Elements of

Functional Analysis

Davide Guidetti and Lorenzo Ntogramatzidis

-

Contents

1 Normed spaces **7**

1.1 Metric spaces . . . 9
1.2 Normed spaces . . . 15
1.2.1 Relation between metric and norm . . . 23
1.3 Basic topological concepts . . . 31
1.4 Convergence and limits . . . 47
1.4.1 Properties of limits . . . 54
1.4.2 Operations on limits . . . 58
1.5 Equivalent norms . . . 66
1.6 Sequences . . . 71
1.6.1 Sequences in $\overline{\mathbb{R}}$ . . . 74
1.6.2 Limit superior and limit inferior . . . 79
1.6.3 Sequences of functions . . . 86
1.7 Subsequences and Bolzano-Weierstrass theorem . . . 98
1.8 Continuity . . . 107
1.8.1 More on uniform convergence and uniform continuity . . . 114
1.9 Norms of matrices . . . 118
1.9.1 Operator norm . . . 120

2 Banach and Hilbert spaces **129**

2.1 Cauchy sequences . . . 129
2.2 Compact sets . . . 135
2.3 Banach spaces . . . 141
2.3.1 Banach spaces and series . . . 148
2.4 Linear operators . . . 157
2.5 Hilbert spaces . . . 164
2.5.1 Sesquilinear forms, quadratic forms and inner products . . . 165
2.5.2 Hilbert spaces: definition and properties . . . 173
2.6 Orthogonal projections in Hilbert spaces . . . 179
2.7 Separable Hilbert spaces . . . 185

3 Measure theory and Lebesgue integral **197**

3.1 Arithmetics in $[0, +\infty]$ . . . 198
3.1.1 Sequences in $[0, +\infty]$ . . . 199
3.1.2 Finite and infinite summations in $[0, +\infty]$ . . . 200
3.2 General considerations about measure . . . 205

3.3 σ-algebras and measure spaces . . . 208
3.3.1 n-dimensional intervals and their volume . . . 214
3.3.2 Hints on the Peano-Jordan measure . . . 222
3.3.3 Outer measure . . . 223
3.4 Lebesgue measure . . . 234
3.4.1 Null measure sets . . . 240
3.4.2 Almost everywhere and for almost all . . . 241
3.5 Introduction on Lebesgue integral . . . 242
3.6 Simple functions . . . 243
3.7 Measurable functions . . . 251
3.8 Integration of non-negative measurable functions . . . 263
3.8.1 Beppo Levi's monotone convergence theorem . . . 265
3.8.2 Fatou's lemma . . . 271
3.9 Signed functions . . . 272
3.10 Relation between Riemann and Lebesgue integral . . . 290
3.11 Reduction of integrals . . . 298
3.11.1 Change of variable theorem . . . 317
3.12 Integral convergence . . . 344
3.13 Differentiation of integral functions . . . 349
3.14 Complements A: Countable sets . . . 355

4 L^p spaces 365
4.1 Identification of measurable functions . . . 365
4.1.1 Operations on equivalent classes . . . 369
4.2 L^p spaces . . . 373
4.3 Completeness of L^p spaces . . . 381

5 Distributions 391
5.1 Support and test functions . . . 394
5.1.1 Density of $\mathcal{D}$ in L^p . . . 408
5.2 Locally summable functions . . . 412
5.3 Linear functionals associated with locally summable functions . . . 417
5.4 Definition of distribution . . . 426
5.4.1 Series . . . 435
5.4.2 Principal value of $1/x$ . . . 436
5.5 Operations on distributions . . . 440
5.5.1 Multiplication of a distribution by a function of class C^∞ . . . 440
5.5.2 Composition of distributions with diffeomorphisms in C^∞ . . . 444
5.5.3 Restriction to an open set . . . 457
5.5.4 Tensor product . . . 459
5.5.5 Derivative of distributions . . . 465
5.5.6 Primitives . . . 486
5.5.7 Derivative of the tensor product . . . 490
5.5.8 Support of a distribution . . . 491
5.6 Convolution . . . 501

5.7 Sufficient conditions for convolvability . . . 503
5.8 Convolution of distributions . . . 507
5.8.1 Derivative of the convolution . . . 514

6 Fourier transform **517**
6.1 Fourier transform in $L^1(\mathbb{R}^n, \mathbb{C})$ . . . 517
6.2 Properties of the Fourier transform . . . 522
6.3 Fourier transform and derivatives . . . 533
6.4 Regularity at infinity of the Fourier transform . . . 542
6.4.1 Fourier transform of the convolution . . . 542
6.5 Schwartz space . . . 543
6.5.1 Slowly increasing functions . . . 551
6.6 Tempered distributions . . . 555
6.7 Fourier transform of tempered distributions . . . 561
6.7.1 Derivatives of tempered distributions . . . 569
6.7.2 Inversion formula . . . 570
6.7.3 Convergence in $\mathcal{S}'(\mathbb{R}^n; \mathbb{C})$ . . . 573
6.7.4 Fourier transform of distributions with compact support . . . 576
6.8 Fourier transform in $L^2(\mathbb{R}^n; \mathbb{C})$ . . . 579
6.8.1 An application to partial differential equations . . . 582
6.9 Sobolev spaces . . . 583
6.9.1 Fourier transform of the convolution for tempered distributions . . . 585

7 Linear partial differential equations **589**
7.1 Fundamental solution of differential operators . . . 589
7.1.1 Existence and uniqueness of solutions . . . 591
7.1.2 Regularity of solutions . . . 592
7.2 Ordinary differential operator . . . 594
7.3 Heat equation . . . 598
7.4 Wave equation . . . 603

8 Laplace transform **607**
8.1 Definition of Laplace transform . . . 607
8.1.1 Gamma function and initial/final value theorems . . . 624
8.2 Laplace transform of distributions . . . 629
8.2.1 Inverse Laplace transform of rational functions . . . 649
8.2.2 Rational functions with real coefficients . . . 652
8.2.3 Residues . . . 654

9 Fourier series **669**
9.1 Introduction . . . 669
9.2 Other orthonormal systems . . . 680
9.3 Fourier series in several real variables . . . 682
9.4 An application . . . 687
9.5 Fourier series for distributions . . . 690

10 Solutions to the Exercises in Chapter 1 **693**

11 Solutions to the Exercises in Chapter 2 **743**

12 Solutions to the Exercises in Chapter 3 **779**

13 Solutions to the Exercises in Chapter 4 **791**

14 Solutions to the Exercises in Chapter 5 **803**

15 Solutions to the Exercises in Chapter 6 **841**

1 Normed spaces

In this chapter, we discuss two notions that play a fundamental role in mathematics: distance and magnitude. We can define the magnitude of a real number x as its absolute value $|x|$, and we can define the distance between two real numbers x and y as the absolute value of their difference $|x - y|$. For example, if $x > 0$ and $y < 0$ we are in the situation represented by the figures

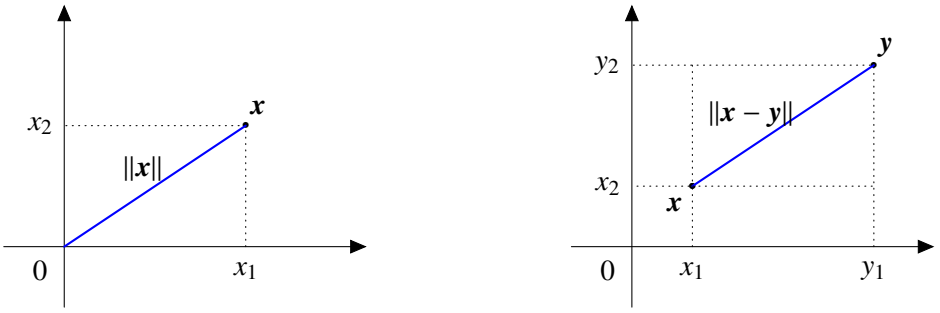
You have studied how the same idea can be used to define magnitude and distance in $\mathbb{R}^2$: given $\boldsymbol{x} = (x_1, x_2)$, the magnitude of $\boldsymbol{x}$ (usually referred to as "length") is its Euclidean norm, given by

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2},$$

and the distance between $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$ is

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

These two quantities can be graphically represented in the Cartesian plane:

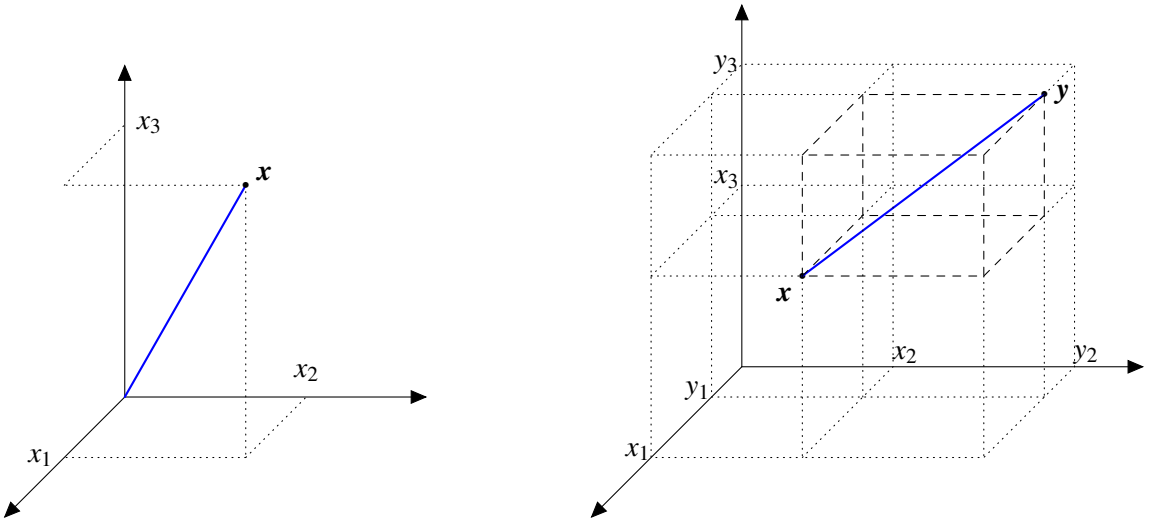


Using the same idea we can define the magnitude of a vector $\boldsymbol{x} = (x_1, x_2, x_3)$ in $\mathbb{R}^3$ as

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2},$$

and the distance between $\boldsymbol{x} = (x_1, x_2, x_3)$ and $\boldsymbol{y} = (y_1, y_2, y_3)$ as

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

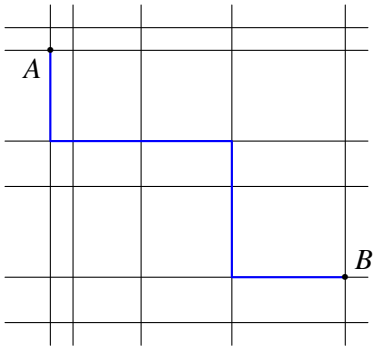


Generalizing this notion to $\mathbb{R}^n$ seems like a trivial task: the magnitude of a vector $\boldsymbol{x} \in \mathbb{R}^n$ is

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2},$$

where $x_1, x_2, \ldots, x_n$ are the components of $\boldsymbol{x}$ in terms of the standard basis, and the distance between two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$ is $\|\boldsymbol{x} - \boldsymbol{y}\|$.

However, in some applications one may want to define magnitude and distance in a different way. For example, the distance that a taxi needs to travel to get from a point A to a point B of a city is typically greater than the Euclidean distance. If all the streets of this city are either mutually parallel or perpendicular (i.e., the city map looks like a grid), we have the situation represented by the figure below, where the blue segments represent the path that a taxi could take to travel from point A to point B:



This begs the question: is it possible to generalize the notion of distance in a way that encompasses this situation? Is there also a corresponding notion of magnitude that we can associate to this distance? Are these generalizations equivalent, in the sense that from a distance we have a corresponding notion of magnitude and *vice-versa*? The generalization of the concept of distance will lead to the notion of *metric*, and the generalization of the concept of magnitude leads to the notion of *norm*. These generalizations will highlight an interesting fact: a norm can always be used to define a distance, but a distance may not correspond to any particular norm.

The next question at this point is: why do we need to be concerned about generalizing notions such as distance and magnitude? To answer this question, recall the definition of convergence and limit. If $f : A \longrightarrow \mathbb{R}$, with $A \subseteq \mathbb{R}$, the real number x_0 is a limit point of A and $\ell \in \mathbb{R}$, then f is said to converge to

ℓ as $x \to x_0$ if

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\quad \forall\, x \in A \setminus \{x_0\}, \qquad |x - x_0| < \delta \ \Rightarrow\ |f(x) - \ell| < \varepsilon.$$

The condition $|x - x_0| < \delta$ says that the distance of x from x_0 is strictly smaller than δ, and $|f(x) - \ell| < \varepsilon$ says that the distance of $f(x)$ from ℓ is strictly smaller than ε. We have similar definitions for the case where x tends to $\pm\infty$ and/or the limit is $\pm\infty$, or even when x tends to x_0 from the left or from the right and $f(x)$ tends to ℓ from the left or from the right.

You might have seen a generalization of the definition of limit for functions $f : \boldsymbol{A} \longrightarrow \mathbb{R}^m$, where $\boldsymbol{A} \subseteq \mathbb{R}^n$. In this case, the definition looks very similar:

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\quad \forall\, x \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\| < \varepsilon.$$

Again, we are considering the Euclidean distance between the points $\boldsymbol{x}$ and $\boldsymbol{x}_0$ and between $f(\boldsymbol{x})$ and $\boldsymbol{\ell}$ (notice that we use bold characters to denote vectors and sets containing vectors). In essence, the definition of limit and convergence hinges on the notion of distance, and generalizing this notion leads to a generalization of the central notion of limit, which then potentially leads to a generalization of the concepts of continuity, derivative, integral and series.

So far in this discussion, we have only considered distances and magnitudes of vectors in $\mathbb{R}^n$, and we know from linear algebra that $\mathbb{R}^n$ has the structure of a real vector space. There are, however, vector spaces that have a richer structure than $\mathbb{R}^n$, and which constitute the main object of our investigation: spaces of functions. For example, given $A \subseteq \mathbb{R}$, the set of continuous functions is a real vector space if we define the addition and the scalar multiplication in the obvious way. Can we define magnitude and distance for vectors (which, in this example, are functions) in this infinite-dimensional vector space? Before we can give an answer to this important question, we need to develop some background material. This is the objective of the coming sections.

1.1 Metric spaces

We begin by introducing a notion of distance or *metric* between points of a set. The properties that any notion of distance is expected to fulfil are the following:

(A) the distance from a point A to a point B is zero if and only if A and B coincide;

(B) the distance from a point A to a point B is the same as the distance from point B to point A;

(C) the distance from A to B is less than or equal to the distance from A to B via any third point C;

(D) the distance between a point with itself is zero;

(E) the distance between two distinct points is positive.

We now see how to turn these intuitive requirements into a formal definition of distance. In particular, we will see that a formal definition only requires properties **(A), (B), (C)** as axioms, because **(D)** and **(E)** follow from them, in the sense that they are a consequence of **(A), (B), (C)**.

Definition 1.1 [Metric/Distance]
Let E be a set. A metric *or* distance *in E is a function*

$$\mathrm{d} : E \times E \longrightarrow \mathbb{R}$$

that satisfies the following axioms:

(i) identity of indiscernibles: $\forall\, x, y \in E, \qquad \mathrm{d}(x, y) = 0 \;\Leftrightarrow\; x = y;$

(ii) symmetry: $\forall\, x, y \in E, \quad \mathrm{d}(x, y) = \mathrm{d}(y, x);$

(iii) triangle inequality: $\forall\, x, y, z \in E, \quad \mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y).$

In this case, the pair (E, d) is called metric space. *If in addition E is a vector space over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$), we say that the pair (E, d) is a* metric vector space.

The triangle inequality property is also called *sub-additivity*. For points in the Euclidean space, the triangle inequality states that the length of one side of a triangle is less than the sum of the lengths of the other two sides. We have anticipated in our previous discussion that properties **(A)**, **(B)** and **(C)** lead to **(D)** and **(E)**. The following theorem shows why.

Theorem 1.1 *Let (E, d) be a metric space. Then:*

- $\forall\, x \in E,\ \mathrm{d}(x, x) = 0;$
- $\forall\, x, y \in E,\ \mathrm{d}(x, y) \geq 0.$

Proof Using the identity of indiscernibles, we have immediately $\mathrm{d}(x, x) = 0$. This proves the first point. Let $x, y \in E$. From the triangle inequality, we have

$$\mathrm{d}(x, y) + \mathrm{d}(y, x) \geq \mathrm{d}(x, x),$$

and by using the symmetry and the result of the previous point we have

$$\mathrm{d}(x, y) + \mathrm{d}(x, y) \geq \mathrm{d}(x, x) = 0,$$

so that we obtain $2\,\mathrm{d}(x, y) \geq 0$, from which the statement follows.

If E and F are two metric spaces with distances d_E and d_F, respectively, a function $f : E \longrightarrow F$ such that

$$\forall\, x, y \in E, \qquad \mathrm{d}_F(f(x), f(y)) = \mathrm{d}_E(x, y) \tag{1.1}$$

is called *isometry*. In other words, an isometry is a function that preserves the distance. An isometry is an injective function (see Exercise 1.1), but it is not necessarily surjective.[1]

[1] The bijective isometries of a metric space into itself form a group, called the *isometry group*.

Example *1.1*

A well-known example of metric in $\mathbb{R}$ is the function

$$\boxed{\begin{array}{rcl} \mathrm{d} : \mathbb{R} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & |x - y| \end{array}}$$

To show that this is indeed a metric, we need to prove that for all $x, y, z \in \mathbb{R}$

(i) $|x - y| = 0 \;\Rightarrow\; x = y$;

(ii) $|x - y| = |y - x|$;

(iii) $|x - y| \leq |x - z| + |z - y|$.

We recall that the absolute value of $x \in \mathbb{R}$ is defined by

$$|x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0. \end{cases}$$

It follows immediately that $|x| = 0$ if and only if $x = 0$. This proves the first point. The second comes immediately from

$$|x - y| = \begin{cases} x - y & \text{for } x \geq y \\ y - x & \text{for } x < y \end{cases} \quad \text{and} \quad |y - x| = \begin{cases} y - x & \text{for } x \leq y \\ x - y & \text{for } x > y, \end{cases}$$

which, in view of the previous point, are the same as

$$|x - y| = \begin{cases} x - y & \text{for } x > y \\ 0 & \text{for } x = y \\ y - x & \text{for } x < y \end{cases} \quad \text{and} \quad |y - x| = \begin{cases} y - x & \text{for } x < y \\ 0 & \text{for } x = y \\ x - y & \text{for } x > y, \end{cases}$$

and therefore they coincide. To prove the last point, we show that the triangle inequality $|x + y| \leq |x| + |y|$ holds for the absolute value. Let $x, y \in \mathbb{R}$.

- if $x \geq 0$ and $y \geq 0$, we have $x + y \geq 0$, so that $|x + y| = x + y = |x| + |y|$;
- if $x \geq 0$ and $y < 0$, we distinguish between the following two cases:
  - if $x + y \geq 0$, we have $|x + y| = x + y$; since $y < 0$, we have $y < -y$, and, using the fact that $x \geq 0$, we obtain $x + y < x + (-y) = |x| + |y|$, which implies $|x + y| < |x| + |y|$;
  - if $x + y < 0$, we have $|x + y| = -(x + y) = (-x) + (-y)$; we have $-x \leq x$, so that $(-x) + (-y) \leq x + (-y) = |x| + |y|$; thus, $|x + y| \leq |x| + |y|$.

  In both cases, we have $|x + y| \leq |x| + |y|$.
- if $x < 0$ and $y \geq 0$, the argument is similar to the previous case;
- if $x < 0$ and $y < 0$, we have $x + y < 0$, so that $|x + y| = -(x + y) = (-x) + (-y) = |x| + |y|$.

Re-writing the triangle inequality as $|a + b| \leq |a| + |b|$, and taking $a = x - z$ and $b = z - y$ leads to $|x - z + z - y| \leq |x - z| + |z - y|$, which is exactly what we had to prove, i.e., $|x - y| \leq |x - z| + |z - y|$.

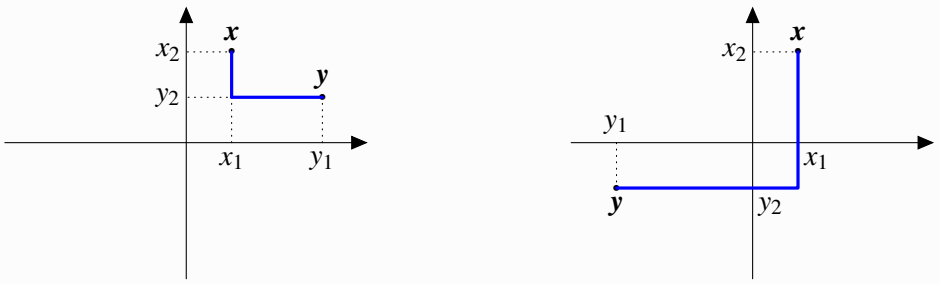
Example *1.2*

Let $n \in \mathbb{N}^*$. We introduce the so-called *taxi-cab metric* in $\mathbb{R}^n$, defined as the function

$$\boxed{\begin{array}{rcl} \mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) & \mapsto & |x_1 - y_1| + |x_2 - y_2| + \ldots + |x_n - y_n| = \displaystyle\sum_{i=1}^{n} |x_i - y_i| \end{array}}$$

The name "taxi-cab metric" comes from the interpretation of this metric as a model for the distance that a taxi-

cab would need to traverse to get from one point to another if the cab was only allowed to move in cardinal directions (north, south, east, west) and not diagonally:



To prove that this is indeed a metric, we need to check the axioms, and we will see that they are all satisfied because we know that they are satisfied for the absolute value:

- identity of indiscernibles: for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, using the fact that the absolute value is a non-negative quantity, we have

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad \Leftrightarrow \quad \forall\, i \in \{1, \ldots, n\},\ |x_i - y_i| = 0 \quad \Leftrightarrow \quad \forall\, i \in \{1, \ldots, n\},\ x_i = y_i \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{y}.$$

- symmetry: for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = |x_1 - y_1| + \ldots + |x_n - y_n| = |y_1 - x_1| + \ldots + |y_n - x_n| = \mathrm{d}(\boldsymbol{y}, \boldsymbol{x}).$$

- sub-additivity: for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$,

$$\begin{aligned} \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) &= |x_1 - y_1| + \ldots + |x_n - y_n| \\ &\leq |x_1 - z_1| + |z_1 - y_1| + \ldots + |x_n - z_n| + |z_n - y_n| = \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y}). \end{aligned}$$

Example *1.3*

Let $n \in \mathbb{N}^*$. We introduce the so-called *Euclidean distance* in $\mathbb{R}^n$, defined as the function

$$\begin{aligned} \mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \end{aligned}$$

In linear algebra you have studied that the norm of a vector $\boldsymbol{x} \in \mathbb{R}^n$ is defined in terms of the dot product as

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}. \tag{1.2}$$

Since in the next section we will define other types of norms for vector spaces, the norm defined by (1.2) will be referred to as *Euclidean norm*. Using this definition, we can write the distance between two points $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ as

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|.$$

We verify that this function is indeed a metric. First, given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, we find

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2} = 0 \quad \Leftrightarrow \quad \forall\, i \in \{1, \ldots, n\},\ x_i = y_i \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{y}$$

The symmetry is trivial, since $(x_i - y_i)^2 = (y_i - x_i)^2$ for all $i \in \{1, \dots, n\}$. We prove the triangle inequality. To this end, we recall the Cauchy-Schwarz inequality for two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$:

$$|\boldsymbol{a} \cdot \boldsymbol{b}| \leq \|\boldsymbol{a}\| \, \|\boldsymbol{b}\|.$$

Trivially, we have also $\boldsymbol{a} \cdot \boldsymbol{b} \leq |\boldsymbol{a} \cdot \boldsymbol{b}| \leq \|\boldsymbol{a}\| \, \|\boldsymbol{b}\|$. We obtain

$$\begin{aligned} \|\boldsymbol{a} + \boldsymbol{b}\|^2 &= \sum_{i=1}^{n} (a_i + b_i)^2 = \sum_{i=1}^{n} (a_i^2 + 2\, a_i\, b_i + b_i^2) \\ &= \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 + 2 \sum_{i=1}^{n} a_i\, b_i \\ &= \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + 2\, \boldsymbol{a} \cdot \boldsymbol{b} \leq \|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2 + 2\, \|\boldsymbol{a}\| \, \|\boldsymbol{b}\| = (\|\boldsymbol{a}\| + \|\boldsymbol{b}\|)^2. \end{aligned}$$

Taking the square root gives $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$, and substituting into it $\boldsymbol{a} = \boldsymbol{x} - \boldsymbol{z}$ and $\boldsymbol{b} = \boldsymbol{z} - \boldsymbol{y}$, we obtain

$$\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{z}\| + \|\boldsymbol{z} - \boldsymbol{y}\| \qquad \Leftrightarrow \qquad \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) \leq \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y}),$$

which is exactly the triangle inequality.

Example *1.4*

Let $n \in \mathbb{N}^*$. The function

$$\begin{aligned} \mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}, \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \max_{1 \leq i \leq n} |x_i - y_i| \end{aligned}$$

is a metric (see Exercise 1.2). This metric is called *maximum metric*.

Example *1.5*

Let E be a non-empty set, and consider the function $\mathrm{d} : E \times E \longrightarrow \mathbb{R}$ defined as

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in E, \qquad \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{y} \\ 1 & \text{if } \boldsymbol{x} \neq \boldsymbol{y}. \end{cases}$$

The pair (E, d) is a metric space. (see Exercise 1.3). This metric is called *discrete metric*. In this metric, all points are equally far apart. This example also shows that every non-empty set E has at least one metric on it.

Remark 1.1 If (E, d) is a metric space, and $F \subseteq E$, it is easy to see that the restriction of the metric d to $F \times F$

$$\mathrm{d}|_{F \times F} : F \times F \longrightarrow \mathbb{R}$$

is still a metric. We say in this case that the metric in F has been "induced" by the metric of (E, d). We can also call $(F, \mathrm{d}|_{F \times F})$ a *metric subspace* of (E, d). For example, $(\mathbb{Q}, |\cdot|)$ is a metric subspace of $(\mathbb{R}, |\cdot|)$, which is in turn a metric subspace of $(\mathbb{C}, |\cdot|)$.

Exercises

Exercise 1.1

Let (E, d_E) and (F, d_F) be two metric spaces. Let $f : E \longrightarrow F$ be an isometry (1.1). Show that f is injective. [Solution: (1.1)]

Exercise 1.2 Let $n \in \mathbb{N}^*$. Show that the function

$$\begin{aligned} \mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}, \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \max_{1 \le i \le n} |x_i - y_i| \end{aligned}$$

is a metric. [Solution: (1.2)]

Exercise 1.3 Let E be a non-empty set, and consider the function $\mathrm{d} : E \times E \longrightarrow \mathbb{R}$ defined as

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in E, \qquad \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{y} \\ 1 & \text{if } \boldsymbol{x} \neq \boldsymbol{y}. \end{cases}$$

Prove that (E, d) is a metric space. [Solution: (1.3)]

Exercise 1.4 Let $\boldsymbol{X}$ be a vector space over a field $\mathbb{F}$ (equal to $\mathbb{R}$ or $\mathbb{C}$). Check whether the 1-distance

$$\mathrm{d}_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

enjoys the following properties:

Translation invariance: $\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w} \in \boldsymbol{X}, \qquad \mathrm{d}_1(\boldsymbol{x} + \boldsymbol{w}, \boldsymbol{y} + \boldsymbol{w}) = \mathrm{d}_1(\boldsymbol{x}, \boldsymbol{y})$.

Homogeneity: $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X},\ \forall\, k \in \mathbb{F}, \quad \mathrm{d}_1(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = |k|\, \mathrm{d}_1(\boldsymbol{x}, \boldsymbol{y})$.

[Solution: (1.4)]

Exercise 1.5 Let E be the set of n-letter words in the alphabet

$$A = \{a_1, a_2, \ldots, a_k\}.$$

This means that $E = \{(x_1, x_2, \ldots, x_n)\} \,|\, x_1, x_2, \ldots, x_n \in A\}$. We can define the distance between two words $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ as the number of positions where the two words have different letters. In other words, $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})$ is the cardinality of the set

$$\{i \in \{1, \ldots, n\} \,|\, x_i \neq y_i\}.$$

Show that (E, d) is a metric space. [Solution: (1.5)]

Exercise 1.6 Check whether the 2-distance

$$\mathrm{d}_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

enjoys the following properties:

Translation invariance: $\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w} \in \boldsymbol{X}, \quad \mathrm{d}_2(\boldsymbol{x}+\boldsymbol{w}, \boldsymbol{y}+\boldsymbol{w}) = \mathrm{d}_2(\boldsymbol{x}, \boldsymbol{y})$;

Homogeneity: $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X},\ \forall\, k \in \mathbb{F}, \quad \mathrm{d}_2(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = |k|\, \mathrm{d}_2(\boldsymbol{x}, \boldsymbol{y})$.

[Solution: (1.6)]

1.2 Normed spaces

The notion of norm is central in mathematics.

Definition 1.2 [NORM]
Let $\boldsymbol{X}$ be a vector space over $\mathbb{F}$ (equal to $\mathbb{R}$ or $\mathbb{C}$). A norm in $\boldsymbol{X}$ is a function

$$\begin{aligned} \|\cdot\| : \boldsymbol{X} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\| \end{aligned}$$

which satisfies the following axioms:

(i) *homogeneity:* $\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \forall\, k \in \mathbb{F},\ \|k\,\boldsymbol{x}\| = |k|\,\|\boldsymbol{x}\|$;

(ii) *sub-additivity:* $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X},\ \|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$;

(iii) *positivity:* $\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \ \|\boldsymbol{x}\| = 0 \ \Rightarrow \boldsymbol{x} = \boldsymbol{0}$.

In this case, the pair $(\boldsymbol{X}, \|\cdot\|)$ is called normed vector space *(or* normed space *in short).*

The sub-additivity property is also called *triangle inequality*. When we say that $(\boldsymbol{X}, \|\cdot\|)$ is a normed space, we mean that:

(i) $\boldsymbol{X}$ is a real or complex vector space (and we will not specify the underlying field if it is clear from the context or if what we are saying does not depend on it);

(ii) $\|\cdot\|$ is a norm in $\boldsymbol{X}$.

We will see that the norm defined in (1.2) is only one particular case of norm which satisfies the axioms in Definition 1.2. Hence, from now on, the symbol $\|\cdot\|$ will be used to denote a generic norm, while the norm in (1.2), for reasons that will become clear in the sequel, will be denoted by $\|\cdot\|_2$.

In several textbooks, the implication in *(iii)* appears as a logical equivalence. However, as the following result shows, the implication $\boldsymbol{x} = \boldsymbol{0} \ \Rightarrow\ \|\boldsymbol{x}\| = 0$ is a consequence of the homogeneity, and therefore it does not need to be imposed as an axiom. Moreover, sometimes an additional condition is included in the definition of norm, which is included in the third condition and justifies the use of the term "positivity": $\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \ \|\boldsymbol{x}\| \geq 0$. However, this is a consequence of *(i)* and *(ii)*, as the following result also shows, and therefore we prefer not to include it in the set of axioms which define a norm.

Theorem 1.2 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Then:*

(i) $\|\boldsymbol{0}\| = 0$;

(ii) $\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \ \|\boldsymbol{x}\| = \|-\boldsymbol{x}\|$;

(iii) $\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \ \|\boldsymbol{x}\| \geq 0$;

(iv) $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X},\ \ \big|\|\boldsymbol{y}\| - \|\boldsymbol{x}\|\big| \leq \|\boldsymbol{y} - \boldsymbol{x}\|$ (REVERSE TRIANGLE INEQUALITY).

Proof

(i) Let $\boldsymbol{x} \in \boldsymbol{X}$. Let $0 \in \mathbb{F}$. From the properties of vector spaces, we know that for every $\boldsymbol{x} \in \boldsymbol{X}$ we have $0\,\boldsymbol{x} = \boldsymbol{0}$, so that using the positive homogeneity we immediately find $\|\boldsymbol{0}\| = \|0\,\boldsymbol{x}\| = |0|\,\|\boldsymbol{x}\| = 0$. This proves the first point.

(ii) Both $\mathbb{R}$ and $\mathbb{C}$ contain an element -1 (the additive inverse of the identity element of the addition), which satisfies $-1\,\boldsymbol{x} = -\boldsymbol{x}$. Using the positive homogeneity we obtain

$$\|-\boldsymbol{x}\| = \|-1\,\boldsymbol{x}\| = |-1|\,\|\boldsymbol{x}\| = \|\boldsymbol{x}\|.$$

(iii) To prove the positivity, we apply the result of the previous point and the sub-additivity to $\boldsymbol{x}$ and $-\boldsymbol{x}$:

$$\|\boldsymbol{0}\| = \|\boldsymbol{x} + (-\boldsymbol{x})\| \leq \|\boldsymbol{x}\| + \|-\boldsymbol{x}\| = \|\boldsymbol{x}\| + \|\boldsymbol{x}\| = 2\,\|\boldsymbol{x}\|.$$

Since $\|\boldsymbol{0}\| = 0$ from the first point, we obtain $0 \leq 2\,\|\boldsymbol{x}\|$, which gives $\|\boldsymbol{x}\| \geq 0$;

(iv) We now prove the last point, which is equivalent to showing that

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \quad -\|\boldsymbol{y} - \boldsymbol{x}\| \leq \|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\|. \tag{1.3}$$

Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. We use the triangle inequality to find

$$\|\boldsymbol{y}\| = \|\boldsymbol{y} - \boldsymbol{x} + \boldsymbol{x}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\| + \|\boldsymbol{x}\|,$$

which implies

$$\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \leq \|\boldsymbol{y} - \boldsymbol{x}\|. \tag{1.4}$$

Interchanging $\boldsymbol{y}$ and $\boldsymbol{x}$ in (1.4) and using the second point gives

$$\|\boldsymbol{x}\| - \|\boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| = \|\boldsymbol{y} - \boldsymbol{x}\|,$$

from which we find

$$\|\boldsymbol{y}\| - \|\boldsymbol{x}\| \geq -\|\boldsymbol{y} - \boldsymbol{x}\|. \tag{1.5}$$

Eq.s (1.4) and (1.5) together yield (1.3).

Example *1.6*

We want to show that for the real vector space $\mathbb{R}$, the absolute value is a norm. In other words, we want to verify that

$$\begin{aligned} |\cdot| : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto |x| = \begin{cases} x & \text{for } x \geq 0 \\ -x & \text{for } x < 0 \end{cases} \end{aligned}$$

satisfies the norm axioms. We prove the positive homogeneity. Let $x \in \mathbb{R}$ and $k \in \mathbb{R}$. We have

$$(x \geq 0 \text{ and } k \geq 0) \quad \text{or} \quad (x \geq 0 \text{ and } k < 0) \quad \text{or} \quad (x < 0 \text{ and } k \geq 0) \quad \text{or} \quad (x < 0 \text{ and } k < 0).$$

We consider each case separately:

- if $x \geq 0$ and $k \geq 0$, we have $k\,x \geq 0$. Thus, $|k\,x| = k\,x = |k|\,|x|$;
- if $x \geq 0$ and $k < 0$, we have $k\,x \leq 0$, so that $|k\,x| = -k\,x = (-k)\,x = |k|\,|x|$;
- if $x < 0$ and $k \geq 0$, the argument is similar to the previous case;

- if $x < 0$ and $k < 0$, we have $k\,x > 0$, and $|k\,x| = k\,x = (-k)(-x) = |k|\,|x|$.

The proof of the sub-additivity was carried out in a previous example, where it was shown precisely that $|x + y| \leq |x| + |y|$. The positivity is obvious from the definition of absolute value.

Example *1.7* The modulus $|\cdot|$ in $\mathbb{C}$ is a norm in $\mathbb{C}$, both when considering $\mathbb{C}$ as a vector space over the field $\mathbb{R}$ and when $\mathbb{C}$ is considered as a vector space over the field $\mathbb{C}$, see Exercise 1.8.

The absolute value in $\mathbb{R}$ is linked to a notion of distance, and allows us to view $\mathbb{R}$ as a metric space. We have also seen that the absolute value is an example of a norm, which is linked to the notion of magnitude. It is not a coincidence that a norm like the absolute value induces a notion of distance. We show that this is always the case: we can immediately associate a distance to a norm in a metric vector space:

$$\begin{aligned} \mathrm{d} : \boldsymbol{X} \times \boldsymbol{X} &\longrightarrow \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \|\boldsymbol{x} - \boldsymbol{y}\|. \end{aligned} \tag{1.6}$$

Indeed, $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = 0$ implies $\boldsymbol{x} = \boldsymbol{y}$ from the positivity property of Definition 1.2 applied to $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{y}$, which proves the identity of indiscernibles. Moreover, $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \|-(\boldsymbol{x} - \boldsymbol{y})\| = \|\boldsymbol{y} - \boldsymbol{x}\| = \mathrm{d}(\boldsymbol{y}, \boldsymbol{x})$, thus proving the symmetry. Finally, $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) \leq \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y})$ can be written as

$$\|\boldsymbol{x} - \boldsymbol{y}\| \leq \|\boldsymbol{x} - \boldsymbol{z}\| + \|\boldsymbol{z} - \boldsymbol{y}\|,$$

which holds from the sub-additivity (writing $\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{z}$ and $\tilde{\boldsymbol{y}} = \boldsymbol{z} - \boldsymbol{y}$). Thus, every normed space is also a metric space, with the distance induced by the norm. It follows also that all concepts that we defined for metric spaces also apply, in particular, to normed spaces.

Remark 1.2 The concept of distance between two points is not the only one that can be constructed starting from a norm. For example, if $(\boldsymbol{X}, \|\cdot\|)$ is a normed space and $\boldsymbol{A}$ and $\boldsymbol{B}$ are two subsets of $\boldsymbol{X}$, given $\boldsymbol{x}_0 \in \boldsymbol{X}$ we can define:

- the distance between the point $\boldsymbol{y}$ and the set $\boldsymbol{A}$ as $\mathrm{d}(\boldsymbol{y}, \boldsymbol{A}) \stackrel{\text{def}}{=} \inf \{\|\boldsymbol{x} - \boldsymbol{y}\| \;\big|\; \boldsymbol{x} \in \boldsymbol{A}\}$;
- the distance between the sets $\boldsymbol{A}$ and $\boldsymbol{B}$ as $\mathrm{d}(\boldsymbol{A}, \boldsymbol{B}) \stackrel{\text{def}}{=} \inf \{\|\boldsymbol{x} - \boldsymbol{y}\| \;\big|\; \boldsymbol{x} \in \boldsymbol{A} \text{ and } \boldsymbol{y} \in \boldsymbol{B}\}$;
- the diameter of a set $\boldsymbol{A}$ as $\mathrm{diam}(\boldsymbol{A}) \stackrel{\text{def}}{=} \sup \{\|\boldsymbol{x} - \boldsymbol{y}\| \;\big|\; \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}\}$.

We have analogous notions for metric spaces, if we replace the conditions $\|\boldsymbol{x} - \boldsymbol{y}\|$ with $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})$.

Remark 1.3 We have seen that a normed space $(\boldsymbol{X}, \|\cdot\|)$ is a metric space with the norm $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$. However, starting from a norm, we can define different metrics. For example, it is not hard to prove that

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\|\boldsymbol{x} - \boldsymbol{y}\|}{1 + \|\boldsymbol{x} - \boldsymbol{y}\|}$$

is also a metric, where the distance between two arbitrary points in $\boldsymbol{X}$ is always less than 1, see Exercise 1.7.

Example *1.8*

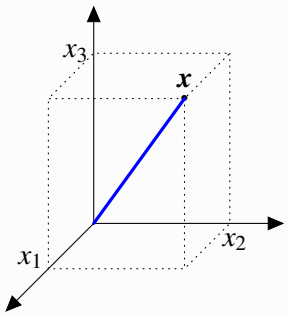
Given $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define the norm as

$$\|\boldsymbol{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

In the next examples we will see that there are other choices of norms that are consistent with Definition 1.2. A simple way to see that this is indeed a norm is to observe that it is induced by the Euclidean distance

$$\begin{aligned} \mathrm{d} : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}. \end{aligned}$$

For this reason, this norm is called *Euclidean norm* of $\mathbb{R}^n$. The blue segment in the figure below represents the Euclidean norm of a vector in $\mathbb{R}^3$:

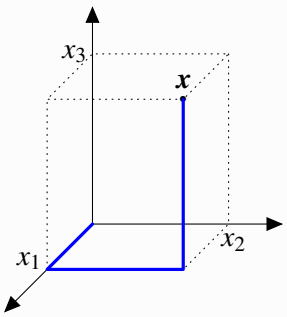


Example *1.9*

Given $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define the 1-norm as

$$\|\boldsymbol{x}\|_1 \stackrel{\text{def}}{=} |x_1| + |x_2| + \ldots + |x_n| = \sum_{k=1}^{n} |x_k|.$$

In the figure below, the sum of the lengths of the blue segments represents the 1-norm of a vector in $\mathbb{R}^3$:



It is easy to check that $\|\cdot\|_1$ is indeed a norm, i.e., it satisfies the conditions in Definition 1.2 (see Exercise 1.10).

Example *1.10*

Given $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define

$$\|\boldsymbol{x}\|_\infty \stackrel{\text{def}}{=} \max_{1 \le i \le n} |x_i|.$$

We want to show that $\|\cdot\|_\infty$ satisfies the conditions in Definition 1.2. We start with the homogeneity. If $k \in \mathbb{R}$, we have

$$\|k\,\boldsymbol{x}\|_\infty = \max_{1 \le i \le n} |k\,x_i| = \max_{1 \le i \le n} |k|\,|x_i| = |k| \max_{1 \le i \le n} |x_i|,$$

where the last equality holds because $|k| \ge 0$ (see Exercise 1.13). We find $\|k\,\boldsymbol{x}\|_\infty = |k|\,\|\boldsymbol{x}\|_\infty$. We now prove the sub-additivity. For $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, we must show that $\|\boldsymbol{x} + \boldsymbol{y}\|_\infty \le \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty$. From

$$\|\boldsymbol{x} + \boldsymbol{y}\|_\infty = \max_{1 \le i \le n} |x_i + y_i| = |x_{i_0} + y_{i_0}|,$$

i.e., $i_0 \in \{1, 2, \ldots, n\}$ is the value of the index for which the maximum is attained. We find

$$|x_{i_0} + y_{i_0}| \le |x_{i_0}| + |y_{i_0}| \le \max_{1 \le i \le n} |x_i| + \max_{1 \le i \le n} |y_i| = \|\boldsymbol{x}\|_\infty + \|\boldsymbol{y}\|_\infty.$$

Finally, we show the positivity. If $\|\boldsymbol{x}\|_\infty = 0$, then $\max_{1 \le i \le n} |x_i| = 0$, from which we find that $x_i = 0$ for all $i \in \{1, 2, \ldots, n\}$, i.e., $\boldsymbol{x} = \boldsymbol{0}$. Another way to prove that $\|\cdot\|_\infty$ is a norm is to observe that it is induced by the maximum metric.

For all $p \ge 1$, it is possible to define in $\mathbb{R}^n$ the so-called p-norm

$$\|\boldsymbol{x}\|_p \stackrel{\text{def}}{=} \Big(\sum_{k=1}^{n} |x_k|^p \Big)^{\frac{1}{p}}.$$

When $p = 1$, we obtain as a particular case the norm $\|\cdot\|_1$. For $p = 2$, we obtain the Euclidean norm (which is therefore also referred to as 2-norm; this explains the use of the subscript 2 in the notation for the Euclidean norm).

Example *1.11*

For all $\boldsymbol{x} \in \mathbb{R}^n$

$$\lim_{p \to +\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty.$$

see Exercise 1.17.

Our first task is to establish that this is indeed a norm. This is accomplished by using a preliminary lemma, known as *Young's inequality*.

Theorem 1.3 [Young's Inequality]
Let $a, b \ge 0$. Let $p, q \in \mathbb{R}$ with $p, q > 1$ and such that $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Then,

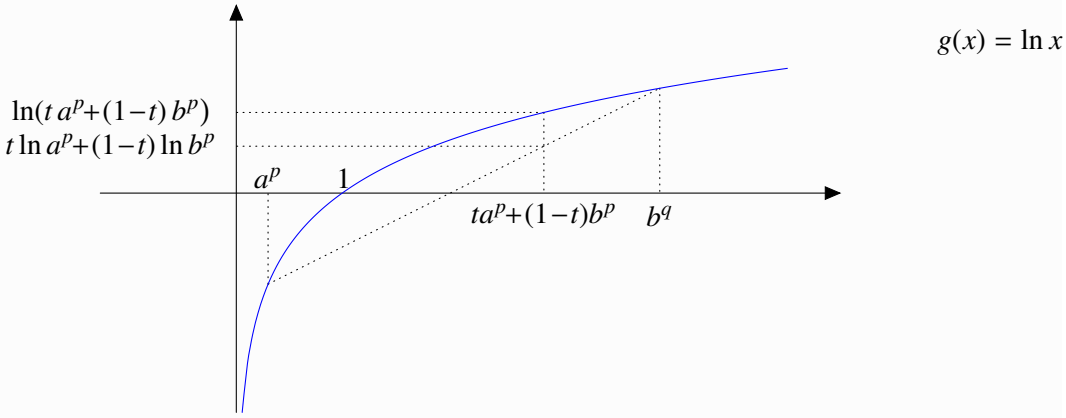
$$a\,b \le \frac{a^p}{p} + \frac{b^q}{q}.$$

Moreover, the inequality becomes an equality if and only if $a^p = b^q$.

Proof For $a = 0$ or $b = 0$ the inequality trivially holds true (we recall that, while a power with real exponent x^α is only defined for $x > 0$ when $\alpha \in \mathbb{R}$, for the sake of convenience it is standard to assume $0^\alpha = 0$). Assume

$a > 0$ and $b > 0$. We can define $t = \frac{1}{p}$, so that from $\frac{1}{p} + \frac{1}{q} = 1$ we have $1 - t = \frac{1}{q}$. Since the logarithm function is concave in $(0, +\infty)$ (why? think about the second-order derivative), for all $x, y > 0$ and for all $t \in [0, 1]$ we have $\ln(t\,x + (1-t)\,y) \geq t \ln x + (1-t) \ln y$. Applying this result with the value $t = \frac{1}{p}$ and with $x = a^p$ and $y = b^q$, we find

$$\ln\left(\frac{a^p}{p} + \frac{b^q}{q}\right) = \ln(t\,a^p + (1-t)\,b^q) \geq t \ln a^p + (1-t) \ln b^q = \frac{1}{p} \ln a^p + \frac{1}{q} \ln b^q = \ln a + \ln b = \ln(a\,b).$$



Taking the exponential on both sides preserves the inequality (the exponential is strictly increasing), and this concludes the proof. Notice that in the latter we have equality if and only if $a^p = b^q$.

If $p = 1$, there are no values of $q \in \mathbb{R}$ satisfying

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{1.7}$$

With some abuse of notation, we say that in this case $q = +\infty$, i.e., we assume for the sake of notational simplicity $\frac{1}{+\infty} = 0$ (and vice-versa when $q = 1$). We say that $p, q \in [1, +\infty]$ are *conjugate exponents* if they satisfy the condition (1.7).

Theorem 1.4 *Let $p \in [1, +\infty]$. Let $\boldsymbol{x} \in \mathbb{R}^n$. The function $\|\cdot\|_p : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by*

$$\|\boldsymbol{x}\|_p \stackrel{def}{=} \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}}$$

defines a norm.

