Basic Functional Analysis
Master 1 UPMC
MM005

Jean-François Babadjian, Didier Smets and Franck Sueur

October 18, 2011

Contents

Chapter 1

Topology

In this chapter we give a few definitions of general topology including compactness and separability. One important particular case of topological spaces are the metric spaces for which most of the definitions can be rephrased in term of sequences. We will also introduce the notion of completeness and we give three theorems using it in a crucial way: the Banach fixed point theorem, Baire's theorem and a theorem about the extension of uniformly continuous functions.

1.1 Basic definitions

1.1.1 General topology

We start with recalling a few basic definitions of general topology.

Definition 1.1.1 (Topology). *Given a set X, we say that a subset τ of $\mathcal{P}(X)$ is a topology on X if*

1. *$\emptyset$ and X are in τ.*
2. *τ is stable by finite intersection.*
3. *τ is stable by union.*

Then we say that (X, τ) is a topological space. The elements of τ are called open sets, and their complementary are the closed sets.

Definition 1.1.2 (Interior and closure). *Given a topological space (X, τ) and a set $A \subset X$, we define*

1. *the interior of A by $A^\circ := \{x \in A : \text{ there exists } U \in \tau \text{ such that } x \in U \subset A\}$;*
2. *the closure of A by $\overline{A} := \{x \in X : \text{ for any } U \in \tau \text{ with } x \in U, \text{ then } U \cap A \neq \emptyset\}$.*

We say that $x \in \overline{A}$ is an adherent point and $x \in A^\circ$ an interior point.

Let us observe that $A^\circ \subset A \subset \overline{A}$.

Definition 1.1.3 (Density). *Given a topological space (X, τ) and a set $A \subset X$, we say that A is dense in X for the topology τ if $\overline{A} = X$.*

Definition 1.1.4 (Limit of a sequence). *Given a topological space (X, τ), we say that a sequence $(x_n)_{n \in \mathbb{N}} \subset X$ converges to x in X if for any open set $U \in \tau$ with $x \in U$, there exists $n_0 \in \mathbb{N}$ such that $x_n \in U$ for all $n \geqslant n_0$.*

Definition 1.1.5 (Continuity). *Given two topological spaces (X_1, τ_1) and (X_2, τ_2), we say that a map $f : X_1 \to X_2$ is continuous at $x_1 \in X_1$ if for all open set $U \in \tau_2$ such that $f(x_1) \in U$, then $f^{-1}(U) \in \tau_1$.*

For extended real-valued functions the following notion of lower semicontinuity is weaker than continuity.

Definition 1.1.6 (**Lower semicontinuity**). *Given a topological space (X, τ) and x_0 in X, we say that a function $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is lower semicontinuous at x_0 if for any $\varepsilon > 0$, there exists a neighborhood $U \in \tau$ of x_0 such that $f(x) \leqslant f(x_0) + \varepsilon$ for all x in U.*

It is not difficult to check that a function is lower semicontinuous if and only if $\{x \in X : \ f(x) > \alpha\}$ is an open set for every $\alpha \in \mathbb{R}$.

1.1.2 Metric spaces

An important case of topological spaces is given by metric spaces that we now introduce.

Definition 1.1.7 (**Distance**). *Given a set X, we say that a function $d : X \times X \to \mathbb{R}^+$ is a distance on X if*

1. *$d(x, y) = 0$ if and only if $x = y$;*
2. *$d(x, y) = d(y, x)$;*
3. *For any x, y, $z \in X$, $d(x, y) \leqslant d(x, z) + d(y, z)$.*

Then we say that (X, d) is a metric space.

For any x in X and any $r > 0$, we denote

$$B(x, r) := \{y \in X : \ d(x, y) < r\} \quad (\text{resp. } \overline{B}(x, r) := \{y \in X : \ d(x, y) \leqslant r\})$$

the open (resp. closed) ball of center x and radius r. A subset is said to be bounded if it is contained in a ball of finite radius.

Proposition 1.1.1. *Given a metric space (X, d), the family τ of all subsets $U \subset X$ such that for each $x \in U$, there exists $r > 0$ satisfying $B(x, r) \subset U$ defines a topology on X.*

The following propositions, whose proofs are left to the reader, highlight the role of the sequences in metric spaces.

Proposition 1.1.2. *Given a metric space (X, d) and the topology τ given by Proposition 1.1.1. The following statements hold:*

1. *A sequence $(x_n)_{n \in \mathbb{N}} \subset X$ converges to x in X if and only if for every $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$, then $d(x, x_n) < \varepsilon$;*
2. *A subset F of X is closed if and only if for every sequence $(x_n)_{n \in \mathbb{N}} \subset F$ converging to x in X, then $x \in F$.*

Proposition 1.1.3. *Let (X_1, d_1) and (X_2, d_2) be two metric spaces, $f : X_1 \to X_2$ and $x_1 \in X_1$. Then the following statements are equivalent:*

1. *f is continuous at x_1;*
2. *For any $\varepsilon > 0$, there exists $\delta > 0$ such that if $x \in X_1$ is such that $d_1(x_1, x) < \delta$, then $d_2(f(x_1), f(x)) < \varepsilon$;*
3. *For any sequence $(x_n)_{n \in \mathbb{N}} \subset X_1$ converging to x_1, the sequence $(f(x_n))_{n \in \mathbb{N}}$ converges to $f(x)$ in X_2.*

Proposition 1.1.4. *Given a metric space (X, d) and x_0 in X. A function $f : X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is lower semicontinuous at x_0 if and only if for any sequence $(x_n)_{n \in \mathbb{N}} \subset X$ converging to x in X, then*

$$f(x) \leqslant \liminf_{n \to \infty} f(x_n).$$

Definition 1.1.8 (**Uniform continuity**). *Let (X_1, d_1) and (X_2, d_2) be two metric spaces. We say that an map $f : X_1 \to X_2$ is uniformly continuous if for any $\varepsilon > 0$, there exists $\delta > 0$ such that if x and $y \in X_1$ satisfy $d_1(x, y) < \delta$, then $d_2(f(x), f(y)) < \varepsilon$.*

It is clear from the definitions above that uniform continuity implies continuity. We will see in Theorem 1.3.1 that the converse statement holds true when the space (X_1, d_1) is compact.

1.2 Completeness

1.2.1 Definition

Completeness is an important notion in general topology and in functional analysis because it enables one to characterize converging sequences without the knowledge of their limit. We first define the Cauchy property.

Definition 1.2.1 (Cauchy sequence). *Given a metric space (X, d), we say that a sequence $(u_n)_{n\in\mathbb{N}} \subset X$ is a Cauchy sequence if for any $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that for all n, $n' \geqslant n_0$ then $d(u_n, u_{n'}) < \varepsilon$.*

Definition 1.2.2 (Completeness). *A metric space (X, d) is complete if any Cauchy sequence converges in X.*

Let us give as a first example the set $\mathbb{R}$ endowed with the usual metric $d(x, y) := |x - y|$. It is also useful to notice that a closed subset of a complete metric space is complete.

1.2.2 Banach fixed point theorem for contraction mapping

An important application of the notion of completeness is given by the following theorem.

Theorem 1.2.1 (Banach, Picard). *Given a complete metric space (X, d) and $f : X \to X$. Assume that f is a contraction, i.e. that there exists a constant $\theta \in (0, 1)$ such that for all x and $y \in X$, then $d(f(x), f(y)) \leqslant \theta d(x, y)$. Then there exists a unique fixed point $x^* \in X$ such that $f(x^*) = x^*$.*

Proof. Let $x_0 \in X$ and let $(x_n)_{n\in\mathbb{N}}$ the associated sequence defined by the relation

$$x_{n+1} = f(x_n). \tag{1.1}$$

By iteration we have

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

For any $n' > n$,

$$d(x_{n'}, x_n) \leqslant \sum_{k=1}^{n'-n} d(x_{n+k}, x_{n+k-1}) \leqslant d(x_1, x_0) \sum_{k=1}^{n'-n} \theta^{n+k-1} \leqslant \frac{\theta^n}{1-\theta} d(x_1, x_0).$$

Therefore $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and by completeness, it converges to an element $x^* \in X$. Since f is continuous, we have $f(x^*) = x^*$ by passing to the limit in (1.1). The uniqueness follows from the contraction assumption. □

The previous theorem is useful in the proof of the Cauchy-Lipschitz theorem in the theory of ordinary differential equations, and also in proof the local inversion theorem.

1.2.3 Baire's theorem

The following theorem was proved by Baire in his 1899 doctoral thesis.

Theorem 1.2.2 (Baire). *In a complete metric space, every intersection of countable collection of dense open sets is dense.*

Proof. Let (X, d) be a complete metric space and $\{U_n\}_{n\in\mathbb{N}}$ be a sequence of dense open sets with the property that $\overline{U_n} = X$ for each $n \in \mathbb{N}$. To prove the result it suffices to show that for any open ball B in X, then $B \cap (\cap_{n\in\mathbb{N}} U_n) \neq \emptyset$. Since U_0 is dense in X, there exists x_0 in X and $r_0 > 0$ such that $\overline{B}(x_0, r_0) \subset B \cap U_0$. By iteration, using the fact that every sets U_n are dense in X, we obtain that there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ a sequence $(r_n)_{n\in\mathbb{N}}$ of positive real numbers with $r_n < r_{n-1}/2$ such that $\overline{B}(x_n, r_n) \subset B(x_{n-1}, r_{n-1}) \cap U_n$. Since for $k \geqslant n$, $x_k \in B(x_n, r_n)$ with $r_n < r_0/2^n$, the sequence $(x_n)_{n\in\mathbb{N}}$ has the Cauchy property. By completeness, there exists $x \in X$ such that $(x_n)_{n\in\mathbb{N}}$ converges to x. We conclude by observing that $x \in B \cap (\cap_{n\in\mathbb{N}} U_n)$ since $x \in \overline{B}(x_n, r_n) \subset U_n \cap B$ for any $n \in \mathbb{N}$. □

1.2.4 Extension of uniformly continuous functions

Theorem 1.2.3 (Extension of uniformly continuous functions). *Given two metric spaces (X_1, d_1) and (X_2, d_2), the latter being complete, a dense subset Y of X_1, and a map $f : Y \to X_2$ which is uniformly continuous. Then there exists a unique uniformly continuous extension $g : X_1 \to X_2$ of f.*

Proof. The uniqueness is straightforward: indeed for any $x \in X_1$, since Y is dense in X_1, there exists a sequence $(x_n)_{n\in\mathbb{N}} \subset Y$ which converges to x. If $g : X_1 \to X_2$ is a uniformly continuous extension of f, then $g(x)$ must be the limit of the sequence $(f(x_n))_{n\in\mathbb{N}}$.

Now to prove the existence of such an extension, observe that the sequence $(f(x_n))_{n\in\mathbb{N}}$ has the Cauchy property, since $(x_n)_{n\in\mathbb{N}}$ has the Cauchy property (because it converges) and f is uniformly continuous. Since (X_2, d_2) is complete, it yields that the sequence $(f(x_n))_{n\in\mathbb{N}}$ converges to an element z in X_2. This z does not depend on the choice of the sequence $(x_n)_{n\in\mathbb{N}}$ in X_1. Indeed, if $(x'_n)_{n\in\mathbb{N}}$ is another sequence converging to x, then $d_1(x_n, x'_n) \to 0$ when $n \to +\infty$, so that, since f is uniformly continuous, $d_2(f(x_n), f(x'_n)) \to 0$ when $n \to +\infty$. Hence the sequence $(f(x'_n))_{n\in\mathbb{N}}$ converges to z as well. Therefore, it makes sense to define $g(x) := z$.

Let us now prove that g is uniformly continuous. Let $\varepsilon > 0$. Since f is uniformly continuous there exists $\delta > 0$ such that for any x, $x' \in Y$ with $d_1(x, x') < \delta$, there holds $d_2(f(x), f(x')) < \varepsilon/3$. Let y and $y' \in X_1$ with $d_1(y, y') < \delta/3$. There exists x, $x' \in Y$ such that $d_1(x, y) < \delta/3$, $d_1(x', y') < \delta/3$, $d_2(f(x), g(y)) < \varepsilon/3$ and $d_2(f(x'), g(y')) < \varepsilon/3$. Therefore, thanks to the triangle inequality, we have $d_1(x, x') < \delta$ and therefore $d_2(f(x), f(x')) < \varepsilon/3$. Using again the triangle inequality, we get $d_2(g(y), g(y')) < \varepsilon$. Hence g is uniformly continuous. □

Some typical applications of the extension of a uniformly continuous function can be found in the study of the convolution product (see Corollary 4.4.1), and in the proof of the inverse Fourier transformation formula (section 9).

1.2.5 Banach spaces and algebra

Let us recall a few definitions:

Definition 1.2.3 (Normed vector space)**.** *A normed vector space over $\mathbb{R}$ is a pair $(V, \|\cdot\|)$ where V is a real vector space and $\|\cdot\|$ is a norm on X that is a function from X to $\mathbb{R}_+$ satisfying*

1. *$\|u\| = 0$ if and only $u = 0$ (positive definiteness),*
2. *for any u in V, for any $\lambda \in \mathbb{R}$, $\|\lambda u\| = |\lambda|\|u\|$ (positive homogeneity),*
3. *for any u, v in V, $\|u + v\| \leqslant \|u\| + \|v\|$ (triangle inequality or subadditivity).*

If the first item above is not satisfied then $\|\cdot\|$ is called a seminorm.

Remark 1.2.1. *We define in a similar way a normed vector space over $\mathbb{C}$ by considering a complex vector space and by extending the second property above to any $\lambda \in \mathbb{C}$.*

Remark 1.2.2. *We can easily associate a distance to the norm of a normed vector space, through the formula $d(u, v) := \|u - v\|$.*

Definition 1.2.4 (Banach space/algebra)**.** *If the topology defined by this distance is complete then we say that $(V, \|\cdot\|)$ is a Banach space. If in addition V is an associative algebra whose multiplication law is compatible with the norm in the sense that $\|u \cdot v\| \leqslant \|u\| \cdot \|v\|$ for any u, v in V, then we say that $(V, \|\cdot\|)$ is a Banach algebra.*

Next proposition shows that in a normed vector space, completeness can be characterized thanks to the series.

Proposition 1.2.1. *Let $(V, \|\cdot\|)$ be a normed vector space. Then $(V, \|\cdot\|)$ is a Banach space if and only if the series normally converging actually converge in $(V, \|\cdot\|)$.*

Let us recall that a for sequence $(u_n)_n$ in V, we say that the series $\sum_n u_n$ is normally converging in V if $\sum_n \|u_n\|$ converges in $\mathbb{R}$.

Proof. Suppose that V is complete, and let $(u_n)_n$ be a sequence in V such that $\sum_n \|u_n\|$ converges in $\mathbb{R}$. Let $\varepsilon > 0$ be given and put $v_n := \sum_{k=0}^n u_k$. Then for $n > m$,

$$\|v_n - v_m\| = \|\sum_{k=m+1}^{n} u_k\| \leqslant \sum_{k=m+1}^{n} \|u_k\| < \varepsilon$$

for all sufficiently large m and n, since $\sum_{k=0}^n \|u_k\| < \infty$. Hence $(v_n)_n$ is a Cauchy sequence and so converges since V is complete, by hypothesis.

Conversely, assume that $\sum_n u_n$ converges in V whenever $\sum_n \|u_n\| < \infty$. Let $(v_n)_n$ be a Cauchy sequence in V. We will show that $(v_n)_n$ converges in V. By iteration there exists a subsequence $(v_{n_k})_k$ of $(v_n)_n$ such that for any k, $\|v_{n_{k+1}} - v_{n_k}\| \leqslant 2^{-k}$. Therefore, setting $u_k := v_{n_{k+1}} - v_{n_k}$, we have:

$$\sum_{k=0}^{N} \|u_k\| \leqslant \sum_{k=0}^{N} \|v_{n_{k+1}} - v_{n_k}\| \leqslant \sum_{k=0}^{N} 2^{-k} < \infty.$$

We therefore get that $\sum_{k=0}^N u_k$ converges to some u in V when $N \to +\infty$. But

$$\sum_{k=0}^{N} u_k = \sum_{k=0}^{N} v_{n_{k+1}} - v_{n_k} = v_{n_{N+1}} - v_{n_0},$$

so that $v_{n_N} \to u + v_{n_0}$ when $N \to +\infty$. Thus the Cauchy sequence $(v_n)_n$ has a convergent subsequence and so must itself converge. □

Let us now give a result to quotient spaces defined as follows.

Definition 1.2.5. *Let X be a vector space, and let M be a vector subspace of X. We define an equivalence relation $\sim$ on X by setting $x \sim y$ if and only if $x - y \in M$. It is straightforward to check that this really is an equivalence relation on X. For x in X we denote $[x]$ the equivalence class containing the element x and we denote the set of equivalence classes by X/M, that we call the quotient space of X by M.*

We define on X/M a sum law by $[x + y] := [x] + [y]$ and a scalar multiplication law by $[\alpha x] := \alpha[x]$.

Let us stress that these definitions are meaningful since M is a linear subspace of X. For example, if $x \sim x'$ and $y \sim y'$ then $x+y \sim x'+y'$, so that the definition is independent of the particular representatives taken from the various equivalence classes. It is then straightforward to get the following.

Proposition 1.2.2. *The quotient space X/M is a linear space.*

Let us now define

$$\|[x]\| := \inf\{\|y\| / \, y \in [x]\}. \tag{1.2}$$

Proposition 1.2.3. *The equality 1.2 defines a seminorm on X/M.*

Proof. Let $\alpha \in \mathbb{R}^*$ and $[x]$ be in X/M. Then

$$\begin{aligned}
\|\alpha[x]\| &:= \|[\alpha x]\| \\
&= \inf\{\|y\| / \, y \in [\alpha x]\} \\
&= \inf\{\|\alpha x + m\| / \, m \in M\} \\
&= \inf\{\|\alpha x + \alpha m\| / \, m \in M\},
\end{aligned}$$

since the mapping $m \mapsto \alpha m$ is a bijection. Thus

$$\|\alpha[x]\| = |\alpha| \inf\{\|x + m\| / \, m \in M\} = |\alpha| \|[x]\|.$$

On the other hand, since $[0] = M$ we have that $\|[0]\| = 0$. Next, we consider the triangle inequality:

$$\begin{aligned}
\|[x] + [y]\| &:= \|[x + y]\| \\
&= \inf\{\|x + y + m\| / \, m \in M\} \\
&= \inf\{\|x + y + m' + m''\| / \, m', m'' \in M\} \\
&\leqslant \inf\{\|x + m\| + \|y + m''\| / \, m', m'' \in M\} \\
&\leqslant \|[x]\| + \|[y]\|.
\end{aligned}$$

□

To see whether or not it is a norm, all that remains is to see if $\|[x]\| = 0$ implies $[x] = 0$. This may be wrong in general but next proposition shows that it is true if M is closed.

Proposition 1.2.4. *Assume that M is a closed linear subspace of the normed space X. Then 1.2 defines a norm on X/M, called the quotient norm.*

Proof. Assume that $\|[x]\| = 0$. Then for any $n \in \mathbb{N}^*$ there exists $m_n \in M$ such that $\|x + m_n\| \leqslant 1/n$. Thus the sequence $(m_n)_n$ is converging to $-x$. Since M is closed this yields that $-x$ is in M, which is a linear space, so x is in M too. $\square$

In the case where M is closed we also have the two following properties.

Proposition 1.2.5. *Let M is a closed linear subspace of a normed space X. Then the canonical projection $\pi : x \in X \mapsto [x] \in X/M$ is continuous.*

Proof. Assume that $(x_n)_n$ is a sequence in X converging to x in X. Then

$$\|\pi(x_n) - \pi(x)\| = \inf\{\|x_n - x + m\| /\, m \in M\} \leqslant \|x_n - x\|,$$

since $0 \in M$. Therefore the sequence $(\pi(x_n))_n$ is converging to $\pi(x)$ in X/M. $\square$

Proposition 1.2.6. *Let M is a closed linear subspace of a Banach space X. Then X/M is a Banach space.*

In order to prove Proposition 1.2.6 it only remains to show that X/M is complete. We are going to use the criterion of Proposition 1.2.1. Actually the method we will use is quite general and we therefore give first a general statement. Then we will go back to the proof of Proposition 1.2.6.

Proposition 1.2.7. *Let X be a Banach space and Y be a normed vector space. Let T be a linear continuous surjective mapping from X to Y. Assume there exists some constant $C > 0$ such that for any y in Y there exists x in X such that*

$$T(x) = y \text{ and } \|x\| \leqslant C\|y\|. \tag{1.3}$$

Then Y is a Banach space.

Proof. Assume that $(y_n)_n$ is a sequence in Y such that $\sum_{k\geqslant 0} \|y_k\| < \infty$. Define the sequence $(w_n)_n$ by setting $w_n := \sum_{k=0}^{n} y_k$. We want to prove that the sequence $(w_n)_n$ is converging in Y so that thanks to Proposition 1.2.1 we will get that Y is complete. Thanks to the assumption there exists a sequence $(x_n)_n$ in X such that $T(x_n) = y_n$ and $\|x_n\| \leqslant C\|y_n\|$. As a consequence we get

$$\sum_{k\geqslant 0} \|x_k\| \leqslant C \sum_{k\geqslant 0} \|y_k\| < \infty.$$

Using Proposition 1.2.1 for X which is assumed to be a Banach space we get that the sequence $(z_n)_n$ defined by $z_n := \sum_{k=0}^{n} x_k$ is convergent to some x in X. Since T is linear we have $T(z_n) = w_n$, and T is also continuous so that $T(z_n)$ is converging to $T(z)$. This proves that the sequence $(w_n)_n$ is converging in Y and therefore the proof of the proposition is done. $\square$

Let us now go back to the proof of Proposition 1.2.6.

Proof. We are going to apply Proposition 1.2.7 with $Y = X/M$ and $T = \pi$. We know that π is linear, continuous, and surjective. Let us show that the last assumption of Proposition 1.2.7 is satisfied with $C = 2$. Let $y = [x]$ be in X/M. If $y = [x] = 0$ then (1.3) holds true with $x = 0$. If $y \neq 0$ then $\|y\| > 0$ and since by definition, $\|y\| = \inf\{\|x\| /\, x \in y\}$ we have that there exists $x \in X$ such that $\|x\| \leqslant 2\|y\|$. $\square$

1.3 Compactness

Several notions of compactness are available. The following one can be formulated in a general setting.

Definition 1.3.1 (Compactness). *We say that a topological space (X, τ) is compact if any open cover has a finite subcover, i.e. for every arbitrary collection $\{U_i\}_{i\in I}$ of open subsets of X such that $X \subset \cup_{i\in I} U_i$, there is a finite subset $J \subset I$ such that $X \subset \cup_{i\in J} U_i$.*

It is a good exercise to prove the following theorem in order to understand the power of the previous definition.

Theorem 1.3.1 (Heine). *Every continuous image of a compact set is compact. Moreover a continuous function on a compact set is uniformly compact.*

In a metric space, compactness can be formulated in terms of sequences. Let us first recall a few facts about the notion of limit points. Let S be a subset of a topological space X. We say that a point $x \in X$ is a limit point of S if every open set containing x also contains a point of S other than x itself. In a metric space, it is equivalent to requiring that every neighbourhood of x contains infinitely many points of S.

Let us also define what we mean by a totally bounded space.

Definition 1.3.2 (Totally boundedness). *We say that a metric space (X, d) is totally bounded if for every $\varepsilon > 0$, there exists a finite cover of X by open balls of radius less than ε.*

Since for every $\varepsilon > 0$, $\cup_{x\in X} B(x, \varepsilon)$ is an open cover of X, it follows from Definitions 1.3.1 and 1.3.2 that a compact metric space is totally bounded.

Proposition 1.3.1. *A metric space is compact if and only if every sequence has a limit point.*

Proof. We start by proving the necessary condition. Let us assume by contradiction that $(x_n)_{n\in\mathbb{N}}$ is a sequence in a compact metric space X without any limit point. Then for every y in X, there exists $r(y) > 0$ such that the ball $B(y, r(y))$ contains only finitely many elements of the sequence. The collection of these balls is a open cover of X, from which we extract a finite subcover. This would yield that the sequence is in the union of a finite number of balls each of them containing only finitely many elements of the sequence which is absurd.

Let us now prove the sufficient condition. We therefore consider a metric space X such that every sequence has a limit point. We first prove that X is totally bounded. Proceeding by contradiction, it would yield the existence of some $\varepsilon > 0$ and some sequence $(x_n)_{n\in\mathbb{N}}$ such that for any m, $n \in \mathbb{N}$ with $m \neq n$, $d(x_m, x_n) > \varepsilon$. Such a sequence cannot have a limit point, which is a contradiction. Hence X is totally bounded.

Let us now prove that X is compact. We consider an open cover $\{U_i\}_{i\in I}$ of X. We define the mapping

$$R : x \in X \mapsto R(x) := \sup\{r > 0 : \ \exists\, i \in I \text{ with } \ B(x, r) \subset U_i\} > 0,$$

which is lower semicontinuous. Indeed, if not, there would exist a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ converging to x in X such that

$$R(x) > \liminf_{n\to\infty} R(x_n),$$

and we could choose ρ and ρ' such that $R(x) > \rho' > \rho > \liminf_n R(x_n)$. Let $i \in I$ be such that $B(x, \rho') \subset U_i$. Since the sequence $(x_n)_{n\in\mathbb{N}}$ converges to x, for n large enough $B(x_n, \rho) \subset B(x, \rho') \subset U_i$, which is against the fact that $\rho > \liminf_n R(x_n)$.

Let us now consider $\varepsilon := \inf_{x\in X} R(x)$ and $(x_n)_{n\in\mathbb{N}}$ a minimizing sequence for R, that is such that

$$\lim_{n\to\infty} R(x_n) = \varepsilon.$$

Then by assumption, the sequence $(x_n)_{n\in\mathbb{N}}$ has a limit point that we call x. Since R is lower semicontinuous, we have

$$0 < R(x) \leqslant \liminf_{n\to\infty} R(x_n) = \varepsilon.$$

We already know that X is totally bounded. Thus for that ε there exist $x_1, \ldots, x_n$ such that

$$X \subset \bigcup_{i=1}^{n} B(x_i, \varepsilon).$$

By definition of ε, for any $i = 1, \ldots, n$ one has $\varepsilon \leqslant R(x_i)$, and thus there exists $j_i \in I$ such that $B(x_i, \varepsilon) \subset U_{j_i}$ and finally

$$X \subset \bigcup_{i=1}^{n} U_{j_i}$$

which gives a finite subcover of X. □

Next proposition gives another criterion of compactness for metric spaces.

Proposition 1.3.2. *A metric space is compact if and only if it is complete and totally bounded.*

Proof. According to Proposition 1.3.1, in a compact space every sequence has a limit point. Since a Cauchy sequence with a limit point must converge to this limit point, we deduce that a compact metric space is complete. Moreover, we have already seen that a compact metric space is totally bounded. Therefore, it only remains to prove the converse statement.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a complete and totally bounded metric space X. We are going to prove that $(x_n)_{n\in\mathbb{N}}$ has a limit point by Cantor's diagonal argument. Since X is totally bounded there exists a ball that we call B_1 of radius 1 which contains a subsequence $(x_n^1)_n$ of $(x_n)_n$. By iteration we obtain that for any $k \geqslant 2$, there exists a ball B_k of radius $1/k$ which contains a subsequence $(x_n^k)_n$ of $(x_n^{k-1})_n$. Then the sequence $(x_n^n)_n$ has the Cauchy property, since for any $k \geqslant 1$, for any $n \geqslant k$, x_n^n is in B_k. Since X is complete, the sequence $(x_n^n)_n$ has a limit, which is a limit point of the sequence $(x_n)_n$. □

1.4 Separability

Definition 1.4.1 (Separability). *We say that a topological space (X, τ) is separable if it contains a countable dense subset, i.e., there exists a sequence $(x_n)_{n\in\mathbb{N}}$ of elements of X such that every nonempty open subset of the space contains at least one element of the sequence.*

Any topological space which is itself finite or countable is separable. An important example of uncountable separable space is the real line (with its usual topology), in which the rational numbers form a countable dense subset.

Proposition 1.4.1. *Every compact metric space is separable.*

Proof. Let (X, d) be a compact metric space. For any $k \in \mathbb{N}^*$, $\cup_{x\in X} B(x, 1/k)$ is a open cover of X. By compactness, there exists $x_1^k, \ldots, x_{n_k}^k \in X$ such that $X = \cup_{j=1}^{n_k} B(x_{n_j}^k, 1/k)$. Then the collection

$$\bigcup_{k\in\mathbb{N}^*} \bigcup_{j=1}^{n_k} \{x_{n_j}^k\}$$

is a countable dense subset of X. □

Let us finish with the following useful criterion for a metric space to be not separable.

Proposition 1.4.2. *If a metric space (X, d) contains a uncountable subset Y such that*

$$\delta := \inf\{d(y, y') : \ y, y' \in Y, \, y \neq y'\} > 0,$$

then X is not separable.

Proof. We argue by contradiction. Let us assume that (X, d) is separable and therefore contains a countable dense subset $(x_n)_{n\in\mathbb{N}}$. We can then define a map by associating to any $y \in Y$ the smallest $n \in \mathbb{N}$ such that $d(y, x_n) < \delta/3$. This map turns out to be injective because if $d(y, x_n) < \delta/3$ and $d(y', x_n) < \delta/3$, then $d(y, y') < 2\delta/3$ which is possible (when y and $y' \in Y$) only if $y = y'$. We deduce that Y is countable which is the absurd. □

Chapter 2

Spaces of continuous functions

2.1 Basic definitions

Definition 2.1.1. *Let be given two metric spaces (X_1, d_1) and (X_2, d_2). We denote by*

$$\mathcal{B}(X_1; X_2) := \{f : X_1 \to X_2 : \ f(X_1) \text{ is a bounded subset of } X_2\}, \tag{2.1}$$
$$\mathcal{C}_b(X_1; X_2) := \{f \in \mathcal{B}(X_1; X_2) \text{ which are continuous}\}. \tag{2.2}$$

For any f_1 and $f_2 \in \mathcal{B}(X_1; X_2)$, we denote the uniform distance by

$$d_u(f_1, f_2) := \sup_{x \in X_1} d_2(f_1(x), f_2(x)). \tag{2.3}$$

Endowed with the distance d_u, $\mathcal{B}(X_1; X_2)$ is a metric space. When (X_1, d_1) is compact, a continuous mapping $f : X_1 \to X_2$ is bounded thanks to Heine's theorem (Theorem 1.3.1). In this case we simply denote $\mathcal{C}(X_1; X_2)$ instead of $\mathcal{C}_b(X_1; X_2)$.

Proposition 2.1.1. *The space $\mathcal{C}_b(X_1; X_2)$ is closed in $\mathcal{B}(X_1; X_2)$.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{C}_b(X_1; X_2)$ converging to f in $\mathcal{B}(X_1; X_2)$. Let us prove that f is continuous at x in X_1. By Proposition 1.1.3, it suffices to consider a sequence $(x_n)_{n\in\mathbb{N}}$ in X_1 converging to x and to prove that $f(x_n)$ converges to $f(x)$. Let $\varepsilon > 0$. There exists n_0 such that for any $n \geqslant n_0$, $d_u(f, f_n) < \varepsilon/3$. Since f_n is continuous there exists $\delta > 0$ such that for any y in X_1 with $d(x, y) < \delta$, then $d(f_n(x), f_n(y)) < \varepsilon/3$. For n large enough, $d(x_n, x) < \delta$, so that by the triangle inequality we get $d(f(x), f(x_n)) < \varepsilon$. Hence f is continuous at x. □

2.2 Completeness

Theorem 2.2.1. *Let (X_1, d_1) and (X_2, d_2) be two metric spaces, the latter being complete. Then $\mathcal{B}(X_1; X_2)$ and $\mathcal{C}_b(X_1; X_2)$ are complete.*

Proof. Since by Proposition 2.1.1 $\mathcal{C}_b(X_1; X_2)$ is closed in $\mathcal{B}(X_1; X_2)$ it suffices to prove that the latter is complete to prove the result. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{B}(X_1; X_2)$. It follows from the definition of d_u that for any x in X_1, the sequence $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in X_2. Since X_2 is complete the sequence $(f_n(x))_{n\in\mathbb{N}}$ has a limit that we call $f(x)$. It then remains to verify that it defines a function f in $\mathcal{B}(X_1; X_2)$ and that $(f_n)_{n\in\mathbb{N}}$ actually converges to f in $\mathcal{B}(X_1; X_2)$. Since $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{B}(X_1; X_2)$, there exists n_0 such that for any n, $n' \geqslant n_0$, $d_u(f_n, f_{n'}) < 1$. Passing to the limit $n \to +\infty$ yields $d_u(f, f_{n'}) < 1$. Since $f_{n'}$ is in $\mathcal{B}(X_1; X_2)$ there exists $x_2 \in X_2$ and $r > 0$ such that $f_{n'}(X_1) \subset B(x_2, r)$. Therefore $f(X_1) \subset B(x_2, r + 1)$ and thus $f \in \mathcal{B}(X_1; X_2)$. To prove that $(f_n)_{n\in\mathbb{N}}$ converges to f in $\mathcal{B}(X_1; X_2)$ it is sufficient to pass to the limit in the Cauchy property. □

Uniform convergence trivially implies pointwise convergence. The following result gives a partial converse statement.

Theorem 2.2.2 (Dini). *Let (X,d) be a compact metric space and $(f_n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{C}_b(X;\mathbb{R})$ such that for every $x \in X$ the sequence $(f_n(x))_{n\in\mathbb{N}}$ is decreasing and bounded from below. If the function defined for every $x \in X$ by*

$$f(x) := \lim_{n\to+\infty} f_n(x) = \inf_{n\in\mathbb{N}} f_n(x)$$

is continuous, then $(f_n)_{n\in\mathbb{N}}$ converges to f in $\mathcal{C}_b(X;\mathbb{R})$.

Proof. We can assume without loss of generality that $f = 0$ otherwise it suffices to consider $f_n - f$ instead of f_n. For every $n \in \mathbb{N}$, the function f_n has a maximum, say in x_n. There exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $(x_{n_k})_{k\in\mathbb{N}}$ converges to x. Then for any $m \in \mathbb{N}$, we have

$$\lim_{k\to+\infty} \max_{x\in X} f_{n_k}(x) = \lim_{k\to+\infty} f_{n_k}(x_{n_k}) \leqslant \lim_{k\to+\infty} f_m(x_{n_k}),$$

since the sequence $(f_n)_{n\in\mathbb{N}}$ is decreasing. Now since f_m is continuous, $\lim_{k\to+\infty} f_m(x_{n_k}) = f_m(x)$. We now let m tends to $+\infty$ to get

$$\lim_{k\to+\infty} \max_{x\in X} f_{n_k}(x) = 0.$$

Since the sequence $(f_n)_{n\in\mathbb{N}}$ is decreasing, we infer that

$$\lim_{n\to+\infty} \max_{x\in X} f_n(x) = 0,$$

and thus $(f_n)_{n\in\mathbb{N}}$ converges to f in $\mathcal{C}_b(X;\mathbb{R})$. □

2.3 Compactness

The following result gives some sufficient conditions for a collection of continuous functions on a compact metric space to be relatively compact (*i.e.* whose closure is compact). In particular this could allow to extract an uniformly convergent subsequence from a sequence of continuous functions. The main condition is the equicontinuity which was introduced at around the same time by Ascoli (1883 − 1884) and Arzelà (1882 − 1883).

Theorem 2.3.1 (Ascoli). *Let (X_1, d_1) be a compact metric space and (X_2, d_2) be a complete metric space. Let $\mathcal{A}$ be a subset of $\mathcal{C}(X_1;X_2)$ such that*

1. *$\mathcal{A}$ is uniformly equicontinuous, i.e., for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $d_1(x,y) < \delta$, then*

$$\sup_{f\in\mathcal{A}} d_2(f(x), f(y)) < \varepsilon;$$

2. *$\mathcal{A}$ is pointwise relatively compact, i.e., for all $x \in X_1$, the set*

$$\overline{\{f(x) : f \in \mathcal{A}\}}$$

   is compact in X_2.

Then $\overline{\mathcal{A}}$ is a compact subset of $\mathcal{C}(X_1;X_2)$.

Proof. Since (X_2, d_2) a complete metric space, then $\mathcal{C}(X_1;X_2)$ is complete as well. Then $\overline{\mathcal{A}}$ as a closed subset of $\mathcal{C}(X_1;X_2)$ is also complete. Therefore, thanks to Proposition 1.3.2, it is sufficient to prove that $\overline{\mathcal{A}}$ is totally bounded, or even that $\mathcal{A}$ is totally bounded.

Let $\varepsilon > 0$ and $\delta > 0$ be as in the uniform equi-continuity property. Since X_1 is compact, there exists $x_1, \ldots, x_n \in X_1$ such that

$$X_1 = \bigcup_{i=1}^{n} B(x_i, \delta).$$

Moreover for any $i = 1, \ldots, n$, the sets $\overline{\{f(x_i) : f \in \mathcal{A}\}}$ are compact in X_2 which leads to the existence of $y_{i,1}, \ldots, y_{i,l_i} \in X_2$ such that

$$\{f(x_i) : \ f \in \mathcal{A}\} \subset \bigcup_{j=1}^{l_i} B(y_{i,l_j}, \varepsilon).$$

Let $E_1 := \{1, \ldots, n\}$, $E_2 := \{(i,j) : 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant l_i\}$ and Γ the set of all maps from E_1 in E_2. Note that the set Γ is finite. For each $\gamma \in \Gamma$, we define

$$A_\gamma := \{f \in \mathcal{A} : d_2(f(x_i), y_{\gamma(i)}) < \varepsilon \text{ for all } 1 \leqslant i \leqslant n\}.$$

By construction $\mathcal{A} = \cup_{\gamma \in \Gamma} A_\gamma$. Let $\gamma \in \Gamma$ be fixed and $f,\ g \in A_\gamma$. Let $x \in X_1$ and $i \in E_1$ such that $x \in B(x_i, \delta)$. We have

$$\begin{aligned} d_2(f(x), g(x)) &\leqslant d_2(f(x), f(x_i)) + d_2(f(x_i), y_{\gamma(i)}) \\ &\quad + d_2(y_{\gamma(i)}, g(x_i)) + d_2(g(x_i), g(x)) \\ &< 4\varepsilon. \end{aligned}$$

Since x is arbitrary, we deduce that $A_\gamma \subset B(f_\gamma, 4\varepsilon)$ for some $f_\gamma \in A_\gamma$, and thus

$$\mathcal{A} \subset \bigcup_{\gamma \in \Gamma} B(f_\gamma, 4\varepsilon),$$

and the proof is complete. □

2.4 Separability

Let us recall that the Weierstrass approximation theorem states that every continuous function defined on a closed interval can be uniformly approximated by polynomial functions. The original version of this result dates back to 1885. Stone considerably generalized the theorem in 1937 and simplified the proof in 1948. Before to state the so-called Stone-Weierstrass theorem, let us observe that when (X,d) is a compact metric space, then $\mathcal{C}(X;\mathbb{R})$, endowed with the uniform norm and the pointwise multiplication, is a Banach algebra. The Stone-Weierstrass theorem provides a characterization of the subalgebras $\mathcal{A}$ of $\mathcal{C}(X;\mathbb{R})$ which are dense in $\mathcal{C}(X;\mathbb{R})$. It turns out that the crucial property for such a subalgebra $\mathcal{A}$ is to separate points, *i.e.*, for any $x,\ y \in X$ with $x \neq y$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. We will focus here our attention on the sufficiency of this condition for subalgebras of $\mathcal{C}(X;\mathbb{R})$ which contain constant functions.

Theorem 2.4.1 (Stone-Weierstrass). *Let (X,d) be a compact metric space, $\mathcal{A}$ a subalgebra of $\mathcal{C}(X;\mathbb{R})$ which contains constant functions and separates points. Then $\mathcal{A}$ is dense in $\mathcal{C}(X;\mathbb{R})$.*

Theorem 2.4.1 implies Weierstrass' original statement since the space of all polynomial functions on a closed interval is a subalgebra of all continuous real-valued functions on this interval which contains the constants and separates points.

To prove Theorem 2.4.1, we first start with the following lemma.

Lemma 2.4.1. *There exists a sequence $(P_n)_{n\in\mathbb{N}}$ of polynomial functions with real coefficients which converge uniformly to the square root function on* $[0,1]$.

Proof. We define a sequence $(P_n)_{n\in\mathbb{N}}$ setting $P_0(x) = 0$ and

$$P_{n+1}(x) = P_n(x) + \frac{1}{2}(x - P_n(x)^2).$$

We first show by iteration that $0 \leqslant P_n(x) \leqslant \sqrt{x}$ for any $x \in [0,1]$. Indeed, the result is obvious when $n = 0$, Assume that that $0 \leqslant P_n(x) \leqslant \sqrt{x}$ for all $x \in [0,1]$ and some $n \in \mathbb{N}$. Then clearly $P_{n+1}(x) \geqslant 0$, while

$$\begin{aligned} P_{n+1}(x) &= P_n(x) + \frac{1}{2}(\sqrt{x} - P_n(x))(\sqrt{x} + P_n(x)) \\ &\leqslant P_n(x) + (\sqrt{x} - P_n(x)) \\ &\leqslant \sqrt{x}, \end{aligned}$$

since $\sqrt{x} + P_n(x) \leqslant 2\sqrt{x} \leqslant 2$ for all $x \in [0,1]$.

Then we infer that the sequence $(P_n)_{n\in\mathbb{N}}$ is increasing and it converges pointwise to $\sqrt{x}$. Using the Dini Theorem (Theorem 2.2.2) we get the desired uniform convergence. □

Corollary 2.4.1. *Under the same hypotheses than in Theorem 2.4.1 we have the following: if f and $g \in \overline{\mathcal{A}}$ then so are* $\max(f,g)$ *and* $\min(f,g)$.

Proof. By continuity of the sum and of the product, we remark that if $\mathcal{A}$ is an algebra, then so is $\overline{\mathcal{A}}$. On the other hand,

$$\max(f,g) := \frac{1}{2}(f+g+|f-g|) \text{ and } \min(f,g) := \frac{1}{2}(f+g-|f-g|).$$

Finally, if $d_u(f,g) > 0$, since

$$P_n\left(\frac{(f-g)^2}{d_u(f,g)^2}\right) d_u(f,g) \to |f-g|$$

uniformly on X, we deduce that $|f-g| \in \overline{\mathcal{A}}$. If rather $d_u(f,g) = 0$, then $|f-g| = 0$ and since $\mathcal{A}$ contains constants, the conclusion follows also in that case. □

Proof of Theorem 2.4.1. The proof will be divided in four main steps:

Step 1: For any x, $y \in X$, and any α, $\beta \in \mathbb{R}$, there exists $f \in \overline{\mathcal{A}}$ such that $f(x) = \alpha$ and $f(y) = \beta$.

Indeed, if $\alpha = \beta$ it suffices to consider the function constant equal to $\alpha = \beta$. On the other hand, if $\alpha \neq \beta$, then by hypothesis there exists $g \in \mathcal{A}$ such that $g(x) \neq g(y)$. Then the function

$$f := \alpha + \frac{\beta - \alpha}{g(y) - g(x)}(g - g(x)),$$

belongs to $\mathcal{A}$ and satisfies $f(x) = \alpha$ and $f(y) = \beta$.

Step 2: Let $h \in \mathcal{C}(X;\mathbb{R})$, $x \in X$ and $\varepsilon > 0$. Then there exists f^x in $\overline{\mathcal{A}}$ such that $f^x(x) = h(x)$ and $f^x(y) < h(y) + \varepsilon$ for any $y \in X$.

