Mathematical methods in Engineering

Davide Guidetti

Chapter 1

Preliminaries

In this chapter we collect some results which are preliminary to the main subjects of the course and probably the student already knows, possibly in a little bit different form. However, it may be useful to have at disposal an easily reachable reference, and this is what we are aiming at.

1.1 The Lebesgue measure in $\mathbb{R}^n$

In this section we want to give the rudiments, concerning the theory of measure and integration in the sense of Lebesgue.

As a subset of $\mathbb{R}^n$ may have infinite measure, it is convenient to extend the standard operations of sum and product between real numbers to the case that a summand or a factor is $+\infty$. So we set

$$[0, +\infty] := [0, +\infty[\cup \{+\infty\}, \tag{1.1.1}$$

where $+\infty$ is an object which does not belong to $[0, +\infty[$. We suppose that

$$a < +\infty \qquad \forall a \in [0, +\infty[. \tag{1.1.2}$$

Next, we set, given $a \in [0, +\infty]$,

$$a + (+\infty) = (+\infty) + a = +\infty \tag{1.1.3}$$

and

$$a \cdot (+\infty) = (+\infty) \cdot a = \begin{cases} +\infty & \text{if} \quad a \neq 0, \\ 0 & \text{if} \quad a = 0. \end{cases} \tag{1.1.4}$$

If $(a_n)_{n \in \mathbb{N}}$ is a sequence with values in $[0, +\infty]$, we define the sum of the series

$$\sum_{n=1}^{\infty} a_n = \begin{cases} \text{usual if the series converges in } \mathbb{R}, \\ +\infty \qquad \text{otherwise.} \end{cases} \tag{1.1.5}$$

We shall sometimes employ also the symbol $-\infty$, and we shall assume it as satisfying

$$-\infty < a \qquad \forall a \in \mathbb{R}. \tag{1.1.6}$$

We set

$$[-\infty, +\infty] := \mathbb{R} \cup \{-\infty, +\infty\}. \tag{1.1.7}$$

Now we construct the Lebesgue measure starting from bounded intervals in $\mathbb{R}^n$, the $n-$dimensional volume of which vol_n we are going to define.

Definition 1.1.1. *A bounded interval in $\mathbb{R}^n$ is the cartesian product of n bounded intervals in $\mathbb{R}$.*

If I is a bounded interval in $\mathbb{R}$, we set

$$\ell(I) := \begin{cases} \sup I - \inf I & \text{if} \quad I \neq \emptyset, \\ 0 & \text{if} \quad I = \emptyset. \end{cases} \tag{1.1.8}$$

If I is a bounded interval in $\mathbb{R}^n$ and $I = I_1 \times ... \times I_n$, we set

$$vol_n(I) := \ell(I_1) \cdot .... \cdot \ell(I_n). \tag{1.1.9}$$

Now we define the outer measure of an arbitrary subset of $\mathbb{R}^n$.

Definition 1.1.2. *Let $A \subseteq \mathbb{R}^n$. We define the outer maesure $L_n^*(A)$ di A as follows:*

$$\begin{aligned} L_n^*(A) := \inf\{\textstyle\sum_{k=1}^{\infty} vol_n(I^k) : I^k n- \text{ dimensional interval } \forall k \in \mathbb{N}, \\ A \subseteq \cup_{k\in\mathbb{N}} I^k\}. \end{aligned} \tag{1.1.10}$$

We observe that, in case $L_n^*(A) = +\infty$, for any countable family $(I^k)_{k\in\mathbb{N}}$ of $n-$dimensional bounded intervals which "covers" A (in the sense that $A \subseteq \cup_{k\in\mathbb{N}} I^k$), it holds $\displaystyle\sum_{k=1}^{\infty} vol_n(I^k) = +\infty$.

Remark 1.1.3. Let A be a set (the elements of which may have arbitrary nature). We shall say that A is **countable** if there exists a bijection $\phi : \mathbb{N} \to A$. This means, in few words, that the elements of A can be listed in a sequence $\{\phi(1), \phi(2), ...\}$. $\mathbb{N}$ is, obviously, countable. It is not difficult to verify that even $\mathbb{Z}$ (the set of integer numbers) is countable: we can define $\phi : \mathbb{N} \to \mathbb{Z}$, $\phi(1) = 0$, $\phi(2) = 1$, $\phi(3) = -1$, and so on. Although it is by no means obvious, one could show that the set $\mathbb{Q}$ of rational numbers is countable.

On the other hand, if a and b are real numbers and $a < b$, the set $]a, b[$ is not countable. More precisely, it does not exist a surjective function from $\mathbb{N}$ to $]a, b[$. So, in some sense, $]a, b[$ has "more elements" than $\mathbb{N}$.

We recall that a set A is **finite** if it is empty, or there exist $n \in \mathbb{N}$ and $\phi : \{k \in \mathbb{N} : k \leq n\} \to A$, which is a bijection. One could show that the natural number n is uniquely determined by A and is called the **cardinality** of A. A is said to be **infinite** if it is not finite. Countable sets are infinite, but they are, in some sense, the smallest infinite sets, in the sense that every infinite set contains a countable one.

The outer measure L_n^* is, to some extent, unsatisfactory. In fact, it would be desirable that, given two arbitrary and disjoint subsets A and B in $\mathbb{R}^n$, it held

$$L_n^*(A \cup B) = L_n^*(A) + L_n^*(B).$$

and this does not happen in general. We can only say that the inequality

$$L_n^*(A \cup B) \leq L_n^*(A) + L_n^*(B),$$

is always true, but, even in the case that A and B are disjoint, it may happen that the strict inequality holds. In order to remedy this inconvenience, we limit ourselves to consider a subclass of subsets of $\mathbb{R}^n$, the so called *sets which are measurable in the sense of Lebesgue*, in which everything works well.

Definition 1.1.4. *Let $A \subseteq \mathbb{R}^n$. We shall say that A is* **measurable** *in the sense of Lebesgue if $\forall E \subseteq \mathbb{R}^n$*

$$L_n^*(E) = L_n^*(E \cap A) + L_n^*(E \cap (\mathbb{R}^n \setminus A)).$$

We shall indicate with $\mathcal{M}_n$ the class of subsets of $\mathbb{R}^n$ which are measurable in the sense of Leabesque. This class of sets is very flexible, with respect to the usual set operations:

Theorem 1.1.5. *(I) $\emptyset$ and $\mathbb{R}^n$ belong to $\mathcal{M}_n$;*

(II) if A and B belong to $\mathcal{M}_n$, then $A \setminus B$ belongs to $\mathcal{M}_n$;

(III) the union of a finite or countable family of elements of $\mathcal{M}_n$ belongs to $\mathcal{M}_n$;

(IV) the intersection of a not empty, finite or countable family of elements of $\mathcal{M}_n$ belongs to $\mathcal{M}_n$.

We define the **Lebesgue measure** L_n as the restriction of L_n^* to $\mathcal{M}_n$. The following facts hold:

Theorem 1.1.6. *Let $\mathcal{I}$ be $\{1, ...n\}$ for some $n \in \mathbb{N}$, or the set of natural numbers $\mathbb{N}$. Then:*

(I) $L_n(\emptyset) = 0$;

(II) if $A_i \in \mathcal{M}_n$ $\forall i \in \mathcal{I}$, one has

$$L_n(\cup_{i\in\mathcal{I}} A_i) \leq \sum_{i\in\mathcal{I}} L_n(A_i);$$

(we observe that $\cup_{i\in\mathcal{I}} A_i$ is misurable by Theorem 1.1.5(III));

(III) if, moreover, the sets A_i are pairwise disjoint, one has

$$L_n(\cup_{i\in\mathcal{I}} A_i) = \sum_{i\in\mathcal{I}} L_n(A_i);$$

(IV) if A and B are elements of $\mathcal{M}_n$ and $A \subseteq B$, the inequality

$$L_n(A) \leq L_n(B)$$

holds;

(V) if , moreover, $L_n(B) < +\infty$, one has

$$L_n(B \setminus A) = L_n(B) - L_n(A).$$

The most interesting result in Theorem 1.1.6 is probably (III): if the sets A_i, pairwise disjoint, are a finite or countable family, the measure of their union coincides with the sum of the measures. Finally, one might wonder whether, given a subset of $\mathbb{R}^n$, there exist criteria to establish if it is measurable. The following result meets this requirement:

Theorem 1.1.7. *(I) If I is a $n-$dimensional interval, then $I \in \mathcal{M}_n$ and $L_n(I) = vol_n(I)$;*

(II) every open subset of $\mathbb{R}^n$ is measurable;

(III) every closed subset if $\mathbb{R}^n$ is measurable.

Exercise 1.1.8. Let A be a finite or countable subset of $\mathbb{R}^n$. Show that A is measurable in the sense of Lebesgue and has measure 0.

Exercise 1.1.9. Let A be a subset of $\mathbb{R}^n$ such that $L_n^*(A) = 0$. Show that A is measurable in the sense of Lebesgue.

1.2 Theory of integration in the sense of Lebesgue

After measure, we pass to consider integration in $\mathbb{R}^n$. We shall begin by integrating the so called nonnegative *simple functions* .

Definition 1.2.1. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$. We shall say that f is a* **simple** *function if there exist $A_1, ..., A_m$ ($m \in \mathbb{N}$) measurable pairwise disjoint subsets of A , the union of which is A, and real numbers $\alpha_1, ..., \alpha_m$ such that,*

$$f(x) = \alpha_i \quad \forall x \in A_i, \quad 1 \le i \le m.$$

Definition 1.2.2. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$ simple according to Definition 1.2.1, with $\alpha_i \ge 0$ $\forall i \in \{1, ..., m\}$. We set*

$$\int_A f(x)dx := \sum_{i=1}^{m} \alpha_i L_n(A_i).$$

One can show that Definition 1.2.2 is well posed, in the sense that it does not depend on the way we decompose A, provided we preserve the property that in each of the parts f is constant. In case $\alpha_i = 0$, by Definition (1.1.4), the product $\alpha_i L_n(A_i)$ gives 0, even when $L_n(A_i) = +\infty$. As a consequence, we obtain that in an arbitrary measurable set, even of infinite measure, the integral of the function identically equal to 0 is 0. In fact, the choice of (1.1.4) was directed towards this.

Now we define a very large class of functions, the so called *measurable functions* .

Definition 1.2.3. *Let $A \in \mathcal{M}_n$, $f : A \to [-\infty, +\infty]$ (see (1.1.7)). We shall say that f is* **measurable** *if there exists a sequence of simple functions $(f_k)_{k \in \mathbb{N}}$ with domain A, such that*

$$\lim_{k \to +\infty} f_k(x) = f(x) \qquad \forall x \in A.$$

It is clear from the definition that every simple function is measurable. However, for example, one can see that

Theorem 1.2.4. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$, continuous. Then f is measurable.*

Even the class of measurable functions is considerably flexible, with respect to standard operations among sets:

Theorem 1.2.5. *Let $A \in \mathcal{M}_n$, $f, g : A \to \mathbb{R}$ measurable, $\phi : \mathbb{R} \to \mathbb{R}$ continuous. Then:*
(I) $f + g$, fg, $\phi \circ f$ are measurable;
(II) if $g(x) \ne 0$ $\forall x \in A$, then $\frac{f}{g}$ is measurable.

From (I), taking $\phi(y) = cy$, it follows, in particular, that, if f is measurable and real valued, the same happens for cf $\forall c \in \mathbb{R}$.

We pass to define the integral of a nonnegative measurable function.

Definition 1.2.6. *Let $A \in \mathcal{M}_n$, $f : A \to [0, +\infty]$, measurable. We set*

$$\int_A f(x)dx$$

$$:= \sup\{\int_A \phi(x)dx : \phi : A \to [0, +\infty[\text{ simple }, 0 \le \phi(x) \le f(x) \quad \forall x \in A\}.$$

One can see that, in case f is simple and nonnegative Definition 1.2.6 coincides with Definition 1.2.2.

Now we state some basic properties of the integral of a nonnegative measurable function.

Theorem 1.2.7. *Let $A \in \mathcal{M}_n$, $f, g : A \to [0, +\infty]$ measurable, $\alpha \in [0, +\infty[$. Then:*

(I) $\int_A f(x)dx \in [0, +\infty]$;

(II) if $f(x) \le g(x)$ $\forall x \in A$, one has also $\int_A f(x)dx \le \int_A g(x)dx$;

(III) $\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx$;

(IV) $\int_A \alpha f(x)dx = \alpha \int_A f(x)dx$;

(V) if $A = A_1 \cup ... \cup A_m$, with A_1,...,A_m measurable and pairwise disjoint, then, for each j, $f_{|A_j}$ is measurable and

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f(x)dx,$$

with $\int_{A_j} f(x)dx := \int_{A_j} f_{|A_j}(x)dx$ (see the following Exercise 1.2.13).

Now we pass to consider the integration of functions with arbitrary sign. We start with the following

Definition 1.2.8. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$. We shall say that f is* **summable** *if it is measurable and*

$$\int_A |f(x)|dx < +\infty$$

(this integral has a meaning by virtue of Theorem 1.2.5 (I)).

We consider now the two functions:

$$\begin{cases} \phi_+ : \mathbb{R} \to \mathbb{R}, \\ \phi_+(x) = \begin{cases} x & \text{if} \quad x \ge 0, \\ 0 & \text{if} \quad x < 0, \end{cases} \end{cases} \tag{1.2.1}$$

$$\begin{cases} \phi_- : \mathbb{R} \to \mathbb{R}, \\ \phi_-(x) = \begin{cases} -x & \text{if} \quad x \le 0, \\ 0 & \text{if} \quad x > 0. \end{cases} \end{cases} \tag{1.2.2}$$

It is easy to verify that ϕ_+ and ϕ_- are continuous and nonnegative, $\phi_+(x) - \phi_-(x) = x$ $\forall x \in \mathbb{R}$, $\phi_+(x) + \phi_-(x) = |x|$ $\forall x \in \mathbb{R}$.

Now let $f : A \to \mathbb{R}$, measurable. We set

$$f_+ := \phi_+ \circ f, \qquad f_- := \phi_- \circ f. \tag{1.2.3}$$

From Theorem 1.2.5 (I) we have that f_+ and f_- are nonnegative and measurable. Moreover,

$$f = f_+ - f_-, \qquad |f| = f_+ + f_-. \tag{1.2.4}$$

From Theorem 1.2.7(III) we deduce that

$$\int_A |f(x)|dx = \int_A f_+(x)dx + \int_A f_-(x)dx. \tag{1.2.5}$$

So, if f is summable,the integrals in the second term of (1.2.5) are both real. Therefore, taking into account the first formula in (1.2.4), the following definition becomes natural:

Definition 1.2.9. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$ summable. We set*

$$\int_A f(x)dx := \int_A f_+(x)dx - \int_A f_-(x)dx.$$

Theorem 1.2.7 admits the following extension to summable functions:

Theorem 1.2.10. *Let $A \in \mathcal{M}_n$, $f, g : A \to \mathbb{R}$ summable, $\alpha \in \mathbb{R}$. Then:*
(I) if $f(x) \leq g(x)$ $\forall x \in A$, one has also $\int_A f(x)dx \leq \int_A g(x)dx$;
(II) $f + g$ is summable and

$$\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx;$$

(III) αf is summable and

$$\int_A \alpha f(x)dx = \alpha \int_A f(x)dx;$$

(IV)if $A = A_1 \cup ... \cup A_m$, with A_1,...,A_m measurable and pairwise disjoint, then $f_{|A_j}$ is summable on A_j for each $j = 1, ..., m$ and

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f_{|A_j}(x)dx \tag{1.2.6}$$

holds; on the other hand, if, $j = 1, ..., m$, $f_{|A_j}$ is summable on A_j, then f is summable on A and (1.2.6) holds.

In many circumstances it is important to have at disposal a definition of integral for complex valued functions.

Definition 1.2.11. *Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{C}$. We shall say that f is* **summable** *if the two real valued functions $Re(f)$ e $Im(f)$ are summable. In this case, we set*

$$\int_A f(x)dx := \int_A Re(f(x))dx + i \int_A Im(f(x))dx.$$

A part of Theorem 1.2.10 can be extended to complex valued functions:

Theorem 1.2.12. *Let $A \in \mathcal{M}_n$, $f, g : A \to \mathbb{C}$ summable, $\alpha \in \mathbb{C}$. Then:*
(I) $f + g$ is summable and

$$\int_A (f(x) + g(x))dx = \int_A f(x)dx + \int_A g(x)dx;$$

(II) αf is summable and

$$\int_A \alpha f(x)dx = \alpha \int_A f(x)dx;$$

(III) if $A = A_1 \cup ... \cup A_m$, with A_1,...,A_m measurable and pairwise disjoint, then $f_{|A_j}$ is summable on A_j for each $j = 1, ..., m$ and

$$\int_A f(x)dx = \sum_{j=1}^{m} \int_{A_j} f_{|A_j}(x)dx \tag{1.2.7}$$

holds; on the other hand, if, for each $j = 1, ..., m$ $f_{|A_j}$ is summable on A_j, then f is summable on A and (1.2.7) holds;
(IV) $|f|$ is measurable, $\int_A |f(x)|dx < +\infty$ and inequality (1.2.8) holds.

Exercise 1.2.13. Let $A, B \in \mathcal{M}_n$ with $A \subseteq B$, $f : B \to [-\infty, +\infty]$ measurable. Show that $f_{|A}$ is measurable in A.

Exercise 1.2.14. Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{R}$ summable. Show that

$$|\int_A f(x)dx| \leq \int_A |f(x)|dx. \tag{1.2.8}$$

Exercise 1.2.15. Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{C}$ summable. Prove that $|f|$ is summable on A (hint: $|f| \leq |Re(f)| + |Im(f)|$).

Exercise 1.2.16. Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{C}$. Show that f is summable if and only if $Re(f)$ e $Im(f)$ are measurable and $|f|$ is summable.

Exercise 1.2.17. Let $A_1, ..., A_m \in \mathcal{M}_n$, pairwise disjoint, $f : \cup_{1 \leq i \leq m} A_i \to [-\infty, +\infty]$ such that $f_{|A_i}$ is measurable on A_i for each $i \in \{1, ..., m\}$. Prove that f is measurable.

Exercise 1.2.18. Let $A \in \mathcal{M}_n$, $f : A \to \mathbb{C}$. Prove that, if f is summable, then $\overline{f}$ is summable (given $z \in \mathbb{C}$, we indicate with $\overline{z}$ the complex conjugate of z). Moreover,

$$\int_A \overline{f(x)}dx = \overline{\int_A f(x)dx}. \tag{1.2.9}$$

1.3 Operational techniques for the computation of integrals

In this section we introduce some basic results, able to make often possible the effective computation of integrals in the sense of Lebesgue. We assume that the reader is acquainted with the theory of integration in the sense of Riemann for functions of one real variable in a closed and bounded interval.

The first result is the following:

Theorem 1.3.1. *Let $A \in \mathcal{M}_n$, with $L_n(A) = 0$, $f : A \to [0, +\infty]$. Then f is measurable and*

$$\int_A f(x)dx = 0. \tag{1.3.1}$$

If $f : A \to \mathbb{C}$, then f is summable and (1.3.1) holds.

Proof See Exercise 1.3.9 □

Now we consider the case that A is a closed and bounded interval in $\mathbb{R}$. We compare integration with respect to Lebesgue's theory with integration with respect to Riemann's theory.

Theorem 1.3.2. *Let a and b be real numbers, with $a < b$, $f : [a, b] \to \mathbb{R}$ which is integrable in the sense of Riemann. Then f is summable in $[a, b]$ and the 'integral in the sense of Lebesgue (Definition 1.2.9) coincides with the integral in the sense of Riemann.*

Now we pass to consider integration in half-open intervals.

Theorem 1.3.3. *Let $-\infty < a < b \leq +\infty$, $f : [a, b[\to [0, +\infty[$ integrable in the sense of Riemann in any interval $[a, c]$ with $c \in]a, b[$. Then f is measurable in $[a, b[$ and*

$$\int_{[a,b[} f(x)dx = \lim_{c \to b} \int_a^c f(x)dx. \tag{1.3.2}$$

holds.

We observe that the limit in (1.3.2) exists (finite or infinite), because the integral function $c \to \int_a^c f(x)dx$ is nondecreasing in $[a, b[$. For a proof, using the monotone convergence theorem, see Exercise 1.5.8.

An analogous result holds for functions defined in intervals of the form $]a, b]$, with $-\infty \leq a < b < +\infty$ (see Exercise 1.3.11).

We pass to real valued functions with arbitrary sign.

Theorem 1.3.4. *Let $-\infty < a < b \leq +\infty[$, let $f : [a, b[\to \mathbb{R}$ be integrable in the sense of Riemann $[a, c[$ for every $c \in]a, b[$ and let $\lim_{c\to b} \int_a^c |f(x)|dx < +\infty$. Then:*

(I) there exists in $\mathbb{R}$ $\lim_{c\to b} \int_a^c f(x)dx$;

(II) f is summable in $[a, b[$;

(III) $\int_{[a,b[} f(x)dx = \lim_{c\to b} \int_a^c f(x)dx$.

Now we consider the multidimensional case, with the classical theorems of Tonelli, Fubini and of the change of variable. We state in advance some notations. We suppose of working in $\mathbb{R}^{m+n}$, with m and n in $\mathbb{N}$ and we indicate with (x, y) the generic element of $\mathbb{R}^{m+n}$, with $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$. Given $x \in R^m$, we set

$$A_x := \{y \in \mathbb{R}^n : (x, y) \in A\}. \tag{1.3.3}$$

Theorem 1.3.5. *(of Tonelli) Let $A \in \mathcal{M}_{m+n}$, $f : A \to [0, +\infty]$ measurable. Let, moreover, $B \in \mathcal{M}_m$, and such that*

$$\{x \in \mathbb{R}^m : A_x \in \mathcal{M}_n, L_n(A_x) > 0\} \subseteq B.$$

We define

$$\begin{cases} g : B \to [0, +\infty], \\ g(x) = \begin{cases} \int_{A_x} f(x, y)dy \text{ if } A_x \neq \emptyset, A_x \in \mathcal{M}_n, f(x, .) \text{ measurabile in } A_x, \\ 0 \text{ otherwise.} \end{cases} \end{cases}$$

Then:

(I) g is measurable in B;

(II) $\int_A f(x, y)dxdy = \int_B g(x)dx$.

Theorem 1.3.6. *(of Fubini) Let $A \in \mathcal{M}_{m+n}$, let $f : A \to \mathbb{C}$ be summable. Next, let $B \in \mathcal{M}_m$ be such that*

$$\{x \in \mathbb{R}^m : A_x \in \mathcal{M}_n, L_n(A_x) > 0\} \subseteq B.$$

We define

$$\begin{cases} g : B \to \mathbb{C}, \\ g(x) = \begin{cases} \int_{A_x} f(x, y)dy \text{ if } A_x \neq \emptyset, A_x \in \mathcal{M}_n, f(x, .) \text{ summable in } A_x, \\ 0 \text{ otherwise.} \end{cases} \end{cases}$$

Then:

(I) g is summable in B;

(II) $\int_A f(x, y)dxdy = \int_B g(x)dx$.

Now we pass to the theorem of the change of variable.

Definition 1.3.7. *Let Ω be an open subset of $\mathbb{R}^n$, $T : \Omega \to \mathbb{R}^n$. We shall say that T is a* **change of variable** *if*

(I) T is injective;

(II) T is of class C^1;

(III) for every $x \in \Omega$ the determinant $detJ_T(x)$ of the Jacobian matrix of T in x is not 0.

Theorem 1.3.8. *(the change of variable theorem) Let $T : \Omega \to \mathbb{R}^n$ be a change of variable, with Ω open in $\mathbb{R}^n$, $A \in \mathcal{M}_n$, $f : A \to [0, +\infty]$ (resp. $f : A \to \mathbb{C}$). We assume that $A \subseteq T(\Omega)$. Then:*

(I) $T^{-1}(A) \in \mathcal{M}_n$;

(II) f is measurable in A (resp. f is summable in A) if and only if $x \to f(T(x))|detJ_T(x)|$ is measurabile (resp. summable) in $T^{-1}(A)$; in such a case,

(III) $\int_A f(y)dy = \int_{T^{-1}(A)} f(T(x))|detJ_T(x)|dx$.

Exercise 1.3.9. Prove Theorem 1.3.1, assuming that every function of domain A is measurable (hint: employ the result of Exercise 1.1.9, starting by considering the case that f is simple and nonnegative).

Exercise 1.3.10. Let $f : [0, 1] \to \mathbb{R}$, $f(x) = 1$ if $x \in \mathbb{Q}$, $f(x) = 0$ otherwise. Verify that f is summable in $[0, 1]$ and has integral 0 (hint: employ the result of Exercise 1.1.8). f is the so called "Dirichlet function", which is not integrable in the sense of Riemann. This example shows that the statement of Theorem 1.3.2 is not invertible.

Exercise 1.3.11. State the analog of Theorem 1.3.3 for an interval of the form $]a, b]$, with $-\infty \leq a < b < +\infty$.

Exercise 1.3.12. Extend Theorem 1.3.4 to the case of complex valued functions.

1.4 Identification of measurable functions coinciding almost everywhere

We start with the following quite suggestive definition:

Definition 1.4.1. *Let $A \in \mathcal{M}_n$. We shall say that a certain property $P(x)$ is valid* **almost everywhere** *in A (a.e.), or for almost every $x \in A$, if $\{x \in A : P(x)$ does not hold $\}$ is measurable and has measure 0.*

Given f and g defined in A and complex valued, we shall say that f *is equivalent to* g, and we shall write

$$f \sim g$$

if $f(x) = g(x)$ a. e. in A. One can easily verify that $\sim$ is an equivalence relation, that is, given three arbitrary functions f, g and h from A to $\mathbb{C}$, one has:

(I) $f \sim f$ (reflexive property);

(II) if $f \sim g$, then $g \sim f$ (simmetric property);

(III) if $f \sim g$ and $g \sim h$, then $f \sim h$ (transitive property).

Given $f : A \to \mathbb{C}$, we define the *equivalence class* $[f]$ of f as

$$[f] := \{g : A \to \mathbb{C} : g \sim f\}. \tag{1.4.1}$$

The following result will be important in the sequel:

Theorem 1.4.2. *Let $A \in \mathcal{M}_n$, $f, g, f_1, f_2, g_1, g_2 : A \to \mathbb{C}$, with $f \sim g$, $f_1 \sim g_1$, $f_2 \sim g_2$, $c \in \mathbb{C}$. Then:*

(I) $f_1 + g_1 \sim f_2 + g_2$;
(II) $cf \sim cg$;
(III) $Re(f) \sim Re(g)$, $Im(f) \sim Im(g)$;
(IV) if f and g are real valued and f is measurable, even g is measurable;
(V) if f is summable, even g is summable and one has

$$\int_A f(x)dx = \int_A g(x)dx.$$

Proof We verify only (IV) and leave the rest to the reader (see Exercise 1.4.4). Let $N := \{x \in A : f(x) \neq g(x)\}$. N is measurable and has measure 0. So g coincides with f in $A \setminus N$ and is measurable in N by Theorem 1.3.1. Therefore, we obtain the conclusion from the result of Exercise 1.2.17. □

From Therem 1.4.2 it follows that, if f and g are complex valued functions with domain A, the following expressions are well defined:

$$[f] + [g] := [f + g] \tag{1.4.2}$$

and

$$\lambda[f] := [\lambda f]. \tag{1.4.3}$$

This happens because, in each case, the second term in (1.4.2)-(1.4.3) is independent of the choice of the single elements in each class.

Exercise 1.4.3. Prove that the relation $\sim$ enjoys the properties $(I) - (III)$.

Exercise 1.4.4. Complete the proof of Theorem 1.4.2.

1.5 Passage to the limit under the sign of integral

Now we state without proof some basic results of passage to the limit under the sign of integral. Let $A \in \mathcal{M}_n$ and, for $k \in \mathbb{N}$, let $f_k : A \to [0, +\infty]$ (resp. $f_k : A \to \mathbb{C}$), $f : A \to [0, +\infty]$ (resp. $f : A \to \mathbb{C}$) be such that, for almost every $x \in A$, $\lim_{k \to +\infty} f_k(x) = f(x)$ holds. Assuming that the functions f_k admit an integral in the sense of Definition 1.2.6 or of Definition 1.2.9, we wonder when it is possible to conclude that f admits an integral and $\lim_{k \to +\infty} \int_A f_k(x)dx = \int_A f(x)dx$ holds. The problem is not trivial, as the following example shows:

Example 1.5.1. Let, for $k \in \mathbb{N}$,

$$\begin{cases} f_k : [0,1] \to \mathbb{R}, \\ f_k(x) = \begin{cases} k^3 x & \text{if} \quad x \in [0, \frac{1}{k}], \\ k^2 - k^3(x - \frac{1}{k}) & \text{if} \quad x \in [\frac{1}{k}, \frac{2}{k}], \\ 0 & \text{if} \quad x \in [\frac{2}{k}, 1]. \end{cases} \end{cases}$$

It is easy to see that $\lim_{k \to +\infty} f_k(x) = 0 \ \forall x \in [0,1]$. This is obvious if $x = 0$. If $x > 0$, it suffices to observe that, if k is large enough, one has $x > \frac{2}{k}$, so that $f_k(x) = 0$. However, $\int_{[0,1]} f_k(x)dx = k$ $\forall k \in \mathbb{N}$, so that $\lim_{k \to +\infty} \int_{[0,1]} f_k(x)dx = +\infty$.

Now we are going to see two classical theorems of passage to the limit. We state before the following notion of convergence for a complex valued sequence, a generalization of which will be examined in the sequel (see the following Section 2.3): let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{C}$ and let $a \in \mathbb{C}$. We shall write $\lim_{n\to+\infty} a_n = a$ if $\forall \epsilon > 0$ there exists $n(\epsilon) \in \mathbb{N}$ such that, $\forall n \in \mathbb{N}$ con $n > n(\epsilon)$, one has $|a_n - a| < \epsilon$.

Theorem 1.5.2. *(of Beppo Levi, or of the monotone convergence) Let $A \in \mathcal{M}_n$ and, for $k \in \mathbb{N}$, let $f_k : A \to [0,+\infty]$ be measurable. Assume, moreover, that $f_k(x) \le f_{k+1}(x)$ $\forall k \in \mathbb{N}$, $\forall x \in A$. We set $f : A \to [0,+\infty]$, $f(x) = \lim_{k\to+\infty} f_k(x)$ (the limit exists in force of the monotonicity of $(f_k(x))_{k\in\mathbb{N}}$). Then f is measurable, with values in $[0,+\infty]$, and*

$$\int_A f(x)dx = \lim_{k\to+\infty} \int_A f_k(x)dx.$$

Example 1.5.3. Let, for $k \in \mathbb{N}$, $f_k : [1,+\infty[\to \mathbb{R}$, $f_k(x) = \frac{x^k e^{-x}}{1+x^k}$. f_k is measurable, because it is continuous, and nonnegative. If $x \ge 1$, one has $x^k \le x^{k+1}$, and, as $y \to \frac{y}{1+y}$ is nondecreasing in $[0,+\infty[$, $\frac{x^k}{1+x^k} \le \frac{x^{k+1}}{1+x^{k+1}}$. So, $f_k(x) \le f_{k+1}(x)$, for every $k \in \mathbb{N}$ and $x \in [1,+\infty[$. We have

$$\lim_{k\to+\infty} f_k(x) = \begin{cases} \frac{1}{2e} & \text{if} \quad x = 1, \\ e^{-x} & \text{if} \quad x > 1. \end{cases}$$

So we set

$$\begin{cases} f : [1,+\infty[\to \mathbb{R}, \\ f(x) = \begin{cases} \frac{1}{2e} & \text{if} \quad x = 1, \\ e^{-x} & \text{if} \quad x > 1. \end{cases} \end{cases}$$

Now, by the monotone convergence theorem,

$$\lim_{k\to+\infty} \int_{[1,+\infty[} f_k(x)dx = \int_{[1,+\infty[} f(x)dx$$

$$= \int_{\{1\}} \frac{1}{2e} dx + \int_{]1,+\infty[} e^{-x} dx$$

$$= \int_{]1,+\infty[} e^{-x} dx = \int_{\{1\}} e^{-x} dx + \int_{]1,+\infty[} e^{-x} dx$$

$$= \int_{[1,+\infty[} e^{-x} dx,$$

where we have used Theorem 1.2.7(V) and Theorem 1.3.1, because $L_1(\{1\}) = 0$. At this point, we apply Theorem 1.3.3 and obtain

$$\int_{[1,+\infty[} e^{-x} dx = \lim_{c\to+\infty} \int_1^c e^{-x} dx$$

$$= \lim_{c\to+\infty} (e^{-1} - e^{-c}) = e^{-1}.$$

We pass to consider complex valued functions.

Theorem 1.5.4. *(of Lebesgue, of the dominated convergence) Let $A \in \mathcal{M}_n$ and, for $k \in \mathbb{N}$, let $f_k : A \to \mathbb{C}$ be measurable and $g : A \to [0, +\infty[$ summable. We assume also that $|f_k(x)| \leq g(x)$ $\forall k \in \mathbb{N}$ and a. e. in A. Let $f : A \to \mathbb{C}$ be such that $f(x) = \lim_{k \to +\infty} f_k(x)$ a. e. in A. Then the functions f_k and f are summable in A and*

$$\int_A f(x)dx = \lim_{k \to +\infty} \int_A f_k(x)dx$$

holds.

Example 1.5.5. Let, for $k \in \mathbb{N}$, $f_k : [0, +\infty[\to \mathbb{R}$, $f_k(x) = \sin(\frac{x}{n})e^{-x}$. f_k is measurable, because it is continuous. We set $g : [0, +\infty[\to \mathbb{R}$, $g(x) = e^{-x}$. g is summable, nonnegative and $|f_k(x)| \leq g(x)$ $\forall k \in \mathbb{N}$, $\forall x \geq 0$. Finally, we have $\lim_{k \to +\infty} f_k(x) = 0$ $\forall x \geq 0$. So, from the dominated convergence theorem, it follows that, for every $k \in \mathbb{N}$, f_k is summable in $[0, +\infty[$ and

$$\lim_{k \to +\infty} \int_{[0,+\infty[} f_k(x)dx = \int_{[0,+\infty[} 0 \quad dx = 0.$$

Exercise 1.5.6. Let $A \in \mathcal{M}_n$, let, for $k \in \mathbb{N}$, $f_k : A \to [0, +\infty]$ be measurable, $s : A \to [0, +\infty]$, $s(x) = \sum_{k=1}^{\infty} f_k(x)$. Prove that f is measurable and

$$\int_A s(x)dx = \sum_{k=1}^{\infty} \int_A f_k(x)dx.$$

Exercise 1.5.7. Let $A \in \mathcal{M}_n$ and, for $k \in \mathbb{N}$, let $f_k : A \to \mathbb{C}$ be measurable. We suppose also that

(a) $L_n(A) < +\infty$;
(b) there exists $M \geq 0$, such that $|f_k(x)| \leq M$ $\forall x \in A$;
(c) there exists $f : A \to \mathbb{C}$, such that $f(x) = \lim_{k \to +\infty} f_k(x)$ a. e. in A.

Prove that the functions f_k and f are summable, and

$$\int_A f(x)dx = \lim_{k \to +\infty} \int_A f_k(x)dx.$$

Exercise 1.5.8. Prove Theorem 1.3.3. Hint: take an arbitrary increasing sequence $(t_k)_{k \in \mathbb{N}}$, converging to b. Define, for $k \in \mathbb{N}$, $f_k : [a, b[\to [0, +\infty[$, $f_k(x) = f(x)$ if $a \leq x \leq t_k$, $f_k(x) = 0$ if $t_k < x < b$. Applying the monotone convergence theorem, prove that f is measurable and

$$\int_{[a,b[} f(x)dx = \lim_{k \to +\infty} \int_{[a,b[} f_k(x)dx = \lim_{k \to +\infty} \int_a^{t_k} f(x)dx.$$

Conclude, using the monotonicity of $t \to \int_a^t f(x)dx$.

Exercise 1.5.9. Prove Theorem 1.3.4. Hint: employ the dominated concergence theorem, taking $g = |f|$

Chapter 2

Normed spaces

2.1 Norms

We assume that the reader is acquainted with the basic elements of linear algebra. For completeness and in order to establish the notation, we start by recalling some of these elements. Let us see, firstly, the notion of linear (vector) space on the field $K = \mathbb{R}$ o $\mathbb{C}$.

Definition 2.1.1. *Let X be a nonempty set, $+$ an operation in X (that is, a function from $X \times X$ yo X), $(\lambda, x) \to \lambda x$ a map from $K \times X$ a X, which we shall denominate "scalar multiplication". We shall say that X, with the sum $+$ and this scalar multiplication), is a linear space over K if the following proprieties hold:*

(ASV1) $\forall x, y \in X$

$$y + x = x + y$$

(commutative propriety of the sum);

(ASV2) $\forall x, y, z \in X$

$$x + (y + z) = (x + y) + z$$

(associative propriety of the sum);

(ASV3) there esists an element $O \in X$ such that, $\forall x \in X$

$$x + O = O + x = x$$

(esistence of a neutral element for the sum);

(ASV4) $\forall x \in X$ *there exists $-x \in X$ such that*

$$x + (-x) = (-x) + x = 0;$$

(existence of an inverse element, with respect to the sum, for every $x \in X$);

(ASV5) $\forall \lambda,\ \mu \in K$, $\forall x \in X$

$$(\lambda + \mu)x = \lambda x + \mu x;$$

(ASV6) $\forall \lambda \in K$, $\forall x, y \in X$

$$\lambda(x + y) = \lambda x + \lambda y;$$

(ASV7) $\forall \lambda,\ \mu \in K$, $\forall x \in X$

$$\lambda(\mu x) = (\lambda \mu)x;$$

(ASV8) $\forall x \in X$

$$1x = x.$$

Remark 2.1.2. As usual, we have indicated with the same symbol $+$ the sum in X and the sum in K. For example, in $(ASV5)$ the first $+$ indicates the sum in K, the second the sum in X.

Example 2.1.3. Given $K = \mathbb{R}$ or $\mathbb{C}$ and $n \in \mathbb{N}$, we indicate with K^n the set of ordered $n-$tuples of elements of K. If $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$ are elements of K^n, we set

$$x + y := (x_1 + y_1, ..., x_n + y_n). \tag{2.1.1}$$

Moreover, if $\lambda \in K$, we set

$$\lambda x := (\lambda x_1, ..., \lambda x_n). \tag{2.1.2}$$

It is well known that K^n, equipped with the sum (2.1.1) and the scalar multiplication (2.1.2) is a linear space over K.

Example 2.1.4. Let A be a nonempty (arbitrary) set. We indicate with $\mathcal{F}(A, K)$ the set of functions with domain A and values in K. Given f and g in $\mathcal{F}(A, K)$ and $\lambda \in K$, we set

$$\begin{cases} f + g : A \to K, \\ (f + g)(a) = f(a) + g(a), \forall a \in A, \end{cases} \tag{2.1.3}$$

$$\begin{cases} \lambda f : A \to K, \\ (\lambda f)(a) = \lambda \cdot f(a), \forall a \in A, \end{cases} \tag{2.1.4}$$

with $\cdot$ product in K. It is easy to see that $\mathcal{F}(A, K)$, with the sum (2.1.3) and the scalar multiplication (2.1.4), is a linear space over K.

Remark 2.1.5. Given a linear space X on the field K and given $Y \subseteq X$, it is said that Y is a **(linear) subspace** of X if, given y_1, y_2 and y elements of Y and $\lambda \in K$, one always has $y_1 + y_2 \in Y$ and $\lambda y \in Y$. It is not difficult to verify that, if Y is a subspace of X, it is a linear space over K with the restrictions of the sum to $Y \times Y$ and of the scalar multiplication to $K \times Y$.

If $A \subseteq \mathbb{R}^n$, the linear space $\mathcal{F}(A, \mathbb{R})$ has several interesting subspaces. We indicate some of them.

$B(A, \mathbb{R})$ is the space of bounded functions form A to $\mathbb{R}$.

$C(A, \mathbb{R})$ is the space of continuous functions from A to $\mathbb{R}$.

$BC(A, \mathbb{R})$ is the space of bounded and continuous functions from A to $\mathbb{R}$.

If A is measurable in the sense of Lebesgue, $\mathcal{M}(A, \mathbb{R})$ is the space of measurable functions from A to $\mathbb{R}$.

If A is measurable in the sense of Lebesgue, $\mathcal{L}^1(A, \mathbb{R})$ is the space of summable functions from A to $\mathbb{R}$.

Instead, we indicate with $\mathcal{L}^1(A)$ the space of (in general) complex valued summable functions, with domain A. In this case, we obtain a linear space over $\mathbb{C}$.

Definition 2.1.6. *Let X be a linear space over K ($= \mathbb{R}$ or $\mathbb{C}$). A* **norm** *in X is a function $\|.\| : X \to [0, +\infty[$, $x \to \|x\|$ $\forall x \in X$, such that*

(I) $\forall \lambda \in K$, $\forall x \in X$

$$\|\lambda x\| = |\lambda| \|x\|;$$

(II) $\forall x, y \in X$

$$\|x + y\| \leq \|x\| + \|y\|;$$

(III) if $x \in X$ and $\|x\| = 0$, then $x = 0$.

Example 2.1.7. Given $x = (x_1, ..., x_n) \in \mathbb{R}^n$, we set,

$$\|x\| := (\sum_{j=1}^{n} x_j^2)^{\frac{1}{2}}.$$

$\|.\|$ is the Euclidean norm in $\mathbb{R}^n$ and it is well known that it fulfills the conditions $(I)-(III)$ in Definition 2.1.6.

It is not the unique norm in $\mathbb{R}^n$. Two different norms are, for $x = (x_1, ..., x_n) \in \mathbb{R}^n$,

$$\|x\|_1 := \sum_{j=1}^{n} |x_j|, \tag{2.1.5}$$

and

$$\|x\|_\infty := \max_{1 \le j \le n} |x_j|. \tag{2.1.6}$$

In order to verify that they are really norms, see the following Exercise 2.1.13.

Example 2.1.8. In analogy with (2.1.6), given $A \subseteq \mathbb{R}^n$ (in fact, one can take an arbitrary nonempty set), if we put $X := B(A, \mathbb{R})$, we can define in X the norm

$$\|x\|_\infty := \sup_A |f| = \sup\{|f(a)| : a \in A\}. \tag{2.1.7}$$

We verify that $\|.\|_\infty$ is a norm in X.

First of all, it is clear that $\|.\|_\infty$ is well defined in X and with values in $[0, +\infty[$. We examine, one by one, the proprieties $(I)-(III)$ in Definition 2.1.6.

(I) If $a \in A$, $|\lambda f(a)| = |\lambda||f(a)| \le |\lambda| \sup_A\{|f(a)| : a \in A\} = |\lambda|\|f\|_\infty$. As this holds for every element a, we deduce

$$\|\lambda f\|_\infty = \sup\{|\lambda f(a)| : a \in A\} \le |\lambda|\|f\|_\infty. \tag{2.1.8}$$

If $\lambda \ne 0$, osserving that $f = \lambda^{-1}(\lambda f)$ and applying (2.1.8), we obtain

$$\|f\|_\infty \le |\lambda^{-1}|\|\lambda f\|_\infty,$$

from which we immediately draw (I). The case $\lambda = 0$ is trivial.

(II) Let f and g be elements of X and $a \in A$. Then

$$|f(a) + g(a)| \le |f(a)| + |g(a)| \le \|f\|_\infty + \|g\|_\infty.$$

It follows that

$$\|f + g\|_\infty = \sup\{|f(a) + g(a)| : a \in A\} \le \|f\|_\infty + \|g\|_\infty.$$

(III) Let $f \in X$ be such that $\|f\|_\infty = 0$. Then $\sup\{|f(a)| : a \in A\} = 0$. Evidently, this implies that $f(a) = 0\ \forall a \in A$. The function which vanishes in every point of A is precisely the neutral element with respect to the sum in the linear space X. It follows that even (III) holds.

Remark 2.1.9. It is easy to verify that in every subspace of $B(A, \mathbb{R})$ (for example, $BC(A, \mathbb{R})$) (2.1.7) defines a norm.

Example 2.1.10. Let $A \in \mathcal{M}_n$ (the classe of measurable subsets in $\mathbb{R}^n$). We put, given $f \in \mathcal{L}^1(A)$,

$$\|f\|_1 := \int_A |f(x)|dx. \tag{2.1.9}$$

If $\|.\|_1$ were a norm, the following fact should hold:

"If $f \in \mathcal{L}^1(A)$ and $\int_A |f(x)|dx = 0$, then $f(x)$ is identically 0."

However, it is easy to see that this is false. For example, if $A = \mathbb{R}^n$, $B \in \mathcal{M}_n$, $B \neq \emptyset$ and $L_n(B) = 0$, we set

$$\begin{cases} f : A \to \mathbb{R}, \\ f(x) = \begin{cases} 1 & \text{if} \quad x \in B, \\ 0 & \text{if} \quad x \notin B. \end{cases} \end{cases}$$

f is a nonnegative simple function, not identically equal to 0 and

$$\|f\|_1 = \int_A f(x)dx = 1 \cdot L_n(B) + 0 \cdot L_n(\mathbb{R}^n \setminus B) = 0.$$

In fact, the following important result can be shown:

Lemma 2.1.11. *Let $A \in \mathcal{M}_n$ and let $f : A \to [0, +\infty]$ be measurable. Then the two following conditions are equivalent:*
(I) $\int_A f(x)dx = 0$;
(II) $f(x) = 0$ a. e. in A, that is, $\{x \in A : f(x) > 0\}$ is measurable and has measure 0.

So, if $\|f\|_1 = 0$, we can only say that $|f(x)| = 0$ a. e. in A, and, consequently, $f(x) = 0$ a. e. in A. In order to overcome this inconvenience, we replace $\mathcal{L}^1(A)$ with the linear space

$$L^1(A) := \{[f] : f \in \mathcal{L}^1(A)\}, \tag{2.1.10}$$

where $[f]$ is the equivalence class of f defined in (1.4.1). $L^1(A)$ is a linear space over $\mathbb{C}$, with the operations of sum and scalar multiplication defined in (1.4.2) and (1.4.3). Given $[f] \in L^1(A)$, we set

$$\|[f]\|_1 := \int_A |f(x)|dx. \tag{2.1.11}$$

The definition is well posed, by Theorem 1.4.2(V).

Let us show that $\|.\|_1$ is a norm in $L^1(A, \mathbb{R})$.

First of all, if $\lambda \in \mathbb{C}$ and $f \in \mathcal{L}^1(A)$,

$$\begin{gathered} \|\lambda[f]\|_1 = \|[\lambda f]\|_1 \\ = \int_A |\lambda f(x)|dx = |\lambda| \int_A |f(x)|dx \\ = |\lambda| \|[f]\|_1. \end{gathered}$$

If f and g are elements of $\mathcal{L}^1(A)$,

$$\begin{gathered} \|[f] + [g]\|_1 = \|[f + g]\|_1 \\ = \int_A |f(x) + g(x)|dx \leq \int_A |f(x)|dx + \int_A |g(x)|dx \end{gathered}$$

$$= \|[f]\|_1 + \|[g]\|_1.$$

Finally, let $\|[f]\|_1 = 0$. Then $\int_A |f(x)|dx = 0$, and so $f(x) = 0$ a. e. . This means that f is equivalent to the neutral element, that is, $[f] = [0]$.

Remark 2.1.12. Although the elements of $L^1(A)$ ($A \in \mathcal{M}_n$) are not functions, but classes of functions coinciding almost everywhere, it is convenient and usual to speak of a certain function f as an element of $L^1(A)$. In reality, we shall always mean the corresponding equivalence class.

Exercise 2.1.13. Check that $\|.\|_1$ e $\|.\|_\infty$, defined in (2.1.5) and (2.1.6), are norms in $\mathbb{R}^n$.

Exercise 2.1.14. Check that, if $\|.\|$ is a norm in X, $\forall x, y \in X$ one has

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Exercise 2.1.15. Let $f : [0, 2\pi] \to \mathbb{R}$, $f(x) = x - \sin(x) - 4\pi$. Calculate $\|f\|_\infty$. Calculate also $\|f\|_1$.

Exercise 2.1.16. Let $f : [-2, 1] \to \mathbf{R}$, $f(x) = x^3$. Calculate $\|f\|_\infty$ and $\|[f]\|_1$.

2.2 Notions of topological type in a normed space

Now we introduce in a generic normed space a series of notions of topological anture already seen in $\mathbb{R}^n$. In this section X will be an arbitrary normed space over $K = \mathbb{R}$ or $\mathbb{C}$. We shall indicate with $\|.\|$ the norm in X.

Let $x \in X$ and $r > 0$. We set

$$B(x, r) := \{y \in X : \|y - x\| < r\}. \tag{2.2.1}$$

We shall denominate $B(x, r)$ the **open ball** with centre x and radius r.

Let $A \subseteq X$ and $x_0 \in X$. We shall say that x_0 belongs to the **interior** of A, and we shall write

$$x_0 \in \mathring{A}, \tag{2.2.2}$$

if there exists $r > 0$, such that $B(x_0, r) \subseteq A$.

Next, given $A \subseteq X$, we shall say that A is **open** if $\mathring{A} = A$.

Given $A \subseteq X$ and $x_0 \in X$, we shall say that x_0 belongs to the **boundary** of A, and we shall write

$$x_0 \in \partial A, \tag{2.2.3}$$

if every open ball $B(x_0, r)$, with $r > 0$, contains both elements of A, and elements not belonging to A.

We shall say that $A \subseteq X$ is **closed** if $\partial A \subseteq A$.

Next, given $A \subseteq X$, we shall denominate **closure** of A, and we shall indicate with $\overline{A}$, the set

$$\overline{A} := A \cup \partial A. \tag{2.2.4}$$

If $A \subseteq X$ and $\overline{A} = X$, we shall say that A is **dense** in X.

Finally we introduce the notion of continuity.

Definition 2.2.1. *Let X and Y be normed spaces, with norms, respectively, $\|.\|_X$ and $\|.\|_Y$, $A \subseteq X$, $f : A \to Y$, $x_0 \in A$. We shall say that f is* **continuous** *in x_0 if $\forall \epsilon \in \mathbb{R}^+$ there exists $r(\epsilon) \in \mathbb{R}^+$, such that , $\forall x \in A$, if $\|x - x_0\|_X < r(\epsilon)$, the inequality*

$$\|f(x) - f(x_0)\|_Y < \epsilon$$

holds.

We shall say that f is continuous if it is continuous in every point of A.

It is convenient to introduce also the notions of limit point and of limit.

Definition 2.2.2. *Let $A \subseteq X$, with X normed space, and $x^0 \in X$. We shall say that x^0 is a limit point of A (and we shall write $x^0 \in L(A)$) if $\forall r > 0$ $B(x^0, r)$ contains some element of A which is distinct from x^0.*

Definition 2.2.3. *Let X and Y be normed spaces, with norms, respectively, $\|.\|_X$ e $\|.\|_Y$. Next, let $A \subseteq X$, $x^0 \in L(A)$ and $l \in Y$. We shall write $\lim_{x \to x^0} f(x) = l$ if $\forall \epsilon \in \mathbb{R}^+$ there esists $r(\epsilon) \in \mathbb{R}^+$ such that, $\forall x \in A$ with $\|x - x_0\|_X < r(\epsilon)$ and $x \neq x^0$, one has*

$$\|f(x) - l\|_Y < \epsilon.$$

Example 2.2.4. We consider the linear space $X := C([0,1], \mathbb{R})$ with the norm $\|.\|_\infty$ (see (2.1.7)). We observe that, by Weierstrass theorem, $C([0,1], \mathbb{R})$ coincides with $BC([0,1], \mathbb{R})$, and so this norm is well defined in X. We set

$$\begin{cases} f : X \to X, \\ f(x)(t) := \int_0^t x(s)ds, \qquad x \in X. \end{cases} \tag{2.2.5}$$

The fundamental theorem of integral calculus guarantees that $f(x)$ is continuous and so it belongs to X. We check that f is continuous.

Let $x_0 \in X$ and $\epsilon > 0$; given $x \in X$, one has, $\forall t \in [0,1]$,

$$|F(x)(t) - F(x_0)(t)| = |\int_0^t (x(s) - x_0(s))ds|$$

$$\leq \int_0^t |x(s) - x_0(s)|ds \leq t\|x - x_0\|_\infty$$

$$\leq \|x - x_0\|_\infty.$$

It follows that

$$\|F(x) - F(x_0)\|_\infty \leq \|x - x_0\|_\infty.$$

So, if we set $r(\epsilon) := \epsilon$, we have that, if $\|x - x_0\|_\infty < r(\epsilon)$, $\|F(x) - F(x_0)\|_\infty < \epsilon$ holds.

Exercise 2.2.5. Prove that open balls are open.

Exercise 2.2.6. Prove that, if $A = B(x_0, r)$, with $x_0 \in X$ and $r > 0$,

$$\partial A = \{x \in X : \|x - x_0\| = r\}.$$

Exercise 2.2.7. Let $A \subseteq X$. Prove that A is open if and only if

$$A \cap \partial A = \emptyset.$$

Exercise 2.2.8. Let $A \subseteq X$. Prove that A is closed if and only if

$$A = \overline{A}.$$

Exercise 2.2.9. Let $A \subseteq X$, $f : A \to \mathbb{C}$. Prove that f is continuous if and only if $Re(f)$ and $Im(f)$ are real valued continuous functions.

(Hint.: for every $\lambda \in \mathbb{C}$ $\max\{|Re(\lambda)|, |Im(\lambda)|\} \leq |\lambda| \leq |Re(\lambda)| + |Im(\lambda)|$).

Exercise 2.2.10. Check that, if X is a normed space with norm $\|.\|$, such a norm is continuous from X to $\mathbb{R}$.

(Hint: employ the result of Exercise 2.1.14).

Exercise 2.2.11. State and prove a theorem of uniqueness of the limit.

Exercise 2.2.12. Let $f : A \subseteq X \to Y$, with X and Y normed spaces. Next, let $x^0 \in A \cap L(A)$. Prove that f is continuous in x^0 if and only if there exists $\lim_{x \to x^0} f(x)$ and it coincides with $f(x^0)$.

2.3 Sequences in a normed space and completeness

We start with the definition of limit of a sequence in a normed space X (with norm $\|.\|$).

Definition 2.3.1. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in the normed space X and let $l \in X$. We shall say that $(x_n)_{n\in\mathbb{N}}$* **converges** *to l, or has* **limit** *l if $\forall \epsilon \in \mathbb{R}^+$ there exists $n(\epsilon) \in \mathbb{N}$ such that, for every $n \in \mathbb{N}$ with $n > n(\epsilon)$, one has*

$$\|x_n - l\| < \epsilon.$$

Remark 2.3.2. Definition 2.3.1 is equivalent to require that

$$\lim_{n\to+\infty} \|x_n - l\| = 0.$$

Remark 2.3.3. One can easily check, with the same argument employed for real values sequences, that, if the limit exists, it is unique (see Exercise 2.3.17).

Remark 2.3.4. The convergence with respect to the norm $\|.\|_\infty$ in $B(A, \mathbb{R})$ (with A nonempty set) is usually called **uniform convergence**.

Example 2.3.5. Let $\delta \in]0, 1[$, $X = C([0, \delta], \mathbb{R})$, with the norm $\|.\|_\infty$. Given $n \in \mathbb{N}$, $t \in [0, \delta]$, we set

$$x_n(t) := t^n, \tag{2.3.1}$$

$$l(t) = 0. \tag{2.3.2}$$

We check that the sequence $(x_n)_{n\in\mathbb{N}}$ converges to l. In fact,

$$\|x_n - l\|_\infty = \sup_{t\in[0,\delta]} |x_n(t) - l(t)|$$

$$= \delta^n \to 0(n \to +\infty).$$

It is quite clear that uniform convergence implies pointwise convergence, in the sense that, if $(x_n)_{n\in\mathbb{N}}$ is a sequence in $B(A, \mathbb{R})$ uniformly convergent to l ($\in B(A, \mathbb{R})$), then for every $t \in A$ one has

$$\lim_{n\to\infty} x_n(t) = l(t).$$

The inverse does not hold. In fact, consider the normed space $B([0,1],\mathbb{R})$ with the norm $\|.\|_\infty$. For $n \in \mathbb{N}$, let $x_n : [0,1] \to \mathbb{R}$, again defined as (2.3.1), and let

$$\begin{cases} m : [0,1] \to \mathbb{R}, \\ m(t) = \begin{cases} 0 & \text{if} \quad t \in [0,1[, \\ 1 & \text{if} \quad t = 1. \end{cases} \end{cases}$$

It is very easy to check that $\lim_{n\to+\infty} x_n(t) = m(t)\ \forall t \in [0,1]$. However, it is not true that $\lim_{n\to+\infty} \|x_n - m\|_\infty = 0$. In fact, for every $n \in \mathbb{N}$,

$$|x_n(t) - m(t)| = \begin{cases} t^n & \text{if} \quad t \in [0,1[, \\ 0 & \text{if} \quad t = 1. \end{cases}$$

So $\|x_n - m\|_\infty = 1\ \forall n \in \mathbb{N}$.

