implies that $\mathcal{F}g$ is equal to zero, and thus so is g , since \mathcal{F} is one to one. \square

Corollary 7.2.2. *Let f and g be two functions in $L^1(\mathbb{R}^d)$ with Fourier transforms also in $L^1(\mathbb{R}^d)$. Then we have*

$$\mathcal{F}(fg) = (2\pi)^{-d} \widehat{f} \star \widehat{g}.$$

Proof. We apply proposition 7.1.4 to \widehat{f} and \widehat{g} . This yields

$$\mathcal{F}(\widehat{f} \star \widehat{g}) = \mathcal{F}^2 f \mathcal{F}^2 g = (2\pi)^{2d} \check{f} \check{g},$$

which can be rewritten as

$$(\widehat{f} \star \widehat{g}) = (2\pi)^{2d} \mathcal{F}^{-1}(\check{f} \check{g}).$$

The Fourier inversion formula in theorem 7.2.1 ensures the result. \square

7.3 Proof of Rellich's theorem

The goal of this section is to prove the Rellich's compactness theorem 6.3.1, which we recall. This proof is presented here for cultural purposes only.

Theorem 7.3.1. *Any bounded subset of $H_0^1(\Omega)$ has compact closure in $L^2(\Omega)$.*

Proof. We will prove that, for any positive number α , we can cover a bounded part A of $H_0^1(\Omega)$ with a finite number of balls with radius α in the L^2 norm. To do so, we start by proving that if χ is a function in $\mathcal{D}(B(0, 1))$ with integral equal to 1, then

$$\left\| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a - a \right\|_{L^2} \leq C_\chi \varepsilon \|\nabla a\|_{L^2} \quad (7.9)$$

where C_χ is a constant that depends on χ . By the Fourier-Plancherel relation (theorem 7.2.1 on page 121) and by theorem 8.2.5, we have

$$\begin{aligned} \left\| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a - a \right\|_{L^2} &= (2\pi)^{\frac{d}{2}} \left\| \mathcal{F}\left(\frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a - a\right) \right\|_{L^2} \\ &= (2\pi)^{\frac{d}{2}} \left\| \widehat{\chi}(\varepsilon \cdot) \widehat{a} - \widehat{a} \right\|_{L^2}. \end{aligned}$$

The mean value inequality and theorem 7.1.1 imply that

$$\begin{aligned} |\widehat{\chi}(\zeta) - 1| &\leq |\zeta| \|D\widehat{\chi}\|_{L^\infty} \\ &\leq |\zeta| \max_{1 \leq j \leq d} \|x_j \chi\|_{L^1(\mathbb{R}^d)}. \end{aligned}$$

Thus, we deduce that

$$\left\| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a - a \right\|_{L^2} \leq C_\chi \varepsilon \| |\cdot| \widehat{a} \|_{L^2}.$$

Proposition 7.1.1 and the Fourier-Plancherel relation then yield inequality (7.9). We will now prove that, for any ε in the interval $]0, 1[$, the set

$$\mathcal{B}_\varepsilon \stackrel{\text{def}}{=} \left\{ \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a, a \in B_{H_0^1(\Omega)}(0, 1) \right\}$$

is a subset of $\mathcal{C}(\overline{\Omega} + B(0, 1))$ with compact closure. To do so, we first observe that, by the Poincaré inequality, we have, for any x in $\overline{\Omega} + B(0, 1)$,

$$\begin{aligned} \left| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a(x) \right| &\leq \left\| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \right\|_{L^2} \|a\|_{L^2} \\ &\leq \frac{C}{\varepsilon^{\frac{d}{2}}} \|\chi\|_{L^2} \|\nabla a\|_{L^2}. \end{aligned} \tag{7.10}$$

Moreover, by the mean value inequality, we have

$$\begin{aligned} \left| \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a(x) - \frac{1}{\varepsilon^d} \chi\left(\frac{\cdot}{\varepsilon}\right) \star a(x') \right| &\leq |x - x'| \left\| \frac{1}{\varepsilon^d} D\left(\chi\left(\frac{\cdot}{\varepsilon}\right) \star a\right) \right\|_{L^\infty} \\ &\leq |x - x'| \varepsilon^{-\frac{d}{2}} \|\chi\|_{L^2} \|\nabla a\|_{L^2}. \end{aligned}$$

For any ε in the interval $]0, 1[$, the set \mathcal{B}_ε is an equicontinuous subset of $\mathcal{C}(\overline{\Omega} + B(0, 1))$. By inequality (7.10) and Ascoli's theorem (see theorem 2.4.1 on page 40), the set \mathcal{B}_ε has compact closure in $\mathcal{C}(\overline{\Omega} + B(0, 1))$. It can thus be covered by a finite number of balls (for the norm inducing uniform convergence on $\overline{\Omega} + B(0, 1)$) with radius

$$\frac{\alpha}{(\mu(\overline{\Omega} + B(0, 1)))^{\frac{1}{2}}}.$$

As, for any bounded function f supported in $\overline{\Omega} + B(0, 1)$, we have

$$\|f\|_{L^2} \leq (\mu(\overline{\Omega} + B(0, 1)))^{\frac{1}{2}} \|f\|_{L^\infty},$$

we have proved that \mathcal{B}_ε has compact closure in $L^2(\Omega)$, owing to (7.9). \square

Chapter 8

Tempered distributions in dimension 1

Introduction

This chapter is one of the endgames of this course. As we saw in chapter 3, identifying a Banach space as the dual of another via a continuous bilinear form (see definition 3.2.1 on page 54) is an important ability that allows one to extract subsequences from any bounded sequence that converge in a weaker sense; we called this sense weak- \star convergence.

When we studied the spaces of p -th power integrable functions in chapter 5, we saw that understanding families of weighted means,

$$\left(\int_{\Omega} f(x)\varphi(x)dx \right)_{\varphi \in A},$$

e.g. with φ in a dense subset A , could sometimes be better illustrations, or yield information that is easier to work with, than the pointwise values of the function f . This way, functions appear as continuous linear functionals on some function space. This is how we can generalise the notion of function. For this chapter, we will assume the point of view of tempered distributions, which has a global vision on the whole space \mathbb{R} . It is particularly adapted to extend the Fourier transform to a larger class of maps, beyond the setting of functions in L^1 or L^2 seen in the previous chapter. For this first contact with distributions, we keep things simple by restricting our study to dimension 1.

In the first section, we will define the concept of tempered distributions and see how it generalises that of locally integrable functions (modulo a small restriction at infinity). For this purpose, theorem 8.1.1 will eloquently show how the knowledge of the family

$$\left(\int_{\Omega} f(x)\varphi(x)dx \right)_{\varphi \in E}$$

for a certain vector space E is equivalent to knowing the function in its usual sense (i.e. its values at all points of \mathbb{R}). After this, several examples of tempered distributions which are not functions in the classical sense will be given. Finally, we will see how the notion of weak- \star convergence leads us to an adapted sense of convergence for distributions (see proposition 8.1.6).

In the second section, we define the operations that are possible on distributions by transposing. This will lead to a striking new fact: all distributions have derivatives (!) and all tempered distributions have a Fourier transform. What follows this is a series of various formulae that can be quite surprising.

In the third section, we show two applications of this theory. The first will involve a classical analysis problem which is difficult to solve in the strict sense of functions, but which requires only a few lines using distributions. Then, we will establish an integral formula which solves the equation

$$u - u'' = f$$

with u equal to 0 at $-\infty$ and $+\infty$.