Proof The case where $p = +\infty$ has already been addressed. Hence, we focus on the case where $p \in [1, +\infty)$. We start from the homogeneity. Let $k \in \mathbb{R}$:

$$\|k\,\boldsymbol{x}\|_p = \Big(\sum_{i=1}^{n} |k\,x_i|^p\Big)^{\frac{1}{p}} = \Big(\sum_{i=1}^{n} |k|^p\,|x_i|^p\Big)^{\frac{1}{p}} = \Big(|k|^p \sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} = |k|\Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} = |k|\,\|\boldsymbol{x}\|_p.$$

We now prove the positivity: we have

$$\|\boldsymbol{x}\|_p = 0 \quad \Rightarrow \quad \Big(\sum_{k=1}^{n} |x_k|^p\Big)^{\frac{1}{p}} = 0 \quad \Rightarrow \quad \sum_{k=1}^{n} |x_k|^p = 0 \quad \Rightarrow \quad \forall\, k \in \{1,\dots,n\},\ |x_k|^p = 0$$
$$\Rightarrow \quad \forall\, k \in \{1,\dots,n\},\ |x_k| = 0 \quad \Rightarrow \quad \forall\, k \in \{1,\dots,n\},\ x_k = 0 \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$

Finally, we now prove the triangle inequality. To this end, we begin proving an intermediate result, known as *Hölder's Inequality*: given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and q such that p and q are conjugate exponents, then

$$\sum_{i=1}^{n} |x_i\, y_i| \le \|\boldsymbol{x}\|_p\, \|\boldsymbol{y}\|_q.$$

If $\boldsymbol{x} = \boldsymbol{0}$, this inequality is trivial: both sides reduce to 0. Likewise, if $\boldsymbol{y} = \boldsymbol{0}$, the statement is true. With no loss of generality, we can therefore assume that both $\boldsymbol{x}$ and $\boldsymbol{y}$ are non-zero. From the positivity, we have in particular that $\|\boldsymbol{x}\|_p \neq 0$ and $\|\boldsymbol{y}\|_q \neq 0$. To use Young's inequality (which requires p and q to be finite), we also need to consider the case where q is equal to $+\infty$ separately. If $p = 1$, then $q = +\infty$, and in this case

$$\sum_{i=1}^{n} |x_i\, y_i| \le \sum_{i=1}^{n} |x_i| \max_{1 \le j \le n} |y_j| = \|\boldsymbol{x}\|_1\, \|\boldsymbol{y}\|_\infty.$$

The case that remains is one where $\boldsymbol{x}$ and $\boldsymbol{y}$ are non-zero and $p, q \in (1, +\infty)$. We can apply Young's inequality with $a = \dfrac{|x_i|}{\|\boldsymbol{x}\|_p}$ and $b = \dfrac{|y_i|}{\|\boldsymbol{y}\|_q}$, where $i \in \{1,\dots,n\}$. We find:

$$\frac{|x_i\, y_i|}{\|\boldsymbol{x}\|_p\, \|\boldsymbol{y}\|_q} \le \frac{|x_i|^p}{p\, \|\boldsymbol{x}\|_p^p} + \frac{|y_i|^q}{q\, \|\boldsymbol{y}\|_q^q}.$$

We can now add over i:

$$\sum_{i=1}^{n} \frac{|x_i\, y_i|}{\|\boldsymbol{x}\|_p\, \|\boldsymbol{y}\|_q} \le \sum_{i=1}^{n} \left(\frac{|x_i|^p}{p\, \|\boldsymbol{x}\|_p^p} + \frac{|y_i|^q}{q\, \|\boldsymbol{y}\|_q^q} \right) = \frac{1}{p\, \|\boldsymbol{x}\|_p^p} \sum_{i=1}^{n} |x_i|^p + \frac{1}{q\, \|\boldsymbol{y}\|_q^q} \sum_{i=1}^{n} |y_i|^q = \frac{1}{p\, \|\boldsymbol{x}\|_p^p} \|\boldsymbol{x}\|_p^p + \frac{1}{q\, \|\boldsymbol{y}\|_q^q} \|\boldsymbol{y}\|_q^q = \frac{1}{p} + \frac{1}{q} = 1.$$

Multiplying both sides by $\|\boldsymbol{x}\|_p\, \|\boldsymbol{y}\|_q$ gives the result.

We now prove the triangle inequality, which, in the specific case of the norm $\|\cdot\|_p$, is also referred to as *Minkowski's inequality*. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. In the case $p = 1$, the triangle inequality holds (see Exercise 1.10). Let us consider $p \in (1, +\infty)$. Using the triangle inequality for the absolute value we find

$$\|\boldsymbol{x} + \boldsymbol{y}\|_p^p = \sum_{i=1}^{n} |x_i + y_i|^p = \sum_{i=1}^{n} |x_i + y_i|\, |x_i + y_i|^{p-1} \le \sum_{i=1}^{n} (|x_i| + |y_i|)\, |x_i + y_i|^{p-1} = \sum_{i=1}^{n} |x_i|\, |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i|\, |x_i + y_i|^{p-1}.$$

Consider the first addend in the right hand-side of the latter: we can apply Hölder's inequality to it, by considering $|x_i + y_i|^{p-1}$ in place of y_i, and considering as q the conjugate exponent of p, i.e.,

$$\frac{1}{p} + \frac{1}{q} = 1 \quad \Rightarrow \quad \frac{1}{q} = \frac{p-1}{p} \quad \Rightarrow \quad q = \frac{p}{p-1}.$$

We obtain

$$\sum_{i=1}^{n} |x_i|\,|x_i+y_i|^{p-1} \le \|\boldsymbol{x}\|_p \Big(\sum_{i=1}^{n} |x_i+y_i|^{q(p-1)}\Big)^{\frac{1}{q}}$$
$$= \|\boldsymbol{x}\|_p \Big(\sum_{i=1}^{n} |x_i+y_i|^{p}\Big)^{1-\frac{1}{p}} = \|\boldsymbol{x}\|_p \frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p^p}{\|\boldsymbol{x}+\boldsymbol{y}\|_p}.$$

We use the same idea for the second addend, and we obtain

$$\sum_{i=1}^{n} |y_i|\,|x_i+y_i|^{p-1} \le \|\boldsymbol{y}\|_p \frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p^p}{\|\boldsymbol{x}+\boldsymbol{y}\|_p},$$

which yields

$$\|\boldsymbol{x}+\boldsymbol{y}\|_p^p \le \|\boldsymbol{x}\|_p \frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p^p}{\|\boldsymbol{x}+\boldsymbol{y}\|_p} + \|\boldsymbol{y}\|_p \frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p^p}{\|\boldsymbol{x}+\boldsymbol{y}\|_p} = (\|\boldsymbol{x}\|_p + \|\boldsymbol{y}\|_p) \frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p^p}{\|\boldsymbol{x}+\boldsymbol{y}\|_p},$$

Diving both sides by $\|\boldsymbol{x}+\boldsymbol{y}\|_p^p$ and multiplying by $\|\boldsymbol{x}+\boldsymbol{y}\|_p$ gives the triangle inequality.

We conclude that the p-norm is a norm.

Remark 1.4 Minkowski's inequality requires the assumption $p \ge 1$. It is possible to prove that, if $p < 1$, this inequality does not hold, and therefore in this case $\|\cdot\|_p$ is not a norm. As an example, let $p = 1/2$, and consider the vectors $\boldsymbol{x} = (1,0)$ and $\boldsymbol{y} = (0,1)$. In this case

$$\|\boldsymbol{x}+\boldsymbol{y}\|_{1/2} = \|(1,1)\|_{1/2} = (1^{1/2}+1^{1/2})^2 = 2^2 = 4$$

and

$$\|\boldsymbol{x}\|_{1/2} = \|(1,0)\|_{1/2} = (1^{1/2})^2 = 1 = \|\boldsymbol{y}\|_{1/2}.$$

Example *1.12*

We can define in $\mathbb{C}^n$ the Euclidean norm $\|\cdot\|$: if $\boldsymbol{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{C}^n$, we define

$$\|\boldsymbol{z}\| = \sqrt{\sum_{k=1}^{n} |z_k|^2},$$

where $|z_k|$ denotes the modulus of the complex number z_k. It is easy to see that $\|\cdot\|$ is indeed a norm, by observing that, if $z_k = a_k + i\,b_k$ for all $k \in \{1,2,\ldots,n\}$, then $|z_k|^2 = |a_k + i\,b_k|^2 = (\sqrt{a_k^2+b_k^2})^2 = a_k^2 + b_k^2$, so that

$$\|\boldsymbol{z}\| = \sqrt{\sum_{k=1}^{n} (a_k^2 + b_k^2)}$$

is the Euclidean norm in $\mathbb{R}^{2n}$ of $(a_1, b_1, \ldots, a_n, b_n) \in \mathbb{R}^{2n}$. For example, if $\boldsymbol{z} = \begin{bmatrix} 2\,i \\ 3-4\,i \end{bmatrix} \in \mathbb{C}^2$, we have[a]

$$\|\boldsymbol{z}\| = \sqrt{4+(9+16)} = \sqrt{29},$$

and this coincides with the Euclidean norm in $\mathbb{R}^4$ of the vector $\boldsymbol{x} = (0,2,3,-4)$, i.e., $\|\boldsymbol{x}\| = \sqrt{0+4+9+16} = \sqrt{29}$.

[a] We can define, in an equivalent way, $\|\boldsymbol{z}\| = \sqrt{\boldsymbol{z}^*\boldsymbol{z}}$, where $\boldsymbol{z}^*$ denotes the conjugate transpose, which gives the

same result:

$$\|z\| = \sqrt{z^* z} = \sqrt{\begin{bmatrix} -2\,i & 3+4\,i \end{bmatrix} \begin{bmatrix} 2\,i \\ 3-4\,i \end{bmatrix}} = \sqrt{4+(9+16)} = \sqrt{29}.$$

1.2.1 Relation between metric and norm

We have seen that a norm always induces a metric, and therefore every normed space is also a metric space. We could ask ourselves if every metric is induced by a norm. The answer is negative, as shown in the following example.

Example *1.13*

Given a non-zero vector space $\boldsymbol{X}$ and the discrete metric $\mathrm{d} : \boldsymbol{X} \times \boldsymbol{X} \longrightarrow \mathbb{R}$ defined as

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \qquad \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{y} \\ 1 & \text{if } \boldsymbol{x} \neq \boldsymbol{y}, \end{cases}$$

we show that the metric space $(\boldsymbol{X}, \mathrm{d})$ is not induced by a norm. By contradiction, suppose that a norm $\|\cdot\| : \boldsymbol{X} \longrightarrow \mathbb{R}$ exists such that, for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$,

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}).$$

Let $\boldsymbol{x} \in \boldsymbol{X} \setminus \{\boldsymbol{0}\}$. From the homogeneity we have

$$\|2\,\boldsymbol{x}\| = 2\,\|\boldsymbol{x}\| = 2\,\|\boldsymbol{x} - \boldsymbol{0}\| = 2\,\mathrm{d}(\boldsymbol{x}, \boldsymbol{0}) = 2.$$

On the other hand,

$$\|2\,\boldsymbol{x}\| = \|3\,\boldsymbol{x} - \boldsymbol{x}\| = \mathrm{d}(3\,\boldsymbol{x}, \boldsymbol{x}) \leq 1.$$

We obtained a contradiction.

Example *1.14*

Consider a finite dimensional vector space $\boldsymbol{V}$ over a field $\mathbb{F}$ (as usual, $\mathbb{R}$ or $\mathbb{C}$). Let n be the dimension of $\boldsymbol{V}$. Let $\mathcal{B} = (\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ be a basis of $\boldsymbol{V}$. We can always construct a norm in this way. Given $\boldsymbol{v} \in \boldsymbol{V}$, let $c_1, c_2, \ldots, c_n \in \mathbb{F}$ be such that $\boldsymbol{v} = c_1\,\boldsymbol{v}_1 + c_2\,\boldsymbol{v}_2 + \ldots + c_n\,\boldsymbol{v}_n$. Then, we can define a norm of $\boldsymbol{v}$ as

$$\|\boldsymbol{v}\| = |c_1| + |c_2| + \ldots + |c_n|,$$

see Exercise 1.15. In other words, if we denote by $[\boldsymbol{v}]_{\mathcal{B}} = (c_1, c_2, \ldots, c_n)$ the coordinate vector of $\boldsymbol{v}$ with respect to the basis $\mathcal{B}$, this norm of $\boldsymbol{v}$ is the 1-norm of its coordinate vector: $\|\boldsymbol{v}\| = \|[\boldsymbol{v}]_{\mathcal{B}}\|_1$. This shows that any finite dimensional real or complex vector space can be considered a normed space with the norm defined in this example, once a basis has been chosen.

Before introducing some more example of norms, this time on infinite dimensional vector spaces, we need to recall the following definitions.

Definition 1.3 [Bounded Set]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{U} \subseteq \boldsymbol{X}$. We say that $\boldsymbol{U}$ is bounded with respect to the norm $\|\cdot\|$ if

$$\exists\, M > 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{U},\ \|\boldsymbol{x}\| \leq M. \tag{1.8}$$

Notice that in the previous definition we can replace the loose inequality sign $\leq$ in $\|\boldsymbol{x}\| \leq M$

with the strict inequality $<$. In other words, (1.8) is equivalent to the condition

$$\exists\, M > 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{U},\ \|\boldsymbol{x}\| < M.$$

(why?).

A function is bounded if its image is bounded, as the following definition says.

Definition 1.4 [Bounded Function]
Let $(\boldsymbol{Y}, \|\cdot\|)$ be a normed space. Let A be a set, and let $f : A \longrightarrow \boldsymbol{Y}$. We say that f is bounded if $f(A)$ is bounded.

Example *1.15*

In analogy with the definition of $\|\cdot\|_\infty$ in $\mathbb{R}^n$, we define a similar norm in the space of bounded functions $B(\boldsymbol{A};\mathbb{R})$, where $\boldsymbol{A} \subseteq \mathbb{R}^n$ (even if we could also simply assume that $\boldsymbol{A}$ is a non-empty set). We consider, in the usual way, the sum of $f, g \in B(\boldsymbol{A};\mathbb{R})$ defined by $(f+g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$ and the scalar product by $(\lambda f)(\boldsymbol{x}) = \lambda f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$, where $\lambda \in \mathbb{R}$. With these two operations, the set $B(\boldsymbol{A};\mathbb{R})$ is a real vector space (why? See Exercise 1.12). In $B(\boldsymbol{A};\mathbb{R})$ we define the norm

$$\|f\|_\infty \stackrel{\text{def}}{=} \sup_{\boldsymbol{x}\in\boldsymbol{A}} |f(\boldsymbol{x})| = \sup\{|f(\boldsymbol{x})| \in \mathbb{R}_+ \mid \boldsymbol{x} \in \boldsymbol{A}\}. \tag{1.9}$$

For example, if $f : \mathbb{R} \longrightarrow \mathbb{R}$ is the inverse tangent (arctan) function, then $\|f\|_\infty = \sup_{x\in\mathbb{R}} |\arctan(x)| = \dfrac{\pi}{2}$. We now prove that (1.9) is indeed a norm. First, we notice that it is a function in $B(\boldsymbol{A};\mathbb{R})$ with values in $[0, +\infty)$.

(i) First, we need to prove $\|k\,f\|_\infty = |k|\,\|f\|_\infty$. This identity is obvious when $k = 0$: in this case, both sides are equal to zero. For every $k \in \mathbb{R}^*$ and every $\boldsymbol{a} \in \boldsymbol{A}$, we have

$$|k\,f(\boldsymbol{a})| = |k|\,|f(\boldsymbol{a})| \le |k|\ \sup\{|f(\boldsymbol{x})| \mid \boldsymbol{x} \in A\} = |k|\,\|f\|_\infty.$$

This holds for every element $\boldsymbol{a}$ of $\boldsymbol{A}$, i.e., $\forall\, \boldsymbol{a} \in \boldsymbol{A}$, $|k\,f(\boldsymbol{a})| \le |k|\,\|f\|_\infty$. This is the same as saying that $|k|\,\|f\|_\infty$ is an upper bound for the set $\{|k\,f(\boldsymbol{a})| \mid \boldsymbol{a} \in \boldsymbol{A}\}$. Now, recalling that the supremum of a set is the least upper bound, and noting that the supremum of $\{|k\,f(\boldsymbol{a})| \mid \boldsymbol{a} \in \boldsymbol{A}\}$ is $\sup_{\boldsymbol{x}\in\boldsymbol{A}} |k\,f(\boldsymbol{x})| = \|k\,f\|_\infty$, we conclude that

$$\|k\,f\|_\infty \le |k|\,\|f\|_\infty. \tag{1.10}$$

Since $k \neq 0$, we can re-write (1.10) with k^{-1} in place of k and with $k\,f$ instead of f, observing that $f = k^{-1}\,(k\,f)$ and we find $\|k^{-1}\,(k\,f)\|_\infty \le |k^{-1}|\,\|k\,f\|_\infty$. We deduce that

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

which is equivalent to $\|k\,f\|_\infty \ge |k|\,\|f\|_\infty$. This inequality, together with (1.10), yields $\|k\,f\|_\infty = |k|\,\|f\|_\infty$. This proves the homogeneity.[a]

(ii) Let $f, g \in B(\boldsymbol{A};\mathbb{R})$ and $\boldsymbol{a} \in \boldsymbol{A}$. Then,

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

Again, this says that $\|f\|_\infty + \|g\|_\infty$ is an upper bound for $\{|f(\boldsymbol{a}) + g(\boldsymbol{a})| \,|\, \boldsymbol{a} \in \boldsymbol{A}\}$. However, the supremum of an ordered set is by definition the least upper bound of that set, so that $\sup\{|f(\boldsymbol{a}) + g(\boldsymbol{a})| \,|\, \boldsymbol{a} \in \boldsymbol{A}\} \leq \|f\|_\infty + \|g\|_\infty$, and the left hand-side is $\|f + g\|_\infty$, and so

$$\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

(iii) Let $f \in B(\boldsymbol{A}; \mathbb{R})$ be such that $\|f\|_\infty = 0$. Then, $\sup\{|f(\boldsymbol{x})| \,|\, \boldsymbol{x} \in \boldsymbol{A}\} = 0$. This implies that $f(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \boldsymbol{A}$. The function which vanishes at every point of $\boldsymbol{A}$ is the additive identity element.

It is easy to verify that in every subspace of $B(\boldsymbol{A}; \mathbb{R})$ (for example, in $BC(\boldsymbol{A}; \mathbb{R})$, which represents the space of bounded and continuous functions from $\boldsymbol{A}$ to $\mathbb{R}$) (1.9) defines a norm.

[a] Alternatively, one could use the same approach taken for the norm $\|\cdot\|_\infty$ for $\mathbb{R}^n$. This requires proving that $\sup_{\boldsymbol{x} \in \boldsymbol{A}} |k| \cdot |f(\boldsymbol{x})| = |k| \cdot \sup_{\boldsymbol{x} \in \boldsymbol{A}} |f(\boldsymbol{x})|$ (see Exercise 1.15, where this point is addressed).

Example *1.16*

Let $a, b \in \mathbb{R}$, with $a < b$. Let

$$C([a, b]; \mathbb{R}) = \{f : [a, b] \longrightarrow \mathbb{R} \,|\, f \text{ is continuous}\}.$$

The set $C([a, b]; \mathbb{R})$ is a real vector space if we consider, in the usual way, the sum of $f, g \in C([a, b]; \mathbb{R})$ defined by $(f + g)(t) = f(t) + g(t)$ for all $t \in [a, b]$ and the scalar product by $(\lambda f)(t) = \lambda f(t)$ for all $t \in [a, b]$, where $\lambda \in \mathbb{R}$, and remembering that the sum of continuous functions is a continuous function, and the product of a continuous function by a real number is a continuous function. Given $f \in C([a, b]; \mathbb{R})$, we can define the norm

$$\|f\|_\infty \stackrel{\text{def}}{=} \max_{x \in [a,b]} |f(x)|. \tag{1.11}$$

This maximum is well defined from the Weierstrass extreme value theorem.[a] This is indeed a norm (see Exercise 1.14). This norm induces the distance

$$\begin{aligned} \mathrm{d} : C([a, b]; \mathbb{R}) \times C([a, b]; \mathbb{R}) &\longrightarrow \mathbb{R} \\ (f, g) &\mapsto \|f - g\|_\infty = \max_{x \in [a,b]} |f(x) - g(x)|. \end{aligned}$$

We can define also the norm

$$\|f\|_1 = \int_a^b |f(t)| \, \mathrm{d}t. \tag{1.12}$$

We verify that $\|f\|_1 = 0$ if and only if $f = 0$. It suffices to show that if $g \in C([a, b]; \mathbb{R})$ is a continuous and non-negative function such that

$$\int_a^b g(t) \, \mathrm{d}t = 0,$$

then $g(t) = 0$ for all $t \in [a, b]$. Suppose by contradiction that there exists $t_0 \in [a, b]$ such that $g(t_0) > 0$. From the continuity of g, there exists $\delta > 0$ such that for all $t \in [t_0 - \delta, t_0 + \delta]$ the inequality $g(t) \geq \frac{g(t_0)}{2}$ is satisfied (can you link this argument to the definition of continuity in terms of ε and δ? What is ε in this case?).

Let us consider a decomposition of the interval $[a, b]$ into $a = \tau_0 < \tau_1 < \ldots < \tau_n = b$ such that $\tau_{j+1} - \tau_j \leq \delta$ for all $j \in \{0, \ldots, n - 1\}$. From the definition of Riemann integral we have

$$0 = \int_a^b g(t) \, \mathrm{d}t \geq \sum_{j=1}^{n} \inf_{t \in [\tau_{j-1}, \tau_j]} g(t)(\tau_j - \tau_{j-1}).$$

It follows that for all $j \in \{1, \dots, n\}$ we have

$$\inf_{t \in [\tau_{j-1}, \tau_j]} g(t)(\tau_j - \tau_{j-1}) = 0,$$

which in turn gives $\inf_{t \in [\tau_{j-1}, \tau_j]} g(t) = 0$. Let $j_0 \in \{1, \dots, n\}$ be such that $t_0 \in [\tau_{j_0-1}, \tau_{j_0}]$, from which we obtain that for all $t \in [\tau_{j_0-1}, \tau_{j_0}]$ we have $g(t) \geq \frac{g(t_0)}{2} > 0$ since $\tau_{j_0} - \tau_{j_0-1} \leq \delta$. It follows that

$$\inf_{t \in [\tau_{j_0-1}, \tau_{j_0}]} g(t) \geq \frac{g(t_0)}{2} > 0,$$

which contradicts what we found above.

The proof of the sub-additivity and of the homogeneity follow directly from the properties of the Riemann integral. In fact

$$\begin{aligned} \|f + g\|_1 &= \int_a^b |(f+g)(t)| \, \mathrm{d}t = \int_a^b |f(t) + g(t)| \, \mathrm{d}t \leq \int_a^b (|f(t)| + |g(t)|) \, \mathrm{d}t \\ &= \int_a^b |f(t)| \, \mathrm{d}t + \int_a^b |g(t)| \, \mathrm{d}t = \|f\|_1 + \|g\|_1, \end{aligned}$$

and, for $k \in \mathbb{R}$,

$$\|k\, f\|_1 = \int_a^b |(k\, f)(t)| \, \mathrm{d}t = \int_a^b |k\, f(t)| \, \mathrm{d}t = \int_a^b |k| \, |f(t)| \, \mathrm{d}t = |k| \int_a^b |f(t)| \, \mathrm{d}t.$$

[a] The extreme value theorem states that if a real-valued function f is continuous and is defined over a compact set, then f must attain a maximum and a minimum, each at least once. A closed interval is an example of compact set in $\mathbb{R}$, as it is bounded and closed. Thus, a function f defined on $[a, b]$ will attain a maximum and a minimum, i.e., there exist numbers $c, d \in [a, b]$ such that $f(c) \geq f(x) \geq f(d)$ for all $x \in [a, b]$. Notice how the closedness of the set is essential: the function $f(x) = 1/x$ is continuous in the interval $(0, 3]$, but certainly does not have a maximum in $(0, 3]$.

Notice the analogy with the norms introduced in $\mathbb{R}^n$: the finite sums are replaced by the integral. The counterpart of the Euclidean norm is

$$\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 \, \mathrm{d}x}. \tag{1.13}$$

Remark 1.5 The space of bounded functions, even assuming the case where the domain is a closed interval $[a, b]$, cannot be equipped with the norm (1.12), nor with the norm (1.13). In fact, boundedness does not imply integrability, at least in the Riemann sense. For example, the function

$$\begin{aligned} f : [0, 1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases} \end{aligned}$$

is bounded and defined over a closed and bounded interval, but it is not Riemann integrable.

Example *1.17*

Consider the space $\mathbb{C}^{m\times n}$ of $m\times n$ matrices with complex entries. Given a matrix $A = (a_{i,j})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ in that space, we define

$$\begin{aligned} \|A\|_1 &= \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}| \\ \|A\|_2 &= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2} \\ \|A\|_\infty &= \max_{\substack{i=1,\dots,m\\ j=1,\dots,n}}|a_{i,j}|. \end{aligned}$$

These are all norms for the vector space $\mathbb{C}^{m\times n}$. Observe that, in essence, these norms are analogous to the vector norms introduced earlier applied to the vector in $\mathbb{C}^{m\cdot n}$

$$(a_{1,1}, a_{1,2}, \dots, a_{1,n}, a_{2,1}, a_{2,2}, \dots, a_{2,n}, \dots, a_{m,1}, a_{m,2}, \dots, a_{m,n}).$$

Remark 1.6 Of the three norms introduced in the previous example for the space $\mathbb{C}^{m\times n}$, the second plays a particularly important role. Indeed, $\|\cdot\|_2$ is referred to as *Frobenius norm*, or *Hilbert-Schmidt norm*, and from now on it will be denoted by $\|A\|_F$. It is easy to see that

$$\|A\|_F = \sqrt{\operatorname{trace}(A^*A)},$$

where A^* is the conjugate transpose of A (and if the matrix is real we can replace the conjugate transpose with the transpose).

Exercises

Exercise 1.7

Show that the distance associated with a norm via (1.6) enjoys two further properties:

Translation invariance: $\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x} \in \boldsymbol{X}$, $\quad \mathrm{d}(\boldsymbol{x}_1 + \boldsymbol{x}, \boldsymbol{x}_2 + \boldsymbol{x}) = \mathrm{d}(\boldsymbol{x}_1, \boldsymbol{x}_2)$;

Homogeneity: $\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$, $\forall\, k \in \mathbb{F}$, $\quad \mathrm{d}(k\,\boldsymbol{x}_1, k\,\boldsymbol{x}_2) = |k|\,\mathrm{d}(\boldsymbol{x}_1, \boldsymbol{x}_2)$.

[Solution: (1.7)]

Exercise 1.8

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let us define, for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ the function

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\|\boldsymbol{x} - \boldsymbol{y}\|}{1 + \|\boldsymbol{x} - \boldsymbol{y}\|}.$$

Show that this is a metric. In particular, show that this metric is non-homogeneous (i.e., it is not true that $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$, $\forall\, k \in \mathbb{F}$, we have $\mathrm{d}(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = |k|\,\mathrm{d}(\boldsymbol{x}, \boldsymbol{y})$) but it is translation invariant (i.e., $\forall\, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w} \in \boldsymbol{X}$ there holds $\mathrm{d}(\boldsymbol{x} + \boldsymbol{w}, \boldsymbol{y} + \boldsymbol{w}) = \mathrm{d}(\boldsymbol{x}, \boldsymbol{y})$). [Solution: (1.8)]

Exercise 1.9

Show that the modulus $|\cdot|$ in $\mathbb{C}$ is a norm in $\mathbb{C}$, both when considering $\mathbb{C}$ as a vector space over the field $\mathbb{R}$ and when $\mathbb{C}$ is considered as a vector space over the field $\mathbb{C}$. [Solution: (1.9)]

Exercise 1.10

Recall that, given $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we have defined

$$\|\boldsymbol{x}\|_1 \stackrel{\text{def}}{=} |x_1| + |x_2| + \ldots + |x_n| = \sum_{k=1}^{n} |x_k|.$$

Show that $\|\cdot\|_1$ is a norm. [Solution: (1.10)]

Exercise 1.11

Prove that the Euclidean norm

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}, \qquad \boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

is a norm in $\mathbb{R}^n$ (Hint: for the subadditivity, reason on $\|\boldsymbol{x} + \boldsymbol{y}\|$, where $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, and take into account the Cauchy-Schwarz inequality $\boldsymbol{x} \cdot \boldsymbol{y} \leq \|\boldsymbol{x}\| \, \|\boldsymbol{y}\|$. You can use the same argument employed to show that the Euclidean distance satisfies the axioms of a metric). [Solution: (1.11)]

Exercise 1.12

Let A be a non-empty set. The set of bounded functions from A to $\mathbb{R}$ can be defined as the set

$$B(A, \mathbb{R}) = \{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M \in \mathbb{R} : \quad \forall\, x \in A,\ |f(x)| \leq M\}.$$

Show that the following ones are equivalent definitions of $B(A, \mathbb{R})$:

(i) $\{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M \geq 0 : \quad \forall\, x \in A,\ |f(x)| \leq M\}$;

(ii) $\{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M > 0 : \quad \forall\, x \in A,\ |f(x)| \leq M\}$;

(iii) $\{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M \in \mathbb{R} : \quad \forall\, x \in A,\ |f(x)| < M\}$.

Show also that $B(A, \mathbb{R})$ is a real vector space with respect to the addition and scalar multiplication defined, respectively, as follows: let $f, g \in B(A, \mathbb{R})$ and $k \in \mathbb{R}$,

$$\begin{array}{rcl} f + g : A & \longrightarrow & \mathbb{R} \\ x & \mapsto & f(x) + g(x) \end{array} \qquad \text{and} \qquad \begin{array}{rcl} k\, f : A & \longrightarrow & \mathbb{R} \\ x & \mapsto & k\, f(x). \end{array}$$

Hint: to avoid checking all the axioms of a vector space, it is sufficient to show that $B(A, \mathbb{R})$ is a subspace of $\mathbb{R}^A = \{f : A \longrightarrow \mathbb{R}\}$. The latter is a vector space because if we have a vector space $\boldsymbol{V}$ over a field $\mathbb{F}$, then given a set A, the set $\boldsymbol{V}^A = \{f : A \longrightarrow \boldsymbol{V}\}$ is a vector space with respect to the operations of addition and multiplication

$$\begin{array}{rcl} f + g : A & \longrightarrow & \boldsymbol{V} \\ x & \mapsto & f(x) + g(x) \end{array} \qquad \text{and} \qquad \begin{array}{rcl} k\, f : A & \longrightarrow & \boldsymbol{V} \\ x & \mapsto & k\, f(x). \end{array}$$

[Solution: (1.12)]

Exercise 1.13

In the proof that $\|\cdot\|_\infty$ is a norm in the normed vector space $\mathbb{R}^n$, we used the fact that, given $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, we have

$$\max_{1 \leq i \leq n} (|\lambda|\, |x_i|) = |\lambda| \max_{1 \leq i \leq n} |x_i|.$$

Prove that this identity indeed holds. [Solution: (1.13)]

Exercise 1.14

Let $a, b \in \mathbb{R}$, with $a < b$. Consider the vector space

$$C([a, b], \mathbb{R}) = \{f : [a, b] \longrightarrow \mathbb{R} \,|\, f \text{ is continuous}\}$$

under the usual addition and scalar multiplication of functions. Let

$$\|f\|_\infty \stackrel{\text{def}}{=} \max_{x\in[a,b]} |f(x)|.$$

Show that this is a norm. [Solution: (1.14)]

Exercise 1.15 The proof of the homogeneity of the norm $\|\cdot\|_\infty$ for the normed vector space $B(A;\mathbb{R})$ in Example 1.15 can be alternatively carried out by generalizing the result of Exercise 1.7 to the case of the supremum. Show that

$$\sup_{x\in A} |k|\,|f(x)| = |k| \sup_{x\in A} |f(x)|,$$

where A is a generic set. [Solution: (1.15)]

Exercise 1.16 Let $\boldsymbol{V}$ be a vector space of dimension n over a field $\mathbb{F}$. Let $(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ be a basis of $\boldsymbol{V}$. Let $\boldsymbol{v} \in \boldsymbol{V}$, and if $\boldsymbol{v} = c_1\,\boldsymbol{v}_1 + \ldots + c_n\,\boldsymbol{v}_n$ for $c_1, \ldots, c_n \in \mathbb{F}$, we define

$$\|\boldsymbol{v}\| = |c_1| + |c_2| + \ldots + |c_n|.$$

Show that $\|\cdot\|$ is a norm. [Solution: (1.16)]

Exercise 1.17 Show that for all $\boldsymbol{x} \in \mathbb{R}^n$

$$\lim_{p\to+\infty} \|\boldsymbol{x}\|_p = \|\boldsymbol{x}\|_\infty.$$

[Solution: (1.17)]

Exercise 1.18 Show that the function

$$\begin{aligned} \|\cdot\| : C([0,1],\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto \max_{0\le x\le 1} |e^x f(x)| \end{aligned}$$

is a norm. [Solution: (1.18)]

Exercise 1.19 Let $w : [0,1] \longrightarrow \mathbb{R}$ be a strictly positive continuous function (i.e., $w(x) > 0$ for all $x \in [0,1]$). Consider the function

$$\begin{aligned} \|\cdot\| : C([0,1],\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto \sup_{0\le x\le 1} w(x)\,|f(x)|. \end{aligned}$$

Show that $\|\cdot\|$ is a norm, and establish whether it is equivalent to the norm $\|\cdot\|_\infty$. [Solution: (1.19)]

Exercise 1.20 Let $(\boldsymbol{V}, \|\cdot\|_v)$ and $(\boldsymbol{W}, \|\cdot\|_w)$ be normed spaces over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). The set $\boldsymbol{V}\times\boldsymbol{W}$ is a vector space over $\mathbb{F}$ with respect to the operations of addition and multiplication given by

$$(\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2) = (\boldsymbol{v}_1 \overset{V}{+} \boldsymbol{v}_2, \boldsymbol{w}_1 \overset{W}{+} \boldsymbol{w}_2), \qquad \boldsymbol{v}_1, \boldsymbol{v}_2 \in \boldsymbol{V},\ \boldsymbol{w}_1, \boldsymbol{w}_2 \in \boldsymbol{W},$$

and

$$\alpha\,(\boldsymbol{v}, \boldsymbol{w}) + (\alpha \overset{V}{\cdot} \boldsymbol{v}, \alpha \overset{W}{\cdot} \boldsymbol{w}), \qquad \boldsymbol{v} \in \boldsymbol{V},\ \alpha \in \mathbb{F},$$

where $\overset{V}{+}$ and $\overset{V}{\cdot}$ are the addition and scalar multiplication of $\boldsymbol{V}$ and $\overset{W}{+}$ and $\overset{W}{\cdot}$ are the addition and scalar multiplication of $\boldsymbol{W}$.

(i) Prove that $\boldsymbol{V} \times \boldsymbol{W}$ is a normed space with respect to the norm $\|(\boldsymbol{v}, \boldsymbol{w})\| \overset{\text{def}}{=} \|\boldsymbol{v}\|_v + \|\boldsymbol{w}\|_w$;

(ii) Prove that $\boldsymbol{V} \times \boldsymbol{W}$ is a normed space with respect to the norm $\|(\boldsymbol{v}, \boldsymbol{w})\| \overset{\text{def}}{=} \max\{\|\boldsymbol{v}\|_v, \|\boldsymbol{w}\|_w\}$.

[Solution: (1.20)]

Exercise 1.21

(i) Show that, if $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|$ is a norm, for all $k > 0$ the function $\boldsymbol{x} \mapsto k\,\|\boldsymbol{x}\|$ is also a norm;

(ii) If $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank} A = m$, if $\boldsymbol{y} \mapsto \|\boldsymbol{y}\|$ is a norm in $\mathbb{R}^m$, is $\boldsymbol{x} \mapsto \|A\,\boldsymbol{x}\|$ a norm in $\mathbb{R}^n$?

(iii) If $\|\boldsymbol{x}\|_{(1)}$ and $\|\boldsymbol{x}\|_{(2)}$ are norms, then $\alpha\,\|\boldsymbol{x}\|_{(1)} + \beta\|\boldsymbol{x}\|_{(2)}$ is a norm for $\alpha, \beta > 0$.

(iv) Which of the following is a norm?

- $\boldsymbol{x} \mapsto \min_i |x_i|$
- $\boldsymbol{x} \mapsto \min_i |x_i| + \max_i |x_i|$
- $(x_1, x_2) \mapsto |x_1 - x_2| + |x_1|$. How can we generalize it to $\mathbb{R}^3$?
- $(x_1, x_2) \mapsto |x_1| + \min\{|x_1|, |x_2|\}$.

[Solution: (1.21)]

Exercise 1.22

Consider the vectors $\boldsymbol{x} = (1, 2, -2, 0)$ and $\boldsymbol{y} = (2, 2, 0, -1)$ in $\mathbb{R}^4$. Determine the distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ by considering:

(i) the taxi-cub metric

(ii) the Euclidean metric

(iii) the discrete metric.

[Solution: (1.22)]

Exercise 1.23

Consider the vector $\boldsymbol{z} = (1 + 2\,i, -1 + i)$ in $\mathbb{C}^2$. Determine $\|\boldsymbol{z}\|_1$, $\|\boldsymbol{z}\|_2$, $\|\boldsymbol{z}\|_\infty$. [Solution: (1.23)]

Exercise 1.24

Consider the vector space $\mathbb{R}^n$, and the function

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \sum_{k=1}^{n} (-1)^k\,|x_k|. \end{aligned}$$

Is this function a norm? [Solution: (1.24)]

Exercise 1.25

Find which of the norm axioms are violated for the function

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, x_2, \ldots, x_n) &\mapsto \left(|x_1|^{\frac{1}{2}} + |x_2|^{\frac{1}{2}} + \ldots + |x_n|^{\frac{1}{2}}\right)^2. \end{aligned}$$

[Solution: (1.25)]

Exercise 1.26

Let $n \in \mathbb{N}^*$. Consider the vector space $\mathbb{R}^n$, and the function

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \sum_{j=1}^{n} j\,|x_j|. \end{aligned}$$

Establish whether this function a norm. [Solution: (1.26)]

Exercise 1.27

Let $\ell, a \in \mathbb{R}$. Show that the following statements are equivalent:

(i) $\ell \leq a$;

(ii) $\forall\, \varepsilon > 0,\ \ell < a + \varepsilon$.

Show that *(i)* is also equivalent to

(iii) $\forall\, \varepsilon > 0,\ \ell \leq a + \varepsilon$.

This means that *(ii)* and *(iii)* are equivalent: explain why. [Solution: (1.27)]

1.3 Basic topological concepts

The following concepts are a generalization of notions that you have seen in previous calculus courses.

Definition 1.5 [Ball]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$ and $r > 0$. We define

$$\mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X}) \stackrel{def}{=} \{\boldsymbol{x} \in \boldsymbol{X} \mid \|\boldsymbol{x} - \boldsymbol{x}_0\| < r\}.$$

The set $\mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$ is called open ball *(or, simply,* ball*) centered at $\boldsymbol{x}_0$ and radius r.*

When there is no ambiguity, we will denote the ball $\mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$ simply as $\mathcal{B}_r(\boldsymbol{x}_0)$.

Example 1.18

If $\boldsymbol{X} = \mathbb{R}$ (so that the underlying field is also $\mathbb{R}$) and the norm $|\cdot|$ is the absolute value (which is the Euclidean norm of $\mathbb{R}$), the open ball $\mathcal{B}_r(x_0)$ is the open interval $(x_0 - r, x_0 + r)$. If $\boldsymbol{X} = \mathbb{R}^2$ and the norm $\|\cdot\|$ is the Euclidean norm, the open ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is a disk in $\mathbb{R}^2$ comprising the points whose distance from the center $\boldsymbol{x}_0 \in \mathbb{R}^2$ is $< r$. Finally, If $\boldsymbol{X} = \mathbb{R}^3$ and the norm $\|\cdot\|$ is, again, the Euclidean norm of $\mathbb{R}^3$, the set $\mathcal{B}_r(\boldsymbol{x}_0)$ is the set of points within a sphere of radius r centered at $\boldsymbol{x}_0$. These points are proved in the solution to Exercise 1.11.

Example 1.19

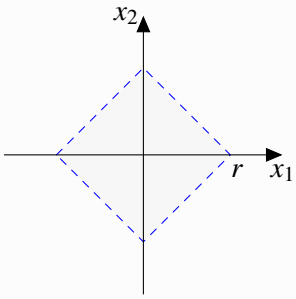
If $\boldsymbol{X} = \mathbb{R}^2$ and the norm $\|\cdot\|$ is the 1-norm $\|\cdot\|_1$, the open ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is a square in $\mathbb{R}^2$ centred at $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2})$. Indeed, considering for the sake of argument $\boldsymbol{x}_0 = (0, 0)$, we have

$$\mathcal{B}_r((0,0)) = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, |x_1| + |x_2| < r\},$$

and the condition $|x_1| + |x_2| < r$ can be written as

$$\begin{cases} x_1 \geq 0 \text{ and } x_2 \geq 0 \\ x_1 + x_2 < r \end{cases} \text{ or } \begin{cases} x_1 \geq 0 \text{ and } x_2 < 0 \\ x_1 - x_2 < r \end{cases} \text{ or } \begin{cases} x_1 < 0 \text{ and } x_2 \geq 0 \\ -x_1 + x_2 < r \end{cases} \text{ or } \begin{cases} x_1 < 0 \text{ and } x_2 < 0 \\ -x_1 - x_2 < r \end{cases}$$

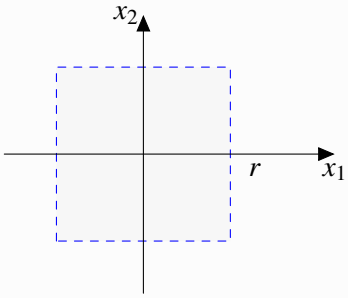
which leads to the square shown in the figure:



Example *1.20* If $\boldsymbol{X} = \mathbb{R}^2$ and the norm $\|\cdot\|$ is the ∞-norm $\|\cdot\|_\infty$, the open ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is a square in $\mathbb{R}^2$ centred at $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2})$. In fact, considering again for simplicity $\boldsymbol{x}_0 = (0,0)$, we have

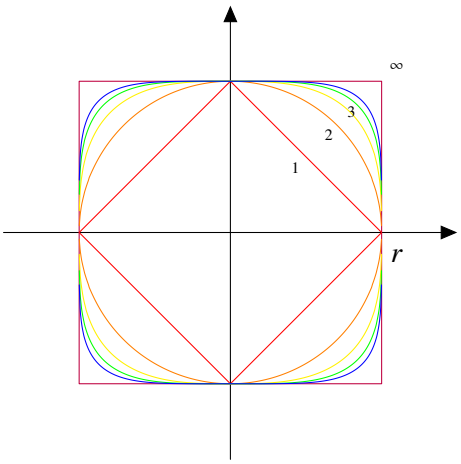
$$\mathcal{B}_r((0,0)) = \{(x_1, x_2) \in \mathbb{R}^2 \mid \max\{|x_1|, |x_2|\} < r\},$$

and the condition $\max\{|x_1|, |x_2|\} < r$ leads immediately to the square shown in the figure:



Generalizing the considerations of the previous examples, we see that in the following figure we represented the sets $\{(x_1, x_2) \in \mathbb{R}^2 \mid \|(x_1, x_2)\|_p = r\}$ for different values of p.

$$\begin{aligned}
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_1) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| + |x_2| < r\} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_2) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt{|x_1|^2 + |x_2|^2} < r\} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_3) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt[3]{|x_1|^3 + |x_2|^3} < r\} \\
&\vdots \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_\infty) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \max\{|x_1| + |x_2|\} < r\}
\end{aligned}$$



Remark 1.7 We recall the notion of convex set. First, consider a normed space $(\boldsymbol{X}, \|\cdot\|)$. Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. We define

$$[\boldsymbol{x}, \boldsymbol{y}] \stackrel{\text{def}}{=} \{\boldsymbol{x} + t\,(\boldsymbol{y} - \boldsymbol{x}) \,\big|\, t \in [0,1]\}.$$

We will call $[\boldsymbol{x}, \boldsymbol{y}]$ the **segment** with **end-points** $\boldsymbol{x}$ and $\boldsymbol{y}$. Notice that when $\boldsymbol{X} = \mathbb{R}$, we have $x, y \in \mathbb{R}$, and if $x \leq y$, the notation $[x, y]$ is also used to denote the closed interval with end-points x and y. However, these two notations are consistent. Given $\boldsymbol{A} \subseteq \boldsymbol{X}$, we say that $\boldsymbol{A}$ is **convex** if, for any pair of points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}$, the

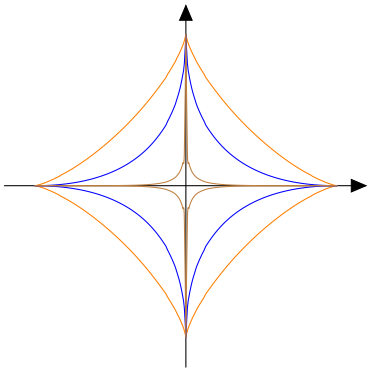
segment $[\boldsymbol{x}, \boldsymbol{y}]$ is in $\boldsymbol{A}$:

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}, \quad [\boldsymbol{x}, \boldsymbol{y}] \subseteq \boldsymbol{A}.$$

Now, given any $\boldsymbol{x}_0 \in \boldsymbol{X}$ and any $r > 0$, we want to show that the ball $\mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$ is convex, i.e., given any point z on the segment $[\boldsymbol{x}, \boldsymbol{y}]$, where $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$, we have $z \in \mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$. For every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$ and $t \in [0, 1]$ we have

$$\begin{aligned} \|z\| = \big\|\boldsymbol{x} + t\,(\boldsymbol{y} - \boldsymbol{x}) - \boldsymbol{x}_0\big\| &= \big\|(1-t)\,\boldsymbol{x} + t\,\boldsymbol{y} - (1-t)\,\boldsymbol{x}_0 - t\,\boldsymbol{x}_0\big\| \\ &= \big\|(1-t)\,(\boldsymbol{x} - \boldsymbol{x}_0) + t\,(\boldsymbol{y} - \boldsymbol{x}_0)\big\| \\ &\leq (1-t)\,\|\boldsymbol{x} - \boldsymbol{x}_0\| + t\,\|\boldsymbol{y} - \boldsymbol{x}_0\| < (1-t)\,r + t\,r = r. \end{aligned}$$

Thus, $[\boldsymbol{x}, \boldsymbol{y}] \subset \mathcal{B}_r(\boldsymbol{x}_0, \|\cdot\|, \boldsymbol{X})$. This argument hinges on the triangle inequality. In the case of the p-norm, such inequality does not hold when $p < 1$ (and therefore, if $p < 1$, the definition that we gave for the p-norm does not represent a norm). In that case, balls are not convex sets. The sets $\{(x_1, x_2) \in \mathbb{R}^2 \,|\, \|(x_1, x_2)\|_p = r\}$ for different values of $p < 1$ are shown below:



Example *1.21*

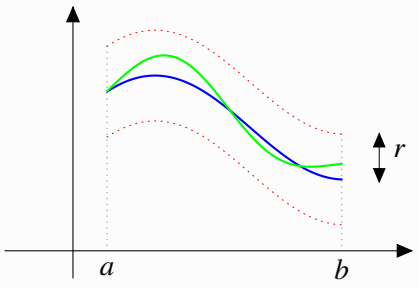
Let $a, b \in \mathbb{R}$, with $a < b$. Consider the real vector space

$$C([a,b]; \mathbb{R}) = \{f : [a,b] \longrightarrow \mathbb{R} \,|\, f \text{ is continuous}\}$$

with the infinity norm defined in (1.11). Given $f \in C([a,b]; \mathbb{R})$ and $r > 0$, the ball

$$\begin{aligned} \mathcal{B}_r(f) &= \{g \in C([a,b]; \mathbb{R}) \;\mid\; \|f - g\|_\infty < r\} \\ &= \{g \in C([a,b]; \mathbb{R}) \;\mid\; \max_{x \in [a,b]} |f(x) - g(x)| < r\} \end{aligned}$$

consists of all continuous functions whose values are within r of the values of f at every $x \in [a,b]$. For example, a function whose graph from f (in blue in the figure) with respect to the max-norm is less that r has a graph which lies inside the dotted red lines $f(x) \pm r$:



We now consider the notion of *neighbourhood*, that we adapt for normed spaces.

Definition 1.6 [Neighbourhood]
Let $(X, \|\cdot\|)$ be a normed space. Let $x_0 \in X$ and let $U \subseteq X$. We say that U is a neighbourhood *of x_0 if there exists $r > 0$ such that $\mathcal{B}_r(x_0) \subseteq U$. We denote by $\mathcal{N}(x_0, X, \|\cdot\|)$ the set of neighborhoods of x_0.*

When there is no ambiguity, we will denote $\mathcal{N}(x_0, X, \|\cdot\|)$ simply as $\mathcal{N}(x_0)$.

Example *1.22*

Let $(X, \|\cdot\|)$ be a normed space. Let $x_0 \in X$. Let $\mathcal{I}$ be a set of indices (typically set of indices is a subset of $\mathbb{N}$, but it can be any set). Consider a set

$$\{U_i \subseteq X \mid i \in \mathcal{I}\}.$$

Let U_i be a neighborhood of x_0 for every $i \in \mathcal{I}$. Then:

- the union $\bigcup_{i \in \mathcal{I}} U_i$ is a neighborhood of x_0
- if $\mathcal{I}$ is finite, the intersection $\bigcap_{i \in \mathcal{I}} U_i$ is a neighborhood of x_0.

See Exercise 1.14.

Definition 1.7 [Interior Point – Interior – Open Set]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Let $x_0 \in X$. We say that x_0 is an **interior point** *of A if there exists $r > 0$ such that the ball $\mathcal{B}_r(x_0)$ is contained in A, i.e., if*

$$\exists r > 0 : \quad \mathcal{B}_r(x_0) \subseteq A. \tag{1.14}$$

The interior *of A is the set of interior points of A, and it is denoted by* $\operatorname{int}(A, X, \|\cdot\|)$*, or, when the underlying normed space is clear from the context, by $\mathring{A}$. We say that A is* **open** *if $A = \mathring{A}$.*

Notice that (1.14) is equivalent to writing $A \in \mathcal{N}(x_0)$. As a consequence, a subset of a normed space is open if and only if every point in the set has a neighborhood that is contained in the set.