A first important consequence of convergence of a sequence is its boundedness:

Theorem 2.3.6. *Let $(a_n)_{n\in\mathbb{N}}$ be convergent in the normed space X. Then $(a_n)_{n\in\mathbb{N}}$ is bounded, in the sense that there exists $M \geq 0$ such that*

$$\|a_n\| \leq M \qquad \forall n \in \mathbb{N}.$$

Proof Let $\lim_{n\to\infty} a_n = l \in X$. There exists $n(1) \in \mathbb{N}$ such that, if $n > n(1)$, one has $\|a_n - l\| < 1$. Consequently, if $n > n(1)$,

$$\|a_n\| = \|(a_n - l) + l\| \leq \|l\| + \|a_n - l\| < \|l\| + 1.$$

So, setting

$$M := \max\{\|a_1\|, ..., \|a_n\|, \|l\| + 1\},$$

one has $\|a_n\| \leq M\ \forall n \in \mathbb{N}$. $\square$

Now we are going to examine some important links between the notion of limit of a sequence and some of the notions introduces in Section 2.2.

Theorem 2.3.7. *Let $A \subseteq X$, with X normed space, and let $x_0 \in X$. Then the following facts are equivalent:*

(I) $x_0 \in \overline{A}$;

(II) there exists a sequence $(a_n)_{n\in\mathbb{N}}$ with values in A, which is convergent to x_0.

Proof We check that (I) implies (II). Let $x_0 \in \overline{A}$. We construct a sequence with values in A, convergent to x_0. If $x_0 \in \overline{A}$, then, either $x_0 \in A$, or $x_0 \in \partial A$. In the first case, we can take $a_n = x_0\ \forall n \in \mathbb{N}$. Instead, let $x_0 \in \partial A$. By definition, for every $n \in \mathbb{N}$ there exists $a_n \in A \cap B(x_0, \frac{1}{n})$. One has

$$\|a_n - x_0\| < \frac{1}{n} \to 0 (n \to +\infty).$$

So, (I) implies (II).

On the other hand, assume that (II) holds. We must show that $x_0 \in \overline{A}$. We consider separately the two cases $x_0 \in A$ and $x_0 \notin A$. In the first case, $x_0 \in \overline{A}$. In the second case, let $r > 0$. As $x_0 \notin A$, $B(x_0, r)$ contains some element not belonging to A: x_0 itself. Next, choosing $n \in \mathbb{N}$ in such a way that $\|a_n - x_0\| < r$, we obtain that $a_n \in A \cap B(x^0, r)$. So $x_0 \in \partial A$. $\square$

Theorem 2.3.8. *Let X and Y be normed spaces, with norms, respectively, $\|.\|_X$ and $\|.\|_Y$, let $A \subseteq X$, $f : A \to Y$, $x_0 \in A$. Then, the following facts are equivalent:*

(I) f is continuous in x_0;

(II) if $(a_n)_{n\in\mathbb{N}}$ is an arbitrary sequence with values in A and convergent in X to x_0, one has that the sequence $(f(a_n))_{n\in\mathbb{N}}$ converges in Y to $f(x_0)$.

Proof We show that (I) implies (II). Let $(a_n)_{n\in\mathbb{N}}$ be with values in A and convergent in X to x_0. Next, let $\epsilon \in \mathbb{R}^+$. As f is continuous in x_0, there exists $r(\epsilon) > 0$ such that, if $a \in A$ and $\|a - x_0\|_X < r(\epsilon)$, then $\|f(a) - f(x_0)\|_Y < \epsilon$. From the definition of convergence of a sequence, one has that there exists $n(r(\epsilon)) \in \mathbb{N}$ such that, if $n > n(r(\epsilon))$, one has $\|a_n - x_0\|_X < r(\epsilon)$. It follows that $\|f(a_n) - f(x_0)\|_Y < \epsilon$. So (II) holds.

On the other hand, we assume that (II) holds . We have to show that f is continuous in x_0. We argue by contradiction, assuming that it is not so. Therefore, there exists $\epsilon > 0$, such that, for every $r > 0$, there exists $a \in A$ with $\|a - x_0\|_X < r$ and $\|f(a) - f(x_0)\|_Y \geq \epsilon$. Let us take $r = 1, \frac{1}{2}, ..., \frac{1}{n}, ...$ and let us choose $a_n \in A \cap B(x_0, \frac{1}{n})$, with $\|f(a_n) - f(x_0)\|_Y \geq \epsilon$. Then, evidently, $(a_n)_{n\in\mathbb{N}}$ converges to x_0, but $(f(a_n))_{n\in\mathbb{N}}$ does not converge to $f(x_0)$. So (II) does not hold, and this is a contradiction. □

Now we are going to introduce and examine the notion of completeness in a normed space. We start with following

Definition 2.3.9. *Let X be a normed space and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. We shall say that $(x_n)_{n\in\mathbb{N}}$ is a* **Cauchy sequence** *if $\forall\epsilon \in \mathbb{R}^+$ there exists $n(\epsilon) \in \mathbb{N}$ such that, for every choice of m and n in $\mathbb{N}$, with $m > n(\epsilon)$ and $n > n(\epsilon)$, one has*

$$\|a_n - a_m\| < \epsilon.$$

Remark 2.3.10. Every convergent sequence is a Cauchy sequence. In fact, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X and let $x_0 \in X$ be such that $\|x_n - x_0\| \to 0$ $(n \to +\infty)$. Next, let $\epsilon > 0$ and $n(\frac{\epsilon}{2}) \in \mathbb{N}$ be such that $\|x_n - x_0\| < \frac{\epsilon}{2}$ if $n > n(\frac{\epsilon}{2})$. If we take n and m in $\mathbb{N}$, both larger than $n(\frac{\epsilon}{2})$, we have

$$\begin{aligned}\|x_n - x_m\| &= \|(x_n - x_0) + (x_0 - x_m)\| \\ &\leq \|x_n - x_0\| + \|x_0 - x_m\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon.\end{aligned}$$

In general, the converse does not hold: in a normed space there may exist Cauchy sequences which are not convergent.

Definition 2.3.11. *Let X be a normed space, with norm $\|.\|$. We shall say that X with such norm is* **complete** *if every Cauchy sequence is convergent. Complete normed spaces are also called* **Banach spaces**.

Now we are going to see some examples of Banach spaces. The first and fundamental result is the following

Theorem 2.3.12. $\mathbb{R}$*, with the absolute value as norm, is complete.*

This result can be esaily extended to $\mathbb{R}^n$:

Theorem 2.3.13. $\mathbb{R}^n$*, with the Euclidean norm, is complete.*

Proof Let $(x^k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $\mathbb{R}^n$ with respect to the Euclidean norm, which we indicate with $\|.\|$. We have to show that the sequence is convergent.

Given $k \in \mathbb{N}$, let $x^k = (x_1^k, ..., x_n^k)$. We fix $\epsilon > 0$. Then, there exists $k(\epsilon) \in \mathbb{N}$ such that, if i and j are natural numbers larger than $k(\epsilon)$, we have $\|x^i - x^j\| < \epsilon$. Let r be a natural number less or equal than n. We consider the sequence in $\mathbb{R}$ $(x_r^k)_{k\in\mathbb{N}}$. If i and j are natural numbers larger than $k(\epsilon)$, one has

$$|x_r^i - x_r^j\| \leq \|x^i - x^j\| < \epsilon.$$

Therefore, for $r = 1, ..., n$, $(x_r^k)_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. As $\mathbb{R}$ is complete, each sequence $(x_r^k)_{k\in\mathbb{N}}$ has a real limit x_r^0. We set

$$x^0 := (x_1^0, ..., x_n^0).$$

Then the sequence $(x^k)_{k\in\mathbb{N}}$ converges to x^0. In fact,

$$\|x^k - x^0\| = [\sum_{r=1}^{n}(x_r^k - x_r^0)^2]^{\frac{1}{2}} \to 0(k \to +\infty).$$

□

Now we consider the norm $\|.\|_\infty$.

Theorem 2.3.14. *Let $A \subseteq \mathbb{R}^n$ be nonempty. Then the space $B(A, \mathbb{R})$, normed with $\|.\|_\infty$, is complete.*

The space $BC(A, \mathbb{R})$ with the norm $\|.\|_\infty$ is complete.

Proof Let $(x^k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $B(A, \mathbb{R})$ with respect to the norm $\|.\|_\infty$. Let $\epsilon > 0$. Then, there exists $k(\epsilon) \in \mathbb{N}$ such that, if i and j are natural numbers larger than $k(\epsilon)$, one has $\|x^i - x^j\|_\infty < \epsilon$. Let a be an arbitrary element of A. We consider the sequence in $\mathbb{R}$ $(x^k(a))_{k\in\mathbb{N}}$. If i and j are natural numbers larger than $k(\epsilon)$, one has

$$|x^i(a) - x^j(a)| \leq \|x^i - x^j\|_\infty < \epsilon.$$

So, $\forall a \in A$, $(x^k(a))_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. As $\mathbb{R}$ is complete, each sequence $(x^k(a))_{k\in\mathbb{N}}$ has a real limit $x^0(a)$. Evidently, x^0 is a function from A to $\mathbb{R}$. We check that $x^0 \in B(A, \mathbb{R})$ and that

$$\|x^k - x^0\|_\infty \to 0(k \to +\infty).$$

Let $\epsilon > 0$ and let $k(\epsilon)$ be as before. Let $j > k(\epsilon)$. Next, let $i > k(\epsilon)$ and $a \in A$. One has

$$|x^i(a) - x^j(a)| \leq \|x^i - x^j\|_\infty < \epsilon. \tag{2.3.3}$$

Passing to the limit as $i \to +\infty$ in (2.3.3), we obtain that $|x^0(a) - x^j(a)| \leq \epsilon \; \forall a \in A$. In particular, from the case $\epsilon = 1$ we obtain, $\forall a \in A$ and for an arbitrary $j > k(1)$,

$$|x^0(a)| = |(x^0(a) - x^j(a)) + x^j(a)|$$

$$\leq |x^0(a) - x^j(a)| + |x^j(a)| \leq 1 + \|x^j\|_\infty.$$

So $x^0 \in B(A, \mathbb{R})$ and $\|x^0\|_\infty \leq 1 + \|x^j\|_\infty$. Moreover, if $j > k(\frac{\epsilon}{2})$, we have

$$\|x^0 - x^j\|_\infty \leq \frac{\epsilon}{2} < \epsilon.$$

From this we obtain the first statement.

Now we check the completeness of $BC(A, \mathbb{R})$. Let $(x^k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $BC(A, \mathbb{R})$. Evidently, it is a Cauchy sequence also in $B(A, \mathbb{R})$ and so, from the first statement, there exists $x^0 \in B(A, \mathbb{R})$ such that $\|x^k - x^0\|_\infty \to 0$ $(k \to \infty)$. To conclude, we have only to verify that $x^0 \in BC(A, \mathbb{R})$. As we already know that x^0 is bounded, we have only to check that it is continuous.

Let $a^0 \in A$. We verify that x^0 is continuous in a_0. Let $\epsilon > 0$. Let $k \in \mathbb{N}$ be such that $\|x^k - x^0\|_\infty < \frac{\epsilon}{3}$. As x^k is continuous, there exists $r > 0$ such that, if $a \in A$ and $\|a - a^0\| < r$, $|x^k(a) - x^k(a^0)| < \frac{\epsilon}{3}$. We such choice of a, one has

$$\begin{aligned} &|x^0(a) - x^0(a^0)| \\ &= |(x^0(a) - x^k(a)) + (x^k(a) - x^k(a^0)) + (x^k(a^0) - x^0(a^0))| \\ &\leq |x^0(a) - x^k(a)| + |x^k(a) - x^k(a^0)| + |x^k(a^0) - x^0(a^0)| \\ &\leq 2\|x^k - x^0\|_\infty + |x^k(a) - x^k(a^0)| < \epsilon. \end{aligned}$$

So x^0 is continuous in a^0. $\square$

One could also prove the following

Theorem 2.3.15. *Let $A \in \mathcal{M}_n$. Then $L^1(A)$, normed with $\|.\|_1$, is complete.*

Example 2.3.16. From Theorem 2.3.14, it follows that $C([0, 1], \mathbb{R})$ is a Banach space with the norm $\|.\|_\infty$. Let X be the linear space of polynomial functions with real coefficients, restricted to $[0, 1]$. A celebrated theorem due to Weierstrass asserts that X is dense in $C([0, 1], \mathbb{R})$. In other words, by virtue of Theorem 2.3.7, we can say that, given an arbitrary continuous function f with domain $[0, 1]$ and real valued, there exists a sequence of polynomial functions uniformly convergent to f. From this it immediately follows that X, with norm $\|.\|_\infty$, is not complete. To see this, it suffices to take a sequence $(x^k)_{k\in\mathbb{N}}$ of polynomial functions uniformly convergent to a continuous function which is not polynomial (for example, to a continuous function which is not differentiable somewhere). This sequence is a Cauchy sequence in X by Remark 2.3.10, but it is not convergent in X.

Exercise 2.3.17. Check that in any normed space the limit of a sequence (if existing) is unique.

Exercise 2.3.18. Check that in any normed space X, if $(x_n)_{n\in\mathbb{N}}$ converges to x_0 and $(y_n)_{n\in\mathbb{N}}$ converges to y_0, then $(x_n+y_n)_{n\in\mathbb{N}}$ converges to x_0+y_0 and, if λ belongs to K, $(\lambda x_n)_{n\in\mathbb{N}}$ converges to λx_0.

Exercise 2.3.19. With reference to Example 2.3.5, check that the sequence $(x_n)_{n\in\mathbb{N}}$ converges to 0 in $L^1([0, 1])$. Observe that $m(x) = 0$ almost everywhere in $[0, 1]$.

Exercise 2.3.20. Let X, Y and Z be three normed spaces, let $A \subseteq X$, $B \subseteq Y$, $f : A \to Y$, $g : B \to Z$, with $f(A) \subseteq B$. Next, let $x_0 \in A$ be such that f is continuous in x_0 and g is continuous in $f(x_0)$. Prove that $g \circ f$ is continuous in x_0. (Hint: apply Theorem 2.3.8).

Exercise 2.3.21. Verify that $\mathbb{C}$, with the absolute value as norm, is complete.

Exercise 2.3.22. Let $A \subseteq X$, with X normed space. Prove that the two following conditions are equivalent:

(I) A is closed;

(II) let $l \in X$ be such that there exists a sequence $(a_n)_{n\in\mathbb{N}}$ with values in A, convergent to l. Then $l \in A$.

Exercise 2.3.23. Study the convergence of the following sequences $(x_k)_{k\in\mathbb{N}}$ in the space $B(A,\mathbb{R})$ indicated:

(a) $A=[0,+\infty[$; $x_k(t)=[2\sinh(2kt)-e^{2kt}]t$;
(b) $A=]0,2]$; $x_k(t)=(\frac{t}{2})^k\ln(\frac{t}{2})$;
(c) $A=\mathbb{R}$; $x_k(t)=\sin(\frac{t}{k})-\cos(\frac{t}{k})$;
(d) $A=[0,2\pi]$; $x_k(t)=\sin(\frac{t}{k})-\cos(\frac{t}{k})$;
(e) $A=[0,+\infty[$; $x_k(t)=\frac{\pi}{2}-\arctan(kt)$;
(f) $A=[1,+\infty[$; $x_k(t)=\frac{\pi}{2}-\arctan(kt)$;
(g) $A=\mathbb{R}^+$; $x_k(t)=\frac{t}{1+k^2t^2}$;
(h) $A=\mathbb{R}^+$; $x_k(t)=\frac{t}{t+k}$;
(i) $A=[0,1]$; $x_k(t)=\frac{t}{t+k}$.

2.4 Spaces with inner product, Hilbert spaces

We start with the definition of inner product.

Definition 2.4.1. *Let X be a linear space over K (as usual, coinciding with $\mathbb{R}$ or $\mathbb{C}$), $<.,.>$ a map from $X\times X$ to K (that is, associating with each ordered pair (x,y) in $X\times X$ an element $<x,y>$ belonging to K). We shall say that $<.,.>$ is an* **inner (scalar) product** *in X if the following conditions hold:*

(IP1) $\forall y\in X$ $x\to<x,y>$ is linear from X to K;
(IP2) $<y,x>=\overline{<x,y>}$ $\forall x,y\in X$;
(IP3) $<x,x>$ is a nonnegative real number $\forall x\in X$;
(IP4) if $x\in X$ and $<x,x>=0$, then $x=0$.

Remark 2.4.2. $(IP1)$ means that $\forall\lambda_1,\lambda_2\in K$, $\forall x^1$, x^2, $y\in X$ one has

$$<\lambda_1x^1+\lambda_2x^2,y>=\lambda_1<x^1,y>+\lambda_2<x^2,y>.$$

Next, from (IPI) and $(IP2)$ it follows that, $\forall\lambda_1,\lambda_2\in K$, $\forall x,y^1,y^2\in X$,

$$<x,\lambda_1y^1+\lambda_2y^2>=$$
$$\overline{<\lambda_1y^1+\lambda_2y^2,x>}=\overline{\lambda_1<y^1,x>+\lambda_2<y^2,x>}=$$
$$=\overline{\lambda_1}<x,y^1>+\overline{\lambda_2}<x,y^2>.$$

It is customary to say that, $\forall x\in X$, $y\to<x,y>$ is antilinear from X to K. In case $K=\mathbb{R}$, this means that a scalar product is linear also in its second term.

Example 2.4.3. Let $X=\mathbb{R}^n$ and $K=\mathbb{R}$. If $x=(x_1,...,x_n)$ and $y=(y_1,...,y_n)$, we set

$$<x,y>:=x_1y_1+...+x_ny_n. \tag{2.4.1}$$

It is easy to check that the mapping defined in (2.4.1) is an inner produch in $\mathbb{R}^n$ (which probably the reader already knows).

Example 2.4.4. Let $X=\mathbb{C}^n$ and $K=\mathbb{C}$. If $x=(x_1,...,x_n)$ and $y=(y_1,...,y_n)$, we set

$$<x,y>:=x_1\overline{y_1}+...+x_n\overline{y_n}. \tag{2.4.2}$$

The mapping defined in (2.4.2) is an inner product in $\mathbb{C}^n$. We leave the easy proof to the reader (see Exercise 2.4.16).

Example 2.4.5. Let $A \in \mathcal{M}_n$ and let $f : A \to \mathbb{C}$ be measurable. We shall say that f belongs to $\mathcal{L}^2(A)$ if

$$\int_A |f(x)|^2 dx < +\infty.$$

If a and b are nonnegative real numbers , we have

$$0 \leq (a-b)^2 = a^2 + b^2 - 2ab.$$

As a consequence, we obtain the inequality

$$ab \leq \frac{1}{2}(a^2 + b^2). \tag{2.4.3}$$

So let f and g be elements of $\mathcal{L}^2(A)$. It $x \in A$, we have

$$\begin{aligned} |f(x)+g(x)|^2 &\leq (|f(x)| + |g(x)|)^2 \\ &= |f(x)|^2 + |g(x)|^2 + 2|f(x)||g(x)| \\ &\leq 2(|f(x)|^2 + |g(x)|^2), \end{aligned}$$

using (2.4.3). Hence

$$\begin{aligned} &\int_A |f(x)+g(x)|^2 dx \\ &\leq \int_A 2(|f(x)|^2 + |g(x)|^2) dx \\ &= 2(\int_A |f(x)|^2 dx + \int_A |g(x)|^2) < +\infty. \end{aligned}$$

Therefore, $f + g \in \mathcal{L}^2(A)$. It is also easy to see that, if $\lambda \in \mathbb{C}$ and $f \in \mathcal{L}^2(A)$, even $\lambda f \in \mathcal{L}^2(A)$. We deduce that $\mathcal{L}^2(A)$ is a linear space over $\mathbb{C}$ with the usual operations of sum and scalar multiplication. Moreover, it is easy to check that, if $f \in \mathcal{L}^2(A)$ and $f \sim g$ in the sense of Definition 1.4.1, then even $g \in \mathcal{L}^2(A)$ (see Exercise 2.4.17).

Definition 2.4.6. *Let $A \in \mathcal{M}_n$. We set*

$$L^2(A) := \{[f] : f \in \mathcal{L}^2(A)\}.$$

$L^2(A)$ is a linear space over $\mathbb{C}$ with the sum and scalar multiplication defined in (1.4.2) and (1.4.3), as it is easy to check.

Let now f and g be elements of $\mathcal{L}^2(A)$. Applying again (2.4.3), we have, $\forall x \in A$,

$$|f(x)\overline{g(x)}| = |f(x)||g(x)| \leq \frac{1}{2}(|f(x)|^2 + |g(x)|^2).$$

It follows that $f\overline{g}$ is summable in A. So let $[f]$ and $[g]$ be elements of $L^2(A)$. We set

$$< [f],[g] > := \int_A f(x)\overline{g(x)} dx. \tag{2.4.4}$$

It is easy to see (applying Theorem 1.4.2) that (2.4.4) does not depend on the choice of the elements in the equivalence classes. We check that $< .,. >$ is an inner product in $L^2(A)$.

We leave to the reader the inspection of $(IP1)$.

We check $(IP2)$. One has

$$< [g], [f] >=$$

$$= \int_A g(x)\overline{f(x)}dx = \int_A \overline{f(x)\overline{g(x)}}dx$$

$$= \overline{\int_A f(x)\overline{g(x)}dx}$$

(applying formula (1.2.9)

$$= \overline{< [f], [g] >}.$$

We check $(IP3)$. if $f \in \mathcal{L}^2(A)$,

$$< [f], [f] >= \int_A |f(x)|^2 dx \geq 0.$$

Finally, we check $(PI4)$. The condition $< [f], [f] >= 0$ is equivalent to $\int_A |f(x)|^2 dx = 0$. From Lemma 2.1.11, it follows that $|f(x)|^2 = 0$ a. e. in A, and so $f(x) = 0$ a.e. in A. We conclude that

$$[f] = [0].$$

Now we examine how it is possible to intoduce a norm in a space with an inner product, which is intimately connected with the inner product itself.

Definition 2.4.7. *Let X be a linear space over K, in which an inner product $< .,. >$ has been introduced. We set, $\forall x \in X$,*

$$\|x\| := \sqrt{< x, x >}.$$

We shall name the mapping $x \to \|x\|$ the **norm associated** *with the inner product $< .,. >$.*

Remark 2.4.8. We observe that Definition 2.4.7 is well posed by virtue of condition $(PI3)$. We shall soon check that $\|.\|$ is really a norm, in the sense of Definition 2.1.6.

In order to prove that $\|.\|$ is a norm, we shall employ the following

Theorem 2.4.9. *Let X be a linear space over K, with inner product $< .,. >$. Next, let $\|.\|$ be as in Definition 2.4.7. Then, $\forall x, y \in X$*

$$| < x, y > | \leq \|x\|\|y\|.$$

(Cauchy-Schwarz inequality)

Proof Let $\alpha \in K$, which we are going to specify later. Then, $\forall t \in \mathbb{R}$, using $(PI1) - (PI3)$,

$$0 \leq < x + t\alpha y, x + t\alpha y >$$

$$=< x, x > +t(\alpha < y, x > +\overline{\alpha} < x, y >) + t^2|\alpha|^2 < y, y >$$

$$= \|x\|^2 + 2tRe(\alpha < y, x >) + t^2|\alpha|^2\|y\|^2,$$

as

$$\alpha < y, x > +\overline{\alpha} < x, y >= \alpha < y, x > +\overline{\alpha < y, x >}$$

$$= 2Re(\alpha < y, x >).$$

We choose

$$\alpha = \begin{cases} \frac{<x,y>}{|<x,y>|} & \text{if} \quad < x, y >\neq 0, \\ 1 & \text{if} \quad < x, y >= 0. \end{cases}$$

We observe that in any case $|\alpha| = 1$ and we have

$$P(t) :=< x + t\alpha y, x + t\alpha y >= \|x\|^2 + 2t| < x, y > | + t^2\|y\|^2 \quad \forall t \in \mathbb{R}.$$

P is a polynomial of degree not exceedin two, with real coefficients, and $P(t) \geq 0\ \forall t \in \mathbb{R}$. This implies that its discriminant is not positive, that is,

$$| < x, y > |^2 - \|x\|^2\|y\|^2 \leq 0,$$

which implies immediately the conclusion. □

Theorem 2.4.10. *Let X be a linear space over K and $< .,. >$ an inner product in in X. Then the norm associated with $< .,. >$ (Definition 2.4.7) is a norm, in the sense of Definition 2.1.6.*

Proof We indicate, as usual, with $\|.\|$ the norm associated with $< .,. >$. Let $\lambda \in K$ and $x \in X$. Then

$$\|\lambda x\| = \sqrt{< \lambda x, \lambda x >} = \sqrt{|\lambda|^2 < x, x >} = |\lambda|\|x\|.$$

Next, if x and y are elements of X, we have

$$\|x + y\|^2 =< x + y, x + y >= \|x\|^2 + 2Re(< x, y >) + \|y\|^2$$

$$\leq \|x\|^2 + 2|Re(< x, y >)| + \|y\|^2 \leq \|x\|^2 + 2| < x, y > | + \|y\|^2$$

$$\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2$$

(by Cauchy-Schwarz inequality)

$$= (\|x\| + \|y\|)^2,$$

from which, taking the square roots, we obtain

$$\|x + y\| \leq \|x\| + \|y\|.$$

Finally, if $\|x\| = 0$, then $< x, x >= 0$, which implies $x = 0$. □

Example 2.4.11. We consider Example 2.4.3. Given $x = (x_1, ..., x_n) \in \mathbb{R}^n$, we obtain from the scalar product the norm

$$\|x\| = \sqrt{< x, x >} = \sqrt{x_1^2 + ... + x_n^2},$$

which is the well known Euclidean norm.

Example 2.4.12. We consider Example 2.4.4. Given $x = (x_1, ..., x_n) \in \mathbb{C}^n$, we obtain from the scalar product the norm

$$\|x\| = \sqrt{< x, x >} = \sqrt{|x_1|^2 + ... + |x_n|^2}. \tag{2.4.5}$$

Example 2.4.13. We consider Example 2.4.5. From the scalar product defined in $L^2(A)$, we obtain the norm

$$\|[f]\| = \sqrt{< [f], [f] >} = \sqrt{\int_A |f(x)|^2 dx}. \tag{2.4.6}$$

In this case, Cauchy-Schwarz inequality gives the following integral inequality, which is valid for f and g in $\mathcal{L}^2(A)$:

$$|\int_A f(x)\overline{g(x)}dx| \leq \sqrt{\int_A |f(x)|^2 dx}\sqrt{\int_A |g(x)|^2 dx}. \tag{2.4.7}$$

We conclude with the definition of Hilbert space.

Definition 2.4.14. *Let X be a normed space, with a norm $\|.\|$ associated with a certain inner product $< .,. >$. We shall say that with such norm X is a* **Hilbert space** *if it is complete, in the sense of Definition 2.3.11.*

We have already seen (Theorem 2.3.13) that $\mathbb{R}^n$, with the Euclidean norm, is complete and is, therefore, a Hilbert space. Even $\mathbb{C}^n$ with the norm 2.4.5, is complete (Exercise 2.4.18). Next, it is possible to prove the following

Theorem 2.4.15. *Let $A \in \mathcal{M}_n$. Then the space $L^2(A)$, with the norm (2.4.6), is complete and so it is a Hilbert space.*

Exercise 2.4.16. Check in detail that the map (2.4.2) is an inner product in $\mathbb{C}^n$. (Hint: start by showing that, if z and v are elements of $\mathbb{C}$, then $\overline{z\overline{v}} = \overline{z}v$. We recall that $z\overline{z} = |z|^2$.)

Exercise 2.4.17. Check that, if $f \in \mathcal{L}^2(A)$ and $f \sim g$, then $g \in \mathcal{L}^2(A)$.

Exercise 2.4.18. Check that $\mathbb{C}^n$, with the norm (2.4.5), is complete. (Hint.: observe that with such norm $\mathbb{C}^n$ coincides with $\mathbb{R}^{2n}$, equipped with the corresponding Euclidean norm.)

2.5 Orthogonal projections in Hilbert spaces

We start with the (probably) most important result in the theory of Hilbert spaces.

Recalling the usual interpretation of the standard scalar product in $\mathbb{R}^n$, we shall say that two elements x and y in a space with inner product $< .,. >$ are orthogonal if $< x, y >= 0$.

Theorem 2.5.1. *Let H be a Hilbert space, $A \subseteq H$ nonempty, convex and closed. Next, let $x \in H$. Then the mapping $y \to \|y - x\|$ has a minimum in A. Such minimum is taken in a unique point $P_A x \in A$.*

We want to give a complete proof of Theorem 2.5.1, also in order to show the role of completeness in the space H.

We start with the following simple

Lemma 2.5.2. *(The parallelogram lemma) Let H be a Hilbert space and let x and y be elements of H. Then*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Proof We have

$$\|x + y\|^2 + \|x - y\|^2 = \|x\|^2 + 2Re < x, y > + \|y\|^2 + \|x\|^2 - 2Re < x, y > + \|y\|^2.$$

□

Remark 2.5.3. The parallelogram lemma is a generalization of the well known fact that, in a parallelogram, the sum of the areas of the squares constructed on the diagonals is double the sum of the areas of the squares constructed on the sides.

Proof of Theorem 2.5.1 We indicate with ν the infimum of the function $y \to \|y-x\|$ restricted to A. Obviously, $\nu \geq 0$. We take a sequence $(y_n)_{n\in\mathbb{N}}$ in A, such that

$$\lim_{n\to+\infty} \|y_n - x\| = \nu. \tag{2.5.1}$$

It is clear that a sequence of this type exists: in fact, by the definition of infimum, for every $n \in \mathbb{N}$ there exists $y_n \in A$, such that

$$\nu \leq \|y_n - x\| \leq \nu + 2^{-n}.$$

Now we show that the sequence $(y_n)_{n\in\mathbb{N}}$ is convergent. As H is complete, it suffices to show that it is a Cauchy sequence. In fact, by the parallelogram lemma, $\forall n, m \in \mathbb{N}$,

$$\|y_n - y_m\|^2 = \|(y_n - x) - (y_m - x)\|^2 = 2(\|y_n - x\|^2 + \|y_m - x\|^2) - \|y_n + y_m - 2x\|^2.$$

As A is convex, $\frac{1}{2}(y_n + y_m) \in A$. It follows that

$$\|y_n + y_m - 2x\|^2 = 4\|\frac{1}{2}(y_n + y_m) - x\|^2 \geq 4\nu^2.$$

We deduce that

$$\|y_n - y_m\|^2 \leq 2(\|y_n - x\|^2 + \|y_m - x\|^2) - 4\nu^2,$$

and, as (2.5.1) holds, there exists $n(\epsilon)$ in $\mathbb{N}$, such that, if both n and m are larger than $n(\epsilon)$,

$$\|y_n - y_m\|^2 < \epsilon^2.$$

So, $(y_n)_{n\in\mathbb{N}}$ is convergent to some element $y_0 \in H$. By Theorem 2.3.7, $y_0 \in \overline{A} = A$, because A is closed. By the result of Exercise 2.2.10,

$$\|y_0 - x\| = \lim_{n\to+\infty} \|y_n - x\| = \nu.$$

We deduce that the mapping $y \to \|y - x\|$ has minimum in A. It remains to show that, if y_0 and y_1 are elements of A, such that

$$\|y_0 - x\| = |y_1 - x\| = \nu,$$

then $y_0 = y_1$. To this aim, we consider the following function

$$\begin{cases} \phi : [0,1] \to \mathbb{R}, \\ \\ \phi(t) = \|(1-t)y_0 + ty_1 - x\|^2. \end{cases}$$

We start by observing that, as A is convex, $\forall t \in [0,1]$ $(1-t)y_0 + ty_1 = y_0 + t(y_1 - y_0) \in A$. We deduce that $\phi(t) \geq \nu^2$ $\forall t \in [0,1]$. On the other hand,

$$\|(1-t)y_0 + ty_1 - x\| = \|(1-t)(y_0 - x) + t(y_1 - x)\| \leq (1-t)\|y_0 - x\| + t\|y_1 - x\| = \nu,$$

so that $\phi(t) = \nu^2$ $\forall t \in [0,1]$. We observe that

$$\phi(t) = \|t(y_1 - y_0) + y_0 - x\|^2 = t^2\|y_1 - y_0\|^2 + 2tRe < y_0 - x, y_1 - y_0 > +\|y_0 - x\|^2,$$

which implies that, as ϕ is constant,

$$0 = \phi''(1) = 2\|y_1 - y_0\|^2.$$

So $y_0 = y_1$ and the proof is complete. □

Remark 2.5.4. In essence, Theorem 2.5.1 states that, if A is convex and closed, for each point $x \in H$, there exists a unique point $P_A x$ in A, which has minimum distance from x. In the case that A is a closed subspace of H, $P_A x$ may be thought as the orthogonal projection of x onto A, on the basis of the following result.

Theorem 2.5.5. *Let H be a Hilbert space, A a closed subspace of H. Next, let $x \in H$. Then, given $y \in A$, the following conditions are equivalent:*
(I) $y = P_A x$;
(II) $< x - y, z >= 0 \ \forall z \in A$.

Proof We check that, if $y = P_A x$, $< x - y, z >= 0 \ \forall z \in A$. Let $z \in A$ and $t \in \mathbb{R}$. As $y + tz \in A \ \forall t \in \mathbb{R}$, the function

$$\begin{cases} P : \mathbb{R} \to \mathbb{R}, \\ P(t) := \|x - y - tz\|^2 \end{cases}$$

has minimum in $t = 0$. One has

$$P(t) = \|x - y\|^2 - 2tRe(< x - y, z >) + t^2 \|z\|^2.$$

So P is a polynomial function. Necessarily, $P'(0) = 0$, which implies

$$Re(< x - y, z >) = 0.$$

If $K = \mathbb{C}$, even $iz \in A$ and so we have also $Re(< x - y, iz >) = 0$; this means that

$$Im(< x - y, z >) = Re(-i < x - y, z >) = Re(< x - y, iz >) = 0.$$

With this we have proved that (I) implies (II).

On the other hand, let $y \in A$ be such that $< x - y, z >= 0 \ \forall z \in A$. Then, given $z \in A$, we have

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2$$

$$= \|x - y\|^2 + 2Re(< x - y, y - z >) + \|y - z\|^2 = \|x - y\|^2 + \|y - z\|^2$$

(because $y - z \in A$)

$$\geq \|x - y\|^2.$$

So $y = P_A x$. $\square$

Now we want to give an explicit formula for P_A in the case that A is a finite dimensional subspace of H. We start with the following preliminary

Definition 2.5.6. *Let $\{x^i : i \in \mathcal{I}\}$, with $\mathcal{I}$ arbitrary set of indexes, a subset of the Hilbert space H. We shall say that $\{x^i : i \in \mathcal{I}\}$ is an* **orthonormal system** *if, $\forall i, j \in \mathcal{I}$,*

$$< x^i, x^j >= \delta_{ij}$$

with δ_{ij} Kronecker's delta.

Remark 2.5.7. In essence, in Definition 2.5.6 we require that the vectors x^i are pairwise orthogonal and have unitary norm. For example, the standard basis $\{e^i : i \in \{1, ..., n\}\}$ of $\mathbb{R}^n$ is an orthonormal system with respect to the standard inner product (see Example 2.4.3). The same happens in $\mathbb{C}^n$ with the inner product described in Example 2.4.4.

As we are going to see, it is helpful to have an orthonormal basis at disposal. To this aim, we state the following

Theorem 2.5.8. *Let A be a finite dimensional subspace of the Hilbert space H. Then A has an orthonormal basis with respect to the inner product in H.*

Proof The proof we are giving is constructive, in the sense that it describes a method (the so called **Gram-Schmidt method**) to construct an orthonormal basis of A starting from an arbitrary basis $\{f^i : 1 \leq i \leq n\}$, with n dimension of A. For simplicity, we are going to consider the case $n = 3$, even if the method is completely general and follows the lines of the particular case we are going to treat.

Let $\{f^1, f^2, f^3\}$ be a basis of A. We set

$$e^1 := \frac{1}{\|f^1\|} f^1.$$

(observe that $f^1 \neq 0$). One has $\|e^1\| = 1$.

We look for an element $\epsilon^2 \in A$ in the form

$$\epsilon^2 = ce^1 + f^2,$$

which is orthogonal to e^1. We get

$$0 =< \epsilon^2, e^1 >= c+ < f^2, e^1 > .$$

So, we have to take

$$c = - < f^2, e^1 > .$$

We observe that, whatever c is, $\epsilon^2 \neq 0$, as e^1 and f^2 are linearly independent. Now we set

$$e^2 := \frac{1}{\|\epsilon^2\|} \epsilon^2.$$

It is clear that e^1 and e^2 are orthogonal and with unitary norm. Moreover, they are both linear combinations of f^1 and f^2.

Finally, we look for an element $\epsilon^3 \in A$ in the form

$$\epsilon^2 = c_1 e^1 + c_2 e^2 + f^3,$$

which is orthogonal to both e^1 and e^2. We obtain

$$0 =< \epsilon^3, e^1 >= c_1+ < f^3, e^1 >,$$

$$0 =< \epsilon^3, e^2 >= c_2+ < f^3, e^2 > .$$

So we have to take

$$c_1 = - < f^3, e^1 >, c_2 = - < f^3, e^2 > .$$

We have that $\epsilon^3 \neq 0$, as f^3 does not depend linearly on e^1 and e^2, because it does not depend on f^1 and f^2. So we set

$$e^3 := \frac{1}{\|\epsilon^3\|} \epsilon^3,$$

and $\{e^1, e^2, e^3\}$ is an orthonormal basis. □

Example 2.5.9. Let $H = L^2([-1,1])$, and let A be the subspace of the equivalence classes containing the restrictions to $[-1,1]$ of polynomial functions of degree less or equal to 2. It is obvious that A is a finite dimensional subspace and a basis of A is $\{[1],[t],[t^2]\}$. In the following, given $f \in \mathcal{L}^2([0,1])$, we shall write f instead of $[f]$, for simplicity of notation. Let $f^1(t) = 1$, $(t \in [-1,1])$. One has

$$\|f^1\| = (\int_{-1}^{1} 1^2 dt)^{\frac{1}{2}} = \sqrt{2}.$$

So we set $e^1(t) := \frac{1}{\sqrt{2}}$ $(t \in [-1,1])$. Following the general method in the proof of Theorem 2.5.8, we set

$$c := - < f^2, e^1 >= - \int_{-1}^{1} \frac{t}{\sqrt{2}} dt = 0.$$

So,

$$\epsilon^2(t) = t, t \in [-1,1].$$

One has

$$\|\epsilon^2\| = (\int_{-1}^{1} t^2 dt)^{\frac{1}{2}} = \sqrt{\frac{2}{3}},$$

so that

$$e^2(t) = \sqrt{\frac{3}{2}} t.$$

Finally, we set

$$c_1 := - < f^3, e^1 >= - \int_{-1}^{1} \frac{t^2}{\sqrt{2}} dt = - \frac{\sqrt{2}}{3},$$

$$c_2 := - < f^3, e^2 >= - \int_{-1}^{1} t^2 (\sqrt{\frac{3}{2}} t) dt = 0.$$

So we have

$$\epsilon^3(t) = t^2 - \frac{1}{3}.$$

Next, one can verify that $\|\epsilon^3\| = \frac{2}{3}\sqrt{\frac{2}{5}}$. So we set

$$e^3(t) = \frac{3}{2}\sqrt{\frac{5}{2}} (t^2 - \frac{1}{3}).$$

$\{e^1, e^2, e^3\}$ is an orthonormal basis of A.

Now, as promised, we are going to exibit a formula for P_A in the case that A is a finite dimensional subspace of H.

Theorem 2.5.10. *Let H be a Hilbert space, A a finite dimensional subspace, and let $\{e^1, ..., e^n\}$ be an orthonormal basis of A. Then A is closed. Moreover, $\forall x \in H$,*

$$P_A x = \sum_{i=1}^{n} < x, e^i > e^i. \tag{2.5.2}$$

Proof The statement that A is closed is left to the reader (Exercise 2.5.14). Concerning (2.5.2), as $P_A x \in A$, it can be written in the form

$$P_A x = c_1 e^1 + ... + c_n e^n,$$

with c_1, ..., c_n elements of the field K. Let $1 \leq j \leq n$. From Theorem 2.5.5 it follows that

$$0 =< x - \sum_{i=1}^{n} c_i e^i, e^j >=< x, e^j > - c_j,$$

and the conclusion follows. □

Example 2.5.11. Let $H = L^2([0,1])$ and let A be the subspace of H of polynomial functions of degree less or equal than 2. Let $f(t) = \sin(t)$. f is an element of H. $P_A f$ may be interpreted as the best approximation of f in A, in the sense that $\int_{[-1,1]} |f(t) - P(t)|^2 dt$, with P polynomial of degree less or equal to 2, takes the minimum value exactly for $P = P_A f$. In Example 2.5.9 we have determined an orthonormal basis $\{e^1, e^2, e^3\}$ of A. Using the calculations already done, we obtain that

$$P_A f(t) = \sum_{i=1}^{3} < f, e^i > e^i(t) =$$

$$= \frac{1}{\sqrt{2}} \int_{-1}^{1} \frac{\sin(s)}{\sqrt{2}} ds + \int_{-1}^{1} \sin(s) \sqrt{\frac{3}{2}} s ds \sqrt{\frac{3}{2}} t$$

$$+ \int_{-1}^{1} \sin(s) \frac{3}{2} \sqrt{\frac{5}{2}} (s^2 - \frac{1}{3}) ds \ \frac{3}{2} \sqrt{\frac{5}{2}} (t^2 - \frac{1}{3}) =$$

$$= 3(\sin(1) - \cos(1))t.$$

We conclude this section with the following

Theorem 2.5.12. *Let $(V_n)_{n \in \mathbb{N}}$ be a sequence of closed subspaces of the Hilbert space H, such that $V_n \subseteq V_{n+1}$ $\forall n \in \mathbb{N}$ and $V := \bigcup_{n \in \mathbb{N}} V_n$ is dense in H. Given $n \in \mathbb{N}$, we indicate with P_n the orthogonal projection onto V_n. Then, $\forall x \in H$*

$$\lim_{n \to +\infty} P_n x = x$$

holds.

Proof Let $\epsilon > 0$. Applying Theorem 2.3.7, we can say that there exists $v_0 \in V$, such that $\|x - v_0\| < \epsilon$. Let $v_0 \in V_{n_0}$, with $n_0 \in \mathbb{N}$. If $n > n_0$, we shall have that, keeping into account that $v_0 \in V_n$,

$$\|x - P_n x\| \leq \|x - v_0\| < \epsilon.$$

□

Exercise 2.5.13. Show that, if A is a closed subspace of H, the operator $P_A : H \to A$ is linear. (Hint.: use Theorem 2.5.5)

Exercise 2.5.14. Show that, if A is a finite dimensional subspace of the Hilbert space H, then A is closed. (Hint: let $\{e^1, ..., e^n\}$ be an orthonormal basis of A. If $x = x_1 e^1 + ... + x_n e^n$, then

$$\|x\| = (|x_1|^2 + ... + |x_n|^2)^{\frac{1}{2}} = \|(x_1, ..., x_n)\|_{\mathbb{C}^n},$$

where $\|.\|_{\mathbb{C}^n}$ is the norm in $\mathbb{C}^n$ introduced in Example 2.4.4. Use the result of Exercise 2.4.18.

Exercise 2.5.15. Construct a basis of the subspace of polynomial functions of degree less or equal to 3, which is orthonormal with respect to the inner product in $L^2([-1,1])$.

Exercise 2.5.16. Let x^1,...,x^n be elements of H, which are pairwise orthogonal. Show that

$$\|x^1 + ... + x^n\|^2 = \|x^1\|^2 + ... + \|x^n\|^2.$$

This formula may be considered a generalization of the classical Pithagoras' theorem.

2.6 Fourier series

We start by recalling that, given $y \in \mathbb{R}$, by definition

$$e^{iy} = \cos(y) + i\sin(y).$$

Observe that one has

$$\overline{e^{ix}} = \cos(x) - i\sin(x) = \\ = \cos(-x) + i\sin(-x) = e^{-ix}.$$

Moreover, the basic formula

$$e^{ix}e^{iy} = e^{i(x+y)} \qquad \forall x, y \in \mathbb{R}, \tag{2.6.1}$$

holds, from which it follows

$$(e^{ix})^n = e^{inx} \qquad \forall n \in \mathbb{N}$$

Next, taking into account that $e^{i0} = 0$, from (2.6.1) we see that

$$(e^{ix})^{-1} = e^{-ix} \qquad \forall x \in \mathbb{R}.$$

Finally, we observe that

$$\cos(x) = Re(e^{ix}) = \frac{e^{ix} + \overline{e^{ix}}}{2} = \frac{e^{ix} + e^{-ix}}{2} \tag{2.6.2}$$

and

$$\sin(x) = Im(e^{ix}) = \frac{e^{ix} - \overline{e^{ix}}}{2i} = \frac{e^{ix} - e^{-ix}}{2i} \tag{2.6.3}$$

(2.6.2) and (2.6.3) are the well known **Euler's formulas**. Let now m and n be integers, with $m \neq n$. One has

$$\begin{gathered} \int_{-\pi}^{\pi} e^{inx}\overline{e^{imx}}dx = \\ = \int_{-\pi}^{\pi} e^{i(n-m)x}dx = \int_{-\pi}^{\pi} \cos((n-m)x)dx + i\int_{-\pi}^{\pi} \sin((n-m)x)dx = \\ = 0. \end{gathered} \tag{2.6.4}$$

On the other hand,

$$\int_{-\pi}^{\pi} e^{inx}e^{-inx}dx = 2\pi. \tag{2.6.5}$$

From (2.6.4) and (2.6.5) we deduce that $\{\frac{e^{inx}}{\sqrt{2\pi}} : n \in \mathbb{Z}\}$ is an orthonormal system in $L^2([-\pi,\pi])$. Given $n \in \mathbb{N}$, we indicate with V_n the subspace of $L^2([-\pi,\pi])$ generated by $\{\frac{e^{ikx}}{\sqrt{2\pi}} : -n \leq k \leq n\}$. If $f \in L^2([-\pi,\pi])$, the orthogonal projection P_nf of f onto V_n is given by

$$P_n f(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(s)\sum_{k=-n}^{n} e^{ik(t-s)}ds \tag{2.6.6}$$

Taking into account the formula of the sum of the term in a geometric progression, one has, for $e^{i(t-s)} \neq 1$,

$$\sum_{k=-n}^{n} e^{ik(t-s)} = e^{-in(t-s)}\sum_{k=0}^{2n} e^{ik(t-s)} = e^{-in(t-s)}\frac{1-e^{i(2n+1)(t-s)}}{1-e^{i(t-s)}}$$
$$= e^{-in(t-s)}\frac{(1-e^{i(2n+1)(t-s)})(1-e^{-i(t-s)})}{|1-e^{i(t-s)}|^2}.$$

One has

$$|1-e^{i(t-s)}|^2 = [1-\cos(t-s)]^2 + \sin^2(t-s)$$
$$= 2 - 2\cos(t-s) = 4\sin^2(\frac{t-s}{2}).$$
$$e^{-in(t-s)}(1-e^{i(2n+1)(t-s)})(1-e^{-i(t-s)}) =$$
$$= e^{in(t-s)} + e^{-in(t-s)} - e^{i(n+1)(t-s)} - e^{-i(n+1)(t-s)}$$
$$= 2[\cos(n(t-s)) - \cos((n+1)(t-s))]$$

(by Euler's formulas)

$$= 4\sin((n+\frac{1}{2})(t-s))\sin(\frac{t-s}{2})$$

(by the prostaphaeresis formulas). We conclude that, if $e^{i(t-s)} \neq 1$, one has

$$\sum_{k=-n}^{n} e^{ik(t-s)} = \frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})}.$$

Extending arbitrarily $\frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})}$ to the points (in finite number in $[-\pi,\pi]$) such that $\sin(\frac{t-s}{2}) = 0$, we have obtained

$$P_n f(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\sin((n+\frac{1}{2})(t-s))}{\sin(\frac{t-s}{2})} f(s)ds. \tag{2.6.7}$$

We set $V := \bigcup_{n\in\mathbb{N}} V_n$. It is possible to show that V is dense in $L^2([-\pi,\pi])$. Applying Theorem 2.5.12, we get the following

Theorem 2.6.1. *Let $f \in L^2([-\pi,\pi])$. Then the sequence $(P_n f)_{n\in\mathbb{N}}$, with $P_n f$ given by(2.6.7), converges to f, for $n \to +\infty$, in $L^2([-\pi,\pi])$.*

We shall usually write

$$f = \sum_{n=-\infty}^{+\infty} \frac{e^{int}}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-ins}ds, \tag{2.6.8}$$

in the sense that $f = \lim_{n\to+\infty}\sum_{k=-n}^{n} \frac{e^{ik.}}{2\pi}\int_{-\pi}^{\pi} f(s)e^{-ins}ds$ in $L^2([-\pi,\pi])$. We observe that, generally speaking, there is no pointwise convergence. Observe, by the way, that, as elements of $L^2([-\pi,\pi])$ are equivalence classes of functions and not functions, the value of $f(t)$, as depending on the specific element in the class, is not defined. One might wonder whether, for example, given $f \in C([-\pi, \pi])$, it holds (point by point) $\lim_{n \to +\infty} P_n f(t) = f(t)$, with $P_n f(t)$ defined in (2.6.7). It is possible to show the following

Theorem 2.6.2. *Let $f \in C([-\pi, \pi])$ be such that*
(I) $f(-\pi) = f(\pi)$;
(II) there exist C and α positive, such that

$$|f(t) - f(s)| \leq C|t - s|^{\alpha} \tag{2.6.9}$$

$\forall t, s \in [-\pi, \pi]$. Then the sequence $(P_n f)_{n \in \mathbb{N}}$ converges to f uniformly in $[-\pi, \pi]$.

Remark 2.6.3. Applying the mean value theorem, it is not diificult to check that, if $Re(f)$ and $Im(f)$ are of class C^1, then assumption (II) of Theorem 2.6.2 is satisfied, with $\alpha = 1$. In general, one could show that, if $f \in C([-\pi, \pi])$ and (I) holds, the conclusion of Theorem 2.6.2 is false.

Remark 2.6.4. Given $f \in L^2([-\pi, \pi]$, we shall call (2.6.8) the classical **Fourier expansion** of f.

In general, if $S := \{e^n : n \in \mathbb{N}\}$ is an orthonormal system in the Hilbert space H, such that the linear space generated by S is dense in H, one has that, $\forall f \in H$,

$$f = \lim_{n \to +\infty} \sum_{k=1}^{n} < f, e^k > e^k. \tag{2.6.10}$$

We shall write

$$f = \sum_{n=1}^{\infty} < f, e^n > e^n \tag{2.6.11}$$

and we shall speak even in this case of the Fourier expansion of f. We shall see a significant example in the following section.

Exercise 2.6.5. Check the assertion in the first part of Remark 2.6.3.

Exercise 2.6.6. Check that, if $f = \sum_{n=1}^{\infty} < f, e^n > e^n$ is a Fourier expansion of $f \in H$ in the sense of (2.6.10), one has

$$\|f\|^2 = \sum_{n=1}^{\infty} | < f, e^n > |^2.$$

(Bessel's identity) (Hint: employ the results of Exercises 2.2.10 and 2.5.16.)

2.7 An application of Fourier expansions to the heat equation

We consider the time evolution of the temperature u in a bar made of a homogeneous material. We know such temperature at time $t = 0$ and a thermostat keeps it constant (for example, at the level 0) at the endpoints of the bar. We identify the bar with an interval $[0, l] \subseteq \mathbb{R}$, with $l > 0$. If $t \geq 0$ and $x \in [0, l]$, we indicate with $u(t, x)$ the temperature at the point x at time $t \geq 0$. We know from physics that u satisfies the following differential equation:

$$D_t u(t, x) = k^2 D_x^2 u(t, x), \quad (t, x) \in]0, +\infty[\times [0, l], \tag{2.7.1}$$

with k positive costant, depending on the material. We have also

$$u(0,x) = f(x), \quad x \in [0,l] \tag{2.7.2}$$

and

$$u(t,0) = u(t,l) = 0, \quad t \geq 0. \tag{2.7.3}$$

Here f is a continuous real valued function, with domain $[0,l]$, such that $f(0) = f(l) = 0$ (this condition assures the compatibility of (2.7.2) and (2.7.3)).

A solution to system (2.7.1)-(2.7.2)-(2.7.3) will be a function u, which is continuous in $[0,+\infty[\times[0,l]$, equipped in $]0,+\infty[\times[0,l]$ of the continuous derivatives $D_t u$, $D_x u$, $D_x^2 u$, satisfying the three conditions (2.7.1)-(2.7.2)-(2.7.3).

In this simple situation, we are going to apply the so called **method of separation of variables**: firstly, we shall look for solutions to (2.7.1) and (2.7.3), which are the product of a function of t by a function of x. Next, we shall construct a solution to the problem, employing expansions involving the particular solutions found.

So, we start to carry out our program, looking for solutions in the form

$$v(t,x) = T(t)X(x), \quad t \geq 0, x \in [0,l], \tag{2.7.4}$$

not identically equal to zero, and satisfying (2.7.1) and (2.7.3). It should be

$$\begin{cases} T'(t)X(x) = k^2 T(t)X''(x), & t > 0, x \in [0,l], \\ X(0) = X(l) = 0. \end{cases} \tag{2.7.5}$$

Let $t_0 > 0$ be such that $T(t_0) \neq 0$. Then, it holds

$$\begin{cases} \lambda X(x) = X''(x), & x \in [0,l], \\ X(0) = X(l) = 0, \end{cases} \tag{2.7.6}$$

with $\lambda := \frac{T'(t_0)}{k^2 T(t_0)}$. In the same way, if $x_0 \in]0,l[$ e $X(x_0) \neq 0$, one has

$$T'(t) = \frac{k^2 X''(x_0)}{X(x_0)} T(t) = \lambda k^2 T(t). \tag{2.7.7}$$

We are interested in solutions which are not identically zero.

Let us consider problem (2.7.6). It is not difficult to check that it has solutions of this type if and only if λ is in the form

$$\lambda = -\frac{n^2\pi^2}{l^2}, n \in \mathbb{N}. \tag{2.7.8}$$

In this case, the solutions are the functions X, which are representable in the form

$$X(x) = C\sin(\frac{n\pi x}{l}), \tag{2.7.9}$$

with $C \in \mathbb{R}$ arbitrary. If $\lambda = -\frac{n^2\pi^2}{l^2}$, we can compute the solutions of (2.7.7), which are the functions of the form

$$T(t) = Ce^{-\frac{(n\pi k)^2 t}{l^2}}, \tag{2.7.10}$$

with $C \in \mathbb{R}$ arbitrary. So, we obtain, for each $n \in \mathbb{N}$, the function

$$v_n(t,x) = e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}), \tag{2.7.11}$$

fulfilling (2.7.1) e (2.7.3). So we look for a solution of (2.7.1)-(2.7.2)-(2.7.3) in the form

$$u(t,x) = \sum_{n=1}^{\infty} c_n v_n(t,x). \tag{2.7.12}$$

With formal computations, we obtain

$$f(x) = u(0,x) = \sum_{n=1}^{\infty} c_n \sin(\frac{n\pi x}{l}). \tag{2.7.13}$$

We set

$$\epsilon_n(x) := \sin(\frac{n\pi x}{l}). \tag{2.7.14}$$

We indicate with $< .,. >$ the inner product in $L^2([0,l])$. If $m \neq n$, one has

$$< \epsilon_n, \epsilon_m >= \int_0^l \sin(\frac{n\pi x}{l})\sin(\frac{m\pi x}{l})dx$$

$$= \frac{1}{2}[\int_0^l \cos(\frac{(n-m)\pi x}{l})dx - \int_0^l \cos(\frac{(n+m)\pi x}{l})dx] = 0,$$

while

$$< \epsilon_n, \epsilon_n >= \int_0^l \sin^2(\frac{n\pi x}{l})dx = \frac{1}{2}\{\int_0^l [1 - \cos(\frac{2n\pi x}{l})]dx\} =$$

$$= \frac{l}{2}.$$

So, if we set

$$e_n(x) = \sqrt{\frac{2}{l}}\sin(\frac{n\pi x}{l}), \tag{2.7.15}$$

we obtain that $\{e_n : n \in \mathbb{N}\}$ is an orthonormal system in $L^2([0,l])$. Now, it is possible to show that the linear space generated by $\{e_n : n \in \mathbb{N}\}$ is dense in $L^2([0,l])$. This implies, in force of Remark 2.6.4, that every element $f \in L^2([0,l])$ can be represented in the form (2.6.11), with the series convergent in $L^2([0,l])$. So, it is natural to take in (2.7.12)

$$c_n = \sqrt{\frac{2}{l}} < f, e_n >= \frac{2}{l}\int_0^l \sin(\frac{n\pi y}{l})f(y)dy, \tag{2.7.16}$$

from which we deduce

$$u(t,x) = \frac{2}{l}\sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l})\int_0^l \sin(\frac{n\pi y}{l})f(y)dy \tag{2.7.17}$$

We observe that, if $t > 0$, the series in the second term of (2.7.17) is absolutely convergent, as, $\forall n \in N$,

$$|e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l})\int_0^l \sin(\frac{n\pi y}{l})f(y)dy|$$

$$\leq e^{-\frac{(n\pi k)^2 t}{l^2}} \int_0^l |f(y)|dy,$$

and the series with such expression as $n-$th term is convergent, by the root test. In addition, it is possible to check that it is correct to differentiate with respect to t and x under the sign of series in (2.7.17) and the following function

$$u(t,x) = \begin{cases} \frac{2}{l}\sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \sin(\frac{n\pi x}{l}) \int_0^l \sin(\frac{n\pi y}{l}) f(y)dy \\ \text{if } t > 0, x \in [0,l], \\ f(x) \\ \text{if } t = 0, x \in [0,l], \end{cases} \tag{2.7.18}$$

is a solution in the specified sense of system (2.7.1)-(2.7.2)-(2.7.3). One might wonder whether it is unique.