8.1 Definition of tempered distributions and examples

As we explained in the introduction of this chapter, the concept of distributions is based on the fact that a function can be entirely described by a collection of its weighted means. We can thus rethink the notion of a function in terms of linear forms on a function space: this function space will be called the space of test functions. How should we choose this space of test functions? If we refer to the L^p spaces we studied, the “smallest” function space we encountered was $\mathcal{D}(\mathbb{R})$, the space of smooth functions with compact support. Indeed, if f is in $L^2(\mathbb{R})$ for example, and

$$\forall \varphi \in \mathcal{D}(\mathbb{R}), \int_{\mathbb{R}} f(x)\varphi(x)dx = 0,$$

then the function f is equal to zero in $L^2(\mathbb{R})$, which means that it is equal to zero almost everywhere. This shows that $\mathcal{D}(\mathbb{R})$ is a good choice for the space of test functions, and this is the one that applies in the general theory of distributions.

For reasons that will become apparent shortly, we would like the space of test functions to be stable under the Fourier transform. As we will see, this is not the case for the space $\mathcal{D}(\mathbb{R})$ of smooth functions with compact support. Indeed, let φ be a function in $\mathcal{D}(\mathbb{R})$. The function

$$\begin{cases} \mathbb{C} & \longrightarrow \mathbb{C} \\ \zeta & \longmapsto \int_{\mathbb{R}} e^{-i\zeta x} \varphi(x) dx \end{cases}$$

is well defined because φ has compact support, and the differentiability theorem for functions of the complex variable defined by integrals yields that it is holomorphic (because it is differentiable) on \mathbb{C} . As the zeros of a holomorphic function are isolated, this function cannot coincide with a function with compact support on the real line unless this function is zero, which, by the Fourier inversion theorem, yields that φ is zero.

So what properties does the Fourier transform of a function in $\mathcal{D}(\mathbb{R})$ have? Corollary 7.1.1 on page 120 implies that

$$\sup_{\xi \in \mathbb{R}^d} (1 + |\xi|)^N \left| \frac{d}{d\xi} \widehat{\varphi}(\xi) \right| < \infty.$$

This leads to defining the following space.

Definition 8.1.1. We denote $\mathcal{S}(\mathbb{R})$, or simply \mathcal{S} , the set of smooth functions f from \mathbb{R} to \mathbb{R} such that

$$\forall n \in \mathbb{N}, \|f\|_{n,\mathcal{S}} \stackrel{\text{def}}{=} \max_{k \leq n} \sup_{x \in \mathbb{R}} (1 + |x|)^n |f^{(k)}(x)| < \infty.$$

Observe first of all that the space $\mathcal{S}(\mathbb{R})$ contains the space $\mathcal{D}(\mathbb{R})$ of smooth functions with compact support. Note also that the Gaussian functions, defined by

$$x \longmapsto e^{-\frac{(x-x_0)^2}{a^2}}$$

are elements of \mathcal{S} .

Tempered distributions will appear as linear functionals that satisfy a form of “continuity” property on this space. The following is the precise definition.

Definition 8.1.2. We call a tempered distribution on \mathbb{R} a linear functional defined on the space $\mathcal{S}(\mathbb{R})$ that satisfies

$$\exists n \in \mathbb{N}, \exists C / \forall \varphi \in \mathcal{S}(\mathbb{R}), |\langle T, \varphi \rangle| \leq C \|\varphi\|_{n, \mathcal{S}}. \quad (8.1)$$

We denote $\mathcal{S}'(\mathbb{R})$ the set of such linear forms.

We are now going to give some examples of tempered distributions.

Definition 8.1.3. We denote $L^1_{\mathcal{M}}(\mathbb{R})$ the set of locally integrable functions for which there exists an integer N such that $(1 + |x|)^{-N} f(x)$ is in $L^1(\mathbb{R})$.

Exercise 8.1.1. Prove that the spaces $L^p(\mathbb{R})$ are all subsets of $L^1_{\mathcal{M}}(\mathbb{R})$.

Exercise 8.1.2. Prove that the functions

$$x \mapsto e^x \quad \text{and} \quad x \mapsto e^{x^2}$$

are not elements of $L^1_{\mathcal{M}}(\mathbb{R})$

Proposition 8.1.1. Let f be a function in $L^1_{\mathcal{M}}(\mathbb{R})$. The linear form defined by

$$\iota \left\{ \begin{array}{l} \mathcal{S}(\mathbb{R}) \longrightarrow \mathbb{C} \\ \varphi \longmapsto \int_{\mathbb{R}} f(x)\varphi(x)dx \end{array} \right.$$

is a distribution on \mathbb{R} .

Proposition 8.1.2. Let a be a point in \mathbb{R} . The linear forms

$$\left\{ \begin{array}{l} \mathcal{S}(\mathbb{R}) \longrightarrow \mathbb{C} \\ \phi \longmapsto \phi(a) \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \mathcal{S}(\mathbb{R}) \longrightarrow \mathbb{C} \\ \phi \longmapsto \sum_{n \in \mathbb{N}} \phi(n) \end{array} \right.$$

are tempered distributions.

Proposition 8.1.3. Let ϕ be a function on \mathbb{C} . Then the function

$$\left\{ \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{C} \\ \varphi \longmapsto \frac{\varphi(x) - \varphi(-x)}{x} \end{array} \right.$$

is integrable, and the linear form defined by

$$\left\langle \text{vp} \frac{1}{x}, \varphi \right\rangle = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\varphi(x) - \varphi(-x)}{x} dx$$

called the Cauchy principal value of $1/x$, is a tempered distribution.

Proof. By the mean value inequality, we have

$$|\varphi(x) - \varphi(-x)| \leq 2|x| \sup_{x \in \mathbb{R}} |\varphi'(x)|.$$

By definition of the semi-norms on \mathcal{S} , we have

$$|\varphi(x) - \varphi(-x)| \leq \frac{2}{1+|x|} \|\varphi\|_{1,\mathcal{S}}.$$

Thus,

$$\begin{aligned} \forall x \in \mathbb{R} / |x| \leq 1, \quad & \left| \frac{\varphi(x) - \varphi(-x)}{x} \right| \leq 2\|\varphi\|_{1,\mathcal{S}}, \\ \text{and } \forall x \in \mathbb{R} / |x| \geq 1, \quad & \left| \frac{\varphi(x) - \varphi(-x)}{x} \right| \leq \frac{2}{|x|^2} \|\varphi\|_{1,\mathcal{S}}. \end{aligned}$$

The function $x \mapsto \frac{\varphi(x) - \varphi(-x)}{x}$ is therefore integrable on \mathbb{R} , and the linear form $\text{vp} \frac{1}{x}$ is continuous on \mathcal{S} . \square

Remark. Let φ be a test function that is identically zero on a neighbourhood of 0. Then,

$$\left\langle \text{vp} \frac{1}{x}, \varphi \right\rangle = \int_{\mathbb{R}} \frac{\varphi(x)}{x} dx.$$

This means that we have defined an “extension”, of sorts, of the function x^{-1} .

Proposition 8.1.4. *Let a be in \mathbb{R} and k be in \mathbb{N} . The linear form defined by*

$$\langle \delta_a^{(k)}, \varphi \rangle \stackrel{\text{def}}{=} (-1)^k \varphi^{(k)}(a)$$

is a tempered distribution.

Proof. It suffices to observe that, by definition of the semi-norms on \mathcal{S} , we have

$$|\varphi^{(k)}(a)| \leq C_a \|\varphi\|_{k,\mathcal{S}},$$

which yields the result. \square

Proposition 8.1.5. *The linear form defined by*

$$\left\langle \text{Pf} \frac{1}{x^2}, \varphi \right\rangle = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\varphi(x) + \varphi(-x) - 2\varphi(0)}{x^2} dx$$

called the Hadamard finite part of $1/x^2$, is a tempered distribution.

