1

A quick look at topological and functional spaces

> *"The unified character of mathematics lies in its very nature; indeed, mathematics is the foundation of all exact natural sciences."*
>
> *David Hilbert (1862-1943)*

Nowadays, *functional analysis*, that is mainly concerned with the study of complete normed vector spaces, occupies a central place in modern mathematical analysis. Initially motivated by the understanding and the study of differental and integral equations arising in applied mathematics, it has largely developed and evolved around the theory of Banach and Hilbert spaces and their rich geometric structure. The importance and the versatility of Hilbert spaces is exampled by the space of Lebesgue square integrable functions. In this context, most functional spaces have infinite dimension and the classical theory focusses on linear operators between these spaces.

To better understand the conceptual breakdown in real analysis offered by the new functional spaces, we introduce the following example, borrowed from [LV02]. Consider for instance the wave equation model, a simplified model describing the transversal oscillations $u = u(x,t)$ of a stretched vibrating string in one dimension of space

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

and supposed to be pegged at its two endpoints, *i.e.,* edowed with the boundary conditions

$$u(0,t) = u(1,t) = 0\,.$$

The physical interpretation suggests also to specify two initial conditions

$$u(x,0) = u_0(x)\,, \quad \text{and} \quad \frac{\partial u}{\partial t}(x,0) = v_0(x)\,,$$

which prescribes the initial position of the string and the initial velocity of the points of the string. The natural setting for finding a solution with finite energy, *i.e.,* such that

$$\int_0^1 \left(\frac{\partial u}{\partial t}\right)^2 < +\infty\,, \quad \text{and} \quad \int_0^1 \left(\frac{\partial u}{\partial x}\right)^2 < +\infty\,,$$

is to use Fourier series and then the solution will have the general form

$$u(x,t) = \sum_{k=1}^{\infty} \big(a_k \cos(k\pi ct) + b_k \sin(k\pi ct) \big) \sin(k\pi x)\,,$$

and the initial and boundary conditions allow to explicit the coefficients a_k and b_k as

$$a_k = 2\int_0^1 u_0(x)\sin(k\pi x)dx\,, \quad \text{and} \quad b_k = \frac{2}{k\pi c}\int_0^1 v_0(x)\sin(k\pi x)dx\,.$$

Such solution involves an infinite sum and is described by a denumerable set of coefficients a_k and b_k. This suggests that the space of solution shall be infinite dimensional. And this statement showed the limits of the yet known results in real analysis. Indeed, the classical Bolzano-Weierstrass theorem about the notion of convergence in a finite dimensional Euclidean space (*i.e.,* every bounded sequence admits a bounded subsequence in $\mathbb{R}^n$) breaks apart. Fortunately, Hilbert spaces were introduced and provided a convenient setting for analyzing this type of problem.

In this chapter we summarize many of the abstract concepts, definitions and theoretical results on the functional spaces that are relevant in functional analysis and important for understanding the properties of the solutions of partial differential equations. These abstract spaces, *metric spaces, normed spaces, inner-product spaces*, are topological algebraic spaces that have all been introduced in the last three decades of the nineteenth century and the first decades of the twentieth century[1]. They ultimately led to a generalization of the notions of functions, continuity, differentiability and integrability. Until then, functions were assumed to be continuous, have derivatives at almost all points and were integrable by existing integration methods.

In Section 1, we recall the elementary topological spaces and the fundamental properties, separability, compactness and completeness. Section 2, Lesbegue integration is introduced as simply as possible, without referring explicitly to the measure theory, essentially in view of presenting L^p spaces. Hilbert spaces are the core of Section 3, in which the projection theorem and Riesz lemma are exposed. In Section 4, distributions are classically discussed in connection with functional analysis to show how this generalization of functions is useful for expressing solutions of partial differential equations. Finally, Section 5, we introduce Sobolev spaces which offer a convenient setting for investigating partial differential equations.

[1] Students and persons interested in biographical notes about the "founders" of functional analysis will read with profit the comprehensive introduction to functional analysis by Karen Saxe [Sax01].

1.1 Elementary topological spaces

Before dealing with Banach and Hilbert spaces, we shall first recall some basic notions and results from metric and topological spaces. The reader must keep in mind that important issues about partial differental equations concern the notions of *convergence* (or *limits*) and *continuity* of functions. For instance, let consider a set of points X with a notion of distance between any two points of X. The convergence of a sequence of points $(x_n)_{n\geq 1} \subset X$ to a point $x \in X$ consists in measuring the distance from x_n to x and looking if this distance tends to 0 as n tends to infinity. There is a general setting for this concept.

1.1.1 Metric spaces

The notion of abstract metric space is due to M. Fréchet[2] (1878-1973).

Definition 1.1.1. *A* metric space *is a couple* (X, d)*, where* X *is a set and* d *is a* metric *or a* distance function *on* X*,* i.e., $d : X \times X \to \mathbb{R}^+$ *is such that*

1. *for any* $x, y \in X$, $d(x, y) \geq 0$ *(*non-negativity*)*
2. *for any* $x, y \in X$, $d(x, y) = 0$ *if and only if* $x = y$ *(*identity*)*
3. *for any* $x, y \in X$, $d(x, y) = d(y, x)$ *(*symmetry*)*
4. *for any* $x, y, z \in X$, $d(x, z) \leq d(x, y) + d(y, z)$ *(*triangle inequality*).*

If identity 2 does not hold, then d is then called a *semi-metric*. Usually, only three conditions are used to define a distance function. Indeed, the first of these conditions is a property that follows from the other three, since:

$$2d(x, y) = d(x, y) + d(y, x) \geq d(x, x) = 0\,.$$

Furthermore, the *inverse triangle inequality* is straightforward to obtain

$$|d(x, y) - d(z, y)| \;\leq\; d(x, z)\,.$$

In this definition of metric space, the nature of the elements in the space is not significant. For most problems in this textbook, metric spaces of functions will be considered, when looking for solutions of partial differential equations.

Sequence spaces and function spaces

In the sequel, we will often deal with the following infinite dimensional metric spaces of real or complex sequences,

1. the metric space ℓ^∞ of all bounded sequences $(x_n)_{n\geq 1}$, for the metric

$$d(x, y) = \sup_i |x_i - y_i|\,;$$

[2] M. Fréchet, *Sur quelques points du calcul fonctionnel*, Rendic. Circ. Mat. Palermo, **22**, 1-74, (1906).

2. the metric space c_0 of all sequences that converges to 0 with the metric of ℓ^∞ (as we observe that $c_0 \subset \ell^\infty$);
3. the metric space ℓ^p $(1 \leq p < \infty)$ consisting of all sequences $(x_n)_{n\geq 1}$ such that $\sum_{i=1}^{\infty} |x_i|^p < \infty$, for the metric

$$d(x,y) = \left(\sum_{i\geq 1} |x_i - y_i|^p \right)^{1/p} \quad ;$$

   In this regard, the space ℓ^1 is the space of all absolutely convergent sequences, *i.e.,* $(x_n)_{n\geq 1}$ is in ℓ^1 if the series $\sum_{i\geq 1} |x_i|$ converges. Probably the most important among all ℓ^p-spaces is the space ℓ^2.

Let Ω be any closed, bounded domain in $\mathbb{R}^n$. A natural measure of the discrepancy between two continuous functions f and g is given by

$$d(f,g) = \sup_{x\in\Omega} |f(x) - g(x)| \,, \quad \text{for all } x \in \Omega \,. \tag{1.1}$$

The subspace of all continuous functions on Ω supplied with the metric (1.1) is denoted $(C^0(\Omega), |\cdot|)$, or simply $C^0(\Omega)$. By extension, the space of all continuous functions on a closed, bounded domain Ω whose derivatives up to order k are continuous are denoted by $C^k(\Omega)$. It is a metric space for the distance function

$$d(f,g) = \sum_{|\alpha|\geq k} \sup_{x\in\Omega} |D^\alpha f(x) - D^\alpha g(x)| \,,$$

where we introduced the classical differential notation

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} \quad \text{with} \quad |\alpha| = \sum_{i=1}^n \alpha_i \,. \tag{1.2}$$

Let $\Omega \subset \mathbb{R}^n$ be a compact domain, Jordan measurable, we can consider another metric on $C^0(\Omega)$

$$d(f,g) = \left(\int_\Omega |f(x) - g(x)|^p dx \right)^{1/p} \,, \quad (p \geq 1) \,,$$

where d defines a metric thanks to the Minkowski inequality (see Section 1.2)

$$\left(\int_\Omega |f(x) + g(x)|^p dx \right)^{1/p} \leq \left(\int_\Omega |f(x)|^p \right)^{1/p} + \left(\int_\Omega |g(x)|^p \right)^{1/p} \,, \quad p \geq 1 \,.$$

The *support* of a function $f : X \to \mathbb{C}^n$, denoted by $\mathrm{supp}(f)$, is the closure of all points x such that $f(x) \neq 0$.

1.1.2 The topology of metric spaces

A specific subset of points in X containing a given point $x \in X$ defines a neighborhood of x. Let (X, d) be a metric space and r a stricly positive scalar value. The set $B_r(x) = \{y \in X : d(x, y) < r\}$ is called the *open ball* around x with radius r. A point $x \in U \subset X$ is called an *interior point* of U if U contains some ball around x and then U is called a *neighborhood* of x. A point x is called a *limit point* of U if $(B_r(x)\backslash\{x\}) \cap U \neq \emptyset$ for every ball around x. A point x is called an *isolated point* of U if there exists a neighborhood of x not containing any other point of U. A set U is *dense* in M if every point of M is a limit point of U.

The *closure* of U, denoted by $\bar{U}$, is the set of points $x \in X$ such that any open ball $B_r(x)$ $(r > 0)$ contains a point of U, *i.e.,* U with its limits points. The *interior* of U, denoted by $\overset{\circ}{U}$ or $int(E)$, is the set of points $x \in X$ such that there exists an open ball $B_r(x)$ $(r > 0)$ which is contained in U, *i.e.,* the set of interior points of U.

The set U is *bounded* if for each $x \in U$, there exists $r > 0$ such that $U \subseteq B_r(x)$. The set U is *totally bounded* if and only if, for any $r > 0$, there exists a finite cover (U_i) of U with balls of radius r (each set U_i in the family is of size r or less). In particular, we have the properties

1. U is open if and only if $U = \overset{\circ}{U}$;
2. U is closed if and only if $\bar{U} = U$ and
3. U is dense in X if and only if $\bar{U} = X$.

A set U containing only interior points is called *open*, it is then a union of open balls. Its complement is called *closed*. A family of sets is called a *cover* of U, if U is contained in the union of these sets. It is an *open cover* if each set is open. If C is a cover of U, a *subcover* of U is a subset of C that still covers U. A set U is *compact* if its of each open cover of U contains a finite subcover. A set U is *sequentially compact* if every sequence of U contains a convergent subsequence.

Proposition 1.1.1. *For any subset of the Euclidean space* $\mathbb{R}^n$*, the following four conditions are equivalent*

1. *Every open cover has a finite subcover.*
2. *Every sequence in the set has a convergent subsequence, the limit point of which belongs to the set.*
3. *Every infinite subset of the set has at least one accumulation point in the set.*
4. *The set is closed and bounded.*

In other spaces, these conditions may or may not be equivalent, depending on the properties of the space. The balls generate a topology on X, making it a *topological space*.

Definition 1.1.2. *A* topological space *is a couple* (X, T) *where* X *is a nonempty set of points and* T *is a collection of open subsets of* X *satisfying the axioms*

1. *The emptyset and* X *are in* T*;*
2. *The intersection of any finite collection of sets in* T *is also in* T*;*
3. *The union of any collection of sets in* T *is also in* T*.*

The collection T *is called a* topology *on* X*.*

If X is any nonempty set, there are usually different choices for the topology T. Two interesting choices are

1. $T = \{\emptyset, X\}$, called the *trivial topology*;
2. $T = P(X)$ (the *power set* of X) that consists of the collection of all subsets of X, called the *discrete topology.*

Given two topologies T_1 and T_2 on X, T_1 is called *coarser* (or *weaker*) than T_2 if and only if $T_1 \subseteq T_2$, and T_2 is then *finer* than T_1. We consider the following useful topological results.

Proposition 1.1.2. *Given a metric space* (X, d)*, the following assertions hold*

1. *The sets* $\emptyset$ *and* X *are both open and closed.*
2. *A set* O *in* (X, d) *is* open *if and only if its complement* $O^c = X \backslash O$ *is closed.*
3. *Any finite intersection of open sets is open.*
4. *Any intersection of closed sets is closed.*
5. *The union of any finite number of closed sets is closed.*
6. *If* E *is a compact subset of* (X, d)*, then* E *is closed.*

1.1.3 Separability, compactness and completeness

A metric space (X, d) is *separable* if it contains a countable dense subset, *i.e.* a subset with a countable number of elements whose closure is the space itself.

Recall that a sequence $(x_n)_{n \geq 1}$ of elements in X is said to be *Cauchy* (or is called a *Cauchy sequence*) if given any $\varepsilon > 0$, there exists an integer n_0 such that for all $m \geq n_0$, $d(x_n, x_m) \leq \varepsilon$, whenever n. Interestingly, any convergent sequence is Cauchy. While in $\mathbb{R}^n$ the converse is true, there exists metric spaces in which the converse does not hold.

A subset U of (X, d) is called *complete* if and only if every Cauchy sequence in U converges to a point of U and (X, d) is *complete* if and only if every Cauchy sequence converges. Hence, we deduce that a *sequentially compact* space is also *complete*.

Lemma 1.1.1. *Suppose* $f \in C^0(\mathbb{C})$ *is defined on a sequentially compact metric space* X*. Then, the following assertions hold*

1. *the modulus of* f *is bounded and* f *attains its bound.*

2. *f is uniformly continuous,* i.e., *given $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$d(x,y) \leq \delta \quad \Rightarrow \quad |f(x) - f(y)| \leq \varepsilon \,.$$

We recall an important result in Euclidean and arbitrary metric spaces.

Theorem 1.1.1 (Heine-Borel).

- *A subset U of $\mathbb{R}^n$ is* compact *if and only if it is* closed *and* bounded.
- *A subset U of a metric space (X,d) is* compact *if and only if it is* complete *and* totally bounded.

and we can summarize

Corollary 1.1.1 (Equivalent forms of compactness). *Let (X,d) be a metric space. The following properties are equivalent*

1. *X is sequentially compact;*
2. *X is complete and totally bounded;*
3. *X is compact.*

Many of the metrics that are of interest for this textbook arise from norms.

1.1.4 Normed spaces

Definition 1.1.3. *A* normed linear space *over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) is a linear space V together with a mapping $\|\cdot\| \to \mathbb{R}$ called a* norm *satisfying*

1. *for any $v \in V$, $\|v\| \geq 0$,* *(*nonnegativity*)*
2. *$\|v\| = 0$ if and only if $v = 0$,* *(*nondegeneracy*)*
3. *for every $v \in V$ and $\lambda \in \mathbb{K}$, $\|\lambda v\| = |\lambda|\,\|v\|$,* *(*multiplicativity*)*
4. *for every $v, w \in V$, $\|v + w\| \leq \|v\| + \|w\|$,* *(*triangle inequality*).*

Norms give always rise to metrics. Indeed, if $(V, \|\cdot\|)$ is a normed space, we define a metric d on V by posing

$$d(v,w) = \|v - w\| \,.$$

Hence, (V,d) is a metric space and all notions associated to metric spaces apply, notably continuity, compactness and completness. But notice that not all metrics come from norms.

Hereafter are some basic examples of normed spaces.

1. $V = \mathbb{C}$ with $\|z\| = |z|$.
2. $V = \mathbb{R}^n$ with $\|x\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}$ for $x = (x_1, \ldots, x_n)$. This is the *Euclidean norm* on $\mathbb{R}^n$. Other norms on $\mathbb{R}^n$ can be also defined

$$\|x\|_1 = |x_1| + \cdots + |x_n| \,, \qquad \|x\|_\infty = \max_{1 \leq i \leq n} (|x_i|) \,.$$

3. Let Ω be a closed, bounded subset of $\mathbb{R}^n$. The function
$$\|f\| = \max_{x\in\Omega} |f(x)|$$
defines a norm on $C^0(\Omega)$.
4. Let $x = (x_n)_{n\geq 1} \in \ell^p$. The function defined by
$$\|x\| = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} \tag{1.3}$$
is a norm in ℓ^p. Furthermore, we observe that the Minkowski inequality is the triangle inequality for this norm, *i.e.,* given $(x_n)_{n\geq 1}, (y_n)_{n\geq 1} \in \ell^p$,
$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{1/p} \leq \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{1/p} .$$

Let $(V, \|\cdot\|)$ be a normed space. A sequence (x_n) of elements of V converges to $x \in V$, if for every $\varepsilon > 0$ there exists n_0 such that for every $n \geq n_0$, $\|x_n - x\| \leq \varepsilon$. Consider the space $C^0(\Omega)$ of all continuous functions defined on a closed bounded set $\Omega \subset \mathbb{R}^n$. A sequence of function $(f_n)_{n\geq 1}$ is *uniformly convergent* to f if for every $\varepsilon > 0$ there exists a constant n_0 such that for all $x \in \Omega$ and for all $n > n_0$, we have $|f_n(x) - f(x)| \leq \varepsilon$. Notice that the norm $\|f\|_\infty = \max_{x\in\Omega} |f(x)|$ defines the uniform convergence and is often called the *uniform norm*.

The normed space V is *imbedded* in the normed space W, and we write $V \to W$ if the following conditions are satisfied

1. V is a vector subspace of W
2. the identity operator I defined on V to W by $Ix = x$ for all $x \in V$ is continuous.

A normed space is *complete* if it is complete for the corresponding metric. A complete normed space is called a *Banach space*. It can then be checked that the uniform norm makes $C^0(\mathbb{C})$ into a Banach space. If $I \subseteq \mathbb{R}$ is a compact interval, then $C^0(I)$ with the maximum norm is a Banach space. Likewise, ℓ^p is a Banach space for $1 \leq p < \infty$, for the norm defined by (1.3).

Lemma 1.1.2. *If S is a closed subspace of a Banach space V and M is a finite dimensional subspace, then $S + M$ is closed.*

A linear map A between two normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ is called a *linear operator*. The *kernel* and the *range* (or *image*) of A are the sets defined as usual by
$$\mathrm{Ker}(A) = \{f \in V\,,\ \ Af = 0\} \quad \text{and} \quad \mathrm{Im}(A) = \{g \in W\,,\ \ g = Af\}\,.$$
The operator A is called *bounded* if the operator norm

$$\|A\| = \sup_{\|f\|_V=1} \|Af\|_W$$

is finite. The set of all bounded linear operators from V to W is denoted by $\mathcal{L}(V, W)$ or $\mathcal{L}(V)$ if $V = W$. It is a normed space with the operator norm and a Banach space if W is a Banach space.