Example *1.23*

Let $(X, \|\cdot\|)$ be a normed space. Let $x_0 \in X$, and $r > 0$. We want to prove that the ball

$$\mathcal{B}_r(x_0) = \{x \in X \mid \|x - x_0\| < r\}$$

is open. To this end, we have to show that given any point $x \in \mathcal{B}_r(x_0)$, we can find a radius $\rho > 0$ such that the ball $\mathcal{B}_\rho(x)$ is contained in $\mathcal{B}_r(x_0)$. Let us take for example

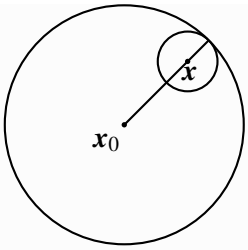
$$\rho = r - \|x - x_0\|,$$

which is > 0 since $\|x - x_0\| < r$. We need to show that every $y \in \mathcal{B}_\rho(x)$ is also in the ball $\mathcal{B}_r(x_0)$. Take $y \in \mathcal{B}_\rho(x)$, so that it satisfies

$$\|y - x\| < \rho = r - \|x - x_0\|. \tag{1.15}$$

We must prove that $\|y - x_0\| < r$. We add and subtract x, and use the triangle inequality:

$$\|y - x_0\| = \|y - x + x - x_0\| = \|(y - x) + (x - x_0)\| \leq \|y - x\| + \|x - x_0\| < r - \|x - x_0\| + \|x - x_0\| = r,$$



where we have used (1.15).

Example *1.24*

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$, and $r > 0$. Let

$$\boldsymbol{A} = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}.$$

This set is called *closed ball* centred at $\boldsymbol{x}_0$ and radius r (in the sequel, this object will be represented with the symbol $\overline{\mathcal{B}_r(\boldsymbol{x}_0)}$). We show that $\boldsymbol{A}$ is not open. Let $\boldsymbol{x} \in \boldsymbol{X}$ be such that $\|\boldsymbol{x} - \boldsymbol{x}_0\| = r$, so that $\boldsymbol{x} \in \boldsymbol{A}$. However, $\boldsymbol{x} \notin \mathring{\boldsymbol{A}}$. In fact, let $\rho > 0$. We must find $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$ such that $\boldsymbol{y} \notin \mathcal{B}_r(\boldsymbol{x}_0)$. We assume that $\boldsymbol{y}$ is in the form

$$\boldsymbol{y} = \boldsymbol{x}_0 + t\,(\boldsymbol{x} - \boldsymbol{x}_0), \qquad t > 0. \tag{1.16}$$

We want to impose the condition $\boldsymbol{y} \notin \mathcal{B}_r(\boldsymbol{x}_0)$, which is equivalent to requiring $\|\boldsymbol{x}_0 - \boldsymbol{y}\| > r$. We have

$$\|\boldsymbol{x}_0 - \boldsymbol{y}\| = t\,\|\boldsymbol{x} - \boldsymbol{x}_0\| = t\,r,$$

and this is $> r$ if we choose $t > 1$. Now we impose the condition $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$, which is equivalent to requiring $\|\boldsymbol{x} - \boldsymbol{y}\| < \rho$, and we want to see if we can find a value $t > 1$ consistent with this requirement: in this case, we will have found $\boldsymbol{y}$ satisfying simultaneously the conditions $\boldsymbol{y} \notin \mathcal{B}_r(\boldsymbol{x}_0)$ and $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$. Using (1.16) we find

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \|(t-1)\,(\boldsymbol{x} - \boldsymbol{x}_0)\| = (t-1)\,r$$

which is $< \rho$ if we choose

$$t < 1 + \frac{\rho}{r}.$$

Therefore, if

$$\boldsymbol{y} = \boldsymbol{x}_0 + \left(1 + \frac{\rho}{2\,r}\right)(\boldsymbol{x} - \boldsymbol{x}_0)$$

we have $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$, but $\boldsymbol{y} \notin \mathcal{B}_r(\boldsymbol{x})$.

Definition 1.8 [Boundary]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $\boldsymbol{x}_0 \in \boldsymbol{X}$. We say that $\boldsymbol{x}_0$ is a **boundary point** *of $\boldsymbol{A}$ if every open ball $\mathcal{B}_r(\boldsymbol{x}_0)$, with $r > 0$, contains both elements of $\boldsymbol{A}$ and elements that are not in $\boldsymbol{A}$, i.e., if*

$$\forall\, r > 0, \qquad \mathcal{B}_r(\boldsymbol{x}_0) \cap \boldsymbol{A} \neq \varnothing \;\; \textit{and} \;\; \mathcal{B}_r(\boldsymbol{x}_0) \cap (\boldsymbol{X} \setminus \boldsymbol{A}) \neq \varnothing. \tag{1.17}$$

We denote by $\partial\boldsymbol{A}$ the set of boundary points of $\boldsymbol{A}$.

The set $\boldsymbol{X} \setminus \boldsymbol{A}$ in (1.17) is the so-called complement of $\boldsymbol{A}$ in $\boldsymbol{X}$, and it is also denoted by $\complement_{\boldsymbol{X}}(\boldsymbol{A})$, or, when the set $\boldsymbol{X}$ is clear from the context, by $\boldsymbol{A}^{\complement}$.

Example *1.25*

Consider the case where $\boldsymbol{X}$ is $\mathbb{R}$ and the norm is the absolute value. Let $A = [0, 1)$. Then, $\partial A = \{0, 1\}$. From this example, it is clear that the elements of ∂A may belong or not to A.

Example *1.26* Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} = \mathcal{B}_r(\boldsymbol{x}_0)$ with $\boldsymbol{x}_0 \in \boldsymbol{X}$ and $r > 0$. We show that

$$\partial \boldsymbol{A} = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| = r\}.$$

In fact, let $\boldsymbol{x} \in \boldsymbol{X}$. We show that, if $\boldsymbol{x} \in \partial \boldsymbol{A}$, then $\|\boldsymbol{x} - \boldsymbol{x}_0\| = r$. If $\|\boldsymbol{x} - \boldsymbol{x}_0\| < r$, in a previous example we showed that there exists $\rho > 0$ such that $\mathcal{B}_\rho(\boldsymbol{x}) \subseteq \boldsymbol{A}$, so that $\boldsymbol{x} \notin \partial \boldsymbol{A}$. Let $\|\boldsymbol{x} - \boldsymbol{x}_0\| > r$, and we show that we can find a value $\rho > 0$ such that $\mathcal{B}_\rho(\boldsymbol{x}) \cap \boldsymbol{A} = \emptyset$. Let $\rho > 0$ and $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$. Then

$$\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y} - \boldsymbol{x}_0\| < \rho + \|\boldsymbol{y} - \boldsymbol{x}_0\|,$$

so that

$$\|\boldsymbol{y} - \boldsymbol{x}_0\| > \|\boldsymbol{x} - \boldsymbol{x}_0\| - \rho,$$

and this quantity is $\geq r$ if $\rho \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| - r$. With this choice of ρ (which is legitimate because $\|\boldsymbol{x} - \boldsymbol{x}_0\| - r$ is strictly positive) we have $\mathcal{B}_\rho(\boldsymbol{x}) \cap \boldsymbol{A} = \emptyset$, so that $\boldsymbol{x} \notin \partial \boldsymbol{A}$.

It follows that if $\boldsymbol{x} \in \partial \boldsymbol{A}$, we must have $\|\boldsymbol{x} - \boldsymbol{x}_0\| = r$. However, this does not conclude the argument: there could still be points satisfying $\|\boldsymbol{x} - \boldsymbol{x}_0\| = r$ which do not belong to $\partial \boldsymbol{A}$. We must show that this cannot occur.

Let $\|\boldsymbol{x} - \boldsymbol{x}_0\| = r$. We show that $\boldsymbol{x} \in \partial \boldsymbol{A}$. Let $\rho > 0$. We must show that $\mathcal{B}_\rho(\boldsymbol{x})$ contains both elements in $\boldsymbol{A}$ and elements that are not in $\boldsymbol{A}$. First, we impose the condition $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$. Let

$$\boldsymbol{y} = \boldsymbol{x}_0 + t\,(\boldsymbol{x} - \boldsymbol{x}_0).$$

We have

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \|(1 - t)\,(\boldsymbol{x} - \boldsymbol{x}_0)\| = |1 - t|\, r,$$

which is $< \rho$ if $|1 - t| < \frac{\rho}{r}$. This is equivalent to

$$1 - \frac{\rho}{t} < t < 1 + \frac{\rho}{r},$$

which is equivalent to requiring $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$. Now, we want to find conditions for which $\boldsymbol{y} \in \boldsymbol{A}$ and for which $\boldsymbol{y} \not\in \boldsymbol{A}$. Recalling that $\boldsymbol{A} = \mathcal{B}_r(\boldsymbol{x}_0)$,

$$\|\boldsymbol{x}_0 - \boldsymbol{y}\| = \|t\,(\boldsymbol{x} - \boldsymbol{x}_0)\| = |t|\, r.$$

Hence, $\boldsymbol{y} \in \boldsymbol{A}$ if $|t| < 1$, whereas $\boldsymbol{y} \notin \boldsymbol{A}$ if $|t| \geq 1$. Hence, if $t = 1$, we have $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x})$ and $\boldsymbol{y} \notin \boldsymbol{A}$, if $t = \max\left\{1 - \frac{\rho}{2r}, \frac{1}{2}\right\}$ we have $\boldsymbol{y} \in \mathcal{B}_\rho(\boldsymbol{x}) \cap \boldsymbol{A}$.

The following result presents an important property of open sets: an open set does not have any point in common with its boundary.

Theorem 1.5 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. The following statements are equivalent:*

(1) *$\boldsymbol{A}$ is open;*

(2) *$\boldsymbol{A} \cap \partial \boldsymbol{A} = \emptyset$.*

Proof Let $\boldsymbol{A}$ be open, and let $\boldsymbol{x} \in \boldsymbol{A}$. Then, $\boldsymbol{x} \in \mathring{\boldsymbol{A}}$. Thus, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{A}$. It follows that $\boldsymbol{x} \notin \partial \boldsymbol{A}$, so that $\boldsymbol{A} \cap \partial \boldsymbol{A} = \emptyset$.

Conversely, assume that $\boldsymbol{A} \cap \partial \boldsymbol{A} = \emptyset$, and we prove that $\boldsymbol{A}$ is open. Let $\boldsymbol{x} \in \boldsymbol{A}$. Then, since $\boldsymbol{A} \cap \partial \boldsymbol{A} = \emptyset$, we have $\boldsymbol{x} \notin \partial \boldsymbol{A}$. It follows that there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{A}$, or $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{X} \setminus \boldsymbol{A}$. The second cannot be true, because $\boldsymbol{x} \in \boldsymbol{A}$. We can conclude that every element of $\boldsymbol{A}$ belongs to $\mathring{\boldsymbol{A}}$, and this says that $\boldsymbol{A}$ is open.

Theorem 1.6 *Let $(X, \|\cdot\|)$ be a normed space. Then*

(i) The union of an arbitrary family of open sets is open;

(ii) The intersection of a finite family of open sets is open.

Proof Let $\mathcal{F} = \{A_i \subseteq X \mid i \in \mathcal{I}\}$ be a family of open sets, where $\mathcal{I} \subseteq \mathbb{N}$ is a set of indices, so that A_i is open for every $i \in \mathcal{I}$. We show that $\bigcup_{i\in\mathcal{I}} A_i$ is open. Let $x_0 \in \bigcup_{i\in\mathcal{I}} A_i$. We show that there exists $r > 0$ such that $\mathcal{B}_r(x_0) \subseteq \bigcup_{i\in\mathcal{I}} A_i$. Since $x_0 \in \bigcup_{i\in\mathcal{I}} A_i$, there exists $i \in \mathcal{I}$ such that $x_0 \in A_i$. Since A_i is open, there exists $r > 0$ such that

$$\mathcal{B}_r(x_0) \subseteq A_i \subseteq \bigcup_{i\in\mathcal{I}} A_i,$$

as required.

We now prove the second point. Let $p \in \mathbb{N}^*$ and let $\mathcal{F} = \{A_1, A_2, \ldots, A_p\}$ be a finite family of open sets. We show that $\bigcap_{i=1}^{p} A_i$ is open. Let $x_0 \in \bigcap_{i=1}^{p} A_i$, so that $x_0 \in A_i$ for all $i \in \{1, \ldots, p\}$. Since A_i is open, there exists $r_i > 0$ such that $\mathcal{B}_{r_i}(x_0) \subseteq A_i$, and therefore, taking $r = \min\{r_1, \ldots, r_p\}$, we find $\mathcal{B}_r(x_0) \subseteq \mathcal{B}_{r_i}(x_0) \subseteq A_i$ for all $i \in \{1, \ldots, p\}$. We conclude that

$$\mathcal{B}_r(x_0) \subseteq \bigcap_{i=1}^{p} A_i,$$

which shows that the intersection is open.

The previous proof fails if we consider the intersection of infinitely many open sets. An example is $X = \mathbb{R}$ and

$$\bigcap_{n\in\mathbb{N}^*} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}.$$

Each interval in this intersection is open, but the result of the intersection is $\{0\}$ (why?), which is not open (why?).

We now define the notion of closed set. Loosely speaking, a closed set is a set that contains its boundary.

Definition 1.9 [Closed Set]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. We say that A is **closed** *if $\partial A \subseteq A$.*

Example *1.27*

Consider the case where $X = \mathbb{R}$ and the norm is the absolute value. From what we have seen, the interval $[0, 1)$ is not closed. On the other hand, $[0, 1]$ is closed, since $\partial([0, 1]) = \{0, 1\} \subset [0, 1]$.

Example *1.28*

Consider a normed space $(X, \|\cdot\|)$. The empty set $\varnothing$ and the vector space X are both simultaneously open and closed.

Let us start by considering X. For every $x_0 \in \mathbb{R}^n$ and for every $r > 0$, the points $x \in \mathcal{B}_r(x_0)$ are also points of X.

The boundary of X is empty for the same reason: given $x_0 \in X$, for every $r > 0$ we have $\mathcal{B}_r(x_0) \cap (X \setminus X) = \mathcal{B}_r(x_0) \cap \varnothing = \varnothing$. Therefore, X is also closed, because $\partial X = \varnothing \subseteq X$ (the empty set is a subset of any set).

We now consider the empty set $\varnothing$, which satisfies the definition of open set vacuously: since it has no

points, every point has a neighborhood contained in the set. We can be more explicit, and consider that, to be open, $\varnothing$ needs to satisfy by definition

$$\forall \boldsymbol{x} \in \varnothing, \quad \exists r > 0 : \ \mathcal{B}_r(\boldsymbol{x}) \subseteq \varnothing,$$

which is the same as the logical implication

$$\boldsymbol{x} \in \varnothing \quad \Rightarrow \quad \exists r > 0 : \ \mathcal{B}_r(\boldsymbol{x}) \subseteq \varnothing. \tag{1.18}$$

The truth table of the implication is

p	q	$p \Rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

We see from this table that if p is false, no matter whether q is true or false, $p \Rightarrow q$ is vacuously true. Therefore, since $\boldsymbol{x} \in \varnothing$ is false, we conclude that (1.18) is true. This shows that $\varnothing$ is open.

To find the boundary of $\varnothing$, we reason in this way. A point $\boldsymbol{x}$ is in $\partial\varnothing$ if for all $r > 0$, we have

$$\mathcal{B}_r(x) \cap \varnothing \neq \varnothing \quad \text{and} \quad \mathcal{B}_r(x) \cap (\boldsymbol{X} \setminus \varnothing) \neq \varnothing.$$

The first is certainly false, because $\mathcal{B}_r(x) \cap \varnothing = \varnothing$. Therefore, there are no points in $\partial\varnothing$, i.e., $\partial\varnothing = \varnothing$.

To see if $\varnothing$ is closed, we ask ourselves if it contains its boundary, which is also equal to $\varnothing$. Since $\varnothing \subseteq \varnothing$, then $\varnothing$ is also closed.

It is possible to prove that, if the normed space is finite dimensional, $\boldsymbol{X}$ and $\varnothing$ are the only subsets of $\boldsymbol{X}$ which are simultaneously closed and open (even if this fact ceases to be true if we re-state all these notions for metric spaces: see Exercise 1.29).

Example *1.29*

Let $\boldsymbol{x}_0 \in \mathbb{R}^n$ and $r > 0$. Let

$$\boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}.$$

To show that $\boldsymbol{A}$ is closed, it is sufficient to prove that $\partial\boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| = r\}$ (try!).

The following result gives an equivalent characterization of closed sets, which is sometimes used as a definition in place of Definition 1.9: a subset of a normed space is closed if and only if its complement is open.

Theorem 1.7 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{A}$ is closed if and only if $\boldsymbol{X} \setminus \boldsymbol{A}$ is open.*

The proof is left as an exercise, see Exercise 1.33.

Definition 1.10 [Closure]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. The set

$$cl(\boldsymbol{A}, \boldsymbol{X}, \|\cdot\|)) \stackrel{def}{=} \boldsymbol{A} \cup \partial\boldsymbol{A}$$

is called **closure** *of $\boldsymbol{A}$. If the normed space $(\boldsymbol{X}, \|\cdot\|)$ is clear from the context, we denote is also by $\overline{\boldsymbol{A}}$.*

Remark 1.8 We recall that $\boldsymbol{A}$ is open if, by definition, $\boldsymbol{A} = \mathring{\boldsymbol{A}}$. It is not difficult to see that $\boldsymbol{A}$ is closed if and only if $\boldsymbol{A} = \overline{\boldsymbol{A}}$, see Exercise 1.32. From this we can conclude that

$$\overline{\mathcal{B}_r(\boldsymbol{x}_0)} = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq r\}.$$

Theorem 1.8 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\partial\boldsymbol{A}$ and $\overline{\boldsymbol{A}}$ are closed subsets of $\boldsymbol{X}$.*

Proof We prove that $\partial\boldsymbol{A}$ is closed, which, by definition of closed set, is equivalent to showing that $\partial(\partial\boldsymbol{A}) \subseteq \partial\boldsymbol{A}$. Let $\boldsymbol{x}_0 \in \partial(\partial\boldsymbol{A})$. We must show that $\boldsymbol{x}_0 \in \partial\boldsymbol{A}$. Let $r > 0$. Then, there exists $\boldsymbol{x}_1 \in \partial\boldsymbol{A} \cap \mathcal{B}_r(\boldsymbol{x}_0)$. Since $\mathcal{B}_r(\boldsymbol{x}_0)$ is open, there exists $r_0 > 0$ such that $\mathcal{B}_{r_0}(\boldsymbol{x}_1) \subseteq \mathcal{B}_r(\boldsymbol{x}_0)$. Since $\boldsymbol{x}_1 \in \partial\boldsymbol{A}$, the set $\mathcal{B}_{r_0}(\boldsymbol{x}_1)$ contains both points of $\boldsymbol{A}$ and points of $\boldsymbol{X} \setminus \boldsymbol{A}$. These same elements belong also to $\mathcal{B}_r(\boldsymbol{x}_0)$. Thus, for every $r > 0$, the ball $\mathcal{B}_r(\boldsymbol{x}_0)$ contains both elements of $\boldsymbol{A}$ and of $\boldsymbol{X} \setminus \boldsymbol{A}$. We conclude that $\boldsymbol{x}_0 \in \partial\boldsymbol{A}$.

We now show that $\overline{\boldsymbol{A}}$ is closed, which, by definition of closed set, is equivalent to showing that $\partial\overline{\boldsymbol{A}} \subseteq \overline{\boldsymbol{A}}$. Let $\boldsymbol{x}_0 \in \partial\overline{\boldsymbol{A}}$. We show that $\boldsymbol{x}_0 \in \overline{\boldsymbol{A}} = \boldsymbol{A} \cup \partial\boldsymbol{A}$, which means that $\boldsymbol{x}_0$ is in $\boldsymbol{A}$ or in $\partial\boldsymbol{A}$ (or both). Only two cases are possible: $\boldsymbol{x}_0 \in \boldsymbol{A}$ or $\boldsymbol{x}_0 \notin \boldsymbol{A}$. Clearly, in the first case $\boldsymbol{x}_0 \in \boldsymbol{A} \cup \partial\boldsymbol{A} = \overline{\boldsymbol{A}}$, and there is nothing else to prove. In the other case $\boldsymbol{x}_0 \notin \boldsymbol{A}$, we must have $\boldsymbol{x}_0 \in \partial\boldsymbol{A}$. Let $r > 0$. Then, $\mathcal{B}_r(\boldsymbol{x}_0)$ contains at least one point that is not in $\boldsymbol{A}$: $\boldsymbol{x}_0$ itself. We show that it also contains some element of $\boldsymbol{A}$. Since $\boldsymbol{x}_0 \in \partial\overline{\boldsymbol{A}}$, the set $\mathcal{B}_r(\boldsymbol{x}_0) \cap \overline{\boldsymbol{A}}$ is not empty, and therefore there exists some vector $\boldsymbol{x}_1 \in \mathcal{B}_r(\boldsymbol{x}_0) \cap \overline{\boldsymbol{A}}$; since $\boldsymbol{x}_1 \in \overline{\boldsymbol{A}} = \boldsymbol{A} \cup \partial\boldsymbol{A}$, we can have $\boldsymbol{x}_1 \in \boldsymbol{A}$ or $\boldsymbol{x}_1 \in \partial\boldsymbol{A}$ (or both). If $\boldsymbol{x}_1 \in \boldsymbol{A}$, we found an element of $\boldsymbol{A}$ in $\mathcal{B}_r(\boldsymbol{x}_0)$. Now, suppose $\boldsymbol{x}_1 \in \partial\boldsymbol{A}$. Let $\rho > 0$ be such that $\mathcal{B}_\rho(\boldsymbol{x}_1) \subseteq \mathcal{B}_r(\boldsymbol{x}_0)$: this is always possible because $\boldsymbol{x}_1 \in \mathcal{B}_r(\boldsymbol{x}_0)$ and $\mathcal{B}_r(\boldsymbol{x}_0)$ is open. Since $\boldsymbol{x}_1 \in \partial\boldsymbol{A}$, the ball $\mathcal{B}_\rho(\boldsymbol{x}_1)$ must contain some element $\boldsymbol{x}_2$ in $\boldsymbol{A}$. This $\boldsymbol{x}_2$ is also contained in $\mathcal{B}_r(\boldsymbol{x}_0)$. We conclude that, in every case, $\mathcal{B}_r(\boldsymbol{x}_0)$ contains some elements of $\boldsymbol{A}$. Hence, $\boldsymbol{x}_0 \in \partial\boldsymbol{A}$, and the proof is concluded.

Theorem 1.9 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Then*

(i) The intersection of an arbitrary family of closed sets is closed;

(ii) The union of a finite family of closed sets is closed.

Proof We use de Morgan's Laws.[a] Consider a family $\mathcal{F} = \{\boldsymbol{A}_i \,|\, i \in \mathcal{I}\}$ of closed sets, where $\mathcal{I} \subseteq \mathbb{N}$ is an index set. We must show that the intersection $\bigcap_{i\in\mathcal{I}} \boldsymbol{A}_i$ is closed, or, equivalently, that the complement $\complement_{\boldsymbol{X}}\big(\bigcap_{i\in\mathcal{I}} \boldsymbol{A}_i\big)$ is open (see Theorem 1.7). We use de Morgan's law

$$\complement_{\boldsymbol{X}}\Big(\bigcap_{i\in\mathcal{I}} \boldsymbol{A}_i\Big) = \bigcup_{i\in\mathcal{I}} \complement_{\boldsymbol{X}}(\boldsymbol{A}_i),$$

and we recall that the union of a family of open sets is open.

Let $p \in \mathbb{N}^*$ and let $\mathcal{F} = \{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p\}$ be a finite family of closed sets. We must show that the union $\bigcup_{i=1}^{p} \boldsymbol{A}_i$ is closed, i.e., that the complement $\complement_{\boldsymbol{X}}\big(\bigcup_{i=1}^{p} \boldsymbol{A}_i\big)$ is open. Again, we use de Morgan's law

$$\complement_{\boldsymbol{X}}\Big(\bigcup_{i=1}^{p} \boldsymbol{A}_i\Big) = \bigcap_{i=1}^{p} \complement_{\boldsymbol{X}}(\boldsymbol{A}_i),$$

and we recall that the finite intersection of a family of open set is open.

[a]Let A, B, X be sets. Let $A \subseteq X$ and $B \subseteq X$. Denoting by $\complement_X(A)$ the complement of A in X, i.e., defining $\complement_X(A) = X \setminus A$, then:

(i) $\complement_X(A \cup B) = \complement_X(A) \cap \complement_X(B)$;

(ii) $\complement_X(A \cap B) = \complement_X(A) \cup \complement_X(B)$.

The previous proof fails if we consider the union of infinitely many closed sets. An example of a non-finite set of closed sets whose union is not closed is

$$\bigcup_{n \in \mathbb{N}^*} \left[\frac{1}{n}, 1\right] = (0, 1],$$

which obviously is not closed.

Theorem 1.10 *Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Then,*

$$\operatorname{int}(X \setminus A) = X \setminus \overline{A}.$$

Proof We have the following chain of equivalent statements:

$$x \in \operatorname{int}(X \setminus A) \iff \exists r > 0 : \mathcal{B}_r(x) \subseteq X \setminus A \iff x \notin A \text{ and } x \notin \partial A \iff x \notin \overline{A} \iff x \in X \setminus \overline{A}.$$

The following theorem asserts that the interior of a set A is the union of all open subsets of A, and, dually, the closure of A is the intersection of all closed sets containing A.

Theorem 1.11 *Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Then,*

$$\begin{aligned} \mathring{A} &= \bigcup \{G \subseteq A \mid G \text{ is open}\} \\ \overline{A} &= \bigcap \{F \supseteq A \mid F \text{ is closed}\}. \end{aligned}$$

Proof We define $S = \bigcup\{G \subseteq A \mid G \text{ is open}\}$. Let $x \in \mathring{A}$. We show that $x \in S$. By definition of interior point, there exists $r > 0$ such that $\mathcal{B}_r(x) \subseteq A$. Since $\mathcal{B}_r(x)$ is open, it is contained in S. Thus, $x \in S$.
Conversely, let $x \in S$. Then, there exists $G \in S$ such that $x \in G$, where $G \subseteq A$ and G is open. It follows that there exists $r > 0$ such that $\mathcal{B}_r(x) \subseteq G \subseteq A$, and this shows that x is an interior point of A, i.e., $x \in \mathring{A}$.

Definition 1.11 [Limit Points and Derived Set]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Let $x_0 \in X$. We say that x_0 is a **limit point** *or* **accumulation point** *for A if, for all $r > 0$, the set $A \cap \mathcal{B}_r(x_0, X, \|\cdot\|)$ contains some elements different from x_0, i.e.,*

$$\forall\, r > 0, \quad (A \cap \mathcal{B}_r(x_0, X, \|\cdot\|)) \setminus \{x_0\} \neq \varnothing.$$

The set of all limit points of A is denoted by $\mathcal{D}(A, X, \|\cdot\|)$, and is called **derived set**. *When there is no ambiguity, we will also use the symbol $\mathcal{D}(A)$.*

It is very easy to see that we can rephrase the condition in Definition 1.11 in terms of neighborhoods: $x_0 \in \mathcal{D}(A)$ if and only if, for any neighborhood U of x_0, we have $(U \cap A) \setminus \{x_0\} \neq \varnothing$.

Example *1.30*

Consider the case $n = 1$. Let $a < b$. Then

$$\mathcal{D}([a,b)) = [a,b].$$

In fact, let $x_0 < a$. Then, $x_0 \notin \mathcal{D}([a,b))$. Indeed, if $0 < r < a - x_0$, we have $(x_0 - r, x_0 + r) \cap [a,b) = \emptyset$.
Let $x_0 > b$. Then, $x_0 \in \mathcal{D}([a,b))$. Indeed, if $0 < r < x_0 - b$, then $(x_0 - r, x_0 + r) \cap [a,b) = \emptyset$.
Let $a \leq x_0 < b$. Then, $x_0 \in \mathcal{D}([a,b))$. Indeed, let $r > 0$. If $r \leq b - x_0$, we can take $x = x_0 + \frac{r}{2}$; then $x \in \mathcal{B}_r(x_0) = (x_0 - r, x_0 + r)$ and it is different from x_0. Moreover,

$$a \leq x_0 < x = x_0 + \frac{r}{2} \leq x_0 + \frac{b - x_0}{2} < x_0 + (b - x_0) = b.$$

It follows that $x \in [a,b)$. If $r > b - x_0$, take $r_1 > 0$ with $r_1 \leq b - x_0$. Then, from the previous argument, there exists x distinct from x_0 and which belongs to $[a,b) \cap \mathcal{B}_{r_1}(x_0)$. Since $\mathcal{B}_{r_1}(x_0) \subseteq \mathcal{B}_r(x_0)$, then x also belongs to $[a,b) \cap \mathcal{B}_r(x_0)$.
Finally, we show that $b \in \mathcal{D}([a,b))$, Let $r > 0$. If $r \leq b - a$, then $x = b - \frac{r}{2}$ is distinct from b, and it belongs to $[a,b) \cap \mathcal{B}_r(x_0) = [a,b) \cap (x_0 - r, x_0 + r)$. If $r > b - a$, we repeat the same argument above: we take $r_1 > 0$ with $r_1 \leq b - a$ and we take x in $[a,b) \cap \mathcal{B}_{r_1}(b) = [a,b) \cap (b - r_1, b + r_1)$ distinct from b. Then, $x \in [a,b) \cap \mathcal{B}_r(b)$.

Theorem 1.12 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{A}$ is closed if and only if $\boldsymbol{A}$ contains all its limit points.*

The proof of this result is left as an exercise (see Exercise 1.28).

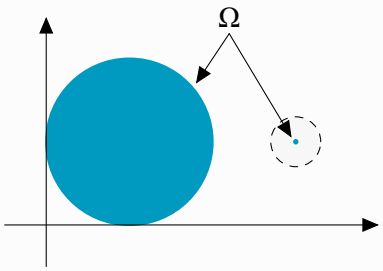
We introduce the notion of isolated point of a set. Intuitively, $\boldsymbol{x}_0$ is an isolated point of $\boldsymbol{A}$ if there exists a neighbourhood of $\boldsymbol{x}_0$ whose intersection with $\boldsymbol{A}$ is just $\boldsymbol{x}_0$. Expressing this condition in terms of a neighborhood or in terms of a ball is clearly the same: a ball centered at $\boldsymbol{x}_0$ is a neighborhood of $\boldsymbol{x}_0$, and a neighborhood of $\boldsymbol{x}_0$ always contains a ball centered at $\boldsymbol{x}_0$ entirely contained in that neighborhood.

Definition 1.12 [Isolated Point]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. We say that $\boldsymbol{x}_0$ is an **isolated point** *of $\boldsymbol{A}$ if there exists $r > 0$ such that*

$$\mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|) \cap \boldsymbol{A} = \{\boldsymbol{x}_0\}.$$

Example *1.31*

$\boldsymbol{A} = \mathcal{B}_1((1,1)) \cup \{(3,1)\}$ in $\mathbb{R}^2$.



The point $\boldsymbol{x}_0 = (3,1)$ is an isolated point: we can find a neighbourhood (dotted circle in the figure) which contains $(3,1)$ but no other element of $\boldsymbol{A}$. The condition $\boldsymbol{x}_0 \in \overline{\boldsymbol{A} \setminus \{\boldsymbol{x}_0\}}$ is not satisfied: $\boldsymbol{A} \setminus \{\boldsymbol{x}_0\} = \mathcal{B}_1((1,1))$, and its closure is $\{(x,y) \in \mathbb{R}^2 \mid \|(x,y) - (1,1)\| \leq 1\}$, which does not contain $\boldsymbol{x}_0$.

The set of accumulation points of $\boldsymbol{A}$ is the closed ball $\{(x,y) \in \mathbb{R}^2 \mid \|(x,y) - (1,1)\| \leq 1\}$.

Definition 1.13 [Adherent Point]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Let $x_0 \in X$. We say that x_0 is an adherent point *or* point of closure *of A if for all $r > 0$*

$$\mathcal{B}_r(x_0, X, \|\cdot\|) \cap A \neq \emptyset.$$

Remark 1.9 Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. The set of adherent points of A coincides with $\overline{A}$, see Exercise 1.25.

Remark 1.10 If $(X, \|\cdot\|)$ is a normed space, and we define $A \subseteq B \subseteq X$, we say that the set A is closed in B if every adherent point of A in B is contained in A, i.e., for all $x_0 \in B$,

$$\exists\, r > 0 : \ \mathcal{B}_r(x_0, X, \|\cdot\|) \cap A \neq \emptyset \qquad \Rightarrow \qquad x_0 \in A.$$

For example, the set $A = [0, 1)$ is closed in $B = (-1, 1)$. In fact, 1 is not in B.

Definition 1.14 [Exterior Point – Exterior]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. Let $x_0 \in X$. We say that x_0 is an exterior point *of A if there exists $r > 0$ such that*

$$\mathcal{B}_r(x_0, X, \|\cdot\|) \cap A = \emptyset.$$

We denote by $\mathrm{ext}(A, X, \|\cdot\|)$ *the set of exterior points of A (or simply by* $\mathrm{ext}(A)$ *when the normed space is clear from the context). The set* $\mathrm{ext}(A, X, \|\cdot\|)$ *is also called the* exterior *of A.*

Definition 1.15 [Dense Subset]
Let $(X, \|\cdot\|)$ be a normed space. Let $A \subseteq X$. We say that A is dense *in X if its closure coincides with X.*

Hence, if A is dense in X, we have $\overline{A} = X$. This implies that every point of X is an adherent point of A. From the definition of adherent point, we see that, if A is dense in X, for every $x \in X$ and for every $r > 0$ there exists $a \in A$ such that $x \in \mathcal{B}_r(a)$. Stated differently, if $\overline{A} = X$

$$\forall\, x \in X, \ \ \forall\, \varepsilon > 0, \ \ \exists\, a \in A : \ \ \|x - a\| < \varepsilon.$$

Intuitively, this means that every point of X has a point of A arbitrarily close to it.

Example *1.32*

The set of rational numbers $\mathbb{Q}$ is dense in $\mathbb{R}$, i.e.,

$$\forall\, x \in \mathbb{R}, \ \forall\, \epsilon > 0, \ \exists\, q \in \mathbb{Q} : \ |x - q| < \epsilon.$$

To see this, we use the so-called Archimede's axiom, which says that given an arbitrary real number $a \in \mathbb{R}$, there exists a natural number n greater than a. Let $x \in \mathbb{R}$ and $\epsilon > 0$. From this axiom, there exists $m \in \mathbb{N}^*$ such that $m > \frac{1}{\varepsilon}$, so that $\frac{1}{m} < \epsilon$. Let $n = \lfloor m\, x \rfloor$, which means that n is the greatest integer less than or equal to $m\, x$. Thus, $n \leq m\, x \leq n + 1$. Dividing through by m we obtain

$$\frac{n}{m} \leq x \leq \frac{n}{m} + \frac{1}{m},$$

and therefore we have also

$$\frac{n}{m} - \frac{1}{m} \leq x \leq \frac{n}{m} + \frac{1}{m}.$$

This inequality can be also written asy $\left|x - \frac{n}{m}\right| < \frac{1}{m}$. Thus, from $\frac{1}{m} < \epsilon$ we have $\left|x - \frac{n}{m}\right| < \epsilon$. This implies that an arbitrary real number has a rational number arbitrarily close to it (or is itself a rational number). In other words, a real number can be approximated with arbitrary precision with a rational number. For example, when we write $\sqrt{2} = 1.4142...$, the dots suggest that, continuing with the decimal digits, we can approach $\sqrt{2}$ with the precision that we want.

The density of $\mathbb{Q}$ in $\mathbb{R}$ can also be used to show that every real number is the limit of a convergent sequence of rational numbers. Even though the notions of convergence of a sequence will be formally introduced in the sequel, the case of sequences of scalars should be familiar from previous Calculus units. Let x be a real number. Let $k \in \mathbb{N}^*$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists $q_k \in \mathbb{Q}$ which satisfies

$$x - \frac{1}{k} < q_k < x + \frac{1}{k}.$$

In this way, we can construct a sequence $(q_k)_{k\in\mathbb{N}^*}$, and we show that it converges to x. Let $\varepsilon > 0$. Let p be a natural number greater than $1/\varepsilon$. For any $k > p$, we have $k > p > 1/\varepsilon$, so that $\frac{1}{k} < \frac{1}{p} < \varepsilon$. We thus have

$$x - \varepsilon < x - \frac{1}{k} < q_k < x + \frac{1}{k} < x + \varepsilon.$$

It follows that for every $k > p$ we have $|q_k - x| < \varepsilon$. Since ε is arbitrary, this proves that $(q_k)_{k\in\mathbb{N}^*}$ converges to x.

Example *1.33*

Consider the subset $\boldsymbol{A} = [0, 1] \times (0, 1)$ of $\mathbb{R}^2$. We have

- $\mathring{\boldsymbol{A}} = (-1, 1) \times (-1, 1)$;
- $\partial \boldsymbol{A} = ([-1, 1] \times \{-1, 1\}) \cup (\{-1, 1\} \times [-1, 1])$;
- $\overline{\boldsymbol{A}} = [-1, 1] \times [-1, 1]$.

Example *1.34*

Consider the subset $\boldsymbol{A} = \mathcal{B}_1(0) \setminus \{0\}$ of $\mathbb{R}^2$. Then:

- $\mathring{\boldsymbol{A}} = \boldsymbol{A}$;
- $\partial \boldsymbol{A} = \{0\} \cup \{(x, y) \in \mathbb{R}^2 \,|\, \|(x, y)\| = 1\}$;
- $\overline{\boldsymbol{A}} = \{(x, y) \in \mathbb{R}^2 \,|\, \|(x, y)\| \leq 1\}$.

Example *1.35*

Consider the subset

$$\boldsymbol{A} = \{(\rho \cos\theta, \rho \sin\theta) \in \mathbb{R}^2 \,|\, 0 \leq \rho \leq 1,\ 0 < \theta < 2\pi\}$$

of $\mathbb{R}^2$. Then:

- $\mathring{\boldsymbol{A}} = \{(\rho \cos\theta, \rho \sin\theta) \in \mathbb{R}^2 \,|\, 0 < \rho < 1,\ 0 < \theta < 2\pi\}$;
- $\partial \boldsymbol{A} = \{(\rho \cos\theta, \rho \sin\theta) \in \mathbb{R}^2 \,|\, \rho = 1,\ 0 < \theta < 2\pi\} \cup ([0, 1] \times \{0\})$;
- $\overline{\boldsymbol{A}} = \{(x, y) \in \mathbb{R}^2 \,|\, \|(x, y)\| \leq 1\}$.

Example *1.36*

Consider the subset $A = [0, 1) \cup \{2\}$ of $\mathbb{R}$. The point 2 is an isolated point of A. The point 1 does not belong to A, but it is an accumulation point for A, and the points of $[0, 1)$ are points of A and are accumulation points of A.

Example *1.37*

A (bounded or unbounded) interval of $\mathbb{R}$ is open if and only if it is an open subset of $\mathbb{R}$, and that it is closed if and only if it is a closed subset of $\mathbb{R}$.

Example *1.38*

The set $\mathbb{N}$, viewed as a subset of $\mathbb{R}$ is made up of just isolated points. All points of $\mathbb{Q}$ are accumulation points for $\mathbb{Q}$, but none of them is internal to $\mathbb{Q}$.

Example *1.39*

Let $a, b \in \mathbb{R}$ be such that $a < b$. Then

- $\overline{(a,b)} = [a,b]$;
- $\overline{(a,b]} = [a,b]$ and $\overline{[a,b)} = [a,b]$;
- $\text{int}([a,b]) = (a,b)$;
- $\text{int}([a,b)) = (a,b)$ and $\text{int}((a,b]) = (a,b)$;
- $\partial[a,b] = \partial(a,b) = \partial(a,b] = \partial[a,b) = \{a,b\}$.

Exercises

Exercise 1.28

Show that, if $\boldsymbol{X} = \mathbb{R}$ (so that the underlying field is also $\mathbb{R}$) and the norm $|\cdot|$ is the absolute value (which is the Euclidean norm of $\mathbb{R}$), the open ball $\mathcal{B}_r(x_0)$ is the open interval $(x_0 - r, x_0 + r)$. Show also that, if $\boldsymbol{X} = \mathbb{R}^2$ and the norm $\|\cdot\|$ is the Euclidean norm, the open ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is a disk in $\mathbb{R}^2$ comprising the points whose distance from the center $\boldsymbol{x}_0 \in \mathbb{R}^2$ is $< r$. Finally, show that if $\boldsymbol{X} = \mathbb{R}^3$ and the norm $\|\cdot\|$ is, again, the Euclidean norm of $\mathbb{R}^3$, the set $\mathcal{B}_r(\boldsymbol{x}_0)$ is the set of points within a sphere of radius r centered at $\boldsymbol{x}_0$. [Solution: (1.28)]

Exercise 1.29

Consider a metric space (E, d). It is possible to define a ball $\mathcal{B}_r(x_0)$ of radius $r > 0$ and center $x_0 \in E$ as

$$\mathcal{B}_r(x_0) \stackrel{\text{def}}{=} \{x \in E \mid \mathrm{d}(x, x_0) < r\}.$$

Geometrically, this ball still represents the set of points whose distance from x is less than r. Show that if d is the discrete metric, every subset A of E is open and closed (Hint: take $0 < r < 1$). [Solution: (1.29)]

Exercise 1.30

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$. Let $\boldsymbol{U}$ be a neighbourhood of $\boldsymbol{x}_0$. Let $\boldsymbol{U} \subseteq \boldsymbol{W} \subseteq \boldsymbol{X}$. Prove that $\boldsymbol{W}$ is also a neighbourhood of $\boldsymbol{x}_0$. [Solution: (1.30)]

Exercise 1.31

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$. Consider a set

$$\{\boldsymbol{U}_i \subseteq \boldsymbol{X} \mid i \in \mathcal{I}\},$$

where $\mathcal{I} \subseteq \mathbb{N}$ is a set of indices, and let $\boldsymbol{U}_i$ be a neighborhood of $\boldsymbol{x}_0$ for every $i \in \mathcal{I}$. Show that:

- the union $\bigcup_{i \in \mathcal{I}} \boldsymbol{U}_i$ is a neighborhood of $\boldsymbol{x}_0$
- if $\mathcal{I}$ is finite, the intersection $\bigcap_{i \in \mathcal{I}} \boldsymbol{U}_i$ is a neighborhood of $\boldsymbol{x}_0$.

Find an example of an infinite set of neighborhoods of a point whose intersection is not a neighborhood of that point. [Solution: (1.31)]

Exercise 1.32

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that $\boldsymbol{A}$ is closed if and only if $\boldsymbol{A} = \overline{\boldsymbol{A}}$. [Solution: (1.32)]

Exercise 1.33

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that $\boldsymbol{A}$ is closed if and only if $\boldsymbol{X} \setminus \boldsymbol{A}$ is open. [Solution: (1.33)]

Exercise 1.34 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Prove that

(i) $\boldsymbol{x}_0 \in \boldsymbol{X}$ is an isolated point of $\boldsymbol{A}$ if and only if $\boldsymbol{x}_0 \in \boldsymbol{A} \setminus \mathcal{D}(\boldsymbol{A})$;

(ii) $\boldsymbol{x}_0 \in \boldsymbol{A}$ is **not** an isolated point of $\boldsymbol{A}$ if and only if $\boldsymbol{x}_0 \in \boldsymbol{A} \cap \mathcal{D}(\boldsymbol{A})$;

(iii) $\boldsymbol{x}_0 \in \boldsymbol{A}$ is either an isolated point of $\boldsymbol{A}$ or a limit point of $\boldsymbol{A}$ (but not both).

[Solution: (1.34)]

Exercise 1.35 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Prove that $\overline{\boldsymbol{X} \setminus \boldsymbol{A}} = \boldsymbol{X} \setminus \mathring{\boldsymbol{A}}$. [Solution: (1.35)]

Exercise 1.36 Consider the set $\boldsymbol{A} = \{(2, 3)\}$, i.e., the set of $\mathbb{R}^2$ which consists of a single point with first and second coordinates equal to 2 and 3, respectively.

(1) Determine if $\boldsymbol{A}$ is open. If not, find the interior of $\boldsymbol{A}$;

(2) Find the boundary of $\boldsymbol{A}$;

(3) Determine if $\boldsymbol{A}$ is closed. If not, find the closure of $\boldsymbol{A}$;

(4) Find the derived of $\boldsymbol{A}$, i.e., the set of its limit points.

[Solution: (1.36)]

Exercise 1.37 Let $\boldsymbol{A} = \mathbb{R}^n \setminus \{\mathbf{0}\}$. Show that the origin $\mathbf{0}$ is a limit point, i.e., $\mathbf{0} \in \mathcal{D}(\boldsymbol{A})$. [Solution: (1.37)]

Exercise 1.38 Consider the normed space $(\mathbb{R}, |\cdot|)$. Given $a \in \mathbb{R}$, show that

$$\mathcal{D}([a, +\infty)) = \mathcal{D}((a, +\infty)) = [a, +\infty).$$

Adapt this solution to find $\mathcal{D}(\boldsymbol{U})$, where

$$\boldsymbol{U} = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 > a^2\},$$

by considering the normed space $(\mathbb{R}^2, \|\cdot\|)$, where $\|\cdot\|$ denotes the Euclidean norm. [Solution: (1.38)]

Exercise 1.39 Consider the set $\boldsymbol{A} = (0, 1) \times \{0\}$ in the normed space $(\mathbb{R}^2, \|\cdot\|)$, where $\|\cdot\|$ denotes the Euclidean norm.

(1) Determine if $\boldsymbol{A}$ is open. If not, find the interior of $\boldsymbol{A}$;

(2) Find the boundary of $\boldsymbol{A}$;

(3) Determine if $\boldsymbol{A}$ is closed. If not, find the closure of $\boldsymbol{A}$.

[Solution: (1.39)]

Exercise 1.40 Consider the normed space $(\mathbb{R}^2, \|\cdot\|)$, where $\|\cdot\|$ is the Euclidean norm. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a continuous function. Show that

$$\boldsymbol{A} = \{(x, y) \in \mathbb{R}^2 \,|\, y < f(x)\}$$

is open. Show that the same is true if we consider any p-norm in $\mathbb{R}^2$. Finally, show that if f is not continuous this result is not true in general. [Solution: (1.40)]

Exercise 1.41 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that

(i) $\boldsymbol{A} \cup \partial\boldsymbol{A} = \mathring{\boldsymbol{A}} \cup \partial\boldsymbol{A}$;

(ii) $\mathring{\boldsymbol{A}} \cap \partial \boldsymbol{A} = \varnothing$.

[Solution: (1.41)]

Exercise 1.42 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that the set of adherent points of $\boldsymbol{A}$ coincides with $\overline{\boldsymbol{A}}$. [Solution: (1.42)]

Exercise 1.43 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that:

(i) $\operatorname{ext}(\boldsymbol{A}) = \boldsymbol{X} \setminus \overline{\boldsymbol{A}}$;

(ii) $\operatorname{ext}(\boldsymbol{A})$ is open;

(iii) $\operatorname{ext}(\boldsymbol{A}) \cap \boldsymbol{A} = \varnothing$ (i.e., $\operatorname{ext}(\boldsymbol{A})$ and $\boldsymbol{A}$ are disjoint);

(iv) if $\boldsymbol{A} \subseteq \boldsymbol{B}$, then $\operatorname{ext}(\boldsymbol{B}) \subseteq \operatorname{ext}(\boldsymbol{A})$;

(v) $\operatorname{ext}(\boldsymbol{A})$ is the union of all open subsets of $\boldsymbol{X}$ that are disjoint from $\boldsymbol{A}$;

(vi) $\operatorname{ext}(\boldsymbol{A})$ is the largest open subset of $\boldsymbol{X}$ that is disjoint from $\boldsymbol{A}$.

[Solution: (1.43)]

Exercise 1.44 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that $\boldsymbol{A}$ is closed if and only if it contains all its limit points. [Solution: (1.44)]

Exercise 1.45 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Show that

$$\mathring{\boldsymbol{A}} = \max\{\boldsymbol{G} \subseteq \boldsymbol{A} \,|\, \boldsymbol{G} \text{ is open}\}.$$

[Solution: (1.45)]

Exercise 1.46 Let $\boldsymbol{A}$ be an open set in $\mathbb{R}^n$. Let $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2}, \ldots, x_{0,n}) \in \boldsymbol{A}$. Show that there exists a sufficiently large value of $k \in \mathbb{N}^*$ such that the hyper-cube

$$\boldsymbol{Q}_k = \prod_{j=1}^{n} \left[x_{0,j} - \frac{1}{k}, x_{0,j} + \frac{1}{k} \right]$$

is entirely contained in $\boldsymbol{A}$. [Solution: (1.46)]

Exercise 1.47 Consider a normed space $(\boldsymbol{X}, \|\cdot\|)$. Suppose that the unit ball centred at the origin $\mathcal{B}_1(\boldsymbol{0}, \boldsymbol{X}, \|\cdot\|)$ contains the segment

$$[\boldsymbol{x}_0, \boldsymbol{y}_0] = \{\lambda \boldsymbol{x}_0 + (1-\lambda)\boldsymbol{y}_0 \,|\, \lambda \in [0,1]\},$$

where $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ are distinct elements of $\boldsymbol{X}$. Show that:

(i) $\|\boldsymbol{x}_0 + \boldsymbol{y}_0\| = \|\boldsymbol{x}_0\| + \|\boldsymbol{y}_0\|$

(ii) $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ are linearly independent.