Proof. The second-order Taylor inequality ensures that

$$|\varphi(x) + \varphi(-x) - 2\varphi(0)| \leq \frac{|x|^2}{2} \sup_{x \in \mathbb{R}} |\varphi''(x)|.$$

Thus, we have

$$\frac{\varphi(x) + \varphi(-x) - 2\varphi(0)}{x^2} \leq \min\left\{1, \frac{1}{|x|^2}\right\} \|\varphi\|_{2,\mathcal{S}}.$$

This yields that $\text{Pf} \frac{1}{x^2}$ is a tempered distribution. \square

Remark. If φ is a test function that is identically zero in a neighbourhood of 0, we have

$$\left\langle \text{Pf} \frac{1}{x^2}, \varphi \right\rangle = \int_{\mathbb{R}} \frac{\varphi(x)}{x^2} dx.$$

This means that we have defined an “extension” of the function x^{-2} .

The following theorem shows that the space $L^1_{\mathcal{M}}(\mathbb{R})$ can be identified as a subspace of the set of tempered distributions.

Theorem 8.1.1. *Let ι be the map defined by*

$$\iota \begin{cases} L^1_{\mathcal{M}}(\mathbb{R}) & \longrightarrow \mathcal{S}'(\mathbb{R}) \\ f & \longmapsto \iota(f) : \varphi \mapsto \int_{\mathbb{R}} f(x)\varphi(x)dx \end{cases}$$

is a linear injection. Moreover, we have the following property: if N is such that $(1+|x|)^{-N}f(x)$ is in $L^1(\mathbb{R})$, then

$$|\langle \iota(f), \varphi \rangle| \leq C \|(1+|\cdot|)^{-N}f\|_{L^1} \|\varphi\|_{N, \mathcal{S}}.$$

Proof. Let f be a function in $L^1_{\mathcal{M}}(\mathbb{R})$ such that $\iota(f) = 0$. We are going to prove that, for any pair of real numbers (a, b) with a strictly less than b , we have

$$\int_a^b |f(x)| dx = 0$$

which is enough to prove the theorem. We consider an even function χ in $\mathcal{D}(-1, 1)$ with integral equal to 1, and the associated sequence of mollifiers, that is the family

$$\chi_{\varepsilon}(x) \stackrel{\text{def}}{=} \varepsilon^{-d} \chi\left(\frac{x}{\varepsilon}\right).$$

Finally, we define a function g by

$$g(x) = \frac{\overline{f(x)}}{|f(x)|} \quad \text{if } f(x) \neq 0, \quad \text{and } 0 \text{ otherwise.}$$

This function is an element of $L^{\infty}(\mathbb{R})$. The function

$$\varphi_{\varepsilon} \stackrel{\text{def}}{=} \chi_{\varepsilon} \star (\mathbf{1}_{[a,b]}g)$$

is a smooth function with compact support on \mathbb{R} , and therefore a function in $\mathcal{S}(\mathbb{R})$. By assumption, we have

$$I_{\varepsilon} \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x)\varphi_{\varepsilon}(x)dx = 0.$$

For any ε in the interval $]0, \delta[$, the function

$$(x, y) \longmapsto \varepsilon^{-d} \chi\left(\frac{x-y}{\varepsilon}\right) \mathbf{1}_{[a,b]}(y)g(y)f(x) = \varepsilon^{-d} \chi\left(\frac{x-y}{\varepsilon}\right) \mathbf{1}_{[a,b]}(y)g(y)f(x) \mathbf{1}_{[a-\delta, b+\delta]}(x)$$

is in $L^1(\mathbb{R} \times \mathbb{R})$, because it is bounded with compact support. Indeed, the modulus of this function is bounded by 1, and, by Fubini’s theorem, theorem 5.1.5 on page 83, for every ε smaller than δ ,

$$\int_{\mathbb{R} \times \mathbb{R}} \varepsilon^{-d} \left| \chi\left(\frac{x-y}{\varepsilon}\right) \right| \mathbf{1}_{[a,b]}(y) |g(y)| \mathbf{1}_{B(0,\delta)}(x) |f(x)| dx dy \leq \|\chi\|_{L^1} \|\mathbf{1}_{[a-\delta, b+\delta]}f\|_{L^1}.$$

The Fubini-Tonelli theorem 5.1.6 on page 84, combined with the evenness of the function χ , yields that

$$\begin{aligned}
I_\varepsilon &= \int_{\mathbb{R} \times \mathbb{R}} \chi_\varepsilon(x-y) \mathbf{1}_{[a,b]}(y) g(y) f(x) dx dy \\
&= \int_{\mathbb{R} \times \mathbb{R}} \chi_\varepsilon(x-y) \mathbf{1}_{[a,b]}(y) g(y) \mathbf{1}_{K+B(0,\delta)}(x) f(x) dx dy \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_\varepsilon(y-x) \mathbf{1}_{[a-\delta,b+\delta]}(x) f(x) dx \right) \mathbf{1}_K(y) g(y) dy \\
&= \int_{\mathbb{R}} (\chi_\varepsilon \star (\mathbf{1}_{[a-\delta,b+\delta]} f))(y) \mathbf{1}_{[a,b]}(y) g(y) dy.
\end{aligned}$$

Theorem 5.4.5 on page 100 states that

$$\lim_{\varepsilon \rightarrow 0} \|(\chi_\varepsilon \star (\mathbf{1}_{[a-\delta,b+\delta]} f)) - \mathbf{1}_{[a-\delta,b+\delta]} f\|_{L^1} = 0$$

As g belongs to $L^\infty(\mathbb{R})$, we have

$$\begin{aligned}
0 &= \lim_{\varepsilon \rightarrow 0} I_\varepsilon \\
&= \int \mathbf{1}_{[a-\delta,b+\delta]}(y) f(y) \mathbf{1}_{[a,b]}(y) g(y) dy \\
&= \int_a^b f(y) \frac{\bar{f}(y)}{|f(y)|} dy \\
&= \int_a^b |f(y)| dy.
\end{aligned}$$

The theorem is proved. □

Remark. This way, any function in $L^1_{\mathcal{M}}$ can be identified as a tempered distribution, via the following definition.

Definition 8.1.4. We say that a tempered distribution T is a function if and only if there exists a function f in $L^1_{\mathcal{M}}$ such that $T = \iota(f)$.

We shall now define the notion of convergence for a sequence of tempered distributions. It is a weak- \star -type convergence in the sense that we defined in chapter 3.

Definition 8.1.5. Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of elements of $\mathcal{S}'(\mathbb{R})$, and T be an element of $\mathcal{S}'(\mathbb{R})$. We say that the sequence $(T_n)_{n \in \mathbb{N}}$ converges to T if and only if

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \quad \lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle = \langle T, \varphi \rangle.$$

We now give an example of a convergent sequence in this sense, that will show how far away this notion of convergence for tempered distributions is from the one we use for functions.

Proposition 8.1.6. The sequence of functions $S_N(x) \stackrel{\text{def}}{=} \sum_{n=1}^N \sin(nx)$ is convergent in $\mathcal{S}'(\mathbb{R})$.