Lemma 1.1.3. *An operator A is bounded if and only if it is continuous.*

An operator in $\mathcal{L}(V, \mathbb{C})$ is called a *bounded linear functional* and the space $V' = \mathcal{L}(V, \mathbb{C})$ is called the *dual space* of V. The space V' is sometimes denoted by V^*.

Since many norms arise from inner product, we shall now define this notion.

1.1.5 Inner product spaces

Definition 1.1.4. *Let V be a linear vector space on $\mathbb{K}$. An inner product on V is a mapping $(\cdot,\cdot) : V \times V \to \mathbb{K}$ satisfying the conditions*

1. $(v, v) \geq 0$ *for all* $v \in V$; (nonnegativity)
2. $(v, v) = 0$ *if and only if* $v = 0$; (nondegeneracy)
3. $(\lambda v, w) = \lambda(v, w)$ *for all* $v, w \in V$ *and* $\lambda \in \mathbb{K}$; (muliplicativity)
4. $(v, w) = \overline{(w, v)}$ *for all* $v, w \in V$; (Hermitian symmetry)
5. $(v, w + u) = (v, w) + (v, u)$ *for all* $u, v, w \in V$. (distributivity)

The inner product is sometimes denoted by $\langle \cdot,\cdot \rangle$ and may be called a *scalar product*. A vector space with an inner product is called an *inner product space* or a *pre-Hilbert space*. Inner products always allow to define norms. Indeed, if $(V, (\cdot,\cdot)$ is an inner product space, we define a norm $\|\cdot\|$ on V by posing

$$\|v\| = \sqrt{(v, v)}\,.$$

Hence, $(V, \|\cdot\|)$ is a normed space.

Hermitian symmetry and linearity in the first variable give

$$(v, \lambda w) = \overline{(\lambda w, v)} = \bar{\lambda}\overline{(w, v)} = \bar{\lambda}(v, w)$$

as well as

$$(v, w + u) = \overline{(w + u, v)} = \overline{(w, v)} + \overline{(u, v)} = (v, w) + (v, u)\,,$$

and thus an inner product is a *sesquilinear form*. Subsequently, an inner produt on a real vector space is a *positive-definite symmetric bilinear form*. Clearly, the complex linear space $\mathbb{C}^n$ with the usual inner product $(z, w) = \sum_{i=1}^n z_i \bar{w}_i$ is an inner product space.

Lemma 1.1.4 (Cauchy-Schwarz inequality). *If $((V, (\cdot,\cdot))$ is an inner product space, then for all $v, w \in V$ we have*

$$|(v, w)| \leq \|v\|\,\|w\|\,.$$

It is easy to see that the equality only occurs if v and w are colinear.

Lemma 1.1.5 (Jordan-von Neumann). *A norm is associated with a inner product if and only if the parallelogram rule holds,* i.e.,

$$2\|v\|^2 + 2\|w\|^2 = \|v+w\|^2 + \|v-w\|^2$$

Strong and weak convergence

Since every inner product space is a normed space for the naturally defined norm $\|x\| = \sqrt{(x,x)}$, the notion of convergence is well defined.

A sequence of vectors $(x_n)_{n\geq 1}$ in an inner product space V is said to *converge strongly* (or to converge in the norm) to a vector v in V if $\|x_n - x\| \to 0$ as $n \to \infty$. A sequence of vectors $(x_n)_{n\geq 1}$ in an inner product space V is said to *converge weakly* to a vector v in V if $(x_n, y) \to (x, y)$ as $n \to \infty$, for all y in V.

Weakly convergent sequences are bounded, *i.e.,* there exists $M > 0$ such that $\|x_n\| \leq M$ for all n. The notion of weak convergence defines a topology on V that is called the *weak topology* on V. From Cauchy-Schwarzs inequality, we can deduce that the weak topology is weaker than the norm topology. Hence, a strongly convergent sequence is also weakly convergent to the same limit, while the converse is not true in general. However, if $(x_n, x) \to)x, x)$ and $\|x_n\| \to \|x\|$, then we have $\|x_n - x\| \to 0$ as $n \to \infty$.

Complete inner product spaces

A complete inner product space is called a *Hilbert space*. Next are some interesting inner product and Hilbert spaces

1. if we are considering continuous functions defined over $\mathbb{C}$, $C^0(I)$, where $I = [a, b]$ is a compact subset of $\mathbb{C}$, is an inner product space endowed with the inner product
$$(f, g) = \int_a^b f(x)\overline{g(x)}dx$$
and then the induced norm is then
$$\|f\|_2 = \left(\int_a^b |f(x)|^2 dx\right)^{1/2},$$
and not the supremum norm $\|f\|_\infty = \sup_{x\in I} |f(x)|$.
2. the space ℓ^2 of complex valued sequences $x = (x_n)_{n\geq 1}$ is a Hilbert space endowed with the inner product
$$(x, y) = \sum_{n=1}^{\infty} x_n \overline{y_n}\,.$$

A vector v in a vector space V is called *normalized* or *unit vector* if $\|v\| = 1$. Two vectors v, w in a vector space V are *orthogonal* ($v \perp w$) if $(v, w) = 0$ and *parallel* if one is a multiple of the other. If v and w are orthogonal, we have the *Pythagorean formula*

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 . \tag{1.4}$$

For v and $w \neq 0$ in an inner product space, the *projection* of v on w is the vector

$$\tilde{v} = \frac{(v, w)}{\|w\|^2} w , \qquad \text{or, if } \|w\| = 1 \quad \tilde{v} = (v, w) w .$$

We recall that a set C in a vector space is *convex* if for any $x, y \in C$, we have $tx + (1 - t)t \in C$, with $t \in [0, 1]$.

Lemma 1.1.6. *Let C be a nonempty convex and complete linear subspace of a inner product space V. If $v \in V$ then there exists a unique $u \in C$ minimizing $\|v - u\|$ and called the* closest point *(or* best approximation*) to v from C. Furthermore, we have $v - u \perp C$,* i.e., *$(v - u, w) = 0$ for any $w \in C$.*

Given an inner product space V and $M \subset V$, we define the *orthogonal space* of M by

$$M^{\perp} = \{v \in V\,;\; (v, m) = 0 \;\text{ for all } m \in M\} .$$

and this space $M^{\perp}$ is often referred to as *M-perp*.

Lemma 1.1.7. *Suppose V is an inner product space and M a complete linear subspace of V. Then, $V = M \oplus M^{\perp}$, the direct sum being orthogonal.*

Proposition 1.1.3. *Suppose V is an inner product space and $M \subset V$. Then $M^{\perp}$ is a linear subspace of V, $M \perp M^{\perp}$ and the intersection $M \cap M^{\perp}$ is either $\{0\}$ or the emptyset.*

Before examining in more detail the properties of Hilbert spaces, we like to introduce the Lesbesgue integral, a fundamental concept for understanding most applications Hilbert space theory.

1.2 The Lebesgue integral

We shall recall that one purpose of this textbook is to introduce the main numerical methods for solving partial differential equations. Since the solutions to these equations involve Hilbert spaces, it is important to show that all differentiable functions defined on a compact interval $I = [a, b]$ belong to a Hilbert space endowed with the inner product

$$(f, g) = \int_a^b f(x)\overline{g(x)}dx .$$

And precisely, the smallest of such spaces is the space of Lebesgue square integrable functions on I. Hence, we will briefly review now this concept of Lebesgue integral[3].

1.2.1 Lebesgue integration

The classical Riemann integral is suitable for dealing with continuous functions defined on bounded subsets of the Euclidean space $\mathbb{R}^n$, or functions with a limited number of discontinuities. However, it cannot handle discontinuous functions. The general setting of the measure theory is suitable for resolving these drawbacks, the integral is then defined from the notion of size in some set V.

Note to the reader.

In this textbook, we are mainly concerned with integration and not measure. Moreover, a complete description of the measure theory is beyond the scope of this book, artly dedicated to undergraduate students. We have thus decided to introduce the Lebesgue integral without referring explicitly to concepts like measure, following a more direct approach pionnered by [DM90]. A reader interested in this topic will consult with profit the books listed in appendix of this chapter. Hence, graduate students and mathematicians could quietly skip this section.

A real valued f defined on $\mathbb{R}$ is called a *step function* if it can be written as a finite linear combination of characteristic functions of semi-open intervals $A_i = [a_i, b_i[\subseteq \mathbb{R}$, *i.e.*,

$$f(x) = \sum_{k=1}^{n} \alpha_k \chi_{A_k}(x)\,, \quad \text{for all x} \in \mathbb{R}\,, \tag{1.5}$$

where $\alpha_k \in \mathbb{R}$ and χ_A is the *characteristic function* of A, such that $\chi_{A_k}(x) = 1$ if $x \in [a_k, b_k[$ and $\chi_{A_k}(x) = 0$ otherwise. In this setting, the intervals A_k are assumed to be disjoint and if we consider besides a minimal number of intervals, then the representation of f is *unique*[4]. It enjoys several properties, notably

1. if f, g are step functions, then $f + g$ and fg are step functions,
2. if f is a step function and $\alpha \in \mathbb{R}$, then αf is a step function,
3. if f is a step function, then $|f|$ is a step function.
4. if f, g are step functions, $\min(f, g)$ and $\max(f, g)$ are step functions.

[3] Named after the French mathematician Henri Lebesgue (1875-1941) who introduced the theory of integration in his doctoral dissertation, *Intégrale, longueur, aire*, University of Nancy, (1902).

[4] Advanced readers will easily compare this definition with the notion of a *measurable simple function* having a finite range, in the measure theory.

The collection of all step functions is a vector space on $\mathbb{R}$. The derivative of a step function is the *Dirac delta* function $\delta(x) = 0$, if $x \neq 0$ and $\delta(x) = +\infty$ otherwise.

We define the integral of a step function as the Riemann integral of this function, *i.e.,*

$$\int f = \sum_{k=1}^{n} \alpha_k (b_k - a_k) = \int_{-\infty}^{\infty} f(x)dx\,.$$

And we observe that this definition is independent of any particular representation of f. We have then

Lemma 1.2.1. *Given f a step function whose support is contained in $\bigcup_{k=1}^{n} A_k$, where the A_k are disjoint semi-open intervals $[a_k, b_k[$. If, for any $M > 0$ $|f| < M$, then*

$$\int |f| \,\leq\, M \sum_{k=1}^{n} (b_k - a_k)\,.$$

Next, we give two results that will be helpful for defining the Lebesgue integral.

Theorem 1.2.1. *1. Let $(f_n)_{n\geq 1}$ be a non-increasing sequence of nonnegative step functions such that $\lim_{n\to\infty} f_n(x) = 0$ for every $x \in \mathbb{R}$. Then,*

$$\lim_{n\to\infty} \int f_n(x)dx = 0\,.$$

2. Let $(f_n)_{n\geq 1}$ and $(g_n)_{n\geq 1}$ be two non decreasing sequences of step functions. If $\lim_{n\to\infty} f_n(x) \leq \lim_{n\to\infty} g_n(x)$ for every $x \in \mathbb{R}$, then

$$\lim_{n\to\infty} \int g_n(x)dx \,\leq\, \lim_{n\to\infty} \int f_n(x)dx\,.$$

And finally, we introduce the expected definition of the Lebesgue integral.

Definition 1.2.1 (Lebesgue integral). *A real-valued function f defined on $\mathbb{R}$ is called* Lebesgue integrable *if there exists a sequence of step functions $(f_n)_{n\geq 1}$ satisfying the following axioms*

1. $\displaystyle\sum_{n=1}^{\infty} \int |f_n(x)|dx < \infty$;
2. $f(x) = \displaystyle\sum_{n=1}^{\infty} f_n(x)$, *for every $x \in \mathbb{R}$ such that* $\displaystyle\sum_{n=1}^{\infty} |f_n(x)| < \infty$.

The integral *of f is then defined by*

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

In this definition, condition 2 shows that f is equal to the sum of series at points where the series converges absolutely. We will show that the set of all points where f does not coincide with $\sum_{n=1}^{\infty} f_n(x)$ is a small set, called a *null set*. Hence, the series converges to $f(x)$ for all x except a null set. This introduces the concept of *convergence almost everywhere*.

The space of all Lebesgue integrable functions defined on $\mathbb{R}$ is denoted by $L^1(\mathbb{R})$. And we will observe that all Riemann integrable functions are Lebesgue integrable. The space $L^1(\mathbb{R})$ is a vector space and the function $\int$ is a linear functional on $L^1(\mathbb{R})$.

Lemma 1.2.2. *If $f \in L^1(\mathbb{R})$ then $|f| \in L^1(\mathbb{R})$ and we have*

$$\left| \int f(x)dx \right| \leq \int |f(x)|dx \,.$$

Corollary 1.2.1. *Given $f, g \in L^1(\mathbb{R})$,*

1. *if $f = \sum_{n=1}^{\infty} f_n$, then $\int |f(x)|\, dx \leq \sum_{n=1}^{\infty} \int |f_n(x)|\, dx$;*
2. $\min(f,g) \in L^1(\mathbb{R})$ *and* $\max(f,g) \in L^1(\mathbb{R})$.

And we have the following result.

Theorem 1.2.2. *If $f \in L^1(\mathbb{R})$, then for t scalar we have*

$$\lim_{t \to 0} \int |f(x+t) - f(x)|dx = 0 \,.$$

Proof. If f is a step function, the results is immediate. Suppose f is an arbitrary Lebesgue integrable function. Then, given $\varepsilon > 0$, if $f = \sum_{n>0}^{\infty} f_n$ there exists n_0 such that $\sum_{n>n_0}^{\infty} \int |f_n| < \varepsilon/3$, and we have

$$\begin{aligned}
\int |f(x+t) - f(x)|dx &\leq \int \left| \sum_{n=1}^{n_0} f_n(x+t) - \sum_{n=1}^{n_0} f_n(x) \right| dx \\
&\quad + \sum_{n>n_0} \int |f_n(x+t)|dx + \sum_{n>n_0} \int |f_n(x)|dx \\
&= \int \left| \sum_{n=1}^{n_0} f_n(x+t) - \sum_{n=1}^{n_0} f_n(x) \right| dx \\
&\quad + 2 \sum_{n>n_0} \int |f_n(x)|dx \\
&< \int \left| \sum_{n=1}^{n_0} f_n(x+t) - \sum_{n=1}^{n_0} f_n(x) \right| dx + 2\varepsilon/3 \,.
\end{aligned}$$

As we know that $\sum_{n=1}^{n_0} f_n$ is a step function, we can then deduce that

$$\lim_{t\to 0}\int\left|\sum_{n=1}^{n_0} f_n(x+t)-\sum_{n=1}^{n_0} f_n(x)\right|dx=0\,,$$

and thus $\int|f(x+t)-f(x)|dx<\varepsilon$ for t sufficiently small. □

Interestingly, $L^1(\mathbb{R})$ can be considered as a Banach space, under specific assumptions, and norms can be defined. To this end, we propose now a definition for the null function previously evoked.

Definition 1.2.2 (null function and null set).

1. *An integrable function f is called a* null function *if $\int|f|=0$. Furthermore, two integrable functions f,g are* equivalent *if $f-g$ is a null function.*
2. *A set $X\subseteq\mathbb{R}$ is called a* null set *or* measure-zero set*, if its characteristic function is a null function.*

Under this definition, any countable set if a null set and a countable union of null sets is also a null set. For example, the set of rational numbers $\mathbb{Q}$ is a null set with respect to $\mathbb{R}^n$, despite being dense in $\mathbb{R}^n$. All subsets of $\mathbb{R}^n$ of dimension smaller than n are null sets in $\mathbb{R}^n$. A classical example of a null set which is not countable is the *Cantor set* [Hal50]. A set is considered null if it is a subset of a null set.

1.2.2 Notions of convergence

In measure theory, a property hold *almost everywhere*, abbreviated *a.e.*, if the set of elements for which this property is not satisfied is a *null set*. Hence, if $f,g\in L^1(\mathbb{R})$ and if the set of elements $x\in\mathbb{R}$ for which $f(x)\neq g(x)$ is a null set, then f *equals* g *almost everywhere*, *i.e.,* $f=g$ a.e.

At this point, we can introduce the equivalence class of $f\in L^1(\mathbb{R})$ as the set of all functions $g\in L^1(\mathbb{R})$ which are equivalent to f, *i.e.,*

$$[f]=\{g\in L^1(\mathbb{R})\,,\ \int|f-g|=0\}\,,$$

and then consider the space $\mathcal{L}^1(\mathbb{R})$ of all equivalence classes of Lebesgue integrable functions, endowed with the norm

$$\|[f]\|=\int|f|\,.$$

Then, the space $(\mathcal{L}^1(\mathbb{R}),\|\cdot\|)$ is a *normed space*. We recall the definition of convergence in a normed space.

Definition 1.2.3. *A sequence of functions $(f_n)\in L^1(\mathbb{R})$ converges in norm to $f\in L^1(\mathbb{R})$, and we denote by $f_n\to f$ i.n., if $\|f_n\to f\|\to 0$.*

According to this definition, if $f_n \to f$ i.n., then $|f_n| \to |f|$ i.n. and $\int f_n \to \int f$. A sequence of function $(f_n) \in L^1(\mathbb{R})$ converges to a function $f \in L^1(\mathbb{R})$ almost everywhere, denoted by $f_n \to f$ a.e., if $f_n(x) \to f(x)$ for every x except a null set.

We introduce two important results regarding $L^1(\mathbb{R})$ spaces. It relates the notions of convergence in norm and convergence almost everywhere.

Theorem 1.2.3 (Riesz).

1. *The space $L^1(\mathbb{R})$ is* complete*;*
2. *Given a sequence (f_n). If $f_n \to f$ in norm, then there exists a subsequence (f_{p_n}) of (f_n) such that $f_{p_n} \to f$ a.e.*

It comes directly that the limit with respect to the convergence in norm can be interchanged with the integration, thus leading to write

$$\int \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int f_n \,.$$

And we shall notice here that this property does not hold when considering the convergence almost everywhere. The next results illustrate the main difference between Lebesgue integration and other integrations.