Indeed, for any $y \in X$, there exists f_y in $\overline{\mathcal{A}}$ such that $f_y(x) = h(x)$ and $f_y(y) = h(y)$. Since h and f_y are continuous, there exists $r_y > 0$ such that for any $z \in B(y, r_y)$,

$$|f_y(z) - f_y(y)| < \frac{\varepsilon}{2} \text{ and } |h(y) - h(z)| < \frac{\varepsilon}{2}.$$

Hence since $f_y(y) = h(y)$, we deduce that $f_y(z) < h(z) + \varepsilon$ for any $z \in B(y, r_y)$. Now using that X is compact, we extract from the cover $\{B(y, r_y)\}_{y \in X}$ a finite subcover $\{B(y_i, r_{y_i})\}_{1 \leqslant i \leqslant l}$. The function

$$f^x := \min_{1 \leqslant i \leqslant l} f_{y_i}$$

belongs to $\overline{\mathcal{A}}$ thanks to Corollary 2.4.1, and satisfies $f^x(x) = h(x)$ and $f^x(y) < h(y) + \varepsilon$ for any $y \in X$.

Step 3: Let $h \in \mathcal{C}(X;\mathbb{R})$ and $\varepsilon > 0$. Then there exists $f \in \overline{\mathcal{A}}$ such that $h(y) - \varepsilon < f(y) < h(y) + \varepsilon$ for any $y \in X$.

Indeed, let $x \in X$. Since the functions f^x (constructed in step 2) and h are continuous, there exists $r'_x > 0$ such that for any $y \in B(x, r'_x)$,

$$|f^x(y) - f^x(x)| < \frac{\varepsilon}{2} \text{ and } |h(y) - h(x)| < \frac{\varepsilon}{2}.$$

Hence since $f^x(x) = h(x)$, we infer that $f^x(y) > h(y) - \varepsilon$ for all $y \in B(x, r'_x)$. From the open cover $\{B(x, r'_x)\}_{x \in X}$ of X we extract a finite subcover $\{B(x_j, r_{x_j})\}_{1 \leqslant j \leqslant m}$ and we introduce the function

$$f := \max_{1 \leqslant j \leqslant m} f^{x_j}$$

which belongs to $\overline{\mathcal{A}}$ from Corollary 2.4.1, and satisfies $h(y) - \varepsilon < f(y) < h(y) + \varepsilon$ for any $y \in X$.

Step 4: It follows from step 3 that $\overline{\mathcal{A}} = \mathcal{C}(X;\mathbb{R})$. □

We now give a several consequences of the Stone-Weierstrass Theorem.

Corollary 2.4.2. *Let F be a bounded and closed subset of $\mathbb{R}^N$ and $f \in \mathcal{C}(F;\mathbb{R})$. Then there exists a sequence $(P_n)_{n \in \mathbb{N}}$ of polynomial functions of N variables with real coefficients which converges uniformly to f on F.*

Proof. The algebra of all polynomial functions with real coefficients clearly separates points and contains all constants. □

Corollary 2.4.3. *Let F be a bounded and closed subset of $\mathbb{R}^N$, then $\mathcal{C}(F;\mathbb{R})$ is separable.*

Proof. The space of all polynomial functions of N variables with rational coefficients is countable. The conclusion follows from Corollary 2.4.2 and of the density of $\mathbb{Q}$ in $\mathbb{R}$. □

Corollary 2.4.4. *Let (X,d) be a compact metric space then, $\mathcal{C}(X;\mathbb{R})$ is separable.*

Proof. Since X is compact, it is separable according to Proposition 1.4.1. Let us denote by $(x_n)_{n\in\mathbb{N}}$ a (countable) sequence dense in X. Let us introduce $\mathbb{R}[(y_n)_{n\in\mathbb{N}}]$ the set of all polynomials with countably many variables. Then we denote by

$$\mathcal{A} := \{f \in \mathcal{C}(X;\mathbb{R}) : \text{ there exists } P \in \mathbb{R}[(y_n)_{n\in\mathbb{N}}] \text{ such that } f(x) = P(d(x_0,x), d(x_1,x), \ldots)\}.$$

Then $\mathcal{A}$ is clearly a subalgebra of $\mathcal{C}(X;\mathbb{R})$ which separates points and contains the constants. As a consequence of the Stone-Weierstrass Theorem we get that $\overline{\mathcal{A}} = \mathcal{C}(X;\mathbb{R})$. Then we check that the set

$$\mathcal{B} := \{f \in \mathcal{C}(X;\mathbb{R}) : \text{ there exists } P \in \mathbb{Q}[(y_n)_{n\in\mathbb{N}}] \text{ such that } f(x) = P(d(x_0,x), d(x_1,x), \ldots)\}$$

is countable and dense in $\mathcal{A}$. □

Let us conclude this chapter with a few comments about Theorem 2.4.1.

The interested reader could observe that Theorem 2.4.1 also holds true when we weaken the assumption that $\mathcal{A}$ contains constant functions into the assumption that $\mathcal{A}$ vanishes at no point: for any x in X, there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$. The separation and nonvanishing conditions are necessary as well as sufficient for the uniform closure to contain all the continuous functions. For instance, if there exists x in X such that for any $f \in \mathcal{A}$, $f(x) = 0$ then any $f \in \overline{\mathcal{A}}$ will also satisfy $f(x) = 0$ because the limit (even a pointwise limit, let alone a uniform limit) of a sequence of functions which are equal to zero at a point will also be equal to zero at that point. A similar comment holds for separation of points.

Although it is not true that an arbitrary continuous function on an arbitrary compact set in $\mathbb{C}$ can be approximated uniformly by holomorphic polynomials, the obstructions to approximability are known. One obstruction is analytic in nature: if f is the uniform limit of holomorphic polynomials on a compact set X, then f must be holomorphic in the interior of X ; while the other obstruction is topological: if a compact set X has holes, then a function f cannot be approximated by holomorphic polynomials if f has singularities hiding in the holes.

The simplest version of Mergelyan's theorem says that if X is a compact subset of $\mathbb{C}$ having no holes (that is, the complement $\mathbb{C} \setminus K$ is connected), and if f is continuous on X and holomorphic in the interior of X, then f can indeed be approximated uniformly on X by holomorphic polynomials. An older and weaker theorem of Runge has the same conclusion, but makes the stronger hypothesis that f is holomorphic in an open neighborhood of X (not just in the interior of X).

On the other hand there is a counterpart of Theorem 2.4.1 for complex-valued functions, for which the assumptions are strengthened to require that A is closed under taking complex conjugates of functions.

Chapter 3

Measure theory and Lebesgue integration

3.1 Measurable spaces and measurable functions

Definition 3.1.1. *A collection $\mathfrak{M}$ of subsets of a set X is said to be a σ-algebra in X if it satisfies the following properties:*

1. $X \in \mathfrak{M}$;
2. *If $A \in \mathfrak{M}$, then $A^c \in \mathfrak{M}$, where $A^c := X \setminus A$ is the complement of A in X;*
3. *if $A = \bigcup_{n=1}^{\infty} A_n$, with $A_n \in \mathfrak{M}$ for each $n \in \mathbb{N}$, then $A \in \mathfrak{M}$.*

Then we say that $(X, \mathfrak{M})$ is a measurable space, and the elements of $\mathfrak{M}$ are called measurable sets.

If (X, τ) is a topological space, we can consider the smallest σ-algebra, denoted by $\mathcal{B}(X)$, containing all open sets. It is called the Borel σ-algebra and its elements are the Borel sets. Note that such a σ-algebra exists. Indeed, the collection of all σ-algebras containing the open sets is not empty since it contains $\mathcal{P}(X)$, and the intersection of a family of σ-algebras remains a σ-algebra. In particular open sets and closed sets are Borel measurable.

Definition 3.1.2. *Let $(X, \mathfrak{M})$ be a measurable space and (Y, τ) be a topological space. A function $f : X \to Y$ is said to be measurable if $f^{-1}(V) \in \mathfrak{M}$ for every $V \in \tau$.*

Note the analogy, on the one hand, between measurable and topological spaces (see Definition 1.1.1), and on the other hand between measurable and continuous functions (see Definition 1.1.5). If $\mathfrak{M} = \mathcal{B}(X)$, then f is called a Borel measurable function. In particular, continuous functions are Borel measurable. An obvious consequence of these definitions is the following proposition:

Proposition 3.1.1. *Let $(X, \mathfrak{M})$ be measurable space and (Y, τ), (Z, τ') be two topological spaces. Consider a measurable function $f : X \to Y$ and a continuous function $g : Y \to Z$. Then the function $g \circ f : X \to Z$ is measurable.*

When the target space is $\mathbb{R}$, every open subset of $\mathbb{R}$ can be written as the countable union of open intervals. Hence since σ-algebras are stable under countable union, we infer that a function $f : X \to \mathbb{R}$ is measurable if and only if the sets $\{f < a\} \in \mathfrak{M}$ for every $a \in \mathbb{R}$. Moreover, since $\{f \leqslant a\} = \bigcap_n \{f < a + 1/n\}$ and $\{f < a\} = \bigcup_n \{f \leqslant a - 1/n\}$, and since σ-algebras are stable under countable intersection, then f is measurable if and only if the sets $\{f \leqslant a\} \in \mathfrak{M}$ for every $a \in \mathbb{R}$.

In practice, it is not always easy to check that a function is measurable. However, we have the following properties of measurable functions:

Proposition 3.1.2. *(i) Let f and $g : X \to \mathbb{R}$ be two measurable functions. Then so are $f + g$, fg, $|f|$, $\min(f, g)$ and $\max(f, g)$.*

(ii) Let $f_n : X \to \mathbb{R}$ be measurable functions, then $\inf_n f_n$, $\sup_n f_n$, $\liminf_n f_n$, $\limsup_n f_n$ are measurable as well.

Proof. To show that fg and $f+g$ are measurable, from Proposition 3.1.1, it is enough to check that the map $\phi : x \mapsto (f(x), g(x))$ is measurable from X to $\mathbb{R}^2$. Indeed, consider an open set $V \subset \mathbb{R}^2$. Then there exist open intervals I_n and J_n of $\mathbb{R}$ such that

$$V = \bigcup_{n=1}^{\infty} (I_n \times J_n).$$

Hence,

$$\phi^{-1}(V) = \bigcup_{n=1}^{\infty} \phi^{-1}(I_n \times J_n) = \bigcup_{n=1}^{\infty} f^{-1}(I_n) \cap g^{-1}(J_n) \in \mathfrak{M},$$

since $f^{-1}(I_n)$ and $g^{-1}(J_n) \in \mathfrak{M}$ by the measurability of f and g, and since $\mathfrak{M}$ is stable by countable union.

Since $\{\min(f, g) < a\} = \{f < a\} \cup \{g < a\} \in \mathfrak{M}$, and $\{\max(f, g) < a\} = \{f < a\} \cap \{g < a\} \in \mathfrak{M}$, we deduce that the functions $\min(f, g)$ and $\max(f, g)$ are also measurable. Moreover, as $|f| = \max(f, 0) - \min(f, 0)$, then $|f|$ is measurable as well.

Similarly, since $\{\inf_n f_n < a\} = \bigcup_{n=1}^{\infty} \{f_n < a\} \in \mathfrak{M}$ and $\{\sup_n f_n \leqslant a\} = \bigcap_{n=1}^{\infty} \{f_n \leqslant a\} \in \mathfrak{M}$, then $\inf_n f_n$ and $\sup_n f_n$ are measurable. Finally by definitions of lower and upper limits

$$\liminf_n f_n := \sup_{k \in \mathbb{N}} \inf_{n \geqslant k} f_k, \quad \limsup_n f_n := \inf_{k \in \mathbb{N}} \sup_{n \geqslant k} f_k$$

so that $\liminf_n f_n$ are $\limsup_n f_n$ are measurable. □

Examples of measurable functions are characteristic functions of a measurable set $A \in \mathfrak{M}$, defined by $\chi_A(x) = 1$ if $x \in A$, and $\chi_A(x) = 0$ if $x \notin A$. Another example are simple functions which are linear combination of characteristic functions:

$$s(x) := \sum_{i=1}^{n} c_i \chi_{A_i}(x),$$

where $c_i \in \mathbb{R}$ and $A_i \in \mathfrak{M}$ for $i = 1, \ldots, n$. The following result states that it is always possible to approximate measurable functions by a sequence of simple functions. It will be instrumental in the definition of the Lebesgue integral.

Theorem 3.1.1. *Let $f : X \to [0, +\infty]$ be a measurable function. There exists a nondeacreasing sequence (s_n) of simple measurable functions such that $s_n(x) \nearrow f(x)$ for each $x \in X$, as $n \to \infty$.*

Proof. For each $n \in \mathbb{N}$ and $k \in \{0, \ldots, n2^n - 1\}$, define the measurable sets

$$E_{n,k} := \left\{ x \in X : \frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n} \right\}, \quad F_n := \{f \geqslant n\}.$$

Now define for each $x \in X$,

$$s_n(x) := \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \chi_{E_{n,k}}(x) + n\chi_{F_n}(x).$$

The sequence (s_n) clearly fulfills the conclusion of the Theorem. □

3.2 Positive measures

Definition 3.2.1. *Let $(X, \mathfrak{M})$ be a measurable space. A set function $\mu : \mathfrak{M} \to [0, +\infty]$ is called a positive measure (or simply a measure) if $\mu(\emptyset) = 0$ and if it is countably additive, i.e.,*

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

for every $A_n \in \mathfrak{M}$ such that $A_n \cap A_m = \emptyset$ if $n \neq m$.

If (X, τ) is a topological space, and μ is a measure over $\mathcal{B}(X)$, then it is called a Borel measure. If further $\mu(K) < \infty$ for every compact set $K \subset X$, then μ is a positive Radon measure.

The following result states the main general properties of positive measures:

Proposition 3.2.1. *Let μ be a positive measure over a measurable space $(X, \mathfrak{M})$. Then*

(i) If A, $B \in \mathfrak{M}$ and $A \subset B$, then $\mu(A) \leqslant \mu(B)$;

(ii) If, for each $n \in \mathbb{N}^*$, $A_n \in \mathfrak{M}$ and $A_n \subset A_{n+1}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n);$$

(iii) If, for each $n \in \mathbb{N}^*$, $A_n \in \mathfrak{M}$ and $A_{n+1} \subset A_n$, and $\mu(A_1) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

Proof. If A, $B \in \mathfrak{M}$ and $A \subset B$, then $B = A \cup (B \setminus A)$. Hence by additivity, $\mu(B) = \mu(A) + \mu(B \setminus A) \geqslant \mu(A)$.

Let $A_n \in \mathfrak{M}$ such that $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}^*$. If there exists $n \in \mathbb{N}^*$ such that $\mu(A_n) = +\infty$ then the result follows. Therefore we assume from now on that $\mu(A_n) < +\infty$ for each $n \in \mathbb{N}^*$. Then define $B_1 = A_1$, and for all $n \geqslant 2$, $B_n := A_n \setminus A_{n-1}$. By construction $B_n \in \mathfrak{M}$ and $B_n \cap B_m = \emptyset$ if $n \neq m$. Moreover

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

Hence by countable additivity,

$$\begin{aligned}\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} \mu(B_n) &= \lim_{n\to\infty} \sum_{k=2}^{n} (\mu(A_k \setminus A_{k-1}) + \mu(A_1) \\ &= \lim_{n\to\infty} \sum_{k=2}^{n} (\mu(A_k) - \mu(A_{k-1})) + \mu(A_1) = \lim_{n\to\infty} \mu(A_n),\end{aligned}$$

which completes the proof of (ii).

Assertion (iii) follows from (ii) since

$$\mu(A_1) - \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu(A_1 \setminus A_n) = \mu\left(\bigcup_{n=1}^{\infty} (A_1 \setminus A_n)\right) = \mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right).$$

□

3.3 Definition and properties of the Lebesgue integral

3.3.1 Lebesgue integral of non negative measurable functions

Definition 3.3.1. *If $s : X \to [0, +\infty)$ is a simple measurable function of the form*

$$s = \sum_{i=1}^{n} c_i \chi_{A_i},$$

where $c_i \geqslant 0$ and $A_i \in \mathfrak{M}$ for $i = 1, \ldots, n$ are pairwise disjoint, and if $E \in \mathfrak{M}$, we define

$$\int_E s \, d\mu := \sum_{i=1}^{n} c_i \mu(A_i \cap E),$$

with the convention that $0 \cdot \infty = 0$.

If $f : X \to [0, +\infty]$ *is a measurable function, and* $E \in \mathfrak{M}$, *we define*

$$\int_E f \, d\mu := \sup \int_E s \, d\mu,$$

where the supremum is taken over all simple measurable functions $s : X \to [0, +\infty)$ *such that* $s \leqslant f$.

Note that if f is a non negative simple measurable function, both definition coincide. Moreover by Theorem 3.1.1 the family of all non negative measurable simple functions less than f is not empty.

The following properties are immediate consequences of the definitions, and the proof is left to the reader. The functions and sets occuring are assumed to be measurable.

Proposition 3.3.1. *1. If* $0 \leqslant f \leqslant g$, *then* $\int_E f \, d\mu \leqslant \int_E g \, d\mu$;

2. *If* $A \subset B$ *and* $f \geqslant 0$, *then* $\int_A f \, d\mu \leqslant \int_B f \, d\mu$;
3. *If* $f \geqslant 0$ *and* c *is a constant,* $0 \leqslant c < \infty$, *then* $\int_E cf \, d\mu = c \int_E f \, d\mu$;
4. *If* $f(x) = 0$ *for all* $x \in E$, *then* $\int_E f \, d\mu = 0$, *even if* $\mu(E) = \infty$;
5. *If* $\mu(E) = 0$, *then* $\int_E f \, d\mu = 0$ *even if* $f \equiv \infty$;

Before extending the definition of the Lebesgue integral to real valued functions, we establish two very important convergence results which are typical of non negative valued functions.

Theorem 3.3.1 (Monotone convergence). *Let* (f_n) *be a sequence of measurable functions on* X, *and suppose that*

1. $0 \leqslant f_n(x) \leqslant f_{n+1}(x)$ *for every* $x \in X$ *and every* $n \in \mathbb{N}^*$;
2. $f_n(x) \to f(x)$ *for each* $x \in X$.

Then f *is measurable and*

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

Proof. By Proposition 3.1.2 the function f is measurable. Moreover since $\int_X f_n \, d\mu \leqslant \int_X f_{n+1} \, d\mu$, then there exists the limit

$$\ell := \lim_{n \to \infty} \int_X f_n \, d\mu,$$

and by monotonicity, $\ell \leqslant \int_X f \, d\mu$.

To prove the converse inequality, consider a simple measurable function $0 \leqslant s \leqslant f$, and a constant $c \in (0, 1)$. Let us define the measurable sets $E_n := \{f_n \geqslant cs\}$ for $n \in \mathbb{N}^*$. Then

$$\int_X f_n \, d\mu \geqslant \int_{E_n} f_n \, d\mu \geqslant c \int_{E_n} s \, d\mu. \tag{3.1}$$

Writting $s := \sum_{i=1}^p c_i \chi_{A_i}$ for some $c_i > 0$ and $A_i \in \mathfrak{M}$, we have by definition of the Lebesgue integral that

$$\int_{E_n} s \, d\mu = \sum_{i=1}^p c_i \mu(A_i \cap E_n).$$

We have $E_n \subset E_{n+1}$ and since $f_n \to f$ pointwise and $c < 1$, $\bigcup_{n=1}^\infty E_n = X$. Hence, for fixed i, the sequence $(A_i \cap E_n)_n$ is increasing and its union is A_i; we therefore deduce from Proposition 3.2.1-2 that

$$\int_{E_n} s \, d\mu \to \sum_{i=1}^p c_i \mu(A_i) = \int_X s \, d\mu.$$

Consequently, passing to the limit in (3.1) as $n \to \infty$, we get that for any $c < 1$,

$$\ell \geqslant c \int_X s \, d\mu,$$

and the conclusion follows sending first $c \to 1^-$, and then by taking the supremum over all simple functions s with $0 \leqslant s \leqslant f$. □

Applying the monotone convergence Theorem to the partial sums of a series of non negative measurable functions, we get the following result.

Corollary 3.3.1. *If $f_n : X \to [0, +\infty]$ are measurable functions for all $n \in \mathbb{N}^*$, and*

$$f(x) = \sum_{n=1}^{\infty} f_n(x),$$

then

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

Remark 3.3.1. As a consequence of Proposition 3.3.1-3 and the previous corollary, we infer that

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

for every non negative measurable functions f and g, and every α, $\beta \geqslant 0$.

Remark 3.3.2. Taking $X = \mathbb{N}^*$ and μ the counting measure ($\mu(A) = \mathrm{Card}(A)$ if $\mathrm{Card}(A) < \infty$, and ∞ otherwise), we deduce that

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

for any doubly indexed sequence (a_{nk}) with $a_{nk} \geqslant 0$ for every n, $k \in \mathbb{N}^*$.

If the sequence miss to converge we get the following lower semicontinuity result.

Lemma 3.3.1 (Fatou). *If $f_n : X \to [0, \infty]$ is are measurable functions for each $n \in \mathbb{N}^*$, then*

$$\int_X \liminf_{n \to \infty} f_n \, d\mu \leqslant \liminf_{n \to \infty} \int_X f_n \, d\mu.$$

Proof. It suffices to apply the monotone convergence Theorem to the sequence $g_n := \inf_{k \geqslant n} f_k$ by observing that $g_n \leqslant f_n$ for each $n \in \mathbb{N}^*$, that the sequence $(g_n)_n$ is increasing and that $\lim_n g_n = \liminf_n f_n$. □

3.3.2 Lebesgue integral of real valued measurable functions

In the sequel, and unless otherwise mentioned, μ is a measure over a measurable space $(X, \mathfrak{M})$.

Definition 3.3.2. *We define $\mathcal{L}^1(X, \mu)$ as the space of all measurable functions $f : X \to \mathbb{R}$ such that*

$$\int_X |f| \, d\mu < \infty.$$

Note that the measurability of f implies that of $|f|$ by Proposition 3.1.2 so that the previous integral is well defined according to Definition 3.3.1. The elements of $\mathcal{L}^1(X, \mu)$ are called Lebesgue integrable functions (with respect to μ).

Definition 3.3.3. *If $f \in \mathcal{L}^1(X, \mu)$ and $E \in \mathfrak{M}$, we define*

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

where $f^+ := \max(f, 0)$ and $f^- := \max(-f, 0)$ are (non negative) measurable functions thanks to Proposition 3.1.2.

Note that the Lebesgue integral of a measurable function f is not affected if we modify the values of f on a set of μ-measure zero. More precisely, let f and g are two measurable functions and $Z \in \mathfrak{M}$ is such that $\mu(Z) = 0$. If $f(x) = g(x)$ for all $x \in X \setminus Z$, then $\int_X f \, d\mu = \int_X g \, d\mu$.

Definition 3.3.4. *If a property P holds outside a set of μ-measure zero, we say that P holds μ almost everywhere (a.e.) in X.*

In particular we have that if $f = g$ μ-a.e. in X, then $\int_X f\,d\mu = \int_X g\,d\mu$. Another application of this terminology is the following result.

Proposition 3.3.2. *Let $f \in \mathcal{L}^1(X, \mu)$, then $|f| < \infty$ μ-a.e. in X.*

Proof. Define $A_n := \{|f| \geqslant n\}$ for $n \in \mathbb{N}^*$. Since f is measurable, then $A_n \in \mathfrak{M}$ for each $n \in \mathbb{N}^*$ and by Proposition 3.3.1, then

$$\mu(A_n) \leqslant \frac{1}{n} \int_X |f|\,d\mu.$$

In particular $\mu(A_1) < \infty$ and (A_n) is a sequence of decreasing measurable sets whose intersection is $\{|f| = \infty\}$. Applying Proposition 3.2.1-3, we deduce that

$$\mu(\{|f| = \infty\}) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = 0.$$

□

The space $\mathcal{L}^1(X, \mu)$ turns out to be a linear space. Indeed,

Theorem 3.3.2. *Suppose that f, $g \in \mathcal{L}^1(X, \mu)$, and α, $\beta \in \mathbb{R}$. Then*

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

Proof. We first observe that $\alpha f + \beta g$ is measurable (by Proposition 3.1.2) and that it belongs to $\mathcal{L}^1(X, \mu)$. Indeed, since $|\alpha f + \beta g| \leqslant |\alpha||f| + |\beta||g|$, then by Proposition 3.3.1-1, we get that

$$\int_X |\alpha f + \beta g|\,d\mu \leqslant \int_X (|\alpha||f| + |\beta||g|)\,d\mu = |\alpha| \int_X |f|\,d\mu + |\beta| \int_X |g|\,d\mu < \infty.$$

Clearly, it suffices to show that

$$\int_X (f + g)\,d\mu = \int_X f\,d\mu + \int_X g\,d\mu, \tag{3.2}$$

and

$$\int_X \alpha f\,d\mu = \alpha \int_X f\,d\mu. \tag{3.3}$$

To prove (3.2), we have $(f + g)^+ - (f + g)^- = f^+ - f^- + g^+ - g^-$, or still $(f + g)^+ + f^- + g^- = (f + g)^- + f^+ + g^+$. Integrating over X and using Remark 3.3.1 leads to

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

Changing back the order of the terms yields (3.2).

The proof of (3.3) follows from Proposition 3.3.1-3 if $\alpha \geqslant 0$. If $\alpha \leqslant 0$, then $-\alpha \geqslant 0$, $-u = u^- - u^+$ and the result follows again from Proposition 3.3.1-3. □

We next prove several convergence results.

Theorem 3.3.3 (Dominated convergence). *Let f_n and f be measurable functions such that $f_n(x) \to f(x)$ for μ-a.e. $x \in X$. Assume that there exists $g \in \mathcal{L}^1(X, \mu)$ such that $|f_n(x)| \leqslant g(x)$ for μ-a.e. $x \in X$ and every $n \in \mathbb{N}^*$. Then $f \in \mathcal{L}^1(X, \mu)$ and*

$$\int_X |f_n - f|\,d\mu \to 0.$$

Proof. By Fatou's Lemma and Theorem 3.3.2,

$$\int_X 2g\,d\mu = \int_X \liminf_{n\to\infty}(2g - |f - f_n|)\,d\mu \leqslant \liminf_{n\to\infty} \int_X (2g - |f - f_n|)\,d\mu,$$

whence

$$\limsup_{n\to\infty} \int_X |f - f_n|\,d\mu \leqslant 0.$$

□

The following result is a kind of converse of the dominated convergence Theorem.

Theorem 3.3.4. *Let f_n and f be functions in $\mathcal{L}^1(X,\mu)$ and*

$$\int_X |f_n - f|\,d\mu \to 0.$$

Then there exists a subsequence (f_{n_k}) and $g \in \mathcal{L}^1(X,\mu)$ such that $|f_{n_k}| \leqslant g$ and $f_{n_k} \to f$ μ-a.e. in X.

Proof. For each $k \in \mathbb{N}^*$, there exists $n_k \in \mathbb{N}^*$ such that $n_k \to \infty$ as $k \to \infty$ and

$$\int_X |f_{n_k} - f|\,d\mu \leqslant \frac{1}{2^k}.$$

Define $g := |f| + \sum_{k=1}^{\infty} |f_{n_k} - f|$. By Corollary 3.3.1,

$$\int_X g\,d\mu = \int_X |f|\,d\mu + \sum_{k=1}^{\infty} \int_X |f_{n_k} - f|\,d\mu \leqslant \int_X |f|\,d\mu + 1$$

so that $g \in \mathcal{L}^1(X,\mu)$. Thus by construction $|f_{n_k}| \leqslant g$ μ-a.e. in X, and since

$$\int_X \sum_{k=1}^{\infty} |f_{n_k} - f|\,d\mu = \sum_{k=1}^{\infty} \int_X |f_{n_k} - f|\,d\mu \leqslant 1,$$

we deduce from Proposition 3.3.2 that $\sum_{k=1}^{\infty} |f_{n_k} - f| < \infty$ μ-a.e. in X and thus $f_{n_k} \to f$ μ-a.e. in X. □

3.4 Modes of convergence

3.4.1 Definitions and relationships

Definition 3.4.1. *Let f_n and f be measurable functions. We say that*

1. *(f_n) converges to f almost everywhere if there exists a measurable set $Z \subset X$ with $\mu(Z) = 0$ such that $f_n(x) \to f(x)$ for each $x \in X \setminus Z$;*

2. *(f_n) converges to f in $\mathcal{L}^1(X,\mu)$ if $\int_X |f_n - f|\,d\mu \to 0$;*

3. *(f_n) converges to f in measure if for any $\varepsilon > 0$, $\mu(\{|f_n - f| > \varepsilon\}) \to 0$;*

4. *(f_n) converges to f almost uniformly if for any $\varepsilon > 0$, there exists a measurable set $E_\varepsilon \subset X$ such that $\mu(X \setminus E_\varepsilon) < \varepsilon$ and*

$$\sup_{x \in E_\varepsilon} |f_n(x) - f(x)| \to 0.$$

Remark 3.4.1. We can easily see that convergence in $\mathcal{L}^1(X,\mu)$ also implies convergence in measure as a consequence of the following inequality

$$\varepsilon\mu(\{|f_n - f| > \varepsilon\}) \leqslant \int_X |f_n - f|\,d\mu.$$

Another link easy to obtain is that convergence almost uniformly implies convergence almost everywhere.

Proposition 3.4.1. *Let f_n and f be measurable functions. If (f_n) converges to f almost uniformly, then it converges to f almost everywhere.*

Proof. If (f_n) converges to f almost uniformly, then for any $\varepsilon > 0$, there exists a measurable set $E_\varepsilon \subset X$ such that $\mu(X \setminus E_\varepsilon) < \varepsilon$ and $\sup_{E_\varepsilon} |f_n - f| \to 0$. Take $\varepsilon = 1/k$ with $k \in \mathbb{N}^*$, and define $Z := \bigcap_{k=1}^\infty (X \setminus E_{1/k})$. Then for any $k \in \mathbb{N}^*$, one has $\mu(Z) \leqslant \mu(X \setminus E_{1/k}) < 1/k \to 0$ so that $\mu(Z) = 0$. Moreover for each $x \in X \setminus Z$, there exists $k \in \mathbb{N}^*$ such that $x \in E_{1/k}$, and in particular, $f_n(x) \to f(x)$ ☐

The converse statement is true in the case of finite measures.

Theorem 3.4.1 (**Egoroff**). *Assume that μ is a finite measure, i.e., $\mu(X) < \infty$, and consider some measurable functions f_n and f. If (f_n) converges almost everywhere to f, then f converges almost uniformly to f.*

Proof. For every $i, j \in \mathbb{N}^*$, define the measurable sets $E_{i,j} := \bigcup_{n=j}^\infty \{|f_n - f| > 2^{-i}\}$. Then $E_{i,j+1} \subset E_{i,j}$, and since $\mu(X) < \infty$, by Proposition 3.2.1-3, we have

$$\lim_{j\to\infty} \mu(E_{i,j}) = \mu\left(\bigcap_{j=1}^\infty E_{i,j}\right) = 0,$$

since (f_n) converges almost everywhere to f. Hence there exists an integer $N(i)$ such that $\mu(E_{i,N(i)}) < \varepsilon/2^i$. Define $E_\varepsilon := X \setminus \bigcup_{i=1}^\infty E_{i,N(i)}$, then $\mu(X \setminus E_\varepsilon) < \varepsilon$. Moreover, if $x \in E_\varepsilon$, then for each $i \in \mathbb{N}^*$, $x \in X \setminus E_{i,N(i)}$ and thus for any $n \geqslant N(i)$, $|f_n(x) - f(x)| \leqslant 2^{-i}$. Consequently, $f_n \to f$ uniformly in E_ε. ☐

Let us now compare convergence almost uniformly and convergence in measure.

Theorem 3.4.2. *Let f_n and f be some measurable functions. Then*

1. *If (f_n) converges to f almost uniformly, then it converges to f in measure;*
2. *If (f_n) converges to f in measure, then there exists a subsequence (f_{n_k}) which converges to f almost uniformly and almost everywhere.*

Proof. 1- If (f_n) converges to f almost uniformly, then for any $\varepsilon > 0$, there exists a measurable set $E_\varepsilon \subset X$ such that $\mu(X \setminus E_\varepsilon) < \varepsilon$ and $\sup_{E_\varepsilon} |f_n - f| \to 0$. Let $\delta > 0$, then

$$\mu(\{|f_n - f| > \delta\}) = \mu(\{|f_n - f| > \delta\} \cap E_\varepsilon) + \mu(\{|f_n - f| > \delta\} \setminus E_\varepsilon) \leqslant \mu(\{|f_n - f| > \delta\} \cap E_\varepsilon) + \varepsilon.$$

But since $f_n \to f$ uniformly in E_ε, there exists $n(\delta, \varepsilon) \in \mathbb{N}^*$ such that for every $n \geqslant n(\delta, \varepsilon)$, $\sup_{E_\varepsilon} |f_n - f| < \delta$ so that $\{|f_n - f| > \delta\} \cap E_\varepsilon = \emptyset$. Hence for $n \geqslant n(\delta, \varepsilon)$, one has $\mu(\{|f_n - f| > \delta\}) < \varepsilon$. Consequently,

$$\limsup_{n\to\infty} \mu(\{|f_n - f| > \delta\}) \leqslant \varepsilon,$$

and letting finally ε tend to zero, we complete the proof of the first statement.

2- Let $\varepsilon > 0$. If $(f_n)_n$ converges in measure to f, then for each $k \in \mathbb{N}^*$, there exists a sequence $n_k \nearrow \infty$ such that

$$\mu\left(\left\{|f_{n_k} - f| > \frac{1}{k}\right\}\right) < \frac{\varepsilon}{2^k}.$$

Define the measurable sets $E_k := \{|f_{n_k} - f| \leqslant 1/k\}$ and $E_\varepsilon := \bigcap_{k=1}^\infty E_k$. Then

$$\mu(X \setminus E_\varepsilon) \leqslant \sum_{k=1}^\infty \mu(X \setminus E_k) \leqslant \varepsilon,$$

and for every $x \in E_\varepsilon$, for each $k \in \mathbb{N}^*$, we have $|f_{n_k}(x) - f(x)| \leqslant \frac{1}{k}$. Thus $f_{n_k} \to f$ almost uniformly and also almost everywhere. ☐

3.4.2 Equi-integrability

Definition 3.4.2. *A sequence of measurable functions $(f_n)_{n\in\mathbb{N}^*}$ is said to be equi-integrable if for any $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\sup_{n\in\mathbb{N}^*} \int_E |f_n|\, d\mu \leqslant \varepsilon,$$

for every measurable set $E \subset X$ satisfying $\mu(E) \leqslant \delta$.

The equi-integrability condition expresses the fact that the sequence (f_n) does not concentrate on sets of arbitrarily small measure. We now give two necessary and sufficient conditions which ensure the equi-integrability of a sequence.

Theorem 3.4.3. *Let $(f_n)_{n\in\mathbb{N}^*}$ be a bounded sequence in $\mathcal{L}^1(X, \mu)$, i.e.,*

$$\sup_{n\in\mathbb{N}^*} \int_X |f_n|\, d\mu < \infty.$$

Then the following conditions are equivalent:

(i) *The sequence (f_n) is equi-integrable;*

(ii)

$$\lim_{t\to\infty} \sup_{n\in\mathbb{N}^*} \int_{\{|f_n|>t\}} |f_n|\, d\mu = 0;$$

(iii) **(De la Vallée Poussin criterion)** *There exists an increasing function $\theta : [0, \infty) \to [0, \infty]$ with $\theta(t)/t \to \infty$ as $t \to \infty$ and such that*

$$\sup_{n\in\mathbb{N}^*} \int_X \theta(|f_n|)\, d\mu < \infty.$$

Proof. Let us define

$$M := \sup_{n\in\mathbb{N}^*} \int_X |f_n|\, d\mu < \infty.$$

Step 1. Assume that (f_n) is equi-integrable and, given $\varepsilon > 0$ let $\delta > 0$ be such that

$$\sup_{n\in\mathbb{N}^*} \int_E |f_n|\, d\mu \leqslant \varepsilon,$$

for every measurable set $E \subset X$ with $\mu(E) \leqslant \delta$. Choose $t_\varepsilon > 0$ such that $\frac{M}{t_\varepsilon} \leqslant \delta$. Then by definition of M, for every $n \in \mathbb{N}^*$ and for all $t \geqslant t_\varepsilon$ we have

$$\mu(\{|f_n| > t\}) \leqslant \frac{1}{t} \int_X |f_n|\, d\mu \leqslant \frac{M}{t} \leqslant \delta,$$

and so

$$\sup_{n\in\mathbb{N}^*} \int_{\{|f_n|>t\}} |f_n|\, d\mu \leqslant \varepsilon,$$

and this validates (ii).

Conversely, suppose that (ii) holds, fix $\varepsilon > 0$, and choose $t_\varepsilon > 0$ such that

$$\sup_{n\in\mathbb{N}^*} \int_{\{|f_n|>t_\varepsilon\}} |f_n|\, d\mu \leqslant \frac{\varepsilon}{2}.$$

Then for every measurable set $E \subset X$ with $\mu(E) \leqslant \frac{\varepsilon}{2t_\varepsilon} =: \delta$ and for all $n \in \mathbb{N}^*$ we have

$$\int_E |f_n|\, d\mu = \int_{E\cap\{|f_n|>t_\varepsilon\}} |f_n|\, d\mu + \int_{E\cap\{|f_n|\leqslant t_\varepsilon\}} |f_n|\, d\mu \leqslant \frac{\varepsilon}{2} + t_\varepsilon \mu(E) \leqslant \varepsilon.$$

Step 2. Assume that (ii) holds and construct an increasing sequence of positive integers (k_i) such that

$$\sup_{n\in\mathbb{N}^*}\int_{\{|f_n|>k_i\}}|f_n|\,d\mu\leqslant\frac{1}{2^i}.$$

Define $b_0=0$ and for each $l\in\mathbb{N}^*$ let b_l be the number of non negative integers i such that $k_i<l$. Note that $b_l\nearrow\infty$ as $l\to\infty$. Define

$$\theta(t)=tb_l\quad\text{ if } t\in[l,l+1).$$

Then

$$\frac{\theta(t)}{t}\geqslant b_{[t]}\to\infty$$

as $t\to\infty$, were $[t]$ denotes the integer part of t. Moreover, for all $n\in\mathbb{N}^*$, By Remark 3.3.2,

$$\begin{aligned}\int_X\theta(|f_n|)\,d\mu &= \sum_{l=1}^{\infty}\int_{\{l\leqslant|f_n|<l+1\}}\theta(|f_n|)\,d\mu\\ &= \sum_{l=1}^{\infty}b_l\int_{\{l\leqslant|f_n|<l+1\}}|f_n|\,d\mu=\sum_{l=1}^{\infty}\sum_{\{i\in\mathbb{N}^*:k_i<l\}}\int_{\{l\leqslant|f_n|<l+1\}}|f_n|\,d\mu\\ &= \sum_{i=1}^{\infty}\sum_{\{l\in\mathbb{N}^*:l>k_i\}}\int_{\{l\leqslant|f_n|<l+1\}}|f_n|\,d\mu\leqslant\sum_{i=1}^{\infty}\int_{\{|f_n|>k_i\}}|f_n|\,d\mu\leqslant\sum_{i=1}^{\infty}\frac{1}{2^i}\leqslant 1.\end{aligned}$$

Conversely, assume that (iii) holds and let

$$C:=\sup_{n\in\mathbb{N}^*}\int_X\theta(|f_n|)\,d\mu<\infty.$$

Since $\theta(t)/t\to\infty$ as $t\to\infty$, for every $\varepsilon>0$, there exists $t_\varepsilon>0$ such that

$$\theta(t)\geqslant\frac{(C+1)t}{\varepsilon}\qquad\text{for all } t\geqslant t_\varepsilon.$$

Then for all $t\geqslant t_\varepsilon$,

$$\sup_{n\in\mathbb{N}^*}\int_{\{|f_n|>t\}}|f_n|\,d\mu\leqslant\frac{\varepsilon}{C+1}\sup_{n\in\mathbb{N}^*}\int_{\{|f_n|>t\}}\theta(|f_n|)\,d\mu\leqslant\varepsilon.$$

Hence (ii) holds. □

We conclude this chapter by a first application of equi-integrability (see also Chapter 6).

Theorem 3.4.4 (Vitali). *Let μ be a finite measure. Let f_n and f be measurable functions such that the sequence (f_n) if equi-integrable and $f_n\to f$ almost everywhere. Then*

$$\int_X|f_n-f|\,d\mu\to 0.$$

Proof. Let ε and δ be such that

$$\sup_{n\in\mathbb{N}^*}\int_E(|f_n|+|f|)\,d\mu\leqslant\varepsilon$$

for every measurable set $E\subset X$ with $\mu(E)\leqslant\delta$. By Egoroff's Theorem, in correspondence with δ, there exists a measurable set $E_\delta\subset X$ such that $\mu(X\setminus E_\delta)<\delta$ and $f_n\to f$ uniformly in E_δ. Hence

$$\int_X|f_n-f|\,d\mu=\int_{E_\delta}|f_n-f|\,d\mu+\int_{X\setminus E_\delta}|f_n-f|\,d\mu\leqslant\mu(X)\sup_{E_\delta}|f_n-f|+\int_{X\setminus E_\delta}(|f_n|+|f|)\,d\mu.$$

Hence letting $n\to\infty$, we get that

$$\limsup_{n\to\infty}\int_X|f_n-f|\,d\mu\leqslant\varepsilon,$$

and since ε is arbitrary we deduce that $f_n\to f$ in $\mathcal{L}^1(X,\mu)$ as claimed. □

3.5 Positive Radon measures

In this section Ω stands for an open subset of $\mathbb{R}^N$ ($N \geqslant 1$). We recall that a positive Radon measure is a Borel measure (*i.e.* a measure on the Borel σ-algebra $\mathcal{B}(\Omega)$ of Ω) which is finite on compact sets. Radon measures give a close relationship between integration and linear functional on the space $\mathcal{C}_c(\Omega)$ of continuous functions with compact support (the closure of the set of points where the function is not zero) in Ω. Indeed if μ is a Radon measure over Ω, then the mapping

$$f \mapsto \int_\Omega f\, d\mu$$

defines a positive linear functional on $\mathcal{C}_c(\Omega)$, *i.e.*,

$$L(\alpha f + \beta g) = \alpha L(f) + \beta L(g) \quad \text{ for all } f,\, g \in \mathcal{C}_c(\Omega) \text{ and all } \alpha,\, \beta \in \mathbb{R},$$
$$L(f) \geqslant 0 \quad \text{ for all } f \in \mathcal{C}_c(\Omega) \text{ with } f \geqslant 0.$$

We will actually see in the Riesz representation Theorem that every positive linear functional on $\mathcal{C}_c(\Omega)$ can be uniquely represented by a positive Radon measure.

Before, we need to introduce several technical results.

Lemma 3.5.1 (Urysohn). *Let K be a compact set and V be a bounded open set such that $K \subset V \subset \overline{V} \subset \Omega$. Then there exists a function $f \in \mathcal{C}_c(\Omega; [0,1])$ such that $f = 1$ on K and $f = 0$ on $\Omega \setminus \overline{V}$.*

Proof. It suffices to take

$$f(x) := \frac{\operatorname{dist}(x, \Omega \setminus \overline{V})}{\operatorname{dist}(x, \Omega \setminus \overline{V}) + \operatorname{dist}(x, K)}.$$

□

Lemma 3.5.2. *Let $V_1, \dots, V_n$ be open sets satisfying $\overline{V_i} \subset \Omega$ for all $i = 1, \dots, n$ and K be a compact set such that $K \subset \bigcup_{i=1}^n V_i$. Then, for each $i = 1, \dots, n$, there exists some functions $f_i \in \mathcal{C}_c(V_i; [0,1])$ such that $\sum_{i=1}^n f_i = 1$ on K.*

Proof. For each $x \in K$, there exists an open ball B_x centered at x and such that $\overline{B_x} \subset V_i$ for some i (depending on x). Hence $K \subset \bigcup_{x \in K} B_x$, and since K is compact one can extract a finite covering $K \subset \bigcup_{j=1}^p B_{x_j}$. Define K_i as the union of those closed balls $\overline{B_{x_j}}$ which are contained in V_i. Then K_i is a compact subset of V_i and by Urysohn's Lemma, there exists a function $g_i \in \mathcal{C}_c(V_i; [0,1])$ such that $g_i = 1$ on K_i. Then the functions

$$f_i(x) := \begin{cases} \dfrac{g_i(x)}{\sum_{j=1}^n g_j(x)} & \text{if} \quad x \in V_i, \\ 0 & \text{if} \quad x \in \Omega \setminus V_i \end{cases}$$

fulfill the conclusion of the lemma. □

We now state the main result of this section.