[Solution: (1.47)]

Exercise 1.48 Given a normed space $(\boldsymbol{X}, \|\cdot\|)$, show that if $\boldsymbol{0} \in \mathcal{B}_r(\boldsymbol{x}_0)$ the only subspace of $\boldsymbol{X}$ contained in $\mathcal{B}_r(\boldsymbol{x}_0)$ is the origin, and if $\boldsymbol{0} \notin \mathcal{B}_r(\boldsymbol{x}_0)$ there are no subspaces of $\boldsymbol{X}$ contained in $\mathcal{B}_r(\boldsymbol{x}_0)$. [Solution: (1.48)]

Exercise 1.49 Given a normed space $(\boldsymbol{X}, \|\cdot\|)$, show that the only subspace of $\boldsymbol{X}$ containing a ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is the space $\boldsymbol{X}$ itself. [Solution: (1.49)]

Exercise 1.50

Show that the set

$$A = \left\{ \frac{1}{n} \,\middle|\, n \in \mathbb{N}^* \right\}$$

has 0 as a limit point, and that 0 is the only limit point of A. [Solution: (1.50)]

1.4 Convergence and limits

We give the definition of limit, which is similar to the one dimensional case.

Definition 1.16 [Convergence at a Point]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed vector spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0$ be a limit point for $\boldsymbol{A}$. Let $\boldsymbol{\ell} \in \boldsymbol{Y}$. We say that f tends to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, or that f converges to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, if

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$$

The meaning of Definition 1.16 is the following: as $\boldsymbol{x}$ approaches $\boldsymbol{x}_0$, the value of the function $f(\boldsymbol{x})$ approaches $\boldsymbol{\ell}$. More precisely, $f(\boldsymbol{x})$ can be made as close to $\boldsymbol{\ell}$ as we desire, by making $\boldsymbol{x}$ sufficiently close to $\boldsymbol{x}_0$. This last statement can be expressed in mathematical terms by saying that, chosen an arbitrary $\varepsilon > 0$, it is always possible to determine a value $\delta > 0$ (which in general depends on ε) such that, if $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, then $f(\boldsymbol{x}) \in \mathcal{B}_\varepsilon(\boldsymbol{\ell})$. We stress the fact that the condition of Definition 1.16 requires that $\boldsymbol{x} \neq \boldsymbol{x}_0$. If this is not the case, then $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X = 0 < \delta$ is always satisfied for $\boldsymbol{x} = \boldsymbol{x}_0$, and therefore using $\boldsymbol{x} = \boldsymbol{x}_0$ in $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$ we obtain $\|f(\boldsymbol{x}_0) - \boldsymbol{\ell}\|_Y < \varepsilon$. Saying that for every $\varepsilon > 0$ the inequality $\|f(\boldsymbol{x}_0) - \boldsymbol{\ell}\|_Y < \varepsilon$ holds is a pompous way of saying that $f(\boldsymbol{x}_0) = \boldsymbol{\ell}$, which is too restrictive.

Definition 1.16 also requires $\boldsymbol{x}_0$ to be a limit point. It seems silly to consider points that are not in the closure of the domain: for example, evaluating the behavior of the function $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ defined by $f(x) = \sqrt{x}$, as x approaches -1 seems unreasonable. But there is a more serious issue: in this case, every value ℓ is a limit. In fact, choosing $\delta < \frac{1}{2}$, the set $A \setminus \{-1\} = \mathbb{R}_+^*$ does not contain points x such that $|x - x_0| < \delta$, and therefore we will not be able to find any x for which the condition in Definition 1.16 is false, which means that any value of ℓ is a limit.

From these considerations, it seems logical to require $\boldsymbol{x}_0$ to belong to the closure of $\boldsymbol{A}$. However, a point of the closure may include isolated points of A. While some authors accept this situation, we prefer to opt for the more standard approach of requiring $\boldsymbol{x}_0$ to be a limit point of $\boldsymbol{A}$.

The notion of convergence can be expressed in a number of alternative ways. The following result shows that we can replace some of the strict inequalities in Definition 1.16 with loose inequalities, and that we can rephrase the inequalities involving the norms of $\boldsymbol{x} - \boldsymbol{x}_0$ and $f(\boldsymbol{x}) - \boldsymbol{\ell}$ in terms of balls. In the same way, we can express the definition of convergence in terms of neighborhoods.

Theorem 1.13 *Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0$ be a limit point for A. Let $\boldsymbol{\ell} \in \boldsymbol{Y}$. The following statements are equivalent:*

(1) *f converges to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$;*

(2) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad 0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$;

(3) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$;

(4) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$;

(5) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$;

(6) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) \ \Rightarrow\ f(\boldsymbol{x}) \in \mathcal{B}_\varepsilon(\boldsymbol{\ell}, \boldsymbol{Y}, \|\cdot\|_Y)$;

(7) $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in (\mathcal{B}_\delta(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad f(\boldsymbol{x}) \in \mathcal{B}_\varepsilon(\boldsymbol{\ell}, \boldsymbol{Y}, \|\cdot\|_Y);$

(8) $\forall\, \boldsymbol{W} \in \mathcal{N}(\boldsymbol{\ell}, \boldsymbol{Y}, \|\cdot\|_Y),\ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}, \qquad \boldsymbol{x} \neq \boldsymbol{x}_0 \ \Rightarrow\ f(\boldsymbol{x}) \in \boldsymbol{W}.$

The proof is left as an exercise. It is worth noting that the balls in **(6)** do not necessarily live in the same space (for example, consider the case where $\boldsymbol{X} = \mathbb{R}^2$ and $\boldsymbol{Y} = \mathbb{R}^3$), and, even if they do, they may be associated with different norms.

The reader already knows that if f converges to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, we say that $\boldsymbol{\ell}$ is the limit of f as $\boldsymbol{x} \to \boldsymbol{x}_0$. However, we are not authorized to talk about "*the* limit" until a result of uniqueness is established. This point is addressed in the following theorem.

Theorem 1.14 [Uniqueness of the Limit]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0$ be a limit point for $\boldsymbol{A}$. Let $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2 \in \boldsymbol{Y}$ be such that f converges to $\boldsymbol{\ell}_1$ as $\boldsymbol{x} \to \boldsymbol{x}_0$ and f converges to $\boldsymbol{\ell}_2$ as $\boldsymbol{x} \to \boldsymbol{x}_0$. Then, $\boldsymbol{\ell}_1 = \boldsymbol{\ell}_2$.

Proof Suppose by contradiction that $\boldsymbol{\ell}_1 \neq \boldsymbol{\ell}_2$, so that

$$\frac{\|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2\|_Y}{3} > 0.$$

Let us take $\varepsilon = \frac{\|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2\|_Y}{3}$. Therefore, we can use Definition 1.16 with this value of ε, and we find that there exists $\delta_1 > 0$ such that $\forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}$

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta_1 \qquad \Rightarrow \qquad \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y < \varepsilon.$$

Likewise, there exists $\delta_2 > 0$ such that $\forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}$

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta_2 \qquad \Rightarrow \qquad \|f(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \varepsilon.$$

Taking $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \min\{\delta_1, \delta_2\}$, the inequalities $\|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y < \varepsilon$ and $\|f(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \varepsilon$ hold simultaneously. Thus, using the triangle inequality

$$3\,\varepsilon = \|\boldsymbol{\ell}_1 - \boldsymbol{\ell}_2\|_Y = \|\boldsymbol{\ell}_1 - f(\boldsymbol{x}) + f(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y \leq \|\boldsymbol{\ell}_1 - f(\boldsymbol{x})\|_Y + \|f(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y = \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y + \|f(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < 2\,\varepsilon.$$

From $3\,\varepsilon < 2\,\varepsilon$, recalling that $\varepsilon > 0$, we obtain $3 < 2$, which is false.

Definition 1.17 [Limit]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $\boldsymbol{x}_0$ be a limit point for A. Let $\boldsymbol{\ell} \in \boldsymbol{Y}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ converge to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$. Then, we say that $\boldsymbol{\ell}$ is the limit *of f as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$ (or, in symbols, as $\boldsymbol{x} \to \boldsymbol{x}_0$), and we write*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \boldsymbol{\ell} \qquad \textit{or} \qquad f(\boldsymbol{x}) \overset{\boldsymbol{x} \to \boldsymbol{x}_0}{\longrightarrow} \boldsymbol{\ell}.$$

We now wish to introduce the notion of divergence at a point. To this end, we recall the definition of $+\infty$ and $-\infty$. These objects are distinct from each other and distinct from every real number:

(1) $+\infty \neq -\infty$

(2) $\forall\, x \in \mathbb{R}, \quad x \neq +\infty$ and $x \neq -\infty$.

If $x \in \mathbb{R}$, then x is called *finite*. If $x = +\infty$ or if $x = -\infty$ (in this case we will also write, in short, if $x = \pm\infty$), it is called *infinite*. We define the *extended real line* as the real line $\mathbb{R}$ with these two additional objects:

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}.$$

We can extend the order relation $\leq$ in $\mathbb{R}$ to $\overline{\mathbb{R}}$ in the following way: given $x, y \in \overline{\mathbb{R}}$, we write $x \leq y$ (or $y \geq x$) if one of the following is true:

(1) $x, y \in \mathbb{R}$, and $x \leq y$ as real numbers;

(2) $y = +\infty$;

(3) $x = -\infty$.

We say that $x < y$ if $x \leq y$ and $x \neq y$. With these definitions, we have the following two fundamental consequences:

(a) $\forall\, x \in \mathbb{R}, \quad x < +\infty$ and $x > -\infty$;

(b) $\forall\, x \in \overline{\mathbb{R}}, \quad -\infty \leq x \leq +\infty$.

The set $\overline{\mathbb{R}}$ with the relation $\leq$ is an ordered set, i.e., the following properties are satisfied:

(1) $\forall\, x \in \overline{\mathbb{R}}, \quad x \leq x$ (Reflexive property)

(2) $\forall\, x, y \in \overline{\mathbb{R}}, \quad x \leq y$ and $y \leq x \quad \Rightarrow \quad x = y$ (Antisymmetric property)

(3) $\forall\, x, y, z \in \overline{\mathbb{R}}, \quad x \leq y$ and $y \leq z \quad \Rightarrow \quad x \leq z$ (Transitive property)

The ordering given by $\leq$ is *total*: this means that we have the additional property

(4) $\forall\, x, y \in \overline{\mathbb{R}}, \quad x \leq y$ or $y \leq x$ (Trichotomy)

The term "trichotomy" means, literally, division in three parts. The meaning comes from the fact that the same property can be stated by saying that for every $x, y \in \overline{\mathbb{R}}$, we have one and only one of the three possibilities: $x < y$, $x = y$ or $x > y$.

In a generic normed space $(\boldsymbol{X}, \|\cdot\|)$, it makes no sense to talk about "plus" or "minus" infinity, as we normally do when dealing with subsets of the real line. We can only define an infinity without sign, that we denote by ∞, which is distinct from any element of $\boldsymbol{X}$, and it allows us to extend the norm $\|\cdot\|$ to a norm in the set $\boldsymbol{X} \cup \{\infty\}$: the norm is the usual one if $\boldsymbol{x} \in \boldsymbol{X} \cup \{\infty\}$ is finite (i.e., if $\boldsymbol{x} \in \boldsymbol{X}$), and it is $+\infty$ if $\boldsymbol{x} = \infty$:

$$\begin{aligned} \|\cdot\| : \boldsymbol{X} \cup \{\infty\} &\longrightarrow \overline{\mathbb{R}} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\| & \text{if } \boldsymbol{x} \in \boldsymbol{X} \\ +\infty & \text{if } \boldsymbol{x} = \infty. \end{cases} \end{aligned}$$

Since we do not have a total order in $\mathbb{R}^n$ for $n > 1$, we cannot expect to have a total order in $\boldsymbol{X} \cup \{\infty\}$.

It is worth noting, at this point, that $(\mathbb{R}, |\cdot|)$ is a particular type of normed space: therefore, for this normed space we can define the element ∞. In other words, we can either equip $\mathbb{R}$ with the two "infinities" $+\infty$ and $-\infty$, or decide to equip it with a single, "signless" infinity ∞.

We keep these two possibilities separate: we do not want to equip the space $\mathbb{R}$ with three infinities $+\infty, -\infty, \infty$, because in such case the property of uniqueness of the limit would fail. While in previous Calculus units the focus was on $+\infty$ and $-\infty$, since here we are addressing the case of normed space, the emphasis will be on ∞.

The divergence is defined with respect to this notion of infinity.

Definition 1.18 [Divergence at a Point]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $\boldsymbol{x}_0$ be a limit point for $\boldsymbol{A}$. We say that $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ is divergent *as $\boldsymbol{x} \to \boldsymbol{x}_0$ if*

$$\forall\, M > 0,\ \exists \delta > 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \quad \Rightarrow \quad \|f(\boldsymbol{x})\|_Y > M$$

We say that f oscillates *as $\boldsymbol{x} \to \boldsymbol{x}_0$ if it is neither convergent nor divergent.*

According to Definition 1.18, a function $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ is divergent as $\boldsymbol{x}$ approaches $\boldsymbol{x}_0$ when the norm $\|f(\boldsymbol{x})\|_Y$ becomes bigger than any assigned M, or, more precisely, if given an arbitrary M, we have $\|f(\boldsymbol{x})\|_Y > M$ if we

choose $\boldsymbol{x}$ to be sufficiently close to $\boldsymbol{x}_0$. In the example of the function $f(x) = 1/x$, if $M = 10^6$, then we have $|f(x)| > 10^6$ if and only if $x \neq 0$ and $\frac{1}{x} > 10^6$, i.e., if and only if $x \neq 0$ and $x < 10^{-6}$. Taking $\delta < 10^{-6}$ therefore guarantees that $|f(x)| > 10^6$. If we choose a bigger M, say $M = 10^{100}$, we will have to take x to be much closer to 0 to ensure $|f(x)| > M$: in this case $\delta < 10^{-100}$.

Definition 1.19 [Convergence and Divergence as $\boldsymbol{x} \to \infty$]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed vector spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{\ell} \in \boldsymbol{Y}$. We say that f converges to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to ∞ if

$$\forall\, \varepsilon > 0,\ \exists N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

Moreover, we say that f is divergent as $\boldsymbol{x}$ tends to ∞ if

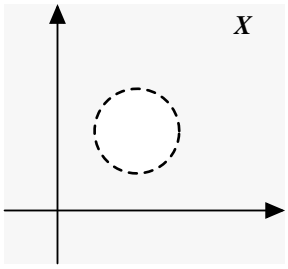
$$\forall\, M > 0,\ \exists N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ \|f(\boldsymbol{x})\|_Y > M$$

Finally, we say that f oscillates *as $\boldsymbol{x}$ tends to ∞ if it is neither convergent nor divergent.*

To be able to talk about "the limit" in all these other cases, we must establish a result of uniqueness for each of these conditions. However, a more elegant option is to establish a unified notion of limit that encompasses Definition 1.16, 1.18 and 1.19, and then prove a result of uniqueness for this general definition. This is possible if we choose **(7)** in Theorem 1.13 as the definition of limit, and we generalize the notion of neighborhood to accommodate the case where the limit is ∞ and the case where $\boldsymbol{x}_0$ is also ∞.[2] We start by introducing the notion of neighborhood of ∞.

Definition 1.20 [Neighborhood of ∞]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ be a normed space. Let $\boldsymbol{U} \subseteq \boldsymbol{X}$. We say that $\boldsymbol{U}$ is a neighborhood of ∞ *if its complement is bounded.*

For example, the complement of a disc or of a square are neighborhoods of ∞.



Now, the notation $\mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X)$ can be extended to the case $\boldsymbol{x}_0 \in \boldsymbol{X} \cup \{\infty\}$. Indeed, if $\boldsymbol{x}_0 \in \boldsymbol{X}$, the notion of neighborhood of $\boldsymbol{x}_0$ is the standard one, and when $\boldsymbol{x}_0 = \infty$ the set $\mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X)$ denotes the neighborhoods of ∞.

In order to have a unified notion of convergence and limit, we also need to generalize the notion of limit point: this can be done by replacing the ball in the definition of limit point with a neighborhood (which can now be a neighborhood of ∞).

[2] We notice that these two infinities are different: the first is the ∞ of $\boldsymbol{Y}$, and the second is the ∞ of $\boldsymbol{X}$. However, there is nothing to gain in distinguishing them notationally.

Definition 1.21 [Generalized Limit Point]
Let $(X, \|\cdot\|_X)$ be a normed space and let $A \subseteq X$. Let $x_0 \in X \cup \{\infty\}$. We say that x_0 is a generalized limit point *for A if, for any neighborhood U of x_0, the set $A \cap U$ contains some elements which are distinct from x_0, i.e.,*

$$(A \cap U) \setminus \{x_0\} \neq \emptyset.$$

We denote by $\tilde{\mathcal{D}}(A, X, \|\cdot\|_X)$, or simply $\tilde{\mathcal{D}}(A)$ if there is no ambiguity, the set of generalized limit points of A.

Clearly, if $x_0 \in X$, then $x_0 \in \tilde{\mathcal{D}}(A)$ if and only if $x_0 \in \mathcal{D}(A)$. We have $\infty \in \tilde{\mathcal{D}}(A)$ if and only if, for all $M > 0$, the set $\{x \in X \mid \|x\|_X > M\} \cap A$ contains elements which are distinct from ∞. Since, for all $M > 0$, we have $\infty \notin \{x \in X \mid \|x\|_X > M\} \cap A$, this occurs if and only if, for all $M > 0$, we have $\{x \in X \mid \|x\|_X > M\} \cap A \neq \emptyset$. In other words, $\infty \in \tilde{\mathcal{D}}(A)$ if and only if, for all $M > 0$, there exists $x \in A$ such that $\|x\|_X > M$. This is the same as requiring that A is not bounded.

We are now ready to give a definition of convergence and limit that encompasses all the previous ones.

Definition 1.22 *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, and let $A \subseteq X$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let $\ell \in Y \cup \{\infty\}$. We say that f converges to ℓ as x tends to x_0 (or, in symbols, as $x \to x_0$) if, for every neighborhood W of ℓ, there exists a neighborhood V of x_0 such that, for all $x \in V \cap A$ different from x_0, we have $f(x) \in W$. In symbols:*

$$\boxed{\forall\, W \in \mathcal{N}(\ell, Y, \|\cdot\|_Y),\ \exists\, V \in \mathcal{N}(x_0, X, \|\cdot\|_X) :\quad \forall\, x \in A \cap V, \qquad x \neq x_0 \ \Rightarrow\ f(x) \in W} \tag{1.19}$$

If $\ell \in Y$, we say that the limit of f as $x \to x_0$ exists in Y. If (1.19) is satisfied with $\ell = \infty$, we say that the limit exists in $Y \cup \{\infty\}$, or that f is divergent as $x \to x_0$, or that f is convergent as $x \to x_0$ with respect to $Y \cup \{\infty\}$. If (1.19) is not satisfied for any $\ell \in Y \cup \{\infty\}$, we say that the limit does not exist with respect to $Y \cup \{\infty\}$.

Again, it is straightforward to see that (1.19) can be written as

$$\forall\, W \in \mathcal{N}(\ell, Y, \|\cdot\|_Y),\ \exists\, V \in \mathcal{N}(x_0, X, \|\cdot\|_X) :\quad \forall\, x \in (A \cap V) \setminus \{x_0\}, \qquad f(x) \in W$$

We only proved the uniqueness of the limit for the case where x_0 and ℓ are finite (see Theorem 1.14). We show that the Definition 1.22 ensures the same property in the general case.

Theorem 1.15 [Uniqueness of the Limit – The General Case]
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $A \subseteq X$. Let $f : A \longrightarrow Y$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let $\ell_1, \ell_2 \in Y \cup \{\infty\}$ be such that f converges to ℓ_1 as $x \to x_0$ and f converges to ℓ_2 as $x \to x_0$. Then, $\ell_1 = \ell_2$.

Proof Assume by contradiction that $\ell_1 \neq \ell_2$. Then, there exists a neighborhood W_1 of ℓ_1 and a neighborhood W_2 of ℓ_2 such that $W_1 \cap W_2 = \emptyset$. Indeed, if ℓ_1 and ℓ_2 are finite, we can take $r < \frac{\|\ell_1 - \ell_2\|_Y}{2}$, and $W_1 = \mathcal{B}_r(\ell_1, Y, \|\cdot\|_Y)$ and $W_2 = \mathcal{B}_r(\ell_2, Y, \|\cdot\|_Y)$ are disjoint. In the case $\ell_1 < \infty$ and $\ell_2 = \infty$, if $\ell_1 \neq 0$ we can take $r \in (0, \|\ell_1\|_Y)$ and $M > \|\ell_1\|_Y + r$, and $W_1 = \mathcal{B}_r(\ell_1, Y, \|\cdot\|_Y)$ and $W_2 = \{x \in X \mid \|x\|_X > M\}$ are disjoint neighborhoods of ℓ_1 and ℓ_2, respectively.

Since f converges to ℓ_1 as $x \to x_0$, there exists a neighborhood V_1 of x_0 such that for all $x \in V_1 \cap A$ we have $f(x) \in W_1$. Likewise, since f converges to ℓ_2 as $x \to x_0$, there exists a neighborhood V_2 of x_0 such that for all $x \in V_2 \cap A$ there holds $f(x) \in W_2$. The set $V_1 \cap V_2$ is a neighborhood of x_0, and since $x_0 \in \tilde{\mathcal{D}}(A)$ we have $\big((V_1 \cap V_2) \cap \tilde{\mathcal{D}}(A)\big) \setminus \{x_0\} \neq \emptyset$. Let $x \in (V_1 \cap V_2) \cap A$. We have simultaneously $f(x) \in W_1$ and $f(x) \in W_2$, which is a contradiction.

Theorem 1.15 ensures that the following definition is well posed.

Definition 1.23 [Limit]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let $\boldsymbol{\ell} \in \boldsymbol{Y} \cup \{\infty\}$. If $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ converges to $\boldsymbol{\ell}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, we say that $\boldsymbol{\ell}$ is the limit *of f as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$ (or, in symbols, as $\boldsymbol{x} \to \boldsymbol{x}_0$), and we write*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \boldsymbol{\ell} \qquad or \qquad f(\boldsymbol{x}) \overset{\boldsymbol{x} \to \boldsymbol{x}_0}{\longrightarrow} \boldsymbol{\ell}. \tag{1.20}$$

Even if the notation (1.20) in Definition 1.23 is unifying, very often, instead of the expressions

$$\lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \boldsymbol{\ell}, \qquad \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \infty, \qquad \lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \infty,$$

one finds, respectively,

$$\lim_{\|\boldsymbol{x}\|_X \to \infty} f(\boldsymbol{x}) = \boldsymbol{\ell}, \qquad \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \|f(\boldsymbol{x})\|_Y = \infty, \qquad \lim_{\|\boldsymbol{x}\|_X \to \infty} \|f(\boldsymbol{x})\|_Y = \infty.$$

To summarize:

(L1) $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \boldsymbol{\ell}$: $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \ \Rightarrow \ \ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$

(L2) $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \infty$: $\forall\, M > 0,\ \exists\, \delta > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \ \Rightarrow \ \ \|f(\boldsymbol{x})\|_Y > M$

(L3) $\lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \boldsymbol{\ell}$: $\forall\, \varepsilon > 0,\ \exists\, N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon;$

(L4) $\lim_{\boldsymbol{x} \to \infty} f(\boldsymbol{x}) = \infty$: $\forall\, M > 0,\ \exists\, N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ \|f(\boldsymbol{x})\|_Y > M.$

The power of Definition 1.22, and in general of the use of neighborhoods to define convergence and limits, is the generality and the flexibility. For example, if we want to adapt this definition to cases where $+\infty$ and $-\infty$ are defined (i.e., when $\boldsymbol{X} = \mathbb{R}$ or $\boldsymbol{Y} = \mathbb{R}$), we only need to specify what we mean by neighborhood of $+\infty$ and of $-\infty$. Hence, if we consider the normed space $(\mathbb{R}, |\cdot|)$, we say that $U \subseteq \mathbb{R}$ is a neighborhood of $+\infty$ if there exists $a \in \mathbb{R}$ such that $(a, +\infty) \subseteq U$ (i.e., it is a subset of $\mathbb{R}$ which contains all sufficiently large real numbers). Likewise, $U \subseteq \mathbb{R}$ is a neighborhood of $-\infty$ if there exists $a \in \mathbb{R}$ such that $(-\infty, a) \subseteq U$. We defined the extended real line as

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\},$$

and we can define a generalized limit point for the normed space $(\mathbb{R}, |\cdot|)$ by allowing $\boldsymbol{x}_0$ to be an element of $\overline{\mathbb{R}}$. We find that $+\infty$ is a generalized limit point of $A \subseteq \mathbb{R}$ if A is not upper bounded, and that $-\infty$ is a generalized limit point of A if A is not lower bounded

With this definition, the condition (1.19) of Definition 1.22 is formally the same in the case where or $(\boldsymbol{X}, \|\cdot\|_X)$ or $(\boldsymbol{Y}, \|\cdot\|_Y)$ are equal to $(\mathbb{R}, |\cdot|)$. Resorting, instead, to the definition using inequalities, we have the following cases:[3]

[3] To be precise, we need to define also generalized limit points to include possibly $\pm\infty$. Indeed, Definition 1.21 only considered the "signless" infinity ∞. The definition is formally the same: consider the normed space $(\mathbb{R}, |\cdot|)$. Let $A \subseteq \mathbb{R}$ and let $x_0 \in \mathbb{R} \cup \{\pm\infty\}$. We say that x_0 is a generalized limit point for A if

$$\forall\, U \in \mathcal{N}(x_0, \mathbb{R}, |\cdot|), \ \ (A \cap U) \setminus \{x_0\} \neq \varnothing, \tag{1.21}$$

where we assume that if $x_0 = +\infty$ the set $\mathcal{N}(x_0, \mathbb{R}, |\cdot|)$ represents the neighborhoods of $+\infty$ and if $x_0 = -\infty$ the set $\mathcal{N}(x_0, \mathbb{R}, |\cdot|)$ is the neighborhoods of $-\infty$ With some abuse of notation, we continue to denote by $\tilde{\mathcal{D}}(A)$ the set of generalized limit points of A. Consider for example the case where $x_0 = +\infty$. A generic neighborhood of $x_0 = +\infty$ can be written as $(M, +\infty)$ for some $M \in \mathbb{R}$, so that (1.21) can be written as

$$\forall\, M \in \mathbb{R}, \ \ (A \cap (M, +\infty)) \setminus \{+\infty\} \neq \varnothing.$$

(i) If $(\boldsymbol{X}, \|\cdot\|_X)$ is $(\mathbb{R}, |\cdot|)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ is a generic normed space:

- if $x_0 = +\infty$ and $\boldsymbol{\ell} \in \boldsymbol{Y}$, the condition (1.19) becomes

$$\forall\, \varepsilon > 0,\ \exists N > 0 :\ \ \forall\, x \in A, \qquad x > N \ \ \Rightarrow \ \ \|f(x) - \boldsymbol{\ell}\|_Y < \varepsilon;$$

- if $x_0 = +\infty$ and $\boldsymbol{\ell} = \infty$, (1.19) becomes

$$\forall\, M > 0,\ \exists N > 0 :\ \ \forall\, x \in A, \qquad x > N \ \ \Rightarrow \ \ \|f(x)\|_Y > M;$$

- if $x_0 = -\infty$ and $\boldsymbol{\ell} \in \boldsymbol{Y}$, (1.19) becomes

$$\forall\, \varepsilon > 0,\ \exists N < 0 :\ \ \forall\, x \in A, \qquad x < N \ \ \Rightarrow \ \ \|f(x) - \boldsymbol{\ell}\|_Y < \varepsilon;$$

- if $x_0 = -\infty$ and $\boldsymbol{\ell} = \infty$, (1.19) becomes

$$\forall\, M > 0,\ \exists N < 0 :\ \ \forall\, x \in A, \qquad x < N \ \ \Rightarrow \ \ \|f(x)\|_Y > M;$$

(ii) If $(\boldsymbol{X}, \|\cdot\|_X)$ is a generic normed space and $(\boldsymbol{Y}, \|\cdot\|_Y)$ is $(\mathbb{R}, |\cdot|)$:

- if $\boldsymbol{x}_0 \in \boldsymbol{X}$ and $\ell = +\infty$, the condition (1.19) becomes

$$\forall\, M > 0,\ \exists \delta > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \ \Rightarrow \ \ f(\boldsymbol{x}) > M;$$

- if $\boldsymbol{x}_0 = \infty$ and $\ell = +\infty$, (1.19) becomes

$$\forall\, M > 0,\ \exists N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ f(\boldsymbol{x}) > M;$$

- if $\boldsymbol{x}_0 \in \boldsymbol{X}$ and $\ell = -\infty$, (1.19) becomes

$$\forall\, M < 0,\ \exists \delta > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \ \Rightarrow \ \ f(\boldsymbol{x}) < M;$$

- if $\boldsymbol{x}_0 = \infty$ and $\ell = -\infty$, (1.19) becomes

$$\forall\, M < 0,\ \exists N > 0 :\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x}\|_X > N \ \ \Rightarrow \ \ f(\boldsymbol{x}) < M;$$

(iii) If $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ are both equal to $(\mathbb{R}, |\cdot|)$:

- if $x_0 = +\infty$ and $\ell = +\infty$, the condition (1.19) becomes

$$\forall\, M > 0,\ \exists N > 0 :\ \ \forall\, x \in A, \qquad x > N \ \ \Rightarrow \ \ f(x) > M;$$

- if $x_0 = +\infty$ and $\ell = -\infty$, the condition (1.19) becomes

$$\forall\, M < 0,\ \exists N > 0 :\ \ \forall\, x \in A, \qquad x > N \ \ \Rightarrow \ \ f(x) < M;$$

- if $x_0 = -\infty$ and $\ell = +\infty$, the condition (1.19) becomes

$$\forall\, M > 0,\ \exists N < 0 :\ \ \forall\, x \in A, \qquad x < N \ \ \Rightarrow \ \ f(x) > M;$$

Since $+\infty$ is not an element of $A \cap (M, +\infty)$, the previous condition can be simplified as

$$\forall\, M \in \mathbb{R}, \ \ A \cap (M, +\infty) \neq \varnothing,$$

which is equivalent to requiring that A not be upper bounded. Similar considerations apply for the case where $x_0 = -\infty$.

- if $x_0 = -\infty$ and $\ell = -\infty$, the condition (1.19) becomes

$$\forall\, M < 0,\ \exists N < 0 :\quad \forall\, x \in A, \qquad\qquad x < N \quad \Rightarrow \quad f(x) < M.$$

You can see from this list how having a general definition of limit that uses the notion of neighborhood has the advantage of being succinct and flexible.

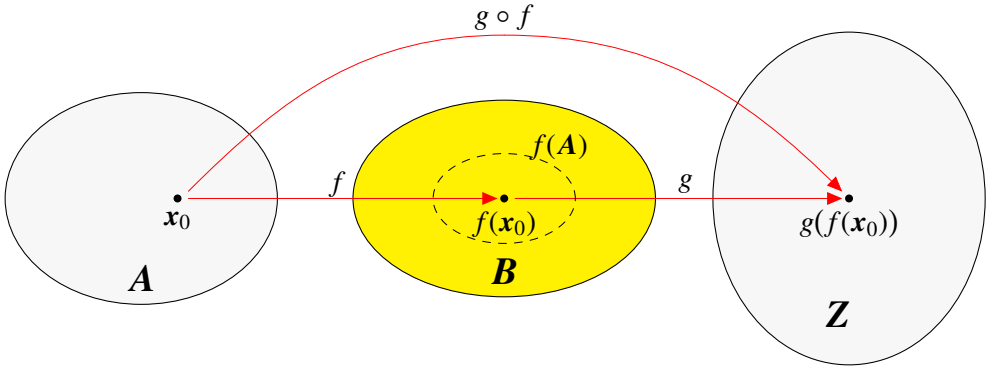
When $\boldsymbol{X}$ is $\mathbb{R}$, we can also define the limit for x that tends to x_0 from the right (and we write $\lim\limits_{x \to x_0^+} f(x)$) or from the left (and we write $\lim\limits_{x \to x_0^-} f(x)$). To this end, we can use the limit of restrictions, i.e.,

$$\lim_{x \to x_0^+} f(x) \stackrel{\text{def}}{=} \lim_{x \to x_0} f|_{A \cap (x_0, +\infty)}(x) \quad \text{and} \quad \lim_{x \to x_0^-} f(x) \stackrel{\text{def}}{=} \lim_{x \to x_0} f|_{(-\infty, x_0) \cap A}(x),$$

where, in the first case, $x_0 \in \mathcal{D}(A \cap (x_0, +\infty))$ and in the second $x_0 \in \mathcal{D}(A \cap (-\infty, x_0))$. Alternatively, we can continue to use Definition 1.22, but change the definition of neighborhood. For example, for the right limit we consider $\boldsymbol{U}$ to be a right-neighborhood of x_0 if there exists $r > 0$ such that $[x_0, x_0 + r) \subseteq \boldsymbol{U}$, and then re-define the notion of limit point using this notion of neighborhood. At that point, Definition 1.22 gives us a notion of right limit. The case of left limits is analogous.

1.4.1 Properties of limits

We present the theorem of the limit of the composition of functions. Consider three normed spaces $(\boldsymbol{X}, \|\cdot\|_X)$, $(\boldsymbol{Y}, \|\cdot\|_Y)$ and $(\boldsymbol{Z}, \|\cdot\|_Z)$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ and $g : \boldsymbol{B} \longrightarrow \boldsymbol{Z}$, where $\boldsymbol{A}$ and $\boldsymbol{B}$ are subsets of $\boldsymbol{X}$ and $\boldsymbol{Y}$ such that $f(\boldsymbol{A}) \subseteq \boldsymbol{B}$ (so that the composition $g \circ f$ is well defined).



We want to give conditions for the existence of the limit

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} (g \circ f)(\boldsymbol{x}).$$

Suppose that the limit $\boldsymbol{y}_0 = \lim\limits_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists, with $\boldsymbol{y}_0 \in \tilde{\mathcal{D}}(\mathcal{B})$, and also $\boldsymbol{\ell} = \lim\limits_{\boldsymbol{y} \to \boldsymbol{y}_0} g(\boldsymbol{y})$. Then, as $\boldsymbol{x}$ approaches $\boldsymbol{x}_0$, then $f(\boldsymbol{x})$ approaches $\boldsymbol{y}_0$, and therefore we expect $g \circ f$ to tend to $\boldsymbol{\ell}$. However, in general this conclusion is false. Indeed, in general, knowing that

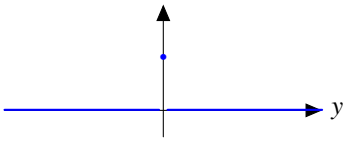
$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \boldsymbol{y}_0 \quad \text{and} \quad \lim_{\boldsymbol{y} \to \boldsymbol{y}_0} g(\boldsymbol{y}) = \boldsymbol{\ell}$$

does **not** allow us to conclude that

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} g(f(\boldsymbol{x})) = \boldsymbol{\ell}.$$

As an example of this phenomenon, consider the function $g : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$g(y) = \begin{cases} 0 & \text{if } y \neq 0 \\ 1 & \text{if } y = 0. \end{cases}$$



Consider also the identically zero function $f : \mathbb{R} \longrightarrow \mathbb{R}$, i.e. $f(x) = 0$ for all $x \in \mathbb{R}$. We have

$$\lim_{x \to 0} f(x) = 0 \quad \text{and} \quad \lim_{y \to 0} g(y) = 0,$$

but $(g \circ f)(x) = 1$ for all $x \in \mathbb{R}$, so that

$$\lim_{x \to 0} g(f(x)) = \lim_{x \to 0} g(0) = \lim_{x \to 0} 1 = 1 \neq 0.$$

The issue comes from the fact that $y_0 = 0$ and the value of the composite function $g \circ f$ is $g(y_0)$ at every point where the value of f is y_0, and the two values $g(y_0)$ and ℓ are different. To avoid this pathology, we can add the requirement

$$\forall\, \boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}, \quad f(\boldsymbol{x}) \neq \boldsymbol{y}_0,$$

as the following result shows. Other conditions, involving different assumptions to avoid this situation, will be presented later on.

Theorem 1.16 [Limit of the Composition]
Let $(\boldsymbol{X}, \|\cdot\|_X)$, $(\boldsymbol{Y}, \|\cdot\|_Y)$ and $(\boldsymbol{Z}, \|\cdot\|_Z)$ be normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $\boldsymbol{B} \subseteq \boldsymbol{Y}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ be such that $f(\boldsymbol{A}) \subseteq \boldsymbol{B}$ and let $g : \boldsymbol{B} \longrightarrow \boldsymbol{Z}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$, $\boldsymbol{y}_0 \in \tilde{\mathcal{D}}(\boldsymbol{B})$ and $\boldsymbol{\ell} \in \boldsymbol{Z} \cup \{\infty\}$. Let

$$f(\boldsymbol{A} \setminus \{\boldsymbol{x}_0\}) \subseteq \boldsymbol{B} \setminus \{\boldsymbol{y}_0\}. \tag{1.22}$$

Finally, let f be convergent at $\boldsymbol{x}_0$ and let g be convergent at $\boldsymbol{y}_0$, and

$$\boldsymbol{y}_0 = \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) \quad \textit{and} \quad \lim_{\boldsymbol{y} \to \boldsymbol{y}_0} g(\boldsymbol{y}) = \boldsymbol{\ell}.$$

Then,

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} (g \circ f)(\boldsymbol{x}) = \boldsymbol{\ell}.$$

Proof We have to prove that

$$\forall\, \boldsymbol{W} \in \mathcal{N}(\boldsymbol{\ell}, \boldsymbol{Z}, \|\cdot\|_Z),\ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}, \qquad \boldsymbol{x} \neq \boldsymbol{x}_0 \ \Rightarrow\ g(f(\boldsymbol{x})) \in \boldsymbol{W}.$$

Let $\boldsymbol{W} \in \mathcal{N}(\boldsymbol{\ell}, \boldsymbol{Z}, \|\cdot\|_Z)$. Since $g(\boldsymbol{y}) \overset{\boldsymbol{y} \to \boldsymbol{y}_0}{\longrightarrow} \boldsymbol{\ell}$,

$$\exists\, \boldsymbol{U} \in \mathcal{N}(\boldsymbol{y}_0, \boldsymbol{Y}, \|\cdot\|_Y) :\ \forall\, \boldsymbol{y} \in \boldsymbol{B} \cap \boldsymbol{U}, \qquad \boldsymbol{y} \neq \boldsymbol{y}_0 \ \Rightarrow\ g(\boldsymbol{y}) \in \boldsymbol{W}. \tag{1.23}$$

Moreover, since $f(\boldsymbol{x}) \overset{\boldsymbol{x} \to \boldsymbol{x}_0}{\longrightarrow} \boldsymbol{y}_0$

$$\exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}, \qquad \boldsymbol{x} \neq \boldsymbol{x}_0 \ \Rightarrow\ f(\boldsymbol{x}) \in \boldsymbol{U}.$$

From (1.22), we find that for all $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{A}$ is such that $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have $f(\boldsymbol{x}) \in \boldsymbol{B} \setminus \{\boldsymbol{y}_0\}$, so that from (1.23) we obtain

$$g(f(\boldsymbol{x})) \in \boldsymbol{W},$$

as needed.

Assumption (1.22) in Theorem 1.16 appears somehow awkward. However, as aforementioned, this is required to rule out situations like the counterexample introduced earlier. Indeed, in that case $f(\mathbb{R} \setminus \{0\}) = 0$, whereas (1.22)

in this case becomes $f(\mathbb{R} \setminus \{0\}) = \mathbb{R} \setminus \{0\}$.[4]

We now present another classic result on the limit of the restriction. We recall that if A, B, C are sets with $A \subseteq B$ and $f : B \longrightarrow C$, the restriction of f to A is the function

$$\begin{aligned} f|_A : A &\longrightarrow C \\ x &\mapsto f(x). \end{aligned}$$

Theorem 1.17 [Limit of the Restriction – I]
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $A, B \subseteq X$ be such that $A \subseteq B$. Let $f : B \longrightarrow Y$. Let $x_0 \in X$. Then:

(i) if $x_0 \in \tilde{\mathcal{D}}(A)$, then $x_0 \in \tilde{\mathcal{D}}(B)$;

(ii) if $x_0 \in \tilde{\mathcal{D}}(A)$, and if $\lim_{x \to x_0} f(x)$ exists (in Y or in $Y \cup \{\infty\}$), then $\lim_{x \to x_0} f|_A(x)$ exists, and the two limits coincide.

Proof The proof of the first point is straightforward: if $x_0 \in \tilde{\mathcal{D}}(A)$, for every neighborhood U of x_0, the set $U \cap A$ contains elements distinct from x_0. Since $U \cap A \subseteq U \cap B$, the same is true for $U \cap B$.
We prove the second point. If $\ell = \lim_{x \to x_0} f(x) \in Y \cup \{\infty\}$, for every neighborhood W of ℓ there exists a neighborhood V of x_0 such that, for all $x \in V \cap B$, if $x \neq x_0$ then $f(x) \in W$. This occurs, in particular, if $x \in V \cap A$ and $x \neq x_0$.

The partial converse of Theorem 1.17 is the following.

Theorem 1.18 [Limit of the Restriction – II]
Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Let $A, B \subseteq X$ be such that $A \subseteq B$. Let $x_0 \in \tilde{\mathcal{D}}(A)$ be such that there exists a neighborhood V of x_0 such that $(V \cap B) \setminus \{x_0\} \subseteq A$. Let the limit $\lim_{x \to x_0} f|_A(x)$ exist in $Y \cup \{\infty\}$ and be equal to ℓ. Then, $\lim_{x \to x_0} f(x)$ exists and it is equal to ℓ.

Proof Let W be a neighborhood of ℓ. Then, there exists a neighborhood U of x_0 such that, for all $x \in U \cap A$, if $x \neq x_0$ then $f(x) \in W$. Let $V' = V \cap U$. The intersection of two neighborhoods of x_0 is a neighborhood of x_0. Then, if $x \in V' \cap B$ and $x \neq x_0$, then $x \in A$ since $x \in V$, and, since $x \in U$, we have also $f(x) \in W$.

The figures below can help us gain some intuition: in **(a)**, the condition of Theorem 1.18 is satisfied: it is

[4] In order to avoid any constraints on the functions in the statement of Theorem 1.16, some authors adopt the different definition of limit

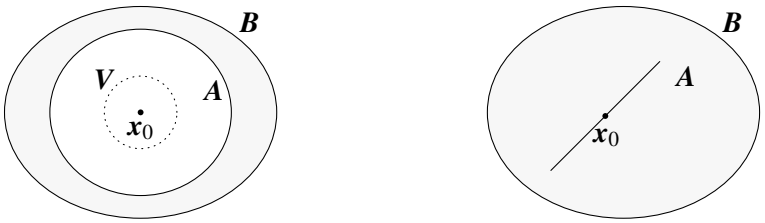
$$\forall\, W \in \mathcal{N}(\ell, Y, \|\cdot\|_Y),\ \exists\, V \in \mathcal{N}(x_0, X, \|\cdot\|_X) :\quad \forall\, x \in A \cap V, \quad f(x) \in W.$$

The notion of limit arising from this definition is often referred to as "deleted limit": with this definition, the pathology described here does not occur, and the theorem of the limit of the composition holds without the need of imposing (1.22). Moreover, in this case x_0 is allowed to be an adherent point of A, and not necessarily a limit point: in the case where x_0 is an isolated point of A, we will have $\lim_{x \to x_0} f(x) = f(x_0)$. There is, however, an important drawback with this definition: the deleted limit (as $x \to 0$) of the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases} \end{aligned}$$

does not exist: why?

possible to find a (punctured) neighborhood around $\boldsymbol{x}_0$ which is entirely contained in $\boldsymbol{A}$: the independent variable $\boldsymbol{x}$ can approach $\boldsymbol{x}_0$ from any direction (in $\boldsymbol{B}$) while remaining in $\boldsymbol{A}$. In **(b)**, the condition is not satisfied: in this case, $\boldsymbol{A}$ is a segment which is internal to $\boldsymbol{B}$, and given a point $\boldsymbol{x}_0$ on this segment, any neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ will contain points that are external to the segment. In this case, the fact that the limit of the restriction exists does not guarantee that the limit $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})$ exists: in fact, the limit of the restriction is defined on the basis of only one possible direction within $\boldsymbol{B}$.



Example *1.40*

By virtue of Theorem 1.17-1.18, given for example $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, the existence of the limit as (x, y) tends to $(0, 0)$ implies the existence of the limit for the restriction of f to any straight line through the origin, with the same value.

Conversely, suppose that the limit exists for the restriction to every line through the origin, and that the limit is the same for every line. Can we say that the limit of f as (x, y) tends to $(0, 0)$ exists? The answer is negative. Consider for example $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined for every $(x, y) \in \mathbb{R}^2$ by

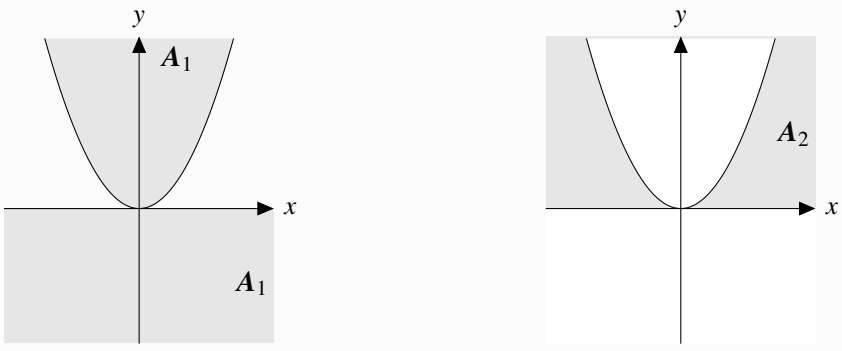
$$f(x, y) = \begin{cases} 1 & \text{if } y \geq x^2 \text{ or } y \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$\boldsymbol{A}_1 = \{(x, y) \in \mathbb{R}^2 \,|\, y \geq x^2 \;\text{ or }\; y \leq 0\}.$$

Then, $(0, 0) \in \mathcal{D}(\boldsymbol{A}_1)$ and, clearly, $\lim_{(x,y) \to (0,0)} f|_{\boldsymbol{A}_1}(x, y) = 1$. Moreover, let

$$\boldsymbol{A}_2 = \{(x, y) \in \mathbb{R}^2 \,|\, 0 < y < x^2\}.$$



Again, $(0, 0) \in \mathcal{D}(\boldsymbol{A}_2)$. In fact, for every $r > 0$, the ball $\mathcal{B}_r((0, 0))$ contains points in the form $(x, x^2/2)$ with $x > 0$ and $x^2 + x^4/4 < r^2$, which belong to $\boldsymbol{A}_2$ and are distinct from $(0, 0)$. We clearly have $\lim_{(x,y) \to (0,0)} f|_{\boldsymbol{A}_2}(x, y) = 0$. It follows that the limit does not exist as (x, y) approaches $(0, 0)$. However, let us consider the restriction of f to a line through the origin. Let $m > 0$ be the angular coefficient of this line, and let

$$\boldsymbol{A}_m = \{(x, m\,x) \in \mathbb{R}^2 \,|\, x \in \mathbb{R} \setminus \{0\}\}.$$

Then, from the definition of f, we have $f(x, m\,x) = 1$ if and only if

$$x \leq 0 \quad \text{or} \quad m\,x \geq x^2,$$

i.e., if and only if

$$x \leq 0 \quad \text{or} \quad \begin{cases} x \geq 0 \\ m\,x \geq x^2 \end{cases}$$

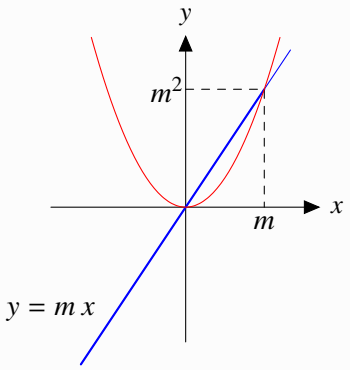
i.e., if and only if

$$x \leq 0 \quad \text{or} \quad \begin{cases} x \geq 0 \\ m \geq x \end{cases}$$

i.e., if and only if

$$x \leq 0 \quad \text{or} \quad 0 \leq x \leq m,$$

i.e., if and only if $x \in (-\infty, m)$ (see the Figure, where $\{(x, y) \in \boldsymbol{A}_m \,|\, f(x, y) = 1\}$ is represented by a thicker line).



If $(x, y) \in \boldsymbol{A}_m$ and $\|(x, y)\| < m$, then necessarily

$$x \leq |x| \leq \|(x, y)\| < m.$$

Thus, $f(x, y) = 1$. We conclude that

$$\lim_{(x,y) \to (0,0)} f|_{\boldsymbol{A}_m}(x, y) = 1.$$

1.4.2 Operations on limits

The following theorem result says that if the limit $\boldsymbol{\ell}$ is finite as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, in a punctured neighborhood of $\boldsymbol{x}_0$ (i.e., a neighborhood from which $\boldsymbol{x}_0$ is deleted) the function is bounded.