Theorem 1.2.4 (Lebesgue's monotone convergence). *Let (f_n) be a nondecreasing sequence of nonnegative integrable functions,* i.e.*, such that for every $k \geq 1$,*

$$0 \leq f_k(x) \leq f_{k+1}(x)\,, \quad \text{for almost every } x \in \mathbb{R}\,.$$

Let f be defined as the pointwise limit of the sequence, $f(x) = \lim_{n\to\infty} f_n(x)$. Then f is integrable and

$$\lim_{n\to\infty} \left(\int f_n \right) = \int f \,.$$

Theorem 1.2.5 (Lebesgue's dominated convergence). *Let (f_n) be a sequence of square integrable functions converging almost everywhere to a function f. Moroever, suppose there exists a square integrable function g such that $|f_n| \leq g$ for all n. Then f is integrable and $f_n \to f$ i.n.,* i.e.,

$$\lim_{n\to\infty} \left(\int f_n \right) = \int f \,.$$

Lemma 1.2.3 (Fatou's lemma). *Let (f_n) be a sequence of nonnegative integrable functions and let $f = \liminf_{n\to\infty} f_n$. Then*

$$\int f \leq \liminf_{n\to\infty} \left(\int f_n \right) \,.$$

1.2.3 Locally integrable functions

Until now, we have dealt with the integration over the whole set $\mathbb{R}$, where the integral $\int f$ was meant for $\int_{-\infty}^{+\infty} f$. However, we need to define the integration over bounded intervals. Let $I = [a, b]$ be an interval and $f : I \to \mathbb{R}$ be a function. We denote by $\int_a^b f$ or $\int_I f$ the integral of f over the interval $[a, b]$. It corresponds to the value of the integral of the product $\int f\chi_{[a,b]}$, where $\chi_{[a,b]}$ represents the characteristic function of $[a, b]$. According to this definition, $\int_a^b f$ corresponds to $\int f$ on $[a, b]$ and zero otherwise. In addition, we have the following conventions

$$\int_1^b f = -\int_b^a f \quad \text{and} \quad \int_a^a f = 0\,.$$

We observe that if $f \in L^1(\mathbb{R})$, then for any interval $[a, b]$ on $\mathbb{R}$, $\int_a^b f$ exists. However, the converse may not hold.

Definition 1.2.4. *A* locally integrable *function is a function f defined on $\mathbb{R}$ such that for any compact interval $[a, b]$, the integral $\int_a^b f$ exists.*

Under this definition, the space $L^1(\mathbb{R})$ is a subspace of the space of locally integrable functions that forms a vector space.

Lemma 1.2.4. *Suppose f is a locally integrable function such that $|f| \leq g$ for some function $g \in L^1(\mathbb{R})$. Then, $f \in L^1(\mathbb{R})$.*

1.2.4 Lebesgue vs. Riemann integration

In this section, we assume the reader to be familiar with the definition of the Riemann integral and its properties. We briefly summarize this notion in order to introduce notations.

We consider a bounded function f defined on the closed, bounded interval $[a, b] \subset \mathbb{R}$. Suppose that $a = x_0 < x_1 < \cdots < x_n = b$ is a *partition* of $[a, b]$ together with a finite sequence of real $t_1, \ldots, t_n$ such that for each $k \geq 1$, $x_{k-1} \leq t_k \leq x_k$. The *lower Riemann sum* and *upper Riemann sum* of f with respect to the partition (x_n) are respectively defined by

$$L_n(f) = \sum_{k=1}^{n} m_k(x_k - x_{k-1})\,, \quad \text{and} \quad U_n(f) = \sum_{k=1}^{n} M_k(x_k - x_{k-1})\,,$$

where $m_k = \inf(f(x); x \in [x_{k-1}, x_k])$ and $M_k = \sup(f(x); x \in [x_{k-1}, x_k])$.

A bounded function f defined on $[a, b]$ is called *Riemann integrable* if $L_n(f) = U_n(f)$ and we denote the Riemann integral by

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

An interesting property is that the Riemann integral is considered as the limit of the Riemann sums of a function when the size of the partition (x_n) tends to zero.

The definition of the Lebesgue integral cannot be seen as a generalization of the Riemann integral, but the following result is interesting in this respect.

Theorem 1.2.6. *If f is a Riemann integrable function on $[a,b]$ then f is Lebesgue integrable on $[a,b]$ and both integrals coincide.*

Proof. For every integer n, we consider a partition of $[a,b]$ into 2^n subintervals each of length $(b-a)/2^n$. Next, we define

$$g_n(x) = \sum_{k=1}^{2^n} m_k \chi_{[x_{k-1}-x_k[}(x) , \quad \text{and} \quad h_n(x) = \sum_{k=1}^{2^n} M_k \chi_{[x_{k-1}-x_k[}(x) ,$$

and observe that $(g_n)_{n\geq 1}$ is an increasing sequence while $(h_n)_{n\geq 1}$ is a decreasing sequence. Denoting $g = \lim_{n\to\infty} g_n(x)$ and $h = \lim_{n\to\infty} h_n(x)$ leads to conclude that g and h are Lebesgue integrable functions such that, for almost every x in $[a,b]$

$$g(x) \leq f(x) \leq h(x) .$$

It is not diffuclt to deduce that, for almost all x in $[a,b]$

$$\lim_{n\to\infty} (h_n(x) - g_n(x)) = h(x) - g(x) .$$

Hence, thanks to the monotone convergence theorem, we have

$$\begin{aligned} 0 \leq \int (h-g) &= \lim_{n\to\infty} \int (h_n - g_n) = \lim_{n\to\infty} \int h_n - \lim_{n\to\infty} \int g_n \\ &= \lim_{n\to\infty} U_n(f) - \lim_{n\to\infty} L_n(f) = 0 . \end{aligned}$$

And it follows that $g = h$ a.e., thus f is Lebesgue integrable. Moreover

$$\int_a^b f(x)dx = \lim_{n\to\infty} \int_a^b g_n(x)dx = \int_a^b g = \int_a^b f ,$$

where the integral on the left denotes the Riemann integral and the integral on the right denotes the Lebesgue integral. □

A version of the fundamental teorem for calculus can be given for the Lebesgue integral.

Theorem 1.2.7 (Fundamental theorem of calculus). *If f is Lebesgue integrable on $[a,b]$ and if we define*

$$F(x) = \int_a^x f$$

then the function F is continuous in $[a,b]$ and differentiable a.e. Furthermore, F is differentiable a.e. and $F'(x) = f(x)$ for almost every x in $[a,b]$.

From this theorem, we can also deduce that the limits

$$F(a) = \lim_{x \to a^+} F(x) \quad \text{and} \quad F(a) = \lim_{x \to b^-} F(x)$$

exists and are finite.

The following result allows to introduce a change of variable in the Lebesgue integral in a similar way as it is performed with the Riemann integral.

Lemma 1.2.5 (Change of variables). *Let g be a nondecreasing differentiable function defined on a bounded interval $[a, b]$ such that g' is integrable over $[a, b]$. We pose*

$$g(a) = \lim_{x \to a^-} g(x) \quad \textit{and} \quad g(b) = \lim_{x \to b^+} g(x)\,.$$

Suppose f is an integrable function over $[g(a), g(b)]$. Then, the product $(f \circ g)g'$ is integrable over $[a, b]$ and we have

$$\int_{g(a)}^{g(b)} f(t)\,dt = \int_a^b (f \circ g)(t) g'(t)\,dt\,.$$

1.2.5 The Lebesgue measure on Euclidean space

Now, we are in good condition for defining more general concepts like measure sets and the Lebesgue measure. The later is a classical manner of assigning length, area and volume to subsets of the Euclidean space. But remember that not all sets are measurable.

Definition 1.2.5. *We introduce the notions of measurable set and of measure as follows*

1. *a set A is called* measurable *if the characteristic function of A is a locally integrable function;*
2. *given a measurable set A. If the characteristic function χ_A is an integrable function, then the* measure $\mu(A)$ *of A is defined by*

$$\mu(A) = \int \chi_A\,,$$

and we assign $\mu(A) = \infty$ if χ_A is not integrable.

Null sets are zero-measure sets, this justify the terminology. More generally, mesure μ is a *countably additive* nonnegative function μ.

Proposition 1.2.1. *Let $(A_k)_{k \geq 1}$ be a sequence of disjoint measureable sets. Then $A = \bigcup_{k \geq 1} A_k$ is measurable and we have*

$$\mu(A) = \mu\Big(\bigcup_{k \geq 1} A_k\Big) = \sum_{k \geq 1} \mu(A_k)\,.$$

The integral over any measurable set Ω can be defined by

$$\int_\Omega f = \int f\,\chi_\Omega\,.$$

Definition 1.2.6 (Measurable function). *A function f is called* measurable *if there exists a sequence of step function $(f_n)_{n\geq 1}$ such that $f_n \to f$ a.e.*

Under this definition, every integrable function is measurable. Furthermore, every locally integrable function is measurable. The measurable functions form a vector space and the following properties hold

1. if f is measurable, then $|f|$ is measurable;
2. if f, g are measurable, then $f+g$, fg are measurable;
3. if $(f_n)_{n\geq 1}$ is a sequence of measurable functions, then the functions

$$(\inf f_n)(x) = \inf_{n\geq 1} f_n(x) \qquad \text{and} \quad (\sup f_n)(x) = \sup_{k\geq 1} f_n(x)$$

$$(\liminf f_n)(x) = \sup_{j\geq l}(\inf_{n\geq j} f_n(x)) \quad \text{and} \quad (\limsup f_n)(x) = \inf_{j\geq l}(\sup_{n\geq j} f_n(x))$$

are measurable functions.

1.2.6 More Lebesgue spaces

It is possible to define arbitrary measure spaces and we will now introduce a few other measure spaces.

The space $L^2(\mathbb{R})$

The space of all locally integrable functions f such that $|f|^2 \in L^1(\mathbb{R})$ is denoted by $L^2(\mathbb{R})$. Functions in $L^2(\mathbb{R})$ are also called *square integrable functions*. The space $L^2(\mathbb{R})$ is a vector space and the product of two square integrable functions is also a function of $L^2(\mathbb{R})$. Furthermore, if we consider the norm defined by $\|f\| = \left(\int |f|^2\right)^{1/2}$, then $(L^2(\mathbb{R}), \|\cdot\|)$ is a complete normed space. The space of square integrable functions that vanish outside an interval $[a,b]$ is denoted by $L^2([a,b])$.

The spaces $L^1(\mathbb{R}^n)$ and $L^2(\mathbb{R}^n)$

We like to consider the Euclidean space $\mathbb{R}^n$ as the measure space. Given $(a_k)_{1\leq k\leq n}$ and $(b_k)_{1\leq k\leq n}$ in $\mathbb{R}^n$, with each $a_k \leq b_k$, we consider subsets of $\mathbb{R}^n$ of the form $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$. For a subset I, we define

$$m(I) = \prod_{k=1}^{n} (b_k - a_k)\,.$$

And we observe that $m(I)$ represents the *length*, *area* and *volume* of I when $n = 1, 2$ or 3, respectively.

We have introduced the notion of *step functions* as real-valued functions that have only a finite number of elements in their range (cf. Definition 1.5). Every such function can be decomposed as a linear combination of characteristic functions. A step function is measurable if and only if each set A_i is a measurable set. Such a function is then called a *simple measurable function*. Every function defined on $\mathbb{R}^n$ can be approximated by simple functions. For a step function f, we write

$$f = \sum_{k=1}^{n} \alpha_k \chi_{I_k} ,$$

and we define

$$\int f = \sum_{k=1}^{n} \alpha_k m(I_k) .$$

Next we introduce the notion of Lebesgue integrable function, that expands naturally Definition 1.2.1.

Definition 1.2.7 (Lebesgue integrable function on $\mathbb{R}^n$). *A real (or complex) valued function f defined on $\mathbb{R}^n$ is called* Lebesgue integrable *if there exists a sequence of step functions $(f_n)_{n\geq 1}$ satisfying the followings axioms*

1. $\displaystyle\sum_{n=1}^{\infty} \int |f_n| < \infty$;
2. $\displaystyle f(x) = \sum_{n=1}^{\infty} f_n(x)$ *for every* $x \in \mathbb{R}^n$ *such that* $\displaystyle\sum_{n=1}^{\infty} |f_n(x)| < \infty$.

The integral *of f is then defined by*

$$\int f = \sum_{n=1}^{\infty} \int f_n .$$

The space of all Lebesgue integrable functions on $\mathbb{R}^n$ is denoted by $L^1(\mathbb{R}^n)$. Likewise, we extend Definition 1.2.4 as follows.

Definition 1.2.8. *A function f defined on $\mathbb{R}^n$ is* locally integrable*, if for every bounded interval I the product $f\chi_I$ is an integrable function.*

By analogy with the previous section, a set $A \subseteq \mathbb{R}^n$ is called *measurable* if the characteristic function of A is a locally integrable function. The *measure* $\mu(A)$ of A is then defined by the value of the integral $\mu(A) = \int \chi_A$ and $\mu(A) = \infty$ is χ_A is locally integrable but not integrable. Finally, a function f defined on $\mathbb{R}^n$ is *measurable* is there exists a sequence of step functions $(f_n)_{n\geq 1}$ such that $f_n \to f$ a.e.

The integral over a measurable set $\Omega \subseteq \mathbb{R}^n$ is defined by the integral of the function f on Ω and 0 everywhere else,

$$\int_{\Omega} f = \int f \chi_{\Omega} \,.$$

The space of locally integrable functions f defined on $\mathbb{R}^n$ such that $|f|^2$ is integrable is denoted by $L^2(\mathbb{R}^n)$. Functions in $L^2(\mathbb{R}^n)$ are called *square integrable*, The space of square integrable functions that vanish outside an subset $\Omega \subseteq \mathbb{R}^n$ is denoted by $L^2(\Omega)$.

The spaces $L^p(\mathbb{R}^n)$

Finally, for a bounded subset $\Omega \subseteq \mathbb{R}^n$, we consider the space $L^p(\Omega)$, for $1 \leq p < \infty$, of all real-valued Lebesgue measurable functions defined on Ω such that

$$\int_{\Omega} |f|^p < \infty \,.$$

And we define, for $p = \infty$, the space $L^\infty(\Omega)$ of all real-valued Lebesgue measurable functions that are *essentially bounded*, *i.e.,* $\operatorname{ess\,sup}_{x \in \Omega} |f(x)| < \infty$, where

$$\operatorname*{ess\,sup}_{x \in \Omega} |f(x)| = \inf\{M \geq 0\,,\ |f(x)| \leq M,\ \text{a.e. in } \Omega\} \,.$$

We observe that the spaces L^p form a sequence of embedded spaces, *i.e.,*

$$L^\infty(\Omega) \subset \cdots \subset L^p(\Omega) \subset \cdots \subset L^2(\Omega) \subset L^1(\Omega) \,.$$

Actually, the space L^p consists of equivalence classes of functions, where two functions belong to the same equivalence class if they coincide almost everywhere. For $1 \leq p < \infty$, the space L^p is a linear space. For any $1 \leq p < \infty$ there exists a unique q such that

$$\frac{1}{p} + \frac{1}{q} = 1 \,, \quad \text{with} \quad q = \infty \,,\ \text{if } p = 1 \,.$$

The number q is called the *Hölder conjugate* of p.

Lemma 1.2.6 (Hólder's inequality). *Suppose $1 < p < \infty$ and $1 < q < \infty$ are Hölder conjugates. Given $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$ then $fg \in L^1(\Omega)$ and*

$$\|fg\|_{L^1(\Omega)} \ = \int_{\Omega} |fg| \leq \|f\|_{L^p(\Omega)} \, \|g\|_{L^q(\Omega)} \,, \qquad \frac{1}{p} + \frac{1}{q} = 1 \,.$$

Theorem 1.2.8. *For $1 \leq p < \infty$, the space $L^p(\Omega)$ is a normed linear space, with the norm defined by*

$$\|f\|_{L^p(\Omega)} = \left(\int_{\Omega} |f|^p \right)^{1/p} \quad \textit{and} \quad \|f\|_{L^\infty(\Omega)}\| = \operatorname*{ess\,sup}_{x \in \Omega} |f(x)| \,. \qquad (1.6)$$

In particular, the triangle inequality holds

$$\|f+g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)}\,,$$

that is called *Minkowski's inequality* in this context. Furthermore, a normed linear space $(V, \|\cdot\|)$ is complete if and only if $\sum_{j\geq 1} f_j$ converges in norm whenever $\sum_{j\geq 1}\|f_j\|$ converges. According to this result, the space $L^p(\Omega)$ is *complete* and thus is a *Banach space* for the norm $\|\cdot\|_{L^p(\Omega)}$, for any value p.

The most important L^p spaces for our purposes are the spaces where the measure μ is the Lebesgue measure on some subset Ω of $\mathbb{R}^n$ or μ is a counting measure on $\mathbb{N}$. In this case, we have already introduced these spaces, denoted by ℓ^p, as the spaces of all bounded sequences $(x_n)_{n\geq 1}$ satisfying

$$\sum_{n\geq 1} |x_n|^p < \infty\,, \quad \text{with} \quad \|(x_n)_{n\geq 1}\|_p = \left(|x_n|^p\right)^{1/p}\,.$$

We close this section on Lebesgue functions by giving an interesting result.

Theorem 1.2.9. *The step functions are dense in* $L^p(\Omega)$*, for each* $1 \leq p < \infty$.

It remains to be stated that the space $L^2(\Omega)$ is a Hilbert space and its norm is induced by an inner product. This can be seen by writing

$$(f,g) = \int_\Omega f\,g\,.$$

This space will be discussed in the next section.

1.3 Hilbert spaces

The work of David Hilbert (1862-1943) on quadratic forms in infinitely many variables in his study of integral equations impelled the theory of Hilbert spaces. Their importance was first recognized years later by John von Neumann (1903-1957) in his work on unbounded Hermitian operators and he is credited for having developed the modern theory of Hilbert spaces.

The space $L^2([a,b])$ is a Hilbert space. We have seen that $L^2([a,b])$ is a normed space, then we still have to show that it is complete. Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $L^2([a,b])$, we have

$$\int_a^b |f_m - f_n|^2 \to 0 \quad \text{as} \quad m,n\to\infty\,.$$

Cauchy-Schwarz's inequality yields

$$\int_a^b |f_m - f_n| \leq \left(\int_a^b 1 \int_a^b |f_m-f_n|^2\right)^{1/2} = \sqrt{b-a}\left(\int_a^b |f_m-f_n|^2\right)^{1/2}$$

and the right-hand side term tends to 0 as $m, n \to 0$. Hence, $(f_n)_{n\geq 1}$ is a Cauchy sequence in $L^1([a,b])$ that converges to a function $f \in L^1([a,b])$. This means that

$$\int_a^b |f - f_n| \to 0 \quad \text{as} \quad n \to \infty\,.$$

Thanks to the Riesz theorem, there exists a subsequence (f_{p_n}) convergent to f a.e. For any ε, we obtain, by letting $n \to \infty$ and for $p_m > p_n$

$$\int_a^b |f_{p_n} - f|^2 \leq \varepsilon\,,$$

by Fatou's lemma. And thus $f \in L^2([a,b])$. Furthermore, we write

$$\int_a^b |f - f_n|^2 \leq \int_a^b |f - f_{p_n}|^2 + \int_a^b |f_{p_n} - f_n|^2 < 2\varepsilon\,,$$

for n sufficiently large, and the completeness is achieved.