Theorem 3.5.1 (Riesz representation Theorem). *Let $L : \mathcal{C}_c(\Omega) \to \mathbb{R}$ be a positive linear functional. Then there exist a σ-algebra $\mathfrak{M}$ containing the Borel σ-algebra $\mathcal{B}(\Omega)$ and a measure μ on $\mathfrak{M}$ such that*

1. $L(f) = \displaystyle\int_\Omega f\, d\mu$, *for every* $f \in \mathcal{C}_c(\Omega)$,
2. $\mu(K) < \infty$ *for every compact set* $K \subset \Omega$,
3. *for every* $E \in \mathfrak{M}$,
$$\mu(E) = \inf\{\mu(V) : E \subset V,\, V \text{ open}\}, \tag{3.4}$$
4. *for every open set E, and every $E \in \mathfrak{M}$ with $\mu(E) < \infty$,*
$$\mu(E) = \sup\{\mu(K) : K \subset E,\, K \text{ compact}\}. \tag{3.5}$$

Moreover this measure is unique in the sense that if $(\mathfrak{M}_1, \mu_1)$ *and* $(\mathfrak{M}_2, \mu_2)$ *satisfy the previous properties then the restrictions of* μ_1 *and* μ_2 *to* $\mathcal{B}(\Omega)$ *are equal.*

Proof. Let us start by proving uniqueness. Assume that μ_1 and μ_2 are two Radon measures satisfying the conclusion of the Riesz representation Theorem. By the regularity property (3.5), it suffices to show that $\mu_1(K) = \mu_2(K)$ for every compact set K. Let $\varepsilon > 0$ and K a compact set. By (3.4), there exists an open set V containing K such that $\mu_2(V) < \mu_2(K) + \varepsilon$. By Urysohn's Lemma one can find a function $f \in \mathcal{C}_c(V; [0,1])$ such that $f = 1$ on K. In particular, $\chi_K \leqslant f \leqslant \chi_V$, hence

$$\mu_1(K) = \int_\Omega \chi_K \, d\mu_1 \leqslant \int_\Omega f \, d\mu_1 = L(f) = \int_\Omega f \, d\mu_2 \leqslant \int_\Omega \chi_V \, d\mu_2 = \mu_2(V) < \mu_2(K) + \varepsilon.$$

Thus $\mu_1(K) \leqslant \mu_2(K)$ and exchanging the roles of μ_1 and μ_2 we deduce that this inequality is actually an equality.

We now turn our attention to the existence. For every open set $V \subset \Omega$, we define

$$\mu(V) := \sup\{L(f) : f \in \mathcal{C}_c(\Omega; [0,1]),\ \mathrm{supp}(f) \subset V\}. \tag{3.6}$$

Clearly, if $V_1 \subset V_2$, then $\mu(V_1) \leqslant \mu(V_2)$ so that we can extend μ to any arbitrary $E \subset \Omega$ by setting

$$\mu(E) := \inf\{\mu(V) : E \subset V,\ V \text{ open}\}.$$

Note that both definitions are consistent for open sets, and that property (3.4) will be automatically satisfied. Moreover, μ is an increasing set function, *i.e.* if $E_1 \subset E_2$, then $\mu(E_1) \leqslant \mu(E_2)$.

Let $\mathfrak{M}_F$ be the family of all sets $E \subset \Omega$ such that $\mu(E) < \infty$ and

$$\mu(E) = \sup\{\mu(K) : K \subset E,\ K \text{ compact}\}.$$

Finally, let $\mathfrak{M}$ be the class of all $E \subset \Omega$ such that $E \cap K \in \mathfrak{M}_F$ for any compact K.

Step 1. *If K is compact, then $K \in \mathfrak{M}_F$, and*

$$\mu(K) = \inf\{L(f) : f \in \mathcal{C}_c(\Omega; [0,1]),\ f = 1 \text{ on } K\}. \tag{3.7}$$

Moreover, if V is open, then V satisfies (3.5). *In particular if $\mu(V) < \infty$, then $V \in \mathfrak{M}_F$.*

Let $f \in \mathcal{C}_c(\Omega; [0,1])$ such that $f = 1$ on K, $\alpha \in (0,1)$, and $V_\alpha := \{f > \alpha\}$. Then $K \subset V_\alpha$ and $\alpha g \leqslant f$ for every $g \in \mathcal{C}_c(V_\alpha; [0,1])$. Hence $\mu(K) \leqslant \mu(V_\alpha) = \sup\{L(g) : g \in \mathcal{C}_c(V_\alpha; [0,1])\} \leqslant \alpha^{-1} L(f)$. Letting $\alpha \to 1$ leads to $\mu(K) \leqslant L(f) < \infty$. Therefore $K \in \mathfrak{M}_F$ since (3.5) is immediate. Next if $\varepsilon > 0$, there exists an open set V containing K with $\mu(V) < \mu(K) + \varepsilon$. By Urysohn's Lemma, there exists a function $f \in \mathcal{C}_c(V; [0,1])$ satisfiying $f = 1$ on K, and thus

$$\mu(K) \leqslant L(f) \leqslant \mu(V) < \mu(K) + \varepsilon$$

which gives (3.7).

Consider now an open set V. Then for any $\alpha < \mu(V)$, there exists $f \in \mathcal{C}_c(V; [0,1])$ such that $\alpha < L(f)$. Hence for every open set W containing $K := \mathrm{supp}(f)$, then $f \in \mathcal{C}_c(W)$ and by definition of μ we have that $L(f) \leqslant \mu(W)$. Taking the infimum over all such W's leads to $L(f) \leqslant \mu(K)$. This shows the existence of a compact set $K \subset V$ with $\alpha < \mu(K)$ which ensures that V satisfies (3.5), and that $V \in \mathfrak{M}_F$ if futher $\mu(V) < \infty$.

Step 2. *For every sets $E_n \subset \Omega$ for $n \in \mathbb{N}^*$, then*

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

We first show that μ is finitely subadditive on open sets, *i.e.*,

$$\mu(V_1 \cup V_2) \leqslant \mu(V_1) + \mu(V_2) \tag{3.8}$$

when V_1 and V_2 are open sets. Let $g \in \mathcal{C}_c(V_1 \cup V_2; [0,1])$. By Lemma 3.5.2, there exists some functions f_1 and f_2 such that $f_i \in \mathcal{C}_c(V_i; [0,1])$ and $f_1 + f_2 = 1$ on $\mathrm{supp}(g)$. Hence $f_i g \in \mathcal{C}_c(V_i; [0,1])$, $g = f_1 g + f_2 g$ so that by linearity of L and by definition of μ,

$$L(g) = L(f_1 g) + L(f_2 g) \leqslant \mu(V_1) + \mu(V_2).$$

Taking the supremum with respect to all g as above gives $\mu(V_1 \cup V_2) \leqslant \mu(V_1) + \mu(V_2)$.

Now if $\mu(E_n) = \infty$ for some n then the result is obvious. Otherwise, if $\mu(E_n) < \infty$ for all n, then for any $\varepsilon > 0$ there exists open sets V_n such that $E_n \subset V_n$ and $\mu(V_n) < \mu(E_n) + 2^{-n}\varepsilon$. Define $V := \bigcup_{n=1}^\infty V_n$, and take $f \in \mathcal{C}_c(V; [0,1])$. In particular, since $\mathrm{supp}(f)$ is compact, one can find finitely many $V_1, \ldots, V_p$ such that $\mathrm{supp}(f) \subset \bigcup_{n=1}^p V_n$. Hence iterating (3.8),

$$L(f) \leqslant \mu\left(\bigcup_{n=1}^p V_n\right) \leqslant \sum_{n=1}^p \mu(V_n) \leqslant \sum_{n=1}^\infty \mu(E_n) + \varepsilon.$$

Since this holds for every $f \in \mathcal{C}_c(V; [0,1])$, and since $\bigcup_{n=1}^\infty E_n \subset V$, it follows that

$$\mu\left(\bigcup_{n=1}^\infty E_n\right) \leqslant \mu(V) \leqslant \sum_{n=1}^\infty \mu(E_n) + \varepsilon,$$

which completes the proof of step 2 since ε is arbitrary.

Step 3. *Suppose that $E = \bigcup_{n=1}^\infty E_n$, where E_n are pairwise disjoint elements of $\mathfrak{M}_F$ for all $n \in \mathbb{N}^*$. Then,*

$$\mu(E) = \sum_{n=1}^\infty \mu(E_n). \tag{3.9}$$

If in addition, $\mu(E) < \infty$, then $E \in \mathfrak{M}_F$.

We first show that μ is finitely additive on compact sets, *i.e.*,

$$\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2) \tag{3.10}$$

if K_1 and K_2 are disjoint compact sets. Let $\varepsilon > 0$, by Urysohn's Lemma, there exists $f \in \mathcal{C}_c(\Omega; [0,1])$ such that $f = 1$ on K_1 and $f = 0$ on K_2. Then by step 1, there exists a function $g \in \mathcal{C}_c(\Omega)$ such that $g = 1$ on $K_1 \cup K_2$ and $L(g) \leqslant \mu(K_1 \cup K_2) + \varepsilon$. Note that $fg = 1$ in K_1 and $(1-f)g = 1$ on K_2. Hence by linearity of L, it follows that

$$\mu(K_1) + \mu(K_2) \leqslant L(fg) + L((1-f)g) = L(g) \leqslant \mu(K_1 \cup K_2) + \varepsilon.$$

Since ε is arbitrary, (3.10) follows from step 2.

If $\mu(E) = \infty$, then (3.9) is an immediate consequence of step 2. Therefore, we can assume that $\mu(E) < \infty$. Let $\varepsilon > 0$, since $E_n \in \mathfrak{M}_F$, there exist compact sets $H_n \subset E_n$ with $\mu(H_n) > \mu(E_n) - 2^{-n}\varepsilon$ for each $n \in \mathbb{N}^*$. Define $K_n := \bigcup_{k=1}^n H_k \subset E$ which is compact. Thus using an induction on (3.10), we deduce that

$$\mu(E) \geqslant \mu(K_n) = \sum_{k=1}^n \mu(H_k) > \sum_{k=1}^n \mu(E_k) - \varepsilon.$$

Since the previous relation holds for every $n \in \mathbb{N}^*$ and every $\varepsilon > 0$, using again step 2, we deduce (3.9). Moreover for n large enough (depending on ε) we have $\mu(E) \leqslant \sum_{k=1}^n \mu(E_k) + \varepsilon \leqslant \mu(K_n) + 2\varepsilon$ which shows that $E \in \mathfrak{M}_F$.

Step 4. *If $E \in \mathfrak{M}_F$ and $\varepsilon > 0$, there is a compact set K and an open set V such that $K \subset E \subset V$ and $\mu(V \setminus K) < \varepsilon$. Moreover, if A and $B \in \mathfrak{M}_F$, then $A \setminus B$, $A \cup B$ and $A \cap B \in \mathfrak{M}_F$.*

Our definitions show that there exist a compact set K and an open set V with $K \subset E \subset V$ to that

$$\mu(V) - \frac{\varepsilon}{2} < \mu(E) < \mu(K) + \frac{\varepsilon}{2} < \infty,$$

where we used the fact that, from step 1, μ is finite on compact sets. Using again step 1, since $V \setminus K$ is open and $\mu(V \setminus K) < \infty$, it therefore belongs to $\mathfrak{M}_F$. We apply step 3 to get that $\mu(K) + \mu(V \setminus K) = \mu(V) < \mu(K) + \varepsilon < \infty$.

Now if A and $B \in \mathfrak{M}_F$ and $\varepsilon > 0$, one can find compact sets K_1 and K_2, and open sets V_1 and V_2 such that $K_1 \subset A \subset V_1$ and $K_2 \subset B \subset V_2$, and $\mu(V_i \setminus K_i) < \varepsilon$ (for $i = 1, 2$). Since

$$A \setminus B \subset V_1 \setminus K_2 \subset (V_1 \setminus K_1) \cup (K_1 \setminus V_2) \cup (V_2 \setminus K_2),$$

from the subadditivity proved in step 2, we infer that $\mu(A \setminus B) \leqslant 2\varepsilon + \mu(K_1 \setminus V_2)$. But since $K_1 \setminus V_2$ is a compact subset of $A \setminus B$, it shows that $A \setminus B \in \mathfrak{M}_F$. Then since $A \cup B = (A \setminus B) \cup B$, an application of step 3 shows that $A \cup B \in \mathfrak{M}_F$. Finally, since $A \cap B = A \setminus (A \setminus B)$, we also have $A \cap B \in \mathfrak{M}_F$.

Step 5. *$\mathfrak{M}$ is a σ-algebra which contains all Borel sets, and all sets $E \subset \Omega$ such that $\mu(E) = 0$. Morever $\mathfrak{M}_F = \{E \subset \mathfrak{M} : \mu(E) < \infty\}$. Finally, μ is a measure on $\mathfrak{M}$ satisfying* (3.4) *and* (3.5).

First step 1 gives that Ω is in $\mathfrak{M}$.

Now let K be an arbitrary compact set in X. If $A \in \mathfrak{M}$, then $A \cap K \in \mathfrak{M}_F$ and thus by step 4, $A^c \cap K = K \setminus (A \cap K) \in \mathfrak{M}_F$. Consequently, $A^c \in \mathfrak{M}$. Now if $A_n \in \mathfrak{M}$ for each $n \in \mathbb{N}^*$, define $B_1 = A_1 \cap K$ and for $n \geqslant 2$, $B_n = (A_n \cap K) \setminus \bigcup_{k=1}^{n-1} B_k$. Then by step 4, (B_n) is made of pairwise disjoint elements of $\mathfrak{M}_F$. Thus by step 3, we have

$$K \cap \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathfrak{M}_F$$

and thus $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$. This shows that $\mathfrak{M}$ is a σ-algebra. If C is closed, then $K \cap C$ is compact and by step 1, it belongs to $\mathfrak{M}_F$. Hence $C \in \mathfrak{M}$. This shows that $\mathfrak{M}$ is a σ-algebra containing all closed sets and consequently all Borel sets. If $E \subset \Omega$ is such that $\mu(E) = 0$, then clearly $E \in \mathfrak{M}$ since μ is an increasing set function.

By steps 1 and 4, we clearly have that $\mathfrak{M}_F \subset \{E \subset \mathfrak{M} : \mu(E) < \infty\}$. Conversely, let $E \in \mathfrak{M}$ such that $\mu(E) < \infty$ and let $\varepsilon > 0$. There exists an open set V containing E such that $\mu(V) < \infty$. Using again steps 1 and 4, one can find a compact set $K \subset V$ such that $\mu(V \setminus K) < \varepsilon$. Since $E \cap K \in \mathfrak{M}_F$, there exists a compact set $H \subset E \cap K$ with $\mu(E \cap K) < \mu(H) + \varepsilon$. Since $E \subset (E \cap K) \cup (V \setminus K)$, it follows that

$$\mu(E) \leqslant \mu(E \cap K) + \mu(V \setminus K) < \mu(H) + 2\varepsilon,$$

which implies that $E \in \mathfrak{M}_F$. Hence properties (3.4) and (3.5) follow.

We next prove that μ is a measure on $(\Omega, \mathfrak{M})$. If (E_n) is a sequence of pairwise disjoint elements of $\mathfrak{M}$, then $E := \bigcup_{n=1}^{\infty} E_n \in \mathfrak{M}$. If $\mu(E) = \infty$, then by Step 2, we have

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n),$$

while if $\mu(E) < \infty$ then $\mu(E_n) < \infty$ for each $n \in \mathbb{N}^*$ and thus E, $E_n \in \mathfrak{M}_F$ for all $n \in \mathbb{N}^*$. Consequently, by step 3, we have the same result which proves that μ is a measure on $\mathfrak{M}$.

Step 6. *Proof of the representation property.* Let $f \in \mathcal{C}_c(\Omega)$. Clearly it is enough to check the inequality $L(f) \leqslant \int_\Omega f\, d\mu$ because by linearity of L,

$$-L(f) = L(-f) \leqslant \int_\Omega (-f)\, d\mu = -\int_\Omega f\, d\mu.$$

Let $K := \operatorname{supp} f$ and $[a, b]$ be an interval which contains the range of f. For $\varepsilon > 0$, let $y_0, \ldots, y_n \in \mathbb{R}$ be such that $y_0 < a < y_1 < \ldots < y_n = b$, and $\max_{1 \leqslant i \leqslant n}(y_i - y_{i-1}) < \varepsilon$. Define

$$E_i := \{y_{i-1} < f \leqslant y_i\} \cap K.$$

Since f is continuous, f is Borel measurable and the sets E_i are therefore disjoint Borel sets whose union is K. There are open sets V_i containing E_i and such that $\mu(V_i) < \mu(E_i) + \varepsilon/n$ and $f(x) < y_i + \varepsilon$ for all $x \in V_i$ and all $i = 1, \ldots, n$. By Theorem 3.5.2, one can find functions $h_i \in \mathcal{C}_c(V_i; [0, 1])$ such that $\sum_{i=1}^{n} h_i = 1$ on K. Hence, $f = \sum_{i=1}^{n} h_i f$ in Ω and by step 1, we infer that

$$\mu(K) \leqslant L\left(\sum_{i=1}^{n} h_i\right) = \sum_{i=1}^{n} L(h_i).$$

Note that $L(h_i) \leqslant \mu(V_i) < \mu(E_i) + \varepsilon/n$, $h_i f \leqslant (y_i + \varepsilon)h_i$ and $y_i - \varepsilon < f(x)$ for $x \in E_i$. Denoting by $M := \max_K |f| < +\infty$ we have

$$\begin{aligned} L(f) &= \sum_{i=1}^n L(h_i f) \leqslant \sum_{i=1}^n (y_i+\varepsilon)L(h_i) = \sum_{i=1}^n (M+y_i+\varepsilon)L(h_i) - M\sum_{i=1}^n L(h_i) \\ &\leqslant \sum_{i=1}^n (M+y_i+\varepsilon)\left(\mu(E_i)+\frac{\varepsilon}{n}\right) - M\mu(K) \\ &\leqslant \sum_{i=1}^n (y_i-\varepsilon)\mu(E_i) + 2\varepsilon\mu(K) + \frac{\varepsilon}{n}\sum_{i=1}^n (M+y_i+\varepsilon) \\ &\leqslant \int_\Omega f\,d\mu + \varepsilon(2\mu(K)+2M+\varepsilon). \end{aligned}$$

Since ε is arbitrary, we proof of the theorem is complete. □

Remark 3.5.1. If we consider the restriction of μ to the Borel σ-algebra $\mathcal{B}(\Omega)$, then μ defines a Radon measure (a Borel measure which is finite on compact sets).

Remark 3.5.2. Step 5 of the proof of Theorem 3.5.1 shows that $\mathfrak{M}$ contains sets E such that $\mu(E) = 0$. We say that the measure space $(\Omega, \mathfrak{M}, \mu)$ is a complete measure space.

Remark 3.5.3. It can be proved that for any $E \in \mathcal{B}(\Omega)$ and any $\varepsilon > 0$, there exist a closed set C and an open set V such that $C \subset E \subset V$, and

$$\mu(V \setminus C) < \varepsilon. \tag{3.11}$$

Indeed, let $E \in \mathcal{B}(\Omega)$ and let (K_n) be an increasing sequence of compact sets whose union is Ω (take *e.g.* $K_n := \{x \in \Omega : |x| \leqslant n \text{ and } \mathrm{dist}(x, \mathbb{R}^N \setminus \Omega) \geqslant 1/n\}$). Since μ is a Radon measure, then $\mu(E \cap K_n) < \infty$ for all $n \geqslant 1$, and there exists an open set V_n such that $E \cap K_n \subset V_n$ and $\mu(V_n) < \lambda(E \cap K_n) + 2^{-n-1}\varepsilon$. In particular, since $\mu(E \cap K_n) < \infty$, then $\mu(V_n \setminus (E \cap K_n) < 2^{-n-1}\varepsilon$. Define the open set $V := \bigcup_{n\geqslant 1} V_n$ which satisfies $E \subset V$, and $V \setminus E \subset \bigcup_{n\geqslant 1} V_n \setminus (E \cap K_n)$. Then we have $\mu(V \setminus E) < \varepsilon/2$. Applying this property to $F := E^c$, there exists an open set W containing F and such that $\mu(W \setminus F) + \varepsilon/2$. Let $C := W^c$ be closed, then $C \subset E$ and $\mu(E \setminus C) = \mu(E \cap W) = \mu(W \setminus (E^c)) = \mu(W \setminus F) < \varepsilon/2$.

The Riesz representation Theorem has many consequences. For instance it enables to prove the existence of the Lebesgue measure and that the Lebesgue integral is a generalization of the Riemann integral. Another application will be given in chapter 6 about the characterization of the dual space of continuous functions vanishing on the boundary. We now give a direct application concerning the regularity of Radon measures which will be useful to prove the density of continuous functions and the separability of Lebesgue spaces in chapter 4.

Theorem 3.5.2. *Any positive Radon measure λ on Ω satisfies*

$$\lambda(E) = \inf\{\lambda(V) : E \subset V,\ V \text{ open}\} \text{ for every Borel set } E \subset \Omega,$$

and

$$\lambda(E) = \sup\{\lambda(K) : K \subset E,\ K \text{ compact}\} \text{ for every Borel set } E \subset \Omega \text{ with } \lambda(E) < \infty.$$

Moreover, for any $E \in \mathcal{B}(\Omega)$ and any $\varepsilon > 0$, there exist a closed set C and an open set V such that $C \subset E \subset V$, and

$$\lambda(V \setminus C) < \varepsilon. \tag{3.12}$$

Proof. Define $L(f) := \int_\Omega f\,d\lambda$ for any $f \in \mathcal{C}_c(\Omega)$. Since $\lambda(K) < \infty$ for any compact set $K \subset \Omega$, L is a positive linear form on $\mathcal{C}_c(\Omega)$, and from the Riesz representation Theorem, there exists a unique Radon measure such that

$$\int_\Omega f\,d\lambda = \int_\Omega f\,d\mu.$$

Since μ satisfies (3.4), (3.5) and (3.11) so that it suffices to prove that $\lambda = \mu$.

Let V be an open subset of Ω. Then $V = \bigcup_{n=1}^{\infty} K_n$ for some compact sets K_n with $n \in \mathbb{N}^*$. By Urysohn's Lemma, we can choose $f_n \in \mathcal{C}_c(V;[0,1])$ such that $f_n = 1$ on K_n. Define $g_n := \max_{1 \leqslant i \leqslant n} f_i$, then $g_n \in \mathcal{C}_c(V;[0,1])$ an it increases to χ_V pointwise in Ω. Hence by the monotone convergence Theorem,

$$\lambda(V) = \lim_{n\to\infty} \int_\Omega g_n \, d\lambda = \lim_{n\to\infty} \int_\Omega g_n \, d\mu = \mu(V).$$

Now let E be a Borel subset of Ω and let $\varepsilon > 0$. By (3.11), there exist an open set V and a closed set C such that $C \subset E \subset V$ and $\mu(V \setminus C) < \varepsilon$. But since $V \setminus C$ in open we infer that $\lambda(V \setminus C) < \varepsilon$. Consequently,

$$\lambda(E) \leqslant \lambda(V) = \mu(V) \leqslant \mu(E) + \varepsilon \leqslant \mu(V) + \varepsilon = \lambda(V) + \varepsilon \leqslant \lambda(E) + 2\varepsilon,$$

so that $\mu(E) = \lambda(E)$ by the arbitrariness of ε, and the proof of (3.12) is complete. □

3.6 Construction of the Lebesgue measure

The Riesz representation Theorem enables one to prove the existence of the Lebesgue measure (the usual volume measure in $\mathbb{R}^N$), and to show that the Lebesgue integral with respect to the Lebesgue measure is a natural extension of the Riemann integral.

Theorem 3.6.1. *There exist a σ-algebra $\mathcal{L}(\mathbb{R}^N)$ (containing the Borel σ-algebra $\mathcal{B}(\mathbb{R}^N)$) and a unique measure $\mathcal{L}^N$ on $\mathcal{L}(\mathbb{R}^N)$ such that*

1. $\mathcal{L}^N([0,1]^N) = 1$*;*
2. *For every $E \in \mathcal{L}(\mathbb{R}^N)$ and every $x \in \mathbb{R}^N$, $\mathcal{L}^N(x+E) = \mathcal{L}^N(E)$;*
3. *$E \in \mathcal{L}(\mathbb{R}^N)$ if and only if there exist a F_σ set A (a countable union of closed sets) and a G_δ set B (a countable intersection of open sets) such that $A \subset E \subset B$ and $\mathcal{L}^N(B \setminus A) = 0$;*
4. *For every $f \in \mathcal{C}_c(\mathbb{R}^N)$,*

$$\int_{\mathbb{R}^N} f \, d\mathcal{L}^N = \int_{\mathbb{R}^N} f(x)\, dx,$$

where the integral in left hand side is the Lebesgue integral of f with respect to the measure $\mathcal{L}^N$, and the integral in the right hand side is the Riemann integral of f.

The measure $\mathcal{L}^N$ is called the *Lebesgue measure*, and $\mathcal{L}(\mathbb{R}^N)$ is the σ-algebra of all *Lebesgue measurable sets.*

Proof. **Step 1.** Define $L : \mathcal{C}_c(\mathbb{R}^N) \to \mathbb{R}$ by

$$L(f) = \int_{\mathbb{R}^N} f(x)\, dx,$$

where the integral is the Riemann integral of f. Note that since f has compact support, then the previous integral is not improper since

$$\int_{\mathbb{R}^N} f(x)\, dx = \int_{\mathrm{Supp}(f)} f(x)\, dx.$$

Clearly L is a positive linear functional on $\mathcal{C}_c(\mathbb{R}^N)$, and according to the Riesz representation Theorem, there exist a σ-algebra $\mathcal{L}(\mathbb{R}^N)$ (containing the Borel σ-algebra $\mathcal{B}(\mathbb{R}^N)$) and a measure $\mathcal{L}^N$ on $\mathcal{L}(\mathbb{R}^N)$ such that $\mathcal{L}^N(K) < \infty$ for every compact set $K \subset \mathbb{R}^N$ and

$$L(f) = \int_{\mathbb{R}^N} f(x)\, dx = \int_{\mathbb{R}^N} f \, d\mathcal{L}^N$$

for every $f \in \mathcal{C}_c(\mathbb{R}^N)$.

Step 2. The conclusion of Theorem 3.5.1 also shows that for any $E \in \mathcal{L}(\mathbb{R}^N)$

$$\mathcal{L}^N(E) = \inf\{\mathcal{L}^N(V) : E \subset V,\ V \text{ open}\} \tag{3.13}$$

and any $E \in \mathcal{L}(\mathbb{R}^N)$ with $\mathcal{L}^N(E) < \infty$

$$\mathcal{L}^N(E) = \sup\{\mathcal{L}^N(K) : K \subset E,\, K \text{ compact}\}.$$

Moreover, from (3.11), for every $E \in \mathcal{L}(\mathbb{R}^N)$ and any $\varepsilon > 0$, there exist a closed set C and an open set V such that $C \subset E \subset V$ and $\mathcal{L}^N(V \setminus C) < \varepsilon$. Hence, for each $n \geqslant 1$, one can find a closed set C_n and an open set V_n such that $C_n \subset E \subset V_n$ and $\mathcal{L}^N(V_n \setminus C_n) < 1/n$. Define $A := \bigcup_{n\geqslant 1} C_n$ and $B := \bigcap_{n\geqslant 1} V_n$. Then A is a F_σ set, B is a G_δ set, $A \subset E \subset B$ and $\mathcal{L}^N(B \setminus A) \leqslant \mathcal{L}^N(V_n \setminus C_n) < 1/n \to 0$. Conversely, assume that $E \subset \mathbb{R}^N$ is such that $A \subset E \subset B$ and $\mathcal{L}^N(B \setminus A) = 0$ for some F_σ set A and some G_δ set B. Then $E = A \cup (E \setminus A)$ where $A \in \mathcal{B}(\mathbb{R}^N)$ and $E \setminus A \subset B \setminus A$ with $B \setminus A \in \mathcal{B}(\mathbb{R}^N)$ and $\mathcal{L}^N(B \setminus A) = 0$. Since the measure $\mathcal{L}^N$ is complete (see Remark 3.5.2) we deduce that $E \setminus A \in \mathcal{L}(\mathbb{R}^N)$ and finally that $E \in \mathcal{L}(\mathbb{R}^N)$.

Step 3. Let us show that for every open cube $Q := \prod_{i=1}^N (a_i, b_i)$ (for $a_i < b_i$ and $i \in \{1, \dots, N\}$), then

$$\mathcal{L}^N(Q) = \prod_{i=1}^N (b_i - a_i). \tag{3.14}$$

Define for each $n \geqslant 1$ and each $x \in \mathbb{R}^N$, $f_n(x) := \prod_{i=1}^N \varphi_n^{a_i,b_i}(x_i)$, where

$$\varphi_n^{a_i,b_i}(x_i) := \begin{cases} 0 & \text{if } \ x_i \notin [a_i, b_i], \\ 1 & \text{if } \ x_i \in \left[a_i + \frac{1}{n}, b_i - \frac{1}{n}\right], \\ n(x_i - a_i) & \text{if } \ x_i \in \left[a_i, a_i + \frac{1}{n}\right], \\ -n(x_i - b_i) & \text{if } \ x_i \in \left[b_i - \frac{1}{n}, b_i\right]. \end{cases}$$

Then $f_n \in \mathcal{C}_c(\mathbb{R}^N)$ for all $n \geqslant 1$ and $\chi_{\overline{Q_n}} \leqslant f_n \leqslant \chi_Q$ where $\overline{Q_n} := \prod_{i=1}^N [a_i + 1/n, b_i - 1/n]$ is a closed cube. Hence integrating the previous chain of inequalities with respect to the Lebesgue measure $\mathcal{L}^N$ leads to

$$\mathcal{L}^N(\overline{Q_n}) \leqslant \int_{\mathbb{R}^N} f_n \, d\mathcal{L}^N = \int_{\mathbb{R}^N} f_n(x)\, dx \leqslant \mathcal{L}^N(Q). \tag{3.15}$$

Since $\overline{Q_n}$ is an increasing sequence of (Lebesgue) measurable sets whose union is $Q = \prod_{i=1}^N (a_i, b_i)$, we deduce from Proposition 3.2.1 (ii) that $\mathcal{L}^N(\overline{Q_n}) \to \mathcal{L}^N(Q)$ as $n \to \infty$. Next by construction of f_n, its Riemann integral can be exactly computed as

$$\int_{\mathbb{R}^N} f_n(x)\, dx = \int_Q f_n(x)\, dx = \prod_{i=1}^N \int_{a_i}^{b_i} \varphi_n^{a_i,b_i}(x_i)\, dx_i = \prod_{i=1}^N \left(b_i - a_i - \frac{1}{n}\right).$$

Taking the limit in 3.15 as $n \to \infty$ yields formula 3.14.

Step 4. Let us show that for any $i \in \{1, \dots, N\}$ and any $a \in \mathbb{R}$,

$$\mathcal{L}^N(\{x_i = a\}) = 0. \tag{3.16}$$

Let us assume without loss of generality that $i = 1$ and $a = 0$. Then for each $k \geqslant 1$, we have $\{x_1 = 0\} \cap Q(0, n) \subset (-1/k, 1/k) \times \prod_{i=2}^N (-n, n)$ where $Q(0, n) = \prod_{i=1}^N (-n, n)$ is the open cube of center 0 and side length $2n$. Using (3.14), we infer that

$$\mathcal{L}^N(\{x_1 = 0\} \cap Q(0, n)) \leqslant \frac{(2n)^{N-1}}{2k}.$$

For fixed $n \geqslant 1$, letting $k \to \infty$ implies that $\mathcal{L}^N(\{x_1 = 0\} \cap Q(0, n)) = 0$. Then taking the limit as $n \to \infty$ and using Proposition 3.2.1 (ii) yields $\mathcal{L}^N(\{x_1 = 0\}) = 0$.

Step 5. Since

$$\prod_{i=1}^N [a_i, b_i] \setminus \prod_{i=1}^N (a_i, b_i)$$

is made of finitely many subsets of hyperplane of the form $\{x_i = a\}$, we deduce from Step 4 that

$$\mathcal{L}^N\left(\prod_{i=1}^N [a_i, b_i] \setminus \prod_{i=1}^N (a_i, b_i)\right) = 0$$

so that

$$\mathcal{L}^N\left(\prod_{i=1}^N [a_i,b_i]\right) = \mathcal{L}^N\left(\prod_{i=1}^N (a_i,b_i)\right) = \prod_{i=1}^N (b_i - a_i).$$

In particular, taking $a_i = 0$ and $b_i = 1$ for all $i \in \{1,\dots,N\}$ leads to $\mathcal{L}^N([0,1]^N) = 1$.

Step 6. Let us show that $\mathcal{L}^N$ is translation invariant. Let $x \in \mathbb{R}^N$ and $V \subset \mathbb{R}^N$ an open set. Since the translation $\tau_x : y \in \mathbb{R}^N \mapsto x + y$ is a homeomorphism (with inverse $(\tau_x)^{-1} = \tau_{-x}$), then $x + V$ is open, and by definition of $\mathcal{L}^N$ on open sets (see (3.6)) we have

$$\mathcal{L}^N(x+V) = \sup\left\{\int_{\mathbb{R}^N} f(y)\,dy : f \in \mathcal{C}_c(\mathbb{R}^N;[0,1]),\ \mathrm{Supp}(f) \subset x+V\right\}.$$

By changing variables in the Riemann integral we obtain that

$$\begin{aligned}
\mathcal{L}^N(x+V) &= \sup\left\{\int_{\mathbb{R}^N} g(x+y)\,dy : g \in \mathcal{C}_c(\mathbb{R}^N;[0,1]),\ \mathrm{Supp}(g) \subset V\right\}\\
&= \sup\left\{\int_{\mathbb{R}^N} g(z)\,dz : g \in \mathcal{C}_c(\mathbb{R}^N;[0,1]),\ \mathrm{Supp}(g) \subset V\right\}\\
&= \mathcal{L}^N(V).
\end{aligned}$$

Next from the above mentioned properties of the translation, if $E \in \mathcal{B}(\mathbb{R}^N)$, then $x + E \in \mathcal{B}(\mathbb{R}^N)$. Applying the outer regularity (3.13) of the Lebesgue measure, we get that

$$\begin{aligned}
\mathcal{L}^N(x+E) &= \inf\{\mathcal{L}^N(V) : x+E \subset V,\ V \text{ open}\}\\
&= \inf\{\mathcal{L}^N(-x+V) : E \subset -x+V,\ V \text{ open}\}\\
&= \inf\{\mathcal{L}^N(U) : E \subset U,\ U \text{ open}\}\\
&= \mathcal{L}^N(E).
\end{aligned}$$

Finally, from step 2, we know that if $E \in \mathcal{L}(\mathbb{R}^N)$ then there exist a F_σ set A, and a G_δ set B such that $A \subset E \subset B$ and $\mathcal{L}^N(B \setminus A) = 0$. Then $x + A \subset x + E \subset x + B$, where $x + A$ is a F_σ and $x + B$ is a G_δ. Moreover, since $x + B \setminus A = (x + B) \setminus (x + A)$ and $B \setminus A \in \mathcal{B}(\mathbb{R}^N)$, then $\mathcal{L}^N((x+B) \setminus (x+A)) = \mathcal{L}^N(x + (B \setminus A)) = \mathcal{L}^N(B \setminus A) = 0$. Consequently, $x + E \in \mathcal{L}(\mathbb{R}^N)$ and

$$\mathcal{L}^N(x+E) = \mathcal{L}^N(x+A) = \mathcal{L}^N(A) = \mathcal{L}^N(E).$$

Step 7. It remains to show the uniqueness of the Lebesgue measure. Let λ be a translation invariant Radon measure such that $\lambda([0,1]) = 1$. We claim that $\lambda = \mathcal{L}^N$.

Let us first show that for each $a \in \mathbb{R}$ and $i \in \{1,\dots,N\}$, then $\lambda(\{x_i = a\}) = 0$. Without loss of generality, assume that $i = 1$ and $a = 0$. Then

$$\lambda(\{x_1 = 0\}) = \lambda\left(\{x_1 = 0\} \cap \bigcup_{n \geqslant 1} [0,n]^N\right) = \lim_{n\to\infty} \lambda\left(\{x_1 = 0\} \cap [0,n]^N\right). \tag{3.17}$$

Define $E_n := \{x_1 = 0\} \cap [0,n]^N$ and note that

$$[0,n]^N = \bigcup_{y_1 \in [-n,n]} (y_1 + E_n) \supset \bigcup_{y_1 \in [-n,n]\cap\mathbb{Q}} (y_1 + E_n),$$

where the (Lebesgue measurable) sets $\{y_1 + E_n\}_{y_1 \in [-n,n]\cap\mathbb{Q}}$ are pairwise disjoint. Hence since λ is a Radon measure, it is finite on compact sets, and the translation invariance yields

$$\sum_{y_1 \in [-n,n]\cap\mathbb{Q}} \lambda(E_n) = \sum_{y_1 \in [-n,n]\cap\mathbb{Q}} \lambda(y_1 + E_n) = \lambda\left(\bigcup_{y_1 \in [-n,n]\cap\mathbb{Q}} (y_1 + E_n)\right) \leqslant \lambda([0,n]^N) < \infty$$

which is possible only if $\lambda(E_n) = 0$. Consequently, from (3.17), we obtain that $\lambda(\{x_1 = 0\}) = 0$.

As a consequence, if $n \in \mathbb{N}^*$, since

$$[0,1)^N = \bigcup_{k\in\{0,\dots,n-1\}^N} \left(\frac{k}{n} + \left[0, \frac{1}{n}\right)^N\right),$$

where the (Lebesgue measurable) sets in the previous union are pairwise disjoint, we deduce that

$$\begin{aligned}
1 = \lambda([0,1]^N) = \lambda([0,1)^N) &= \lambda\left(\bigcup_{k\in\{0,\dots,n-1\}^N} \left(\frac{k}{n} + \left[0, \frac{1}{n}\right)^N\right)\right) \\
&= \sum_{k\in\{0,\dots,n-1\}^N} \lambda\left(\frac{k}{n} + \left[0, \frac{1}{n}\right)^N\right) = \sum_{k\in\{0,\dots,n-1\}^N} \lambda\left(\left[0, \frac{1}{n}\right)^N\right) = n^N \mathcal{L}^N\left(\left[0, \frac{1}{n}\right)^N\right).
\end{aligned}$$

Hence $\lambda([0,1/n)^N) = n^{-N} = \mathcal{L}^N([0,1/n)^N)$.

Next we show that λ coincides with $\mathcal{L}^N$ on cubes. Let $Q := \prod_{i=1}^N [a_i, b_i]$, and assume first that a_i and $b_i \in \mathbb{Q}$ with $a_i < b_i$ for $i \in \{1,\dots,N\}$. Then there exist integers $n \in \mathbb{N}$, α_i and $\beta_i \in \mathbb{Z}$ such that $a_i = \alpha_i/n$ and $b_i = \beta_i/n$. Hence

$$Q = \left(\frac{\alpha_i}{n}, \dots, \frac{\alpha_N}{n}\right) + \prod_{i=1}^N \left[0, \frac{q_i}{n}\right],$$

where $q_i = \beta_i - \alpha_i \in \mathbb{N}$. Thanks to the translation invariance of λ, we deduce that

$$\lambda(Q) = \lambda\left(\prod_{i=1}^N \left[0, \frac{q_i}{n}\right)\right).$$

On the other hand,

$$\lambda\left(\prod_{i=1}^N \left[0, \frac{q_i}{n}\right)\right) = \lambda\left(\prod_{i=1}^N \bigcup_{k_i\in\mathbb{N}, |k_i|\leqslant q_i-1} \left[\frac{k_i}{n}, \frac{k_i+1}{n}\right)\right) = \lambda\left(\bigcup_{k\in\mathbb{K}} \left(\frac{k}{n} + \left[0, \frac{1}{n}\right)\right)\right),$$

where $\mathbb{K} := \{k \in \mathbb{N}^N : 0 \leqslant k_i \leqslant q_i - 1 \text{ for all } i \in \{1,\dots,N\}\}$. Thus using again the translation invariance of λ yields

$$\begin{aligned}
\lambda\left(\prod_{i=1}^N \left[0, \frac{q_i}{n}\right)\right) = \sum_{k\in\mathbb{K}} \lambda\left(\left(\frac{k}{n} + \left[0, \frac{1}{n}\right)\right)\right) &= \sum_{k\in\mathbb{K}} \lambda\left(\left[0, \frac{1}{n}\right)\right) \\
&= \frac{1}{n^N}\mathrm{Card}(\mathbb{K}) = \frac{1}{n^N}\prod_{i=1}^N q_i = \prod_{i=1}^N (b_i - a_i).
\end{aligned}$$

Finally, we obtain that $\lambda(Q) = \mathcal{L}^N(Q)$.

If now a_i and $b_i \in \mathbb{R}$, then there exist sequences $(a_i^n)_{n\geqslant 1}$ and $(b_i^n)_{n\geqslant 1} \subset \mathbb{Q}$ such that $a_i^n \searrow a_i$ and $b_i^n \nearrow b_i$ as $n \to \infty$, for each $i \in \{1,\dots,N\}$. Since $\prod_{i=1}^N [a_i^n, b_i^n]$ is an increasing sequence of closed cubes whose union is the open cube $\prod_{i=1}^N (a_i, b_i)$, we deduce that

$$\begin{aligned}
\lambda\left(\prod_{i=1}^N [a_i, b_i]\right) = \lambda\left(\prod_{i=1}^N (a_i, b_i)\right) &= \lim_{n\to\infty} \lambda\left(\prod_{i=1}^N [a_i^n, b_i^n]\right) \\
&= \lim_{n\to\infty} \prod_{i=1}^N (b_i^n - a_i^n) = \prod_{i=1}^N (b_i - a_i) = \mathcal{L}^N\left(\prod_{i=1}^N [a_i, b_i]\right).
\end{aligned}$$

To show that $\lambda(V) = \mathcal{L}^N(V)$ for every open set, we use the fact that any open set in $\mathbb{R}^N$ can be written as the countable union of disjoint cubes of the form $\prod_{i=1}^N [b_i - a_i)$ where a_i and $b_i \in \mathbb{R}$ for $i \in \{1,\dots,N\}$. Finally since λ is a Radon measure, we use the outer regularity of both λ and $\mathcal{L}^N$ (see Theorem 3.5.2) to show that $\lambda(E) = \mathcal{L}^N(E)$ for every Borel set E. $\square$

Chapter 4

Lebesgue spaces

An important class of examples of Banach spaces is given by the Lebesgue spaces L^p of all measurable functions whose absolute value raised to the p-th power has finite integral. In this chapter, and unless otherwise mentioned, μ is a measure over a measurable space $(\Omega, \mathfrak{M})$.