Lemma 1.1 *Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let f be convergent at $\boldsymbol{x}_0$, and let $\boldsymbol{\ell} = \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) \in \boldsymbol{Y}$. Then, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that $f|_{(\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}}$ is bounded, i.e.,*

$$\exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) : \ \exists\, M > 0 : \ \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \quad \|f(\boldsymbol{x})\|_Y < M.$$

Proof From our assumptions, for every $\varepsilon > 0$, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that for all $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{A}$ with $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. Thus, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, using the triangle inequality we find

$$\|f(\boldsymbol{x})\|_Y = \|f(\boldsymbol{x}) - \boldsymbol{\ell} + \boldsymbol{\ell}\|_Y \leq \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y + \|\boldsymbol{\ell}\|_Y < \varepsilon + \|\boldsymbol{\ell}\|_Y.$$

Taking $M \geq \varepsilon + \|\boldsymbol{\ell}\|_Y$ gives the desired result.

Theorem 1.19 *Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f, g : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let f and g converge to $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2 \in \boldsymbol{Y} \cup \{\infty\}$, respectively, as $\boldsymbol{x} \to \boldsymbol{x}_0$. Then:*

(i) if $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2 \in \boldsymbol{Y}$, the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x}) + g(\boldsymbol{x}))$ exists in $\boldsymbol{Y}$ and it is equal to $\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2$;

(ii) if $\boldsymbol{\ell}_1 \in \boldsymbol{Y}$ and $\boldsymbol{\ell}_2 = \infty$, the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x}) + g(\boldsymbol{x}))$ exists in $\boldsymbol{Y} \cup \{\infty\}$ and it is equal to ∞.

Proof *(i)* Let $\varepsilon > 0$. There exists a neighborhood $\boldsymbol{V}_1$ such that, if $\boldsymbol{x} \in \boldsymbol{V}_1 \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, then $\|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y < \frac{\varepsilon}{2}$. Likewise, there exists a neighborhood $\boldsymbol{V}_2$ such that, if $\boldsymbol{x} \in \boldsymbol{V}_2 \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, then $\|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \frac{\varepsilon}{2}$. The intersection of neighborhoods $\boldsymbol{V}_1 \cap \boldsymbol{V}_2$ of $\boldsymbol{x}_0$ is a neighborhood of $\boldsymbol{x}_0$, and for every $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{V}_2) \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$ both inequalities $\|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y < \frac{\varepsilon}{2}$ and $\|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \frac{\varepsilon}{2}$ hold true, so that using the triangle inequality we find

$$\begin{aligned} \|(f(\boldsymbol{x}) + g(\boldsymbol{x})) - (\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2)\|_Y &= \|(f(\boldsymbol{x}) - \boldsymbol{\ell}_1) + (g(\boldsymbol{x}) - \boldsymbol{\ell}_2)\|_Y \\ &\leq \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y + \big\|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(ii) We have to prove that

$$\forall\, H > 0, \quad \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0) : \quad \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \quad \|f(\boldsymbol{x}) + g(\boldsymbol{x})\|_Y > H.$$

Let $H > 0$. Since f converges to $\boldsymbol{\ell}_1$, which is finite, we can apply Lemma 1.1, which ensures that

$$\exists\, \boldsymbol{V}_1 \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) : \quad \exists\, M > 0 : \quad \forall\, \boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \quad \|f(\boldsymbol{x})\|_Y < M.$$

Since $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} g(\boldsymbol{x}) = \infty$,

$$\forall\, N > 0, \quad \exists\, \boldsymbol{V}_2 \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_X) : \quad \forall\, \boldsymbol{x} \in (\boldsymbol{V}_2 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \quad \|g(\boldsymbol{x})\|_Y > N.$$

We can take $N = M + H$, and we still denote by $\boldsymbol{V}_2$ the corresponding neighborhood. The intersection $\boldsymbol{V}_1 \cap \boldsymbol{V}_2$ is still a neighborhood of $\boldsymbol{x}_0$, and for all $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{V}_2) \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, using the reversed triangle inequality (see Theorem 1.2), we obtain

$$\|f(\boldsymbol{x}) + g(\boldsymbol{x})\|_Y \geq \big|\|f(\boldsymbol{x})\|_Y - \|g(\boldsymbol{x})\|_Y\big| = \|g(\boldsymbol{x})\|_Y - \|f(\boldsymbol{x})\|_Y \geq (H + M) - M = H.$$

Clearly, the case where $\boldsymbol{\ell}_1 = \infty$ and $\boldsymbol{\ell}_2 \in \boldsymbol{Y}$ can be brought back to the second case of Theorem 1.19 by interchanging the roles of f and g (normed spaces are vector spaces, and therefore the addition is commutative). The case where both $\boldsymbol{\ell}_1$ and $\boldsymbol{\ell}_2$ are equal to ∞ is not covered in the statement of Theorem 1.19. The reason is that in this situation we cannot draw a conclusion about the limit of the sum. We express this fact by saying that $\infty + \infty$ is an **indeterminate form**. To convince ourselves that we cannot establish a result that gives information about the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x}) + g(\boldsymbol{x}))$ under the only assumption that both functions be divergent as $\boldsymbol{x} \to \boldsymbol{x}_0$, consider the normed space $(\mathbb{R}, |\cdot|)$ and the functions $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ and $x_0 = \infty$. Then:

(i) if $f(x) = x$ and $g(x) = -x$, the sum converges to 0;

(ii) if $f(x) = 2\,x$ and $g(x) = -x$, the limit of then sum is ∞;

(iii) if $f(x) = 2\,x$ and $g(x) = x$, the limit of then sum is ∞;

(iv) if $f(x) = x + \sin x$ and $g(x) = -x$, the sum does not converge.

In this example, since the normed space is $\mathbb{R}$, one can consider the infinity with the sign. In this case, *(iii)* is not an indeterminate form. In particular, $(+\infty) + (+\infty)$ is not an indeterminate form, but $(+\infty) + (-\infty)$ is still an

indeterminate form.

We can write the result of Theorem 1.19 as

$$\lim_{x \to x_0} (f(x) + g(x)) = \lim_{x \to x_0} f(x) + \lim_{x \to x_0} g(x),$$

by introducing in $Y \cup \{\infty\}$ the new operation

$$\forall\, x \in Y, \qquad x + \infty = \infty + x = \infty.$$

Theorem 1.20 *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed spaces over the field $\mathbb{F}$ (where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$). Let $A \subseteq X$. Let $f : A \longrightarrow Y$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let f converge to $\ell \in Y \cup \{\infty\}$ as $x \to x_0$. Let $\lambda \in \mathbb{F}$. Then:*

(i) if $\ell \in Y$, the limit $\lim_{x \to x_0} (\lambda f)(x)$ exists in Y and it is equal to $\lambda \ell$;

(ii) if $\lambda = 0$ and $\ell = \infty$, the limit $\lim_{x \to x_0} (\lambda f)(x)$ exists in Y and it is equal to 0.

(iii) if $\lambda \neq 0$ and $\ell = \infty$, the limit $\lim_{x \to x_0} (\lambda f)(x)$ exists in $Y \cup \{\infty\}$ and it is equal to ∞.

Proof *(i)* We have to show that for all $\varepsilon > 0$ there exists a neighborhood V of x_0 such that, for all $x \in V \cap A$ and $x \neq x_0$, we have $\|\lambda f(x) - \lambda \ell\|_Y < \varepsilon$. The statement is trivially true if $\lambda = 0$. Suppose $\lambda \neq 0$. Let $\varepsilon > 0$. From the convergence of f to ℓ as $x \to x_0$, there exists a neighborhood V of x_0 such that, for all $x \in V \cap A$ and $x \neq x_0$, we have

$$\|f(x) - \ell\|_Y < \frac{\varepsilon}{|\lambda|}.$$

With this choice of x, using the homogeneity of the norm we find

$$\|\lambda f(x) - \lambda \ell\|_Y = |\lambda| \, \|f(x) - \ell\|_Y = |\lambda| \, \frac{\varepsilon}{|\lambda|} = \varepsilon,$$

as required;

(ii) if $\lambda = 0$, clearly λf is the identically zero function, and therefore the limit is 0;

(iii) We have to show that for all $N > 0$ there exists a neighborhood V of x_0 such that, for all $x \in V \cap A$ and $x \neq x_0$, we have $\|\lambda f(x)\|_Y > N$. Since $\lim_{x \to x_0} f(x) = \infty$, there exists a neighborhood V of x_0 such that, for all $x \in V \cap A$ and $x \neq x_0$, we have $\|f(x)\|_Y > \frac{N}{|\lambda|}$. For such values of x we have

$$\|\lambda f(x)\|_Y = |\lambda| \, \|f(x)\|_Y > |\lambda| \, \frac{N}{|\lambda|} = N.$$

We can write the result of Theorem 1.20 as

$$\lim_{x \to x_0} (\lambda f(x)) = \lambda \lim_{x \to x_0} f(x),$$

by introducing in $Y \cup \{\infty\}$ the new operation

$$\forall\, \lambda \in \mathbb{F}, \qquad \lambda \cdot \infty = \begin{cases} \infty & \text{if } \lambda \neq 0 \\ 0 & \text{if } \lambda = 0 \end{cases}$$

We observe that the counterpart of the results on the limit of the product and quotient of two functions cannot be established in the case of normed spaces in general. A definition for the product can be given, for example, in two particular cases:

- when Y is a field;

- by assuming that $\boldsymbol{Y}$ is an inner product vector space, and considering the inner product of $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$.

The same issue arises for the quotient, because in general we cannot divide by vectors. Again, in the particular case where $\boldsymbol{Y}$ is a field, a definition can be recovered.

Theorem 1.21 *Let $(\boldsymbol{X}, \|\cdot\|_x)$ and $(\mathbb{F}, |\cdot|)$ be two normed spaces, where $\mathbb{F}$ is a field (either $\mathbb{R}$ or $\mathbb{C}$). Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f, g : \boldsymbol{A} \longrightarrow \mathbb{F}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let f and g converge to $\ell_1, \ell_2 \in \mathbb{F} \cup \{\infty\}$, respectively, as $\boldsymbol{x} \to \boldsymbol{x}_0$. Then:*

(i) if $\ell_1, \ell_2 \in \mathbb{F}$, the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x})\, g(\boldsymbol{x}))$ exists in $\mathbb{F}$ and it is equal to $\ell_1\, \ell_2$;

(ii) If $\ell_1 \neq 0$ and $\ell_2 = \infty$, the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x})\, g(\boldsymbol{x}))$ exists in $\mathbb{F} \cup \{\infty\}$ and it is equal to ∞;

(iii) If $\ell_1 = \ell_2 = \infty$, the limit $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} (f(\boldsymbol{x})\, g(\boldsymbol{x}))$ exists in $\mathbb{F} \cup \{\infty\}$ and it is equal to ∞.

Proof *(i)* From Lemma 1.1, there exists a neighborhood $\boldsymbol{V}_1$ of $\boldsymbol{x}_0$ and there exists $M > 0$ such that, for all $\boldsymbol{x} \in \boldsymbol{V}_1 \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have $|g(\boldsymbol{x})| < M$. For $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, using the triangle inequality, we obtain

$$\begin{aligned} |f(\boldsymbol{x})\, g(\boldsymbol{x}) - \ell_1\, \ell_2| &= |f(\boldsymbol{x})\, g(\boldsymbol{x}) - \ell_1\, \ell_2 - \ell_1\, g(\boldsymbol{x}) + \ell_1\, g(\boldsymbol{x})| \\ &= |(f(\boldsymbol{x}) - \ell_1)\, g(\boldsymbol{x}) + \ell_1\, (g(\boldsymbol{x}) - \boldsymbol{\ell}_2)| \\ &\leq |f(\boldsymbol{x}) - \ell_1|\, |g(\boldsymbol{x})| + |\ell_1|\, |g(\boldsymbol{x}) - \ell_2|. \end{aligned}$$

Let $\varepsilon, \eta > 0$. There exists a neighborhood $\boldsymbol{V}_2$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}_2$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$ we have $|f(\boldsymbol{x}) - \ell_1| < \eta$, and there exists a neighborhood $\boldsymbol{V}_3$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}_3$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$ there holds $|g(\boldsymbol{x}) - \ell_2| < \eta$. The set $\boldsymbol{V} = \boldsymbol{V}_1 \cap \boldsymbol{V}_2 \cap \boldsymbol{V}_3$ is also a neighborhood of $\boldsymbol{x}_0$, and if $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$

$$|f(\boldsymbol{x})\, g(\boldsymbol{x}) - \ell_1\, \ell_2| \leq |f(\boldsymbol{x}) - \ell_1|\, |g(\boldsymbol{x})| + |\ell_1|\, |g(\boldsymbol{x}) - \ell_2| \leq (M + |\ell_1|)\, \eta.$$

It is always possible to choose $\eta > 0$ so that $(M + |\ell_1|)\, \eta < \varepsilon$: it suffices to choose $\eta < \frac{\varepsilon}{M+|\ell_1|}$. Hence, $\boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$ we have

$$|f(\boldsymbol{x})\, g(\boldsymbol{x}) - \ell_1\, \ell_2| < \varepsilon;$$

(ii) We now assume that $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = \ell_1 \in \mathbb{F} \setminus \{0\}$ and $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} g(\boldsymbol{x}) = \infty$. We have to show that for all $M > 0$ there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $|f(\boldsymbol{x})\, g(\boldsymbol{x})| > M$. Let $M > 0$. Since $\ell_1 \neq 0$, the set $\boldsymbol{W} = \mathbb{F} \setminus \mathcal{B}_{\frac{|\ell_1|}{2}}(0)$ is a neighborhood of ℓ_1. Hence, there exists a neighborhood $\boldsymbol{V}_1$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $f(\boldsymbol{x}) \in \boldsymbol{W}$, i.e., for all $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$,

$$|f(\boldsymbol{x})| \geq \frac{|\ell_1|}{2}.$$

Since g is divergent as $\boldsymbol{x} \to \boldsymbol{x}_0$, there exists a neighborhood $\boldsymbol{V}_2$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V}_2 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have

$$|g(\boldsymbol{x})| > 2\, \frac{M}{|\ell_1|}.$$

The set $\boldsymbol{V}_1 \cap \boldsymbol{V}_2$ is a neighborhood of $\boldsymbol{x}_0$, and for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have, using the properties of the absolute value,

$$|f(\boldsymbol{x})\, g(\boldsymbol{x})| = |f(\boldsymbol{x})|\, |g(\boldsymbol{x})| > \frac{|\ell_1|}{2} \cdot 2\, \frac{M}{|\ell_1|} = M.$$

(iii) Let $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} g(\boldsymbol{x}) = \infty$. We have to show that for all $M > 0$ there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $|f(\boldsymbol{x})\, g(\boldsymbol{x})| > M$. Let $M > 0$. The set $\boldsymbol{W} = \mathbb{F} \setminus \mathcal{B}_1(0)$ is a neighborhood of ∞, so that there exists a neighborhood $\boldsymbol{V}_1$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $f(\boldsymbol{x}) \in \boldsymbol{W}$, i.e., for all $\boldsymbol{x} \in (\boldsymbol{V}_1 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $|f(\boldsymbol{x})| > 1$. Since g is divergent as $\boldsymbol{x} \to \boldsymbol{x}_0$,

there exists a neighborhood $\boldsymbol{V}_2$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V}_2 \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $|g(\boldsymbol{x})| > M$. The set $\boldsymbol{V}_1 \cap \boldsymbol{V}_2$ is a neighborhood of $\boldsymbol{x}_0$, and for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we find

$$|f(\boldsymbol{x})\, g(\boldsymbol{x})| = |f(\boldsymbol{x})|\, |g(\boldsymbol{x})| > 1 \cdot M = M.$$

We can write the result of Theorem 1.21 as

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} (f(\boldsymbol{x})\, g(\boldsymbol{x})) = (\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}))(\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x}))$$

if we define

- $x \cdot \infty = \infty \cdot x = \infty$ for $x \in \mathbb{F} \setminus \{0\}$;
- $\infty \cdot \infty = \infty$.

Theorem 1.22 *Let $(\boldsymbol{X}, \|\cdot\|_x)$ and $(\mathbb{F}, |\cdot|)$ be two normed spaces, where $\mathbb{F}$ is a field (either $\mathbb{R}$ or $\mathbb{C}$). Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f, g : A \longrightarrow \mathbb{F}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let f and g converge to $\ell_1, \ell_2 \in \mathbb{F} \cup \{\infty\}$, respectively, as $\boldsymbol{x} \to \boldsymbol{x}_0$. Assume also that $g(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in A$. Then:*

(i) if $\ell_1, \ell_2 \in \mathbb{F}$ and $\ell_2 \neq 0$, the limit $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \dfrac{f(\boldsymbol{x})}{g(\boldsymbol{x})}$ exists in $\mathbb{F}$ and it is equal to $\dfrac{\ell_1}{\ell_2}$;

(ii) if $\ell_1 \neq 0$ and $\ell_2 = 0$, the limit $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \dfrac{f(\boldsymbol{x})}{g(\boldsymbol{x})}$ exists in $\mathbb{F} \cup \{\infty\}$ and it is equal to ∞;

(iii) if $\ell_1 \in \mathbb{F}$ and $\ell_2 = \infty$, the limit $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \dfrac{f(\boldsymbol{x})}{g(\boldsymbol{x})}$ exists in $\mathbb{F}$ and it is equal to 0.

Proof We first assume that $f(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{A}$. Consider the following cases:

(a) Let $\ell_2 \in \mathbb{F} \setminus \{0\}$. Let $\boldsymbol{x} \in \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have

$$\left|\frac{1}{g(\boldsymbol{x})} - \frac{1}{\ell_2}\right| = \frac{|g(\boldsymbol{x}) - \ell_2|}{|g(\boldsymbol{x})|\, |\ell_2|}.$$

Let $\eta_1 > 0$ and $\boldsymbol{x} \in \boldsymbol{A}$ such that $|g(\boldsymbol{x}) - \ell_2| < \eta_1$. Then,

$$|\ell_2| = |\ell_2 - g(\boldsymbol{x}) + g(\boldsymbol{x})| \leq |g(\boldsymbol{x})| + |\ell_2 - g(\boldsymbol{x})| \leq |g(\boldsymbol{x})| + \eta_1.$$

We find $|g(\boldsymbol{x})| \geq |\ell_2| - \eta_1$. We can choose $\eta_1 = \frac{|\ell_2|}{2}$. Therefore, there exists a neighborhood $\boldsymbol{V}_1$ of $\boldsymbol{x}_0$ such that, if $\boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}_1$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, then

$$|g(\boldsymbol{x})| \geq \frac{|\ell_2|}{2}.$$

With this choice of $\boldsymbol{x}$, we obtain

$$\left|\frac{1}{g(\boldsymbol{x})} - \frac{1}{\ell_2}\right| \leq 2\, \frac{|g(\boldsymbol{x}) - \ell_2|}{|\ell_2|^2}.$$

Let $\varepsilon > 0$. There exists a neighborhood $\boldsymbol{V}_2$ of $\boldsymbol{x}_0$ such that for all $\boldsymbol{x} \in \boldsymbol{V}_2 \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have $|g(\boldsymbol{x}) - \ell_2| < \frac{|\ell_2|^2 \varepsilon}{2}$. Then, if $\boldsymbol{V} = \boldsymbol{V}_1 \cap \boldsymbol{V}_2$, for all $\boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$ we have

$$\left|\frac{1}{g(\boldsymbol{x})} - \frac{1}{\ell_2}\right| \leq 2\, \frac{|g(\boldsymbol{x}) - \ell_2|}{|\ell_2|^2} < \varepsilon.$$

(b) Let $\ell_2 = 0$. We prove that $\frac{1}{g(\boldsymbol{x})}$ tends to ∞ as $\boldsymbol{x} \to \boldsymbol{x}_0$. Let $\boldsymbol{x} \in \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$, we have

$$\forall\, \varepsilon > 0, \quad \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_x) : \quad \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad |g(\boldsymbol{x})| < \varepsilon. \tag{1.24}$$

Any $\varepsilon > 0$ can be written as $1/M$, where $M > 0$ and *vice-versa*. It follows that (1.24) is equivalent to

$$\forall\, M > 0,\quad \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_x) :\quad \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad |g(\boldsymbol{x})| < \frac{1}{M}, \tag{1.25}$$

which is in turn equivalent to

$$\forall\, M > 0,\quad \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_x) :\quad \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad \frac{1}{|g(\boldsymbol{x})|} > M, \tag{1.26}$$

which says that $\frac{1}{g}$ is divergent as $\boldsymbol{x} \to \boldsymbol{x}_0$.

(c) Let $\ell_2 = \infty$. In this case, $\frac{1}{g(\boldsymbol{x})}$ tends to 0 as $\boldsymbol{x} \to \boldsymbol{x}_0$, as one can prove using the same argument of the previous point.

Notice that the cases $(\ell_1, \ell_2) = (0, 0)$ and $(\ell_1, \ell_2) = (\infty, \infty)$, are not covered by Theorem 1.22: these are indeterminate forms. By introducing, for $\boldsymbol{x} \in \boldsymbol{X}$, the rules

- $\frac{\boldsymbol{x}}{\infty} = 0$
- $\boldsymbol{x} \neq 0 \quad \Rightarrow \quad \frac{\boldsymbol{x}}{0} = \infty$
- $\frac{\infty}{0} = \infty$,

the result of Theorem 1.22 can be restated by saying that if the functions f and g are not both going to zero nor both going to ∞ as $\boldsymbol{x} \to \boldsymbol{x}_0$, we have

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \frac{\lim\limits_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x})}{\lim\limits_{\boldsymbol{x} \to \boldsymbol{x}_0} g(\boldsymbol{x})}.$$

We conclude this section with other classical results that we generalize to normed spaces. Since these involve the order relation $\leq$, we restrict our attention to functions whose codomain is $\mathbb{R}$.

Theorem 1.23 [Conservation of Sign]
Let $(\boldsymbol{X}, \|\cdot\|_x)$ be a normed space and consider the normed space $(\mathbb{R}, |\cdot|)$. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$. Let f converge to $\ell \in \mathbb{R}$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$. Then,

(i) if $\ell > 0$, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $f(\boldsymbol{x}) > 0$;

(ii) if $\ell < 0$, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, we have $f(\boldsymbol{x}) < 0$.

Proof Consider the first case. The set $(0, +\infty)$ is a neighborhood of ℓ, and therefore, from the definition of convergence, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}$, there holds $f(\boldsymbol{x}) \in (0, +\infty)$, i.e., $f(\boldsymbol{x}) > 0$. The second follows using the same argument.

The statements *(i-ii)* in Theorem 1.23 can be captured simultaneously by saying that, if the limit is finite and non-zero, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that

$$\forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad \operatorname{sgn} f(\boldsymbol{x}) = \operatorname{sgn} \ell,$$

where sgn is the *sign* or *signum* function $\operatorname{sgn} : \mathbb{R} \longrightarrow \mathbb{R}$ defined, for $x \in \mathbb{R}$, by

$$\operatorname{sgn} x = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

It is worth observing that the statement of Theorem 1.23 can be generalized to include functions that are divergent, provided that we endow $\mathbb{R}$ with the infinity with the sign. In fact, in this case, if $\ell = \lim\limits_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = +\infty$, with

the same proof we can see that there exists a neighborhood V of x_0 such that $f(x) > 0$ for all $x \in (V \cap A) \setminus \{x_0\}$. Obviously, in this case the extension is straightforward since $+\infty > 0$. Clearly, if $\ell = -\infty$, there will be a neighborhood V of x_0 such that $f(x) < 0$ for all $x \in (V \cap A) \setminus \{x_0\}$. This extension cannot be adapted to the case where $\mathbb{R}$ is endowed with the signless infinity ∞, because ∞ is neither positive nor negative. The same issue arises in the following results. The first can be considered the partial converse of Theorem 1.23.

Theorem 1.24 [Conservation of Sign – II]
Let $(X, \|\cdot\|_X)$ be a normed space and consider the normed space $(\mathbb{R}, |\cdot|)$. Let $A \subseteq X$. Let $f : A \longrightarrow \mathbb{R}$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let f converge to $\ell \in \mathbb{R}$ as x tends to x_0. Then,

(i) if there exists a neighborhood V of x_0 such that, for all $x \in (V \cap A) \setminus \{x_0\}$, we have $f(x) > 0$, then $\lim_{x \to x_0} f(x) \geq 0$;

(ii) if there exists a neighborhood V of x_0 such that, for all $x \in (V \cap A) \setminus \{x_0\}$, we have $f(x) < 0$, then $\lim_{x \to x_0} f(x) \leq 0$.

Proof Let $\ell = \lim_{x \to x_0} f(x)$. Let V be a neighborhood of x_0 such that, for all $x \in V \cap A$, we have $f(x) > 0$. Suppose by contradiction that $\ell < 0$. Then, in view of Theorem 1.23, there exists a neighborhood U of x_0 such that, for all $x \in U \cap A$, we have $f(x) < 0$. Let $x \in (U \cap V \cap A) \setminus \{x_0\}$. Then, $f(x) \geq 0$ and $f(x) < 0$ simultaneously, which is a contradiction.

The result of Theorem 1.24 can be adapted to the case where the limit is $+\infty$ or $-\infty$.

Theorem 1.25 [Comparison Theorem]
Let $(X, \|\cdot\|_X)$ be a normed space and consider the normed space $(\mathbb{R}, |\cdot|)$. Let $A \subseteq X$. Let $f, g : A \longrightarrow \mathbb{R}$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let f and g have limit in $\mathbb{R}$ as $x \to x_0$ and be such that

$$\forall\, x \in A, \quad f(x) \leq g(x).$$

Then,

$$\lim_{x \to x_0} f(x) \leq \lim_{x \to x_0} g(x).$$

Proof The statement follows immediately by applying Theorem 1.23 to the function $h : A \longrightarrow \mathbb{R}$ defined by $h = g - f$, which is non-negative.

Again, the result of Theorem 1.25 can be adapted to the case where the limit is $+\infty$ or $-\infty$.

Theorem 1.26 [Squeeze Theorem]
Let $(X, \|\cdot\|_X)$ be a normed space and consider the normed space $(\mathbb{R}, |\cdot|)$. Let $A \subseteq X$. Let $f, g, h : A \longrightarrow \mathbb{R}$ be such that $g(x) \leq f(x) \leq h(x)$ for all $x \in A$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let the limits of g and h exist in $\mathbb{R}$ as $x \to x_0$ and be equal. Then, the limit of f exists in $\mathbb{R}$ as $x \to x_0$ and

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \lim_{x \to x_0} h(x).$$

Proof Let $\lim_{x \to x_0} g(x) = \lim_{x \to x_0} h(x) = \ell \in \mathbb{R}$. We prove that $f \xrightarrow{x_0} \ell$. Let $\varepsilon > 0$. It is sufficient to determine a neighborhood V of x_0 such that for all $x \in (V \cap A) \setminus \{x_0\}$ we have $\ell - \varepsilon < f(x) < \ell + \varepsilon$. Since $\lim_{x \to x_0} g(x) = \ell$, there exists a neighborhood V_1 of x_0 such that for all $x \in (V_1 \cap A) \setminus \{x_0\}$ there holds $\ell - \varepsilon < g(x) < \ell + \varepsilon$.

Since $\lim_{x \to x_0} h(x) = \ell$, there exists a neighborhood V_2 of x_0 such that for all $x \in (V_2 \cap A) \setminus \{x_0\}$ there holds $\ell - \varepsilon < h(x) < \ell + \varepsilon$. Let $V = V_1 \cap V_2$. Then, if $x \in (V \cap A) \setminus \{x_0\}$, we have $\ell - \varepsilon < g(x) \leq f(x) \leq h(x) < \ell + \varepsilon$, so that $\ell - \varepsilon < f(x) < \ell + \varepsilon$.

Exercises

Exercise 1.51 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $A \subseteq X$. Let $x_0 \in \mathcal{D}(A)$. Consider a function $f : A \longrightarrow Y$ and let $\ell \in \mathbb{R}^m$. Show that the condition

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon \tag{1.27}$$

can be equivalently expressed as:

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A, \qquad 0 < \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon.$$

[Solution: (1.51)]

Exercise 1.52 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $A \subseteq X$. Let $x_0 \in \mathcal{D}(A)$. Consider a function $f : A \longrightarrow Y$. Prove the equivalence of the two convergence conditions to $\ell \in Y$:

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon \tag{1.28}$$

and

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y \leq \varepsilon. \tag{1.29}$$

[Solution: (1.52)]

Exercise 1.53 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $A \subseteq X$. Let $x_0 \in \mathcal{D}(A)$. Consider a function $f : A \longrightarrow Y$. Prove the equivalence of the two convergence conditions to $\ell \in Y$:

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon \tag{1.30}$$

and

$$\forall\, \varepsilon > 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X \leq \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon. \tag{1.31}$$

[Solution: (1.53)]

Exercise 1.54 Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two normed vector spaces. Let $A \subseteq X$. Let $x_0 \in \mathcal{D}(A)$. Consider a function $f : A \longrightarrow Y$. Prove the following conditions are **not** equivalent to the condition of limit:

(i) $\forall\, \varepsilon \geq 0,\ \exists \delta > 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon$

(ii) $\forall\, \varepsilon > 0,\ \exists \delta \geq 0 :\ \forall\, x \in A \setminus \{x_0\}, \qquad \|x - x_0\|_X < \delta \ \Rightarrow\ \|f(x) - \ell\|_Y < \varepsilon$

[Solution: (1.54)]

Exercise 1.55 Let $(X, \|\cdot\|_X)$ and $(\mathbb{F}, |\cdot|)$ be two normed spaces, where $\mathbb{F}$ is a field (either $\mathbb{R}$ or $\mathbb{C}$). Let $A \subseteq X$. Let $f : A \longrightarrow \mathbb{F}$. Let $x_0 \in \tilde{\mathcal{D}}(A)$. Let f be divergent as $x \to x_0$. Assume also that $f(x) \neq 0$ for all $x \in A$. Show that the function $\frac{1}{f}$ is convergent as $x \to x_0$, and the limit is zero. [Solution: (1.55)]

1.5 Equivalent norms

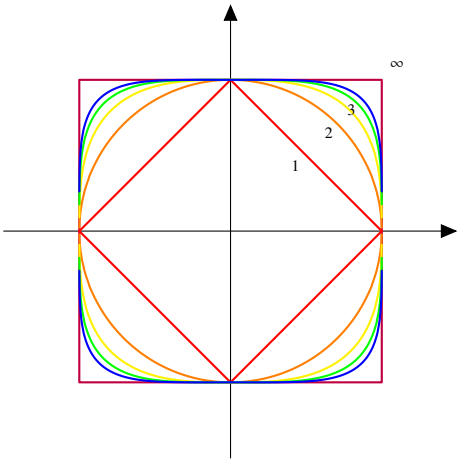
The notions of convergence and continuity can be both conveniently expressed in terms of neighborhoods. In particular, their definitions involve two neighborhoods: one in the space $\boldsymbol{X}$ where the domain of the function A lives, and the other in the codomain $\boldsymbol{Y}$. In turn, the shapes of these neighborhoods depend on which norms have been chosen in $\boldsymbol{X}$ and $\boldsymbol{Y}$. This begs the question: if, say, we can endow $\boldsymbol{Y}$ with two different norms $\|\cdot\|_Y^{(1)}$ and $\|\cdot\|_Y^{(2)}$, and we established that the limit a specific function with respect to $(\boldsymbol{Y}, \|\cdot\|_Y^{(1)})$ is equal to $\boldsymbol{\ell}$ as $\boldsymbol{x} \to \boldsymbol{x}_0$, can we say that the same function also converges with respect to the normed space $(\boldsymbol{Y}, \|\cdot\|_Y^{(2)})$ and, if so, can we guarantee that the limit is the same? The correct answer to these questions is: it depends. In fact, we will see that if the norms $\|\cdot\|_Y^{(1)}$ and $\|\cdot\|_Y^{(2)}$ are *equivalent*, in the sense that will be specified in the sequel, then the answer is affirmative: the properties of convergence/limit and continuity do not change if we change the norm with an equivalent norm. This is not necessarily the case if the norms are not equivalent.

Before presenting the definition of equivalent norm, we want to gain some insight. Let $n \in \mathbb{N}^*$. Consider the normed space $\mathbb{R}^n$. For every $p \in [1, +\infty]$, we can define the open ball centered at $\boldsymbol{x}_0 \in \mathbb{R}^n$ of radius $r > 0$ with respect to the norm $\|\cdot\|_p$ as

$$\mathcal{B}_r(\boldsymbol{x}_0, \mathbb{R}^n, \|\cdot\|_p) = \{\boldsymbol{x} \in \boldsymbol{X} \mid \|\boldsymbol{x} - \boldsymbol{x}_0\|_p < r\}.$$

For the sake of brevity, $\mathcal{B}_r(\boldsymbol{x}_0, \mathbb{R}^n, \|\cdot\|_p)$ will be referred to as a $\|\cdot\|_p$-ball. The different shapes of the unit ball of $\mathbb{R}^2$ centred at the origin for different values of p is shown in the figure:

$$\begin{aligned}
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_1) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| + |x_2| < 1\} && \text{(red)} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_2) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt{|x_1|^2 + |x_2|^2} < 1\} && \text{(orange)} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_3) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt[3]{|x_1|^3 + |x_2|^3} < 1\} && \text{(yellow)} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_4) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \sqrt[4]{|x_1|^4 + |x_2|^4} < 1\} && \text{(green)} \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_5) &= \{(x_1, x_2) \in \mathbb{R}^5 \mid \sqrt[5]{|x_1|^5 + |x_2|^5} < 1\} && \text{(blue)} \\
&\vdots \\
\mathcal{B}_1((0,0), \mathbb{R}^2, \|\cdot\|_\infty) &= \{(x_1, x_2) \in \mathbb{R}^2 \mid \max\{|x_1| + |x_2|\} < 1\} && \text{(purple)}
\end{aligned}$$



Likewise, $\mathcal{B}_1((0,0,0), \mathbb{R}^3, \|\cdot\|_1)$ is a regular octahedron (composed of eight equilateral triangles, four of which meet at each vertex), $\mathcal{B}_1((0,0,0), \mathbb{R}^3, \|\cdot\|_2)$ is a sphere and $\mathcal{B}_1((0,0,0), \mathbb{R}^3, \|\cdot\|_\infty)$ is a cube.

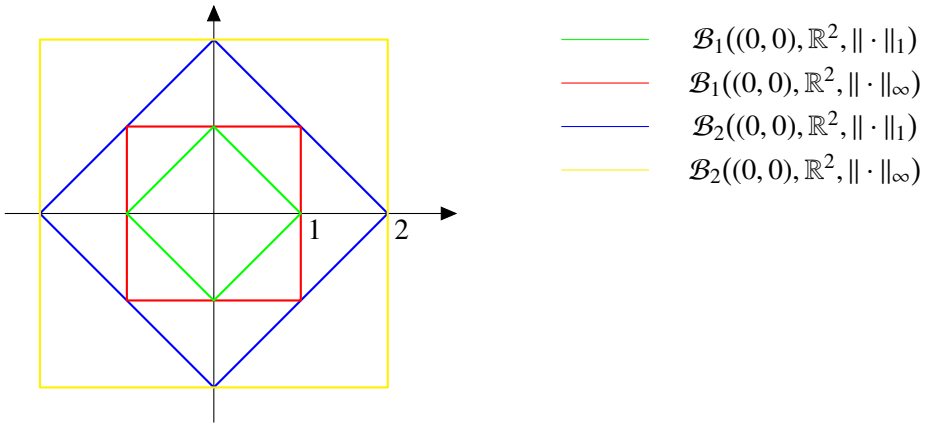
To help ourselves gain more insight, consider for example the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ in $\mathbb{R}^n$. We observe that for every $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i| \leq n \cdot \max_{1 \leq i \leq n} |x_i| = n \|\boldsymbol{x}\|_\infty \quad \text{and} \quad \|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} |x_i| \leq \sum_{i=1}^{n} |x_i| = \|\boldsymbol{x}\|_1.$$

Hence, there exists two constants c_1 and c_2 such that

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \qquad \|\boldsymbol{x}\|_1 \leq c_1 \|\boldsymbol{x}\|_\infty \quad \text{and} \quad \|\boldsymbol{x}\|_\infty \leq c_2 \|\boldsymbol{x}\|_1,$$

because we have seen that we can take $c_1 = n$ and $c_2 = 1$. In the case $n = 2$ the geometric interpretation is easy to visualize: the unit ball associated with the norm $\|\cdot\|_\infty$ is contained in the ball of radius 2 associated with the norm $\|\cdot\|_1$, and the unit ball of the norm $\|\cdot\|_1$ is contained in the unit ball of the norm $\|\cdot\|_\infty$.



The same situation occurs if the radii 1 and 2 are substituted with r and $2r$, respectively. More in general, we are interested in the case where, given two arbitrary values $p, q \in [1, +\infty]$, every $\|\cdot\|_p$-ball contains a concentric $\|\cdot\|_q$-ball (of different radius). This observation leads to the notion of equivalent norms.

Definition 1.24 [Equivalent Norms]
Let $\boldsymbol{X}$ be a real or complex vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two norms in $\boldsymbol{X}$. We say that $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are **equivalent** *if there exist $c_1, c_2 > 0$ such that*

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \le c_1 \|\boldsymbol{x}\|_{(2)} \ \text{ and } \ \|\boldsymbol{x}\|_{(2)} \le c_2 \|\boldsymbol{x}\|_{(1)}. \tag{1.32}$$

Exercise 1.56

Let $\boldsymbol{X}$ be a real vector space and let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two equivalent norms in $\boldsymbol{X}$, i.e., there exist $c_1, c_2 > 0$ which satisfy

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \le c_1 \|\boldsymbol{x}\|_{(2)} \ \text{ and } \ \|\boldsymbol{x}\|_{(2)} \le c_2 \|\boldsymbol{x}\|_{(1)}. \tag{1.33}$$

Show that:

(i) if $c_1, c_2 > 0$ satisfy (1.33), then $c_2 \ge \frac{1}{c_1}$;

(ii) (1.33) is equivalent to the condition

$$\exists\, M, m > 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad m\|\boldsymbol{x}\|_{(1)} \le \|\boldsymbol{x}\|_{(2)} \le M\|\boldsymbol{x}\|_{(1)}. \tag{1.34}$$

[Solution: (1.56)]

Exercise 1.57

Show that, as the name suggests, the equivalence of norms is an equivalence relation in the set of norms in a normed vector space. In other words, if we denote by $\sim$ the equivalence in the definition of equivalent norms, show that:

(i) the reflexive property holds: for every norm $\|\cdot\|$ in $\boldsymbol{X}$, we have $\|\cdot\| \sim \|\cdot\|$;

(ii) the symmetric property holds: for every two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ in $\boldsymbol{X}$, we have that $\|\cdot\|_1 \sim \|\cdot\|_2$ implies $\|\cdot\|_2 \sim \|\cdot\|_1$;

(iii) the transitive property holds: for every three norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_3$ in $\boldsymbol{X}$, we have that $\|\cdot\|_1 \sim \|\cdot\|_2$ and $\|\cdot\|_2 \sim \|\cdot\|_3$ imply $\|\cdot\|_1 \sim \|\cdot\|_3$.

[Solution: (1.57)]

Example *1.41*

Let $n \in \mathbb{N}^*$. Let $1 \le p < \infty$. Generalizing the argument in the beginning of this section, we show that the norms $\|\cdot\|_p$ and $\|\cdot\|_\infty$ in $\mathbb{R}^n$ are equivalent, i.e., we prove that there exist $c_1, c_2 > 0$ that satisfy (1.32). To this

end, we show that

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_p \le n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty,$$

so that (1.32) is satisfied with $c_1 = n^{\frac{1}{p}}$ and $c_2 = 1$. Indeed, for all $i \in \{1, \dots, n\}$ we have $|x_i| \le \|\boldsymbol{x}\|_\infty$, so that

$$\|\boldsymbol{x}\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \le \left(\sum_{k=1}^n \|\boldsymbol{x}\|_\infty^p\right)^{\frac{1}{p}} = \left(n \|\boldsymbol{x}\|_\infty^p\right)^{\frac{1}{p}} = n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty.$$

We now show that for every $\boldsymbol{x} \in \mathbb{R}^n$ we have $\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{x}\|_p$. In fact, for every $i \in \{1, \dots, n\}$

$$|x_i| \le (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}} = \|\boldsymbol{x}\|_p,$$

so that

$$\|\boldsymbol{x}\|_\infty = \max_{1 \le k \le n} \{|x_k|\} \le \|\boldsymbol{x}\|_p.$$

Example *1.42*

We prove that the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, by showing that

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|_2 \le \|\boldsymbol{x}\|_1 \le \sqrt{n}\, \|\boldsymbol{x}\|_2.$$

First, we notice that

$$(|x_1| + |x_2| + \dots + |x_n|)^2 = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 + \sum_{\substack{i,j \in \{1,\dots,n\} \\ i > j}} |x_i\, x_j|,$$

so that

$$\|\boldsymbol{x}\|_2^2 = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \le (|x_1| + |x_2| + \dots + |x_n|)^2 = \|\boldsymbol{x}\|_1^2.$$

To prove the second inequality, we use Cauchy-Schwartz theorem for the dot product: $|\boldsymbol{x} \cdot \boldsymbol{y}|^2 \le (\boldsymbol{x} \cdot \boldsymbol{x})(\boldsymbol{y} \cdot \boldsymbol{y})$, and

$$\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \dots + |x_n| = \sum_{i=1}^n |x_i| \cdot 1 \le \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^n 1^2\right)^{\frac{1}{2}} = \sqrt{n}\, \|\boldsymbol{x}\|_2.$$

It is easy to generalize the results of the previous two examples to the vector space $\mathbb{C}^n$ with the corresponding norms. In this case, for all $p \ge 1$ and all $\boldsymbol{z} \in \mathbb{C}^n$ we have

$$\|\boldsymbol{z}\|_\infty \le \|\boldsymbol{z}\|_p \le n^{\frac{1}{p}} \|\boldsymbol{z}\|_\infty$$
$$\|\boldsymbol{z}\|_2 \le \|\boldsymbol{z}\|_1 \le \sqrt{n}\, \|\boldsymbol{z}\|_2.$$

Geometrically, two norms are equivalent if and only if an open ball with respect to either one of the norms contains an open ball with respect to the other norm, as the following result shows.

Theorem 1.27 *Let $\boldsymbol{X}$ be a vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two equivalent norms in $\boldsymbol{X}$. Let $c_1, c_2 > 0$ be such that*

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \le c_1 \|\boldsymbol{x}\|_{(2)} \quad \text{and} \quad \|\boldsymbol{x}\|_{(2)} \le c_2 \|\boldsymbol{x}\|_{(1)}.$$

Let $r > 0$. Then, for every $\boldsymbol{x}_0 \in \boldsymbol{X}$, there hold

$$\mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)}) \subseteq \mathcal{B}_{c_1 r}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)}) \quad \text{and} \quad \mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)}) \subseteq \mathcal{B}_{c_2 r}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)}),$$

where

$$\mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(i)}) = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\|_{(i)} < r\}, \quad i \in \{1, 2\}.$$

Proof Consider the first. Let $\boldsymbol{x} \in \mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)})$. Then, $\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(2)} < r$. We must show that $\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(1)} < c_1\, r$. From the definition of equivalent norms, we have immediately

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(1)} \leq c_1\, \|\boldsymbol{x} - \boldsymbol{x}_0\|_{(2)} < c_1\, r,$$

as required. Try to prove the second in a similar way.

The result in Theorem 1.27 shows that the concepts of convergence/limit and the notion of continuity that we have introduced do not change if the norm considered in the development of the theory is replaced with an equivalent norm.

Theorem 1.28 *Let $\boldsymbol{X}$ be a vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two equivalent norms in $\boldsymbol{X}$. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$. Let $\boldsymbol{U} \subseteq \boldsymbol{X}$. Then, if $\boldsymbol{U}$ is a $\|\cdot\|_{(1)}$-neighbourhood of $\boldsymbol{x}_0$, then $\boldsymbol{U}$ is a $\|\cdot\|_{(2)}$-neighbourhood of $\boldsymbol{x}_0$.*

Proof Since $\boldsymbol{U}$ is a $\|\cdot\|_{(1)}$-neighborhood of $\boldsymbol{x}_0$, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)}) \subseteq \boldsymbol{U}$. We show that there exists $\rho > 0$ such that $\mathcal{B}_\rho(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)}) \subseteq \mathcal{B}_r(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)})$. In other words, we must be able to find $\rho > 0$ such that $\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(2)} < \rho$ implies $\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(1)} < r$. Since the norms are equivalent

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_{(1)} \leq\ c_1\, \|\boldsymbol{x} - \boldsymbol{x}_0\|_{(2)} < c_1\, \rho.$$

and therefore we can take $\rho < \frac{r}{c_1}$. It follows that $\boldsymbol{U}$ is a $\|\cdot\|_{(2)}$-neighborhood of $\boldsymbol{x}_0$.

It follows that if $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are equivalent norms in $\boldsymbol{X}$,

$$\forall\, \boldsymbol{x}_0 \in \boldsymbol{X}, \qquad \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)}) = \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)}). \tag{1.35}$$

The next, obvious question is: do we have a similar result for neighborhoods of ∞? The answer is a corollary of the following simple result.

Theorem 1.29 *Let $\boldsymbol{X}$ be a vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two equivalent norms in $\boldsymbol{X}$. Let A be bounded in $(\boldsymbol{X}, \|\cdot\|_{(1)})$. Then, A is bounded in $(\boldsymbol{X}, \|\cdot\|_{(2)})$.*

Proof Let $c_1, c_2 > 0$ satisfy the condition

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \leq c_1\, \|\boldsymbol{x}\|_{(2)} \ \text{ and } \ \|\boldsymbol{x}\|_{(2)} \leq c_2\, \|\boldsymbol{x}\|_{(1)}.$$

Since A is bounded in $(\boldsymbol{X}, \|\cdot\|_{(1)})$, there exists $M > 0$ such that, for all $\boldsymbol{x} \in A$, we have $\|\boldsymbol{x}\|_{(1)} \leq M$. We need to show that there exists $N > 0$ such that, for all $\boldsymbol{x} \in A$, we have $\|\boldsymbol{x}\|_{(2)} \leq N$. Since

$$\|\boldsymbol{x}\|_{(2)} \leq c_1\, \|\boldsymbol{x}\|_{(1)} \leq c_2\, M,$$

we can choose $N = c_2\, M$.

Since a neighborhood of ∞ is, by definition, a set whose complement is bounded, and since bounded sets do not change when we change the norm to an equivalent one, neighborhoods of ∞ also do not change when adopting an equivalent norm. Thus, we can generalize (1.35) by writing

$$\forall\, \boldsymbol{x}_0 \in \boldsymbol{X} \cup \{\infty\}, \qquad \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(1)}) = \mathcal{N}(\boldsymbol{x}_0, \boldsymbol{X}, \|\cdot\|_{(2)}).$$

The problem that arises naturally at this point is how to check whether two norms of the same vector space are equivalent or not. We would ideally hope that any two norms be equivalent: this would mean that all the notions developed for a normed vector space would continue to be valid independently of the norm. Unfortunately, we will see that this is not the case in general. Nevertheless, in the particular (but important) case of finite-dimensional vector spaces, all the norms are equivalent.

Exercises

Exercise 1.58 Let $1 \leq p \leq \infty$ and $1 \leq q \leq \infty$. Show that the norms $\|\cdot\|_p$ and $\|\cdot\|_q$ are equivalent in $\mathbb{R}^n$ (Hint: recall that $\|\cdot\|_p$ and $\|\cdot\|_\infty$ in $\mathbb{R}^n$ are equivalent). [Solution: (1.58)]

Exercise 1.59 Let

$$C^1([a,b],\mathbb{R}) = \{f : [a,b] \longrightarrow \mathbb{R} \mid f \text{ is differentiable and } f' \text{ is continuous}\}.$$

Prove that

$$\|f\|_1 \stackrel{\text{def}}{=} \max_{a \leq x \leq b} |f(x)| + \max_{a \leq x \leq b} |f'(x)|$$

and

$$\|f\|_2 \stackrel{\text{def}}{=} |f(a)| + \max_{a \leq x \leq b} |f'(x)|$$

are two norms. Show that the two norms are equivalent (Hint: use the fundamental theorem of calculus). [Solution: (1.59)]

Exercise 1.60 Let $\boldsymbol{X}$ be a (real or complex) vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be equivalent norms in $\boldsymbol{X}$. Show that the bounded sets are the same for each norm. [Solution: (1.60)]

Exercise 1.61 Consider a real vector space $\boldsymbol{X}$ with the norms $\|\cdot\|_{(1)}, \|\cdot\|_{(2)}, \ldots, \|\cdot\|_{(q)}$ over a field $\mathbb{F}$. Let $\alpha_1, \ldots, \alpha_q \geq 0$ be such that they are not all equal to zero. Show that a linear combination

$$\alpha_1 \|\cdot\|_{(1)} + \alpha_2 \|\cdot\|_{(2)} + \ldots + \alpha_q \|\cdot\|_{(q)}$$

is a norm for $\boldsymbol{X}$. [Solution: (1.61)]

Exercise 1.62 Which of the following functions $p : \mathbb{R}^3 \longrightarrow \mathbb{R}$ define a norm in $\mathbb{R}^3$?