1.3.1 Orthonormal bases

A *basis* of a vector space E is a linearly independent subset B of E spanning E, *i.e.*, such that any vector $x \in E$ can be written as $x = \sum_{k=1}^n \alpha_k x_n$, where $x_k \in B$ and the α_k are scalars. In inner product spaces, the reasons for which bases are so important are twofold. Infinite sums are considered instead of finite linear combinations and the notion of orthogonality replaces the linear independence property.

Let V be a Hilbert space V endowed with an inner product $(\cdot,\cdot)$. A sequence (or a set of vectors) $(v_n)_{n\geq 1}$ is called an *orthonormal sequence* if $(v_k, v_j) = \delta_{k,j}$, for $1 \leq k, j < \infty$, where δ_{jk} denotes the Kronecker delta symbol, *i.e.* δ_{jk} equals one if $j = k$ and zero otherwise. If the sequence is infinite, then it converges weakly to 0.

For example, the set of functions $f_n(x) = \exp(inx)/\sqrt{2\pi}$, for $n \in \mathbb{Z}$ is an orthonormal sequence for the space $L^2([-\pi,\pi])$ endowed with the L^2 inner product

$$(f, g) = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx\,.$$

Likewise, the set of Legendre polynomials $P_n(x)$ defined by the Rodrigues formula

$$P_0(x) = 1\,, \quad P_n(x) = \sqrt{\frac{2n+1}{2}}\frac{1}{2^n n!}\frac{d^n}{dx^n}\left((x^2-1)^n\right)\,, \quad n \geq 1\,,$$

forms an orthonormal system in the space $L^2([-1,1])$

The next result generalizes the Pythagorean formula (1.4). It can be established by induction.

Lemma 1.3.1 (Pythagorean formula). *Let $(f_k)_{1\leq k\leq n}$ be a set of orthogonal vectors in an inner product space V. Then,*

$$\left\|\sum_{k=1}^{n} f_k\right\|^2 = \sum_{k=1}^{n} \|f_k\|^2 \,.$$

Suppose $(f_n)_{n\geq 1}$ is an orthonormal sequence in an inner product space V. Then, an interesting question would be to find (complex) numbers α_n such that, for all $f \in V$,

$$f = \sum_{n=1}^{\infty} \alpha_n f_n \,.$$

Unfortunately, this cannot be achieved in general. Nevertheless, we have the following results.

Lemma 1.3.2. *Suppose that $(f_n)_{n\geq 1}$ is an orthonormal sequence in an inner product space V and that $f = \sum_{n=1}^{\infty} \alpha_n f_n$. Then, $\alpha_n = (f, f_n)$ for each n.*

We call the term $\sum_{n=1}^{\infty}(f, f_n)f_n$ the *Fourier series* of f with respect to the orthonormal sequence $(f_n)_{n\geq 1}$ and (f, f_n) are the *Fourier coefficients* of f with respect to $(f_n)_{n\geq 1}$. The following result gives details on the size of these coefficients.

Theorem 1.3.1 (Bessel's inequality). *Suppose that $(f_n)_{n\geq 1}$ is an orthonormal sequence in an inner product space V. Then, for every $f \in V$ we have*

$$\sum_{n=1}^{\infty} |(f, f_n)|^2 \;\leq\; \|f\|^2 \,. \tag{1.7}$$

In particular, this inequality shows that the series of nonnegative numbers $\sum_{n=1}^{\infty} |(f, f_n)|^2$ converges for every $f \in V$. This property means that the sequence $(f, f_n)_{n\geq 1}$ is an element of the Hilbert space ℓ^2 of square-summable sequences.

Let $(f_n)_{n\geq 1}$ be an orthonormal sequence in V. If for any $f \in V$, there exists coefficients α_n such that $f = \sum_{k=1}^{\infty} \alpha_n f_n$, then the sequence $(f_n)_{n\geq 1}$ is called a *complete orthonormal* sequence in V or an *orthonormal basis* for V. According to this definition, an orthonormal sequence $(f_n)_{n\geq 1}$ in a Hilbert space V is complete if for every $f \in V$ we have

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

Actually, this equality means that

$$\lim_{n\to\infty} \left\| f - \sum_{n=1}^{\infty}(f, f_n) f_n \right\| = 0 \,,$$

with respect to the norm $\|\cdot\|$ in V.

Theorem 1.3.2 (Parseval's theorem). *Suppose that $(f_n)_{n\geq 1}$ is an orthonormal sequence in an inner product space V. Then, $(f_n)_{n\geq 1}$ is a* complete orthonormal *sequence if and only if for every $f \in V$ we have*

$$\sum_{n=1}^{\infty} |(f, f_n)|^2 \;=\; \|f\|^2 \,. \tag{1.8}$$

The completeness of V is in general sufficient to ensure the convergence of the series $\sum_{n=1}^{\infty} |(f, f_n)|^2$ as stated next.

Theorem 1.3.3. *Let $(f_n)_{n\geq 1}$ be a complete orthonormal sequence in a Hilbert space V and let $(\alpha_n)_{n\geq 1}$ be a sequence of real or complex numbers. Then, the series $\sum_{n=1}^{\infty} \alpha_n f_n$ converges if and only if $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. Moreover, in that case we have*

$$\left\| \sum_{n=1}^{\infty} \alpha_n f_n \right\|^2 = \sum_{n=1}^{\infty} |\alpha_n|^2 \,.$$

And we have an important characterization of complete orthonormal sequences.

Lemma 1.3.3. *An orthonormal sequence $(f_n)_{n\geq 1}$ in a Hilbert space V is* complete *if and only if*

$$(f, f_n) = 0\,, \;\; \textit{for all } n \geq 1 \quad \Rightarrow \quad f = 0\,.$$

A Hilbert space is called *separable* if it contains a complete orthonormal sequence. Finite dimensional Hilbert spaces are separable.

If V is separable, the construction of an orthonormal basis is easy. Indeed, there exists a countable total set $(f_n)_{n\geq 1}$. The *Gram-Schmidt orthogonalization* procedure (cf. Section 4.2.5) allows to construct an orthonormal set $(u_n)_{n\geq 1}$ such that $\mathrm{span}(u_n) = \mathrm{span}(f_n)$ for any n.

Theorem 1.3.4. *Every separable inner product space has a countable orthonormal basis.*

Corollary 1.3.1. *If V is separable, then every orthonormal basis is countable.*

A bijective linear operator $A \in \mathcal{L}(V, W)$, where V and W are Hilbert spaces, is called *unitary* if A preserves the inner products (or the norms)

$$(Ag, Af) = (g, f)\,, \quad g, f \in V\,.$$

Hence, V and W are called *unitarily equivalent.* Let V be an infinite dimensional Hilbert space, and let $(f_n)_{n\geq 1}$ be any orthogonal basis. The map $A : V \to \ell^2(\mathbb{N})$, $f \mapsto ((f_k, f))_{k\geq 1}$ is unitary.

Lemma 1.3.4. *Any separable infinite dimensional Hilbert space is unitarily equivalent to $\ell^2(\mathbb{N})$.*

1.3.2 The projection theorem and Riesz lemma

In this section, we consider closed vector subpaces of a Hilbert space V, that are Hilbert spaces since a closed subspace of a complete normed space is complete.

Let S be a nonempty subset of a Hilbert space V. We recall that $u \in H$ is *orthogonal* to S if $(u, v) = 0$ for every $v \in S$. Then, the *orthogonal complement* of S is the set of all elements in V orthogonal to S, denoted by $S^\perp$. By continuity of the inner product, it follows that $S^\perp$ is a closed linear subspace and by linearity that $\overline{\mathrm{span}(S)}^\perp = S^\perp$. Obviously, $\{0\}^\perp = V$ and $V^\perp = \{0\}$. Moreover, is S is a closed subspace of V, we have $S^{\perp\perp} = S$.

Lemma 1.3.5. *Let S be a subspace of a Hilbert space V. Then S is* dense *if and only if $S^\perp = \{0\}$.*

A fundamental property of Hilbert spaces is that the distance of a point to a closed convex set if always attained. This result is especially important in the approximation theory.

Theorem 1.3.5 (Closest point). *Let C be a closed convex subset of a Hilbert space V. For every element $g \in V$, there exists a unique closest element (or a* best approximation*) $f \in C$ to g minimizing $\|g - f\|$,* i.e. *such that*

$$\|g - f\| = \inf_{h\in C} \|g - h\| .$$

Proof. We prove first the existence of such point. Let $(f_n)_{n\geq 1}$ be a sequence in C such that

$$\lim_{n\to\infty} \|g - f_n\| = \inf_{h\in C} \|g - h\| .$$

Posing $d = \inf_{h\in C} \|g - f_n\|$, we know that $\frac{1}{2}(f_m + f_n) \in C$ and thus we have

$$\|g - \frac{1}{2}(f_m + f_n)\| \geq d\,, \qquad \text{for all } m, n \geq 1\,.$$

From the parallelogram formula, we show that

$$\||f_m - f_n\|^2 = 2\|g - f_m\|^2 + 2\|g - f_n\|^2 - 4\|g - \frac{1}{2}(f_m - f_n)\|^2\,,$$

and since $2\|g - f_m\|^2 + 2\|g - f_n\|^2$ tends to $4d^2$ when $n, m \to \infty$, we conclude that $\|f_m - f_n\|^2$ tends to 0 when $m, n \to \infty$ and thus, $(f_n)_{n\geq 1}$ is a Cauchy sequence. The limit $f = \lim_{n\to\infty} f_n$ exists in C since V is complete and C is closed. We have then

$$\|g - f\| = \|g - \lim n \to \infty f_n\| = \lim_{n\to\infty} \|g - f_n\| = d\,.$$

The uniqueness of f can be easily obtained by contradiction. □

This theorem gives an existence and uniqueness result which is often used in optimization problems. However, it is of limited practical usefulness for finding this optimal point. The following theorem provides a useful characterization.

Corollary 1.3.2. *Let C be a closed convex subset of a real Hilbert space V. For $f \in C$ and $g \in V$, the following assertions are equivalent*

1. $\|g - f\| = \inf_{h \in C} \|g - h\|$
2. $(g - f, h - f) \leq 0$, *for all $h \in C$.*

Theorem 1.3.6 (Projection theorem). *Let S be a closed subspace of a Hilbert space V. Then every $g \in V$ can be uniquely decomposed as $g = f + h$ where $f \in S$ and $h \in S^\perp$, and thus we write symbolically*

$$H = S \oplus S^\perp \,.$$

The space V is the *direct sum* of S and its orthogonal complement $S^\perp$. In other words, to every $g \in V$, we can assign a unique element f which is the element in S closest to f. This property allows us to consider the operator $P_s g = f$ called the *orthogonal projection* corresponding to S. Note that we have

$$P_S^2 = P_S \,, \quad \text{and} \quad (P_S f, g) = (f, P_S g) \,, \text{ for every } f, g \in V \,.$$

Clearly, we have also $P_{S^\perp} g = g - P_S g = h$, with $h \in S^\perp$.

1.3.3 Bounded linear operators on Hilbert spaces

Linear operators on normed and Hilbert spaces play an important role in applied mathematics. We turn now to *linear functionals* $A : V \to \mathbb{C}$ and we will consider also bilinear functionals and quadratic forms.

We recall that a linear operator $A : X \to Y$, where X, Y are normed linear spaces, is called *bounded* if the operator norm is finite, *i.e.,* if there exists $M > 0$ such that

$$\|Af\|_Y \leq M \|f\|_X \,, \quad \text{for all } f \in X \,.$$

A consequence of the linearity of the operator is that *continuity* can be checked at a single point only.

Lemma 1.3.6. *Consider a linear operator $A : X \to Y$, where X, Y are normed linear spaces. The operator A is* continuous *at every point if it is continuous at a single point.*

And according to Lemma 1.1.3, a linear operator $A : X \to Y$ between two normed linear spaces is continuous if and only if it is bounded on X.

One of the most important operator for our purposes is obviously the *differential operator*, defined on the space of differentiable functions on an interval $[a, b] \subset \mathbb{R}$, by

$$(Df)(x) = \frac{df(x)}{dx} = f'(x)\,.$$

However this operator is unbounded.

The Cauchy-Schwarz inequality shows that the linear functional $A : f \mapsto (g, f)$ is bounded for the norm $\|g\|$. Interestingly, in a Hilbert space, every bounded linear functional can be written in this way.

Theorem 1.3.7 (Riesz lemma). *Let A be a bounded linear functional on a Hilbert space V. Then, there is a unique $g \in H$ such that $Af = (g, f)$ for all f in V. Moreover, we have $\|A\| = \|g\|$.*

Proof. If $A \equiv 0$ then we set $g = 0$. Otherwise, consider the kernel of A, $\mathrm{Ker}(A) = \{f \in V\,, Af = 0\}$. Since A is linear and bounded, $\mathrm{Ker}(A)$ is a closed proper subspace of V, and let $\tilde{g}$ be a unit vector of $\mathrm{Ker}(A)^{\perp}$. For every $f \in V$, we write $Af\tilde{g} - A\tilde{g}f \in \mathrm{Ker}(A)$ and thus we have

$$0 = (\tilde{g}, Af\tilde{g} - A\tilde{g}f) = Af - A\tilde{g}\,(\tilde{g}, f)\,.$$

Consequently, we set $g = (A\tilde{g})\tilde{g}$. The uniqueness is obtained by considering g_1, g_2 two such vectors. Then, for any $f \in V$,

$$(g_1 - g_2, f) = (g_1, f) - (g_2, f) = Af - Af = 0$$

which means that $g_1 - g_2 \in V^{\perp} = \{0\}$. Finally, the equality $\|A\| = \|g\|$ results from the definition of $\|A\|$ and the Cauchy-Schwarz inequality. □

This result indicates that a Hilbert space is equivalent to its dual space, $H = H'$. In other words, any continuous linear functional given on a Hilbert space can be uniquely identified with an element in the same space, *i.e.,* H and H' are isomorphic.

We have two results involving bilinear functionals defined on a Hilbert space.

Lemma 1.3.7. *Let A be a bounded operator on a Hilbert space V. Then, the bilinear functional $a(\cdot,\cdot)$ (*i.e.*, linear in each variable) defined by $a(u, v) = (Au, v)$ is bounded and $\|A\| = \|a\|$.*

Lemma 1.3.8. *Let $a(\cdot,\cdot)$ be a bounded bilinear functional on a Hilbert space V. There exists a unique bounded operator A on V such that*

$$a(f, g) = (Ag, g) \quad \textit{for all } f, g \in V\,.$$

A bilinear functional $a(\cdot,\cdot)$ on a normed space V is called *elliptic* or *coercive* if there exists a constant $\alpha > 0$ such that

$$a(f, f) \;\geq\; \|f\|^2\,, \quad \text{for all } f \in V\,.$$

The following theorem is a generalization of the Riesz lemma, it has important applications to boundary value problems (cf. Chapter 2).

Theorem 1.3.8 (Lax-Milgram theorem). *Let $a(\cdot,\cdot)$ be a bounded elliptic bilinear functional on a Hilbert space V. For every bounded linear functional A on V there exists a unique $u \in H$ such that*

$$a(f,g) = Ag\,, \qquad \textit{for all } g \in V\,.$$

Proof. Lemma 1.3.8 states that there exists a bounded operator A such that

$$a(f,g) = (Af,g)\,, \quad \text{for all } f,g, \in V\,.$$

Since the bilinear functional a is elliptic, we have then

$$\alpha\|f\|^2 \,\leq\, a(f,f) = (Af,f) \,\leq\, \|Af\|\|f\|\,,$$

and then we write

$$\|Af\| \,\geq\, \alpha\|f\|\,, \quad \text{for all } f \in V\,.$$

This indicates that A is bounded (or continuous). Moreover, A is a one-to-one map. Indeed, consider $f_1, f_2 \in V$. If $Af_1 = Af_2$ then $A(f_1 - f_2) = 0$ and thus

$$\|f_1 - f_2\| \,\leq\, \frac{1}{\alpha}\|A(f_1 - f_2)\| = 0\,.$$

It remains to show that the range of A is V. Let $(f_n)_{n\geq 1}$ be a sequence in V. If $\|Af_n - g\|$ tends to 0 for some $g \in V$ then

$$\|f_n - f_m\| \,\leq\, \frac{1}{\alpha}\|Af_n - Af_m\| \to 0\,, \qquad \text{as } m,n \to \infty\,.$$

Hence, $(f_n)_{n\geq 1}$ is a Cauchy sequence. The space V is complete, there exists $f \in V$ such that $\|f_n - f\| \to 0$ and thus $\|Af_n - Af\| \to 0$ since A is continuous. We deduce that $Af = g$ and $g \in \text{Im}(A)$. The range of A is closed. Suppose $\text{Im}(A) \neq V$, then there is an element $g \neq 0 \in H$ such that $g \perp \text{Im}(A)$, *i.e.,*

$$(Af,g) = 0\,, \qquad \text{for all } f \in V\,.$$

In particular, we have

$$0 = |(Af,f)| = |a(f,f)| \,\geq\, \alpha\|f\|^2\,,$$

which contradicts the assumption $g \neq 0$ and completes the proof. □

The Lax-Milgram theorem is especially used to establish the existence and uniqueness of weak solutions of boundary-value problems as will be seen in the next Chapter.

A bounded linear operator A on a Hilbert space V is said to be *invertible* if there exists a bounded operator A^{-1} such that $AA^{-1} = A^{-1}A = id_V$ and A^{-1} is then called the *inverse* of A. In finite dimensional spaces, if the operator is injective and an isometry it is invertible. Notice that A may have a linear space inverse on V without being invertible. The following result provides a criterion for invertibilty.

Lemma 1.3.9. *Let A be a bounded linear operator on a Hilbert space V, with $\|A\| < 1$. Then the operator $I - A$ is* invertible *and its inverse is given by*

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n \,.$$

Corollary 1.3.3. *Consider two linear operators A, T defined on a Hilbert space V and suppose that T is invertible and that $\|T - A\| < \|T^{-1}\|^{-1}$. Then A is invertible.*

By analogy with linear algebra (see Chapter 4), we will define the *spectrum* of a linear operator A in the set $\mathcal{L}(V)$ as the set of complex numbers λ such that $A - \lambda I$ is not invertible in V, where I is the identity operator. And we observe that the spectrum of an operator on a finite-dimensional vector space is precisely the set of eigenvalues. It can be shown that the spectrum of every element A in $\mathcal{L}(V)$ is a compact and nonempty set and is contained in the disk $\{\lambda \in \mathbb{C}\,,\ |\lambda| \leq \| A\|\}$.

