Special Mathematics Lecture

Introduction to functional analysis

Nagoya University, Spring 2023

Lecturer: Serge Richard

Goals of these lectures notes:

Provide the necessary background information for understanding the main ideas of distribution theory, Lebesgue integral, and the basics of operator theory.

These notes correspond to 14 lectures lasting 90 minutes each.

Website for this course:
http://www.math.nagoya-u.ac.jp/~richard/SMLspring2023.html

Comments or corrections are welcome:
richard@math.nagoya-u.ac.jp

Contents

1 Distribution theory **2**
1.1 Test functions and distributions . . . 2
1.2 Differentiation of distributions . . . 6
1.3 Other operations with distributions . . . 7
1.4 Continuous extension of distributions . . . 9
1.5 Fourier transform, Schwartz functions, and tempered distributions . . . 10

2 Lebesgue theory of integration **14**
2.1 Reminder on Riemann integration . . . 14
2.2 Lebesgue measure . . . 16
2.3 Lebesgue measurable functions . . . 18
2.4 Lebesgue integral . . . 20
2.5 Dominated convergence theorem . . . 23
2.6 L^p-spaces . . . 24
2.7 Approximation in L^p-spaces . . . 28

3 Operator theory on Hilbert spaces **31**
3.1 Hilbert space . . . 31
3.2 Bounded linear operators . . . 35
3.3 Special classes of bounded linear operators . . . 37
3.4 Vector valued functions and operator valued functions . . . 40
3.5 Unbounded and self-adjoint operators . . . 41
3.6 Resolvent and spectrum . . . 43
3.7 Examples of self-adjoint operators . . . 45
3.8 Spectral theorem . . . 47

Index **52**

Bibliography **55**

About these lectures notes:
These notes and the corresponding lectures have been mainly inspired by the books [Di], [Ne], and [Am]. Other references will be mentioned in due time.

Chapter 1

Distribution theory

1.1 Test functions and distributions

Our first goal is to give a meaning to the expression

$$\int_{\mathbb{R}} f(x)\delta(x)\,\mathrm{d}x = f(0). \tag{1.1.1}$$

What is exactly the Dirac delta function δ, how should one understand this equality, and can one generalize this in a more general context ? The shortest answer to the first question is: *δ is a continuous linear functional on smooth functions with compact support*. We shall subsequently give a meaning to this answer, and further extend it.

Let us start by recalling a few definitions borrowed from any course of calculus for functions with several real variables. We fix a natural number $n \geq 1$. For the sake of generality, we shall denote by $\mathbb{K}$ either $\mathbb{R}$ or $\mathbb{C}$. In the first part of these notes, we could deal with real functions only, but subsequently, complex functions would appear naturally. The symbol $\mathbb{K}$ covers both situations.

The canonical basis of $\mathbb{R}^n$ is denoted by $\{E_j\}_{j=1}^n$ with $E_j = (0, 0, \dots, 0, 1, 0, \dots, 0)^t$, where the only entry 1 is at the position j, and where ^{t} denotes the transpose of a horizontal vector. We shall denote any point in $\mathbb{R}^n$ by X, with $X = (x_1, x_2, \dots, x_n)$. For any $f : \mathbb{R}^n \to \mathbb{K}$ and for $j \in \{1, \dots, n\}$ we consider the derivative of f with respect to the variable x_j evaluated at $X \in \mathbb{R}^n$ by

$$[\partial_j f](X) \equiv [D_j f](X) \equiv [\nabla_j f](X) = \lim_{\varepsilon \to 0} \frac{f(X + \varepsilon E_j) - f(X)}{\varepsilon}$$

whenever the limit exists. If the limit exists for all $X \in \mathbb{R}^n$, then the map

$$\partial_j f : \mathbb{R}^n \ni X \mapsto [\partial_j f](X) \in \mathbb{K}$$

is a new function called the *partial derivative of f with respect to the variable x_j*. We can then iterate the process and consider for any $m \in \mathbb{N} := \{0, 1, 2, \dots\}$ the new function $\partial_j^m f := \partial_j \partial_j \dots \partial_j f$, where the derivative is taken m times. More generally, for any *multi-index*[1] $\alpha \in \mathbb{N}^n$, meaning that

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \quad \text{with} \quad \alpha_j \in \mathbb{N},$$

[1] If $\alpha, \beta \in \mathbb{N}^n$, observe that $(\alpha + \beta) \in \mathbb{N}^n$, where the addition is defined componentwise. We also set $|\alpha| := \sum_{j=1}^n \alpha_j$, which is a natural number.

we set

$$\partial^\alpha f := \partial_1^{\alpha_1}\partial_2^{\alpha_2}\dots\partial_n^{\alpha_n} f, \tag{1.1.2}$$

whenever all limits involved in the definition of this expression exist. It is then natural to define:

Definition 1.1.1 (Smooth functions)**.** *, A function $f : \mathbb{R}^n \to \mathbb{K}$ is* smooth*, or is a C^∞-*function*, if for any multi-indices $\alpha \in \mathbb{N}^n$ the function $\partial^\alpha f : \mathbb{R}^n \to \mathbb{K}$ exists and is continuous. The set of all smooth functions is denoted by $C^\infty(\mathbb{R}^n)$.*

Let us emphasize an important fact for such functions: If $f \in C^\infty(\mathbb{R}^n)$ and for any $j, k \in \{1, \dots, n\}$ one has

$$\partial_j[\partial_k f] = \partial_k[\partial_j f]$$

which means that the partial derivatives in (1.1.2) can be taken in any order. For that reason and for smooth functions, there is no loss of generality in ordering the partial derivatives in any convenient order.

Before the next definition, let us briefly recall some basic definitions from topology on $\mathbb{R}^n$. We denote by $\mathcal{B}_r(X)$ the open ball in $\mathbb{R}^n$ centered at X and for radius $r > 0$, namely $\mathcal{B}_r(X) := \{Y \in \mathbb{R}^n \mid \|X - Y\| < r\}$, where $\|X\|$ denotes the Euclidean norm and is defined by

$$\|X\| := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \equiv \Big(\sum_{j=1}^n x_j^2\Big)^{1/2}.$$

Let Ω be a subset of $\mathbb{R}^n$. The set Ω is *open* if whenever $X \in \Omega$ there exists $r > 0$ such that $\mathcal{B}_r(X) \subset \Omega$. It is easily observed that for any $r > 0$ and $X \in \mathbb{R}^n$, the open ball $\mathcal{B}_r(X)$ is an open set. On the other hand, the set Ω is *closed* if its complement is open, meaning that $\mathbb{R}^n \setminus \Omega$ is an open set.

We say that $X \in \mathbb{R}^n$ belongs to the *boundary* $\partial\Omega$ of Ω if for any $r > 0$, $\mathcal{B}_r(X) \cap \Omega \neq \emptyset$ and $\mathcal{B}_r(X) \cap (\mathbb{R}^n \setminus \Omega) \neq \emptyset$. In other words, any ball centred at a point of the boundary of Ω has a non-empty intersection with Ω and with the complement of Ω. Note that a set Ω is closed precisely when $\partial\Omega \subset \Omega$ while Ω is an open set precisely when $\partial\Omega \cap \Omega = \emptyset$. Furthermore, the *interior* of Ω is defined by $\{X \in \Omega \mid \mathcal{B}_r(X) \subset \Omega \text{ for some } r > 0\}$ and is often denoted by Ω^o. Clearly, the equality $\partial\Omega \cap \Omega^o = \emptyset$ always holds. We also denote by $\overline{\Omega}$ the *closure* of Ω, namely $\overline{\Omega} = \Omega \cup \partial\Omega = \Omega^o \cup \partial\Omega$. It is readily observed that the closure of Ω is always a closed set.

We are now ready for the next two definitions.

Definition 1.1.2 (Support)**.** *For any $f : \mathbb{R}^n \to \mathbb{K}$, the* support of *$f$ is defined by*

$$\operatorname{supp}(f) := \overline{\{X \in \mathbb{R}^n \mid f(X) \neq 0\}},$$

where the overline means the closure of the set in $\mathbb{R}^n$.

Definition 1.1.3 (Test function)**.** *A smooth function $f : \mathbb{R}^n \to \mathbb{K}$ with bounded support is called a* test function*. In other words, if $f \in C^\infty(\mathbb{R}^n)$ and* $\operatorname{supp}(f) \subset \mathcal{B}_r(0)$ *for $r \in \mathbb{R}$ large enough, then f is a test function. The set of all test function is denoted by $\mathcal{D}(\mathbb{R}^n)$.*

Observe that if $f \in \mathcal{D}(\mathbb{R}^n)$, then $\partial^\alpha f$ also belongs to $\mathcal{D}(\mathbb{R}^n)$, for any $\alpha \in \mathbb{N}^n$. On the other hand, it is not completely trivial to show that test functions exist. Let us exhibit one:

Example 1.1.4. *For any $X \in \mathbb{R}^n$ set*

$$f(X) := \begin{cases} \exp\big(-\frac{1}{1-\|X\|^2}\big) & \text{if } \|X\| < 1 \\ 0 & \text{if } \|X\| \geq 1. \end{cases}$$

Then $f \in \mathcal{D}(\mathbb{R}^n)$ and the support of f corresponds to $\overline{\mathcal{B}_1(0)} = \{Y \in \mathbb{R}^n \mid \|Y\| \leq 1\}$.

Exercise 1.1.5. *Exhibit other elements of $\mathcal{D}(\mathbb{R}^n)$. If necessary, you can study and use the notion of convolution, see*

https://en.wikipedia.org/wiki/Convolution

Let us provide a few properties of $\mathcal{D}(\mathbb{R}^n)$. It is a *vector space*, meaning that it is stable for the addition of two elements and for the multiplication by any element of $\mathbb{K}$. It is also an *algebra*, meaning that the product of two elements is still an element of $\mathcal{D}(\mathbb{R}^n)$. In addition, it is an *ideal* in $C^\infty(\mathbb{R}^n)$, meaning that the product of $f \in \mathcal{D}(\mathbb{R}^n)$ and $g \in C^\infty(\mathbb{R}^n)$ is again an element of $\mathcal{D}(\mathbb{R}^n)$, namely $fg \in \mathcal{D}(\mathbb{R}^n)$.

In calculus, we often consider the convergence of sequences, namely if $(x_j)_{j\in\mathbb{N}} \subset \mathbb{K}$ is a sequence of numbers (meaning $x_j \in \mathbb{K}$ for any $j \in \mathbb{N}$), then it is natural to wonder if this sequence admits a limit as $j \to \infty$. More precisely, does there exist $x_\infty \in \mathbb{K}$ such that for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with $|x_j - x_\infty| < \varepsilon$ for all $j > N$. If x_∞ exists, it is called the limit of the sequence, and we write $x_j \to x_\infty$ in $\mathbb{K}$ as $j \to \infty$. Clearly, the same notion holds for a sequence in $\mathbb{K}^n$, namely for a sequence $(X_j)_{j\in\mathbb{N}} \subset \mathbb{K}^n$ with $X_j \in \mathbb{K}^n$, this sequence admits a limit in $\mathbb{K}^n$ if there exists $X_\infty \in \mathbb{K}^n$ such that for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ with $\|X_j - X_\infty\| < \varepsilon$ for all $j > N$. If the limit exists, we write $X_j \to X_\infty$ in $\mathbb{K}^n$ as $j \to \infty$.

What about the generalization for a sequence of functions in $\mathcal{D}(\mathbb{R}^n)$?

Definition 1.1.6 (Convergence in $\mathcal{D}(\mathbb{R}^n)$)**.** *A sequence $(f_j)_{j\in\mathbb{N}} \subset \mathcal{D}(\mathbb{R}^n)$ of test functions converges to $f_\infty \in \mathcal{D}(\mathbb{R}^n)$ if the following two conditions are satisfied:*

1. *For any $\alpha \in \mathbb{N}^n$ one has*

$$\sup_{X\in\mathbb{R}^n} \left|\partial^\alpha f_j(X) - \partial^\alpha f_\infty(X)\right| \to 0 \quad \text{as} \quad j \to \infty,$$

2. *There exists $r \in \mathbb{R}$ large enough such that* $\operatorname{supp}(f_j) \subset B_r(0)$ *for all $j \in \mathbb{N}$.*

In this case, we write $f_j \to f_\infty$ in $\mathcal{D}(\mathbb{R}^n)$ as $j \to \infty$.

Note that the first condition can be equivalently written as $\left\|\partial^\alpha f_j - \partial^\alpha f_\infty\right\|_\infty \to 0$ as $j \to \infty$, where the following norm is used for any $g : \mathbb{R}^n \to \mathbb{K}$

$$\|g\|_\infty := \sup_{X\in\mathbb{R}^n} |g(X)|.$$

This norm is usually called *the* sup-*norm*, or the L^∞-*norm*. We are now ready to define the second main concept of this section:

Definition 1.1.7 (Distribution on $\mathbb{R}^n$)**.** *A* distribution on $\mathbb{R}^n$ *(or simply a* distribution*) is a continuous linear function on $\mathcal{D}(\mathbb{R}^n)$ with values in $\mathbb{K}$. More precisely, a map $T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K}$ is a distribution on $\mathbb{R}^n$ if it satisfies:*

1. *$T(f_1 + \lambda f_2) = T(f_1) + \lambda T(f_2)$ for any $f_1, f_2 \in \mathcal{D}(\mathbb{R}^n)$ and $\lambda \in \mathbb{K}$,*
2. *Whenever the sequence $(f_j)_{j\in\mathbb{N}} \subset \mathcal{D}(\mathbb{R}^n)$ converges to $f_\infty \in \mathcal{D}(\mathbb{R}^n)$, then the sequence $\big(T(f_j)\big)_{j\in\mathbb{N}}$ converges to $T(f_\infty)$ in $\mathbb{K}$ as $j \to \infty$.*

The set of all distributions on $\mathbb{R}^n$ is denoted by $\mathcal{D}'(\mathbb{R}^n)$.

It is easily observed that the set $\mathcal{D}'(\mathbb{R}^n)$ is also a vector space. We shall also use the alternative notation

$$T(f) \equiv \langle T, f\rangle \tag{1.1.3}$$

which emphasizes the linearity in both arguments (but be aware that it is linear in both argument, not antilinear in one of them).

Let us now present a few examples of distributions.

Examples 1.1.8. *1) Let us set*[2]

$$L^1_{loc}(\mathbb{R}^n) := \Big\{h : \mathbb{R}^n \to \mathbb{K} \mid \int_{B_r(Y)} |h(X)|\,\mathrm{d}X < \infty \quad \textit{for any } r > 0, Y \in \mathbb{R}^n\Big\},$$

and call it the set of locally integrable functions on $\mathbb{R}^n$. *For any $h \in L^1_{loc}(\mathbb{R}^n)$ we define $T_h \in \mathcal{D}'(\mathbb{R}^n)$ acting on $f \in \mathcal{D}(\mathbb{R}^n)$ as*

$$T_h(f) := \int_{\mathbb{R}^n} h(X)\, f(X)\,\mathrm{d}X \tag{1.1.4}$$

and call T_h a regular distribution. *Thus, any element of $L^1_{loc}(\mathbb{R}^n)$ can be identified with a distribution through the map $h \mapsto T_h$.*

2) For any $Y \in \mathbb{R}^n$ we define $\delta_Y \in \mathcal{D}'(\mathbb{R}^n)$ acting on $f \in \mathcal{D}(\mathbb{R}^n)$ as

$$\delta_Y(f) := f(Y). \tag{1.1.5}$$

By an abuse of notation we often write $\delta_Y(f) = \int_{\mathbb{R}^n} \delta_Y(X)\, f(X)\,\mathrm{d}X$, but this notation is misleading, only (1.1.5) is correct. It should be emphasized that δ_Y is not a function, it is a distribution.

3) For any $\alpha \in \mathbb{N}^n$ and $Y \in \mathbb{R}^n$ we also define $\delta^\alpha_Y \in \mathcal{D}'(\mathbb{R}^n)$ acting on $f \in \mathcal{D}(\mathbb{R}^n)$ as

$$\delta^\alpha_Y(f) := (-1)^{|\alpha|}[\partial^\alpha f](Y). \tag{1.1.6}$$

Note that notion of derivative of a distribution will be generalized subsequently.

Exercise 1.1.9. *In the framework of the previous examples, show that T_h, δ_Y, and δ^α_Y belong to $\mathcal{D}'(\mathbb{R}^n)$.*

It is clear that the expressions (1.1.4) to (1.1.6) can be defined for more general functions f. In this sense, we would like to extend the applicability of T_h, δ_Y and δ^α_Y to a larger set of functions. For that purpose, the following statement is important, and we shall come back to it later. It is also very useful for checking when a map is a distribution (since the second condition in Definition 1.1.7 is usually difficult to check).

Theorem 1.1.10. *A map $T : \mathcal{D}(\mathbb{R}^n) \to \mathbb{K}$ belongs to $\mathcal{D}'(\mathbb{R}^n)$ if and only if T is linear and if for any $Y \in \mathbb{R}^n$ and any $r > 0$ there exist $c > 0$ and $m \in \mathbb{N}$ such that*

$$|T(f)| \leq c \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_\infty \tag{1.1.7}$$

for all $f \in \mathcal{D}(\mathbb{R}^n)$ with $\mathrm{supp}(f) \subset \overline{\mathcal{B}_r(Y)}$.

If m can be made independent of Y and r, the smallest possible value for m is called the *order* of T. In this case, the constant c can still depend on Y and r, see also Definition 1.4.1.

Exercise 1.1.11. *Determine the order of the distributions T_h, δ_Y and δ^α_Y introduced in Examples 1.1.8.*

Exercise 1.1.12 (♥). *Define the support of a distribution, and determine the support of the distributions T_h, δ_Y and δ^α_Y introduced in Examples 1.1.8. You can use [Di, Sec. 2.3] or*

https://en.wikipedia.org/wiki/Support_(mathematics)

[2] In Remark 2.6.5, we shall provide a better definition of $L^1_{loc}(\mathbb{R}^n)$, and in particular give a meaning to the integral sign. In addition, we shall see that this set corresponds to equivalence classes of functions, and that the value of a function at a specific point is not important. In particular, there is no loss of generality in changing the value of any function in $L^1_{loc}(\mathbb{R}^n)$ at a finite or countable number of points.

1.2 Differentiation of distributions

Let us firstly recall the formula for the integration by parts in $\mathbb{R}$, namely

$$\int_a^b f'(x)\,g(x)\,\mathrm{d}x = f(x)\,g(x)|_a^b - \int_a^b f(x)\,g'(x)\,\mathrm{d}x.$$

In particular, if $a = -\infty$, $b = \infty$ and if $f, g \in \mathcal{D}(\mathbb{R})$ then one has

$$\int_a^b f'(x)\,g(x)\,\mathrm{d}x = -\int_a^b f(x)\,g'(x)\,\mathrm{d}x.$$

Similarly, if we consider functions $f, g \in \mathcal{D}(\mathbb{R}^n)$ and for any $j \in \{1, \dots, n\}$, then

$$\int_{\mathbb{R}^n} [\partial_j f](X)\,g(X)\,\mathrm{d}X = -\int_{\mathbb{R}^n} f(X)\,[\partial_j g](X)\,\mathrm{d}X$$

and more generally

$$\int_{\mathbb{R}^n} [\partial^\alpha f](X)\,g(X)\,\mathrm{d}X = (-1)^{|\alpha|}\int_{\mathbb{R}^n} f(X)\,[\partial^\alpha g](X)\,\mathrm{d}X$$

for any $\alpha \in \mathbb{N}^n$. It is then natural to set:

Definition 1.2.1 (Differentiation of a distribution)**.** *For any $T \in \mathcal{D}'(\mathbb{R}^n)$ and for any $\alpha \in \mathbb{N}^n$, we define $\partial^\alpha T$ acting on $f \in \mathcal{D}(\mathbb{R}^n)$ as*

$$[\partial^\alpha T](f) := (-1)^{|\alpha|}T(\partial^\alpha f).$$

*$\partial^\alpha T$ is called the α-*derivative of T.

The following statement justifies the previous definition. Its proof is left as an exercise.

Lemma 1.2.2. *In the setting of the previous definition, $\partial^\alpha T$ belongs to $\mathcal{D}'(\mathbb{R}^n)$.*

With the notation introduced in (1.1.3) the definition of the derivative of a distribution reads

$$\langle \partial^\alpha T, f\rangle = \langle T, (-1)^{|\alpha|}\partial^\alpha f\rangle.$$

Let us stress a surprising consequence of the previous definition and lemma: Any $h \in L^1_{loc}(\mathbb{R}^n)$ can be differentiated an arbitrary number of times ! Indeed, if one identifies any $h \in L^1_{loc}(\mathbb{R}^n)$ with the distribution T_h, the previous definition means that $\partial^\alpha T_h$ is well defined, for any $\alpha \in \mathbb{N}^n$. In summary,

In the sense of distribution, any function in $L^1_{loc}(\mathbb{R}^n)$ can be differentiate arbitrarily. ☺

However, if the function h is regular enough, one has:

Exercise 1.2.3. *Show that if h is sufficiently differentiable, then $\partial^\alpha T_h = T_{\partial^\alpha h}$ for $\alpha \in \mathbb{N}^n$.*

Examples 1.2.4. *1) Let us define the* Heaviside function $H : \mathbb{R} \to \mathbb{R}$ *on* $x \in \mathbb{R}$ *by*

$$H(x) := \begin{cases} 0 & \text{if} \quad x < 0 \\ 1 & \text{if} \quad x \geq 0. \end{cases}$$

Clearly, H belongs to $L^1_{loc}(\mathbb{R})$, and thus T_H belongs to $\mathcal{D}'(\mathbb{R})$. What is then ∂T_H ? For answering this question, observe that for $f \in \mathcal{D}(\mathbb{R})$ one has

$$[\partial T_H](f) = -T_H(f') = -\int_{-\infty}^{\infty} H(x)\,f'(x)\,\mathrm{d}x$$

$$= -\int_0^\infty f'(x)\,\mathrm{d}x = -f(x)|_0^\infty = f(0) = \delta_0(f).$$

As a consequence, one has obtained that $\partial T_H = \delta_0$, also written carelessly $\partial H = \delta_0$. Note that this kind of results can be extended to other discontinuous functions.

2) Consider now the principal value distribution $\mathrm{Pv}\frac{1}{x} \in \mathcal{D}'(\mathbb{R})$ *defined on $f \in \mathcal{D}(\mathbb{R})$ by*

$$\mathrm{Pv}\frac{1}{x}(f) := \lim_{\varepsilon\searrow 0}\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \frac{1}{x} f(x)\,\mathrm{d}x = \lim_{\varepsilon\searrow 0}\Big(\int_{-\infty}^{-\varepsilon}\frac{1}{x} f(x)\,\mathrm{d}x + \int_{\varepsilon}^{\infty}\frac{1}{x} f(x)\,\mathrm{d}x\Big). \tag{1.2.1}$$

Then, let us observe that

$$\begin{aligned}\lim_{\varepsilon\searrow 0}\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \frac{1}{x} f(x)\,\mathrm{d}x &= \lim_{\varepsilon\searrow 0}\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \ln(|x|)' f(x)\,\mathrm{d}x \\ &= -\lim_{\varepsilon\searrow 0}\Big(\int_{\mathbb{R}\setminus(-\varepsilon,\varepsilon)} \ln(|x|) f'(x)\,\mathrm{d}x + O(\varepsilon|\ln(|\varepsilon|)|)\Big) \\ &= -\int_{\mathbb{R}} \ln(|x|) f'(x)\,\mathrm{d}x,\end{aligned}$$

which means that

$$\mathrm{Pv}\frac{1}{x}(f) = -\int_{\mathbb{R}} \ln(|x|) f'(x)\,\mathrm{d}x = [\partial T_{\ln(|\cdot|)}](f).$$

In careless terms, it means that $\partial \ln(|\cdot|) \equiv \ln(|\cdot|)' = \mathrm{Pv}\frac{1}{x}$. This can be seen as an illustration that any $L^1_{loc}(\mathbb{R})$-function can be differentiated in the distribution sense.

Exercise 1.2.5. *Check that $\mathrm{Pv}\frac{1}{x}$ defined in* (1.2.1) *is a distribution, and compute the exact expression for $O(\varepsilon|\ln(|\varepsilon|)|)$ in the above computation.*

Exercise 1.2.6 (♥). *For $n = 3$, consider the function $h : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}$ given by $h(X) := \frac{1}{\|X\|}$. Check that $h \in L^1_{loc}(\mathbb{R}^3)$, that $T_h \in \mathcal{D}'(\mathbb{R}^3)$, and that the following equality holds:*

$$\Delta T_h \equiv (\partial_1^2 + \partial_2^2 + \partial_3^2)T_h = -4\pi\delta_0.$$

In careless terms, this equality reads $\Delta\frac{1}{\|\cdot\|} = -4\pi\delta_0$, or even more carelessly $\Delta\frac{1}{\|x\|} = -4\pi\delta_0(x)$. ☹

The fact that regular distributions can be arbitrarily differentiated is central in the theory of distributions. Indeed, the following deep result holds:

Theorem 1.2.7 (Local structure theorem)**.** *For any $T \in \mathcal{D}'(\mathbb{R}^n)$ and for any bounded open subset $\Omega \subset \mathbb{R}^n$, there exist a continuous function $h : \mathbb{R}^n \to \mathbb{K}$ and a multi-index $\alpha \in \mathbb{N}^n$ such that $T(f) = [\partial^\alpha T_h](f)$ for all $f \in \mathcal{D}(\mathbb{R}^n)$ with* $\mathrm{supp}(f) \subset \Omega$*. Note that h and α depend on Ω.*

We refer to [Di, Thm. 5.6] for a proof, which involves the Hahn-Banach theorem.