- ☐ $p(x_1, x_2, x_3) = |x_1| + 2|x_2| + 3|x_3|$
- ☐ $p(x_1, x_2, x_3) = |x_1| + |x_2| + |x_3| + 1$
- ☐ $p(x_1, x_2, x_3) = |x_1| + |x_2| - |x_3|$
- ☐ $p(x_1, x_2, x_3) = |x_1 - x_2 + x_3|$
- ☐ $p(x_1, x_2, x_3) = |x_1 - x_2|$
- ☐ $p(x_1, x_2, x_3) = \max\{|x_1|, |x_2|, |x_3|\} + |x_1 + x_2|$
- ☐ $p(x_1, x_2, x_3) = \min\{|x_1|, |x_2|, |x_3|\}$
- ☐ $p(x_1, x_2, x_3) = \max\{x_1, x_2, x_3\}$
- ☐ $p(x_1, x_2, x_3) = |x_1| + |x_2|^2 + |x_3|^3$

[Solution: (1.62)]

Exercise 1.63 Consider a real or complex vector space $\boldsymbol{X}$. A function $p : \boldsymbol{X} \longrightarrow \mathbb{R}$ is a *seminorm* if

(i) it is homogeneous, i.e., for all $k \in \mathbb{F}$ and all $\boldsymbol{x} \in \boldsymbol{X}$ we have $p(k\,\boldsymbol{x}) = |k|\, p(\boldsymbol{x})$;

(ii) it is subadditive, i.e., for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$, the triangle inequality $p(\boldsymbol{x} + \boldsymbol{y}) \leq p(\boldsymbol{x}) + p(\boldsymbol{y})$ holds.

Show that for a seminorm there holds

- $p(\boldsymbol{0}) = 0$
- $p(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \boldsymbol{X}$.

It is not true, however, that $p(\boldsymbol{x}) = 0$ implies $\boldsymbol{x} = \mathbf{0}$. Differently from a norm, we do not have the positivity. With this in mind:

(i) given a norm $\|\cdot\|$ in $\mathbb{R}^m$, and given a matrix $A \in \mathbb{R}^{m\times n}$, show that

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|A\,\boldsymbol{x}\| \end{aligned}$$

is a seminorm in $\mathbb{R}^n$. Show that p is a norm in $\mathbb{R}^n$ if and only if $\operatorname{rank} A = n$;

(ii) Show that the sum of a seminorm and a norm is a norm.

[Solution: (1.63)]

Exercise 1.64

Let $\boldsymbol{X}$ be a normed vector space. Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be equivalent norms in $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k\in\mathbf{N}}$ be a sequence in $\boldsymbol{X}$ and let $\boldsymbol{x} \in \boldsymbol{V}$. Verify that

$$\|\boldsymbol{x}_k - \boldsymbol{x}\|_{(1)} \overset{k\to\infty}{\longrightarrow} 0 \quad \Leftrightarrow \quad \|\boldsymbol{x}_k - \boldsymbol{x}\|_{(2)} \overset{k\to\infty}{\longrightarrow} 0.$$

[Solution: (1.64)]

1.6 Sequences

Roughly speaking, a sequence is a collection of elements $x_0, x_1, x_2, \ldots$ in a set X. If, instead of a generic set X, we consider a vector space $\boldsymbol{X}$, we write the elements of the sequence as $\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$, by following the convention of denoting vector spaces/subspaces using bold capital letters and vectors using bold characters. This way of writing the elements of a sequence suggests that the subscript is free to vary in $\mathbb{N}$. We can therefore interpret the subscript as the value that the independent variable takes at specific values in $\mathbb{N}$. This is saying that sequences are, in essence, functions from $\mathbb{N}$ to elements in a set X – or in a vector space $\boldsymbol{X}$. For these functions, we will see that it makes sense to define the limit as the independent variable tends to $+\infty$.

Definition 1.25 [Sequence]
Let X be a set. A sequence with values in X is a function

$$\begin{aligned} x : \mathbb{N} &\longrightarrow X \\ k &\mapsto x_k. \end{aligned} \tag{1.36}$$

Notice that the letter x was previously used to denote the independent variable. In the notation introduced in Definition 1.25, the independent variable is k, and x is the name of the function. As already mentioned, if the set X of Definition 1.25 is a normed space, the function x defined in (1.36) is denoted using the bold character $\boldsymbol{x}$. Again, if instead of a generic set X we consider a vector space $\boldsymbol{X}$, the elements of the sequence (1.36) for different values of $k \in \mathbb{N}$ are denoted by $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots$ (instead of $\boldsymbol{x}(0), \boldsymbol{x}(1), \ldots$). Traditionally, the notation $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is used to represent a sequence instead of (1.36). As already mentioned, in this notation the subscript k plays the role of independent variable.[5] We also observe that in the notation $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ we are not explicitly saying what the codomain is. If we need to specify the codomain, we will say that "$(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a sequence of elements of X", or that "$(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is an X-valued sequence" (and we will talk about "real sequence" when X is $\mathbb{R}$). Clearly, even if not explicitly assigned, the codomain will have to be a set containing the image (or range) of $\boldsymbol{x}$, i.e., $\{\boldsymbol{x}_k \,|\, k \in \mathbb{N}\}$.

As already mentioned, for sequences the limit as $k \to +\infty$ is well defined. Indeed, the set $\mathbb{N}$ does not have limit

[5] It is very common to find n in place of k as the index of a sequence. However, we prefer using the letter k for the independent variable, because the letter n is, for us, "reserved" for the superscript of $\mathbb{R}^n$. Another common notation for a sequence is $(\boldsymbol{x}_k)_{k=0}^{\infty}$, or also $\{\boldsymbol{x}_k\}_{k\in\mathbb{N}}$.

points, i.e., $\mathcal{D}(\mathbb{N}) = \varnothing$.[6] On the other hand, $\mathbb{N}$ is not upper bounded, and therefore $+\infty$ is a generalized limit point of $\mathbb{N}$. In symbols, $\tilde{\mathcal{D}}(\mathbb{N}) = \{+\infty\}$. This means, in particular, that given a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$, it makes sense to ask ourselves if the limit

$$\lim_{k\to+\infty} \boldsymbol{x}_k \tag{1.37}$$

exists. More precisely, given the normed space $(\mathbb{R}, |\cdot|)$ and another normed space $(\boldsymbol{X}, \|\cdot\|)$, as well as a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ with values in $\boldsymbol{X}$, we can define the limit (1.37) by considering $A = \mathbb{N} \subset \mathbb{R}$ and $\boldsymbol{x}_0 = +\infty \in \tilde{\mathcal{D}}(\mathbb{N})$. In particular, given $\hat{\boldsymbol{x}} \in \boldsymbol{X} \cup \{\infty\}$, we can re-write (1.19) as

$$\forall\, \boldsymbol{W} \in \mathcal{N}(\hat{\boldsymbol{x}}),\ \exists\, \boldsymbol{V} \in \mathcal{N}(+\infty):\ \ \forall\, k \in \mathbb{N} \cap \boldsymbol{V}, \qquad k \neq +\infty \ \Rightarrow\ \boldsymbol{x}_k \in \boldsymbol{W}, \tag{1.38}$$

where, for the sake of brevity, we have denoted $\mathcal{N}(\hat{\boldsymbol{x}}, \boldsymbol{X}, \|\cdot\|_x)$ by $\mathcal{N}(\hat{\boldsymbol{x}})$ and $\mathcal{N}(+\infty, \mathbb{R}, |\cdot|)$ by $\mathcal{N}(+\infty)$. Since $+\infty \notin \mathbb{N}$, the condition $k \neq +\infty$ is automatically satisfied, and therefore we can re-write (1.38) as

$$\forall\, \boldsymbol{W} \in \mathcal{N}(\hat{\boldsymbol{x}}),\ \exists\, \boldsymbol{V} \in \mathcal{N}(+\infty):\ \ \forall\, k \in \mathbb{N} \cap \boldsymbol{V}, \qquad \boldsymbol{x}_k \in \boldsymbol{W}. \tag{1.39}$$

Neighborhoods of $+\infty$ are sets in the form $(p, +\infty)$ for some $p \in \mathbb{R}$. We can re-write (1.39) as

$$\forall\, \boldsymbol{W} \in \mathcal{N}(\hat{\boldsymbol{x}}),\ \exists\, p \in \mathbb{R}:\ \ \forall\, k \in \mathbb{N} \cap (p, +\infty), \qquad \boldsymbol{x}_k \in \boldsymbol{W}, \tag{1.40}$$

which can also be written as

$$\forall\, \boldsymbol{W} \in \mathcal{N}(\hat{\boldsymbol{x}}),\ \exists\, p \in \mathbb{R}:\ \ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \boldsymbol{x}_k \in \boldsymbol{W}. \tag{1.41}$$

In particular, the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent to $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ if (and only if)

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{R}:\ \ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|_x < \varepsilon, \tag{1.42}$$

and it diverges if (and only if)

$$\forall\, M > 0,\ \exists\, p \in \mathbb{R}:\ \ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \|\boldsymbol{x}_k\|_x > M. \tag{1.43}$$

Again, a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is oscillating if it is neither convergent nor divergent. Notice that in (1.42) and (1.43) one can equivalently assume p to be a natural number (why?).

Remark 1.11 We can generalize the definition of a sequence to include sequences with domain $A \subset \mathbb{N}$ (i.e., A strictly contained in $\mathbb{N}$). This is very convenient because, for example, we want to be able to handle sequences like $x_k = \frac{1}{k}$, which can be written as $(x_k)_{k\in\mathbb{N}^*}$. However, we are still interested in finding limit of sequences, and therefore this generalization is carried out provided that $+\infty$ remains a generalized limit point of A. In this case, we assume that A is obtained removing from $\mathbb{N}$ a finite subset of $\mathbb{N}$. This guarantees that there exists $p \in \mathbb{N}$ such that $\{k \in \mathbb{N} \,|\, k \geq p\} \subseteq \mathbb{N}$ (why?). If we want to generalize all the results on limits of sequences, we should be considering sequences in the form $(x_k)_{k\in\mathbb{N}_p}$, where $\mathbb{N}_p = \{k \in \mathbb{N} \,|\, k \geq p\}$ for some $p \in \mathbb{N}$.

Lastly, by convention, we will write $k \to \infty$ instead of $k \to +\infty$, so that (1.37) becomes

$$\lim_{k\to\infty} \boldsymbol{x}_k,$$

even if the symbol ∞ used here should not be confused with the "signless" ∞ introduced earlier: it is used for notational convenience.

Consider the normed space $(\mathbb{R}, |\cdot|)$, and a sequence $(x_k)_{k\in\mathbb{N}}$, with values in the real vector space $\mathbb{R}$. We recall that $(x_k)_{k\in\mathbb{N}}$ is *monotonic* or *monotone* if one of the following conditions are satisfied:[7]

[6] Indeed, let $x_0 \in \mathbb{R}$. If $x_0 < 0$, we take $r > 0$, and $r \leq -x_0$. Then, $\mathbb{N} \cap \mathcal{B}_r(x_0) = \mathbb{N} \cap (x_0 - r, x_0 + r) = \varnothing$. Let $x_0 \in \mathbb{N}$ and $0 < r \leq 1$. Then, $\mathbb{N} \cap \mathcal{B}_r(x_0) = \mathbb{N} \cap (x_0 - r, x_0 + r) = \{x_0\}$. This intersection does not contain elements which are distinct from x_0.
Let $x_0 > 0$ and $x_0 \notin \mathbb{N}$. Then, there exists $p \in \mathbb{N}$ such that $p < x_0 < p+1$. We take $0 < r < \min\{x_0 - p, p+1-x_0\}$. Then, $\mathbb{N} \cap \mathcal{B}_r(x_0) = \varnothing$.

[7] We avoid using the adjectives "decreasing" because some authors use it for the non-decreasing case and some others for the strictly decreasing case, and similarly for the increasing case.

- $\forall\, k \in \mathbb{N},\quad x_{k+1} \geq x_k$ (NON-DECREASING);
- $\forall\, k \in \mathbb{N},\quad x_{k+1} > x_k$ (STRICTLY INCREASING);
- $\forall\, k \in \mathbb{N},\quad x_{k+1} \leq x_k$ (NON-INCREASING);
- $\forall\, k \in \mathbb{N},\quad x_{k+1} < x_k$ (STRICTLY DECREASING).

We now show that if $(x_k)_{k\in\mathbb{N}}$ is monotonic, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}\cup\{\infty\}$; hence, in this case, the sequence cannot oscillate.

Theorem 1.30 *Let $(x_k)_{k\in\mathbb{N}}$ be a non-decreasing sequence with values in $\mathbb{R}$. Then, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}\cup\{\infty\}$. Moreover,*

- *if the sequence $(x_k)_{k\in\mathbb{N}}$ is upper bounded, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}$ and it coincides with $\sup_{k\in\mathbb{N}} x_k$;*
- *if the sequence $(x_k)_{k\in\mathbb{N}}$ is not upper bounded, the limit $\lim_{k\to\infty} x_k$ is ∞.*

Proof Suppose that $(x_k)_{k\in\mathbb{N}}$ is upper bounded. Let $S = \sup_{k\in\mathbb{N}} x_k$. We show that $\lim_{k\to\infty} x_k = S$. Let $\varepsilon > 0$. Hence, $S-\varepsilon$ is not an upper bound of the sequence. It follows that there exists $p\in\mathbb{N}$ such that $x_p > S-\varepsilon$. Let $k\in\mathbb{N}$ be such that $k > p$. Then,

$$S - \varepsilon < x_p \leq x_k.$$

On the other hand, from the definition of supremum, we have also $x_k \leq S < S+\varepsilon$, so that we obtain $|x_k - S| < \varepsilon$ for all $k > p$.
Now, we assume that the sequence is not upper bounded. Let $M\in\mathbb{R}$ be arbitrary. Since M is not an upper bound for $\{x_k \,|\, k\in\mathbb{N}\}$, there exists $p\in\mathbb{N}$ such that $M < x_p$. It follows that, if $k > p$, we have $M < x_p \leq x_k$.

A similar result holds also for non-increasing sequences. Its proof can be carried out in a similar way.

Theorem 1.31 *Let $(x_k)_{k\in\mathbb{N}}$ be a non-increasing sequence with values in $\mathbb{R}$. Then, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}\cup\{\infty\}$. Moreover,*

- *if the sequence $(x_k)_{k\in\mathbb{N}}$ is lower bounded, $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}$ and it coincides with $\inf_{k\in\mathbb{N}} x_k$;*
- *if the sequence $(x_k)_{k\in\mathbb{N}}$ is not lower bounded, the limit $\lim_{k\to\infty} x_k$ is ∞.*

We can also consider the definition of limit with respect to $\overline{\mathbb{R}}$. In this case, if $(x_k)_{k\in\mathbb{N}}$ is non-decreasing and not upper bounded, the limit is $+\infty$, and if $(x_k)_{k\in\mathbb{N}}$ is non-increasing and not lower bounded, the limit is $-\infty$.

The following result shows that convergence implies boundedness: this is a consequence of Lemma 1.1, but we will present an *ad-hoc* proof for clarity.

Theorem 1.32 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed vector space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a convergent sequence in $\boldsymbol{X}$. Then, $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is bounded.*

Proof Let $\hat{\boldsymbol{x}}\in\boldsymbol{X}$ be the limit of the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$, i.e.,

$$\lim_{k\to\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}.$$

From the definition of convergence, taking $\varepsilon = 1$, there exists $p\in\mathbb{N}$ such that, if $k > p$, we have $\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| < 1$.

Thus, if $k > p$,

$$\|\boldsymbol{x}_k\| = \|(\boldsymbol{x}_k - \hat{\boldsymbol{x}}) + \hat{\boldsymbol{x}}\| \leq \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| + \|\hat{\boldsymbol{x}}\| < 1 + \|\hat{\boldsymbol{x}}\|.$$

Setting $M = \max\{\|\boldsymbol{x}_1\|, \|\boldsymbol{x}_2\|, \dots, \|\boldsymbol{x}_p\|, 1 + \|\hat{\boldsymbol{x}}\|\}$, we obtain $\|\boldsymbol{x}_k\| \leq M$ for all $k \in \mathbb{N}$.

As already mentioned, it is easy to verify that the notion of limit does not change by going from one norm to an equivalent one. In particular, when dealing with a finite dimensional normed space, it is not necessary to specify what norm we are using to define a limit. The same observation is true when considering the notion of boundedness. The choice of the norm must be made explicit only if we are considering quantities that effectively depend on the norm, such as the distance of two points or the diameter of a set. Other results that continue to hold true in this more general setting are those that link the operation of limit to algebraic operations, by considering obviously only the operations that are well posed in a normed space: the sum of vectors and the product of vectors by scalars. In particular, we have a property of linearity.

Theorem 1.33 [Linearity]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. The set of convergent sequences of $\boldsymbol{X}$ is a vector space, and the function lim *in $\boldsymbol{X}$ that maps every convergent sequence to its limit is a linear operator (i.e., a linear transformation).*

Exercises

Exercise 1.65

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed vector space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence in $\boldsymbol{X}$ and let $\hat{\boldsymbol{x}} \in \boldsymbol{X}$. Verify that, if $\lim_{k\to\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}$, then we have also

$$\lim_{k\to\infty} \|\boldsymbol{x}_k\| = \|\hat{\boldsymbol{x}}\|.$$

(Hint: Use the reverse triangle inequality $\big|\|\boldsymbol{u}\| - \|\boldsymbol{v}\|\big| \leq \|\boldsymbol{u} - \boldsymbol{v}\|$). [Solution: (1.65)]

Exercise 1.66

Let $(x_k)_{k\in\mathbb{N}}$ be a non-decreasing sequence of $\mathbb{R}_+$. Let $(x_k)_{k\in\mathbb{N}}$ be convergent and let

$$\lim_{k\to\infty} x_k = 0.$$

Show that $x_k = 0$ for all $k \in \mathbb{N}$. [Solution: (1.66)]

Exercise 1.67

Consider a sequence $(x_k)_{k\in\mathbb{N}}$ of $\mathbb{R}$, bounded above, and let $a = \sup_{k\in\mathbb{N}} |x_k| \in \mathbb{R}$. Let $\ell = \lim_{k\to\infty} |x_k|$. Show that $\ell \leq a$. [Solution: (1.67)]

1.6.1 Sequences in $\overline{\mathbb{R}}$

We recall the definition of the real extended line $\overline{\mathbb{R}} = [-\infty, +\infty]$ as the set

$$[-\infty, +\infty] \stackrel{\text{def}}{=} \mathbb{R} \cup \{-\infty, +\infty\},$$

where the elements $-\infty$ and $+\infty$ are distinct from each other and from any other real number. We recall also that, given $a \in \overline{\mathbb{R}}$, we say that

- a is *finite* if $a \in \mathbb{R}$;
- a is *infinite* if $a = \pm\infty$.

We start by introducing an ordering in the set of extended real numbers. If $a, b \in \overline{\mathbb{R}}$, we say that $a \leq b$ (or, equivalently, $b \geq a$) if one of the following statements is true:

(i) if $a, b \in \mathbb{R}$, then $a \leq b$ in the usual sense;

(ii) $b = +\infty$;

(iii) $a = -\infty$.

We also say that $a < b$ (or, equivalently, $b > a$) if $a \leq b$ and $a \neq b$. It easy to see that this order relation is total, and this means that the following four properties hold true:

(i) $a \leq a$ (REFLEXIVITY)

(ii) $a \leq b$ and $b \leq c$ imply $a \leq c$ (TRANSITIVITY)

(iii) $a \leq b$ and $b \leq a$ imply $a = b$ (ANTI-SYMMETRY)

(iv) $a < b$ or $b > a$ or $a = b$ (LAW OF TRICHOTOMY).

We now introduce operations in $\overline{\mathbb{R}}$, such as addition, multiplication, subtraction and division. This is a delicate step, because these operations will not always obey the familiar rules of algebra; for this reason, they have to be handled with care.

Given $a, b \in \overline{\mathbb{R}}$, we define the sum $a + b$ in terms of the usual sum when a and b are finite, and we assign the sum when a or b are infinite:

$$a + b = \begin{cases} a + b & \text{if } a, b \in \mathbb{R} \\ +\infty & \text{if } (a \neq -\infty \text{ and } b = +\infty), \text{ or if } (a = +\infty \text{ and } b \neq -\infty) \\ -\infty & \text{if } (a \neq +\infty \text{ and } b = -\infty), \text{ or if } (a = -\infty \text{ and } b \neq +\infty) \\ 0 & \text{if } (a = +\infty \text{ and } b = -\infty), \text{ or if } (a = -\infty \text{ and } b = +\infty). \end{cases}$$

Notice that we assigned the value 0 to the sums $(+\infty) + (-\infty)$ and $(-\infty) + (+\infty) = 0$. This value is conventional: in the use we will make of the set $[-\infty, +\infty]$ in measure theory, we will avoid this occurrence. However, assigning a value to $a + b$ for any values of $a, b \in [-\infty, +\infty]$ allows us to view the operation $+$ as a function $+ : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \longrightarrow \overline{\mathbb{R}}$.

The addition in $\overline{\mathbb{R}}$ thus defined is commutative, and its identity element is 0 for every $a \in \overline{\mathbb{R}}$. This operation is not associative. Indeed,

$$\big((+\infty) + (+\infty)\big) + (-\infty) = (+\infty) + (-\infty) = 0 \quad \text{but} \quad (+\infty) + \big((+\infty) + (-\infty)\big) = (+\infty) + 0 = +\infty.$$

We can use this new operation of sum in $\overline{\mathbb{R}}$ to define the additive inverse in $\overline{\mathbb{R}}$. If $a \in \overline{\mathbb{R}}$,

$$-a \overset{\text{def}}{=} \begin{cases} -a & \text{if } a \in \mathbb{R} \\ +\infty & \text{if } a = -\infty \\ -\infty & \text{if } a = +\infty. \end{cases}$$

We can define the product in $\overline{\mathbb{R}}$ in the following way. If $a, b \in \overline{\mathbb{R}}$, then:

$$a \cdot b = \begin{cases} a \cdot b & \text{if } a, b \in \mathbb{R} \\ +\infty & \text{if } (a > 0 \text{ and } b = +\infty), \text{ or if } (a = +\infty \text{ and } b > 0), \\ & \text{or if } (a < 0 \text{ and } b = -\infty), \text{ or if } (a = -\infty \text{ and } b < 0) \\ -\infty & \text{if } (a > 0 \text{ and } b = -\infty), \text{ or if } (a = -\infty \text{ and } b > 0) \\ & \text{or if } (a < 0 \text{ and } b = +\infty), \text{ or if } (a = +\infty \text{ and } b < 0) \\ 0 & \text{if } (a = 0 \text{ and } b = +\infty), \text{ or if } (a = +\infty \text{ and } b = 0) \\ & \text{or if } (a = 0 \text{ and } b = -\infty), \text{ or if } (a = -\infty \text{ and } b = 0). \end{cases}$$

The operation of product in $[-\infty, +\infty]$ thus defined is commutative and associative, and its identity element is 1. The multiplicative inverse is defined in this way: if $a \in [-\infty, +\infty]$, let

$$\frac{1}{a} \overset{\text{def}}{=} \begin{cases} \frac{1}{a} & \text{if } a \in \mathbb{R}^* \\ +\infty & \text{if } a = 0 \\ 0 & \text{if } a = +\infty \text{ or if } a = -\infty. \end{cases}$$

We define the operation of *subtraction* that operates in this way: if $a, b \in \overline{\mathbb{R}}$, we write

$$a - b \stackrel{\text{def}}{=} a + (-b).$$

The element $a - b \in \overline{\mathbb{R}}$ is called *difference* of a and b. Likewise, we define a new operation called *division* that operates in this way: if $a, b \in \overline{\mathbb{R}}$, we write

$$\frac{a}{b} \stackrel{\text{def}}{=} a \frac{1}{b}.$$

The element $\frac{a}{b} \in \overline{\mathbb{R}}$ is called *quotient* of a and b.

In this way, we have defined 4 operations in $\overline{\mathbb{R}}$, called addition, multiplication, subtraction and division, and we refer to their respective results as sum, product, difference and quotient. By construction, these operations behave exactly as the corresponding operations in $\mathbb{R}$ when their operands are finite. We have seen that the associative rule of the addition does not hold if one of the addends is infinite. It is also possible to show that the multiplication is not distributive over the addition; in other words, if $a, b, c \in \overline{\mathbb{R}}$, in general $a\,(b + c) = a\,b + a\,c$ does not hold (see Exercise 1.72).

Another standard rule of algebra that does not hold in $\overline{\mathbb{R}}$ is the so-called *cancellation law*. You know that, if $a, b, c \in \mathbb{R}$, we have

$$a + b = c + b \quad \Rightarrow \quad a = c.$$

However, this is not true in general if, instead, $a, b, c \in \overline{\mathbb{R}}$. Indeed, $a + (+\infty) = c + (+\infty)$ (which is certainly satisfied for any finite values of a and c) does not imply $a = c$.

Given a set $A \subseteq \overline{\mathbb{R}}$, we can define its maximum and minimum in the following way. We say that $M \in \overline{\mathbb{R}}$ is maximum of A if $M \in A$ and for all $x \in A$, we have $x \leq M$. Likewise, $m \in \overline{\mathbb{R}}$ is minimum of A if $m \in A$ and for all $x \in A$, we have $x \geq m$. The maximum, when it exists, is unique; the same is true also for the minimum, and the proof is the same as in the standard case.[8] The uniqueness of maximum and minimum allows us to talk about "the maximum" of a set instead of "a maximum". We can also denote them unambiguously by $\max A$ and $\min A$.

Let A be a subset of $\overline{\mathbb{R}}$. We say that $U \in \overline{\mathbb{R}}$ is an *upper bound* of A if $x \leq U$ for all $x \in A$. Likewise, we say that $L \in \overline{\mathbb{R}}$ is an *lower bound* of A if $x \geq L$ for all $x \in A$. For example if $A = \mathbb{R}$ is thought of as a subset of $\overline{\mathbb{R}}$, then $+\infty$ is an upper bound for A (and the same is true if we consider $A = \overline{\mathbb{R}}$).

If $A = \varnothing$, every element of $\overline{\mathbb{R}}$ is an upper bound (and also a lower bound) of A.

Let $A \subseteq \overline{\mathbb{R}}$. Let $\alpha \in \overline{\mathbb{R}}$. We say that

- α is *supremum* of A if α is the minimum of the set of upper bounds of A, i.e.,
  (a) $\forall\, x \in A, \quad x \leq \alpha$;
  (b) $\forall\, M \in [-\infty, +\infty]$ such that M is an upper bound of A, we have $\alpha \leq M$.
- α is *infimum* of A if if α is the maximum of the set of lower bounds of A, i.e.,
  (a) $\forall\, x \in A, \quad \alpha \leq x$;
  (b) $\forall\, L \in [-\infty, +\infty]$ such that L is a lower bound of A, we have $L \leq \alpha$.

[8] Indeed, if $M_1, M_2 \in \overline{\mathbb{R}}$ both satisfy the conditions to be a maximum of A, then:

$$M_1 \in A \quad \text{and} \quad \forall\, x \in A, \quad x \leq M_1$$
$$M_2 \in A \quad \text{and} \quad \forall\, x \in A, \quad x \leq M_2.$$

Choosing $x = M_2$ in the first and choosing $x = M_1$ in the second yield, respectively, $M_2 \leq M_1$ and $M_1 \leq M_2$, so that $M_1 = M_2$.

We recall that the set of real numbers $\mathbb{R}$ with the ordering $\leq$ is **complete**: given $A \subseteq \mathbb{R}$ with $A \neq \emptyset$,

- if A is bounded from above, the supremum exists (in $\mathbb{R}$);
- if A is bounded from below, the infimum exists (in $\mathbb{R}$).

We observe that in $\overline{\mathbb{R}}$ the supremum and infimum always exist (as elements of $\overline{\mathbb{R}}$). Indeed, given $A \subseteq \overline{\mathbb{R}}$, we have:

- if A is non-empty and there exists $u \in \mathbb{R}$ such that $x \leq u$ for all $x \in A$ (so that $+\infty \notin A$), then $\sup A$ exists in $\mathbb{R}$ and is the usual one;
- if A is non-empty and there does not exist $u \in \mathbb{R}$ such that $x \leq u$ for all $x \in A$ (but we could have $+\infty \in A$ or $+\infty \notin A$), then $\sup A$ exists in $\overline{\mathbb{R}}$ and it is equal to $+\infty$;
- if A is empty, then $\sup A = -\infty$. Indeed, when $A = \emptyset$, the set of upper bounds is $\overline{\mathbb{R}}$. The minimum of this set is $-\infty$. Hence, with this definition $\sup \emptyset = -\infty$.

Example *1.43* If $A = \mathbb{N}$, the supremum is not defined in the usual sense, but we can write $\sup A = +\infty$ if we view A as a subset of $\overline{\mathbb{R}}$, because the set of upper bounds of $\mathbb{N}$ is $\{+\infty\}$, whose minimum is $+\infty$.

Likewise,

- if A is non-empty and there exists $\ell \in \mathbb{R}$ such that $x \geq \ell$ for all $x \in A$ (so that $-\infty \notin A$), then $\inf A$ exists in $\mathbb{R}$ and is the usual one;
- if A is non-empty and there does not exist $\ell \in \mathbb{R}$ such that $x \geq \ell$ for all $x \in A$ (but we could have $-\infty \in A$ or $-\infty \notin A$), then $\inf A$ exists in $\overline{\mathbb{R}}$ and it is equal to $-\infty$;
- if A is empty, then $\inf A = +\infty$, since the set of lower bounds is $\overline{\mathbb{R}}$.

It is straightforward to see that for every non-empty set $A \subseteq \overline{\mathbb{R}}$ we have $\sup A \geq \inf A$. The only exception to this inequality is the case where $A = \emptyset$, because as aforementioned $\sup \emptyset = -\infty$ and $\inf \emptyset = +\infty$.

From this discussion, it turns out that $\overline{\mathbb{R}}$ is complete with respect to the ordering $\leq$ considered here: every subset of $\overline{\mathbb{R}}$ admits supremum and infimum.

From now on, unless explicitly specified, we will be using the terms "supremum" and "infimum" according to this new definition, as elements of $\overline{\mathbb{R}}$, even when applied to subsets of $\mathbb{R}$.

Example *1.44* Let $A = \mathbb{R}_+ \cup \{+\infty\} = [0, +\infty]$. The set of upper bounds of A is $\{+\infty\}$, so that $\sup A = +\infty$. The set of lower bounds is $[-\infty, 0]$, and therefore $\inf A = 0$.

We can now define the supremum and infimum of a sequence of elements in $\overline{\mathbb{R}}$. Let $A = \mathbb{N} \setminus F$, where F is a finite set. A sequence of elements in $\overline{\mathbb{R}}$ is a function $x : A \longrightarrow \overline{\mathbb{R}}$. We define its supremum and infimum in the usual way as

$$\sup_{k \in A} x_k \stackrel{\text{def}}{=} \sup\{x_k \in \overline{\mathbb{R}} \,|\, k \in A\}$$
$$\inf_{k \in A} x_k \stackrel{\text{def}}{=} \inf\{x_k \in \overline{\mathbb{R}} \,|\, k \in A\}.$$

If, for example, $A = \{k \in \mathbb{N} \,|\, k \geq 5\}$, we will write $\sup_{k \geq 5} x_k$.

Example *1.45* Consider the sequence $(x_k)_{k \in \mathbb{N}}$ where $x_k = k$ for all $k \in \mathbb{N}$. Then,

$$\sup_{k \in \mathbb{N}} x_k = \sup\{x_k \,|\, k \in \mathbb{N}\} = \sup \mathbb{N} = \{+\infty\}$$
$$\inf_{k \in \mathbb{N}} x_k = \inf \mathbb{N} = 0.$$

We have the following simple results.

Theorem 1.34 *Let $A = \mathbb{N} \setminus F$, where F is a finite set. Let $(x_k)_{k\in A}$ be a sequence of real numbers or of elements in $\overline{\mathbb{R}}$. Let $\sup_{k\in A} x_k$ and $\inf_{k\in A} x_k$ be the supremum and the infimum of $(x_k)_{k\in A}$, respectively. Then:*

(1) *for every $k \in A$, we have $\inf_{h\in A} x_h \leq x_k \leq \sup_{h\in A} x_h$;*

(2) *if $M \in \overline{\mathbb{R}}$ is an upper bound for $(x_k)_{k\in A}$, then $\sup_{k\in A} x_k \leq M$;*

(3) *for every $y \in \overline{\mathbb{R}}$ such that $y < \sup_{h\in A} x_h$, there exists $k \in A$ such that $y < x_k \leq \sup_{h\in A} x_h$.*

Proof The first is trivial: the index set A is non-empty because F is finite, and therefore every upper bound of A is $\geq$ than any lower bound. The second is an immediate consequence of the fact that the supremum is the least upper bound. Finally, the third follows from a simple argument by contradiction: if such $k \in A$ did not exist, y would be an upper bound, and since $y < \sup_{h\in A} x_h$, it would be smaller than the least upper bound.

Given a sequence $(x_k)_{k\in\mathbb{N}}$ of elements in $\overline{\mathbb{R}}$, we define the limit as follows.

Definition 1.26 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\overline{\mathbb{R}}$. Let $\hat{x} \in \mathbb{R}$. We write*

$$\lim_{k\to\infty} x_k = \hat{x}$$

if there exists $N \in \mathbb{N}$ such that:

(A) *$x_k \in \mathbb{R}$ for all $k > N$*

(B) *for every $\varepsilon > 0$ there exists $p \in \mathbb{N}$ and $p > N$ such that*

$$k > p \quad \Rightarrow \quad |x_k - \hat{x}| < \varepsilon.$$

We write

$$\lim_{k\to\infty} x_k = +\infty$$

if for every $M \in \mathbb{R}$ there exists $p \in \mathbb{N}$ such that for every $k \in \mathbb{N}$,

$$k > p \quad \Rightarrow \quad x_k > M.$$

We write

$$\lim_{k\to\infty} x_k = -\infty$$

if for every $M \in \mathbb{R}$ there exists $p \in \mathbb{N}$ such that for every $k \in \mathbb{N}$,

$$k > p \quad \Rightarrow \quad x_k < M.$$

We can treat the case where the limit is $+\infty$ or $-\infty$ by assuming $\hat{x} \in [-\infty, +\infty]$, and considering this case to be a convergence in $[-\infty, +\infty]$. In other words, if $\hat{x} \in [-\infty, +\infty]$, we can write

$$\lim_{k\to\infty} x_k = \hat{x}$$

to capture both cases of Definition 1.26 simultaneously. In this way, we can view divergence as a particular type of convergence: in this case we talk about convergence in $[-\infty, +\infty]$.

Recall that if $(x_k)_{k\in\mathbb{N}}$ is a non-decreasing sequence of real numbers, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R} \cup \{\pm\infty\}$. Moreover, we have seen in the first set of notes that

- if $(x_k)_{k\in\mathbb{N}}$ is upper bounded, the limit $\lim_{k\to\infty} x_k$ exists in $\mathbb{R}$ and it coincides with $\sup_{k\in\mathbb{N}} x_k$;
- if $(x_k)_{k\in\mathbb{N}}$ is not upper bounded, the limit $\lim_{k\to\infty} x_k$ is $+\infty$.

However, considering the supremum as defined in this section, the previous two cases are unified: if $(x_k)_{k\in\mathbb{N}}$ is non-decreasing, the limit $\lim_{k\to\infty} x_k$ exists and is equal to $\sup_{k\in\mathbb{N}} x_k$. Likewise, if $(x_k)_{k\in\mathbb{N}}$ is non-increasing, the limit $\lim_{k\to\infty} x_k$ exists and is equal to $\inf_{k\in\mathbb{N}} x_k$.

Exercises

Exercise 1.68 Let $A, B \subset \overline{\mathbb{R}}$ be non-empty sets. Show that:

(i) $\inf A \le \sup A$;

(ii) $A \subseteq B$ implies $\sup A \le \sup B$ and $\inf A \ge \inf B$.

[Solution: (1.68)]

Exercise 1.69 Show that the ordering $\le$ in $\overline{\mathbb{R}}$ is a total order. In other words, prove that for all $a, b, c \in \overline{\mathbb{R}}$, the following properties hold:

(i) $a \le a$ (REFLEXIVE PROPERTY)

(ii) $a \le b$ and $b \le c$ imply $a \le c$ (TRANSITIVE PROPERTY)

(iii) $a \le b$ and $b \le a$ imply $a = b$ (ANTI-SYMMETRIC PROPERTY)

(iv) $a \le b$ or $b \le a$ (TOTAL ORDERING).

[Solution: (1.69)]

Exercise 1.70 Prove the commutative property of the addition in $\overline{\mathbb{R}}$, and identify all the cases where the associativity is not satisfied, to show that associativity holds in any subset of $(-\infty, +\infty]$ (in other words, the lack of associativity is related to one or more addends being equal to $-\infty$). [Solution: (1.70)]

Exercise 1.71 Prove the commutative and the associativity property of the multiplication in $\overline{\mathbb{R}}$. [Solution: (1.71)]

Exercise 1.72 Identify all the cases where the distributivity of the multiplication over the addition

$$a\,(b + c) = a\,b + a\,c \tag{1.44}$$

in $\overline{\mathbb{R}}$ fails. Establish if the distributivity holds in $[0, +\infty]$. [Solution: (1.72)]

1.6.2 Limit superior and limit inferior

Consider an arbitrary sequence $(x_k)_{k\in\mathbb{N}}$ of $\overline{\mathbb{R}}$. We construct a new sequence $(y_k)_{k\in\mathbb{N}}$ as follows:

$$\forall\, k \in \mathbb{N}, \qquad y_k = \sup\{x_h \in \overline{\mathbb{R}} \,|\, h \ge k\}.$$

Hence, for each $k \in \mathbb{N}$, the term y_k is the supremum of all the elements in the sequence from x_k onward, and the sequence $(y_k)_{k\in\mathbb{N}}$ is obtained by taking the supremum of successively truncated tails of the sequence $(x_k)_{k\in\mathbb{N}}$. In

particular,

$$\begin{aligned} y_0 &= \sup\{x_h \in \overline{\mathbb{R}} \,|\, h \geq 0\} = \sup\{x_0, x_1, \ldots\} \\ y_1 &= \sup\{x_h \in \overline{\mathbb{R}} \,|\, h \geq 1\} = \sup\{x_1, x_2, \ldots\} \\ y_2 &= \sup\{x_h \in \overline{\mathbb{R}} \,|\, h \geq 2\} = \sup\{x_2, x_3, \ldots\} \\ &\vdots \end{aligned}$$

In this way, we obtain a non-increasing sequence $(y_k)_{k\in\mathbb{N}}$.[9] The limit superior is defined as the limit $(y_k)_{k\in\mathbb{N}}$:

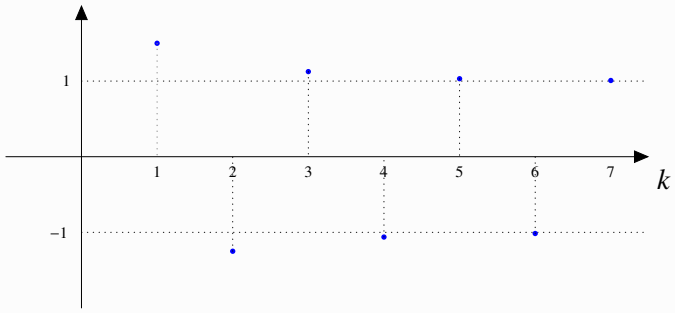
$$\limsup_{k\to\infty} x_k \overset{\text{def}}{=} \lim_{k\to\infty} y_k.$$

Notice that the limit in the right hand-side always exists in $\overline{\mathbb{R}}$ because $(y_k)_{k\in\mathbb{N}}$ is non-increasing, and it is equal to $\inf_{k\in\mathbb{N}} y_k$. Hence, we have also

$$\begin{aligned} \limsup_{k\to\infty} x_k &= \lim_{k\to\infty} \sup_{h\geq k} x_h \\ &= \inf_{k\in\mathbb{N}} \sup_{h\geq k} x_h. \end{aligned}$$

Example *1.46*

Consider the sequence $(x_k)_{k\in\mathbb{N}^*}$ described by the term $x_k = (-1)^{k+1}\left(1 + \frac{1}{2^k}\right)$, which can be represented by the graph



The first terms of the sequence are

$$1.5, \quad -1.25, \quad 1.125, \quad -1.0625, \quad 1.03125, \quad \ldots$$

Then, $(y_k)_{k\in\mathbb{N}}$ is the sequence

$$1.5, \quad 1.125, \quad 1.125, \quad 1.03125, \quad 1.03125, \quad \ldots$$

and the infimum is 1, so that the limit superior of the original sequence is 1. Notice that the limit superior exists even if the sequence does not converge.

Similarly we can define $(z_k)_{k\in\mathbb{N}}$ as follows:

$$\forall\, k \in \mathbb{N}, \qquad z_k = \inf\{x_h \in \overline{\mathbb{R}} \,|\, h \geq k\}.$$

and define the limit inferior of the sequence $(x_k)_{k\in\mathbb{N}}$ as

$$\liminf_{k\to\infty} x_k = \lim_{k\to\infty} z_k,$$

[9] Indeed, if $A \subseteq B$, it is easy to see that $\sup A \leq \sup B$, see Exercise 1.68.

where the limit in the right hand-side exists because $(z_k)_{k\in\mathbb{N}}$ is non-decreasing. This also implies that the limit is equal to $\sup_{k\in\mathbb{N}} z_k$. We can therefore write

$$\begin{aligned}\liminf_{k\to\infty} x_k &= \lim_{k\to\infty} \inf_{h\geq k} x_h \\ &= \sup_{k\in\mathbb{N}} \inf_{h\geq k} x_h.\end{aligned}$$

Example *1.47*

Consider the sequence $(x_k)_{k\in\mathbb{N}^*}$ of the previous example, whose terms are

$$1.5, \quad -1.25, \quad 1.125, \quad -1.0625, \quad 1.03125$$

The sequence $(z_k)_{k\in\mathbb{N}}$ is the sequence

$$-1.25, \quad -1.25, \quad -1.0625, \quad -1.0625, \quad -1.015625, \quad \ldots$$

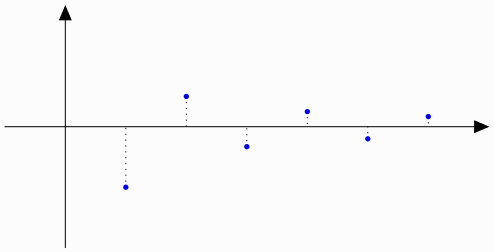
and the supremum is -1, so that the limit inferior of the original sequence is -1.

Example *1.48*

Let $(x_k)_{k\in\mathbb{N}^*}$ be a sequence described by $x_k = \frac{1}{k}(-1)^k$, so that its elements are given by

$$-1, \quad \frac{1}{2}, \quad -\frac{1}{3}, \quad \frac{1}{4}, \quad -\frac{1}{5}, \quad \frac{1}{6}, \quad -\frac{1}{7}, \quad \frac{1}{8}, \quad -\frac{1}{9}, \quad \frac{1}{10}, \quad \ldots$$

$$\begin{aligned} x : \mathbb{N}^* &\longrightarrow \mathbb{R} \\ k &\mapsto \tfrac{1}{k}(-1)^k \end{aligned}$$



In this case $(y_k)_{k\in\mathbb{N}^*}$ is the sequence

$$\frac{1}{2}, \quad \frac{1}{2}, \quad \frac{1}{4}, \quad \frac{1}{4}, \quad \frac{1}{6}, \quad \frac{1}{6}, \quad \ldots$$

whose infimum is 0, so that the limit superior is also equal to 0. Likewise, $(z_k)_{k\in\mathbb{N}^*}$ is the sequence

$$-1, \quad -\frac{1}{3}, \quad -\frac{1}{3}, \quad -\frac{1}{5}, \quad -\frac{1}{5}, \quad -\frac{1}{7}, \quad -\frac{1}{7}, \quad \ldots$$

whose supremum is 0, so that the limit inferior is 0. In this case, the two limits coincide. Moreover, in this case the sequence is convergent, and the limit is equal to the limit superior and to the limit inferior. We will see that this is a general result: the limit superior coincides with the limit inferior if and only if the sequence is convergent. In this case, the three coincide.

Example *1.49*

Let $(x_k)_{k\in\mathbb{N}}$ be a sequence described by $x_k = k(-1)^{k+1}$, so that its elements are given by

$$0, \quad 1, \quad -2, \quad 3, \quad -4, \quad 5, \quad -6, \quad 7, \quad -8, \quad 9, \quad -10, \quad \ldots$$

In this case $(y_k)_{k\in\mathbb{N}}$ is the sequence

$$+\infty, \quad +\infty, \quad +\infty, \quad +\infty, \quad \ldots$$

and therefore the limit superior is $+\infty$. Likewise, $(z_k)_{k\in\mathbb{N}}$ is the sequence

$$-\infty, \quad -\infty, \quad -\infty, \quad -\infty, \quad \dots$$

so that the limit inferior is $-\infty$.

We now examine some important properties of the limit superior and inferior.

Theorem 1.35 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements of $[-\infty,+\infty]$. Then,*

$$\liminf_{k\to\infty} x_k \le \limsup_{k\to\infty} x_k.$$

Proof We have

$$\inf_{h\ge k} x_h \le \sup_{h\ge k} x_h,$$

and taking the limit on both sides preserves the inequality.

Our aim now is to show that if a sequence $(x_k)_{k\in\mathbb{N}}$ converges to $\hat{x}$, then $\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k = \hat{x}$, and, conversely, if $\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k = \tilde{x}$, then the sequence $(x_k)_{k\in\mathbb{N}}$ converges to $\tilde{x}$. To prove this result, we will use the following lemma.

Lemma 1.2 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements of $[-\infty,+\infty]$. Let $\ell_+ = \limsup_{k\to\infty} x_k$ and $\ell_- = \liminf_{k\to\infty} x_k$. Then:*

(i) $\forall\, M > \ell_+, \ \exists\, p\in\mathbb{N} : \ \forall\, k\in\mathbb{N}, \quad k > p \ \Rightarrow \ x_k < M;$

(ii) $\forall\, M < \ell_+, \ \forall\, p\in\mathbb{N}, \ \exists\, k\in\mathbb{N}, \quad k > p \ \Rightarrow \ x_k > M;$

(iii) $\forall\, M < \ell_-, \ \exists\, p\in\mathbb{N} : \ \forall\, k\in\mathbb{N}, \quad k > p \ \Rightarrow \ x_k > M;$

(iv) $\forall\, M > \ell_-, \ \forall\, p\in\mathbb{N} : \ \exists\, k\in\mathbb{N}, \quad k > p \ \Rightarrow \ x_k < M.$

Proof We prove the first. If $\ell_+ = +\infty$, the statement is obvious (vacuously), because there are no elements of $\mathbb{R}$ strictly greater than $+\infty$. Let $\ell_+ \neq +\infty$ and $M > \ell_+$. By definition of limit superior, $\lim_{k\to\infty} \sup_{h\ge k} x_h < M$, so that there exists $p\in\mathbb{N}$ such that for all $k\in\mathbb{N}$ with $k > p$, we have $\sup_{h\ge k} x_h < M$. Indeed, let $(a_k)_{k\in\mathbb{N}}$ be a convergent sequence and let $\hat{a}$ denote its limit. Suppose that $\hat{a} < M$. Since the sequence converges to $\hat{a}$, we can choose ε to satisfy $\hat{a} < \hat{a} + \varepsilon < M$, and there must exist $p\in\mathbb{N}$ such that for all $k > p$ we have $|a_k - \hat{a}| < M$, which gives $a_k - \hat{a} < \varepsilon$, so that $a_k < \hat{a} + \varepsilon < M$. This shows that $\sup_{h\ge k} x_h < M$, and as a consequence for all $k > p$ we have $x_k < M$. This proves the first claim.

Consider the second. If $\ell_+ = -\infty$, the statement is trivial (again, vacuously). Let $\ell_+ \neq -\infty$. Let $M < \ell_+$ and let $p\in\mathbb{N}$ be arbitrary. Using the definition of limit superior,

$$\ell_+ = \inf_{s\in\mathbb{N}} \sup_{h\ge s} x_h > M \quad \Rightarrow \quad \forall\, s\in\mathbb{N}, \ M < \sup_{h\ge s} x_h.$$

Taking in particular $s = p$ gives $M < \sup_{h\in A} x_h$, where $A = \{h\in\mathbb{N}\,|\,h\ge p\}$. We can use **(3)** in Theorem 1.34 to see that there exists $k\in A$ (so that $k\ge p$) such that $M < x_k < \sup_{h\ge p} x_h$.

The third and fourth statements can be proved in a similar way.

Theorem 1.36 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements of $[-\infty,+\infty]$. Then, $(x_k)_{k\in\mathbb{N}}$ is convergent in $[-\infty,+\infty]$ if and only if $\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k$. In this case, we have*

$$\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k = \lim_{k\to\infty} x_k.$$

Proof Let $\liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k$. Let $\ell = \liminf_{k\to\infty} x_k = \limsup_{k\to\infty} x_k$. From *(i)* in Lemma 1.2, for every $M \in [-\infty,+\infty]$ greater than ℓ, there exists $p \in \mathbb{N}$ such that, for all $k > p$, we have $x_k < M$. From *(iii)* in Lemma 1.2, for every $N \in [-\infty,+\infty]$ smaller than ℓ, there exists $q \in \mathbb{N}$ such that, for all $k > q$, we have $x_k > N$. Taking $k > \max\{p,q\}$, we have $N < x_k < M$. From the arbitrariness of N and M, we conclude that $x_k \xrightarrow{k\to\infty} \ell$.