1.3.4 Compact linear operators

An important class of operators is the class of *compact operators*. A linear operator A defined on a normed space V is called *compact* if for every sequence $(f_n)_{n\geq 1}$, the sequence $(Af_n)_{n\geq 1}$ has a convergent subsequence in V whenever f_n is bounded. The set of all compact operators is often denoted by $\mathcal{C}(V)$.

Lemma 1.3.10. *Every compact linear operator is bounded. A linear combination of compact operators is bounded.*

Lemma 1.3.11. *Let V be a Hilbert space and let $(A_n)_{n\geq 1}$ be a convergent sequence of compact operators. Then the limit A is a compact operator.*

One of the most important example of compact operator is the *integral operator*, defined by

$$(Kf)(x) = \int_a^b K(x,y) f(y) dy \,.$$

The function K is called the *kernel* of the operator. For instance, if

$$\int_a^b \int_a^b |K(x,y)|^2 dy dx < \infty$$

then K is a bounded operator on $L^2([a,b])$ and

$$\|K\| \,\leq\, \left(\int_a^b \int_a^b |K(x,y)|^2 dy dx \right)^{1/2} \,.$$

Theorem 1.3.9 (Arzelà-Ascoli). *Consider a uniformly equicontinuous sequence of functions $(f_n)_{n\geq 1}$ on a compact interval,* i.e., *such that for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$|f_n(x) - f_n(y)| \leq \varepsilon \quad \textit{if} \quad |x - y| < \delta.$$

If the sequence if $(f_n)_{n\geq 1}$ bounded, then it admits a uniformly convergent subsequence.

A linear operator A defined on an inner product space V is called *symmetric* if its domain is dense and if, for every f, g in V

$$(g, Af) = (Ag, f)\,.$$

Bounded symmetric operators are also called *Hermitian.* A complex number z is called *eigenvalue* of a symmetric linear operator A if there is a nonzero vector $u \in V$ such that $Au = zu$. The vector u is then called a corresponding *eigenvector.*

Lemma 1.3.12. *Let A be a symmetric linear operator. Then all eigenvalues are real and all corresponding eigenvectors are orthogonal.*

Theorem 1.3.10. *Any symmetric compact operator A has an eigenvalue λ_1 such that $|\lambda_1| = \|A\|$.*

Proof. We pose $\lambda = \|A\|$, with $\lambda \neq 0$ and we observe that

$$\|A\|^2 = \sup_{\|f\|=1} \|Af\|^2 = \sup_{\|f\|=1} (Af, f) = \sup_{\|f\|=1} (f, A^2 f)\,,$$

then there exists a sequence f_n such that

$$\lim_{n\to\infty} (f_n, A^2 f_n) = \lambda^2\,.$$

The operator A being compact, we assume that $\lim_{n\to\infty} A^2 f_n = \lambda^2 f$, *i.e.,* $A^2 f_n$ is convergent and thus we write

$$\begin{aligned}\|(A^2 - \lambda^2) f_n\|^2 &= \|A^2 f_n\|^2 - 2\lambda^2 (f_n, A^2 f_n) + \lambda^4 \\ &\leq 2\lambda^2(\lambda^2 - (f_n, A^2 f_n))\end{aligned}$$

and thus $\lim_{n\to\infty}(A^2 f_n - \lambda^2 f_n) = 0$. And we deduce that $\lim_{n\to\infty} f_n = f$. We notice that u is a normalized eigenvector of A^2 since $(A^2 - \lambda^2)u = 0$. This equation can be factorized as $(A + \lambda)(A - \lambda)u = 0$ which is equivalent to $(A - \lambda)u = v$ and $(A + \lambda)v = 0$. Hence, v is an eigenvector corresponding to $-\lambda$ or $v = 0$. This shows that $u \neq 0$ is an eigenvector of λ. □

If the operator A is bounded, the largest eigenvalue in modulus cannot be larger than $\|A\|$.

Compact operators are very similar to matrices, in the sense that the spectrum of a compact operator is similar to the spectrum of a finite matrix.

Theorem 1.3.11. *Let A be a compact symmetric operator on a Hilbert space V. Then the spectrum of A is either a finite set or a sequence of real eigenvalues converging to 0. In this case, the corresponding normalized eigenvectors u_j form an orthonormal set and every $f \in V$ can be decomposed as*

$$f = \sum_{k=1}^{\infty} (u_k, f) u_k + g \,,$$

where $g \in \mathrm{Ker}(A)$.

1.3.5 Adjoint and self-adjoint operators

The spectral theory of continuous self-adjoint linear operators on a Hilbert space generalizes the usual spectral decomposition of a matrix. We introduce a *conjugation* for operators on Hilbert spaces.

Let A be a bounded linear operator on a Hilbert space V. The operator $A^* : V \to V$ defined by

$$(Af, g) = (f, A^* g) \,, \quad \text{for all } f, g \in V \,,$$

is called the *adjoint operator* of A.

Proposition 1.3.1. *The adjoint operator satisfies the following properties*

1. $(A + B)^* = A^* + B^*$;
2. $(\lambda A)^* = \bar{\lambda} A^*$;
3. $(AB)^* = B^* A^*$;
4. $(A^*)^* = A$;

Lemma 1.3.13. *The adjoint operator A^* of a bounded operator A is bounded. Furthermore, the following identities hold*

$$\|A\| = \|A^*\| \,, \qquad \textit{and} \qquad \|A^* A\| = \|A A^*\| = \|A\|^2 \,.$$

A special case is when operators A and A^* coincide. If $A = A^*$, *i.e.*, if we have $(Af, g) = (f, Ag)$, for all $f, g \in V$, then the operator A is called *self-adjoint* or *Hermitian*. If $A = -A^*$, the operator A is called *anti-Hermitian*. For instance, let A be the operator on $L^2([a, b])$ defined by

$$(Af)(x) = x f(x) \,, \quad \text{for all } x \,.$$

We have then

$$(Af, g) = \int_a^b x f(x) \overline{g(x)} \, dx = \int_a^b f(t) \overline{x g(x)} \, dx = (f, Ag) \,,$$

and thus A is self-adjoint.

Let $a(\cdot, \cdot)$ be a bounded bilinear functional on a Hilbert space V and let A be an operator on V such that $a(f, g) = (Af, g)$ for all $f, g \in V$. Then A is self-adjoint if and only if $a(\cdot, \cdot)$ is symmetric.

Lemma 1.3.14. *Let A be a bounded operator on a Hilbert space V. Then the operators A^*A and $A + A^*$ are self-adjoint. Furthermore, the product of two self-adjoint operators is self-adjoint if and only if the operators commute.*

We consider the *projection operator* P_S (cf. Section 1.3.2) which maps every element of H onto its orthogonal projection in the subspace S.

Lemma 1.3.15. *For every subspace S of a Hilbert space V, the projection operator P_S is a Hermitian (self-adjoint) operator and is an* idempotent operator*, i.e., such that $P_S^2 = P_S$. Conversely, any idempotent operator is a projection operator onto a given subspace.*

The following property of Hermitian operators is related to the spectral properties of these operators.

Proposition 1.3.2. *If A is a Hermitian operator in a Hilbert space V, then*

$$\|A\| = \sup_{\|f\|=1} |(Af, f)| \,.$$

A bounded operator A on a Hilbert space V is called *isometric* if $\|Af\| = \|f\|$ for every $f \in H$. It is called *unitary* if $A^*A = AA^* = I$ on the Hilbert space V. Every unitary operator is isometric, *i.e.,* it preserves the distance between any two elements in V. Furthermore, if A is unitary so are A^{-1} and A^*. For example, consider the operator A on $L^2([0, 1])$ defined by

$$(Af)(x) = f(1 - x) \,.$$

It is easy to show that $A = A^* = A^{-1}$ and thus A is unitary.

Lemma 1.3.16. *A bounded operator A on a Hilbert space V is isometric if and only if $A^*A = I$ on V.*

An operator A defined on a Hilbert space V is called *positive* if it is *Hermitian* and if $(Af, f) \geq 0$, for all $f \in V$. Consider for instance the integral operator K on $L^2([a, b])$ defined by

$$(Kf)(x) = \int_a^b K(x, y) f(y) dy$$

where $K(\cdot, \cdot)$ is a positive continuous functions defined on $[a, b]$. Then, K is a positive operator since we have, for all $f \in L^2([a, b])$

$$(Kf, f) = \int_a^b \int_a^b K(x, y) f(y) \overline{f(y)} \, dy dx = \int_a^b \int_a^b K(x, y) |f(y)|^2 \, dy dx \geq 0 \,.$$

Lemma 1.3.17. *Let A be a bounded operator on a Hilbert space V. Then the operators A^*A and AA^* are* positive*. Furthermore, if A is invertible then A^{-1} is positive.*

A *square root* of a positive operator A on a Hilbert space V is a self-adjoint operator B such that $B^2 = A$. It can be shown that every positive operator has a unique positive square root.

We define the *resolvent operator* of an operator A on a Hilbert space V as the operator $A_\lambda = (A - \lambda I)^{-1}$. As for bounded operators, a complex number λ is called a *regular value* for A if the domain of the resolvent operator is the whole space V. The set of complex numbers λ which are not regular is called the *spectrum* of A. The set of all eigenvectors corresponding to an eigenvalue λ is called the *eigenspace* of λ and different spaces corresponding to distinct eigenvalues of a Hermitian or unitary operator are orthogonal to each other. All eigenvalues of a Hermitian operator are real. Furthermore, all eigenvalues of a positive (resp strictly positive) operator are nonnegative (resp. positive). All eigenvalues of a unitary operator on a Hilbert space V are complex numbers of unit module.

Lemma 1.3.18. *Let P be an invertible operator on a Hilbert space V. Then, for any operator A defined on V, the operators A and PAP^{-1} have the same eigenvalues.*

Proposition 1.3.3. *If A is a bounded Hermitian operator defined on a Hilbert space V then the eigenvalues of A are such that*

$$|\lambda| \leq \sup_{\|f\|=1} (Af, f) .$$

Furthermore, if A is a compact Hermitian operator on V, then there is an element g such that $\|g\| = 1$ and

$$|(Ag, g)| = \sup_{\|f\|=1} |(Af, f)| .$$

Theorem 1.3.12 (Spectral theorem). *Let V be an infinite dimensional Hilbert space and A be a nonzero, compact and Hermitian operator on V. Then there exists a sequence of nonzero real-valued eigenvalues $(\lambda_n)_{n\geq 1}$ and a corresponding sequence of eigenvectors $(u_n)_{n\geq 1}$ of A forming an orthonormal basis. Moreover, for each f in V, we have*

$$Af = \sum_{k=1}^{\infty} \lambda_k (f, u_k) u_k .$$

Furthermore, let P_n be the projection operator onto the space spanned by $(u_n)_{n\geq 1}$. Then, for all f in V we have

$$f = \sum_{n=1}^{\infty} P_n f , \qquad \text{and} \qquad A = \sum_{n=1}^{\infty} \lambda_n P_n .$$

1.3.6 The Fourier transform

In Section 1.2, we introduced Lebesgue L^p spaces and related main properties. The space L^2 of square integrable functions is undoubtedly the most important L^p-space. It has a key role in Fourier analysis[5]. Here, we start by considering the Fourier transform in $L^2(\mathbb{R})$ from the historical point of view of Fourier. Given an arbitrary function f defined on $[-\pi, \pi]$, if f can be decomposed as

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(kx) + b_k \sin(kx)$$

then the coefficients must be given by, for all $k \geq 1$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx)\, dx \quad \text{and} \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx)\, dx\,.$$

Rapidly, the questions of the existence of such decomposition, of the convergence of the series, and the type of convergence, of the limit arise. We show now that all these questions can be addressed in the context of an inner product space or a Hilbert space.

Given $f \in L^1(\mathbb{R})$, its *Fourier transform* is the function $\hat{f}$ defined by

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-ikx)\, dx \tag{1.9}$$

where dx denotes the ordinary Lebesgue measure. It may not be obvious that this definition is justifiable for all functions in $L^1(\mathbb{R})$, although the integral is well-defined for all real k. Nevertheless, we shall notice that the function $\exp(-ikx)$ is continuous and bounded and thus the product $f(x)\exp(ikx)$ is locally integrable for any $k \in \mathbb{R}$. Furthermore, we have

$$|f(x) \exp(-ikx)| \leq |f(x)|\,,$$

and thus the integral (1.9) is well defined for all $k \in \mathbb{R}$.

For example, consider the function $f(x) = \exp(-\alpha|x|)$ depicted in Figure 1.1 (left-hand side), for $\alpha = 1$. We have then

$$\begin{aligned}
\hat{f} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-\alpha|x|) \exp(-ikx)\, dx \\
&= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} \exp((\alpha - ik)x)\, dx + \int_{0}^{\infty} \exp(-(\alpha + ik)x)\, dx \right) \\
&= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\alpha - ik} + \frac{1}{\alpha + ik} \right) \\
&= \frac{1}{\sqrt{2\pi}} \frac{2\alpha}{\alpha^2 + k^2} = \frac{\sqrt{2}\alpha}{\sqrt{\pi}(\alpha^2 + k^2)}\,,
\end{aligned}$$

[5] Named after the French mathematician Joseph Fourier (1768-1830) known for investigating the decomposition of general functions as infinite sums of simpler functions, the Fourier series, and their application to problems of heat flow.

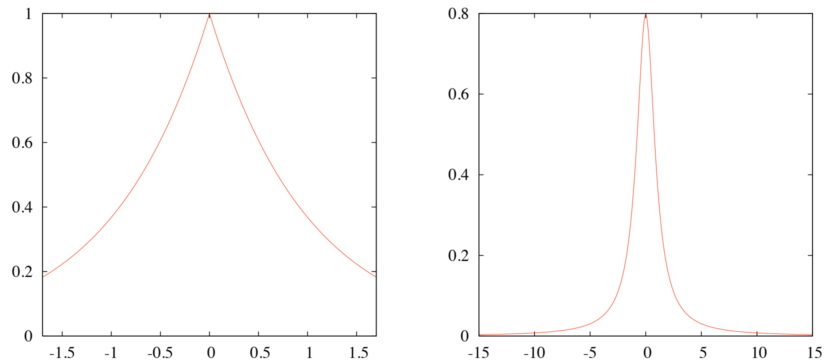

Fig. 1.1. *A function $f \in L^1(\mathbb{R})$ (left) and its Fourier transform (right).*

and the Fourier transform $\hat{f}$ is represented in Figure 1.1 (right-hand side). The following properties come almost directly from the definition of the Fourier transform.

Lemma 1.3.19. *Given $f, g \in L^1(\mathbb{R})$ and $\alpha \in \mathbb{C}$, we have*

1. $\|\hat{f}\|_\infty \leq \|f\|_1$, *where* $\|f\|_1 = \int_{-\infty}^{\infty} |f(x)|\, dx$;
2. $\hat{f}$ *is* uniformly continuous;
3. $\widehat{f+g} = \hat{f} + \hat{g}$ *and* $\widehat{\alpha f} = \alpha \hat{f}$.

The Fourier transform of integrable functions has additional properties.

Lemma 1.3.20. *Let $(f_n)_{n\geq 1}$ be a sequence in $L^1(\mathbb{R})$ such that $\|f_n\| \to 0$ when $n \to \infty$. Then, the sequence $(\hat{f}_n)_{n\geq 1}$ converges uniformly to $\hat{f}$ on $\mathbb{R}$.*

Theorem 1.3.13 (Riemann-Lebesgue). *If $f \in L^1(\mathbb{R})$ then $|\hat{f}(k)| \to 0$ as $|k| \to\infty$ and thus $\hat{f} \in C_c(\mathbb{R})$, where $C_c(\mathbb{R})$ denotes the space of continuous functions on $\mathbb{R}$ vanishing at infinity (*i.e.*, compactly supported functions).*

Proof. The proof is found in most classical analysis books. We remind that $\exp(-ikx) = \exp(-ikx - \pi)$ and thus by substitution, we write

$$\begin{aligned}\hat{f}(k) &= \frac{1}{2}\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty} f(x)\, dx - \int_{-\infty}^{\infty} f(x - \pi/k)\exp(-ikx)\, dx\right)\\ &= \frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(f(x) - f(x-\pi/k))\exp(-ikx)\, dx\,.\end{aligned}$$

And we deduce then

$$|\hat{f}(k)| \leq \frac{1}{2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}|f(x) - f(x-\pi/k))|\, dx\,.$$

Theorem 1.2.2 allows to conclude that $\lim_{|k|\to\infty} |\hat{f}(k)| = 0$. □

This theorem is notably used to prove the validity of asymptotic approximations for integrals. From these results, we observe that the Fourier transform is a continuous linear operator from $L^1(\mathbb{R})$ to the normed space $(C_c(\mathbb{R}), \|\cdot\|_\infty)$.

Convolution theorem

Before going further on the analysis of the Fourier transform for functions in $L^2(\mathbb{R})$, we need to introduce the convolution of two functions. If f, g are functions in $L^1(\mathbb{R})$, their *convolution* is the product $f * g$ defined by

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

Proposition 1.3.4. *Convolution satisfies the same algebraic properties as classical multiplication,* i.e., *for $f, g \in L^1(\mathbb{R})$*

1. *$f * (\alpha g + \beta h) = \alpha(f * g) + \beta(f * h)$, for any α, β;*
2. *$f * g = g * f$;*
3. *$f * (g * h) = (f * g) * h$.*

And we have the following result.

Theorem 1.3.14 (Convolution theorem). *Given $f, g \in L^1(\mathbb{R})$, we have $\widehat{f * g} = \hat{f}\hat{g}$.*

Corollary 1.3.4. *Suppose $f, g \in L^1(\mathbb{R})$ and α is a real number. Then*

1. *if $g(x) = f(x)\exp(i\alpha x)$ then $\hat{g}(k) = \hat{f}(k-\alpha)$;*
2. *if $g(x) = f(x-\alpha)$ then $\hat{g}(k) = \hat{f}(k)\exp(-i\alpha k)$;*
3. *if $h = f * g$ then $\hat{h}(k) = \hat{f}(k)\underline{\hat{g}}(k)$;*
4. *if $g(x) = \overline{f(-x)}$ then $\hat{g}(k) = \overline{\hat{f}}(k)$;*
5. *if $g(x) = \overline{f(x/\alpha)}$ for $\alpha > 0$ then $\hat{g}(k) = \alpha\hat{f}(\alpha k)$;*
6. *if $g(x) = -ixf(x)$ then $\hat{f}$ is differentiable and $\hat{f}'(k) = \hat{g}(k)$.*

This shows that the Fourier transform converts the convolution to a pointwise product as it converts the multiplication by a character to a translation. We recall that a function φ is a *character* of $\mathbb{R}$ if $|\varphi(x)| = 1$ and if, for all $x, y \in \mathbb{R}$, $\varphi(x+y) = \varphi(x)\varphi(y)$.