So, let v be another solution to the same problem. It is immediately seen that $z := u - v$ is a solution to (2.7.1)-(2.7.2)-(2.7.3) with $f = 0$. In conclusion, we are reduced to the following problem : let $z \in C([0,+\infty[\times[0,l])$, equipped with the continuous derivatives $D_t z$, $D_x z$, $D_x^2 z$ in $]0,+\infty[\times[0,l]$, and such that

$$D_t z(t,x) = k^2 D_x^2 z(t,x), \quad \forall (t,x) \in]0,+\infty[\times[0,l], \tag{2.7.19}$$

$$z(0,x) = 0, \quad \forall x \in [0,l] \tag{2.7.20}$$

and

$$z(t,0) = z(t,l) = 0, \quad \forall t \geq 0. \tag{2.7.21}$$

Are we allowed to say that z is identically zero? To this aim, let us introduce the function $E : [0,+\infty[\to \mathbb{R}$, defined as follows:

$$E(t) := \int_0^l z(t,x)^2 dx. \tag{2.7.22}$$

From the regularity assumptions on z, one could deduce that E is continuous in $[0,+\infty[$ and differentiable in $]0,+\infty[$ and that, for $t > 0$, it holds

$$E'(t) = 2\int_0^l z(t,x) D_t z(t,x) dx.$$

It follows

$$E'(t) = 2k^2 \int_0^l z(t,x) D_x^2 z(t,x) dx =$$

(integrating by parts and using (2.7.21))

$$= -2k^2 \int_0^l (D_x z(t,x))^2 dx \leq 0.$$

It follows that E is nonincreasing in $[0,+\infty[$. So, $\forall t \geq 0$,

$$0 \leq E(t) \leq E(0) = 0.$$

We conclude that $E(t) = 0$ $\forall t \geq 0$. Applying Lemma 2.1.11, we can say that $\forall t \geq 0$ $z(t,x) = 0$ a. e. in $[0,l]$. So, using the fact that z is continuous (see Exercise 2.7.2), we deduce that z is identically zero.

We can summarize what we have seen in the following

Theorem 2.7.1. *Consider the problem (2.7.1)-(2.7.2)-(2.7.3), with* $f \in C([0,l],\mathbb{R})$*,* $f(0) = f(l) = 0$*. Then such problem has a unique solution* u *which is continuous and real valued in* $[0,+\infty[\times[0,l]$*, equipped with the continuous derivatives* $D_t u$*,* $D_x u$*,* $D_x^2 u$ *in* $]0,+\infty[\times[0,l]$*. Such solution can be represented in the form (2.7.17).*

Exercise 2.7.2. Let $f \in C([0,l])$ be such that $f(x) = 0$ a. e. in $[0,l]$. Show that $f(x) = 0$ $\forall x \in [0,l]$.

Exercise 2.7.3. Consider the problem (2.7.1)-(2.7.2), with the further condition

$$D_x u(t,0) = D_x u(t,l) = 0, \qquad t \geq 0 \tag{2.7.23}$$

replacing (2.7.3). Following the arguments employed for problem (2.7.1)-(2.7.2)-(2.7.3), study the system (2.7.1)-(2.7.2)-(2.7.23). Use the fact that the linear space generated by $\{\frac{1}{\sqrt{l}}\} \cup \{\sqrt{\frac{2}{l}}\cos(\frac{n\pi.}{l}) : n \in \mathbb{N}\}$ is dense in $L^2([0,l])$. In such a way, prove (at least, partially) the following

Theorem 2.7.4. *Consider the problem (2.7.1)-(2.7.2)-(2.7.23), with* $f \in C([0,l],\mathbb{R})$*. Then such problem has a unique solution* u*, which is continuous and real valued in* $[0,+\infty[\times[0,l]$*, equipped with the continuous derivatives* $D_t u$*,* $D_x u$*,* $D_x^2 u$ *in* $]0,+\infty[\times[0,l]$*. Such solution can be represented in the form*

$$u(t,x) = \begin{cases} \frac{2}{l}[\frac{1}{2}\int_0^l f(y)dy + \sum_{n=1}^{\infty} e^{-\frac{(n\pi k)^2 t}{l^2}} \cos(\frac{n\pi x}{l}) & \\ \times \int_0^l \cos(\frac{n\pi y}{l}) f(y)dy] & \text{se } t > 0, x \in [0,l], \\ f(x) & \text{se } t = 0, x \in [0,l]. \end{cases} \tag{2.7.24}$$

Chapter 3

Functions of one complex variable

3.1 Holomorphic functions

The absolute value is a norm in $\mathbb{C}$, thinking of it as a linear space over itself. We recall that, as a set, $\mathbb{C} = \mathbb{R}^2$ and we observe that the complex absolute value coincides with the Euclidean norm in $\mathbb{R}^2$. So, employing well known results concerning the theory of several real variables, we can say that:

Lemma 3.1.1. *Given $A \subseteq \mathbb{C}$, $f : A \to \mathbb{C}$ and $z \in A$, f is continuous in z if and only if $Re(f)$ and $Im(f)$ are continuous in z.*

If $z \in L(A)$, $\lim_{v \to z} f(v) = l$ is equivalent to

$$\lim_{v \to z} Re(f(v)) = Re(l), \lim_{v \to z} Im(f(v)) = Im(l).$$

We are going to introduce the notion of complex derivative.

Definition 3.1.2. *Let $A \subseteq \mathbb{C}$, open, $f : A \to \mathbb{C}$, $z^0 \in A$. We shall say that f is* **differentiable in complex sense** *in z^0 if there exists in $\mathbb{C}$*

$$f'(z^0) := \lim_{z \to z^0} \frac{f(z) - f(z^0)}{z - z_0} = \lim_{h \to 0} \frac{f(z^0 + h) - f(z^0)}{h}. \tag{3.1.1}$$

We shall call $f'(z^0)$ the (complex) derivative of f in z^0.

Definition 3.1.3. *Let $A \subseteq \mathbb{C}$, open, $f : A \to \mathbb{C}$. We shall say that f is* **holomorphic** *in A if it is differentiable in complex sense in each point of A and, moreover, the function $z \to f'(z)$ is continuous in A.*

Example 3.1.4. Let $n \in \mathbb{N}$, $f : \mathbb{C} \to \mathbb{C}$, $f(z) = z^n$. Then f is holomorphic. In fact, $\forall z \in \mathbb{C}$, one has

$$\lim_{h \to 0} \frac{(z+h)^n - z^n}{h} = \lim_{h \to 0} [(z+h)^{n-1} + (z+h)^{n-2} z + ... + (z+h) z^{n-2} + z^{n-1}]$$
$$= n z^{n-1}.$$

Example 3.1.5. Let $f : \mathbb{C} \to \mathbb{C}$, $f(z) = Re(z)$. Then f is not differentiable in complex sense in any point of $\mathbb{C}$. In fact, if $z \in \mathbb{C}$,

$$\lim_{h \to 0, h \in \mathbb{R}} \frac{Re(z+h) - Re(z)}{h} = 1,$$

while
$$\lim_{h\to 0, h\in\mathbb{R}} \frac{Re(z+ih)-Re(z)}{ih} = 0.$$

Here is a list of properties already seen for the standard derivative:

Theorem 3.1.6. *Let $A \subseteq \mathbb{C}$ open, $f, g : A \to \mathbb{C}$, $z^0 \in A$. Next, let f and g be differentiable in complex sense in z^0. Then:*
(I) f is continuous in z^0;
(II) $f+g$ is differentiable in z^0 and
$$(f+g)'(z^0) = f'(z^0) + g'(z^0);$$
(III) fg is differentiable in z^0 and
$$(fg)'(z^0) = f'(z^0)g(z^0) + f(z^0)g'(z^0);$$
(IV) if $g(z) \neq 0$ $\forall z \in A$, $\frac{f}{g}$ is differentiable in z^0 and
$$(\frac{f}{g})'(z^0) = \frac{f'(z^0)g(z^0) - f(z^0)g'(z^0)}{g(z^0)^2}.$$

Theorem 3.1.7. *Let A and B be open subsets of $\mathbb{C}$, let $f : A \to \mathbb{C}$, $g : B \to \mathbb{C}$ be such that $f(A) \subseteq B$, $z^0 \in A$. Next, let f be differentiable in z^0 and let g be differentiable in $f(z^0)$. Then $g \circ f$ is differentiable in z^0 and*
$$(g \circ f)'(z^0) = g'(f(z^0))f'(z^0).$$

The next, extremely important, result states a basic link between the real and the imaginary part of a holomorphic function.

Theorem 3.1.8. *Let $A \subseteq \mathbb{C}$, A open and let $f : A \to \mathbb{C}$. Then, the following conditions are equivalent:*
(I) f is holomorphic in A;
*(II) we set $u := Re(f)$, $v := Im(f)$, and we think of them as functions from A open subset of $\mathbb{R}^2$ to $\mathbb{R}$. Then u and v belong to $C^1(A)$ and the following (**Cauchy-Riemann**) conditions hold:*
$$D_x u = D_y v, \tag{3.1.2}$$
$$D_y u = -D_x v. \tag{3.1.3}$$

Proof Let $f : A \to \mathbb{C}$ be holomorphic. Next, let $z \in A$. If $z = (x, y)$ ($x + iy$ in algebraic notation), we have
$$f'(z) = \lim_{h\to 0, h\in\mathbb{R}} \frac{f(z+h)-f(z)}{h} =$$
$$= \lim_{h\to 0, h\in\mathbb{R}} (\frac{u(x+h,y)-u(x,y)}{h} + i\frac{v(x+h,y)-v(x,y)}{h}).$$
From Lemma 3.1.1 it follows the existence of $D_x u(x, y)$ and of $D_x v(x, y)$. Moreover, we have
$$D_x u(x,y) = Re(f'(z)), D_x v(x,y) = Im(f'(z)). \tag{3.1.4}$$

Next, we have

$$f'(z) = \lim_{h\to 0, h\in\mathbb{R}} \frac{f(z+ih)-f(z)}{ih} =$$

$$= \lim_{h\to 0, h\in\mathbb{R}} \left(\frac{v(x,y+h)-v(x,y)}{h} - i\frac{u(x,y+h)-u(x,y)}{h}\right).$$

From Lemma 3.1.1 it follows the existence of $D_y u(x,y)$ and of $D_y v(x,y)$. Moreover, we have

$$D_y u(x,y) = -Im(f'(z)), D_y v(x,y) = Re(f'(z)). \tag{3.1.5}$$

Comparing (3.1.4) with (3.1.5), we obtain that (I) implies (II).

On the other hand, let us assume that (II) holds. We set

$$\alpha := D_x u(z) = D_y v(z),$$

$$\beta := D_x v(z) = -D_y u(z).$$

We indicate with $h = (h_1, h_2) = h_1 + ih_2$ the generic element of $\mathbb{C}$. We consider the first order Taylor expansions of u and v around z: if $z + h \in A$, we have

$$u(z+h) = u(x+h_1, y+h_2) = u(z) + \alpha h_1 - \beta h_2 + r(h),$$

$$v(z+h) = v(x+h_1, y+h_2) = v(z) + \beta h_1 + \alpha h_2 + s(h),$$

with

$$\lim_{h\to 0} \frac{r(h)}{|h|} = \lim_{h\to 0} \frac{s(h)}{|h|} = 0.$$

It follows

$$\frac{f(z+h)-f(z)}{h} = \frac{u(z+h)-u(z)+i(v(z+h)-v(z))}{h}$$

$$= \frac{\alpha h_1 - \beta h_2 + r(h) + i(\beta h_1 + \alpha h_2 + s(h))}{h} =$$

$$= \frac{(\alpha + i\beta)(h_1 + ih_2) + r(h) + is(h)}{h}$$

$$= \alpha + i\beta + \frac{r(h)+is(h)}{h}.$$

From

$$\left|\frac{r(h)+is(h)}{h}\right| \le \frac{|r(h)|}{|h|} + \frac{|s(h)|}{|h|} \to 0 (h \to 0)$$

it follows the existence of $f'(z)$ and the identity

$$f'(z) = \alpha + i\beta.$$

□

Remark 3.1.9. Let $f : A \to \mathbb{C}$ be holomorphic, with A open and pathwise connected in $\mathbb{C}$. From the Cauchy Riemann conditions it immediately follows that the vector fields

$$U(x,y) := (u(x,y), -v(x,y)) \tag{3.1.6}$$

and

$$V(x,y) := (v(x,y), u(x,y)) \tag{3.1.7}$$

are closed in A.

On the contrary, if U and V, defined in (3.1.6) and (3.1.7), are of class C^1 and closed, $f := u + iv$ is holomorphic.

We conclude this section with some important examples of holomorphic functions.

Example 3.1.10. (Complex exponential function) Let $z \in \mathbb{C}$, $z = x + iy$, with x and y real. We set

$$e^z := e^x e^{iy} = e^x(\cos(y) + i\sin(y)). \tag{3.1.8}$$

The following basic properties are easy to check. We leave the proof to the reader (see Exercise 3.1.15).

$$\begin{array}{lll} (I) & \forall z, v \in \mathbb{C} & e^z e^v = e^{z+v} \\ (II) & \forall z \in \mathbb{C} & |e^z| = e^{Re(z)} > 0; \\ (III) & \forall z \in \mathbb{C} & (e^z)^{-1} = e^{-z}. \end{array} \tag{3.1.9}$$

Setting $f : \mathbb{C} \to \mathbb{C}$, $f(z) = e^z$, it is easy to verify that f is holomorphic, as it fulfills the Cauchy-Riemann conditions. In fact, if $u := Re(f)$, $v := Im(f)$, one has, given $z = (x,y)$,

$$u(x,y) = e^x \cos(y),$$

$$v(x,y) = e^x \sin(y),$$

from which

$$D_x u(x,y) = e^x \cos(y) = D_y v(x,y),$$

$$D_y u(x,y) = -e^x \sin(y) = -D_x v(x,y).$$

Moreover, one has

$$f'(z) = D_x u(z) + iD_x v(z) = e^x \cos(y) + ie^x \sin(y) = f(z).$$

Example 3.1.11. (Logarithm functions) Let A be open and pathwise connected in $\mathbb{C} \setminus \{0\}$. We have to exclude 0 because of propriety (II) of the exponential function. A logarithm function is a function $f : A \to \mathbb{C}$ continuous and such that $e^{f(z)} = z \ \forall z \in A$. It is possible to show that, if f is a logarithm function, it is holomorphic and $f'(z) = \frac{1}{z} \ \forall z \in A$. Here we limit ourselves to consider one specific function of this type.

Let $A := \mathbb{C} \setminus \{x \in \mathbb{R} : x \leq 0\}$. We set

$$f : A \to \mathbb{C},$$

$$f(z) = \ln(|z|) + iArg(z),$$

with $Arg(z)$ unique element of $arg(z)$ which is larger than $-\pi$ and less than π. From Exercise 3.1.16, one has $e^{f(z)} = z \ \forall z \in A$. If $\theta := Arg(z)$, necessarily $\cos(\theta) = \frac{Re(z)}{|z|}$, $\sin(\theta) = \frac{Im(z)}{|z|}$, $\theta \in]-\pi, \pi[$. We consider the particular case $Re(z) > 0$. Then $\theta \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ and

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} = \frac{Im(z)}{Re(z)}.$$

holds. So, if $z = x + iy$, with $x > 0$, we have

$$f(z) = \ln(\sqrt{x^2 + y^2}) + i \arctan(\frac{y}{x}).$$

If $u := Re(f)$ and $v := Im(f)$, it follows that, for $x > 0$,

$$D_x u(x, y) = \frac{x}{x^2 + y^2} = D_y v(x, y),$$

$$D_y u(x, y) = \frac{y}{x^2 + y^2} = -D_x v(x, y),$$

implying that f fulfills the Cauchy-Riemann conditions, at least in $\{z \in \mathbb{C} : Re(z) > 0\}$. Moreover,

$$f'(z) = D_x u(z) + iD_x v(z) = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2} = \frac{1}{x + iy} = \frac{1}{z}.$$

Example 3.1.12. Let $A \subseteq \mathbb{C} \setminus \{0\}$ be open and pathwise connected, g a logarithm function in A. If $\alpha \in \mathbb{R}$, we set

$$f_\alpha(z) := e^{\alpha g(z)}.$$

We observe that f_α is a possible version of z^α. We take into account that its definition depends on the chosen logarithm function.

By examples 3.1.10 and 3.1.11 and Theorem 3.1.7, f_α is holomorphic and we have, $\forall z \in A$,

$$f'_\alpha(z) =$$

$$= \alpha\frac{e^{\alpha g(z)}}{z} = \alpha\frac{e^{\alpha g(z)}}{e^{g(z)}} = \alpha e^{(\alpha-1)g(z)}$$

$$= \alpha f_{\alpha-1}(z).$$

We have written that, in general, f_α depends on the logarithm function g. However, let $\alpha \in \mathbb{Z}$. If g_1 e g_2 are logarithm functions in A, we have that $g_1(z) - g_2(z) = 2k\pi i$ $\forall z \in A$, for some $k \in \mathbb{Z}$ (see the following Exercise 3.1.17). It follows that

$$\frac{e^{\alpha g_1(z)}}{e^{\alpha g_2(z)}} = e^{\alpha(g_1(z)-g_2(z))} = e^{2k\alpha\pi i} = 1,$$

because $k\alpha \in \mathbb{Z}$. So, in case $\alpha \in \mathbb{Z}$, f_α does not depend on the logarithm function.

Example 3.1.13. We extend the trigonometric functions "sine" and "cosine" to $\mathbb{C}$. In force of Euler's formulas, it is natural to set, given $z \in \mathbb{C}$:

$$\sin(z) := \frac{e^{iz} - e^{-iz}}{2i}, \tag{3.1.10}$$

$$\cos(z) := \frac{e^{iz} + e^{-iz}}{2}. \tag{3.1.11}$$

It is immediate to verify that these extensions are holomorphic in $\mathbb{C}$. For some of their properties, we refer also to Exercise 3.1.18.

Exercise 3.1.14. Let A be open and pathwise connected in $\mathbb{C}$, $f, g : A \to \mathbb{C}$ holomorphic. Prove that:

(I) if $Re(f)$ is costant, then f is constant;

(II) if $Im(f)$ is costant, then f is constant;

(III) if $f'(z) = g'(z) = 0\ \forall z \in A$, then $f - g$ is costant.

(Hint.: employ the Cauchy-Riemann conditions and well known results concerning vector fields.)

Exercise 3.1.15. Check the validity of the properties of the complex exponential function stated in Example 3.1.10.

Exercise 3.1.16. Let $v \in \mathbb{C} \setminus \{0\}$. Prove that the equation

$$e^z = v \tag{3.1.12}$$

has infinite solutions in $\mathbb{C}$: precisely, the complex numbers which are representable in the form

$$z = \ln(|v|) + i\theta, \tag{3.1.13}$$

with $\theta \in arg(v)$, wher $\arg(v) := \{\theta \in \mathbb{R} : e^{i\theta} = \frac{v}{|v|}\}$.

Observe that $e^z = 1$ if and only if $z = 2k\pi i$, for some $k \in \mathbb{Z}$.

Exercise 3.1.17. Let A be open and pathwise connected in $\mathbb{C} \setminus \{0\}$, let $f, g : A \to \mathbb{C}$ two logarithm functions. Check that there exists $k \in \mathbb{Z}$, such that

$$f(z) - g(z) = 2k\pi i$$

$\forall z \in A$. (Hint.: employ the results of Exercises 3.1.14 (III) and 3.1.16.)

Exercise 3.1.18. Check that, $\forall z, v \in \mathbb{C}$,

(I)

$$\sin'(z) = \cos(z), \cos'(z) = -\sin(z); \tag{3.1.14}$$

(II)

$$\sin^2(z) + \cos^2(z) = 1; \tag{3.1.15}$$

(III)

$$\sin(-z) = -\sin(z), \cos(-z) = \cos(z); \tag{3.1.16}$$

(IV)

$$\sin(v + z) = \sin(v)\cos(z) + \cos(v)\sin(z); \tag{3.1.17}$$

(V)

$$\cos(v + z) = \cos(v)\cos(z) - \sin(v)\sin(z); \tag{3.1.18}$$

(VI) $\sin(z) = 0$ if and only if $z = k\pi$ for some $k \in \mathbb{Z}$;

(VII) $\cos(z) = 0$ if and only if $z = k\frac{\pi}{2}$ for some $k \in \mathbb{Z}$, k odd.

Exercise 3.1.19. Characterize

(I) $\{z \in \mathbb{C} : \sin(z) = 2\}$;

(II) $\{z \in \mathbb{C} : \cos(z) = 3\}$.

The fact that these sets are not empty implies that, differently from the case of $\mathbb{R}$, in $\mathbb{C}$ the ranges of sin and cos are not contained in $[-1, 1]$!

3.2 Complex integrals

Many important properties of holomorphic functions are expressed in terms of complex integrals. In this section we are going to collect the main facts, concerning them, that we shall use.

Definition 3.2.1. *Let $J = [a, b]$ be a closed and bounded interval in $\mathbb{R}$. A continuous map $\alpha : [a, b] \to \mathbb{C}$ will be called a* **continuous path** *in $\mathbb{C}$. If α is of class C^1 (see the following Remark 3.2.2), we shall say that α is a* **path of class** C^1*. We shall say that α is a* **piecewise** C^1 **path** *if it is a continuous path and there exists $\{a = a_0 < .... < a_k = b\}$ subset of $[a, b]$, such that, if $j = 1, ..., k$, $\alpha_{|[a_{j-1},a_j]}$ is of class C^1.*

If α is a continuous path, its range will be called the **support** *of α and will be indicated with $supp(\alpha)$.*

If $\alpha : [a, b] \to \mathbb{C}$ is a continuous path, such that $\alpha(a) = \alpha(b)$, we shall say that it is **closed**.

Remark 3.2.2. With reference to Definition 3.2.1, we recall that $\mathbb{C} = \mathbb{R}^2$. So, if $\alpha(t) := (\alpha_1(t), \alpha_2(t))$ $(t \in [a, b])$, the statement "α is of class C^1" is equivalent to the statement that α_1 and α_2, which are, respectively, its real and inmaginary part, are of class C^1. Moreover, $\forall t \in [a, b]$, one has

$$\alpha'(t) = (\alpha_1'(t), \alpha_2'(t)) = \alpha_1'(t) + i\alpha_2'(t). \tag{3.2.1}$$

Of course, $\alpha'(t)$ is defined, as usual, as

$$\lim_{h\to 0} \frac{\alpha(t + h) - \alpha(t)}{h}.$$

Example 3.2.3. Let $z^0 \in C$, $r > 0$, $\alpha : [0, 2\pi] \to \mathbb{C}$, $\alpha(t) = z^0 + re^{it}$. α is a continuous path, having as support the circumference of centre z^0 and radius r. One has $\alpha(t) = Re(z^0) + r\cos(t) + i(Im(z^0) + r\sin(t))$. So α is of class C^1. Moreover, $\forall t \in [0, 2\pi]$,

$$\alpha'(t) = -r\sin(t) + ir\cos(t) = ire^{it}.$$

We pass to the definition of complex integral.

Definition 3.2.4. *Let $\alpha : [a, b] \to \mathbb{C}$ be a continuous path of class C^1, $f : \alpha([a, b]) \to \mathbb{C}$. If $t \to f(\alpha(t))\alpha'(t)$ is summable in $[a, b]$, we define the complex integral $\int_\alpha f(z)dz$ as $\int_{[a,b]} f(\alpha(t))\alpha'(t)dt$.*

Example 3.2.5. Let $z^0 \in \mathbb{C}$, $r > 0$, $\alpha : [0, 2\pi] \to \mathbb{C}$, $\alpha(t) := z^0 + re^{it}$. Let $n \in \mathbb{Z}$. We compute $\int_\alpha (z - z^0)^n dz$. By definition,

$$\int_\alpha (z - z^0)^n dz$$

$$= \int_{[0,2\pi]} (z^0 + re^{it} - z^0)^n rie^{it}dt = r^{n+1}i \int_0^{2\pi} e^{i(n+1)t}dt$$

If $n \neq -1$, one has

$$\int_0^{2\pi} e^{i(n+1)t}dt = \int_0^{2\pi} \cos((n + 1)t)dt + i \int_0^{2\pi} \sin((n + 1)t)dt =$$

$$= [\frac{\sin((n + 1)t)}{n + 1}]_{t=0}^{t=2\pi} + i[-\frac{\cos((n + 1)t)}{n + 1}]_{t=0}^{t=2\pi} = 0.$$

In case $n = -1$, one has

$$\int_0^{2\pi} e^{i(n+1)t}dt = \int_0^{2\pi} 1dt = 2\pi.$$

In conclusion,

$$\int_\alpha (z - z^0)^n dz = \begin{cases} 0 & \text{if} \quad n \neq -1, \\ 2\pi i & \text{if} \quad n = -1. \end{cases}$$

It is convenient to extend the definition of complex integral to the case that α is piecewise of class C^1:

Definition 3.2.6. *Let $\alpha : [a, b] \to \mathbb{C}$ be a piecewise C^1 path, with $a = a_0 < a_1 < ... < a_k = b$, and $\alpha_{|[a_0,a_1]}$, ..., $\alpha_{|[a_{k-1},a_k]}$ of class C^1. We shall indicate with α^j $(1 \leq j \leq k)$ the restriction of α to $[a_{j-1}, a_j]$. If $f : \alpha([a, b]) \to \mathbb{C}$ is such that all the integrals $\int_{\alpha^1} f(z)dz$, ..., $\int_{\alpha^k} f(z)dz$ are defined, we set*

$$\int_\alpha f(z)dz = \int_{\alpha^1} f(z)dz + ... + \int_{\alpha^k} f(z)dz.$$

It is possible to check that this definition is independent of $\{a_0, a_1, \ldots, a_{k-1}, a_k\}$. .

Remark 3.2.7. Now we recall that, if A is a pathwise connected open subset of $\mathbb{R}^2$, $f : A \to \mathbb{R}^2$ is a vector field and $\alpha : [a, b] \to \mathbb{R}^2$ is a path of class C^1 and support in A, the following curvilinear integral of second type is defned:

$$\int f \cdot d\alpha := \int_{[a,b]} f(\alpha(t)) \cdot \alpha'(t)dt, \tag{3.2.2}$$

with $\cdot$ standard inner product in $\mathbb{R}^2$.

Let $f(z) = (u(z), v(z)) = u(z) + iv(z)$ and $\alpha(t) = (\alpha_1(t), \alpha_2(t)) = \alpha_1(t) + i\alpha_2(t)$. Then, concerning the complex integral $\int_\alpha f(z)dz$, one has

$$\int_\alpha f(z)dz =$$

$$= \int_{[a,b]} f(\alpha(t))\alpha'(t)dt = \int_{[a,b]} [u(\alpha(t)) + iv(\alpha(t))][\alpha_1'(t) + i\alpha_2'(t)]dt =$$

$$= \int_{[a,b]} [u(\alpha(t))\alpha_1'(t) - v(\alpha(t))\alpha_2'(t)]dt + i \int_{[a,b]} [v(\alpha(t))\alpha_1'(t) + u(\alpha(t))\alpha_2'(t)]dt$$

$$= \int U \cdot d\alpha + i \int V \cdot d\alpha,$$

with $U(z) := (u(z), -v(z))$ and $V(z) := (v(z), u(z))$. We recall (Remark 3.1.9) that, if f is holomorphic, then the vector fields U and V are closed.

From Remark 3.2.7, employing well known properties of curvilinear integrals of second type, we can draw some basic properties of complex integrals. We start with the following sequence of statements (see also Exercise 3.2.23):

Theorem 3.2.8. *Let $\alpha : [a, b] \to \mathbb{C}$ be a path which is piecewise of class C^1, $f, g : \alpha([a, b]) \to \mathbb{C}$, such that the integrals $\int_\alpha f(z)dz$ and $\int_\alpha g(z)dz$ are defined, r and s complex numbers. Then:*

(I) the integral $\int_\alpha [rf(z) + sg(z)]dz$ is defined and coincides with $r \int_\alpha f(z)dz + s \int_\alpha g(z)dz$;

(II) let $c \in]a,b[$. We indicate with α^1 and α^2 the restrictions of α to (respectively) $[a,c]$ and $[c,b]$. Then the integrals $\int_{\alpha^1} f(z)dz$ and $\int_{\alpha^2} f(z)dz$ and defined and the identity

$$\int_\alpha f(z)dz = \int_{\alpha^1} f(z)dz + \int_{\alpha^2} f(z)dz$$

holds.

Now we consider the influence of changes of parameter on complex integrals. We begin with a definition:

Definition 3.2.9. *Let $\alpha : [a,b] \to \mathbb{C}$ and $\beta : [c,d] \to \mathbb{C}$ be continuous paths. We shall say that they are* **equivalent** *if there exists $u : [c,d] \to [a,b]$, of class C^1, with $u'(t) \neq 0$ $\forall t \in [c,d]$ and surjective onto $[a,b]$, such that*

$$\beta(t) = \alpha(u(t)) \qquad \forall t \in [c,d].$$

We shall say that α and β are **positively equivalent** *if $u'(t) > 0$ $\forall t \in [c,d]$.*

Remark 3.2.10. Definition 3.2.9 is well known in the theory of curves. One can check that it is an equivalence relation between paths. The condition $u'(t) \neq 0$ $\forall t \in [c,d]$ implies that, either $u'(t) > 0$ $\forall t \in [c,d]$, or $u'(t) < 0$ $\forall t \in [c,d]$, so that u is strictly monotone. In the first case the passage from α to β preserves the sense of motion, in the second it inverts it.

From Remark 3.2.7 and corresponding properties for curvilinear integrals we deduce the following

Theorem 3.2.11. *Let α and β be paths which are piecewise C^1 and equivalent. Let $f : \alpha([a,b]) = \beta([c,d]) \to \mathbb{C}$ (see Exercise 3.2.24) be such that $\int_\alpha f(z)dz$ is defined. Then:*

(I) $\int_\beta f(z)dz$ is defined as well. Let u be the map from $[c,d]$ to $[a,b]$ of class C^1 described in Definition 3.2.9. Then:

(II) if $u'(t) > 0$ $\forall t \in [c,d]$, one has

$$\int_\beta f(z)dz = \int_\alpha f(z)dz;$$

(III) if $u'(t) < 0$ $\forall t \in [c,d]$, one has

$$\int_\beta f(z)dz = -\int_\alpha f(z)dz.$$

We recall that, if α is a continuous path, $\alpha : [a,b] \to \mathbb{C}$, its length $l(\alpha)$ is defined as

$$l(\alpha) := \sup\{\sum_{j=1}^{k} |\alpha(t_j) - \alpha(t_{j-1})| : a = t_0 < t_1 < ... < t_{k-1} < t_k = b\}. \qquad (3.2.3)$$

It is well known that, if α is of class C^1, one has

$$l(\alpha) = \int_a^b |\alpha'(t)|dt. \qquad (3.2.4)$$

The following estimate is rather useful:

Theorem 3.2.12. *Let $\alpha : [a,b] \to \mathbb{C}$ be piecewise C^1, $f : \alpha([a,b]) \to \mathbb{C}$ bounded, and such that $\int_\alpha f(z)dz$ is defined. Then*

$$|\int_\alpha f(z)dz| \leq \sup|f| \cdot l(\alpha).$$

Proof Limiting ourselves to the case that α is of class C^1, we have

$$|\int_\alpha f(z)dz|$$

$$= |\int_{[a,b]} f(\alpha(t))\alpha'(t)dt| \leq \int_{[a,b]} |f(\alpha(t))||\alpha'(t)|dt$$

(by Theorem 1.2.12 (IV))

$$\leq \sup|f| \cdot \int_{[a,b]} |\alpha'(t)|dt =$$

$$= \sup|f| \cdot l(\alpha).$$

□

Now we introduce some important definitions and results, concerning closed paths and vector fields.

Definition 3.2.13. *Let A be an open subset of $\mathbb{C}$, $\alpha, \beta : [a,b] \to \mathbb{C}$ continuous and closed paths, with support in A. We shall say that α and β are $A-$***homotopic** *if there exists $F : [0,1] \times [a,b] \to A$ such that:*

(I) F is continuous;

(II) $F(0,t) = \alpha(t)$, $F(1,t) = \beta(t)$ $\forall t \in [a,b]$;

(III) $F(s,a) = F(s,b)$ $\forall s \in [0,1]$.

A map F with properties $(I) - (III)$ is called a **homotopy** *between α e β.*

Remark 3.2.14. α and β are $A-$homotopic if it is possible to modify α with continuity in such a way to carry it to β, without going out of A. The deformation must be such that, in each instant, the path is always closed. (this conditions is expressed rigorously by (III)).

Example 3.2.15. Let $z^0 \in \mathbb{C}$, $A := \mathbb{C} \setminus \{z^0\}$, $0 < r_0 < r_1$, $\alpha^j : [0,2\pi] \to \mathbb{C}$, $\alpha^j(t) := z^0 + r_j e^{it}$ $(j \in \{0,1\})$. α^0 and α^1 are $A-$homotopic. It suffices to take:

$$\begin{cases} F : [0,1] \times [0,2\pi] \to A, \\ F(s,t) = z^0 + [r_0 + s(r_1 - r_0)]e^{it}, \quad (s,t) \in [0,1] \times [0,2\pi]. \end{cases}$$

Example 3.2.16. Let $A := \mathbb{C}$, let $\alpha : [a,b] \to \mathbb{C}$ be a closed path, $\beta : [a,b] \to \mathbb{C}$, such that $\beta(t) = z^0$ $\forall t \in [a,b]$, with $z^0 \in \mathbb{C}$. We shall say that β is a **punctual** path. α and β are A homotopic. To see this, it suffices to take

$$\begin{cases} F : [0,1] \times [a,b] \to A, \\ F(s,t) = (1-s)\alpha(t) + sz^0, \quad (s,t) \in [0,1] \times [a,b]. \end{cases}$$

So, every closed path is $\mathbb{C}-$homotopic to a punctual path. If A is an open subset of $\mathbb{C}$, we shall say that A is **simply connected** if every closed path with support in A is $A-$homotopic to a punctual path.

It is possible to prove the following important theorem:

Theorem 3.2.17. *Let A be a pathwise connected open subset in $\mathbb{R}^2$, $F : A \to \mathbb{R}^2$ a closed vector filed, α and β closed, piecewise C^1 paths with support in A and $A-$homotopic. Then*

$$\int F \cdot d\alpha = \int F \cdot d\beta.$$

Remark 3.2.18. We recall that, in the particular case that F has a potential, $\int F \cdot d\alpha = 0$ for every closed piecewise C^1 path with support in A

Theorem 3.2.17 has some important consequences, concerning complex integrals of holomorphic functions:

Corollary 3.2.19. *Let A be an open, pathwise connected subset of C, α and β closed piecewise C^1 paths with support in A and $A-$homotopic, $f : A \to \mathbb{C}$ holomorphic. Then*

$$\int_\alpha f(z)dz = \int_\beta f(z)dz.$$

Proof From Remark 3.2.7,

$$\int_\alpha f(z)dz = \int U \cdot d\alpha + i \int V \cdot d\alpha,$$

with U and V closed vector fields. The conclusion follows immediately from Theorem 3.2.17. □

Corollary 3.2.20. *Let A be an open simply connected subset of $\mathbb{C}$, α a closed piecewise C^1 path with support in A, $f : A \to \mathbb{C}$ holomorphic. Then*

$$\int_\alpha f(z)dz = 0.$$

Proof It suffices to apply Corollary 3.2.19, taking as β a punctual path.
□

We conclude this section with a classical formula.

We shall employ the following notation: if $z^0 \in \mathbb{C}$, $r > 0$, we set

$$\begin{cases} C_r : [0, 2\pi] \to \mathbb{C}, \\ C_r(t) = z^0 + re^{it}, t \in [0, 2\pi]. \end{cases} \tag{3.2.5}$$

Theorem 3.2.21. *(Cauchy's integral formula) Let A be an open subset of $\mathbb{C}$, $z^0 \in A$, $r > 0$ such that $\{z \in \mathbb{C} : |z - z^0| \le r\} \subseteq A$, $f : A \to \mathbb{C}$ holomorphic. Let α be a piecewise C^1, closed path, $A \setminus \{z^0\}-$ homotopic to $C_r(z_0)$. Then*

$$\int_\alpha \frac{f(z)}{z - z^0} dz = 2\pi i f(z^0).$$

Proof As $z \to \frac{f(z)}{z-z^0}$ is holomorphic in $A \setminus \{z^0\}$, by Corollary 3.2.19,

$$\int_\alpha \frac{f(z)}{z - z^0} dz = \int_{C_r(z^0)} \frac{f(z)}{z - z^0} dz.$$

Moreover, $C_r(z^0)$ is $A \setminus \{z^0\}-$homotopic to $C_\rho(z^0)$ for every $\rho \in]0, r]$. It follows that

$$\int_{C_r(z^0)} \frac{f(z)}{z - z^0} dz = \lim_{\rho \to 0} \int_{C_\rho(z^0)} \frac{f(z)}{z - z^0} dz$$

$$= \lim_{\rho \to 0} \int_0^{2\pi} f(z^0 + \rho e^{it}) i dt = 2\pi i f(z^0) + \lim_{\rho \to 0} \int_0^{2\pi} [f(z^0 + \rho e^{it}) - f(z^0)] i dt.$$

From Theorem 3.2.12 we obtain

$$|\int_0^{2\pi} [f(z^0 + \rho e^{it}) - f(z^0)] i dt| \leq \max_{t \in [0,2\pi]} |f(z^0 + \rho e^{it}) - f(z^0)| 2\pi \to 0 (\rho \to 0),$$

because f is continuous in z^0. The conclusion follows.

□

Remark 3.2.22. $A-$homotopic paths are defined (by definition) in the same interval. Employing Theorem 3.2.11, it may be useful to generalize slightly Theorem 3.2.21, requiring only that α is positively equivalent to some closed, piecewise C^1 path, which is $A \setminus \{z^0\}-$ homotopic to $C_r(z^0)$. This little bit more general version does not require that α and β are defined in $[0, 2\pi]$.

For example, let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic. We consider the following piecewise C^1 path α:

$$\alpha : [0, 4] \to \mathbb{C},$$

$$\alpha(t) = \begin{cases} 1 + t(i-1) & \text{if} \quad t \in [0,1], \\ i + (t-1)(-1-i) & \text{if} \quad t \in [1,2], \\ -1 + (t-2)(-i+1) & \text{if} \quad t \in [2,3], \\ -i + (t-3)(1+i) & \text{if} \quad t \in [3,4]. \end{cases}$$

α describes once in counterclockwise sense the boundary of the square with vertexes 1, i, -1, $-i$. α is positively equivalent to $\beta : [0, 2\pi] \to \mathbb{C}$, $\beta(s) = \alpha(\frac{2s}{\pi})$. It is intuitively clear that β is $\mathbb{C} \setminus \{0\}-$homotopic to $C_1(0)$ as well. So we have

$$f(0) = \frac{1}{2\pi i} \int_\alpha \frac{f(z)}{z} dz.$$

Moreover, employing again Theorems 3.2.11 and 3.2.8(II), we can also write:

$$\int_\alpha \frac{f(z)}{z} dz = \int_0^1 \frac{f(1 + t(i-1))}{1 + t(i-1)} (i-1) dt + \int_0^1 \frac{f(i + t(-1-i))}{i + t(-1-i)} (-1-i) dt$$

$$+ \int_0^1 \frac{f(-1 + t(-i+1))}{-1 + t(-i+1)} (-i+1) dt + \int_0^1 \frac{f(-i + t(1+i))}{-1 + t(1+i)} (1+i) dt.$$

Exercise 3.2.23. Prove Theorem 3.2.8.

Exercise 3.2.24. Check that equivalent paths have the same support.

3.3 Analytic functions

We start with some basic results, concerning series with complex terms, which will be crucial for the main subject of the section. These results are extensions of well known facts, which we have already seen in the case of real terms.

Let $(a_n)_{n \in \mathbb{N}}$ be a sequence of complex numbers. Starting from it, we can construct another sequence, called sequence of partial sums or series associated with $(a_n)_{n \in \mathbb{N}}$, setting, for $n \in \mathbb{N}$,

$$s_n := a_1 + ... + a_n. \qquad (3.3.1)$$

We shall indicate the series with the notation $\sum_{n=1}^{\infty} a_n$. If the sequence $(s_n)_{n\in\mathbb{N}}$ (that is, the series) has a limit in $\mathbb{C}$ for $n \to +\infty$, we shall say that it is convergent and we shall call such limit the sum of the series.

If $a_n = x_n + iy_n$, with x_n and y_n real numbers for each $n \in \mathbb{N}$, one has, for every n,

$$s_n = \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} x_k + i \sum_{k=1}^{n} y_k.$$

As the convergence of a sequence with complex terms is equivalent to the convergence of the two sequences with real terms of the real and imaginary parts, $\sum_{n=1}^{\infty} a_n$ is convergent in $\mathbb{C}$ if and only if the series $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ are convergent in $\mathbb{R}$.

The notion of absolute convergence is extensible to series with terms in $\mathbb{C}$:

Definition 3.3.1. *Let $a_n \in \mathbb{C}$ $\forall n \in \mathbb{N}$. We say that the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent if the series $\sum_{n=1}^{\infty} |a_n|$ is convergent in $\mathbb{R}$.*

In the next theorem, we shall list some extensions to $\mathbb{C}$ of well known properties of series with terms in $\mathbb{R}$.

Theorem 3.3.2. *Let $(a_n)_{n\in\mathbb{N}}$ be a series with terms in $\mathbb{C}$, with $a_n = x_n + iy_n$ (x_n and y_n real numbers) $\forall n \in \mathbb{N}$. Then*

(I) if the series $\sum_{n=1}^{\infty} a_n$ is convergent, necessarily $\lim_{n\to+\infty} a_n = 0$;

(II) if the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, it is also convergent.

Proof (I) If the series $\sum_{n=1}^{\infty} a_n$ is convergent, even $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ are convergent. Applying the corresponding result for series with real terms, one has $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} y_n = 0$. So

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} (x_n + iy_n) = 0.$$

(II) One has, $\forall n \in \mathbb{N}$,

$$|x_n| = |Re(a_n)| \le |a_n|, |y_n| = |Im(a_n)| \le |a_n|.$$

Recalling well known properties of series with real nonnegative terms, we can say that, if the series $\sum_{n=1}^{\infty} a_n$ is assolutely convergent, the same happens for $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$. So the series $\sum_{n=1}^{\infty} x_n$ and $\sum_{n=1}^{\infty} y_n$ are convergent, from which the conclusion follows.

□

We pass to the definitions of power series and analytic function.

Definition 3.3.3. *A* **power series** *is a series in the form*

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

depending on the complex parameter z. Here z_0 and a_n ($\forall n \in \mathbb{N}_0$) are fixed complex numbers.

Definition 3.3.4. *Let A be an open subset of $\mathbb{C}$ and $f : A \to \mathbb{C}$. We shall say that f is* **analytic** *in A if, $\forall z_0 \in A$ there exist $r > 0$ and a power series in the form $\sum_{n=0}^{\infty} a_n(z - z_0)^n$, such that:*

a) $B(z_0, r) = \{z \in \mathbb{C} : |z - z_0| < r\} \subseteq A$;

b) the series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ is convergent for every $z \in B(z_0, r)$;

c) $\forall z \in B(z_0, r)$ one has $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$.

Let us examine some properties of power series. We introduce preliminarily the notion of radius of convergence.

Definition 3.3.5. *Let $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ be a power series. Its* **radius of convergence** *is defined as*

$$\sup\{|z-z_0| : \sum_{n=0}^{\infty} a_n(z-z_0)^n \text{ is convergent}\}.$$

Remark 3.3.6. Observe first that a power series converges, at least, for $z = z_0$. Obviously, the radius of convergence of a power series is an element of $[0, +\infty]$. In particular, saying that it is $+\infty$ is equivalent to saying that, for each r in $\mathbb{R}^+$, there exists $z \in \mathbb{C}$ such that $|z - z_0| > r$ and the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ is convergent.

The main property of power series is described by the following

Theorem 3.3.7. *(Lemma of Abel) Let $z_1 \in \mathbb{C}$ be such that the series $\sum_{n=0}^{\infty} a_n(z_1-z_0)^n$ is convergent. Then, $\forall z \in \mathbb{C}$ such that $|z-z_0| < |z_1-z_0|$ the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ is absolutely convergent.*

Proof Let $z \in \mathbb{C}$ be such that $|z - z_0| < |z_1 - z_0|$ (observe that the existence of z implies that $|z_1 - z_0| > 0$). By Theorems 3.3.2(I) and 2.3.6, there exists $M \geq 0$ such that

$$|a_n||z_1 - z_0|^n \leq M \qquad \forall n \in \mathbb{N}_0.$$

It follows that

$$|a_n(z-z_0)^n| = |a_n||z_1-z_0|^n(\frac{|z-z_0|}{|z_1-z_0|})^n \leq M(\frac{|z-z_0|}{|z_1-z_0|})^n.$$

As $\frac{|z-z_0|}{|z_1-z_0|} < 1$, the series $\sum_{n=0}^{\infty} M(\frac{|z-z_0|}{|z_1-z_0|})^n$ is convergent. The conclusion follows immediately. □

From the lemma of Abel, we immediately deduce the following

Theorem 3.3.8. *Let $\rho \in [0, +\infty]$ be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$. Then:*

(I) if $\rho = 0$ the series converges only if $z = z_0$;

(II) if $0 < \rho < +\infty$, the series converges absolutely if $|z - z_0| < \rho$, it does not converge if $|z - z_0| > \rho$;

(III) if $\rho = +\infty$, the series converges absolutely for every $z \in \mathbb{C}$.

Proof We limit ourselves to the case (II). The other cases can be treated analogously, with some simplifications.

Let $|z - z_0| < \rho$. Then, there exists $z_1 \in \mathbb{C}$ such that $|z - z_0| < |z_1 - z_0|$ and the series $\sum_{n=0}^{\infty} a_n(z_1-z_0)^n$ is convergent (othewise, it would be $\rho \leq |z - z_0|$). So , by the lemma of Abel, the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ is assolutely convergent. On the contrary, if $|z - z_0| > \rho$, the conclusion follows from the definition of radius of convergence,

□

For future use we state the following

Corollary 3.3.9. *Let $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ be a power series with radius of convergence $\rho \in [0, +\infty]$, and let $z_1 \in \mathbb{C}$. Then:*

(I) if $\sum_{n=0}^{\infty} a_n(z_1-z_0)^n$ is absolutely convergent, $|z_1 - z_0| \leq \rho$;

(II) if $\sum_{n=0}^{\infty} a_n(z_1-z_0)^n$ is not absolutely convergent, $|z_1 - z_0| \geq \rho$.

Proof See Exercise 3.3.21. □

Example 3.3.10. We consider the power series $\sum_{n=0}^{\infty} n!z^n$. If $z \in \mathbb{R}^+$, it follows from the ratio test that the series is not convergent. From Corollary 3.3.9 we deduce that the radius of convergence is 0.

Example 3.3.11. We consider the power series $\sum_{n=0}^{\infty} z^n$. If $z \in [0,1[$, the series is convergent, while it is not convergent if $z \in [1,+\infty]$. Consequently, the radius of convergence is necessarily 1. Arguing as in the case of $z \in \mathbb{R}$, one can see that, if $|z| < 1$, the sum of the series is $(1-z)^{-1}$.

Example 3.3.12. We consider the power series $\sum_{n=0}^{\infty} \frac{z^n}{n!}$. From the ratio test, it follows immediately that the series is absolutely convergent for every $z \in \mathbb{C}$. So, in this case the radius of convergence is $+\infty$.

Remark 3.3.13. If the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ is $\rho \in]0,+\infty[$ and $|z-z_0| = \rho$, Theorem 3.3.8 does not say if the series is convergent. Let us examine some examples.

We start with the series considered in Example 3.3.11, with radius of convergence 1. If $|z| = 1$, one has $|z^n| = 1 \; \forall n \in \mathbb{N}_0$. Consequently, it is not satisfied the condition $\lim_{n\to+\infty} z^n = 0$, which is necessary for convergence (see Theorem 3.3.2(I)).

Next, let us consider the power series $\sum_{n=1}^{\infty} \frac{z^n}{n}$. If $z = 1$, the series is not convergent. Consequently, the radius of convergence ρ is less or equal to 1. On the other hand, if $z \in [0,1[$, the series is convergent by the ratio test. So we conclude that the radius of convergence is again 1. Observe that, by Leibniz test, the series is convergent for $z = -1$. In fact, it is possible to show that there is convergence for every $z \in \mathbb{C}$, with absolute value equal to 1 and $z \neq 1$.

Finally, let us consider the power series $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$. Still, the radius of convergence is 1 (this can be seen observing that, if $r \in \mathbb{R}^+$, there is convergence if $r \in [0,1]$, there is not convergence if $r > 1$). In general, if $|z| = 1$, one has $|\frac{z^n}{n^2}| = \frac{1}{n^2}$. So, there is absolute convergence for every $z \in \mathbb{C}$ such that $|z| = 1$.

We consider the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$. Differentiating it formally term by term, we obtain $\sum_{n=1}^{\infty} na_n(z-z_0)^{n-1}$, or also, setting $n-1 = m$, $\sum_{m=0}^{\infty}(m+1)a_{m+1}(z-z_0)^m$. We shall call the power series $\sum_{n=0}^{\infty}(n+1)a_{n+1}(z-z_0)^n$ **formal derivative** . We wonder what one can say, concerning the convergence of the formal derivative. The following result holds:

Theorem 3.3.14. *The formal derivative of a power series has the same radius of convergent as the original series.*

Proof Let us indicate with ρ the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$, with ρ' the radius of convergence of its formal derivative.

We start by showing that $\rho' \leq \rho$. This is obvious if $\rho' = 0$. Otherwise, let $z \in \mathbb{C}$ be such that $|z-z_0| < \rho'$. Then, for every $n \in \mathbb{N}$, we have

$$|a_n(z-z_0)^n| = \frac{|z-z_0|}{n}|na_n(z-z_0)^{n-1}| \leq |z-z_0||na_n(z-z_0)^{n-1}|.$$

As the series $\sum_{n=1}^{\infty} |na_n(z-z_0)^{n-1}|$ is convergent, this implies that the series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ is convergent, and so, $\rho' \leq \rho$.

On the other hand, let us check that $\rho \leq \rho'$. Again, this is obvious if $\rho = 0$. So we suppose that $\rho \in]0,+\infty]$. Let $z \in \mathbb{C}$, with $|z-z_0| < \rho$. We show that the series $\sum_{n=0}^{\infty}(n+1)a_{n+1}(z-z_0)^n$ is absolutely convergent. We fix $z_1 \in \mathbb{C}$ such that $|z-z_0| < |z_1-z_0| < \rho$. Then the series

$\sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ is convergent. Employing an argument already used, we obtain that there exists $M \geq 0$ such that $|a_n(z_1 - z_0)^n| \leq M$ $\forall n \in \mathbb{N}_0$. It follows that, $\forall n \in \mathbb{N}_0$,

$$|(n+1)a_{n+1}(z - z_0)^n| = |a_{n+1}(z_1 - z_0)^{n+1}||\frac{n+1}{z_1 - z_0}||\frac{z - z_0}{z_1 - z_0}|^n \leq$$

$$\leq \frac{M}{|z_1 - z_0|}|(n+1)|\frac{z - z_0}{z_1 - z_0}|^n.$$

The series $\sum_{n=0}^{\infty} \frac{M}{|z_1 - z_0|}|(n+1)|\frac{z-z_0}{z_1-z_0}|^n$ is convergent, taking into account that $|\frac{z-z_0}{z_1-z_0}| < 1$, as a consequence of the ratio test.

So, the formal derivative is absolutely convergent whenever $|z - z_0| < \rho$. It follows that $\rho \leq \rho'$.

□

Theorem 3.3.14 makes the following result, that we limit ourselves to state, plausible:

Theorem 3.3.15. *Let $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ be a power series with radius of convergence $\rho > 0$. We set $f : B(z_0, \rho) \to \mathbb{C}$ ($f : \mathbb{C} \to \mathbb{C}$ if $\rho = +\infty$), $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$. Then f is holomorphic. Moreover, for every $z \in B(z_0, \rho)$ ($z \in \mathbb{C}$ if $\rho = +\infty$) one has*

$$f'(z) = \sum_{n=1}^{\infty} na_n(z - z_0)^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}(z - z_0)^n.$$

Corollary 3.3.16. *Let $A \subseteq \mathbb{C}$ be open, $f : A \to \mathbb{C}$ analytic. Then:*

(I) f is holomrphic, it has complex derivatives of any order and these derivatives are holomorphic;

(II) if, for some $r > 0$, $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ in $B(z_0, r)$, then

$$a_n = \frac{f^{(n)}(z_0)}{n!} \qquad \forall n \in \mathbb{N}_0.$$

Proof The proof follows almost immediately from Theorem 3.3.15. In fact, if, for some $r > 0$, $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ in $B(z_0, r)$, the radius of convergence of $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ is greater or equal than r. It follows from Theorem 3.3.15 that f is holomorphic in $B(z_0, r)$ and

$$f'(z) = \sum_{n=1}^{\infty} na_n(z - z_0)^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}(z - z_0)^n \quad \forall z \in B(z_0, r). \tag{3.3.2}$$

Applying now Theorems 3.3.14 and 3.3.15 to the formal derivative, we deduce that we can differentiate f' in $B(z_0, r)$ and obtain that $\forall z \in B(z_0, r)$ one has

$$f''(z) = \sum_{n=1}^{\infty}(n+1)na_{n+1}(z - z_0)^{n-1} = \sum_{n=0}^{\infty}(n+2)(n+1)a_{n+2}(z - z_0)^n \tag{3.3.3}$$

and so on.

Concerning (II), one has, first of all, $a_0 = f(z_0)$ and, from (3.3.2) and (3.3.3),

$$f'(z_0) = a_1, f''(z_0) = 2a_2.$$

Differentiating again, (II) can be obtained in general.