Conversely, suppose that $(x_k)_{k\in\mathbb{N}}$ is convergent in $[-\infty,+\infty]$. Let $\hat{x}$ denote the limit. Let $\ell_+ = \limsup_{k\to\infty} x_k$ and $\ell_- = \liminf_{k\to\infty} x_k$. We prove that $\hat{x} \geq \ell_+$. Suppose by contradiction that $\hat{x} < \ell_+$. Let $b \in [-\infty,+\infty]$ be such that $\hat{x} < b < \ell_+$. Since $\lim_{k\to\infty} x_k = \hat{x}$, there exists $p \in \mathbb{N}$ such that, for all $k \in \mathbb{N}$ with $k > p$, we have $x_k < b$. However, from Lemma 1.2, since $b < \ell_+ = \limsup_{k\to\infty} x_k$, there exists $q \in \mathbb{N}$ such that, for all $k \in \mathbb{N}$ with $k > q$ there holds $x_k > b$. This is a contradiction. In a similar way, we can show that $\hat{x} \leq \ell_-$. Therefore, $\ell_+ \leq \hat{x} \leq \ell_-$. On the other hand, since $\hat{x} \leq \ell_+$, we have also $\ell_- = \ell_+$.

Exercises

Exercise 1.73 Let $A, B \subset \overline{\mathbb{R}}$ be non-empty sets. Show that:

(i) $\inf A \leq \sup A$;

(ii) $A \subseteq B$ implies $\sup A \leq \sup B$ and $\inf A \geq \inf B$.

[Solution: (1.73)]

Exercise 1.74 Let $(a_k)_{k\in\mathbb{N}^*}$ be a sequence described by $a_k = k$. Find its limit superior and inferior. [Solution: (1.74)]

Exercise 1.75 Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\overline{\mathbb{R}}$. Let $b \in \overline{\mathbb{R}}$. Show that

(i) $\limsup_{k\to\infty}(a_k) = -\liminf_{k\to\infty}(-a_k)$;

(ii) $\limsup_{k\to\infty}(b + a_k) = b + \limsup_{k\to\infty}(a_k)$;

(iii) $\liminf_{k\to\infty}(b + a_k) = b + \liminf_{k\to\infty}(a_k)$;

(iv) $\liminf_{k\to\infty} a_k \leq \limsup_{k\to\infty} a_k$;

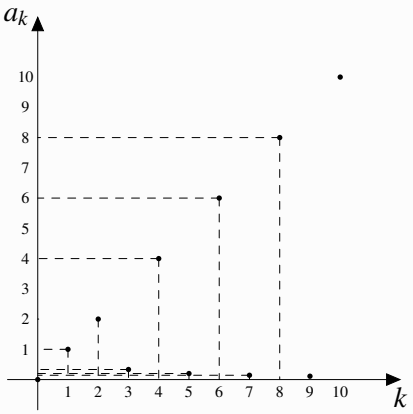
(v) $(a_k)_{k\in\mathbb{N}}$ is convergent if and only if $\liminf_{k\to\infty} a_k = \limsup_{k\to\infty} a_k$.

[Hint: for the first, show that $\inf_{k\in\mathbb{N}}(-a_k) = -\sup_{k\in\mathbb{N}}(a_k)$. Hint: for the second: show that $\sup_{k\in\mathbb{N}}(b + a_k) = b + \sup_{k\in\mathbb{N}} a_k$. To prove *(iv)*, show that given a bounded set A, in general $\inf A \leq \sup A$, and recall that if a_k for all $k \in \mathbb{N}$, then $\lim_{k\to\infty} a_k \geq 0$.]. [Solution: (1.75)]

Exercise 1.76 Let $(a_k)_{k\in\mathbb{N}}$ be the oscillating sequence whose k-th term is $a_k = (-1)^k$. Find the limit superior and the limit inferior. [Solution: (1.76)]

Exercise 1.77 Find the limit superior and the limit inferior of the sequence $(a_k)_{k\in\mathbb{N}^*}$ described by

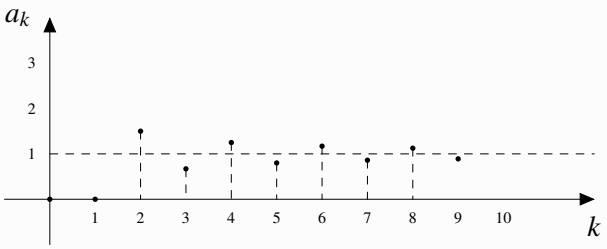
$$\begin{aligned} a : \mathbb{N}^* &\longrightarrow \mathbb{R} \\ k &\mapsto \begin{cases} k & \text{if } k \text{ is even} \\ \frac{1}{k} & \text{if } k \text{ is odd.} \end{cases} \end{aligned}$$



[Solution: (1.77)]

Exercise 1.78 Find the limit superior and the limit inferior of the sequence $(a_k)_{k\in\mathbb{N}^*}$ described by

$$\begin{aligned} a : \mathbb{N}^* &\longrightarrow \mathbb{R} \\ k &\mapsto 1 + (-1)^k \tfrac{1}{k} \end{aligned}$$



[Solution: (1.78)]

Exercise 1.79 Let $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ be bounded sequences in $\mathbb{R}$. Show that

$$\limsup_{k\to\infty}(a_k + b_k) \leq \limsup_{k\to\infty} a_k + \limsup_{k\to\infty} b_k.$$

[Solution: (1.79)]

Exercise 1.80 Let $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ be bounded sequences in $\mathbb{R}$. Show that

$$\limsup_{k\to\infty} a_k = -\liminf_{k\to\infty}(-a_k),$$

and conclude that

$$\limsup_{k\to\infty}(a_k - b_k) \leq \limsup_{k\to\infty} a_k - \liminf_{k\to\infty} b_k.$$

[Solution: (1.80)]

Exercise 1.81 Let $(a_k)_{k\in\mathbb{N}}$ and $(b_k)_{k\in\mathbb{N}}$ be bounded sequences in $\mathbb{R}$. Using the results of the previous two examples, show that if at least one of the two sequences converges, then

$$\limsup_{k\to\infty}(a_k+b_k)=\limsup_{k\to\infty}a_k+\limsup_{k\to\infty}b_k.$$

[Solution: (1.81)]

Exercise 1.82 Consider the sequence $(x_k)_{k\in\mathbb{N}^*}$ with $x_k=3+(-1)^k\dfrac{k}{k+8}$. Find

$$\limsup_{k\to\infty}x_k \qquad \text{and} \qquad \liminf_{k\to\infty}x_k,$$

and use these values to find if the sequence converges in $\mathbb{R}$. [Solution: (1.82)]

Exercise 1.83 Consider a sequence $(x_k)_{k\in\mathbb{N}}$ of real numbers, and a subsequence $(x_{k_j})_{j\in\mathbb{N}}$ extracted from it. Tick the boxes that correspond to statements that are true in general:

- ☐ If $(x_k)_{k\in\mathbb{N}}$ is convergent, then $(x_{k_j})_{j\in\mathbb{N}}$ is convergent and the limits coincide;
- ☐ If $(x_{k_j})_{j\in\mathbb{N}}$ is convergent, then $(x_k)_{k\in\mathbb{N}}$ is convergent and the limits coincide;
- ☐ If $(x_k)_{k\in\mathbb{N}}$ is bounded, then $(x_{k_j})_{j\in\mathbb{N}}$ is bounded;
- ☐ If $(x_{k_j})_{j\in\mathbb{N}}$ is bounded, then $(x_k)_{k\in\mathbb{N}}$ is bounded;
- ☐ If $(x_k)_{k\in\mathbb{N}}$ oscillates, then $(x_{k_j})_{j\in\mathbb{N}}$ does not converge;
- ☐ If $(x_k)_{k\in\mathbb{N}}$ is bounded, then $(x_{k_j})_{j\in\mathbb{N}}$ converges;
- ☐ $(x_k)_{k\in\mathbb{N}}$ always has a limit inferior and limit superior in $\mathbb{R}$;
- ☐ $(x_k)_{k\in\mathbb{N}}$ always has a limit inferior and limit superior in $\overline{\mathbb{R}}$;
- ☐ If $\limsup_{k\to\infty}x_k=\liminf_{k\to\infty}x_k$, then $(x_k)_{k\in\mathbb{N}}$ is convergent in $\overline{\mathbb{R}}$;
- ☐ If $(x_k)_{k\in\mathbb{N}}$ is convergent in $\mathbb{R}$, then $\limsup_{k\to\infty}x_k=\liminf_{k\to\infty}x_k$.

[Solution: (1.83)]

Exercise 1.84 Let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be bounded sequences in $\mathbb{R}$. Let $x_k, y_k \geq 0$ for all $k\in\mathbb{N}$.

(1) Show that

$$\limsup_{k\to\infty}(x_k\,y_k)\leq(\limsup_{k\to\infty}x_k)(\limsup_{k\to\infty}y_k); \tag{1.45}$$

(2) Show, via and example, that it is possible to have the strict inequality in (1.45);

(3) Show that the assumption $x_k, y_k \geq 0$ for all $k\in\mathbb{N}$ is essential for this result to be true.

(4) If one of the two sequences converges and the product is not in the form $0\cdot\infty$ or $\infty\cdot 0$, show that

$$\limsup_{k\to\infty}(x_k\,y_k)=(\limsup_{k\to\infty}x_k)(\limsup_{k\to\infty}y_k).$$

[Solution: (1.84)]

1.6.3 Sequences of functions

Consider a set A, and a vector space $\boldsymbol{Y}$. We can define a sequence $(f_k)_{k\in\mathbb{N}}$ as in Definition 1.25, by considering $X = \boldsymbol{Y}^A$, i.e., where, for all $k \in \mathbb{N}$, each f_k is a function from A to $\boldsymbol{Y}$. This sequence is said to be a sequence of functions. In this section we examine the two fundamental notions of convergence for sequences of functions.

Pointwise convergence

When we say that the sequence of functions $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the function f we mean that, given any point $\boldsymbol{x} \in A$, the sequence $(f_k(x))_{k\in\mathbb{N}}$ of vectors converges to the vector $f(x)$. Pointwise convergence is only one of various senses in which a sequence of functions can converge to a particular function.

Definition 1.27 [Pointwise Convergence]
Let $(\boldsymbol{Y}, \|\cdot\|)$ be a normed space. Let A be a set. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions from A to $\boldsymbol{Y}$. Let $f : A \longrightarrow \boldsymbol{Y}$. We say that $(f_k)_{k\in\mathbb{N}}$ converges pointwise to f if

$$\forall\, x \in A, \quad f_k(x) \overset{k\to\infty}{\longrightarrow} f(x)$$

We can express the same condition of the definition as $\lim_{k\to\infty} f_k(x) = f(x)$ for all $x \in A$, or, by using the definition of convergence, by using the condition

$$\forall\, x \in A,\ \forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N},\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \|f_k(x) - f(x)\| < \varepsilon \tag{1.46}$$

Given a sequence $(f_k)_{k\in\mathbb{N}}$ of functions from A to $\boldsymbol{Y}$ as in Definition 1.27, we say that $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent if there exists $f : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges pointwise to f. From the uniqueness of the limit, if $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent, there is one and only one function $f : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges pointwise to f. Such function is called *limit* (or *pointwise limit*) of $(f_k)_{k\in\mathbb{N}}$, and it is denoted by

$$\lim_{k\to\infty} f_k.$$

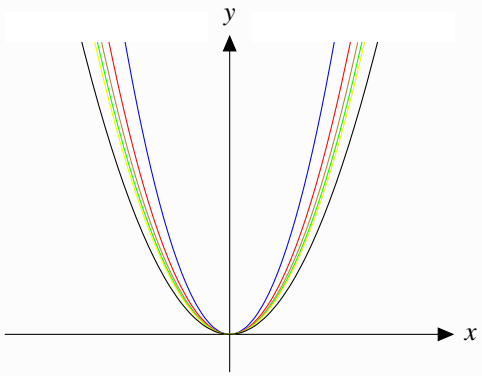
Example 1.50

For all $k \in \mathbb{N}^*$, the graph of each functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto \left(1 + \frac{1}{k}\right) x^2 \end{aligned}$$

is a parabola:

$k = 1$	blue
$k = 2$	red
$k = 3$	brown
$k = 4$	green
$k = 5$	yellow
$k \to \infty$	black



We observe that, for every $x \in \mathbb{R}$, we have a sequence of real numbers $\left(\left(1+\frac{1}{k}\right)x^2\right)_{k\in\mathbb{N}^*}$, which converges to x^2. Thus, this sequence of functions converges pointwise to

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto x^2. \end{aligned}$$

To see this using the condition (1.46), we fix $x \in \mathbb{R}$ and $\varepsilon > 0$. We look for $p \in \mathbb{R}$ such that, if $k > p$, then $|f_k(x) - f(x)| < \varepsilon$. This latter inequality, in the present case, is

$$\left|\left(1+\frac{1}{k}\right)x^2 - x^2\right| < \varepsilon \quad \Leftrightarrow \quad \frac{1}{k}x^2 < \varepsilon \quad \Leftrightarrow \quad k > \frac{x^2}{\varepsilon}.$$

Hence, we can take $p = \frac{x^2}{\varepsilon}$ (or any real number greater or equal to $\frac{x^2}{\varepsilon}$). Notice that p depends on both ε and x: this consideration will become crucial in the sequel.

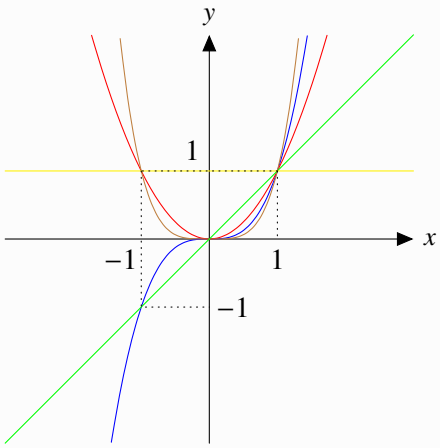
Example *1.51*

For all $k \in \mathbb{N}$, let

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto x^k. \end{aligned}$$

The sequence $(f_k)_{k\in\mathbb{N}}$ is a sequence of functions from $\mathbb{R}$ to $\mathbb{R}$. Given $x > 1$, we have $\lim_{k\to\infty} x^k = +\infty$. Thus, $(f_k(x))_{k\in\mathbb{N}}$ is not convergent in $\mathbb{R}$ (and it is not even convergent in the extended real line $\overline{\mathbb{R}}$ in this case, because for $x \leq -1$ the limit $\lim_{k\to\infty} x^k$ does not exist).[a] It follows that $(f_k)_{k\in\mathbb{N}}$ is not pointwise convergent.

$k = 0$ yellow
$k = 1$ green
$k = 2$ red
$k = 3$ blue
$k = 4$ brown



[a] If $x_0 > 1$, the sequence x_0^k converges in $\overline{\mathbb{R}}$ but not in $\mathbb{R}$, and the limit is $+\infty$. If $x_0 \in (-1, 1]$, the sequence x_0^k converges in $\mathbb{R}$, and the limit is 0; if $x_0 = -1$, then x_0^k does not converge with respect to any topological space. Finally, if $x_0 < -1$, the sequence x_0^k only converges with respect to the Riemann sphere $\mathbb{R} \cup \{\infty\}$, and the limit is ∞, but it does not converge in $\overline{\mathbb{R}}$.

Example *1.52*

For all $k \in \mathbb{N}$, let

$$\begin{aligned} f_k : [0,1] &\longrightarrow \mathbb{R}, \\ x &\mapsto x^k. \end{aligned}$$

The sequence $(f_k)_{k\in\mathbb{N}^*}$ is a sequence of functions from $[0, 1]$ to $\mathbb{R}$. For $x \in [0, 1)$, the sequence $(f_k(x))_{k\in\mathbb{N}} = (x^k)_{k\in\mathbb{N}^*}$ is convergent with respect to $\mathbb{R}$ and $\lim_{k\to\infty} x^k = 0$. For $x = 1$. the sequence $(f_k(x))_{k\in\mathbb{N}} = (1)_{k\in\mathbb{N}}$ is

convergent in $\mathbb{R}$ and $\lim_{k\to\infty} 1^k = 1$. It follows that $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent and

$$f = \lim_{k\to\infty} f_k : [0,1] \longrightarrow \mathbb{R},$$
$$x \mapsto \begin{cases} 0 & x \in [0,1) \\ 1 & x = 1. \end{cases}$$

Interestingly, even if all the functions f_k of the sequence $(f_k(x))_{k\in\mathbb{N}}$ are continuous, the pointwise limit is not a continuous function.

Example *1.53*

Consider a differentiable function $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$. For all $k \in \mathbb{N}$, let

$$\varphi_k : \mathbb{R} \longrightarrow \mathbb{R},$$
$$x \mapsto \frac{\varphi\left(x + \frac{1}{k}\right) - \varphi(x)}{\frac{1}{k}}$$

The sequence $(\varphi_k)_{k\in\mathbb{N}^*}$ is pointwise convergent to φ' since for every $x \in \mathbb{R}$ we have $\lim_{k\to\infty} \varphi_k(x) = \varphi'(x)$.

Example *1.54*

For all $k \in \mathbb{N}$, let

$$f_k : \mathbb{R} \longrightarrow \mathbb{R},$$
$$x \mapsto \arctan(x - k).$$

The sequence $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent to $-\pi/2$ in $\mathbb{R}$. Indeed, if x is fixed and k goes to infinity, $x - k$ goes to $-\infty$ and therefore the inverse tangent tends to $-\pi/2$.

The pointwise converges preserves, at the limit, some of the properties that the functions of the sequence enjoy. For example, given a set $A \subseteq \mathbb{R}$ and a sequence of non-decreasing functions $(f_k)_{k\in\mathbb{N}}$ from A to $\mathbb{R}$ which is pointwise convergent to a certain function $f : A \longrightarrow \mathbb{R}$, given $x, y \in A$ with $x < y$, by Theorem 1.25

$$\forall\, k \in \mathbb{N}, \quad f_k(x) \le f_k(y) \quad \Rightarrow \quad \lim_{k\to\infty} f_k(x) \le \lim_{k\to\infty} f_k(y),$$

which is the same as $f(x) \le f(y)$, so that f is non-decreasing as well. The same argument can be used if $(f_k)_{k\in\mathbb{N}}$ is non-increasing, to show that the pointwise limit is a non-increasing function.

Exercise 1.85

Let $A \subseteq \mathbb{R}$. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of non-negative functions $f_k : A \longrightarrow \mathbb{R}$, so that for all $k \in \mathbb{N}$ and all $x \in A$, there holds $f(x) \ge 0$. Let $(f_k)_{k\in\mathbb{N}}$ be pointwise convergent to a function $f : A \longrightarrow \mathbb{R}$. Show that f is also non-negative, i.e., $f(x) \ge 0$ for all $x \in A$. [Solution: (1.85)]

Exercise 1.86

Let $I \subseteq \mathbb{R}$ be an interval containing at least two points. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of convex functions $f_k : I \longrightarrow \mathbb{R}$, so that for all $k \in \mathbb{N}$, for all $\lambda \in [0,1]$ and all $x, y \in I$, we have $f_k(\lambda x + (1-\lambda) y) \le \lambda f_k(x) + (1-\lambda) f_k(y)$. Let $(f_k)_{k\in\mathbb{N}}$ be pointwise convergent to a function $f : I \longrightarrow \mathbb{R}$. Show that f is also convex, i.e., $f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y)$ for all $x, y \in I$ and all $\lambda \in [0,1]$. [Solution: (1.86)]

The properties considered so far involve one, two, three points each time, so that to conclude the argument it is sufficient a piece of information that involves the points one by one. But this begs the following question: does the pointwise convergence preserve properties whose formulation requires considering an infinite number of points each time, like the boundedness, the continuity, the differentiability and the integrability? The following example will help us investigate this issue.

Example *1.55*

For all $k \in \mathbb{N}^*$, let

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto \frac{x}{x^4 + \frac{1}{k}}. \end{aligned}$$

The sequence $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent to the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto \begin{cases} \frac{1}{x^3} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases} \end{aligned}$$

All the functions f_k are bounded, continuous, differentiable and Riemann integrable on $\mathbb{R}$ (to see that it is bounded it suffices to observe that it has finite global maximum and global minimum because as $x \to \pm\infty$ the function vanishes), but the pointwise limit F of the sequence $(f_k)_{k\in\mathbb{N}^*}$ does not possess any of these properties.

Consider again the example of the integration. The pointwise limit of f of a sequence $(f_k)_{k\in\mathbb{N}}$ of integrable functions in the interval $[a, b]$ may not be itself integrable (not even as an improper integral). However, it can happen that f is integrable, and in this case we would like to know whether the operation of limit as $k \to \infty$ and the operation of integration can be interchanged, i.e., we ask ourselves if it is legitimate to write

$$\lim_{k\to\infty} \int_a^b f_k(x)\,\mathrm{d}x = \int_a^b \left(\lim_{k\to\infty} f_k(x)\right) \mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

For example, given a sequence of numbers $(c_k)_{k\in\mathbb{N}}$, the sequence of functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} c_k & \text{if } 0 < x < \frac{1}{k} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is pointwise convergent to the identically zero function of $\mathbb{R}$. Hence, the integral of f over any interval is 0, regardless of the sequence $(c_k)_{k\in\mathbb{N}}$, whereas

$$\int_0^1 f_k(x)\,\mathrm{d}x = \int_0^{1/k} c_k\,\mathrm{d}x = \frac{c_k}{k}.$$

For different choices of the sequence $(c_k)_{k\in\mathbb{N}}$, the limit of the integrals can:

- diverge (for example when $c_k = k^2$ for all $k \in \mathbb{N}$),
- oscillate (for example when $c_k = (-1)^k k$ for all $k \in \mathbb{N}$),
- converge to 0 (for example when $c_k = 1$ for all $k \in \mathbb{N}$), or
- converge to a different value (for example when $c_k = k$ for all $k \in \mathbb{N}$).

It seems therefore important to introduce a different (and stronger) type of convergence, that plays an important role.

Uniform convergence

With this type of convergence, we will be able to answer some of the issues raised in the previous section.

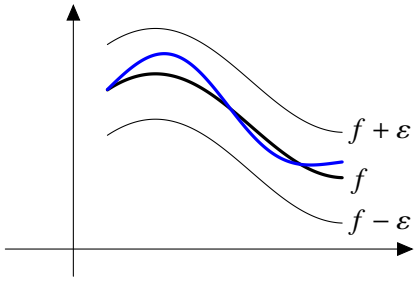
Definition 1.28 [Uniform Convergence to a Function]
Let $(\boldsymbol{Y}, \|\cdot\|)$ be a normed space. Let A be a non-empty set. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions from A to

Y. *Let* $f : A \longrightarrow \boldsymbol{Y}$. *We say that* $(f_k)_{k\in\mathbb{N}}$ converges uniformly to f *if*

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \forall\, x \in A, \qquad \|f_k(x) - f(x)\| < \varepsilon \tag{1.47}$$

If $(f_k)_{k\in\mathbb{N}}$ is a sequence of functions from A to $\boldsymbol{Y}$ as in Definition 1.28, we say that $(f_k)_{k\in\mathbb{N}}$ *is uniformly convergent* if there exists $f : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly.

From a geometric point of view, if $A \subseteq \mathbb{R}$ and $\boldsymbol{Y} = \mathbb{R}$, the sequence $(f_k)_{k\in\mathbb{N}}$ converges uniformly to f when, once we assign a strip centred on the graphic of f within the graphics of the functions $f - \varepsilon$ and $f + \varepsilon$, for a sufficiently large k the graphic of f_k is contained in this strip.



Comparing the condition for pointwise convergence

$$\forall\, x \in A,\ \forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N},\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \|f_k(x) - f(x)\| < \varepsilon$$

with that of uniform convergence

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \forall\, x \in A, \qquad \|f_k(x) - f(x)\| < \varepsilon$$

we see that the difference is in the position of the quantifiers. In particular, in the definition of pointwise convergence, the number p is assigned after ε and x have been chosen. Thus, in general, p depends on ε and x. In the uniform convergence, p is assigned after only ε has been fixed. Thus, in the uniform convergence p depends only on ε. This consideration suggests that uniform convergence is a stronger property than pointwise convergence: if a sequence of functions $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly, it is also pointwise convergent to f. The following theorem makes this claim more precise.

Theorem 1.37 [Pointwise and Uniform Convergence]
Let $(\boldsymbol{Y}, \|\cdot\|)$ *be a normed space. Let* A *be a non-empty set. Let* $(f_k)_{k\in\mathbb{N}}$ *be a sequence of functions from* A *to* $\boldsymbol{Y}$. *Let* $f : A \longrightarrow \boldsymbol{Y}$. *Then, if* $(f_k)_{k\in\mathbb{N}}$ *converges to* f *uniformly, it is pointwise convergent to* f.

Proof Suppose that $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly. Let $\varepsilon > 0$ and $x \in A$. There exists $p \in \mathbb{N}$ such that for all $k \in \mathbb{N}$ with $k > p$, for all $\xi \in A$ there holds $\|f_k(\xi) - f(\xi)\| < \varepsilon$. Thus, taking in particular $\xi = x$, if $k \in \mathbb{N}$ and $k > p$, we obtain $\|f_k(x) - f(x)\| < \varepsilon$.

Theorem 1.38 *Let* $(\boldsymbol{Y}, \|\cdot\|)$ *be a normed space. Let* A *be a non-empty set. Let* $(f_k)_{k\in\mathbb{N}}$ *be a sequence of functions from* A *to* $\boldsymbol{Y}$. *Let* $(f_k)_{k\in\mathbb{N}}$ *be uniformly convergent. Then:*

- $(f_k)_{k\in\mathbb{N}}$ *is pointwise convergent;*

- *there exists one and only one $f : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly;*
- *the unique function $f : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly is equal to the pointwise limit $f = \lim_{k\to\infty} f_k$.*

Proof The first statement follow immediately from Theorem 1.37. The second follows from the uniqueness of the limit: if there exist $f, g : A \longrightarrow \boldsymbol{Y}$ such that $(f_k)_{k\in\mathbb{N}}$ converges to f and g uniformly, then, from Theorem 1.37, $(f_k)_{k\in\mathbb{N}}$ converges to f and g pointwise. Let $x \in A$. From $\lim_{k\to\infty} f_k(x) = f(x) = g(x)$ (because the limit is unique), we find that f and g are equal at every point, so that $f = g$. The last statement is a consequence of the first two: suppose that $(f_k)_{k\in\mathbb{N}}$ converges to g uniformly, and the pointwise limit of $(f_k)_{k\in\mathbb{N}}$ is f, but that $f \neq g$. However, g is also the pointwise limit of $(f_k)_{k\in\mathbb{N}}$. As in the previous point, the uniqueness of the limit implies that $f = g$.

Example *1.56*

Consider the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$, where for all $k \in \mathbb{N}^*$

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R}, \\ x &\mapsto \begin{cases} 1 & \text{if } 0 < x \leq 1/k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

To see that the sequence is pointwise convergent to the zero function, we show that

$$\forall\, x \in \mathbb{R},\ \forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N},\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ |f_k(x)| < \varepsilon.$$

We assign $x \in \mathbb{R}$ and $\varepsilon > 0$. If $x \leq 0$ or $\varepsilon \geq 1$, we can take an arbitrary $p \in \mathbb{N}$: indeed, if $x \leq 0$ we have $f_k(x) = 0$, so that $|f_k(x)| < \varepsilon$ is always satisfied. If $\varepsilon \geq 1$, we always have $|f_k(x)| \leq 1$ from the definition of f_k. If $x > 0$ and $\varepsilon < 1$, we must take $p \geq 1/x$, so that for all $k > p \geq 1/x$ we have $x > 1/k$, which guarantees that $f_k(x) = 0$.
We proved that $(f_k)_{k\in\mathbb{N}^*}$ converges to the zero function pointwise. If $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent, the uniform limit must be equal to the uniform limit. In other words, to establish if $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent it suffices to test the condition

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \in \mathbb{N}^*, \qquad k > p \ \Rightarrow\ \forall\, x \in \mathbb{R}, \quad |f_k(x)| < \varepsilon.$$

Suppose that $\varepsilon \in (0, 1)$, $\forall k \in \mathbb{N}^*$ the condition $|f_k(x)| < \varepsilon$ is not satisfied if $x \in \left(0, \frac{1}{k}\right)$ because for such value of k we have $|f_k(x)| = 1$.

The following result provides a convenient instrument to establish if a sequence of functions $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to a function f.

Theorem 1.39 *Let $(\boldsymbol{Y}, \|\cdot\|)$ be a normed space. Let A be a non-empty set. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions from A to $\boldsymbol{Y}$. Let $f : A \longrightarrow \boldsymbol{Y}$. The sequence of functions $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly if and only if the sequence $\left(\sup_{x\in A} \|f_k(x) - f(x)\|\right)_{k\in\mathbb{N}}$ tends to 0 as $k \to \infty$.*

Proof From the definition of uniform convergence, $(f_k(x))_{k\in\mathbb{N}}$ converges uniformly to f if and only if

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \in \mathbb{N}, \qquad k \geq p \ \Rightarrow\ \forall\, x \in A, \quad \|f_k(x) - f(x)\| < \varepsilon.$$

If $\varepsilon > 0$ and $k \in \mathbb{N}$, the relation

$$\forall\, x \in A, \qquad \|f_k(x) - f(x)\| < \varepsilon \tag{1.48}$$

is equivalent to

$$\sup_{x\in A}\|f_k(x)-f(x)\|\leq\varepsilon. \tag{1.49}$$

Indeed, (1.49) immediately implies (1.48), since for every $x\in A$ we have $\|f_k(x)-f(x)\|\leq\sup_{x\in A}\|f_k(x)-f(x)\|$. We show that (1.48) implies (1.49). Now, (1.48) says that ε is an upper bound for $\{\|f_k(x)-f(x)\|\,\big|\,x\in A\}$, and $\sup_{x\in A}\|f_k(x)-f(x)\|$ is the least upper bound of the same set. It follows that $(f_k(x))_{k\in\mathbb{N}}$ converges to f uniformly if and only if

$$\forall\,\varepsilon>0,\ \exists\,p\in\mathbb{N}:\ \forall\,k\in\mathbb{N},\qquad k>p\ \Rightarrow\ \sup_{x\in A}\|f_k(x)-f(x)\|\leq\varepsilon.$$

This implies that $\sup_{x\in A}\|f_k(x)-f(x)\|\overset{k\to\infty}{\longrightarrow}0$.

We observe that the sequence $\left(\sup_{x\in A}\|f_k(x)-f(x)\|\right)_{k\in\mathbb{N}}$ in Theorem 1.39 is a sequence of elements in the extended real line $\overline{\mathbb{R}}=[-\infty,+\infty]$.

Remark 1.12 The convergence with respect to the norm $\|\cdot\|_\infty$ in the set of bounded functions $B(A;\mathbb{R})$ from A (which is a non-empty set) to $\mathbb{R}$ is the uniform convergence. In fact, for $f\in B(A;\mathbb{R})$, we recall that we defined

$$\|f\|_\infty=\sup_{x\in A}|f(x)|.$$

Therefore, convergence of a sequence $(f_k)_{k\in\mathbb{N}}$ of $B(A;\mathbb{R})$ to $f\in B(A;\mathbb{R})$ with respect to $\|\cdot\|_\infty$ can be expressed via the condition

$$\forall\,\varepsilon>0,\ \exists\,p\in\mathbb{N}:\ \forall\,k\in\mathbb{N},\qquad k>p\ \Rightarrow\ \|f_k-f\|_\infty<\varepsilon,$$

i.e.,

$$\forall\,\varepsilon>0,\ \exists\,p\in\mathbb{N}:\ \forall\,k\in\mathbb{N},\qquad k>p\ \Rightarrow\ \sup_{x\in A}|f_k(x)-f(x)|<\varepsilon,$$

which is equivalent to saying that the sequence $(\sup_{x\in A}|f_k(x)-f(x)|)_{k\in\mathbb{N}}$ tends to 0 as $k\to\infty$. This is a necessary and sufficient condition for the sequence $(f_k)_{k\in\mathbb{N}}$ to converge to f uniformly.

Example *1.57*

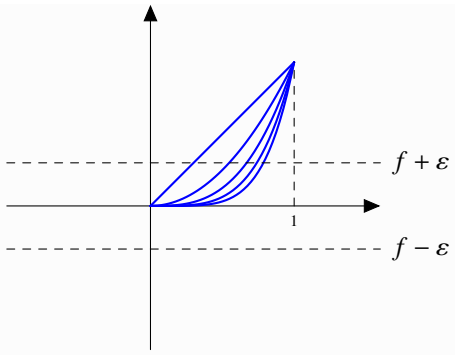
For every $k\in\mathbb{N}^*$, let

$$\begin{aligned}f_k:[0,1]&\longrightarrow\mathbb{R}\\ x&\mapsto x^k.\end{aligned}$$

The sequence $(f_k)_{k\in\mathbb{N}^*}$ is a sequence of functions from $[0,1]$ to $\mathbb{R}$. We have seen that $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent and that the pointwise limit is the function

$$\begin{aligned}f:[0,1]&\longrightarrow\mathbb{R}\\ x&\mapsto\begin{cases}0 & x\in[0,1)\\ 1 & x=1.\end{cases}\end{aligned}$$

We want to prove that $(f_k)_{k\in\mathbb{N}^*}$ is not uniformly convergent. To this end, it suffices to show that $(f_k)_{k\in\mathbb{N}^*}$ does not uniformly converge to f. This claim can be intuitively understood by some geometric considerations: we choose $\varepsilon<1$ and we consider the strip between the graphics of $f-\varepsilon$ and $f+\varepsilon$. If $(f_k)_{k\in\mathbb{N}^*}$ uniformly converged to f, for a sufficiently large k the graphic of f_k would be contained in this strip. This, however, is not the case: as k increases, the graphic of f_k for $x\in[0,1)$ gets increasingly closer to the x-axis, but since it must reach the point $(1,1)$, at some point it will escape from that strip.



$f(x) = x^k, \quad k \in \{1, 2, 3, 4, 5\}$

We prove the same result in a more rigorous way, by using the result of Theorem 1.39. For all $k \in \mathbb{N}^*$ we have

$$\begin{aligned} f_k - f : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} x^k & x \in [0,1) \\ 0 & x = 1. \end{cases} \end{aligned}$$

Therefore,

$$\sup_{x \in [0,1]} |f_k(x) - f(x)| = \sup_{x \in [0,1)} x^k = 1.$$

In fact, for every $x \in [0,1)$, we have $x^k < 1$. Thus, 1 is an upper bound for $\{x^k \mid x \in [0,1)\}$. To show that it is the least upper bound, we can reason by contradiction. Suppose that U is an upper bound for $\{x^k \mid x \in [0,1)\}$ and that it is strictly smaller than 1, i.e., $U < 1$. However, U cannot be an upper bound, because it is possible to find elements in $\{x^k \mid x \in [0,1)\}$ which are strictly greater than U: it suffices to take $x > \sqrt[k]{U} \in [0,1)$, and this gives $x^k > U$, with $x \in [0,1)$.

Obviously the constant sequence $(1)_{k \in \mathbb{N}^*}$ is convergent, and $\lim_{k \to \infty} 1 = 1$. Therefore, $(f_k)_{k \in \mathbb{N}^*}$ does not uniformly converge to f.

Example *1.58*

For every $k \in \mathbb{N}^*$, let

$$\begin{aligned} f_k : [0,1) &\longrightarrow \mathbb{R} \\ x &\mapsto x^k. \end{aligned}$$

The sequence $(f_k)_{k \in \mathbb{N}^*}$ is a sequence of functions from $[0,1)$ to $\mathbb{R}$. It is straightforward to see that $(f_k)_{k \in \mathbb{N}^*}$ is pointwise convergent and that $\lim_{k \to \infty} f_k = 0$, where 0 is the function

$$\begin{aligned} [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto 0. \end{aligned}$$

Using the same notation of the previous example, $f_k - f : [0,1) \longrightarrow \mathbb{R}$ is defined by $f_k(x) - f(x) = x^k$ for $x \in [0,1)$, and $\sup_{x \in [0,1)} x^k = 1$. Thus, $(f_k)_{k \in \mathbb{N}^*}$ does not converge uniformly.

Example *1.59*

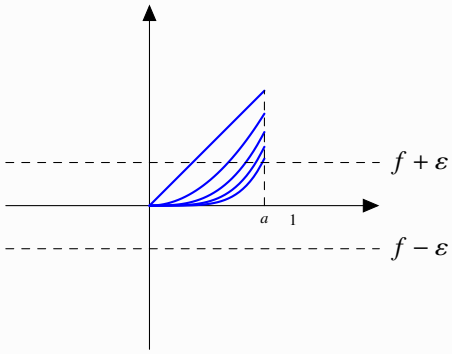
Let $a \in [0,1)$. For every $k \in \mathbb{N}^*$, let

$$\begin{aligned} f_k : [0,a] &\longrightarrow \mathbb{R} \\ x &\mapsto x^k. \end{aligned}$$

It is straightforward to see that $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent and that $\lim_{k\to\infty} f_k = 0$, where 0 is the function

$$\begin{aligned}[0,a] &\longrightarrow \mathbb{R}\\ x &\mapsto 0.\end{aligned}$$

We now show that $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent to 0. This claim is geometrically intuitive: once we choose $\varepsilon > 0$, considering the strip between the graphics of constant functions $-\varepsilon$ and $+\varepsilon$, for sufficiently large k the functions f_k are contained in this strip.



$$f(x) = x^k, \quad k \in \{1,2,3,4,5\}$$

We prove this result in a more rigorous way by using again Theorem 1.39. For all $k \in \mathbb{N}^*$ we have

$$\sup_{x\in[0,a]} |f_k(x)| = \sup_{x\in[0,a]} x^k = a^k.$$

The sequence $(a^k)_{k\in\mathbb{N}^*}$ is convergent and $\lim_{k\to\infty} a^k = 0$. Thus, $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent to 0.

The point a can be chosen arbitrarily close to 1. In loose terms, this means that by removing an arbitrarily small segment it can happen that a sequence which is not uniformly convergent becomes uniformly convergent.

Example *1.60*

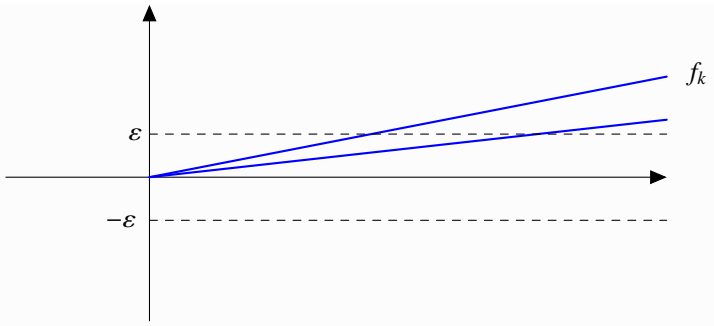
For every $k \in \mathbb{N}^*$, let

$$\begin{aligned}f_k : \mathbb{R}_+ &\longrightarrow \mathbb{R}\\ x &\mapsto \frac{x}{k}.\end{aligned}$$

The sequence $(f_k)_{k\in\mathbb{N}^*}$ is a sequence of functions from $\mathbb{R}$ to $\mathbb{R}$. It is straightforward to see that $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent and that $\lim_{k\to\infty} f_k$ is the function

$$\begin{aligned}\mathbb{R}_+ &\longrightarrow \mathbb{R}\\ x &\mapsto 0.\end{aligned}$$

We now show that $(f_k)_{k\in\mathbb{N}^*}$ is not uniformly convergent. To this end, it suffices to show that $(f_k)_{k\in\mathbb{N}^*}$ is not uniformly convergent to the zero function. The claim is intuitive geometrically. Indeed, if we choose $\varepsilon > 0$ and we consider the strip between the graphics of the constant functions $-\varepsilon$ and $+\varepsilon$, no matter how large we choose k since the graphic of f_k is a half-line from the origin $(0,0)$, at some point the graphic will escape the strip.



We prove this result in a more formal way by using Theorem 1.39. For all $k \in \mathbb{N}^*$ we have

$$\sup_{x \in \mathbb{R}_+} |f_k(x)| = \sup_{x \in \mathbb{R}_+} \frac{x}{k} = +\infty.$$

The sequence $(+\infty)_{k \in \mathbb{N}^*}$ does not go to 0 as k tends to ∞, and so $(f_k)_{k \in \mathbb{N}^*}$ does not converge uniformly to the zero function.

Example *1.61*

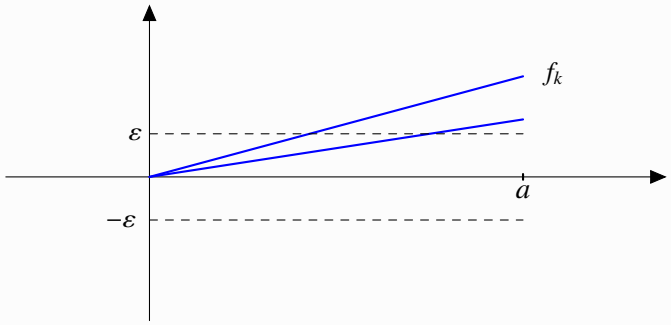
Let $a \in \mathbb{R}_+^*$. For every $k \in \mathbb{N}^*$, let

$$\begin{aligned} f_k : [0, a] &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{x}{k}. \end{aligned}$$

It is straightforward to see that $(f_k)_{k \in \mathbb{N}^*}$ is pointwise convergent and that $\lim_{k \to \infty} f_k$ is again the function

$$\begin{aligned} [0, a] &\longrightarrow \mathbb{R} \\ x &\mapsto 0. \end{aligned}$$

We now show that $(f_k)_{k \in \mathbb{N}^*}$ is uniformly convergent. To this end, it suffices to show that $(f_k)_{k \in \mathbb{N}^*}$ is uniformly convergent to the zero function. The claim is intuitive geometrically. Indeed, if we choose $\varepsilon > 0$ and we consider the strip between the graphics of the constant functions $-\varepsilon$ and $+\varepsilon$, when k is sufficiently large the graphic will be contained in the strip. It is easy to see that it is sufficient to choose $\frac{1}{k} < \frac{\varepsilon}{a}$.



We obtain the same result using Theorem 1.39. For all $k \in \mathbb{N}^*$ we have

$$\sup_{x \in [0,a]} |f_k(x)| = \sup_{x \in [0,a]} \frac{x}{k} = \frac{a}{k}.$$

The sequence $(\frac{a}{k})_{k \in \mathbb{N}^*}$ is convergent, and $\lim_{k \to \infty} \frac{a}{k} = 0$. Thus, $(f_k)_{k \in \mathbb{N}^*}$ converges uniformly to 0.

As we have seen, properties like boundedness, continuity, integrability and differentiability of a sequence are not preserved under the operation of pointwise limit. The situation is different if, instead, we consider the uniform

limit. For example, the uniform limit of a sequence of bounded functions is a bounded function (see Exercise 1.87). We will take into account some of the other notions (and in particular continuity and Riemann integrability) in the sequel.

Exercises

Exercise 1.87 Let A be a subset of $\mathbb{R}$. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of bounded functions $f_k : A \longrightarrow \mathbb{R}$. Let $(f_k)_{k\in\mathbb{N}}$ converge uniformly to the function $f : A \longrightarrow \mathbb{R}$. Show that f is bounded. [Solution: (1.87)]

Exercise 1.88 Let $c > 0$. Establish if the sequence of functions $(f_k)_{k\in\mathbf{N}}$ described by

$$\begin{aligned} f_k : [c, +\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\sin(k\,x)}{1 + k\,x} \end{aligned}$$

converges pointwise/uniformly. [Solution: (1.88)]

Exercise 1.89 Consider, for all $k \in \mathbb{N}^*$, the functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} k & \text{if } 0 < x < \frac{1}{k} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists. Does $(f_k)_{k\in\mathbb{N}}$ converge uniformly to its pointwise limit? [Solution: (1.89)]

Exercise 1.90 Consider, for all $k \in \mathbb{N}^*$, the functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \frac{1}{k} & \text{if } 0 < x < k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists. Does $(f_k)_{k\in\mathbb{N}}$ converge uniformly to its pointwise limit? [Solution: (1.90)]

Exercise 1.91 Consider the sequence of functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{k\,(1 + x^2)}. \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists, and determine whether $(f_k)_{k\in\mathbb{N}}$ converges uniformly or not to that limit. Repeat the same exercise for the sequence

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{x\,(x + k)}{k}. \end{aligned}$$

Discuss what changes in this second case if we restrict this sequence to the interval $[-a, a]$. [Solution: (1.91)]

Exercise 1.92

Consider the sequence of functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto k\left(x + \frac{1}{k}\right)^2 - k\,x^2. \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists, and determine whether $(f_k)_{k\in\mathbb{N}}$ converges uniformly or not to that limit. [Solution: (1.92)]

Exercise 1.93

Consider the sequence of functions

$$\begin{aligned} f_k : \mathbb{R}_+ &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{k\,x^2}{k + x}. \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists, and determine whether $(f_k)_{k\in\mathbb{N}}$ converges uniformly or not to that limit. Does anything change by reducing the domain to $[0, a]$? [Solution: (1.93)]

Exercise 1.94

Consider the sequence of functions

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto |\cos x|^k. \end{aligned}$$

Find the pointwise limit of $(f_k)_{k\in\mathbb{N}}$, if it exists, and determine whether $(f_k)_{k\in\mathbb{N}}$ converges uniformly or not to that limit. [Solution: (1.94)]

Exercise 1.95

Consider the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ described as follows:

$$\begin{aligned} f_k : [0,3] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{if } 0 \le x \le 1/k \\ k^2\,x - k & \text{if } 1/k < x \le 2/k \\ 3\,k - k^2\,x & \text{if } 2/k < x \le 3/k \\ 0 & \text{if } 3/k < x \le 3. \end{cases} \end{aligned}$$

(i) Show that for all $k \in \mathbb{N}^*$, the function f_k is continuous;

(ii) Establish if $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent, and, if it is, determine the pointwise limit f;

(iii) For every $k \in \mathbb{N}^*$, find

$$\int_0^3 f_k(x)\,\mathrm{d}x;$$

(iv) From the result of the previous point, what can be inferred about the uniform convergence of $(f_k)_{k\in\mathbb{N}^*}$?

[Solution: (1.95)]

Exercise 1.96

Evaluate the convergence of the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ from $[0, 1]$ to $\mathbb{R}$ defined by

$$f_k(t) = e^{\frac{x}{k}}$$

in the normed space $C([0, 1]; \mathbb{R})$ with the norm $\|\cdot\|_\infty$. [Solution: (1.96)]

Exercise 1.97

Consider the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ from $\mathbb{R}$ to $\mathbb{R}$ defined for all $t \in \mathbb{R}$ by

$$f_k(t) = \begin{cases} \frac{1}{k} & \text{if } |t| \leq k \\ \frac{(k+1)-|t|}{k} & \text{if } k \leq |t| \leq k+1 \\ 0 & \text{if } |t| \geq k+1 \end{cases}$$

in the normed space $(C(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$.

- Establish if $(f_k)_{k\in\mathbb{N}^*}$ is pointwise convergent to a function in $(C(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$;
- Establish if $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent to a function in $(C(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$.