We can also observe that the Fourier transform of $(f(x+\alpha) - f(x))/\alpha$ is $\hat{f}(k)(\exp(i\alpha k) - 1)/\alpha$. And this suggest that the Fourier transform of f' if $ik\hat{f}(k)$, if $f, f' \in L^1$ and if f is the integral of f'. This interesting feature is extremely useful in the analysis of differential equations.

Inversion theorem

It is surely interesting to return from the Fourier transforms to the functions, *i.e.,* to have an *inversion formula*. Hence, if $f, \hat{f} \in L^1(\mathbb{R})$ we would expect a formula like

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) \exp(ixk)\, dk\,. \tag{1.10}$$

If $\hat{f} \in L^1(\mathbb{R})$, the righ-hand side is well defined but a proof of this identity (which is true indeed) is difficult and not straightforward. However, when both f and $\hat{f}$ are integrable, the following inverse equality holds for almost every x Notice that even if $\hat{f} \in C_0(\mathbb{R})$ the integral 1.10 will not converge for general $f \in L^1(\mathbb{R})$. However, this integral converges absolutely for a dense subspace as will be seen later. Actually, we have the following results.

Theorem 1.3.15 (Inversion theorem). *If $f, \hat{f} \in L^1(\mathbb{R})$ and if*

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) \exp(ixk)\, dk\,, \quad x \in \mathbb{R}\,,$$

then $g \in C_c(\mathbb{R})$ and $f(x) = g(x)$ a.e.

Corollary 1.3.5 (Uniqueness theorem). *If $f \in L^1(\mathbb{R})$ and $\hat{f}(k) = 0$ for all $k \in \mathbb{R}$ then $f(x) = 0$ a.e.*

Plancherel theorem

Since the Lebesgue measure of $\mathbb{R}$ is infinite, the set $L^2(\mathbb{R})$ of square integrable functions is not a subset of $L^1(\mathbb{R})$. Hence, we cannot extend the definition of the Fourier transform (1.9) directly to every function $f \in L^2(\mathbb{R})$. However, if $f \in L^1 \cap L^2$ then $\hat{f} \in L^2$ and the following *Plancherel identity*, holds

$$\|\hat{f}\|_2 = \|f\|_2\,.$$

The isometric mapping of $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ into $L^2(\mathbb{R})$ has a unique extension to a linear mapping from $L^2(\mathbb{R})$ to $L^2(\mathbb{R})$ and this extension will be called the *Fourier transform* of every $f \in L^2(\mathbb{R})$. We have the following theorem.

Theorem 1.3.16 (Plancherel theorem). *For every $f \in L^2(\mathbb{R})$ there exists a function $\hat{f} \in L^2(\mathbb{R})$ such that*

1. *if $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ then $\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-ikx)\, dx$ is the Fourier transform of f;*
2. *for every $f \in L^2(\mathbb{R})$, we have $\|\hat{f}\|_2 = \|f\|_2$ (Plancherel identity);*
3. *the mapping $f \to hatf$ is a Hilbet space isomorphism of $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$;*

4. the following symmetric relations exist between f and $\hat{f}$

$$\|\hat{f} - \varphi\|_2 \to 0 \text{ as } n \to \infty, \text{ where } \varphi(k) = \frac{1}{\sqrt{2\pi}} \int_{-n}^{n} f(x) \exp(ikx)\, dx \text{ and}$$

$$\|f - \psi\|_2 \to 0 \text{ as } n \to \infty, \text{ where } \psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-n}^{n} \hat{f}(k) \exp(ixk)\, dk.$$

This result indicates that the Fourier transform of any $L^2(\mathbb{R})$ function f is another square integrable function. The first two properties determine the mapping $f \to \hat{f}$ uniquely, since $L^1 \cap L^2$ is dense in L^2. The fourth property is called the *inversion theorem* in $L^2(\mathbb{R})$.

Theorem 1.3.17 (Parseval identity). *If $f, g \in L^2(\mathbb{R})$ then we have*

$$\int_{-\infty}^{\infty} f(x)\overline{g(x)}\, dx \;=\; \int_{-\infty}^{\infty} \hat{f}(k)\overline{\hat{g}(k)}\, dk\,.$$

The Fourier transform can be extended to functions in $L^1(\mathbb{R}^n)$ as follows

$$\hat{f}(k) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(x) \exp(-i(k,x))\, dx\,,$$

where x and k are n-dimensional vectors and (k,x) denotes here the usual dot product of the vectors, *i.e.,* $(k,x) = \sum_{j=1}^{n} k_j x_j$. All the basic properties listed above still hold.

1.4 Distributions

In this section, we shall provide a condensed introduction of the main concepts and properties of distribution theory which will be required in the next section and the following chapters. The reader is referred to the book of L. Schwartz [Sch50] for a detailled analysis of the topics introduced here. The current form of distribution theory, which is largely based on Fourier transformation, is due to L. Schwartz[6], is now an essential part in the analysis of partial differential equations.

From Section 1.2, we know that the Lebesgue L^p spaces contain functions non regular and non continuous functions for which derivatives are not defined in the classial sense. Nonetheless, the classical derivatives exist almost everywhere. Thus, it seems reasonable to generalize the notion of derivative to be independent of zero-measure sets. This is the purpose of the concept of *weak derivative* introduced by J. Leray (1906-1998) and S.L. Sobolev. An example of singular function is the *delta* function δ, introduced by P. Dirac (1902-1984), characterized by the following identity

[6] French mathematician (1915-2002) who introduced distributions as *objects* which generalize functions. The theory of *generalized function* was initiated in 1935 by the Russian mathematician S.L. Sobolev (1908-1989).

$$\int_{-\infty}^{\infty} f(x)\delta(x)\,dx = f(0) \tag{1.11}$$

for any continuous function f. No classical function satisfies this property. Indeed, from this definition, it results that $\delta(x) = 0$ for $x \neq 0$ and $\int_{-\infty}^{\infty} \delta(x)\,dx = 1$, for $f(x) = 1$. Hence, the intuition would guide us to see the delta function as being the limit of a sequence of ordinary functions f_n such as Gaussian functions

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} f(x) f_n(x)\,dx = f(0)\,, \quad \text{with} \quad f_n(x) = \frac{n}{\sqrt{2\pi}} \exp\left(-\frac{(nx)^2}{2}\right).$$

The purpose of the distribution theory is to provide a convenient setting for defining delta-like functions as continuous functionals in a space of regular test functions.

1.4.1 Test functions

We remind that the *support* of a function $f : \mathbb{R}^n \to \mathbb{R}$, denoted by $\mathrm{supp}(f)$, is the closure of all points x such that $f(x) \neq 0$. A function f on $\mathbb{R}^n$ has a *compact suppport* or is *compactly supported* if its support is a compact (*i.e.,* closed bounded) subset of $\mathbb{R}^n$. The space of all functions $f \in C^k(\mathbb{R}^n)$ having a compact support is denoted $C_c^k(\mathbb{R}^n)$ or sometimes $\mathcal{D}^k(\mathbb{R}^n)$. A *test function* is an infinitely differentiable function on $\mathbb{R}^n$ with compact support. The space $C_c^\infty(\mathbb{R}^n)$ of all test functions is denoted by $\mathcal{D}(\mathbb{R}^n)$ or simply by $\mathcal{D}$. It results from the definition that a test function f is a infinitely differentiable function vanishing outside a bounded set. Moreover, the space $\mathcal{D}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$.

In solving boundary value problems, distributions on an open subset Ω of $\mathbb{R}^n$ are needed. The concept of test function can be extended to open sets quite naturally. We consider the space $\mathcal{D}(\Omega)$ of test functions on Ω as the space of all functions $f \in C^\infty(\Omega)$ with support contained in a compact subset of Ω.

The existence of a test function is not obvious. A classical example is given by the function $\varphi(x) = f(|x|^2 - 1)$ where $|x|^2 = x_1^2 + \cdots + x_n^2$ and

$$f(t) = \exp(1/t)\,, \text{ if } t < 0 \quad \text{and} \quad f(t) = 0\,, \text{ if } t \geq 0\,.$$

By a suitable translation and scaling, we show that

$$\varphi \in C_C^\infty(\mathbb{R}^n)\,, \quad \int \varphi(x)\,dx = 1\,, \quad \varphi \geq 0\,, \quad \mathrm{supp}(\varphi) = \{x; |x| \leq 1\}\,.$$

Using this example, we can obtain new test functions by taking an integrable function u and defining the convolution

$$u_\varepsilon(x) = \int u(x - \varepsilon y)\varphi(y)\,dy = \varepsilon^{-n} \int u(y)\varphi((x-y)/\varepsilon)\,dy\,.$$

Other examples of test functions are $\varphi(\alpha x+\beta)$, $\alpha,\beta\in\mathbb{R}$, $\varphi^{(k)(x)}$ for $k>0$, and $f(x)\varphi(x)$ where f is an arbitrary infinitely differentiable function. The existence of test functions allows t prove the following result.

Proposition 1.4.1. *If f, g are continuous functions in $\mathbb{R}^n$ and if the following identity holds*

$$\int_{-\infty}^{\infty} f(x)\varphi(x)\,dx = \int_{-\infty}^{\infty} g(x)\varphi(x)\,dx\,, \qquad \text{for every } \varphi\in C_c^{\infty}(\mathbb{R}^n)\,,$$

then $f=g$.

This result explains the name of *test function*. Instead of testing the equality at any point $x\in\mathbb{R}^n$, the functions f and g are tested using infinitely differentiable functions with compact support.

A sequence of test functions $(\varphi_n)_{n\geq 1}\in\mathcal{D}(\mathbb{R}^n)$ is said to *converges to order* m to a function $\varphi\in\mathcal{D}(\mathbb{R}^n)$ if

1. the function φ_n and φ all have supports within a common bounded set $K\subset\mathbb{R}^n$;
2. $D^k\varphi_n\to D^k\varphi$ uniformly for all x for all orders $k=0,\ldots,m$, where D^k denotes the partial derivatives of order k (cf. Equation (1.2)).

If the convergence to order m is achieved for all $m\geq 0$ then $(\varphi_n)_{n\geq 1}$ is said to converge to φ. This definition still holds when replacing $\mathbb{R}^n$ by an open set $\Omega\subset\mathbb{R}^n$.

1.4.2 Distributions

The main idea is to identify arbitrary functions with linear functionals in the space of test functions. A *distribution of order* m is a linear functional T on $\mathcal{D}(\mathbb{R}^n)$ which is continuous to order m, *i.e.,* such that

1. $T(a\varphi+b\psi)=aT(\varphi)+bT(\psi)$, for every $a,b\in\mathbb{C}$ and $\varphi,\psi\in\mathcal{D}(\mathbb{R}^n)$,
2. $T(\varphi_n)\to T(\varphi)$ whenever $\varphi_n\to\varphi$ is any sequence of functions in $\mathcal{D}(\mathbb{R}^n)$ convergent to order m.

A linear function T that is continuous with respect to sequences $(\varphi_n)_{n\geq 1}$ that are convergent to all orders is simply called a *distribution* on $\mathbb{R}^n$. The space of all distributions is the dual space of $\mathcal{D}(\mathbb{R}^n)$ and is denoted by $\mathcal{D}'(\mathbb{R}^n)$ or simply $\mathcal{D}'$.

The space L^1 being a Hilbert space, we have seen that this space and its dual can be identified through the inner product $(\cdot,\cdot)$, thanks to the Riesz representation theorem 1.3.7

$$u: L^1(\mathbb{R}^n)\times L^1(\mathbb{R}^n)\to\mathbb{R}\,,\quad (u,\varphi)\mapsto\langle u,\varphi\rangle=(u,\varphi)_{L^1(\mathbb{R}^n)}=\int_{-\infty}^{\infty} u\varphi\,.$$

The bilinear form $(u, \varphi) \mapsto \langle u, \varphi \rangle$ is continuous and still makes sense if the functions φ are taken in a subspace V of L^1. Hence, every locally integrable functions f on $\mathbb{R}^n$ can be identified with a distribution T as

$$\langle T, \varphi \rangle = (f, \varphi) = \int_{-\infty}^{\infty} f\varphi\,.$$

Formally we introduce the following definition.

Definition 1.4.1. *A* distribution *T in $\mathbb{R}^n$ is a linear form on $\mathcal{D}(\mathbb{R}^n)$ such that to every compact set $K \subset \mathbb{R}^n$, there exist constants C and m such that*

$$|\langle T, \varphi \rangle| = |T(\varphi)| \,\leq\, C \sum_{|k| \leq m} \sup |D^k \varphi|\,, \qquad \varphi \in C_c^\infty(K)\,.$$

Conventionally, a function $f \in L^p(\mathbb{R}^n)$ will be identified to its associated distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ and the pairing between f and a test function φ is written as

$$(f, \varphi) = \langle T, \varphi \rangle\,.$$

A distribution $T \in \mathcal{D}'$ is called *regular* if there exists a locally integrable function f such that

$$\langle T, \varphi \rangle = \int_{-\infty}^{\infty} f\varphi\,, \quad \text{for every } \varphi \in \mathcal{D}\,.$$

The distribution is called *singular* if it is not regular.

We provide an important example of such distribution. Given $a \in \mathbb{R}$ we define the distribution δ_a on $\mathcal{D}(\mathbb{R})$ by

$$\langle \delta_a, \varphi \rangle = \delta_a(\varphi) = \varphi(a)\,, \quad \text{for every } \varphi \in \mathcal{D}(\mathbb{R})\,.$$

The map $\delta_a : \mathcal{D}(\mathbb{R}) \to \mathbb{R}$ is linear and is continuous. Moreover, we have

$$|\langle \delta_a, \varphi \rangle| \,\leq\, \|\varphi\|_{L^\infty(\mathbb{R})}\,, \quad \text{for every } \varphi \in \mathcal{D}(\mathbb{R})\,.$$

Hence, δ_a is a distribution, called the *Dirac mass* at point a. But, it cannot be represented by any locally integrable function (cf. the introduction of this chapter). The Dirac delta is thus a singular distribution.

An equivalent form of the definition 1.4.1 is given by the following theorem.

Theorem 1.4.1. *A linear form T on $\mathcal{D}(\mathbb{R}^n)$ is a* distribution *if and only if $T(\varphi_n) \to 0$ when $n \to \infty$ for every sequence $(\varphi_n)_{n \geq 1}$ in $\mathcal{D}(\mathbb{R}^n)$ such that*

1. *$D^\alpha \varphi_n \to 0$ uniformly when $n \to \infty$, for every multiindex α;*
2. *there is a fixed compact subset of $\mathbb{R}^n$ containing the support of all φ_n.*

A sequence satisfying these two conditions is said to converge to *0 in $\mathcal{D}(\mathbb{R}^n)$.*

1.4.3 Operations on distributions

We define now several usual operations on distributions that are already familiar for classical infinitely differentiable functions.

If T and S are distributions of order m on $\mathbb{R}^n$, it is easy to see that $T+S$ and αT are distributions of order m, for all $\alpha \in \mathbb{R}$. This shows that $\mathcal{D}'^m(\mathbb{R}^n)$ is a vector space. However, it has been proved by Schwartz that an associative multiplication of two distributions cannot be defined, *i.e.,* the product of arbitrary distributions is not a distribution. For example, suppose we define $\langle ST, u\rangle = \langle S, u\rangle\langle T, u\rangle$, this would clearly not be linear in u. However, if ψ is a function in $C^m(\mathbb{R}^n)$ and T a distribution of order m then ψT can be defined as a distribution of order m, simply by posing

$$\langle \psi T, \varphi\rangle = \langle T, \psi\varphi\rangle$$

since $\psi\varphi$ is in $C_c^m(\mathbb{R}^n)$ for every $\varphi \in C_c^m(\mathbb{R}^n)$. And we observe that this definition still holds if ψ is not compactly supported.

Suppose Ω is an open subset in $\mathbb{R}^n$ and consider a distribution T on $\mathbb{R}^n$. We define the *restriction* of T to the open set Ω, denoted by $T_{|\Omega}$, as

$$\langle T_{|\Omega}, \varphi\rangle = \langle T, \varphi\rangle\,, \qquad \text{for every } \varphi \in C_c^\infty(\Omega)\,.$$

Furthermore, if f s a locally integrable function we have:

$$\int_\Omega f_{|\Omega}(x)\varphi(x)\,dx = \int_{-\infty}^{\infty} f_{|\Omega}(x)\varphi(x)\,dx\,, \qquad \text{for every } \varphi \in C_c^\infty(\Omega)\,.$$

And we observe that this definition coincide with the notion of restriction for classical functions. However, the restriction of a distribution is only defined for an open set.

Differentiation

One of the strongest advantage of the distribution theory is that the differentiation operation is always defined and every distribution has all derivatives that are distributions. Let consider a differentiable function $f \in C^1(\mathbb{R}^n)$ and $\varphi \in \mathcal{D}^1(\mathbb{R}^n)$, then using integration by parts

$$\begin{aligned}\left\langle \frac{\partial f}{\partial x_i}, \varphi \right\rangle &= \int_{-\infty}^{\infty} \frac{\partial f}{\partial x_i}(x)\varphi(x)\,dx\,,\\ &= [\varphi f]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\partial \varphi}{\partial x_i}(x) f(x)\,dx = -\left\langle f, \frac{\partial \varphi}{\partial x_i}\right\rangle\,.\end{aligned}$$

This result is still valid if we set f to be a distribution. More precisely, let $T \in \mathcal{D}'(\mathbb{R}^n)$ be a distribution on $\mathbb{R}^n$. Then the *derivative* $\partial T/\partial x_k$ is defined by

$$\left\langle \frac{\partial T}{\partial x_k}, \varphi \right\rangle = -\left\langle T, \frac{\partial \varphi}{\partial x_k} \right\rangle .$$

More generally, given a multiindex α, then the linear functional denoted by $D^\alpha T$ is defined by

$$\langle D^\alpha T, \varphi\rangle = \langle T, (-1)^{|\alpha|} D^\alpha \varphi\rangle \;=\; (-1)^{|\alpha|}\langle T, D^\alpha \varphi\rangle .$$

Proposition 1.4.2. *If T is a distribution on $\mathbb{R}^n$ then $D^\alpha T$ is a distribution on $\mathbb{R}^n$ for any multiindex α.*

Proof. Linearity is obvious to show. To prove that $D^\alpha T$ is continuous, we use the definition of convergence of a sequence of functions $(\varphi_n)_{n\geq 1}$ to φ in $\mathcal{D}$ which requires that all derivatives up to and including orde α converge uniformly on a compact subset. Hence, $D^\alpha \varphi_n$ tends to $D^\alpha \varphi$ in $\mathcal{D}$ and then $D^\alpha T$ is a continuous fonctional and thus $D^\alpha T$ is a distribution. □

Set H to be the Heaviside function of a single variable x (*i.e.,* the characteristic function of $\mathbb{R}^+$), $H(x) = 1$ if $x \geq 0$, and $H(x) = 0$ if $x < 0$. It is clear that H is not differentable in the usual sense, but it is obviously a locally integrable function and it generates a regular distribution. Indeed, for any test function $\varphi \in \mathcal{D}(\mathbb{R})$, we have

$$\begin{aligned} \left\langle \frac{\partial H}{\partial x}, \varphi \right\rangle = \left\langle \frac{dH}{dx}, u \right\rangle &= -\int_{-\infty}^{+\infty} \varphi'(x) H(x)\, dx \\ &= -\int_0^{+\infty} \varphi'(x)\, dx \;=\; \varphi(0) = \langle \delta_{x=0}, \varphi\rangle . \end{aligned}$$

Hence, the derivative of the Heaviside function H in the sense of distributions is the Dirac mass and we will write as for functions $\delta_0 = H'$.