Proof. We are looking for a tempered distribution S such that, for any function f in \mathcal{S} , we have

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N \int_{\mathbb{R}} \sin(nx) \varphi(x) dx = \langle S, \varphi \rangle.$$

Integrating by parts twice, we get that, for any integer n greater than or equal to 1,

$$\int_{\mathbb{R}} \sin(nx)\varphi(x)dx = -\frac{1}{n^2} \int_{\mathbb{R}} \sin(nx)\varphi''(x)dx. \quad (8.2)$$

We define a function Σ by

$$\Sigma(x) = -\sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nx)$$

This is a continuous, bounded function on \mathbb{R} . We then set, for f in \mathcal{S} ,

$$\langle \mathcal{S}, \varphi \rangle \stackrel{\text{def}}{=} \int_{\mathbb{R}} \Sigma(x)\varphi''(x)dx.$$

This defines a tempered distribution. Moreover, for any function φ , by (8.2), we have

$$\lim_{N \rightarrow \infty} \sum_{n=1}^N \int_{\mathbb{R}} \sin(nx)\varphi(x)dx = \int_{\mathbb{R}} \Sigma(x)\varphi''(x)dx = \langle \mathcal{S}, \varphi \rangle.$$

The proposition is proved. □

8.2 Operations on tempered distributions

The big idea is the following: we have seen that the space \mathcal{S}' is much “bigger” than the space \mathcal{S} ; on the functions in \mathcal{S} , we can perform a large number of operations, including derivation, the Fourier transform, convolution with a function with controlled growth. For the last one, we need to define the notion of continuous linear map from \mathcal{S} to \mathcal{S} .

Definition 8.2.1. *A linear map A from \mathcal{S} to itself is said to be continuous if and only if*

$$\forall k \in \mathbb{N}, \exists (C_k, n_k) \in]0, \infty[\times \mathbb{N} / \forall \phi \in \mathcal{S}, \|A\phi\|_{k, \mathcal{S}} \leq C_k \|f\|_{n_k, \mathcal{S}}.$$

Let us introduce the following two function spaces.

Definition 8.2.2. *The space of functions of moderate (or slow) growth, denoted \mathcal{O}_M , is the set of functions f that are smooth on \mathbb{R} and that satisfy*

$$\forall k \in \mathbb{N}, \exists N \in \mathbb{N}, \exists C \in \mathbb{R}^+ / \forall x \in \mathbb{R}, |f^{(k)}(x)| \leq (1 + |x|)^N.$$

We denote $L_{\mathcal{S}}^1$ the space of locally integrable functions that satisfy, for any integer N , that the function $(1 + |x|)^N f(x)$ is in L^1 .

Polynomials are excellent examples of functions in \mathcal{O}_M . Locally integrable functions that converge to zero faster than any negative power of $|x|$ as x tends to infinity are excellent examples of functions in $L_{\mathcal{S}}^1$. For example, the function

$$x \mapsto \frac{1}{x^\alpha} e^{-|x|^\beta},$$

with α in $]0, 1[$ and β positive, belongs to $L_{\mathcal{S}}^1$.

Proposition 8.2.1. *The following linear maps are continuous from \mathcal{S} to \mathcal{S} in the sense of definition 8.2.1:*

- the map $\varphi \mapsto \check{\varphi}(x) \stackrel{\text{def}}{=} \varphi(-x)$;
- the map $\varphi \mapsto (-1)^k \varphi^{(k)}$;
- the Fourier transform, as defined in definition 7.1.1 on page 117;
- the map M_f , which denotes multiplication by a function f in \mathcal{O}_M ;
- the map $f \star$, which denotes convolution by a function f in $L^1_{\mathcal{S}}$, i.e.

$$(f \star \varphi)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(y) \varphi(x-y) dy.$$

Proof. That the first map is continuous is obvious. For the second map, we observe immediately that, by definition of the norms $\|\cdot\|_{k,\mathcal{S}}$,

$$\|\varphi^{(\ell)}\|_{k,\mathcal{S}} \leq C \|\varphi\|_{k+\ell,\mathcal{S}}.$$

To prove that the multiplication map is continuous, it suffices to apply Leibniz's law, which states that

$$(f\varphi)^{(k)} = \sum_{\ell \leq k} C_k^\ell f^{(k-\ell)} \varphi^{(\ell)}.$$

As the function f belongs to \mathcal{O}_M , there exists an integer N such that, for any integer ℓ less than k , we have

$$\forall x \in \mathbb{R}, |f^{(\ell)}(x)| \leq C(1+|x|)^N.$$

We then deduce that

$$\|f\varphi\|_{k,\mathcal{S}} \leq C_f \|\varphi\|_{k+N,\mathcal{S}}.$$

Now, we prove the continuity of the Fourier transform. Note that theorem 7.1.1 on page 119 yields that

$$M(\widehat{\varphi})' = -iM\mathcal{F}(M\varphi) = -\mathcal{F}((M\varphi)') = -\widehat{\varphi} - \mathcal{F}(M\varphi'),$$

where M is the multiplication operator by the identity function. As a result,

$$(1+|\xi|)|\widehat{\varphi}(\xi)| \leq 2\|\varphi\|_{L^1(\mathbb{R})} + \|M\varphi'\|_{L^1(\mathbb{R})} \leq 2\|f\|_{3,\mathcal{S}}.$$

We admit the general proof for the semi-norms with index k on $\mathcal{S}(\mathbb{R})$.

Finally, we study the convolution operator. Derivating under the integral sign, we have

$$(f \star \varphi)^{(k)}(x) = \int_{\mathbb{R}} f(y) \varphi^{(k)}(x-y) dy.$$

As $|x| \leq 2 \max\{|x-y|, |y|\}$, we have

$$(1+|x|)^k \leq 2^k \left((1+|x-y|)^k + (1+|y|)^k \right).$$

Thus, we write

$$\begin{aligned} (1+|x|)^k |(f \star \varphi)^{(k)}(x)| &\leq 2^k \int_{\mathbb{R}} |f(y)| (1+|x-y|)^k |\partial^\alpha \varphi(x-y)| dy \\ &\quad + 2^k \int_{\mathbb{R}} (1+|y|)^k |f(y)| |\varphi^{(k)}(x-y)| dy. \end{aligned}$$

We then deduce that

$$\|f \star \varphi\|_{k,\mathcal{S}} \leq C \|(1+|\cdot|)^k f\|_{L^1} \|\varphi\|_{k,\mathcal{S}},$$

hence the continuity of convolution. \square

We can define the transposes of these operations (which often have some of the same properties) on the space \mathcal{S}' . These provide spectacular extensions of the notions they represent. All of this relies on the following theorem.

Theorem 8.2.1. *Let A be a continuous linear map from $\mathcal{S}(\mathbb{R})$ to $\mathcal{S}(\mathbb{R})$. Then, the linear map tA ,*

$${}^tA \begin{cases} \mathcal{S}'(\mathbb{R}) & \longrightarrow \mathcal{S}'(\mathbb{R}) \\ T & \longmapsto {}^tAT \end{cases} \text{ defined by } \langle {}^tAT, \varphi \rangle = \langle T, A\varphi \rangle.$$

is well defined. Moreover, it is continuous in the following sense: if $(T_n)_{n \in \mathbb{N}}$ is a sequence of tempered distributions that converges to a tempered distribution T , then the sequence $({}^tAT_n)_{n \in \mathbb{N}}$ converges to tAT .

Proof. By definition, T is a continuous linear functional on $\mathcal{S}(\mathbb{R})$, and A is a continuous linear map from $\mathcal{S}(\mathbb{R})$ to itself. So the composition ${}^tAT \stackrel{\text{def}}{=} T \circ A$ is a continuous linear form on $\mathcal{S}(\mathbb{R})$, and is therefore a tempered distribution.

Moreover, let $(T_n)_{n \in \mathbb{N}}$ be a sequence of tempered distributions that converges to T in the sense of definition 8.1.2, which means that, for any function θ in $\mathcal{S}(\mathbb{R})$,

$$\lim_{n \rightarrow \infty} \langle T_n, \theta \rangle = \langle T, \theta \rangle.$$

Applying this to $\theta = A\varphi$, we get that

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle {}^tAT_n, \varphi \rangle = \langle {}^tAT, \varphi \rangle.$$

The theorem is proved. □

For a first application of this theorem, we extend the operation $\check{\cdot}$ to tempered distributions by the formula

$$\forall \varphi \in \mathcal{S}, \langle \check{T}, \varphi \rangle \stackrel{\text{def}}{=} \langle T, \check{\varphi} \rangle$$

Now let us define the derivatives of a tempered distribution.