1.3 Other operations with distributions

In addition to the differentiation of a distribution, let us consider another operation on distributions, namely *the multiplication of a distribution by a smooth function*. We consider $T \in \mathcal{D}'(\mathbb{R}^n)$ and let $g \in C^\infty(\mathbb{R}^n)$. We then define gT acting on $f \in \mathcal{D}(\mathbb{R}^n)$ by

$$[gT](f) := T(gf). \tag{1.3.1}$$

Observe that the r.h.s. is well defined since the product gf belongs to $\mathcal{D}(\mathbb{R}^n)$. Then, one easily checks that gT is indeed an element of $\mathcal{D}'(\mathbb{R}^n)$. In addition, if T is a regular distribution, namely if there exists $h \in L^1_{loc}(\mathbb{R}^n)$ such that $T = T_h$, then one has

$$[gT_h](f) = T_h(gf) = \int_{\mathbb{R}^n} h(X)[gf](X)\,\mathrm{d}X = \int_{\mathbb{R}^n} h(X)g(X)f(X)\,\mathrm{d}X = \int_{\mathbb{R}^n} [gh](X)f(X)\,\mathrm{d}X = T_{gh}(f)$$

which means that $gT_h = T_{gh}$ with $gh \in L^1_{loc}(\mathbb{R}^n)$.

The following result can also be easily obtained, the proof is left as an exercise.

Lemma 1.3.1. *For any $g \in C^\infty(\mathbb{R}^n)$ one has $g\delta_0 = g(0)\delta_0$. If $n = 1$, one also has $g\delta_0' = g(0)\delta_0' - g'(0)\delta_0$.*

Exercise 1.3.2. *By using Leibniz rule, generalize the previous result for any $n \in \{1, 2, 3, \dots\}$.*

It is then natural to wonder if distributions can be multiplied ? If S and T are elements of $\mathcal{D}'(\mathbb{R}^n)$, what would be the meaning of ST ? In general, one can not give a meaning to this product. On the other hand, there exists a notion of convolution product of two distributions, but not all distributions can be convoluted. We do not develop this topic and refer to [Di, Chap. 6].

Exercise 1.3.3 (♥). *Summarize the notion of convolution product of two distributions (when it is defined) and provide a few examples, see [Di, Chap. 6].*

In Definition 1.1.6, the notion of convergence was defined for a sequence $(f_j)_{j\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n)$. Let us now consider a similar notion but in $\mathcal{D}'(\mathbb{R}^n)$.

Definition 1.3.4 (Convergence in $\mathcal{D}'(\mathbb{R}^n)$)**.** *A sequence $(T_j)_{j\in\mathbb{N}} \subset \mathcal{D}'(\mathbb{R}^n)$ of distributions converges to $T_\infty \in \mathcal{D}'(\mathbb{R}^n)$ if $\lim_{j\to\infty} T_j(f) = T_\infty(f)$ for all $f \in \mathcal{D}(\mathbb{R}^n)$. In this case, we write $T_j \to T_\infty$ in $\mathcal{D}'(\mathbb{R}^n)$ as $j \to \infty$.*

In fact, a deeper result exists:

Theorem 1.3.5. *Let $(T_j)_{j\in\mathbb{N}} \subset \mathcal{D}'(\mathbb{R}^n)$ be a sequence of distributions and assume that $\lim_{j\to\infty} T_j(f) \in \mathbb{C}$ exists for all $f \in \mathcal{D}(\mathbb{R}^n)$. Then, there exists $T_\infty \in \mathcal{D}'(\mathbb{R}^n)$ such that $\lim_{j\to\infty} T_j(f) = T_\infty(f)$ for all $f \in \mathcal{D}(\mathbb{R}^n)$.*

Exercise 1.3.6 (♥). *Study the previous result, as presented in [Di, Thm. 10.10]. It is based on the notion of Fréchet spaces and on the Banach-Steinhaus theorem.*

A simple consequence of Definition 1.3.4 is the following lemma, whose proof is left as an exercise.

Lemma 1.3.7. *If $(T_j)_{j\in\mathbb{N}} \subset \mathcal{D}'(\mathbb{R}^n)$ is a sequence of distribution converging to $T_\infty \in \mathcal{D}'(\mathbb{R}^n)$, then for any $\alpha \in \mathbb{N}^n$, the sequence $(\partial^\alpha T_j)_{j\in\mathbb{N}} \subset \mathcal{D}'(\mathbb{R}^n)$ converges to $\partial^\alpha T_\infty \in \mathcal{D}'(\mathbb{R}^n)$.*

The notion of convergence of distributions is also very useful for defining the distribution δ_0 (also called *Dirac delta function*) as a limit of regular distributions.

Exercise 1.3.8. *Consider $h : \mathbb{R}^n \to \mathbb{K}$ satisfying $\int_{\mathbb{R}^n} |h(X)|\,\mathrm{d}X < \infty$, and assume that $\int_{\mathbb{R}^n} h(X)\,\mathrm{d}X = 1$. For $j \in \mathbb{N}$, set $h_j(X) := j^n h(jX)$. Then, prove that $T_{h_j} \to \delta_0$ in $\mathcal{D}'(\mathbb{R}^n)$ as $j \to \infty$. Equivalently, for $\varepsilon > 0$ one often sets $h_\varepsilon(X) := \frac{1}{\varepsilon^n} h\left(\frac{X}{\varepsilon}\right)$. Show that $T_{h_\varepsilon} \to \delta_0$ in $\mathcal{D}'(\mathbb{R}^n)$ as $\varepsilon \searrow 0$.*

Exercise 1.3.9. *For $j \in \mathbb{N}$, consider $h_j(x) := \frac{\sin(jx)}{x}$ for any $x \in \mathbb{R} \setminus \{0\}$ and $h_j(0) := j$. Show that $T_{h_j} \to \pi\delta_0$ in $\mathcal{D}'(\mathbb{R})$ as $j \to \infty$. Why is this situation not covered by the previous exercise ?*

Exercise 1.3.10. *For $\varepsilon > 0$ consider the function $h_{\pm\varepsilon}(x) := \frac{1}{x \pm i\varepsilon}$ for all $x \in \mathbb{R}$. Show that $T_{h_{\pm\varepsilon}}$ converges to $\mathrm{Pv}\frac{1}{x} \mp i\pi\delta_0$ in $\mathcal{D}'(\mathbb{R})$ as $\varepsilon \searrow 0$, where the Principal value distribution $\mathrm{Pv}\frac{1}{x}$ has been introduced in Examples 1.2.4.*

Exercise 1.3.11. *1) For $j \in \mathbb{N}$ consider the function defined for $x \in \mathbb{R}$ by*

$$D_j(x) := \begin{cases} \sum_{k=-j}^{j} \mathrm{e}^{ikx} & \text{if } x \in (-\pi,\pi) \\ 0 & \text{if } x \notin (-\pi,\pi) \end{cases} = \begin{cases} \frac{\sin\left((j+\frac{1}{2})x\right)}{\sin\left(\frac{x}{2}\right)} & \text{if } x \in (-\pi,\pi) \\ 0 & \text{if } x \notin (-\pi,\pi). \end{cases}$$

Then show that T_{D_j} converges to $2\pi\delta_0$ in $\mathcal{D}'(\mathbb{R})$ as $j \to \infty$. This is often written

$$\lim_{j\to\infty} \int_{-\pi}^{\pi} D_j(x) f(x)\,\mathrm{d}x = 2\pi f(0)$$

for any $f \in \mathcal{D}(\mathbb{R})$. D_j is called the Dirichlet kernel.

2) By defining the Cesaro summation of Dirichlet kernel*, namely for $k \in \{1, 2, \dots\}$ and $x \in \mathbb{R}$*

$$F_k(x) := \frac{1}{k}\sum_{j=0}^{j-1} D_j(x) = \begin{cases} \frac{1}{k}\left(\frac{\sin\left(\frac{kx}{2}\right)}{\sin\left(\frac{x}{2}\right)}\right)^2 & \text{if } x \in (-\pi,\pi) \\ 0 & \text{if } x \notin (-\pi,\pi). \end{cases}$$

Then show that T_{F_k} converges to $2\pi\delta_0$ in $\mathcal{D}'(\mathbb{R})$ as $k \to \infty$. This is often written

$$\lim_{k\to\infty} \int_{-\pi}^{\pi} F_k(x) f(x)\,\mathrm{d}x = 2\pi f(0)$$

for any $f \in \mathcal{D}(\mathbb{R})$. D_k is called the Fejér kernel.

1.4 Continuous extension of distributions

As already mentioned in Section 1.1, one is often interested in extending the applicability of a distribution to a larger set of functions, and not only to $\mathcal{D}(\mathbb{R}^n)$. This can be done with some care and for some distributions.

The following definition complement the content of Theorem 1.1.10.

Definition 1.4.1 (Summable distribution)**.** *A distribution $T \in \mathcal{D}'(\mathbb{R}^n)$ is* summable *if there exists $m \in \mathbb{N}$ and $c > 0$ such that*

$$|T(f)| \le c \sum_{|\alpha|\le m} \|\partial^\alpha f\|_\infty \tag{1.4.1}$$

for all $f \in \mathcal{D}(\mathbb{R}^n)$. The smallest m satisfying this condition is called the summability-order *of T.*

By Theorem 1.1.10, any distribution is locally a summable distribution, but in the above definition the value of m should hold globally. It turns out that summable distributions of summability-order m can be continuously extended to $C_b^m(\mathbb{R}^n) \equiv BC^m(\mathbb{R}^n)$, the set of all m-times continuously differentiable functions with all derivatives $\partial^\alpha f$ bounded, for any $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$. More precisely, one has:

Lemma 1.4.2. *Let $T \in \mathcal{D}'(\mathbb{R}^n)$ be a summable distribution of summability-order m, then T extends uniquely to a continuous linear functional on $C_b^m(\mathbb{R}^n)$ which has the bounded convergence property of order m.*

Let us stress that the continuity property mentioned in this lemma corresponds exactly to the inequality (1.4.1). For the bounded convergence property, it means that if we consider a sequence $(g_j)_{j\in\mathbb{N}} \subset C_b^m(\mathbb{R}^n)$ with $\sup_j \|\partial^\alpha g_j\|_\infty < \infty$ for any $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$, and if for any $Y \in \mathbb{R}^n$, any $r > 0$ and any $\alpha \in \mathbb{N}^n$ with $|\alpha| \le m$ one has

$$\lim_{j\to\infty} \sup_{X\in\mathcal{B}_r(Y)} |\partial^\alpha g_j(X)| = 0,$$

then $\lim_{j\to\infty} T(g_j) = 0$. This property is quite natural, and very convenient for the computations. Indeed, consider $f \in C_b^m(\mathbb{R}^n)$ and let $h \in \mathcal{D}(\mathbb{R}^n)$ with $h(0) = 1$. Set then $f_j(X) := h(\frac{X}{j})f(X)$. Observe that the sequence $(g_j)_{j\in\mathbb{N}}$ with $g_j := f - f_j$ satisfies the conditions mentioned above, from which one infers that

$$T(f) = \lim_{j\to\infty} T(f_j).$$

Note that the above formula corresponds to an approximation of f by functions which have better properties, it is not an approximation of T by a sequence of distributions. For fix $f \in C_b^m(\mathbb{R}^n)$, approximating T by a sequence $(T_j)_{j\in\mathbb{N}} \subset \mathcal{D}'(\mathbb{R}^n)$ converging to T in $\mathcal{D}'(\mathbb{R}^n)$ is more delicate. Indeed, if $T_j(g) \to T(g)$ as $j \to \infty$ for all $g \in \mathcal{D}(\mathbb{R}^n)$, it does not automatically implies that $T_j(f)$ is well defined for $f \in C_b^m(\mathbb{R}^n)$, and therefore it does not imply that $T_j(f) \to T(f)$ as $j \to \infty$. Such convergence holds only for specific sequences $(T_j)_{j\in\mathbb{N}}$.

Exercise 1.4.3. *Illustrate the last observation of the previous paragraph based on sequences of distributions converging to the Dirac delta function.*

As a final remark, observe that the constant function $\mathbf{1} : \mathbb{R}^n \ni x \mapsto 1 \in \mathbb{K}$ belongs to $C_b^m(\mathbb{R}^n)$ for any $m \in \mathbb{N}$ but does not belong to $\mathcal{D}(\mathbb{R}^n)$. Thus, for any summable distribution, one has $T(\mathbf{1}) < \infty$. If T is a regular distribution, with $T = T_h$ and $h \in L^1_{loc}(\mathbb{R}^n)$, then it means that h has to satisfy $\int_{\mathbb{R}^n} |h(X)|\, \mathrm{d}X < \infty$. Being summable is a very strong requirement on the distribution T. In fact, it can be shown that any summable distribution is of the form $T = \sum_\alpha \partial^\alpha h_\alpha$ for $h_\alpha \in L^1(\mathbb{R}^n) := \{g : \mathbb{R}^n \to \mathbb{K} \mid \int_{\mathbb{R}^n} |g(X)|\, \mathrm{d}X < \infty\}$, and where the sum is a finite sum.

1.5 Fourier transform, Schwartz functions, and tempered distributions

The Fourier transform is a very useful transformation which appears in several contexts. We introduce it and provide some related constructions in the framework of distribution theory.

For $f \in L^1(\mathbb{R}^n)$ and $\xi \in \mathbb{R}^n$ we set

$$[\mathcal{F}f](\xi) \equiv \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-i\xi\cdot X} f(X)\, \mathrm{d}X \tag{1.5.1}$$

and call it *the Fourier transform of f*. Note that the letter k (or K with our current convention) is often used by physicists instead of ξ. Note also that various normalizations are possible, and that coefficients 2π can appear at several places in equivalent definitions.

Let us list some properties of the Fourier transformation, for $f, g \in L^1(\mathbb{R}^n)$:

1. $\mathcal{F}$ is a linear map on $L^1(\mathbb{R}^n)$,
2. $|\hat{f}(\xi)| \leq \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} |f(X)|\, \mathrm{d}X$,
3. $\hat{f}$ belongs to $C_0(\mathbb{R}^n)$, meaning that $\hat{f}$ is a continuous function on $\mathbb{R}^n$ satisfying $\lim_{\|\xi\|\to\infty} \hat{f}(\xi) = 0$,
4. $\mathcal{F}(f * g) \equiv \widehat{f * g} = \hat{f}\,\hat{g}$, where the convolution of f and g is defined by

$$[f * g](X) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(X - Y)g(Y)\, \mathrm{d}Y, \tag{1.5.2}$$

5. If $\partial_j f$ exists and belongs to $L^1(\mathbb{R}^n)$, then

$$[\mathcal{F}(-i\partial_j f)](\xi) \equiv [\widehat{-i\partial_j f}](\xi) = \xi_j \hat{f}(\xi). \tag{1.5.3}$$

As a consequence of 1. and 3., observe that $\mathcal{F}$ is a linear map from $L^1(\mathbb{R}^n)$ to $C_0(\mathbb{R}^n)$.

Exercise 1.5.1. *Prove these relations, or at least some of them. If necessary, consider $f, g \in \mathcal{D}(\mathbb{R}^n)$ instead of $f, g \in L^1(\mathbb{R}^n)$.*

The following statement plays a very important role in quantum mechanics. Its precise meaning will become clearer in the chapter on Lebesgue integral.

Theorem 1.5.2. *The Fourier transform $\mathcal{F}$ extends continuously from $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ to a bijective map from $L^2(\mathbb{R}^n)$ to itself, with inverse map given by*

$$[\mathcal{F}^{-1} f](X) \equiv \check{f}(X) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{i\xi\cdot X} f(\xi)\,\mathrm{d}\xi$$

for any $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. In addition, the equality $\|f\|_{L^2(\mathbb{R}^n)} = \|\hat{f}\|_{L^2(\mathbb{R}^n)}$ holds, with

$$\|f\|_{L^2(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n} |f(X)|^2\,\mathrm{d}X\right)^{1/2} \quad \text{and} \quad \|\hat{f}\|_{L^2(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n} |\hat{f}(\xi)|^2\,\mathrm{d}\xi\right)^{1/2}.$$

Note that the equality between the two norms is often referred to as Parseval theorem or the Parseval–Plancherel identity.

Let us stress that the expression (1.5.1) is not well defined for $f \in L^2(\mathbb{R}^n)$, since the integral might not converge absolutely. The way to define the Fourier transform on L^2-functions is through a limiting process, namely

$$[\mathcal{F} f](\xi) := \lim_{L\to\infty} \frac{1}{(2\pi)^{n/2}} \int_{[-L,L]^n} \mathrm{e}^{-i\xi\cdot X} f(X)\,\mathrm{d}X,$$

and still, this should not be understood for all ξ but only for almost every $\xi \in \mathbb{R}^n$. Clarification will be given in the chapter on Lebesgue integral.

Our next aim is to give a meaning to the Fourier transform of a distribution. Unfortunately, it is meaningless to define $[\mathcal{F}T](f)$ by $T(\mathcal{F}f)$ for any $T \in \mathcal{D}'(\mathbb{R}^n)$ and $f \in \mathcal{D}(\mathbb{R}^n)$ simply because $\mathcal{F}f \notin \mathcal{D}(\mathbb{R}^n)$ in general.

Exercise 1.5.3. *Show that $\mathcal{F}\mathcal{D}(\mathbb{R}^n) \not\subset \mathcal{D}(\mathbb{R}^n)$. For that purpose, it is enough to exhibit a counterexample.*

In order to give a meaning to an equality of the from $[\mathcal{F}T](f) := T(\mathcal{F}f)$, we need to define a set which is stable under the Fourier transform, and define distributions on this set. For $\alpha \in \mathbb{N}^n$ we shall use the notation X^α for the polynomial $X^\alpha := x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$.

Definition 1.5.4 (Schwartz function, Schwartz space)**.** *A function $f \in C^\infty(\mathbb{R}^n)$ is called a* Schwartz function *if for any $\alpha, \beta \in \mathbb{N}^n$ one has*

$$\left\|X^\beta \partial^\alpha f\right\|_\infty \equiv \sup_{X\in\mathbb{R}^n} \left|X^\beta [\partial^\alpha f](X)\right| < \infty.$$

The set of all Schwartz functions is denoted by $\mathcal{S}(\mathbb{R}^n)$.

Exercise 1.5.5. *Exhibit some functions which belong to $\mathcal{S}(\mathbb{R}^n)$ but which are not in $\mathcal{D}(\mathbb{R}^n)$.*

Observe that the following inclusions hold:

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset C_b^\infty(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n) \subset L^1_{loc}(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n).$$

In addition, one observes that $\mathcal{S}(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, which means that the Fourier transform can be defined on any $f \in \mathcal{S}(\mathbb{R}^n)$. We mention below three properties of $\mathcal{S}(\mathbb{R}^n)$ which can be easily proved. The last one is a generalization of (1.5.3).

1. $\mathcal{S}(\mathbb{R}^n)$ is a vector space, but it is not an ideal in $C^\infty(\mathbb{R}^n)$,
2. If $f \in \mathcal{S}(\mathbb{R}^n)$, then for any $\alpha,\beta \in \mathbb{N}^n$ the function $X^\beta \partial^\alpha f$ is still an element of $\mathcal{S}(\mathbb{R})$,
3. For $f \in \mathcal{S}(\mathbb{R}^n)$ and any $\alpha,\beta \in \mathbb{N}^n$ the following equality holds:
$$\mathcal{F}\Big(X^\beta (-i\partial)^\alpha f\Big) = (i\partial)^\beta X^\alpha \hat{f}. \tag{1.5.4}$$

Exercise 1.5.6. *Prove the above three properties.*

A slightly more involved property is contained in the subsequent statement, whose proof is left as an exercise.

Lemma 1.5.7. $\mathcal{F}(\mathcal{S}(\mathbb{R}^n)) = \mathcal{S}(\mathbb{R}^n)$.

For the space $\mathcal{S}(\mathbb{R}^n)$ a slightly different notion of convergence can be defined.

Definition 1.5.8 (Convergence in $\mathcal{S}(\mathbb{R}^n)$)**.** *A sequence $(f_j)_{j\in\mathbb{N}} \subset \mathcal{S}(\mathbb{R}^n)$ converges to $f_\infty \in \mathcal{S}(\mathbb{R}^n)$ if for any $\alpha,\beta \in \mathbb{N}^n$ one has $\big\|X^\beta \partial^\alpha (f_j - f_\infty)\big\|_\infty \to 0$ as $j \to \infty$. In this case, we write $f_j \to f_\infty$ in $\mathcal{S}(\mathbb{R}^n)$ as $j \to \infty$.*

Thus, the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is going to replace the set of test functions $\mathcal{D}(\mathbb{R}^n)$ for defining a new type of distributions.

Definition 1.5.9 (Tempered distribution)**.** *A* tempered distribution *T is a distribution which is continuous on $\mathcal{S}(\mathbb{R}^n)$, which means that whenever $f_j \to f_\infty$ in $\mathcal{S}(\mathbb{R}^n)$ as $j \to \infty$, then $T(f_j) \to T(f_\infty)$ as $j \to \infty$. The set of all tempered distributions is denoted by $\mathcal{S}'(\mathbb{R}^n)$.*

By definition, one has $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$. In addition, these two sets are not equal because there exist elements of $\mathcal{D}'(\mathbb{R}^n)$ which do not belong to $\mathcal{S}'(\mathbb{R}^n)$. For example, observe that the regular distribution $T_h \in \mathcal{D}'(\mathbb{R})$ defined by $h(x) := \mathrm{e}^{x^2}$ does not belong to $\mathcal{S}'(\mathbb{R})$. On the other hand, one readily observes that if $T \in \mathcal{S}'(\mathbb{R}^n)$, then $\partial^\alpha T \in \mathcal{S}'(\mathbb{R}^n)$, for any $\alpha \in \mathbb{N}^n$.

A result similar to Theorem 1.1.10 can also be proved. Note that in the present situation no localization is necessary.

Theorem 1.5.10. *The distribution T is tempered if and only if there exist $c > 0$ and $m \in \mathbb{N}$ such that*
$$|T(f)| \le c \sum_{|\alpha|,|\beta|\le m} \big\|X^\beta \partial^\alpha f\big\|_\infty \tag{1.5.5}$$
for all $f \in \mathcal{S}(\mathbb{R}^n)$.

We can now define the Fourier transform of these distributions, since they have been constructed for that purpose:

Definition 1.5.11 (Fourier transform of tempered distributions)**.** *For any $T \in \mathcal{S}'(\mathbb{R}^n)$, we set $[\mathcal{F}T](f) := T(\mathcal{F}f)$ for any $f \in \mathcal{S}(\mathbb{R}^n)$.*

Clearly, $\mathcal{F}T$ belongs to $\mathcal{S}'(\mathbb{R}^n)$ for any $T \in \mathcal{S}'(\mathbb{R}^n)$. In fact, the map $\mathcal{F} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is a bijective map, since $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is bijective. We gather in the following exercises a few equalities about tempered distributions.

Exercise 1.5.12. *Prove the following equalities (and give a meaning to them if necessary):*

1. *$\mathcal{F}\delta_0 = T_{(2\pi)^{-n/2}}$, meaning that $\mathcal{F}\delta_0$ is equal to the regular distribution based on the constant function $X \mapsto \frac{1}{(2\pi)^{n/2}}1$. This is often written simply $\mathcal{F}\delta_0 = \frac{1}{(2\pi)^{n/2}}$,*

2. $\mathcal{F}T_{\boldsymbol{1}} = (2\pi)^{n/2}\delta_0$,

3. $\mathcal{F}T_h = T_{\hat{h}}$ *for any* $h \in L^2(\mathbb{R}^n)$.

Exercise 1.5.13. *Consider the function* $h : \mathbb{R} \to \mathbb{R}$ *with* $h(x) = x^2$. *Check that* $T_h \in \mathcal{S}'(\mathbb{R})$ *and determine* $\mathcal{F}T_h$. *Similar question for* $h : \mathbb{R}^n \ni X \mapsto X^2 \in \mathbb{R}$ *and* $T_h \in \mathcal{S}'(\mathbb{R}^n)$, *what is* $\mathcal{F}T_h$ *?*

Chapter 2

Lebesgue theory of integration

In this chapter, we provide the necessary notions for understanding the theory of integration of Lebesgue. This approach extends the construction of Riemann integrals, usually studied in a first or second course of calculus. Our main reference for this chapter is [Ne]. Note that we concentrate on integrals on $\mathbb{R}$, but the construction is similar for integrals on $\mathbb{R}^n$. Also, we start with real valued functions, but later we shall consider functions with values in $\mathbb{C}$.