□

The inverse of Corollary 3.3.16 is also true:

Theorem 3.3.17. *Let $A \subseteq \mathbb{C}$ be open, $f : A \to \mathbb{C}$ holomorphic. Then f is analytic in A.*

Incomplete proof Let $z_0 \in A$ and $r > 0$ be such that $\{z \in \mathbb{C} : |z - z_0| \leq r\} \subseteq A$. Next, let $z \in \mathbb{C}$ be such that $|z - z_0| < r$. Then it is clear that $C_r(z_0)$ is $A \setminus \{z\}-$homotopic to $C_\rho(z)$, if $\rho > 0$ is sufficiently small. From Cauchy's integral formula we deduce

$$f(z) = \frac{1}{2\pi i} \int_{C_r(z_0)} \frac{f(v)}{v - z} dv. \tag{3.3.4}$$

If $|v - z_0| = r$, one has

$$\frac{1}{v - z} = \frac{1}{(v - z_0) - (z - z_0)} = \frac{1}{v - z_0} \frac{1}{1 - \frac{z - z_0}{v - z_0}} =$$

$$= \frac{1}{v - z_0} \sum_{n=0}^{\infty} (\frac{z - z_0}{v - z_0})^n = \sum_{n=0}^{\infty} (v - z_0)^{-n-1} (z - z_0)^n,$$

where we have used the fact that $|\frac{z - z_0}{v - z_0}| < 1$. Replacing the first term of this sequence of identities with the last in (3.3.4), and assuming that it is possible to carry the series outside the integral (which should be proved), we obtain

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n, \tag{3.3.5}$$

with

$$a_n = \frac{1}{2\pi i} \int_{C_r(z_0)} f(v)(v - z_0)^{-n-1} dv. \tag{3.3.6}$$

□

Remark 3.3.18. From Corollary 3.3.16 we already know that a_n (defined in (3.3.6)) coincides with $\frac{f^{(n)}(z_0)}{n!}$. Examining the proof of Theorem 3.3.17, we deduce also that the power series $\sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n$ has radius of convergence at least equal to r, for every $r > 0$ such that $\{z \in \mathbb{C} : |z - z_0| \leq r\} \subseteq A$ and that, for such values of r, if $|z - z_0| < r$,

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n \tag{3.3.7}$$

holds.

Remark 3.3.19. Owing to Theorems 3.3.7 and 3.3.17, the classes of analytic and of holomorphic functions coincide. As analytic functions are equipped of derivatives of any order and the derivatives are all analytic, we draw the (at first sight surprising) fact that every holomorphic function has, in fact, complex derivatives of any order and the derivatives are holomorphic. This phenomenon has nothing corresponding for functions of one real variable: it is like saying that every function of class C^1 is automatically of class C^∞!

Example 3.3.20. Let $f : \mathbb{C} \to \mathbb{C}$, $f(z) = e^z$. We know from Example 3.1.10 that f is holomorphic. So, by Theorem 3.3.17, f is analytic in $\mathbb{C}$. We recall that $f'(z) = f(z)$, $\forall z \in \mathbb{C}$.

Therefore, $f^{(n)}(z) = e^z$ $\forall n \in \mathbb{N}_0$, $\forall z \in \mathbb{C}$. In particular, $f^{(n)}(0) = 1$ $\forall n \in \mathbb{N}_0$. We obtain, keeping into account that the domain of f is $\mathbb{C}$ and applying Remark 3.3.18, that

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \qquad \forall z \in \mathbb{C}. \tag{3.3.8}$$

From formulas (3.1.10) and (3.1.11), it follows that, for every $z \in \mathbb{C}$,

$$\cos(z) =$$
$$= \frac{e^{iz} + e^{-iz}}{2} = \frac{1}{2} \sum_{n=0}^{\infty} \frac{(iz)^n + (-iz)^n}{n!}.$$

If $n \in \mathbb{N}_0$,

$$(iz)^n + (-iz)^n = [1 + (-1)^n] i^n z^n.$$

So, $(iz)^n + (-iz)^n = 0$ if n is odd. On the contrary, if $n = 2k$, with $k \in \mathbb{N}_0$, we have

$$(iz)^n + (-iz)^n = 2i^{2k} z^{2k} = 2(-1)^k z^{2k}.$$

It follows that

$$\frac{1}{2} \sum_{n=0}^{\infty} \frac{(iz)^n + (-iz)^n}{n!} = \frac{1}{2} \sum_{k=0}^{\infty} 2(-1)^k \frac{z^{2k}}{(2k)!},$$

implying the classical formula

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} \qquad \forall z \in \mathbb{C}. \tag{3.3.9}$$

Analogously, one can prove

$$\sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^{2k+1} z^{2k+1}}{(2k+1)!} \qquad \forall z \in \mathbb{C}. \tag{3.3.10}$$

For this, see Exercise 3.3.22.

Exercise 3.3.21. Prove Corollary 3.3.9.

Exercise 3.3.22. Prove formula (3.3.10.

Exercise 3.3.23. Show that

$$\cosh(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \qquad \forall x \in \mathbb{R}.$$

Exercise 3.3.24. Show that

$$\sinh(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \qquad \forall x \in \mathbb{R}.$$

Exercise 3.3.25. Let log be the logarithm function considered in Example 3.1.11, with domain $\mathbb{C} \setminus \{x \in \mathbb{R} : x \leq 0\}$. Prove that, if $|z - 1| < 1$ the formula

$$\log(z) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(z-1)^n}{n}$$

holds (Hint: observe that $f^{(n)}(z) = (-1)^{n-1} \frac{(n-1)!}{z^n}$ $(n \in \mathbb{N})$.

3.4 Isolated singularities and Laurent expansions

Holomorphic functions allow some series expansion even in some neighbourhood of an isolated singularity. In order to clarify the result we are interested in, we start by precising what we mean with the term "isolated singularity".

Definition 3.4.1. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, $z_0 \in \mathbb{C}$. We shall say that f has an isolated singularity in z_0 if $z_0 \notin A$, but there exists $r > 0$ such that $B(z_0, r) \setminus \{z_0\} \subseteq A$.*

Let now $a_n \in \mathbb{C}$ for each $n \in \mathbb{Z}$. If the two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=1}^{\infty} a_{-n}$ are convergent, we set

$$\sum_{n=-\infty}^{+\infty} a_n := \sum_{n=0}^{+\infty} a_n + \sum_{n=1}^{+\infty} a_{-n}. \tag{3.4.1}$$

The result we are interested in is the following:

Theorem 3.4.2. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, z_0 an isolated singularity of f. Then there exist, uniquely determined, two power series $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ and $\sum_{n=1}^{\infty} a_{-n}v^n$ such that:*

(I) if $r > 0$ and $B(z_0, r) \setminus \{z_0\} \subseteq A$, the series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ has radius of convergence at least equal to r;

(II) $\sum_{n=1}^{\infty} a_{-n}v^n$ has radius of convergence $+\infty$;

(III) if $r > 0$ is such that $B(z_0, r) \setminus \{z_0\} \subseteq A$, $\forall z \in B(z_0, r) \setminus \{z_0\}$ one has

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n(z - z_0)^n. \tag{3.4.2}$$

Incomplete proof Let $r > 0$ be such that $B(z_0, r) \setminus \{z_0\} \subseteq A$ and $z \in B(z_0, r) \setminus \{z_0\}$. We fix r_1 and r_2 positive, so that

$$r_1 < |z - z_0| < r_2 < r.$$

We assume (for example) that $Im(z) > Im(z_0)$. Let $0 < \epsilon < Arg(z)$ (see Example 3.1.11). Let α be a piecewise C^1 path, describing the following sequence of curves:

$a)$ the segment with endpoints $z_0 + r_1 e^{i\epsilon}$ and $z_0 + r_2 e^{i\epsilon}$;

$b)$ the arch of circumference of radius r_2 and centre z_0, with endpoints $z_0 + r_2 e^{i\epsilon}$ and $z_0 + r_2$, in counterclockwise sense on the circumference itself;

$c)$ the segment with endpoints $z_0 + r_2$ e $z_0 + r_1$;

$d)$ the arch of circumference of radius r_1 and centre z_0 with endpoints $z_0 + r_1$ and $z_0 + r_1 e^{i\epsilon}$, in clockwise sense on the circumference itself; clearly, this path is $A \setminus \{z\}-$homotopic to $C_r(z)$, for every $r > 0$ sufficiently small. So, by Cauchy's integral formula,

$$f(z) = \frac{1}{2\pi i} \int_{\alpha} \frac{f(v)}{v - z} dv.$$

Letting ϵ go to 0, the pieces $a)$ e $c)$ tend to eliminate each other, the piece $b)$ tends to coincide with the circumference $C_{r_2}(z_0)$, the piece $d)$ tends to coincide with the circumference of centre z_0 and radius r_1, described in clockwise sense. These arguments make plausible, letting ϵ go to 0, the formula

$$f(z) = \frac{1}{2\pi i} \int_{C_{r_2}(z_0)} \frac{f(v)}{v - z} dv - \frac{1}{2\pi i} \int_{C_{r_1}(z_0)} \frac{f(v)}{v - z} dv. \tag{3.4.3}$$

The first integral in the second term of (3.4.3) can be treated like the integral (3.3.4): setting, for $n \in \mathbb{N}_0$,

$$a_n = \frac{1}{2\pi i}\int_{C_{r_2}(z_0)} f(v)(v-z_0)^{-n-1}dv, \tag{3.4.4}$$

we have

$$\frac{1}{2\pi i}\int_{C_{r_2}(z_0)} \frac{f(v)}{v-z}dv = \sum_{n=0}^{\infty} a_n(z-z_0)^n.$$

Concerning the second integral in (3.4.3), if $|v-z_0| = r_1$, one has

$$\frac{1}{v-z} = \frac{1}{(v-z_0)-(z-z_0)} = -\frac{1}{z-z_0}\frac{1}{1-\frac{v-z_0}{z-z_0}} =$$

$$= -\frac{1}{z-z_0}\sum_{n=0}^{\infty}(\frac{v-z_0}{z-z_0})^n = -\sum_{n=0}^{\infty}(v-z_0)^n(z-z_0)^{-n-1},$$

where we have used the fact that $|\frac{v-z_0}{z-z_0}| < 1$. So, replacing the first formula of this chain with the last in (3.4.3), and assuming (this should be proved) that we can carry the series outside the integral, we obtain

$$-\frac{1}{2\pi i}\int_{C_{r_1}(z_0)} \frac{f(v)}{v-z}dv = \sum_{n=1}^{\infty} a_{-n}(z-z_0)^{-n}, \tag{3.4.5}$$

with

$$a_{-n} = \frac{1}{2\pi i}\int_{C_{r_1}(z_0)} f(v)(v-z_0)^{n-1}dv, \tag{3.4.6}$$

for every $n \in \mathbb{N}$.

We stop here, without entering into further details of the proof.

□

Remark 3.4.3. With reference to the (incomplete) proof of Theorem 3.4.2, we can say (applying Corollary 3.2.19) that, if $R > 0$ is such that $B(z_0, R) \setminus \{z_0\} \subseteq A$, for every $z \in B(z_0, R) \setminus \{z_0\}$ the formula

$$\sum_{n=-\infty}^{+\infty} a_n(z-z_0)^n, \tag{3.4.7}$$

holds, with

$$a_n = \frac{1}{2\pi i}\int_{C_r(z_0)} f(v)(v-z_0)^{-n-1}dv \qquad \forall n \in \mathbb{Z}, \tag{3.4.8}$$

where r is an arbitrary element of $]0, R[$.

The expansion (3.4.7) is called the **Laurent expansion** of f around the point z_0.

Now we pass to classify isolated singular points of holomorphic functions:

Definition 3.4.4. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, z_0 an isolated singularity of f. Let (3.4.7) be the Laurent expansion of f around z_0. We shall say that:*

a) z_0 is a **removable singularity** *of f if $a_n = 0$ $\forall n \in \mathbb{Z}$, $n < 0$;*

b) z_0 is a **polar singularity** *of f if $\{n \in \mathbb{Z} : n < 0$ and $a_n \neq 0\}$ is finite, but not empty;*

c) z_0 is an **essential singularity** *of f if $\{n \in \mathbb{Z} : n < 0$ and $a_n \neq 0\}$ is infinite.*

In case b), it is also said that z_0 is a **pole** *of f. We shall say that the pole is of order n_0 if n_0 is the maximum natural number n, such that $a_{-n} \neq 0$. The poles of order 1 are also called simple poles.*

Remark 3.4.5. In case a), there exists $R > 0$ such that, if $|z - z_0| < R$, one has $f(z) = \sum_{n=0}^{+\infty} a_n(z - z_0)^n$. So it is clear that, extending f to z_0, setting $f(z_0) = a_0$, we obtain a holomorphic function in $A \cup \{z_0\}$. From this the term "removable" comes.

Remark 3.4.6. If $\sum_{n=-\infty}^{+\infty} a_n(z - z_0)^n$ is the Laurent expansion of f around the isolated singularity z_0, the power series $\sum_{n=1}^{\infty} a_{-n}v^n$ has convergence radius $+\infty$. Consequently, the series $\sum_{n=1}^{\infty} a_{-n}(z - z_0)^{-n}$ converges for every $z \in \mathbb{C} \setminus \{z_0\}$ and its sum gives a function which is holomorphic in this set.

Let us examine some examples:

Example 3.4.7. Let $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $f(z) = \frac{e^z - 1}{z}$. For every $z \in \mathbb{C} \setminus \{0\}$ we have

$$f(z) = \frac{1}{z}\left(\sum_{n=0}^{\infty} \frac{z^n}{n!} - 1\right) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} = \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!}$$

So 0 is a removable singularity of f. The function

$$g : \mathbb{C} \to \mathbb{C},$$

$$g(z) = \begin{cases} \frac{e^z - 1}{z} & \text{se} \quad z \neq 0, \\ 1 & \text{se} \quad z = 0 \end{cases}$$

is holomorphic in the whole $\mathbb{C}$.

Example 3.4.8. Let $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $f(z) = \frac{e^z}{z^2}$. For every $z \in \mathbb{C} \setminus \{0\}$, we have

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{n-2}}{n!} = \sum_{n=-2}^{\infty} \frac{z^n}{(n+2)!}.$$

So $a_n = 0$ if $n < -2$, but $a_{-2} = a_{-1} = 1$. Therefore, 0 is a pole of order 2 for f.

Example 3.4.9. Let $f : \mathbb{C} \setminus \{0\} \to \mathbb{C}$, $f(z) = e^{\frac{1}{z}}$. For every $z \in \mathbb{C} \setminus \{0\}$, one has

$$f(z) = \sum_{n=0}^{\infty} \frac{z^{-n}}{n!} = \sum_{n=-\infty}^{0} \frac{z^n}{(-n)!}.$$

So $a_n = \frac{1}{(-n)!}$ for every $n \in \mathbb{Z}$, $n \leq 0$. We conclude that 0 is an essential singularity of f.

3.5 The residue theorem

In this section we shall illustrate the so called residue theorem, a basic tool for the computation of complex integrals, which has many important applications. We start with the definition of residue:

Definition 3.5.1. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, $z_0 \in \mathbb{C}$ an isolated singularity of f. We shall call* **residue** *of f in z_0, and we shall indicate with $Res(f, z_0)$, the term a_{-1} in the Laurent expansion of f around z_0.*

We describe a simple method of calculation of the residue in the case of a polar singularity.

Lemma 3.5.2. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, $z_0 \in \mathbb{C}$ an isolated singularity of f. We suppose that z_0 is a pole of order less or equal to n ($n \in \mathbb{N}$). Then*

$$Res(f, z_0) = \frac{1}{(n-1)!} \lim_{z \to z_0} (\frac{d}{dz})^{n-1}[(z - z_0)^n f(z)].$$

Incomplete proof We limit ourselves to the case $n = 3$. Then there exists $r > 0$ such that, if $0 < |z - z_0| < r$, we have

$$f(z) = a_{-3}(z - z_0)^{-3} + a_{-2}(z - z_0)^{-2} + a_{-1}(z - z_0)^{-1} + a_0 + a_1(z - z_0) + ...,$$

hence

$$(z - z_0)^3 f(z) = a_{-3} + a_{-2}(z - z_0) + a_{-1}(z - z_0)^2 + a_0(z - z_0)^3 + a_1(z - z_0)^4 + ...,$$

$$(\frac{d}{dz})^2[(z - z_0)^3 f](z) = 2a_{-1} + 6a_0(z - z_0) + 12a_1(z - z_0)^2 + ...,$$

which implies the conclusion.

□

As a premise to the following example, we put here some simple results of practical usefulness. We start with a complex version of De L'Hopital's theorem:

Lemma 3.5.3. *Let $z_0 \in \mathbb{C}$, $r > 0$, let $f, g : B(z_0, r) \setminus \{z_0\} \to \mathbb{C}$ be holomorphic and such that $\lim_{z \to z_0} f(z) = \lim_{z \to z_0} g(z) = 0$, $g(z) \neq 0$ $\forall z \in B(z_0, r) \setminus \{z_0\}$. Moreover, we assume that $g'(z) \neq 0$ $\forall z \in B(z_0, r) \setminus \{z_0\}$ and there exists in $\mathbb{C}$ $\lim_{z \to z_0} \frac{f'(z)}{g'(z)}$. Then, there exists $\lim_{z \to z_0} \frac{f(z)}{g(z)}$ and coincides with $\lim_{z \to z_0} \frac{f'(z)}{g'(z)}$.*

Incomplete proof It is possible to check that the natural extensions of f and g to $B(z_0, r)$ are holomorphic in $B(z_0, r)$. Suppose (calling again, for simplicity, these extensions f and g) that $g'(z_0) \neq 0$. Then, $\forall z \in B(z_0, r)$, $f(z) = \sum_{n=1}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n$ and $g(z) = \sum_{n=1}^{\infty} \frac{g^{(n)}(z_0)}{n!}(z - z_0)^n$. As $g'(z_0) \neq 0$, we have

$$\lim_{z \to z_0} \frac{f'(z)}{g'(z)} = \frac{f'(z_0)}{g'(z_0)}.$$

On the other hand, if $z \neq z_0$,

$$\frac{f(z)}{g(z)} = \frac{(z - z_0) \sum_{n=0}^{\infty} \frac{f^{(n+1)}(z_0)}{(n+1)!}(z - z_0)^n}{(z - z_0) \sum_{n=0}^{\infty} \frac{g^{(n+1)}(z_0)}{(n+1)!}(z - z_0)^n}$$

$$= \frac{\sum_{n=0}^{\infty} \frac{f^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n}{\sum_{n=0}^{\infty} \frac{g^{(n+1)}(z_0)}{(n+1)!}(z-z_0)^n} \to \frac{f'(z_0)}{g'(z_0)}(z \to z_0).$$

□

Lemma 3.5.2 requires an estimate of the order of the pole. We illustrate a simple result in this direction. If $f : A \to \mathbb{C}$ is holomorphic (with A open subset of $\mathbb{C}$) and $z_0 \in A$, we shall say that f has in z_0 a **zero of order** m ($m \in \mathbb{N}_0$) if $f^{(k)}(z_0) = 0$ per $k \in \mathbb{N}_0$, $k < m$, while $f^{(m)}(z_0) \neq 0$. Observe that, if $f(z_0) \neq 0$, then f has in z_0 a zero of order 0.

Lemma 3.5.4. *Let A be an open subset of $\mathbb{C}$, $f, g : A \to \mathbb{C}$ holomorphic, $z_0 \in A$. Suppose that z_0 is zero of order m for f, of order n for g, with m and n elements of $\mathbb{N}_0$. Next, let $h : \{z \in A : g(z) \neq 0\} \to \mathbb{C}$, $h(z) = \frac{f(z)}{g(z)}$. Then:*
(I) if $m \geq n$, h has in z_0 a removable singularity;
(II) if $m < n$, h has in z_0 a pole of order $n - m$.

Proof If $r > 0$ is sufficiently small, we have, for $|z - z_0| < r$,

$$f(z) = \sum_{k=m}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k =$$

$$= (z-z_0)^m \sum_{r=0}^{\infty} \frac{f^{(r+m)}(z_0)}{(r+m)!}(z-z_0)^r := (z-z_0)^m k(z)$$

and

$$g(z) = \sum_{k=n}^{\infty} \frac{g^{(k)}(z_0)}{k!}(z-z_0)^k =$$

$$(z-z_0)^n \sum_{r=0}^{\infty} \frac{g^{(r+n)}(z_0)}{(r+n)!}(z-z_0)^r := (z-z_0)^n l(z).$$

The functions k and l are holomorphic in $\{z \in \mathbb{C} : |z - z_0| < r\}$, $k(z_0) = \frac{f^{(m)}(z_0)}{m!} \neq 0$ and $l(z_0) = \frac{g^{(n)}(z_0)}{n!} \neq 0$. Lowering (if necessary) r, we may assume $l(z) \neq 0$ if $|z - z_0| < r$. The function $\frac{k}{l}$ allows a certain expansion $\sum_{k=0}^{\infty} c_k(z-z_0)^k$ in $B(z_0, r)$ with $c_0 \neq 0$. So, if $0 < |z - z_0| < r$, one has

$$h(z) = (z-z_0)^{m-n} \sum_{k=0}^{\infty} c_k(z-z_0)^k = \sum_{k=0}^{\infty} c_k(z-z_0)^{k+m-n} =$$

$$= \sum_{j=m-n}^{\infty} c_{j-m+n}(z-z_0)^j,$$

which implies the conclusion.

□

Example 3.5.5. Let $A := \{z \in \mathbb{C} : \sin(z) \neq 0\}$, $f : A \to \mathbb{C}$, $f(z) = \frac{e^z - 1}{\sin^2(z)}$. Employing the result in Exercise 3.1.18(VI), we have immediately that 0 is an isolated singularity for f. We determine $Res(f, 0)$. Lemma 3.5.2 requires an upper estimate of the order of the pole. The function $z \to \sin^2(z)$ has in 0 a zero of order 2, while the function $z \to e^z - 1$ has in 0 a simple zero (that is, of order 1).

So, by virtue of Lemma 3.5.4, f has in 0 a pole of order 1 (or simple). Hence, by Lemma 3.5.2, we have

$$Res(f,0) = \lim_{z\to 0} zf(z) = \lim_{z\to 0} \frac{z(e^z-1)}{\sin^2(z)}.$$

In order to compute this limit, we try to apply Lemma 3.5.3. Differentiating at the numerator and at the denominator, we obtain $\frac{e^z-1+ze^z}{2\sin(z)\cos(z)}$, which is again in the undetermined form $\frac{0}{0}$. Differentiang another time, we get $\frac{2e^z+ze^z}{2(\cos^2(z)-\sin^2(z))}$, which tends to 1, as $z\to 0$. We conclude that $Res(f,0)=1$.

As a premise to the main result, we put the notion of **index** of a closed path with respect to a point.

Definition 3.5.6. *Let α be a closed, piecewise C^1 path in $\mathbb{C}$ and let $z_0 \in \mathbb{C}$, with z_0 not belonging to the support of α. We define the index of α with respect to z_0 and indicate with $ind(\alpha, z_0)$ the complex number*

$$ind(\alpha, z_0) := \frac{1}{2\pi i}\int_\alpha (z-z_0)^{-1}dz. \tag{3.5.1}$$

It is possible to prove the following properties of the index:

Theorem 3.5.7. *Let α be a a closed, piecewise C^1 path in $\mathbb{C}$. Then*

(I) $\forall z \in \mathbb{C}\setminus supp(\alpha)$ $ind(\alpha, z_0) \in \mathbb{Z}$;

(II) if $A \subseteq \mathbb{C}\setminus supp(\alpha)$ and A is arcwise connected, then $ind(\alpha, z)$ is the same for every element z of A;

(III) if $A \subseteq \mathbb{C}\setminus supp(\alpha)$ and A is arcwise connected and unbounded, then $ind(\alpha, z) = 0$ $\forall z \in A$.

Remark 3.5.8. In practice, $ind(\alpha, z_0)$ indicates the number of times α "turns around" z_0, counting 1 for each lap in counterclockwise sense, -1 for each lap in clockwise sense.

Example 3.5.9. Let $n \in \mathbb{N}$ and $\alpha : [0, 2\pi] \to \mathbb{C}$, $\alpha(t) = e^{int}$. Then $supp(\alpha) = \{z \in \mathbb{C} : |z| = 1\}$. Let $z_0 \in \mathbb{C}$, with $|z_0| \neq 1$. By Theorem 3.5.7(III) applied to $A := \{z \in \mathbb{C} : |z| > 1\}$, we have in this case $ind(\alpha, z_0) = 0$. On the contrary, assume that $|z_0| < 1$. Then, applying Theorem 3.5.7(II) to $A := \{z \in \mathbb{C} : |z| < 1\}$, we can say that

$$ind(\alpha, z_0) = ind(\alpha, 0) = \frac{1}{2\pi i}\int_\alpha z^{-1}dz =$$

$$= \frac{1}{2\pi i}\int_0^{2\pi} e^{-int} ine^{int}dt = n.$$

In fact, α turns around z_0 n times in counterclockwise sense. Let us consider, instead, $\beta : [0, 2\pi] \to \mathbb{C}$, $\beta(t) = e^{i(2\pi - t)} = e^{-it}$, turning around (for example) 0 once in clockwise sense. Then we have

$$ind(\beta, 0) = \frac{1}{2\pi i}\int_\beta z^{-1}dz =$$

$$= \frac{1}{2\pi i}\int_0^{2\pi} e^{it}(-i)e^{-it}dt = -1.$$

We shall need also the following

Lemma 3.5.10. *Let A be an open subset of $\mathbb{C}$, $f : A \to \mathbb{C}$ holomorphic, α a piecewise C^1 path with support in A. Suppose that thre exists $F : A \to \mathbb{C}$ holomorphic, such that $F'(z) = f(z)$ $\forall z \in A$. Then*

$$\int_\alpha f(z)dz = 0.$$

Incomplete proof Suppose that α is of class C^1. Then, if $\alpha : [a, b] \to \mathbb{C}$,

$$\int_\alpha f(z)dz = \int_a^b f(\alpha(t))\alpha'(t)dt =$$

(applying the fundamental theorem of integral calculus)

$$= [F(\alpha(t)]_{t=a}^{t=b} = 0,$$

because $\alpha(a) = \alpha(b)$.

□

We pass to the main result of this section, providing (as will be clear in the following) an important tool for computation:

Theorem 3.5.11. *(The residue theorem) Let A be a simply connected open subset of $\mathbb{C}$, let z_1,...,z_n be elements of A, $f : A \setminus \{z_1, ..., z_n\} \to \mathbb{C}$ holomorphic, α a closed, piecewise C^1 path, with support in $A \setminus \{z_1, ..., z_n\}$. Then*

$$\int_\alpha f(z)dz = 2\pi i \sum_{i=1}^{n} Res(f, z_i)ind(\alpha, z_i). \tag{3.5.2}$$

Incomplete proof We consider, for $i = 1, ..., n$, the Laurent expansion of f around z_i. If this expansion is $\sum_{n=-\infty}^{+\infty} a_{i,n}(z - z_i)^n$, we set

$$S_i(z) := \sum_{n=-\infty}^{-1} a_{i,n}(z - z_i)^n = \sum_{n=1}^{+\infty} a_{i,-n}(z - z_i)^{-n}.$$

We have already observed (see Remark 3.4.6) that S_i is holomorphic in $\mathbb{C} \setminus \{z_i\}$. Let us consider now the function $z \to f(z) - S_1(z) - ... - S_n(z)$, holomorphic in $A \setminus \{z_1, ..., z_n\}$. We observe that each of the points $z_1, ..., z_n$ is a removable singularity for this function. To see this, let us take (for example) $i = 1$. Then $f(z) - S_1(z) = \sum_{n=0}^{+\infty} a_{i,n}(z - z_i)^n$ for z sufficiently close to z_1,while S_2, ..., S_n are regular in z_1. So, applying Corollary 3.2.20, we have

$$\int_\alpha f(z)dz =$$

$$= \int_\alpha (f(z) - S_1(z) - ... - S_n(z))dz + \int_\alpha (S_1(z) + ... + S_n(z))dz$$

$$= \int_\alpha (S_1(z) + ... + S_n(z))dz.$$

For $i = 1, ..., n$, one has

$$\int_\alpha S_i(z)dz = \int_\alpha \sum_{n=1}^{+\infty} a_{i,-n}(z - z_i)^{-n}dz$$

$$= \sum_{n=1}^{+\infty} a_{i,-n} \int_\alpha (z - z_i)^{-n} dz$$

(we do not justify this passage). Now let us consider, for $n \in \mathbb{N}$, the integral $\int_\alpha (z - z_i)^{-n} dz$. If $n \neq 1$, we have

$$(z - z_i)^{-n} = F'(z),$$

with $F(z) = (1 - n)^{-1}(z - z_i)^{1-n}$. Applying Lemma 3.5.3, we conclude that, if $n \neq 1$, $\int_\alpha (z - z_i)^{-n} dz = 0$. So we have that

$$\int_\alpha S_i(z)dz = a_{i,-1} \int_\alpha (z - z_i)^{-1} dz =$$

$$= 2\pi i Res(f, z_i) ind(\alpha, z_i).$$

□

Example 3.5.12. Let $\alpha = C_2(2i)$. We want to compute

$$\int_\alpha \frac{1}{z^4 - 8z^2 - 9} dz. \tag{3.5.3}$$

The complex zeros of the polynomial $P(z) := z^4 - 8z^2 - 9$ are 3, -3, i e $-i$. So we set $f : \mathbb{C} \setminus \{3, -3, i, -i\} \to \mathbb{C}$, $f(z) = \frac{1}{z^4 - 8z^2 - 9}$. f is holomorphic and, applying the residue theorem, we deduce that the integral in (3.5.3) coincides with

$$2\pi i[Res(f, 3)ind(\alpha, 3) + Res(f, -3)ind(\alpha, -3)+$$

$$+Res(f, i)ind(\alpha, i) + Res(f, -i)ind(\alpha, -i)].$$

Now we observe that $|3 - 2i| = |-3 - 2i| = \sqrt{13} > 2$ and $|-i - 2i| = 3 > 2$. So, by Theorem 3.5.7(III), we have

$$ind(C_2(2i), 3) = ind(C_2(2i), -3) = ind(C_2(2i), -i) = 0.$$

On account of Remark 3.5.8, we can also write

$$ind(C_2(2i), i) = 1.$$

So the integral in (3.5.3) coincides with $2\pi i Res(f, i)$. In order to compute $Res(f, i)$, we observe that $P'(i) = -20i \neq 0$, so that P has in i a simple zero. It follows from Lemma 3.5.4 that f has a simple pole in i. So

$$Res(f, i) = \lim_{z \to i} (z - i)f(z) = \lim_{z \to i} \frac{1}{4z^3 - 16z} = \frac{i}{20}.$$

We conclude that the integral in (3.5.3) equals $2\pi i \frac{i}{20} = -\frac{\pi}{10}$.

Example 3.5.13. We employ the residue theorem to compute

$$\int_{\mathbb{R}} \frac{1}{x^4 + 1} dx. \tag{3.5.4}$$

First of all, we observe that the function we want to integrate is continuous and so measurable in $\mathbb{R}$. Moreover, it is positive. So the integral exists in the sense of Definition 1.2.6. We can also apply the monotone convergence theorem to deduce that (3.5.4) coincides with

$$\lim_{n\to+\infty}\int_{-n}^{n}\frac{1}{x^4+1}dx.$$

In fact,

$$\int_{-n}^{n}\frac{1}{x^4+1}dx=\int_{\mathbb{R}}\frac{\chi_n(x)}{x^4+1}dx,$$

where we have indicated with χ_n the characteristic function of the interval $[-n,n]$, and for each $n\in\mathbb{N}$ and $\forall x\in\mathbb{R}$,

$$\frac{\chi_n(x)}{x^4+1}\le\frac{\chi_{n+1}(x)}{x^4+1}$$

and

$$\lim_{n\to+\infty}\frac{\chi_n(x)}{x^4+1}=\frac{1}{x^4+1}.$$

Now we consider the closed, piecewise C^1 path, obtained describing firstly the interval $[-n,n]$, then the semi-circumference $\{ne^{it}\,:\,t\in[0,\pi]\}$. The function $z\to\frac{1}{z^4+1}$ is holomorphic in $\mathbb{C}\setminus\{e^{i\frac{(2k+1)\pi}{4}}\,:0\le k\le 3,k\in\mathbb{Z}\}$. If $n>1$, none of the complex numbers of the form $e^{i\frac{(2k+1)\pi}{4}}$ belongs to the support of α_n. It is also clear that

$$ind(\alpha_n,e^{i\frac{\pi}{4}})=ind(\alpha_n,e^{i\frac{3\pi}{4}})=1,$$

while

$$ind(\alpha_n,e^{i\frac{5\pi}{4}})=ind(\alpha_n,e^{i\frac{7\pi}{4}})=0.$$

So, it follows from the residue theorem that

$$\int_{\alpha_n}\frac{1}{z^4+1}dz=2\pi i[Res(f,e^{i\frac{\pi}{4}})+Res(f,e^{i\frac{3\pi}{4}})],$$

with $f(z)=\frac{1}{z^4+1}$. It is not difficult to check that f has in $e^{i\frac{\pi}{4}}$ a pole of order 1. So, applying Lemmata 3.5.2 and 3.5.3, we can write

$$\begin{aligned}Res(f,e^{i\frac{\pi}{4}})&=\lim_{z\to e^{i\frac{\pi}{4}}}\frac{z-e^{i\frac{\pi}{4}}}{z^4+1}\\&=\lim_{z\to e^{i\frac{\pi}{4}}}\frac{1}{4z^3}=\frac{e^{-i\frac{3\pi}{4}}}{4}.\end{aligned}$$

In the same way,

$$\begin{aligned}Res(f,e^{i\frac{3\pi}{4}})&=\lim_{z\to e^{i\frac{3\pi}{4}}}\frac{z-e^{i\frac{3\pi}{4}}}{z^4+1}\\&=\lim_{z\to e^{i\frac{3\pi}{4}}}\frac{1}{4z^3}=\frac{e^{-i\frac{9\pi}{4}}}{4}.\end{aligned}$$

So

$$\int_{\alpha_n}\frac{1}{z^4+1}dz=2\pi i(\frac{e^{-i\frac{3\pi}{4}}}{4}+\frac{e^{-i\frac{9\pi}{4}}}{4})=\frac{\pi}{\sqrt{2}}.$$

On the other hand,

$$\int_{\alpha_n} \frac{1}{z^4+1} dz =$$

$$= \int_{-n}^{n} \frac{1}{x^4+1} dx + \int_{C_n^+(0)} \frac{1}{z^4+1} dz,$$

with

$$\begin{cases} C_n^+(0) : [0,\pi] \to \mathbb{C}, \\ C_n^+(0)(t) = ne^{it}. \end{cases} \tag{3.5.5}$$

If $|z| = n$ e $n > 1$, one has, applying the result of Exercise 2.1.14,

$$|z^4+1| \geq ||z^4|-1|| = n^4 - 1 > 0,$$

so that

$$|f(z)| \leq \frac{1}{n^4-1}$$

if $|z| = n > 1$. So, applying Theorem 3.2.12, we have

$$|\int_{C_n^+(0)} \frac{1}{z^4+1} dz| \leq \frac{n\pi}{n^4-1} \to 0 (n \to +\infty).$$

Summing up, we have, for $n \geq 2$,

$$\int_{-n}^{n} \frac{1}{x^4+1} dx + \int_{C_n^+(0)} \frac{1}{z^4+1} dz = \frac{\pi}{\sqrt{2}},$$

so that, at the limit for $n \to +\infty$,

$$\int_{\mathbb{R}} \frac{1}{x^4+1} dx = \frac{\pi}{\sqrt{2}}.$$

Example 3.5.14. We want to compute

$$\int_{\mathbb{R}^+} \frac{x^\alpha}{1+x^2} dx, \tag{3.5.6}$$

with $\alpha \in \mathbb{R}$.

Whatever α is, if we set $f : \mathbb{R}^+ \to \mathbb{R}$, $f(x) = \frac{x^\alpha}{1+x^2}$, f is measurable, because it is continuous, and nonnegative. So the integral is always defined. We would like to know, first of all, for which values of α it is a real number. One has

$$\int_{\mathbb{R}^+} f(x) dx = \int_{]0,1]} f(x) dx + \int_{]1,+\infty[} f(x) dx. \tag{3.5.7}$$

The function $g(x) := x^\beta$ is summable in $]0,1]$ if and only if $\beta > -1$. If $0 < x \leq 1$, we have

$$f(x) \leq x^\alpha \leq 2f(x).$$

So the first integral in (3.5.7) is finite if and only if $\alpha > -1$. Moreover, the function $g(x) := x^\beta$ is summable in $[1, +\infty[$ if and only if $\beta < -1$. If $x \geq 1$, we have

$$f(x) \leq x^{\alpha-2}$$

and

$$x^{\alpha-2} = f(x)(1+\frac{1}{x^2}) \le 2f(x).$$

So the second integral in (3.5.7) is finite if and only if $\alpha - 2 < -1$, that is, $\alpha < 1$. We conclude that the integral in (3.5.6) si finite if and only if

$$-1 < \alpha < 1. \tag{3.5.8}$$

From now on, we shall assume that condition (3.5.8) is satisfied.

Now we consider, for a certain $n \in \mathbb{N}$, the closed, piecewise C^1 path α_n obtained describing once, in counterclockwise sense, first the interval $[\frac{1}{n}, n]$, then, in the order, the semi-circumference $\{z \in \mathbb{C} : |z| = n, Im(z) \ge 0\}$, $[-n, -\frac{1}{n}]$, the semi-circumference $\{z \in \mathbb{C} : |z| = \frac{1}{n}, Im(z) \ge 0\}$. We prolounge the function f as a holomorphic function in a proper open subset of $\mathbb{C}$. To this aim, we start by considering the logarithm function log defined in $A := \mathbb{C} \setminus \{iy : y \in \mathbb{R}, y \le 0\}$ in the following way:

$$log(z) = \ln(|z|) + i\theta, \theta \in arg(z) \cap]-\frac{\pi}{2}, \frac{3\pi}{2}[. \tag{3.5.9}$$

Next, we set, for each $\beta \in \mathbb{R}$ (recalling Example 3.1.12)

$$z^\beta := e^{\beta log(z)} \tag{3.5.10}$$

Now we consider the function

$$\begin{cases} g : A \setminus \{i\} \to \mathbb{C}, \\ \quad g(z) = \frac{z^\alpha}{1+z^2}. \end{cases}$$

g is holomorphic and its restriction to $\mathbb{R}^+$ is f. Applying now the residue theorem, we have, for $n \ge 2$,

$$\int_{\alpha_n} g(z)dz = 2\pi i Res(g, i).$$

We can easily check that g has a simple pole in i. So, applying again Lemma 3.5.3 and recalling Example 3.1.12, we have

$$Res(g, i) = \lim_{z\to i}(z-i)g(z)$$

$$= \lim_{z\to i} \frac{z^\alpha + (z-i)\alpha z^{\alpha-1}}{2z} = \frac{i^\alpha}{2i} = \frac{e^{\alpha log(i)}}{2i}$$

$$= \frac{e^{i\alpha\frac{\pi}{2}}}{2i},$$

hence

$$\int_{\alpha_n} g(z)dz = \pi e^{i\alpha\frac{\pi}{2}}.$$

On the other hand,

$$\int_{\alpha_n} g(z)dz =$$

$$= \int_{\frac{1}{n}}^{n} \frac{x^\alpha}{1+x^2}dx + \int_{C_n^+(0)} \frac{z^\alpha}{1+z^2}dz + \int_{-n}^{-\frac{1}{n}} \frac{x^\alpha}{1+x^2}dx - \int_{C_{\frac{1}{n}}^+(0)} \frac{z^\alpha}{1+z^2}dz.$$

If $x < 0$, one has

$$x^\alpha = e^{\alpha log(x)} = e^{\alpha(ln(-x)+i\pi)} = (-x)^\alpha e^{i\alpha\pi},$$

hence

$$\int_{-n}^{-\frac{1}{n}} \frac{x^\alpha}{1+x^2}dx = e^{i\alpha\pi}\int_{-n}^{-\frac{1}{n}} \frac{(-x)^\alpha}{1+x^2}dx = e^{i\alpha\pi}\int_{\frac{1}{n}}^{n} \frac{x^\alpha}{1+x^2}dx.$$

Moreover, if $|z| = R$, one has

$$|z^\alpha| = |e^{\alpha log(z)}| = e^{Re(\alpha log(z))} = e^{\alpha ln(R)} = R^\alpha.$$

If $R > 1$, we have also

$$|1+z^2| = |z^2-(-1)| \geq ||z^2|-1| = R^2-1,$$

while, if $R < 1$,

$$|1+z^2| = |1-(-z^2)| \geq |1-|-z^2|| = 1-R^2.$$

Applying again Theorem 3.2.12, we obtain

$$|\int_{C_n^+(0)} \frac{z^\alpha}{1+z^2}dz| \leq \frac{n^\alpha}{n^2-1}\pi n \to 0(n\to+\infty)$$

and

$$|\int_{C_{\frac{1}{n}}^+(0)} \frac{z^\alpha}{1+z^2}dz| \leq \frac{n^{-\alpha}}{1-n^{-2}}\pi n^{-1} \to 0(n\to+\infty),$$

on account of (3.5.8). We conclude, letting n go to $+\infty$, that

$$(1+e^{i\alpha\pi})\int_{\mathbb{R}^+} \frac{x^\alpha}{1+x^2}dx = \pi e^{i\alpha\frac{\pi}{2}},$$

hence

$$\int_{\mathbb{R}^+} \frac{x^\alpha}{1+x^2}dx = \frac{\pi e^{i\alpha\frac{\pi}{2}}}{1+e^{i\alpha\pi}} = \frac{\pi}{2\cos(\frac{\alpha\pi}{2})}.$$

We examine another example, which will be useful in the following.

Example 3.5.15. Given $\xi \in \mathbb{R}$, we compute

$$\int_{\mathbb{R}} e^{-x^2-ix\xi}dx. \tag{3.5.11}$$

First of all, we observe that

$$|e^{-x^2-ix\xi}| = e^{-x^2}$$

One has $e^{-x^2} = o(x^{-2})$ for $x \to \pm\infty$. So, there exists $M \in \mathbb{R}^+$ such that, if $|x| > M$, $e^{-x^2} < x^{-2}$. Hence,

$$\int_{\mathbb{R}} e^{-x^2}dx \leq \int_{-M}^{M} e^{-x^2}dx + \int_{\{|x|>M\}} x^{-2}dx < +\infty.$$

Therefore, on account of the result of Exercise 1.2.16, the integrand function in (3.5.11) is summable for every $\xi \in \mathbb{R}$. We start by considering the case $\xi = 0$. We set, by convenience,

$$I := \int_{\mathbb{R}} e^{-x^2}dx.$$

Applying Theorem 1.3.8 with $T : \mathbb{R}^+ \times]0, 2\pi[\to \mathbb{R}^2$, $T(\rho, \theta) = (\rho\cos(\theta), \rho\sin(\theta))$, one has

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}dxdy = \int_{\mathbb{R}^+\times]0,2\pi[} e^{-\rho^2}\rho d\rho d\theta$$

$$= 2\pi\int_{\mathbb{R}^+} e^{-\rho^2}\rho d\rho = 2\pi \lim_{c\to+\infty}\int_0^c e^{-\rho^2}\rho d\rho$$

$$= 2\pi \lim_{c\to+\infty}\frac{1}{2}(1-e^{-c^2}) = \pi.$$

On the other hand, by the theorem of Tonelli,

$$\int_{\mathbb{R}^2} e^{-x^2-y^2}dxdy = I^2.$$

As $I > 0$, necessarily $I = \sqrt{\pi}$. Next, we consider the case $\xi \neq 0$. We start by observing that

$$x^2 + ix\xi = (x^2 + 2x\frac{i\xi}{2} - \frac{\xi^2}{4}) + \frac{\xi^2}{4}$$

$$= (x + \frac{i\xi}{2})^2 + \frac{\xi^2}{4}.$$

It follows that

$$\int_{\mathbb{R}} e^{-x^2-ix\xi}dx = e^{-\frac{\xi^2}{4}}\int_{\mathbb{R}} e^{-(x+\frac{i\xi}{2})^2}dx.$$

Suppose now, for example, $\xi > 0$. We consider, given $n \in \mathbb{N}$, a closed, piecewise C^1 path α_n, clockwise oriented, having as support the boundary of the rectangle with vertexes$-n+\frac{i\xi}{2}$, $n+\frac{i\xi}{2}$, n, $-n$. On account of Corollary 3.2.20, one has

$$\int_{\alpha_n} e^{-z^2}dz = 0.$$

On the other hand,

$$\int_{\alpha_n} e^{-z^2}dz =$$

$$= \int_{-n}^{n} e^{-(x+\frac{i\xi}{2})^2}dx - \int_{-n}^{n} e^{-x^2}dx - \int_0^{\xi} e^{-(n+iy)^2}idy + \int_0^{\xi} e^{-(-n+iy)^2}idy$$

If $y \in [0, \xi]$, we have

$$|e^{-(n+iy)^2}| = e^{y^2-n^2} \leq e^{\xi^2-n^2}.$$

It follows that

$$|\int_0^{\xi} e^{-(n+iy)^2}idy| \leq \xi e^{\xi^2-n^2} \to 0 (n \to +\infty).$$

Arguing in a similar way, we see also that

$$\lim_{n\to+\infty}\int_0^{\xi} e^{-(-n+iy)^2}idy = 0.$$

So, letting n go to $+\infty$, we obtain

$$0 = \int_{\mathbb{R}} e^{-(x+\frac{i\xi}{2})^2}dx - \sqrt{\pi},$$

hence

$$\int_{\mathbb{R}} e^{-(x+\frac{i\xi}{2})^2} dx = \sqrt{\pi}.$$

With analogous arguments, one can see that this formula holds even if $\xi < 0$. So, we can conclude that for every $\xi \in \mathbb{R}$ the integral in (3.5.11) is $\sqrt{\pi}e^{-\frac{\xi^2}{4}}$.

Exercise 3.5.16. Compute the residues of the following functions in the specified point z_0:

(I) $\frac{\sin(z)}{z^3}$, $z_0 = 0$;

(II) $(2-z)^{-1}e^{\frac{1}{z}}$, $z_0 = 2$;

(III) $\frac{\sin(z)}{\cos(z)}$, $z_0 = \frac{\pi}{2}$;

(IV) $\frac{z^\alpha}{(z^\beta-1)^2}$, $z_0 = 1$, con $z^\alpha = f_\alpha(z)$ e $z^\beta = f_\beta(z)$ (see Example 3.1.12), taking, as logarithm function, $g(z) = \ln(|z|) + iArg(z)$ for $Re(z) > 0$, where $Arg(z)$ is the element of the argument of z in $]-\pi,\pi[$;

(V) $\frac{z^3}{sin^5(z)}$, $z_0 = 0$.

Exercise 3.5.17. Compute

(I) $\int_0^{2\pi} \frac{1}{\cos(x)+2} dx$;

(II) $\int_0^{2\pi} \frac{1}{\cos(x)+\sin(x)+2} dx$;

(III) $\int_0^{2\pi} \frac{1}{\cos^2(x)+1} dx$;

(IV) $\int_0^{2\pi} \frac{\cos(x)}{\cos(x)+\sin(x)+2} dx$.

(Hint: employ Euler's formulas and transform the given integrals into complex integrals in $C_1(0)$).

Exercise 3.5.18. Compute

(I) $\int_{\mathbb{R}} \frac{1}{x^2+2x+2} dx$;

(II) $\int_{\mathbb{R}} \frac{1}{(x^2+1)^2} dx$;

(III) $\int_{\mathbb{R}} \frac{1}{x^6+1} dx$;

(IV) $\int_{\mathbb{R}} \frac{1}{x^8+1} dx$;

(V) $\int_{\mathbb{R}^+} \frac{x^2}{x^4+1} dx$.

Exercise 3.5.19. Compute, given $\alpha \in]-1,1[$,

$$\int_{\mathbb{R}^+} \frac{x^\alpha \ln(x)}{1+x^2} dx. \tag{3.5.12}$$

Start by showing that the integrand function is summable in $\mathbb{R}^+$, employing the fact that $\ln(x) = o(x^\epsilon)$ for $x \to +\infty$ and $\ln(x) = o(x^{-\epsilon})$ for $x \to 0$, for every $\epsilon > 0$.

Exercise 3.5.20. Computing, checking preliminarily their existence,

(I) $\int_{\mathbb{R}^+} \frac{\ln(x)}{x^4+1} dx$;

(II) $\int_{\mathbb{R}^+} \frac{\ln^2(x)}{x^2+1} dx$;

(III) $\int_{\mathbb{R}^+} \frac{x^\alpha}{x^4+1} dx$ $(-1 < \alpha < 3)$.

3.6 Holomorphic functions and harmonic functions

We start by introducing, in $\mathbb{R}^n (n \in \mathbb{N})$ the Laplace operator Δ:

$$\Delta := \sum_{j=1}^{n} \frac{\partial^2}{\partial x_j^2} = \frac{\partial^2}{\partial x_1^2} + ... \frac{\partial^2}{\partial x_n^2}. \tag{3.6.1}$$

The differential operator Δ plays a remarkable role in mathematical physics.

Definition 3.6.1. *Let A be an open subset of $\mathbb{R}^n$, $u \in C^2(A)$. We shall say that u is* **harmonic** *in A if*

$$\Delta u(x) = 0 \qquad \forall x \in A.$$

Remark 3.6.2. Examples of harmonic functions are polynomial functions of degree less or equal to one. Other examples are $u : \mathbb{R}^2 \to \mathbb{R}$, $u(x, y) = e^x \cos(y)$ and $v : \mathbb{R}^2 \to \mathbb{R}$, $v(x, y) = x^2 - y^2$, as one can easily check.

In case $n = 2$, there exists a remarkable connection between harmonic and holomorphic functions:

Theorem 3.6.3. *Let A be an open subset in $\mathbb{C}$, f $A \to \mathbb{C}$ holomorphic, $u : A \to \mathbb{R}$, $u(z) = u(x, y) = Re(f(z))$. Then u is harmonic in A.*

On the other hand, let A be a simply connected open subset of $\mathbb{R}^2$ (which we identify with $\mathbb{C}$) and let $u : A \to \mathbb{R}$ be harmonic. Then there exists $f : A \to \mathbb{C}$ holomorphic, such that $u = Re(f)$.

Proof Let $f : A \to \mathbb{C}$ be holomorphic, $u = Re(f)$. We recall (see Remark 3.3.19) that u is of class C^∞, as f has complex derivatives of any order, which are all holomorphic. We check that $\Delta u(x, y) = 0$ $\forall (x, y) \in A$. Let $v := Im(f)$. Then, applying the Cauchy-Riemann conditions and the theorem of Schwarz, we have, $\forall (x, y) \in A$,

$$\Delta u(x, y) = \frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y)$$

$$= \frac{\partial^2 v}{\partial x \partial y}(x, y) - \frac{\partial^2 v}{\partial y \partial x}(x, y) = 0.$$

On the other hand, let A be an open simply connected subset of $\mathbb{R}^2$ and let $u : A \to \mathbb{R}$ be harmonic. We want to construct $f : A \to \mathbb{C}$ holomorphic, such that $u = Re(f)$. To this aim, owing to Theorem 3.1.8, it suffices to determine $v \in C^1(A)$ such that

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{3.6.2}$$

and set

$$\begin{cases} f : A \to \mathbb{C}, \\ f(z) = u(z) + iv(z), \quad z \in A. \end{cases}$$

Let us consider the vector field $F : A \to \mathbb{R}^2$,

$$F(x, y) := (-\frac{\partial u}{\partial y}(x, y), \frac{\partial u}{\partial x}(x, y)).$$

As u is harmonic, F is closed and so it has a potential, because A is simply conncected. Therefore, there exists $v \in C^1(A)$ fulfilling (3.6.2). With this, the result is completely proved.

□

Remark 3.6.4. If $f : A \to \mathbb{C}$ is holomorphic, (A open subset of $\mathbb{C}$), as $Im(f) = Re(-if)$, even $Im(f)$ is harmonic in A.

Remark 3.6.5. We have already observed that the real and the imaginary part of a holomorphic function are of class C^∞. So, on account of the fact that open balls are simply connected, in force of Theorem 3.6.3, given $u : A \to \mathbb{R}$ harmonic in A open subset of $\mathbb{R}^2$, we can construct in every open ball in A a holomorphic function, the real part of which is u. So we can conclude that, at least in the case $n = 2$, harmonic functions are of class C^∞. In fact, it is possible to show that this is true in any dimension.

3.7 Maximum principle and Dirichlet problem for the Laplace equation in the standard circle of $\mathbb{R}^2$

In this section we illustrate some results concerning the Dirichlet problem for the Laplace equation. Such problem can be formulated as follows: let A be an open subset in $\mathbb{R}^n$ ($n \in \mathbb{N}$) and let $g : \partial A \to \mathbb{R}$ be continuous. We look for functions $u : \overline{A} \to \mathbb{R}$, continuous in $\overline{A}$ and of class C^2 in A, such that

$$\begin{cases} \Delta u(x) = 0, & \forall x \in A, \\ u(x') = g(x'), & \forall x' \in \partial A. \end{cases} \tag{3.7.1}$$

We begin with a first important result, the so called **maximum principle**.

Theorem 3.7.1. *(Maximum principle) Let A be an open bounded subset of $\mathbb{R}^n$, $u \in C(\overline{A}) \cap C^2(A)$ real valued, such that*

$$\Delta u(x) \geq 0 \qquad \forall x \in A. \tag{3.7.2}$$

Then

$$\max_{\overline{A}} u = \max_{\partial A} u. \tag{3.7.3}$$

Proof First of all, we observe that, as A is bounded, $\overline{A}$ and ∂A are closed and bounded (see Exercises 3.7.4 and 3.7.5). So, by the theorem of Weierstrass, $\max_{\overline{A}} u$ and $\max_{\partial A} u$ exist. As $\partial A \subseteq \overline{A}$, we can certainly say that

$$\max_{\partial A} u \leq \max_{\overline{A}} u. \tag{3.7.4}$$

We want to reverse inequality (3.7.4). To this aim, we start by considering the less general case

$$\Delta u(x) > 0 \qquad \forall x \in A. \tag{3.7.5}$$

We show that, in this case, u cannot have points of maximum in A, so that it follows that its points of maximum are necessarily in ∂A and so (3.7.3) holds. We argue by contradiction, assuming that (3.7.5) holds and there exists $x^0 \in A$, which is a point of maximum for u. Let us consider the quadratic form

$$Q(h) := \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 u}{\partial x_i x_j}(x^0) h_i h_j, \qquad h = (h_1, ..., h_n) \in \mathbb{R}^n. \tag{3.7.6}$$

We recall that Q is negative semidefinite (see D. Guidetti, "Analisi Matematica B", Teorema 2.5.2). It follows that, if $k \in \{1, ..., n\}$ and e^k is the $k-$th element of the standard basis of $\mathbb{R}^n$,

$$0 \geq Q(e^k) = \frac{\partial^2 u}{\partial x_k^2}(x^0).$$

So $\Delta u(x^0) \leq 0$, in contradiction with (3.7.5).

Let us consider now the general case (3.7.2). We fix $v \in C^2(\mathbb{R}^n)$, real valued, such that $\Delta v(x) > 0$ $\forall x \in \mathbb{R}^n$. A suitable choice is (for example), $v : \mathbb{R}^n \to \mathbb{R}$, $v(x_1, ..., x_n) = x_1^2$, verifying $\Delta v(x) = 2$ $\forall x \in \mathbb{R}^n$. Given $\epsilon > 0$, we set

$$u_\epsilon := u + \epsilon v. \tag{3.7.7}$$

We have

$$\Delta u_\epsilon(x) = \Delta u(x) + \epsilon \Delta v(x) > 0 \qquad \forall x \in A.$$

So, employing the particular case already treated, we get

$$\max_{\overline{A}} u_\epsilon = \max_{Fr(A)} u_\epsilon.$$

This identity implies that, whatever is $x \in \overline{A}$, we have

$$u(x) + \epsilon v(x) \leq \max_{\partial A}(u + \epsilon v). \tag{3.7.8}$$

If $x' \in \partial A$, it holds

$$u(x') + \epsilon v(x') \leq \max_{\partial A} u + \epsilon \max_{\partial A} v,$$

hence, whatever $x \in \overline{A}$ is,

$$u(x) + \epsilon v(x) \leq \max_{\partial A} u + \epsilon \max_{\partial A} v. \tag{3.7.9}$$

Passing to the limit for $\epsilon \to 0^+$ in (3.7.9), we obtain that, whatever $x \in \overline{A}$ is, we have

$$u(x) \leq \max_{\partial A} u. \tag{3.7.10}$$

This implies the conclusion. □

From Theorem 3.7.1 the following uniqueness result can be easily obtained:

Corollary 3.7.2. *Let A be an open, bounded subset in $\mathbb{R}^n$, $u_0, u_1 \in C(\overline{A}) \cap C^2(A)$ real valued, such that $\Delta u_0(x) = \Delta u_1(x)$ $\forall x \in A$, $u_0(x') = u_1(x')$ $\forall x' \in \partial A$. Then $u_0(x) = u_1(x)$ $\forall x \in \overline{A}$.*

Proof We set $u : \overline{A} \to \mathbb{R}$, $u(x) = u_0(x) - u_1(x)$. Then $u \in C(\overline{A}) \cap C^2(A)$, it is real valued, $\Delta u(x) = 0$ $\forall x \in A$, $u(x') = 0$ $\forall x' \in Fr(A)$. Owing to the maximum principle, one has

$$u(x) \leq 0 \qquad \forall x \in \overline{A}. \tag{3.7.11}$$

On the other hand, $\Delta(-u)(x) = 0$ $\forall x \in A$, $-u(x') = 0$ $\forall x' \in Fr(A)$ hold also. So, again by the maximum principle, we have $-u(x) \leq 0$ $\forall x \in \overline{A}$, that is,

$$u(x) \geq 0 \qquad \forall x \in \overline{A}. \tag{3.7.12}$$

From (3.7.11) and (3.7.12) we obtain the conclusion. □

Let us consider now problem (3.7.1) in case $A = \{z \in \mathbb{C} : |z| < 1\} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ (the "standard circle").