4.1 First definitions and properties

Definition 4.1.1. *Let $1 \leqslant p < \infty$. We define*

$$\mathcal{L}^p(\Omega,\mu) := \left\{ u : \Omega \to \mathbb{R} \text{ measurable } : \|u\|_{\mathcal{L}^p(\Omega,\mu)} := \left(\int_\Omega |u|^p d\mu \right)^{1/p} < +\infty \right\},$$

and

$$\mathcal{L}^\infty(\Omega,\mu) := \left\{ u : \Omega \to \mathbb{R} \text{ measurable } : \|u\|_{\mathcal{L}^\infty(\Omega,\mu)} := \operatorname*{ess\,sup}_{x\in\Omega} |u(x)| < +\infty \right\},$$

where

$$\operatorname*{ess\,sup}_{x\in\Omega} |u(x)| := \inf\{C \in \mathbb{R} : |u(x)| \leqslant C \ \mu\text{-a.e. in } \Omega\}.$$

The ℓ^p spaces ($1 \leqslant p < \infty$) are a special case of $\mathcal{L}^p$ spaces, when Ω is the set $\mathbb{N}$ of nonnegative integers, and the measure μ is the counting measure on $\mathbb{N}$.

In general the map $\mathcal{L}^p(\Omega,\mu) \ni u \mapsto \|u\|_{\mathcal{L}^p(\Omega,\mu)}$ does not define a norm over $\mathcal{L}^p(\Omega,\mu)$ (it is actually a semi norm). However we have the following result.

Proposition 4.1.1. *Let $f : \Omega \to [0,\infty]$ be a measurable function such that*

$$\int_\Omega f \, d\mu = 0.$$

Then $f(x) = 0$ for μ-a.e. $x \in \Omega$.

Proof. Define the measurable sets $E_n := \{f \geqslant 1/n\}$. Note that (E_n) is an increasing sequence of measurable sets, and $\{f > 0\} = \bigcup_{n=1}^\infty E_n$. Hence by Proposition 3.3.1,

$$\frac{1}{n}\mu(E_n) \leqslant \int_{E_n} f \, d\mu \leqslant \int_\Omega f \, d\mu = 0,$$

and thus $\mu(E_n) = 0$ for every $n \in \mathbb{N}$. Then by Proposition 3.2.1

$$\mu(\{f > 0\}) = \mu\left(\bigcup_{n=1}^\infty E_n \right) = \lim_{n\to\infty} \mu(E_n) = 0,$$

which proves that $f = 0$ μ-a.e. in Ω. $\square$

Indeed from the previous proposition, if $u \in \mathcal{L}^p(\Omega,\mu)$ is such that $\|u\|_{\mathcal{L}^p(\Omega,\mu)} = 0$ then $u(x) = 0$ for μ-a.e. $x \in \Omega$.

Given two measurable functions $u,\ v : \Omega \to [-\infty,+\infty]$, we say that u is equivalent to v, and we write $u \sim v$, if $u(x) = v(x)$ for μ-a.e. $x \in \Omega$. Note that $\sim$ is an equivalence relation in the class of measurable functions. The spaces $\mathcal{L}^p(\Omega,\mu)$ can be made into a normed vector space, denoted $L^p(\Omega,\mu)$, by taking their quotient space with respect to this equivalence relation. If $u \in \mathcal{L}^p(\Omega;\mu)$, we denote by $[u]$ its equivalence class. With an abuse of notations, from now on we identify a measurable function u to its equivalence class $[u]$.

Definition 4.1.2. *Let $u \in L^p(\Omega,\mu)$, we denote*

$$
\begin{aligned}
\|u\|_p &:= \left(\int_\Omega |u|^p d\mu\right)^{1/p} \quad \text{if } 1 \leqslant p < \infty,\\
\|u\|_\infty &:= \operatorname*{ess\,sup}_{x\in\Omega} |u(x)| \quad \text{if } p = \infty.
\end{aligned}
$$

Thanks to this identification, the Lebesgue spaces $L^p(\Omega,\mu)$ are normed spaces. Before proving this property we state two important inequalities which will be instrumental in the sequel.

Proposition 4.1.2 (Hölder's inequality). *Let $1 \leqslant p \leqslant \infty$ and $p' \geqslant 1$ its conjugate exponent defined by $1/p + 1/p' = 1$ (by convention $p' = 1$ if $p = \infty$, and $p' = \infty$ if $p = 1$). If $u \in L^p(\Omega,\mu)$ and $v \in L^{p'}(\Omega,\mu)$ then $uv \in L^1(\Omega,\mu)$ and*

$$\|uv\|_1 \leqslant \|u\|_p \|v\|_{p'}.$$

Proof. Since the logarithm is concave on $\mathbb{R}_+^*$, for $a, b > 0$ and $1 \leqslant p, p' < \infty$ with $1/p + 1/p' = 1$, we have

$$\log(ab) = \log(a) + \log(b) = \frac{1}{p}\log(a^p) + \frac{1}{p'}\log(a^{p'}) \leqslant \log\left(\frac{1}{p}a^p + \frac{1}{p'}b^{p'}\right).$$

Taking the exponential we obtain Young's inequality:

$$ab \leqslant \frac{1}{p}a^p + \frac{1}{p'}b^{p'}.$$

Using this with $a = \lambda u$ and $b = \lambda^{-1} v$ and integrating over Ω yields

$$\int_\Omega |uv|\, d\mu \leqslant \frac{1}{p}\lambda^p \|u\|_p^p + \frac{1}{p'}\lambda^{-p'}\|v\|_{p'}^{p'}.$$

The last term is minimized by choosing $\lambda = \|v\|_{p'}^{1/p} \|u\|_p^{(1-p)/p}$.

If either $p = \infty$ or $p' = \infty$, the inequality is immediate. ☐

Remark 4.1.1. From Hölder's inequality, it follows that if $\mu(\Omega) < \infty$ and $1 \leqslant q \leqslant p \leqslant \infty$ then $L^\infty(\Omega,\mu) \subset L^q(\Omega,\mu) \subset L^p(\Omega,\mu) \subset L^1(\Omega,\mu)$. However, if $\mu(\Omega) = \infty$, nothing can be said in general.

Proposition 4.1.3 (Minkowski inequality). *For $1 \leqslant p \leqslant \infty$, and for any $u,\ v \in L^p(\Omega,\mu)$, we have*

$$\|u + v\|_p \leqslant \|u\|_p + \|v\|_p.$$

Proof. We start with $p = \infty$. By definition of the essential supremum, $|u(x)| \leqslant \|u\|_\infty$ and $|v(x)| \leqslant \|v\|_\infty$ for μ-a.e. $x \in \Omega$, hence

$$|u(x) + v(x)| \leqslant |u(x)| + |v(x)| \leqslant \|u\|_\infty + \|v\|_\infty,$$

for μ-a.e. in Ω. Using again the definition of the essential supremum of $u + v$, we get $\|u + v\|_\infty \leqslant \|u\|_\infty + \|v\|_\infty$.

Similarly, for $p = 1$, we have

$$\|u + v\|_1 = \int_\Omega |u + v|\, d\mu \leqslant \int_\Omega (|u| + |v|)\, d\mu \leqslant \|u\|_1 + \|v\|_1.$$

For $1 < p < \infty$, by Hölder's inequality, we infer that

$$\begin{aligned}
\|u+v\|_p^p &= \int_\Omega |u+v|^p \, d\mu = \int_\Omega |u+v||u+v|^{p-1} \, d\mu \\
&\leqslant \int_\Omega |u||u+v|^{p-1} \, d\mu + \int_\Omega |v||u+v|^{p-1} \, d\mu \\
&\leqslant (\|u\|_p + \|v\|_p) \left(\int_\Omega (|u+v|^{p-1})^{p/(p-1)} \right)^{(p-1)/p} \\
&= (\|u\|_p + \|v\|_p) \|u+v\|_p^{p-1}.
\end{aligned}$$

□

We are now in position to state that the Lebesgue spaces are normed spaces:

Proposition 4.1.4. *For $1 \leqslant p \leqslant \infty$, the map $L^p(\Omega, \mu) \ni u \mapsto \|u\|_p$ defines a norm over $L^p(\Omega, \mu)$.*

Proof. From Proposition 4.1.1, if $\|u\|_p = 0$, then $u = 0$ μ-a.e. in Ω, and thus $u = 0$ in $L^p(\Omega, \mu)$. Clearly, $\|\lambda u\|_p = |\lambda| \|u\|_p$ for every $\lambda \in \mathbb{R}$ and $u \in L^p(\Omega, \mu)$. Finally the triangle inequality is a consequence of the Minkowski inequality. □

4.2 Completeness

We saw in Proposition 4.1.4 that the Lebesgue spaces $L^p(\Omega, \mu)$ are normed spaces. We will prove now that they are actually Banach spaces as a consequence of the following completeness result.

Theorem 4.2.1 (Fréchet – Riesz). *For $1 \leqslant p \leqslant \infty$, the space $L^p(\Omega, \mu)$ is complete.*

Proof. It is enough to prove that any normally converging series in $L^p(\Omega, \mu)$ is converging in $L^p(\Omega, \mu)$ (see Proposition 1.2.1). Let $(u_n)_n \subset L^p(\Omega; \mu)$ be such that $\sum_{n=1}^\infty \|u_n\|_p < \infty$. Then the sequence of partial sums

$$v_n := \sum_{k=1}^n |u_k|$$

is increasing, and from the Minkowski inequality,

$$\int_\Omega v_n^p \, d\mu = \|v_n\|_p^p \leqslant \left(\sum_{k=1}^\infty \|u_k\|_p \right)^p < \infty$$

if $p < \infty$, while

$$v_n(x) \leqslant \sum_{k=1}^\infty \|u_k\|_\infty$$

for μ-a.e. $x \in \Omega$ if $p = \infty$. We deduce from the monotone convergence Theorem that (v_n) converges μ-a.e. in Ω to the function $\sum_{k=1}^\infty |u_k|$ which belongs to $L^p(\Omega, \mu)$. In particular, by Proposition 3.3.2, we deduce that the series $\sum_{k=1}^\infty |u_k(x)|$ converges for μ-a.e. $x \in \Omega$, and from the completeness of $\mathbb{R}$ that the series $\sum_{k=1}^\infty u_k(x)$ converges as well. Therefore $\sum_{k=1}^\infty u_k \in L^p(\Omega, \mu)$ and for $n \geqslant m$, we have

$$\left\| \sum_{k=1}^n u_k - \sum_{k=1}^m u_k \right\|_p = \left\| \sum_{k=m+1}^n u_k \right\|_p \leqslant \sum_{k=m+1}^n \|u_k\|_p \,,$$

where we used again the Minkowski inequality. Letting $n \to \infty$, we infer that

$$\left\| \sum_{k=1}^\infty u_k - \sum_{k=1}^m u_k \right\|_p \leqslant \sum_{k=m+1}^\infty \|u_k\|_p \,,$$

and since $\sum_{n=1}^\infty \|u_n\|_p < \infty$, then $\sum_{k=m+1}^\infty \|u_k\|_p \to 0$ as $m \to \infty$, and the conclusion follows. □

It follows from Theorem 3.4.3 that if (f_n) is a sequence bounded both in $\mathcal{L}^1(X, \mu)$ and in $\mathcal{L}^p(\Omega, \mu)$ for some $1 < p < \infty$ then the sequence (f_n) is equi-integrable.

4.3 Density and separability

Theorem 4.3.1. *For any $1 \leqslant p \leqslant \infty$ the space of all measurable simple functions are dense in $L^p(\Omega, \mu)$.*

Proof. Let $u \in L^p(\Omega, \mu)$. decomposing $u = u_+ - u_-$, we can assume without loss of generality that $u \geqslant 0$. Let (s_n) be the sequence of simple functions constructed in Theorem 3.1.1.

If $p = \infty$, then for any $n \geqslant \|u\|_\infty$, we have $|s_n(x) - u(x)| \leqslant 2^{-n}$ for μ-a.e. $x \in \Omega$, and thus $\|s_n - u\|_\infty \leqslant 2^{-n} \to 0$.

If $1 \leqslant p < \infty$, since $s_n(x) \nearrow u(x)$ for μ-a.e. $x \in \Omega$, we get that $|u(x) - s_n(x)|^p \to 0$ for μ-a.e. $x \in \Omega$ and $|u(x) - s_n(x)|^p \leqslant 2^p u(x)^p$ for μ-a.e. $x \in \Omega$. Hence by the dominated convergence Theorem $\|u - s_n\|_p \to 0$. □

From now on, we assume that Ω is an open subset of $\mathbb{R}^N$, with $N \geqslant 1$, and that μ is a positive Radon measure.

Theorem 4.3.2. *For any $1 \leqslant p < \infty$ the space $\mathcal{C}_c(\Omega)$ is dense in $L^p(\Omega, \mu)$.*

Proof. Let $K_n := \{x \in \Omega : \mathrm{dist}(x, \Omega^c) \geqslant 1/n \text{ and } \|x\|_{\mathbb{R}^n} \leqslant n\}$. Then (K_n) is an increasing sequence of compact sets such that $\cup_n K_n = \Omega$. Let $\varepsilon > 0$. By the dominated convergence Theorem, for n large enough we have that $\|u - \chi_{K_n} u\|_{L^p(\Omega,\mu)} \leqslant \varepsilon$. Since μ is a positive Radon measure it is sufficient to consider the case where $\mu(\Omega)$ is finite. By Theorem 4.3.1 if is enough to show that any simple function can be approximated in $L^p(\Omega, \mu)$ by a continuous function with compact support in Ω. Then by linearity it suffices to consider characteristic functions. Let $E \subset \Omega$ be a Borel set, by Theorem 3.5.2, there exist a compact set K and an open set V such that $K \subset E \subset V$ and $\mu(V \setminus K) < \varepsilon$. Then from Urysohn's Lemma there exists a function $v \in \mathcal{C}_c(V; [0, 1])$ such that $v = 1$ on K. Hence

$$\int_\Omega |v - \chi_E|^p \, d\mu \leqslant \mu(V \setminus K) < \varepsilon,$$

which completes the proof of the Theorem. □

Corollary 4.3.1. *For any $1 \leqslant p < \infty$ the space $L^p(\Omega, \mu)$ is separable.*

Proof. With the notations of the previous proof, we have that $\mathcal{C}_c(\Omega) \subset \bigcup_{n=1}^\infty \mathcal{C}(K_n)$. We already proved in Corollary 2.4.4 that $\mathcal{C}(K_n)$ are separable with respect to the uniform convergence. But since μ is a Radon measure, it is finite on compact sets and thus uniform convergence in K_n implies convergence in $L^p(K_n, \mu)$ because

$$\int_{K_n} |u - v|^p \, d\mu \leqslant \mu(K_n) \sup_{x \in \Omega} |u(x) - v(x)|^p.$$

As a consequence, $\mathcal{C}(K_n)$ is separable in the topology of $L^p(\Omega, \mu)$. Finally, since a countable union of separable spaces is separable, we infer that $\mathcal{C}_c(\Omega)$ is separable in the topology of $L^p(\Omega, \mu)$. The conclusion follows from Theorem 4.3.2 □

The space $L^\infty(\Omega, \mu)$ is more specific to the Radon measure μ. Indeed, if μ is a Dirac mass, then the space $L^\infty(\Omega, \mu)$ can be identified to $\mathbb{R}$ which is therefore separable. However, this property fails in general.

Theorem 4.3.3. *If $\mu = \mathcal{L}^N$ is the Lebesgue measure, then $L^\infty(\Omega, \mathcal{L}^N)$ is not separable.*

Proof. If $x \in \Omega$ and $R > 0$ are such that $B(x, R) \subset \Omega$, then the family $\{\chi_{B(x,r)} : 0 < r < R\}$ is uncountable and if $0 < r \neq r' < R$, then $\|\chi_{B(x,r)} - \chi_{B(x,r')}\|_\infty = 1$. The conclusion follows from the non separability criterion (Proposition 1.4.2). □

4.4 Convolution

From now on we will exclusively work with the Lebesgue measure $\mu = \mathcal{L}^N$. The corresponding Lebesgue spaces $L^p(\Omega, \mathcal{L}^N)$ will be simply denoted by $L^p(\Omega)$.

4.4.1 Definition and Young's inequality

Definition 4.4.1. *Let u and $v \in \mathcal{C}_c(\mathbb{R}^N)$, the convolution product $u * v$ is defined by*

$$(u * v)(x) := \int_{\mathbb{R}^N} u(x-y)v(y)\, dy$$

for all $x \in \mathbb{R}^N$.

Next Lemma asserts that $u * v$ is $\mathcal{C}_c(\mathbb{R}^N)$. It In fact if either u or v is more regular, one can expect more regularity on the convolution product. We recall that for any open set $\Omega \subset \mathbb{R}^N$,

$$\mathcal{D}(\Omega) := \mathcal{C}_c^\infty(\Omega) := \{u \in \mathcal{C}^\infty(\Omega) : \operatorname{supp}(u) \text{ is a compact subset of } \Omega\}.$$

Lemma 4.4.1. *The following properties hold true.*

1. *If $u, v \in \mathcal{C}_c(\Omega)$ then $u * v \in \mathcal{C}_c(\Omega)$ with $\operatorname{supp}(u * v) \subset \overline{\operatorname{supp}(u) + \operatorname{supp}(v)}$.*
2. *If $u \in \mathcal{C}_c(\Omega)$ and $v \in \mathcal{C}_c^\infty(\Omega)$, then $u * v \in \mathcal{C}_c^\infty(\Omega)$ and for any multi-index $\alpha \in \mathbb{N}^N$, $\partial^\alpha(u * v) = u * (\partial^\alpha v)$.*
3. *For every $1 \leqslant p \leqslant \infty$,*
$$\|u * v\|_p \leqslant \|u\|_p \|v\|_1. \tag{4.1}$$

Proof. For the first part, it follows from the definition and the dominated convergence Theorem that the function $x \mapsto (u * v)(x)$ is continuous. Moreover $\operatorname{supp}(u * v) \subset \overline{\operatorname{supp}(u) + \operatorname{supp}(v)}$ so that $u * v \in \mathcal{C}_c(\mathbb{R}^N)$.

For the second part it suffices to consider the case $|\alpha| = 1$, the general case following by an induction argument. By changing variables, we infer that

$$\frac{(u * v)(x + h\alpha) - (u * v)(x)}{h} = \int_{\mathbb{R}^N} u(y) \frac{v(x + h\alpha - y) - v(x - y)}{h}\, dy$$

and the conclusion follows from the dominated convergence Theorem.

For the third part, we write thanks to Fubini's Theorem

$$\begin{aligned}
\int_{\mathbb{R}^N} |(u * v)(x)|\, dx &= \int_{\mathbb{R}^N} \left| \int_{\mathbb{R}^N} u(x-y)v(y)\, dy \right| dx \\
&\leqslant \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |u(x-y)||v(y)|\, dy\, dx \\
&= \int_{\mathbb{R}^N} \left(\int_{\mathbb{R}^N} |u(x-y)|\, dx \right) |v(y)|\, dy \\
&= \|u\|_1 \|v\|_1,
\end{aligned}$$

which complete the case $p = 1$. If now $1 < p < \infty$, thanks to Hölder's inequality and Fubini's Theorem, we get

$$\begin{aligned}
\int_{\mathbb{R}^N} |(u * v)(x)|^p\, dx &= \int_{\mathbb{R}^N} \left| \int_{\mathbb{R}^N} u(x-y)v(y)\, dy \right|^p dx \\
&\leqslant \int_{\mathbb{R}^N} \left| \left(\int_{\mathbb{R}^N} |u(x-y)|^p |v(y)|\, dy \right)^{1/p} \left(\int_{\mathbb{R}^N} |v(y)|\, dy \right)^{1-1/p} \right|^p dx \\
&= \|v\|_1^{p-1} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |u(x-y)|^p |v(y)|\, dy\, dx \\
&= \|u\|_p^p \|v\|_1^p.
\end{aligned}$$

Finally, if $p = \infty$, we simply write

$$|(u * v)(x)| \leqslant \int_{\mathbb{R}^N} |u(x-y)||v(y)|\, dy \leqslant \|u\|_\infty \|v\|_1$$

for all $x \in \mathbb{R}^N$. □

Corollary 4.4.1. *The convolution product $(u,v) \mapsto u * v$ from $\mathcal{C}_c(\mathbb{R}^N) \times \mathcal{C}_c(\mathbb{R}^N)$ to $\mathcal{C}_c(\mathbb{R}^N)$ extends uniquely into a continuous (multilinear) application from $L^p(\mathbb{R}^N) \times L^1(\mathbb{R}^N)$ to $L^p(\mathbb{R}^N)$ (for all $1 \leqslant p \leqslant \infty$). Moreover, if $u \in L^p(\mathbb{R}^N)$ is such that $u = 0$ a.e. in $\mathbb{R}^N \setminus K$, for some compact set $K \subset \mathbb{R}^N$, and if $v \in \mathcal{C}_c^\infty(\mathbb{R}^N)$, then $u * v \in \mathcal{C}_c^\infty(\mathbb{R}^N)$ and $\operatorname{supp}(u * v) \subset K + \operatorname{supp}(v)$.*

Proof. For $1 \leqslant p < \infty$, the existence and uniqueness of the continuous extension follows from the general extension result proved in Theorem 1.2.3 and the density of $\mathcal{C}_c(\mathbb{R}^N)$ into $L^p(\mathbb{R}^n)$ (see Theorem 4.3.2). The uniform continuous character (on bounded sets of $L^1(\mathbb{R}^N) \times L^p(\mathbb{R}^N)$) is ensured by the estimate (4.1) proved in Lemma 4.4.1.

For $p = \infty$, it suffices to observe that for $u \in L^\infty(\mathbb{R}^N)$ and $v \in L^1(\mathbb{R}^N)$, for every $x \in \mathbb{R}^N$ the function $y \mapsto u(x-y)v(y)$ (defined for a.e. $y \in \mathbb{R}^N$) belongs to $L^1(\mathbb{R}^N)$ thanks to Hölder's inequality, and that $|(u*v)(x)| \leqslant \|u\|_\infty \|v\|_1$.

The second part follows easily from Lemma 4.4.1. □

4.4.2 Mollifier

Definition 4.4.2. *Let $\rho \in \mathcal{C}_c^\infty(\mathbb{R}^N; [0, +\infty))$ be such that $\operatorname{supp}(\rho) \subset B(0,1)$ and $\int_{\mathbb{R}^N} \rho(x)\,dx = 1$. For all $n \in \mathbb{N}$, we set $\rho_n(x) := n^N \rho(nx)$ so that $\int_{\mathbb{R}^N} \rho_n(x)\,dx = 1$ and $\operatorname{supp}(\rho_n) \subset B(0, \frac{1}{n})$. We say that the sequence $(\rho_n)_{n\in\mathbb{N}}$ is a mollifier.*

Lemma 4.4.2. *The following properties hold true.*

1. *If $u \in \mathcal{C}_c(\mathbb{R}^N)$, then $u * \rho_n$ converges uniformly to u as $n \to \infty$.*
2. *If $u \in L^p(\mathbb{R}^N)$ for $1 \leqslant p < \infty$, then $u * \rho_n \to u$ in $L^p(\mathbb{R}^N)$ as $n \to 0$.*

Proof. Assume that $u \in \mathcal{C}_c(\mathbb{R}^N)$. Since $\int_{\mathbb{R}^N} \rho_n(x)\,dx = 1$ and $\operatorname{supp}(\rho_n) \subset B(0, \frac{1}{n})$, then

$$
\begin{aligned}
|(u*\rho_n)(x) - u(x)| &= \left| \int_{\mathbb{R}^N} (u(x-y) - u(x))\rho_n(y)\,dy \right| \\
&\leqslant \int_{B(0,\frac{1}{n})} |u(x-y) - u(x)|\rho_n(y)\,dy \\
&\leqslant \sup\{|u(z_1) - u(z_2)| : z_1,\, z_1 \in \mathbb{R}^N,\, |z_1 - z_2| \leqslant 1/n\} \to 0
\end{aligned}
$$

as $n \to \infty$ since u is uniformly continuous. Hence $u * \rho_n$ converges uniformly to u.

For this same u, we use Hölder's inequality with $f_x(y) := (u(x-y) - u(x))(\rho_n(y))^{\frac{1}{p}}$ and $g(y) := (\rho_n(y))^{\frac{1}{p'}}$, we get

$$
|(u*\rho_n)(x) - u(x)| = \left| \int_{\mathbb{R}^N} f_x(y) g(y)\,dy \right| \quad \leqslant \quad \|f_x\|_p \,\|g\|_{p'}.
$$

But

$$
\|f_x\|_p = (\int_{\mathbb{R}^N} |u(x-y) - u(x)|^p \rho_n(y)\,dy)^{\frac{1}{p}} \text{ and } \|g\|_{p'} = (\int_{\mathbb{R}^N} \rho_n(y)\,dy)^{\frac{1}{p'}} = 1,
$$

so that by Fubini's Theorem, we get

$$
\begin{aligned}
\int_{\mathbb{R}^N} |(u*\rho_n)(x) - u(x)|^p\,dx &\leqslant \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} |u(x-y) - u(x)|^p \rho_n(y)\,dy\,dx \\
&= \int_{B(0,\frac{1}{n})} \left(\int_{\mathbb{R}^N} |u(x-y) - u(x)|^p\,dx \right) \rho_n(y)\,dy \\
&\leqslant \sup_{y \in B(0,\frac{1}{n})} \|\tau_y u - u\|_p^p,
\end{aligned}
$$

where $\tau_y u(x) := u(x-y)$ is the translation of u by y. On $B(0,1) + \operatorname{supp}(u)$ the uniform convergence implies the convergence in $L^p(\mathbb{R}^N)$, hence $\sup_{y\in B(0,\frac{1}{n})} \|\tau_y u - u\|_p^p$ and thus $\|u * \rho_n - u\|_p$ converge to 0 as $n \to \infty$.

The general case $u \in L^p(\mathbb{R}^N)$ follows from the density of $\mathcal{C}_c(\mathbb{R}^N)$ in $L^p(\mathbb{R}^N)$ (Theorem 4.3.2). More precisely, for any $\varepsilon > 0$, there exists $\tilde{u} \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\|\tilde{u} - u\|_p \leqslant \varepsilon$. Moreover, according to the previous case, there exists n_0 such that for $n \geqslant n_0$, $\|u * \rho_n - u\|_p \leqslant \varepsilon$. As a consequence, we have

$$\|\tilde{u} * \rho_n - \tilde{u}\|_p \leqslant \|u * \rho_n - u\|_p + \|u * \rho_n - \tilde{u} * \rho_n\|_p + \|u - \tilde{u}\|_p \leqslant 3\varepsilon.$$

For the second term of the right hand side we used (4.1). $\square$

Remark 4.4.1. *Let us observe that proceeding as in the proof above we can show that for any $1 \leqslant p < \infty$, for any $u \in L^p(\mathbb{R}^N)$, $\sup_{y \in B(0,\delta)} \|\tau_y u - u\|_p$ goes to 0 when δ goes to 0.*

Corollary 4.4.2. *For every open set $\Omega \subset \mathbb{R}^N$, the space $\mathcal{C}_c^\infty(\Omega)$ is dense in $L^p(\Omega)$.*

Proof. From Theorem 4.3.2, for every $u \in L^p(\Omega)$ and every $\varepsilon > 0$, there exist $v \in \mathcal{C}_c(\Omega)$ such that

$$\left(\int_\Omega |u - v|^p \, dx\right)^{1/p} < \frac{\varepsilon}{2}.$$

Extend v by zero outside Ω and denote by $\tilde{v}$ this extension. Then $\tilde{v} \in \mathcal{C}_c(\mathbb{R}^N)$ and $\mathrm{supp}(\tilde{v}) \subset \Omega$. By Lemma 4.4.2, we can choose $n_0 \in \mathbb{N}$ large enough so that for all $n \geqslant n_0$,

$$\left(\int_{\mathbb{R}^N} |\tilde{v} - (\tilde{v} * \rho_n)|^p \, dx\right)^{1/p} < \frac{\varepsilon}{2}.$$

By Corollary 4.4.1, we have

$$\mathrm{supp}(v * \rho_n) \subset \overline{\mathrm{supp}(v) + \mathrm{supp}(\rho_n)} \subset \overline{\mathrm{supp}(v) + B(0, 1/n)}.$$

Hence for $n > \max\{n_0, 1/\mathrm{dist}(\mathrm{supp}(v), \mathbb{R}^N \setminus \Omega))$ we have $\mathrm{supp}(v * \rho_n) \subset \Omega$ so that $v * \rho_n \in \mathcal{C}_c^\infty(\Omega)$ and

$$\left(\int_\Omega |u - (\tilde{v} * \rho_n)|^p \, dx\right)^{1/p} < \varepsilon.$$

$\square$

4.5 A compactness result

This regularization procedure enables one to transpose to Lebesgue spaces the Ascoli criterion for continuous functions.

Theorem 4.5.1 (Riesz - Fréchet - Kolmogorov). *Let $\Omega \subset \mathbb{R}^N$ be an open set and ω be a bounded open subset of Ω such that $\overline{\omega} \subset \Omega$. Let $A \subset L^p(\Omega)$ for $1 \leqslant p < \infty$ such that*

(i) $\displaystyle\sup_{u \in A} \|u\|_{L^p(\Omega)} < \infty$;

(ii) $\displaystyle\sup_{y \in B(0,\delta),\, u \in A} \|\tau_y u - u\|_{L^p(\omega)} \to 0$ *as* $\delta \to 0$.

Then $A_{|\omega}$ (the restriction to ω of functions in A) has compact closure in $L^p(\omega)$.

Proof. We first observe that $\tau_y u$ is well defined on ω as long as $\delta < \mathrm{dist}(\omega, \mathbb{R}^N \setminus \Omega)$ so that (ii) actually makes sense. Upon extending the elements of A by 0 outside Ω, we can assume without loss of generality that $\Omega = \mathbb{R}^N$.

Since $L^p(\Omega)$ is complete, so is the closure of $A_{|\omega}$. According to Theorem 1.3.2, it is therefore enough to prove that it is possible to cover $A_{|\omega}$ by a finite number of open balls in $L^p(\omega)$ of radius at most ε, for $\varepsilon > 0$. Let us start by selecting $n \in \mathbb{N}$ large enough so that

$$\sup_{y \in B(0,\frac{1}{n}),\, u \in A} \|\tau_y u - u\|_{L^p(\omega)} < \frac{\varepsilon}{3}.$$

Let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of mollifiers and let $B := \{(u*\rho_n)_{|\overline{\omega}} : u \in A\}$, we will prove that the set $B \subset \mathcal{C}(\overline{\omega})$ fulfills the assumptions of Ascoli's Theorem. First of all, for any $x \in \overline{\omega}$, we have

$$\sup_{u\in A} |u*\rho_n(x)| \leqslant \sup_{u\in A} \int_{\mathbb{R}^N} |u(x-y)|\rho_n(y)\,dy \leqslant \sup_{u\in A} \|u\|_{L^p(\mathbb{R}^N)} \|\rho_n\|_{L^{p'}(\mathbb{R}^N)} < \infty,$$

where we used Hölder's inequality and $1/p + 1/p' = 1$. On the other hand,

$$\begin{aligned}
\sup_{u\in A,\, x_1,\, x_2\in\overline{\omega}} |u*\rho_n(x_1) - u*\rho_n(x_2)| &= \sup_{u\in A,\, x_1,\, x_2\in\overline{\omega}} |u*(\rho_n - \tau_{x_1-x_2}\rho_n)(x_1)| \\
&\leqslant \sup_{u\in A} \|u\|_{L^1(\overline{\omega}+\overline{B}(0,\frac{1}{n}))} \sup_{x_1,\, x_2\in\overline{\omega}} \|\rho_n - \tau_{x_1-x_2}\rho_n\|_{L^\infty(\overline{\omega}+\overline{B}(0,\frac{1}{n}))} \\
&\leqslant C \sup_{u\in A} \|u\|_{L^p(\mathbb{R}^N)} \|\rho_n\|_{\mathrm{Lip}} |x_1 - x_2|.
\end{aligned}$$

It follows from Ascoli's Theorem that B has compact closure in $\mathcal{C}(\overline{\omega})$. Let $\sigma > 0$ to be fixed later. There exist finitely many balls of radius $\sigma\varepsilon$ whose union covers B in $\mathcal{C}(\overline{\omega})$. But for u and $v \in \mathcal{C}(\overline{\omega})$,

$$\|u-v\|_{L^p(\omega)} = \left(\int_\omega |u-v|^p\,dx\right)^{1/p} \leqslant \mathcal{L}^N(\omega)^{1/p} d_u(u,v).$$

Let us choose $\sigma := \frac{\mathcal{L}^N(\omega)^{-1/p}}{3}$ and we deduce that B is contained in a finite union of ball in $L^p(\omega)$ whose radius is at most $\varepsilon/3$. Finally, it follows from the proof of Lemma 4.4.2 that for $u \in A$,

$$\|u*\rho_n - u\|_{L^p(\omega)} \leqslant \sup_{y\in B(0,\frac{1}{n})} \|\tau_y u - u\|_{L^p(\omega)} \leqslant \frac{\varepsilon}{3}.$$

Finally Minkowski's inequality yields that $A_{|\omega}$ is contained in a finite union of balls of radius at most ε, and the proof is complete. □

Chapter 5

Continuous linear maps

5.1 Space of continuous linear maps

This section is devoted to $\mathcal{L}_c(X,Y)$ the space of the bounded linear operators between normed linear spaces X and Y. Let us recall that a linear operator T is bounded if one of the following assertion is satisfied:

1. T is bounded on every ball,
2. T is bounded on some ball,
3. T is continuous at every point,
4. T is continuous at some point.
5. T is uniformly continuous.
6. T is Lipschitz.

Theorem 5.1.1. *If X and Y are some normed linear spaces, then $\mathcal{L}_c(X,Y)$ is a normed linear space with the norm*

$$\begin{aligned} \|T\|_{\mathcal{L}_c(X,Y)} &:= \sup_{x\neq x'\in X} \frac{\|Tx-Tx'\|_Y}{\|x-x'\|_X} \\ &= \sup_{x\in X\setminus 0} \frac{\|Tx\|_Y}{\|x\|_X}, \\ &= \sup_{\|x\|_X\leqslant 1} \|Tx\|_Y, \\ &= \sup_{\|x\|_X=1} \|Tx\|_Y. \end{aligned}$$

If moreover Y is a Banach space then $\mathcal{L}_c(X,Y)$ is a Banach space.

Proof. It is easy to check that $\mathcal{L}_c(X,Y)$ is a normed linear space, and the only issue is to show that it is complete. When no confusion is possible we will simply denote $\|\cdot\|$ instead of $\|\cdot\|_{\mathcal{L}_c(X,Y)}$. Suppose that (T_n) is a Cauchy sequence in $\mathcal{L}_c(X,Y)$. Then for each $x\in X$, the sequence (T_nx) is Cauchy in the complete space Y, so there exists $Tx\in Y$ with $T_nx\to Tx$. Clearly $T: X\to Y$ is linear. Is it bounded? The real sequence $\|T_n\|$ is Cauchy, hence bounded, say $\|T_n\|\le K$. It follows that $\|T\|\le K$, and so $T\in\mathcal{L}_c(X,Y)$. To conclude the proof, we need to show that $\|T_n-T\|\to 0$. We have

$$\begin{aligned} \|T_n-T\| = \sup_{\|x\|\le 1}\|T_nx-Tx\|_Y = \sup_{\|x\|\le 1}\lim_{m\to\infty}\|T_nx-T_mx\|_Y \\ = \limsup_{m\to\infty}\sup_{\|x\|\le 1}\|T_nx-T_mx\|_Y \le \limsup_{m\to\infty}\|T_n-T_m\|_Y. \end{aligned}$$

Thus $\limsup_{n\to\infty}\|T_n-T\|=0$. $\square$

If $T \in \mathcal{L}_c(X,Y)$ and $U \in \mathcal{L}_c(Y,Z)$, then $UT = U \circ T \in \mathcal{L}_c(X,Z)$ and

$$\|UT\|_{\mathcal{L}_c(X,Z)} \leq \|U\|_{\mathcal{L}_c(Y,Z)} \|T\|_{\mathcal{L}_c(X,Y)}.$$

In particular, $\mathcal{L}_c(X) := \mathcal{L}_c(X,X)$ is a algebra, *i.e.*, it has an additional "multiplication" operation which makes it a non-commutative algebra, and this operation is continuous.

The dual space of X is $X' := \mathcal{L}_c(X,\mathbb{R})$ (or $\mathcal{L}_c(X,\mathbb{C})$ for complex vector spaces). According to the previous proposition it is a Banach space (whether X is or not).

Definition 5.1.1. *Given some normed linear spaces X and Y, and a sequence $(u_n)_n$ a sequence in $\mathcal{L}_c(X,Y)$, we say that*

1. *$(u_n)_n$ converges strongly to u in $\mathcal{L}_c(X,Y)$ if $\|u_n - u\|_{\mathcal{L}_c(X,Y)} \to 0$ when $n \to \infty$,*
2. *$(u_n)_n$ converges weakly* to u in $\mathcal{L}_c(X,Y)$ if for any $x \in X$, $(u_n(x))_n$ converges to $u(x)$ in Y.*

5.2 Uniform boundedness principle–Banach-Steinhaus theorem

We start this section with the following result.

Proposition 5.2.1. *Let X be a normed vector space and Y be a Banach space. Consider a dense subset A of X and $(u_n)_n$ a sequence in $\mathcal{L}_c(X,Y)$ such that*

1. *$\sup_n \|u_n\|_{\mathcal{L}_c(X,Y)} < \infty$,*
2. *for any x in A, $(u_n(x))_n$ converges.*

Then there exists a unique u in $\mathcal{L}_c(X,Y)$ such that $(u_n)_n$ converges weakly to u in $\mathcal{L}_c(X,Y)$. Moreover*

$$\|u\|_{\mathcal{L}_c(X,Y)} \leqslant \liminf_{n\to\infty} \|u_n\|_{\mathcal{L}_c(X,Y)}. \tag{5.1}$$

Proof. We can assume without loss of generality that $\sup_n \|u_n\| > 0$, otherwise the proof is straightforward. Let x in X and $\varepsilon > 0$. Since A is a dense subset of X, there exists $x' \in A$ such that $\|x - x'\|_X \leqslant \delta := \frac{\varepsilon}{3 \sup_n \|u_n\|_{\mathcal{L}_c(X,Y)}}$. Since $(u_n(x'))_n$ converges, it has the Cauchy property and there exists $n_\varepsilon \in \mathbb{N}$ such that $\|u_n(x') - u_m(x')\|_Y \leqslant \frac{\varepsilon}{3}$ for any $n, m \geqslant n_\varepsilon$. For any $n, m \geqslant n_\varepsilon$, we have

$$\begin{aligned}
\|u_n(x) - u_m(x)\|_Y &\leqslant \|u_n(x) - u_n(x')\|_Y + \|u_n(x') - u_m(x')\|_Y + \|u_m(x') - u_m(x)\|_Y, \\
&\leqslant (\|u_n\|_{\mathcal{L}_c(X,Y)} + \|u_m\|_{\mathcal{L}_c(X,Y)})\|x - x'\|_X + \frac{\varepsilon}{3}, \\
&\leqslant 2\delta \sup_n \|u_n\|_{\mathcal{L}_c(X,Y)} + \frac{\varepsilon}{3} \\
&\leqslant \varepsilon.
\end{aligned}$$

Therefore the sequence $(u_n(x))_n$ has the Cauchy property. Since Y is a Banach space, it converges. Let us denote by u the (weak) limit of $(u_n)_n$ thus obtained. It is straighforward that u is linear. Moreover for any $x \in X$,

$$\|u(x)\|_Y = \lim_{n\to\infty} \|u_n(x)\|_Y \leqslant \liminf_{n\to\infty} \|u_n\|_{\mathcal{L}_c(X,Y)} \|x\|_X,$$

which yields (5.1). □

The two following results are some consequences of the Baire theorem 1.2.2.

Theorem 5.2.1 (Uniform boundedness principle). *Let X be a Banach space and Y be a normed vector space. Suppose that B is a collection of continuous linear operators from X to Y. The uniform boundedness principle states that if for all x in X we have*

$$\sup_{u\in B} \|u(x)\|_Y < \infty. \tag{5.2}$$

Then

$$\sup_{u\in B} \|u\|_{\mathcal{L}_c(X,Y)} < \infty. \tag{5.3}$$

Proof. We assume that $\sup_{u\in B}\|u\|_{\mathcal{L}_c(X,Y)} = +\infty$ and we are going to prove that there exists x in X such that

$$\sup_{u\in B}\|u(x)\|_Y = +\infty. \tag{5.4}$$

For $u \in B$ and $n \in \mathbb{N}^*$, we set $\Theta_{u,n} := \{x \in X : \|u(x)\|_Y > n\}$ and $\Theta_n := \bigcup_{u\in B}\Theta_{u,n}$, which are all open. We are going to prove that Θ_n is dense in X for all $n \in \mathbb{N}^*$. Let x in $X \setminus \Theta_n$, *i.e.*, $\sup_{u\in B}\|u(x)\|_Y \leqslant n$. Since we assume that $\sup_{u\in B}\|u\|_{\mathcal{L}_c(X,Y)} = \infty$, then for each $\varepsilon > 0$, there exists v in B such that $\|v\|_{\mathcal{L}_c(X,Y)} > 2n/\varepsilon$. Hence there must exist $x' \in X$ such that $\|x'\|_X = 1$ and $\|v(x')\|_X > 2n/\varepsilon$. Therefore

$$\|v(x + \varepsilon x')\|_Y \geqslant -\|v(x)\|_Y + \varepsilon\|v(x')\|_Y > n,$$

so that $x + \varepsilon x' \in \Theta_n$. Hence Θ_n is dense. Now using the Baire theorem 1.2.2 we get that $\bigcap_{n\in\mathbb{N}^*} \Theta_n$ is dense in X. In particular $\bigcap_{n\in\mathbb{N}^*} \Theta_n$ is not empty, and thus there exists $x \in \bigcap_{n\in\mathbb{N}^*} \Theta_n$ which implies 5.4. □

We infer from Theorem 5.2.1 the following corollary.

Corollary 5.2.1 (Banach-Steinhaus). *Let X be a Banach space and Y be a normed vector space. If $(u_n)_n$ is a sequence of $\mathcal{L}_c(X,Y)$ which converges weakly* to u, then u is in $\mathcal{L}_c(X,Y)$ and*

$$\|u\|_{\mathcal{L}_c(X,Y)} \leqslant \liminf_{n\to\infty}\|u_n\|_{\mathcal{L}_c(X,Y)}.$$

In the particular case where Y is $\mathbb{R}$ or $\mathbb{C}$, we denote by $X' := \mathcal{L}_c(X,Y)$, and we call X' the *topological dual space* of X. Applying the uniform boundedness principle yields the following result:

Corollary 5.2.2. *Let X be a Banach space. Then any weakly* converging sequence in X' is bounded.*

One advantage of the weak* convergence is that the following partial converse is available:

Theorem 5.2.2. *Let X be a separable Banach space. Then any bounded sequence in X' admits a weakly* converging subsequence.*

Proof. Let $(u_n)_n$ be a bounded sequence of X'. Let $(x_k)_{k\in\mathbb{N}}$ a dense sequence in X. Using Cantor's diagonal argument there exists a subsequence $(u_{n_j})_j$ of $(u_n)_n$ such that $(u_{n_j}(x_k))_j$ converges for any $k \in \mathbb{N}$. We then apply Proposition 5.2.1. □

5.3 Geometry of Banach spaces and identification of their dual

We will slightly restrict our attention to spaces which enjoy two properties described in the definitions below.

Definition 5.3.1 (Uniformly convex normed vector space). *We say that a normed vector space X is uniformly convex if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for any x, $y \in X$ with $\|x\|_X \leqslant 1$, $\|y\|_X \leqslant 1$ and $\left\|\frac{x+y}{2}\right\|_X > 1 - \delta$, then $\left\|\frac{x-y}{2}\right\|_X < \varepsilon$.*

In particular the unit sphere of a uniformly convex normed vector space contains no segments. As a first application, let us give the following proposition.