2.1 Reminder on Riemann integration

Our goal is to recall the meaning of the Riemann integral $\int_a^b f(x)\,\mathrm{d}x$ for suitable functions f defined on the closed interval $[a, b]$. For $n \in \mathbb{N}$, we firstly consider a n-partition $\mathcal{P}$ of $[a, b]$, namely

$$\mathcal{P} = \{x_0, x_1, \dots, x_{n-1}, x_n\} \quad \text{with} \quad a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b. \tag{2.1.1}$$

Note that a *regular partition* means that all subintervals of the partition are of the same length. We also consider the set of bounded functions on $[a, b]$ defined by

$$\mathcal{L}^\infty([a, b]) := \Big\{f : [a, b] \to \mathbb{R} \mid \|f\|_\infty := \sup_{x\in[a,b]} |f(x)| < \infty\Big\}.$$

Given a n-partition $\mathcal{P}$ of $[a, b]$ and a function $f \in \mathcal{L}^\infty([a, b])$ we set

$$L(f, \mathcal{P}) := \sum_{j=1}^{n} \Big(\inf_{x\in[x_{j-1},x_j]} f(x)\Big)(x_j - x_{j-1})$$

and

$$U(f, \mathcal{P}) := \sum_{j=1}^{n} \Big(\sup_{x\in[x_{j-1},x_j]} f(x)\Big)(x_j - x_{j-1}).$$

Here, $L(f, \mathcal{P})$ stands for *lower Riemann sum*, while $U(f, \mathcal{P})$ stands for *upper Riemann sum*. Observe that the following inequalities hold:

$$(b - a) \inf_{x\in[a,b]} f(x) \le L(f, \mathcal{P}) \le U(f, \mathcal{P}) \le (b - a) \sup_{x\in[a,b]} f(x)$$

where $\inf_{x\in[a,b]} f(x)$ and $\sup_{x\in[a,b]} f(x)$ are well defined since f is a bounded function on $[a,b]$.

Observe then that if $\mathcal{P}'$ is a finer partition of $[a,b]$, meaning that $\mathcal{P}\subset\mathcal{P}'$ ($\mathcal{P}'$ contains the points of $\mathcal{P}$ and additional points, thus it contains more subdivisions of $[a.b]$), then one has

$$L(f,\mathcal{P}) \leq L(f,\mathcal{P}') \leq U(f,\mathcal{P}') \leq U(f,\mathcal{P}).$$

By considering then all possible partitions of $[a,b]$, for arbitrary n, we can consider the

$$\textit{lower Riemann integral } \sup_{\mathcal{P}} L(f,\mathcal{P})$$

and the

$$\textit{upper Riemann integral } \inf_{\mathcal{P}} U(f,\mathcal{P}).$$

One is naturally led to the following definition.

Definition 2.1.1 (Riemann integral function)**.** *A function $f\in\mathcal{L}^\infty([a,b])$ is* Riemann integrable *on $[a,b]$ if*

$$\sup_{\mathcal{P}} L(f,\mathcal{P}) = \inf_{\mathcal{P}} U(f,\mathcal{P}).$$

In this case, we write $\int_a^b f(x)\,\mathrm{d}x$ for this value.

A standard example of a function which is not Riemann integrable is given by

$$f:[0,1]\to\mathbb{R},\ f(x)=\begin{cases}1 & \text{if } x\in\mathbb{Q},\\ 0 & \text{if } x\in\mathbb{R}\setminus\mathbb{Q}.\end{cases} \tag{2.1.2}$$

Exercise 2.1.2. *Show that the previous function is not Riemann integrable.*

Exercise 2.1.3. *Consider the function $h:[0,10]\to\mathbb{R}$ defined by $h(x)=1$ if $x\in[\sqrt{2},2\sqrt{2}]$ and $h(x)=0$ otherwise. By using regular partitions of $[0,10]$, show that the function h is Riemann integrable on $[0,10]$.*

Let us still state one standard result related to Riemann integrability. Proof can be found in [Ne, Thm. 0.2.4] or done as an exercise.

Theorem 2.1.4. *A bounded function $f:[a,b]\to\mathbb{R}$ is Riemann integrable if and only if for any $\varepsilon>0$ there exists a partition $\mathcal{P}$ of $[a,b]$ such that $U(f,\mathcal{P})-L(f,\mathcal{P})\leq\varepsilon$.*

Based on this, it is possible to prove that all continuous functions on $[a,b]$ are Riemann integrable. More precisely, one has:

Theorem 2.1.5 (Fundamental theorem of calculus)**.** *Let $f:[a,b]\to\mathbb{R}$ be a continuous function, let $x\in[a,b]$ and set $F(x):=\int_a^x f(y)\mathrm{d}y$. Then $F:[a,b]\to\mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b) with $F'(x)=f(x)$ for any $x\in(a,b)$.*

Not only continuous functions are Riemann integrable, for example monotone functions are also Riemann integrable, as proved in the following exercise.

Exercise 2.1.6 (Monotone functions are Riemann integrable)**.** *Consider a function $f:[a,b]\to\mathbb{R}$ which is increasing, and let $\mathcal{P}_n$ be the regular partition of $[a,b]$ with n subintervals.*

1. *Write $U(f,\mathcal{P}_n)$ and $L(f,\mathcal{P}_n)$ as precisely as possible,*
2. *Compute $U(f,\mathcal{P}_n)-L(f,\mathcal{P}_n)$ (if you don't see anything, try with $n=2$ and $n=3$),*
3. *By using Theorem 2.1.4, conclude that f is Riemann integrable.*

2.2 Lebesgue measure

We shall now start the study of Lebesgue integrals by starting with the notion of Lebesgue measure. In this section, the construction is done in $\mathbb{R}^n$, we shall come back to $\mathbb{R}$ only in the subsequent section.

Definition 2.2.1 (Closed box or interval)**.** *Given any $A = (a_1, \dots, a_n)$ and $B = (b_1, \dots, b_n)$ in $\mathbb{R}^n$, a* closed box *or an* interval *in $\mathbb{R}^n$ consists in the set*

$$I := \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid a_j \leq x_j \leq b_j \textit{ for all } j \in \{1, \dots, n\}\}.$$

By imposing that $a_j < b_j$ for any $j \in \{1, \dots, n\}$, the volume $v(I)$ of I is defined by $v(I) := \prod_{j=1}^n (b_j - a_j) > 0$.

In the sequel, we shall cover subsets of $\mathbb{R}^n$ by a collection of closed boxes, and get an upper estimate on its volume.

Definition 2.2.2 (Covering)**.** *For $\Omega \subset \mathbb{R}^n$, the set $S := \{I_j\}_j$ is a* covering *of Ω if $\Omega \subset \cup_j I_j$. The family S can be finite or countably infinite. We set*

$$\sigma(S) := \sum_j v(I_j) \ \in (0, \infty].$$

Based on this notion of covering, we define the first notion of volume for Ω.

Definition 2.2.3 (Lebesgue outer measure)**.** *For $\Omega \subset \mathbb{R}^n$, the* Lebesgue outer measure *of Ω is defined by*

$$m^*(\Omega) := \inf \{\sigma(S) \mid S \textit{ covering of } \Omega\}.$$

Recall that an infimum is not not always realized. However, for any $\varepsilon > 0$ there exists a covering S of Ω with $m^*(\Omega) \leq \sigma(S) \leq m^*(\Omega) + \varepsilon$.

Exercise 2.2.4. *Show that $m^*(I) = v(I)$ for any closed box I, see for example [Ne, Prop. 1.1.11].*

The following properties can be proved for the outer measure:

1. If $\Omega_1 \subset \Omega_2$, then $m^*(\Omega_1) \leq m^*(\Omega_2)$,
2. $m^*(\Omega_1 \cup \Omega_2) \leq m^*(\Omega_1) + m^*(\Omega_2)$,
3. $m^*(U_j \Omega_j) \leq \sum_j m^*(\Omega_j)$ for a finite or countable family.

Exercise 2.2.5. *Prove the above properties. Even if these properties look natural, proving them is not so simple. Some inspiration can be obtained from [Ne, Prop. 1.1.8 & 1.1.9].*

Unfortunately, the Lebesgue outer measure has also some unpleasant feature. There exist sets Ω_1, Ω_2 with $\Omega_1 \cap \Omega_2 = \emptyset$ but with $m^*(\Omega_1 \cup \Omega_2) < m^*(\Omega_1) + m^*(\Omega_2)$. Even if these sets are difficult to construct, a measure should not have such a bad property. The following result will play an important role for avoiding such singular behavior.

Theorem 2.2.6. *Let $\Omega \subset \mathbb{R}^n$ with $m^*(\Omega) < \infty$. Then for any $\varepsilon > 0$ there exists an open set Λ with $\Omega \subset \Lambda$ and such that*

$$m^*(\Lambda) \leq m^*(\Omega) + \varepsilon.$$

If $m^*(\Omega) = \infty$, the statement holds with $\Lambda = \mathbb{R}^n$.

One infers from the above properties that

$$m^*(\Lambda) = m^*(\Omega \cup (\Lambda \setminus \Omega)) \leq m^*(\Omega) + m^*(\Lambda \setminus \Omega)$$

which is equivalent to

$$m^*(\Lambda) - m^*(\Omega) \leq m^*(\Lambda \setminus \Omega).$$

However, this inequality does not provide any useful information on $m^*(\Lambda \setminus \Omega)$. In order to have a good estimate on this set, we need to impose it !

Definition 2.2.7 (Lebesgue measurability, Lebesgue measure)**.** *A set $\Omega \subset \mathbb{R}^n$ is* Lebesgue measurable *if for any $\varepsilon > 0$ there exists an open set Λ with $\Omega \subset \Lambda$ and such that*

$$m^*(\Lambda \setminus \Omega) \leq \varepsilon.$$

For any Lebesgue measurable set Ω, we define $m(\Omega) := m^*(\Omega)$ and call it the Lebesgue measure *of Ω.*

The following properties can now be proved:

1. If Ω is an open set, then Ω is Lebesgue measurable,
2. If $m^*(\Omega) = 0$, then Ω is Lebesgue measurable,
3. If $\Omega := \cup_j \Omega_j$ is a finite or countable union of Lebesgue measurable sets, then Ω is Lebesgue measurable, and $m(\Omega) \leq \sum_j m(\Omega_j)$,
4. If $\Omega := \cap_j \Omega_j$ is a finite or countable intersection of Lebesgue measurable sets, then Ω is Lebesgue measurable,
5. Any closed set is Lebesgue measurable, in particular any closed box I is Lebesgue measurable, with $m(I) = v(I)$.

Exercise 2.2.8. *Prove some of the above statements. Inspiration can be found in [Ne, Sec. 1.2].*

With Lebesgue measurable sets, the unpleasant feature mentioned before can not take place. Recall that a family of sets $\{\Omega_j\}_j$ are pairwise disjoint if $\Omega_j \cap \Omega_k = \emptyset$ for any $j \neq k$.

Theorem 2.2.9. *Let $\{\Omega_j\}$ be a finite or countable family of pairwise disjoint subsets of $\mathbb{R}^n$ which are Lebesgue measurable. Then*

$$m(\cup_j \Omega_j) = \sum_j m(\Omega_j).$$

Another useful result for Lebesgue measurable sets is about the complement, as shown in [Ne, 1.2.18]. We shall write Ω^{C} for the complement of the set Ω, namely

$$\Omega^{\mathrm{C}} := \{X \in \mathbb{R}^n \mid X \notin \Omega\}.$$

Theorem 2.2.10. *If Ω is a Lebesgue measurable set, then its complement Ω^{C} is also a Lebesgue measurable set.*

Let us close this section by saying that sets which are not Lebesgue measurable exist, but they are not easy to exhibit. For example, on $\mathbb{R}$ let us define the equivalence relation: for $x, y \in \mathbb{R}$, we set $x \sim y$ if $x - y \in \mathbb{Q}$, and write $[x] := \{y \in \mathbb{R} \mid y \sim x\}$. Consider then the set $\Omega \subset \mathbb{R}$ obtained by choosing exactly one representative for each equivalence class (the axiom of choice is involved in this process). Then Ω is not Lebesgue measurable. The proof of this statement is not easy, and we refer to [Ne, Sec. 1.3] for additional information on non Lebesgue measurable sets.

2.3 Lebesgue measurable functions

Even if Riemann integral can deal with non continuous functions, it was mainly developed for continuous functions. Lebesgue integral deals with more general functions, and gives the same value whenever a function is Riemann integrable. In fact, the main difference is going to be the definition of the partition of the interval $[a, b]$.

Definition 2.3.1 (Lebesgue measurable function)**.** *A function $f : [a, b] \to \mathbb{R}$ is a* Lebesgue measurable function*, or simply is* Lebesgue measurable*, if for any $s \in \mathbb{R}$, the sets*

$$\{x \in [a, b] \mid f(x) > s\} \tag{2.3.1}$$

are Lebesgue measurable sets.

There should be no confusion between the notion of Lebesgue measurable sets and the notion of Lebesgue measurable functions. In fact, these two concepts are very closely related, as shown in the following exercise. For its statement, let us define a *characteristic function* on a set $\Omega \subset \mathbb{R}$ and for any $x \in \mathbb{R}$ by

$$\chi_\Omega(x) := \begin{cases} 1 & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega \end{cases}. \tag{2.3.2}$$

Clearly, the function defined in Exercise 2.1.3 is a characteristic function. Characteristic functions can also be defined on $\mathbb{R}^n$, with a straightforward definition.

Exercise 2.3.2. *Show that the function χ_Ω is a Lebesgue measurable function if and only if the set Ω is a Lebesgue measurable set.*

The following generalization is going to play an important role in the sequel. It is left as an exercise to show that these functions are Lebesgue measurable functions.

Definition 2.3.3 (Simple functions)**.** *Let $\{\Omega_j\}_{j=1}^n$ be a family of pairwise disjoint measurable subsets of $[a, b]$ covering $[a, b]$, and let $\{c_j\}_{j=1}^n$ be a family of real numbers. A* simple function *on $[a, b]$ is defined for any $x \in [a, b]$ by*

$$\varphi(x) := \sum_{j=1}^{n} c_j \, \chi_{\Omega_j}(x).$$

Let us now mention that the strict inequality used in (2.3.1) is not so strict, as shown in the following exercise:

Exercise 2.3.4. *Let f be defined on $[a, b]$. Show that the following statements are equivalent:*

1. *The function f is Lebesgue measurable,*
2. *For any $s \in \mathbb{R}$ the sets $\{x \in [a, b] \mid f(x) \geq s\}$ are Lebesgue measurable sets,*
3. *For any $s \in \mathbb{R}$ the sets $\{x \in [a, b] \mid f(x) \leq s\}$ are Lebesgue measurable sets,*
4. *For any $s \in \mathbb{R}$ the sets $\{x \in [a, b] \mid f(x) < s\}$ are Lebesgue measurable sets,*

For the proof, Theorem 2.2.10 can be useful.

In the following statement, we gather some properties of the set of Lebesgue measurable functions. Note that the proof is not so straightforward, as for continuous functions for example. Nevertheless, the proof can be worked out as a exercise, see also [Ne, Thm. 2.1.5 & 2.1.6].

Proposition 2.3.5. *Let f, g be Lebesgue measurable functions on $[a, b]$ and let $\lambda \in \mathbb{R}$.*

1. *The sum $f + \lambda g$ is a Lebesgue measurable function,*
2. *The product fg is a Lebesgue measurable function,*
3. *If $g(x) \neq 0$ for all $x \in [a, b]$, then f/g is a Lebesgue measurable function.*

Coming back to the Riemann integral, it is not difficult to observe that if $f : [a, b] \to \mathbb{R}$ is Riemann integrable, changing the value of f on a finite number of points of $[a, b]$ does not change the value of its Riemann integral. More precisely, if $g : [a, b] \to \mathbb{R}$ satisfies $f(x) = g(x)$ for all $x \in [a, b]$ except on a finite number of $x \in [a, b]$, then g is Riemann integrable, and $\int_a^b g(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x$. Clearly, one can not do better than this, otherwise the function presented in (2.1.2) would be Riemann integrable. For Lebesgue integral, more flexibility will be available.

Definition 2.3.6 (Almost everywhere)**.** *Consider $f, g : [a, b] \to \mathbb{R}$.*

1. *We write $f = g$ a.e. if the set $\{x \in [a, b] \mid f(x) \neq g(x)\}$ has Lebesgue measure 0,*
2. *We write $f \leq g$ a.e. if the set $\{x \in [a, b] \mid f(x) > g(x)\}$ has Lebesgue measure 0.*

In both cases, we say that the relation holds almost everywhere*.*

Note that one can define similarly $f < g$ a.e., $f \geq g$ a.e., and $f > g$ a.e. The next statement reveals the importance of this concept, and already provides a glimpse about the generality we are dealing with.

Proposition 2.3.7. *Let $f : [a, b] \to \mathbb{R}$ be a Lebesgue measurable function, and let $g = f$ a.e. Then g is also a Lebesgue measurable function on $[a, b]$.*

Exercise 2.3.8. *Prove the previous statement, see also [Ne, Prop. 2.1.9].*

Let us now define the notions of lim sup and lim inf, which appear in several contexts. The very nice feature of these limits is that they always exist, which is not the case for the ordinary limit.

Definition 2.3.9 (lim sup and lim inf)**.** *Consider a sequence of functions $(f_j)_{j\in\mathbb{N}}$ with $f_j : \Omega \to \mathbb{R}$ where Ω is an arbitrary subset of $\mathbb{R}$. We assume that the sequence is* pointwise bounded*, meaning that $\sup_{j\in\mathbb{N}}|f_j(x)| < \infty$ for any $x \in \Omega$ (but this supremum can be x-dependent). For any $x \in \Omega$, we set*

$$f^*(x) \equiv \limsup_{j\to\infty} f_j(x) := \lim_{j\to\infty} \big\{ \sup_{k\geq j} f_k(x) \big\}$$

and call the corresponding function $f^* : \Omega \to \mathbb{R}$ the lim sup *function of the sequence. We also set*

$$f_*(x) \equiv \liminf_{j\to\infty} f_j(x) := \lim_{j\to\infty} \big\{ \inf_{k\geq j} f_k(x) \big\}$$

and call the corresponding function $f_* : \Omega \to \mathbb{R}$ the lim inf *function of the sequence.*

Let us firstly check that these definitions are meaningful:

Exercise 2.3.10. *Show that $\limsup_{j\to\infty} f_j(x)$ and $\liminf_{j\to\infty} f_j(x)$ always exist, and observe that the following inequality always hold:*

$$f_*(x) \leq f^*(x).$$

Show also that $f_*(x) = f^*(x)$ if and only if $\lim_{j\to\infty} f_j(x)$ exists.

A slightly more involved (and very useful) result can also be proved:

Theorem 2.3.11. *Let $f_j : [a,b] \to \mathbb{R}$ be Lebesgue measurable functions, and assume that $(f_j)_{j\in\mathbb{N}}$ is a pointwise bounded sequence. Then the two functions f^* and f_* are also Lebesgue measurable on $[a,b]$.*

As an immediate consequence of this result and of the content of the previous exercise, one has:

Corollary 2.3.12. *Let $f_j : [a,b] \to \mathbb{R}$ be Lebesgue measurable functions, and assume that $(f_j)_{j\in\mathbb{N}}$ is a pointwise bounded sequence. Assume that for each $x \in [a,b]$, the functions f_j have a limit as $j \to \infty$, namely $\exists f : [a,b] \to \mathbb{R}$ such that $\lim_{j\to\infty} f_j(x) = f(x)$ for all $x \in [a,b]$. Then the function f is a Lebesgue measurable function on $[a,b]$.*

2.4 Lebesgue integral

We can now define the notion of Lebesgue integral, mimicking the process outlined for Riemann integral. The link with Lebesgue measurable functions will then be established soon.

Definition 2.4.1 (Lebesgue measurable partition)**.** *A* Lebesgue measurable partition $\mathcal{P}$ *of* $[a,b]$ *consists in a finite collection* $\{\Omega_j\}_{j=1}^n \subset [a,b]$ *satisfying*

1. *For each j, the subset Ω_j of $[a,b]$ is Lebesgue measurable,*
2. $\cup_{j=1}^n \Omega_j = [a,b]$*,*
3. $m(\Omega_j \cap \Omega_k) = 0$ *for any $j \neq k$.*

Observe that the third condition does not imply $\Omega_j \cap \Omega_k = \emptyset$, but it means that if this intersection is not empty, then it is of measure 0. Clearly, the partitions introduced in (2.1.1) define Lebesgue measurable partitions by setting $\Omega_j := [x_{j-1}, x_j]$. However, such partitions are quite special in the set of all Lebesgue measurable partitions.

Given a Lebesgue measurable partition $\mathcal{P}$ of $[a,b]$ and a function $f \in \mathcal{L}^\infty([a,b])$ we set

$$L(f,\mathcal{P}) := \sum_{j=1}^n \Big(\inf_{x\in\Omega_j} f(x)\Big) m(\Omega_j)$$

and

$$U(f,\mathcal{P}) := \sum_{j=1}^n \Big(\sup_{x\in\Omega_j} f(x)\Big) m(\Omega_j).$$

Then, by analogy to the construction of Section 2.1, one defines:

Definition 2.4.2 (Lebesgue integral function)**.** *A function $f \in \mathcal{L}^\infty([a,b])$ is* Lebesgue integrable *on $[a,b]$ if*

$$\sup_{\mathcal{P}} L(f,\mathcal{P}) = \inf_{\mathcal{P}} U(f,\mathcal{P}), \tag{2.4.1}$$

where the supremum and the infimum are taken over all Lebesgue measurable partitions. If the equality holds, then we write $\int_a^b f(x)\,\mathrm{d}x$ for this value.

Since the set of partitions introduced in (2.1.1) is included in the set of all Lebesgue measurable partitions, one directly infers from this definition that Riemann integrable functions are Lebesgue integrable. Indeed, if the equality is already realized with a subset of partitions, then it is also realized with more partitions. In addition, one easily realizes that the equality (2.4.1) has a better chance to exist by considering a large set of partitions instead of a restricted set of partitions. As a consequence, there exist more Lebesgue integrable functions than Riemann integrable functions.

Let us now state the link between Lebesgue integrable functions and Lebesgue measurable functions. Clearly, the concepts are different, but one has:

Theorem 2.4.3. *For any $f \in \mathcal{L}^\infty([a,b])$, f is a Lebesgue measurable function if and only if f is a Lebesgue integrable function.*

We provide below half of the proof, since it provides a key difference between Lebesgue integrals and Riemann integrals. For the second half of the proof, we refer to [Ne, Thm. 2.2.13], which can be worked out as an exercise.

Proof. We show that if f is bounded and Lebesgue measurable on $[a,b]$, then f is also Lebesgue integrable. First of all, observe that since f is bounded, there exists $M > 0$ such that $|f(x)| < M$ for all $x \in [a,b]$. Let us then consider a n-partition of $[-M, M]$, in the sense of (2.1.1), namely $\{y_0, y_1, \dots, y_n\}$ with $y_0 = -M$, $y_n = M$ and $y_{j-1} < y_j$ for $j \in \{1, 2, \dots, n\}$.

For any fixed ε, we can also choose n and $\{y_j\}_{j=0}^n$ such that $y_j - y_{j-1} < \frac{\varepsilon}{b-a}$. For $j \in \{1, \dots, n\}$, let us now set

$$\Omega_j := \{x \in [a,b] \mid y_{j-1} \le f(x) < y_j\} \equiv f^{-1}([y_{j-1}, y_j)).$$

Observe then that

$$\Omega_j = \{x \in [a,b] \mid f(x) < y_j\} \setminus \{x \in [a,b] \mid f(x) < y_{j-1}\},$$

and by Exercise 2.3.4, one infers that Ω_j is Lebesgue measurable. In addition, $\cup_{j=1}^n \Omega_j = [a,b]$, and $\Omega_j \cap \Omega_k = \emptyset$ for any $j \neq k$. As a consequence, the family $\{\Omega_j\}_{j=1}^n$ defines a Lebesgue measurable partition $\mathcal{P}$ in the sense of Definition 2.4.1, with the additional property that

$$\sum_{j=1}^n m(\Omega_j) = m\Big(\bigcup_{j=1}^n \Omega_j\Big) = m([a,b]) = b - a.$$

Let us finally compute

$$\begin{aligned}
U(f,\mathcal{P}) - L(f,\mathcal{P}) &\le \sum_{j=1}^n y_j m(\Omega_j) - \sum_{j=1}^n y_{j-1} m(\Omega_j) \\
&= \sum_{j=1}^n (y_j - y_{j-1}) m(\Omega_j) \\
&\le \frac{\varepsilon}{b-a} \sum_{j=1}^n m(\Omega_j) \\
&= \varepsilon.
\end{aligned}$$

One concludes from the previous inequality that f is Lebesgue integrable, see also Theorem 2.1.4. ◻

Clearly, the theory would not be satisfactory without the following statements. Proofs are left as possible exercises.

Lemma 2.4.4. *Let $f \in \mathcal{L}^\infty([a,b])$ with $f = 0$ a.e. Then f is Lebesgue integrable, and one has $\int_a^b f(x)\,\mathrm{d}x = 0$.*

Corollary 2.4.5. *Let $f, g \in \mathcal{L}^\infty([a,b])$ with $f = g$ a.e. and assume that f is Lebesgue integrable. Then g is also Lebesgue integrable, and $\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x$.*

Let us state a related result, which is in fact used for the second part of the proof of Theorem 2.4.3.