A sequence of regular distributions $(T_n)_{n\geq 1}$ (generated by a sequence of functions $(f_n)_{n\geq 1}$) is said to *converge weakly to* a distribution T if

$$\langle T_n, \varphi\rangle \to \langle T, \varphi\rangle , \qquad \text{for every } \varphi \in \mathcal{D} .$$

This notion is called the *weak distributional convergence*. For example, the sequence of functions on $\mathbb{R}$ defined by

$$f_n(x) = \frac{n}{\pi(1 + (nx)^2)} , \quad n \geq 1 ,$$

is distributionally convergent to the Dirac delta distribution δ (cf. Figure 1.2).

Lemma 1.4.1. *Let $T_n \to T$ in $\mathcal{D}'(\mathbb{R})$. Then, for every multiindex α*

$$D^\alpha T_n \to D^\alpha T .$$

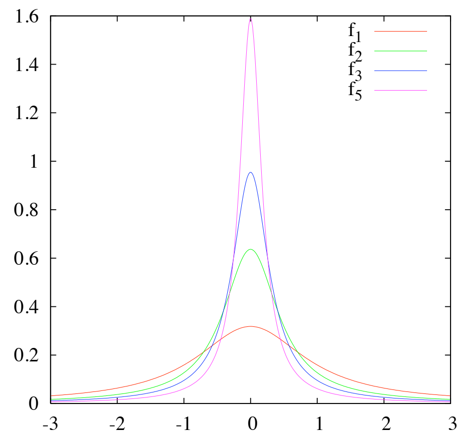

Fig. 1.2. *Convergence of a sequence of distributions to the Dirac delta function* δ.

Let T be a distribution in $\mathcal{D}'(\mathbb{R})$. The *antiderivative* (*primitive*) of T is a distribution S on $\mathbb{R}$ such that $S' = T$. For every distribution T defined on $\mathbb{R}$, there exists an antiderivative distribution S defined to within a constant term. Indeed, consider a test function $\varphi_0 \in \mathcal{D}(\mathbb{R})$ such that

$$\int_{-\infty}^{\infty} \varphi_0(x)\, dx = 1\,.$$

Then, every function $\varphi \in \mathcal{D}(\mathbb{R})$ can be uniquely decomposed as

$$\varphi = C\varphi_0 + \varphi_1\,, \quad \text{with } C = \int_{-\infty}^{\infty} \varphi(x)\, dx\,, \text{ and } \int_{-\infty}^{\infty} \varphi_1(x)\, dx = 0\,.$$

Given $F \in \mathcal{D}'(\mathbb{R})$ we define a functional T on $\mathcal{D}$ by

$$\langle S, \varphi \rangle = \langle S, C\varphi_0 + \varphi_1 \rangle = CK - \langle T, \psi \rangle$$

where K is a constant and ψ is the test function

$$\psi(x) = \int_{-\infty}^{x} \varphi_1(x)\, dx\,.$$

Then $S' = T$ and S is a distribution. We have then the following result.

Lemma 1.4.2. *Given $T \in \mathcal{D}'(\mathbb{R})$. The distributions for which $T' = 0$ are the constant functions.*

Lemma 1.4.3. *Consider $\Omega =]0,1[$. Then, for every funtion $f \in L^1(\mathbb{R})$, we have*

$$\frac{d}{dx}\left(\int_0^x f\right) = f\,, \quad \text{in } \mathcal{D}'(\Omega)\,.$$

1.4.4 Fourier transform

In Section 1.3.6 we have recalled the notion of Fourier transform applicable to any function f such that $|f|$ is integrable over $\mathbb{R}$. The *Fourier transform* of a distribution T can be defined as the distribution $\tilde{T}$ given by

$$\mathcal{F}(\langle T, \varphi\rangle) = \langle T, \hat{\varphi}\rangle\,,$$

for all test function φ (and we used the convenient notation $\mathcal{F}(\cdot)$ for the Fourier transform of an expression). If the inverse Fourier transform is defined on distributions by $\mathcal{F}^{-1}(\langle T, \varphi\rangle) = \langle T, \hat{\varphi}^{-1}\rangle$ then

$$\mathcal{F}^{-1}\mathcal{F}\langle T, \varphi\rangle = \langle T, \varphi\rangle\,.$$

However, the Fourier theorem requires the function φ to be of bounded variation for the result to hold. If φ is a function of bounded support, the function $\hat{\varphi}$ cannot be of bounded support in general. Hence the term $\langle T, \hat{\varphi}\rangle$ may not be well-defined. We need to introduce another function space to define the Fourier transform of distributions.

A function φ is called *rapidly decreasing* as $|x| \to\infty$ if the equality

$$\lim_{|x|\to\infty} |x|^m \varphi(x) = 0$$

holds for an arbitrary m. Such a function approches 0 as $|x| \to\infty$ faster than any inverse power $|x|^{-n}$. The space of test functions $\mathcal{S}(\mathbb{R})$ consists of functions rapidly decreasing as $|x| \to \infty$ together with derivatives of every order up to order p, *i.e.,*

$$\mathcal{S}(\mathbb{R}) = \{\varphi\,;\ \mathrm{supp}_{x\in\mathbb{R}}\,|x^m\varphi^{(p)}(x)| < \infty\,,\ \text{for all } m,p \in \mathbb{N}^*\}\,.$$

$\mathcal{S}(\mathbb{R})$ is called the space of rapidly decreasing functions. A sequence of functions $(\varphi_n)_{n\geq 1}$ of $\mathcal{S}(\mathbb{R})$ is said to *converge* to φ if the equality

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |x^m(\varphi_n^{(p)}(x) - \varphi^{(p)}(x))| = 0\,, \quad 0 \leq |p| \leq m\,,$$

is satisfied for an arbitrary m. The space of linear functions on $\mathcal{S}(\mathbb{R})$ is denoted $\mathcal{S}'(\mathbb{R})$ and these functions are called *tempered distributions*. Since every test function is a rapidly decreasing function then $\mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$. If T is a tempered distribution, then the Fourier transform is well-defined, *i.e.,* the Fourier transform of a rapidly decreasing function is another rapidly decreasing function.

For a locally integrable function f on $\mathbb{R}$ identified with a distribution T, we have

$$\begin{aligned}
\mathcal{F}(\langle T, \varphi\rangle) = \mathcal{F}(f, \varphi) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \exp(-ikx)\varphi(x)\,dx\right) f(k)dk \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(x) \left(\int_{-\infty}^{\infty} \exp(-ikx) f(k)\,dk\right) dx \\
&= \int_{-\infty}^{\infty} \varphi(x)\hat{f}(x)\,dx = \langle T, \hat{\varphi}\rangle\,.
\end{aligned}$$

With this definition, the Fourier transform of the Dirac distribution can be defined by

$$\mathcal{F}(\langle \delta_a, \varphi \rangle) = \langle \delta_a, \hat{\varphi} \rangle = \hat{\varphi}(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp(-iax)\varphi(x)\,dx\,.$$

1.5 Sobolev spaces

In this section, we introduce Sobolev spaces of integer order and we provide some of their basic properties. Our intent is to deal later with partial differential equations and linear operators and we will see that Sobolev spaces are designed to find the correct spaces for solutions. In this section, we are considering functions for which all the derivatives, in the distributional sense, belong to L^2 space and such that most of the classical derivation rules can be applied. Before turning to Sobolev spaces, we briefly recall the simpler concept of Hölder spaces.

1.5.1 Hölder spaces

We remind that there exists a condition for checking the regularity of real-valued functions, stronger than the regular continuity, *i.e.,* the *Lipschitz continuity*[7]. This condition is central to the Cauchy-Lipschitz theorem (or the Picard-Lindelöf theorem) which states and ensures the existence and uniqueness of the solution to an initial value differential problem. More specifically, a continuous real-valued function f defined on an open subset Ω of $\mathbb{R}^n$ is said to be *Lipschitz continuous* if there exists a nonnegative real constant k such that the following estimate holds

$$|f(x) - f(y)| \leq k\,\|x - y\|\,, \qquad \text{for all } x, y \in \Omega\,.$$

The function f is called *locally Lipschitz continuous* if for every $x \in \Omega$ there exists a neighborhood U of x such that $f|_U$ is Lipschitz continuous. A real-valued function f is said to be *Hölder continuous* if there exist nonnegative real constants k and m such that

$$|f(x) - f(y)| \leq k\,\|x - y\|^m\,, \qquad \text{for all } x, y \in \Omega\,.$$

The number m is called the *Hölder exponent* and we observe that if $m = 1$, then the function satisfies the Lipschitz condition while it is simply bounded if $m = 0$.

Hölder spaces, composed of functions satisfying a Hölder condition, are especially relevant for solving partial differential equations. In particular, the topological Hölder vector space $C^{k,m}(\Omega)$ consists of all functions f in $C^k(\Omega)$,

[7] Named after the German mathematician Rudolf O.S. Lipschitz (1832-1903).

i.e., having derivatives up to order k, for which the k^{th} partial derivatives are Hölder continuous with exponent $0 < m < 1$. Furthermore, if the Hölder coefficient

$$|f|_{C^{0,m}(\Omega)} = \sup_{x,y\in\Omega} \frac{|f(x) - f(y)|}{\|x - y\|^m}$$

is finite, then f is said to be *uniformly Hölder continuous* with exponent m in Ω and then the Hölder coefficient defines a *seminorm*, denoted by $[\cdot]_{C^{0,m}(\bar{\Omega})}$. Hölder spaces $C(\bar{\Omega})$ are usually endowed with the norm

$$\|f\|_{C^{k,m}(\bar{\Omega})} = \sum_{|\alpha|\leq k} \|D^\alpha f\|_{C^0(\bar{\Omega})} + \sum_{|\alpha|=k} [D^\alpha f]_{C^{0,m}(\bar{\Omega})} \,.$$

where the norm $\|\cdot\|_{C^0}$ is classically defined by

$$\|f\|_{C^0(\bar{\Omega})} = \sup_{x\in\Omega} |f(x)| \,,$$

and the m^{th}-Hölder norm is

$$\|f\|_{C^{0,m}(\bar{\Omega})} = \|f\|_{C^0(\bar{\Omega})} + [f]_{C^{0,m}(\bar{\Omega})} \,.$$

It can be shown that the Hölder space $C^{k,m}(\bar{\Omega})$ is a Banach space. It does not make difference whether we use Ω or its closure in the previous definitions.

1.5.2 Sobolev spaces of integer order

We remind that a function $f \in L^2(\Omega)$, for an open subset $\Omega \subset \mathbb{R}^n$, is identified to its associated distribution also denoted by $f \in \mathcal{D}'(\Omega)$ (for the sake of simplicity) and the pairing between f and a test function φ is written as $(f, \varphi) = \langle f, \varphi\rangle$. Likewise, any function f in $C^{k,m}(\bar{\Omega})$ can be identified to a distribution.

Notion of weak derivative

The notion of *weak derivative* allows to generalize the concept of derivative to functions that are only integrable, *i.e.*, function s which belong to $L^1(Omega)$ space.

Let consider a continuous function $f \in C^k(\Omega)$, where k is a nonnegative integer, α be a multiindex of order $|\alpha| = k$. Then, if φ is a test function, integrating by parts α times leads to

$$\int_\Omega f(x) D^\alpha \varphi(x)\, dx = (-1)^{|\alpha|} \int_\Omega D^\alpha f(x) \varphi(x)\, dx \,.$$

We notice that the boundary terms vanishe since the function φ is compactly supported. Furthermore, this identity is stil valid if the function f is only locally integrable. But then, the expression $D^\alpha f$ has to be considered in the distributional sense seen before. To this end, we provide the following definition.

Definition 1.5.1 (Weak derivative). *Suppose f, g are locally integrable functions for some open subset Ω and α is a multiindex. The function g is said to be the α^{th}*-weak derivative *of f if*

$$\int_\Omega f(x) D^\alpha g(x)\, dx = (-1)^{|\alpha|} \int_\Omega g(x)\varphi(x)\, dx \,,$$

for all test functions $\varphi \in C_c^\infty(\Omega)$.

If f has a weak derivative, it is written as $D^\alpha f$ and its weak derivative is then uniquely defined up to a zero-measure set. If f is sufficiently smooth to have a continuous partial derivative $D^\alpha f$ in the classical sense, then $D^\alpha f$ is also a partial derivative of f in the distributional sense. But the converse may not be true, *i.e.,* $D^\alpha f$ may exist in the distributional sense without existing in the classical sense.

Spaces H^m

The Sobolev space $H^m(\mathbb{R}^n)$, for integer m, is defined as the completion (or closure) of the linear space $C_c^\infty(\mathbb{R}^n)$ of infinitely differentiable functions with compact support, with respect to the norm

$$\|f\|_{H^m} = \left(\sum_{|\alpha| \le m} \|D^\alpha f(x)\|^2_{L^2(\Omega)} \right)^{1/2} = \left(\sum_{|\alpha| \le m} \int_{\mathbb{R}^n} |D^\alpha f(x)|^2 \right)^{1/2} .$$

We note that the space $H^0(\mathbb{R}^n)$ coincide with the space $L^2(\mathbb{R}^n)$.

Let Ω be a bounded domain in $\mathbb{R}^n$. The space $C^\infty(\bar{\Omega})$ of infinitely differentiable functions is not complete with respect to the norm $\|\cdot\|_{H^m}$. Then, the space $H^m(\Omega)$ is defined accordingly as the completion of the space $C^\infty(\bar{\Omega})$ of functions that are infinitely differentiable in the closure of Ω with respect to the norm

$$\|f\|_{H^m(\Omega)} = \left(\sum_{|\alpha| \le m} \int_\Omega |D^\alpha f(x)|^2 \right)^{1/2} .$$

The functions in spaces $H^m(\Omega)$ have derivatives in the distributional sense of order $|\alpha| \le m$ which belong to $H^{m-|\alpha|}(\Omega)$. It is then worth noticing that the space $H^m(\Omega)$ consists of functions $f \in L^2(\Omega)$ such that $D^\alpha f \in L^2(\Omega)$ for any multiindex α with $|\alpha| \le m$, *i.e.,*

$$H^m(\Omega) = \{ f \in L^2(\Omega) \,,\ \ D^\alpha f \in L^2(\Omega) \,,\ 0 \le |\alpha| \le m \} .$$

It has the noticeable property of being a Hilbert space[8] for the inner product defined by

[8] This explains the notation H for such spaces.

$$(f,g)_{H^m(\Omega)} = \sum_{|\alpha|\le m} \int_\Omega D^\alpha f\, D^\alpha g\,,$$

and the associated norm $\|\cdot\|_{H^m(\Omega)}$ defined above. The space $H^m(\Omega)$ is complete since the space $L^2(\Omega)$ is complete. And for $m > s$, there is a continuous imbedding $H^m(\Omega) \subset H^s(\Omega)$.

For example, we consider the Sobolev space of order 1 on Ω, denoted by $H^1(\Omega)$, for the inner product

$$(f,g)_{H^1(\Omega)} = \int_\Omega \left(f(x)g(x)\,dx + Df(x)Dg(x)\right)\,dx\,,$$

and the corresponding norm

$$\|f\|_{H^1(\Omega)} = \left(\int_\Omega |f(x)|^2\,dx + \int_\Omega |Df(x)|^2\,dx\right)^{1/2}.$$

For $m \ge 1$ we define the seminorm

$$|f|_{H^m(\Omega)} = \left(\sum_{|\alpha|=m} \|D^\alpha f\|^0_{L^2(\Omega)}\right)^{1/2},$$

and we notice that

$$\|f\|_{H^m(\Omega)} = \left(\|f\|^2_{H^{m-1}(\Omega)} + |f|^2_{H^m(\Omega)}\right)^{1/2}.$$

The space $H^m(\mathbb{R}^n)$ is also interesting as it can be described using Fourier transform of tempered distributions (cf. Section 1.4.4). A function $f \in L^2(\mathbb{R}^n)$ belongs to $H^m(\mathbb{R}^n)$ if and only if $(1+|k|^2)^{m/2}\hat{f}(k) \in L^2(\mathbb{R}^n)$. The norm $\|f\|_{H^m}$ is then given by

$$\|f\|^2_{H^m} = (2\pi)^{-n} \int (1+|k|^2)^{m/2} |\hat{f}(k)|^2\,dk\,.$$

It is then possible to show that the space $\mathcal{S}(\mathbb{R}^n)$ is dense in $H^m(\mathbb{R}^n)$ and thus that $C_c^\infty(\mathbb{R}^n)$ is dense in $H^m(\mathbb{R}^n)$, for any integer value $m \ge 0$. This justifies that $H^m(\mathbb{R}^n)$ can be defined as the completion of $C_c^\infty(\mathbb{R}^n)$ in the norm $\|\cdot\|_{H^m}$.

We shall also mention that the spaces H^m provide a convenient way of attesting the regularity of a function. Here, the regularity (or the degree of smoothness) of a function $f \in H^m$ correspond to the number of times the function f is weakly differentiable before its derivatives are no longer in L^2.