Proposition 5.3.1. *Let X be a uniformly convex Banach space, and let f in $X' \setminus \{0\}$. Then there exists a unique $u \in X$ such that $\|u\|_X = 1$ and $f(u) = \|f\|_{X'}(:= \max_{\|v\|_X=1}\|f(v)\|)$.*

Proof. Let $(u_n)_n$ be a maximizing sequence. We assume without loss of generality that $f(u_n) \geqslant 0$. Let $\varepsilon > 0$ and δ be given as in Definition 5.3.1. Since the sequence $(u_n)_n$ maximizes $|f(v)|$ over $\{v \in X : \|v\|_X = 1\}$, there exists $N \in \mathbb{N}$ such that for any $n \geqslant N$, $f(u_n) > (1-\delta)\|f\|_{X'}$. Therefore for any $n, m \geqslant N$, we have

$$\frac{1}{2}(f(u_n) + f(u_m)) > (1-\delta)\|f\|_{X'}.$$

But since

$$\frac{1}{2}(f(u_n) + f(u_m)) = f\left(\frac{1}{2}(u_n + u_m)\right) \leqslant \|f\|_{X'}\left\|\frac{1}{2}(u_n + u_m)\right\|,$$

it follows, by combining both previous inequalities and dividing by $\|f\|_{X'}$, that

$$\left\|\frac{1}{2}(u_n+u_m)\right\| > 1-\delta. \tag{5.5}$$

Moreover, since $\|u_n\|_X = 1$, then $\frac{f(u_n)}{\|f\|_{X'}} \leqslant 1$, and using the uniform convexity of X, we can infer from (5.5) that $\|u_n - u_m\|_X < 2\varepsilon$, that is the sequence $(u_n)_n$ satisfies the Cauchy property. Since X is complete, there exists $u \in X$ such that $u_n \to u$. Using the continuity of the norm and of the function f we get that $f(u) = \|f\|_{X'}$.

Uniqueness is proved by proceeding in the same way. ☐

Definition 5.3.2 (Smooth normed vector space). *We say that a normed vector space X is smooth if for any linearly independent elements x and y, the function $t \in \mathbb{R} \mapsto \|x+ty\|_X$ is differentiable for all $t \in \mathbb{R}$.*

Let us stress that this is equivalent to the fact that for any $u \in X \setminus \{0\}$, then the limit

$$\lim_{\varepsilon\to 0} \frac{\|u+\varepsilon v\|_X - \|u\|_X}{\varepsilon}$$

exists for any $v \in X$.

Proposition 5.3.2. *Let X be a smooth normed vector space over $\mathbb{R}$. Then the norm is Fréchet-differentiable outside 0, i.e., for any $u \in X \setminus \{0\}$, there exists $L \in X'$, that we denote $D\|\cdot\|_X(u)$, such that for any $v \in X$,*

$$\lim_{\varepsilon\to 0} \frac{\|u+\varepsilon v\|_X - \|u\|_X}{\varepsilon} = L(v).$$

Moreover

$$\|D\|\cdot\|_X(u)\|_{X'}\|u\|_X = D\|\cdot\|_X(u)(u) = \|u\|_X. \tag{5.6}$$

Proof. Since X is a smooth normed vector space, for any $u \in X \setminus \{0\}$, for any $v \in X$,

$$D\|\cdot\|_X(u)(v) := \lim_{\varepsilon\to 0} \frac{\|u+\varepsilon v\|_X - \|u\|_X}{\varepsilon}$$

exists. We first prove that $D\|\cdot\|_X(u)$ is linear. For any $\lambda \in \mathbb{R}\setminus\{0\}$ and any u, $v \in X$,

$$\frac{\|u+\varepsilon\lambda v\|_X - \|u\|_X}{\varepsilon} = \lambda \frac{\|u+\varepsilon\lambda v\|_X - \|u\|_X}{\lambda\varepsilon}$$

so that

$$D\|\cdot\|_X(u)(\lambda v) = \lambda D\|\cdot\|_X(u)(v).$$

Now, using the triangle inequality, we have for any v_1, v_2 in X, that

$$\frac{\|u+\varepsilon(v_1+v_2)\|_X - \|u\|_X}{\varepsilon} \leqslant \frac{\|u+2\varepsilon v_1\|_X - \|u\|_X}{2\varepsilon} + \frac{\|u+2\varepsilon v_2\|_X - \|u\|_X}{2\varepsilon}$$

so that, passing to the limit $\varepsilon \to 0$,

$$D\|\cdot\|_X(u)(v_1+v_2) \leqslant D\|\cdot\|_X(u)(v_1) + D\|\cdot\|_X(u)(v_2). \tag{5.7}$$

Applying now (5.7) to $-v_1$ and $-v_2$ implies that the previous inequality is actually an equality. Hence the map $D\|\cdot\|_X(u)$ is linear from X to $\mathbb{R}$.

We next show that it is continuous. Indeed using the triangle inequality we have that for any $v \in X$,

$$\left|\frac{\|u+\varepsilon v\|_X - \|u\|_X}{\varepsilon}\right| \leqslant \|v\|_X,$$

so that

$$|D\|\cdot\|_X(u)(v)| \leqslant \|v\|_X.$$

Therefore $D\|\cdot\|_X(u)$ is a continuous linear map from X to $\mathbb{R}$ with norm $\|D\|\cdot\|_X(u)\|_{X'} \leqslant 1$. Moreover since $D\|\cdot\|_X(u)(u) = \|u\|$, then $\|D\|\cdot\|_X(u)\|_{X'} = 1$ which yields (5.6). ☐

Theorem 5.3.1 (Identification of the dual of a smooth and uniformly convex Banach space). *Let X be a smooth and uniformly convex Banach space, and let f in $X' \setminus \{0\}$. Then there exists a unique $u \in X$ such that $\|u\|_X = 1$ and $f = \|f\|_{X'} D\| \cdot \|_X(u)$.*

Proof. Let u be given by Proposition 5.3.1, and let $v \in X$. The function

$$g : \varepsilon \mapsto f(u + \varepsilon v) - \|f\|_{X'} \|u + \varepsilon v\|_X$$

is nonpositive, vanishes for $\varepsilon = 0$, and is derivable at $\varepsilon = 0$, since f is linear continuous and X is smooth. Therefore $g'(0) = 0$, *i.e.*

$$f(v) - \|f\|_{X'} D\| \cdot \|_X(u)(v) = 0.$$

Let us now prove the uniqueness part of Theorem 5.3.1. If $u \in X$ is such that $\|u\|_X = 1$ and $f = \|f\|_{X'} D\| \cdot \|_X(u)$, then in particular $f(u) = \|f\|_{X'} D\| \cdot \|_X(u)(u) = \|f\|_{X'}$, according to (5.6). It then suffices to recall the uniqueness part of Proposition 5.3.1. □

An important corollary is the following constructive version of Hahn-Banach theorem which allows the extension of continuous linear forms defined on a subspace of a smooth uniformly convex Banach space to the whole space without increasing the norm.

Corollary 5.3.1 (Hahn-Banach). *Let X be a smooth uniformly convex Banach space, Y be a subspace of X and $f \in Y'$. There exists an unique extension $\tilde{f}$ of f to X such that $\|\tilde{f}\|_{X'} = \|f\|_{Y'}$.*

Proof. Let f in Y'. Only the case $f \neq 0$ deserves a proof. According to the Theorem of extension of uniformly continuous functions 1.2.3 there exists a unique extension $\overline{f}$ of f to the closure $\overline{Y}$ of Y in X and $\|\overline{f}\|_{\overline{Y}'} = \|f\|_{Y'}$. Since $\overline{Y}$ is a closed subspace of the Banach space X, then $\overline{Y}$ is a smooth uniformly convex Banach space. Thanks to Theorem 5.3.1 there exists a unique $u \in \overline{Y}$ such that $\|u\|_{\overline{Y}} = 1$ and $\overline{f} = \|\overline{f}\|_{\overline{Y}'} D\| \cdot \|_{\overline{Y}}(u)$. Let us now consider $\tilde{f} := \|f\|_{Y'} D\| \cdot \|_X(u)$ in X'. Then $\tilde{f}$ is an extension of $\overline{f}$ and $\|\tilde{f}\|_{X'} = \|f\|_{Y'}$, according to (5.6).

It remains to prove the uniqueness part. Let $\check{f}$ an extension of f to X such that $\|\check{f}\|_{X'} = \|f\|_{Y'}$. Then evaluating $\check{f}$ for the value u above yields $\check{f}(u) = \overline{f}(u) = \|f\|_{Y'} = \|\check{f}\|_{X'}$. According to (the proof of) Theorem 5.3.1, $\check{f} = \|\check{f}\|_{X'} D\| \cdot \|_X(u) = \tilde{f}$. □

The previous result can actually be extended to a more general setting with a non-constructive proof using Zorn's lemma.

Theorem 5.3.2 ((General) Hahn-Banach). *If f is a bounded linear functional on a subspace of a normed linear space, then f extends to the whole space with preservation of norm.*

Note that, unlike the previous result, Theorem 5.3.2 does not contain any uniqueness part. It is named for Hans Hahn and Stefan Banach who proved this theorem independently in the late 1920s, though it was proved earlier (in 1912) by Eduard Helly. A key consequence is that the dual space of a nontrivial normed linear space is itself nontrivial. (Note: the norm is important for this. There exist topological vector spaces, e.g., L^p for $0 < p < 1$, with no non-zero continuous linear functionals.) Note that there are virtually no hypotheses beyond linearity and existence of a norm. In fact for some purposes a weaker version is useful. For X a vector space, we say that $p : X \to \mathbb{R}$ is sublinear if $p(x + y) \leq p(x) + p(y)$ and $p(\alpha x) = \alpha p(x)$ for $x, y \in X$, $\alpha \geq 0$.

Theorem 5.3.3. *Let X be a vector space, $p : X \to \mathbb{R}$ a sublinear functional, S a subspace of X, and $f : S \to \mathbb{R}$ a linear function satisfying $f(x) \leq p(x)$ for all $x \in S$, then f can be extended to X so that the same inequality holds for all $x \in X$.*

Proof. It suffices to extend f to the space spanned by S and one element $x_0 \in X \setminus S$, preserving the inequality, since if we can do that we can complete the proof with Zorn's lemma.

We need to define $f(x_0)$ such that $f(tx_0 + s) \leq p(tx_0 + s)$ for all $t \in \mathbb{R}$, $s \in S$. The case $t = 0$ is known and it is easy to use homogeneity to restrict to $t = \pm 1$. Thus we need to find a value $f(x_0) \in \mathbb{R}$ such that

$$f(s) - p(-x_0 + s) \leq f(x_0) \leq p(x_0 + s) - f(s) \quad \text{for all } s \in S.$$

Now it is easy to check that for any s_1, $s_2 \in S$, $f(s_1) - p(-x_0 + s_1) \leq p(x_0 + s_2) - f(s_2)$, and so such an x_0 exists. □

Chapter 6

Duality in the Lebesgue spaces and bounded measures

We now go back to the Lebesgue spaces studied in chapter 4 in light of the results obtained in chapter 5. In the present chapter, μ is a positive Radon measure on an open subset Ω of $\mathbb{R}^N$.

6.1 Uniform convexity and smoothness of the norm

Proposition 6.1.1. *For every $1 < p < \infty$, the space $L^p(\Omega, \mu)$ is uniformly convex.*

The proof rests essentially on the strict convexity of the map $s \mapsto s^p$. More precisely we will rely on the following lemma which is left to the reader.

Lemma 6.1.1. *For any $\varepsilon > 0$, there exists $\delta > 0$ such that for any and s, t in $\mathbb{C}$ with $|s| \leqslant 1$, $|t| \leqslant 1$ and $|s-t| \geqslant 2\varepsilon$, then*

$$|\frac{s+t}{2}|^p \leqslant (1-\delta)\frac{|s|^p + |t|^p}{2}.$$

Nevertheless, one needs to pass from pointwise inequalities to uniform integral inequalities.

Proof. Let $\varepsilon > 0$. It suffices to prove that there exists $\delta > 0$ such that for any and u, v in $L^p(\Omega, \mu)$ with $\|u\|_{L^p(\Omega,\mu)} \leqslant 1$, $\|v\|_{L^p(\Omega,\mu)} \leqslant 1$ and $\|u - v\|_{L^p(\Omega,\mu)} \geqslant 2\varepsilon$, then

$$\left\| \frac{u+v}{2} \right\|_{L^p(\Omega,\mu)} \leqslant 1 - \delta.$$

Let us define

$$A := \left\{ x \in \Omega : \ |u(x) - v(x)|^p > \frac{\varepsilon^p}{2}(|u(x)|^p + |v(x)|^p) \right\}.$$

Let us define $r(x) := \max(|u(x)|, |v(x)|)$. By definition of A, $r(x) > 0$ for any $x \in A$, and

$$\left| \frac{u(x)}{r(x)} - \frac{v(x)}{r(x)} \right| \geqslant \frac{\varepsilon}{2^{\frac{1}{p}}}.$$

Thanks to Lemma 6.1.1 we deduce the existence of $\delta > 0$ such that

$$\left| \frac{u(x) + v(x)}{2} \right|^p \leqslant (1-\delta) \left(\frac{|u(x)|^p + |v(x)|^p}{2} \right)$$

for any $x \in A$. On the other hand, for any $x \in \Omega \setminus A$, by convexity of the map $s \mapsto |s|^p$,

$$\left| \frac{u(x) + v(x)}{2} \right|^p \leqslant \frac{|u(x)|^p + |v(x)|^p}{2}.$$

Integrating over the corresponding domains yields

$$\begin{aligned}\int_\Omega \frac{|u(x)|^p+|v(x)|^p}{2}\,d\mu - \int_\Omega \left|\frac{u(x)+v(x)}{2}\right|^p\,d\mu \\ &\geqslant \int_A \frac{|u(x)|^p+|v(x)|^p}{2}\,d\mu - \int_A \left|\frac{u(x)+v(x)}{2}\right|^p\,d\mu\\ &\geqslant \int_A \frac{|u(x)|^p+|v(x)|^p}{2}\,d\mu - (1-\delta)\int_A \frac{|u(x)|^p+|v(x)|^p}{2}\,d\mu\\ &\geqslant \frac{\delta}{2}\max(\|u\|^p_{L^p(A,\mu)},\|v\|^p_{L^p(A,\mu)}).\end{aligned}$$

Since $\|u\|_{L^p(\Omega,\mu)} \leqslant 1$, $\|v\|_{L^p(\Omega,\mu)} \leqslant 1$ the left hand side in the inequality above is less than $1-\|\frac{u+v}{2}\|^p_{L^p(\Omega,\mu)}$. On the other hand, by definition of A,

$$\left(\int_{\Omega\setminus A}|u-v|^p d\mu\right)^{1/p} \leqslant \frac{\varepsilon}{2^{1/p}}\left(\int_{\Omega\setminus A}|u|^p d\mu + \int_{X\setminus A}|v|^p d\mu\right)^{1/p} \leqslant \varepsilon.$$

As a consequence,

$$\|u-v\|_{L^p(A,\mu)} \geqslant \|u-v\|_{L^p(\Omega,\mu)} - \|u-v\|_{L^p(\Omega\setminus A,\mu)} \geqslant \varepsilon,$$

so that, using the triangle inequality, we get

$$\max(\|u\|_{L^p(A,\mu)},\|v\|_{L^p(A,\mu)}) \geqslant \frac{\varepsilon}{2}.$$

As a consequence,

$$1-\left\|\frac{u+v}{2}\right\|^p_{L^p(\Omega,\mu)} \geqslant \frac{\delta}{2}(\frac{\varepsilon}{2})^p$$

so that

$$\left\|\frac{u+v}{2}\right\|_{L^p(\Omega,\mu)} \leqslant \left(1-\frac{\delta}{2}(\frac{\varepsilon}{2})^p\right)^{1/p} \leqslant 1-\delta',$$

for $\delta' = \frac{\delta}{2p}(\frac{\varepsilon}{2})^p$ and the conclusion follows. $\square$

Proposition 6.1.2. *For every $1<p<\infty$, the space $L^p(\Omega,\mu)$ is smooth and for any u in $L^p(\Omega,\mu)\setminus 0$ and for any v in $L^p(\Omega,\mu)$,*

$$D\|\cdot\|_p(u)(v) = \|u\|_p^{1-p}\int_\Omega |u|^{p-2}uv\,d\mu.$$

Proof. Note that $u\mapsto \|u\|_{L^p(\Omega,\mu)}$ is the composition of $u\mapsto \|u\|^p_{L^p(\Omega,\mu)}$ and the map $g: s\mapsto s^{1/p}$. Hence it is enough to prove that for any u in $L^p(\Omega,\mu)\setminus 0$ and for any v in $L^p(\Omega,\mu)$, the function

$$\varepsilon \mapsto \int_\Omega |u+\varepsilon v|^p d\mu$$

admits a derivative in 0. Note that for μ-a.e. $x\in\Omega$,

$$\lim_{\varepsilon\to 0}\frac{|u(x)+\varepsilon v(x)|^p-|u(x)|^p}{\varepsilon} = p|u(x)|^{p-2}u(x)v(x),$$

and by the mean value Theorem, for every $\varepsilon\in(0,1)$, there exists $t_\varepsilon\in(0,\varepsilon)$ such that

$$\frac{|u(x)+\varepsilon v(x)|^p-|u(x)|^p}{\varepsilon} \leqslant p|u(x)+t_\varepsilon v(x)|^{p-1}|v(x)| \leqslant C(|v(x)|^p+|u(x)|^{p-1}|v(x)|) =: f(x),$$

for some constant $C>0$ depending only on p. Since by Hölder's inequality,

$$\begin{aligned}\int_\Omega f\,d\mu &\leqslant C\left(\int_\Omega |v|^p\,d\mu + \int_\Omega |u|^{p-1}|v|\,d\mu\right)\\ &\leqslant C\left(\|v\|^p_{L^p(\Omega,\mu)} + \|v\|_{L^p(\Omega,\mu)}\|u\|^{p-1}_{L^p(\Omega,\mu)}\right) < \infty.\end{aligned}$$

Hence from the dominated convergence Theorem

$$\lim_{\varepsilon\to 0}\frac{\|u+\varepsilon v\|^p_{L^p(\Omega,\mu)}-\|u\|^p_{L^p(\Omega,\mu)}}{\varepsilon}=p\int_\Omega |u|^{p-2}uv\,d\mu.$$

By the chain rule formula, we obtain that

$$\begin{aligned}D(g\circ\|\cdot\|^p_{L^p})(u).v &= Dg(\|u\|^p_{L^p}).\Big(D\|\cdot\|^p_{L^p})(u).v\Big)\\ &= g'(\|u\|^p_{L^p})\Big(D\|\cdot\|^p_{L^p})(u).v\Big)\\ &= \|u\|^{1-p}_{L^p(\Omega,\mu)}\int_\Omega |u|^{p-2}uv\,d\mu.\end{aligned}$$

□

Remark 6.1.1. For $\mu=\mathcal{L}^N$, the previous results fail for $p=1$ or ∞, *i.e.*, neither $L^1(\Omega,\mathcal{L}^N)$ nor $L^\infty(\Omega,\mathcal{L}^N)$ are uniformly convex nor smooth.

6.2 Duality in the Lebesgue spaces

Theorem 6.2.1. *For $1<p<\infty$, the dual of the space $L^p(\Omega,\mu)$ is isometrically isomorphic to $L^{p'}(\Omega,\mu)$ with $p':=p/(p-1)$. More precisely, for any $f\in(L^p(\Omega,\mu))'$, there exists a unique v in $L^{p'}(\Omega,\mu)$ such that*

$$f(u)=\int_\Omega uv\,d\mu,$$

for all $u\in L^p(\Omega,\mu)$, and

$$\|f\|_{(L^p(\Omega,\mu))'}=\|v\|_{L^{p'}(\Omega,\mu)}.$$

Proof. Thanks to Theorem 5.3.1 for any $f\in(L^p(\Omega,\mu))'\setminus 0$, there exists a unique $w\in L^p(\Omega,\mu)$ with $\|w\|_{L^p(\Omega,\mu)}=1$ such that for all $u\in L^p(\Omega,\mu)$,

$$f(u)=\|f\|_{(L^p(\Omega,\mu))'}D\|\cdot\|(w)(u),$$

that is, using now Proposition 6.1.2,

$$f(u)=\|f\|_{(L^p(\Omega,\mu))'}\|w\|_p^{1-p}\int_\Omega u|w|^{p-2}w\,d\mu=\int_\Omega uv\,d\mu,$$

with $v:=\|f\|_{(L^p(\Omega,\mu))'}\|w\|_p^{1-p}|w|^{p-2}w$. Since $w\in L^p(\Omega,\mu)$ with $\|w\|_{L^p(\Omega,\mu)}=1$, we have that $v\in L^{p'}(\Omega,\mu)$ with $\|v\|_{L^{p'}(\Omega,\mu)}=\|f\|_{(L^p(\Omega,\mu))'}$.

It remains now to prove that if v in $L^{p'}(\Omega,\mu)$ is such that $\int_\Omega uv\,d\mu=0$, for all $u\in L^p(\Omega,\mu)$, then $v=0$. To this purpose it suffices to take $u=|v|^{\frac{2-p}{p-1}}v$. □

Taking Theorem 6.2.1 into account the notion of weak convergence in $L^p(\Omega,\mu)$, seen as the dual of $L^{p'}(\Omega,\mu)$ (since $(p')'=p$), reads as follows.

Definition 6.2.1 (Weak convergence in $L^p(\Omega,\mu)$). *Let $1<p<\infty$, we say that a sequence $(u_n)\subset L^p(\Omega,\mu)$ converges weakly to u in $L^p(\Omega,\mu)$ if for any v in $L^{p'}(\Omega,\mu)$,*

$$\int_\Omega u_n v\,d\mu\to\int_\Omega uv\,d\mu$$

as $n\to\infty$.

Since for $1<p<\infty$, $L^p(\Omega,\mu)$ is a separable Banach space, it follows from Theorem 5.2.2 that any bounded sequence (u_n) in $L^p(\Omega,\mu)$ has a subsequence which converges weakly.

Theorem 6.2.2. *The dual of the space $L^1(\Omega,\mu)$ is isometrically isomorphic to $L^\infty(\Omega,\mu)$. More precisely, for any $f \in (L^1(\Omega,\mu))'$, there exists a unique $v \in L^\infty(\Omega,\mu)$ such that*

$$f(u) = \int_\Omega uv\, d\mu, \tag{6.1}$$

for all $u \in L^1(\Omega,\mu)$), and

$$\|f\|_{(L^1(\Omega,\mu))'} = \|v\|_{L^\infty(\Omega,\mu)}.$$

Proof. Let $f \in (L^1(\Omega,\mu))'$. Let $\phi \in L^2(\Omega,\mu)$ a strictly positive function such that for any K compact of Ω, there holds

$$\inf_{x\in K} \phi(x) > 0.$$

Then the application $f_\phi : w \in L^2(\Omega,\mu) \mapsto f(\phi w)$ is well defined since

$$\begin{aligned} |f(\phi w)| &\leqslant \|f\|_{(L^1(\Omega,\mu))'}\|\phi w\|_{L^1(\Omega,\mu)}, \\ &\leqslant \|f\|_{(L^1(\Omega,\mu))'}\|\phi\|_{L^2(\Omega,\mu)}\|w\|_{L^2(\Omega,\mu)}. \end{aligned}$$

Thus f_ϕ is in $(L^2(\Omega,\mu))'$ and using Theorem 6.2.1 there exists a unique v_ϕ in $L^2(\Omega,\mu)$ such that

$$\|f_\phi\|_{(L^2(\Omega,\mu))'} = \|v_\phi\|_{L^2(\Omega,\mu)},$$

and for any w in $L^2(\Omega,\mu)$,

$$f_\phi(w) = \int_\Omega v_\phi w\, d\mu,$$

i.e.

$$f(\phi w) = \int_\Omega \frac{v_\phi}{\phi}\phi w\, d\mu.$$

Let us denote $v := \frac{v_\phi}{\phi}$. We are going to prove that v is in $L^\infty(\Omega,\mu)$ with $\|v\|_{L^\infty(\Omega,\mu)} \leqslant \|f\|_{(L^1(\Omega,\mu))'}$. Let us assume ab absurdo that there exists $\varepsilon > 0$ and $A \subset \Omega$, measurable with $\mu(A) > 0$ such that $|v(x)| > \|f\|_{(L^1(\Omega,\mu))'} + \varepsilon$, for any $x \in A$. We then consider

$$w(x) = 1_A(x)\mathrm{sgn}(v(x)),$$

where $\mathrm{sgn}(u) = 1$ if $u > 0$, $\mathrm{sgn}(u) = -1$ if $u < 0$, and $\mathrm{sgn}(u) = 0$ if $u = 0$. Then

$$f(\phi w) = \int_\Omega \frac{v_\phi}{\phi}\phi w\, d\mu = \int_A |v|\phi\, d\mu \geqslant (\|f\|_{(L^1(\Omega,\mu))'} + \varepsilon)\int_A \phi\, d\mu,$$

and

$$f(\phi w) \leqslant \|f\|_{(L^1(\Omega,\mu))'}\|\phi w\|_{L^1(\Omega,\mu)} = \|f\|_{(L^1(\Omega,\mu))'}\int_A \phi\, d\mu,$$

what is a contradiction, since $\int_A \phi\, d\mu > 0$.

Finally, the fact that (6.1) holds for any $u \in L^1(\Omega,\mu)$ follows from the density of the smooth compactly supported functions in $L^1(\Omega,\mu)$. □

Definition 6.2.2 (Weak convergence in $L^1(\Omega,\mu)$). *We say that a sequence $(u_n) \subset L^1(\Omega,\mu)$ converges weakly to u in $L^1(\Omega,\mu)$ if for any $v \in L^\infty(\Omega,\mu)$,*

$$\int_\Omega u_n v\, d\mu \to \int_\Omega uv\, d\mu$$

as $n \to \infty$.

Let us stress that unlike in $L^p(\Omega,\mu)$ (for $1 < p < \infty$), bounded sequence in $L^1(\Omega,\mu)$ are not necessarily weakly relatively compact in that space. Indeed, if may happen that the accumulation points for the weak topology of bounded sequences in $L^1(\Omega,\mu)$ are outside $L^1(\Omega,\mu)$. For instance the Dirac mass can be obtained as a weak limit of sequences of $L^1(\Omega,\mu)$ functions, the so-called approximations to the identity.

6.3 Bounded Radon measures

Let us consider $\Omega \subset \mathbb{R}^N$ and denote by $\mathcal{C}_0(\Omega)$ the closure of the space $\mathcal{C}_c(\Omega)$ for the uniform topology.

Proposition 6.3.1. *A function $f \in \mathcal{C}_0(\Omega)$ if and only if for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \Omega$ such that $|f| < \varepsilon$ on $\Omega \setminus K_\varepsilon$.*

Proof. Let $\varepsilon > 0$ and a compact set $K \subset \Omega$ such that $|f| < \varepsilon$ on $\Omega \setminus K$. By Urysohn's Lemma one can find $g \in \mathcal{C}_c(\Omega; [0,1])$ such that $g = 1$ on K. Set $h = fg$, then $h \in \mathcal{C}_c(\Omega)$ and $\|f - h\|_\infty < \varepsilon$. Thus $f \in \mathcal{C}_0(\Omega)$.

Conversely, consider $f \in \mathcal{C}_0(\Omega)$, then there exists a sequence $(f_n) \subset \mathcal{C}_c(\Omega)$ such that $f_n \to f$ uniformly in Ω. Let $\varepsilon > 0$ and $n_\varepsilon \in \mathbb{N}$ such that $\|f_{n_\varepsilon} - f\|_\infty < \varepsilon/2$, and define $K := \{x \in \Omega / \ |f_{n_\varepsilon}| \geqslant \varepsilon/2\}$. Then K is a compact subset of Ω and for all $x \in \Omega \setminus K$, $|f| \leqslant |f - f_{n_\varepsilon}| + |f_{n_\varepsilon}| < \varepsilon$. $\square$

Definition 6.3.1. *The space of all bounded Radon measures over Ω, denoted by $\mathcal{M}(\Omega)$, is the dual space of $\mathcal{C}_0(\Omega)$.*

Thanks to the Riesz representation Theorem (Theorem 3.5.1) we can characterize the space of bounded Radon.

Theorem 6.3.1. *For every $L \in \mathcal{M}(\Omega)$ there exist two positive Radon measures λ^+ and λ^- on Ω such that*

$$L(u) = \int_\Omega u \, d\lambda^+ - \int_\Omega u \, d\lambda^-.$$

Proof. We claim that for every $L \in \mathcal{M}(\Omega)$, there exist some positive linear forms L^+ and L^- on $\mathcal{C}_0(\Omega)$ such that $L(u) = L^+(u) - L^-(u)$ for every $u \in \mathcal{C}_0(\Omega)$. With this result, the conclusion of the Theorem follows as an immediate consequence of the Riesz representation Theorem (Theorem 3.5.1).

We now prove the claim. To this aim, let us define the cone $\mathcal{C}^+ := \{u \in \mathcal{C}_0(\Omega) : u \geqslant 0\}$, and for $u \in \mathcal{C}^+$,

$$L^+(u) := \sup\{L(v) : v \in \mathcal{C}^+, \, v \leqslant u\}.$$

Step 1: *L^+ is positive and finite on $\mathcal{C}^+$.* Let $u \in \mathcal{C}^+$. As $0 \in \mathcal{C}^+$, $L^+(u) \geqslant 0$. Let now $v \in \mathcal{C}^+$ be such that $0 \leqslant v \leqslant u$. By continuity of L, $L(v) \leqslant \|L\|\|v\|_\infty \leqslant \|L\|\|u\|_\infty$, and then taking the sup with respect to v yields $0 \leqslant L^+(u) \leqslant \|L\|\|u\|_\infty < \infty$.

Step 2: *L^+ is additive on $\mathcal{C}^+$.* Let u_1 and $u_2 \in \mathcal{C}^+$ and $v \in \mathcal{C}^+$ be such that $0 \leqslant v \leqslant u_1 + u_2$. We decompose v as $v = \min(u_1, v) + \max(v - u_1, 0)$, where $\min(u_1, v) \leqslant u_1$ and $\max(v - u_1, 0) \leqslant u_2$. Since $\min(u_1, v)$ and $\max(v - u_1, 0) \in \mathcal{C}^+$, then

$$L(v) = L(\min(u_1, v)) + L(\max(v - u_1, 0)) \leqslant L^+(u_1) + L^+(u_2),$$

hence taking the sup over all v leads to

$$L^+(u_1 + u_2) \leqslant L^+(u_1) + L^+(u_2).$$

To prove the converse inequality, let $\varepsilon > 0$. By definition of L^+, there exists v_1 and $v_2 \in \mathcal{C}^+$ such that $0 \leqslant v_i \leqslant u_i$ and $L^+(u_i) \leqslant L(v_i) + \varepsilon$ for $i = 1, 2$. As $0 \leqslant v_1 + v_2 \leqslant u_1 + u_2$, it follows that

$$L^+(u_1 + u_2) \geqslant L(v_1 + v_2) = L(v_1) + L(v_2) \geqslant L^+(u_1) + L^+(u_2) - 2\varepsilon,$$

and it suffices to let $\varepsilon \to 0$.

Step 3: *Definition and additivity of L^+ on $\mathcal{C}_0(\Omega)$.* Let $u \in \mathcal{C}_0(\Omega)$. We decompose u as the difference of it positive and negative parts $u = u^+ - u^-$ with $u^\pm \in \mathcal{C}^+$. We set $L^+(u) = L^+(u^+) - L^+(u^-)$. Now if u and $v \in \mathcal{C}_0(\Omega)$, then $(u+v)^+ - (u+v)^- = u^+ - u^- + v^+ - v^-$ so that $(u+v)^+ + u^- + v^- = (u+v)^- + u^+ + v^+$. Hence by the additivity of L^+ on $\mathcal{C}^+$,

$$L^+((u+v)^+) + L^+(u^-) + L^+(v^-) = L^+((u+v)^-) + L^+(u^+) + L^+(v^+),$$

and thus changing back the order of the terms yields $L^+(u + v) = L^+(u) + L^+(v)$.

Step 4: *L^+ is continuous on $\mathcal{C}_0(\Omega)$.* Let $u \in \mathcal{C}_0(\Omega)$. Since L^+ is positive, then $L^+(|u| \pm u) \geqslant 0$, hence, by additivity of L^+, $L^+(|u|) \geqslant \pm L^+(u)$, *i.e.*, $|L^+(u)| \leqslant L^+(|u|)$. Let now u_1 and $u_2 \in \mathcal{C}_0(\Omega)$, then by steps 3 and 1,

$$|L^+(u_1) - L^+(u_2)| = |L^+(u_1 - u_2)| \leqslant L^+(|u_1 - u_2|) \leqslant \|L\| \|u_1 - u_2\|_\infty.$$

Step 5: *L^+ is a linear form on $\mathcal{C}_0(\Omega)$.* The additivity of L^+ shows that for all $n \in \mathbb{N}$, $L^+(nu) = nL^+(u)$. But since $(-u)^\pm = u^\mp$, then $L^+(-u) = -L^+(u)$ and the previous identity extends to any $n \in \mathbb{Z}$. Now if $r = p/q \in \mathbb{Q}$ with $p, q \in \mathbb{Z}$ and $q \neq 0$, then $L^+(qru) = qL^+(ru) = L^+(pu) = pL^+(u)$, hence $L^+(ru) = rL^+(u)$. The continuity of L^+ and the density of $\mathbb{Q}$ in $\mathbb{R}$ implies that $L^+(\alpha u) = \alpha L^+(u)$ for all $\alpha \in \mathbb{R}$.

Step 6: *L^- is a positive linear form on $\mathcal{C}_0(\Omega)$.* Define $L^- := L^+ - L$. Then L^- is clearly a linear form. Moreover, since by definition of L^+, $L^+(u) \geqslant L(u)$ for all $u \in \mathcal{C}^+$, then L^- is positive. □

Remark 6.3.1. We use the notation

$$L(u) = \int_\Omega u \, d\lambda,$$

with the *signed measure* $\lambda := \lambda^+ - \lambda^-$.

Definition 6.3.2 (Weak* convergence in $\mathcal{M}(\Omega)$). *We say that a sequence $(\lambda_n) \subset \mathcal{M}(\Omega)$ converges weakly* to λ in $\mathcal{M}(\Omega)$ if for any v in $\mathcal{C}_0(\Omega)$,*

$$\int_\Omega v \, d\lambda_n \to \int_\Omega v \, d\lambda$$

as $n \to \infty$.

Since, by Corollary 2.4.4, the space $\mathcal{C}_c(\Omega)$ is separable, it follows that $\mathcal{C}_0(\Omega)$ is separable as well. Hence, from Theorem 5.2.2 that any bounded sequence of bounded Radon measures has a subsequence which converges weakly* to a bounded Radon measure.

If μ is a positive Radon measure, we observe that the space $L^1(\Omega, \mu)$ can be injected into $\mathcal{M}(\Omega)$ through the map

$$u \in L^1(\Omega, \mu) \mapsto Tu \in \mathcal{M}(\Omega),$$

with

$$Tu : v \in \mathcal{C}_0(\Omega) \mapsto \int_\Omega uv \, d\mu.$$

As a consequence if (u_n) is a uniformly bounded sequence in $L^1(\Omega, \mu)$, then one can extract a subsequence weakly* converges in $\mathcal{M}(\Omega)$ to a bounded Radon measure, *i.e.*, there exist $(u_{n_k}) \subset (u_n)$ and $\lambda \in \mathcal{M}(\Omega)$ such that for every $v \in \mathcal{C}_0(\Omega)$,

$$\int_\Omega u_{n_k} v \, d\mu \to \int_\Omega v \, d\lambda.$$

The following result gives a complete characterization of sequences which are weakly converging in $L^1(\Omega, \mu)$.

Theorem 6.3.2 (Dunford-Pettis). *Assume that Ω is an open subset of $\mathbb{R}^N$ such that $\mu(\Omega) < \infty$. Let (u_n) be a bounded sequence in $L^1(\Omega, \mu)$.*

(i) If $u_n \rightharpoonup f$ weakly in $L^1(\Omega, \mu)$ for some $f \in L^1(X, \mu)$, then the sequence (f_n) is equi-integrable.

(ii) If (f_n) is equi-integrable, then there exist a subsequence $(f_{n_j}) \subset (f_n)$ and $f \in L^1(\Omega, \mu)$ such that $f_{n_j} \rightharpoonup f$ weakly in $L^1(\Omega, \mu)$.

Note that the equi-integrability property (see Definition 3.4.2) ensures that the sequence does concentrate on sets of arbitrarily small measure. If $\mu(\Omega) = \infty$ one must further ensure that the mass of (u_n) does not go to infinity. So in addition to (i), one must further require that for each $\varepsilon > 0$ there exists a compact set E_ε such that

$$\sup_{n \in \mathbb{N}^*} \int_{\Omega \setminus E_\varepsilon} |u_n| \, d\mu < \varepsilon.$$

Chapter 7

Hilbert analysis

A Hilbert space, named after David Hilbert, is a vector space possessing the structure of an inner product which is complete for the norm associated to its inner product. It generalizes the notion of Euclidean space. In particular the Pythagorean theorem and parallelogram law hold true in a Hilbert space.

7.1 Inner product space

In the real case an inner product on a a vector space is a positive definite, symmetric bilinear form on $X \times X \to \mathbb{R}$. In the complex case it is positive definite, Hermitian symmetric, sesquilinear form $X \times X \to \mathbb{C}$.

Definition 7.1.1 (Inner product). *Let X be a vector space on $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We say that a map $(\cdot, \cdot)$ from $X \times X$ to K is a inner product if*

1. $\forall u,\ v,\ w \in X$, $\forall \alpha, \beta$ *in* K,
$$(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w), \tag{7.1}$$
2. $\forall u,\ v \in X$, $(u, v) = \overline{(v, u)}$,
3. $\forall u \in X \setminus \{0\}$, $(u, u) > 0$.

Endowed with $(\cdot, \cdot)$, X is a inner product space (or pre-Hilbert space).

Lemma 7.1.1 (Cauchy-Schwarz inequality). *Let X be a pre-Hilbert space. Then*

$$\forall u, v \in X,\ |(u, v)| \leqslant \sqrt{(u, u)}\sqrt{(v, v)}. \tag{7.2}$$

Proof. Let $P(\lambda) := (u + \lambda v, u + \lambda v)$ for every $\lambda \in \mathbb{R}$. Since $P \geqslant 0$, the discriminant of the quadratic equation $P(\lambda) = 0$ is nonpositive. This yields the Cauchy-Schwarz inequality. □

An inner product gives rise to a norm. An inner product space is thus a special case of a normed linear space. A complete inner product space is a Hilbert space, a special case of a Banach space.

Lemma 7.1.2. *Let X be a pre-Hilbert space. Then the map $u \in X \mapsto \sqrt{(u, u)}$ defined a norm on X.*

Proof. The main point is to prove the triangle inequality, what can be done thanks to the Cauchy-Schwarz inequality. □

The polarization identity expresses the norm of an inner product space in terms of the inner product. For real inner product spaces it is

$$(u, v) = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2).$$

For complex spaces it is

$$(u, v) = \frac{1}{4}(\|u + v\|^2 + i\|u + iv\|^2 - \|u - v\|^2 - i\|u - iv\|^2).$$

In inner product spaces we also have the parallelogram law:

$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

This gives a criterion for a normed space to be an inner product space. Any norm coming from an inner product satisfies the parallelogram law and, conversely, if a norm satisfies the parallelogram law, we can show (but not so easily) that the polarization identity defines an inner product, which gives rise to the norm.

Lemma 7.1.3. *Let X be a pre-Hilbert space. Then X is smooth and uniformly convex. In addition the scalar product is a continuous bilinear mapping from $X \times X$ to $\mathbb{C}$.*

Proof. We first observe that for any $u \in X \setminus \{0\}$, for any $v \in X$, we have

$$\|u + \varepsilon v\|^2 = \|u\|^2 + 2\varepsilon \Re(u, v) + \varepsilon^2 \|v\|^2,$$

so that the mapping $\varepsilon \mapsto \|u + \varepsilon v\|^2$ is differentiable at 0 and thus

$$D\| \cdot \|(u) \cdot v = \Re(\|u\|^{-1} u, v).$$

That X is uniformly convex follows from the parallelogram equality.

Finally, in order to prove that the scalar product is continuous from $X \times X$ to $\mathbb{C}$, we first notice that Minkowski's inequality yields the continuity of the norm. Then the polarization identities allow to conclude. □

Definition 7.1.2 (Orthonormal sequence). *A family $(e_i)_{i \in I}$ in X is said an orthonormal sequence if for any $i, j \in I$, $(e_i, e_j) = \delta_{i,j}$.*

Lemma 7.1.4. *Let X be a inner product space and $(e_n)_{n \in \mathbb{N}}$ be an orthonormal sequence in X. Then, for any $u \in X$,*

1. *for any $k \in \mathbb{N}$,*
$$\|u\|^2 = \sum_{n=0}^{k} |(u, e_n)|^2 + \|u - \sum_{n=0}^{k} (u, e_n) e_n\|^2,$$

2. **Bessel's inequality:**
$$\sum_{n=0}^{\infty} |(u, e_n)|^2 \leqslant \|u\|^2.$$

Proof. We have

$$\begin{aligned}
\|u - \sum_{n=0}^{k} (u, e_n) e_n\|^2 &= (u - \sum_{n=0}^{k} (u, e_n) e_n, u - \sum_{n=0}^{k} (u, e_n) e_n) \\
&= \|u\|^2 - 2\Re(u, \sum_{n=0}^{k} (u, e_n) e_n) + \sum_{0 \leqslant n, m \leqslant k} (u, e_n) \overline{(u, e_m)} (e_m, e_n) \\
&= \|u\|^2 - 2 \sum_{n=0}^{k} |(u, e_n)|^2 + \sum_{0 \leqslant n \leqslant k} |(u, e_n)|^2 \\
&= \|u\|^2 - \sum_{n=0}^{k} |(u, e_n)|^2,
\end{aligned}$$

what yields (1). To obtain (2) it suffices to pass to the limit in the inequality

$$\sum_{n=0}^{k} |(u, e_n)|^2 \leqslant \|u\|^2.$$

□

Remark 7.1.1. Note that in general the sequence $u - \sum_{n=0}^{k}(u, e_n)e_n$ does not converge to 0 when n goes to infinity.

Remark 7.1.2. Let $\mathcal{E}$ be an orthonormal set of arbitrary cardinality. It follows from Bessel's inequality that for $\epsilon > 0$ and $u \in X$, $\{ e \in \mathcal{E} \,:\, (u, e) \geq \epsilon \}$ is finite, and hence that $\{ e \in \mathcal{E} \,:\, (u, e) > 0 \}$ is countable. We can thus extend Bessel's inequality to an arbitrary orthonormal set:

$$\sum_{e \in \mathcal{E}} (u, e)^2 \leqslant \|u\|^2,$$

where the sum is just a countable sum of positive terms.

7.2 Hilbert spaces

Definition 7.2.1 (Hilbert space)**.** *A Hilbert space is an inner product space which is complete.*

Theorem 7.2.1. *Let X be a Hilbert space and $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in X. The series $\sum_{n\in\mathbb{N}} \alpha_n e_n$ converges in X if and only if the sequence $(\alpha_n)_{n\in\mathbb{N}}$ belongs to $\ell^2(\mathbb{N})$. Moreover when the series $\sum_{n\in\mathbb{N}} \alpha_n e_n$ converges in X, then*

$$\left\| \sum_{n\in\mathbb{N}} \alpha_n e_n \right\|^2 = \sum_{\in\mathbb{N}} |\alpha_n|^2. \tag{7.3}$$

Proof. Since the space X is complete, the series $\sum_{n\in\mathbb{N}} \alpha_n e_n$ converges in X if and only if it satisfies the Cauchy property. But according to the Pythagore equality, there holds for any $n, m \in \mathbb{N}$ with $n > m$,

$$\left\| \sum_{k=m}^{n} \alpha_k e_k \right\|^2 = \sum_{k=m}^{n} |\alpha_k|^2.$$

Therefore $\sum_{n\in\mathbb{N}} \alpha_n e_n$ satisfies the Cauchy property in X if and only if the series $\sum_{k=0}^{n} |\alpha_k|^2$ satisfies the Cauchy property in $\mathbb{R}$. Since $\mathbb{R}$ is also complete, this yields the first part of the result. To prove (7.3) it is sufficient to use the continuity of the norm and the Pythagore equality:

$$\left\| \sum_{n\in\mathbb{N}} \alpha_n e_n \right\|^2 = \left\| \lim_{k\to\infty} \sum_{n=0}^{k} \alpha_n e_n \right\|^2 = \lim_{k\to\infty} \left\| \sum_{n=0}^{k} \alpha_n e_n \right\|^2 = \lim_{k\to\infty} \sum_{n=0}^{k} |\alpha_n|^2 = \sum_{n\in\mathbb{N}} |\alpha_n|^2.$$

$\square$

Combining Lemma 7.1.4 (2) and Theorem 7.2.1 we obtain the following result.