Lemma 2.4.6. *Let $f \in \mathcal{L}^\infty([a,b])$ be Lebesgue measurable, and assume that $f \geq 0$ a.e. on $[a,b]$ and that $\int_a^b f(x)\,\mathrm{d}x = 0$. Then $f = 0$ a.e. on $[a,b]$.*

We the information collected so far, it is now rather to complement Exercise 2.1.2, see also [Ne, Example 2.2.6 & 2.2.9].

Exercise 2.4.7. *Show that the function defined in (2.1.2) is Lebesgue integrable, while it has already been shown that it is not Riemann integrable. Compute its Lebesgue integral.*

So far, only bounded function have be considered for the Lebesgue integral. The operation can be extended to unbounded functions, and to unbounded domain, in a way similar to the improper Riemann integrals. We briefly sketch the processes.

Let $f : [a,b] \to \mathbb{R}$ with $f(x) \geq 0$ for all $x \in [a,b]$ and assume that $f \notin \mathcal{L}^\infty([a,b])$. For any $N \in \mathbb{N}$ we set $f_N := f(x)$ if $f(x) \leq N$ while $f_N(x) := 0$ if $f(x) > N$. We say that f is Lebesgue integrable on $[a,b]$ if $f_N \in \mathcal{L}^\infty([a,b])$ is Lebesgue integrable for all $N \in \mathbb{N}$ and if $\lim_{N\to\infty}\big(\int_a^b f_N(x)\,\mathrm{d}x\big) < \infty$. In this case, we simply write

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{N\to\infty}\Big(\int_a^b f_N(x)\,\mathrm{d}x\Big).$$

If f is not positive, we firstly set

$$f_+(x) := \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) < 0 \end{cases} \qquad \text{and} \qquad f_-(x) := \begin{cases} 0 & \text{if } f(x) \geq 0 \\ -f(x) & \text{if } f(x) < 0. \end{cases}$$

The function f_+ is called the positive part of f, while the function f_- is called the negative part of f. Clearly, we have $f = f_+ - f_-$. If both f_+ and f_- are Lebesgue integrable in the sense mentioned above, then we set

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f_+(x)\,\mathrm{d}x - \int_a^b f_-(x)\,\mathrm{d}x.$$

The set of all Lebesgue integral functions on $[a,b]$ are denoted by $\mathcal{L}([a,b])$. These functions can be bounded or not bounded. Let us also mention that the various properties shown for bounded Lebesgue integrable functions extend to arbitrary elements of $\mathcal{L}([a,b])$.

Note that when f is unbounded from below and from above, this approach (separating the positive and the negative part) is the only possible one, otherwise some meaningless situation can easily take place.

Exercise 2.4.8. *Determine if the function $f : [-1,1] \to \mathbb{R}$ defined by $f(x) := 1/x$ if $x \neq 0$ and $f(0) := 0$, belongs to $\mathcal{L}([-1,1])$? Similarly, determine if the function $g : [0,1] \to \mathbb{R}$ defined by $g(x) := 1/\sqrt{x}$ if $x \neq 0$ and $g(0) = 0$ belongs to $\mathcal{L}([0,1])$? Determine also and with a full proof if these functions define improper Riemann integrals on these intervals, and compare the two proofs.*

Let us still state one useful result. For it, let us stress that the notion of Lebesgue measurable functions was introduced without requiring the boundedness of the functions. For the proof, we refer to [Ne, Lem. 2.4.5].

Proposition 2.4.9. *Let $g \in \mathcal{L}([a,b])$, and let f be a Lebesgue measurable function on $[a,b]$ satisfying $|f(x)| \leq g(x)$ for a.e. $x \in [a,b]$. Then $f \in \mathcal{L}([a,b])$.*

Exercise 2.4.10. *Exhibit a Lebesgue measurable function which is not a Lebesgue integrable function. Because of Theorem 2.4.3, note that such a function can not be bounded.*

So far, we have only integrated functions on intervals, for more general Lebesgue measurable sets $\Omega \subset [a, b]$ and for $f \in \mathcal{L}([a, b])$ we set

$$\int_\Omega f(x)\,\mathrm{d}x := \int_a^b [f\chi_\Omega](x)\,\mathrm{d}x$$

where χ_Ω is the characteristic function introduced in 2.3.2. This definition is based on the fact that the set of Lebesgue integrable functions is stable under multiplication.

In the context of the previous definition, a natural statement holds:

Lemma 2.4.11. *Let $f \in \mathcal{L}([a, b])$, and consider a family $\{\Omega_j\}_{j\in\mathbb{N}}$ with $\Omega_j \subset [a, b]$, Lebesgue measurable and satisfying $\Omega_j \subset \Omega_{j+1}$. Assume also that $\cup_j\Omega_j = [a, b]$. Then one has*

$$\lim_{j\to\infty} \int_{\Omega_j} f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

For unbounded domain of integration, the approach is similar to the improper Riemann integrals, but surprisingly the final result is different. More precisely, for $f : [a, \infty) \to \mathbb{R}$ with $f \geq 0$ a.e. we consider the product $f\chi_{[a,b]}$ for any $b > 0$ and we say that $f \in \mathcal{L}([a, \infty))$ if $f\chi_{[a,b]}$ belongs to $\mathcal{L}([a, b])$ for any $b > a$ and if

$$\lim_{b\to\infty} \Big(\int_a^b f(x)\,\mathrm{d}x \Big) < \infty.$$

If f is not positive, then the previous condition has to hold separately for the positive part f_+ of f, and for the negative part f_- of f, and one then sets

$$\int_a^\infty f(x)\,\mathrm{d}x = \lim_{b\to\infty} \Big(\int_a^b f_+(x)\,\mathrm{d}x \Big) - \lim_{b\to\infty} \Big(\int_a^b f_-(x)\,\mathrm{d}x \Big).$$

Clearly, a similar construction is valid for $\mathcal{L}((-\infty, b])$ and for $\mathcal{L}(\mathbb{R})$, if the two limits are considered separately in the latter case.

The following exercise provides an illustration that Lebesgue integrals on a half-line do not correspond to improper Riemann integrals.

Exercise 2.4.12. *Exhibit a function on $[1, \infty)$ with is not Lebesgue integrable but which is improper Riemann integrable.*

2.5 Dominated convergence theorem

In this section, we state one of the most useful result in analysis, the dominated convergence theorem. It allows the exchange of a limit and of an integral, under suitable conditions.

Let us start by introducing the notion of pointwise convergence. Let $(f_j)_{j\in\mathbb{N}}$ be a sequence of real functions on $\Omega \subset \mathbb{R}$.This sequence converges pointwise if for any $x \in \Omega$ the limit $\lim_{j\to\infty} f_j(x)$ converges. Note that this convergence is taking place in $\mathbb{R}$, and by writing it precisely with ε and N, the N will depend on ε and on x.

Examples 2.5.1. *1. The sequence $(f_j)_{j\in\mathbb{N}}$, with $f_j : [0,1] \to \mathbb{R}$ defined by $f_j(x) := x^j$, converges pointwise to the function f_∞ given by*

$$f_\infty(x) := \begin{cases} 0 & \text{if } x \in [0,1), \\ 1 & \text{if } x = 1. \end{cases}$$

2. The sequence $(f_j)_{j\in\mathbb{N}^*}$, with $f_j : [0,1] \to \mathbb{R}$ defined by $f_j := j\chi_{(0,1/j]}$, converges pointwise to the 0-function, namely $f_\infty = 0$.

Let us observe that in the first example, each f_j is a continuous function on $[0,1]$ while the limiting function f_∞ is no more continuous. For the second example, observe that for any $j \in \mathbb{N}^*$ one has $\int_0^1 f_j(x)\,\mathrm{d}x = 1$ while $\int_0^1 f_\infty(x)\,\mathrm{d}x = 0$. Thus, some operations are not continuous with respect to the pointwise convergence of functions. However, observe that in both examples, the functions f_j and the limit f_∞ are Lebesgue integrable. This fact plays a key role in the following statement, but note that an additional condition is necessary.

Theorem 2.5.2 (Dominated convergence theorem)**.** *Let $(f_j)_{j\in\mathbb{N}}$ be a sequence of Lebesgue measurable functions on $[a,b]$ such that $\lim_{j\to\infty} f_j(x) = f(x)$ for* a.e. $x \in [a,b]$. *Suppose also that there exists $g \in \mathcal{L}([a,b])$ with*

$$|f_j(x)| \leq g(x)$$

for a.e. $x \in [a,b]$ *and for every $j \in \mathbb{N}$. Then $f_j \in \mathcal{L}([a,b])$, $f \in \mathcal{L}([a,b])$, and*

$$\lim_{j\to\infty} \int_a^b f_j(x)\,\mathrm{d}x = \int_a^b \Big(\lim_{j\to\infty} f_j(x)\Big)\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Thus, this statement says that a limit over j and an integral can be exchanged, once the existence of a *dominating function g* is guaranted. Note that in the applications, finding the dominating function can be a hard work. The proof of this theorem is rather involved, but relies on the two results already stated, namely Proposition 2.4.9 and Lemma 2.4.11. We refer to [Ne, Thm. 2.4.6] for the proof.

Exercise 2.5.3. *Show that the dominated convergence theorem does not apply in the limits considered in Section 1.3. If it were applicable, then the Dirac delta distribution would be a Lebesgue integrable function !*

Exercise 2.5.4. *Study the examples of application of the dominated convergence theorem provided in Examples* 2.4.7, 2.4.8, *and* 2.4.9 *of [Ne]. Alternatively, exhibit other applications of this statement.*

Let us still state one important result which is in fact equivalent to the dominated convergence theorem.

Theorem 2.5.5 (Monotone convergence theorem)**.** *Let $(f_j)_{j\in\mathbb{N}}$ be a sequence of nonnegative functions belonging to $\mathcal{L}([a,b])$. Suppose that for* a.e. $x \in [a,b]$, *the sequence $(f_j(x))_{j\in\mathbb{N}}$ is increasing, and that $\lim_{j\to\infty} f_j(x) = f(x)$ for* a.e. $x \in [a,b]$. *Then,*

1. *If $f \in \mathcal{L}([a,b])$, then $\lim_{j\to\infty} \int_a^b f_j(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x$,*
2. *If $f \notin \mathcal{L}([a,b])$, then $\lim_{j\to\infty} \int_a^b f_j(x)\,\mathrm{d}x = \infty$.*

2.6 L^p-spaces

In the previous section, we have considered measurability or integrability of functions on $[a,b]$, or on $[a,\infty)$, $(-\infty,b]$, and $\mathbb{R}$. However, it is not difficult to check that these notions extend to more general bounded and closed subsets of $\mathbb{R}^n$ or to $\mathbb{R}^n$ itself. In fact, extending Lebesgue integrals to $\mathbb{R}^n$ is even easier than extending

Riemann integrals to $\mathbb{R}^n$, since the index of the partitions are simpler. The dominated convergence theorem or the monotone convergence theorem also extend to this more general setting.

In the rest of this chapter, we shall consider functions on Lebesgue measurable and bounded subsets of $\mathbb{R}^n$. For simplicity, we shall denote such sets by $\Omega \subset \mathbb{R}^n$, and keep in mind that Ω could be an open and bounded set in $\mathbb{R}^n$, or a closed and bounded set in $\mathbb{R}^n$. Extensions to unbounded subsets of $\mathbb{R}^n$ are also possible with a limiting procedure.

In this framework, let us recall a very useful result available for functions on $\mathbb{R}^n$: Fubini's theorem. We present it below for functions defined on a rectangle $I := I_1 \times I_2 \subset \mathbb{R}^2$, with I_j intervals, but its generalization to more complicated domains in $\mathbb{R}^n$ is rather straightforward.

Theorem 2.6.1 (Fubini's theorem)**.** *For any $f \in \mathcal{L}(I)$ with $I := I_1 \times I_2 \subset \mathbb{R}^2$, the following properties hold:*

1. *For almost every $x \in I_1$, the function $I_2 \ni y \mapsto f(x, y) \in \mathbb{R}$ is Lebesgue measurable and integrable on I_2, with integral denoted by $\int_{I_2} f(x, y)\,\mathrm{d}y$,*
2. *The function $I_1 \ni x \mapsto \int_{I_2} f(x, y)\,\mathrm{d}y \in \mathbb{R}$ is Lebesgue measurable and integrable on I_1, with integral denoted by $\int_{I_1} \big(\int_{I_2} f(x, y)\mathrm{d}y\big)\mathrm{d}x$,*
3. *The following equality holds*
$$\int_I f(x, y)\,\mathrm{d}x\mathrm{d}y = \int_{I_1} \Big(\int_{I_2} f(x, y)\mathrm{d}y\Big)\mathrm{d}x,$$
where the l.h.s. means the Lebesgue integral in $\mathbb{R}^2$, defined as in Definition 2.4.2 but for a function of two variables,
4. *By exchanging x and y in the previous statement, the following equality also holds*
$$\int_I f(x, y)\,\mathrm{d}x\mathrm{d}y = \int_{I_2} \Big(\int_{I_1} f(x, y)\mathrm{d}x\Big)\mathrm{d}y.$$

For the sake of generality, we shall also consider functions on $\Omega \subset \mathbb{R}^n$ with values in $\mathbb{K}$, and not only in $\mathbb{R}$. Recall that $\mathbb{K}$ means either $\mathbb{R}$ or $\mathbb{C}$. In the complex case, observe that any $f : \Omega \to \mathbb{C}$ can always be written as $f = f_r + i f_i \equiv \Re(f) + i\Im(f)$, where $f_r \equiv \Re(f) = \frac{f+\overline{f}}{2}$ corresponds to the real part of f, and where $f_i \equiv \Im(f) = \frac{f-\overline{f}}{2i}$ corresponds to the imaginary part of f. In this case, a complex valued function f is Lebesgue measurable if and only if f_r and f_i are Lebesgue measurable, and f belongs to $\mathcal{L}(\Omega)$ if and only if f_r and f_i belong to $\mathcal{L}(\Omega)$, as real valued (bounded or unbounded) functions. The theory developed so far applies thus to f_r and to f_i, and together, these two functions define the complex valued function f. In particular, if f is Lebesgue integrable, one has
$$\int_\Omega f(X)\,\mathrm{d}X = \int_\Omega f_r(X)\,\mathrm{d}X + i\int_\Omega f_i(X)\,\mathrm{d}X.$$

Let us now start the construction of L^p-spaces, starting with $L^1(\Omega)$ for $\Omega \subset \mathbb{R}^n$ as mentioned above. Recall that if $f = g$ a.e., then $\int_\Omega f(X)\,\mathrm{d}X = \int_\Omega g(X)\,\mathrm{d}X$, see Corollary 2.4.5 when $\Omega = [a, b]$. For that reason, we would like to put such functions in an equivalence class: For any $f, g \in \mathcal{L}(\Omega)$, we write $f \sim g$ whenever $f = g$ a.e. The property of this relation is summarized in the following exercise.

Exercise 2.6.2. *Prove that the relation $\sim$ defines an* equivalence relation*, namely the following three properties are satisfied for any $f, g, h \in \mathcal{L}(\Omega)$:*

1. $f \sim f$ (reflexivity),
2. *If* $f \sim g$ *then* $g \sim f$ (symmetry),
3. *If* $f \sim g$ *and* $g \sim h$*, then* $f \sim h$ (transitivity).

Before the next definition, recall that a *norm* on a complex vector space Ξ corresponds to a map $\Xi \ni \xi \mapsto \|\xi\| \in [0, \infty)$ satisfying

1. $\|\lambda\xi\| = |\lambda| \, \|\xi\|$ for any $\lambda \in \mathbb{C}, \xi \in \Xi$,
2. $\|\xi_1 + \xi_2\| \leq \|\xi_1\| + \|\xi_2\|$ for any $\xi_1, \xi_2 \in \Xi$,
3. $\|\xi\| = 0$ if and only if $\xi = 0$.

Definition 2.6.3 (L^1-space)**.** *The set $L^1(\Omega)$ is defined by $\mathcal{L}(\Omega)/\sim$, namely the elements of $L^1(\Omega)$ consist of equivalence classes of Lebesgue integrable functions which are equal almost everywhere. The set $L^1(\Omega)$ is endowed with the norm*

$$\|f\|_1 \equiv \|f\|_{L^1(\Omega)} = \int_\Omega |f(X)| \mathrm{d}X,$$

where $|f(X)|$ denotes the absolute value of $f(X)$ if f is real valued, or the modulus of $|f(X)|$ if the function f is complex valued.

Observe that considering equivalence classes was necessary, otherwise the condition $\|f\|_1 = 0$ would not imply that $f = 0$, it would simply imply that $f = 0$ a.e. Whenever one deals with equivalence class, the notation $[f]$ should be preferred, meaning that we consider the set of all functions which are equivalent to f. However, this notation is not used, and one keeps writing f, g for the elements of $L^1(\Omega)$. Despite this notation, we should keep in mind that these elements are equivalence classes, which means that $f \in L^1(\Omega)$ is not defined everywhere but only almost everywhere. Changing f on a set of Lebesgue measure 0 does not affect f, we still get the same element of $L^1(\Omega)$.

Exercise 2.6.4. *Show that the map $f \mapsto \|f\|_1$ defines a norm on $L^1(\Omega)$.*

One very important property of $L^1(\Omega)$ is called the *completeness*. Completeness means that any *Cauchy sequence*[1] converges in $L^1(\Omega)$. We do not prove this statement, since it is not an easy proof, but refer to [Ne, Thm. 3.1.13]. Note that the standard example of complete space is $\mathbb{R}$, while $\mathbb{Q}$ is not complete. For these two examples, the norm considered is simply the absolute value. The set $\mathbb{C}$ of complex numbers is also complete, when endowed with the norm defined by the modulus. A vector space endowed with a norm and complete with this norm is called a *Banach space*. It means that $L^1(\Omega)$, or $\mathbb{R}$ or $\mathbb{C}$ are Banach spaces.

Remark 2.6.5. *Having defined $L^1(\Omega)$, one can now give a precise definition to $L^1_{loc}(\mathbb{R}^n)$: It consists in the set of all functions $f : \mathbb{R}^n \to \mathbb{K}$ such that for any bounded and Lebesgue measurable set Ω, the function $f\chi_\Omega$ belongs to $L^1(\Omega)$. It means that locally, the functions f look like element of $L^1(\Omega)$, and therefore are Lebesgue integrable on Ω. However, globally they don't have to be integrable.*

In the next exercise, we emphasize the difference between a pointwise convergence and the convergence in the L^1-norm.

[1] A Cauchy sequence in $L^1(\Omega)$ is a sequence $(f_j)_{j\in\mathbb{N}} \subset L^1(\Omega)$ satisfying the condition: for any $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|f_j - f_{j'}\|_1 \leq \epsilon$ for all $j, j' \geq N$. Note that any convergent sequence is Cauchy, but the converse is not true.

Exercise 2.6.6. *For $x \in [0,1]$ consider the functions*

$$f_j(x) := \begin{cases} 2j^2 x & \text{if } 0 \le x \le \frac{1}{2j} \\ -2j^2(x - \frac{1}{j}) & \text{if } \frac{1}{2j} < x \le \frac{1}{j} \\ 0 & \text{otherwise.} \end{cases}$$

Check that this sequence $(f_j)_{j \in \mathbb{N}^*}$ converges pointwise to the function $f_\infty = 0$, but that $\lim_j \to \infty \|f_j - f_\infty\|_1 = \frac{1}{2}$. Thus, the sequence $(f_j)_{j\in\mathbb{N}^*}$ does not converge to f_∞ in the L^1-norm. In fact, observe that $\|f_j\|_1 = \frac{1}{2}$ for any $j \in \mathbb{N}^*$

Let us now generalize the previous construction, and replace 1 by any $p \ge 1$. Comments will be provided after the definition.

Definition 2.6.7 (L^p-space)**.** *For any $p \ge 1$ the set $L^p(\Omega)$ is defined as*

$$\Big\{ f : \Omega \to \mathbb{K} \mid f \text{ is Lebesgue measurable, and } |f|^p \in \mathcal{L}(\Omega) \Big\} \Big/ \sim$$

endowed with the norm

$$\|f\|_p := \Big(\int_\Omega |f(X)|^p \,\mathrm{d}X \Big)^{\frac{1}{p}}.$$

As for $L^1(\Omega)$, we do not consider single functions but equivalence class of functions which are equal almost everywhere. Functions are taking values in $\mathbb{K}$, and in this case the measurability condition has to hold for its real part and for its imaginary part. The last condition is a integrability condition for $|f|^p$ with is a positive valued function, and therefore this condition reads $\int_\Omega |f(X)|^p \,\mathrm{d}X < \infty$.

Exercise 2.6.8. *Show that the set $L^p(\Omega)$ is a complex vector space.*

Proving that the map $f \mapsto \|f\|_p$ is a norm is not an easy task. Clearly, only the condition $\|f+g\|_p \le \|f\|_p + \|g\|_p$ is really involved. It is based on one technical lemma and on one very useful inequality.

Lemma 2.6.9. *Let $\alpha, \beta \in (0,1)$ with $\alpha + \beta = 1$. Then for any $a, b \ge 0$ one has*

$$ab \le \alpha a^{\frac{1}{\alpha}} + \beta b^{\frac{1}{\beta}}.$$

Theorem 2.6.10 (Hölder inequality)**.** *Let $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, and consider $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$. Then the product fg belongs to $L^1(\Omega)$ and the following inequality holds*

$$\|fg\|_1 \le \|f\|_p \, \|g\|_q. \tag{2.6.1}$$

Based on these two results, the following statement can be proved. Observe that it corresponds to the missing argument for showing that $\|\cdot\|_p$ is indeed a norm on $L^p(\Omega)$.

Theorem 2.6.11 (Minkowski's inequality)**.** *For $p \ge 1$ and for any $f, g \in L^p(\Omega)$ one has*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p.$$

Exercise 2.6.12. *Prove some of the previous statements, getting some inspiration from [Ne, Lem. 3.2.6], [Ne, Thm. 3.2.5], and [Ne, Thm. 3.2.7].*

As for $L^1(\Omega)$, it turns out that $L^p(\Omega)$ are Banach spaces. In addition, $L^2(\Omega)$ has an additional property: it has a *scalar product*. This additional property comes from the following observation: One solution of the identity $\frac{1}{p} + \frac{1}{q} = 1$ is $(p, q) = (2, 2)$, which leads by (2.6.1) to $\|fg\|_1 \leq \|f\|_2 \, \|g\|_2$. We shall come back to L^2-spaces in the next chapter.

Let us conclude this section with one more space: $L^\infty(\Omega)$. In order to define it, we need to introduce one more notion.

Definition 2.6.13 (Essential sup and essential inf)**.** *Let $f : \Omega \to \mathbb{R}$ be a measurable function. We define the* essential supremum *of f by*

$$\operatorname{ess\,sup} f := \inf\{M \mid f \leq M \ \text{ a.e.}\}$$

and the essential infimum *of f by*

$$\operatorname{ess\,inf} f := \sup\{m \mid f \geq m \ \text{ a.e.}\}.$$

Clearly, these two quantities correspond to the supremum and to the infimum of f, up to a set of Lebesgue measure 0. Based on these notions, we can define:

Definition 2.6.14 (L^∞-space)**.** *The space $L^\infty(\Omega)$ consists in*

$$\Big\{f : \Omega \to \mathbb{K} \mid f \text{ is Lebesgue measurable, and } \operatorname{ess\,sup}|f| < \infty\Big\} \Big/ \sim$$

endowed with the norm

$$\|f\|_\infty := \operatorname{ess\,sup}|f|.$$

In this case, proving that the map $f \mapsto \|p\|_\infty$ and proving that $L^\infty(\Omega)$ is a Banach space, are much simpler than in the case $p \geq 1$. The former property can be proved as an exercise.

2.7 Approximation in L^p-spaces

Approximating real numbers by fractional numbers is a quite natural operation. Indeed, working with all digits of π is probably not possible, and an approximation by a fraction, as $\frac{314}{100}$, is much more convenient. This is possible because the set $\mathbb{Q}$ is dense in $\mathbb{R}$. As a consequence, the approximation can be as accurate as necessary.

For L^p-spaces the same idea applies. Most of the elements of $L^p(\Omega)$ are not continuous and can be quite irregular. In addition, keep in mind that an element of $L^p(\Omega)$ is defined only almost everywhere: it can be modified on a set of Lebesgue measure 0. Thus, approximating an arbitrary element of $L^p(\Omega)$ by a more regular element is of interest. This is the aim of this section.

The first step in the construction is to approximate any element of L^p by a bounded element.