Definition 8.2.3. *Let k be an element of \mathbb{N} . The k -th order derivation on \mathcal{S}' is the transpose of the continuous linear map from $\mathcal{S}(\mathbb{R})$ to itself defined by $(-1)^k \frac{d^k}{dx^k}$, which, in terms of a tempered distribution T , is written as*

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle T^{(k)}, \varphi \rangle \stackrel{\text{def}}{=} \langle T, (-1)^k \varphi^{(k)} \rangle.$$

First note that, for any function f of class C^1 with compact support in \mathbb{R} , we have, by integration by parts,

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \int_{\mathbb{R}} f'(x)\varphi(x)dx = - \int_{\mathbb{R}} f(x)\varphi'(x)dx.$$

So these two notions of derivation coincide for a function of class C^1 .

Now, note that the function Σ which appeared in the proof of proposition 8.1.6 is none other than

$$-\left(\frac{d}{dx}\right)^2 \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(n \cdot),$$

which is the second-order derivative in the sense of distributions of the uniformly continuous, bounded function

$$x \mapsto -\sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nx).$$

We are now going to study a very simple example that illustrates how deriving in the distribution sense detects more kinds of variation than derivation in the classical sense. Indeed, we now know that we can differentiate functions that are not of class C^1 , for example functions that are differentiable everywhere except at one point of \mathbb{R} . The following property holds for the Heaviside function.

Proposition 8.2.2. *We have: $\frac{d}{dx} \mathbf{1}_{\mathbb{R}^+} = \delta_0$.*

Proof. We study, for any given φ in $\mathcal{S}(\mathbb{R})$, the integral

$$-\int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}^+}(x) \varphi(x) dx.$$

Integrating by parts, we get

$$\begin{aligned} -\int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}^+}(x) \varphi'(x) dx &= -\lim_{A \rightarrow \infty} \int_0^A \varphi'(x) dx \\ &= \lim_{A \rightarrow \infty} (\varphi(0) - \varphi(A)) \\ &= \varphi(0). \end{aligned}$$

Looking back at the definition of the Dirac mass, we see that the proposition is proved. \square

To further illustrate this notion, we compute the derivative of the function $x \mapsto \log|x|$.

Proposition 8.2.3. *Let $\text{vp} \frac{1}{x}$ be the tempered distribution defined in proposition 8.1.3. In the sense of distributions, we have the following:*

$$\frac{d}{dx} \log|x| = \text{vp} \frac{1}{x}.$$

Proof. We study, for any given φ in $\mathcal{S}(\mathbb{R})$, the integral

$$I(\varphi) = -\int_{\mathbb{R}} \log|x| \varphi'(x) dx.$$

By the dominated convergence theorem, we have

$$I(\varphi) = \lim_{\varepsilon \rightarrow 0} I_{\varepsilon}(\varphi) \stackrel{\text{def}}{=} -\int_{[-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] \setminus [-\varepsilon, \varepsilon]} \log|x| \varphi'(x) dx.$$

Integrating by parts, we get

$$\begin{aligned} I_{\varepsilon}(\varphi) &= -\log \varepsilon \varphi(-\varepsilon) + \log\left(\frac{1}{\varepsilon}\right) \varphi\left(-\frac{1}{\varepsilon}\right) + \int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \frac{\varphi(x)}{x} dx \\ &\quad + \log \varepsilon \varphi(\varepsilon) - \log\left(\frac{1}{\varepsilon}\right) \varphi\left(\frac{1}{\varepsilon}\right) + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{\varphi(x)}{x} dx \\ &= \log \varepsilon (\varphi(\varepsilon) - \varphi(-\varepsilon)) - \log \varepsilon \varphi\left(-\frac{1}{\varepsilon}\right) + \log \varepsilon \varphi\left(\frac{1}{\varepsilon}\right) + \int_{-\frac{1}{\varepsilon}}^{-\varepsilon} \frac{\varphi(x)}{x} dx + \int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{\varphi(x)}{x} dx. \end{aligned}$$

By changing the variable, $x = -y$, we get that

$$\int_{[-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] \setminus [-\varepsilon, \varepsilon]} \frac{\varphi(x)}{x} dx = - \int_{[-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] \setminus [-\varepsilon, \varepsilon]} \frac{\varphi(x)}{x} dx,$$

hence

$$I_\varepsilon(\varphi) = \log \varepsilon (\varphi(\varepsilon) - \varphi(-\varepsilon)) - \log \varepsilon \varphi\left(-\frac{1}{\varepsilon}\right) + \log \varepsilon \varphi\left(\frac{1}{\varepsilon}\right) + \frac{1}{2} \int_{[-\frac{1}{\varepsilon}, \frac{1}{\varepsilon}] \setminus [-\varepsilon, \varepsilon]} \frac{\varphi(x) - \varphi(-x)}{x} dx.$$

As the function φ belongs to \mathcal{S} , we have

$$\lim_{\varepsilon \rightarrow 0} \log \varepsilon (\varphi(\varepsilon) - \varphi(-\varepsilon)) - \log \varepsilon \varphi\left(-\frac{1}{\varepsilon}\right) + \log \varepsilon \varphi\left(\frac{1}{\varepsilon}\right) = 0.$$

Moreover, proposition 8.1.3 states, in particular, that the function

$$x \mapsto \frac{\varphi(x) - \varphi(-x)}{x}$$

is integrable on \mathbb{R} . The dominated convergence theorem then ensures that

$$- \int_{\mathbb{R}} \log |x| \varphi'(x) dx = \frac{1}{2} \int_{\mathbb{R}} \frac{\varphi(x) - \varphi(-x)}{x} dx,$$

which is the statement we set out to prove. \square

Exercise 8.2.1. Prove that

$$\frac{d}{dx} \text{vp} \frac{1}{x} = -\text{Pf} \frac{1}{x^2}.$$

We are now going to define the notion of antiderivative (on \mathbb{R}) of a tempered distribution. We start by proving the following result.

Theorem 8.2.2. Let T be a tempered distribution such that $T' = 0$. Then T is a constant function, that is

$$\langle T, \varphi \rangle = c \int_{\mathbb{R}} \varphi(x) dx.$$

Proof. The result relies on the following lemma.

Lemma 8.2.1. Let $\mathcal{S}_0(\mathbb{R})$ be the set of functions of $\mathcal{S}(\mathbb{R})$ with integral equal to 0. The map

$$\mathbb{P} \left\{ \begin{array}{l} \mathcal{S}_0(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R}) \\ \varphi \longmapsto x \mapsto \int_{-\infty}^x \varphi(y) dy \end{array} \right.$$

is a continuous linear map from $\mathcal{S}_0(\mathbb{R})$ to $\mathcal{S}(\mathbb{R})$ that satisfies $(\mathbb{P}\varphi)' = \varphi$ for any φ in $\mathcal{S}_0(\mathbb{R})$.