Corollary 7.2.1. *Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence in a Hilbert space X, and let $u \in X$. Then the series*

$$\sum_{n\in\mathbb{N}} (u, e_n) e_n$$

converges in X.

Given an orthonormal sequence $(e_n)_{n\in\mathbb{N}}$ in X, we define the linear mapping

$$\Phi : u \in X \mapsto ((u, e_n))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}).$$

Notice that the range of Φ is contained in $\ell^2(\mathbb{N})$ according to (1). Combining Lemma 7.1.3 and Corollary 7.2.1 we get that if $u \in X$ satisfies $u = \sum_{n\in\mathbb{N}} \alpha_n e_n$, then $\Phi(u) = (\alpha_n)_{n\in\mathbb{N}}$.

Definition 7.2.2 (Hilbert basis). *An orthonormal sequence $(e_n)_{n\in\mathbb{N}}$ in X is said total or a Hilbert basis if Φ is injective, i.e., if $(u, e_n) = 0$ for every $n \in \mathbb{N}$ implies that $u = 0$.*

Let us observe that it follows from the continuity of the scalar product that for every $u \in X$, $\Phi(u) = \Phi(\sum_{n\in\mathbb{N}} (u, e_n) e_n)$.

Theorem 7.2.2. *Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal family in an Hilbert space X. Then the following statements are equivalent:*

(i) The family $(e_n)_{n\in\mathbb{N}}$ is total;

(ii) For every $u \in X$, $u = \sum_{n\in\mathbb{N}}(u, e_n)e_n$;

(iii) For every $u \in X$, $\|u\|^2 = \sum_{n\in\mathbb{N}} |(u, e_n)|^2$.

Proof. Let us first assume (i). Since Φ is injective by assumption, (ii) follows. Let us now assume that (ii) holds. Then, using Pythagore's equality, we obtain that for every $u \in X$,

$$\|u\|^2 = \left\|\lim_{k\to\infty}\sum_{n=0}^{k}(u, e_n)e_n\right\|^2 = \lim_{k\to\infty}\left\|\sum_{n=0}^{k}(u, e_n)e_n\right\|^2 = \lim_{k\to\infty}\sum_{n=0}^{k}|(u, e_n)|^2 = \sum_{n\in\mathbb{N}}|(u, e_n)|^2.$$

Finally we assume that (iii) holds. Then, if $u \in X$ is such that $(u, e_n) = 0$ for every $n \in \mathbb{N}$ then clearly $u = 0$. □

As a consequence of Zorn's lemma, every Hilbert space admits an orthonormal basis; furthermore, any two orthonormal bases of the same space have the same cardinality, called the Hilbert dimension of the space.

7.3 Projection on a closed convex set

An essential property of Hilbert space is that the distance of a point to a closed convex set is alway attained.

Theorem 7.3.1. *Let X be a Hilbert space, K a closed convex subset, and $u \in X$. Then there exists a unique $\bar{u} \in K$ such that*

$$\|u - \bar{u}\| = \inf_{v\in K}\|u - v\|.$$

Moreover $\bar{u}$ is the unique element of K which satisfies $\Re(u - \bar{u}, v - \bar{u}) \leqslant 0$ for any $v \in K$.

Proof. Translating, we may assume that $u = 0$, and so we must show that there is a unique element of K of minimal norm. Let $d = \inf_{v\in K}\|v\|$ and chose $u_n \in K$ with $\|u_n\| \to d$. Then the parallelogram law gives

$$\left\|\frac{u_n - u_m}{2}\right\|^2 = \frac{1}{2}\|u_n\|^2 + \frac{1}{2}\|u_m\|^2 - \left\|\frac{u_n + u_m}{2}\right\|^2 \leqslant \frac{1}{2}\|u_n\|^2 + \frac{1}{2}\|u_m\|^2 - d^2,$$

where we have used convexity to infer that $(u_n + u_m)/2 \in K$. Thus (u_n) is a Cauchy sequence and so has a limit $\bar{u}$, which must belong to K, since K is closed. Since the norm is continuous, $\|\bar{u}\| = \lim_n \|u_n\| = d$.

For uniqueness, note that if $\|\bar{u}\| = \|\tilde{u}\| = d$, then $\|(\bar{u} + \tilde{u})/2\| = d$ and the parallelogram law gives

$$\|\bar{u} - \tilde{u}\|^2 = 2\|\bar{u}\|^2 + 2\|\tilde{u}\|^2 - \|\bar{u} + \tilde{u}\|^2 = 2d^2 + 2d^2 - 4d^2 = 0.$$

Let now prove the characterization of $\bar{u}$ through obtuse angles. Let v be in K, $\lambda \in (0, 1)$ and let $z := (1 - \lambda)\bar{u} + \lambda v$ which is in K by convexity. Therefore

$$\|u - \bar{u}\|^2 \leqslant \|u - z\|^2 = \|(u - \bar{u}) - \lambda(v - \bar{u})\|^2 = \|u - \bar{u}\|^2 + 2\lambda\Re(u - \bar{u}, \bar{u} - v) + \lambda^2\|v - \bar{u}\|^2.$$

Thus,

$$2\Re(u - \bar{u}, v - \bar{u}) - \lambda\|v - \bar{u}\|^2 \leqslant 0$$

and then, by letting λ tend to 0^+, we obtain that $\Re(u - \bar{u}, v - \bar{u}) \leqslant 0$ for any $v \in K$.

Conversely if $\bar{u}$ is an element of K which satisfies $\Re(u - \bar{u}, v - \bar{u}) \geqslant 0$ for any $v \in K$, then we have

$$\|(1 - \lambda)\bar{u} + \lambda v - u\|^2 \geqslant \|u - \bar{u}\|^2 + \lambda^2\|v - \bar{u}\|^2.$$

Letting λ goes to 1 yields $\|v - u\|^2 \geqslant \|v - \bar{u}\|^2$. □

The unique nearest element to u in K is often denoted $P_K u$, and referred to as the projection of u onto K. It satisfies $P_K \circ P_K = P_K$, the definition of a projection. This terminology is especially used when K is a closed linear subspace of X, in which case P_K is a linear projection operator.

Theorem 7.3.2. *Let X be a Hilbert space, Y a closed subspace, and $x \in X$. Then there exists a continuous linear mapping P_Y from X onto Y with $\|P_Y\| \leqslant 1$ such that for any $v \in Y$,*

$$\|u - P_Y u\| = \inf_{v \in Y} \|u - v\|.$$

Moreover $P_Y u$ is the unique element of Y which satisfies $(u - P_Y u, v) = 0$ for any $v \in Y$.

We say that P_Y is the orthogonal projection onto Y.

Proof. The existence of P_Y is given by the previous theorem. We now prove the characterization of $P_Y u$ as the unique element of Y which satisfies $(u - \bar{u}, v) = 0$ for any $v \in Y$. Using the characterization of the previous theorem with $v + P_Y u$ instead of v, we have that $P_Y u$ satisfies $\Re(u - P_Y u, v) \leqslant 0$ for any $v \in Y$. Using this last inequality with $-v$, iv and $-iv$ instead of v yields $(u - P_Y u, v) = 0$ for any $v \in Y$. The converse is straightforward: if an element $\bar{u}$ in Y satisfies $(u - \bar{u}, v) = 0$ for any $v \in Y$ then it satisfies the characterization of the previous theorem so it is the projection of u onto Y.

From this characterization we infer that P_Y is linear. Now to prove that P_Y is continuous with $\|P_Y\| \leqslant 1$ it suffices to apply the Cauchy-Schwarz inequality to the characterization. □

If S is any subset of a inner product space X, let

$$S^\perp = \{\, u \in X : \, (u, s) = 0 \text{ for all } s \in S \,\}.$$

Then $S^\perp$ is a closed subspace of X. We obviously have $S \cap S^\perp = 0$ and $S \subset S^{\perp\perp}$. Furthermore if $S_1 \subset S_2$ then $S_2^\perp \subset S_1^\perp$.

Lemma 7.3.1. *If X is a Hilbert space and S is a closed subspace of X, then $X = S \bigoplus S^\perp$.*

Proof. We have that $S \cap S^\perp = \{0\}$ since $u \in S \cap S^\perp$ implies $\|u\|^2 = (u, u) = 0$. In addition for any u in X, $u = P_S u + (u - P_S u)$ provides a decomposition in $S \bigoplus S^\perp$, according to the previous theorem. □

Corollary 7.3.1. *If X is a Hilbert space and S is a subspace of X, then $\overline{S} = X$ if and only if $S^\perp = \{0\}$.*

Proof. Suppose that $\overline{S} = X$ and let u be in $S^\perp$. Then there exists $(u_n)_n$ in S converging to u. For any n, we have $(u_n, u) = 0$, and since the scalar product is continuous, passing to the limit yields $\|u\|^2 = (u, u) = 0$.

Conversely if we assume now that $S^\perp = \{0\}$ then from the previous lemma applied to $\overline{S}$ we infer that $X = \overline{S} \bigoplus \overline{S}^\perp$. But $\overline{S}^\perp = \{0\}$ so that $\overline{S} = X$. □

7.4 Duality and weak convergence

The identification of the dual space of Hilbert spaces is easy.

Theorem 7.4.1 (Riesz Representation Theorem). *If X is a real Hilbert space, define $j : X \to X'$ by $j_y(x) = (x, y)$. This map is a linear isometry of X onto X'. For a complex Hilbert space it is a conjugate linear isometry (it satisfies $j_{\alpha y} = \bar{\alpha} j_y$).*

Proof. It is easy to see that j is an isometry of X into X' and the main issue is to show that any $f \in X'$ can be written as j_y for some y. We may assume that $f \neq 0$, so $\ker(f)$ is a proper closed subspace of X. Let $y_0 \in [\ker(f)]^\perp$ be of norm 1 and set $y = \overline{f(y_0)} y_0$. For all $x \in X$, we clearly have that $f(y_0)x - f(x)y_0 \in \ker(f)$, so

$$j_y(x) = (x, y) = (x, \overline{f(y_0)} y_0) = (f(y_0)x, y_0) = (f(x)y_0, y_0) = f(x).$$

□

Via the map j we can define an inner product on X', so it is again a Hilbert space. The Riesz map j is used to identify X and X' so that a sequence $(u_n)_n$ in X weakly converges to u if and only for any v in X, $(u_n, v) \to (u, v)$. Then as a consequence of Corollary 5.2.1 (respectively Theorem 5.2.2), we have the following results:

Proposition 7.4.1. *Let X be a Hilbert space and $(u_n)_n$ be a weakly converging sequence in X. Then $(u_n)_n$ is bounded.*

Theorem 7.4.2. *Let X be a Hilbert space and $(u_n)_n$ be a bounded sequence in X. Then there exists a subsequence $(u_{n_k})_k$ which weakly converges to some u in X.*

Proof. Let us introduce $Y := \overline{\text{Vect}\ (u_n)_{n\in\mathbb{N}}}$ which is, for the topology induced by X, a separable Hilbert space. Therefore since $(u_n)_n$ is a bounded sequence in Y, there exists a subsequence $(u_{n_k})_k$ which weakly converges to u in Y. Let us now consider the orthogonal projection P on Y. We have, for any v in X,

$$\begin{aligned}
(u_n, v) &= (u_n, Pv) + (u_n, (Id - P)v) \\
&= (u_n, Pv) \quad \text{since } (Id - P)v \in Y^\perp \\
&\to (u, Pv) \quad \text{when } n \to +\infty \\
& \quad = (u, v).
\end{aligned}$$

□

7.5 Convexity and optimization

Theorem 7.5.1 (Banach-Saks). *Let X be a Hilbert space and $(u_n)_n$ be a sequence in X which weakly converges to u. then there exists a subsequence $(u_{n_k})_k$ whose Cesaro means strongly converge to u, i.e.*

$$\frac{u_{n_0} + \cdots + u_{n_k}}{k+1} \to u \tag{7.4}$$

when $k \to +\infty$.

Proof. Without loss of generality we can assume that $u = 0$, otherwise we consider the sequence $(u_n - u)_n$ instead of $(u_n)_n$. We choose $n_0 = 0$. Let $k \geqslant 1$ and assume that $(n_j)_{0\leqslant j\leqslant k-1}$ have been determined. Since $(u_n)_n$ weakly converges to 0, for any j such that $0 \leqslant j \leqslant k-1$, there exists $n'_j \in \mathbb{N}$ such that

$$|(u_{n_j}, u_n)| \leqslant k^{-1} \tag{7.5}$$

for any $n \geqslant n'_j$. We set $n_k = \max(n_0, ..., n_{k-1}, n'_0, ..., n'_{k-1}) + 1$. Thus the sequence $(n_k)_k$ is increasing and for any $k \geqslant 1$, for any $j \in \mathbb{N}$ such that $0 \leqslant j < k$, $|(u_{n_j}, u_{n_k})| \leqslant k^{-1}$.

Let us now verify that the Cesaro means $(v_k := \sum_{j=0}^{k} \frac{u_{n_j}}{k+1})_k$ strongly converge to 0. To this purpose it is useful to note that, as a consequence of the Banach-Steinhaus theorem, we have that $\sup_{n\in\mathbb{N}} \|u_n\| \leqslant K$. Then

$$\begin{aligned}
\|v_k\|^2 &= (\sum_{j=0}^{k} \frac{u_{n_j}}{k+1}, \sum_{l=0}^{k} \frac{u_{n_l}}{k+1}) \\
&= \sum_{j=0}^{k} \frac{\|u_{n_j}\|^2}{(k+1)^2} + 2\Re(\sum_{j=1}^{k}\sum_{l=0}^{j-1} \frac{(u_{n_j}, u_{n_l})}{(k+1)^2}), \\
&\leqslant \frac{K^2}{k+1} + \frac{2}{(k+1)^2}\sum_{j=1}^{k}\sum_{l=0}^{j-1}\frac{1}{j} \\
&\leqslant \frac{K^2}{k+1} + \frac{2(k-1)}{(k+1)^2}
\end{aligned}$$

which converges to 0 when $k \to +\infty$. □

Theorem 7.5.2. *Let X be a Hilbert space, C be a convex closed subset of X. Then C is weakly closed in X. Moreover, if f is a continuous and convex function from C to $\mathbb{R}$ then f is weakly lower semi-continuous. As a consequence if C is bounded then f admits a minimun in C. If in addition f is strictly convex then this minimun is unique.*

Proof. We start by proving that C is weakly closed in X. Let us consider a sequence $(u_n)_n$ in C which weakly converges to u in X. We are going to prove that u is actually in C. Thanks to the Banach-Saks theorem, there exists a subsequence $(u_{n_k})_k$ whose Cesaro means $(v_k := \sum_{j=0}^{k} \frac{u_{n_j}}{k+1})_k$ strongly converge to u. Since C is convex, the v_k are also in C, and since C is closed, their limit, u, is also in C.

Let us now consider a continuous and convex function f from C to $\mathbb{R}$ and a sequence $(u_n)_n$ be a sequence in C which weakly converge to u. In order to prove that f is weakly lower semi-continuous, we will prove that

$$f(u) \leqslant \liminf f(u_n). \tag{7.6}$$

By definition of $\liminf$, there exists a subsequence (v_n) of (u_n) such that $\lim f(v_n) = \liminf f(u_n)$.

Of course the subsequence (v_n) also weakly converges to u so that using the Banach-Saks theorem, we obtain a subsequence (w_n) of (v_n) such that $(\sum_{j=0}^{k} \frac{w_j}{k+1})_k$ strongly converge to u. Since f is strongly continuous, the sequence $(f(\sum_{j=0}^{k} \frac{w_j}{k+1}))_k$ strongly converges to $f(u)$. Now f being convex, we have

$$f(\sum_{j=0}^{k} \frac{w_j}{k+1}) \leqslant \sum_{j=0}^{k} \frac{f(w_j)}{k+1}. \tag{7.7}$$

As a subsequence of (v_n), the sequence (w_n) satisfies $\lim f(w_n) = \liminf f(u_n)$ so that by the Cesaro theorem the right hand side of (7.7) converges to $\liminf f(u_n)$ so that passing to the limit in (7.7) yields the inequality (7.6).

Let us now assume that C is bounded so that if $(u_n)_n$ is a minimizing sequence of f in C then $(u_n)_n$ is bounded and thus it admits a subsequence weakly converging to some u. Then the (weak) lower semicontinuity of f implies that u is a minimizer of f over C. □

7.6 Spectral decomposition of symmetric compact operators

This section is devoted to the spectral decomposition of symmetric compact operators. This can be seen as an extension of the spectral decomposition of symmetric matrices to operators acting on infinite-dimensional spaces. As for matrices the first step is to prove the existence of an eigenvalue. Yet the classical argument which hinges on the D'Alembert-Gauss theorem, applied to the characteristic polynomial, is not available anymore. In infinite dimensions. It will be replaced by a variational argument. Actually we first give a minimization characterization of the operator norm.

The first step in this direction is the following general fact:

Proposition 7.6.1. *Let A in $\mathcal{L}_c(X)$. Then*

$$\|A\| = \sup_{\{u,v \in X:\ \|u\|=\|v\|=1\}} |(Au, v)|.$$

Proof. Thanks to the Cauchy-Schwarz inequality we have that for any $u, v \in X$,

$$|(Au, v)| \leqslant \|Au\| \|v\| \leqslant \|A\| \|u\| \|v\|$$

so that

$$\|A\| \geqslant \sup_{\{u,v \in X:\ \|u\|=\|v\|=1\}} |(Au, v)|.$$

Moreover $\|Au\| = |(Au, v)|$ with $v = \|Au\|^{-1} Au$ if $Au \neq 0$ or any v with $\|v\| = 1$ otherwise. Then

$$\|A\| := \sup_{\{u \in X: \|u\|=1\}} \|Au\| \leqslant \sup_{\{u,v \in X:\ \|u\|=\|v\|=1\}} |(Au, v)|.$$

□

Let us introduce now the notion of symmetry to go farther.

Definition 7.6.1. *Let X be an inner product space, and A in $\mathcal{L}_c(X)$. We say that A is symmetric if*

$$\forall u, v \in X, \quad (Au, v) = (u, Av). \tag{7.8}$$

Let us notice that for a symmetric operator A, (Au, u) is in $\mathbb{R}$ for any $u \in X$.

Lemma 7.6.1. *Let X be a pre-Hilbert space, and A in $\mathcal{L}_c(X)$ symmetric. Then*

$$\|A\| = \sup_{\{u \in X : \|u\|=1\}} |(Au, u)|. \tag{7.9}$$

Proof. Let us denote

$$\alpha := \sup_{\{u \in X:\ \|u\|=1\}} |(Au, u)| \text{ and } \beta := \sup_{\{u,v \in X:\ \|u\|=\|v\|=1\}} |(Au, v)|.$$

Taking into account Proposition 7.6.1 it is sufficient to prove that $\alpha \geqslant \beta$. Let u, v in X with $\|u\| = \|v\| = 1$. We can assume without any loss of generality that $(Au, v) > 0$, since if this is not satisfied, we can multiply v by a complex number of modulus one. We have

$$\begin{aligned}
(Au, v) &= \frac{1}{4}((A(u+v), u+v) - (A(u-v), u-v)), \\
&\leqslant \frac{1}{4}(|(A(u+v), u+v)| + |(A(u-v), u-v)|, \\
&\leqslant \frac{\alpha}{4}(\|u+v\|^2 + \|u-v\|^2), \\
&\leqslant \frac{\alpha}{2}(\|u\|^2 + \|v\|^2) = \alpha,
\end{aligned}$$

so that taking the sup over u, v in X with $\|u\| = \|v\| = 1$ completes the proof. □

Let us now turn our attention to the eigenvalue, whose definition is given now.

Definition 7.6.2 (Eigenvalue). *Let X be a normed vector space and A in $\mathcal{L}_c(X)$. We say that $\lambda \in \mathbb{C}$ is an eigenvalue of A if there exists $v \in X \setminus \{0\}$ such that $Av = \lambda v$. The set of all the eigenvalues of A is called the spectrum of A and is denoted $sp(A)$.*

Proposition 7.6.2. *Let X be a normed vector space and A in $\mathcal{L}_c(X)$. Then*

1. *for any $\lambda \in sp(A)$, then $|\lambda| \leqslant \|A\|$.*
2. *if X is inner product space and if A in $\mathcal{L}_c(X)$ is symmetric, then $sp(A) \subset \mathbb{R}$.*

Proof. First if $v \in X \setminus \{0\}$ is such that $Av = \lambda v$. Then $|\lambda|\|v\| = \|Av\| \leqslant \|A\|\|v\|$, so that $|\lambda| \leqslant \|A\|$. Now if X is an inner product space and A in $\mathcal{L}_c(X)$ is symmetric, then

$$\lambda(v, v) = (\lambda v, v) = (Av, v) = (v, Av) = (v, \lambda v) = \overline{\lambda}(v, v)$$

so that $\lambda = \overline{\lambda}$, that is $\lambda \in \mathbb{R}$. □

The following result says that among the extremal values authorized by the previous analysis, one of them is almost an eigenvalue.

Lemma 7.6.2. *Let X be an inner product space and if A in $\mathcal{L}_c(X)$ is symmetric, then there exists $\lambda_1 \in \{-\|A\|, \|A\|\}$ and a sequence $(u_n)_n$ in X, with $\|u_n\| = 1$ for any n, such that $\|Au_n - \lambda_1 u_n\| \to 0$ when $n \to \infty$.*

Proof. Since the case $A = 0$ is straightforward we assume that $A \neq 0$. Let $(u_n)_n$ be a maximizing sequence of $\|A\| = \sup_{\|u\|=1} |(Au, u)|$, that is a sequence $(u_n)_n$ in X, with $\|u_n\| = 1$ for any n, such that $|(Au_n, u_n)|$ converges to $\|A\|$. There exists a subsequence that we still denote $(u_n)_n$ such that $((Au_n, u_n))_n$ converges to $\lambda_1 \in \{-\|A\|, \|A\|\}$. Then

$$\begin{aligned}
\|Au_n - \lambda_1 u_n\|^2 = (Au_n - \lambda_1 u_n, Au_n - \lambda_1 u_n) &= \|Au_n\|^2 - 2\lambda_1 (Au_n, u_n) + \lambda_1^2 \\
&\leqslant \|A\|^2 - 2\lambda_1 (Au_n, u_n) + \lambda_1^2 \to 0.
\end{aligned}$$

□

To go farther we will require the following additional hypothesis:

Definition 7.6.3 (Compact operator). *Let X be a normed vector space. We say that A in $\mathcal{L}_c(X)$ is compact if the image of any bounded subset of X is a relatively compact subset of X.*

Let us stress that this is equivalent to require that the image of the unit ball in X under A is relatively compact in X or that for any bounded sequence in X the sequence of their image has a convergent subsequence.

Proposition 7.6.3 (First eigenvalue of a compact operator). *Let X be an inner product space and let A in $\mathcal{L}_c(X)$ be symmetric compact. Then there exists $\lambda_1 \in \{-\|A\|, \|A\|\}$ which is an eigenvalue of A.*

Proof. Thanks to Lemma 7.6.2 there exists $\lambda_1 \in \{-\|A\|, \|A\|\}$ and a sequence $(u_n)_n$ in X, with $\|u_n\| = 1$ for any n, such that $\|Au_n - \lambda_1 u_n\| \to 0$ when $n \to \infty$. Since A is compact, there exists a subsequence $(u_{n_k})_k$ of $(u_n)_n$ such that $(Au_{n_k})_k$ has a limit. These two last facts imply that $(u_{n_k})_k$ converges to some u in X, which is an eigenvector for λ_1. □

Remark 7.6.1. In finite dimensions it is possible to give a simple variational proof of Proposition 7.6.3: the closed unit ball of $\mathbb{R}^n$ is compact so that the continuous function $u \mapsto (Au, u)$ has a maximum, say u^*. By Lagrange's multiplier theorem (associated with the equality constraint $\|u\|^2 = 1$), it must satisfies $Au^* = \lambda u^*$, for some λ in $\mathbb{R}$.

Lemma 7.6.3. *Let X be an inner product space and A in $\mathcal{L}_c(X)$ symmetric. Let $e_1, \dots, e_n$ be eigenvectors of A. Then $(\text{Vect } (e_1, \dots, e_n))^\perp$ is stable by A.*

Proof. Let u in $(\text{Vect } (e_1, \dots, e_n))^\perp$ and $i \in \{1, \dots, n\}$. Then, using the fact that A is symmetric, we get

$$(Au, e_i) = (u, Ae_i) = (u, \lambda e_i) = \overline{\lambda}(u, e_i) = 0,$$

where λ is the eigenvalue associated to e_i. This proves that $(\text{Vect } (e_1, \dots, e_n))^\perp$ is stable by A. □

Now we proceed by iteration, by restriction to the orthogonal of the eigenspaces already determined.

Theorem 7.6.1 (Poincaré principle). *Let X be an inner product space, and $A \in \mathcal{L}_c(X)$ be a symmetric and compact linear map with an infinite dimensional range. Then there exists a sequence of eigenvalues $(\lambda_n)_{n\in\mathbb{N}^*}$ and a orthonormal family $(e_n)_{n\in\mathbb{N}^*}$ of eigenvectors of A such that*

1. *the sequence $(|\lambda_n|)_{n\in\mathbb{N}^*}$ is strictly positive and decreasing,*
2. *the sequence $(\lambda_n)_{n\in\mathbb{N}^*}$ tends to 0,*
3. *we have, for any $n \in \mathbb{N}^*$, that*

$$|\lambda_n| = \sup_{u\in F_n} |(Au, u)| \text{ where } F_n := \{u \in X : \ \|u\| = 1, (u, e_1) = \cdots = (u, e_{n-1}) = 0\}.$$

Proof. We already have the existence of λ_1 and e_1 thanks to Proposition 7.6.3. Let us now assume that $n \geqslant 2$ and that we have constructed λ_1,..., λ_{n-1} and $e_1, \dots, e_{n-1}$. Let us denote

$$Y_{n-1} := (\text{Vect } (e_1, \dots, e_{n-1}))^\perp. \tag{7.10}$$

According to Lemma 7.6.3 the subspace Y_{n-1} is stable by A. Thanks to Proposition 7.6.3 the restriction of A to Y_{n-1} admits an eigenvalue λ_n such that

$$|\lambda_n| = \|A|_{Y_{n-1}}\| = \sup_{\{u\in Y_{n-1}:\ \|u\|_{Y_{n-1}}=1\}} |(A|_{Y_{n-1}}u, u)| = \sup_{u\in F_{n-1}} |(Au, u)|, \tag{7.11}$$

and an associated eigenvector e_n of norm one. This eigenvalue λ_n does not vanish since A has an infinite dimensional range. Since e_n is in Y_{n-1}, we have $(e_n, e_i) = 0$ for $i = 1, \dots, n-1$. Since the sequence of sets Y_k decreases with k, we obtain that the sequence $(|\lambda_k|)_k$ is decreasing. It only remains to prove that the sequence $(\lambda_n)_{n\in\mathbb{N}^*}$ tends to 0. Let us assume by contradiction that the sequence $(|\lambda_n^{-1}|)_{n\in\mathbb{N}^*}$ is bounded. Then, since A is compact, the sequence $(e_n = A(\lambda_n^{-1}e_n))_{n\in\mathbb{N}^*}$ should has a convergent subsequence. But this is impossible since $\|e_n - e_m\| = \sqrt{2}$ for any $n \neq m$. □

The spectral theorem also provides a decomposition, called the spectral decomposition, of the underlying vector space on which the operator acts.

Theorem 7.6.2 (Spectral decomposition). *1. Let X be an inner product space, and $A \in \mathcal{L}_c(X)$ be a symmetric and compact linear map with an infinite dimensional range. Then for any u in X,*

$$Au = \sum_{n \geqslant 1} \lambda_n (u, e_n) e_n.$$

2. If X is a Hilbert space, and $A \in \mathcal{L}_c(X)$ be a symmetric and compact linear map, then there exists an orthonormal basis of X consisting of eigenvectors of A. More specifically, the orthogonal complement of the kernel of A admits, either a finite orthonormal basis of eigenvectors of a countably infinite orthonormal basis $(e_n)_{n\in\mathbb{N}^*}$ of eigenvectors of A with corresponding eigenvalues $(|\lambda_n|)_{n\in\mathbb{N}^*}$ strictly positive and decreasing to 0 when n goes to infinity.

3. If moreover X is separable, then there exists a countable orthonormal basis of X consisting of eigenvectors of A.

Proof. Let u be in X. We have, for any $k \geqslant 2$, by definition of Y_k, that $u - \sum_{n=1}^{k-1} (u, e_n) e_n$ is in Y_{k-1}. Therefore

$$\left\| Au - \sum_{n=1}^{k-1} \lambda_n (u, e_n) e_n \right\| = \left\| A(u - \sum_{n=1}^{k-1} (u, e_n) e_n) \right\| \leqslant \|A|_{Y_{k-1}}\| \left\| u - \sum_{n=1}^{k-1} (u, e_n) e_n) \right\|.$$

But using (7.11), we have $\|A_{Y_{k-1}}\| = |\lambda_k|$, and on the other hand Pythagore's theorem provides $\|u - \sum_{n=1}^{k-1} (u, e_n) e_n)\| \leqslant \|u\|$, so that

$$\|Au - \sum_{n=1}^{k-1} \lambda_n (u, e_n) e_n\| \leqslant |\lambda_k| \|u\| \to 0.$$

Hence for any u in X,

$$Au = \sum_{n \geqslant 1} \lambda_n (u, e_n) e_n.$$

Now if X is a Hilbert space we can use Corollary 7.2.1 to get that $\sum_{n \geqslant 1} (u, e_n) e_n$ converges in X. Using that A is continuous, we obtain

$$A(\sum_{n \geqslant 1} (u, e_n) e_n) = \sum_{n \geqslant 1} (u, e_n) A(e_n) = \sum_{n \geqslant 1} \lambda_n (u, e_n) e_n = Au.$$

Therefore $u - \sum_{n \geqslant 1} (u, e_n) e_n$ is in the kernel of A. But the kernel of A is a closed subspace of X, it is therefore a Hilbert space and admits (under Zorn's lemma) a Hilbert basis. Putting together this basis and the family $(e_n)_{n\in\mathbb{N}^*}$ yields a Hilbert basis consisting of eigenvectors of A. If moreover X is separable, then this basis of X is countable. The case where the range of A has a finite dimension is even more simple. □

There are many extensions of Theorem 7.6.1 and 7.6.2: first one can drop the hypothesis that the operator is compact, assuming only that the operator is continuous. Such operators may have no eigenvalues, the way the operator can be diagonalized is therefore to be understood as a unitary conjugation to a multiplication operator.

Let us first sketch this point of view on the previous case of compact operators. Suppose even, for concreteness, that X is an infinite dimensional separable Hilbert space and that $(e_n)_{n\in\mathbb{N}^*}$ is an orthonormal basis of X consisting of eigenvectors of a compact symmetric operator A. Then the map $U : X \mapsto \ell^2$ given by

$$U \left(\sum_{n\in\mathbb{N}^*} c_n e_n \right) = (c_n)_{n\in\mathbb{N}^*},$$

is an isometric isomorphism. Moreover when we use this map to transfer the action of T to ℓ^2, *i.e.*, when we consider the operator UTU^{-1} on ℓ^2, we see that this operator is simply multiplication by the bounded sequence $(\lambda_n)_{n\in\mathbb{N}^*}$.

The spectral theorem also holds for normal operators on a Hilbert space (let us recall that an operator is said normal if it commutes with its hermitian adjoint).

The spectral theorem can even be extended for (self-adjoint or normal) unbounded operators, such as differential operators. To give an example, any constant coefficient differential operator is unitarily equivalent to a multiplication operator. Indeed the unitary operator that implements this equivalence is the Fourier transform.

Finally let us mention that spectral theory also deals with operators acting on Banach spaces.

Chapter 8

Fourier series

8.1 Functions on the torus

Let X be any space. A function $u : \mathbb{R} \to X$ is said to be T-periodic ($T > 0$) if

$$u(x + T) = u(x)$$

for all $x \in \mathbb{R}$. If u is T-periodic, then it can be seen as a function defined on the torus $\mathbb{T}$ which is defined as the quotient space $\mathbb{R}/(T\mathbb{Z})$, naturally endowed of group and complete metric space structures.

For matters connected to integration theory, it is convenient to identify $\mathbb{T}$ to the interval $[-T/2, T/2)$ (modulo the choice of a cut point), this bijection being bi-continuous. Then we have

$$\mathcal{C}^k(\mathbb{T}; \mathbb{C}) \simeq \{u \in \mathcal{C}^k(\mathbb{R}; \mathbb{C}) : u \text{ is } T\text{-periodic}\},$$

for $k \in \mathbb{N}$, and

$$L^p(\mathbb{T}; \mathbb{C}) \simeq \{u \in L^p([-T/2, T/2); \mathbb{C}) \text{ extended to } \mathbb{R} \text{ by } T\text{-periodicity}\},$$

for $1 \leqslant p \leqslant \infty$, where the symbol $\simeq$ stands for canonical bijection.

In the sequel, we will only consider the case $T = 1$.

8.2 Fourier coefficients of $L^1(\mathbb{T}; \mathbb{C})$-functions

Definition 8.2.1. *Let $u \in L^1(\mathbb{T}; \mathbb{C})$ and $k \in \mathbb{Z}$. We define the k-th Fourier coefficient of u as the complex number*

$$\hat{u}(k) := \int_{-1/2}^{1/2} u(s) e^{-2i\pi ks} ds.$$

The sequence $(\hat{u}(k))_{k \in \mathbb{Z}}$ is called the sequence of Fourier coefficients of u, and the series

$$\sum_{k \in \mathbb{Z}} \hat{u}(k) e^{2i\pi kt}$$

is the Fourier series of u.

The goal of this chapter is to establish a Fourier inversion formula, *i.e.*,

$$u(t) = \sum_{k \in \mathbb{Z}} \hat{u}(k) e^{2i\pi kt}.$$

To this aim, we need to answer to two questions: on the one hand, in what sense does the Fourier series converge (pointwise, uniformly, integral, ...); and on the other hand, is the Fourier series equal (and in what sense) to the initial function u?

Let us start with the first question. We observe that the sequence $(e^{2i\pi k\cdot})_{k\in\mathbb{Z}}$ is orthonormal in $L^2(\mathbb{T};\mathbb{C})$. Indeed, for every integers $k \neq j$,

$$\begin{aligned}
\left(e^{2i\pi k\cdot}, e^{2i\pi j\cdot}\right)_{L^2([-1/2,1/2))} &= \int_{-1/2}^{1/2} e^{2i\pi kt}\overline{e^{2i\pi jt}}\, dt \\
&= \int_{-1/2}^{1/2} e^{2i\pi(k-j)t}\, dt \\
&= \frac{1}{2i\pi(k-j)} \left[e^{2i\pi(k-j)t}\right]_{-1/2}^{1/2} = 0,
\end{aligned}$$

while if $k \in \mathbb{Z}$,

$$\left(e^{2i\pi k\cdot}, e^{2i\pi k\cdot}\right)_{L^2([-1/2,1/2))} = \int_{-1/2}^{1/2} e^{2i\pi 0t}\, dt = 1.$$

Moreover, since for any $t \in [-1/2, 1/2)$, $\left|e^{2i\pi kt}\right| = 1$, the norm of $e^{2i\pi k\cdot}$ in $\mathcal{C}(\mathbb{T};\mathbb{C})$ is equal to 1. From the results obtained in chapters 1 and 7, we deduce the following theorem.

Theorem 8.2.1. *Let $(\alpha_k)_{k\in\mathbb{Z}}$ be a sequence of complex numbers. Then*

1. *If $\sum_{k\in\mathbb{Z}} |\alpha_k| < \infty$, then the series $\sum_{k\in\mathbb{Z}} \alpha_k e^{2i\pi kt}$ converges (uniformly) in $\mathcal{C}(\mathbb{T};\mathbb{C})$.*
2. *The series $\sum_{k\in\mathbb{Z}} \alpha_k e^{2i\pi kt}$ converges in $L^2(\mathbb{T};\mathbb{C})$ if and only if $(\alpha_k)_{k\in\mathbb{Z}}$ is in $\ell^2(\mathbb{Z};\mathbb{C})$.*

Proof. The first item is a consequence of the fact that $\mathcal{C}(\mathbb{T};\mathbb{C})$ is a Banach space, while the second one is an application of Theorem 7.2.1 in Chapter 7 together with the fact that $L^2(\mathbb{T};\mathbb{C})$ is a Hilbert space. □

8.3 Fourier inversion formula

Let us come to the second problematic, namely the equality between the Fourier series of a function (when it converges) and the function itself. The following result gives a criterion.

Theorem 8.3.1. *If $u \in \mathcal{C}(\mathbb{T};\mathbb{C})$ is such that $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$, then the Fourier series $\sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt}$ converges uniformly to u in $\mathcal{C}(\mathbb{T};\mathbb{C})$.*

Proof. Fist of all, Theorem 8.2.1 ensures that the series $\sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt}$ converges uniformly over $\mathbb{T}$. Thus, it remains to identify the limit with u. We observe that by definition of $\hat{u}(k)$,

$$\sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt} = \sum_{k\in\mathbb{Z}} \int_{-1/2}^{1/2} u(s)e^{2i\pi k(t-s)}\, ds.$$

A priori, it is not possible to exchange the integral and the sum because the series $\sum_{k\in\mathbb{Z}} u(s)e^{2i\pi k(t-s)}$ diverges at each points s where $u(s) \neq 0$. For $0 < r < 1$, let us introduce the functions

$$U_r(t) := \sum_{k\in\mathbb{Z}} \int_{-1/2}^{1/2} u(s)r^{|k|}e^{2i\pi k(t-s)} ds.$$

For the same reasons than before, the series in the right hand side of the previous equality converges uniformly. Moreover,

$$\sum_{k\in\mathbb{Z}} \int_{-1/2}^{1/2} |u(s)|r^{|k|}|e^{2i\pi k(t-s)}|\, ds \leqslant \|u\|_\infty \sum_{k\in\mathbb{Z}} r^{|k|} = \|u\|_\infty \left(\frac{2}{1-r} - 1\right).$$

By Lebesgue's dominated convergence theorem, we can now commute integral and series (this is why we plugged the extra r) and get

$$U_r(t) = \int_{-1/2}^{1/2} u(s)P_r(t-s)\, ds,$$

where

$$P_r(\tau) := \sum_{k\in\mathbb{Z}} r^{|k|} e^{2i\pi k\tau}$$

for any $\tau \in [-1/2, 1/2)$. Observe that P_r can simply be written as

$$\begin{aligned}
P_r(\tau) &= \sum_{k\in\mathbb{N}} r^k e^{2i\pi k\tau} + \sum_{k\in\mathbb{N}} r^k e^{-2i\pi k\tau} - 1 \\
&= 2\mathrm{Re}\left(\sum_{k\in\mathbb{N}} r^k e^{2i\pi k\tau}\right) - 1 = 2\mathrm{Re}\sum_{k\in\mathbb{N}} (re^{2i\pi\tau})^k - 1 \\
&= 2\mathrm{Re}\left(\frac{1}{1-re^{2i\pi\tau}}\right) - 1 = \frac{1-r^2}{1-2r\cos(2\pi\tau)+r^2}
\end{aligned}$$

so that

$$0 < P_r(\tau) < \frac{2}{1-r}$$

for every $\tau \in [-1/2, 1/2)$ and any $r \in (0,1)$. Note also that

$$\int_{-1/2}^{1/2} P_r(\tau)\, d\tau = \sum_{k\in\mathbb{Z}} r^{|k|} \int_{-1/2}^{1/2} e^{2i\pi k\tau}\, d\tau = 1 \tag{8.1}$$

because all terms in the previous are zero except that for $k = 0$, and for each $\delta > 0$,

$$\sup_{\tau\in[-1/2,1/2)\setminus[-\delta,\delta]} P_r(\tau) \leqslant \frac{1-r^2}{1-2r\cos(2\pi\delta)+r^2} \tag{8.2}$$

tends to zero as $r \to 1$. Consequently, thanks to (8.1), one can write

$$\begin{aligned}
|U_r(t) - u(t)| &= \left|\int_{-1/2}^{1/2} (u(t-s) - u(t)) P_r(s)\, ds\right| \\
&\leqslant \left|\int_{-\delta}^{\delta} (u(t-s) - u(t)) P_r(s)\, ds\right| + \left|\int_{[-1/2,1/2)\setminus[-\delta,\delta]} (u(t-s) - u(t)) P_r(s)\, ds\right| \\
&\leqslant \sup_{x,y\in\mathbb{T},\, |x-y|\leqslant\delta} |u(x) - u(y)| + 2\|u\|_\infty \int_{[-1/2,1/2)\setminus[-\delta,\delta]} P_r(s)\, ds,
\end{aligned}$$

and we deduce from (8.2) that

$$\limsup_{r\to 1^-} \|U_r - u\|_\infty \leqslant \sup_{x,y\in\mathbb{T},\, |x-y|\leqslant\delta} |u(x) - u(y)|.$$

Since the left hand side of the previous inequality is independent of δ, we can take the limit as $\delta \to 0$ in the right hand side, and we finally obtain

$$\lim_{r\to 1^-} \|U_r - u\|_\infty = 0$$

which means that U_r uniformly converges to u on $[-1/2, 1/2)$.

We now show the uniform convergence of U_r to U_1 (*i.e.* the Fourier series of u), from which, by uniqueness of the limit, the proof will be complete. Let $\varepsilon > 0$, since $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$, there exists $N \in \mathbb{N}$ such that

$$\sum_{|k|\geqslant N} |\hat{u}(k)| < \frac{\varepsilon}{2}.$$

On the other hand,

$$\lim_{r\to 1^-} \sum_{|k|<N} \hat{u}(k) r^{|k|} e^{2i\pi kt} = \sum_{|k|<N} \hat{u}(k) e^{2i\pi kt},$$

the convergence being uniform for $t \in \mathbb{T}$. Hence,

$$\limsup_{r\to 1^-} \|U_r - U_1\|_\infty \leqslant \limsup_{r\to 1^-} \left\| \sum_{|k|\geqslant N} \hat{u}(k)(1 - r^{|k|})e^{2i\pi kt} \right\| \leqslant 2 \sum_{|k|\geqslant N} |\hat{u}(k)| < \varepsilon.$$

Since ε is arbitrary, the conclusion follows. □

The previous result is applicable provided the condition $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$ is satisfied. In particular, it holds for functions smooth enough as the next proposition shows.

Proposition 8.3.1. *Let $u \in \mathcal{C}^n(\mathbb{T}; \mathbb{C})$ for some $n \in \mathbb{N}$. Then for every $j \in \{0, \dots, n\}$ and every $k \in \mathbb{Z}$,*

$$\widehat{u^{(j)}}(k) = (2i\pi k)^j \hat{u}(k),$$

where $u^{(j)}$ denotes the j-th derivative of u. Consequently if $u \in \mathcal{C}^2(\mathbb{T}; \mathbb{C})$ then $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$.