Theorem 3.7.3. *Let $A = \{z \in \mathbb{C} : |z| < 1\}$, $g \in C(\partial A; \mathbb{R})$. Then there exists a unique $u \in C(\overline{A}; \mathbb{R}) \cap C^2(A)$ solving (3.7.1).*

Partial proof The uniqueness follows from Corollary 3.7.2.

Concerning the existence, we shall construct heuristically a certain solution of the problem. We shall not verify in all the details that the function we are going to construct is really a solution.

So, let us suppose that a solution u with the required properties exists. By Theorem 3.6.3, there exists a holomorphic function $f : A \to \mathbb{C}$, such that $u = Re(f)$. Applying Theorem 3.3.17 and Remark 3.3.18, we may say that, for every $z \in \mathbb{C}$, with $|z| < 1$, we have

$$f(z) = \sum_{n=0}^{\infty} \alpha_n z^n, \tag{3.7.13}$$

with the radius of convergence of the power series $\sum_{n=0}^{\infty} \alpha_n z^n$ at least equal to 1. So, if $r \in [0, 1[$ and $\theta \in]-\pi, \pi]$, we have

$$\begin{array}{rcl} u(re^{i\theta}) & = & Re(f(re^{i\theta})) \\ & = & Re(\sum_{n=0}^{\infty} \alpha_n r^n e^{in\theta}) \\ & = & \sum_{n=0}^{\infty} a_n r^n \cos(n\theta) + \sum_{n=1}^{\infty} b_n r^n \sin(n\theta), \end{array} \tag{3.7.14}$$

where we have put

$$a_n := Re(\alpha_n), \qquad n \in \mathbb{N}_0, \tag{3.7.15}$$

$$b_n := -Im(\alpha_n), \qquad n \in \mathbb{N}. \tag{3.7.16}$$

Now we have to determine the coefficients $(a_n)_{n \in \mathbb{N}_0}$ and $(b_n)_{n \in \mathbb{N}}$. Arguing formally, for $r = 1$, we get:

$$\begin{array}{rcl} g(e^{i\theta}) & = & u(\cos(\theta), \sin(\theta)) \\ & = & \sum_{n=0}^{\infty} a_n \cos(n\theta) + \sum_{n=1}^{\infty} b_n \sin(n\theta) \\ & = & a_0 + \sum_{n=1}^{\infty} \frac{a_n - ib_n}{2} e^{in\theta} + \sum_{n=1}^{\infty} \frac{a_n + ib_n}{2} e^{-in\theta}. \end{array} \tag{3.7.17}$$

The last expression in (3.7.17) is the Fourier series expansion of $\theta \to g(e^{i\theta})$. So

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{it}) dt, \tag{3.7.18}$$

and, for each $n \in \mathbb{N}$,

$$a_n + ib_n = \frac{1}{\pi} \int_{-\pi}^{\pi} e^{int} g(e^{it}) dt,$$

hence

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nt) g(e^{it}) dt, \tag{3.7.19}$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nt) g(e^{it}) dt, \tag{3.7.20}$$

From (3.7.14), (3.7.18), (3.7.19), (3.7.20), we obtain that, for every $r \in [0, 1[$ and $\theta \in]-\pi, \pi]$, carrying the series inside the integral

$$\begin{array}{rcl} u(re^{i\theta}) & = & \frac{1}{\pi} \int_{-\pi}^{\pi} \{\frac{1}{2} + \sum_{n=1}^{\infty} r^n [\cos(n\theta)\cos(nt) + \sin(n\theta)\sin(nt)] g(e^{it}) dt \\ & = & \frac{1}{\pi} \int_{-\pi}^{\pi} [\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(n(\theta - t))] g(e^{it}) dt \end{array} \tag{3.7.21}$$

If $r \in [0,1[$, $s \in \mathbb{R}$, one has

$$\begin{array}{rcl} \frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(ns) & = & \frac{1}{2} + Re\{\sum_{n=1}^{\infty} r^n e^{ins}\} \\ & = & \frac{1}{2} + rRe\{\frac{e^{is}}{1-re^{is}}\} \\ & = & \frac{1}{2} + r\frac{\cos(s)-r}{1-2r\cos(s)+r^2} \\ & = & \frac{1-r^2}{2(1-2r\cos(s)+r^2)}. \end{array}$$

So we have obtained the following classical formula:

$$u(re^{i\theta}) = \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(\theta-t)+r^2} g(e^{it})dt. \tag{3.7.22}$$

Now, it would be possible to show that the function $u : \overline{A} \to \mathbb{R}$, such that

$$u(re^{i\theta}) = \begin{cases} \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(\theta-t)+r^2} g(e^{it})dt & \text{se} \quad 0 \le r < 1, \\ g(e^{i\theta}) & \text{se} \quad r = 1 \end{cases}$$

belongs to $C(\overline{A};\mathbb{R}) \cap C^2(A)$ and is harmonic in A. Here we shall limit ourselves ti verify that

$$\lim_{r\to 1} u(r,0) = g(1,0). \tag{3.7.23}$$

To this aim, let us observe preliminarly that, for every $r \in [0,1[$, one has

$$\frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(s)+r^2} ds = 1. \tag{3.7.24}$$

In fact, assuming that it is possible to invert the series with the integral, we get

$$\begin{array}{c} \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(s)+r^2} ds \\ = \frac{1}{\pi}\int_{-\pi}^{\pi} (\frac{1}{2} + \sum_{n=1}^{\infty} r^n \cos(ns))ds \\ = \frac{1}{\pi}(\pi + \sum_{n=1}^{\infty} r^n \int_{-\pi}^{\pi} \cos(ns))ds) \\ = 1. \end{array}$$

Moreover,

$$\frac{1-r^2}{2\pi(1-2r\cos(s)+r^2)} > 0 \qquad \forall r \in [0,1[, s \in]-\pi,\pi],$$

being

$$1-2r\cos(s)+r^2 \ge 1-2r+r^2 = (1-r)^2.$$

So,

$$\begin{array}{rcl} |u(r,0)-g(1,0)| & = & |\frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(t)+r^2}(g(e^{it})-g(1,0))dt| \\ & \le & \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi} \frac{1}{1-2r\cos(t)+r^2}|g(e^{it})-g(1,0)|dt. \end{array} \tag{3.7.25}$$

Let us fix $\epsilon \in \mathbb{R}^+$. As g is continuous, there exists $\delta \in]0,\frac{\pi}{2}[$, such that, if $|t| < \delta$, one has

$$|g(e^{it}) - g(1,0)| < \frac{\epsilon}{2}.$$

From (3.7.25), it follows

$$
\begin{aligned}
|u(r,0)-g(1,0)| \leq \ & \tfrac{\epsilon}{2}\tfrac{1-r^2}{2\pi}\int_{-\delta}^{\delta}\tfrac{1}{1-2r\cos(t)+r^2}dt \\
& +2\max_{\partial A}|g|\tfrac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\tfrac{1}{1-2r\cos(t)+r^2}dt \\
& +\int_{\delta}^{\pi}\tfrac{1}{1-2r\cos(t)+r^2}dt).
\end{aligned} \tag{3.7.26}
$$

We have

$$
\frac{1-r^2}{2\pi}\int_{-\delta}^{\delta}\frac{1}{1-2r\cos(t)+r^2}dt \leq \frac{1-r^2}{2\pi}\int_{-\pi}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt = 1.
$$

Moreover, if $|t| \geq \delta$, $\cos(t) \leq \cos(\delta) < 1$ holds. It follows that

$$
1-2r\cos(t)+r^2 \geq 1-2r\cos(\delta)+r^2
$$

With elementary methods, one can show that

$$
\min_{[0,1]}(1-2r\cos(\delta)+r^2) = \sin^2(\delta) > 0.
$$

So

$$
\begin{aligned}
& 2\max_{\partial A}|g|\frac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\frac{1}{1-2r\cos(t)+r^2}dt + \int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt) \\
& = 2\max_{\partial A}|g|\frac{1-r^2}{\pi}\int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt \\
& \leq 2\max_{\partial A}|g|\frac{1-r^2}{\pi}\frac{\pi-\delta}{\sin^2(\delta)},
\end{aligned}
$$

converging to 0 as $r \to 1$.

We conclude that there exists $r(\epsilon) \in [0,1[$, such that, if $r(\epsilon) < r < 1$, one has

$$
2\max_{\partial A}|g|\frac{1-r^2}{2\pi}(\int_{-\pi}^{-\delta}\frac{1}{1-2r\cos(t)+r^2}dt + \int_{\delta}^{\pi}\frac{1}{1-2r\cos(t)+r^2}dt) < \frac{\epsilon}{2},
$$

and so, if $r(\epsilon) < r < 1$,

$$
|u(r,0)-g(1,0)| < \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon.
$$

□

Exercise 3.7.4. Let $A \subseteq \mathbb{R}^n$, A bounded. Show that $\overline{A}$ and ∂A are bounded.

Exercise 3.7.5. Let $A \subseteq \mathbb{R}^n$. Show that ∂A is closed.

3.8 Conformal mappings and the Dirichlet problem for the Laplace equation in $\mathbb{R}^2$

Now we want to extend Theorem 3.7.3 to more general open sets. To this aim, we shall employ certain changes of variables, known in literature as ”conformal mappings”.

Definition 3.8.1. *Let A and B be open subsets of $\mathbb{C}$. A* **conformal mapping** *from B to A is a function $\phi : B \to A$ such that:*

(a) ϕ is a bijection between B e A;

(b) ϕ is holomorphic;

(c) $\phi'(z) \neq 0$ $\forall z \in B$.

We shall say that B is conformally equivalent to A if there exists a conformal mapping from B to A.

The following result, which we shall not prove, holds:

Theorem 3.8.2. *Let A and B be open subsets of $\mathbb{C}$ and let ϕ be a conformal mapping from B to A. Then the inverse mapping ϕ^{-1} is conformal from A to B.*

Remark 3.8.3. The only item which is not obvious, in the proof of Theorem 3.8.2, is that ϕ^{-1} is holomorphic. Supposing that this is true, the derivative of ϕ^{-1} can be easily computed: it suffices to observe that $\phi^{-1}(\phi(z)) = z$ $\forall z \in B$. So, applying the rule of differentiation of composed functions, we obtain

$$(\phi^{-1})'(\phi(z))\phi'(z) = 1 \qquad \forall z \in B,$$

hence

$$(\phi^{-1})'(\phi(z)) = \frac{1}{\phi'(z)} \qquad \forall z \in B, \tag{3.8.1}$$

which is a natural extension of the formula of differentiation of inverse functions seen in Analisis A.

Remark 3.8.4. The term "conformal mapping" refers to the fact that these mappings "preserve the angles between curves". To give a precise meaning to this statement, let us consider two paths of class C^1 α and β, with domain $[-\delta, \delta]$ ($\delta > 0$) and support in $\mathbb{R}^2$. Suppose that $\alpha(0) = \beta(0)$ and the vectors $\alpha'(0)$ and $\beta'(0)$ are both nonzero. So we can define the "angle between α and β relative to $t = 0$" as the angle of measure $\theta \in [0, \pi]$ lying between the tangent vectors $\alpha'(0)$ and $\beta'(0)$. Recalling the geometric interpretation of the scalar product in $\mathbb{R}^2$, we have

$$\theta = \arccos\left(\frac{\alpha'(0) \cdot \beta'(0)}{\|\alpha'(0)\|\|\beta'(0)\|}\right),$$

where we have indicated with $\|.\|$ the Euclidean norm in $\mathbb{R}^2$, coinciding with the absolute value in $\mathbb{C}$. We observe that, if $z = (z_1, z_2)$ and $v = (v_1, v_2)$ are both elements of $\mathbb{R}^2$, if we indicate with zv their product in $\mathbb{C}$, we have

$$z \cdot v = z_1 v_1 + z_2 v_2 = Re(z\overline{v}).$$

So,

$$\theta = \arccos\left(\frac{Re(\alpha'(0)\overline{\beta'(0)})}{|\alpha'(0)||\beta'(0)|}\right).$$

Let now B be an open subset of $\mathbb{C}$, containing the supports of α and β and let $\phi : B \to A$ be a conformal mapping, with A open in $\mathbb{C}$. We consider the paths $\phi \circ \alpha$ and $\phi \circ \beta$ and the angle θ' between them, relative to $t = 0$. As $(\phi \circ \alpha)'(0) = \phi'(\alpha(0))\alpha'(0)$ and

$$(\phi \circ \beta)'(0) = \phi'(\beta(0))\beta'(0) = \phi'(\alpha(0))\beta'(0),$$

we have

$$\begin{aligned}\theta' &= \arccos\left(\frac{Re(\phi'(\alpha(0))\alpha'(0)\overline{\phi'(\alpha(0))\beta'(0)})}{|\phi'(\alpha(0)|^2|\alpha'(0)||\beta'(0)|}\right)\\ &= \arccos\left(\frac{Re(\alpha'(0)\overline{\beta'(0)})}{|\alpha'(0)||\beta'(0)|}\right)\\ &= \theta.\end{aligned}$$

Example 3.8.5. Let $B := \{z \in \mathbb{C} : |z| < 1\}$, $\phi : B \to \mathbb{C}$, $\phi(z) = \frac{1+z}{1-z}$. ϕ is holomorphic and, for every $z \in B$,

$$\phi'(z) = \frac{2}{(1-z)^2} \neq 0.$$

Let $v \in \mathbb{C}$. Let us consider the equation

$$\phi(z) = v, \qquad z \in B. \tag{3.8.2}$$

If $v \neq -1$, the equation (3.8.2) has in $\mathbb{C}$ the unique solution $z = \frac{v-1}{v+1}$. Such solution belongs to B if and only if

$$|v-1| < |v+1|,$$

which is equivalent ro $Re(v) > 0$. So ϕ is a conformal mapping between B and $\{v \in \mathbb{C} : Re(v) > 0\}$.

Remark 3.8.6. It is not difficult to verify that the relation "being conformally equivalent" is of equivalence between open subsets of $\mathbb{C}$ (Exercise 3.8.16).

At this point it is natural to wonder whether, given two open subsets of $\mathbb{C}$, they are conformally equivalent. A first famous step in this direction is the following classical theorem, due to Riemann:

Theorem 3.8.7. *Let A be a nonempty open subset of $\mathbb{C}$. Then A is conformally equivalent to $\{z \in \mathbb{C} : |z| < 1\}$ if and only if it is simply connected and does not coincide with $\mathbb{C}$.*

For a proof, see W. Rudin "Real and complex analysis", Chapter 14.

Theorem 3.8.7 tells us that, if A is a proper simply connected open subset of $\mathbb{C}$, there exists a conformal mapping ϕ from A to $\{z \in \mathbb{C} : |z| < 1\}$, but says nothing about a continuous extension of ϕ to $\overline{A}$. So, we introduce the two following definitions:

Definition 3.8.8. *Let $E \subseteq \mathbb{C}$ and $F \subseteq \mathbb{C}$. A* **homeomorphism** *from E to F is a function $\phi : E \to F$, injective and surjective, continuous, with inverse ϕ^{-1} continuous.*

If there exists a homeomorphism from E to F, we shall say that E is homeomorphic to F.

Remark 3.8.9. It is easy to see that the relation "being homeomorphic to" is of equivalence between subsets of $\mathbb{C}$. (Exercise 3.8.16). Moreover, it is obvious that conformal mappings are homeomorphisms.

Definition 3.8.10. *Let A be an open subset of $\mathbb{C}$, $\beta \in \partial A$. We shall say that it is a* **simple point** *of ∂A if it enjoys the following property: for an arbitrary sequence (α_n) in A, converging to β, it is possible to construct a continuous path $\gamma : [0,1] \to \mathbb{C}$ and a sequence $(t_n)_{n\in\mathbb{N}}$, with $0 \leq t_1 < t_2 < ...$, $t_n \to 1$, such that $\gamma(t_n) = \alpha_n$ $\forall n \in \mathbb{N}$, $\gamma([0,1[) \subseteq A$ e $\gamma(1) = \beta$.*

Remark 3.8.11. It is easy to check that, if A is a convex open subset of $\mathbb{C}$, every point of ∂A is simple for ∂A (see Exercise 3.8.18)). On the contrary, let us give an example of a point of the boundary which is not simple.

Let

$$A := \{z \in \mathbb{C} : |z| < 1\} \setminus \{z \in \mathbb{R} : z \in [0,1[\}.$$

We set $\beta = 1/2$ and, for $n \in \mathbb{N}$,

$$\alpha_n := \begin{cases} 1/2 + i/n & \text{if } n \text{ is odd,} \\ 1/2 - i/n & \text{if } n \text{ is even.} \end{cases}$$

Let $\gamma : [0,1[\to \mathbb{C}$ be continuous and let $(t_n)_{n\in\mathbb{N}}$ be a sequence satisfying $0 \le t_1 < t_2 < ..., t_n \to 1$, such that $\gamma(t_n) = \alpha_n \ \forall n \in \mathbb{N}$, $\gamma([0,1[) \subseteq A$. Then, necessarily, for each $n \in \mathbb{N}$, there exists τ_n between t_n and t_{n+1}, such that

$$\gamma_1(\tau_n) \le 0. \tag{3.8.3}$$

This implies that the condition $\lim_{t\to 1}\gamma(t) = \beta$ cannot hold. In fact, in this case, there would exist $\delta > 0$, such that, for every $t \in [0,1[$ satisying $t > 1 - \delta$, it would be $|\gamma(t) - 1/2| < 1/2$, which implies $\gamma_1(t) > 0$. As the sequence $(t_n)_{n\in\mathbb{N}}$ converges to 1, for every $\delta \in]0,1[$, for n sufficiently large, it would hold $\tau_n > 1 - \delta$, and so $\gamma_1(\tau_n) > 0$, in contradiction with (3.8.3).

Now we are able to state the following result (for a proof, see Rudin, *"Real and complex analysis"*, Theorem 14.19):

Theorem 3.8.12. *Let A be an open, bounded, simply connected subset of $\mathbb{C}$, whose boundary points are all simple, let ϕ be a conformal mapping from A to $\{z \in \mathbb{C} : |z| < 1\}$. Then ϕ is extensible to a homeomorphism between $\overline{A}$ and $\{z \in \mathbb{C} : |z| \le 1\}$, the restriction of which to ∂A is a homeomorphism between ∂A and $\{z \in \mathbb{C} : |z| = 1\}$.*

Now we come back to Dirichlet's problem for the Laplace equation. The following lemma explains our interest for conformal mappings:

Lemma 3.8.13. *Let B and A be open subsets of $\mathbb{C}$, let $u \in C^2(A;\mathbb{R})$, let $\phi : B \to \mathbb{C}$ be holomorphic, with $\phi(B) \subseteq A$. Then, for every $z \in A$,*

$$\Delta(u \circ \phi)(z) = \Delta u(\phi(z))|\phi'(z)|^2.$$

In particular, if u is harmonic, even $u \circ \phi$ is harmonic.

Proof The proof is an exercise of differentiation.

Let us indicate with ϕ_1 and ϕ_2 the components of ϕ. Then, for every $z \in B$, we have

$$\frac{\partial(u \circ \phi)}{\partial x}(z) = D_1 u(\phi(z))\frac{\partial \phi_1}{\partial x}(z) + D_2 u(\phi(z))\frac{\partial \phi_2}{\partial x}(z),$$

$$\begin{aligned} \tfrac{\partial^2(u\circ\phi)}{\partial x^2}(z) \quad = \quad & D_1^2 u(\phi(z))\tfrac{\partial\phi_1}{\partial x}(z)^2 + 2D_{12}u(\phi(z))\tfrac{\partial\phi_1}{\partial x}(z)\tfrac{\partial\phi_2}{\partial x}(z) \\ & + D_2^2 u(\phi(z))\tfrac{\partial\phi_2}{\partial x}(z)^2 + D_1 u(\phi(z))\tfrac{\partial^2\phi_1}{\partial x^2}(z) \\ & + D_2 u(\phi(z))\tfrac{\partial^2\phi_2}{\partial x^2}(z). \end{aligned} \tag{3.8.4}$$

Analogously, we have

$$\begin{aligned} \tfrac{\partial^2(u\circ\phi)}{\partial y^2}(z) \quad = \quad & D_1^2 u(\phi(z))\tfrac{\partial\phi_1}{\partial y}(z)^2 + 2D_{12}u(\phi(z))\tfrac{\partial\phi_1}{\partial y}(z)\tfrac{\partial\phi_2}{\partial y}(z) \\ & + D_2^2 u(\phi(z))\tfrac{\partial\phi_2}{\partial y}(z)^2 + D_1 u(\phi(z))\tfrac{\partial^2\phi_1}{\partial y^2}(z) \\ & + D_2 u(\phi(z))\tfrac{\partial^2\phi_2}{\partial y^2}(z). \end{aligned} \tag{3.8.5}$$

From the Cauchy Riemann conditions, we obtain

$$\frac{\partial \phi_1}{\partial x}(z)^2 + \frac{\partial \phi_1}{\partial y}(z)^2 = \frac{\partial \phi_1}{\partial x}(z)^2 + \frac{\partial \phi_2}{\partial x}(z)^2 = |\phi'(z)|^2,$$

$$\frac{\partial \phi_1}{\partial x}(z)\frac{\partial \phi_2}{\partial x}(z) + \frac{\partial \phi_1}{\partial y}(z)\frac{\partial \phi_2}{\partial y}(z) = 0,$$

$$\frac{\partial \phi_2}{\partial x}(z)^2 + \frac{\partial \phi_2}{\partial y}(z)^2 = \frac{\partial \phi_1}{\partial y}(z)^2 + \frac{\partial \phi_1}{\partial x}(z)^2 = |\phi'(z)|^2.$$

Moreover, ϕ_1 e ϕ_2 are harmonic by Theorem 3.6.3 and Remark 3.6.4. So. summing (3.8.4) and (3.8.5), we obtain the conclusion. □

Now we are able to state and prove the following

Theorem 3.8.14. *Let A be a bounded, simply connected, open subset of R^2, whose boundary points are all simple, in the sense of Definition 3.8.10. Let $g \in C(\partial A, \mathbb{R})$. Then the Dirichlet problem for the Laplace equation (3.7.1) has a unique solution $u \in C(\overline{A}; \mathbb{R}) \cap C^2(A)$.*

Proof The uniqueness follows from Corollary 3.7.2.

Let us prove the existence. We indicate with B $\{z \in \mathbb{C} : |z| < 1\}$. By Theorem 3.8.12, there exists a conformal mapping ϕ between A and B, which is extensible to a homeomorphism between $\overline{A}$ and $\overline{B}$, which we continue to indicate with ϕ. Let $h : \partial B \to \mathbb{R}$, $h(z) = g(\phi^{-1}(z))$, $z \in \partial B$. $h \in C(\partial B, \mathbb{R})$, as it is the composition of continuous functions. Let us consider the problem

$$\begin{cases} \Delta v(z) = 0, & z \in B, \\ v(z') = h(z'), & z' \in Fr(B). \end{cases}$$

By Theorem 3.7.3, such problem has a unique solution $v \in C(\overline{B}, \mathbb{R}) \cap C^2(B)$. We define $u : \overline{A} \to \mathbb{R}$, $u(z) = v(\phi(z))$. Then $u \in C(\overline{A}, \mathbb{R}) \cap C^2(A)$. By Lemma 3.8.13, u is harmonic in A. Moreover, if $z \in \partial A$, as $\phi(z) \in \partial B$, we have

$$u(z) = v(\phi(z)) = h(\phi(z)) = g(z).$$

The proof is complete. □

Remark 3.8.15. Apart very simple cases (see, for example, Exercises 3.8.19 and 3.8.20), we are not able to write down an explicit expression of a conformal mapping between $B := \{z \in \mathbb{C} : |z| < 1\}$ and A, simply connected open subset of $\mathbb{C}$, not coinciding with it. A particularly important case, for applications, is that A is a bounded polygon. Employing the result of Exercise 3.8.19, we can replace B with the halfspace $\{z \in \mathbb{C} : Im(z) > 0\}$. In this case, conformal mappings between $\{z \in \mathbb{C} : Im(z) > 0\}$ and A can be explicitly exibited, in the form of complex primitives of certain holomorphic functions. These mappings are called "Schwarz-Christoffel" mappings (see, for example, T. W. Gamelin *"Complex Analysis"* (Springer) Chapter XI.

Exercise 3.8.16. 1) Check that the relation"being homeomorphic" is of equivalence between subsets of $\mathbb{C}$.

2) Check that the relation "being conformally equivalent to" is of equivalence between open subsets of $\mathbb{C}$.

Exercise 3.8.17. Let E and F be homeomorphic subsets of $\mathbb{C}$. Check that:

1) if E is closed and bounded, F is closed and bounded;

2) if E is arcwise connected, F is arcwise connected;

3) if E is simply connected, F is simply connected.

Check (constructing a counterexample) that, if E is bounded, F is not necessarily bounded.

Exercise 3.8.18. Check that, if A is a convex open subset of $\mathbb{C}$, all points of ∂A are simple.

(Hint.: let $\beta \in \partial A$ and let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence with values in A, converging to β. Set, for $n \in \mathbb{N}$, $t_n = 1 - 2^{-n}$ and, if $t \in [1 - 2^{-n}, 1 - 2^{-n-1}]$,

$$\gamma(t) = \alpha_n + \frac{\alpha_{n+1} - \alpha_n}{t_{n+1} - t_n}(t - t_n).$$

Observe that, in the interval $[t_n, t_{n+1}]$, γ describes the segment with endpoints α_n and α_{n+1}. Check, using the fact that $\lim_{n\to+\infty} \alpha_n = \beta$, that $\lim_{t\to 1} \gamma(t) = \beta$.)

Exercise 3.8.19. Construct, starting from Example 3.8.5, a conformal mapping between $\{z \in \mathbb{C} : |z| < 1\}$ and $\{z \in \mathbb{C} : Im(z) > 0\}$. Write down also the expression of the inverse mapping. (Hint: construct preliminarly a conformal mapping between $\{z \in \mathbb{C} : Re(z) > 0\}$ and $\{z \in \mathbb{C} : Im(z) > 0\}$)

Exercise 3.8.20. Let $z_0 \in \mathbb{C}$ and $r > 0$. Then the mapping $\phi(z) := \frac{z-z_0}{r}$ is conformal between $\{z \in \mathbb{C} : |z - z_0| < r\}$ e $\{z \in \mathbb{C} : |z| < 1\}$. Use this fact to write an explicit expression for the solution the Dirichlet's problem

$$\begin{cases} \Delta u(z) = 0, & |z - z_0| < r, \\ u(z) = g(z), & |z - z_0| = r, \end{cases}$$

with $g \in C(\{z \in \mathbb{C} : |z - z_0| = r\}, \mathbb{R})$.

Chapter 4

The Fourier transform

4.1 The Fourier transform in $L^1(\mathbb{R}^n)$

We recall (see Example 2.1.10) that $L^1(\mathbb{R}^n)$ is the linear space of equivalence classes of summable functions. It is a Banach space with the norm (2.1.9). In the following, for simplicity, given $f \in \mathcal{L}^1(\mathbb{R}^n)$, we shall write f instead of $[f]$ and we shall think of the elements of $L^1(\mathbb{R}^n)$ as summable functions, identifying elements coinciding almost everywhere. Of course, case by case, the several results we are going to illustrate shall have to be invariant with respect to the equivalence relation.

We pass to introduce the definition of the **Fourier transform** of $f \in L^1(\mathbb{R}^n)$.

Definition 4.1.1. *Let $f \in L^1(\mathbb{R}^n)$. We shall call Fourier transform of f and we shall indicate with the notation $\hat{f}$ or $\mathcal{F}f$ the function*

$$\left\{ \begin{array}{c} \hat{f} : \mathbb{R}^n \to \mathbb{C}, \\ \hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i<x,\xi>} f(x)dx, \end{array} \right. \tag{4.1.1}$$

with $< .,. >$ "standard" inner product in $\mathbb{R}^n$ (see Example 2.4.3).

Remark 4.1.2. For every $\xi \in \mathbb{R}^n$ we have $|e^{-i<x,\xi>} f(x)| = |f(x)|$. It follows that the integral in (4.1.1) is well defined.

Definition 4.1.3. *Let $A \subseteq \mathbb{R}^n$. We indicate with $BC(A)$ the linear space of functions $f : A \to \mathbb{C}$ which are continuous and bounded. We introduce in $BC(A)$ the norm $\|.\|_\infty$, defined as follows:*

$$\|f\|_\infty := \sup\{|f(a)| : a \in A\}. \tag{4.1.2}$$

Remark 4.1.4. One can check, arguing as in the case of $BC(A, \mathbb{R})$, that, with the norm $\|.\|_\infty$, $BC(A)$ is a Banach space (see Exercise 4.1.18).

Theorem 4.1.5. *(Riemann-Lebesgue's theorem) For every $f \in L^1(\mathbb{R}^n)$ the Fourier transform $\hat{f}$ belongs to $BC(\mathbb{R}^n)$. Moreover, we have*

$$\lim_{\|\xi\| \to +\infty} \hat{f}(\xi) = 0, \tag{4.1.3}$$

in the sense that $\forall \epsilon > 0$ there exists $\delta(\epsilon) > 0$ such that, if $\xi \in \mathbb{R}^n$ and $\|\xi\| > \delta(\epsilon)$, $|\hat{f}(\xi)| < \epsilon$ holds.

Incomplete proof From Theorem 1.2.12(IV) it follows that

$$|\hat{f}(\xi)| \leq \int_{\mathbb{R}^n} |f(x)|dx = \|f\|_1. \tag{4.1.4}$$

So $\hat{f}$ is bounded. To show that it is continuous, we can apply Theorem 2.3.12.

Let $(\xi^k)_{k\in\mathbb{N}}$ be a sequence in $\mathbb{R}^n$ such that $\lim_{k\to+\infty} \xi^k = \xi^0 \in \mathbb{R}^n$. For every $k \in \mathbb{N}$, we have

$$|e^{-i<x,\xi^k>}f(x)| = |f(x)| \qquad \forall x \in \mathbb{R}^n.$$

So we can apply the dominated convergence theorem, to conclude that

$$\lim_{k\to+\infty} \hat{f}(\xi^k) = \lim_{k\to+\infty} \int_{\mathbb{R}^n} e^{-i<x,\xi^k>}f(x)dx = \int_{\mathbb{R}^n} e^{-i<x,\xi^0>}f(x)dx = \hat{f}(\xi^0).$$

Hence $\hat{f}$ is continuous.

We do not prove (4.1.3). □

Corollary 4.1.6. *We indicate with $\mathcal{F}$ the mapping*

$$\begin{cases} \mathcal{F} : L^1(\mathbb{R}^n) \to BC(\mathbb{R}^n), \\ \mathcal{F}f = \hat{f}. \end{cases} \tag{4.1.5}$$

Then $\mathcal{F}$ is linear and continuous from $L^1(\mathbb{R}^n)$ to $BC(\mathbb{R}^n)$.

Proof We leave to the reader the trivial proof of linearity (Exercise 4.1.19). We show that $\mathcal{F}$ is continuous. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^1(\mathbb{R}^n)$ such that $\lim_{k\to+\infty} f_k = f$ in $L^1(\mathbb{R}^n)$. We have to show that $\lim_{k\to+\infty} \mathcal{F}f_k = \mathcal{F}f$ in $BC(\mathbb{R}^n)$. From (4.1.4) it follows that $\|\mathcal{F}g\|_\infty \leq \|g\|_1$ $\forall g \in L^1(\mathbb{R}^n)$. This implies that

$$\|\mathcal{F}f - \mathcal{F}f_k\|_\infty = \|\mathcal{F}(f - f_k)\|_\infty \leq \|f - f_k\|_1 \to 0$$

for $k \to +\infty$. □

Example 4.1.7. A consequence of Example 3.5.15 is that, if $f(x) = e^{-x^2}$, then

$$\hat{f}(\xi) = \sqrt{\pi}e^{-\frac{\xi^2}{4}}.$$

Example 4.1.8. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{1+x^2}$. $f \in L^1(\mathbb{R})$. We compute its Fourier transform. We have to calculate

$$\int_R \frac{e^{-ix\xi}}{1+x^2}dx \tag{4.1.6}$$

with ξ varying in $\mathbb{R}$. We start by observing that, if $z \in \mathbb{C}$,

$$|e^{-iz\xi}| = e^{\xi Im(z)}.$$

So, if $\xi \geq 0$, we have $|e^{-iz\xi}| \leq 1$ if $Im(z) \leq 0$, if $\xi \leq 0$, we have $|e^{-iz\xi}| \leq 1$ if $Im(z) \geq 0$. Suppose that $\xi \geq 0$. Given $n \in \mathbb{N}$, $n \geq 2$, we consider the closed path α_n oriented in clockwise sense, describing once, first $[-n, n]$, then the semicircumference $\{z \in \mathbb{C} : |z| = n, Im(z) \leq 0\}$.As the function $f(z) := \frac{e^{-iz\xi}}{1+z^2}$ is holomorphic in $\mathbb{C} \setminus \{i, -i\}$, we have

$$\int_{\alpha_n} f(z)dz = 2\pi i Res(f, -i)Ind(\alpha_n, -i) = -2\pi i Res(f, -i).$$

It is easy to verify that $-i$ is a simple pole for f. So

$$Res(f,-i) = \lim_{z\to -i}(z+i)f(z) = \lim_{z\to -i}\frac{e^{-iz\xi}}{z-i} = i\frac{e^{-\xi}}{2}.$$

It follows

$$\int_{\alpha_n} f(z)dz = \pi e^{-\xi}.$$

On the other hand,

$$\int_{\alpha_n} f(z)dz = \int_{-n}^{n}\frac{e^{-ix\xi}}{1+x^2}dx - \int_{C_r^-(0)} f(z)dz,$$

with

$$\begin{cases} C_r^-(0) : [\pi, 2\pi] \to \mathbb{C}, \\ C_r^-(0)(t) = re^{it}, \end{cases} \tag{4.1.7}$$

for $r > 0$. From the usual Theorem 3.2.12 and from the previous remarks, we obtain

$$|\int_{C_r^0(0)} f(z)dz| \le \frac{1}{n^2-1}\pi n \to 0 \qquad (n\to+\infty).$$

So, passing to the limit with $n \to +\infty$, we obtain

$$\int_{\mathbb{R}}\frac{e^{-ix\xi}}{1+x^2}dx = \pi e^{-\xi}$$

if $\xi \ge 0$.

The case $\xi < 0$ can be treated analogously, integrating on a path describing, in counterclockwise sense, $[-n, n]$ and the semi circumference $\{z \in \mathbb{C} : |z| = n, Im(z) \ge 0\}$, in order to take advantage of the boundedness of $e^{-iz\xi}$ in $\{z \in \mathbb{C} : Im(z) \ge 0\}$. We leave to the reader the completion of the computation. We limit ourselves to say that the following formula holds:

$$\int_{\mathbb{R}}\frac{e^{-ix\xi}}{x^2+1}dx = \pi e^{-|\xi|} \qquad \forall \xi\in\mathbb{R}. \tag{4.1.8}$$

Example 4.1.9. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \chi_+(x)e^{-x}$, with χ_+ characteristic function of $\mathbb{R}^+$. We have to calculate

$$\int_{\mathbb{R}^+} e^{-x(1+i\xi)}dx \tag{4.1.9}$$

with $\xi \in \mathbb{R}$. Applying Theorem 1.3.4, we obtain

$$\int_{\mathbb{R}^+} e^{-x(1+i\xi)}dx = \lim_{c\to+\infty}\int_0^c e^{-x(1+i\xi)}dx = \lim_{c\to+\infty}\frac{1-e^{-c(1+i\xi)}}{1+i\xi}.$$

From $|e^{-c(1+i\xi)}| = e^{-c} \to 0$ for $c \to +\infty$, we obtain

$$\hat{f}(\xi) = \frac{1}{1+i\xi} \qquad \forall\xi\in\mathbb{R}.$$

If $\xi \ge 1$, we have

$$|\hat{f}(\xi)| = \frac{1}{\sqrt{1+\xi^2}} = \frac{1}{\xi}\frac{1}{\sqrt{1+\frac{1}{\xi^2}}} \ge \frac{1}{\sqrt{2}\xi}.$$

As $\int_{[1,+\infty[}\frac{1}{\sqrt{2}\xi}d\xi = +\infty$, $\hat{f}$ is not summable.

This example shows that, given $f \in L^1(\mathbb{R}^n)$, $\hat{f} \in L^1(\mathbb{R}^n)$ does not necessarily hold.

Now we examine the interaction between the Fourier transform and derivatives. We start by precising the following convention: given $f \in L^1(\mathbb{R}^n)$, we shall say that f is continuous if there exists g continuous such that $g \sim f$. This means that the equivalence class $[f]$ contains a necessarily unique continuous function (remember Exercise 2.7.2 in the one-dimensional case). Analogous conventions will be supposed to hold when we write that f is of class C^1, etc..

Proposition 4.1.10. *Let $f \in L^1(\mathbb{R}^n)$ and of class C^1. Suppose that, for some $j \in \{1, ..., n\}$, $D_j f \in L^1(\mathbb{R}^n)$. Then $\forall \xi \in \mathbb{R}^n$*

$$\mathcal{F}(D_j f)(\xi) = i\xi_j \hat{f}(\xi).$$

Incomplete proof We prove the result in case $n = 1$, supposing also that $\lim_{x \to \pm\infty} f(x) = 0$. Integrating by parts, we have $\forall \xi \in \mathbb{R}$,

$$\hat{f'}(\xi) = \int_{\mathbb{R}} e^{-ix\xi} f'(x) dx = \lim_{n \to +\infty} \int_{-n}^{n} e^{-ix\xi} f'(x) dx =$$

$$= \lim_{n \to +\infty} (e^{-in\xi} f(n) - e^{in\xi} f(-n) + i\xi \int_{-n}^{n} e^{-ix\xi} f(x) dx) = i\xi \hat{f}(\xi).$$

□

Proposition 4.1.11. *Let $f \in L^1(\mathbb{R}^n)$, be such that, for some $j \in \{1, ..., n\}$, $x_j f \in L^1(\mathbb{R}^n)$. Then $\hat{f}$ has the partial derivative $D_j \hat{f}(\xi)$ in every point $\xi \in \mathbb{R}^n$; moreover, $\forall \xi \in \mathbb{R}^n$,*

$$D_j \hat{f}(\xi) = -i\mathcal{F}(x_j f)(\xi).$$

Sketch of the proof Differentiating formally with respect to ξ_j, we have

$$D_j \hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i<x,\xi>} (-ix_j) f(x) dx.$$

□

Now we introduce some extremely useful notations.

We shall call **multiindex** any element of $\mathbb{N}_0^n$, with $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$. If $\alpha := (\alpha_1, ..., \alpha_n)$ and $\beta := (\beta_1, ..., \beta_n)$, we set

$$\alpha + \beta := (\alpha_1 + \beta_1, ..., \alpha_n + \beta_n). \tag{4.1.10}$$

We shall write

$$\alpha \leq \beta \tag{4.1.11}$$

to indicate that $\alpha_j \leq \beta_j$ $\forall j \in \{1, ..., n\}$ and

$$|\alpha| := \alpha_1 + ... + \alpha_n. \tag{4.1.12}$$

We shall call $|\alpha|$ the **weight** of the multiindex α.

If $x = (x_1, ..., x_n) \in \mathbb{C}^n$ and $\alpha = (\alpha_1, ..., \alpha_n)$ is a multiindex, we set

$$x^\alpha := x_1^{\alpha_1} ... x_n^{\alpha_n}, \tag{4.1.13}$$

with the convention that $0^0 := 1$. For example, if $x = (0, 2i)$ and $\alpha = (0, 3)$, we have

$$x^\alpha = (2i)^3 = -8i.$$

Finally, if A is an open subset of $\mathbb{R}^n$, $f \in C^m(A)$, for some $m \in \mathbb{N}$ and $|\alpha| \le m$, we set

$$D^\alpha f := D_1^{\alpha_1}...D_n^{\alpha_n} f, \tag{4.1.14}$$

where D_j indicates the partial derivative with respect to the variable x_j. We recall that, as a consequence of the theorem of Schwarz, the derivatives with respect to the single variables in (4.1.14) can be mixed without changing the final result. For example, for $n = 3$, if $\alpha = (1, 2, 0)$, we have

$$D^\alpha f = D_1 D_2^2 f.$$

Observe that, if β is another multiindex such that$|\alpha + \beta| \le m$, the formula

$$D^\alpha (D^\beta f) = D^{\alpha+\beta} f \tag{4.1.15}$$

holds. Now we examine generalizations of Propositions 4.1.10 and 4.1.11.

Corollary 4.1.12. (I) *Let $f \in L^1(\mathbb{R}^n)$ and of class C^m $(m \in \mathbb{N})$. Suppose that, for every multiindex α, with $|\alpha| \le m$, $D^\alpha f \in L^1(\mathbb{R}^n)$. Then $\forall \xi \in \mathbb{R}^n$*

$$\mathcal{F}(D^\alpha f)(\xi) = (i\xi)^\alpha \hat{f}(\xi).$$

(II) *Let $f \in L^1(\mathbb{R}^n)$ be such that, for some $m \in \mathbb{N}$, $\|x\|^m f \in L^1(\mathbb{R}^n)$, with $\|.\|$ Euclidean norm. Then $\hat{f} \in C^m(\mathbb{R}^n)$; moreover, for every multiindex α, such that $|\alpha| \le m$, one has $(-ix)^\alpha f \in L^1(\mathbb{R}^n)$ and $\forall \xi \in \mathbb{R}^n$,*

$$D^\alpha \hat{f}(\xi) = \mathcal{F}((-ix)^\alpha f)(\xi).$$

Proof The result follows applying several times Propositions 4.1.10 e 4.1.11.

Concerning (II), observe that, if $|\alpha| \le m$, then $(-ix)^\alpha f \in L^1(\mathbb{R}^n)$. In fact, $\forall x \in \mathbb{R}^n$, $\forall j \in \{1, ..., n\}$, $|x_j| \le \|x\|$, hence $|(-ix)^\alpha f(x)| \le \|x\|^{|\alpha|}|f(x)|$. If $\|x\| \le 1$, one has $\|x\|^{|\alpha|}|f(x)| \le |f(x)|$. If $\|x\| \ge 1$, $\|x\|^{|\alpha|}|f(x)| \le \|x\|^m |f(x)|$ holds. So, in any case

$$|(-ix)^\alpha f(x)| \le (1 + \|x\|^m)|f(x)|.$$

$\square$

Remark 4.1.13. In force of Riemann-Lebesgue's theorem, with the assumptions of Corollary 4.1.12(II), if $|\alpha| \le m$, one has that $D^\alpha \hat{f} \in BC(\mathbb{R}^n)$ and satisfies the condition (4.1.3).

We conclude this section with the fundamental **inversion formula**.

Theorem 4.1.14. *Let $f \in L^1(\mathbb{R}^n)$ be such that $\hat{f} \in L^1(\mathbb{R}^n)$. Then*

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi>} \hat{f}(\xi) d\xi = \frac{1}{(2\pi)^n} \hat{\hat{f}}(-x) \qquad a.e. \quad in \quad \mathbb{R}^n. \tag{4.1.16}$$

Incomplete proof We show the theorem with the further assumption that $f \in BC(\mathbb{R}^n)$.

From the dominated convergence theorem, we have

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi>} \hat{f}(\xi) d\xi =$$

$$\lim_{k \to +\infty} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi> - \frac{\|\xi\|^2}{k}} \hat{f}(\xi) d\xi =$$

$$= \lim_{k\to+\infty} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi>-\frac{\|\xi\|^2}{k}} (\int_{\mathbb{R}^n} e^{-i<y,\xi>} f(y)dy)d\xi.$$

As

$$|e^{i<x,\xi>-\frac{\|\xi\|^2}{k}} e^{-i<y,\xi>} f(y)| = e^{-\frac{\|\xi\|^2}{k}} |f(y)|$$

and, applying the theorem of Tonelli,

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-\frac{\|\xi\|^2}{k}} |f(y)|d\xi dy = (\int_{\mathbb{R}} e^{-\frac{t^2}{2}} dt)^n \int_{\mathbb{R}^n} |f(y)|dy < +\infty,$$

it follows from the theorem of Fubini

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi>-\frac{\|\xi\|^2}{k}} (\int_{\mathbb{R}^n} e^{-i<y,\xi>} f(y)dy)d\xi =$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (\int_{\mathbb{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}} d\xi) f(y)dy$$

Again by the theorem of Fubini, we have

$$\int_{\mathbb{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}} d\xi = \prod_{j=1}^n \int_{\mathbb{R}} e^{i(x_j-y_j)\xi_j-\frac{\xi_j^2}{k}} d\xi_j =$$

$$= k^{\frac{n}{2}} \prod_{j=1}^n \int_{\mathbb{R}} e^{i(x_j-y_j)\sqrt{k}t-t^2} dt = (k\pi)^{\frac{n}{2}} e^{-\frac{k\|x-y\|^2}{4}},$$

applying the result of Example 3.5.15. So

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (\int_{\mathbb{R}^n} e^{i<x-y,\xi>-\frac{\|\xi\|^2}{k}} d\xi) f(y)dy =$$

$$= (\frac{k}{4\pi})^{\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\frac{k\|x-y\|^2}{4}} f(y)dy = \pi^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\|z\|^2} f(x - \frac{2z}{\sqrt{k}})dz.$$

Now, we pass to the limit, letting k go to $+\infty$. As f is continuous, we have that, $\forall z \in \mathbb{R}^n$,

$$\lim_{k\to+\infty} e^{-\|z\|^2} f(x - \frac{2z}{\sqrt{k}}) = e^{-\|z\|^2} f(x).$$

Moreover, for every $k \in \mathbb{N}$ and whatever is the choice of x and z in $\mathbb{R}^n$, we have

$$|e^{-\|z\|^2} f(x - \frac{2z}{\sqrt{k}})| \leq g(z) := \sup_{\mathbb{R}^n} |f| e^{-\|z\|^2}.$$

As g is summable, we can (applying the dominated convergence theorem), pass to the limit in the integral,, to conclude that

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i<x,\xi>} \hat{f}(\xi)d\xi = \pi^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-\|z\|^2} dz f(x) =$$

$$= f(x).$$

With this the proof is complete. □

Corollary 4.1.15. *The mapping $\mathcal{F}$, defined in (4.1.5), is injective.*

Proof As $\mathcal{F}$ is linear, it suffices to show that its kernel has only the element 0 of the linear space $L^1(\mathbb{R}^n)$. This follows immediately from the inversion formula (4.1.16).
□

Remark 4.1.16. Suppose that the assumptions of Theorem 4.1.14 are fulfilled. Then

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i<x,\xi>}\hat{f}(\xi)d\xi = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{-i<x,\xi>}\hat{f}(-\xi)d\xi,$$

which means that $f = \hat{Gf}$, with $Gf(\xi) = \frac{1}{(2\pi)^n}\hat{f}(-\xi)$.

The following result will be useful in the sequel.

Theorem 4.1.17. *Let f and g be elements of $L^1(\mathbb{R}^n)$. Then*

$$\int_{\mathbb{R}^n}\hat{f}(\xi)g(\xi)d\xi = \int_{\mathbb{R}^n} f(x)\hat{g}(x)dx. \tag{4.1.17}$$

Proof First of all, the integrals in (4.1.17) are defined. In fact, for example, $\hat{f}$ is bounded. So, $\forall\xi \in \mathbb{R}^n$,

$$|\hat{f}(\xi)g(\xi)| \leq \|\hat{f}\|_\infty |g(\xi)|.$$

Moreover, as $(x,\xi) \to e^{-i<x,\xi>}f(x)g(\xi)$ is summable in $\mathbb{R}^n \times \mathbb{R}^n$, by the theorem of Fubini, we have

$$\int_{\mathbb{R}^n}\hat{f}(\xi)g(\xi)d\xi =$$

$$\int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} e^{-i<x,\xi>}f(x)dx)g(\xi)d\xi = \int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-i<x,\xi>}f(x)g(\xi)dxd\xi =$$

$$= \int_{\mathbb{R}^n} f(x)(\int_{\mathbb{R}^n} e^{-i<x,\xi>}g(\xi)d\xi)dx$$

$$= \int_{\mathbb{R}^n} f(x)\hat{g}(x)dx.$$

□

Exercise 4.1.18. Check that the normed space $BC(A)$ of Definition 4.1.3 is a Banach space.

Exercise 4.1.19. Check that $\mathcal{F}$ is linear from $L^1(\mathbb{R}^n)$ to $BC(\mathbb{R}^n)$.

Exercise 4.1.20. Let X be Y normed spaces on the same field. with norms (respectively) $\|.\|_X$ and $\|.\|_Y$. Let $T : X \to Y$ be linear and such that there exists $L \geq 0$ so that

$$\|Tx\|_Y \leq L\|x\|_X \qquad \forall x \in X. \tag{4.1.18}$$

Show that T is continuous.

Exercise 4.1.21. Let X and Y be normed spaces on the same field, with norms (respectively) $\|.\|_X$ and $\|.\|_Y$. Let $T : X \to Y$ be linear. Show that T is continuous if and only if it is continuous in 0.

Exercise 4.1.22. Let $f \in L^1(\mathbb{R}^n)$. Prove what follows, applying the change of variable theorem:

(I) let $A : \mathbb{R}^n \to \mathbb{R}^n$ be linear and invertible; then $f \circ A \in L^1(\mathbb{R}^n)$ and, $\forall \xi \in \mathbb{R}^n$,

$$\mathcal{F}(f \circ A)(\xi) = \frac{\hat{f}((A^{-1})^T \xi)}{|det(A)|}$$

holds, where we have indicated with B^T the transpose of the matrix B;

(II) if $f(-x) = -f(x)$ a. e. in $\mathbb{R}^n$, then $\hat{f}(-\xi) = -\hat{f}(\xi)$ $\forall \xi \in \mathbb{R}^n$;

(III) if $f(-x) = f(x)$ a. e. in $\mathbb{R}^n$, then $\hat{f}(-\xi) = \hat{f}(\xi)$ $\forall \xi \in \mathbb{R}^n$;

(IV) if $x^0 \in \mathbb{R}^n$ and $g(x) = f(x - x^0)$ a. e. in $\mathbb{R}^n$, then $g \in L^1(\mathbb{R}^n)$ and $\hat{g}(\xi) = e^{-i<x^0,\xi>}\hat{f}(\xi)$ $\forall \xi \in \mathbb{R}^n$.

Exercise 4.1.23. Compute the Fourier transforms of the following functions:

(I) $f(x) = \frac{1}{x^2+2x+2}$;

(II) $f(x) = \frac{1}{(x^2+1)^2}$;

(III) $f(x) = \frac{1}{x^4+1}$;

(IV) $f(x) = \frac{x}{(x^2+1)^2}$;

(V) $f(x,y) = \frac{1}{(x^2+1)(y^2+1)}$.

Exercise 4.1.24. Let $f \in L^1(\mathbb{R}^n)$ and $g(x) = \overline{f(x)}$ a. e. Check that, $\forall \xi \in \mathbb{R}^n$,

$$\hat{g}(\xi) = \overline{\hat{f}(-\xi)}.$$

4.2 The class $\mathcal{D}(\Omega)$

Now we digress a little, in order to introduce a class of functions which will play an auxiliary, but rather important role.

So we start by specifying that, if Ω is an open subset of $\mathbb{R}^n$, we shall indicate with $C^\infty(\Omega)$ the class of complex valued functions with domain Ω, which are equipped of all the derivatives of any order and such derivatives are continuous in Ω.

Definition 4.2.1. *Let Ω be an open subset of $\mathbb{R}^n$. Let $f \in C^\infty(\Omega)$. We shall write that $f \in \mathcal{D}(\Omega)$ if there exists $K \subseteq \Omega$, closed and bounded, such that $f(x) = 0$ if $x \in \Omega \setminus K$.*

It is easy to check that, if $f \in \mathcal{D}(\Omega)$, all its derivatives $D^\alpha f \in \mathcal{D}(\Omega)$, for any multiindex α. Moreover, it is easy to verify that, if f and g are elements of $\mathcal{D}(\Omega)$, even $f + g \in \mathcal{D}(\Omega)$. This follows from the fact that, if K and L are closed and bounded subsets of Ω, if f vanished outside K and g vanishes outside L, then $f + g$ vanishes outside $K \cup L$. It is possible to show that $K \cup L$ is closed and bounded. Finally, it is clear that, if $f \in \mathcal{D}(\Omega)$ and $\alpha \in \mathbb{C}$, then $\alpha f \in \mathcal{D}(\Omega)$. So $\mathcal{D}(\Omega)$ is a linear space of functions, with the usual operations of sum and product by a scalar.

The construction of explicit nontrivial elements of $\mathcal{D}(\Omega)$ is not obvious. To this aim, we begin with the following

Lemma 4.2.2. *Let $g : \mathbb{R} \to \mathbb{R}$ be defined as follows:*

$$g(x) = \begin{cases} e^{-\frac{1}{x}} & if \quad x > 0, \\ 0 & se \quad x \leq 0. \end{cases} \tag{4.2.1}$$

Then $g \in C^\infty(\mathbb{R})$.

Incomplete proof It is immediately seen that g is continuous, as $\lim_{x\to 0^+} e^{-\frac{1}{x}} = \lim_{y\to+\infty} e^{-y} = 0$. Moreover, we have $g'(x) = 0$ if $x < 0$ and $D_-g(0) = 0$. If $x > 0$, $g'(x) = e^{-\frac{1}{x}}x^{-2}$ holds, while

$$D_+g(0) = \lim_{x\to 0^+} \frac{g(x)}{x} = \lim_{x\to 0^+} e^{-\frac{1}{x}}x^{-1} = \lim_{y\to+\infty} e^{-y}y = 0.$$

So g is differentiable in every point of $\mathbb{R}$. g is of class C^1, as

$$\lim_{x\to 0^+} g'(x) = \lim_{x\to 0^+} e^{-\frac{1}{x}}x^{-2} = 0.$$

The argument can be iterated, using the fact that, for every $n \in \mathbb{N}$, if $x > 0$, $g^{(n)}(x)$ has the form $e^{-\frac{1}{x}}P_n(\frac{1}{x})$, with P_n polynomial function. □

Example 4.2.3. We construct a first example of element of $\mathcal{D}(\mathbb{R})$ which is not identically zero. Let $\epsilon > 0$. We set

$$\begin{cases} f_\epsilon : \mathbb{R} \to \mathbb{R}, \\ f_\epsilon(x) = g(\epsilon^2 - x^2), \end{cases} \tag{4.2.2}$$

with g defined in (4.2.1). f_ϵ is of class C^∞, as composition of functions of class C^∞, and $f_\epsilon(x) > 0$ if $x \in]-\epsilon, \epsilon[$, while $f_\epsilon(x) = 0$ if $|x| \geq \epsilon$. Referring to Definition 4.2.1, we can take $K = [-\epsilon, \epsilon]$. Clearly, f_ϵ is summable and
$\int_{\mathbb{R}} f_\epsilon(x)dx \in \mathbb{R}^+$. Setting

$$\phi_\epsilon(x) := \frac{f_\epsilon(x)}{\int_{\mathbb{R}} f_\epsilon(y)dy}, \tag{4.2.3}$$

we obtain an example of element of di $\mathcal{D}(\mathbb{R})$ which is zero outside $[-\epsilon, \epsilon]$ and with integral 1.