Proof. Let x be less than or equal to -1 . We have

$$\begin{aligned} |(\mathbb{P}\varphi)(x)| &\leq \|\varphi\|_{N+2, \mathcal{S}} \int_{-\infty}^x \frac{1}{(1+|y|)^{N+2}} dy \\ &\leq |x|^{-N} \|\varphi\|_{N+2, \mathcal{S}} \int_{-\infty}^0 \frac{1}{(1+|y|)^2} dy. \end{aligned}$$

If x is greater than or equal to 1, we use the fact that the integral of φ is equal to zero to write that

$$(\mathbb{P}\varphi)(x) = - \int_x^{+\infty} \varphi(y) dy,$$

and we use the same reasoning. We conclude by observing that $(\mathbb{P}\varphi)' = \varphi$. \square

Back to the proof of theorem 8.2.2. For the rest of the proof, we consider a function φ_1 in $\mathcal{S}(\mathbb{R})$ with integral equal to 1. Let p be the projection of $\mathcal{S}(\mathbb{R})$ onto $\mathcal{S}_0(\mathbb{R})$ parallel to the vector line generated by φ_1 , that is the map

$$p(\varphi) = \varphi - \left(\int_{-\infty}^{+\infty} \varphi(y) dy \right) \varphi_1. \quad (8.3)$$

By lemma 8.2.1, we have $(\mathbb{P}(p(\varphi)))' = p(\varphi)$. As we assume T' to be zero, we deduce that

$$\langle T, p(\varphi) \rangle = 0 = \langle T, \varphi \rangle - \left(\int_{-\infty}^{+\infty} \varphi(y) dy \right) \langle T, \varphi_1 \rangle.$$

The theorem is proved. □

We now apply these ideas in order to find antiderivatives of distributions in dimension 1.

Proposition 8.2.4. *Let T be a tempered distribution on \mathbb{R} . There exists a tempered distribution S such that $S' = T$, and the difference between two distributions satisfying this relation is constant.*

Proof. Let φ_1 be a function in $\mathcal{S}(\mathbb{R})$ with integral equal to 1. We will show that $S \stackrel{\text{def}}{=} -{}^t\mathbb{P}T$ satisfies $S' = T$. We observe that, for any function φ in $\mathcal{S}(\mathbb{R})$, the integral of φ' is equal to 0, thus

$$\mathbb{P} \circ \frac{d}{dx} = \text{Id}_{\mathcal{S}(\mathbb{R})}.$$

By transposing, we deduce that

$${}^t\left(\mathbb{P} \circ \frac{d}{dx}\right) = \text{Id}_{\mathcal{S}'(\mathbb{R})}.$$

As ${}^t(AB) = {}^tB{}^tA$, we get that

$${}^t\left(\frac{d}{dx}\right) \circ {}^t\mathbb{P} = \text{Id}_{\mathcal{S}'(\mathbb{R})}.$$

The fact that ${}^t\left(\frac{d}{dx}\right) = -\frac{d}{dx}$ yields the result. □

Remark. A more explicit proof can be done by simply defining the tempered distribution S by

$$\langle S, \varphi \rangle \stackrel{\text{def}}{=} -\langle T, (\mathbb{P} \circ p)\varphi \rangle \quad \text{with} \quad (\mathbb{P} \circ p)\varphi(x) \stackrel{\text{def}}{=} \int_{-\infty}^x \left(\varphi(y) - \left(\int_{-\infty}^{\infty} \varphi(t) dt \right) \varphi_1(y) \right) dy,$$

where φ_1 is a function in $\mathcal{S}(\mathbb{R})$ with integral equal to 1, for example $\sqrt{\pi}e^{-x^2}$.

The relation between derivative and antiderivative can be precisely described in the following theorem.

Theorem 8.2.3. *Let T be a tempered distribution on \mathbb{R} . If its derivative (in the sense of distributions) T' is a function in $L^1_{\mathcal{M}}(\mathbb{R})$, then T is a continuous function, and we have*

$$T(x) = \int_0^x T'(y) dy + C.$$

Proof. The proof relies on the following lemma that we admit for the moment.

Lemma 8.2.2. Let f be a function in $L^1_{\mathcal{M}}(\mathbb{R})$. Then, in the sense of distributions, we have

$$\frac{d}{dx} \int_0^x f(y)dy = f(x)$$

Set $F(x) = \int_a^x T'(y)dy$. By lemma 8.2.2, we have $F' = T'$. By proposition 8.2.2, we also have

$$T - F = C.$$

This yields the theorem, providing we can prove lemma 8.2.2 of course. \square

Proof of lemma 8.2.2. By definition of derivation in the distribution sense, we need to prove that if

$$F(x) \stackrel{\text{def}}{=} \int_0^x f(y)dy,$$

then we have

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \quad - \int_{\mathbb{R}} F(x)\varphi'(x)dx = \int_{\mathbb{R}} f(x)\varphi(x)dx.$$

Observe first that if

$$\mathcal{I}^+ \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2, 0 \leq y \leq x\} \quad \text{and} \quad \mathcal{I}^- \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}^2, x \leq y \leq 0\},$$

then the functions

$$\mathbf{1}_{\mathcal{I}^+}(x, y)f(y)\varphi'(x) \quad \text{and} \quad \mathbf{1}_{\mathcal{I}^-}(x, y)f(y)\varphi'(x)$$

are integrable on \mathbb{R}^2 . Indeed, by definition of $L^1_{\mathcal{M}}(\mathbb{R})$,

$$\begin{aligned} I(f, \varphi) &\stackrel{\text{def}}{=} \int_{\mathbb{R}^2} \mathbf{1}_{\mathcal{I}^+}(x, y)|f(y)| |\varphi'(x)| dx dy \\ &\leq \int_{\mathbb{R}^2} \mathbf{1}_{\mathcal{I}^+}(x, y)(1 + |y|)^M |\tilde{f}(y)| |\varphi'(x)| dx dy \quad \text{with } \tilde{f} \in L^1(\mathbb{R}). \end{aligned}$$

By Fubini's theorem for non-negative functions, we have

$$I(f, \varphi) \leq C_{M,a} \int_{\mathbb{R}} (1 + |x|)^{M+1} |\varphi'(x)| dx \leq C_{M,a} \|\varphi\|_{M+1, \mathcal{S}}.$$

The same bound for $\mathbf{1}_{\mathcal{I}^-}(x, y)f(y)\varphi'(x)$ is proved the same way. Note that, thanks to Fubini's theorem, we can write that

$$\begin{aligned} - \int_{\mathbb{R}} F(x)\varphi'(x)dx &= - \int_{\mathbb{R}} \left(\int_y^\infty \varphi'(x)dx \right) \mathbf{1}_{[0, \infty[}(y)f(y)dy \\ &\quad + \int_{\mathbb{R}} \left(\int_\infty^y \varphi'(x)dx \right) \mathbf{1}_{]-\infty, 0]}(y)f(y)dy \\ &= \int_{\mathbb{R}} \varphi(y) \mathbf{1}_{[0, \infty[}(y)f(y)dy + \int_{\mathbb{R}} \varphi(y) \mathbf{1}_{]-\infty, 0]}(y)f(y)dy \\ &= \int_{\mathbb{R}} \varphi(y)f(y)dy. \end{aligned}$$

This ends the proof of lemma 8.2.2, and therefore also that of theorem 8.2.3. \square

Now we see how to multiply a tempered distribution by a function in \mathcal{O}_M .

Definition 8.2.4. Let θ be a function in \mathcal{O}_M . The θ multiplier operator on \mathcal{S}' is the transpose of the θ multiplier operator on \mathcal{S} , which means that, for T in \mathcal{S}' ,

$$\langle \theta T, \varphi \rangle \stackrel{\text{def}}{=} \langle T, \theta \varphi \rangle.$$

It is obvious that if T is a function in $L^1_{\mathcal{M}}$, then multiplication in this sense coincides with the usual product in the function sense.

We prove a generalisation of Leibniz's law.