Proof. Since $u^{(i)} = (u^{(j-1)})'$ for $j \geqslant 1$, it is enough to consider the case $j = 1$, proceeding by induction for the other cases. Let us assume that $u \in \mathcal{C}^1(\mathbb{T}; \mathbb{C})$ for some $n \in \mathbb{N}$. As $u' \in \mathcal{C}(\mathbb{T}; \mathbb{C}) \subset L^1(\mathbb{T}; \mathbb{C})$, we have

$$\widehat{u'}(k) = \int_{-1/2}^{1/2} u'(s)e^{-2i\pi ks}\, ds,$$

and an integration by parts ensures that

$$\widehat{u'}(k) = -\int_{-1/2}^{1/2} u(s)(-2i\pi k)e^{-2i\pi ks}\, ds = (2i\pi k)\hat{u}(k),$$

where we used the fact that the function $s \mapsto u(s)e^{-2i\pi ks}$ is 1-periodic.

If $u \in \mathcal{C}^2(\mathbb{T}; \mathbb{C})$, we get that for each $k \in \mathbb{Z} \setminus \{0\}$,

$$\hat{u}(k) = -\frac{1}{4\pi^2 k^2}\widehat{u''}(k)$$

so that

$$|\hat{u}(k)| \leqslant \frac{1}{4\pi^2 k^2} \int_{-1/2}^{1/2} |u''(s)||e^{-2i\pi ks}|\, ds \leqslant \frac{\|u''\|_\infty}{4\pi^2 k^2}$$

and thus $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$. □

We will now extend the Fourier inversion formula to $L^2(\mathbb{T}; \mathbb{C})$. Let us recall that since $\{e^{2i\pi kt}\}_{k\in\mathbb{Z}}$ is an orthonormal family, Theorem 7.2.2 of Chapter 7 gives us some equivalences from which we deduce the following result.

Theorem 8.3.2. *The Fourier transform $\Phi : L^2(\mathbb{T}; \mathbb{C}) \to \ell^2(\mathbb{T}; \mathbb{C})$, $u \mapsto (\hat{u}(k))_{k\in\mathbb{Z}}$ is an isometrical isomorphism (a linear one to one mapping preserving the norm). In other words, there holds the Parseval identity*

$$\|u\|_2^2 = \sum_{k\in\mathbb{Z}} |\hat{u}(k)|^2$$

for every $u \in L^2(\mathbb{T}; \mathbb{C})$, and

$$u(t) = \sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt}$$

for almost every $t \in [-1/2, 1/2)$, the series in the right hand side converging in the sense of the $L^2(\mathbb{T}; \mathbb{C})$-norm.

Proof. It follows from Theorem 8.3.1 and the second part of Proposition 8.3.1 that for each $u \in \mathcal{C}^2(\mathbb{T}; \mathbb{C})$,

$$u(t) = \sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt}$$

for every $t \in [-1/2, 1/2)$, where the series is converging uniformly over $[-1/2, 1/2)$ (and thus also in the sense of the $L^2(\mathbb{T};\mathbb{C})$-norm because $[-1/2, 1/2)$ has finite Lebesgue measure). As a consequence, the restriction of the linear mapping from $L^2(\mathbb{T};\mathbb{C})$ to $L^2(\mathbb{T};\mathbb{C})$ defined by

$$u \mapsto \sum_{k\in\mathbb{Z}} \hat{u}(k) e^{2i\pi kt}$$

to $\mathcal{C}^2(\mathbb{T};\mathbb{C})$ is equal to the identity. As $\mathcal{C}^2(\mathbb{T};\mathbb{C})$ is dense in $L^2(\mathbb{T};\mathbb{C})$, the uniqueness of the continuous extension ensures that

$$u(t) = \sum_{k\in\mathbb{Z}} \hat{u}(k) e^{2i\pi kt}$$

for every $u \in L^2(\mathbb{T};\mathbb{C})$. In the previous equality, the series is converging in the sense of the $L^2(\mathbb{T};\mathbb{C})$-norm, and the functions u and $t \mapsto \sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt}$ coincide as $L^2(\mathbb{T};\mathbb{C})$ functions, *i.e.*, almost everywhere in $[-1/2, 1/2)$.

Since

$$\hat{u}(k) = \int_{-1/2}^{1/2} u(s)\overline{e^{2i\pi ks}}ds = (u, e^{2i\pi ks})_{L^2(\mathbb{T};\mathbb{C})},$$

it follows from Theorem 7.2.2 that Φ is an isometry which is nothing else than Parseval identity. □

To conclude this chapter, let us mention two easy applications of the previous theorem and Theorem 8.3.1.

8.4 Functional inequalities

In a broader framework, both following functional inequalities play an instrumental role in the analysis of partial differential equations.

Theorem 8.4.1 (Poincaré-Wirtinger inequality). *Let $u \in \mathcal{C}^1(\mathbb{T};\mathbb{C})$ with zero average (i.e. such that $\int_{-1/2}^{1/2} u(s)ds = 0$). Then*

$$\|u\|_{L^2(\mathbb{T};\mathbb{C})} \leqslant \frac{1}{2\pi}\|u'\|_{L^2(\mathbb{T};\mathbb{C})}.$$

Proof. The zero average condition implies that $\hat{u}(0) = 0$ (it is actually an equivalence). Since $\mathcal{C}(\mathbb{T};\mathbb{C}) \subset L^2(\mathbb{T};\mathbb{C})$, applying the Parseval equality to u and u' together with Proposition 8.3.1, we get that

$$\begin{aligned}
\|u\|^2_{L^2(\mathbb{T};\mathbb{C})} &= \sum_{k\in\mathbb{Z}} |\hat{u}(k)|^2 = \sum_{k\in\mathbb{Z}\setminus\{0\}} |\hat{u}(k)|^2 \\
&= \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{4\pi^2k^2}|\widehat{u'}(k)|^2 \leqslant \frac{1}{4\pi^2}\sum_{k\in\mathbb{Z}\setminus\{0\}} |\widehat{u'}(k)|^2 \\
&\leqslant \frac{1}{4\pi^2}\|u'\|^2_{L^2(\mathbb{T};\mathbb{C})}.
\end{aligned}$$

□

Theorem 8.4.2 (Sobolev inequality). *Let $u \in \mathcal{C}^2(\mathbb{T};\mathbb{C})$ with zero average (i.e. such that $\int_{-1/2}^{1/2} u(s)ds = 0$). Then*

$$\|u\|_{L^\infty(\mathbb{T};\mathbb{C})} \leqslant \frac{1}{\sqrt{12}}\|u'\|_{L^2(\mathbb{T};\mathbb{C})}.$$

Proof. Since $u \in \mathcal{C}^2(\mathbb{T};\mathbb{C})$, using Proposition 8.3.1, $\sum_{k\in\mathbb{Z}} |\hat{u}(k)| < \infty$, and thanks to Theorem 8.3.1 we get that

$$\sum_{k\in\mathbb{Z}} \hat{u}(k)e^{2i\pi kt} = u(t)$$

for every $t \in [-1/2, 1/2)$, where the series in the left hand side is converging uniformly. Therefore

$$|u(t)| \leqslant \sum_{k\in\mathbb{Z}} |\hat{u}(k)|.$$

As before, the average condition implies that $\hat{u}(0) = 0$, and by Proposition 8.3.1 and the Cauchy-Schwarz inequality, we get

$$|u(t)| \leqslant \sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{2\pi k}|\widehat{u'}(k)| \leqslant C\sqrt{\sum_{k\in\mathbb{Z}\setminus\{0\}} |\widehat{u'}(k)|^2},$$

with $C := \sqrt{\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{1}{(2\pi k)^2}} = \frac{1}{\sqrt{12}}$. The Parseval equality yields in turn that

$$|u(t)| \leqslant \frac{1}{\sqrt{12}}\|u'\|_{L^2(\mathbb{T};\mathbb{C})}$$

for every $t \in [-1/2, 1/2)$. $\square$

8.5 Adaptation for T-periodic functions

The results obtained so far in the previous sections can be extended to T-periodic functions with $T > 0$ non necessarily equal to 1. Indeed if $u = u(t)$ is T-periodic, then the function $v(t) := u(t/T)$ becomes 1-periodic. We thus work on the function v and once the conclusion is obtained, we rewrite everything in terms of u. The family $\{e^{2i\pi kt}\}_{k\in\mathbb{Z}}$ is then the transformed of the family $\{e^{\frac{2i\pi kt}{T}}\}_{k\in\mathbb{Z}}$ on $[-T/2, T/2)$. When $T \to \infty$, the frequency $\{k/T\}_{k\in\mathbb{Z}}$ tend to become dense in the real line, and formally, for $T = \infty$, (*i.e.* functions on $\mathbb{R}$ without any periodicity) one should consider the uncountable family $\{e^{2i\pi st}\}_{s\in\mathbb{R}}$. This is the object of the next chapter where we will make this argument rigorous.

Chapter 9

Fourier transform of integrable and square integrable functions

Fourier transform describes which frequencies are present in a complex-valued function of real variables. It therefore could be thought as an extension of the theory of Fourier series to non periodic functions. In this chapter we will study the Fourier transform of L^1 and L^2 functions.

9.1 Fourier transform of integrable functions

Definition 9.1.1. *Let $u \in L^1(\mathbb{R}^N;\mathbb{C})$. The Fourier transform $\hat{u}$ of u is the function defined for $\xi \in \mathbb{R}^N$ by*

$$\hat{u}(\xi) := \int_{\mathbb{R}^N} u(x)e^{-2i\pi x\cdot\xi}\, dx,$$

where $x \cdot y$ denotes the scalar product of x and y in $\mathbb{R}^N$.

We observe that this pointwise definition of $\hat{u}$ makes sense everywhere thanks to the comparison Theorem.

Proposition 9.1.1. *The Fourier transform is a continuous linear map from $L^1(\mathbb{R}^N;\mathbb{C})$ to $L^\infty(\mathbb{R}^N;\mathbb{C})$.*

Proof. The linearity follows from that of the Lebesgue integral. For the continuity we have for every $u \in L^1(\mathbb{R}^N;\mathbb{C})$ and any $\xi \in \mathbb{R}^N$,

$$|\hat{u}(\xi)| \leqslant \int_{\mathbb{R}^N} |u(x)||e^{-2i\pi x\cdot\xi}|\, dx = \|u\|_1,$$

so that $\|\hat{u}\|_\infty \leqslant \|u\|_1$. $\square$

If we formally derive $\hat{u}$ with respect to ξ_i, with $i \in \{1,\dots,N\}$, we get that

$$\frac{\partial\hat{u}}{\partial\xi_i}(\xi) = -2i\pi\int_{\mathbb{R}^N} x_i u(x)e^{-2i\pi x\cdot\xi}dx.$$

This computation is justified provided $x \mapsto x_i u(x)$ belongs to $L^1(\mathbb{R}^N;\mathbb{C})$. In the same way in order to compute $\frac{\partial^2\hat{u}}{\partial\xi_i^2}(\xi)$ we need to assume that $x \mapsto x_i^2 u(x) \in L^1(\mathbb{R}^N;\mathbb{C})$. This suggests that the regularity of $\hat{u}$ is connected to the decreasing character of u at infinity.

Conversely, if u is smooth enough, we can write for $\xi_i \neq 0$,

$$\hat{u}(\xi) = \int_{\mathbb{R}^N} u(x)e^{-2i\pi x\cdot\xi}\, dx = \frac{1}{2i\pi\xi_i}\int_{\mathbb{R}^N} \frac{\partial u}{\partial x_i}(x)e^{-2i\pi x\cdot\xi}\, dx.$$

This suggests in turn that the decreasing character of $\hat{u}$ is closely related to the regularity of u. We introduce the *Schwartz space* of all infinitely differentiable functions that are rapidly decreasing at infinity along with all partial derivatives.

Definition 9.1.2 (Schwartz space). *A function $u : \mathbb{R}^N \to \mathbb{C}$ belongs to the Schwartz space, denoted by $\mathcal{S}(\mathbb{R}^N)$, if any of its partial derivatives decreases at infinity more rapidly than any polynomial. More precisely, $u \in \mathcal{S}(\mathbb{R}^N)$ if and only if for any multi-indexes α and $\beta \in \mathbb{N}^N$, there exists a constant $C = C(u, \alpha, \beta) > 0$ such that*

$$\sup_{x \in \mathbb{R}^N} |x^\beta \partial^\alpha u(x)| \leqslant C.$$

We recall that for all multi-index $\alpha := (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$, we denote

$$\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \cdots \partial {x_N}^{\alpha_N}}, \quad x^\alpha := x_1^{\alpha_1} \cdots x_N^{\alpha_N},$$

where $|\alpha| := \alpha_1 + \cdots + \alpha_N$ is the length of α. Denoting $\alpha! := \alpha_1! \cdots \alpha_N!$, we get the Leibnitz' rule

$$\partial^\alpha (u\, v) = \sum_{\beta + \gamma = \alpha} \frac{\alpha!}{\beta! \gamma!} \partial^\beta u \, \partial^\gamma v$$

for any u, $v \in \mathcal{C}^\infty(\mathbb{R}^N; \mathbb{C})$ and for all $\alpha \in \mathbb{N}^N$.

The following result can be easily checked by the reader.

Proposition 9.1.2. *For all u, $v \in \mathcal{S}(\mathbb{R}^N)$, all $\alpha \in \mathbb{N}^N$ and all $P \in \mathbb{C}[X_1, \dots, X_N]$, then uv, $\partial^\alpha u$ and Pu belong to $\mathcal{S}(\mathbb{R}^N)$.*

Let us also notice that $\mathcal{C}_c^\infty(\mathbb{R}^N; \mathbb{C}) \subset \mathcal{S}(\mathbb{R}^N)$, and that $x \mapsto e^{-|x|^2} \in \mathcal{S}(\mathbb{R}^N) \setminus \mathcal{C}_c^\infty(\mathbb{R}^N; \mathbb{C})$.

The Schwartz space is a good framework to justify the above formal computations. Namely

Proposition 9.1.3. *For all $u \in \mathcal{S}(\mathbb{R}^N)$ and all $\alpha \in \mathbb{N}^N$,*

$$\begin{aligned} \partial^\alpha \hat{u}(\xi) &= \widehat{(-2i\pi x)^\alpha u}(\xi), \\ \widehat{\partial^\alpha u}(\xi) &= (2i\pi\xi)^\alpha \hat{u}(\xi), \end{aligned}$$

for any $\xi \in \mathbb{R}^N$.

Proof. As mentioned before, it suffices in the first case to derive the definition of $\hat{u}$, and in the second case to integrate by parts. In both cases, the fact that $u \in \mathcal{S}(\mathbb{R}^N)$ enables to justify the operation (domination property, and vanishing boundary term). □

Let us recall that $\mathcal{C}_0(\mathbb{R}^N; \mathbb{C})$ denotes the space of all functions vanishing at infinity, that is the closure of $\mathcal{C}_c(\mathbb{R}^N; \mathbb{C})$ in $\mathcal{C}_b(\mathbb{R}^N; \mathbb{C})$ (see Proposition 6.3.1 for a characterization of that space). We can now extend Proposition 9.1.1 as follows.

Theorem 9.1.1 (Riemann-Lebesgue). *The Fourier transform is a continuous linear mapping from $L^1(\mathbb{R}^N; \mathbb{C})$ to $\mathcal{C}_0(\mathbb{R}^N; \mathbb{C})$.*

Proof. Let $u \in L^1(\mathbb{R}^N; \mathbb{C})$. By density there exists a sequence $(u_n)_{n \in \mathbb{N}} \subset \mathcal{C}_c^\infty(\mathbb{R}^N; \mathbb{C}) \subset \mathcal{S}(\mathbb{R}^N)$ which converges to u in $L^1(\mathbb{R}^N; \mathbb{C})$. Using Proposition 9.1.1, we deduce that $\hat{u}_n \to \hat{u}$ in $L^\infty(\mathbb{R}^N; \mathbb{C})$. Thanks to Propositions 6.3.1 and 9.1.3, for each $n \in \mathbb{N}$, one has $\hat{u}_n \in \mathcal{C}_0(\mathbb{R}^N; \mathbb{C})$. Indeed if $\varepsilon > 0$, and $\alpha \in \mathbb{N}^N$ is such that $|\alpha| = 1$, then

$$|\hat{u}_n(\xi)| \leqslant \frac{1}{2\pi|\xi|} \|\widehat{\partial^\alpha u_n}\|_\infty \leqslant \frac{1}{2\pi|\xi|} \|\partial^\alpha u_n\|_1 < \varepsilon$$

for every $\xi \notin K_\varepsilon$, where $K_\varepsilon := \{\xi \in \mathbb{R}^N : |\xi| \leqslant 2\pi/\varepsilon \|\partial^\alpha u_n\|_1\}$ is a compact set. Hence, since $\mathcal{C}_0(\mathbb{R}^N; \mathbb{C})$ is closed with respect to the uniform convergence, we get that $\hat{u} \in \mathcal{C}_0(\mathbb{R}^N; \mathbb{C})$. □

The next result will be useful in the sequel.

Proposition 9.1.4. *If u and $v \in L^1(\mathbb{R}^N; \mathbb{C})$, then*

$$\int_{\mathbb{R}^N} \hat{u} v \, dx = \int_{\mathbb{R}^N} u \hat{v} \, dx.$$

Proof. We first remark that if u and $v \in L^1(\mathbb{R}^N;\mathbb{C})$, then by the Riemann-Lebesgue Theorem (Theorem 9.1.1), $\hat{u}$ and $\hat{v} \in \mathcal{C}_0(\mathbb{R}^N;\mathbb{C})$ so that $\hat{u}v$ and $u\hat{v} \in L^1(\mathbb{R}^N;\mathbb{C})$ by Hölder's inequality, and the above integrals are well defined. We next use the Fubini and Tonelli Theorems to get that

$$\begin{aligned}\int_{\mathbb{R}^N} \hat{u}(x)v(x)\,dx &= \int_{\mathbb{R}^N}\left(\int_{\mathbb{R}^N} u(y)e^{-2i\pi x\cdot y}\,dy\right)v(x)\,dx\\ &= \int_{\mathbb{R}^N}\left(\int_{\mathbb{R}^N} v(x)e^{-2i\pi x\cdot y}\,dx\right)u(y)\,dy\\ &= \int_{\mathbb{R}^N}\hat{v}(y)u(y)\,dy.\end{aligned}$$

□

The Fourier transform enjoys some good properties with respect to the groups of translations and dilatations.

Definition 9.1.3. *If $u : \mathbb{R}^N \to \mathbb{C}$, for $a \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}\setminus\{0\}$, we define the translation of u by a, and the dilatation of u by λ as*

$$(\tau_a u)(x) := u(x-a),$$
$$(\delta_\lambda u)(x) := u\left(\tfrac{x}{\lambda}\right),$$

for every $x \in \mathbb{R}^N$.

Lemma 9.1.1. *If $u \in L^1(\mathbb{R}^N;\mathbb{C})$, $a \in \mathbb{R}^N$ and $\lambda \in \mathbb{R}\setminus\{0\}$, we have*

$$\widehat{\tau_a u}(\xi) = e^{-2i\pi a\cdot\xi}\hat{u}(\xi),$$
$$\widehat{\delta_\lambda u}(\xi) = |\lambda|^N\delta_{1/\lambda}\hat{u}(\xi),$$

for every $\xi \in \mathbb{R}^N$.

Proof. By definition,

$$\begin{aligned}\widehat{\tau_a u}(\xi) &= \int_{\mathbb{R}^N} u(x-a)e^{-2i\pi x\cdot\xi}\,dx = \int_{\mathbb{R}^N} u(y)e^{-2i\pi(y+a)\cdot\xi}\,dy\\ &= e^{-2i\pi a\cdot\xi}\int_{\mathbb{R}^N} u(y)e^{-2i\pi y\cdot\xi}\,dy = e^{-2i\pi a\cdot\xi}\hat{u}(\xi).\end{aligned}$$

Similarly

$$\begin{aligned}\widehat{\delta_\lambda u}(\xi) &= \int_{\mathbb{R}^N} u\left(\frac{x}{\lambda}\right)e^{-2i\pi x\cdot\xi}\,dx = |\lambda|^N\int_{\mathbb{R}^N} u(y)e^{-2i\pi(\lambda y)\cdot\xi}\,dy\\ &= |\lambda|^N\int_{\mathbb{R}^N} u(y)e^{-2i\pi y\cdot(\lambda\xi)}\,dy = |\lambda|^N\hat{u}(\lambda\xi)\\ &= |\lambda|^N\delta_{1/\lambda}\hat{u}(\xi).\end{aligned}$$

□

Corollary 9.1.1. *The function $u : x \mapsto e^{-\pi|x|^2}$ is kept invariant through the Fourier transform.*

Proof. Indeed, let $\alpha \in \mathbb{N}^N$ be a multi-index with length $|\alpha| = 1$. Thanks to the properties of the exponential, we have

$$\partial^\alpha u = (-2\pi x)^\alpha u.$$

Since $u \in \mathcal{S}(\mathbb{R}^N)$, we can take the Fourier transform on both sides of the previous equality and get, taking into account Proposition 9.1.3, $(2i\pi\xi)^\alpha\hat{u} = (-i\partial)^\alpha\hat{u}$ so that

$$\partial^\alpha\hat{u} = (-2\pi\xi)^\alpha\hat{u}.$$

Hence

$$\partial^\alpha\left(\frac{\hat{u}}{u}\right)=\frac{(\partial^\alpha \hat{u})u-(\partial^\alpha u)\hat{u}}{u^2}=0.$$

Since α is any multi-index of length 1, we infer that $\frac{\hat{u}}{u}$ is constant. As

$$\left(\frac{\hat{u}}{u}\right)(0)=\frac{\hat{u}(0)}{u(0)}=\hat{u}(0)=\int_{\mathbb{R}^N} e^{-\pi|x|^2}\,dx.$$

Now to calculate the integral above, we observe that, thanks to the Fubini principle, we have

$$\int_{\mathbb{R}^N} e^{-\pi|x|^2}\,dx=(\int_{\mathbb{R}} e^{-\pi x^2}\,dx)^N=(\int_{\mathbb{R}^2} e^{-\pi|x|^2}\,dx)^{\frac{N}{2}}=1,$$

by using the polar coordinates. The conclusion follows. □

If $t>0$, we deduce from Lemma 9.1.1 and Corollary 9.1.1 that

$$\widehat{e^{-\pi t^2|x|^2}}=\delta_{1/t}\widehat{(e^{-\pi|x|^2})}=t^{-N}e^{-\pi|x|^2/t^2}.$$

We now arrive to the main result of this section which is the analogue of Theorem 8.3.1 in chapter 8.

Theorem 9.1.2 (Fourier inversion formula). *Let $u\in L^1(\mathbb{R}^N;\mathbb{C})\cap\mathcal{C}_b(\mathbb{R}^N;\mathbb{C})$ be such that $\hat{u}\in L^1(\mathbb{R}^N;\mathbb{C})$. Then, for every $x\in\mathbb{R}^N$, $\hat{\hat{u}}(x)=u(-x)$.*

Proof. By definition,

$$\hat{\hat{u}}(x)=\int_{\mathbb{R}^N}\hat{u}(\xi)e^{-2i\pi\xi\cdot x}\,d\xi.$$

Unfortunately, it is not possible to apply Proposition 9.1.4 because the function $\xi\mapsto e^{-2i\pi\xi\cdot x}$ does not belong to $L^1(\mathbb{R}^N;\mathbb{C})$. However, since $\hat{u}\in L^1(\mathbb{R}^N;\mathbb{C})$, it follows from Lebesgue's dominated convergence Theorem that

$$\hat{\hat{u}}(x)=\lim_{t\to 0^+}\int_{\mathbb{R}^N}\hat{u}(\xi)e^{-2i\pi\xi\cdot x}e^{-\pi t^2|\xi|^2}\,d\xi$$

because $e^{-\pi t^2|\xi|^2}\to 1$ pointwise as $t\to 0^+$, and $|\hat{u}(\xi)e^{-2i\pi\xi\cdot x}e^{-\pi t^2|\xi|^2}|\leqslant|\hat{u}(\xi)|$ for every $\xi\in\mathbb{R}^N$, with $|\hat{u}|\in L^1(\mathbb{R}^N)$. Now since $\xi\mapsto e^{-\pi t^2|\xi|^2}\in L^1(\mathbb{R}^N;\mathbb{C})$, we immediately obtain from Lemma 9.1.1 and Proposition 9.1.4 that

$$\begin{aligned}\int_{\mathbb{R}^N}\hat{u}(\xi)e^{-2i\pi\xi\cdot x}e^{-\pi t^2|\xi|^2}\,d\xi &= \int_{\mathbb{R}^N}\widehat{\tau_x u}(\xi)e^{-\pi t^2|\xi|^2}\,d\xi\\ &= \int_{\mathbb{R}^N}\tau_x u(\xi)\widehat{e^{-\pi t^2|\xi|^2}}\,d\xi\\ &= \int_{\mathbb{R}^N}u(\xi-x)t^{-N}e^{-\pi|\xi|^2/t^2}\,d\xi\\ &= \int_{\mathbb{R}^N}u(ty-x)e^{-\pi|y|^2}\,dy.\end{aligned}$$

Consequently,

$$\hat{\hat{u}}(x)=\lim_{t\to 0^+}\int_{\mathbb{R}^N}u(ty-x)e^{-\pi|y|^2}\,dy=u(-x)\int_{\mathbb{R}^N}e^{-\pi|y|^2}\,dy=u(-x),$$

where we used once more Lebesgue's dominated convergence Theorem which is licit since

$$|u(ty-x)e^{-\pi|y|^2}|\leqslant\|u\|_\infty e^{-\pi|y|^2}$$

and $y\mapsto e^{-\pi|y|^2}\in L^1(\mathbb{R}^N)$. □

Corollary 9.1.2. *The Fourier transform is a linear one to one map from $\mathcal{S}(\mathbb{R}^N)$ to $\mathcal{S}(\mathbb{R}^N)$.*

Proof. Let $u \in \mathcal{S}(\mathbb{R}^N)$. Let us show that for any multi-indexes α and $\beta \in \mathbb{N}^N$, one has

$$\sup_{\xi \in \mathbb{R}^N} |\xi^\beta \partial^\alpha \hat{u}(\xi)| < \infty.$$

Thanks to Proposition 9.1.3, $\partial^\alpha \hat{u}(\xi) = \widehat{(-2i\pi x)^\alpha u}(\xi)$ and thus

$$\begin{aligned} \xi^\beta \partial^\alpha \hat{u}(\xi) &= \xi^\beta \widehat{(-2i\pi x)^\alpha u}(\xi) \\ &= \frac{1}{(2i\pi)^{|\beta|}} (2i\pi\xi)^\beta \widehat{(-2i\pi x)^\alpha u}(\xi) \\ &= \frac{1}{(2i\pi)^{|\beta|}} \widehat{\partial^\beta((-2i\pi x)^\alpha u)}. \end{aligned}$$

Since $u \in \mathcal{S}(\mathbb{R}^N)$, by Proposition 9.1.2, $x \mapsto \partial^\beta((-2i\pi x)^\alpha u(x)) \in \mathcal{S}(\mathbb{R}^N)$ as well, and in particular, it also belongs to $L^1(\mathbb{R}^N;\mathbb{C})$. As a consequence of the Riemann-Lebesgue Theorem (Theorem 9.1.1), we infer that $\widehat{\partial^\beta((-2i\pi x)^\alpha u)} \in L^\infty(\mathbb{R}^N;\mathbb{C})$. We thus proved that if $u \in \mathcal{S}(\mathbb{R}^N)$, then $\hat{u} \in \mathcal{S}(\mathbb{R}^N)$. Therefore, we can apply the Fourier inversion formula in $\mathcal{S}(\mathbb{R}^N)$ to get that

$$\hat{\hat{\hat{\hat{u}}}} = u \quad \text{for all } u \in \mathcal{S}(\mathbb{R}^N),$$

and consequently $u \mapsto \hat{u}$ is one to one. $\square$

Remark 9.1.1. According to Theorem 9.1.2 and Corollary 9.1.2, it follows that the inverse Fourier transform can be written as

$$u(x) = \int_{\mathbb{R}^N} \hat{u}(\xi) e^{2i\pi x\cdot\xi}\, d\xi$$

for all $u \in \mathcal{S}(\mathbb{R}^N)$.

One advantage of the Fourier transform is that it maps a convolution product into a usual product of functions.

Corollary 9.1.3. *If u and $v \in \mathcal{S}(\mathbb{R}^N)$, then $u * v \in \mathcal{S}(\mathbb{R}^N)$ and*

(i) $\widehat{u * v} = \hat{u}\hat{v}$;

(ii) $\hat{u} * \hat{v} = \widehat{uv}$.

Proof. We start by proving (ii). By definition of the convolution product, we have thanks to Proposition 9.1.4, Lemma 9.1.1 and Theorem 9.1.2,

$$\begin{aligned} (\hat{u} * \hat{v})(z) &= \int_{\mathbb{R}^N} \hat{u}(z-y)\hat{v}(y)\, dy = \int_{\mathbb{R}^N} (\tau_z \delta_{-1} \hat{u})(y)\hat{v}(y)\, dy \\ &= \int_{\mathbb{R}^N} (\widehat{\tau_z \delta_{-1} \hat{u}})(y) v(y)\, dy = \int_{\mathbb{R}^N} e^{-2i\pi y\cdot z} \widehat{\delta_{-1}\hat{u}}(y) v(y)\, dy \\ &= \int_{\mathbb{R}^N} e^{-2i\pi y\cdot z} u(y)v(y)\, dy = \widehat{uv}(z). \end{aligned}$$

Above we used that $\widehat{\delta_{-1}\hat{u}} = u$, which is obtained by applying the Fourier inversion formula to $\hat{u}$.

By invertibility of the Fourier transform on $\mathcal{S}(\mathbb{R}^N)$, $u = \hat{f}$ and $v = \hat{g}$ for some f and $g \in \mathcal{S}(\mathbb{R}^N)$. Consequently, using (ii), we infer that $\widehat{u * v} = \widehat{\hat{f} * \hat{g}} = \widehat{\widehat{fg}} = (\delta_{-1} f)(\delta_{-1} g) = \hat{u}\hat{v}$ because, thanks to Theorem 9.1.2, $\delta_{-1} f = \hat{\hat{f}} = \hat{u}$ and similarly for g. Hence we get (i).

Thanks to (i), we have $u * v = \delta_{-1}\widehat{\hat{u}\hat{v}}$. Since $\mathcal{S}(\mathbb{R}^N)$ is stable with respect to the Fourier transform, the multiplication, and obviously δ_{-1}, it follows that $u * v \in \mathcal{S}(\mathbb{R}^N)$ which completes the proof. $\square$

Corollary 9.1.4. *If $u \in \mathcal{S}(\mathbb{R}^N)$, then*

$$\|\hat{u}\|_2 = \|u\|_2.$$

Proof. Indeed, we have thanks to Proposition 9.1.4 and the Fourier inversion formula (Theorem 9.1.2),

$$\begin{aligned}
\|\hat{u}\|_2^2 &= \int_{\mathbb{R}^N} \hat{u}(\xi)\overline{\hat{u}(\xi)}\, d\xi \\
&= \int_{\mathbb{R}^N} \hat{u}(\xi)\overline{\int_{\mathbb{R}^N} u(x)e^{-2i\pi x\cdot\xi}\, dx}\, d\xi \\
&= \int_{\mathbb{R}^N} \hat{u}(\xi)\int_{\mathbb{R}^N} \overline{u}(x)e^{-2i\pi x\cdot(-\xi)}\, dx\, d\xi \\
&= \int_{\mathbb{R}^N} \hat{u}(\xi)\widehat{\overline{u}}(-\xi)\, d\xi = \int_{\mathbb{R}^N} \hat{u}(\xi)\widehat{\delta_{-1}\overline{u}}(\xi)\, d\xi \\
&= \int_{\mathbb{R}^N} u(\xi)\widehat{\widehat{\delta_{-1}\overline{u}}}(\xi)\, d\xi = \int_{\mathbb{R}^N} u(\xi)\delta_{-1}\delta_{-1}\overline{u}(\xi)\, d\xi \\
&= \int_{\mathbb{R}^N} u(\xi)\overline{u}(\xi)\, d\xi = \|u\|_2^2.
\end{aligned}$$

□

We will next extend the Fourier transform from $L^1(\mathbb{R}^N;\mathbb{C})$ to $L^2(\mathbb{R}^N;\mathbb{C})$. Let us remark that neither $L^2(\mathbb{R}^N;\mathbb{C}) \not\subset L^1(\mathbb{R}^N;\mathbb{C})$ nor $L^1(\mathbb{R}^N;\mathbb{C}) \not\subset L^2(\mathbb{R}^N;\mathbb{C})$.

9.2 Fourier transform of L^2 functions

Thanks to Corollary 9.1.4, the restriction of the Fourier transform to $\mathcal{S}(\mathbb{R}^N) \subset L^2(\mathbb{R}^N;\mathbb{C})$ is linear and continuous (it is actually an isomorphism) from $\mathcal{S}(\mathbb{R}^N) \subset L^2(\mathbb{R}^N;\mathbb{C})$ to $\mathcal{S}(\mathbb{R}^N) \subset L^2(\mathbb{R}^N;\mathbb{C})$. As $\mathcal{S}(\mathbb{R}^N)$ is dense in $L^2(\mathbb{R}^N;\mathbb{C})$ for the norm $\|\cdot\|_2$ (indeed $\mathcal{C}_c^\infty(\mathbb{R}^N;\mathbb{C}) \subset \mathcal{S}(\mathbb{R}^N)$ and from Corollary 4.4.2, $\mathcal{C}_c^\infty(\mathbb{R}^N;\mathbb{C})$ is dense in $L^2(\mathbb{R}^N;\mathbb{C})$ for the norm $\|\cdot\|_2$), we deduce from the extension theorem 1.2.3 that there exists a unique continuous extension

$$\begin{array}{rcl}
\mathcal{F} : L^2(\mathbb{R}^N;\mathbb{C}) & \longrightarrow & L^2(\mathbb{R}^N;\mathbb{C}) \\
u & \longmapsto & \mathcal{F}(u),
\end{array}$$

such that $\mathcal{F}(u) = \hat{u}$ for each $u \in \mathcal{S}(\mathbb{R}^N)$. We call $\mathcal{F}(u)$ the Fourier transform on $L^2(\mathbb{R}^N;\mathbb{C})$. It is not clear, a priori, that

$$\mathcal{F}(u) = \hat{u}$$

for $u \in (L^1(\mathbb{R}^N;\mathbb{C}) \cap L^2(\mathbb{R}^N;\mathbb{C})) \setminus \mathcal{S}(\mathbb{R}^N)$. We will check this property later on. Let us start with the

Theorem 9.2.1. *The Fourier transform $\mathcal{F}$ is an isometric isomorphism from $L^2(\mathbb{R}^N;\mathbb{C})$ to $L^2(\mathbb{R}^N;\mathbb{C})$. Moreover, for all $u \in L^2(\mathbb{R}^N;\mathbb{C})$,*

$$\mathcal{F}(\mathcal{F}(u)) = \delta_{-1}u.$$

Proof. The conservation of the L^2-norm and the Fourier inversion formula hold in $\mathcal{S}(\mathbb{R}^N)$ (see Theorem 9.1.2 and Corollary 9.1.4). It thus suffices to appeal to the density of $\mathcal{S}(\mathbb{R}^N)$ into $L^2(\mathbb{R}^N;\mathbb{C})$, and to pass to the limit using the continuity of $\mathcal{F}$ with respect to the L^2-norm. □

The next result extends Proposition 9.1.4 to $L^2(\mathbb{R}^N;\mathbb{C})$-functions.

Proposition 9.2.1. *Let u and $v \in L^2(\mathbb{R}^N;\mathbb{C})$, then*

$$\int_{\mathbb{R}^N} \mathcal{F}(u)v\, dx = \int_{\mathbb{R}^N} u\mathcal{F}(v)\, dx.$$

Proof. Here again, we use the density of $\mathcal{S}(\mathbb{R}^N)$ into $L^2(\mathbb{R}^N;\mathbb{C})$. Indeed, let (u_n) and $(v_n) \subset \mathcal{S}(\mathbb{R}^N)$ be such that $u_n \to u$ and $v_n \to v$ in $L^2(\mathbb{R}^N;\mathbb{C})$. Thanks to Proposition 9.1.4, we infer that

$$\int_{\mathbb{R}^N} \hat{u}_n v_n\, dx = \int_{\mathbb{R}^N} u_n \hat{v}_n\, dx.$$

But since for each $n \in \mathbb{N}$, u_n and $v_n \in \mathcal{S}(\mathbb{R}^N)$, then $\hat{u}_n = \mathcal{F}(u_n)$ and $\hat{v}_n = \mathcal{F}(v_n)$. Finally it suffices to pass to the limit as $n \to \infty$ in the previous equality, using the continuity of $\mathcal{F}$ in $L^2(\mathbb{R}^N;\mathbb{C})$ and the Cauchy-Schwarz inequality. □

Corollary 9.2.1. *If $u \in L^1(\mathbb{R}^N;\mathbb{C}) \cap L^2(\mathbb{R}^N;\mathbb{C})$, then $\mathcal{F}(u) = \hat{u}$.*

Proof. Let $\chi_n \in \mathcal{C}_c^\infty(\mathbb{R}^N;[0,1])$ be such that $\chi_n = 1$ in $B(0,n)$. If $v \in \mathcal{C}_c^\infty(\mathbb{R}^N;\mathbb{C})$, then $\chi_n v \in \mathcal{C}_c^\infty(\mathbb{R}^N;\mathbb{C}) \subset \mathcal{S}(\mathbb{R}^N)$. According to Propositions 9.1.4 and 9.2.1, we get that

$$\int_{\mathbb{R}^N} \hat{u}\chi_n \bar{v}\, dx = \int_{\mathbb{R}^N} u\widehat{\chi_n v}\, dx = \int_{\mathbb{R}^N} u\mathcal{F}(\chi_n v)\, dx = \int_{\mathbb{R}^N} \mathcal{F}(u)\chi_n v\, dx,$$

so that

$$\int_{B(0,n)} (\hat{u} - \mathcal{F}(u))v\, dx = 0$$

for any $v \in \mathcal{C}_c^\infty(\mathbb{R}^N;\mathbb{C})$. Take $v = \rho_j * \overline{\hat{u} - \mathcal{F}(u)}$, where $(\rho_j)_{j\in\mathbb{N}}$ is a sequence of mollifiers as in section 4.4.2. Taking the limit as $j \to \infty$, and using Lemma 4.4.2 together with Theorem 3.3.4, we infer that $\hat{u} = \mathcal{F}(u)$ almost everywhere in $B(0,n)$. Finally since n is arbitrary, both functions actually coincide almost everywhere in $\mathbb{R}^N$. □

We conclude this section by stating an analogous result than Corollary 9.1.4 for $L^2(\mathbb{R}^N;\mathbb{C})$-functions.

Corollary 9.2.2 (Plancherel identity). *If $u \in L^2(\mathbb{R}^N;\mathbb{C})$, then*

$$\|u\|_2 = \|\mathcal{F}(u)\|_2,$$

and more generality, for u, $v \in L^2(\mathbb{R}^N;\mathbb{C})$, we have

$$\int_{\mathbb{R}^N} u\,\bar{v}\, dx = \int_{\mathbb{R}^N} \mathcal{F}(u)\overline{\mathcal{F}(v)}\, dx.$$

Proof. We proceed as usual by density. For $u \in L^2(\mathbb{R}^N;\mathbb{C})$, consider a sequence $(u_n) \subset \mathcal{S}(\mathbb{R}^N)$ such that $u_n \to u$ in $L^2(\mathbb{R}^N;\mathbb{C})$. Then, by Theorem 9.2.1, $\hat{u}_n = \mathcal{F}(u_n) \to \mathcal{F}(u)$ in $L^2(\mathbb{R}^N;\mathbb{C})$, and thanks to Corollary 9.1.4, we get that

$$\|u\|_2 = \lim_{n\to\infty} \|u_n\|_2 = \lim_{n\to\infty} \|\hat{u}_n\|_2 = \lim_{n\to\infty} \|\mathcal{F}(u_n)\|_2 = \|\mathcal{F}(u)\|_2.$$

Concerning the second statement, it suffice to observe that for any u and $v \in L^2(\mathbb{R}^N;\mathbb{C})$,

$$\begin{aligned}
\int_{\mathbb{R}^N} u\bar{v}\, dx &= \frac{\|u+v\|_2^2 - \|u-v\|_2^2}{2} \\
&= \frac{\|\mathcal{F}(u+v)\|_2^2 - \|\mathcal{F}(u-v)\|_2^2}{2} \\
&= \frac{\|\mathcal{F}(u)+\mathcal{F}(v)\|_2^2 - \|\mathcal{F}(u)-\mathcal{F}(v)\|_2^2}{2} \\
&= \int_{\mathbb{R}^N} \mathcal{F}(u)\overline{\mathcal{F}(v)}\, dx.
\end{aligned}$$

□

Now that we proved that the extension $\mathcal{F}$ correspnds to the $\hat{}$ for $u \in L^1(\mathbb{R}^N;\mathbb{C}) \cap L^2(\mathbb{R}^N;\mathbb{C})$, we can use the same notation $\hat{u}$ for $u \in L^1(\mathbb{R}^N;\mathbb{C}) \cup L^2(\mathbb{R}^N;\mathbb{C})$, and also for $u \in L^1(\mathbb{R}^N;\mathbb{C}) + L^2(\mathbb{R}^N;\mathbb{C})$ since any function can be written as the sum of a $L^1(\mathbb{R}^N;\mathbb{C})$ and a $L^2(\mathbb{R}^N;\mathbb{C})$ function.

9.3 Application to the heat equation

The model version of the heat equation (with initial datum u_0) can be written as

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) - \Delta u(t,x) = 0 & \text{for all } (t,x) \in \mathbb{R}^+ \times \mathbb{R}^N, \\[2ex] u(0,x) = u_0(x) & \text{for all } x \in \mathbb{R}^N, \end{cases}$$

where $u : \mathbb{R}^+ \times \mathbb{R}^N \to \mathbb{R}$, and $\Delta u(t,x) = \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}(t,x)$.