Lemma 2.7.1. *Let $p \geq 1$, let Ω be Lebesgue measurable and bounded, and consider $f \in L^p(\Omega)$. For any $\varepsilon > 0$ there exists $f_\varepsilon \in L^p(\Omega)$ with f_ε bounded, such that $\|f - f_\varepsilon\|_p \leq \varepsilon$.*

Proof. Clearly, we can assume that $f \notin \mathcal{L}^\infty(\Omega)$, otherwise there is nothing to prove. For any $N \in \mathbb{N}$ and $X \in \Omega$ let us set

$$f_N(X) := \begin{cases} -N & \text{if } f(X) < -N \\ f(X) & \text{if } |f(X)| \leq N \\ N & \text{if } f(X) > N. \end{cases}$$

Clearly, f_N is a bounded function. In addition, for almost every $X \in \Omega$ one has $\lim_{N\to\infty} f_N(X) = f(X)$, or equivalently $\lim_{N\to\infty} |f_N(X) - f(X)| = 0$. This implies that $\lim_{N\to\infty} |f_N(X) - f(X)|^p = 0$, for almost every $X \in \Omega$. In addition, let us observe that

$$|f_N(X) - f(X)|^p \leq (|f_N(X)| + |f(X)|)^p \leq (2|f(X)|)^p = 2^p |f(X)|^p.$$

Since $|f(\cdot)|^p$ belongs to $L^1(\Omega)$, we can use it as a dominating function for the dominated convergence theorem, from which we infer that

$$\lim_{N\to\infty} \int_\Omega |f_N(X) - f(X)|^p \,\mathrm{d}X = \int_\Omega \lim_{N\to\infty} |f_N(X) - f(X)|^p \,\mathrm{d}X = \int_\Omega 0 \,\mathrm{d}X = 0,$$

or equivalently $\lim_{N\to\infty} \|f_N - f\|_p = 0$. Thus, given $\varepsilon > 0$ we can determine $N \in \mathbb{N}$ such that $\|f_N - f\|_p \leq \varepsilon$. By setting $f_\varepsilon := f_N$ for this N, we infer that f_ε is bounded and that $\|f - f_\varepsilon\|_p \leq \varepsilon$. □

Recall that simple functions have been introduced in Definition 2.3.3, and observe that there is no problem for extending this definition to more general Lebesgue measurable and bounded sets $\Omega \subset \mathbb{R}^n$. Then, our next aim is to approximate bounded L^p-functions by simple functions.

Lemma 2.7.2. *Let $p \geq 1$, let Ω be a Lebesgue measurable and bounded subset of $\mathbb{R}^n$, and consider $f \in L^p(\Omega)$ with f bounded. For any $\varepsilon > 0$ there exists a simple function φ such that $\|f - \varphi\|_p \leq \varepsilon$.*

Observe that the statement is trivial if the Lebesgue measure of Ω is 0. Therefore, we shall assume that $m(\Omega) > 0$ in the following proof.

Proof. The proof is quite similar to the proof of Theorem 2.4.3. Since f is bounded, there exists $M > 0$ such that $|f(X)| < M$ for all $X \in \Omega$. Let us then consider a n-partition of $[-M, M]$, namely $\{y_0, y_1, \dots, y_n\}$ with $y_0 = -M$, $y_n = M$ and $y_{j-1} < y_j$ for $j \in \{1, 2, \dots, n\}$. For any fixed ε, we can also choose n and $\{y_j\}_{j=0}^n$ such that $y_j - y_{j-1} < \frac{\varepsilon}{V^{\frac{1}{p}}}$, where $V := m(\Omega)$, the Lebesgue measure of the set Ω.

For $j \in \{1, \dots, n\}$, let us now set

$$\Omega_j := \{X \in \Omega \mid y_{j-1} \leq f(X) < y_j\} \equiv f^{-1}([y_{j-1}, y_j)).$$

Observe then that $\cup_{j=1}^n \Omega_j = \Omega$, and $\Omega_j \cap \Omega_k = \emptyset$ for any $j \neq k$. In addition,

$$\Omega_j = \{X \in \Omega \mid f(X) < y_j\} \setminus \{X \in \Omega \mid f(X) < y_{j-1}\},$$

which implies that Ω_j is Lebesgue measurable. We can then set

$$\varphi := \sum_{j=1}^n y_{j-1} \chi_{\Omega_j}.$$

Clearly, φ is a simple function, and one has $\varphi \leq f$. Moreover, observe that

$$|f(X) - \varphi(X)| \leq \frac{\varepsilon}{V^{\frac{1}{p}}} \qquad \text{for any } X \in \Omega_j.$$

It only remains to compute

$$\|f - \varphi\|_p = \Big(\int_\Omega \big|f(X) - \varphi(X)\big|^p \,\mathrm{d}X \Big)^{\frac{1}{p}}$$

$$
\begin{aligned}
&= \Big(\sum_{j=1}^{n}\int_{\Omega_j}\big|f(X)-\varphi(X)\big|^p\,\mathrm{d}X\Big)^{\frac{1}{p}}\\
&\le \Big(\sum_{j=1}^{n}\int_{\Omega_j}\Big|\frac{\varepsilon}{V^{\frac{1}{p}}}\Big|^p\,\mathrm{d}X\Big)^{\frac{1}{p}}\\
&= \frac{\varepsilon}{V^{\frac{1}{p}}}\Big(\int_{\Omega}1\,\mathrm{d}X\Big)^{\frac{1}{p}}\\
&= \varepsilon,
\end{aligned}
$$

since $\int_\Omega 1\,\mathrm{d}X = m(\Omega) = V$. This concludes the proof. □

The next step in our construction is to approximate simple functions by continuous functions. Let us emphasize that the next statement is not correct for $p = \infty$. Anyway, when writing $p \ge 1$, the special case $p = \infty$ is not included.

Lemma 2.7.3. *Let $p \ge 1$ and consider a simple function φ defined on the Lebesgue measurable and bounded set $\Omega \subset \mathbb{R}^n$. Then, for any $\varepsilon > 0$, there exists a continuous and bounded function g_ε defined on Ω such that $\|\varphi - g_\varepsilon\|_p \le \varepsilon$.*

We do not provide the proof of this statement, but refer to [Ne, Lem. 3.3.3 & Corol. 3.3.4]. Note that the proof consists by considering firstly the characteristic function on a Lebesgue measurable set, and then by extending the result to arbitrary simple functions.

Exercise 2.7.4. *Prove the previous lemma about the approximation of simple functions by continuous functions.*

Finally, by taking the previous three statements into account, we can easily deduce the main result of this section.

Theorem 2.7.5. *Let $p \ge 1$, let $\Omega \subset \mathbb{R}^n$ be Lebesgue measurable and bounded and consider $f \in L^p(\Omega)$. For any $\varepsilon > 0$ there exists a continuous and bounded function $g_\varepsilon \in L^p(\Omega)$ such that $\|f - g_\varepsilon\|_p \le \varepsilon$.*

The proof is left as an exercise, and consists in an $\varepsilon/3$ argument. Let us mention that the previous statement can be seen as a density result: it says that the set of continuous functions in dense in $L^p(\Omega)$ for the norm $\|\cdot\|_p$.

Let us finally mention that these results extend to unbounded set Ω. For example, it can be shown that any element of $L^p(\mathbb{R}^n)$ can be approximated by a continuous function with bounded support. It even turns out that the set of test functions $\mathcal{D}(\mathbb{R}^n)$ or the of Schwartz functions $\mathcal{S}(\mathbb{R}^n)$ are also dense in $L^p(\mathbb{R}^n)$. However, be aware that these results are not correct for $L^\infty(\mathbb{R}^n)$.

Chapter 3

Operator theory on Hilbert spaces

This chapter is mainly based on the first two chapters of the book [Am]. Its content is quite standard and can be found in several reference books.

3.1 Hilbert space

Definition 3.1.1 (Hilbert space)**.** A (complex) Hilbert space $\mathcal{H}$ *is a vector space on* $\mathbb{C}$ *with a strictly positive scalar product (or inner product) which is complete for the associated norm*[1] *and which admits a countable orthonormal basis. The scalar product is denoted by* $\langle\cdot,\cdot\rangle$ *and the corresponding norm by* $\|\cdot\|$.

In particular, note that for any $f, g, h \in \mathcal{H}$ and $\lambda \in \mathbb{C}$ the following properties hold:

1. $\langle f, g\rangle = \overline{\langle g, f\rangle}$,
2. $\langle f, g + \lambda h\rangle = \langle f, g\rangle + \lambda\langle f, h\rangle$, but $\langle f + \lambda g, h\rangle = \langle f, h\rangle + \overline{\lambda}\langle g, h\rangle$,
3. $\|f\|^2 = \langle f, f\rangle \geq 0$, and $\|f\| = 0$ if and only if $f = 0$.

Note that $\overline{\langle g, f\rangle}$ means the complex conjugate of $\langle g, f\rangle$. Note also that the linearity in the second argument in 2. is a matter of convention, some authors define the linearity in the first argument. In 3. the norm of f is defined in terms of the scalar product $\langle f, f\rangle$. We emphasize that the scalar product can also be defined in terms of the norm of $\mathcal{H}$, this is the content of the *polarisation identity*:

$$4\langle f, g\rangle = \|f + g\|^2 - \|f - g\|^2 - i\|f + ig\|^2 + i\|f - ig\|^2. \tag{3.1.1}$$

From now on, the symbol $\mathcal{H}$ will always denote a Hilbert space.

Examples 3.1.2.
1. *$\mathcal{H} = \mathbb{C}^n$ with $\langle\alpha,\beta\rangle = \sum_{j=1}^n \overline{\alpha_j}\beta_j$ for any $\alpha,\beta \in \mathbb{C}^n$,*
2. *$\mathcal{H} = \ell^2(\mathbb{Z}) := \{a = (a_j)_{j\in\mathbb{Z}} \subset \mathbb{C} \mid \sum_{j\in\mathbb{Z}} |a_j|^2 < \infty\}$, with $\langle a, b\rangle = \sum_{j\in\mathbb{Z}} \overline{a_j}b_j$ for any $a, b \in \ell^2(\mathbb{Z})$,*
3. *$\mathcal{H} = \ell^2(\mathbb{Z}^n) := \{a = (a_j)_{j\in\mathbb{Z}^n} \subset \mathbb{C} \mid \sum_{j\in\mathbb{Z}^n} |a_j|^2 < \infty\}$, with $\langle a, b\rangle = \sum_{j\in\mathbb{Z}^n} \overline{a_j}b_j$ for any $a, b \in \ell^2(\mathbb{Z}^n)$,*
4. *$\mathcal{H} = L^2(\mathbb{R}^n)$ with $\langle f, g\rangle = \int_{\mathbb{R}^n} \overline{f(X)}\,g(X)\,\mathrm{d}X$ for any $f, g \in L^2(\mathbb{R}^n)$.*

[1] Recall that $\mathcal{H}$ is said to be complete if any Cauchy sequence in $\mathcal{H}$ has a limit in $\mathcal{H}$. More precisely, $(f_j)_{j\in\mathbb{N}} \subset \mathcal{H}$ is a Cauchy sequence if for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|f_j - f_{j'}\| < \varepsilon$ for any $j, j' \geq N$. Then $\mathcal{H}$ is complete if for any such sequence there exists $f_\infty \in \mathcal{H}$ such that $\lim_{j\to\infty} \|f_j - f_\infty\| = 0$.

Note that all finite dimensional Hilbert spaces can be identified with $\mathbb{C}^n$ for some $n \in \mathbb{N}$. Thus, the example 1. corresponds to the most general finite dimensional Hilbert space. In 4., the space $L^2(\mathbb{R}^n)$ is a special instance of the space $L^p(\mathbb{R}^n)$ introduced in the previous chapter, namely

$$L^2(\mathbb{R}^n) := \Big\{ f : \mathbb{R}^n \to \mathbb{C} \mid f \text{ Lebesgue measurable and } \int_{\mathbb{R}^n} |f(X)|^2 \,\mathrm{d}X < \infty \Big\}.$$

Note however that we simply write $\|f\|$ and not more precise notation $\|f\|_2$. If $f, g \in L^2(\mathbb{R}^n)$, then the product fg is integrable because of Hölder inequality, see Theorem 2.6.10.

Remark 3.1.3. *For any $K \in \mathbb{R}^n$, the function $\mathbb{R}^n \ni X \mapsto \exp(-iK \cdot X) \in \mathbb{C}$ does not belong to $L^2(\mathbb{R}^n)$. On the other hand, if $\varphi \in \mathcal{D}(\mathbb{R}^n)$ or if $\varphi \in \mathcal{S}(\mathbb{R}^n)$, then the function*

$$\mathbb{R}^n \ni X \mapsto \int_{\mathbb{R}^n} \exp(-iK \cdot X)\varphi(K)\,\mathrm{d}K \in \mathbb{C}$$

belong to $L^2(\mathbb{R}^n)$.

Let us recall some useful inequalities valid in any Hilbert space: For any $f, g \in \mathcal{H}$ one has

$$|\langle f, g\rangle| \leq \|f\|\,\|g\| \qquad \text{Schwarz inequality,} \tag{3.1.2}$$

$$\|f + g\| \leq \|f\| + \|g\| \qquad \text{triangle inequality,} \tag{3.1.3}$$

$$\|f + g\|^2 \leq 2\|f\|^2 + 2\|g\|^2, \tag{3.1.4}$$

$$\big|\|f\| - \|g\|\big| \leq \|f - g\|. \tag{3.1.5}$$

Exercise 3.1.4. *Prove these inequalities. You can get some inspiration from [Am, p. 3-4].*

There exist two natural notions of convergences in $\mathcal{H}$.

Definition 3.1.5 (Strong and weak convergence)**.** *Let $(f_j)_{j\in\mathbb{N}}$ be a sequence in $\mathcal{H}$.*

1. *The sequence is* strongly convergent *to $f_\infty \in \mathcal{H}$ if $\lim_{j\to\infty} \|f_j - f_\infty\| = 0$. In this case, we write* $\mathrm{s-}\lim_{j\to\infty} f_j = f_\infty$,
2. *The sequence is* weakly convergent *to $f_\infty \in \mathcal{H}$ if for any $g \in \mathcal{H}$ one has $\lim_{j\to\infty}\langle g, f_j - f_\infty\rangle = 0$. In this case, we write* $\mathrm{w-}\lim_{j\to\infty} f_j = f_\infty$.

Exercise 3.1.6. *In the Hilbert space $\ell^2(\mathbb{N})$, consider the sequence $(f_j)_{j\in\mathbb{N}}$ given by $f_j(k) = \delta_{jk}$, meaning $f_j(k) = 1$ if $j = k$ and $f_j(k) = 0$ if $j \neq k$. Show that this sequence weakly converges to 0 but does not converge strongly. More generally, let $\{e_j\}_{j\in\mathbb{N}}$ be an orthonormal basis of an infinite dimensional Hilbert space. Show that* $\mathrm{w-}\lim_{j\to\infty} e_j = 0$, *but that* $\mathrm{s-}\lim_{j\to\infty} e_j$ *does not exist.*

Clearly, a strongly convergent sequence is also weakly convergent. The converse is not true. The next statement provides the exact link between these two notions of convergence.

Lemma 3.1.7. *Consider a sequence $(f_j)_{j\in\mathbb{N}} \subset \mathcal{H}$. One has*

$$\mathrm{s-}\lim_{j\to\infty} f_j = f_\infty \quad\Longleftrightarrow\quad \mathrm{w-}\lim_{j\to\infty} f_j = f_\infty \ \text{ and } \ \lim_{j\to\infty} \|f_j\| = \|f_\infty\|.$$

Proof. Assume first that $\text{s-}\lim_{j\to\infty} f_j = f_\infty$. By Schwarz inequality one infers that for any $g \in \mathcal{H}$:

$$|\langle g, f_j - f_\infty\rangle| \leq \|f_j - f_\infty\|\|g\| \to 0 \qquad \text{as } j \to \infty,$$

which means that $\text{w-}\lim_{j\to\infty} f_j = f_\infty$. In addition, by (3.1.5) one also gets

$$\big|\|f_j\| - \|f_\infty\|\big| \leq \|f_j - f_\infty\| \to 0 \quad \text{as} \quad j \to \infty,$$

and thus $\lim_{j\to\infty} \|f_j\| = \|f_\infty\|$.

For the reverse implication, observe first that

$$\|f_j - f_\infty\|^2 = \|f_j\|^2 + \|f_\infty\|^2 - \langle f_j, f_\infty\rangle - \langle f_\infty, f_j\rangle. \tag{3.1.6}$$

If $\text{w-}\lim_{n\to\infty} f_j = f_\infty$ and $\lim_{j\to\infty} \|f_j\| = \|f_\infty\|$, then the right-hand side of (3.1.6) converges to $\|f_\infty\|^2 + \|f_\infty\|^2 - \|f_\infty\|^2 - \|f_\infty\|^2 = 0$, so that $\text{s-}\lim_{j\to\infty} f_j = f_\infty$. □

Let us also note that if $\text{s-}\lim_{j\to\infty} f_j = f_\infty$ and $\text{s-}\lim_{j\to\infty} g_j = g_\infty$ then one has

$$\lim_{j\to\infty} \langle f_j, g_j\rangle = \langle f_\infty, g_\infty\rangle$$

by a simple application of Schwarz inequality. In the next definition, we introduce the notion of subspace of a Hilbert space.

Definition 3.1.8 (Subspace, closed subspace)**.** A subspace $\mathcal{M}$ *of a Hilbert space* $\mathcal{H}$ *is a linear subset of* $\mathcal{H}$*, or more precisely* $\forall f, g \in \mathcal{M}$ *and* $\lambda \in \mathbb{C}$ *one has* $f + \lambda g \in \mathcal{M}$*. The subspace* $\mathcal{M}$ *is* closed *if whenever a sequence* $(f_j)_{j\in\mathbb{N}} \subset \mathcal{M}$ *converges to* $f_\infty \in \mathcal{H}$*, then* $f_\infty \in \mathcal{M}$*. In other words, a closed space contains its limit points.*

Note that if $\mathcal{M}$ is closed, then $\mathcal{M}$ is a Hilbert space in itself, with the scalar product and norm inherited from $\mathcal{H}$. For the next examples, we recall that $f, g \in \mathcal{H}$ are said to be *orthogonal* if $\langle f, g\rangle = 0$. In this case, we write $f \perp g$.

Examples 3.1.9.

1. *If* $f_1, \dots, f_j \in \mathcal{H}$*, then* $\mathrm{Vect}(f_1, \dots, f_j)$ *is the closed vector space generated by the linear combinations of* $f_1, \dots f_j$*.* $\mathrm{Vect}(f_1, \dots, f_j)$ *is a closed subspace.*
2. *The sets* $\mathcal{D}(\mathbb{R}^n)$ *or* $\mathcal{S}(\mathbb{R}^n)$ *are subspaces of* $L^2(\mathbb{R}^n)$*, but they are not closed. Their closure is equal to* $L^2(\mathbb{R}^n)$ *itself. For that reason, we say that these subspaces are dense in* $L^2(\mathbb{R}^n)$*.*
3. *If* $\mathcal{M}$ *is a subset of* $\mathcal{H}$*, then*
$$\mathcal{M}^\perp := \{f \in \mathcal{H} \mid \langle f, g\rangle = 0, \forall g \in \mathcal{M}\} \tag{3.1.7}$$
is a closed subspace of $\mathcal{H}$*.*

Exercise 3.1.10. *Check that in the above example the set* $\mathcal{M}^\perp$ *is a closed subspace of* $\mathcal{H}$*. The subspace* $\mathcal{M}^\perp$ *is called the* orthocomplement *of* $\mathcal{M}$ *in* $\mathcal{H}$*.*

Let us be more precise about the notion of dense subset: A subset $\mathcal{M}$ in a Hilbert space $\mathcal{H}$ (or more generally in a normed space) is *dense* in $\mathcal{H}$ if for any $f \in \mathcal{H}$ and any $\varepsilon > 0$ there exists $g \in \mathcal{M}$ with $\|f - g\| \leq \varepsilon$.

Exercise 3.1.11. *Check that a subspace* $\mathcal{M} \subset \mathcal{H}$ *is dense in* $\mathcal{H}$ *if and only if* $\mathcal{M}^\perp = \{0\}$*.*

The following result is important in the setting of Hilbert spaces. Its proof is not complicated but a little bit long, we thus refer to [Am, Prop. 1.7].

Proposition 3.1.12 (Projection Theorem). *Let $\mathcal{M}$ be a closed subspace of a Hilbert space $\mathcal{H}$. Then, for any $f \in \mathcal{H}$ there exist a unique $f_1 \in \mathcal{M}$ and a unique $f_2 \in \mathcal{M}^\perp$ such that $f = f_1 + f_2$.*

Let us close this section with the so-called Riesz Lemma. For that purpose, recall first that the dual $\mathcal{H}^*$ of the Hilbert space $\mathcal{H}$ consists in the set of all bounded linear functionals on $\mathcal{H}$, *i.e.* $\mathcal{H}^*$ consists in all mappings $\varphi : \mathcal{H} \to \mathbb{C}$ satisfying for any $f, g \in \mathcal{H}$ and $\lambda \in \mathbb{C}$

1. $\varphi(f + \lambda g) = \varphi(f) + \lambda\varphi(g)$, (linearity)
2. $|\varphi(f)| \leq c\|f\|$, (boundedness)

where c is a constant independent of f. One then sets

$$\|\varphi\|_{\mathcal{H}^*} := \sup_{0 \neq f \in \mathcal{H}} \frac{|\varphi(f)|}{\|f\|}.$$

Clearly, if $g \in \mathcal{H}$, then g defines an element φ_g of $\mathcal{H}^*$ by setting $\varphi_g(f) := \langle g, f\rangle$. Indeed φ_g is linear and one has

$$\|\varphi_g\|_{\mathcal{H}^*} := \sup_{0 \neq f \in \mathcal{H}} \frac{1}{\|f\|}|\langle g, f\rangle| \leq \sup_{0 \neq f \in \mathcal{H}} \frac{1}{\|f\|}\|g\|\,\|f\| = \|g\|.$$

In fact, note that $\|\varphi_g\|_{\mathcal{H}^*} = \|g\|$ since $\frac{1}{\|g\|}\varphi_g(g) = \frac{1}{\|g\|}\|g\|^2 = \|g\|$.

The following statement shows that any element $\varphi \in \mathcal{H}^*$ can be obtained from an element $g \in \mathcal{H}$. It corresponds thus to a converse of the previous construction.

Lemma 3.1.13 (Riesz Lemma). *For any $\varphi \in \mathcal{H}^*$, there exists a unique $g \in \mathcal{H}$ such that for any $f \in \mathcal{H}$*

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

In addition, g satisfies $\|\varphi\|_{\mathcal{H}^*} = \|g\|$.

Since the proof is quite standard, we only sketch it and leave the details to any motivated student (you ?), see also [Am, Prop. 1.8].

Sketch of the proof. If $\varphi \equiv 0$, then one can set $g := 0$ and observe trivially that $\varphi = \varphi_g$.

If $\varphi \neq 0$, let us first define $\mathcal{M} := \{f \in \mathcal{H} \mid \varphi(f) = 0\}$ and observe that $\mathcal{M}$ is a subspace of $\mathcal{H}$. One also observes that $\mathcal{M} \neq \mathcal{H}$ since otherwise $\varphi \equiv 0$. Thus, let $h \in \mathcal{H}$ such that $\varphi(h) \neq 0$ and decompose $h = h_1 + h_2$ with $h_1 \in \mathcal{M}$ and $h_2 \in \mathcal{M}^\perp$ by Proposition 3.1.12. One infers then that $\varphi(h_2) = \varphi(h) \neq 0$.

For arbitrary $f \in \mathcal{H}$ one can consider the element $f - \frac{\varphi(f)}{\varphi(h_2)}h_2 \in \mathcal{H}$ and observe that $\varphi(f - \frac{\varphi(f)}{\varphi(h_2)}h_2) = 0$. One deduces that $f - \frac{\varphi(f)}{\varphi(h_2)}h_2$ belongs to $\mathcal{M}$, and since $h_2 \in \mathcal{M}^\perp$ one infers that

$$\varphi(f) = \frac{\varphi(h_2)}{\|h_2\|^2}\langle h_2, f\rangle.$$

One can thus set $g := \frac{\overline{\varphi(h_2)}}{\|h_2\|^2}h_2 \in \mathcal{H}$ and easily obtain the remaining parts of the statement. □

As a consequence of the previous statement, one often identifies $\mathcal{H}^*$ with $\mathcal{H}$ itself.

Exercise 3.1.14. *Check that this identification is not linear but* anti-linear.

3.2 Bounded linear operators

First of all, let us recall that a linear map B between two complex vector spaces $\mathcal{M}$ and $\mathcal{N}$ satisfies $B(f+\lambda g) = Bf + \lambda Bg$ for all $f, g \in \mathcal{M}$ and $\lambda \in \mathbb{C}$.

Definition 3.2.1 (Bounded linear operators)**.** *A map $B : \mathcal{H} \to \mathcal{H}$ is a* bounded linear operator *if $B : \mathcal{H} \to \mathcal{H}$ is a linear map, and if there exists $c > 0$ such that $\|Bf\| \leq c\|f\|$ for all $f \in \mathcal{H}$. The set of* all bounded linear operators on $\mathcal{H}$ *is denoted by $\mathcal{B}(\mathcal{H})$.*

Note that the set of bounded linear operators is sometimes denoted by $\mathcal{L}(\mathcal{H})$. For any $B \in \mathcal{B}(\mathcal{H})$, one sets

$$\begin{aligned} \|B\| &:= \inf\{c > 0 \mid \|Bf\| \leq c\|f\| \ \forall f \in \mathcal{H}\} \\ &= \sup_{0 \neq f \in \mathcal{H}} \frac{\|Bf\|}{\|f\|}. \end{aligned} \tag{3.2.1}$$

and call it *the norm of B*. Observe that the same notation is used for the norm of an element of $\mathcal{H}$ and for the norm of an element of $\mathcal{B}(\mathcal{H})$, but this does not lead to any confusion.