Proposition 8.2.5. If θ belongs to \mathcal{O}_M and T belongs to \mathcal{S}' , then we have, for any integer k ,

$$(\theta T)^{(k)} = \sum_{\ell \leq k} C_k^\ell \theta^{(k-\ell)} T^{(\ell)}.$$

Proof. As in the classical function setting, we only need to prove that $(\theta T)' = \theta' T + \theta T'$, formula that we then iterate. By definition of derivation and multiplication by θ , we have

$$\begin{aligned} \langle (\theta T)', \varphi \rangle &= -\langle \theta T, \varphi' \rangle \\ &= -\langle T, \theta \varphi' \rangle. \end{aligned}$$

Leibniz's law for smooth functions implies that

$$\langle (\theta T)', \varphi \rangle = -\langle T, (\theta \varphi)' \rangle + \langle T, \theta' \varphi \rangle.$$

Again, by definition of derivation and multiplication, we deduce that

$$\begin{aligned} \langle (\theta T)', \varphi \rangle &= \langle T', \theta \varphi \rangle + \langle T, \theta' \varphi \rangle \\ &= \langle \theta T' + \theta' T, \varphi \rangle. \end{aligned}$$

The proposition is proved. □

Now we define the convolution of a tempered distribution with a function in $L^1_{\mathcal{S}}$.

Definition 8.2.5. Let f be a function in $L^1_{\mathcal{S}}$. We define the convolution of f and tempered distribution as the transpose of the convolution of \check{f} and an element of \mathcal{S} , which means

$$\langle f \star T, \varphi \rangle \stackrel{\text{def}}{=} \langle T, \check{f} \star \varphi \rangle.$$

Let us check that we extend the notion of convolution defined in theorem 5.4.1 on page 97. Let g be a function in $L^1(\mathbb{R})$. We have

$$\begin{aligned} \langle \iota(g), \check{f} \star \varphi \rangle &= \int_{\mathbb{R}} g(x) (\check{f} \star \varphi)(x) dx \\ &= \int_{\mathbb{R}} g(x) \left(\int_{\mathbb{R}} \check{f}(x-y) \varphi(y) dy \right) dx. \end{aligned}$$

Fubini's theorem implies that

$$\begin{aligned} \langle \iota(g), \check{f} \star \varphi \rangle &= \int_{\mathbb{R} \times \mathbb{R}} g(x) f(y-x) \varphi(x) dx \\ &= \int_{\mathbb{R}} g(x) \left(\int_{\mathbb{R}} f(x-y) \varphi(y) dy \right) dx \\ &= \int_{\mathbb{R}} (g \star f)(y) \varphi(y) dy, \end{aligned}$$

hence the result.

The relation between convolution and derivation is described by the following proposition.

Proposition 8.2.6. *Let f be in $L^1_{\mathcal{S}}$, and T be in \mathcal{S}' . For any integer k , we have*

$$(f \star T)^{(k)} = f \star T^{(k)}.$$

Proof. By definition of derivation and convolution on \mathcal{S}' , we have

$$\begin{aligned} \langle (f \star T)^{(k)}, \varphi \rangle &= \langle f \star T, (-1)^k \varphi^{(k)} \rangle \\ &= \langle T, \check{f} \star (-1)^k \varphi^{(k)} \rangle. \end{aligned}$$

By the differentiation theorem for functions defined by integrals, we have

$$\check{f} \star \varphi^{(k)} = (\check{f} \star \varphi)^{(k)}.$$

Once again using the definition of derivation and convolution on \mathcal{S}' , we deduce that

$$\begin{aligned} \langle (f \star T)^{(k)}, \varphi \rangle &= \langle T^{(k)}, \check{f} \star \varphi \rangle \\ &= \langle f \star T^{(k)}, \varphi \rangle, \end{aligned}$$

hence the proposition. □

A fundamental and valuable point in this theory is the extension of the Fourier transform to tempered distributions. In the previous chapter, we saw that the Fourier transform was originally defined on $L^1(\mathbb{R})$, but could be extended to $L^2(\mathbb{R})$ by density. The fact that the space \mathcal{S} is invariant under Fourier transform (theorem 7.2.1 on page 121) allows us to extend the Fourier transform to $\mathcal{S}'(\mathbb{R})$ by duality.

Definition 8.2.6. *The Fourier transform on \mathcal{S}' is the transpose of the Fourier transform on \mathcal{S} , that is*

$$\forall \varphi \in \mathcal{S}, \langle \mathcal{F}_{\mathcal{S}'} T, \varphi \rangle = \langle T, \widehat{\varphi} \rangle.$$

The following proposition shows that this is indeed an extension of the Fourier transform.

Proposition 8.2.7. *For any function f in $L^1(\mathbb{R})$, we have*

$$\mathcal{F}_{\mathcal{S}'}(\iota(f)) = \iota(\widehat{f}).$$

Proof. Reprising the notations used in theorem 8.1.1, we have, for any function φ in \mathcal{S} , that

$$\begin{aligned} \langle \mathcal{F}_{\mathcal{S}'} \iota(f), \varphi \rangle &= \langle \iota(f), \widehat{\varphi} \rangle \\ &= \int_{\mathbb{R}} f(x) \widehat{\varphi}(x) dx. \end{aligned}$$

Given that the function

$$(x, y) \mapsto f(x) e^{-i(x|\xi)} \varphi(\xi)$$

is integrable on $\mathbb{R} \times \mathbb{R}$, we deduce that

$$\begin{aligned} \langle \mathcal{F}_{\mathcal{S}'} \iota(f), \varphi \rangle &= \int_{\mathbb{R}} \widehat{f}(\xi) \varphi(\xi) d\xi \\ &= \langle \iota(\widehat{f}), \varphi \rangle, \end{aligned}$$

hence the proposition. □

We now generalise the basic formulae on the Fourier transform to tempered distributions.

Theorem 8.2.4. *For any tempered distribution T , we have*

$$\mathcal{F}_{\mathcal{S}'}^2 T = 2\pi\check{T}, \quad (\mathcal{F}_{\mathcal{S}'}(T))' = -i\mathcal{F}_{\mathcal{S}'}(MT) \quad \text{and} \quad \mathcal{F}_{\mathcal{S}'}(T') = iM\mathcal{F}_{\mathcal{S}'}T.$$

Proof. For any φ in \mathcal{S} , we have $\mathcal{F}^2\varphi = 2\pi\check{\varphi}$. Transposing this relation, we have, for any T in \mathcal{S}' ,

$$({}^t\mathcal{F})^2 T = \mathcal{F}_{\mathcal{S}'}^2 T = 2\pi\check{T},$$

which yields the first formula. For φ in \mathcal{S} , we know by theorem 7.1.1 that

$$\mathcal{F}(i\varphi') = M\widehat{\varphi} \quad \text{and} \quad \mathcal{F}(M\varphi) = -i\frac{d}{d\xi}\widehat{\varphi}.$$

The last two equalities are obtained by transposing these. □

Remark. Henceforth, we no longer use different notations for the Fourier transform on functions in $L^1(\mathbb{R})$ and the one on tempered distributions. So, for any T in $\mathcal{S}'(\mathbb{R})$, we indifferently denote its Fourier transform \widehat{T} or $\mathcal{F}(T)$.

Some applications of the above theorem 8.2.4 consist of computations of Fourier transforms of functions that are not in L^1 . For example, let us consider the constant function 1, and the tempered distribution $\text{vp}\frac{1}{x}$ from proposition 8.1.3.