Example 4.2.4. Let a, b, c, d be real numbers, with $a < c < d < b$. Next, let $\epsilon > 0$ be such that $a < c - 2\epsilon$ and $d + 2\epsilon < b$. We consider the function $f :]a, b[\to \mathbb{R}$,

$$f(x) := \int_{c-\epsilon}^{d+\epsilon} \phi_\epsilon(y - x)dy,$$

with ϕ_ϵ defined in (4.2.3). One can check that it is possible to differentiate under the sign of integral and that $f \in C^\infty(]a, b[)$. Now we observe that, if $x \in [c, d]$, as ϕ_ϵ is zero outside $[-\epsilon, \epsilon]$, we have that $\phi(y - x)$ is zero outside $[c - \epsilon, d + \epsilon]$. So

$$f(x) = \int_{c-\epsilon}^{d+\epsilon} \phi_\epsilon(y - x)dy = \int_{\mathbb{R}} \phi_\epsilon(y - x)dy = \int_{\mathbb{R}} \phi_\epsilon(y)dy = 1.$$

If, alternatively, $x < c - 2\epsilon$, $\phi_\epsilon(y - x) = 0$ $\forall y \in [c - \epsilon, d + \epsilon]$. So, in this case, $f(x) = 0$. Analogously, one can see that $f(x) = 0$ if $x > d + 2\epsilon$. Therefore, summing up, $f \in \mathcal{D}(]a, b[)$, as it is identically zero outside $[c - 2\epsilon, d + 2\epsilon]$ and identically 1 in $[c, d]$.

It is clear that $\mathcal{D}(\Omega) \subseteq L^1(\Omega) \cap L^2(\Omega)$, as, if $f \in \mathcal{D}(\Omega)$ and f is zero outside K closed and bounded,

$$\int_\Omega |f(x)|dx \leq \max_K |f| L_n(K) < +\infty$$

holds, because every bounded and measurable set has finite Lebesgue measure (see Exercise 4.2.7). Finally, it is useful to know that $\mathcal{D}(\Omega)$ is dense in $L^p(\Omega)$, for $p \in \{1, 2\}$:

Theorem 4.2.5. *Let Ω be an open subset of $\mathbb{R}^n$. Then $\mathcal{D}(\Omega)$ is dense in both the spaces $L^1(\Omega)$ and $L^2(\Omega)$.*

Exercise 4.2.6. Verify that, if A and B are bounded and closed subset of $\mathbb{R}^n$, $A \cup B$ is closed and bounded (Hint: show that $\partial(A \cup B) \subseteq \partial A \cup \partial B$).

Exercise 4.2.7. Verify that every bounded and measurable subset A of $\mathbb{R}^n$ has finite measure. (Hint.: verify preliminarily that there exists $L > 0$ such that $A \subseteq [-L, L]^n$)

Exercise 4.2.8. Let, for $n \in \mathbb{N}$ and $j \in \{1, ..., n\}$, $a_j < c_j < d_j < b_j$. Construct $f \in \mathcal{D}(\prod_{j=1}^n]a_j, b_j[)$ identically equal to 1 in $\prod_{j=1}^n [c_j, d_j]$.

4.3 The Fourier transform in $L^2(\mathbb{R}^n)$

We have seen in Section 4.1 that the Fourier transform in the space $L^1(\mathbb{R}^n)$ can be directly defined by integration. Nevertheless, in this space it does not enjoy optimal properties. For example, we have seen that it does not map $L^1(\mathbb{R}^n)$ into itself, (see Example 4.1.9), while the inversion formula (Theorem 4.1.14) holds under somewhat complicated assumptions. Now we want to extend the Fourier transform to $L^2(\mathbb{R}^n)$. If, on one side, the definition is less direct, on the other side we shall see that in this space it enjoys much better properties.

We begin with the following

Lemma 4.3.1. *Let $f, g \in \mathcal{D}(\mathbb{R}^n)$. Then:*
(I) $\hat{f} \in L^1(\mathbb{R}^n)$;
(II)

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}dx; \tag{4.3.1}$$

(Parseval identity);
(III) $\hat{f} \in L^2(\mathbb{R}^n)$ and

$$\|\hat{f}\|_2 = (2\pi)^{\frac{n}{2}} \|f\|_2, \tag{4.3.2}$$

where we indicate with $\|.\|_2$ the norm in $L^2(\mathbb{R}^n)$.

Proof We set

$$\Delta f := D_1^2 f + ... + D_n^2 f.$$

(we recall that Δ is the Laplace operator, see Section 3.6). For every $m \in \mathbb{N}$ $\Delta^m f \in \mathcal{D}(\mathbb{R}^n)$. By Corollary 4.1.12 (I), we have, if $\xi \in \mathbb{R}^n$,

$$\mathcal{F}(\Delta^m f)(\xi) = (-1)^m \|\xi\|^{2m} \hat{f}(\xi).$$

This implies that $\forall m \in \mathbb{N}$ there exists $C(m) \geq 0$ such that

$$|\hat{f}(\xi)| \leq C(m)\|\xi\|^{-m}$$

$\forall \xi \in \mathbb{R}^n \setminus \{0\}$. From this it follows that $\hat{f}$ is summable in $\{\xi \in \mathbb{R}^n : \|\xi\| \geq 1\}$. As it is clearly summable in $\{\xi \in \mathbb{R}^n : \|\xi\| \leq 1\}$, we can conclude that $\hat{f}$ is summable and so the inversion formula (4.1.16) holds. So we have

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx =$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\hat{f}}(-x)\overline{g(x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\hat{f}}(x)\overline{g(-x)}dx.$$

We set $h(x) := \overline{g(-x)}$. For $\xi \in \mathbb{R}^n$, one has

$$\hat{h}(\xi) = \int_{\mathbb{R}^n} e^{-i<x,\xi>}\overline{g(-x)}dx =$$

$$= \overline{\int_{\mathbb{R}^n} e^{-i<x,\xi>}g(x)dx} = \overline{\hat{g}(\xi)}.$$

Therefore, it follows from Theorem 4.1.17 that

$$\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{\hat{f}}(x)\overline{g(-x)}dx = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi.$$

In conclusion, we have proved Parseval identity (4.3.1).

(III) follows immediately from (II), taking $g = f$.

□

Now we are able to prove the following

Lemma 4.3.2. *Let $f \in L^2(\mathbb{R}^n)$. Let us consider a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n)$ such that $\lim_{k\to\infty} f_k = f$ in $L^2(\mathbb{R}^n)$. Then the sequence $(\hat{f}_k)_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R}^n)$ and the limit does not depend on the chosen sequence (obviously if it is convergent to f).*

Proof By Theorem 4.2.5, there exists a sequence $(f_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n)$, such that $\lim_{k\to\infty} f_k = f$ in $L^2(\mathbb{R}^n)$. We consider the sequence $(\hat{f}_k)_{k\in\mathbb{N}}$. As $L^2(\mathbb{R}^n)$ is complete, in order to verify that it is convergent, it suffices to check that it is a Cauchy sequence. So, let $\epsilon > 0$; as $(f_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, there exists $n(\epsilon) \in \mathbb{N}$ such that, if j and k are integers larger than $n(\epsilon)$, $\|f_j - f_k\|_2 < \frac{\epsilon}{(2\pi)^{\frac{n}{2}}}$ holds. So it follows from Lemma 4.3.1 (III) that

$$\|\hat{f}_k - \hat{f}_j\|_2 = \|\mathcal{F}(f_k - f_j)\|_2 < \epsilon.$$

Now we verify that the limit of the sequence $(\hat{f}_k)_{k\in\mathbb{N}}$ does not depend on the choice of $(f_k)_{k\in\mathbb{N}}$. Let $(g_k)_{k\in\mathbb{N}}$ be another sequence in $\mathcal{D}(\mathbb{R}^n)$ such that $\lim_{k\to\infty} g_k = f$ in $L^2(\mathbb{R}^n)$. We indicate, for the moment, with l_1 and l_2 the limits of the sequences $(\hat{f}_k)_{k\in\mathbb{N}}$ and $(\hat{g}_k)_{k\in\mathbb{N}}$, respectively. From the continuity of the norm (Exercise 2.2.10) it follows that

$$\|l_1 - l_2\|_2 = \lim_{k\to\infty} \|\hat{f}_k - \hat{g}_k\|_2$$

$$= \lim_{k\to\infty} \|\mathcal{F}(f_k - g_k)\|_2 = (2\pi)^{\frac{n}{2}} \lim_{k\to\infty} \|f_k - g_k\|_2 = (2\pi)^{\frac{n}{2}} \|f - f\|_2 = 0.$$

So $l_1 = l_2$.

□

Now we are able to define the Fourier transform of $f \in L^2(\mathbb{R}^n)$:

Definition 4.3.3. *Let $f \in L^2(\mathbb{R}^n)$. We set*

$$\mathcal{F}f = \hat{f} := \lim_{k\to\infty} \mathcal{F}f_k,$$

where $(f_k)_{k\in\mathbb{N}}$ is an arbitrary sequence in $\mathcal{D}(\mathbb{R}^n)$, converging to f in $L^2(\mathbb{R}^n)$.

In the next statement, we list the main properties of the Fourier transform in $L^2(\mathbb{R}^n)$.

Theorem 4.3.4. (*I*) *The mapping* $\mathcal{F} : f \to \mathcal{F}f$ *in Definition 4.3.3 is linear from* $L^2(\mathbb{R}^n)$ *into itself;*

(II) for f *and* g *in* $L^2(\mathbb{R}^n)$, *Parseval identity (4.3.1) holds;*

(III) $\forall f \in L^2(\mathbb{R}^n)$ *the identity (4.3.2) holds;*

(IV) $\mathcal{F}$ *is continuous from* $L^2(\mathbb{R}^n)$ *into itself;*

(V) if $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, *the Fourier transform in the sense of* $L^2(\mathbb{R}^n)$ *coincides with the Fourier transform in the sense of* $L^1(\mathbb{R}^n)$*;*

(VI) $\mathcal{F}$ *is a bijection from* $L^2(\mathbb{R}^n)$ *to itself; the inverse mapping* $\mathcal{F}^{-1}$ *fulfills the formula*

$$\mathcal{F}^{-1}f(\xi) = \frac{1}{(2\pi)^n}\hat{f}(-\xi) \qquad q.d.. \tag{4.3.3}$$

□

Incomplete proof We leave to the reader the proof of (I) (remember the result in Exercise 2.3.19).

(II) follows easily from the result in Exercise 4.3.7. In fact, if f and g are elements of $L^2(\mathbb{R}^n)$, there exist sequences $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n)$, such that we have $\lim_{k\to+\infty} f_k = f$ and $\lim_{k\to+\infty} g_k = g$ in $L^2(\mathbb{R}^n)$. So

$$\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx =< f, g >=$$

$$= \lim_{k\to+\infty} < f_k, g_k >= \lim_{k\to+\infty} \int_{\mathbb{R}^n} f_k(x)\overline{g_k(x)}dx =$$

$$= \lim_{k\to+\infty} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}_k(\xi)\overline{\hat{g_k}(\xi)}d\xi =$$

$$= \lim_{k\to+\infty} \frac{1}{(2\pi)^n} < \hat{f}_k, \hat{g_k} >= \frac{1}{(2\pi)^n} < \hat{f}, \hat{g} >=$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi)\overline{\hat{g}(\xi)}d\xi.$$

(III) follows immediately from (II).

(IV) follows from (III) and the result of Exercise 4.1.20.

We do not prove (V).

(VI) The injectivity of $\mathcal{F}$ follows from (III): if $\mathcal{F}f = 0$, by (III) $f = 0$. In order to show the subjectivity, we introduce the operator

$$Gf(\xi) := (2\pi)^{-n}\mathcal{F}f(-\xi), \quad f \in L^2(\mathbb{R}^n).$$

As $\mathcal{F}$ maps $L^2(\mathbb{R}^n)$ into itself linearly and continuously, it is readily seen that the same happens for G (as the mapping $g \to g(-\cdot)$ maps $L^2(\mathbb{R}^n)$ into itself linearly and continuously). Moreover, as we already know that, if $f \in \mathcal{D}(\mathbb{R}^n)$, $\mathcal{F}f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, it follows from (V) and Remark 4.1.16 that

$$\mathcal{F}(Gf) = f, \quad \forall f \in \mathcal{D}(\mathbb{R}^n). \tag{4.3.4}$$

With the usual argument we can extend (4.3.4) to any element f in $L^2(\mathbb{R}^n)$ and get the subjectivity of $\mathcal{F}$. Moreover, as we already know that $\mathcal{F}$ is injective, we deduce that

$$G = \mathcal{F}^{-1}.$$

□

Remark 4.3.5. As $G = \mathcal{F}^{-1}$, we have also that

$$G\mathcal{F}f = f \quad \forall f \in L^2(\mathbb{R}^n),$$

so that

$$f = (2\pi)^{-n}\mathcal{F}^2 f(-\cdot), \quad \forall f \in L^2(\mathbb{R}^n).$$

Example 4.3.6. From Example 4.1.9, if $f(x) = \chi_+(x)e^{-x}$, we have $\hat{f}(\xi) = g(\xi) = \frac{1}{1+i\xi}$. It is easy to check that f and g belong to $L^2(\mathbb{R})$. So $g = \mathcal{F}f$. It follows from Remark 5.5.10 that

$$\mathcal{F}g = \mathcal{F}^2 f = 2\pi f(-\cdot)$$

or

$$\mathcal{F}g(x) = \begin{cases} 2\pi e^x & \text{if} \quad x \le 0, \\ 0 & \text{if} \quad x > 0. \end{cases}$$

Exercise 4.3.7. Let H be a Hilbert space with inner product $< .,. >$. Next, let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be sequences in H, such that $\lim_{n\to+\infty} x_n = x$, $\lim_{n\to+\infty} y_n = y$. Prove that

$$\lim_{n\to+\infty} < x_n, y_n > = < x, y > .$$

Exercise 4.3.8. Prove that $\mathcal{F}[g(-\cdot)] = (\mathcal{F}g)(-\cdot)\ \forall g \in L^2(\mathbb{R}^n)$.

4.4 Weak derivatives, convolution and Fourier transform

We begin with some results and preliminary definitions.

Definition 4.4.1. *Let Ω be an open subset of $\mathbb{R}^n$ and let $f : \Omega \to \mathbb{C}$ be measurable. We shall say that $[f] \in L^1_{loc}(\Omega)$ (or that $[f]$ is* **locally summable** *in Ω), if, for every $H \subseteq \Omega$, with H closed and bounded, $f_{|H}$ is summable in H.*

Remark 4.4.2. Concerning Definition 4.4.1, we shall, as usual, say that "f is locally summable", in alternative to "$[f]$ is locally summable". It is not difficult to check that this definition is invariant with respect to modifications of f in subsets of measure zero.

Remark 4.4.3. It is almost obvious that $L^1(\Omega) \subseteq L^1_{loc}(\Omega)$. In addition, it is not difficult to check that, if $f \in L^2(\Omega)$, then $f \in L^1_{loc}(\Omega)$. In fact, owing to the usual inequality (2.4.3), we have

$$|f(x)| = |f(x)| \cdot 1 \le \frac{1}{2}(|f(x)|^2 + 1) \qquad \forall x \in \Omega.$$

So, if $H \subseteq \Omega$, with H closed and bounded, one has

$$\int_H |f(x)|dx \le \frac{1}{2}\int_H (|f(x)|^2 + 1)dx$$

$$\leq \frac{1}{2}\left(\int_\Omega |f(x)|^2 dx + L_n(H)\right) < +\infty.$$

It is also easy to see that, if $f \in C(\Omega)$, then $[f] \in L^1_{loc}(\Omega)$ (here it is convenient to distinguish f from $[f]$). In fact, if $H \subseteq \Omega$, with H closed and bounded, we have

$$\int_H |f(x)|dx \leq \max_H |f| \cdot L_n(H) < +\infty.$$

Now we state (without proof) an extremely useful result.

Theorem 4.4.4. *Let Ω be an open subset of $\mathbb{R}^n$, let $[f]$ and $[g]$ be locally summable in Ω. Then the following facts are equivalent:*
(a) $[f] = [g]$ (that is, $f(x) = g(x)$ a.e.);
(b) $\forall v \in \mathcal{D}(\Omega)$ one has

$$\int_\Omega f(x)v(x)dx = \int_\Omega g(x)v(x)dx.$$

After these preparations, we are ready to introduce the definition of weak derivative.

Definition 4.4.5. *Let Ω be an open subset of $\mathbb{R}^n$, let $f \in L^1_{loc}(\Omega)$ and let α be a multiindex. We shall denominate $\alpha-$***weak derivative** *of f a locally summable function g such that $\forall v \in \mathcal{D}(\Omega)$*

$$\int_\Omega g(x)v(x)dx = (-1)^{|\alpha|}\int_\Omega f(x)D^\alpha v(x)dx \tag{4.4.1}$$

holds.

Remark 4.4.6. An immediate consequence of Theorem 4.4.4 is that, if f has a $\alpha-$weak derivative, this is uniquely determined a.e.. We shall indicate it with $\partial^\alpha f$.

Definition 4.4.5 is motivated by the following

Theorem 4.4.7. *Let Ω be an open subset of $\mathbb{R}^n$ and let $f \in C^m(\Omega)$. Then, if $\alpha \in \mathbb{N}_0^n$ and $|\alpha| \leq m$, f has a weak derivative $\partial^\alpha f$ and $\partial^\alpha f(x) = D^\alpha f(x)$ a. e.*

Incomplete proof We consider the case that Ω is a regular open subset of $\mathbb{R}^3$. We assume that $\partial\Omega = S_1 \cup \dots S_p$, with $S_1, \dots, S_p$, regular surfaces with corresponding normal unit outer vectors $\nu_1, \dots, \nu_p$. We consider first the case $m = 1$ and $\alpha = e^j$. Then, if $v \in \mathcal{D}(\Omega)$, we can extend it to an element of $C^\infty(\mathbb{R}^3)$, just putting $v(x) = 0$ if $x \notin \Omega$. From

$$D_j f \cdot v = D_j(fv) - f \cdot D_j v,$$

employing Gauss-Green's formula, we get

$$\int_\Omega D_j f(x)v(x)dx = \int_{\overline{\Omega}} D_j f(x)v(x)dx = \int_{\overline{\Omega}} D_j(fv)(x)dx - \int_\Omega f(x)D_j v(x)dx$$

$$= \sum_{i=1}^p \int_{S_i} f(x)v(x)\nu_i^j(x)d\sigma - \int_\Omega f(x)D_j v(x)dx$$

$$= -\int_\Omega f(x)D_j v(x)dx.$$

In general, we can iterate the foregoing argument: if $f \in C^m(\Omega)$, $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, with $|\alpha| \leq m$, and $v \in \mathcal{D}(\Omega)$, we have

$$\int_\Omega D^\alpha f(x)v(x)dx = \int_\Omega D_1(D_1^{\alpha_1-1}D_2^{\alpha_2}D_3^{\alpha_3}f)(x)v(x)dx$$

$$= -\int_\Omega (D_1^{\alpha_1-1}D_2^{\alpha_2}D_3^{\alpha_3}f)(x)D_1 v(x)dx = (-1)^{\alpha_1}\int_\Omega (D_2^{\alpha_2}D_3^{\alpha_3}f)(x)D_1^{\alpha_1}v(x)dx$$

$$= (-1)^{\alpha_1+\alpha_2+\alpha_3}\int_\Omega f(x)(D_1^{\alpha_1}D_2^{\alpha_2}D_3^{\alpha_3}v)(x)dx.$$

□

We illustrate an example of existence of the weak derivative, in a case where the function is not everywhere differentiable.

Example 4.4.8. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$. f is not differentiable in a classical sense in 0. However it has a weak (first order) derivative ∂f. Such derivative coincides almost everywhere with the function "sign" (sgn). In fact, let $v \in \mathcal{D}(\mathbb{R})$. Suppose that $v(x) = 0$ if $x \notin [c, d]$, for certain c and d in $\mathbb{R}$. We may assume $c < 0 < d$. Then

$$-\int_{\mathbb{R}} |x|v'(x)dx = -\int_c^d |x|v'(x)dx =$$

$$= \int_c^0 xv'(x)dx - \int_0^d xv'(x)dx$$

$$= [xv(x)]_{x=c}^{x=0} - \int_c^0 v(x)dx - [xv(x)]_{x=0}^{x=d} + \int_0^d v(x)dx =$$

$$= -\int_{[c,0[} v(x)dx + \int_{]0,d]} v(x)dx =$$

$$= \int_{\mathbb{R}} sgn(x)v(x)dx.$$

We examine the interaction between weak derivatives and Fourier transform. On account of Corollary 4.1.12(I), the following result is quite natural:

Theorem 4.4.9. *Let $f \in L^2(\mathbb{R}^n)$ and $\alpha \in \mathbb{N}_0^n$. Then, if f has the weak derivative $\partial^\alpha f$ and such derivative belongs to $L^2(\mathbb{R}^n)$, one has $\mathcal{F}(\partial^\alpha u) = (i\xi)^\alpha \mathcal{F} f$.*

On the other hand, suppose that $(i\xi)^\alpha \hat{f} \in L^2(\mathbb{R}^n)$. Then f has the weak derivative $\partial^\alpha f$. Moreover, $\partial^\alpha f \in L^2(\mathbb{R}^n)$ and $\partial^\alpha f = \mathcal{F}^{-1}((i\xi)^\alpha \hat{f})$.

Incomplete proof We prove only the second statement. We put $g := \mathcal{F}^{-1}((i\xi)^\alpha \hat{f})$. Next, let $v \in \mathcal{D}(\mathbb{R}^n)$. If we employ the identity of Parseval and the result of Exercise 4.1.24, we obtain

$$\int_{\mathbb{R}^n} g(x)v(x)dx = \int_{\mathbb{R}^n} g(x)\overline{\overline{v(x)}}dx =$$

$$= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} (i\xi)^\alpha \hat{f}(\xi)\hat{v}(-\xi)d\xi =$$

$$= (-1)^{|\alpha|} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi)(-i\xi)^\alpha \hat{v}(-\xi)d\xi =$$

$$= (-1)^{|\alpha|} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi)\mathcal{F}(D^\alpha v)(-\xi)d\xi,$$

using Corollary 4.1.12(I). Again employing the result of Exercise 4.1.24 and the identity of Parseval, we can see that the latest integral coincides with

$$(-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x)D^\alpha v(x)dx,$$

hence we get the conclusion. □

We pass to the notion of convolution.

Definition 4.4.10. *Let f and g be elements of $L^1_{loc}(\mathbb{R}^n)$. We shall say that f and g are* **convolvable** *if for almost every $x \in \mathbb{R}^n$ the function $y \to f(x-y)g(y)$ is summable in $\mathbb{R}^n$. In this case, we denominate* **convolution** *of f and g and indicate with the symbol $f * g$, the function*

$$\begin{cases} f * g : \mathbb{R}^n \to \mathbb{C}, \\ (f*g)(x) = \begin{cases} \int_{\mathbb{R}^n} f(x-y)g(y)dy \\ \text{if } y \to f(x-y)g(y) \text{ is summable }, \\ 0 \text{ otherwise.} \end{cases} \end{cases} \tag{4.4.2}$$

The following result holds:

Theorem 4.4.11. *(I) Let f and g be elements of $L^1(\mathbb{R}^n)$. Then they are convolvable and the convolution $f * g$ belongs to $L^1(\mathbb{R}^n)$. Moreover,*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1. \tag{4.4.3}$$

(II) Let $f \in L^1(\mathbb{R}^n)$ and let $g \in L^2(\mathbb{R}^n)$. Then they are convolvable and the convolution $f * g$ belongs to $L^2(\mathbb{R}^n)$. Moreover,

$$\|f * g\|_2 \leq \|f\|_1 \|g\|_2. \tag{4.4.4}$$

(III) In each of the cases (I) and (II), the formula

$$\mathcal{F}(f * g) = \mathcal{F}f\mathcal{F}g. \tag{4.4.5}$$

holds.

Incomplete proof Let f and g belong to $L^1(\mathbb{R}^n)$. Supposing that they are convolvable, and that the convolution is measurable, we have, applying the theorem of Tonelli,

$$\int_{\mathbb{R}^n} |(f*g)(x)|dx \leq \int_{\mathbb{R}^n} (\int_{\mathbb{R}^n} |f(x-y)||g(y)|dy)dx =$$

$$= \int_{\mathbb{R}^n} |g(y)|(\int_{\mathbb{R}^n} |f(x-y)|dx)dy = \int_{\mathbb{R}^n} |g(y)|(\int_{\mathbb{R}^n} |f(z)|dz)dy =$$

$$= \|f\|_1 \|g\|_1.$$

Moreover, for every $\xi \in \mathbb{R}^n$,

$$\mathcal{F}(f*g)(\xi) = \int_{\mathbb{R}^n} e^{-i<x,\xi>}(\int_{\mathbb{R}^n} f(x-y)g(y)dy)dx =$$

$$= \int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} e^{-i<x-y,\xi>} f(x-y)e^{-i<y,\xi>}g(y)dy)dx =$$

$$= \int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} e^{-i<x-y,\xi>} f(x-y)dx)e^{-i<y,\xi>}g(y)dy =$$

$$= \int_{\mathbb{R}^n}(\int_{\mathbb{R}^n} e^{-i<z,\xi>} f(z)dz)e^{-i<y,\xi>}g(y)dy =$$

$$\mathcal{F}f(\xi)\mathcal{F}g(\xi).$$

We omit the case $g \in L^2(\mathbb{R}^n)$. $\square$

Exercise 4.4.12. Let Ω be an open subset of $\mathbb{R}^n$ and let α and β be multiindexes. We suppose that there exist in $L^1_{loc}(\Omega)$ the weak derivatives $\partial^\alpha f$ and $\partial^\beta(\partial^\alpha f)$. Check that f has the weak derivative $\partial^{\alpha+\beta} f$ and

$$\partial^{\alpha+\beta} f = \partial^\beta(\partial^\alpha f).$$

holds.

Exercise 4.4.13. Prove that the convolution in $L^1(\mathbb{R}^n)$ enjoys the commutative and associative properties. These property allow, given f_1,...,f_m in $L^1(\mathbb{R}^n)$, to consider the convolution $f_1 * ... * f_m$, where the order of application of the single operations has no influence on the final result.

(Hint: employ the fact that the Fourier transform of the convolution is the product of the Fourier transforms, together with the injectivity of $\mathcal{F}$.)

4.5 Some applications of the Fourier transform to problems of differential equations

The Fourier transform is useful to study a lot of problems in the fields (of example) of partial differential equations and probability. Here we illustrate some simple applications to the first of these subjects.

Example 4.5.1. (**The Helmoltz equation in** $\mathbb{R}^n$) Let $\lambda \in \mathbb{C}$. We consider the problem

$$(\lambda - \Delta)u(x) = f(x), \qquad x \in \mathbb{R}^n, \tag{4.5.1}$$

where we indicate (as usual) with Δ the Laplace operator

$$\Delta := \sum_{j=1}^{n} D_j^2 = D_1^2 + ... + D_n^2. \tag{4.5.2}$$

We assume that $\lambda \in \mathbb{R}^+$. We introduce some notable classes of (classes of equivalence of) functions:

Definition 4.5.2. *Let $m \in \mathbb{N}$. We indicate with $H^m(\mathbb{R}^n)$ the set of elements f in $L^2(\mathbb{R}^n)$ equipped with weak derivatives $\partial^\alpha f$ belonging to $L^2(\mathbb{R}^n)$ for every multiindex α such that $|\alpha| \leq m$.*

These classes of functions are a particular case of the so called "Sobolev spaces", of great importance in modern analysis.

The following statement holds:

Theorem 4.5.3. *Let $\lambda \in \mathbb{R}^+$. Then, for every $f \in L^2(\mathbb{R}^n)$ problem (4.5.1) has a unique solution u in the class $H^2(\mathbb{R}^n)$. u solves the problem, in the sense that*

$$\lambda u - \sum_{j=1} \partial_j^2 u = f,$$

where we have indicated with $\partial_j^2 u$ the corresponding weak derivative of u.

Proof By Theorem 4.4.9, if a solution u exists in $H^2(\mathbb{R}^n)$, it must hold

$$(\lambda + \|\xi\|^2)\mathcal{F}(u) = \mathcal{F}(f),$$

hence

$$u = \mathcal{F}^{-1}((\lambda + \|\xi\|^2)^{-1}\mathcal{F}(f)). \tag{4.5.3}$$

We observe that u defined in (4.5.3) is an element of $L^2(\mathbb{R}^n)$. In fact,

$$\int_{\mathbb{R}^n} |((\lambda + \|\xi\|^2)^{-1}\mathcal{F}(f)(\xi)|^2 d\xi$$

$$\leq \lambda^{-2} \int_{\mathbb{R}^n} |\mathcal{F}(f)(\xi)|^2 d\xi = \frac{(2\pi)^n}{\lambda^2} \int_{\mathbb{R}^n} |f(x)|^2 dx < +\infty$$

by Theorem 4.3.4(III).

Moreover, $u \in H^2(\mathbb{R}^n)$. To see this, it suffices to apply Theorem 4.4.9. In fact, let $\alpha \in \mathbb{N}_0^n$, with $|\alpha| \leq 2$. Then

$$(i\xi)^\alpha \mathcal{F}u(\xi) = \frac{(i\xi)^\alpha}{\lambda + \|\xi\|^2}\mathcal{F}f(\xi).$$

The function $\xi \to \frac{(i\xi)^\alpha}{\lambda+\|\xi\|^2}\mathcal{F}f(\xi)$ belongs to $L^2(\mathbb{R}^n)$, as

$$|\frac{(i\xi)^\alpha}{\lambda + \|\xi\|^2}\mathcal{F}f(\xi)| \leq \frac{\|\xi\|^{|\alpha|}}{\lambda + \|\xi\|^2}|\mathcal{F}f(\xi)|. \tag{4.5.4}$$

The first factor in the second term of (4.5.4) is bounded as a function of ξ, because $|\alpha| \leq 2$. So

$$\frac{\|\xi\|^{|\alpha|}}{\lambda + \|\xi\|^2}|\mathcal{F}f(\xi)| \leq C|\mathcal{F}f(\xi)|,$$

for some C positive.

So, if $|\alpha| \leq 2$, u has the weak derivative $\partial^\alpha u$ in $L^2(\mathbb{R}^n)$; hence, it belongs to $H^2(\mathbb{R}^n)$.

Example 4.5.4. (**Cauchy problem for the heat (diffusion) equation**) We look for a function u of the variables (t, x), with $t \geq 0$, $x \in \mathbb{R}^n$, satisfying the following conditions:

$$\begin{cases} D_t u(t,x) = \Delta_x u(t,x), & t > 0, x \in \mathbb{R}^n, \\ u(0,x) = f(x), & x \in \mathbb{R}^n. \end{cases} \tag{4.5.5}$$

where we have indicated with Δ_x the Laplace operator with respect to the (space) variables $x = (x_1, ..., x_n)$.

We argue formally, setting

$$U(t,\xi) := \int_{\mathbb{R}^n} e^{-i<x,\xi>} u(t,x) dx,$$

for $t > 0$ and $\xi \in \mathbb{R}^n$. In essence, we apply (formally) the Fourier tranform with respect to x to $u(t, .)$ for every $t \geq 0$. Changing the order of application of the transform and the time derivative D_t, we obtain

$$\begin{cases} D_t U(t,\xi) = -\|\xi\|^2 U(t,\xi), & t > 0, \xi \in \mathbb{R}^n, \\ U(0,\xi) = \hat{f}(\xi), & \xi \in \mathbb{R}^n, \end{cases} \tag{4.5.6}$$

hence

$$U(t,\xi) = e^{-t\|\xi\|^2}\hat{f}(\xi), \qquad t \geq 0, \xi \in \mathbb{R}^n. \tag{4.5.7}$$

Proceeding formally, we recall that, owing to Theorem 4.4.11(III), the product of the Fourier transforms is the Fourier transform of the convolution. So we compute the inverse Fourier transform of $\xi \to e^{-t\|\xi\|^2}$, per $t > 0$. We have, if $x \in \mathbb{R}^n$,

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i<x,\xi>}e^{-t\|\xi\|^2}d\xi =$$

$$= \frac{1}{(2\pi)^n}\prod_{j=1}^{n}\int_{\mathbb{R}} e^{ix_j\xi_j}e^{-t\xi_j^2}d\xi_j = \frac{1}{(4\pi^2 t)^{\frac{n}{2}}}\prod_{j=1}^{n}\int_{\mathbb{R}} e^{i\frac{x_j\eta_j}{\sqrt{t}}}e^{-\eta_j^2}d\eta_j =$$

$$= \frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{\|x\|^2}{4t}},$$

applying the result of Example 3.5.15.

So we have obtained the following "formal" solution of problem (4.5.5):

$$u(t,x) = \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{\|x-y\|^2}{4t}}f(y)dy & \text{if} \quad t>0, \\ f(x) & \text{if} \quad t=0. \end{cases} \tag{4.5.8}$$

Now we state a precise result. We need the definition of "classical solution":

Definition 4.5.5. *Let $f \in C(\mathbb{R}^n)$. A* **classical solution** *of (4.5.5) is a function u with domain $[0,+\infty[\times\mathbb{R}^n$, continuous, equipped in $]0,+\infty[\times\mathbb{R}^n$ of the derivatives $D_t u$, and $D_x^\alpha u$ for $|\alpha| \le 2$, continuous in $]0,+\infty[\times\mathbb{R}^n$.*

The following fact holds:

Theorem 4.5.6. *Let $f \in BC(\mathbb{R}^n)$. Then problem (4.5.5) has a unique classical solution u such that $u(t,.)$ is bounded in $\mathbb{R}^n$ $\forall t \ge 0$. Such solution u can be represented in the form (4.5.8).*

We limit ourselves to remark that, if $f \in BC(\mathbb{R}^n)$, the function in (4.5.8) is well defined, as, for every $t>0$ and for every $x \in \mathbb{R}^n$ the function $y \to e^{-\frac{\|x-y\|^2}{4t}}f(y)$ is summable in $\mathbb{R}^n$.

Example 4.5.7. (**Cauchy problem for the wave equation in space dimension** 1) We look for a function u of the variables (t,x), with $t \ge 0$, $x \in \mathbb{R}$, satisfying the following conditions:

$$\begin{cases} D_t^2 u(t,x) = D_x^2 u(t,x), & t>0, x\in\mathbb{R}, \\ u(0,x) = f(x), & x \in \mathbb{R}, \\ D_t u(0,x) = g(x), & x \in \mathbb{R}. \end{cases} \tag{4.5.9}$$

We operate formally, putting, as usual,

$$U(t,\xi) := \int_{\mathbb{R}} e^{-ix\xi}u(t,x)dx,$$

for $t>0$ and $\xi \in \mathbb{R}$. Applying the Fourier transform with respect to x to $u(t,.)$ for every $t \ge 0$, reversing the order of application between the transform and the time derivatives, we obtain

$$\begin{cases} D_t^2 U(t,\xi) = -\xi^2 U(t,\xi), & t>0, \xi\in\mathbb{R}, \\ U(0,\xi) = \hat{f}(\xi), & \xi \in \mathbb{R}, \\ D_t U(0,\xi) = \hat{g}(\xi), & \xi \in \mathbb{R}, \end{cases} \tag{4.5.10}$$

hence, at least for $\xi \ne 0$,

$$U(t,\xi) = \cos(t\xi)\hat{f}(\xi) + \frac{\sin(t\xi)}{\xi}\hat{g}(\xi). \tag{4.5.11}$$

Now we want to apply the inverse Fourier transform. Always proceeding formally, we have

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{ix\xi}\cos(t\xi)\hat{f}(\xi)d\xi$$

$$= \frac{1}{2}(\frac{1}{2\pi}\int_{\mathbb{R}} e^{i(x+t)\xi}\hat{f}(\xi)d\xi + \frac{1}{2\pi}\int_{\mathbb{R}} e^{i(x-t)\xi}\hat{f}(\xi)d\xi) =$$

$$= \frac{f(x+t)+f(x-t)}{2}.$$

Concerning the second summand in (4.5.11), it is easy to check that, if $a > 0$ and χ_a is the characteristic function of $[-a, a]$, we have, for $\xi \neq 0$,

$$\hat{\chi_a}(\xi) = 2\frac{\sin(a\xi)}{\xi}.$$

So, at least formally, the inverse Fourier transform of $\frac{\sin(t\xi)}{\xi}\hat{g}(\xi)$ is

$$\frac{1}{2}(\chi_t * g)(x) = \frac{1}{2}\int_{x-t}^{x+t} g(y)dy.$$

In conclusion, we have obtained

$$u(t,x) = \frac{f(x+t)+f(x-t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} g(y)dy. \tag{4.5.12}$$

Formula (4.5.12) is known as **formula of D'Alembert** . We proceed as in the case of the heat equation.

Definition 4.5.8. *Let $f, g \in C(\mathbb{R}^n)$. A* **classical solution** *of (4.5.9) is a function u in $C([0,+\infty[\times\mathbb{R})$, equipped in $]0,+\infty[\times\mathbb{R}$ with the continuous derivatives $D_t u$,$D_t^2 u$ e $D_x^\alpha u$ for $|\alpha| \leq 2$, with $D_t u$ continuously extensible to $[0,+\infty[\times\mathbb{R}^n$.*

The following result holds:

Theorem 4.5.9. *Let $f \in C^2(\mathbb{R})$, $g \in C^1(\mathbb{R})$. Then problem (4.5.9) has a unique classical solution. u. Such solution can be represented in the form (4.5.12).*

Remark 4.5.10. Differently from the case of the heat equation (see Theorem 4.5.6), for the wave equation a result of existence and uniqueness holds without assumptions of boundedness of the data or the solution.

Remark 4.5.11. In the case of the heat equation, the information given by the initial condition has infinite speed of propagation: in fact, assume, for example, that the datum f in (4.5.5) is such that $f(x) > 0$ if $\|x\| < \delta$, $f(x) = 0$ if $\|x\| \geq \delta$, for some $\delta \in \mathbb{R}^+$. From (4.5.8) we deduce that, for every $t \in \mathbb{R}^+$, $u(t,x) > 0$ $\forall x \in \mathbb{R}^n$.

On the contrary,in the case of the wave equation the information given by the initial conditions has finite speed of propagation: assume, for example, that $f(x) \neq 0$ only if $|x| < \delta$, while (for simplicity) $g = 0$. Let $x > \delta$. Then, if $t \in \mathbb{R}^+$, $u(t,x) = \frac{f(x-t)}{2}$. So, in order that $u(t,x) \neq 0$, it is necessary that $x - t \in (-\delta, \delta)$, which implies $x - t < \delta$ or $t > x - \delta$.

Exercise 4.5.12. Write explicitly in integral form the solution of problem (4.5.1) in space dimension one. Observe that the solution is not unique, if we do not require the belonging to $L^2(\mathbb{R})$.

Exercise 4.5.13. Study the nonhomogeneous heat equation

$$\begin{cases} D_t u(t,x) = \Delta_x u(t,x) + f(t,x), & t > 0, x \in \mathbb{R}^n, \\ u(0,x) = u_0(x), & x \in \mathbb{R}^n. \end{cases} \tag{4.5.13}$$

Applying formally the Fourier transform with respect to the space variables x, deduce the following expression of the solution:

$$u(t,x) = \\ \frac{1}{(4\pi t)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{\|x-y\|^2}{4t}} u_0(y)dy + \int_0^t \frac{1}{(4\pi(t-s))^{\frac{n}{2}}} (\int_{\mathbb{R}^n} e^{-\frac{\|x-y\|^2}{4(t-s)}} f(s,y)dy)ds. \tag{4.5.14}$$

It would be possible to prove that, if $u_0 \in BC(\mathbb{R}^n)$, $f \in BC([0,+\infty[\times\mathbb{R}^n)$ and has the derivative $D_t f \in BC([0,+\infty[\times\mathbb{R}^n)$, then (4.5.14) is the unique classical solution to (4.5.13) (in the sense of Definition 4.5.5) such that for every $t \geq 0$ $u(t,.)$ is bounded in $\mathbb{R}^n$.

Exercise 4.5.14. Check that, if $f \in C^2(\mathbb{R})$ and $g \in C^1(\mathbb{R})$, then (4.5.12) is really a classical solution of (4.5.9).

Chapter 5

Elements of calculus of probability

5.1 Probability spaces

In the study of several phenomena in natural sciences, physics, sociology, medicine, etc., one often encounters situations in which it is practically impossible to predict (in deterministic way) how a certain process is going to evolve. However, in many of these situations, the researcher is able to evaluate what is the development which seems the most plausible to expect. In these cases, we have to deal with phenomena which are usually defined as random.

In order to construct mathematical models in this context, the method we are going to describe has proved to be convenient: one starts by introducing a certain set Ω, representing all possible outcomes which could occur, or all possible results of a certain experiment (here the world "experiment" is intended in a generalized sense). One associates with any element of a certain class of concrete situations a suitable subset of Ω. In such a way, it is determined a certain family $\mathcal{A}$ of subsets of Ω. Then, we assign to any element A of this family a real number $P(A)$ between 0 and 1, called the "probability of A", indicating to what extent we are expecting the concrete situation corresponding to A is going to occur. Of course, the different situations are related. This suggests to require that the probability measure P should satisfty some general properties. Moreover, the classical operations among sets in the class $\mathcal{A}$ have natural interpretations as corresponding concrete situations. For example, intersection in $\mathcal{A}$ represents the contemporary happening of these situations, while union stands for the fact that, at least, one of them occurs. Let us examine some examples.

Example 5.1.1. Let us suppose that we want to construct a model of the experiment formed by the launch of a dice with the faces numbered from one to six and perfectly balanced. The set of all possible results is

$$\Omega := \{1, 2, 3, 4, 5, 6\}.$$

Then, we can identify, for example, the concrete situation "the result is even" with the set $A := \{2, 4, 6\}$ and the situation "the result is less or equal to 4" with $B = \{1, 2, 3, 4\}$. $A \cup B = \{1, 2, 3, 4, 6\}$ corresponds to "the result is even or less or equal to 4", while $A \cap B = \{2, 4\}$ stands for "the result is even and less or equal to 4".

Example 5.1.2. We suppose of having at disposal a device (for example, an electronic one) and of considering the time when it stops working, if we switch it on at time $t = 0$. Then, we represent such time by a nonnegative real number, employing an appropriate unit of measure. We indicate with Ω the set of all possible results, that is, the interval $[0, +\infty[$. Then its subset $[1, 2]$ can be identified with the concrete situation that the device stops working in some instant between 1 and 2.

We have mentioned the fact that the main operations in sets have a natural interpretation as concrete situations. Then, if we indicate with $\mathcal{A}$ the class of subsets of Ω of which we want to define a probability, we shall require that $\mathcal{A}$ is closed with respect to such operations. The assumption that is usually made is that $\mathcal{A}$ is a $\sigma-$algebra. In the following, it will be convenient to put, given $A \subseteq \Omega$, with Ω "sample space",

$$A^c := \Omega \setminus A. \tag{5.1.1}$$

We shall call A^c **complement** of A in Ω.

Definition 5.1.3. *Let Ω be a set. We indicate with $\mathcal{P}(\Omega)$ the* **power set** *of Ω, that is , the set of subsets of Ω. Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$. We shall say that $\mathcal{A}$ is a $\sigma-$algebra if it fulfills the following conditions:*

(I) $\Omega \in \mathcal{A}$;

(II) if $A_n \in \mathcal{A}$ $\forall n \in \mathbb{N}$, then $\bigcup\limits_{n \in \mathbb{N}} A_n \in \mathcal{A}$;

(III) if $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$.

From the definition, it follows that a $\sigma-$algebra is closed with respect to the main operations in sets:

Theorem 5.1.4. *Let $\mathcal{A}$ be a $\sigma-$algebra in the set Ω. Then:*

(I) $\emptyset \in \mathcal{A}$;

(II) if $n \in \mathbb{N}$ and $A_i \in \mathcal{A}$ for $i = 1, ..., n$, $\bigcup\limits_{i=1}^{n} A_i \in \mathcal{A}$;

(III) if $\mathcal{I} = \mathbb{N}$, or $\mathcal{I} = \{1, ..., n\}$, for some $n \in \mathbb{N}$, and $A_i \in \mathcal{A}$ $\forall i \in \mathcal{I}$, then $\bigcap\limits_{i \in \mathcal{I}} A_i \in \mathcal{A}$;

(IV) if A and B are elements of $\mathcal{A}$, $A \setminus B \in \mathcal{A}$.

Incomplete proof (I) It suffices to observe that $\emptyset = \Omega^c$ and to employ (I) and (III) in Definition 5.1.3.

(II) follows from (II) of Definition 5.1.3 and from (I), observing that, if we put $A_i := \emptyset$ for each $i \in \mathbb{N}$ with $i > n$, we have

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i \in \mathbb{N}} A_i.$$

We omit the proof of (III) (see Exercise 5.1.11).

(IV) follows from the fact that

$$A \setminus B = A \cap B^c$$

and from (III).

□

Now we consider the number $P(A)$, which should express the probability of the concrete situation corresponding to $A \in \mathcal{A}$. P is clearly a function from $\mathcal{A}$ to $[0, 1]$. We shall require that P is a probability measure, in the following sense:

Definition 5.1.5. *Let Ω be a nonempty set and let $\mathcal{A}$ be a $\sigma-$algebra in Ω. A* **probability measure** *in $\mathcal{A}$ is a function $P : \mathcal{A} \to [0, 1]$, such that*

(I) $P(\Omega) = 1$;

(II) if $A_n \in \mathcal{A}$ $\forall n \in \mathbb{N}$ and the sets A_n are pairwise disjoint for different values of n, then

$$P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

Here also, we can draw from the definition further properties of probability measures:

Theorem 5.1.6. *Let Ω be a nonempty set, let $\mathcal{A}$ be a $\sigma-$algebra in Ω and P a probability measure in $\mathcal{A}$. Then:*

(I) $P(\emptyset) = 0$;

(II) if $n \in \mathbb{N}$, $A_i \in \mathcal{A}$ for $i = 1, ..., n$ and the sets A_i are pairwise disjoint,

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i);$$

(III) if A and B are elements of $\mathcal{A}$ and $A \subseteq B$, we have

$$P(B \setminus A) = P(B) - P(A).$$

Proof (I) Applying property (II) in Definition 5.1.5 with $A_n = \emptyset \ \forall n \in \mathbb{N}$, we obtain immediately that, from $P(\emptyset) > 0$, it follows $P(\emptyset) = +\infty$. This is incompatible with the definition of probability measure.

(II) follows from (I) and (II) in Definition 5.1.5, observing (as we have already done) that, if we put $A_i = \emptyset$ for $i > n$, we have $\bigcup_{i \in \mathbb{N}} A_i = \bigcup_{i=1}^{n} A_i$.

(III) One has $B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \emptyset$. It follows from (II) that

$$P(B) = P(A) + P(B \setminus A),$$

hence we get the conclusion.

□

Remark 5.1.7. From (III) of Theorem 5.1.6, it follows that, if $A \subseteq B$, then $P(A) \leq P(B)$.

Now we are able to introduce the definition of probability space:

Definition 5.1.8. *A* **probability space** *is a triple $(\Omega, \mathcal{A}, P)$, where Ω is a nonempty set, $\mathcal{A}$ is a $\sigma-$algebra in Ω, P is a probability measure, with domain $\mathcal{A}$.*

The elements of $\mathcal{A}$ will be called **events**.

We pass to some concrete examples.

Example 5.1.9. (Finite probability spaces) Let Ω be a set with a finite number ($n \in \mathbb{N}$) of elements. We assume that they are $\{\omega_1, ..., \omega_n\}$. We take the σ-algebra $P(\Omega)$ of all subsets of Ω. We want to define a probability measure in $P(\Omega)$. So we start by considering the singletons $\{\omega_1\}$, ...,$\{\omega_n\}$. If we assign to $\{\omega_j\}$ $(1 \leq j \leq n)$ the probability $p_j \in [0, 1]$, the property (II) of Theorem 5.1.6 imposes that

$$1 = P(\Omega) = p_1 + ... + p_n. \tag{5.1.2}$$

On the other hand, if the condition (5.1.2) is satisfied, it is not difficult to prove that there exists a unique probability measure in $P(\Omega)$ such that $P(\{\omega_j\}) = p_j$ per $j = 1, ..., n$. Evidently, it will hold, if $A \subseteq \Omega$,

$$P(A) = \sum_{\omega_j \in A} p_j. \tag{5.1.3}$$

So, we reexamine Example 5.1.1. We deal with the fact that we have to assign a probability p_j to the fact that the outcome of the launch is j, with $1 \leq j \leq 6$. The conditions which should be satisfied are $p_j \in [0,1]$ for $1 \le j \le n$, and $\sum_{j=1}^{6} p_j = 1$. Then, it is clear that, if we think that the dice is well balanced, in such a way that it guarantees the equiprobability of the single results, the only possibility is to put $p_j = \frac{1}{6}$ for each $j = 1, ..., 6$. On the contrary, if we think that this equiprobability does not hold, we shoud assign values p_j, possibly different from each other, but preserving the condition (5.1.2). In the case of a balanced dice, for every $A \subseteq \{1, ..., 6\}$ the condition (5.1.3) imposes that

$$P(A) = \frac{\sharp(A)}{6},$$

where $\sharp(A)$ indicates the cardinality (number of elements) of the set A.

Another example in the same order if ideas is the following: we consider the result of the launch of two balanced dice. We can take as sample space Ω the set of ordered pairs $\{(i,j) : 1 \le i, j \le 6\}$, which is made of 36 elements. We obtain a probability space, setting, for each $A \subseteq \Omega$,

$$P(A) := \frac{\sharp(A)}{36}.$$

For example, we consider the event "the total outcome is 2". Then we identify it with $A = \{(1,1)\}$, hence $P(A) = \frac{1}{36}$. On the other hand, if we indicate with B the event "the total outcome is 7", we have $B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, hence $P(B) = \frac{6}{36} = \frac{1}{6}$.

Example 5.1.10. Now we construct a probabilistic model related to Example 5.1.2. We set $\Omega := \mathbb{R}^+$ and indicate with $\mathcal{A}$ the class of subsets of Ω which are measurable in the sense of Lebesgue. From Theorem 1.1.5, it immediately follows that $\mathcal{A}$ is a $\sigma-$algebra. Concerning the probability measure, we fix $g : \mathbb{R}^+ \to [0, +\infty[$ measurable, such that

$$\int_{\mathbb{R}^+} g(t)dt = 1 \tag{5.1.4}$$

and set, given $A \in \mathcal{A}$,

$$P(A) := \int_A g(t)dt. \tag{5.1.5}$$

It is not difficult to check that P is a probability measure, Here we limit ourselves to prove property (II) in Definition 5.1.5. Let $(A_n)_{n\in\mathbb{N}}$ be a sequence of pairwise disjoint elements in $\mathcal{A}$. We indicate with χ_n the characteristic function of A_n. Evidently, if $A := \bigcup_{n\in\mathbb{N}} A_n$, we have that, for every $x \in \Omega$ the characteristic function χ_A of A satisfies

$$\chi_A(x) = \sum_{n=1}^{\infty} \chi_n(x).$$

So

$$P(A) = \int_A g(t)dt = \int_{\mathbb{R}^+} g(t)\chi_A(t)dt =$$

$$= \int_{\mathbb{R}^+} \sum_{n=1}^{\infty} g(t)\chi_n(t)dt =$$

(applying the result of Exercise 1.5.6)

$$= \sum_{n=1}^{\infty} \int_{\mathbb{R}^+} g(t)\chi_n(t)dt = \sum_{n=1}^{\infty} \int_{A_n} g(t)dt =$$

$$= \sum_{n=1}^{\infty} P(A_n).$$

For example, let $\lambda > 0$. We take $g(t) := \lambda e^{-\lambda t}$. With this choice, if A stands for the event "the device stops working in some instant between 1 e 2" (in the fixed time unit) , we shall identify it with the interval $]1,2[$. So

$$P(A) = \int_{]1,2[} \lambda e^{-\lambda t} dt = e^{-\lambda} - e^{-2\lambda}.$$

Exercise 5.1.11. Prove (III) in Theorem 5.1.4. The crucial point is the formula

$$\bigcap_{i \in \mathcal{I}} A_i = (\bigcup_{i \in \mathcal{I}} A_i^c)^c. \tag{5.1.6}$$

Exercise 5.1.12. Construct a probabilistic model of the double launch of a balanced dice. Calculate the probability of obtaining at least one six.

Exercise 5.1.13. Let A and B be events in the probability space $(\Omega,\, \mathcal{A},\, P)$. Check that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

5.2 Elements of combinatorics

In this section we introduce some simple elements of combinatorics, allowing to determine the cardinality of certain finite sets, in order to compute the probability of some events in case of finite probability spaces.

The main result is the following

Theorem 5.2.1. *Let N and K be finite sets, with (respectively) n and k elements. Here n and k are natural numbers and $k \leq n$. Then:*

(I) the number of injective mappings from K to N is $\frac{n!}{(n-k)!}$;

(II) the number of subsets of N with k elements is $\binom{n}{k}$, with

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}.$$

Proof Concerning (I), let $K = \{a_1, ..., a_k\}$. If we want to list all the injective mappings $f : K \to N$, we can choose $f(a_1)$ in n different ways, $f(a_2)$ in $n-1$ different ways ($f(a_2)$ may be any elemet distinct from $f(a_1)$), ..., $f(a_k)$ in $n-k+1$ different ways. So the number of injective mappings from K to N is $n(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$.

Concerning (II), let K' be a generic subset of N with k elements. The injective mappings from K to N the range of which is K' are evidently the bijections from K to K'. Applying (I), we can say that there are $k!$. So, multiplying $k!$ by the number of subsets of N with k elements, we obtain the cardinality of the set of injective mappings from K to N, that is $\frac{n!}{(n-k)!}$. Hence we get the conclusion.

□

Example 5.2.2. What is the probability of getting a winning triplet at lottery, with a single bet?

In each single lottery drum five numbers between one and ninety are weekly drawn. The number of subsets of five elements in a set with ninety elements is $\binom{90}{5}$. So the probability that a single family of five elements coincides with the family of drawn elements is $\frac{1}{\binom{90}{5}}$. Lets fix a certain triplet. The number of subsets of five elements containing this triplet is clearly $\binom{87}{2}$ (there remain to be fixed two elements, which can be chosen in a set of 87 elements). So the probability that the triplet is contained in the set of drawn numbers is

$$\frac{\binom{87}{2}}{\binom{90}{5}} = \frac{87!}{85!2} \frac{85!5!}{90!} \cong 8,5 \times 10^{-5}.$$

Example 5.2.3. A box contains five red balls and ten white balls. Five of them are drawn at random (each drawn ball is kept outside the box). What is the probability of drawing exactly three red balls?

Let Ω be the family of subsets with five elements in the set of $10 + 5 = 15$ balls. We have $\sharp(\Omega) = \binom{15}{5}$. The number of elemets of Ω with three red balls and two white balls is $\binom{5}{3}\binom{10}{2}$. So the probability we are trying to compute is

$$\frac{\binom{5}{3}\binom{10}{2}}{\binom{15}{5}} \cong 0,15.$$

We conclude this section with a well known formula:

Corollary 5.2.4. *(Newton's binomial theorem) Let a and b be complex numbers and $n \in \mathbb{N}$. Then*

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^{n-j} b^j.$$

Proof Expanding the product $(a+b) \cdot ... \cdot (a+b)$ (n factors) taking one of the summands a or b in each factor, we obtain a sum with terms of the form $c_j a^{n-j} b^j$, with $0 \leq j \leq n$. c_j stands for the number of possible choices, obtained taking $n - j$ times a and j times b. Now, we can associate with each choice where we have taken b j times the family of factors where we have taken b, forming a subset of j elements in a set of n elements. We conclude, applying Theorem 5.2.1(II), that $c_j = \binom{n}{j}$.
□

Exercise 5.2.5. Compute the probability that, taking n persons at random, $(2 \leq n \leq 365)$ at least two have the same birthday.

(Hint: neglect people who were born of February 29th in a leap year. Try to compute the probability of the complementing event. Quite surprisingly, one can see that already with only 23 persons the searched probability is larger than $\frac{1}{2}$!)

5.3 Conditional probability and independence

Definition 5.3.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, and let A and B be elements of $\mathcal{A}$, with $P(A) > 0$. We define the* **conditional probability** *$P(B|A)$ of B given A as follows:*

$$P(B|A) := \frac{P(B \cap A)}{P(A)}.$$

Remark 5.3.2. Intuitively, $P(B|A)$ is a measure of the probability that B occurs, in the case that A occurs. Observe that, if A and B are incompatible (that is, $A \cap B = \emptyset$), we have $P(B|A) = 0$. Moreover, $P(A|A) = 1$.

Example 5.3.3. The forty per cent of a certain population is made of smokers (S), the sixty per cent is made of nonsmokers (N). It is known that the twentyfive per cent of the smokers is affected by a certain chronic disease of the respiratory system; on the other hand, only the seven per cent of nonsmokers is affected by this disease. What is the probability that an individual, who is affected by the disease, is a smoker?