Sobolev spaces $W^{m,p}$

For $1 \le p \le \infty$ and m a nonnegative integer, the Sobolev space $W^{m,p}(\Omega)$ is composed of all locally integrable real-valued (or complex-valued) functions

f such that for each multiindex $|\alpha| \leq m$, $D^\alpha f$ exists in the distributional (or weak) sense in $L^p(\Omega)$, *i.e.,*

$$W^{m,p} = \{f \in L^p(\Omega)\,,\ \ D^\alpha f \in L^p(\Omega)\,,\ |\alpha| \leq m\}\,.$$

We denote by $W_0^{m,p}(\Omega)$ the closure of $C_c^\infty(\Omega)$ in the space $W^{m,p}(\Omega)$. We have then the chain of imbeddings

$$W_0^{m,p}(\Omega) \subset W^{m,p}(\Omega) \subset L^p(\Omega)\,.$$

We observe that for $p = 2$, $H^m(\Omega) = W^{m,2}(\Omega)$, for every $k \geq 0$ and we will then write classically $H_0^m(\Omega) = W_0^{m,2}(\Omega)$. Furthermore, we have clearly $W^{0,p}(\Omega) = L^p(\Omega)$.

The space $W^{1,p}(\Omega)$ is endowed with the standard norm

$$\begin{aligned}
\|f\|_{W^{1,p}(\Omega)} &= \left(\|f\|^p_{L^p(\Omega)} + \|Df\|^p_{L^p(\Omega)}\right)^{1/p} \\
&= \left(\int_\Omega |f(x)|^p + |Df(x)|^p\, dx\right)^{1/p}, \quad \text{if } p < \infty \\
\|f\|_{W^{1,\infty}(\Omega)} &= \max(\|f\|_{L^\infty(\Omega)}, \|Df\|_{L^\infty(\Omega)})
\end{aligned}$$

Noticing that Lipschitz continuous functions are differentiable a.e., leads to conclude that $W^{1,\infty}(\Omega) = C^{0,1}(\Omega)$.

Likewise, for $m > 1$, the standard norm on $W^{k,p}$ spaces is defined by

$$\begin{aligned}
\|f\|_{W^{m,p}(\Omega)} &= \left(\sum_{|\alpha|\leq m} \|D^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p} = \left(\sum_{|\alpha|\leq m} \int_\Omega |D^\alpha f|^p\, dx\right)^{1/p} \\
\|f\|_{W^{m,\infty}(\Omega)} &= \max_{|\alpha|\leq m} \|D^\alpha f\|_{L^\infty(\Omega)} = \sum_{|\alpha\leq m} \operatorname*{ess\,sup}_\Omega |D^\alpha f|\,.
\end{aligned}$$

Lemma 1.5.1. *Equipped with the standard norms defined above, the space* $W^{m,p}(\Omega)$ *is a* Banach space. *Furthermore, for finite p,* $W^{m,p}(\Omega)$ *is a* separable *space. In particular,* $W^{m,2}(\Omega)$ *is a separable Hilbert space with the inner product*

$$(f,g)_{H^m} = \sum_{|\alpha|\leq m} \int_\Omega D^\alpha f(x)\, D^\alpha g(x)\, dx\,.$$

The next lemma give some properties of weak derivatives and $W^{m,p}$ spaces.

Lemma 1.5.2. *Consider* $f, g \in W^{m,p}(\Omega)$ *and a multiindex* $|\alpha| \leq m$. *Then*

1. $D^\alpha f \in W^{m-|\alpha|,p}(\Omega)$;
2. $\lambda f + \mu g \in W^{m,p}(\Omega)$ *and* $D^\alpha(\lambda f + \mu g) = \lambda D^\alpha f + \mu D^\alpha g$, *for each* $\lambda, \mu \in \mathbb{R}$;
3. $D^\beta(D^\alpha f) = D^{\alpha+\beta} f$, *for all multiindices* $|\alpha| + |\beta| \leq m$;
4. *if* $\omega \subset \Omega$ *is an open subset, then* $f \in W^{m,p}(\omega)$.

Imbeddings

We have already mentioned that there exist the following imbedding $H^m(\Omega) \subset H^s(\Omega)$ for integers $m > s$. Likewise, $W^{m,p}(\Omega) \subset W^{s,p}(\Omega)$ for $m \geq s$ and any $1 \leq p < \infty$. If Ω is a bounded set then we know that $L^p(\Omega) \subset L^q(\Omega)$ for any $p \geq q$. It follows then that $W^{m,p}(\Omega) \subset W^{s,q}(\Omega)$, whenever $p \geq q$ and $m \geq s$. We have the next results.

Lemma 1.5.3. *Suppose Ω is a bounded set with a piecewise regular boundary. Then there exist the compact imbeddings*

$$W^{m,p}(\Omega) \subset C_b(\Omega)\,, \quad \text{for } m > n/p$$
$$W^{m,p}(\Omega) \subset C_b^k(\Omega)\,, \quad \text{for } m > n/p + k\,,$$

where $C_b(\Omega)$ denotes the space of bounded continuous functions on Ω with respect to the supremum norm and $C_b^k(\Omega)$ is the space of functions of $C^k(\Omega)$ having bounded derivatives up to order k.

Corollary 1.5.1. *There exists a continuous imbedding*

$$W^{m,p}(\Omega) \subset L^q(\Omega)\,, \quad \text{for } m \geq 0\,, \quad m \geq n\left(\frac{1}{p} - \frac{1}{q}\right)\,.$$

1.5.3 Trace theorems

Next we discuss the notion of trace. Suppose Γ denotes the piecewise smooth boundary of Ω. We consider the restriction operator $T : C^1(\bar{\Omega}) \to L^p(\Gamma)$, $Tf = f_{|\Gamma}$. The problem is that a function $f \in W^{1,p}(\Omega)$ is only defined a.e. on Ω and since Γ has n-dimensional Lebesgue measure zero, the restriction T has no meaning as such. The notion of *trace operator* has been invented to overcome this problem. The domain of T is a subset of $W^{1,p}(\Omega)$ and since $C^1(\bar{\Omega})$ is dense in $W^{1,p}(\Omega)$, T admits a continuous extension [Ada75].

Lemma 1.5.4 (Trace of $W^{1,p}$ functions). *Suppose Ω is bounded and Γ is C^1 and let $1 \leq p < \infty$. Then there exists a continuous linear operator $T : W^{1,p}(\Omega) \to L^p(\Gamma)$ such that*

1. *$(Tf)(x) = f(x)$ for all $x \in \Gamma$ if $f \in W^{1,p} \cap C^0(\bar{\Omega})$.*
2. *there exists a constant $C > 0$ such that*

$$\|Tf\|_{L^p(\Gamma)} \leq C\,\|f\|_{W^{1,p}(\Omega)}\,,$$

for all $f \in W^{1,p}(\Omega)$.

The operator T is called the *trace operator* and Tf is called the *trace* of f on Γ.

Corollary 1.5.2. *Suppose Ω is bounded and Γ is C^1. There exists a continuous linear operator $T_0 : H^1(\Omega) \to L^2(\Gamma)$ such that*

1. $(T_0 f)(x) = f(x)$, *for all* $x \in \Gamma$ *if* $f \in H^1 \cap C^0(\bar{\Omega})$.
2. *there exists a constant* $C > 0$ *such that*

$$\|T_0 f\|_{L^2(\Gamma)} \leq C \|f\|_{H^1(\Omega)} , \tag{1.12}$$

for all $f \in H^1(\Omega)$.

Notice that the trace operator is not defined for functions in $L^2(\Omega)$.

1.5.4 Fractional order spaces and dual spaces

It is possible to define Sobolev spaces $W^{m,p}(\Omega)$ for arbitrary real number m. In particular, fractional order Sobolev spaces $H^m(\mathbb{R}^n)$ can be defined using the Fourier transform. Such spaces are of interest when considering for instance the image of the trace operator defined above, $\mathrm{Im}(T_0)$. The image is a subset of $L^2(\Gamma)$, dense in $L^2(\Gamma)$ that will be denoted $H^{1/2}(\Gamma)$. The space $H^{1/2}(\Gamma)$ is a Hilbert space for the norm

$$\|f\|_{H^{1/2}(\Gamma)} = \inf_{g \in H^1(\Omega), T_0 g = f} \|g\|_{H^1(\Omega)} .$$

With this definition, the estimate (1.12) can be improved as follows

$$\|T_0 f\|_{H^{1/2}(\Gamma)} \leq \|f\|_{H^1(\Omega)} ,$$

for all $f \in H^1(\Omega)$. Next, we turn to some particular spaces for which we provide some properties.

The space $H_0^1(\Omega)$ can also be defined as follows

$$H_0^1(\Omega) = \{f \in H^1(\Omega) , \ T_0 f = 0\} ,$$

and we observe that $H_0^1(\mathbb{R}^n) \neq H^1(\mathbb{R}^n)$.

Lemma 1.5.5. *The space* $C_c^\infty(\Omega)$ *is dense in* $H_0^1(\Omega)$.

By analogy, we consider the subspace of $H^m(\Omega)$ defined by

$$H_0^m(\Omega) = \{f \in H^m(\Omega) , \ D^\alpha f = 0 \text{ on } \Gamma\}$$

for all multiindex $|\alpha| < m$.

An important result, the Poincaré-Friedrichs inequality states that the seminorm

$$|f|_{H^m(\Omega)} = \left(\int_\Omega \sum_{|\alpha| = m} |D^\alpha f|^2 \, dx \right)^{1/2} \tag{1.13}$$

is a norm on $H_0^m(\Omega)$ equivalent to the norm defined previously

$$\|f\|_{H^m(\Omega)} = \left(\int_\Omega \sum_{|\alpha| \leq m} |D^\alpha f|^2 \, dx \right)^{1/2} .$$

And we have the result

Lemma 1.5.6 (Poincaré-Friedrichs inequalities).

1. *Suppose* Ω *is bounded in one direction. Then there exists a constant* $C > 0$ *such that*
$$\|f\|_{L^2(\Omega)} \leq C\,|f|_{H^1(\Omega)}\,, \quad \textit{for all } f \in H^1_0(\Omega)\,.$$
2. *If* Ω *is bounded of diameter* C *(or contained in a cube of side length* C*), then*
$$|f|_{H^m(\Omega)} \leq \|f\|_{H^m(\Omega)} \leq (1+C)^m\,|f|_{H^m(\Omega)}\,,$$
for all $f \in H^m_0(\Omega)$.

We denote by $H^{-1}(\Omega)$ the dual space of $H^1_0(\Omega)$; this space has an importance in the analysis of partial differential equations as will be seen in next chapter. The definition of $H^{-1}(\Omega)$ implies that any function f in $H^{-1}(\Omega)$ is a bounded linear functional on $H^1_0(\Omega)$. The space $H^{-1}(\Omega)$ is a Hilbert space for the dual norm
$$\|f\|_{H^{-1}(\Omega)} = \sup_{u \in H^1_0(\Omega)} \left(\frac{\langle f, u\rangle}{\|u\|_{H^1(\Omega)}} \right)\,,$$
where $\langle \cdot,\cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H^1_0(\Omega)$.

Since $\mathcal{D}(\Omega)$ is dense in $H^1_0(\Omega)$, the functions in $H^{-1}(\Omega)$ can be identified to distributions and this justifies the notation $\langle \cdot,\cdot \rangle$. For this reason, $H^{-1}(\Omega)$ is often called a *distribution space*. For example, given $f \in L^2(\Omega)$ then $Df \in H^{-1}(\Omega)^n$.

Lemma 1.5.7 (Characterization of H^{-1}). *Suppose* $f \in H^{-1}(\Omega)$*. The* f *can be nonuniquely decomposed as follows*
$$\langle f, u\rangle = \int_\Omega f_0 u + \sum_{i=1}^n f_i \frac{\partial u}{\partial x_i}\,dx\,, \quad \textit{with } u \in H^1_0(\Omega)\,, \tag{1.14}$$
where every $f_i \in L^2(\Omega)$*, for* $0 \leq i \leq n$*. Furthermore, if* f *satisfies (1.14) for every* $f_i \in L^2(\Omega)$*, then we have*
$$\|f\|_{H^{-1}(\Omega)} = \inf \left(\int_\Omega \sum_{i=0}^n |f_i|^2\,dx \right)^{1/2}\,,$$
We will use the notation
$$f = f_0 + \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}\,,$$
whenever (1.14) holds.

Likewise, we denote by $H^{-2}(\Omega)$ the dual space of $H^2_0(\Omega)$ and we recall that
$$H^2_0(\Omega) = \{f \in H^2(\Omega)\,,\ T_0 f = 0 \text{ and } T_1 f = 0\}\,,$$

where the trace operator T_1 is defined by $T_1 f = Df \cdot n$, and n is the outer unit normal to Ω. We can show that for any $f \in H^2(\Omega)$, $T_1 f$ is in $L^2(\Gamma)$. Since $\mathcal{D}(\Omega)$ is dense in $H_0^2(\Omega)$, then $H^{-2}(\Omega)$ is also a distribution space.

We denote by $(H^1(\Omega))'$ the dual space of $H^1(\Omega)$. It is a Hilbert space for the norm

$$\|f\|_{(H^1(\Omega))'} = \sup_{u \in H^1(\Omega), u \neq 0} \frac{f(u)}{\|u\|_{H^1(\Omega)}} .$$

1.5.5 Generalization of integration by parts formulas

In this paragraph, we remind the classical identities commonly used for writing the weak formulation of partial differential equations. From now on, Ω is supposed to be a bounded domain with Lipschitz continuous boundary denoted $\partial\Gamma$. The outer unit normal to Ω is defined a.e on $\partial\Omega$ and will be denoted by $n(x) = (n_1, \ldots, n_n)^t(x)$, *i.e.* n_i is the i^{th} component of n.

Using the density of $\mathcal{D}(\bar{\Omega})$ in $H^1(\Omega)$, we write the Green formula for functions in $H^1(\Omega)$.

Lemma 1.5.8 (Green's formula in $H^1(\Omega)$). *For every $f, g \in H^1(\Omega)$, we have*

$$\int_\Omega f \frac{\partial g}{\partial x_i}\, dx = -\int_\Omega \frac{\partial f}{\partial x_i} g\, dx + \int_{\partial\Omega} f g n_i\, d\sigma .$$

Using the trace operator introduced above, we provide the following result.

Lemma 1.5.9 (Green's formula in $H^2(\Omega)$). *For every $f \in H^2(\Omega)$ and $g \in H^1(\Omega)$, the following identity holds*

$$\int_\Omega \Delta f\, g\, dx = -\int_\Omega Df\, Dg\, dx + \int_{\partial\Omega} \frac{\partial f}{\partial n}\, g\, d\sigma ,$$

where Δ denotes the Laplacian operator *and $\partial f / \partial n = Df \cdot n$.*

And we conclude this section by giving the following result.

Proposition 1.5.1. *Suppose $f \in H^1(\Omega)$ such that $\Delta f \in L^2(\Omega)$. Then, $T_1 f$ defines an element of $H^{-1/2}(\Omega)$.*

1.6 Exercises and Problems

References

[Ada75] Adams R.A., *Sobolev Spaces*, Academic press Inc., New York, (1975).

[Aub79] Aubin J.-P., *Applied Functional Analysis*, John Wiley, New York, (1979).

[Ber95] Bartle R.G., *The elements of integration and Lebesgue measure*, Wiley Classics Library, John Wiley & Sons, New York, (1995).

[Bou04] Bourbaki N., *Integration I, Chapters 1-6.*, in Elements of Mathematics, Springer-Verlag, Berlin, (2004).

[Bra00] Bracewell R.N., *The Fourier Transform and its Applications*, 3rd ed., Mac Graw Hill, Boston, (2000).

[Bre05] Brezis H., *Analyse fonctionnelle: théorie et applications*, Dunod, (2005).

[But06] Butz T., *Fourier Transformation for Pedestrians*, Birkhäuser, (2006).

[Con94] Conway J.B., *A Course in Functional Analysis*, Graduate Texts in Mathematics, **96**, 2nd ed., Springer-Verlag, New York, (1994).

[DL85] Dautray R., Lions J.L., *Mathematical Analysis and Numerical Methods for Science and Technology. Vol 2: Functional and Variational Methods*, Springer-Verlag, (1985).

[DM90] Debnath L., Mikusiński, *Introduction to Hilbert Spaces with Applications*, Academic press Inc., New York, (1990).

[EMT04] Eidelman Y., Milman V., Tsolomitis A., *Functional Analysis. An Introduction*, Graduate Studies in Mathematics, **66**, AMS, Providence, (2004).

[Fol92] Folland G.B., *Fourier Analysis and its Applications*, The Wadsworth & Brooks / Cole mathematics series, Pacific Grove, (1992).

[Hal50] Halmos P.R., *Measure Theory*, Van Nostrand Cie, New York, (1950).

[Hor76] Hormander L., *Linear Partial Differential Operators*, Grundlehren der mathematischen Wissenschaften, **116**, 4th ed., Springer-Verlag, Berlin, (1976).

[KF99] Kolmogorov A.N., Fomin, S.V., *Elements of the Theory of Functions and Functional Analysis*, Dover Publications, (1999).

[Kre89] Kreyszig E., *Introductory Functional Analysis with Applications*, John Wiley, New York, (1989).

[Lan93] Lang S., *Real and Functional Analysis*, Graduate Texts in Mathematics, **142**, 3rd ed., Springer, New York, (1993).

[Lax02] Lax P., *Functional Analysis*, Wiley-Interscience, (2002).

[LV02] Lebedev L.P., Vorovich I.I., *Functional Analysis in Mechanics*, Springer Monographs in Mathematics, Springer, New York (2002).

[Mor00] Morrison T.J., *Functional Analysis. An introduction to Banach space theory*, Pure and Applied Mathematics, John Wiley & Sons Inc., New York, (2000).

[Rud70] Rudin W., *Real and complex analysis*, Mc Graw Hill, New York, (1970).

[Rud76] Rudin W., *Principles of mathematical analysis*, Mc Graw Hill, New York, (1976).

[Rud91] Rudin W., *Functional analysis*, 2nd edition, Mc Graw Hill, New York, (1991).

[Sax01] Saxe K., *Beginning Functional Analysis*, Springer, New York, (2001).

[Sch50] Schwartz L., *Théorie des distributions I-II*, Hermann, Paris, (1950-51).

[Sog93] Sogge C.D., *Fourier inegrals in classical analysis*, Cambridge Tracts in Mathematics, **105**, Cambridge University press, (1993).

[VGK+72] Vilenkin N.Y., Gorin E.A., Kostyuchenko A.G., Krasnosel'skii M.A., Krein S.G. (eds), *Functional Analysis*, Wolters-Noordhoff B.V., (1972).
[Yos80] Yosida K., *Functional Analysis*, 6th ed., Springer-Verlag, New York, (1980).