Proposition 8.2.8. *We have the following formulae:*

$$\mathcal{F}(\delta_0) = \mathbf{1}, \quad \mathcal{F}(\mathbf{1}) = 2\pi\delta_0 \quad \text{and} \quad \mathcal{F}\left(\text{vp}\frac{1}{x}\right)(\xi) = -i\pi\frac{\xi}{|\xi|}.$$

Proof. The first formula is a simple consequence of the fact that

$$\langle \mathcal{F}(\delta_0), \varphi \rangle = \widehat{\varphi}(0) = \int_{\mathbb{R}} \varphi(x) dx.$$

The second one comes from the Fourier inversion formula on \mathcal{S}' , which is

$$\mathcal{F}^2\delta_0 = \mathcal{F}(\mathbf{1}) = 2\pi\delta_0.$$

Let us now take a close look at the principal value of $1/x$. By definition of this distribution, for any function φ in $\mathcal{S}(\mathbb{R})$, we have

$$\begin{aligned} \left\langle \text{vp}\frac{1}{x}, x\varphi \right\rangle &= \frac{1}{2} \int_{-\infty}^{\infty} \frac{x\varphi(x) + x\varphi(-x)}{x} dx \\ &= \frac{1}{2} \int_{\mathbb{R}} (\varphi(x) + \varphi(-x)) dx \\ &= \int_{\mathbb{R}} \varphi(x) dx \\ &= \langle \mathbf{1}, \varphi \rangle. \end{aligned}$$

Thus, $x \text{vp}\frac{1}{x} = 1$. Using theorem 8.2.4, we deduce that

$$-i\frac{d}{d\xi}\mathcal{F}\left(\text{vp}\frac{1}{x}\right) = \mathcal{F}(\mathbf{1}).$$

But we know that $\mathcal{F}\mathbf{1} = 2\pi\delta_0$, so we have

$$\frac{d}{d\xi}\widehat{T} = -2i\pi\delta_0.$$

Thus, by proposition 8.2.2 and lemma 8.2.2, we have

$$\widehat{T} = -2i\pi H(\xi) + C,$$

where C is a scalar constant. To get the value of this constant, we observe that the principal value of $1/x$ is an odd distribution, in the sense that

$$\check{\text{vp}}\frac{1}{x} = -\text{vp}\frac{1}{x}.$$

We let the reader prove that the Fourier transform of an odd tempered distribution is an odd tempered distribution. Thus, the distribution \widehat{T} must be odd, so the constant must be $C = i\pi$. The result is proved. \square

We now study the relations between convolution, multiplication and the Fourier transform, with an aim to extend (i.e. transpose) the result of proposition 7.1.4 on page 121 and corollary 7.2.2 on page 123.

Theorem 8.2.5. *Let θ be a function in \mathcal{S} , and T be a tempered distribution. We have*

$$\mathcal{F}(\theta \star T) = \widehat{\theta}\widehat{T} \quad \text{and} \quad \mathcal{F}(\theta T) = (2\pi)^{-1}\widehat{\theta} \star \widehat{T}.$$

Proof. By definition of the Fourier transform and of convolution, we have, for any function φ in \mathcal{S} ,

$$\begin{aligned} \langle \mathcal{F}(\theta \star T), \varphi \rangle &= \langle \theta \star T, \widehat{\varphi} \rangle \\ &= \langle T, \check{\theta} \star \widehat{\varphi} \rangle \\ &= \langle \widehat{T}, \mathcal{F}^{-1}(\check{\theta} \star \widehat{\varphi}) \rangle. \end{aligned}$$

By the Fourier inversion theorem and the formula for computing Fourier transforms of the convolution of two functions (see proposition 7.1.4 on page 121), we have

$$\begin{aligned} \mathcal{F}^{-1}(\check{\theta} \star \widehat{\varphi})(\xi) &= (2\pi)^{-1}\mathcal{F}(\check{\theta} \star \widehat{\varphi})(-\xi) \\ &= (2\pi)^{-1} \int_{\mathbb{R} \times \mathbb{R}} e^{i(x|\xi)} \theta(y-x) \widehat{\varphi}(y) dy dx \\ &= (2\pi)^{-1} \int_{\mathbb{R} \times \mathbb{R}} e^{-i(y-x|\xi)} \theta(y-x) e^{i(y|\xi)} \widehat{\varphi}(y) dy dx \\ &= \widehat{\theta}(\xi) \mathcal{F}^{-1}(\widehat{\varphi})(\xi) \\ &= \widehat{\theta}(\xi) \varphi(\xi). \end{aligned}$$

By definition of multiplication for tempered distributions, we get that

$$\langle \mathcal{F}(T \star \theta), \varphi \rangle = \langle \widehat{T}, \widehat{\theta}\varphi \rangle = \langle \widehat{\theta}\widehat{T}, \varphi \rangle,$$

which yields the first formula. To establish the second one, we apply the first to $\mathcal{F}(\check{\theta})$ and $\mathcal{F}(\check{T})$, so we get

$$\mathcal{F}(\mathcal{F}(\check{\theta}) \star \mathcal{F}(\check{T})) = \mathcal{F}^2 \check{\theta} \mathcal{F}^2 \check{T} = (2\pi)^2 \theta T.$$

Taking the Fourier transform of this equality yields

$$\mathcal{F}^2(\mathcal{F}(\check{\theta}) \star \mathcal{F}(\check{T})) = (2\pi)^2 \mathcal{F}(\theta T),$$

and the Fourier inversion formula ensures the result. \square

8.3 Two applications

The following theorem is an application of the previous result. It solves a classical harmonic analysis problem.

Theorem 8.3.1. *Pick φ in $\mathcal{S}(\mathbb{R})$. Then,*

$$\left\| \varphi \star \text{vp} \frac{1}{x} \right\|_{L^2} = \pi \|\varphi\|_{L^2}.$$

Proof. We use the Fourier transform. By proposition 8.2.5, we have

$$\mathcal{F}\left(\varphi \star \text{vp} \frac{1}{x}\right) = \widehat{\varphi} \mathcal{F}\left(\text{vp} \frac{1}{x}\right) = -i\pi \frac{\xi}{|\xi|} \widehat{\varphi}(\xi).$$

The Fourier-Plancherel relation in theorem 7.2.1 on page 121 ensures that

$$\begin{aligned} \left\| \varphi \star \text{vp} \frac{1}{x} \right\|_{L^2} &= (2\pi)^{\frac{1}{2}} \left\| \mathcal{F}\left(\varphi \star \text{vp} \frac{1}{x}\right) \right\|_{L^2} \\ &= (2\pi)^{\frac{1}{2}} \left\| i\pi \frac{\xi}{|\xi|} \widehat{\varphi}(\xi) \right\|_{L^2} \\ &= \pi \|\varphi\|_{L^2}, \end{aligned}$$

which proves the theorem. □

Let us show an application of this theory to get an explicit solution to a differential equation.

Theorem 8.3.2. *Let f be a tempered distribution. There exists a unique tempered distribution u that solves*

$$u - u'' = f$$

in the space $\mathcal{S}'(\mathbb{R})$. This solution is given by the formula

$$u = \frac{1}{2} e^{-|\cdot|} \star f.$$

Proof. To prove this, we use theorem 8.2.4 to state that

$$\mathcal{F}((\text{Id} - \Delta)u) = (1 + |\xi|^2) \widehat{u}.$$

The Fourier inversion formula in theorem 8.2.4 then yields

$$(\text{Id} - \Delta)u = f \iff (1 + |\xi|^2) \widehat{u} = \widehat{f}.$$

The function $\xi \mapsto (1 + |\xi|^2)^{\pm 1}$ is a smooth function of slow growth, hence

$$\begin{aligned} \widehat{u} &= \frac{1 + |\xi|^2}{1 + |\xi|^2} \widehat{u} \\ &= \frac{1}{1 + |\xi|^2} \widehat{f}. \end{aligned}$$

So the solution of the equation can be written as

$$u = \mathcal{F}^{-1} \left(\frac{1}{1 + |\xi|^2} \widehat{f} \right).$$

When $d = 1$, we use the result of exercise 7.1 on page 119, which was

$$\frac{1}{2} \mathcal{F}(e^{-|\cdot|})(\xi) = \frac{1}{1 + |\xi|^2},$$

to end the proof of the theorem. □