We have to compute $P(S|M) = \frac{P(S \cap M)}{P(M)}$. We know that $P(M|S) = \frac{1}{4}$. We have

$$P(M) = P(S \cap M) + P(N \cap M) =$$
$$= P(M|S)P(S) + P(M|N)P(N) = \frac{1}{4} \cdot \frac{2}{5} + \frac{7}{100} \cdot \frac{3}{5} = \frac{71}{500}.$$

Moreover,

$$P(S \cap M) = P(M|S)P(S) = \frac{1}{4} \cdot \frac{2}{5} = \frac{1}{10}.$$

It follows that

$$P(S|M) = \frac{\frac{1}{10}}{\frac{71}{500}} = \frac{50}{71} \cong 0,70.$$

Remark 5.3.4. Let A_1,..., A_n and B be events with positive probability in $(\Omega, \mathcal{A}, P)$. Suppose that A_1,...,A_n make a partition of Ω, in the sense that they are pairwise disjoint and their union is Ω. Let $j \in \{1, ..., n\}$. Observe that B is the disjoint union of the events $B \cap A_k$ $(1 \leq k \leq n)$. Then

$$\begin{gathered} P(A_j|B) = \\ = \frac{P(A_j \cap B)}{P(B)} = \frac{P(A_j \cap B)}{P(A_j)} \frac{P(A_j)}{P(B)} = P(B|A_j) \frac{P(A_j)}{\sum_{k=1}^{n} P(B \cap A_k)} = \\ = \frac{P(B|A_j)P(A_j)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}. \end{gathered} \tag{5.3.1}$$

The identity we have drawn in (5.3.1) is the so called **Bayes' formula**. Employing this formula, we can obtain more quickly the result in Example 5.3.3. In fact, we have

$$\begin{gathered} P(S|M) = \\ = \frac{P(M|S)P(S)}{P(M|S)P(S) + P(M|N)P(N)} = \\ = \frac{\frac{1}{4} \cdot \frac{2}{5}}{\frac{1}{4} \cdot \frac{2}{5} + \frac{7}{100} \frac{3}{5}} = \frac{50}{71}. \end{gathered}$$

Bayes' formula is known also as the formula of the probability of causes, because in applications the events A_j are potential causes of B and we want to evaluate which is the most probable one.

Definition 5.3.5. *Let A and B be events in the probability space $(\Omega, \mathcal{A}, P)$. We shall say that A and B are* **independent** *if $P(A \cap B) = P(A)P(B)$.*

Remark 5.3.6. If $P(A) > 0$, the independence between A and B is equivalent to

$$P(B|A) = P(B).$$

The sense of Definition 5.3.5 is the following: the probability of the occurrence of A does not change if B occurs. The inverse is also true, in case $P(B) > 0$.

Example 5.3.7. We consider a deck of 40 cards, with the usual four suits. We draw one card at random. Let A be the event "the drawn card is an ace", B the event "the drawn card is a heart". We check that A and B are independent.

One has $P(A) = \frac{4}{40} = \frac{1}{10}$, $P(B) = \frac{10}{40} = \frac{1}{4}$. $A \cap B$ is the event "the ace of hearts is drawn". So

$$P(A \cap B) = \frac{1}{40} = \frac{1}{10} \cdot \frac{1}{4} = P(A)P(B).$$

Therefore, A and B are independent.

Example 5.3.8. Consider the launch of three balanced dice. So

$$\Omega = \{(i, j, k) : i, j, k \in \{1, 2, 3, 4, 5, 6\}\}.$$

Any single ordered triplet has probability $\frac{1}{6^3} = \frac{1}{216}$. Let A be the event "the sum $i + j + k$ equals 6". A is identifiable with the set of triplets

$$\{(1, 1, 4), (1, 2, 3), (1, 3, 2), (1, 4, 1), (2, 1, 3), (2, 2, 2),$$
$$(2, 3, 1), (3, 1, 2), (3, 2, 1), (4, 1, 1)\}.$$

So

$$P(A) = \frac{10}{216} = \frac{5}{108} \cong 0,046.$$

Let B be the event "i,j,k are pairwise distinct". The cardinality of B coincides with the number of injective functions from a set of 3 elements to a set of 6 elements, and equals, on account of Theorem 5.2.1,

$$\frac{6!}{(6-3)!} = \frac{6!}{3!} = 120.$$

So

$$P(B) = \frac{120}{216} = \frac{5}{9} \cong 0,56.$$

We have

$$A \cap B = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}.$$

From this,

$$P(A \cap B) = \frac{6}{6^3} = \frac{1}{36} \cong 0,028.$$

On the other hand,

$$P(A)P(B) = \frac{5}{108} \cdot \frac{5}{9} = \frac{25}{972} \cong 0,026.$$

So, A and B are not independent. From $P(A)P(B) < P(A \cap B)$, we draw $P(B) < P(B|A)$. This means that the occurrence of A makes the occurrence of B more probable.

It is of interest to define the independence of families made of more than two events:

Definition 5.3.9. *Let $\{A_i : i \in \mathcal{I}\}$ be a family of events in the probability space $(\Omega, \mathcal{A}, P)$, depending on the parameter i in $\mathcal{I}$. We shall say that the events A_i are independent if, whatever is the choice of $i_1, ..., i_n$ ($n \in \mathbb{N}$) in $\mathcal{I}$, pairwise distinct, one has*

$$P(A_{i_1} \cap ... \cap A_{i_n}) = P(A_{i_1}) \cdot ... \cdot P(A_{i_n}).$$

Example 5.3.10. *(Bernoulli processes)* Now we want to describe a simple mathematical model, which is applicable in many concrete situations.

Suppose of repeating n times a certain experiment, always with the same conditions. We introduce the assumption that results in different tests have no influence to each other. We are interested only in two complementary aspects of the experiment, which we call "success" (S) and "failure" (F). Let p $(\in [0,1])$ the probability of S in each test. Consequently, the probability of F will be $q := 1-p$. So we set

$$\Omega := \{\omega = \{\omega_1, ..., \omega_n) : \omega_1, ..., \omega_n \in \{S, I\}\}. \tag{5.3.2}$$

We observe that Ω is finite and made of 2^n elements. Following the general scheme in Example 5.1.9 and so posing $\mathcal{A} = \mathcal{P}(\Omega)$, we have to assign a probability p_ω to each $n-$tuple $\omega^0 = (\omega_1^0, ..., \omega_n^0)$, in such a way that $\sum\limits_{\omega\in\Omega} p_\omega = 1$. For each $j = 1, ..., n$, we set

$$A_j := \{\omega \in \Omega : \omega_j = \omega_j^0\}. \tag{5.3.3}$$

It is natural to put

$$P(A_j) = \begin{cases} p & \text{se} \quad \omega_j^0 = S, \\ q & \text{se} \quad \omega_j^0 = I. \end{cases} \tag{5.3.4}$$

Observing that

$$\{\omega^0\} = A_1 \cap ... \cap A_n$$

and on account of the fact that different tests have no influence to each other, we assume that A_1,...,A_n are independent, so that

$$p_{\omega^0} := P(A_1) \cdot ... \cdot P(A_n) = p^m q^{n-m}, \tag{5.3.5}$$

where m stands for the number of successes in the sequence $\omega_1^0, ..., \omega_n^0$.

Now we check that the condition $\sum\limits_{\omega\in\Omega} p_\omega = 1$ holds. To this aim, we observe, firstly, that the $n-$tuples containing exactly m successes $(0 \le m \le n)$ can be associated with the subsets of m elements of a set with n elements (we associate with each sequence ω the set $\{i_1, ..., i_m\}$, with $1 \le i_1 < ... < i_m \le n$ such that $\omega_j = S$ if $j \in \{i_1, ..., i_m\}$). So, in force of Theorem 5.2.1 (II), the $n-$tuples containing exactly m successes are $\binom{n}{m}$. It follows, applying Newton's formula, that

$$\sum_{\omega\in\Omega} p_\omega = \sum_{m=0}^{n} \binom{n}{m} p^m q^{n-m} = (q+p)^n = 1^n = 1.$$

In conclusion, we have, for $A \subseteq \Omega$,

$$P(A) = \sum_{\omega\in A} p_\omega. \tag{5.3.6}$$

With position (5.3.6), we could also verify that (5.3.4) holds and that the sets A_j $(1 \le j \le n)$ defined in (5.3.3) are independent (see Exercises 5.3.13 and 5.3.14).

Remark 5.3.11. We consider again the scheme described in Example 5.3.10. For $0 \le m \le n$, we indicate with B_m the event "we have exactly m successes". From what we have seen, we have

$$P(B_m) = \binom{n}{m} p^m q^{n-m}. \tag{5.3.7}$$

We wonder: which is the most probable number of successes, or, in different words, for which m $P(B_m)$ is maximum ? To answer this question, we consider the inequality

$$P(B_m) \leq P(B_{m+1}), \quad 0 \leq m \leq n-1. \tag{5.3.8}$$

Applying formula (5.3.7), it is easy to check that (5.3.8) is equivalent to

$$m \leq np - q, \tag{5.3.9}$$

and the strict inequality holds if $m < np - q$. We observe preliminarly that, as $0 \leq p \leq 1$ and $q = 1-p$, $np - q \leq n$ and $np - q = n$ if and only if $p = 1$ and $q = 0$. Moreover:

(I) if $np - q < 0$, we have $P(B_0) > P(B_1) > ... > P(B_n)$; the most probable result is $m = 0$;

(II) suppose that $0 \leq np - q < n$ and $np - q \notin \mathbb{Z}$; we indicate with m_0 the integer part of $np - q$ $(0 \leq m_0 < n)$; then $P(B_0) < ... < P(B_{m_0}) < P(B_{m_0+1}) > ... > P(B_n)$; in this case, the most probable result is $m_0 + 1$;

(III) suppose that $0 \leq np - q < n$ and $np - q = m_0 \in \mathbb{Z}$; then $P(B_0) < ... < P(B_{m_0}) = P(B_{m_0+1}) > ... > P(B_n)$; in this case, the most probable results are (with the same probability) m_0 and $m_0 + 1$;

(IV) suppose that $np - q = n$; as already observed, this is equivalent to $p = 1$ and $q = 0$; in this case, $P(B_0) = ... = P(B_{n-1}) = 0$, while $P(B_n) = 1$; evidently, the most probable result is $m = n$.

Example 5.3.12. We introduce a first example of Bernoulli process. Suppose of throwing 50 times a perfectly balanced coin. We consider "success" the result "head" (H), "failure" the result "tail" (T). Then we have $n = 50$, $p = q = \frac{1}{2}$. If $0 \leq m \leq 50$, the probability of obtaining exactly m heads (in 50 launches) is $\binom{50}{m}2^{-50}$. In this case,

$$np - q = 50 \cdot \frac{1}{2} - \frac{1}{2} = \frac{49}{2} = 24,5.$$

Following the arguments in Remark 5.3.11, the most probable number of heads is 25. The probability of obtaining exactly 25 heads is

$$\binom{50}{25} 2^{-50} \cong 0,11.$$

Exercise 5.3.13. Prove that (5.3.5) and (5.3.6) imply (5.3.4).

(Hint: suppose that, for example, $\omega_j^0 = S$. For each $i = 1, ..., n$, let us indicate with C_i the subset of elements ω such that $\omega_j = S$ and $\omega_k = S$ for i elements k in the sequence $1, ..., n$. If $\omega \in C_i$, one has $p_\omega = p^i q^{n-i}$. One can verify that C_i has $\binom{n-1}{i-1}$ elements. Consequently

$$\begin{array}{rclcl} P(A_j) & = & \sum_{i=1}^{n} \binom{n-1}{i-1} p^i q^{n-i} & = & \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} q^{n-1-k} \\ & = & p(q+p)^{n-1} & = & p. \end{array}$$

Exercise 5.3.14. Prove that the events $A_1, ..., A_n$ defined in (5.3.3) are independent.

(Hint: we have to show that, if $1 \leq j_1 < ... < j_r \leq n$ $(2 \leq r \leq n)$, one has

$$P(A_{j_1} \cap ... \cap A_{j_r}) = P(A_{j_1}) \cdot ... \cdot P(A_{j_r}) = p^s q^{r-s},$$

where s indicates the cardinality of $\{j \in \{j_1, ..., j_r\} : \omega_j^0 = S\}$. For each $i = s, ..., n-(r-s) = n-r+s$, we indicate with C_i the subset of the elements $\omega \in \Omega$ such that $\omega_j = \omega_j^0$ if $j \in \{j_1, ..., j_r\}$

and $\omega_j = S$ for i elements j of the sequence $1, ..., n$. If $\omega \in C_i$, one has $p_\omega = p^i q^{n-i}$. One can verify that C_i has $\binom{n-r}{i-s}$ elements. Consequently,

$$\begin{aligned} P(A_{j_1} \cap ... \cap A_{j_r}) &= \sum_{i=s}^{n-r+s} \binom{n-r}{i-s} p^i q^{n-i} &= \sum_{k=0}^{n-r} \binom{n-r}{k} p^{k+s} q^{n-s-k} \\ &= p^s q^{r-s}(q+p)^{n-r} &= p^s q^{r-s}. \end{aligned}$$

Exercise 5.3.15. At the roulette we gamble on the numbers 3, 13, 22. Suppose that we know that the game has been stacked in such a way that only odd numbers may occur. What is the probability that one of the three numbers comes out? Remember that the possible results (apart the cheat) are the integers from 0 to 36.

Exercise 5.3.16. Five false coins are mixed with nine genuine. One of the coins is drawn at random.

(I) Compute the probability that a false coin is drawn.

If we draw two coins (keeping an already drawn one outside of the box), compute the probability that

(II) one is false and one is genuine;
(III) they are both false;
(IV) they are both genuine.

Exercise 5.3.17. A box contains A white stones and B black stones. A second box contains C white stones and D black stones ($A, B, C, D \in \mathbb{N}$). A stone is drawn from the first box and is transferred into the second one. Then a stone is drawn from the second. Calculate the probability of the following events:

(I) the first drawn stone is white;
(II) the first drawn stone is black;
(III) the second drawn stone is white, in case the first drawn stone is white;
(IV) the second drawn stone is white, in case the first drawn stone is black.

Exercise 5.3.18. Two stones are drawn from a box, containing four red and two white. The first drawn stone is kept outside the box. Compute the probability of the following events:

(I) both the drawn stones are white;
(II) both the drawn stones are red;
(III) the drawn stones have the same colour;
(IV) at least one of the drawn stones is red.

Exercise 5.3.19. The american senate consists of two senators for each of the fifty states. A committee of fifty senators is drawn at random. What is the probability that it contains both the senators of Alaska? Compute the probability that it contains both the senators of Alaska under the assumption that it contains at least one.

Exercise 5.3.20. We launch a balanced dice twice. Let A, B, C be the events:

A:"the result in the first launch is odd";
B: "the result in the second launch is odd";
C: "the sum of the two results is odd".
Check that A, B, C are pairwise independent, but A, B, C are not.

Exercise 5.3.21. A draft corrector examines a manuscript, containing twenty mistakes: the probability that it recognizes one of them, when he meets it, is $\frac{3}{4}$. Compute that most probable number of uncorrected mistakes after one reading. Compute the same number after two readings.

Exercise 5.3.22. The bolts produced by a factory are defective with probability 1/10. They are sold in boxes containing four pieces each. What is the probability that a box contains more than two defective bolts?

Exercise 5.3.23. It is known that the 20 per cent of travellers who have booked a place in a flight are not present at the departure So an air company accepts up to 55 reservations for a flight with 50 places. What is the probability at (at least) one traveler who has got a reservation and is present at the departure, is not able to go? (Assume that the travellers are present or not at the departure independently of each other).

Exercise 5.3.24. A bag contains nine coins with head and tail and one coin with two heads. One of these coins is drawn at random. We launch it six times and get six times head. What is the probability that we have drawn the coin with two heads?

Exercise 5.3.25. A test to single out a certain disease is positive in the 99 per cent of the cases if it is applied to somebody who is affected, in the 2 per cent of the cases if it is applied to somebody who is not affected. On account of the fact that the percentage of affected people is considered as approximately equal to $0,1$, compute the probability that a person with a positive test is really affected.

5.4 Random variables

Definition 5.4.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$. We shall say that X is a* **simple random variable** *if there exist A_1,...,A_m ($m \in \mathbb{N}$) elements of $\mathcal{A}$ pairwise incompatible, the union of which is Ω, and real numbers $\alpha_1, ..., \alpha_m$ so that*

$$X(\omega) = \alpha_i \quad \forall \omega \in A_i, \quad 1 \leq i \leq m.$$

We shall say that X is a **real random variable** *(rrv) if there exists a sequence of simple random variables $(X_k)_{k \in \mathbb{N}}$ with domain Ω, such that*

$$\lim_{k \to +\infty} X_k(\omega) = X(\omega) \qquad \forall \omega \in \Omega.$$

Remark 5.4.2. The definitions of simple and real random variables have a strong analogy with the definitions of simple and measurable functions. As a matter of fact, we can observe that, in case Ω is a Lebesgue measurable subset of $\mathbb{R}^n$ and $\mathcal{A}$ is the class of Lebesgue measurable subsets of Ω, simple random variables are exactly simple functions, real random variables are exactly measurable functions.

Definition 5.4.3. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}^n$, with $n \in \mathbb{N}$, $X(\omega) = (X_1(\omega), \dots, X_n(\omega))$. We shall say that X is a $n-$***dimensional random variable** *if $X_1, \dots, X_n$ are rrv.*

Example 5.4.4. Let us consider the launch of a balanced dice with six faces. Let $\Omega := \{1, 2, 3, 4, 5, 6\}$, $\mathcal{A} = \mathcal{P}(\Omega)$, $P(A) := \sharp(A)/6$, $\forall A \in \mathcal{A}$. We set

$$X : \Omega \to \mathbb{R},$$
$$X(\omega) = \begin{cases} 0 & \text{if } \omega \text{ is even,} \\ 1 & \text{if } \omega \text{ is odd} \end{cases}$$

X is a simple random variable, as, if we set $A_1 := \{2, 4, 6\}$ and $A_2 = \{1, 3, 5\}$, A_1 and A_2 are incompatible events with union Ω, $X(\omega) = 1$ if $\omega \in A_1$, $X(\omega) = 0$ if $\omega \in A_2$.

Example 5.4.5. We launch a dart against a round target, with radius r ($r \in \mathbb{R}^+$). We schematize the experiment, setting

$$\begin{gathered}\Omega := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le r^2\},\\ \mathcal{A} := \{A \subseteq \Omega : A \in \mathcal{M}_2\},\\ P(A) := \tfrac{L_2(A)}{\pi r^2}, \qquad A \in \mathcal{A}.\end{gathered} \tag{5.4.1}$$

This choice of P means, evidently, that we think that the probability that the dart touches A is proportional to its area $L_2(A)$.

It is easy to see that $(\Omega, \mathcal{A}, P)$ is a probability space. Given $\omega \in \Omega$, we indicate with $X(\omega)$ the distance of ω from the origin $(0,0)$. Obviously, $X(\omega) = \|\omega\|$, where we have indicated with $\|.\|$ the euclidean norm. By virtue of Remark 5.4.2 and Theorem 1.2.4, X is a rrv.

Another example of rrv connected with the same experiment is

$$\begin{gathered}Y : \Omega \to \mathbb{R},\\ Y(\omega_1, \omega_2) = |\omega_1|,\end{gathered} \tag{5.4.2}$$

measuring the distance of the point ω centered by the dart from the axis ω_1.

Finally, if we set

$$\begin{gathered}Z : \Omega \to \mathbb{R}^2,\\ Z(\omega) = (X(\omega), Y(\omega)),\end{gathered} \tag{5.4.3}$$

we obtain a bidimensional random variable.

We introduce now some notations: let Ω, R, be sets, $X : \Omega \to R$; if $B \subseteq R$, we set

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}. \tag{5.4.4}$$

In case $R = \mathbb{R}$, we set, given $a \in \mathbb{R}$,

$$\{X < a\} := \{\omega \in \Omega : X(\omega) < a\}. \tag{5.4.5}$$

Analogous notations will be employed for other kinds of inequality.

The next theorem, of relevant importance, has purely set-theoretical contents:

Theorem 5.4.6. *Let Ω, R be sets, , $X : \Omega \to R$, $\{B_i : i \in \mathcal{I}\}$ a family of subsets of R, depending on the parameter $i \in \mathcal{I}$, $B \subseteq R$. Then:*
(I) $\{X \in \cup_{i\in\mathcal{I}} B_i\} = \cup_{i\in\mathcal{I}}\{X \in B_i\}$;
(II) $\{X \in \cap_{i\in\mathcal{I}} B_i\} = \cap_{i\in\mathcal{I}}\{X \in B_i\}$;
(III) $\{X \in B^c\} = \{X \in B\}^c = \Omega \setminus \{X \in B\}$, $B^c := R \setminus B$.

Proof (I) Saying that $\omega \in \{X \in \cup_{i\in\mathcal{I}} B_i\}$ means that $X(\omega) \in \cup_{i\in\mathcal{I}} B_i$ and so that there exists i_0 in $\mathcal{I}$ such that $X(\omega) \in B_{i_0}$. Therefore, $\omega \in \{X \in B_{i_0}\} \subseteq \cup_{i\in\mathcal{I}}\{X \in B_i\}$. On the other hand, if $\omega \in \cup_{i\in\mathcal{I}}\{X \in B_i\}$, there exists $i_0 \in \mathcal{I}$ such that $\omega \in \{X \in B_{i_0}\}$. So $X(\omega) \in B_{i_0}$, hence $X(\omega) \in \cup_{i\in\mathcal{I}} B_i$, and we get the conclusion.

(II) Saying that $\omega \in \{X \in \cap_{i\in\mathcal{I}} B_i\}$ means that $X(\omega) \in \cap_{i\in\mathcal{I}} B_i$. This is equivalent to $X(\omega) \in B_i$ for every $i \in \mathcal{I}$, to $\omega \in \{X \in B_i\}$ for every i and so to $\omega \in \cap_{i\in\mathcal{I}}\{X \in B_i\}$.

(III) $\omega \in \{X \in B^c\}$ if and only if $X(\omega) \in B^c$. This is equivalent to $\omega \in \Omega$ and $X(\omega) \notin B$, which is the same as $\{X \in B\}^c$.

□

Given an $n-$dimensional random variable X and $B \subseteq \mathbb{R}^n$, we wonder whether the probability of $\{X \in B\}$ is defined. Here we are going to face the following problem: under what conditions, given $B \subseteq \mathbb{R}^n$, $\{X \in B\}$ is an event, that is, $\{X \in B\}$ belongs to the $\sigma-$algebra $\mathcal{A}$? We shall see that there exists a large class of subsets of $\mathbb{R}^n$ for which this holds: the so called Borel subsets of $\mathbb{R}^n$.

We begin with a preliminary result:

Theorem 5.4.7. *Let Ω and $\mathcal{I}$ be sets, $\mathcal{I} \neq \emptyset$ and $\forall i \in \mathcal{I}$ let $\mathcal{A}_i$ be a $\sigma-$algebra in Ω. Then $\cap_{i\in\mathcal{I}}\mathcal{A}_i$ is a $\sigma-$algebra in Ω.*

Proof The proof is really trivial. Here we limit ourselves to prove that, if $A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$ for every $n \in \mathbb{N}$, then $\cup_{n\in\mathbb{N}}A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$. Let $i \in \mathcal{I}$. Then, for every $n \in \mathbb{N}$, $A_n \in \mathcal{A}_i$. As $\mathcal{A}_i$ is a $\sigma-$algebra, $\cup_{n\in\mathbb{N}}A_n \in \mathcal{A}_i$. This holds for every i. So $\cup_{n\in\mathbb{N}}A_n \in \cap_{i\in\mathcal{I}}\mathcal{A}_i$.
□

Definition 5.4.8. *Let $n \in \mathbb{N}$; we indicate with $\mathcal{B}(\mathbb{R}^n)$ the intersection of all $\sigma-$algebras in $\mathbb{R}^n$ containing the class of open subsets.*

Remark 5.4.9. By virtue of Theorem 5.4.7, $\mathcal{B}(\mathbb{R}^n)$ is a $\sigma-$algebra in $\mathbb{R}^n$ and may be thought as the smallest $\sigma-$algebra in $\mathbb{R}^n$, containing all open subsets. The elements of $\mathcal{B}(\mathbb{R}^n)$ are called **Borel subsets** of $\mathbb{R}^n$. It is possible to prove that, in any normed space, closed subsets are exactly the complements of open subsets. So even closed subsets are Borel, by virtue of properties of $\sigma-$algebras (see Theorem 5.1.4).

Observe that $\mathcal{M}_n$ is a $\sigma-$algebra containing open subsets. So Borel subsets are measurable in the sense of Lebesgue. In fact, $\mathcal{B}(\mathbb{R}^n)$ is strictly contained in $\mathcal{M}_n$. However, even in the case of $\mathcal{B}(\mathbb{R}^n)$ we can say that it contains all subsets of $\mathbb{R}^n$ appearing in applications.

The following result holds:

Theorem 5.4.10. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $X : \Omega \to \mathbb{R}^n$ ($n \in \mathbb{N}$) be an $n-$dimensional randon variable. Then, $\forall B \in \mathcal{B}(\mathbb{R}^n)$, $\{X \in B\}$ is an event, that is, $\{X \in B\} \in \mathcal{A}$.*

Definition 5.4.11. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$ a rrv. The* **distribution function** *F_X of X is the following*

$$\begin{cases} F_X : \mathbb{R} \to \mathbb{R}, \\ F_X(t) = P(X \leq t), \quad t \in \mathbb{R}. \end{cases} \tag{5.4.6}$$

Remark 5.4.12. In Definition 5.4.11 we have written $P(X \leq t)$ instead of $P(\{X \leq t\})$, or $P(\{\omega \in \Omega : X(\omega) \leq t\})$. In general, we shall employ the notations $P(X \in B)$, $P(X \geq t)$, etc. in alternative to $P(\{X \in B\})$, $P(\{X \geq t\})$, etc., whenever this simplification does not seem to cause confusion. We observe that $\{x \in \mathbb{R} : x \leq t\}$ is a Borel subset of $\mathbb{R}$, because it is closed. So, in force of Theorem 5.4.10, Definition 5.4.11 is well posed.

Example 5.4.13. Consider Example 5.4.4. It is easy to verify that

$$\{X \leq t\} = \begin{cases} \emptyset & \text{if} \quad t < 0, \\ \{2,4,6\} & \text{if} \quad 0 \leq t < 1, \\ \Omega & \text{if} \quad t \geq 1. \end{cases}$$

So

$$F_X(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ 1/2 & \text{if} \quad 0 \leq t < 1, \\ 1 & \text{if} \quad t \geq 1. \end{cases} \tag{5.4.7}$$

Example 5.4.14. We determine the distribution function of the rrv in Example 5.4.5. Let $t \in \mathbb{R}$. We have

$$\{X \le t\} = \begin{cases} \emptyset & \text{if} \quad t < 0, \\ \{\omega \in \Omega : \|\omega\| \le t\} & \text{if} \quad 0 \le t \le r, \\ \Omega & \text{if} \quad t > r. \end{cases}$$

So

$$F_X(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ (t/r)^2 & \text{if} \quad 0 \le t < r, \\ 1 & \text{if} \quad t \ge r. \end{cases} \tag{5.4.8}$$

Before stating, and at least partially, showing the main properties of the distribution function of a rrv, we introduce without proof two important properties of probability measures.

Theorem 5.4.15. *Let $(\Omega, \mathcal{A}, P)$ be a probability space. Next, let $(A_n)_{n\in\mathbb{N}}$ be a sequence of events. Then.*

(I) if $A_n \subseteq A_{n+1}$ $\forall n \in \mathbb{N}$ and $A = \cup_{n\in\mathbb{N}} A_n$, then

$$P(A) = \lim_{n\to+\infty} P(A_n);$$

(II) if $A_{n+1} \subseteq A_n$ $\forall n \in \mathbb{N}$ and $A = \cap_{n\in\mathbb{N}} A_n$, then

$$P(A) = \lim_{n\to+\infty} P(A_n).$$

Remark 5.4.16. In each of the two cases of Theorem 5.4.15, $\lim\limits_{n\to+\infty} P(A_n)$ exists because the sequence $(P(A_n))_{n\in\mathbb{N}}$ is monotonic (nondecreasing in the first case, non increasing in the second).

Theorem 5.4.17. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$ a rrv. Then:*

(I) $\forall t \in \mathbb{R}$, $0 \le F_X(t) \le 1$;

(II) $\forall a, b \in \mathbb{R}$, con $a < b$,

$$P(a < X \le b) = F_X(b) - F_X(a);$$

(III) F_X is monotonic nondecreasing;

(IV) F_X is right continuous, that is, $\forall t \in \mathbb{R}$, $\lim\limits_{s\to t^+} F_X(s) = F_X(t)$;

(V) $\lim\limits_{t\to-\infty} F_X(t) = 0$ and $\lim\limits_{t\to+\infty} F_X(t) = 1$.

Proof (I) is obvious.

(II) One has $\{X \le a\} \subseteq \{X \le b\}$ and $\{X \le b\} \setminus \{X \le a\} = \{a < X \le b\}$. So (II) follows from Theorem 5.1.6(III).

(III) If $a < b$, from (II) we have $F_X(b) - F_X(a) \ge 0$.

(IV) Let $t \in \mathbb{R}$. Then $\{X \le t\} = \cap_{n\in\mathbb{N}}\{X \le t + 1/n\}$. In fact, if $\omega \in \Omega$ and $X(\omega) \le t$, then $X(\omega) \le t + 1/n$ $\forall n \in \mathbb{N}$. On the other hand, from $X(\omega) \le t + 1/n$ $\forall n \in \mathbb{N}$, it follows $X(\omega) \le t$. In fact, if $X(\omega) > t$, taking n sufficiently large, we get $t + 1/n < X(\omega)$. As $\forall n \in \mathbb{N}$ $\{X \le t + 1/(n+1)\} \subseteq \{X \le t + 1/n\}$, from Theorem 5.4.17(II) we obtain

$$F_X(t) = P(X \le t) = \lim_{n\to+\infty} P(X \le t + 1/n) = \lim_{n\to+\infty} F_X(t + 1/n). \tag{5.4.9}$$

Let now $\epsilon \in \mathbb{R}^+$. By (5.4.9), there exists $n_0 \in \mathbb{N}$, such that $F_X(t + 1/n_0) < F_X(t) + \epsilon$. So, if $0 < s < 1/n_0$, from (III), we have

$$F_X(t) - \epsilon < F_X(t) \le F_X(s) \le F_X(t + 1/n_0) < F_X(t) + \epsilon.$$

(V) For every $n \in \mathbb{N}$, $\{X \leq -n-1\} \subseteq \{X \leq -n\}$. As $\cap_{n\in\mathbb{N}}\{X \leq -n\} = \emptyset$, one has, applying again Theorem 5.4.15(II),

$$\lim_{n\to+\infty} F_X(-n) = \lim_{n\to+\infty} P(X \leq -n) = 0. \tag{5.4.10}$$

Then the first limit in (V) follows from (5.4.10) and (III).

Moreover, for every $n \in \mathbb{N}$, $\{X \leq n\} \subseteq \{X \leq n+1\}$. As $\cup_{n\in\mathbb{N}}\{X \leq n\} = \Omega$, one has, applying Theorem 5.4.15(I),

$$\lim_{n\to+\infty} F_X(n) = \lim_{n\to+\infty} P(X \leq n) = 1. \tag{5.4.11}$$

So the second limit in (V) follows from (5.4.11) and (III).

□

Remark 5.4.18. Example 5.4.13 shows that the distribution function, which is always right continuous, may be discontinuous. In that case, one has, for example,

$$\lim_{s\to 0^-} F_X(s) = 0 \neq F_X(0) = 1/2.$$

See, concerning this, also Exercise 5.4.53.

The distribution function is defined only for rrv-s. Concerning $n-$dimensional random variables ($n \in \mathbb{N}$), we introduce the notion of distribution law:

Definition 5.4.19. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}^n$ a $n-$dimensional random variable. We define* **distribution law** *of X the function*

$$\begin{cases} Q_X : \mathcal{B}(\mathbb{R}^n) \to \mathbb{R}, \\ Q_X(B) = P(X \in B), \quad B \in \mathcal{B}(\mathbb{R}^n). \end{cases} \tag{5.4.12}$$

Theorem 5.4.20. *The distribution law of a $n-$dimensional random variable X is a probability measure in the $\sigma-$algebra $\mathcal{B}(\mathbb{R}^n)$.*

Proof First of all, it is clear from (5.4.12) that $Q_X(B) \in [0,1]$ $\forall B \in \mathcal{B}(\mathbb{R}^n)$. Moreover,

$$Q_X(\mathbb{R}^n) = P(X \in \mathbb{R}^n) = P(\Omega) = 1.$$

Finally, let $\{B_k : k \in \mathbb{N}\}$ be a sequence of pairwise disjoint Borel subsets of $\mathbb{R}^n$. It is clear that the events $\{X \in B_k\}$ ($k \in \mathbb{N}$) are pairwise incompatible. So, by virtue of Theorem 5.4.6(I), one has

$$\begin{array}{rcl} Q_X(\cup_{k\in\mathbb{N}} B_k) & = & P(X \in \cup_{k\in\mathbb{N}} B_k) \\ = P(\cup_{k\in\mathbb{N}}\{X \in B_k\}) & = & \sum_{k=1}^{\infty} P(X \in B_k) \\ & = & \sum_{k=1}^{\infty} Q_X(B_k). \end{array}$$

□

Example 5.4.21. If $X : \Omega \to \mathbb{R}^n$ is a random variable with finite range $\{\alpha^1, ..., \alpha^r\}$, we have, for each $B \in \mathcal{B}(\mathbb{R}^n)$,

$$Q_X(B) = P(X \in B) = \sum_{\alpha^j \in B} P(X = \alpha^j). \tag{5.4.13}$$

The result of Example 5.4.21 can be generalized, introducing the notion of discrete random variable:

Definition 5.4.22. *Let X be a $n-$dimensional random variable. We shall say that X is* **discrete** *if its range is finite or countable.*

If X is a discrete random variable, with range $\{\alpha^j : j \in \mathbb{N}\}$, one can easily check that, for every $B \in \mathcal{B}(\mathbb{R}^n)$, formula (5.4.13) holds. In the second term of (5.4.13), we shall have a finite sum or a series, in dependence of the fact that the range is finite or countable.

Example 5.4.23. Suppose of throwing a balanced dice, until we get a "head". Suppose also that results in different launches are independent. We indicate with H the single result "head", with T the single result "tail". We set

$$\Omega := \{H, TH, TTH, TTTH, ...\}. \tag{5.4.14}$$

Ω is the set whose elements are all possible finite sequences $\underbrace{TT...T}_{n}H$, with $n \in \mathbb{N}$, to which we add the sequence with the only term H and the infinite sequence $TTTTTT....$ Let $\omega \in \Omega$. We put:

$$p_\omega := \begin{cases} 2^{-n-1} & \text{if} \quad \omega = \underbrace{TT...T}_{n}H, \\ 1/2 & \text{if} \quad \omega = H, \\ 0 & \text{if} \quad \omega \text{ and the sequence identically equal to T.} \end{cases} \tag{5.4.15}$$

We put also $\mathcal{A} = \mathcal{P}(\Omega)$ and, if $A \in \mathcal{A}$,

$$P(A) := \sum_{\omega \in A} p_\omega. \tag{5.4.16}$$

We observe that

$$P(\Omega) = \sum_{n=0}^{\infty} 2^{-n-1} = \sum_{n=1}^{\infty} 2^{-n} = \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$$

and it is not difficult to check that P is a probability measure. We indicate with X the rrv assigning to each sequence its number of launches. So $X(H) = 1$, $X(TH) = 2$, etc.. If ω is the sequence identically equal to T, we set $X(\omega) = +\infty$. With a small abuse of language, (X has $+\infty$ in its range; however, the event $\{X = +\infty\}$ has probability zero), we shall say that X is a rrv. Let now $B =]10, +\infty[$. Evidently,

$$\begin{aligned} Q_X(B) &= P(X \in]10, +\infty[) = P(X > 10) = \textstyle\sum_{j=11}^{\infty} 2^{-j} \\ &= \textstyle\sum_{i=1}^{\infty} 2^{-(10+i)} = 2^{-10}. \end{aligned}$$

We calculate the probability that X is odd. In this case, $B = \{1, 3, 5, ...\}$. Then

$$\begin{aligned} Q_X(B) &= P(X \in \{1, 3, 5, ...\}) = \sum_{j \in \{1,3,5,...\}} 2^{-j} \\ &= \textstyle\sum_{i=0}^{\infty} 2^{-(2i+1)} = \frac{1}{2} \sum_{i=0}^{\infty} 4^{-i} \\ &= \frac{1}{2} \frac{1}{1-\frac{1}{4}} = \frac{2}{3}. \end{aligned}$$

Definition 5.4.24. *Let X be a $n-$dimensional random variable, $f \in \mathcal{L}^1(\mathbb{R}^n)$, almost everywhere nonnegative. We shall say that f is a* **density** *of X if $\forall B \in \mathcal{B}(\mathbb{R}^n)$*

$$Q_X(B) = P(X \in B) = \int_B f(t)dt.$$

Remark 5.4.25. It is possible to check that two densities of the same random variable coincide almost everywhere. We observe also that, if f is a density of X,

$$\int_{\mathbb{R}^n} f(t)dt = Q_X(\mathbb{R}^n) = 1.$$

Moreover, if $L_n(B) = 0$,

$$Q_X(B) = \int_B f(t)dt = 0.$$

In particular, if X has a density, for every $\alpha \in \mathbb{R}^n$, $P(X = \alpha) = 0$.

Remark 5.4.26. Let X be a rrv, with the density f. Then, $\forall t \in \mathbb{R}$,

$$F_X(t) = P(X \leq t) = \int_{]-\infty,t]} f(s)ds. \tag{5.4.17}$$

It is true also the following inverse statement: if there exists $f \in \mathcal{L}^1(\mathbb{R})$, almost everywhere nonnegative, fulfilling (5.4.17), then f is a density of X. So, if $B \in \mathcal{B}(\mathbb{R})$, one has

$$Q_X(B) = P(X \in B) = \int_B f(s)ds.$$

So, we consider Example 5.4.14. We set

$$f(t) = \begin{cases} 0 & \text{if} \quad t \in]-\infty, 0[\cup]r, +\infty[, \\ \frac{2t}{r^2} & \text{if} \quad t \in [0, r]. \end{cases}$$

Then it is clear that, for every $t \in \mathbb{R}$,

$$F_X(t) = \int_{]-\infty,t]} f(s)ds.$$

Therefore, X has f as a density.

Example 5.4.27. We consider the interval $\Omega = [0, a[$ $(a > 0)$. We choose at random a point ω in Ω. With this, we mean that the probability that the chosen point belongs to a certain subset B of Ω, measurable in the sense of Lebesgue, is proportional to $L_1(B)$. So we put

$$\begin{array}{ll} \mathcal{A} := \{A \subseteq \Omega : A \in \mathcal{M}_1\}, & \\ P(A) = cL_1(A), & A \in \mathcal{A}, \end{array} \tag{5.4.18}$$

with $c \in \mathbb{R}^+$. c must be such that $P(\Omega) = 1$. Therefore,

$$1 = P(\Omega) = P([0, a[) = ca,$$

hence $c = a^{-1}$. So

$$P(A) = \frac{L_1(A)}{a}$$

for every $A \in \mathcal{A}$.

Let us consider the rrv X, such that $\forall \omega \in \Omega$ $X(\Omega)$ is the quotient of the distances of ω from 0 and from a. So

$$\begin{cases} X : \Omega \to \mathbb{R}, \\ X(\omega) = \frac{\omega}{a-\omega}, \quad \omega \in \Omega. \end{cases}$$

Let us determine the distribution function F_X. If $t \in \mathbb{R}$, we have:

$$\{X \le t\} = \begin{cases} \emptyset & \text{if} \quad t < 0, \\ [0, \frac{at}{1+t}] & \text{if} \quad t \ge 0, \end{cases}$$

hence

$$F_X(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ \frac{t}{1+t} & \text{if} \quad t \ge 0. \end{cases}$$

We set

$$f(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ \frac{1}{(1+t)^2} & \text{if} \quad t \ge 0. \end{cases}$$

We have

$$F_X(t) = \int_{]-\infty,t]} f(s)ds$$

for every $t \in \mathbb{R}$, so we can deduce that f is a density of X.

Example 5.4.28. We refer to Examples 5.1.2 e 5.1.10. We set

$$X(\omega) = \omega \quad (\omega \in \mathbb{R}^+).$$

X is the rrv, specifying the time duration of the device (or the instant it stops working). Evidently, for every $t \in \mathbb{R}$,

$$\{X \le t\} = \begin{cases} \emptyset & \text{if} \quad t \le 0, \\ ]0, t] & \text{if} \quad t > 0, \end{cases}$$

hence

$$F_X(t) = \begin{cases} 0 & \text{if} \quad t \le 0, \\ \int_{]0,t]} g(s)ds & \text{if} \quad t > 0. \end{cases}$$

So, if we set

$$\begin{cases} \tilde{g} : \mathbb{R} \to \mathbb{R}, \\ \tilde{g}(t) = \begin{cases} g(t) & \text{if } t > 0, \\ 0 & \text{if } t \le 0, \end{cases} \end{cases}$$

we have that $\tilde{g}$ is a density of X. In case $g(t) = \lambda e^{-\lambda t}$ ($\lambda \in \mathbb{R}^+$), it is said that X has an exponential distribution. Let us justify this choice of g. It is a consequence of the following assumption: that the probability that the device stops working in the time interval $]t, t+h]$ $(t, h \in \mathbb{R}^+)$ if it is still working at time t depends only on h. Precisely,

$$P(\{t < X \le t+h\}|\{X > t\}) = \phi(h), \tag{5.4.19}$$

with $\phi : \mathbb{R}^+ \to \mathbb{R}$. One has

$$P(\{t < X \le t+h\}|\{X > t\}) = \frac{P(t < X \le X+h)}{P(X > t)} = \frac{F_X(t+h) - F_X(t)}{1 - F_X(t)},$$

hence

$$\frac{F_X(t+h) - F_X(t)}{h} = \frac{\phi(h)}{h}[1 - F_X(t)]. \tag{5.4.20}$$

Suppose now that F_X is differentiable in $\mathbb{R}^+$ and that $\lim_{h\to 0} \frac{\phi(h)}{h} = \lambda \in \mathbb{R}^+$. Passing to the limit as $h \to 0$ in (5.4.20), we obtain that F_X satisfies in $\mathbb{R}^+$ the ordinary differential equation

$$F_X'(t) = \lambda[1 - F_X(t)]. \tag{5.4.21}$$

It easily follows that F_X has the form

$$F_X(t) = Ce^{-\lambda t} + 1,$$

for some $C \in \mathbb{R}$, $C < 0$. From the fact that $F_X(0) = 0$ and assuming that F_X is continuous in $\mathbb{R}$, we obtain $C = -1$. So

$$F_X(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ 1 - e^{-\lambda t} & \text{if } t \geq 0. \end{cases} \tag{5.4.22}$$

From (5.4.22) it is easy to check that X has the following density f:

$$f(t) = \begin{cases} 0 & \text{if} \quad t < 0, \\ \lambda e^{-\lambda t} & \text{if } t \geq 0. \end{cases} \tag{5.4.23}$$

To conclude, we might wonder what is the meaning of condition (5.4.19). Intuitively, it is the following: that, in some sense, the aging of the device is negligible. In fact, this condition implies, in particular, that the probability that a working device at time t is still working at time $2t$ coincides with the probability that a working device at time $2t$ is still working at time $3t$.

Now we pass to the basic notion of expectation of a rrv. We begin with the case of a simple random variable.

Definition 5.4.29. *Let X be a simple random variable in the probability space $(\Omega, \mathcal{A}, P)$. Suppose that $\Omega = A_1 \cup ... \cup A_m$, with A_1,...,A_m pairwise disjoint, and that $X(\omega) = \alpha_j$ if $\omega \in A_j$ $(1 \leq j \leq m)$. Then we call* **expectation** *(or average) of X and indicate with $E(X)$ the real number $\sum_{j=1}^{m} \alpha_j P(A_j)$.*

Remark 5.4.30. One can check that the definition of $E(X)$ does not depend on the chosen decomposition $\{A_j : 1 \leq j \leq m\}$ of Ω. We observe also that, if $X(\omega) = \alpha \ \forall \alpha \in \omega$, $E(X) = \alpha$.

Example 5.4.31. Let X be the rrv introduced in Example 5.4.4. One has

$$E(X) = P(A_2) = \frac{1}{2}.$$

Example 5.4.32. Let us consider a Bernoulli process, with probability of success and failure in each single test equal to p and q respectively (see Example 5.3.10). Let X be the rrv assigning to each sequence $\omega = (\omega_1, ..., \omega_n)$ the number of successes. So, if, for $m = 0, ..., n$, B_m is the event introduced in Remark 5.3.11, we have, on the basis of (5.3.7),

$$E(X) = \sum_{m=1}^{n} m \binom{n}{m} p^m q^{n-m}.$$

We pass to the definition of expectation of a rrv.

Definition 5.4.33. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $X : \Omega \to \mathbb{R}$ a rrv such that $X(\omega) \geq 0$ $\forall \omega \in \Omega$. We set*

$$\begin{aligned} E(X) \ &:= \sup\{E(Y) : Y : \Omega \to [0, +\infty[\text{ simple random variable}, \\ &0 \leq Y(\omega) \leq X(\omega) \quad \forall \omega \in \Omega\}. \end{aligned} \tag{5.4.24}$$

Finally, let $X : \Omega \to \mathbb{R}$ *be a generic rrv. We set*

$$X_+ := \phi_+ \circ X, \qquad X_- := \phi_- \circ X, \tag{5.4.25}$$

with $\phi_\pm$ *defined in (1.2.1)-(1.2.2). It is possible to check that* $X_\pm$ *are nonnegative real random variables. If* $E(X_+) < +\infty$ *and* $E(X_-) < +\infty$*, we call expectation of* X *the real number*

$$E(X) := E(X_+) - E(X_-). \tag{5.4.26}$$

Remark 5.4.34. Definition 5.4.33 recalls the definitions of integral 1.2.6 and 1.2.9. In fact, both probability measures and Lebesgue measures can be considered as particular cases in an abstract theory of integration. So, it will not be surprising to see that expectation and integral have similar properties.

Theorem 5.4.35. *Let* $(\Omega, \mathcal{A}, P)$ *be a probability space ,* $X, Y : \Omega \to \mathbb{R}$ *real random variables with a well defined expectation. Then:*
(I) $X + Y$ *has a well defined expectation; moreover,* $E(X + Y) = E(X) + E(Y)$*;*
(II) if $\alpha \in \mathbb{R}$*,* αX *has a well defined expectation and* $E(\alpha X) = \alpha E(X)$*;*
(III) if $X(\omega) \leq Y(\omega)$ $\forall \omega \in \Omega$*,* $E(X) \leq E(Y)$*;*
(IV) in particular, if $X(\omega) \geq 0$ $\forall \omega \in \Omega$*,* $E(X) \geq 0$*.*

In the particular case of simple random variables, we propose the proof of Theorem 5.4.35 as an exercise (see Exercise 5.4.56).

We pass to some useful results for the computation of the expectation of a rrv.

Theorem 5.4.36. *Let* X *be a discrete rrv in the probability space* $(\Omega, \mathcal{A}, P)$*. Suppose that che* $\Omega = \cup_{j \in \mathbb{N}} A_j$*, with* A_j *pairwise disjoint events and* $X(\omega) = \alpha_j \in \mathbb{R}$ $\forall \omega \in A_j$*. Then* X *has a well defined expectation if and only if the series* $\sum_{j=1}^{\infty} \alpha_j P(A_j)$ *is absolutely convergent. In this case,*

$$E(X) = \sum_{j=1}^{\infty} \alpha_j P(A_j).$$

Example 5.4.37. Let us consider the probability space in Example 5.4.23. Again, let X be the rrv counting the number of launches, which is necessary to get "head". We have already seen that the range of X is the set of natural numbers plus $+\infty$, with probability 0. Moreover, we have that, $\forall j \in \mathbb{N}$, $P(X = j) = 2^{-j}$. As the series $\sum_{j=1}^{\infty} j 2^{-j}$ is convergent (this can be seen employing, for example, the quotient test), X has a well defined expectation and

$$E(X) = \sum_{j=1}^{\infty} j 2^{-j}. \tag{5.4.27}$$

Let us compute the sum of the series in (5.4.27). Considering the power series $\sum_{j=0}^{\infty} z^j$, which has radius of convergence 1 and is such that, $\forall z \in \mathbb{C}$, with $|z| < 1$,

$$\sum_{j=0}^{\infty} z^j = (1 - z)^{-1}. \tag{5.4.28}$$

Applying Theorem 3.3.15, we can differentiate in each of the two terms in (5.4.28) and get

$$\sum_{j=1}^{\infty} j z^{j-1} = (1-z)^{-2}, \tag{5.4.29}$$

which is valid if $|z| < 1$. Multiplying by z, we obtain immediately, if $|z| < 1$,

$$\sum_{j=1}^{\infty} j z^{j} = z(1-z)^{-2}. \tag{5.4.30}$$

Then, taking $z = 1/2$, we obtain

$$E(X) = \sum_{j=1}^{\infty} j 2^{-j} = 2.$$

Concerning random variables with density, the following result holds:

Theorem 5.4.38. *Let X be a rrv, with density f. Then expectation is defined for X if and only if $t \to tf(t)$ is summable in $\mathbb{R}$. In such a case,*

$$E(X) = \int_{\mathbb{R}} tf(t)dt.$$

Example 5.4.39. Let us consider a rrv, with an exponential distribution (see Example 5.4.28). Suppose that its density is $f(t) = \lambda e^{-\lambda t}\chi_+(t)$, where we have indicated with χ_+ the characteristic function of $[0, +\infty[$. Applying Theorem 5.4.38, we obtain

$$E(X) = \int_{[0,+\infty[} \lambda t e^{-\lambda t} dt = \lambda^{-1}.$$

So the parameter λ stands for the inverse of the expectation of X.

Theorem 5.4.38 has the following important generalization:

Theorem 5.4.40. *Let X be a $n-$dimensional random variable, with density f and let $g : \mathbb{R}^n \to \mathbb{R}$ be continuous. Then $g(X) := g \circ X$ is a rrv, the expectation of which is defined if and only if $x \to g(x)f(x)$ is summable in $\mathbb{R}^n$. In such a case,*

$$E(g(X)) = \int_{\mathbb{R}^n} g(x)f(x)dx.$$

Incomplete proof We check only that the conclusion holds, in case $g : \mathbb{R}^n \to \mathbb{R}$ is a simple function, with $g(x) = \sum_{j=1}^{m} \beta_j \chi_{B_j}(x)$, with B_1,...,B_m pairwise disjoint Borel subsets of $\mathbb{R}^n$, and β_1,...,β_m real numbers. Observe that such a g is not (generally speaking) continuous, but every continuous function can be suitably approximated with functions of this type.

We set, for $j = 1, ..., m$,

$$A_j := \{X \in B_j\}.$$

We observe that the events A_j are pairwise incompatible. So, clearly, $g(X)$ coincides with the simple random variable $\sum_{j=1}^{m} \beta_j \chi_{A_j}$. From this, we get

$$\begin{aligned} E(g(X)) &= \textstyle\sum_{j=1}^{m} \beta_j P(A_j) \\ &= \textstyle\sum_{j=1}^{m} \beta_j \int_{B_j} f(x)dx \\ &= \textstyle\int_{\mathbb{R}^n} \sum_{j=1}^{m} \beta_j \chi_{B_j}(x) f(x)dx \\ &= \textstyle\int_{\mathbb{R}^n} g(x)f(x)dx. \end{aligned}$$

□

Example 5.4.41. We consider Example 5.4.5. Let $X : \Omega \to \mathbb{R}^2$, $X(\omega) = \omega$ (We recall that $\Omega \subseteq \mathbb{R}^2$). X is a bidimensional random variable. If $B \in \mathcal{B}(\mathbb{R}^2)$, we have

$$\{X \in B\} = \Omega \cap B,$$

hence

$$P(X \in B) = P(\Omega \cap B) = \frac{L_2(\Omega \cap B)}{\pi r^2} = \int_B f(x)dx,$$

with

$$f(x) = \begin{cases} (\pi r^2)^{-1} & \text{if} \quad x \in \Omega, \\ 0 & \text{if} \quad x \notin \Omega. \end{cases}$$

Let X_1 be the first component of X. Let us observe that $X_1(\omega)^2 = g(X(\omega))$, with $g : \mathbb{R}^2 \to \mathbb{R}$, $g(x_1, x_2) = x_1^2$. By Theorem 5.4.40, we have

$$\begin{aligned} E(X_1^2) &= \textstyle\int_{\mathbb{R}^2} x_1^2 f(x_1, x_2) dx_1 dx_2 \\ &= (\pi r^2)^{-1} \textstyle\int_\Omega x_1^2 dx_1 dx_2 \\ &= (\pi r^2)^{-1} \textstyle\int_0^r (\int_0^{2\pi} (\rho \cos(\theta))^2 \rho d\theta) d\rho \\ &= (\pi r^2)^{-1} \textstyle\int_0^r \rho^3 d\rho \int_0^{2\pi} \cos^2(\theta) d\theta \\ &= \tfrac{r^2}{4}. \end{aligned}$$

Definition 5.4.42. *Let X be a rrv, such that $E(X^2) < +\infty$. We define* **variance** *of X the number $\sigma^2(X)$ defined as*

$$\sigma^2(X) := E((X - E(X))^2).$$

Remark 5.4.43. If $E(X^2) < +\infty$, then expectation is defined for X. In fact, from the usual elementary inequality (2.4.3), we have

$$|X| = |X| \cdot 1 \leq \frac{1}{2}(X^2 + 1),$$

hence

$$E(X_+) + E(X_-) = E(|X|) \leq \frac{1}{2}(E(X^2) + 1).$$

We observe also that

$$(X - E(X))^2 = X^2 - 2E(X)X + E(X)^2,$$

from which we draw the useful formula

$$\sigma^2(X) = E(X^2) - 2E(X)^2 + E(X)^2 = E(X^2) - E(X)^2. \tag{5.4.31}$$

Remark 5.4.44. Intuitively, variance indicates how much a random variable is scattered, with respect to its expectation. For example, it is easy to check that a constant random variable has variance zero. (see Exercise 5.4.57). On the other hand, if X is a random variable, the range of which is $\{c, -c\}$ $(c \in \mathbb{R}^+)$, with $P(X = c) = P(X = -c) = \frac{1}{2}$, one has

$$\sigma^2(X) = c^2.$$

Example 5.4.45. Let X be a rrv with density f. Then, applying Theorem 5.4.40, we can say that

$$E(X^2) = \int_{\mathbb{R}} t^2 f(t) dt.$$

If this expectation is finite, variance is defined for X and

$$\sigma^2(X) = \int_{\mathbb{R}} t^2 f(t) dt - (\int_{\mathbb{R}} t f(t) dt)^2.$$

Definition 5.4.46. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, let $\mathcal{I}$ be a family of indexes and, for every $i \in \mathcal{I}$, let X_i be a $n-$dimensional random variable, with domain Ω. We shall say that the family $\{X_i : i \in \mathcal{I}\}$ is* **independent** *(or that the random variables $\{X_i : i \in \mathcal{I}\}$ are independent) if, whatever is the choice of $i_1, ..., i_m$ in $\mathcal{I}$ pairwise distinct ($m \in \mathbb{N}$ arbitrary) and of B_1,..., B_m Borel subsets of $\mathbb{R}^n$, one has that the events $\{X_{i_1} \in B_1\}$,...,$\{X_{i_m} \in B_m\}$ are independent. In other words,*

$$P(X_{i_1} \in B_1, ..., X_{i_m} \in B_m) = P(X_{i_1} \in B_1) \cdot ... \cdot P(X_{i_m} \in B_m).$$

Example 5.4.47. Let us consider a Bernoulli process (see Example 5.3.10). We indicate with Ω the set of the $n-$tuples $\omega = (\omega_1, ..., \omega_i, ..., \omega_n)$, with $\omega_1, ..., \omega_n \in \{S, F\}$. Let, for $i = 1, ..., n$, X_i be a rrv, depending only on the result in the $i-$th test. In other words, let us assume that

$$X_i(\omega_1, ..., \omega_i, ..., \omega_n) = f_i(\omega_i),$$

with $f_i(S) = \alpha_i$, $f_i(I) = \beta_i$. For the sake of simplicity, let us suppose that, for each $i = 1, ..., n$, $\alpha_i = 1$, $\beta_i = 0$. Let us check that the rrv-s $\{X_i : 1 \leq i \leq n\}$ are indipendent. We take (arbitrarily), $r \in \{1, \ldots, n\}$ and $1 \leq j_1 < ... < j_r \leq n$. For each $i \in \{1, \ldots, r\}$, let $B_i \in \mathcal{B}(\mathbb{R})$. We check that

$$P(X_{j_1} \in B_1, ..., X_{j_r} \in B_r) = P(X_{j_1} \in B_1) \cdot ... \cdot P(X_{j_r} \in B_r). \tag{5.4.32}$$

Let us observe, first of all, that one has, for each $i \in \{1, ..., r\}$:

$$\{X_{j_i} \in B_i\} = \begin{cases} \emptyset & \text{if } B_i \cap \{0,1\} = \emptyset, \\ \{\omega \in \Omega : \omega_{j_i} = S\} & \text{if } B_i \cap \{0,1\} = \{1\}, \\ \{\omega \in \Omega : \omega_{j_i} = F\} & \text{if } B_i \cap \{0,1\} = \{0\}, \\ \Omega & \text{if } \{0,1\} \subseteq B_{j_i}. \end{cases} \tag{5.4.33}$$

So, if, as usual, we indicate with p the probability of S in a single test and with q the probability of F, we have, for each $i = 1, ..., r$:

$$P(X_{j_i} \in B_i) = \begin{cases} 0 & \text{if } B_i \cap \{0,1\} = \emptyset, \\ p & \text{if } B_i \cap \{0,1\} = \{1\}, \\ q & \text{if } B_i \cap \{0,1\} = \{0\}, \\ 1 & \text{if } \{0,1\} \subseteq B_i. \end{cases} \tag{5.4.34}$$

Next, we observe that, if, for some i, $B_i \cap \{0,1\} = \emptyset$, the second tern in (5.4.32) is zero. In this case, we have also

$$\{X_{j_1} \in B_1\} \cap ... \cap \{X_{j_r} \in B_r\} = \emptyset,$$

so that (5.4.32) holds. Let us suppose, on the contrary, that for each i, $B_i \cap \{0,1\} \neq \emptyset$. We set:

$$\mathcal{I}_1 := \{i \in \{1, ..., r\} : B_i \cap \{0,1\} = \{1\}\},$$

$$\mathcal{I}_2 := \{i \in \{1, ..., r\} : B_i \cap \{0,1\} = \{0\}\},$$

$$\mathcal{I}_3 := \{i \in \{1, ..., r\} : \{0,1\} \subseteq B_i\}.$$

The product in the second term of (5.4.32) gives $p^{\sharp(\mathcal{I}_1)} q^{\sharp(\mathcal{I}_2)}$. For convenience, we put

$$E := \{X_{j_1} \in B_1\} \cap ... \cap \{X_{j_r} \in B_r\}.$$

We observe that E consists of the elements $\omega = (\omega_1, \ldots, \omega_n)$ of Ω such that $\omega_{j_i} = S$ if $i \in \mathcal{I}_1$, $\omega_{j_i} = F$ if $i \in \mathcal{I}_2$. The number of "successes" in ω is greater of equal to $\sharp(\mathcal{I}_1)$ and does not exceed $n - \sharp(\mathcal{I}_2)$. Moreover, if $\sharp(\mathcal{I}_1) \leq h \leq n - \sharp(\mathcal{I}_2)$, each element of E with exactly h "successes" can be associated to a subset of $\{1, \ldots, n\} \setminus \{j_i : i \in \mathcal{I}_1 \cup \mathcal{I}_2\}$ with $h - \sharp(\mathcal{I}_1)$ elements, precisely $\{j \in \{1, \ldots, n\} \setminus \{j_i : i \in \mathcal{I}_1 \cup \mathcal{I}_2\} : \omega_j = S\}$. We deduce that

$$P(E) = \textstyle\sum_{h=\sharp(\mathcal{I}_1)}^{n-\sharp(\mathcal{I}_2)} \binom{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)}{h-\sharp(\mathcal{I}_1)} p^h q^{n-h}$$

$$= \textstyle\sum_{k=0}^{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)} \binom{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)}{k} p^{\sharp(\mathcal{I}_1)+k} q^{n-\sharp(\mathcal{I}_1)-k}$$

$$= p^{\sharp(\mathcal{I}_1)} q^{\sharp(\mathcal{I}_2)} \textstyle\sum_{k=0}^{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)} \binom{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)}{k} p^k q^{n-\sharp(\mathcal{I}_1)-\sharp(\mathcal{I}_2)-k}$$

$$= p^{\sharp(\mathcal{I}_1)} q^{\sharp(\mathcal{I}_2)}.$$

We conclude this section with some important properties of independent families of random variables.