Examples 3.2.2. *1) If $\mathcal{H} = \mathbb{C}^n$, check that any square matrix $A \in M_n(\mathbb{C})$ defines a bounded linear operator.*

2) If $\mathcal{H} = L^2(\mathbb{R}^n)$ and $m : \mathbb{R}^n \to \mathbb{C}$ is a continuous and bounded function, heck that the multiplication operator *M defined by $[Mf](X) := m(X) f(X)$ for any $f \in \mathcal{H}$ and $X \in \mathbb{R}^n$, defines a bounded linear operator on $\mathcal{H}$.*

Exercise 3.2.3. *Let $\mathcal{M}_1, \mathcal{M}_2$ be two dense subspaces of $\mathcal{H}$, and let $B \in \mathcal{B}(\mathcal{H})$. Show that*

$$\|B\| = \sup_{f \in \mathcal{M}_1, g \in \mathcal{M}_2 \text{ with } \|f\| = \|g\| = 1} |\langle f, Bg \rangle|. \tag{3.2.2}$$

Exercise 3.2.4. *Show that $\mathcal{B}(\mathcal{H})$ is a normed algebra, namely it is an algebra with a norm satisfying the inequality*

$$\|AB\| \leq \|A\| \, \|B\| \tag{3.2.3}$$

for any $A, B \in \mathcal{B}(\mathcal{H})$.

An additional structure can be added to $\mathcal{B}(\mathcal{H})$: an involution.

Lemma 3.2.5. *For any $B \in \mathcal{B}(\mathcal{H})$, there exists a unique $B^* \in \mathcal{B}(\mathcal{H})$ satisfying for any $f, g \in \mathcal{H}$:*

$$\langle B^* f, g \rangle := \langle f, Bg \rangle. \tag{3.2.4}$$

The operator B^* is called the adjoint of B

Exercise 3.2.6. *Let $\mathcal{H} = L^2(\mathbb{R}^n)$ and let $\varphi \in L^\infty(\mathbb{R}^n)$ be an essentially bounded function defined on $\mathbb{R}^n$. We set M_φ for the operator of multiplication by φ, namely $[M_\varphi f](X) := \varphi(X) f(X)$ for all $f \in \mathcal{H}$ and* a.e. *$X \in \mathbb{R}^n$. Check that the adjoint of M_φ is given by the operator $M_{\overline{\varphi}}$, the multiplication operator by the function $\overline{\varphi}$ (the complex conjugated function of φ).*

Exercise 3.2.7. *For any $B \in \mathcal{B}(\mathcal{H})$ show that:*

1. *B^* is uniquely defined by* (3.2.4) *and satisfies $B^* \in \mathcal{B}(\mathcal{H})$ with $\|B^*\| = \|B\|$. The existence can be proved with Riesz lemma,*
2. *$(B^*)^* = B$,*
3. *$\|B^* B\| = \|B\|^2$,*

4. *If $A \in \mathcal{B}(\mathcal{H})$, then $(AB)^* = B^*A^*$.*

Remark 3.2.8. *A complete normed algebra endowed with an involution for which the property 3. holds is called a C^*-algebra. In particular $\mathcal{B}(\mathcal{H})$ is a C^*-algebra. Such algebras have a well-developed and deep theory, see for example [Mu]. However, we shall not go further in this direction in this course.*

Exercise 3.2.9. *Find out the correct norm on $M_n(\mathbb{C})$ (the set of $n \times n$ matrices) which makes $M_n(\mathbb{C})$ a C^*-algebra. Note that if $B = (b_{jk})_{j,k=1}^n$, the matrix B^* is given by $B^* = (\overline{b_{kj}})_{j,k=1}^n$.*

We have already considered two distinct topologies on $\mathcal{H}$, namely the strong and the weak topology. On $\mathcal{B}(\mathcal{H})$ there exist several topologies, but we shall consider only three of them.

Definition 3.2.10 (Convergence in $\mathcal{B}(\mathcal{H})$)**.** *Consider a sequence $(B_j)_{j\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$, and let $B_\infty \in \mathcal{B}(\mathcal{H})$.*

1. *This sequence is is* uniformly convergent *to B_∞ if $\lim_{j\to\infty} \|B_j - B_\infty\| = 0$,*
2. *This sequence is* strongly convergent *to B_∞ if for any $f \in \mathcal{H}$ one has $\lim_{j\to\infty} \|(B_j - B_\infty)f\| = 0$,*
3. *This sequence is* weakly convergent *to B_∞ if for any $f, g \in \mathcal{H}$ one has $\lim_{j\to\infty}\langle g, (B_j - B_\infty)f\rangle = 0$.*

In these cases, one writes respectively $u - \lim_{j\to\infty} B_j = B_\infty$, $s - \lim_{j\to\infty} B_j = B_\infty$ and $w - \lim_{j\to\infty} B_j = B_\infty$.

Clearly, uniform convergence implies strong convergence, and strong convergence implies weak convergence. The reverse statements are not true. Note that if $(B_j)_{j\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ is weakly convergent, then the sequence $(B_j^*)_{j\in\mathbb{N}}$ of its adjoint operators is also weakly convergent. However, the same statement does not hold for a strongly convergent sequence. Finally, we shall not prove but often use that $\mathcal{B}(\mathcal{H})$ is also weakly and strongly closed. In other words, any weakly (or strongly) Cauchy sequence in $\mathcal{B}(\mathcal{H})$ converges in $\mathcal{B}(\mathcal{H})$.

Exercise 3.2.11. *Exhibit an example of a strongly convergent sequence of bounded operators which is not convergent in norm. Exhibit a sequence of bounded operators which is weakly convergent but not strongly convergent.*

Exercise 3.2.12. *Let $(A_j)_{j\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ and $(B_j)_{j\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ be two strongly convergent sequence in $\mathcal{B}(\mathcal{H})$, with limits A_∞ and B_∞ respectively. Show that the sequence $(A_jB_j)_{j\in\mathbb{N}}$ is strongly convergent to the element $A_\infty B_\infty$.*

Let us now move to the notion of invertibility, similar to the invertibility of a matrix.

Definition 3.2.13 (Invertibility, bounded invertibility)**.** *An operator $B \in \mathcal{B}(\mathcal{H})$ is* invertible *if the equation $Bf = 0$ only admits the solution $f = 0$ (injectivity condition). In such a case, there exists a linear map $B^{-1} : \mathrm{Ran}(B) \to \mathcal{H}$ which satisfies $B^{-1}Bf = f$ for any $f \in \mathcal{H}$, and $BB^{-1}g = g$ for any $g \in \mathrm{Ran}(B)$. If B is invertible and $\mathrm{Ran}(B) = \mathcal{H}$, then $B^{-1} \in \mathcal{B}(\mathcal{H})$ and B is said to be* invertible in $\mathcal{B}(\mathcal{H})$ *(or* boundedly invertible*).*

Note that the two conditions B invertible and $\mathrm{Ran}(B) = \mathcal{H}$ imply $B^{-1} \in \mathcal{B}(\mathcal{H})$ is a consequence of the Closed graph Theorem. However, if $\mathrm{Ran}(B) \neq \mathcal{H}$, the operator B^{-1} might not be bounded, as shown in the following exercise.

Exercise 3.2.14. *In the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$, consider the bounded operator of multiplication M by the function $m : \mathbb{R} \to \mathbb{R}$ with $m(x) = \tanh(x)$, namely $[Mf](x) = \tanh(x)f(x)$ for any $f \in \mathcal{H}$ and any $x \in \mathbb{R}$. Show that the inverse of M is given by the operator M^{-1} defined by $[M^{-1}f](x) := \frac{1}{\tanh(x)}f(x)$ but that this operator is not a bounded operator. In other words, show that there is no $c < \infty$ such that $\|M^{-1}f\| \leq c\|f\|$ for all $f \in \mathcal{H}$.*

In the sequel, we shall use the notation $\mathbf{1} \in \mathcal{B}(\mathcal{H})$ for the operator defined on any $f \in \mathcal{H}$ by $\mathbf{1}f = f$, and $\mathbf{0} \in \mathcal{B}(\mathcal{H})$ for the operator defined by $\mathbf{0}f = 0$.

The next statement is very useful in applications, and holds in a much more general context. Its proof is classical and can be found in every textbook.

Lemma 3.2.15 (Neumann series)**.** *If $B \in \mathcal{B}(\mathcal{H})$ and $\|B\| < 1$, then the operator $(\mathbf{1} - B)$ is invertible in $\mathcal{B}(\mathcal{H})$, with*

$$(\mathbf{1} - B)^{-1} = \sum_{n=0}^{\infty} B^n,$$

and with $\left\|(\mathbf{1} - B)^{-1}\right\| \leq (1 - \|B\|)^{-1}$. The series converges in the uniform norm of $\mathcal{B}(\mathcal{H})$.

Note that we have used the identity $B^0 = \mathbf{1}$.

3.3 Special classes of bounded linear operators

In this section we provide information on some subsets of $\mathcal{B}(\mathcal{H})$. We start with some operators which will play an important role in the sequel.

Definition 3.3.1 (Self-adjoint operator)**.** *An operator $B \in \mathcal{B}(\mathcal{H})$ is called* self-adjoint *or* Hermitian *if $B^* = B$, or equivalently if for any $f, g \in \mathcal{H}$ one has*

$$\langle f, Bg \rangle = \langle Bf, g \rangle. \tag{3.3.1}$$

For these operators the computation of their norm can be simplified (see also Exercise 3.2.3) :

Exercise 3.3.2. *If $B \in \mathcal{B}(\mathcal{H})$ is self-adjoint and if $\mathcal{M}$ is a dense subspace in $\mathcal{H}$, show that*

$$\|B\| = \sup_{f \in \mathcal{M},\ \|f\|=1} |\langle f, Bf \rangle|. \tag{3.3.2}$$

Exercise 3.3.3. *In the framework of multiplication operators introduced in Exercise 3.2.6, show that M_φ is a self-adjoint operator if and only if φ is a real valued function.*

A special set of self-adjoint operators is provided by the set of orthogonal projections:

Definition 3.3.4 (Projection)**.** *An element $P \in \mathcal{B}(\mathcal{H})$ is an* orthogonal projection *if $P = P^2 = P^*$.*

It not difficult to check that there is a one-to-one correspondence between the set of closed subspaces of $\mathcal{H}$ and the set of orthogonal projections in $\mathcal{B}(\mathcal{H})$. Indeed, any orthogonal projection P defines a closed subspace $\mathcal{M} := P\mathcal{H}$. Conversely by taking the projection Theorem (Proposition 3.1.12) into account one infers that for any closed subspace $\mathcal{M}$ one can define an orthogonal projection P with $P\mathcal{H} = \mathcal{M}$.

Remark 3.3.5. *If $\mathcal{H} = L^2(\mathbb{R}^n)$, we can often think about a projection as a characteristic function. Indeed, if Ω is a closed subset of $\mathbb{R}^n$ and if χ_Ω denotes the characteristic on the subset Ω, then the multiplication operator M_{χ_Ω} defines an orthogonal projection, see Exercise 3.2.6 for the notation. In this case, the closed subspace $\mathcal{M}$ associated with this projection is the set of $f \in L^2(\mathbb{R}^n)$ such that* $\operatorname{supp}(f) \subset \Omega$*. In other words, the closed subspace $\mathcal{M}$ correspond to $L^2(\Omega)$, seen as a subspace of $L^2(\mathbb{R}^n)$.*

Observe that if P, Q are two orthogonal projections, the products PQ and QP are not orthogonal projections in general. In the sequel, we might simply say projection instead of orthogonal projection. However, let us stress that in other contexts a projection is often an operator P satisfying only the relation $P^2 = P$. We gather in the next exercise some easy relations between orthogonal projections and the underlying closed subspaces. For that purpose we use the notation $P_{\mathcal{M}}, P_{\mathcal{N}}$ for the orthogonal projections on the closed subspaces $\mathcal{M}$ and $\mathcal{N}$ of $\mathcal{H}$.

Exercise 3.3.6. *Show the following relations:*

1. *If $P_{\mathcal{M}}P_{\mathcal{N}} = P_{\mathcal{N}}P_{\mathcal{M}}$, then $P_{\mathcal{M}}P_{\mathcal{N}}$ is a projection and the associated closed subspace is $\mathcal{M} \cap \mathcal{N}$,*
2. *If $\mathcal{M} \subset \mathcal{N}$, then $P_{\mathcal{M}}P_{\mathcal{N}} = P_{\mathcal{N}}P_{\mathcal{M}} = P_{\mathcal{M}}$,*
3. *If $\mathcal{M} \perp \mathcal{N}$, then $P_{\mathcal{M}}P_{\mathcal{N}} = P_{\mathcal{N}}P_{\mathcal{M}} = \mathbf{0}$, and $P_{\mathcal{M}\oplus\mathcal{N}} = P_{\mathcal{M}} + P_{\mathcal{N}}$,*
4. *If $P_{\mathcal{M}}P_{\mathcal{N}} = \mathbf{0}$, then $\mathcal{M} \perp \mathcal{N}$.*

Let us now consider unitary operators, and then more general isometries and partial isometries. For that purpose, we recall that $\mathbf{1}$ denotes the identify operator in $\mathcal{B}(\mathcal{H})$.

Definition 3.3.7 (Unitary operator)**.** *An element $U \in \mathcal{B}(\mathcal{H})$ is a* unitary operator *if $UU^* = \mathbf{1}$ and if $U^*U = \mathbf{1}$.*

Let us emphasize that if $\mathcal{H}$ is of infinite dimension, then the two conditions $UU^* = \mathbf{1}$ and $U^*U = \mathbf{1}$ are not equivalent. For $\mathcal{H} = \mathbb{C}^n$ and when $U \in M_n(\mathbb{C})$, then these two conditions are equivalent. In the general situation, note that if U is unitary, then U is invertible in $\mathcal{B}(\mathcal{H})$ with $U^{-1} = U^*$. Indeed, observe first that $Uf = 0$ implies $f = U^*(Uf) = U^*0 = 0$. Secondly, for any $g \in \mathcal{H}$, one has $g = U(U^*g)$, and thus $\mathrm{Ran}(U) = \mathcal{H}$. Finally, the equality $U^{-1} = U^*$ follows from the unicity of the inverse.

More generally, one sets:

Definition 3.3.8 (Isometry)**.** *An element $V \in \mathcal{B}(\mathcal{H})$ is called an* isometry *if the equality*

$$V^*V = \mathbf{1} \tag{3.3.3}$$

holds.

Clearly, a unitary operator is an instance of an isometry. For isometries the following properties can easily be obtained.

Proposition 3.3.9. *a) Let $V \in \mathcal{B}(\mathcal{H})$ be an isometry. Then*

1. *V preserves the scalar product, namely $\langle Vf, Vg\rangle = \langle f, g\rangle$ for any $f, g \in \mathcal{H}$,*
2. *V preserves the norm, namely $\|Vf\| = \|f\|$ for any $f \in \mathcal{H}$,*
3. *If $\mathcal{H} \neq \{0\}$ then $\|V\| = 1$,*
4. *VV^* is the projection on $\mathrm{Ran}(V)$,*
5. *V is invertible (in the sense of Definition 3.2.13),*
6. *The adjoint V^* satisfies $V^*f = V^{-1}f$ if $f \in \mathrm{Ran}(V)$, and $V^*g = 0$ if $g \perp \mathrm{Ran}(V)$.*

b) An element $W \in \mathcal{B}(\mathcal{H})$ is an isometry if and only if $\|Wf\| = \|f\|$ for all $f \in \mathcal{H}$.

Exercise 3.3.10. *Provide a proof for the previous proposition (as well as the proof of the next proposition).*

More generally one can still set:

Definition 3.3.11 (Partial isometry)**.** *An element $W \in \mathcal{B}(\mathcal{H})$ is called a* partial isometry *if the following equality holds:*

$$W^*W = P \tag{3.3.4}$$

with P an orthogonal projection.

Again, unitary operators or isometries are special examples of partial isometries. As before the following properties of partial isometries can be easily proved.

Proposition 3.3.12. *Let $W \in \mathcal{B}(\mathcal{H})$ be a partial isometry as defined in* (3.3.4)*. Then*

1. *One has $WP = W$ and $\langle Wf, Wg\rangle = \langle Pf, Pg\rangle$ for any $f, g \in \mathcal{H}$,*
2. *If $P \neq \mathbf{0}$ then $\|W\| = 1$,*
3. *$Q := WW^*$ is the projection on* $\mathrm{Ran}(W)$*, and $QW = W$.*

For a partial isometry W one usually calls *initial set projection* the projection defined by $P := W^*W$ and by *final set projection* the projection defined by $Q := WW^*$.

Let us now introduce a last subset of bounded operators, namely the ideal of *compact operators*. For that purpose, consider first any family $\{g_j, h_j\}_{j=1}^N \subset \mathcal{H}$ and for any $f \in \mathcal{H}$ one sets

$$Af \equiv \Big(\sum_{j=1}^N |h_j\rangle\langle g_j|\Big)f := \sum_{j=1}^N \langle g_j, f\rangle h_j. \tag{3.3.5}$$

Then $A \in \mathcal{B}(\mathcal{H})$, and $\mathrm{Ran}(A) \subset \mathrm{Vect}(h_1, \dots, h_N)$. Such an operator A is called a *finite rank operator*, and the notation used is called the *bra-ket* notation. In fact, any operator $B \in \mathcal{B}(\mathcal{H})$ with $\dim(\mathrm{Ran}(B)) < \infty$ is a finite rank operator.

Exercise 3.3.13. *For the operator A defined in* (3.3.5)*, give an upper estimate for $\|A\|$ and compute A^*.*

Definition 3.3.14 (Compact operator)**.** *An element $B \in \mathcal{B}(\mathcal{H})$ is a* compact operator *if there exists a family $\{A_j\}_{j\in\mathbb{N}}$ of finite rank operators such that* $\lim_{j\to\infty} \|B - A_j\| = 0$*. The set of all compact operators is denoted by $\mathcal{K}(\mathcal{H})$.*

Exercise 3.3.15. *Check that a projection P is a compact operator if and only if $P\mathcal{H}$ is of finite dimension.*

The following proposition contains some basic properties of $\mathcal{K}(\mathcal{H})$. Its proof can be obtained by playing with families of finite rank operators.

Proposition 3.3.16. *The following properties hold:*

1. *$B \in \mathcal{K}(\mathcal{H}) \iff B^* \in \mathcal{K}(\mathcal{H})$,*
2. *$\mathcal{K}(\mathcal{H})$ is a $*$-algebra, complete for the norm $\|\cdot\|$,*
3. *If $B \in \mathcal{K}(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$, then AB and BA belong to $\mathcal{K}(\mathcal{H})$.*

As a consequence, $\mathcal{K}(\mathcal{H})$ is a C^*-algebra and an ideal of $\mathcal{B}(\mathcal{H})$. In fact, compact operators have the nice property of improving some convergences, as shown in the next statement.

Proposition 3.3.17. *Let $K \in \mathcal{K}(\mathcal{H})$).*

1. *If $\{f_j\}_{j\in\mathbb{N}} \subset \mathcal{H}$ is a weakly convergent sequence with limit $f_\infty \in \mathcal{H}$, then the sequence $\{Kf_j\}_{j\in\mathbb{N}}$ strongly converges to Kf_∞,*

2. *If the sequence $\{B_j\}_{j\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ strongly converges to $B_\infty \in \mathcal{B}(\mathcal{H})$, then the sequences $\{B_jK\}_{j\in\mathbb{N}}$ and $\{KB_j^*\}_{j\in\mathbb{N}}$ converge in norm to $B_\infty K$ and KB_∞^*, respectively.*

3.4 Vector valued functions and operator valued functions

Instead of function defined on (a,b) and with value in $\mathbb{R}$, in $\mathbb{C}$, or in $\mathbb{C}^n$, it is natural to consider more generally vector valued functions $\varphi : (a,b) \to \mathcal{H}$, namely functions taking values in an arbitrary Hilbert space. Various notions of continuity and differentiability can then be defined, as for example:

Definition 3.4.1 (Continuity and differentiability of vector valued functions)**.** *Let $I := (a,b)$ with $a < b$ and consider a vector-valued function $\varphi : I \to \mathcal{H}$.*

1. *φ is* strongly continuous *on I if for any $t \in I$ one has $\lim_{\varepsilon\to 0} \|\varphi(t+\varepsilon) - \varphi(t)\| = 0$,*
2. *φ is* weakly continuous *on I if for any $t \in I$ and any $g \in \mathcal{H}$ one has*
$$\lim_{\varepsilon\to 0} \langle g, \varphi(t+\varepsilon) - \varphi(t)\rangle = 0,$$
3. *φ is* strongly differentiable *on I if there exists another vector-valued function $\varphi' : I \to \mathcal{H}$ such that for any $t \in I$ one has*
$$\lim_{\varepsilon\to 0} \Big\| \tfrac{1}{\varepsilon}(\varphi(t+\varepsilon) - \varphi(t)) - \varphi'(t) \Big\| = 0,$$
4. *φ is* weakly differentiable Weakly differentiable *on I if there exists another vector-valued function $\varphi' : I \to \mathcal{H}$ such that for any $t \in I$ and $g \in \mathcal{H}$ one has*
$$\lim_{\varepsilon\to 0} \langle g, \tfrac{1}{\varepsilon}(\varphi(t+\varepsilon) - \varphi(t)) - \varphi'(t)\rangle = 0,$$

The map φ' is called the strong derivative*, respectively the* weak derivative*, of φ.*

Note that functions from (a,b) to $\mathcal{H}$ define a vector space, but the multiplication of two such functions is not defined (one can not multiply to elements of $\mathcal{H}$). It is thus not an algebra.

Similarly, we can consider operator valued functions $\Psi : (a,b) \to \mathcal{B}(\mathcal{H})$ and define various notions of continuity and differentiability. For example, for the continuity one can set:

Definition 3.4.2 (Continuity of operator valued functions)**.** *Let $I := (a,b)$ with $a < b$ and consider a vector-valued function $\Psi : I \to \mathcal{B}(\mathcal{H})$.*

1. *The map Ψ is* continuous in norm *on I if for all $t \in I$*
$$\lim_{\varepsilon\to 0} \big\| \Psi(t+\varepsilon) - \Psi(t) \big\| = 0,$$
2. *The map Ψ is* strongly continuous *on I if for any $f \in \mathcal{H}$ and all $t \in I$*
$$\lim_{\varepsilon\to 0} \big\| \Psi(t+\varepsilon)f - \Psi(t)f \big\| = 0,$$
3. *The map Ψ is* weakly continuous *on I if for any $f, g \in \mathcal{H}$ and all $t \in I$*
$$\lim_{\varepsilon\to 0} \langle g, (\Psi(t+\varepsilon) - \Psi(t))f\rangle = 0.$$

One writes respectively $u - \lim_{\varepsilon\to 0} \Psi(t+\varepsilon) = \Psi(t)$, $s - \lim_{\varepsilon\to 0} \Psi(t+\varepsilon) = \Psi(t)$ *and* $w - \lim_{\varepsilon\to 0} \Psi(t+\varepsilon) = \Psi(t)$.

The same type of definition holds for the differentiability. Note also that such functions from (a, b) to $\mathcal{B}(\mathcal{H})$ define an algebra, since $\Psi(t)\Phi(t)$ corresponds to the product in $\mathcal{B}(\mathcal{H})$, for any $\Psi, \Phi : (a, b) \to \mathcal{B}(\mathcal{H})$ and for any $t \in (a, b)$.

For functions $\varphi : (a, b) \to \mathcal{H}$ or $\Psi : (a, b) \to \mathcal{B}(\mathcal{H})$ one can define Riemann type's integrals, by considering finer partitions of the interval (a, b). The convergence of these sums can be defined with the different notions of continuity recalled above. Usually, the strong convergence is the one considered, but for $\Psi : (a, b) \to \mathcal{B}(\mathcal{H})$ the uniform convergence of the Riemann sums is also interesting for the applications. Note that $\int_a^n \varphi(t)\,\mathrm{d}t$ corresponds to an element of $\mathcal{H}$, while $\int_a^b \Psi(t)\,\mathrm{d}t$ corresponds to an element of $\mathcal{B}(\mathcal{H})$.

Let us conclude this section by mentioning a few useful results for integrals of vector valued functions or of operator valued functions:

Proposition 3.4.3. *For* $\varphi, \varphi_1, \varphi_2 : (a, b) \to \mathcal{H}$ *and for* $\lambda \in \mathbb{C}$

1. $\int_a^b (\varphi_1(t) + \lambda\varphi_2(t))\,\mathrm{d}t = \int_a^b \varphi_1(t)\,\mathrm{d}t + \lambda \int_a^b \varphi_2(t)\,\mathrm{d}t,$
2. $\left\| \int_a^b \varphi(t)\,\mathrm{d}t \right\| \leq \int_a^b \|\varphi(t)\|\,\mathrm{d}t,$

whenever these integrals exist. The same results hold for $\Psi, \Psi_1, \Psi_2 : (a, b) \to \mathcal{B}(\mathcal{H})$ *replacing* $\varphi, \varphi_1, \varphi_2$, *and also whenever these integrals exist.*

3.5 Unbounded and self-adjoint operators

In this section, we define an extension of the notion of bounded linear operators. Obviously, the following definitions and results are also valid for bounded linear operators.