Let us consider, at least formally, the Fourier transform of u with respect to the x variable:

$$v(t,\xi) := \int_{\mathbb{R}^N} u(t,x) e^{-2i\pi x\cdot\xi}\, dx.$$

Using Proposition 9.1.3, we obtain after having applied the Fourier transform to the heat equation

$$\begin{cases} \dfrac{\partial v}{\partial t}(t,\xi) = -4\pi^2|\xi|^2 v(t,\xi) & \text{for all } (t,\xi) \in \mathbb{R}^+ \times \mathbb{R}^N, \\ v(0,\xi) = \hat{u}_0(\xi) & \text{for all } \xi \in \mathbb{R}^N. \end{cases}$$

The advantage of this last formulation is that, for fixed ξ (considered as a parameter), it is not anymore a partial differential equation but an ordinary differential equation whose solution is simply given by

$$v(t,\xi) = \hat{u}_0(\xi) e^{-4\pi^2|\xi|^2 t}.$$

We observe that, setting $\rho(\xi) := e^{-\pi|\xi|^2}$, then

$$\begin{aligned} e^{-4\pi^2|\xi|^2 t} &= \delta_{1/\sqrt{4\pi t}}\rho \\ &= \delta_{1/\sqrt{4\pi t}}\hat{\rho} \\ &= (4\pi t)^{-N/2}\widehat{\delta_{\sqrt{4\pi t}}\rho}, \end{aligned}$$

where we used the fact that $\hat{\rho} = \rho$ according to Corollary 9.1.1. Hence, if $u_0 \in \mathcal{S}(\mathbb{R}^N)$, using Corollary 9.1.3,

$$v(t,\xi) = \hat{u}_0(\xi)(4\pi t)^{-N/2}\widehat{\delta_{\sqrt{4\pi t}}\rho}(\xi) = (4\pi t)^{-N/2}\widehat{u_0 * \delta_{\sqrt{4\pi t}}\rho}(\xi).$$

But Corollary 9.1.2 ensures that the Fourier transform is one to one on $\mathcal{S}(\mathbb{R}^N)$, hence,

$$u(t,x) = (4\pi t)^{-N/2} u_0 * \delta_{\sqrt{4\pi t}}\rho(x) = (4\pi t)^{-N/2}\int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}}\, dy.$$

Now that we have "guessed" the form of the solution, we can state the following:

Theorem 9.3.1. *Let $u_0 \in L^1(\mathbb{R}^N)$. Then the function $u : (0,\infty) \times \mathbb{R}^N \to \mathbb{R}$ defined by*

$$u(t,x) := (4\pi t)^{-N/2}\int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}}\, dy$$

is infinitely differentiable on $(0,\infty) \times \mathbb{R}^N$. Moreover, it satisfies

$$\frac{\partial u}{\partial t}(t,x) - \Delta u(t,x) = 0 \textit{ for all } (t,x) \in (0,\infty) \times \mathbb{R}^N,$$

and

$$\lim_{t\to 0^+} \|u(t,\cdot) - u_0\|_1 = 0.$$

Proof. We remark that the function $(t,x,y) \mapsto (4\pi t)^{-N/2} e^{-\frac{|x-y|^2}{4t}}$ admits partial derivatives of any order with respect to t and/or x which are bounded (and thus belong to $L^\infty(\mathbb{R}^N)$ with respect to the y variable). Since $u_0 \in L^1(\mathbb{R}^N)$ we are in position to apply the Lebesgue's dominated convergence theorem which enables to derivate under the integral sign, and ensures that $u \in \mathcal{C}^\infty((0,\infty) \times \mathbb{R}^N)$. Moreover, we can easily check that

$$\frac{\partial u}{\partial t}(t,x) - \Delta u(t,x) = 0$$

for all $(t,x) \in (0,\infty) \times \mathbb{R}^N$.

We now study the convergence to the initial datum. We observe that, for any $t > 0$,

$$(4\pi t)^{-N/2}\int_{\mathbb{R}^N} e^{-\frac{|y|^2}{4t}}\, dy = 1 \tag{9.1}$$

so that

$$|u(t,x) - u_0(x)| \leqslant (4\pi t)^{-N/2} \int_{\mathbb{R}^N} |u_0(x-y) - u_0(x)| e^{-\frac{|y|^2}{4t}}\, dy$$

and thus, integrating with respect to x and applying Fubini's Theorem leads to

$$\begin{aligned}
\int_{\mathbb{R}^N} |u(t,x) - u_0(x)|\, dx &\leqslant \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} |u_0(x-y) - u_0(x)| \frac{e^{-\frac{|y|^2}{4t}}}{(4\pi t)^{N/2}}\, dy\, dx \\
&\leqslant \int_{B(0,\delta)} \left(\int_{\mathbb{R}^N} |\tau_y u_0(x) - u_0(x)|\, dx \right) \frac{e^{-\frac{|y|^2}{4t}}}{(4\pi t)^{N/2}}\, dy \\
&\quad + \int_{\mathbb{R}^N \setminus B(0,\delta)} \left(\int_{\mathbb{R}^N} |\tau_y u_0(x) - u_0(x)|\, dx \right) \frac{e^{-\frac{|y|^2}{4t}}}{(4\pi t)^{N/2}}\, dy \\
&\leqslant \sup_{y \in B(0,\delta)} \|\tau_y u_0 - u_0\|_1 + 2\|u_0\|_1 \int_{\mathbb{R}^N \setminus B(0,\delta)} \frac{e^{-\frac{|y|^2}{4t}}}{(4\pi t)^{N/2}}\, dy.
\end{aligned}$$

By the continuity of the translation in $L^1(\mathbb{R}^N)$ (cf. Remark 4.4.1), one has

$$\sup_{y \in B(0,\delta)} \|\tau_y u_0 - u_0\|_1 \to 0$$

as $\delta \to 0$, while a change of variable yields for every $\delta > 0$,

$$\int_{\mathbb{R}^N \setminus B(0,\delta)} \frac{e^{-\frac{|y|^2}{4t}}}{(4\pi t)^{N/2}}\, dy = \int_{\mathbb{R}^N \setminus B(0,\delta/\sqrt{4\pi t})} e^{-\pi |z|^2}\, dz \to 0$$

as $t \to 0$, since $z \mapsto e^{-\pi|z|^2} \in L^1(\mathbb{R}^N)$. Finally taking first the limit as $\delta \to 0$, and then as $t \to 0$ leads to

$$\lim_{t \to 0^+} \|u(t,\cdot) - u_0\|_1 = 0,$$

and the proof is complete. □

Let us observe that another consequence of the property (9.1) is that for any $t > 0$,

$$\int_{\mathbb{R}^N} u(t,\cdot) dx = \int_{\mathbb{R}^N} u_0. \tag{9.2}$$

Chapter 10

Tempered distributions and Sobolev spaces

Distributions are objects that generalize functions. Whereas a derivative of a function does not always exist in the classical sense, distributions admit some derivatives which are themselves some distributions. They are therefore very useful in order to solve partial differential equations. One famous distributions is the delta distribution introduced by Dirac in 1927. In the late 1940s Laurent Schwartz developed a comprehensive theory of distributions, capitalizing on some earlier works by Sobolev. The basic idea is to identify functions with abstract linear functionals on a space of smooth test functions. Operations on distributions can then be understood by moving them to the test function. Here we will restrict ourselves to the tempered distributions, which are sufficient to deal with Fourier transform in generality.

10.1 Tempered distributions

10.1.1 First definitions

Let us recall the definition of the Schwartz space

$$\mathcal{S}(\mathbb{R}^N) := \{u \in \mathcal{C}^\infty(\mathbb{R}^N;\mathbb{C}) : \ \forall \alpha,\, \beta \in \mathbb{N}^N,\ x^\alpha \partial^\beta u \in L^\infty(\mathbb{R}^N;\mathbb{C})\}.$$

The space $\mathcal{S}(\mathbb{R}^N)$ is a complete metric space. It may be possible to write explicitly the distance between two elements of $\mathcal{S}(\mathbb{R}^N)$ but it will not be useful for the considerations we will have in the sequel. We rather write the definition of a converging sequence in that topology.

Definition 10.1.1. We say that a sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^N)$ converges to u in $\mathcal{S}(\mathbb{R}^N)$ if for all multi-indexes α and $\beta \in \mathbb{N}^N$, then $\|x^\alpha \partial^\beta (u_n - u)\|_\infty \to 0$ as $n \to +\infty$.

We are now in position to introduce the tempered distributions.

Definition 10.1.2 (Tempered distributions). We denote by $\mathcal{S}'(\mathbb{R}^N)$ the space of *tempered distributions* which is defined as the space of all sequentially continuous linear maps on $\mathcal{S}(\mathbb{R}^N)$. In other words, $T \in \mathcal{S}'(\mathbb{R}^N)$ if and only if $T(u_n) \to T(u)$ for all sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^N)$ converging to u in $\mathcal{S}(\mathbb{R}^N)$.

We usually write $\langle T, u\rangle$ instead of $T(u)$.

It is possible to associate to any integrable function (and actually much more functions) a tempered distribution through an integral:

Lemma 10.1.1. *Let*

$$Z := \{u : \mathbb{R}^N \to \mathbb{C} \text{ measurable such that there exists } k_u \in \mathbb{N} : (1+|x|^2)^{-k_u} u \in L^1(\mathbb{R}^N;\mathbb{C})\}.$$

Then the mapping

$$\begin{array}{rcl} i : Z & \longrightarrow & \mathcal{S}'(\mathbb{R}^N) \\ u & \longmapsto & T_u, \end{array}$$

where

$$\langle T_u, v\rangle := \int_{\mathbb{R}^N} uv\,dx \quad \textit{for every } v \in \mathcal{S}(\mathbb{R}^N)$$

is a well defined canonical continuous linear injection from Z to $\mathcal{S}'(\mathbb{R}^N)$.

Proof. It suffices to remark that $uv = (1+|x|^2)^{-k_u}u(1+|x|^2)^{k_u}v$ is the product of a $L^1(\mathbb{R}^N;\mathbb{C})$ function (by definition of Z) and a $L^\infty(\mathbb{R}^N;\mathbb{C})$ function (because $v \in \mathcal{S}(\mathbb{R}^N)$, which implies that the map T_u is well defined. The continuity follows from the same reasoning, while the injectivity is a consequence of the uniqueness part of the Riesz theorem. □

Clearly, $\mathcal{S}(\mathbb{R}^N) \subset Z$.

We shall use the following notion of convergence of tempered distributions:

Definition 10.1.3. We say that a sequence $(T_n)_{n\in\mathbb{N}} \subset \mathcal{S}'(\mathbb{R}^N)$ converges to T in $\mathcal{S}'(\mathbb{R}^N)$ if $\langle T_n, u\rangle \to \langle T, u\rangle$ for any $u \in \mathcal{S}(\mathbb{R}^N)$.

If $\mathcal{S}(\mathbb{R}^N)$ was a normed space, then $\mathcal{S}'(\mathbb{R}^N)$ would have been the topological dual of $\mathcal{S}(\mathbb{R}^N)$, and the convergence in $\mathcal{S}'(\mathbb{R}^N)$ would have been the weak* convergence. Unfortunately, It is not the case but many things remain nevertheless true.

10.1.2 Transpose

Many operations on distributions will be defined by means of the transpose that we now define.

Definition 10.1.4 (Transpose). Let $L : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N)$ be a sequentially continuous linear map, *i.e.*, $Lu_n \to Lu$ in $\mathcal{S}(\mathbb{R}^N)$ whenever $u_n \to u$ in $\mathcal{S}(\mathbb{R}^N)$. The *transpose* map of L, denoted L^t, is defined on $\mathcal{S}'(\mathbb{R}^N)$ by

$$\langle L^t T, u\rangle = \langle T, Lu\rangle$$

for all $u \in \mathcal{S}(\mathbb{R}^N)$ and all $T \in \mathcal{S}'(\mathbb{R}^N)$.

Note that L^tT defines well a tempered distribution. Indeed, if $(u_n)_{n\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^N)$ is a sequence converging to u in $\mathcal{S}(\mathbb{R}^N)$, then since L is sequentially continuous, it follows that $Lu_n \to Lu$ in $\mathcal{S}(\mathbb{R}^N)$, and since $T \in \mathcal{S}'(\mathbb{R}^N)$, we have that $\langle T, Lu_n\rangle \to \langle T, Lu\rangle$. Hence by definition of the transpose $\langle L^tT, u_n\rangle \to \langle L^tT, u\rangle$ so that $L^tT \in \mathcal{S}'(\mathbb{R}^N)$.

Proposition 10.1.1. *The map $L^t : \mathcal{S}'(\mathbb{R}^N) \to \mathcal{S}'(\mathbb{R}^N)$ is linear and continuous.*

Proof. The linearity is obvious. For what concerns continuity, assume that $T_n \to T$ in $\mathcal{S}'(\mathbb{R}^N)$. Then for each $u \in \mathcal{S}(\mathbb{R}^N)$, we have $Lu \in \mathcal{S}(\mathbb{R}^N)$ and thus, $\langle T_n, Lu\rangle \to \langle T, Lu\rangle$. Hence by definition of the transpose, $\langle L^tT_n, u\rangle \to \langle L^tT, u\rangle$. This shows that $L^tT_n \to L^tT$ in $\mathcal{S}'(\mathbb{R}^N)$. □

We are going to define elementary operations on tempered distributions (multiplication, derivation, Fourier transform, convolution) as transpose of their analogous continuous linear maps on $\mathcal{S}(\mathbb{R}^N)$. To this aim, let us first prove the following

Proposition 10.1.2. *The following linear maps are continuous from $\mathcal{S}(\mathbb{R}^N)$ to $\mathcal{S}(\mathbb{R}^N)$:*

1. *Translation: $u \mapsto \tau_a u(x) := u(x-a)$, $a \in \mathbb{R}^N$;*
2. *Dilatation: $u \mapsto \delta_\lambda u(x) := u(x/\lambda)$, $\lambda \in \mathbb{R}\setminus\{0\}$;*
3. *Derivation: $u \mapsto \partial^\alpha u$, $\alpha \in \mathbb{N}^N$;*
4. *Multiplication: $u \mapsto \chi u$, $\chi \in \Theta_M(\mathbb{R}^N)$, where $\Theta_M(\mathbb{R}^N) := \{\chi \in \mathcal{C}^\infty(\mathbb{R}^N;\mathbb{C}) : \forall\, \alpha \in \mathbb{N}^N, \exists k_\alpha \in \mathbb{N}$ such that $(1+|x|^2)^{-k_\alpha}\partial^\alpha\chi \in L^\infty(\mathbb{R}^N;\mathbb{C})\}$ is the space of all tempered functions on $\mathbb{R}^N$;*
5. *Fourier transform: $u \mapsto \hat{u}$;*
6. *Convolution: $u \mapsto \chi * u$, where $\chi \in \mathcal{S}(\mathbb{R}^N)$.*

Proof. Only the three last statements need to be proved, the other ones being obvious. Concerning 4), according to Leibniz' formula, we have for all α, $\beta \in \mathbb{N}^N$,

$$x^\beta \partial^\alpha(\chi u) = x^\beta \sum_{\gamma+\delta=\alpha} \frac{\alpha!}{\gamma!\delta!} \partial^\gamma \chi \partial^\beta u.$$

For fixed γ, we have $|\partial^\gamma \chi| \leqslant C_\gamma (1+|x|^2)^{k_\gamma}$ so that

$$|x^\beta \partial^\gamma \chi| \leqslant C_\gamma (1+|x|^2)^{k_\gamma + |\beta|}.$$

The conclusion follows from the definition of the convergence in $\mathcal{S}(\mathbb{R}^N)$ and from the fact that the sum is finite.

Concerning 5), we observe that

$$\begin{aligned} \xi^\beta \partial^\alpha \hat{u} &= \xi^\beta \widehat{(-2i\pi x)^\alpha u} \\ &= \frac{1}{(2i\pi)^{|\beta|}} (2i\pi\xi)^\beta \widehat{(-2i\pi x)^\alpha u} \\ &= \frac{1}{(2i\pi)^{|\beta|}} \widehat{\partial^\beta((-2i\pi x)^\alpha u)}. \end{aligned}$$

If $u_n \to u$ in $\mathcal{S}(\mathbb{R}^N)$, then from 3) and 4) we obtain that $\partial^\beta((-2i\pi x)^\alpha u_n) \to \partial^\beta((-2i\pi x)^\alpha u)$ in $\mathcal{S}(\mathbb{R}^N)$ and thus in $L^1(\mathbb{R}^N;\mathbb{C})$ as well. From the Riemann-Lebesgue Theorem (Theorem 9.1.1) the Fourier transform maps continuously $L^1(\mathbb{R}^N;\mathbb{C})$ into $\mathcal{C}_0(\mathbb{R}^N;\mathbb{C})$ (and thus also $L^\infty(\mathbb{R}^N;\mathbb{C})$). It ensures that $\widehat{\partial^\beta((-2i\pi x)^\alpha u_n)} \to \widehat{\partial^\beta((-2i\pi x)^\alpha u)}$ in $L^\infty(\mathbb{R}^N;\mathbb{C})$, whence the conclusion.

It remains to prove 6). For that, let us remark that form Theorem 9.1.2 and Corollary 9.1.3, one has

$$u * \chi = \delta_{-1}\widehat{\widehat{u * \chi}} = \delta_{-1}\widehat{\hat{u}\hat{\chi}}$$

and the conclusion follows from 2), 4) and 5). □

The following formulas are satisfied in $\mathcal{S}(\mathbb{R}^N)$:

Proposition 10.1.3. *For every u and $v \in \mathcal{S}(\mathbb{R}^N)$, one has*

1. $\displaystyle\int_{\mathbb{R}^N} (\tau_a u) v \, dx = \int_{\mathbb{R}^N} u(\tau_{-a} v)\, dx$ *for all* $a \in \mathbb{R}^N$;
2. $\displaystyle\int_{\mathbb{R}^N} (\delta_\lambda u) v \, dx = |\lambda|^N \int_{\mathbb{R}^N} u(\delta_{1/\lambda} v)\, dx$ *for all* $\lambda \in \mathbb{R}\setminus\{0\}$;
3. $\displaystyle\int_{\mathbb{R}^N} (\partial^\alpha u) v \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^N} u(\partial^\alpha v)\, dx$ *for all* $\alpha \in \mathbb{N}^N$;
4. $\displaystyle\int_{\mathbb{R}^N} (\chi u) v \, dx = \int_{\mathbb{R}^N} u(\chi v)\, dx$ *for all* $\chi \in \Theta_M(\mathbb{R}^N)$;
5. $\displaystyle\int_{\mathbb{R}^N} \hat{u} v \, dx = \int_{\mathbb{R}^N} u \hat{v}\, dx$;
6. $\displaystyle\int_{\mathbb{R}^N} (\chi * u) v \, dx = \int_{\mathbb{R}^N} u((\delta_{-1}\chi) * v)\, dx$ *for all* $\chi \in \mathcal{S}(\mathbb{R}^N)$.

A way to interpret the properties of Proposition 10.1.3 uses the injection $i : \mathcal{S}(\mathbb{R}^N) \to \mathcal{S}'(\mathbb{R}^N)$ introduced in Lemma 10.1.1. Indeed 1) implies that the transpose of τ_a (: $\mathcal{S}(\mathbb{R}^N) \to \mathcal{S}(\mathbb{R}^N)$), $\tau_a^t : \mathcal{S}'(\mathbb{R}^N) \to \mathcal{S}'(\mathbb{R}^N)$ restricted to $i(\mathcal{S}(\mathbb{R}^N))$ is nothing but $i \circ \tau_{-a}$. We proceed similarly for the other properties and we get the following definitions

Definition 10.1.5. *The following linear maps are continuous from $\mathcal{S}'(\mathbb{R}^N)$ to $\mathcal{S}'(\mathbb{R}^N)$:*

1. *Translation:* $\langle \tau_a T, v\rangle := \langle T, \tau_{-a} v\rangle$;

2. *Multiplication:* $\langle \delta_\lambda T, v\rangle := |\lambda|^N \langle T, \delta_{1/\lambda} v\rangle$;

3. *Derivation:* $\langle \partial^\alpha T, v\rangle := (-1)^{|\alpha|}\langle T, \partial^\alpha v\rangle$;

4. *Multiplication:* $\langle \chi T, v\rangle := \langle T, \chi v\rangle$ *for every* $\chi \in \Theta_M(\mathbb{R}^N)$;

5. *Fourier transform:* $\langle \hat{T}, v\rangle := \langle T, \hat{v}\rangle$;

6. *Convolution:* $\langle \chi * T, v\rangle := \langle T, (\delta_{-1}\chi) * v\rangle$ *for every* $\chi \in \mathcal{S}(\mathbb{R}^N)$.

Let us observe that a tempered distribution can derived as much as we want, the result being an element of $\mathcal{S}'(\mathbb{R}^N)$.

Theorem 10.1.1. *The Fourier transform is a continuous linear one to one mapping from* $\mathcal{S}'(\mathbb{R}^N)$ *to* $\mathcal{S}'(\mathbb{R}^N)$. *Moreover, we have the Fourier inversion formula*

$$\hat{\hat{T}} = \delta_{-1} T$$

for every $T \in \mathcal{S}'(\mathbb{R}^N)$.

Proof. By definition we have for all $v \in \mathcal{S}(\mathbb{R}^N)$,

$$\begin{aligned} \langle \hat{\hat{T}}, v\rangle &= \langle \hat{T}, \hat{v}\rangle = \langle T, \hat{\hat{v}}\rangle \\ &= \langle T, \delta_{-1} v\rangle = \langle \delta_{-1} T, v\rangle, \end{aligned}$$

where we use the Fourier inversion formula (see Theorem 9.1.2) for functions in $\mathcal{S}(\mathbb{R}^N)$. The fact that the Fourier transform is one to one follows from the inversion formula in $\mathcal{S}'(\mathbb{R}^N)$. □

The formulas relating derivation, convolution, Fourier transform, etc... which are valid in $\mathcal{S}(\mathbb{R}^N)$ can be transposed almost immediately in $\mathcal{S}'(\mathbb{R}^N)$.

Proposition 10.1.4. *For every* $T \in \mathcal{S}'(\mathbb{R}^N)$, *we have*

1. $\widehat{\tau_a T} = e^{-2i\pi a\cdot\xi}\hat{T}$ *for all* $a \in \mathbb{R}^N$, *where the right hand side is the product of the tempered distribution* $\hat{T}$ *with the tempered function* $\xi \mapsto e^{-2i\pi a\cdot\xi}$;

2. $\widehat{\delta_\lambda T} = |\lambda|^N \delta_{1/\lambda}\hat{T}$ *for all* $\lambda \in \mathbb{R}\setminus\{0\}$;

3. $\widehat{\partial^\alpha T} = (-2i\pi\xi)^\alpha \hat{T}$ *for all* $\alpha \in \mathbb{N}^N$;

4. $\widehat{\chi T} = \hat{\chi} * \hat{T}$ *for all* $\chi \in \mathcal{S}(\mathbb{R}^N)$;

5. $\widehat{\chi * T} = \hat{\chi}\hat{T}$ *for all* $\chi \in \mathcal{S}(\mathbb{R}^N)$;

6. $\partial^\alpha(\chi * T) = \partial\chi * T = \chi * \partial^\alpha T$ *for all* $\alpha \in \mathbb{N}^N$ *and all* $\chi \in \mathcal{S}(\mathbb{R}^N)$.

The Dirac at a point $a \in \mathbb{R}^N$, as a bounded Radon measure, is also a tempered distribution since $\mathcal{S}(\mathbb{R}^N) \subset \mathcal{C}_0(\mathbb{R}^N)$. It is defined by

$$\langle \delta_a, v\rangle := v(a)$$

for all $v \in \mathcal{S}(\mathbb{R}^N)$.

Proposition 10.1.5. *The Fourier transform of* δ_a *is the tempered function* $\xi \mapsto e^{-2i\pi a\cdot\xi}$. *In particular* $\hat{\delta}_0 = 1$, *and* δ_0 *acts as an identity element for the convolution product:*

$$\chi * \delta_0 = i(\chi) \sim \chi,$$

where $\sim$ *means the identification of a function with its associated distribution.*

Proof. Let $v \in \mathcal{S}(\mathbb{R}^N)$, we have

$$\begin{aligned}
\langle \hat{\delta}_a, v\rangle &= \langle \delta_a, \hat{v}\rangle = \hat{v}(a)\\
&= \int_{\mathbb{R}^N} v(x)e^{-2i\pi a\cdot x}\,dx = \langle i(e^{-2i\pi a\cdot x}), v\rangle \sim \langle e^{-2i\pi a\cdot x}, v\rangle.
\end{aligned}$$

Hence

$$\begin{aligned}
\chi * \delta_0 &= \delta_{-1}\left(\widehat{\widehat{\chi * \delta_0}}\right) = \delta_{-1}\left(\widehat{\hat{\chi}\hat{\delta}_0}\right)\\
&= \delta_{-1}\left(\widehat{\hat{\chi}i(1)}\right) = i(\delta_{-1}\hat{\hat{\chi}}) = i(\chi).
\end{aligned}$$

□

10.1.3 Fundamental solution of a differential operator with constant coefficients

Definition 10.1.6. *A differential operator with constant coefficients on $\mathbb{R}^N$ is an operator*

$$L := \sum_{i\in I} c_i\partial^{\alpha_i},$$

where I has finite cardinality, $c_i \in \mathbb{C}$ and $\alpha_i \in \mathbb{N}$. We say that $T \in \mathcal{S}'(\mathbb{R}^N)$ is a fundamental solution of the differential operator L if $L(T) = \delta_0$ in $\mathcal{S}'(\mathbb{R}^N)$.

Corollary 10.1.1. *Let L be a differential operator with constant coefficients in $\mathbb{R}^N$, and T be a fundamental solution of L. Then for every $f \in \mathcal{S}(\mathbb{R}^N)$, if we set $u := f * T$, we have $L(u) = f$. In other words, $f * T$ is a solution of the partial differential equation $L(u) = f$.*

Proof. From Proposition 10.1.4, we have

$$L(u) = \sum_{i\in I} c_i\partial^{\alpha_i}(f*T) = f * \sum_{i\in I} c_i\partial^{\alpha_i}T = f * L(T) = f*\delta_0 = f.$$

□

Example 10.1.1. Let us consider the differential operator L on $\mathbb{R}^{N+1}$ associated to the heat equation:

$$L(T) := \frac{\partial T}{\partial t} - \Delta T = \frac{\partial T}{\partial t} - \sum_{i=1}^N \frac{\partial^2 T}{\partial x_i^2}.$$

The reader can try to show as an exercise that the tempered distribution associated to the function of Z defined by

$$G(x,t) := \begin{cases} \dfrac{1}{(4\pi t)^{N/2}} e^{-\frac{|x|^2}{4t}} & \text{if} \quad t > 0,\\ 0 & \text{if} \quad t \leqslant 0 \end{cases}$$

is a fundamental solution for L on $\mathbb{R}^{N+1}$.
Hint: Observe that $G \in \mathcal{C}^\infty((\mathbb{R}^N\times\mathbb{R})\setminus\{0,0\}) \cap L^1(\mathbb{R}^N\times\mathbb{R})$ and that for each $(x,t) \neq (0,0)$,

$$\left(\frac{\partial G}{\partial t} - \Delta G\right)(x,t) = 0.$$

Then write that for all $\chi \in \mathcal{S}(\mathbb{R}^N\times\mathbb{R})$,

$$\begin{aligned}
\left\langle \frac{\partial G}{\partial t} - \Delta G, \chi\right\rangle &= \int_{\mathbb{R}^N\times\mathbb{R}} G(x,t)\left(\frac{\partial \chi}{\partial t} - \Delta\chi\right)(x,t)\,dx\,dt\\
&= \lim_{\varepsilon\to 0}\int_{(\mathbb{R}^N\times\mathbb{R})\setminus\Lambda_\varepsilon} G(x,t)\left(\frac{\partial \chi}{\partial t} - \Delta\chi\right)(x,t)\,dx\,dt,
\end{aligned}$$

where $\lambda_\varepsilon := B(0,\varepsilon)\times[-\varepsilon^3,\varepsilon^3]$ is a small cylinder surrounding the origin. Note that $G \in \mathcal{C}^\infty((\mathbb{R}^N\times\mathbb{R})\setminus\Lambda_\varepsilon)$, perform an integration by parts, and identify the limit as $\varepsilon \to 0$ of the boundary terms.

10.2 Sobolev spaces

10.2.1 Definition

Definition 10.2.1. *For any $k \in \mathbb{N}$, we define the spaces*

$$W^{k,2}(\mathbb{R}^N) := \{u \in \mathcal{S}'(\mathbb{R}^N) : \partial^\alpha u \in L^2(\mathbb{R}^N) \text{ for all } \alpha \in \mathbb{N} \text{ with } |\alpha| \leqslant k\}$$

and

$$H^k(\mathbb{R}^N) := \{u \in \mathcal{S}'(\mathbb{R}^N) : (1 + |\xi|^2)^{k/2}\hat{u} \in L^2(\mathbb{R}^N)\},$$

endowed with the scalar products

$$(u, v)_{W^{k,2}(\mathbb{R}^N)} := \sum_{|\alpha| \leqslant k} (\partial^\alpha u, \partial^\alpha v)_{L^2(\mathbb{R}^N)}$$

and

$$(u, v)_{H^k(\mathbb{R}^N)} := \big((1 + |\xi|^2)^{k/2}\hat{u}, (1 + |\xi|^2)^{k/2}\hat{v}\big)_{L^2(\mathbb{R}^N)},$$

and with the associated norms.

It turns out that the spaces $W^{k,2}(\mathbb{R}^N)$ and $H^k(\mathbb{R}^N)$ are algebraically and topologically identical.

Lemma 10.2.1. *We have $W^{k,2}(\mathbb{R}^N) = H^k(\mathbb{R}^N)$ and their norms are equivalent.*

Proof. Thanks to the Plancherel identity, Corollary 9.2.2, that for each $\alpha \in \mathbb{N}^N$ with $|\alpha| \leqslant k$, then $\partial^\alpha u \in L^2(\mathbb{R}^N)$ if and only if $(2i\pi\xi)^\alpha \hat{u} \in L^2(\mathbb{R}^N)$. On the other hand, by simple algebraic manipulation, it can be easily seen that for fixed $k \in \mathbb{N}^N$, there exists a constant $C_k > 0$ such that

$$\frac{1}{C_k}(1 + |\xi|^2)^{k/2} \leqslant \sum_{|\alpha| \leqslant k} (2i\pi\xi)^\alpha \leqslant C_k(1 + |\xi|^2)^{k/2} \tag{10.1}$$

for any $\xi \in \mathbb{R}^N$. Hence $\partial^\alpha u \in L^2(\mathbb{R}^N)$ if and only if $(1+|\xi|^2)^{k/2}\hat{u} \in L^2(\mathbb{R}^N)$, so that $W^{k,2}(\mathbb{R}^N) = H^k(\mathbb{R}^N)$ algebraically. The equivalence of the norms (and thus the topological equality) follows from (10.1). □

The Sobolev spaces $H^k(\mathbb{R}^N)$ play an intermediate role between the Lebesgue space $L^2(\mathbb{R}^N)$ (where function are missing of regularity) and the Schwartz space $\mathcal{S}(\mathbb{R}^N)$ (which is not normed, and thus not a Hilbert space). Indeed, we have both following important results.

Theorem 10.2.1. *For each $k \in \mathbb{N}$, $H^k(\mathbb{R}^N)$ is a Hilbert space.*

Proof. It is enough to show that $H^k(\mathbb{R}^N)$ is complete. Let $(u_n)_{n\geqslant 1} \subset H^k(\mathbb{R}^N)$ be a Cauchy sequence in $H^k(\mathbb{R}^N)$. By definition, it follows that $\xi \mapsto (1 + |\xi|^2)^{k/2}\hat{u}_n(\xi)$ is a Cauchy sequence in $L^2(\mathbb{R}^N)$. Since the latter is complete, there exists $v_\infty \in L^2(\mathbb{R}^N)$ such that $(1 + |\xi|^2)^{k/2}\hat{u}_n \to v_\infty$ in $L^2(\mathbb{R}^N)$. As the function $\xi \mapsto (1 + |\xi|^2)^{-k/2}$ is tempered then by Definition 10.1.5, $(1 + |\xi|^2)^{-k/2}v_\infty \in \mathcal{S}'(\mathbb{R}^N)$ and by the bijectivity of the Fourier transform in $\mathcal{S}'(\mathbb{R}^N)$ (see Theorem 10.1.1), there exists $u_\infty \in \mathcal{S}'(\mathbb{R}^N)$ such that $\hat{u}_\infty = (1 + |\xi|^2)^{-k/2}v_\infty$. Hence $(1 + |\xi|^2)^{k/2}\hat{u}_\infty = v_\infty \in L^2(\mathbb{R}^N)$ so that $u_\infty \in H^k(\mathbb{R}^N)$. Moreover, $(1 + |\xi|^2)^{k/2}\hat{u}_n \to (1 + |\xi|^2)^{k/2}\hat{u}_\infty$ in $L^2(\mathbb{R}^N)$, and thus, by definition $u_n \to u_\infty$ in $H^k(\mathbb{R}^N)$. □

10.2.2 A few properties

Theorem 10.2.2. *If $k > m + N/2$, then $H^k(\mathbb{R}^N) \subset \mathcal{C}_0^m(\mathbb{R}^N)$ algebraically and topologically. We write $H^k(\mathbb{R}^N) \hookrightarrow \mathcal{C}_0^m(\mathbb{R}^N)$.*

Let us recall that $\mathcal{C}_0^m(\mathbb{R}^N) := \{u \in \mathcal{C}^m(\mathbb{R}^N) : \partial^\alpha u \in \mathcal{C}_0(\mathbb{R}^N) \text{ for each } |\alpha| \leqslant m\}$.

Proof. From the Riemann-Lebesgue Theorem (Theorem 9.1.1), we know that the Fourier transform maps continuously $L^1(\mathbb{R}^N)$ into $\mathcal{C}_0(\mathbb{R}^N)$. Hence, according the Fourier inversion formula, Theorem 9.1.2, it suffices to check that for each $\alpha \in \mathbb{N}^N$ with $|\alpha| \leqslant m$, we have $\widehat{\partial^\alpha u} \in L^1(\mathbb{R}^N)$ when $u \in H^k(\mathbb{R}^N)$. Indeed, if it is the case, then $\partial^\alpha u = \delta_{-1}\widehat{\widehat{\partial^\alpha u}} \in \mathcal{C}_0(\mathbb{R}^N)$.

Let $u \in H^k(\mathbb{R}^N)$, we have

$$|\widehat{\partial^\alpha u}(\xi)| = |(2i\pi\xi)^\alpha \hat{u}(\xi)| = \left|\frac{(2i\pi\xi)^\alpha}{(1+|\xi|^2)^{k/2}}(1+|\xi|^2)^{k/2}\hat{u}(\xi)\right|. \tag{10.2}$$

Since $k > m + N/2$ and $|\alpha| \leqslant m$, letting $\varepsilon := k - m - N/2 > 0$, we have

$$\left|\frac{(2i\pi\xi)^\alpha}{(1+|\xi|^2)^{k/2}}\right| \leqslant C(1+|\xi|^2)^{-k/2+|\alpha|/2} \leqslant C(1+|\xi|^2)^{-N/4-\varepsilon/2} \in L^2(\mathbb{R}^N),$$

for some constant $C > 0$ depending only on N, k and m. On the other hand, since $(1+|\xi|^2)^{k/2}\hat{u} \in L^2(\mathbb{R}^N)$ we deduce from (10.2) and the Cauchy-Schwarz inequality that $\widehat{\partial^\alpha u} \in L^1(\mathbb{R}^N)$. $\square$

In particular, in 1-space dimension ($N = 1$), then $H^1(\mathbb{R}) \hookrightarrow \mathcal{C}_0(\mathbb{R})$.

For later applications, it will be convenient to restrict ourself to an open subset $\Omega \subset \mathbb{R}^N$. We will not define $H^k(\Omega)$ but only $H^k_0(\Omega)$, which roughly speaking, stands for those $H^k(\mathbb{R}^N)$ functions vanishing outside Ω (this simplied way of saying hides problems of extending smoothly Sobolev functions on the boundary of Ω). Let us see the rigorous definition of "functional analysis" nature.

Definition 10.2.2. *The space $H^k_0(\Omega)$ is the closure of $\mathcal{C}^\infty_c(\Omega)$ is $H^k(\mathbb{R}^N)$.*

It follows from the definition of $H^k_0(\Omega)$ that it is a closed subset of $H^k(\mathbb{R}^N)$ endowed with the norm $\|\cdot\|_{H^k}$. It is consequently complete as well.

Proposition 10.2.1 (Locality). *If $u \in H^k_0(\Omega)$, then for each $\alpha \in \mathbb{N}^N$ such that $|\alpha| \leqslant k$,*

$$\|\partial^\alpha u\|_{L^2(\mathbb{R}^N)} = \|\partial^\alpha u\|_{L^2(\Omega)}.$$

Proof. The property is obviously true for functions in $\mathcal{C}^\infty_c(\Omega)$. If $u \in H^k_0(\Omega)$ and $(u_n)_{n\geqslant 1} \subset \mathcal{C}^\infty_c(\Omega)$ converging to u in $H^k(\mathbb{R}^N)$. Then $\partial^\alpha u_n \to \partial^\alpha u$ in $L^2(\mathbb{R}^N)$ and consequently,

$$\left|\|\partial^\alpha u\|_{L^2(\mathbb{R}^N\setminus\Omega)} - \|\partial^\alpha u_n\|_{L^2(\mathbb{R}^N\setminus\Omega)}\right| \leqslant \|\partial^\alpha(u-u_n)\|_{L^2(\mathbb{R}^N\setminus\Omega)} \leqslant \|\partial^\alpha(u-u_n)\|_{L^2(\mathbb{R}^N)},$$

which goes to 0 when n goes to infinity. $\square$

Theorem 10.2.3 (Compactness). *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set. Then the injection of $H^1_0(\Omega)$ in $L^2(\Omega)$ is compact.*

Proof. Let $u \in H^1_0(\Omega)$ be such that $\|u\|_{H^1(\mathbb{R}^N)} \leqslant 1$, and $h \in \mathbb{R}^N$. Thanks to the Plancherel identity, Corollary 9.2.2, one has

$$\begin{aligned} \|\tau_h u - u\|_{L^2(\mathbb{R}^N)} &= \|\widehat{\tau_h u} - \hat{u}\|_{L^2(\mathbb{R}^N)} \\ &= \|e^{-2i\pi\xi\cdot h}\hat{u} - \hat{u}\|_{L^2(\mathbb{R}^N)} \\ &= \|(e^{-2i\pi\xi\cdot h} - 1)\hat{u}\|_{L^2(\mathbb{R}^N)}. \end{aligned}$$

It is easy to check that there exists a constant $C > 0$ such that

$$|e^{-2i\pi\xi\cdot h} - 1| = |e^{-2i\pi\xi\cdot h} - e^{-2i\pi 0\cdot h}| \leqslant C|h||\xi|.$$

Hence,

$$\begin{aligned} \|\tau_h u - u\|_{L^2(\mathbb{R}^N)} &\leqslant C|h|\||\xi|\hat{u}\|_{L^2(\mathbb{R}^N)} \\ &\leqslant C|h|\|(1+|\xi^2)^{1/2}\hat{u}\|_{L^2(\mathbb{R}^N)} \\ &= C|h|\|u\|_{H^1(\mathbb{R}^N)} \leqslant C|h|. \end{aligned}$$

As a consequence, according to the Riesz-Fréchet-Kolmogorov Theorem (Theorem 4.5.1), we deduce that $\{u \in H^1_0(\Omega) : \|u\|_{H^1(\mathbb{R}^N)} \leqslant 1\}$ is compact in $L^2(\Omega)$. $\square$

Corollary 10.2.1 (Poincaré inequality). *If $\Omega \subset \mathbb{R}^N$ is a bounded open set, then there exists a constant $C > 0$ such that*

$$\|u\|_{L^2(\Omega)} \leqslant C\|\nabla u\|_{L^2(\Omega)}$$

for every $u \in H_0^1(\Omega)$.

Proof. If the conclusion was not true, for any $n \geqslant 1$, there would exists $v_n \in H_0^1(\Omega)$ such that $\|v_n\|_{L^2(\Omega)} > n\|\nabla v_n\|_{L^2(\Omega)}$. Define $u_n := v_n/\|v_n\|_{L^2(\Omega)}$, then $\|u_n\|_{L^2(\Omega)} = 1$ and $\|\nabla u_n\|_{L^2(\Omega)} \to 0$ as $n \to \infty$. In particular the sequence $(u_n)_{n\geqslant 1}$ would be bounded in $H_0^1(\Omega)$, and thanks to the compact imbedding of $H_0^1(\Omega)$ into $L^2(\Omega)$ (Theorem 10.2.3), there would exist a subsequence of $(u_n)_{n\geqslant 1}$ (still denoted by $(u_n)_{n\geqslant 1}$) and $u \in H_0^1(\Omega)$ such that $u_n \to u$ in $L^2(\Omega)$, and in particular by Proposition 10.2.1,

$$\|u\|_{L^2(\mathbb{R}^N)} = \|u\|_{L^2(\Omega)} = \lim_{n\to\infty} \|u_n\|_{L^2(\Omega)} = 1.$$

On the other hand, since $\nabla u_n \to 0$ in $L^2(\Omega)$ then $\nabla u = 0$ a.e. in Ω, and thus also a.e. in $\mathbb{R}^N$ by the localization property (Proposition 10.2.1). This would imply that u is constant in $\mathbb{R}^N$ which is absurd since $\|u\|_{L^2(\mathbb{R}^N)} = 1 < \infty$. $\square$

It follows from the Poincaré inequality that the norm on $H_0^1(\Omega)$ is equivalent to

$$\|u\|_{H_0^1(\Omega)} := \|\nabla u\|_{L^2(\Omega)},$$

and that $H_0^1(\Omega)$ endowed with the scalar product

$$(u, v)_{H_0^1(\Omega)} := \int_\Omega \nabla u \cdot \overline{\nabla v}\, dx$$

is a Hilbert space.

10.2.3 Dirichlet problem

We conclude this chapter by an application concerning the existence of solutions of the Dirichlet problem.

Theorem 10.2.4. *Let $\Omega \subset \mathbb{R}^N$ be a bounded open set, and $f \in L^2(\Omega)$. Then, there exists a unique $u \in H_0^1(\Omega)$ such that*

$$-\Delta u = f$$

in the sense of distributions in Ω, i.e., for any $\chi \in \mathcal{C}_c^\infty(\Omega)$,

$$\langle -\Delta u, \chi\rangle = \langle f, \chi\rangle.$$

Proof. The map

$$\begin{array}{rcl} L : H_0^1(\Omega) & \longrightarrow & \mathbb{C}, \\ u & \longmapsto & \displaystyle\int_\Omega u(x)f(x)\, dx \end{array}$$

is a continuous linear map. Indeed by the Cauchy-Schwarz and Poincaré inequalities,

$$L(u) \leqslant \|u\|_{L^2(\Omega)}\|f\|_{L^2(\Omega)} \leqslant C\|\nabla u\|_{L^2(\Omega)}\|f\|_{L^2(\Omega)} \leqslant C\|f\|_{L^2(\Omega)}\|u\|_{H_0^1(\Omega)}.$$

Hence by the Riesz representation Theorem in Hilbert spaces (Theorem 7.4.1), there exists a unique $u \in H_0^1(\Omega)$ such that

$$L(v) = (u, v)_{H_0^1(\Omega)}$$

for all $v \in H_0^1(\Omega)$, or in other words,

$$\int_\Omega \nabla u \cdot \nabla v\, dx = \int_\Omega fv\, dx.$$

In particular, since $\mathcal{C}_c^\infty(\Omega) \subset L^2(\Omega)$, the previous relation holds for any $v \in \mathcal{C}_c^\infty(\Omega)$ as well. Next, thanks to the definition of the derivation in the sense of distributions, Definition 10.1.5, and the fact that $\mathcal{C}_c^\infty(\Omega) \subset \mathcal{S}(\mathbb{R}^N)$, we deduce that

$$\langle -\Delta u, \chi\rangle = \langle \nabla u, \nabla \chi\rangle$$

for any $\chi \in \mathcal{C}_c^\infty(\Omega)$. Moreover, since $\nabla u \in L^2(\mathbb{R}^N)$

$$\langle \nabla u, \nabla \chi \rangle = \int_\Omega \nabla u \cdot \nabla \chi \, dx,$$

where we used the locality property, Proposition 10.2.1. Gathering everything, we infer that

$$\langle -\Delta u, \chi \rangle = \langle f, \chi \rangle$$

for any $\chi \in \mathcal{C}_c^\infty(\Omega)$. The uniqueness of the solution follows from that of the Riesz representation Theorem 7.4.1. □

We say that $u \in H_0^1(\Omega)$ is a weak solution (or solution in the sense of distributions) of the equation

$$-\Delta u = f \text{ in } \Omega.$$

The next step, which is out the scope of this course, consists in studying the regularity of weak solutions, and to ask wether weak solutions are actually classical solutions.

Bibliography

[1] H. Brézis: *Analyse fonctionnelle. Théorie et applications*, Masson, 1993.

[2] E. Di Benedetto: *Real analysis*, Birkhaüser, Boston (2002)

[3] W. Rudin: *Real and complex analysis*, McGraw-Hill Int. Ed., third edition (1987).

[4] W. Rudin: *Functional analysis*, McGraw-Hill Int. Ed., second edition (1991).

[5] M. Willem: *Analyse harmonique réelle* , Hermann, 1997.