Theorem 5.4.48. *Let X_1,..., X_n be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$, with densities, respectively, f_1,...,f_n. Then the $n-$dimensional random variable $X = (X_1, ..., X_n)$ has the density*

$$(f_1 \otimes ... \otimes f_n)(x_1, ..., x_n) := f_1(x_1) \cdot ... \cdot f_n(x_n). \tag{5.4.35}$$

Incomplete proof We consider only the case $n = 2$. It is possible to show that $f_1 \otimes f_2$ is measurable in $\mathbb{R}^2$. From the theorem of Tonelli, it immediately follows that its integral is one. Let $B \in \mathcal{B}(\mathbb{R}^2)$. We consider only the case

$$B = B_1 \times B_2,$$

with B_1 and B_2 Borel subsets of $\mathbb{R}$. Then one has, employing the fact that X_1 and X_2 are independent and, again, the theorem of Tonelli,

$$\begin{aligned} P(X \in B) = \ & P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1) \cdot P(X_2 \in B_2) \\ & = \textstyle\int_{B_1} f_1(x_1)dx_1 \cdot \int_{B_2} f_2(x_2)dx_2 \\ & = \textstyle\int_B (f_1 \otimes f_2)(x_1, x_2)dx_1dx_2. \end{aligned}$$

□

Theorem 5.4.49. *Let X_1,..., X_n be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$, with densities f_1,...,f_n respectively. Then the rrv $X_1 + ... + X_n$ has the density $f_1 * ... * f_n$ (see Exercise 4.4.13 for the convolution of more than two functions).*

Incomplete proof We treat only the case $n = 2$. By virtue of Remark 5.4.26, it suffices to show that $\forall t \in \mathbb{R}$,

$$P(X_1 + X_2 \leq t) = \int_{]-\infty,t]} (f_1 * f_2)(x)dx. \tag{5.4.36}$$

We set $B_t := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \leq t\}$. $B_t \in \mathcal{B}(\mathbb{R}^2)$ because it is closed. Then, applying Theorem 5.4.48 and the theorem ofi Tonelli, we have

$$\begin{aligned} P(X_1 + X_2 \leq t) \ & = P((X_1, X_2) \in B_t) = \textstyle\int_{B_t} f_1(x_1)f_2(x_2)dx_1dx_2 \\ & = \textstyle\int_{\mathbb{R}} (\int_{]-\infty,t-x_2]} f_1(x_1)dx_1) f_2(x_2)dx_2 \\ & = \textstyle\int_{\mathbb{R}} (\int_{]-\infty,t]} f_1(x - x_2)dx) f_2(x_2)dx_2 \\ & = \textstyle\int_{]-\infty,t]} (\int_{\mathbb{R}} f_1(x - x_2)f_2(x_2)dx_2)dx \\ & = \textstyle\int_{]-\infty,t]} (f_1 * f_2)(x)dx. \end{aligned}$$

□

Example 5.4.50. We consider again Example 5.4.28 (which was based on Examples 5.1.2 and 5.1.10). Suppose that, when the device stops working, a second one starts, completely similar to the first. We indicate with Y the duration of the second device, which is governed by the same distribution law. Then Y has the density $f(t) = \lambda e^{-\lambda t}\chi_+(t)$, with χ_+ characteristic function of $[0, +\infty[$. Let us suppose that the durations of the two devices are independent. In this case, it is natural to assume that X and Y are independent rrv-s. Let us consider the rrv $X + Y$, indicating the total duration of the two devices. Then, by Theorem 5.4.49, $X+Y$ has the density $g := f * f$. If $t \in \mathbb{R}$,

$$\begin{aligned} (f * f)(t) &= \lambda^2 \int_{\mathbb{R}} e^{-\lambda(t-s)}\chi_+(t-s)e^{-\lambda s}\chi_+(s)ds \\ &= \lambda^2 t e^{-\lambda t}\chi_+(t). \end{aligned} \tag{5.4.37}$$

Theorem 5.4.51. *Let X and Y be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$. Let us suppose that each of them has a defined expectation. Then also XY has a defined expectation and $E(XY) = E(X)E(Y)$.*

Incomplete proof Suppose that X e Y have densities, respectively, f and g. The case of X and Y simple random variables is treated in Exercise 5.4.60. We set $Z : \Omega \to \mathbb{R}^2$, $Z(\omega) = (X(\omega), Y(\omega))$. By Theorem 5.4.48, Z has the density $f \otimes g$. Then, by the theorem of Fubini, as $XY = F \circ Z$, with $F : \mathbb{R}^2 \to \mathbb{R}$, $F(x_1, x_2) = x_1 x_2$, one has, applying Theorem 5.4.40,

$$\begin{aligned} E(XY) &= \int_{\mathbb{R}^2} x_1 x_2 f(x_1) g(x_2) dx_1 dx_2 \\ &= \int_{\mathbb{R}} x_1 f(x_1) dx_1 \cdot \int_{\mathbb{R}} x_2 g(x_2) dx_2 = E(X)E(Y). \end{aligned}$$

□

Theorem 5.4.52. *Let $X_1, ..., X_n$ be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$. Suppose that, for each of them, the variance $\sigma^2(X_j)$ is defined.Then it is also defined the variance of $X_1 + ... + X_n$ and*

$$\sigma^2(X_1 + ... + X_n) = \sigma^2(X_1) + ... + \sigma^2(X_n). \tag{5.4.38}$$

Incomplete proof We check (5.4.38) in case $n = 2$. One has, employing Theorem 5.4.51,

$$\begin{aligned} \sigma^2(X_1 + X_2) &= E((X_1 + X_2)^2) - E(X_1 + X_2)^2 \\ &= E(X_1^2 + 2X_1X_2 + X_2^2) - [E(X_1) + E(X_2)]^2 \\ &= E(X_1^2) + 2E(X_1)E(X_2) + E(X_2^2) - E(X_1)^2 \\ &\quad -2E(X_1)E(X_2) - E(X_2)^2 \\ &= \sigma^2(X_1) + \sigma^2(X_2). \end{aligned}$$

□

Exercise 5.4.53. Prove the following further properties of the distribution function F_X of a rrv X: let a and b be real numbers, with $a < b$; then:

(I) $P(a \leq X \leq b) = F_X(b) - F_X(a) + P(X = a)$;
(II) $P(a < X < b) = F_X(b) - F_X(a) - P(X = b)$;
(III) $P(a \leq X < b) = F_X(b) - F_X(a) - P(X = b) + P(X = a)$;
(IV) there exists $\lim_{t \to a^-} F_X(t)$ and coincides with $P(X < a)$;
(V) $F_X(a) - \lim_{t \to a^-} F_X(t) = P(X = a)$;
(VI) F_X is continuous in a if and only if $P(X = a) = 0$.

Exercise 5.4.54. Let us suppose of throwing a balanced dice until we get six, with independent results in different launches. Construct a probabilist model of this experiment. Let X be the random variable counting the number of launches. Compute $P(X > 10)$.

Exercise 5.4.55. Let X be a rrv. Check that $E(X)$ is defined in the sense of (5.4.26) if and only if $E(|X|) < +\infty$.

Exercise 5.4.56. Prove Theorem 5.4.35 in the case that X e Y are simple random variables. Use the fact that it is always possible to split Ω into a finite number of events $\{A_j : 1 \leq j \leq n\}$, pairwise incompatible, in such a way that, in each A_j, X and Y are constant.

Exercise 5.4.57. Check that a constant rrv has variance zero.

Exercise 5.4.58. Compute the variance of the rrv in Example 5.4.4.

Exercise 5.4.59. Compute the variance of the rrv X in Example 5.4.37. Draw preliminarily, using again Theorem 3.3.15, the formula

$$\sum_{j=1}^{\infty} j^2 z^j = \frac{z + z^2}{(1-z)^3},$$

valid for every $z \in \mathbb{C}$, with $|z| < 1$.

Exercise 5.4.60. Let X and Y be simple and independent random variables in the probability space a $(\Omega, \mathcal{A}, P)$. Check that $E(XY) = E(X)E(Y)$.

(Hint: suppose that $\Omega = \cup_{j=1}^m A_j$, with the events A_j pairwise incompatible, and that $X(\omega) = \alpha_j$ if $\omega \in A_j$. We may assume that for each $j = 1, ..., m$, $A_j = \{X = \alpha_j\}$. We observe that $X = \sum_{j=1}^{m} \alpha_j \chi_{A_j}$. Analogously, let $\Omega = \cup_{k=1}^n B_k$, with the events B_k pairwise incompatible and, for each $k = 1, ..., n$, $B_k = \{Y = \beta_k\}$. Observe that $Y = \sum_{k=1}^{n} \beta_k \chi_{B_k}$. Then one has

$$XY = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k \chi_{A_j} \chi_{B_k} = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k \chi_{A_j \cap B_k}.$$

On account of

$$P(X = \alpha_j, Y = \beta_k) = P(X = \alpha_j) P(Y = \beta_k),$$

because X and Y are independent, one has

$$\begin{aligned} E(XY) &= \textstyle\sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k E(\chi_{A_j \cap B_k}) \\ &= \textstyle\sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k P(X = \alpha_j, Y = \beta_k) \\ &= \textstyle\sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j \beta_k P(X = \alpha_j) P(Y = \beta_k) \\ &= \textstyle\sum_{j=1}^{m} \alpha_j P(A_j) \cdot \sum_{k=1}^{n} \beta_k P(B_k) \\ &= E(X) \cdot E(Y).) \end{aligned}$$

Exercise 5.4.61. Let X and Y be simple and independent random variables in the probability space $(\Omega, \mathcal{A}, P)$. Check that

$$\sigma^2(X + Y) = \sigma^2(X) + \sigma^2(Y).$$

Exercise 5.4.62. The final step in a long calculation requires the computation of the sum of three integers X_1, X_2, X_3. Suppose the following:

(a) the computations of X_1, X_2, X_3 are independent;

(b) in the calculation of each X_i $(1 \leq i \leq 3)$, the probability that it is correct is $p(\in]0, 1[)$;

(c) a possible mistake may be only a difference of one in excess or in default;

(d) the probability of a mistake by excess coincides with the probability of a mistake by default.

Calculate the probability that $X_1+X_2+X_3$ is correct (on account of possible compensations).

Exercise 5.4.63. A rrv X has a Poisson distribution if it can take only nonnegative integer values and, $\forall k \in \mathbb{N}_0$,

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!},$$

for some $\lambda \in \mathbb{R}^+$. Compute $E(X)$ e $\sigma^2(X)$.

Exercise 5.4.64. A rrv X admits a uniform density in $[a, b]$ $(a, b \in \mathbb{R},\ a < b)$, if such density is constant in $[a, b]$ and zero outside it. In such a case, compute $E(X)$ and $\sigma^2(X)$.

Exercise 5.4.65. Compute the variance of a rrv with exponential density $f(t) = \lambda e^{-\lambda t}\chi_+(t)$, with $\lambda \in \mathbb{R}^+$.

Exercise 5.4.66. Let X be a rrv such that $E(X^2) < +\infty$, and let a,$b \in \mathbb{R}$, $Y := aX + b$. Check that

$$E(Y) = aE(X) + b, \tag{5.4.39}$$

$$\sigma^2(Y) = a^2\sigma^2(X). \tag{5.4.40}$$

Exercise 5.4.67. Let $m \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. We set

$$\begin{cases} f : \mathbb{R} \to \mathbb{R}, \\ f(t) = \frac{e^{-(t-m)^2/(2\sigma^2)}}{\sqrt{2\pi}\sigma}. \end{cases} \tag{5.4.41}$$

Check the following:

(I) $\forall m \in \mathbb{R}$, $\forall \sigma \in \mathbb{R}^+$, $\int_{\mathbb{R}} f(t)dt = 1$;

(II) if X is a rrv, admitting the density f, then $E(X) = m$ and $\sigma^2(X) = \sigma^2$.

In such a case, we shall say that X is a rrv with **normal** distribution.

Exercise 5.4.68. Let X be a rrv with uniform density in $[0, 1]$. Check that the following rrv-s admit a density and compute it:

(I) $3X + 1$;

(II) X^2;

(III) e^X.

Exercise 5.4.69. Let X and Y be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$, with density $t \to 2e^{-2t}\chi_{\mathbb{R}^+}$. Determine the distribution function of $2X - Y$.

(Hint: observe that, for every $t \in \mathbb{R}$, $\{\omega \in \Omega : 2X(\omega) - Y(\omega) \leq t\} = \{\omega \in \Omega : (X(\omega), Y(\omega)) \in \{(x_1, x_2) \in \mathbb{R}^2 : 2x_1 - x_2 \leq t\}\}$).

5.5 Law of large numbers and central limit theorem

In this section we illustrate two important results of asymptotic nature, that is, results describing what happens when a certain natural parameter n goes to $+\infty$. Their interest and usefulness lies in the fact that they allow, in presence of a large number of tests, to predict, with a reasonable margin of error, a certain average behavior.

We begin with a very simple, but often quite useful, inequality.

Theorem 5.5.1. *(Chebyscev inequality) Let X be a nonnegative rrv in the probability space $(\Omega, \mathcal{A}, P)$, with $E(X) < +\infty$. Then, $\forall \epsilon \in \mathbb{R}^+$,*

$$P(X \geq \epsilon) \leq E(X)/\epsilon.$$

Proof Consider the following simple random variable:

$$\begin{cases} Y : \Omega \to \mathbb{R}, \\ Y(\omega) = \begin{cases} \epsilon & \text{if } X(\omega) \geq \epsilon, \\ 0 & \text{otherwise.} \end{cases} & \omega \in \Omega. \end{cases}$$

Then $Y(\omega) \leq X(\omega)$ $\forall \omega \in \Omega$. It follows that

$$\epsilon P(X \geq \epsilon) = E(Y) \leq E(X),$$

hence we get the conclusion.

□

We pass to the following basic fact:

Lemma 5.5.2. *Let X and Y be rrv-s with the same distribution law (not necessarily defined in the same probability space). Then, if X admits expectation, the same happens for Y and $E(X) = E(Y)$. If X admits variance, the same happens for Y and $\sigma^2(X) = \sigma^2(Y)$.*

Incomplete proof We limit ourselves to observe that, in case X (and so Y) admits a density, the conclusion follows from Theorem 5.4.38 and from Example 5.4.45.

□

We pass to the first important result of asymptotic nature.

Theorem 5.5.3. *(Weak law of large numbers) Let X_1,...,X_n be independent rrv-s in the probability space $(\Omega, \mathcal{A}, P)$. Suppose that they admit the same expectation μ and the same variance σ^2. We set*

$$\overline{X}_n := (X_1 + ... + X_n)/n. \tag{5.5.1}$$

Then, $\forall \epsilon > 0$,

$$P(|\overline{X}_n - \mu| \geq \epsilon) \leq \sigma^2/(n\epsilon^2). \tag{5.5.2}$$

Proof We observe, firstly, that

$$E(\overline{X}_n) = \mu.$$

So, by Chebyscev inequality, (Theorem 5.5.1), one has

$$P(|\overline{X}_n - \mu| \geq \epsilon) = P((\overline{X}_n - \mu)^2 \geq \epsilon^2) \leq \sigma^2(\overline{X}_n)/\epsilon^2. \tag{5.5.3}$$

From the result of Exercise 5.4.66 and from Theorem 5.4.52, we have

$$\sigma^2(\overline{X}_n) = \sigma^2(X_1 + ... + X_n)/n^2 = \sigma^2/n. \tag{5.5.4}$$

Then the conclusion follows from (5.5.3) and (5.5.4).

□

Remark 5.5.4. The intuitive meaning of the law of large numbers is the following: the rrv-s $\overline{X}_n$ tend to coincide with the expectation μ (which is a real number) as $n \to +\infty$, if the expectations and the variances of the independent random variables X_n remain unchanged. Observe that $\overline{X}_n$ is an "average" of X_1,...,X_n.

Example 5.5.5. Let us consider the repeated launch of a balanced dice. Let, for every $n \in \mathbb{N}$, S_n be the number of times that we get 6 in n launches. We want to estimate $P(252 < S_{1764} < 336)$.

Let $\Omega := \{(\omega_1, ..., \omega_{1764}) : \omega_j \in \{1, 2, 3, 4, 5, 6\}\}$. We indicate, for each $j \in \{1, ..., 1764\}$, with X_j the rrv such that $X_j(\omega) = 1$ if $\omega_j = 6$, $X_j(\omega) = 0$ otherwise. Evidently, $S_{1764} = \sum_{j=1}^{1764} X_j$. For each $j = 1, ..., 1764$, the $X_j - s$ are independent random variables (by Example 5.4.47) with the same distribution law. In fact, they are simple, with $P(X_j = 1) = 1/6$, $P(X_j = 0) = 5/6$, hence $E(X_j) = 1/6$. Moreover

$$\sigma^2(X_j) = E(X_j^2) - (1/6)^2 = E(X_j) - 1/36 = 1/6 - 1/36 = 5/36.$$

Employing the notations in Theorem 5.5.3, we have $\overline{X}_{1764} = S_{1764}/1764$. Then, for every $\epsilon \in \mathbb{R}^+$, in force of the weak law of large numbers, one has

$$P(|\overline{X}_{1764} - 1/6| \geq \epsilon) \leq \frac{5/36}{1764\epsilon^2} \tag{5.5.5}$$

Now,

$$\begin{aligned} \{252 < S_{1764} < 336\} &= \{\tfrac{1}{7} < \overline{X}_{1764} < \tfrac{4}{21}\} \\ &= \{-\tfrac{1}{42} < \overline{X}_{1764} - \tfrac{1}{6} < \tfrac{1}{42}\} \end{aligned}$$

So, using (5.5.5) with $\epsilon = 1/42$, we obtain

$$\begin{aligned} P(252 < S_{1764} < 336) &= P(|\overline{X}_{1764} - \tfrac{1}{6}| < \tfrac{1}{42}) \\ &= 1 - P(|\overline{X}_{1764} - \tfrac{1}{6}| \geq \tfrac{1}{42}) \\ &\geq 1 - \tfrac{42^2 \times 5/36}{1764} = \tfrac{31}{36} \cong 0,86. \end{aligned}$$

Now we pass to illustrate the second important result of this section, the so called **central limit theorem**. We begin with some definitions and remarks.

Definition 5.5.6. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of rrv-s, not necessarily defined in the same probability space. We indicate with F_n and F respectively the distribution functions of X_n and X. Then we shall say that the sequence $(X_n)_{n\in\mathbb{N}}$ converges in law to X, and we shall write*

$$X_n \overset{L}{\to} X \tag{5.5.6}$$

if $\forall t \in \mathbb{R}$ such that F is continuous in t, one has

$$\lim_{n\to\infty} F_n(t) = F(t).$$

Remark 5.5.7. As F is nondecreasing (Theorem 5.4.17 (III)), it is possible to prove that the set of its discontinuity points is, at most, countable. We recall (Exercise 1.1.8) that a countable set has Lebesgue measure zero. So, convergence in law implies pointwise convergence almost everywhere of the sequence of the distribution functions $(F_n)_{n\in\mathbb{N}}$.

We observe also that, if X and Y are rrv-s with the same distribution law and

$$X_n \overset{L}{\to} X ,$$

holds, even

$$X_n \overset{L}{\to} Y .$$

is true.

Definition 5.5.8. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $X : \Omega \to \mathbb{C}$. We shall say that X is a* **complex random variable** *(crv) if $Re(X)$ and $Im(X)$ are rrv-s.*

If $E(Re(X))$ and $E((Im(X))$ are well defined and real, we call **expectation** *of X the complex number*

$$E(X) := E(Re(X)) + iE(Im(X)). \tag{5.5.7}$$

Definition 5.5.9. *Let X be a rrv in the probability space $(\Omega, \mathcal{A}, P)$. We define the* **characteristic function** *ϕ_X of X as*

$$\begin{cases} \phi_X : \mathbb{R} \to \mathbb{C}, \\ \phi_X(\xi) = E(e^{i\xi X}), \quad \xi \in \mathbb{R}. \end{cases} \tag{5.5.8}$$

Remark 5.5.10. By virtue of Theorem 5.4.40, for every $\xi \in \mathbb{R}$, the function with domain Ω $\omega \to e^{i\xi X(\omega)}$ is a crv, because

$$e^{i\xi X(\omega)} = \cos(\xi X(\omega)) + i \sin(\xi X(\omega)).$$

Obviously, as sin and cos are bounded, the expectations $E(\cos(\xi X))$ and $E(\sin(\xi X))$ are well defined. It follows that any rrv admits its characteristic function.

By the way, the expression "characteristic function" does not seem particularly appropriate, as it is also used to indicate the function $f(x)$ such that $f(x) = 1$ in some subset of the domain, $f(x) = 0$ otherwise. However, in probability texts the meaning is as in Definition 5.5.8. In these texts, characteristic functions in the older sense are named "indicatrix functions".

Remark 5.5.11. If X is a rrv admitting the density f, one has, $\forall \xi \in \mathbb{R}$, applying Theorem 5.4.40,

$$\begin{aligned} \phi_X(\xi) &= E(e^{i\xi X}) = E(\cos(\xi X)) + iE(\sin(\xi X)) \\ &= \int_{\mathbb{R}} \cos(t\xi) f(t)dt + i \int_{\mathbb{R}} \sin(t\xi) f(t)dt = \int_{\mathbb{R}} e^{it\xi} f(t)dt \\ &= \hat{f}(-\xi), \end{aligned}$$

where we have indicated with $\hat{f}$ the Fourier transform of f.

Here we state, without proof, a very useful result, connecting convergence in law to pointwise convergence of characteristic functions.

Theorem 5.5.12. *(P. Levy) Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of rrv-s, not necessarily defined in the same probability space, and X a rrv. Then, the following are equivalent:*

(I)

$$X_n \overset{L}{\to} X;$$

(II) $\forall \xi \in \mathbb{R}$, one has

$$\lim_{n\to+\infty} \phi_{X_n}(\xi) = \phi_X(\xi).$$

Definition 5.5.13. *A rrv X admits a* **standard normal distribution** *if it admits the density $f(t) = e^{-t^2/2}/\sqrt{2\pi}$ $(t \in \mathbb{R})$.*

Remark 5.5.14. Applying the result in Exercise 5.4.67, we can say that, if X admits standard normal distribution, it has expectation zero and variance one.

Now we are able to state and partially prove the following classical result:

Theorem 5.5.15. *(Central limit theorem) Let, for $n \in \mathbb{N}$, $(\Omega_n, \mathcal{A}_n, P_n)$ be a sequence of probability spaces, X_{n1},...,X_{nn} rrv-s with domain Ω_n. Suppose that:*

(I) for every $n \in \mathbb{N}$, $k \in \{1, ..., n\}$, the rrv-s X_{nk} have all the same distribution law, with expectation $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}^+$ (independent of n and k, on account of Lemma 5.5.2);

(II) for every $n \in \mathbb{N}$, the rrv-s X_{n1},...,X_{nn} are independent.

We set, again for $n \in \mathbb{N}$,

$$S_n^* := (X_{n1} + ... + X_{nn} - n\mu)/(\sigma\sqrt{n}), \tag{5.5.9}$$

with $\sigma := \sqrt{\sigma^2}$.

Then, the sequence $(S_n^*)_{n\in\mathbb{N}}$ converges in law to a rrv with standard normal distribution.

Incomplete proof We prove the theorem adding the further (unnecessary) assumption that the rrv-s X_{nk} $(n \in \mathbb{N}, k \in \{1, ..., n\})$ admit a density f.

We set, for $t \in \mathbb{R}$, $g(t) := e^{-t^2/2}/\sqrt{2\pi}$. It can be easily checked that, for every $\xi \in \mathbb{R}$, one has

$$\hat{g}(\xi) = e^{-\xi^2/2}.$$

Then, on account of the theorem of P. Levy, and of Remark 5.5.11, we can try to show that, $\forall \xi \in \mathbb{R}$,

$$\lim_{n\to+\infty} \phi_{S_n^*}(\xi) = e^{-\xi^2/2}. \tag{5.5.10}$$

holds. We set, for $n \in \mathbb{N}$,

$$S_n := X_{n1} + ... + X_{nn}. \tag{5.5.11}$$

By Theorem 5.4.49, S_n admits the density

$$f_n := f * ... * f \quad (n \text{ factors }). \tag{5.5.12}$$

We check that S_n^* admits the density

$$f_n^*(t) := \sqrt{n}\sigma f_n(\sqrt{n}\sigma t + n\mu), t \in \mathbb{R}. \tag{5.5.13}$$

In fact, for every $t \in \mathbb{R}$,

$$\begin{aligned} P_n(S_n^* \leq t) &= P_n(S_n \leq n\mu + \sqrt{n}\sigma t) = \int_{]-\infty, n\mu+\sqrt{n}\sigma t]} f_n(s)ds \\ &= \int_{]-\infty,t]} \sqrt{n}\sigma f_n(\sqrt{n}\sigma s + n\mu)ds. \end{aligned}$$

Therefore, applying again Remark 5.5.11, one has, $\forall \xi \in \mathbb{R}$,

$$\begin{array}{ll} \phi_{S_n^*}(\xi) & = \sqrt{n}\sigma \int_{\mathbb{R}} e^{it\xi} f_n(\sqrt{n}\sigma t + n\mu)dt = \int_{\mathbb{R}} e^{i(s-n\mu)\xi/(\sqrt{n}\sigma)} f_n(s)ds \\ & = e^{-i\sqrt{n}\mu\xi/\sigma} \hat{f}(-\xi/(\sqrt{n}\sigma))^n, \end{array} \tag{5.5.14}$$

where in the final passage we have used Theorem 4.4.11 (III), on account of which, for every $\eta \in \mathbb{R}$, $\hat{f}_n(\eta) = \hat{f}(\eta)^n$. So we have to show that, for every $\xi \in \mathbb{R}$,

$$\lim_{n\to+\infty} e^{-i\sqrt{n}\mu\xi/\sigma} \hat{f}(-\xi/(\sqrt{n}\sigma))^n = e^{-\xi^2/2}. \tag{5.5.15}$$

To this aim, we begin by examining the Fourier transform $\hat{f}$. As the rrv-s X_{nk} admit expectation and variance, by Theorem 5.4.40, the functions $t^j f$ $(j \in \{0,1,2\})$ are summable in $\mathbb{R}$. Then it follows, by Corollary 4.1.12 (II), that $\hat{f}$ is of class C^2. So we consider the Taylor expansion of $\hat{f}$ around 0 (in general, $\hat{f}$ is complex valued, but a result which is analogous to Theorem 3.7.1 in "Analisi matematica A" holds even in this case). So, we have

$$\hat{f}(\eta) = \hat{f}(0) + \hat{f}'(0)\eta + \hat{f}''(0)\eta^2/2 + r(\eta),$$

with $r(\eta) = o(\eta^2)$ as $\eta \to 0$. As f is a density, we have

$$\hat{f}(0) = \int_{\mathbb{R}} f(t)dt = 1. \tag{5.5.16}$$

From Corollary 4.1.12 and from Theorem 5.4.40, we obtain also

$$\hat{f}'(0) = -i \int_{\mathbb{R}} t f(t)dt = -i\mu \tag{5.5.17}$$

and

$$\begin{array}{ll} \hat{f}''(0) & = -\int_{\mathbb{R}} t^2 f(t)dt = -E(X_{nk}^2) = -[E(X_{nk}^2) - \mu^2] - \mu^2 \\ & -\sigma^2 - \mu^2. \end{array} \tag{5.5.18}$$

Then, from (5.5.16)-(5.5.18), it follows that

$$\hat{f}(\eta) = 1 - i\mu\eta - (\sigma^2 + \mu^2)\eta^2/2 + r(\eta), \tag{5.5.19}$$

with $r(\eta) = o(\eta^2)$ as $\eta \to 0$ and, for a fixed $\xi \in \mathbb{R}$,

$$\hat{f}(-\xi/(\sqrt{n}\sigma)) = 1 + i\mu\xi/(\sqrt{n}\sigma) - (\sigma^2 + \mu^2)\xi^2/(2n\sigma^2) + o(n^{-1})(n \to +\infty). \tag{5.5.20}$$

Now we indicate with log the logarithm function with domain $\mathbb{C}\setminus]-\infty, 0]$, such that

$$\log(z) = \ln(|z|) + iArg(z),$$

with $Arg(z) \in]-\pi, \pi[\cap arg(z)$. log admits the Taylor expansion

$$\log(1+z) = z - z^2/2 + o(z^2)(z \to 0). \tag{5.5.21}$$

So, as $\lim_{n\to+\infty} \hat{f}(-\xi/(\sqrt{n}\sigma)) = 1$, from (5.5.20)-(5.5.21) we obtain, as $n \to +\infty$,

$$\begin{array}{c} \log(\hat{f}(-\xi/(\sqrt{n}\sigma))) \\ = \log(1 + i\mu\xi/(\sqrt{n}\sigma) - (\sigma^2 + \mu^2)\xi^2/(2n\sigma^2) + o(n^{-1})) \\ = i\mu\xi/(\sqrt{n}\sigma) - (\sigma^2 + \mu^2)\xi^2/(2n\sigma^2) \\ -[i\mu\xi/(\sqrt{n}\sigma) + o(n^{-1/2})]^2/2 + o(n^{-1}) \\ = i\mu\xi/(\sqrt{n}\sigma) - \xi^2/(2n) + o(n^{-1}). \end{array} \tag{5.5.22}$$

Therefore,

$$e^{-i\sqrt{n}\mu\xi/\sigma}\hat{f}(-\xi/(\sqrt{n}\sigma))^n \begin{aligned}[t] &= e^{-i\sqrt{n}\mu\xi/\sigma}e^{n\log(\hat{f}(-\xi/(\sqrt{n}\sigma))} \\ &= e^{-\xi^2/2+o(1)} \end{aligned} \tag{5.5.23}$$

which tends to $e^{-\xi^2/2}$ as $n \to +\infty$.

In this way (5.5.15) is proved. □

Remark 5.5.16. The central limit theorem is, at least at first sight, rather surprising, as the fact that S_n^* converges in law ro a rrv with standard normal distribution is largely independent of the distribution law of the rrv-s X_{nk}. As we shall see in the following examples, the interest of the theorem lies primarely in the fact that it is a rather efficient tool of calculation, thanks to the tables of the standard normal distribution, which can be found in many books. Such tables indicate the values of

$$\Phi(t) := (2\pi)^{-1/2}\int_{-\infty}^{t} e^{-s^2/2}ds \tag{5.5.24}$$

for t in a discrete subset of $\mathbb{R}^+$.

Remark 5.5.17. Under the assumptions of Theorem 5.5.15, it is possible to show that

$$\lim_{n\to+\infty} P(S_n^* \le t) = \Phi(t)$$

uniformly in $t \in \mathbb{R}$.

Example 5.5.18. An insurance company must establish the price X of a certain accident policy. Let us suppose that the company expects 10.000 customers and that the compensation (in case the accident really happens) is of 1.000 euros. We know that the estimated probability that a single customer has an accident is $6 \cdot 10^{-3}$. The company wants to fix the price X in such a way that the probability of making profits of, at least , 10.000 euros is not less than 9/10.

Let k be the number of customers who have an accident. If X is the price of the policy, the net profit is, evidently,

$$10.000 \cdot X - 1.000 \cdot k. \tag{5.5.25}$$

In order that such profit is what desired, the inequality

$$10.000 \cdot X - 1.000 \cdot k \ge 10.000, \tag{5.5.26}$$

must hold, that is,

$$k \le 10(X-1). \tag{5.5.27}$$

Let k_0 ($\in \{0, ..., 10.000\}$). Then the probability that the number of accidents does not overcome k_0 (we assume that accidents for distinct customers happen independently) is, on account of (5.3.7),

$$\sum_{k=0}^{k_0} \binom{10.000}{k} (6\cdot 10^{-3})^k (1 - 6\cdot 10^{-3})^{10.000-k}, \tag{5.5.28}$$

which is quite difficult to estimate. So we follow another way.

Let, for each $j \in \{1, ..., 10.000\}$, X_j be the rrv, the value of which is one if the $j-$th customer has an accident, zero otherwise. Let us suppose that the rrv-s X_j are independent, and with the same distribution law. Then, if k indicates the total number of accidents, we shall have

$$k = \sum_{j=1}^{10.000} X_j. \tag{5.5.29}$$

So, for every j,

$$\mu := E(X_j) = P(X_j = 1) = 6 \cdot 10^{-3}, \tag{5.5.30}$$

$$\sigma^2 := \sigma^2(X_j) = E(X_j^2) - E(X_j)^2 = 6 \cdot 10^{-3} - 3,6 \cdot 10^{-5} \cong 6 \cdot 10^{-3}. \tag{5.5.31}$$

We set

$$\sigma := \sqrt{\sigma^2} \cong 7,7 \cdot 10^{-2}. \tag{5.5.32}$$

Finally, we set

$$n := 10.000 \tag{5.5.33}$$

and

$$S_n^* := (k - n\mu)/(\sqrt{n}\sigma). \tag{5.5.34}$$

Then, if $k_0 \in \{0, ..., 10.000\}$,

$$\begin{aligned} P(k \leq k_0) \quad &= P(S_n^* \leq (k_0 - n\mu)/(\sqrt{n}\sigma)) \\ &\cong P(S_n^* \leq (k_0 - 60)/(100 \cdot 7,7 \cdot 10^{-2})) \\ &= P(S_n^* \leq (k_0 - 60)/7,7) \\ &:= P(S_n^* \leq k_1). \end{aligned} \tag{5.5.35}$$

From the central limit theorem we have that

$$P(S_n^* \leq k_1) \cong \Phi(k_1), \tag{5.5.36}$$

with Φ defined in (5.5.24). k_1 should be such that

$$\Phi(k_1) \geq 9/10. \tag{5.5.37}$$

This happens if

$$k_1 \geq 1,29, \tag{5.5.38}$$

which implies, on account of (5.5.35),

$$k_0 \geq 1,29 \cdot 7,7 + 60 \cong 69,9. \tag{5.5.39}$$

So, it should be, remembering (5.5.27),

$$10(X - 1) \geq 69,9, \tag{5.5.40}$$

that is,

$$X \geq 7,99. \tag{5.5.41}$$

Example 5.5.19. We throw a balanced coin 10.000 times. We estimate the probability that the number of "heads" is between 4,950 and 5,050.

Let, for each $j = 1, ..., 10.000$, X_j be a rrv, the value of which is one if the $j-$th launch gives "head", zero if the $j-$th launch gives "tail". It is reasonable to assume that the rrv-s X_j are independent. Moreover, they have the same distribution law, with expectation $\mu = 1/2$, variance $\sigma^2 = 1/4$. If $n = 10.000$, $S_n := \sum_{j=1}^n X_j$ is the total number of "heads". We want to estimate $P(4.950 < S_n \leq 5.050)$. We set, as usual,

$$S_n^* := (S_n - n\mu)/(\sqrt{n}\sigma) = (S_n - 5.000)/50. \tag{5.5.42}$$

Then, applying the central limit theorem, we have

$$\begin{aligned} P(4.950 < S_n \leq 5.050) \quad &= P(-1 < S_n^* \leq 1) = P(S_n^* \leq 1) - P(S_n^* \leq -1) \\ &\cong \Phi(1) - \Phi(-1). \end{aligned}$$

From the tables, we get

$$\Phi(1) \cong 0,8413.$$

Moreover, if $t \geq 0$, one has

$$\begin{aligned}\Phi(-t) &= (2\pi)^{-1/2} \int_{-\infty}^{-t} e^{-s^2/2} ds = (2\pi)^{-1/2} \int_{t}^{+\infty} e^{-s^2/2} ds \\ &= 1 - \Phi(t).\end{aligned}$$

So,

$$\Phi(-1) = 1 - \Phi(1) \cong 0,1587.$$

We conclude that

$$P(4.950 < S_n \leq 5.050) \quad \cong 0,8413 - 0,1587 = 0,6826.$$

Exercise 5.5.20. A balanced coin is thrown n times. We indicate with X_n the number of "heads" in these n launches. Determine n such that $P(0,4 < X_n < 0,6)$ is larger than $0,9$.

Exercise 5.5.21. Let, for $t \in \mathbb{R}$, $g(t) = e^{-t^2/2}/\sqrt{2\pi}$. Check that, for every $\xi \in \mathbb{R}$, one has

$$\hat{g}(\xi) = e^{-\xi^2/2}.$$

5.6 Markov chains

In this section we shall touch on an interesting type of discrete stochastic process: Markov chains. We could loosely say that a discrete stochastic process is a sequence $(X_n)_{n \in \mathbb{N}_0}$ of rrv-s, defined in some (fixed) probability space $(\Omega, \mathcal{A}, P)$. Intuitively, n stands for a discrete time parameter and $(X_n)_{n \in \mathbb{N}_0}$ describes the random evolution in time of a certain process.

We begin with the following

Definition 5.6.1. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, let $(X_n)_{n \in \mathbb{N}_0}$ be a sequence of rrv-s in $(\Omega, \mathcal{A}, P)$ and $S := \{s_1, ..., s_m\}$ be a finite subset of $\mathbb{R}$. We shall say that $(X_n)_{n \in \mathbb{N}_0}$ is a* **stationary Markov chain** *, with set of the states S, if the following conditions are fulfilled:*

(I) for every $n \in \mathbb{N}_0$ X_n has range in S;

(II) let j_1,...,j_p integers, with $0 \leq j_1 < ... < j_{p-1} < j_p$ ($p \in \mathbb{N}$, $p \geq 2$). If s_{i_1},...,s_{i_p} are elements of S (not necessarily pairwise distinct) and $P(X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) > 0$, then

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = P(X_{j_p} = s_{i_p} | X_{j_{p-1}} = s_{i_{p-1}}); \qquad (5.6.1)$$

(III) if s_i and s_j are elements of S (not necessarily distinct), $n \in \mathbb{N}$ and $P(X_n = s_j) > 0$,

$$P(X_{n+1} = s_i | X_n = s_j) = a_{ij},$$

with a_{ij} independent of n.

Remark 5.6.2. Concerning Definition 5.6.1, the set S is the set of the states that the system may take in its time evolution.

Item (II) in Definition 5.6.1 is the so called **Markov property**, which is the main feature of these processes. Its intuitive meaning is the following: that the knowledge of the state of the process in times preceding j_{p-1} (j_1,...,j_{p-2}) does not give more information, concerning the state of the process at time j_p, than the simple knowledge of the state of the process at time j_{p-1}.

Finally, item (III) represents the stationarity of the process, in the sense that the probability that $X_{n+1} = s_i$ if $X_n = s_j$ does not depend on n and so it does not change in time.

Remark 5.6.3. If the conditions (I)-(III) of Definition 5.6.1 are satisfied, we can associate with the process the matrix $m \times m$ $A = (a_{ij})_{1 \le i,j \le m}$, such that, if, at some time n, $P(X_n = s_j) > 0$, then

$$a_{ij} = P(X_{n+1} = s_i | X_n = s_j). \tag{5.6.2}$$

A is a **stochastic matrix**, that is, a matrix with real nonnegative terms, with the some of these terms in each column equal to one. In fact, if $P(X_n = s_j) > 0$, as

$$\cup_{1 \le i \le m}(\{X_{n+1} = s_i\} \cap \{X_n = s_j\}) = \{X_n = s_j\},$$

one has

$$\begin{aligned} \textstyle\sum_{i=1}^m a_{ij} &= \textstyle\sum_{i=1}^m P(X_{n+1} = s_i | X_n = s_j) = \sum_{i=1}^m \frac{P(X_{n+1}=s_i, X_n=s_j)}{P(X_n=s_j)} \\ &= 1. \end{aligned}$$

Example 5.6.4. (Player's ruin) Two guys play a series of games (of any type), in such a way that the results in the single matches do not influence each other. They are also independent of the initial situation. Suppose the following:

(I) the probability that the first player wins a single game is p ($\in [0,1]$), the probability that the second player wins a single game is $q = 1 - p$;

(II) in each game, the prize is one euro: each player puts one euro on the table, and the winner takes the total sum;

(III) the capitals of the player at the beginning are, respectively, a and b euros ($a, b \in \mathbb{N}_0$);

(III) the sequence of games stops if one of the two player loses all the money.

Let us indicate with X_n the capital of the first player after n games ($n \in \mathbb{N}_0$). X_n may take any value between 0 and $a + b$. So we put

$$S := \{0, ..., a + b\}. \tag{5.6.3}$$

We determine $P(X_0 = s_0, X_1 = s_1, \dots, X_n = s_n)$, for every $n \in \mathbb{N}_0$, with $s_0, \dots, s_n \in \{0, \dots, a + b\}$. It is natural to set

$$P(X_0 = s_0) = \begin{cases} 1 & \text{if } s_0 = a, \\ \\ 0 & \text{otherwise .} \end{cases} \tag{5.6.4}$$

Next, assume that we have determined $P(X_0 = s_0, X_1 = s_1, \dots, X_n = s_n)$. Then we set

$$\begin{aligned} &P(X_0 = s_0, X_1 = s_1, \dots, X_n = s_n, X_{n+1} = s_{n+1}) \\ &= P(X_0 = s_0, X_1 = s_1, \dots, X_n = s_n) \cdot a_{s_{n+1}, s_n}, \end{aligned} \tag{5.6.5}$$

with

$$a_{ij} = \begin{cases} 1 & \text{if } i = j = 0, \\ p & \text{if } 1 \le j \le a + b - 1, i = j + 1, \\ q & \text{if } 1 \le j \le a + b - 1, i = j - 1, \\ 1 & \text{if } i = j = a + b, \\ 0 & \text{otherwise.} \end{cases} \tag{5.6.6}$$

It is intuitively clear that $(X_n)_{n \in \mathbb{N}_0}$ is a Markov chain. Here we limit ourselves to check that, $\forall n \in \mathbb{N}$, $\forall i, j, k$ in $\{0, \dots, a + b\}$, if $P(X_n = i, X_{n+1} = j) > 0$, then

$$P(X_{n+2} = k | X_n = i, X_{n+1} = j) = P(X_{n+2} = k | X_{n+1} = j) = a_{kj}.$$

In fact, if $P(X_n = i, X_{n+1} = j) > 0$, we have also $P(X_{n+1} = j) > 0$ and

$$P(X_{n+1} = j) = \sum_{j_0=0}^{a+b} \cdots \sum_{j_n=0}^{a+b} P(X_0 = j_0, \ldots, X_n = j_n, X_{n+1} = j)$$

$$= \sum_{j_0=0}^{a+b} \cdots \sum_{j_n=0}^{a+b} P(X_0 = j_0, \ldots, X_n = j_n) a_{j,j_n}.$$

Analogously,

$$P(X_{n+1} = j, X_{n+2} = k) = \sum_{j_0=0}^{a+b} \cdots \sum_{j_n=0}^{a+b} P(X_0 = j_0, \ldots, X_n = j_n, X_{n+1} = j, X_{n+2} = k)$$

$$= \sum_{j_0=0}^{a+b} \cdots \sum_{j_n=0}^{a+b} P(X_0 = j_0, \ldots, X_n = j_n) a_{j,j_n} a_{kj}.$$

We deduce that

$$P(X_{n+2} = k | X_{n+1} = j) = \frac{P(X_{n+1} = j, X_{n+2} = k)}{P(X_{n+1} = j)} = a_{kj}.$$

On the other hand,

$$P(X_n = i, X_{n+1} = j) = \sum_{j_0=0}^{a+b} \cdots \sum_{j_{n-1}=0}^{a+b} P(X_0 = j_0, \ldots, X_{n-1} = j_{n-1}, X_n = i, X_{n+1} = j)$$

$$= \sum_{j_0=0}^{a+b} \cdots \sum_{j_{n-1}=0}^{a+b} P(X_0 = j_0, \ldots, X_{n-1} = j_{n-1}) a_{i,j_{n-1}} a_{j_i},$$

$$P(X_n = i, X_{n+1} = j, X_{n+2} = k)$$

$$= \sum_{j_0=0}^{a+b} \cdots \sum_{j_{n-1}=0}^{a+b} P(X_0 = j_0, \ldots, X_{n-1} = j_{n-1}, X_n = i, X_{n+1} = j, X_{n+2} = k)$$

$$= \sum_{j_0=0}^{a+b} \cdots \sum_{j_{n-1}=0}^{a+b} P(X_0 = j_0, \ldots, X_{n-1} = j_{n-1}) a_{i,j_{n-1}} a_{j_i} a_{kj},$$

so that, again,

$$P(X_{n+2} = k | X_n = i, X_{n+1} = j) = \frac{P(X_n = i, X_{n+1} = j, X_{n+2} = k)}{P(X_n = i, X_{n+1} = j)} = a_{kj}.$$

So we have a Markov chain. The terms of matrix A, descriibed in general in Remark 5.6.3, will be, of course, those in (5.6.6).

Now we ask some questions:

(I) what is the probability that the second player loses all the money?

(II) What is the probability that the first player loses all the money?

(III) What is the probability that the series of games is endless ?

The first case can be identifies with the event $A_a := \bigcup_{n\in\mathbb{N}_0}\{X_n = a + b\}$, the second with the event $B_a := \bigcup_{n\in\mathbb{N}_0}\{X_n = 0\}$, the third with $C_a := (A_a \cup B_a)^c$. We have written a as a subscript of A and B because, in order to answer questions (I)-(III), it will be convenient to consider the case that the money at disposal of the first player at the beginning is a generic $s \in \{0, ..., a + b\}$. The analogs of A_a, B_a and C_a will be A_s, B_s, C_s. Let us indicate with p_s, q_s, r_s the probabilities of A_s, B_s, C_s in the respective probability spaces (that is, when we assume that, at the beginning, the first player has a euros). We observe, first of all, that

$$\begin{array}{lll} p_0 = 0, & q_0 = 1, & r_0 = 0, \\ p_{a+b} = 1, & q_{a+b} = 0, & r_{a+b} = 0. \end{array} \tag{5.6.7}$$

Moreover,

$$p_s + q_s + r_s = 1 \quad \forall s \in \{0, ..., a+b\}. \tag{5.6.8}$$

Suppose now that we are in the case $X_0 = s \in \{1, ..., a+b-1\}$. We indicate with β the event "the first player wins the first game", with γ the event "the second player wins the first game" Then

$$\begin{aligned} p_s = P(A_s) &= P(A_s \cap \beta) + P(A_s \cap \gamma) = \tfrac{P(A_s \cap \beta)}{P(\beta)} P(\beta) + \tfrac{P(A_s \cap \gamma)}{P(\beta)} P(\gamma) \\ &= pP(A_s|\beta) + qP(A_s|\gamma). \end{aligned} \tag{5.6.9}$$

If the first player wins the first game, his capital goes from s to $s+1$. So we are in the same situation of the case that, at the beginning, the first player has $s+1$ euros. So, from (5.6.9) it is clear that, for each $s = 1, ..., a+b-1$, one has

$$p_s = p \cdot p_{s+1} + q \cdot p_{s-1}, \tag{5.6.10}$$

hence, as $p + q = 1$,

$$p(p_{s+1} - p_s) = q(p_s - p_{s-1}), 1 \le s \le a+b-1. \tag{5.6.11}$$

Suppose now that $0 < p < 1$. Then, from (5.6.11) it follows

$$p_{s+1} - p_s = \frac{q}{p}(p_s - p_{s-1}), 1 \le s \le a+b-1 \tag{5.6.12}$$

and so, from (5.6.7), for each $s = 1, ..., a+b-1$,

$$p_{s+1} - p_s = \frac{q}{p}(p_s - p_{s-1}) = (\frac{q}{p})^2(p_{s-1} - p_{s-2}) = ... = (\frac{q}{p})^s p_1. \tag{5.6.13}$$

From (5.6.13) and (5.6.7), we obtain

$$\begin{aligned} 1 = p_{a+b} - p_0 &= \textstyle\sum_{s=0}^{a+b-1}(p_{s+1} - p_s) = \sum_{s=0}^{a+b-1}(\frac{q}{p})^s p_1 \\ &= \begin{cases} (a+b)p_1 & \text{if } p = q = 1/2, \\ \frac{1-(q/p)^{a+b}}{1-q/p} p_1 & \text{if } p \neq q. \end{cases} \end{aligned} \tag{5.6.14}$$

It follows that

$$p_1 = \begin{cases} \frac{1}{a+b} & \text{if } p = q = 1/2, \\ \frac{1-q/p}{1-(q/p)^{a+b}} & \text{if } p \neq q \end{cases} \tag{5.6.15}$$

and, for each $s = 1, ..., a+b$,

$$\begin{aligned} p_s &= \textstyle\sum_{j=0}^{s-1}(p_{j+1} - p_j) = \sum_{j=0}^{s-1}(\frac{q}{p})^j p_1 \\ &= \begin{cases} \frac{s}{a+b} & \text{if } p = q = 1/2, \\ \frac{1-(q/p)^s}{1-(q/p)^{a+b}} & \text{if } p \neq q. \end{cases} \end{aligned} \tag{5.6.16}$$

Inverting the roles of the players, we get also

$$q_s = \begin{cases} \frac{a+b-s}{a+b} & \text{if } p = q = 1/2, \\ \frac{1-(p/q)^{a+b-s}}{1-(p/q)^{a+b}} & \text{if } p \neq q. \end{cases} \tag{5.6.17}$$

From (5.6.16)-(5.6.17), it immediately follows that, for each $s \in \{0, ..., a+b\}$,

$$p_s + q_s = 1, \tag{5.6.18}$$

hence

$$r_s = 0. \tag{5.6.19}$$

We come back to the general case, that is, we assume only that the conditions (I)-(II) of Definition 5.6.1. are satisfied. We consider the stochastic matrix A defined in (5.6.2). As it is a square matrix ($m \times m$), its powers A^l ($l \in \mathbb{N}_0$), obtained by the standard multiplication of matrixes, are well defined. We introduce also the following notation: if $l \in \mathbb{N}_0$, $1 \leq i, j \leq m$, we indicate with $a_{ij}^{(l)}$ the term of place (i, j) in A^l.Then the following holds:

Theorem 5.6.5. *Suppose that the assumptions of Definition 5.6.1 are satisfied. Let A be the matrix defined in (5.6.2). Next, let j_1,...,j_p be nonnegative integers, such that $0 \leq j_1 < ... < j_{p-1} < j_p$ ($p \geq 2$), s_{i_1},...,s_{i_p} elements of S (not necessarily pairwise distnct) and let $P(X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) > 0$. Finally, let $l = j_p - j_{p-1}$. Then*

$$P(X_{j_p} = s_{i_p} | X_{j_1} = s_{i_1}, ..., X_{j_{p-1}} = s_{i_{p-1}}) = a_{i_p, i_{p-1}}^{(l)}. \tag{5.6.20}$$

Incomplete proof By virtue of the Markov property, it is sufficient to consider the case $p = 2$. So we write j instead of j_1 and $j + l$ instead of j_2. We limit ourselves to prove the result in case $l = 2$. So we have

$$\begin{aligned} P(X_{j+2} = s_{i_2} | X_j = s_{i_1}) &= \frac{P(X_j = s_{i_1}, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} \\ &= \sum_{k=1}^{m} \frac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})}. \end{aligned} \tag{5.6.21}$$

Let k be such that $P(X_j = s_{i_1}, X_{j+1} = s_k) > 0$. Then

$$\begin{aligned} \frac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} &= \frac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1}, X_{j+1} = s_k)} \frac{P(X_j = s_{i_1}, X_{j+1} = s_k)}{P(X_j = s_{i_1})} \\ &= P(X_{j+2} = s_{i_2} | X_j = s_{i_1}, X_{j+1} = s_k) \\ &\times P(X_{j+1} = s_k | X_j = s_{i_1}) \\ &= a_{i_2, k} a_{k, i_1}. \end{aligned} \tag{5.6.22}$$

The identity in formula (5.6.22) holds also in case $P(X_j = s_{i_1}, X_{j+1} = s_k) = 0$. In fact, in this case,

$$\frac{P(X_j = s_{i_1}, X_{j+1} = s_k, X_{j+2} = s_{i_2})}{P(X_j = s_{i_1})} = 0,$$

while

$$\begin{aligned} a_{k, i_1} &= P(X_{j+1} = s_k | X_j = s_{i_1}) = \frac{P(X_j = s_{i_1}, X_{j+1} = s_k)}{P(X_j = s_{i_1})} \\ &= 0. \end{aligned}$$

From (5.6.21) we obtain immediately

$$P(X_{j+2} = s_{i_2} | X_j = s_{i_1}) = \sum_{k=1}^{m} a_{i_2, k} a_{k, i_1} = a_{i_2, i_1}^{(2)}. \tag{5.6.23}$$

□