Definition 3.5.1 (Linear operator)**.** *A* linear operator *on* $\mathcal{H}$ *is a pair* $(A, \mathcal{D}(A))$, *where* $\mathcal{D}(A)$ *is a subspace of* $\mathcal{H}$ *and* A *is a linear map from* $\mathcal{D}(A)$ *to* $\mathcal{H}$. $\mathcal{D}(A)$ *is called the* domain *of* A. *One says that the operator* $(A, \mathcal{D}(A))$ *is* densely defined *if* $\mathcal{D}(A)$ *is dense in* $\mathcal{H}$.

Note that one often just says the *linear operator* A, but that its domain $\mathcal{D}(A)$ is implicitly taken into account. For such an operator, its range $\mathrm{Ran}(A)$ is defined by

$$\mathrm{Ran}(A) := A\,\mathcal{D}(A) = \{f \in \mathcal{H} \mid f = Ag \text{ for some } g \in \mathcal{D}(A)\}.$$

In addition, one defines the kernel $\mathrm{Ker}(A)$ of A by

$$\mathrm{Ker}(A) := \{f \in \mathcal{D}(A) \mid Af = 0\}.$$

Let us also stress that the sum $A + B$ for two linear operators is *a priori* only defined on the subspace $\mathcal{D}(A) \cap \mathcal{D}(B)$, and that the product AB is *a priori* defined only on the subspace $\{f \in \mathcal{D}(B) \mid Bf \in \mathcal{D}(A)\}$. These two sets can be very small.

Example 3.5.2. *Let* $\mathcal{H} := L^2(\mathbb{R})$ *and consider the operator* X *defined by* $[Xf](x) = xf(x)$ *for any* $x \in \mathbb{R}$. *Clearly,* $\mathcal{D}(X) = \{f \in \mathcal{H} \mid \int_{\mathbb{R}} |xf(x)|^2\,\mathrm{d}x < \infty\} \subsetneq \mathcal{H}$. *In addition, by considering the family of functions* $\{f_y\}_{y\in\mathbb{R}} \subset \mathcal{D}(X)$ *with* $f_y(x) := 1$ *in* $x \in [y, y+1]$ *and* $f_y(x) = 0$ *if* $x \notin [y, y+1]$, *one easily observes that* $\|f_y\| = 1$ *but* $\sup_{y\in\mathbb{R}} \|Xf_y\| = \infty$, *which can be compared with* (3.2.1). *As a consequence, the linear operator* X *can not be bounded.*

Clearly, a linear operator A can be defined on several domains. For example the operator X of the previous example is well-defined on the Schwartz space $\mathcal{S}(\mathbb{R})$, or on the set $C_c(\mathbb{R})$ of continuous functions on $\mathbb{R}$ with bounded support, or on the space $\mathcal{D}(X)$ mentioned in the previous example. More generally, one has:

Definition 3.5.3 (Extension, restriction)**.** *For any pair of linear operators $(A, \mathcal{D}(A))$ and $(B, \mathcal{D}(B))$ satisfying $\mathcal{D}(A) \subset \mathcal{D}(B)$ and $Af = Bf$ for all $f \in \mathcal{D}(A)$, one says that $(B, \mathcal{D}(B))$ is an* extension *of $(A, \mathcal{D}(A))$ to $\mathcal{D}(B)$, or that $(A, \mathcal{D}(A))$ is the* restriction *of $(B, \mathcal{D}(B))$ to $\mathcal{D}(A)$.*

Let us now note that if $(A, \mathcal{D}(A))$ is densely defined and if there exists $c \in \mathbb{R}$ such that $\|Af\| \leq c\|f\|$ for all $f \in \mathcal{D}(A)$, then there exists a natural continuous extension $\overline{A}$ of A with $\mathcal{D}(\overline{A}) = \mathcal{H}$. This extension satisfies $\overline{A} \in \mathcal{B}(\mathcal{H})$ with $\|\overline{A}\| \leq c$, and is called the *closure* of the operator A.

Exercise 3.5.4. *Work on the details of this extension.*

If there is no $c > 0$ such that $\|Af\| \leq c\|f\|$ for all $f \in \mathcal{D}(A)$, then the linear operator A is *unbounded*. In this case, some extensions are more convenient than others.

Definition 3.5.5 (Closed operator)**.** *An linear operator $(A, \mathcal{D}(A))$ is* closed *if for any sequence $\{f_n\}_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ with $\text{s-}\lim_{n\to\infty} f_n = f_\infty \in \mathcal{H}$ and such that $\{Af_n\}_{n\in\mathbb{N}}$ is strongly Cauchy, then one has $f_\infty \in \mathcal{D}(A)$ and $\text{s-}\lim_{n\to\infty} Af_n = Af_\infty$.*

Let us now come back to the notion of the adjoint of an operator. This concept is slightly more subtle for unbounded operators than in the bounded case.

Definition 3.5.6 (Adjoint operator)**.** *Let $(A, \mathcal{D}(A))$ be a densely defined linear operator on $\mathcal{H}$. The* adjoint A^* *of A is the operator defined by*

$$\mathcal{D}(A^*) := \Big\{ f \in \mathcal{H} \mid \exists f^* \in \mathcal{H} \text{ with } \langle f^*, g\rangle = \langle f, Ag\rangle \text{ for all } g \in \mathcal{D}(A) \Big\}$$

and $A^* f := f^*$ for all $f \in \mathcal{D}(A^*)$.

Let us note that the density of $\mathcal{D}(A)$ is necessary to ensure that A^* is well defined. Indeed, if f_1^*, f_2^* satisfy for all $g \in \mathcal{D}(A)$

$$\langle f_1^*, g\rangle = \langle f, Ag\rangle = \langle f_2^*, g\rangle,$$

then $\langle f_1^* - f_2^*, g\rangle = 0$ for all $g \in \mathcal{D}(A)$, and this equality implies $f_1^* = f_2^*$ only if $\mathcal{D}(A)$ is dense in $\mathcal{H}$. Note also that once $(A^*, \mathcal{D}(A^*))$ is defined, one has

$$\langle A^* f, g\rangle = \langle f, Ag\rangle \qquad \forall f \in \mathcal{D}(A^*) \text{ and } \forall g \in \mathcal{D}(A).$$

Some relations between A and its adjoint A^* are gathered in the following lemma.

Lemma 3.5.7. *Let $(A, \mathcal{D}(A))$ be a densely defined linear operator on $\mathcal{H}$. Then*

1. *One has* $\operatorname{Ker}(A^*) = \operatorname{Ran}(A)^\perp$,
2. *If $(B, \mathcal{D}(B))$ is an extension of $(A, \mathcal{D}(A))$, then $(A^*, \mathcal{D}(A^*))$ is an extension of $(B^*, \mathcal{D}(B^*))$.*

Proof. 1. Let $f \in \operatorname{Ker}(A^*)$, *i.e.* $f \in \mathcal{D}(A^*)$ and $A^* f = 0$. Then, for all $g \in \mathcal{D}(A)$, one has

$$0 = \langle A^* f, g\rangle = \langle f, Ag\rangle$$

meaning that $f \in \operatorname{Ran}(A)^\perp$. Conversely, if $f \in \operatorname{Ran}(A)^\perp$, then for all $g \in \mathcal{D}(A)$ one has

$$\langle f, Ag\rangle = 0 = \langle 0, g\rangle$$

meaning that $f \in \mathcal{D}(A^*)$ and $A^* f = 0$, by the definition of the adjoint of A.

2. Consider $f \in \mathcal{D}(B^*)$ and observe that $\langle B^* f, g\rangle = \langle f, Bg\rangle$ for any $g \in \mathcal{D}(B)$. Since $(B, \mathcal{D}(B))$ is an extension of $(A, \mathcal{D}(A))$, one infers that $\langle B^* f, g\rangle = \langle f, Ag\rangle$ for any $g \in \mathcal{D}(A)$. Now, this equality means that $f \in \mathcal{D}(A^*)$ and that $A^* f = B^* f$, or more explicitly that A^* is defined on the domain of B^* and coincide with this operator on this domain. This means precisely that $(A^*, \mathcal{D}(A^*))$ is an extension of $(B^*, \mathcal{D}(B^*))$. □

Let us now introduce the analogue of the bounded self-adjoint operators but in the unbounded setting. These operators play a key role in quantum mechanics and their study is very well developed.

Definition 3.5.8 (Self-adjoint)**.** *A densely defined linear operator* $(A, \mathcal{D}(A))$ *is* self-adjoint *if* $\mathcal{D}(A^*) = \mathcal{D}(A)$ *and* $A^* f = Af$ *for all* $f \in \mathcal{D}(A)$.

Note that a self-adjoint operator is always closed (more generally any adjoint $(A^*, \mathcal{D}(A^*))$ is closed). Recall also that in the bounded case, a self-adjoint operator was characterized by the equality

$$\langle Af, g\rangle = \langle f, Ag\rangle \tag{3.5.1}$$

for any $f, g \in \mathcal{H}$. In the unbounded case, such an equality still holds if $f, g \in \mathcal{D}(A)$. However, let us emphasize that (3.5.1) does not completely characterize a self-adjoint operator. In fact, a densely defined operator $(A, \mathcal{D}(A))$ satisfying (3.5.1) is called a *symmetric* operator, and self-adjoint operators are special instances of symmetric operators (but not all symmetric operators are self-adjoint). In fact, for a symmetric operator the adjoint operator $(A^*, \mathcal{D}(A^*))$ is an extension of $(A, \mathcal{D}(A))$, but the equality of these two operators holds only if $(A, \mathcal{D}(A))$ is self-adjoint. The theory of extension of symmetric operators is rich and nice: some of them admit no self-adjoint extension while others admit a unique self-adjoint extension, others admit a finite family of self-adjoint extensions, while some admit an infinite number of families of self-adjoint extensions. They can be characterized and classified. Note finally that for any symmetric operator the scalar $\langle f, Af\rangle$ is real for any $f \in \mathcal{D}(A)$.

3.6 Resolvent and spectrum

We come now to the important notion of the spectrum of an operator. As already mentioned in the previous section we shall often speak about a linear operator A, its domain $\mathcal{D}(A)$ being implicitly taken into account. Recall also that the notion of a closed linear operator has been introduced in Definition 3.5.5.

The notion of the inverse of a bounded linear operator has already been introduced in Definition 3.2.13. By analogy we say that any linear operator A is *invertible* if $\mathrm{Ker}(A) = \{0\}$. In this case, the inverse A^{-1} gives a bijection from $\mathrm{Ran}(A)$ onto $\mathcal{D}(A)$. More precisely $\mathcal{D}(A^{-1}) = \mathrm{Ran}(A)$ and $\mathrm{Ran}(A^{-1}) = \mathcal{D}(A)$. It can then be checked that if A is closed and invertible, then A^{-1} is also closed. Note also if A is closed and if $\mathrm{Ran}(A) = \mathcal{H}$ then $A^{-1} \in \mathcal{B}(\mathcal{H})$, in which case one says that A is *boundedly invertible* or *invertible in* $\mathcal{B}(\mathcal{H})$.

The next definition provides the generalization to the notion of eigenvalues familiar for matrices.

Definition 3.6.1 (Resolvent set, spectrum)**.** *For a closed linear operator A its* resolvent set $\rho(A)$ *is defined by*

$$\begin{aligned}\rho(A) :=& \{z \in \mathbb{C} \mid (A - z) \textit{ is invertible in } \mathcal{B}(\mathcal{H})\}\\ =& \{z \in \mathbb{C} \mid \mathrm{Ker}(A - z) = \{0\} \textit{ and } \mathrm{Ran}(A - z) = \mathcal{H}\}.\end{aligned}$$

For $z \in \rho(A)$ the operator $(A - z)^{-1} \in \mathcal{B}(\mathcal{H})$ is called the resolvent *of A at the point z. The* spectrum $\sigma(A)$ *of A is defined as the complement of $\rho(A)$ in $\mathbb{C}$,* i.e.

$$\sigma(A) := \mathbb{C} \setminus \rho(A). \tag{3.6.1}$$

Exercise 3.6.2. *In the framework of Exercise 3.2.6, namely $\mathcal{H} = L^2(\mathbb{R}^n)$, $\varphi \in L^\infty(\mathbb{R}^n)$, and M_φ the operator of multiplication by φ, determine the spectrum of M_φ.*

The following statement summarized several properties of the resolvent set and of the resolvent of a closed linear operator.

Proposition 3.6.3. *Let A be a closed linear operator on a Hilbert space $\mathcal{H}$. Then*

1. *The resolvent set $\rho(A)$ is an open subset of $\mathbb{C}$,*
2. *If $z_1, z_2 \in \rho(A)$ then the* first resolvent equation *holds, namely*

$$(A - z_1)^{-1} - (A - z_2)^{-1} = (z_1 - z_2)(A - z_1)^{-1}(A - z_2)^{-1} \tag{3.6.2}$$

3. *If $z_1, z_2 \in \rho(A)$ then the operators $(A - z_1)^{-1}$ and $(A - z_2)^{-1}$ commute.*

Exercise 3.6.4. *Provide the proof of the previous proposition.*

As a consequence of the previous proposition, the spectrum of a closed linear operator is always closed. In particular, $z \in \sigma(A)$ if $A - z$ is not invertible or if $\mathrm{Ran}(A - z) \neq \mathcal{H}$. The first situation corresponds to the definition of an eigenvalue:

Definition 3.6.5 (Eigenvalue and eigenfunction)**.** *For a closed linear operator A, a value $z \in \mathbb{C}$ is an* eigenvalue *of A if there exists $f \in \mathcal{D}(A)$, $f \neq 0$, such that $Af = zf$. In such a case, the element f is called an* eigenfunction *or an* eigenvector *of A associated with the eigenvalue z. The dimension of the vector space generated by all eigenfunctions associated with an eigenvalue z is called the* multiplicity *of z. The set of all eigenvalues of A is denoted by $\sigma_p(A)$.*

Exercise 3.6.6. *Show that $\sigma_p(A) \subset \sigma(A)$.*

Exercise 3.6.7. *Provide some examples of multiplication operators having eigenvalues. Can you state a general condition for a multiplication operator to have eigenvalues ?*

Let us still provide two properties of the spectrum of an operator in the special cases of a bounded operator or of a self-adjoint operator.

Exercise 3.6.8. *By using the Neumann series, show that for any $B \in \mathcal{B}(\mathcal{H})$ its spectrum is contained in the ball in the complex plane of center* 0 *and of radius $\|B\|$.*

Theorem 3.6.9. *Let A be a self-adjoint operator in $\mathcal{H}$.*

1. *Any eigenvalue of A is real,*
2. *More generally, the spectrum of A is real,* i.e. $\sigma(A) \subset \mathbb{R}$,
3. *Eigenvectors associated with different eigenvalues are orthogonal to one another.*

Proof. a) Assume that there exists $z \in \mathbb{C}$ and $f \in \mathcal{D}(A)$, $f \neq 0$ such that $Af = zf$. Then one has

$$z\|f\|^2 = \langle f, zf\rangle = \langle f, Af\rangle = \langle Af, f\rangle = \langle zf, f\rangle = \bar{z}\|f\|^2.$$

Since $\|f\| \neq 0$, one deduces that $z \in \mathbb{R}$.

b) Let us consider $z = \lambda + i\varepsilon$ with $\lambda, \varepsilon \in \mathbb{R}$ and $\varepsilon \neq 0$, and show that $z \in \rho(A)$. Indeed, for any $f \in \mathcal{D}(A)$ one has

$$\|(A - z)f\|^2 = \|(A - \lambda)f - i\varepsilon f\|^2$$

$$= \langle (A-\lambda)f - i\varepsilon f, (A-\lambda)f - i\varepsilon f\rangle$$
$$= \|(A-\lambda)f\|^2 + \varepsilon^2\|f\|^2.$$

It follows that $\|(A-z)f\| \geq |\varepsilon|\|f\|$, and thus $A-z$ is invertible.

Now, for any for any $g \in \mathrm{Ran}(A-z)$ let us observe that

$$\|g\| = \big\|(A-z)(A-z)^{-1}g\big\| \geq |\varepsilon|\big\|(A-z)^{-1}g\big\|.$$

Equivalently, it means for all $g \in \mathrm{Ran}(A-z)$, one has

$$\big\|(A-z)^{-1}g\big\| \leq \frac{1}{|\varepsilon|}\|g\|. \tag{3.6.3}$$

Let us finally observe that $\mathrm{Ran}(A-z)$ is dense in $\mathcal{H}$. Indeed, by Lemma 3.5.7 one has

$$\mathrm{Ran}(A-z)^\perp = \mathrm{Ker}((A-z)^*) = \mathrm{Ker}(A^* - \bar{z}) = \mathrm{Ker}(A-\bar{z}) = \{0\}$$

since all eigenvalues of A are real. Thus, the operator $(A-z)^{-1}$ is defined on the dense domain $\mathrm{Ran}(A-z)$ and satisfies the estimate (3.6.3). As explained just before the Exercise 3.5.4, it means that $(A-z)^{-1}$ continuously extends to an element of $\mathcal{B}(\mathcal{H})$, and therefore $z \in \rho(A)$.

c) Assume that $Af = \lambda f$ and that $Ag = \mu g$ with $\lambda, \mu \in \mathbb{R}$ and $\lambda \neq \mu$, and $f, g \in \mathcal{D}(A)$, with $f \neq 0$ and $g \neq 0$. Then

$$\lambda\langle f, g\rangle = \langle Af, g\rangle = \langle f, Ag\rangle = \mu\langle f, g\rangle,$$

which implies that $\langle f, g\rangle = 0$, or in other words that f and g are orthogonal. □

3.7 Examples of self-adjoint operators

In this section we briefly present some examples of self-adjoint operators which often appear in the literature. Most of them are related to quantum mechanics. Indeed, any physical system is described with such an operator. Self-adjoint operators are the natural generalization of Hermitian matrices. Obviously, the following list of examples is only very partial, and many other operators should be considered as well. This material is standard and can be found for example in the books [Am] and [Te].

Recall firstly that operators of multiplication have been introduced in Exercise 3.2.6 and that some of their properties have been studied in Exercises 3.2.14, 3.3.3, 3.6.2, and 3.6.7. We summarize below the main information about these operators, with an emphasize when they are self-adjoint. Note that several notations are used for the multiplication operator defined by a function φ, as for example M_φ, $\varphi(\mathrm{X})$, or $\varphi(\mathrm{Q})$.

For any measurable complex function φ on $\mathbb{R}^n$, the multiplication operator $\varphi(\mathrm{X})$ on $L^2(\mathbb{R}^n)$ is defined by

$$[\varphi(\mathrm{X})f](X) = \varphi(X)f(X) \qquad \forall X \in \mathbb{R}^n$$

for any

$$f \in \mathcal{D}(\varphi(\mathrm{X})) := \Big\{g \in L^2(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} |\varphi(X)|^2 |g(X)|^2 \mathrm{d}X < \infty\Big\}.$$

Clearly, $\varphi(\mathrm{X})$ belongs to $\mathcal{B}(L^2(\mathbb{R}^n))$ if and only if $\varphi \in L^\infty(\mathbb{R}^n)$, and in this case $\mathcal{D}(\varphi(\mathrm{X})) = L^2(\mathbb{R}^n)$ and $\|\varphi(\mathrm{X})\| = \|\varphi\|_\infty$, see Definition 2.6.14 for the space $L^\infty(\mathbb{R}^n)$ and for the norm $\|\cdot\|_\infty$.

If $\varphi \in L^\infty(\mathbb{R}^n)$, one easily observes that $\varphi(\mathrm{X})^* = \overline{\varphi}(\mathrm{X})$, and thus $\varphi(\mathrm{X})$ is self-adjoint if and only if φ is a real function. If φ is real but does not belong to $L^\infty(\mathbb{R}^n)$, one can show that $(\varphi(\mathrm{X}), \mathcal{D}(\varphi(\mathrm{X})))$ defines a self-adjoint operator in $L^2(\mathbb{R}^n)$, see also [Pe, Example 5.1.15]. For example, for any $j \in \{1, \dots, n\}$ the operator X_j defined by $[\mathrm{X}_j f](X) = x_j f(X)$ is a self-adjoint operator with domain $\mathcal{D}(\mathrm{X}_j)$. Note that the n operators $(\mathrm{X}_1, \dots, \mathrm{X}_n)$ are often referred to as the *position operators* in $L^2(\mathbb{R}^n)$. More generally, for any multi-index $\alpha \in \mathbb{N}^n$ one also sets

$$\mathrm{X}^\alpha = \mathrm{X}_1^{\alpha_1} \dots \mathrm{X}_n^{\alpha_n}$$

and this expression defines a self-adjoint operator on its natural domain. Other useful multiplication operators are defined for any $s > 0$ by the functions

$$\mathbb{R}^n \ni X \mapsto \langle X \rangle^s := \Big(1 + \sum_{j=1}^n x_j^2\Big)^{s/2} \in \mathbb{R}.$$

The corresponding operators $(\langle \mathrm{X} \rangle^s, \mathcal{H}_s(\mathbb{R}^n))$, with

$$\mathcal{H}_s(\mathbb{R}^n) := \Big\{f \in L^2(\mathbb{R}^n) \mid \langle \mathrm{X} \rangle^s f \in L^2(\mathbb{R}^n)\Big\} = \Big\{f \in L^2(\mathbb{R}^n) \mid \int_{\mathbb{R}^n} \langle X \rangle^{2s} |f(X)|^2 \mathrm{d}X < \infty\Big\},$$

are again self-adjoint operators on $L^2(\mathbb{R}^n)$. Note that one usually calls $\mathcal{H}_s(\mathbb{R}^n)$ *the weighted Hilbert space with weight s* since it is naturally a Hilbert space with the scalar product $\langle f, g \rangle_s := \int_{\mathbb{R}^n} \overline{f(X)} g(X) \langle X \rangle^{2s} \mathrm{d}X$.

We shall now introduce a new type of operators on $L^2(\mathbb{R}^n)$, the *differential operators*. For that purpose, recall that the Fourier transform has been introduced in (1.5.1), namely

$$[\mathcal{F} f](\xi) \equiv \hat{f}(\xi) := \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \mathrm{e}^{-i\xi \cdot X} f(X)\, \mathrm{d}X,$$

for any $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ and $\xi \in \mathbb{R}^n$. In addition, it has been mentioned in Theorem 1.5.2 that $\mathcal{F}$ extends to a bijective map from $L^2(\mathbb{R}^n)$ to itself. In fact, $\mathcal{F}$ is a unitary operator, as defined in Definition 3.3.7.

Let us use again the multi-index notation and set for any $\alpha \in \mathbb{N}^n$

$$(-i\partial)^\alpha = (-i\partial_1)^{\alpha_1} \dots (-i\partial_n)^{\alpha_n} = (-i)^{|\alpha|} \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n}$$

with $|\alpha| = \alpha_1 + \cdots + \alpha_n$. With this notation at hand, and as already mentioned in (1.5.4) the following relations hold for any f in the Schwartz space $\mathcal{S}(\mathbb{R}^n)$ and any $\alpha \in \mathbb{N}^n$:

$$\mathcal{F}(-i\partial)^\alpha f = \mathrm{X}^\alpha \mathcal{F} f,$$

or equivalently $(-i\partial)^\alpha f = \mathcal{F}^* \mathrm{X}^\alpha \mathcal{F} f$. Thus, keeping these relations in mind, one defines for any for any $\alpha \in \mathbb{N}^n$ the operator $\big((-i\partial)^\alpha, D((-i\partial)^\alpha)\big)$, where

$$f \in \mathcal{D}((-i\partial)^\alpha) \Longleftrightarrow \mathcal{F} f \in \mathcal{D}(\mathrm{X}^\alpha) \Longleftrightarrow f \in \mathcal{F}^* \mathcal{D}(\mathrm{X}^\alpha).$$

In particular, for $j \in \{1, \dots, n\}$ we can consider the self-adjoint operator $\mathrm{D}_j := \mathcal{F}^* \mathrm{X}_j \mathcal{F}$ with domain $\mathcal{F}^* \mathcal{D}(\mathrm{X}_j)$. Similarly, for any $s > 0$, one also defines the operator $\langle \mathrm{D} \rangle^s := \mathcal{F}^* \langle \mathrm{X} \rangle^s \mathcal{F}$ with domain

$$\mathcal{H}^s(\mathbb{R}^n) := \Big\{f \in L^2(\mathbb{R}^n) \mid \langle \mathrm{X} \rangle^s \mathcal{F} f \in L^2(\mathbb{R}^n)\Big\} \equiv \Big\{f \in L^2(\mathbb{R}^n) \mid \langle \mathrm{X} \rangle^s \hat{f} \in L^2(\mathbb{R}^n)\Big\}. \tag{3.7.1}$$

Note that the space $\mathcal{H}^s(\mathbb{R}^n)$ is called *the Sobolev space of order s*, and the operators $(\mathrm{D}_1, \dots, \mathrm{D}_n)$ are usually called *the momentum operators*[2].

[2]In physics textbooks, the position operators are often denoted by $(\mathrm{Q}_1, \dots, \mathrm{Q}_n)$ while $(\mathrm{P}_1, \dots, \mathrm{P}_n)$ denote the momentum operators.

Exercise 3.7.1. *Show that the following relations hold on the Schwartz space* $\mathcal{S}(\mathbb{R}^n)$*:* $[i\mathrm{X}_j, \mathrm{X}_k] = \boldsymbol{0} = [\mathrm{D}_j, \mathrm{D}_k]$ *for any* $j, k \in \{1, \dots, d\}$ *while* $[i\mathrm{D}_j, \mathrm{X}_k] = \boldsymbol{1}\delta_{jk}$*.*

We can now introduce the usual *Laplace operator* $-\Delta$ acting on any $f \in \mathcal{S}(\mathbb{R}^n)$ as

$$-\Delta f = -\sum_{j=1}^{n} \partial_j^2 f = \sum_{j=1}^{n} (-i\partial_j)^2 f = \sum_{j=1}^{n} \mathrm{D}_j^2 f. \tag{3.7.2}$$

This operator admits a self-adjoint extension with domain $\mathcal{H}^2(\mathbb{R}^n)$, *i.e.* $(-\Delta, \mathcal{H}^2(\mathbb{R}^n))$ is a self-adjoint operator in $L^2(\mathbb{R}^n)$. However, let us stress that the expression (3.7.2) is not valid (pointwise) on all the elements of the domain $\mathcal{H}^2(\mathbb{R}^n)$. On the other hand, one has $-\Delta = \mathcal{F}^*\mathrm{X}^2\mathcal{F}$, with $\mathrm{X}^2 = \sum_{j=1}^n \mathrm{X}_j^2$, from which one easily infers that $\sigma(-\Delta) = [0, \infty)$. Indeed, this follows from the content of Exercise 3.6.2 together with the invariance of the spectrum through the conjugation by a unitary operator.

More generally, for any measurable function φ on $\mathbb{R}^n$ let us now set $\varphi(\mathrm{D}) \equiv \varphi(\mathrm{P}) := \mathcal{F}^*\varphi(\mathrm{X})\mathcal{F}$, with domain $\mathcal{D}(\varphi(\mathrm{D})) = \{f \in L^2(\mathbb{R}^n) \mid \hat{f} \in \mathcal{D}(\varphi(\mathrm{X}))\}$, and as before this operator is self-adjoint in $L^2(\mathbb{R}^n)$ if φ is real valued. Then, if one defines the convolution of two (suitable) functions on $\mathbb{R}^n$ as in (1.5.2), namely

$$[k * f](X) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^n} k(Y) f(X - Y)\,\mathrm{d}Y$$

and if one takes the equality $\check{g} * f = \mathcal{F}^*(g\hat{f})$ into account, one infers that the operator $\varphi(\mathrm{D})$ corresponds to *a convolution operator*, namely

$$\varphi(\mathrm{D}) f = \check{\varphi} * f. \tag{3.7.3}$$

Obviously, the meaning of such an equality depends on the class of functions f and φ considered.

Some examples of functions φ (called h subsequently) which are often considered in the literature are the functions defined by $h(\xi) = \xi^2$, $h(\xi) = |\xi|$ or $h(\xi) = \sqrt{1 + \xi^2} - 1$. In these cases, the operator $h(\mathrm{D}) = -\Delta$ corresponds to *the free Laplace operator*, the operator $h(\mathrm{D}) = |\mathrm{D}|$ is *the relativistic Schrödinger operator without mass*, while the operator $h(\mathrm{D}) = \sqrt{1 - \Delta} - 1$ corresponds to *the relativistic Schrödinger operator with mass*. In these three cases, one has $\sigma(h(\mathrm{D})) = [0, \infty)$ while $\sigma_p(h(\mathrm{D})) = \emptyset$, see Definition 3.6.5.

3.8 Spectral theorem

In this section, we provide one formulation of the spectral theorem. It is the generalization of the diagonalization of Hermitian matrices. We firstly consider non-decreasing functions defined on $\mathbb{R}$ taking values in the set $\mathcal{P}(\mathcal{H})$ of orthogonal projections on a Hilbert space $\mathcal{H}$. Such functions were introduced in Section 3.4.

Definition 3.8.1 (Spectral family)**.** A spectral family*, or* a resolution of the identity*, is a family* $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ *of orthogonal projections in* $\mathcal{H}$ *satisfying:*

(i) The family is non-decreasing, i.e. $E_\lambda E_\mu = E_{\min\{\lambda,\mu\}}$,

(ii) The family is strongly right continuous, i.e. $E_\lambda = E_{\lambda+0} = s - \lim_{\varepsilon\searrow 0} E_{\lambda+\varepsilon}$,

(iii) $s - \lim_{\lambda\to-\infty} E_\lambda = \boldsymbol{0}$ *and* $s - \lim_{\lambda\to\infty} E_\lambda = \boldsymbol{1}$*,*

It is important to observe that the condition (i) implies that the elements of the families are *commuting*, *i.e.* $E_\lambda E_\mu = E_\mu E_\lambda$. We also define *the support of the spectral family Spectral support* as the following subset of $\mathbb{R}$:

$$\mathrm{supp}\{E_\lambda\} = \Big\{\mu \in \mathbb{R} \mid E_{\mu+\varepsilon} - E_{\mu-\varepsilon} \neq \mathbf{0},\ \forall \varepsilon > 0\Big\}.$$

Given such a family, one firstly defines

$$E((a,b]) := E_b - E_a, \qquad a,b \in \mathbb{R}, \tag{3.8.1}$$

and extends this definition to all sets $V \in \mathcal{A}_B$ [3]. As a consequence of the construction, note that

$$E\Big(\bigcup_k V_k\Big) = \sum_k E(V_k) \tag{3.8.2}$$

whenever $\{V_k\}$ is a countable family of disjoint elements of $\mathcal{A}_B$. Thus, one ends up with a projection-valued map $E : \mathcal{A}_B \to \mathcal{P}(\mathcal{H})$ which satisfies $E(\emptyset) = \mathbf{0}$, $E(\mathbb{R}) = \mathbf{1}$, $E(V_1)E(V_2) = E(V_1 \cap V_2)$ for any Borel sets V_1, V_2. In addition,

$$E((a,b)) = E_{b-0} - E_a, \qquad E([a,b]) = E_b - E_{a-0},$$

where $E_{\lambda-0} := \lim_{\varepsilon\searrow 0} E_{\lambda-\varepsilon}$, and therefore $E(\{a\}) = E_a - E_{a-0}$.

Definition 3.8.2 (Spectral measure)**.** *The map $E : \mathcal{A}_B \to \mathcal{P}(\mathcal{H})$ defined by* (3.8.1) *is called* the spectral measure *associated with the family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$. This spectral measure is* bounded from below *if there exists $\lambda_- \in \mathbb{R}$ such that $E_\lambda = \mathbf{0}$ for all $\lambda < \lambda_-$. Similarly, this spectral measure is* bounded from above *if there exists $\lambda_+ \in \mathbb{R}$ such that $E_\lambda = \mathbf{1}$ for all $\lambda > \lambda_+$.*

Let us note that for any spectral family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ and any $f \in \mathcal{H}$ one can associate a *measure*[4] m_f on $\mathbb{R}$ which satisfies

$$m_f(V) = \|E(V)f\|^2 = \langle E(V)f, f\rangle \tag{3.8.3}$$

for any $V \in \mathcal{A}_B$. Note in particular that $m_f(\mathbb{R}) = \|f\|^2$.

Our next aim is to define integrals of the form

$$\int_a^b \varphi(\lambda)\, E(\mathrm{d}\lambda) \tag{3.8.4}$$

for a continuous function $\varphi : [a,b] \to \mathbb{C}$ and for any spectral family $\{E_\lambda\}_{\lambda\in\mathbb{R}}$. Such integrals can be defined by considering a partition $a = x_0 < x_1 < ... < x_n = b$ of $[a,b]$ and a collection $\{y_j\}$ with $y_j \in (x_{j-1}, x_j)$ and by defining the operator

$$\sum_{j=1}^n \varphi(y_j) E((x_{j-1}, x_j]). \tag{3.8.5}$$

It turns out that by considering finer and finer partitions of $[a,b]$, the corresponding expression (3.8.5) strongly converges to an element of $\mathcal{B}(\mathcal{H})$ which is independent of the successive choice of partitions. The resulting operator is denoted by (3.8.4).

The following statement contains usual results which can be obtained in this context. The proof is not difficult, but one has to deal with several partitions of intervals. We refer to [Am, Prop. 4.10] for a detailed proof.

Proposition 3.8.3 (Spectral integrals)**.** *Let $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ be a spectral family, let $-\infty < a < b < \infty$ and let $\varphi : [a,b] \to \mathbb{C}$ be continuous. Then one has*

[3] We denote by $\mathcal{A}_B$ the set of all *Borel sets* on $\mathbb{R}$, also called the *Borel algebra*. It consists of the collection of all sets obtained from the open sets (or equivalently from closed sets) through the operations of countable union, countable intersection, and relative complement.

[4] A *measure* on $\mathbb{R}$ is a map $m : \mathcal{A}_B \to [0,\infty]$ satisfying the following $m(\emptyset) = 0$ and the *countable additivity* property, namely if $\{V_j\}_{j\in\mathbb{N}} \subset \mathcal{A}_B$ are pairwise disjoint, then $m(\sum_j V_j) = \sum_j m(V_j)$.

1. $\left\| \int_a^b \varphi(\lambda)\, E(\mathrm{d}\lambda) \right\| = \sup_{\mu \in [a,b] \cap \mathrm{supp}\{E_\lambda\}} |\varphi(\mu)|$,

2. $\left(\int_a^b \varphi(\lambda)\, E(\mathrm{d}\lambda) \right)^* = \int_a^b \overline{\varphi}(\lambda)\, E(\mathrm{d}\lambda)$,

3. For any $f \in \mathcal{H}$, $\left\| \int_a^b \varphi(\lambda)\, E(\mathrm{d}\lambda) f \right\|^2 = \int_a^b |\varphi(\lambda)|^2\, m_f(\mathrm{d}\lambda)$,

4. If $\psi : [a,b] \to \mathbb{C}$ *is continuous, then*

$$\int_a^b \varphi(\lambda)\, E(\mathrm{d}\lambda) \cdot \int_a^b \psi(\lambda)\, E(\mathrm{d}\lambda) = \int_a^b \varphi(\lambda)\psi(\lambda)\, E(\mathrm{d}\lambda).$$

Let us now observe that if the support $\mathrm{supp}\{E_\lambda\}$ is bounded, then one can consider

$$\int_{-\infty}^{\infty} \varphi(\lambda)\, E(\mathrm{d}\lambda) = s - \lim_{M\to\infty} \int_{-M}^{M} \varphi(\lambda)\, E(\mathrm{d}\lambda). \tag{3.8.6}$$

Similarly, by taking property (iii) of the previous proposition into account, one observes that this limit can also be taken if $\varphi \in C_b(\mathbb{R})$. On the other hand, if φ is not bounded on $\mathbb{R}$, the r.h.s. of (3.8.6) is not necessarily well defined. In fact, if φ is not bounded on $\mathbb{R}$ and if $\mathrm{supp}\{E_\lambda\}$ is not bounded either, then the r.h.s. of (3.8.6) is an unbounded operator and can only be defined on a dense domain of $\mathcal{H}$.

Lemma 3.8.4. *Let* $\varphi : \mathbb{R} \to \mathbb{C}$ *be continuous, and let us set*

$$\mathcal{D}_\varphi := \Big\{ f \in \mathcal{H} \mid \int_{-\infty}^{\infty} |\varphi(\lambda)|^2\, m_f(\mathrm{d}\lambda) < \infty \Big\}.$$

Then the pair $\left(\int_{-\infty}^{\infty} \varphi(\lambda)\, E(\mathrm{d}\lambda), \mathcal{D}_\varphi \right)$ *defines a closed linear operator on* $\mathcal{H}$*. This operator is self-adjoint if and only if* φ *is a real function.*

The proof of this statement is not complicated, but use all the material introduced so far. It can be considered as a challenging exercise to provide it.

A function φ of special interest is the function defined by the identity function id, namely $\mathrm{id}(\lambda) = \lambda$.

Definition 3.8.5 (Self-adjoint operator associated with a spectral family)**.** *For any spectral family* $\{E_\lambda\}_{\lambda\in\mathbb{R}}$*, the operator* $\left(\int_{-\infty}^{\infty} \lambda\, E(\mathrm{d}\lambda), \mathcal{D}_{\mathrm{id}} \right)$ *with*

$$\mathcal{D}_{\mathrm{id}} := \Big\{ f \in \mathcal{H} \mid \int_{-\infty}^{\infty} \lambda^2\, m_f(\mathrm{d}\lambda) < \infty \Big\}$$

is called the self-adjoint operator associated with $\{E_\lambda\}$.

By this procedure, any spectral family defines a self-adjoint operator on $\mathcal{H}$. *The spectral Theorem* corresponds to the converse statement:

Theorem 3.8.6 (Spectral Theorem)**.** *With any self-adjoint operator* $(A, \mathcal{D}(A))$ *on a Hilbert space* $\mathcal{H}$ *one can associate a unique spectral family* $\{E_\lambda\}$*, called* the spectral family of A*, such that* $\mathcal{D}(A) = \mathcal{D}_{\mathrm{id}}$ *and* $A = \int_{-\infty}^{\infty} \lambda\, E(\mathrm{d}\lambda)$.

In summary, there is a bijective correspondence between self-adjoint operators and spectral families. This theorem extends the fact that any $n \times n$ hermitian matrix is diagonalizable. The proof of this theorem is not trivial and is rather lengthy. In the sequel, we shall assume it, and state various consequences of this theorem.

Based on this one-to-one correspondence it is now natural to set the following definition:

Definition 3.8.7 (Bounded functional calculus)**.** *Let A be a self-adjoint operator in $\mathcal{H}$ and $\{E_\lambda\}$ be the corresponding spectral family. For any bounded and continuous function $\varphi : \mathbb{R} \to \mathbb{C}$ one sets $\varphi(A) \in \mathcal{B}(\mathcal{H})$ for the operator defined by*

$$\varphi(A) := \int_{-\infty}^{\infty} \varphi(\lambda)\, E(\mathrm{d}\lambda). \tag{3.8.7}$$

For the next statement, we set $C_b(\mathbb{R})$ for the set of all continuous and bounded complex functions on $\mathbb{R}$.

Proposition 3.8.8. *a) For any $\varphi \in C_b(\mathbb{R})$ one has*

(i) $\varphi(A) \in \mathcal{B}(\mathcal{H})$ and $\|\varphi(A)\| = \sup_{\lambda\in\sigma(A)} |\varphi(\lambda)|$,

(ii) $\varphi(A)^* = \overline{\varphi}(A)$, and $\varphi(A)$ is self-adjoint if and only if φ is real,

(iii) $\varphi(A)$ is unitary if and only if $|\varphi(\lambda)| = 1$.

b) If $\varphi \in C(\mathbb{R})$, then (3.8.7) *defines a closed operator $\varphi(A)$ with domain*

$$\mathcal{D}(\varphi(A)) = \{f \in \mathcal{H} \mid \int_{-\infty}^{\infty} |\varphi(\lambda)|^2\, m_f(\mathrm{d}\lambda) < \infty\}. \tag{3.8.8}$$

In the point (iii) above, one can consider the function $\varphi_t \in C_b(\mathbb{R})$ defined by $\varphi_t(\lambda) := \mathrm{e}^{-i\lambda t}$ for any fixed $t \in \mathbb{R}$. Then, if one sets $U_t := \varphi_t(A)$ one first observes that $U_t U_s = U_{t+s}$. Indeed, one has

$$\begin{aligned} U_t U_s &= \int_{-\infty}^{\infty} \mathrm{e}^{i\lambda t} E(\mathrm{d}\lambda) \int_{-\infty}^{\infty} \mathrm{e}^{-i\lambda s} E(\mathrm{d}\lambda) = \int_{-\infty}^{\infty} \mathrm{e}^{-i\lambda t}\mathrm{e}^{-i\lambda s} E(\mathrm{d}\lambda) \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{-i\lambda(t+s)} E(\mathrm{d}\lambda) = U_{t+s}. \end{aligned}$$

In addition, by an application of the dominated convergence theorem 2.5.2, one infers that the map $\mathbb{R} \ni t \mapsto U_t \in \mathcal{B}(\mathcal{H})$ is strongly continuous. Indeed, since $\left|\mathrm{e}^{-i\lambda(t+\varepsilon)} - \mathrm{e}^{-i\lambda t}\right|^2 \leq 4$ one has

$$\|U_{t+\varepsilon} f - U_t f\|^2 = \int_{-\infty}^{\infty} \left|\mathrm{e}^{-i\lambda(t+\varepsilon)} - \mathrm{e}^{-i\lambda t}\right|^2 m_f(\mathrm{d}\lambda) \ \to\ 0 \quad \text{as} \quad \varepsilon \to 0.$$

As a consequence, such a family $\{U_t\}_{t\in\mathbb{R}}$ is called *a strongly continuous unitary group*. Note that since $\mathrm{e}^{-i\lambda t} = \sum_{k=0}^{\infty} \frac{(-i\lambda t)^k}{k!}$ one also infers that whenever A is a bounded operator

$$U_t = \sum_{k=0}^{\infty} \frac{(-itA)^k}{k!} \tag{3.8.9}$$

with a norm converging series. On the other hand, if A is not bounded, then this series converges on elements $f \in \cap_{k=0}^{\infty} \mathcal{D}(A^k)$. In particular, it converges strongly on elements of $\mathcal{H}$ which have bounded support with respect to the corresponding spectral measure.

Let us now mention that the above construction is only one part of a one-to-one relation between strongly continuous unitary groups and self-adjoint operators. The proof of the following statement can be found for example in [Am, Prop. 5.1].

Theorem 3.8.9 (Stone's Theorem)**.** *There exists a bijective correspondence between self-adjoint operators on $\mathcal{H}$ and strongly continuous unitary groups on $\mathcal{H}$. More precisely, if A is a self-adjoint operator on $\mathcal{H}$, then*

$\{e^{-itA}\}_{t\in\mathbb{R}}$ is a strongly continuous unitary group, while if $\{U_t\}_{t\in\mathbb{R}}$ is a strongly continuous unitary group, one sets

$$\mathcal{D}(A) := \Big\{f \in \mathcal{H} \mid \exists\ s - \lim_{t\to 0} \frac{1}{t}[U_t - 1]f\Big\}$$

and for $f \in \mathcal{D}(A)$ one sets $Af = s - \lim_{t\to 0} \frac{i}{t}[U_t - 1]f$, and then $(A, \mathcal{D}(A))$ is a self-adjoint operator.

Let us close with section with two important observations. First of all, the map $\varphi \mapsto \varphi(A)$ can be extended from continuous and bounded φ to bounded and measurable functions φ. This extension can be realized by considering integrals in the weak form. In particular, this extension is necessary for defining $\varphi(A)$ whenever φ is the characteristic function on some Borel set V.

The second observation is going to provide an alternative formula for $\varphi(A)$ in terms of the unitary group $\{e^{-itA}\}_{t\in\mathbb{R}}$. Indeed, assume that the inverse Fourier transform $\check{\varphi}$ of φ belongs to $L^1(\mathbb{R})$, then the following equality holds

$$\varphi(A) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \check{\varphi}(t)\, e^{-itA}\, dt. \tag{3.8.10}$$

Indeed, observe that

$$\langle f, \varphi(A)f\rangle = \int_{\mathbb{R}} \varphi(\lambda) m_f(d\lambda) = \int_{\mathbb{R}} m_f(d\lambda) \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\lambda t}\check{\varphi}(t) dt.$$

By application of Fubini's theorem 2.6.1 one can interchange the order of integrations and obtain

$$\begin{aligned}\langle f, \varphi(A)f\rangle =& \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dt\check{\varphi}(t) \int_{\mathbb{R}} e^{-i\lambda t} m_f(d\lambda)\\ =& \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dt\,\check{\varphi}(t)\langle f, e^{-itA} f\rangle = \Big\langle f, \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} dt\,\check{\varphi}(t) e^{-itA} f\Big\rangle,\end{aligned}$$

and one gets (3.8.10) by applying the polarisation identity (3.1.1).

The theories we have developed in the three chapters of this course have applications in several fields, like in physics, in engineering, in quantum mechanics, in quantum field theory, in chemistry, *etc*. What about applications in your domain of interest ?

Index

C^∞, 3
L^1, 26
L^∞, 28
L^p, 27
$\mathcal{B}(\mathcal{H})$, 35
$\mathcal{D}'(\mathbb{R}^n)$, 4
$\mathcal{D}(\mathbb{R}^n)$, 3
$\mathcal{H}$, 31
$\mathcal{S}'(\mathbb{R}^n)$, 12
$\mathcal{S}(\mathbb{R}^n)$, 11
α-derivative, 6
$\rho(A)$, 43
$\sigma(A)$, 43
$\sigma_p(A)$, 44

a.e., 19
Adjoint, 35, 42
Algebra, 4
Almost everywhere, 19
Anti-linear, 34

Banach space, 26
Borel set, 48
Boundary, 3
Bounded from above, 48
Bounded from below, 48
Bounded functional calculus, 50
Bounded linear operator, 35
Boundedly invertible, 36, 43
Bra-ket notation, 39

Cauchy sequence, 26
Cesaro summation, 9
Characteristic function, 18
Closed box, 16
Closed operator, 42
Closed set, 3
Closed subspace, 33
Closure, 3, 42
Compact operator, 39
Completeness, 26
Continuous in norm, 40
Continuous unitary group, 50
Convergence of distributions, 8
Convergence of test functions, 4
Convolution operator, 47
Countable additivity, 48
Covering, 16

Dense, 33
Differential operator, 46
Differentiation of a distribution, 6
Dirac delta distribution, 8
Dirichlet kernel, 9
Distribution, 4
Domain, 41
Dominated convergence theorem, 24
Dominating function, 24

Eigenfunction, 44
Eigenvalue, 44
Eigenvector, 44
Equivalence relation, 26
Essential infimum, 28
Essential supremum, 28
extension, 42

Fejér kernel, 9
Final set projection, 39
Finite rank operator, 39
First resolvent equation, 44
Fourier transform, 10
Fubini's theorem, 25

Hölder inequality, 27
Heaviside function, 6

Hermitian, 37
Hilbert space, 31

Ideal, 4
In the sense of distribution, 6
Initial set projection, 39
Interior, 3
Invertible, 36, 43
Isometry, 38

Laplace operator, 47
Lebesgue integrable, 20
Lebesgue measurable function, 18
Lebesgue measurable partition, 20
Lebesgue outer measure, 16
Lesbesgue measurable, 17
Lesbesgue measure, 17
Lim inf, 19
Lim sup, 19
Linear operator, 41
Locally integrable, 5
Lower Riemann integra, 15
Lower Riemann sum, 14

Measure, 48
Minkowski's inequality, 27
Momentum operator, 46
Multi-index, 2
Multiplication operator, 35–37, 44, 45
Multiplicity, 44

Neumann series, 37
Norm, 26, 35

Open set, 3
Operator valued function, 40
Order of a distribution, 5
Orthocomplement, 33
Orthogonal, 33
Orthogonal projection, 37

Parseval theorem, 11
Parseval-Plancherel identity, 11
Partial derivative, 2
Partial isometry, 39
Pointwise bounded, 19
Polarisation identity, 31
Position operator, 46
Principal value distribution, 7
Projection, 37

Reflexivity, 26
Regular distribution, 5
Regular partition, 14
Resolvent, 43
Resolvent set, 43
restriction, 42
Riemann integral, 15

Scalar product, 28
Schrödinger operator, 47
Schwartz function, 11
Schwartz space, 11
Schwarz inequality, 32
Self-adjoint, 37, 43
Separable, 31
Simple function, 18
Smooth function, 3
Sobolev space, 46
Spectral family, 47
Spectral measure, 48
Spectral theorem, 49
Spectrum, 43
Strong convergence, 32, 36
Strongly continuous, 40
Strongly differentiable, 40
Subspace, 33
Summability order, 9
Summable distribution, 9
Sup-norm, 4
Support of a function, 3
Symmetric, 43
Symmetry, 26

Tempered distribution, 12
Test functions, 3
Transitivity, 26
Triangle inequality, 32

Unbounded operator, 42
Uniform convergence, 36
Unitary operator, 38
Upper Riemann integral, 15
Upper Riemann sum, 14

Vector space, 4

Vector valued function, 40

Weak convergence, 32, 36
Weakly continuous, 40
Weighted Hilbert space, 46

Bibliography

[Am] W. Amrein, *Hilbert space methods in quantum mechanics*, Fundamental Sciences, EPFL Press, Lausanne, distributed by CRC Press, Boca Raton, FL, 2009.

[Di] G. van Dijk, *Distribution theory: Convolution, Fourier transform, and Laplace transform*, De Gruyter Graduate Lectures, De Gruyter, Berlin, 2013.

[Mu] G.J. Murphy, *C*-algebras and operator theory*, Academic Press, Inc., Boston, MA, 1990.

[Ne] G. Nelson, *A user-friendly introduction to Lebesgue measure and integration*, Student Mathematical Library 78, American Mathematical Society, Providence, RI, 2015.

[Pe] G. Pedersen, *Analysis now*, Graduate texts in mathematics 118, Springer, 1989.

[Te] G. Teschl, *Mathematical methods in quantum mechanics, with applications to Schrödinger operators*, Graduate Studies in Mathematics 99, American Mathematical Society, Providence, RI, 2009.