[Solution: (1.97)]

1.7 Subsequences and Bolzano-Weierstrass theorem

The concept of subsequence corresponds to the following intuitive idea. Consider an infinite number of objects in a list:

$$\boldsymbol{x}_0 \quad \boldsymbol{x}_1 \quad \boldsymbol{x}_2 \quad \boldsymbol{x}_3 \quad \boldsymbol{x}_4 \quad \boldsymbol{x}_5 \quad \boldsymbol{x}_6 \quad \boldsymbol{x}_7 \quad \boldsymbol{x}_8 \quad \boldsymbol{x}_9 \quad \boldsymbol{x}_{10} \quad \boldsymbol{x}_{11} \quad \boldsymbol{x}_{12} \quad \boldsymbol{x}_{13} \quad \dots \tag{1.50}$$

Suppose we want to restrict (1.50) to a subset N' of $\mathbb{N}$, which still contains an infinite number of elements: for example, we could choose N' to be the set of prime numbers:

$$\begin{array}{cccccccccccccc} \boldsymbol{x}_0 & \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 & \boldsymbol{x}_4 & \boldsymbol{x}_5 & \boldsymbol{x}_6 & \boldsymbol{x}_7 & \boldsymbol{x}_8 & \boldsymbol{x}_9 & \boldsymbol{x}_{10} & \boldsymbol{x}_{11} & \boldsymbol{x}_{12} & \boldsymbol{x}_{13} & \dots \\ & & \uparrow & \uparrow & & \uparrow & & \uparrow & & & & \uparrow & & \uparrow & \dots \\ & & \boldsymbol{x}_{k_0} & \boldsymbol{x}_{k_1} & & \boldsymbol{x}_{k_2} & & \boldsymbol{x}_{k_3} & & & & \boldsymbol{x}_{k_4} & & \boldsymbol{x}_{k_5} & \dots \end{array}$$

We can therefore construct a new sequence by defining its terms by

$$\boldsymbol{y}_0 = \boldsymbol{x}_{k_0} = \boldsymbol{x}_2, \quad \boldsymbol{y}_1 = \boldsymbol{x}_{k_1} = \boldsymbol{x}_3, \quad \boldsymbol{y}_2 = \boldsymbol{x}_{k_2} = \boldsymbol{x}_5, \quad \boldsymbol{y}_3 = \boldsymbol{x}_{k_3} = \boldsymbol{x}_7, \quad \boldsymbol{y}_4 = \boldsymbol{x}_{k_4} = \boldsymbol{x}_{11}, \quad \boldsymbol{y}_5 = \boldsymbol{x}_{k_5} = \boldsymbol{x}_{13}, \quad \dots$$

We observe that, with this notation, k has become a function $k : \mathbb{N} \longrightarrow \mathbb{N}$. More precisely, it is a sequence, and this is why we are denoting $k(i)$ as k_i for $i \in \mathbb{N}$. The sequence $(k_i)_{i\in\mathbb{N}}$ satisfies, in this particular case,

$$k_0 = 2, \quad k_1 = 3, \quad k_2 = 5, \quad k_3 = 7, \quad k_4 = 11, \quad k_5 = 13, \quad \dots$$

What we immediately notice is that this sequence $(k_i)_{i\in\mathbb{N}}$ is strictly increasing. In this way, we have constructed a new sequence $(\boldsymbol{y}_i)_{i\in\mathbb{N}}$ where, for every $i \in \mathbb{N}$, we have $\boldsymbol{y}_i = \boldsymbol{x}_{k_i}$. This new sequence is called *subsequence* of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$, and it is the result of the composition of the function $\boldsymbol{x}$ with the function k. Indeed, $\boldsymbol{y}_i = \boldsymbol{x}_{k_i}$ for $i \in \mathbb{N}$ is the same as writing $\boldsymbol{y}_i = (\boldsymbol{x} \circ k)(i)$ for $i \in \mathbb{N}$. Hence, a subsequence is simply the composition of a sequence with a strictly increasing sequence.

Definition 1.29 [Subsequence]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let A be a subset of $\mathbb{N}$ obtained by removing from $\mathbb{N}$ a finite set. Let $(\boldsymbol{x}_k)_{k\in A}$ and $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ be two sequences with values in $\boldsymbol{X}$. We say that $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ is a subsequence of $(\boldsymbol{x}_k)_{k\in A}$ if there exists a strictly increasing function $k : \mathbb{N} \longrightarrow A$ such that

$$\boldsymbol{y} = \boldsymbol{x} \circ k.$$

In this case, we also say that $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ is extracted *from the sequence $(\boldsymbol{x}_k)_{k\in A}$.*

Example *1.62*

Three subsequences of the sequence $(k^2)_{k\in\mathbb{N}}$ are:

- $(4\,k^2)_{k\in\mathbb{N}}$. In this case

$$\begin{array}{cccccccccc} x_0=0 & x_1=1 & x_2=4 & x_3=9 & x_4=16 & x_5=25 & x_6=36 & x_7=49 & x_8=64 & \dots \\ \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & \dots \\ x_{k_0} & & x_{k_1} & & x_{k_2} & & x_{k_3} & & x_{k_4} & \dots \end{array}$$

  so that $k_0=0$, $k_1=2$, $k_2=4$, $k_3=6$, …, i.e., $k_i=2\,i$ for all $i\in\mathbb{N}$. Hence, $y_i=x_{k_i}=x_{2\,i}$.

- $(k^8)_{k\in\mathbb{N}}$. In this case, the subsequence is obtained as the composition of the original sequence with the (strictly increasing) sequence $k:\mathbb{N}\longrightarrow\mathbb{N}$ defined by $k_i=i^4$;

- $((k+1)^2)_{k\in\mathbb{N}}$. The terms of the subsequence are $1,4,9,16,\dots$, and therefore the subsequence is the composition of the original sequence with the (strictly increasing) sequence $k:\mathbb{N}\longrightarrow\mathbb{N}$ defined by $k_i=i+1$.

It is easy to see that, if a sequence $(\boldsymbol{x}_k)_{k\in A}$ is bounded, i.e., if there exists $M>0$ such that for all $k\in A$ we have $\|\boldsymbol{x}_k\|\leq M$, any subsequence $(\boldsymbol{x}_{k_i})_{i\in\mathbb{N}}$ extracted from it is also bounded. Indeed, for all $i\in\mathbb{N}$ we have $k_i\in\mathbb{N}$, and therefore $\|\boldsymbol{x}_{k_i}\|\leq M$.

Exercise 1.98

In general, a strictly increasing sequence of real numbers is not necessarily divergent. For example $(\arctan k)_{k\in\mathbb{N}}$ is strictly increasing, but it converges to $\pi/2$. However, if the strictly increasing sequence is a sequence of natural numbers, it will diverge. To see this, consider a strictly increasing sequence $(k_j)_{j\in\mathbb{N}}$ of natural numbers. Show that

$$\forall\, j\in\mathbb{N},\quad k_j\geq j.$$

[Solution: (1.98).

An interesting question at this point is: if a sequence converges, what happens to its subsequences? The following result addresses this issue.

Theorem 1.40 *Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space. Let A be a subset of $\mathbb{N}$ obtained by removing from $\mathbb{N}$ a finite set. If a sequence $(\boldsymbol{x}_k)_{k\in A}$ tends to $\hat{\boldsymbol{x}}\in\boldsymbol{X}\cup\{\infty\}$, every subsequence extracted from it tends to $\hat{\boldsymbol{x}}$.*

Proof Let $(\boldsymbol{x}_{k_i})_{i\in\mathbb{N}}$ be a subsequence of $(\boldsymbol{x}_k)_{k\in A}$. We must show that for every neighborhood $\boldsymbol{W}$ of $\hat{\boldsymbol{x}}$ there exists $i^\star$ such that, for all $i>i^\star$, we have $\boldsymbol{x}_{k_i}\in\boldsymbol{W}$. Let $\boldsymbol{W}$ be a neighborhood of $\hat{\boldsymbol{x}}$. Let $p\in\mathbb{N}$ be such that

$$\forall\, k>p,\quad \boldsymbol{x}_k\in\boldsymbol{W}. \tag{1.51}$$

Since $(k_i)_{i\in\mathbb{N}}$ is a strictly increasing sequence of natural numbers, we can use the previous exercise to deduce that there exists $i^\star$ such that for all $i>i^\star$ we have $k_i>p$. This is the $i^\star$ we needed: indeed, from (1.51), since $k_i>p$, we obtain $\boldsymbol{x}_{k_i}\in\boldsymbol{W}$.

The most important theorem about the existence of convergent sequences is the Bolzano-Weierstrass theorem. We begin by considering sequences of real numbers. The Bolzano-Weierstrass theorem states that from every bounded real sequence it is possible to extract a convergent subsequence.

Theorem 1.41 [Bolzano-Weierstrass Theorem]
Let $(x_k)_{k\in\mathbb{N}}$ be a bounded sequence of $\mathbb{R}$. Then, there exists a strictly increasing function $k:\mathbb{N}\longrightarrow\mathbb{N}$ such that the subsequence $x\circ k$ is convergent (or, stated differently, there exists a strictly increasing sequence $(k_j)_{j\in\mathbb{N}}$ such that the subsequence $(x_{k_j})_{j\in\mathbb{N}}$ of the sequence $(x_k)_{k\in\mathbb{N}}$ is convergent).

Proof Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of real numbers. Assume that it is bounded, i.e., there exists $M > 0$ such that $|x_k| \le M$ for all $k \in \mathbb{N}$. Consider the set

$$A = \{y \in \mathbb{R} \mid \exists N \in \mathbb{N} : \quad \forall k > N,\ x_k \ge y\}.$$

In other words, A is the set of all values y that are smaller than x_k only for finitely many values of k. We show that A is non-empty and bounded from above.

It is easy to see that A is non-empty because $-M \in A$. In fact, if we take $y = -M$, in the definition of A we can take $N = 0$, and it is indeed true that for all $k > 0$ we have $x_k \ge -M$. It is equally easy to see that A is bounded from above, from the definition of A: for every $y \in A$, we have $y \le M$.

Since A is non-empty and bounded from above, it has a finite supremum, i.e., $\sup A \in \mathbb{R}$ (from the axiom of completeness of real numbers). Let $\hat{x} = \sup A$. We now construct a sequence which converges to $\hat{x}$:

- We take $b_1 \in A$ such that $\hat{x} - 1 < b_1 \le \hat{x}$: this is always possible because $A \cap (\hat{x} - 1, \hat{x}]$ is non-empty (recall that $\hat{x} = \sup A$). Since $b_1 \in A$, by definition of A,

$$\exists N_1 \in \mathbb{N} : \ \forall k > N_1, \quad x_k \ge b_1. \tag{1.52}$$

  Now, observe that $\hat{x} + 1 \notin A$ (otherwise $\hat{x}$ would not be the supremum of A), and therefore, from the way A is defined, it is false that $\exists N \in \mathbb{N}$ such that for all $h > N$ we have $x_h \ge \hat{x} + 1$. In other words, for all $N \in \mathbb{N}$, there exists $h > N$ such that $x_h < \hat{x} + 1$. We can take in particular the value $N = N_1$ defined in (1.52), and denote the corresponding h by k_1. It follows that there exists $k_1 > N_1$ such that

$$x_{k_1} \ge b_1 \quad \text{and} \quad x_{k_1} < \hat{x} + 1.$$

  Recalling that $b_1 > \hat{x} - 1$, we obtain $\hat{x} - 1 < b_1 \le x_{k_1} < \hat{x} + 1$, from which we find $\hat{x} - 1 < x_{k_1} < \hat{x} + 1$, and this inequality can be written as $|x_{k_1} - \hat{x}| < 1$;

- We take $b_2 \in A$ such that $\hat{x} - \frac{1}{2} < b_1 \le \hat{x}$: this is always possible because $A \cap \left(\hat{x} - \frac{1}{2}, \hat{x}\right]$ is non-empty, since $\hat{x} = \sup A$. Since $b_2 \in A$, by definition of A, there exists $N_2 \in \mathbb{N}$ such that, for all $k > N_2$, we have $x_k \ge b_2$. It is not restrictive to assume $k > k_1$, where k_1 was found at the previous step, i.e.,

$$\exists N_2 \in \mathbb{N} : \ \forall k > \max\{k_1, N_2\}, \quad x_k \ge b_2. \tag{1.53}$$

  Now, observe that $\hat{x} + \frac{1}{2} \notin A$, and therefore, from the way A is defined, it is false that $\exists N \in \mathbb{N}$ such that for all $h > N$ we have $x_h \ge \hat{x} + \frac{1}{2}$. In other words, for all $N \in \mathbb{N}$, there exists $h > N$ such that $x_h < \hat{x} + \frac{1}{2}$. We can take in particular $N = \max\{k_1, N_2\}$, where N_2 is defined in (1.53), and denote the corresponding h by k_2. It follows that

$$k_2 > k_1 \quad \text{and} \quad x_{k_2} \ge b_2 \quad \text{and} \quad x_{k_2} < \hat{x} + \frac{1}{2}.$$

  Recalling that $b_2 > \hat{x} - \frac{1}{2}$, we obtain $\hat{x} - \frac{1}{2} < b_2 \le x_{k_2} < \hat{x} + \frac{1}{2}$, from which we find $\hat{x} - \frac{1}{2} < x_{k_2} < \hat{x} + \frac{1}{2}$, and this inequality can be written as $|x_{k_2} - \hat{x}| < \frac{1}{2}$;

- Assume that, for some $j \in \mathbb{N}^*$, we have determined $1 \le k_1 < \ldots < k_j$ such that $|x_{k_r} - \hat{x}| < \frac{1}{r}$ for each $r = 1, \ldots, j$. We take $b_{j+1} \in A \cap \left(\hat{x} - \frac{1}{j+1}, \hat{x}\right]$. There exists N_{j+1} such that $x_k \ge b_{j+1}$ if $k > N_{j+1}$. On the other hand, $\hat{x} + \frac{1}{j+1} \notin A$. So we can find $k_{j+1} > \max\{k_j, N_{j+1}\}$ such that $b_{j+1} \le x_{k_{j+1}} < \hat{x} + \frac{1}{j+1}$, implying $|x_{k_{j+1}} - \hat{x}| < \frac{1}{j+1}$. In this way we construct a subsequence $(x_{k_j})_{j\in\mathbb{N}}$ of $(x_k)_{k\in\mathbb{N}}$ such that $|x_{k_j} - \hat{x}| < \frac{1}{j}$ for any $j \in \mathbb{N}^*$. This shows that the limit of x_{k_j} as $j \to \infty$ is $\hat{x}$, as required.

The following result is a generalization of Bolzano-Weierstrass theorem: it exploits the idea that the composition of a strictly increasing sequence of real numbers with a strictly increasing sequence of natural numbers is strictly increasing (see Exercise 1.100).

Theorem 1.42 *Let $q \in \mathbb{N}^*$. Let $(x_k^1)_{k\in\mathbb{N}}$, $(x_k^2)_{k\in\mathbb{N}}$, ..., $(x_k^q)_{k\in\mathbb{N}}$ be q bounded sequences of $\mathbb{R}$. Then, there exists an increasing function $\ell : \mathbb{N} \longrightarrow \mathbb{N}$ such that the subsequences $x^1 \circ \ell, x^2 \circ \ell, \ldots, x^q \circ \ell$ are all convergent in $\mathbb{R}$.*

Proof A precise proof of this result would require the use of induction. However, the mechanism is very simple, and we limit ourselves to considering the case $q = 3$. Using Bolzano-Weierstrass theorem (Theorem 1.41) starting from the first sequence $(x_k^1)_{k\in\mathbb{N}}$, we can extract a convergent subsequence $(x_{k_i}^1)_{i\in\mathbb{N}}$, where $(k_i)_{i\in\mathbb{N}}$ is a strictly increasing sequence of $\mathbb{N}$. We extract from $(x_k^2)_{k\in\mathbb{N}}$ the subsequence $(x_{k_i}^2)_{i\in\mathbb{N}}$ by choosing exactly its indices that ensure the convergence of $(x_{k_i}^1)_{i\in\mathbb{N}}$. This obviously results in a bounded subsequence. From this subsequence, using again Bolzano-Weierstrass theorem (Theorem 1.41), we can extract a convergent subsequence $(x_{k_{i_j}}^2)_{j\in\mathbb{N}}$, where $(k_{i_j})_{j\in\mathbb{N}}$ is a strictly increasing sequence of $\mathbb{N}$. We consider the bounded sequence obtained choosing these indices in $(x_k^3)_{k\in\mathbb{N}}$, i.e., the sequence $(x_{k_{i_j}}^3)_{j\in\mathbb{N}}$ is bounded, and therefore we can re-apply Bolzano-Weierstrass theorem (Theorem 1.41) to obtain a convergent subsequence $(x_{k_{i_{j_h}}}^3)_{h\in\mathbb{N}}$, where $(k_{i_{j_h}})_{h\in\mathbb{N}}$ is a strictly increasing sequence of $\mathbb{N}$. The sequences $(x_{k_{i_{j_h}}}^1)_{h\in\mathbb{N}}$, $(x_{k_{i_{j_h}}}^2)_{h\in\mathbb{N}}$ and $(x_{k_{i_{j_h}}}^3)_{h\in\mathbb{N}}$ are all convergent by Theorem 1.40.

We now show how to use Theorem 1.42 to generalize Bolzano-Weierstrass theorem (Theorem 1.41) to complex sequences. We recall that a sequence $(z_k)_{k\in\mathbb{N}}$ of complex numbers converges to the complex number $\hat{z}$ if and only if the sequences $(\mathfrak{Re}\{z_k\})_{k\in\mathbb{N}}$ and $(\mathfrak{Im}\{z_k\})_{k\in\mathbb{N}}$ converge to $\mathfrak{Re}\{\hat{z}\}$ and $\mathfrak{Im}\{\hat{z}\}$, respectively (see Exercise 1.99).

Theorem 1.43 [Bolzano-Weierstrass Theorem – The Complex Case]
Let $(z_k)_{k\in\mathbb{N}}$ be a bounded sequence of $\mathbb{C}$. Then, there exists a strictly increasing function $k : \mathbb{N} \longrightarrow \mathbb{N}$ such that the subsequence $z \circ k$ is convergent.

Proof Since $(z_k)_{k\in\mathbb{N}}$ is bounded, the real sequences $(\mathfrak{Re}\{z_k\})_{k\in\mathbb{N}}$ and $(\mathfrak{Im}\{z_k\})_{k\in\mathbb{N}}$ are also bounded (why?). Since $(\mathfrak{Re}\{z_k\})_{k\in\mathbb{N}}$ and $(\mathfrak{Im}\{z_k\})_{k\in\mathbb{N}}$ are real sequences, we can apply Theorem 1.42 to them and obtain two subsequences $(\mathfrak{Re}\{z_{k_j}\})_{j\in\mathbb{N}}$ and $(\mathfrak{Im}\{z_{k_j}\})_{j\in\mathbb{N}}$ which are convergent. Hence, the sequence $(z_{k_j})_{j\in\mathbb{N}}$ is also convergent.

Just like real sequences, we can consider subsequences of a given sequence with elements in a generic normed space. It is still true that every subsequence of a convergent sequence converges to the same limit. However, Bolzano-Weierstrass theorem (Theorem 1.41) can only be extended to finite-dimensional normed spaces. In preparation for this result, we present a result that we have already anticipated, which plays an important role: all norms in a finite dimensional vector space are equivalent.

Theorem 1.44 *Let $\boldsymbol{X}$ be a* **finite dimensional** *vector space over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ be two norms in $\boldsymbol{X}$. Then, $\|\cdot\|_{(1)}$ and $\|\cdot\|_{(2)}$ are equivalent.*

Proof Let n denote the dimension of $\boldsymbol{X}$, and let $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n)$ be a basis of $\boldsymbol{X}$. We express every vector $\boldsymbol{x} \in \boldsymbol{X}$ through its coordinates $x_1, x_2, \ldots, x_n$ with respect to this basis and we consider the norm given by the formula

$$\|\boldsymbol{x}\| = |x_1| + |x_2| + \ldots + |x_n|. \tag{1.54}$$

Since the equivalence of norms is an equivalence relation, it is sufficient to show that every norm in $\boldsymbol{X}$ is equivalent to the norm defined in (1.54). Let $\|\cdot\|_\star$ be a norm in $\boldsymbol{X}$. Using the triangle inequality and the

homogeneity we have

$$\begin{aligned}\|\boldsymbol{x}\|_\star = \|x_1\,\boldsymbol{b}_1 + x_2\,\boldsymbol{b}_2 + \ldots + x_n\,\boldsymbol{b}_n\|_\star &\le \|x_1\,\boldsymbol{b}_1\|_\star + \|x_2\,\boldsymbol{b}_2\|_\star + \ldots + \|x_n\,\boldsymbol{b}_n\|_\star \\ &= |x_1|\,\|\boldsymbol{b}_1\|_\star + |x_2|\,\|\boldsymbol{b}_2\|_\star + \ldots + |x_n|\,\|\boldsymbol{b}_n\|_\star \le c_1\,\|\boldsymbol{x}\|, \end{aligned} \tag{1.55}$$

where $c_1 = \max\limits_{1\le i\le n} \|\boldsymbol{b}_i\|_\star$. We need to prove the opposite inequality. In other words, we need to prove that

$$\exists\, c_2 > 0: \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|\boldsymbol{x}\| \le c_2\,\|\boldsymbol{x}\|_\star. \tag{1.56}$$

We reason by contradiction: we assume that such c_2 does not exist. The negation of (1.56) is the statement

$$\forall\, c_2 > 0, \quad \exists\, \boldsymbol{x} \in \boldsymbol{X}: \quad \|\boldsymbol{x}\| > c_2\,\|\boldsymbol{x}\|_\star. \tag{1.57}$$

Notice that $\boldsymbol{x}$ in (1.57) is allowed to depend on c_2. It is easy to see that we can restrict c_2 to be a natural number, and we obtain an equivalent condition (why? see Exercise 1.101). Therefore, we can re-write (1.57) as

$$\forall\, k \in \mathbb{N}^*: \quad \exists\, \boldsymbol{x}_k \in \boldsymbol{X}, \quad \|\boldsymbol{x}_k\| > k\,\|\boldsymbol{x}_k\|_\star. \tag{1.58}$$

This means that there exists a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}^*}$ of vectors of $\boldsymbol{X}$ such that

$$\|\boldsymbol{x}_k\| > k\,\|\boldsymbol{x}_k\|_\star \quad \forall\, k \in \mathbb{N}^*. \tag{1.59}$$

There holds $\|\boldsymbol{x}_k\| > 0$. In fact, if we had $\|\boldsymbol{x}_k\| = 0$, then from the norm axioms $\boldsymbol{x}_k = \boldsymbol{0}$, which in turn implies that $\|\boldsymbol{x}_k\|_\star = 0$, and from (1.58) we would get $0 > 0$. Dividing through by $k\,\|\boldsymbol{x}_k\|$ both sides of (1.59) gives $\frac{1}{k} > \frac{\|\boldsymbol{x}_k\|_\star}{\|\boldsymbol{x}_k\|}$ for all $k \in \mathbb{N}^*$. Now, consider the new vectors $\boldsymbol{\xi}_k = \boldsymbol{x}_k/\|\boldsymbol{x}_k\|$. Using the homogeneity, we see that $\|\boldsymbol{\xi}_k\|_\star = \left\|\frac{\boldsymbol{x}_k}{\|\boldsymbol{x}_k\|}\right\|_\star = \frac{1}{\|\boldsymbol{x}_k\|}\,\|\boldsymbol{x}_k\|_\star$, so that combining the two together we obtain

$$\|\boldsymbol{\xi}_k\|_\star < \frac{1}{k} \qquad \text{and} \qquad \|\boldsymbol{\xi}_k\| = 1, \quad \forall\, k \ge 1. \tag{1.60}$$

We denote by $\xi_{k,i}$ the i-th component of $\boldsymbol{\xi}_k$, so that

$$\boldsymbol{\xi}_k = \xi_{k,1}\,\boldsymbol{b}_1 + \xi_{k,2}\,\boldsymbol{b}_2 + \ldots + \xi_{k,n}\,\boldsymbol{b}_n.$$

The sequences $(\xi_{k,i})_{k\in\mathbb{N}}$ are real or complex bounded sequences. Using the generalization of Bolzano-Weierstrass Theorem (Theorem 1.42), we can construct the n subsequences $(\xi_{k_j,i})_j$ $(i = 1, 2, \ldots, n)$, each one converging to its limit ℓ_i, which involve the same indices k_j. Now we introduce the vector $\boldsymbol{\xi} = \ell_1\,\boldsymbol{b}_1 + \ell_2\,\boldsymbol{b}_2 + \ldots + \ell_n\,\boldsymbol{b}_n$. We will find a contradiction by showing that $\boldsymbol{\xi}$ needs to be, simultaneously, equal to zero and different from zero. Using the triangle inequality and (1.55), and from the first of (1.60), we find

$$\begin{aligned}\|\boldsymbol{\xi}\|_\star &= \|\boldsymbol{\xi} - \boldsymbol{\xi}_{k_j} + \boldsymbol{\xi}_{k_j}\|_\star \\ &\le \|\boldsymbol{\xi} - \boldsymbol{\xi}_{k_j}\|_\star + \|\boldsymbol{\xi}_{k_j}\|_\star \le c_1\,\|\boldsymbol{\xi} - \boldsymbol{\xi}_{k_j}\| + \frac{1}{k_j} = c_1 \sum_{i=1}^{n} |\ell_i - \xi_{k_j,i}| + \frac{1}{k_j},\end{aligned}$$

and the last term tends to zero as $j \to \infty$. Thus, $\|\boldsymbol{\xi}\|_\star = 0$, so that $\boldsymbol{\xi} = \boldsymbol{0}$. On the other hand, from the second of (1.60) and from the reverse triangle inequality (see *(iv)* in Theorem 1.2) we find

$$\Big|\|\boldsymbol{\xi}\| - 1\Big| = \Big|\|\boldsymbol{\xi}\| - \|\boldsymbol{\xi}_{k_j}\|\Big| \le \|\boldsymbol{\xi} - \boldsymbol{\xi}_{k_j}\| = \sum_{i=1}^{n} |\ell_i - \xi_{k_j,i}|,$$

and the last term tends to zero as $j \to \infty$. Thus, $\|\boldsymbol{\xi}\| = 1$, and $\boldsymbol{\xi}$ cannot be the zero vector.

The result of Theorem 1.44 fails if the vector space is infinite dimensional, as the following result shows.

Example 1.63

Consider the vector space $C([a,b];\mathbb{R})$ of continuous functions defined over the closed interval $[a,b]$ (where $a,b\in\mathbb{R}$ satisfy $a<b$). We show that the norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are not equivalent. Let $f\in C([a,b];\mathbb{R})$. It is easy to see that

$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x \le (b-a)\max_{a\le x\le b}|f(x)| = (b-a)\,\|f\|_\infty,$$

and we can take $b-a$ as the value of one of the two constants in Definition 1.24. However, there does not exist $c>0$ such that $\|f\|_\infty \le c\,\|f\|_1$ for all $f\in C([a,b];\mathbb{R})$. Suppose for the sake of argument that $a=0$ and $b=1$, and consider the sequence of functions $(f_k)_{k\in\mathbb{N}}$ defined, for $k\in\mathbb{N}$, by

$$\begin{aligned} f_k : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto x^k. \end{aligned}$$

We have $f_k\in C([0,1];\mathbb{R})$ for all $k\in\mathbb{N}$. For all $k\in\mathbb{N}$ we have

$$\|f_k\|_1 = \int_0^1 x^k\,\mathrm{d}x = \left[\frac{x^{k+1}}{k+1}\right]_0^1 = \frac{1}{k+1} \quad \text{and} \quad \|f_k\|_\infty = \max_{x\in[0,1]}|x^k| = 1.$$

For the two norms to be equivalent, there should exist $c>0$ such that $1\le c\,\frac{1}{k+1}$ for all $k\in\mathbb{N}$, i.e., $c\ge k+1$ for all $k\in\mathbb{N}$, which is a contradiction: there is no real number greater than any natural number.

As an example of application of the previous result, we consider again sequences $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ in a normed space $(\boldsymbol{X},\|\cdot\|)$. To fix the ideas, suppose that $\boldsymbol{X}=\mathbb{R}^2$, and consider the sequence $\boldsymbol{x}_k=(x_{k,1},x_{k,2})=\left(3+\frac{1}{k},\frac{\sin k}{k^2}\right)$. This sequence tends to the vector $\hat{\boldsymbol{x}}=(3,0)$. Hence, the convergence of a sequence of vectors implies the convergence of each component and the limit of the vector is the vector of the limits of each component and, moreover, the converse is also true. We now want to extend this result, which appears to be natural for sequences in $\mathbb{R}^n$, to the case of a generic finite dimensional normed space.

Theorem 1.45 *Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $\boldsymbol{X}$ be finite dimensional, and let n denote the dimension of $\boldsymbol{X}$. Let $\mathcal{B}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\ldots,\boldsymbol{b}_n)$ be a basis of $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of $\boldsymbol{X}$. Let $\hat{\boldsymbol{x}}\in\boldsymbol{X}$. Let*

$$\begin{aligned} \hat{\boldsymbol{x}} &= \hat{x}_1\,\boldsymbol{b}_1+\hat{x}_2\,\boldsymbol{b}_2+\ldots+\hat{x}_n\,\boldsymbol{b}_n \\ \boldsymbol{x}_k &= x_{k,1}\,\boldsymbol{b}_1+x_{k,2}\,\boldsymbol{b}_2+\ldots+x_{k,n}\,\boldsymbol{b}_n \qquad \forall\, k\in\mathbb{N}, \end{aligned}$$

where $x_1,\ldots,x_n\in\mathbb{F}$ and where $x_{k,1},\ldots,x_{k,n}\in\mathbb{F}$ for all $k\in\mathbb{N}$. Then, the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$ if and only if, for all $i\in\{1,2,\ldots,n\}$, the sequence $(x_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$.

Proof Since $\boldsymbol{X}$ is finite dimensional, all the norms defined for $\boldsymbol{X}$ are equivalent to each other. Thus, we only need to prove the statement for one particular norm, and the result will extend to all other norms. Let us choose the norm

$$\|\hat{\boldsymbol{x}}\| = |\hat{x}_1|+|\hat{x}_2|+\ldots+|\hat{x}_n|.$$

Then,

$$\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\| = |x_{k,1}-\hat{x}_1|+|x_{k,2}-\hat{x}_2|+\ldots+|x_{k,n}-\hat{x}_n|. \tag{1.61}$$

We prove that if the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$, then the sequence $(x_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$ for all $i\in\{1,2,\ldots,n\}$. From (1.61), for all $i\in\{1,2,\ldots,n\}$ we have

$$0\le|x_{k,i}-\hat{x}_i|\le\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\|.$$

By virtue of the Squeeze Theorem (Theorem 1.26), if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$, then $\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\|\xrightarrow{k\to\infty}0$, and

therefore $|x_{k,i} - \hat{x}_i| \xrightarrow{k\to\infty} 0$ for all $i \in \{1, 2, \dots, n\}$.
Conversely, suppose that $(x_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$ for all $i \in \{1, 2, \dots, n\}$, and we prove that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$. In other words, we show that for every $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that, for all $k > p$, we have $\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| < \varepsilon$. Let $\varepsilon > 0$ and let $i \in \{1, 2, \dots, n\}$. Since $(x_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$, there exists $p_i \in \mathbb{N}$ such that, for all $k > p_i$, we have $|x_{k,i} - \hat{x}_i| < \varepsilon/n$. Hence, defining $p = \max\{p_1, \dots, p_n\}$, for all $k > p$ we have from (1.61)

$$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| = |x_{k,1} - \hat{x}_1| + |x_{k,2} - \hat{x}_2| + \dots + |x_{k,n} - \hat{x}_n| < \underbrace{\frac{\varepsilon}{n} + \frac{\varepsilon}{n} + \dots + \frac{\varepsilon}{n}}_{n \text{ times}} = \varepsilon,$$

as required.

We are now ready to address the generalization of Bolzano-Weierstrass theorem to finite-dimensional normed spaces: if $(\boldsymbol{X}, \|\cdot\|)$ is a finite-dimensional normed space, every bounded sequence of $\boldsymbol{X}$ has a convergent subsequence.

Theorem 1.46 [Bolzano-Weierstrass Theorem – The General Case]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a finite-dimensional normed space over a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a bounded sequence of $\boldsymbol{X}$. Then, there exists a strictly increasing function $k : \mathbb{N} \longrightarrow \mathbb{N}$ such that the subsequence $\boldsymbol{x} \circ k$ is convergent.

Proof Given a basis $\mathcal{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n)$ of $\boldsymbol{X}$, we can express every vector $\boldsymbol{\xi} \in \boldsymbol{X}$ through its coordinates $\xi_1, \xi_2, \dots, \xi_n$ with respect to $\mathcal{B}$. We can consider again the norm $\|\boldsymbol{\xi}\| = |\xi_1| + |\xi_2| + \dots + |\xi_n|$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a bounded sequence of $\boldsymbol{X}$, and for every $k \in \mathbb{N}$ let us denote by $x_{k,1}, x_{k,2}, \dots, x_{k,n}$ the coordinates of $\boldsymbol{x}_k$ with respect to $\mathcal{B}$. Since for every $k \in \mathbb{N}$ we have $\|\boldsymbol{x}_k\| = |x_{k,1}| + |x_{k,2}| + \dots + |x_{k,n}|$, the coordinates $x_{k,i}$ satisfy

$$|x_{k,i}| \le \|\boldsymbol{x}_k\| \quad \forall\, k \in \mathbb{N},\ \forall\, i \in \{1, 2, \dots, n\}.$$

Hence, the coordinates of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ define n real or complex bounded sequences. We can then apply Theorem 1.42, and we obtain a strictly increasing sequence $(k_j)_{j\in\mathbb{N}}$ of $\mathbb{N}$ such that the corresponding n convergent subsequences $(x_{k_j,i})_{j\in\mathbb{N}}$ ($i \in \{1, 2, \dots, n\}$) converge to the limits $\ell_1, \ell_2, \dots, \ell_n$, respectively. From

$$\left\| \boldsymbol{x}_{k_j} - \sum_{i=1}^{n} \ell_i\, \boldsymbol{b}_i \right\| = \left\| \sum_{i=1}^{n} x_{k_j,i}\, \boldsymbol{b}_i - \sum_{i=1}^{n} \ell_i\, \boldsymbol{b}_i \right\| = \left\| \sum_{i=1}^{n} (x_{k_j,i} - \ell_i)\, \boldsymbol{b}_i \right\| = \sum_{i=1}^{n} |x_{k_j,i} - \ell_i|,$$

we deduce that the subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ converges to $\displaystyle\sum_{i=1}^{n} \ell_i\, \boldsymbol{b}_i$.

A stronger (but harder to prove) result holds: a normed space $(\boldsymbol{X}, \|\cdot\|)$ is finite-dimensional if and only if from every bounded sequence in $\boldsymbol{X}$ we can extract a convergent subsequence. We restrict ourselves to proposing an example of an infinite dimensional normed space where there exists at least a bounded sequence from which it is not possible to extract a convergent subsequence.

Example *1.64*

Consider the vector space $C([0, 1]; \mathbb{R})$ with the norm $\|\cdot\|_\infty$. First of all, we observe that, as for the set of bounded functions from a non-empty set to $\mathbb{R}$, the convergence associated with this norm is the uniform convergence. Let us take the sequence of functions $(f_k)_{k\in\mathbb{N}}$ defined by

$$\begin{aligned} f_k : [0, 1] &\longrightarrow \mathbb{R} \\ x &\mapsto x^k. \end{aligned}$$

This sequence is bounded, since for every $k \in \mathbb{N}$ we have

$$\|f_k\|_\infty = \sup_{x \in [0,1]} |x^k| = 1,$$

but it does not have convergence subsequences. In fact, if by contradiction there exists a convergent subsequence $(f_{k_j})_{j \in \mathbb{N}}$, a continuous function f exists to which such subsequence must converge. However, the given sequence converges pointwise to the function

$$\begin{aligned} g : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1. \end{cases} \end{aligned}$$

It follows that the subsequence considered is pointwise convergent to g as well. On the other hand, the same subsequence converges pointwise to f, which implies $f = g$. This is a contradiction, since f is continuous and g is discontinuous.

Exercises

Exercise 1.99 Show that a sequence $(z_k)_{k \in \mathbb{N}}$ of complex numbers converges to the complex number $\hat{z}$ if and only if the sequences $(\Re\mathrm{e}\{z_k\})_{k \in \mathbb{N}}$ and $(\Im\mathrm{m}\{z_k\})_{k \in \mathbb{N}}$ converge to $\Re\mathrm{e}\{\hat{z}\}$ and $\Im\mathrm{m}\{\hat{z}\}$, respectively.
(Hint: for the **if** part, remember the triangle inequality. For the **only if** part, recall that for every complex number z we have $\Re\mathrm{e}\{z\} \le |\Re\mathrm{e}\{z\}| \le |z|$ and $\Im\mathrm{m}\{z\} \le |\Im\mathrm{m}\{z\}| \le |z|$.) [Solution: (1.99)]

Exercise 1.100 Let $(a_k)_{k \in \mathbb{N}}$ be a strictly increasing sequence of real numbers, and let $(\beta_k)_{k \in \mathbb{N}}$ be a strictly increasing sequence of natural numbers. Show that the sequence $a \circ \beta$ is strictly increasing. [Solution: (1.100)]

Exercise 1.101 Prove that the statement

$$\forall\, c \in \mathbb{R}_+^*,\ \exists\, \boldsymbol{x} \in \boldsymbol{X} :\quad \|\boldsymbol{x}\| > c\, \|\boldsymbol{x}\|_\star$$

is equivalent to

$$\forall\, k \in \mathbb{N}^*,\ \exists\, \boldsymbol{x} \in \boldsymbol{X} :\quad \|\boldsymbol{x}\| > k\, \|\boldsymbol{x}\|_\star.$$

[Solution: (1.101)]

Exercise 1.102 Make, if possible, one example for each of the following cases:

- a sequence $(x_k)_{k \in \mathbb{N}}$ of real numbers that oscillates, but has a convergent subsequence;
- a sequence $(x_k)_{k \in \mathbb{N}}$ of real numbers that is not bounded, but has a convergent subsequence;
- a sequence $(x_k)_{k \in \mathbb{N}}$ of real numbers that diverges, from which it is impossible to extract a convergent subsequence;
- a sequence $(x_k)_{k \in \mathbb{N}}$ of real numbers that is bounded, and does not have a convergent subsequence.

(Hint: one of those examples cannot be made, as it would violate an important theorem: which one?) [Solution: (1.102)]

Exercise 1.103 Let $(\boldsymbol{X}, \|\cdot\|)$ be a finite-dimensional normed vector space. Let $n \in \mathbb{N}^*$ be the dimension of $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ be a sequence of vectors of $\boldsymbol{X}$. Let $\mathcal{B}$ be a basis of $\boldsymbol{X}$. Then, the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ converges to $\hat{\boldsymbol{x}}$ if and only if, for

every $i \in \{1, 2, \dots, n\}$, the series $\sum_{k=0}^{\infty} x_k^i$ of the i-th coordinates of the elements $\boldsymbol{x}_k$ with respect to the basis $\mathcal{B}$ of $\boldsymbol{X}$ converges to the i-th coordinate $\hat{x}_i$ of $\hat{\boldsymbol{x}}$. [Solution: (1.103)]

We finish by listing a set of properties that can be deduced by applying Theorem 1.45 to sequences of $\mathbb{R}^n$. In the case of $\mathbb{R}^n$, some properties of the convergence of sequences involve the inner product.

Theorem 1.47 *Let $n \in \mathbb{N}^*$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ and $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ be two sequences of $\mathbb{R}^n$, respectively convergent to the limits $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$. Then,*

$$\lim_{k\to\infty} (\boldsymbol{x}_k \cdot \boldsymbol{y}_k) = \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}}.$$

Proof Using the Cauchy-Schwarz inequality we find immediately

$$\begin{aligned} |\boldsymbol{x}_k \cdot \boldsymbol{y}_k - \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}}| &= |\boldsymbol{x}_k \cdot \boldsymbol{y}_k - \boldsymbol{x}_k \cdot \hat{\boldsymbol{y}} + \boldsymbol{x}_k \cdot \hat{\boldsymbol{y}} - \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}}| \\ &\leq |\boldsymbol{x}_k \cdot \boldsymbol{y}_k - \boldsymbol{x}_k \cdot \hat{\boldsymbol{y}}| + |\boldsymbol{x}_k \cdot \hat{\boldsymbol{y}} - \hat{\boldsymbol{x}} \cdot \hat{\boldsymbol{y}}| \\ &= |\boldsymbol{x}_k \cdot (\boldsymbol{y}_k - \hat{\boldsymbol{y}})| + |(\boldsymbol{x}_k - \hat{\boldsymbol{x}}) \cdot \hat{\boldsymbol{y}}| \\ &\leq \|\boldsymbol{x}_k\| \, \|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\| + \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \, \|\hat{\boldsymbol{y}}\|. \end{aligned}$$

The right hand-side of the latter goes to 0 as $k \to \infty$: indeed, in the first term of the sum, $\|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\|$ goes to 0, and the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent, and therefore bounded. The second term of the sum clearly goes to 0 since $\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|$ goes to 0.

Theorem 1.48 *Let $n \in \mathbb{N}^*$. A sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of $\mathbb{R}^n$ converges to $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ if and only if, for every $\boldsymbol{y} \in \mathbb{R}^n$, there holds*

$$\lim_{k\to\infty} (\boldsymbol{x}_k \cdot \boldsymbol{y}) = \hat{\boldsymbol{x}} \cdot \boldsymbol{y}.$$

Proof One implication can be obtained from Theorem 1.47 by taking $\boldsymbol{y}_k = \boldsymbol{y}$. The other can be obtained by choosing as $\boldsymbol{y}$ the vectors of the standard basis of $\mathbb{R}^n$ and applying Theorem 1.45.

Example *1.65*

Consider the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of $\mathbb{R}^2$ given by

$$\boldsymbol{x}_k = \left(\frac{1}{k}, \frac{1}{k^2}\right).$$

This sequence converges to $(0, 0)$. Consider for example the Euclidean norm. We have indeed

$$\|\boldsymbol{x}_k\|^2 = k^{-2} + k^{-4}.$$

Theorem 1.45 applied to the standard basis ensures that the real sequences $\left(\frac{1}{k}\right)_{k\in\mathbb{N}}$ and $\left(\frac{1}{k^2}\right)_{k\in\mathbb{N}}$ both tend to 0, and Theorem 1.48 guarantees that, given an arbitrary vector $\boldsymbol{y} = (y_1, y_2) \in \mathbb{R}^2$, the sequence

$$\left(\frac{y_1}{k} + \frac{y_2}{k^2}\right)_{k\in\mathbb{N}}$$

is also convergent to zero.

Example *1.66*

Consider the sequence $(A_k)_{k\in\mathbb{N}}$ of $m \times n$ matrices with real entries $a_{k,i,j}$. Regardless of the norm considered in the space $\mathbb{R}^{m\times n}$, the convergence of this sequence to a matrix $A \in \mathbb{R}^{m\times n}$, with entries $a_{i,j}$, is equivalent to the

mn conditions

$$\lim_{k\to\infty} a_{k,i,j} = a_{i,j} \qquad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n.$$

In fact, it suffices to apply Theorem 1.45 to the standard basis of $\mathbb{R}^{m\times n}$: for $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, n\}$, the ij-th matrix of the basis has 1 in position (i, j) and 0 elsewhere.

1.8 Continuity

We can now proceed with a definition of continuity.

Definition 1.30 [Continuity]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. We say that f is continuous at $\boldsymbol{x}_0$ if

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$$

Moreover, the function f is said to be continuous if it is continuous at $\boldsymbol{x}_0$ for every $\boldsymbol{x}_0 \in \boldsymbol{A}$.

Notice that the condition in Definition 1.30 can be equivalently written as

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \cap \mathcal{B}_\delta(\boldsymbol{x}_0, \|\cdot\|_X, \boldsymbol{X}), \quad \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon,$$

where we recall that $\mathcal{B}_\delta(\boldsymbol{x}_0, \|\cdot\|_X, \boldsymbol{X}) = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta\}$.

The equivalent characterizations of convergence given in Theorem 1.13 can be adapted, *mutatis mutandis*, for the definition of continuity. In particular, under the assumptions of Definition 1.30, f is continuous at $\boldsymbol{x}_0$ if and only if

$$\forall\, \boldsymbol{W} \in \mathcal{N}(f(\boldsymbol{x}_0), \|\cdot\|_Y, \boldsymbol{Y}),\ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0, \|\cdot\|_X, \boldsymbol{X}) :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \cap \boldsymbol{V}, \quad f(\boldsymbol{x}) \in \boldsymbol{W}.$$

Notice that, however, here $\boldsymbol{W}$ and $\boldsymbol{V}$ are never neighborhoods of the point ∞: in fact, in the definition of continuity it is required that $\boldsymbol{x}_0 \in \boldsymbol{A}$ and $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$, so that $\boldsymbol{x}_0$ is a vector of $\boldsymbol{X}$ and $f(\boldsymbol{x}_0)$ is a vector of $\boldsymbol{Y}$. Notice also that, differently from the definition of convergence, $\boldsymbol{x}_0$ is allowed to be an isolated point of $\boldsymbol{A}$. We will see in Theorem 1.49 that, in this case, Definition 1.30 becomes vacuous.

We have the usual relationship between continuity and limit when $\boldsymbol{x}_0$ is also a limit point: f is continuous at $\boldsymbol{x}_0$ if and only if

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0).$$

The precise statement is as follows.

Theorem 1.49 [Convergence and Continuity]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. Then:

(i) If $\boldsymbol{x}_0 \in A \setminus \mathcal{D}(\boldsymbol{A})$ (i.e., if $\boldsymbol{x}_0$ is an isolated point of $\boldsymbol{A}$), then f is continuous at $\boldsymbol{x}_0$;

(ii) if $\boldsymbol{x}_0 \in \boldsymbol{A} \cap \mathcal{D}(\boldsymbol{A})$ (i.e., if $\boldsymbol{x}_0$ is a point of A but not isolated), then f is continuous at $\boldsymbol{x}_0$ if and only if f is convergent at $\boldsymbol{x}_0$, and $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0)$.

Proof We prove the first. Since $\boldsymbol{x}_0 \in \boldsymbol{A} \setminus \mathcal{D}(\boldsymbol{A})$ is equivalent to saying that $\boldsymbol{x}_0$ is an isolated point of A, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}_0) \cap \boldsymbol{A} = \{\boldsymbol{x}_0\}$. We have to show that for all $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$, there holds $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$. Let $\varepsilon > 0$. Let us choose $\delta = r$. Clearly, $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$ implies $\boldsymbol{x} = \boldsymbol{x}_0$, so that $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y = \|f(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)\|_Y = 0 < \varepsilon$, as required.

We now prove the second claim. Let $\boldsymbol{x}_0 \in \boldsymbol{A} \cap \mathcal{D}(\boldsymbol{A})$. Let f be continuous at $\boldsymbol{x}_0$. Then, for all $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$, there holds $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$. This is obviously still true if, in particular, we choose $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$ and $\boldsymbol{x} \neq \boldsymbol{x}_0$. Therefore, the condition of Definition 1.16 is satisfied with $\boldsymbol{\ell} = f(\boldsymbol{x}_0)$. Conversely, let $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = f(\boldsymbol{x}_0)$. From Definition 1.16, given $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$ with $\boldsymbol{x} \neq \boldsymbol{x}_0$, the inequality $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$ holds. However, this inequality is also trivially satisfied for $\boldsymbol{x} = \boldsymbol{x}_0$, and therefore it holds for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$. Hence, the condition of Definition 1.30 is satisfied.

We have already seen a result on the limit of the composition. We now introduce another useful result.

Theorem 1.50 *Let $(\boldsymbol{X}, \|\cdot\|_X)$, $(\boldsymbol{Y}, \|\cdot\|_Y)$ and $(\boldsymbol{Z}, \|\cdot\|_Z)$ be normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $\boldsymbol{B} \subseteq \boldsymbol{Y}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ and $g : \boldsymbol{B} \longrightarrow \boldsymbol{Z}$. Let $f(\boldsymbol{A}) \subseteq \boldsymbol{B}$. Let $\boldsymbol{x}_0 \in \tilde{\mathcal{D}}(\boldsymbol{A})$, and suppose that $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \boldsymbol{y}_0$. Moreover, let $\boldsymbol{y}_0 \in \boldsymbol{B}$ and let g be continuous at $\boldsymbol{y}_0$. Then, $g \circ f$ is convergent at $\boldsymbol{x}_0$ and*

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} (g \circ f)(\boldsymbol{x}) = g(\boldsymbol{y}_0).$$

Proof We have to prove that for every neighborhood $\boldsymbol{U}$ of $g(\boldsymbol{y}_0)$, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{A}$, if $\boldsymbol{x} \neq \boldsymbol{x}_0$ then $g(f(\boldsymbol{x})) \in \boldsymbol{U}$. Let $\boldsymbol{U}$ be a neighborhood of $g(\boldsymbol{y}_0)$. From the continuity of g at $\boldsymbol{y}_0$, there exists a neighborhood $\boldsymbol{W}$ of $\boldsymbol{y}_0$ such that, for all $\boldsymbol{y} \in \boldsymbol{W} \cap \boldsymbol{B}$, we have $g(\boldsymbol{y}) \in \boldsymbol{U}$. Since f converges to $\boldsymbol{y}_0$ as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$, there exists a neighborhood $\boldsymbol{V}$ of $\boldsymbol{x}_0$ such that, for all $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{A}$, if $\boldsymbol{x} \neq \boldsymbol{x}_0$ we have $f(\boldsymbol{x}) \in \boldsymbol{W}$. On the other hand, we have also $f(\boldsymbol{x}) \in \boldsymbol{B}$, so that $f(\boldsymbol{x}) \in \boldsymbol{W} \cap \boldsymbol{B}$, and therefore we can take $\boldsymbol{y} = f(\boldsymbol{x})$, and we find that for all $\boldsymbol{x} \in \boldsymbol{V} \cap \boldsymbol{A}$, if $\boldsymbol{x} \neq \boldsymbol{x}_0$ we have $g(\boldsymbol{y}) = g(f(\boldsymbol{x})) \in \boldsymbol{U}$, as required.

We have the following consequences.

Theorem 1.51 *Let $(\boldsymbol{X}, \|\cdot\|_X)$, $(\boldsymbol{Y}, \|\cdot\|_Y)$ and $(\boldsymbol{Z}, \|\cdot\|_Z)$ be normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $\boldsymbol{B} \subseteq \boldsymbol{Y}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ and $g : \boldsymbol{B} \longrightarrow \boldsymbol{Z}$ be such that $f(\boldsymbol{A}) \subseteq \boldsymbol{B}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. If f is continuous at $\boldsymbol{x}_0$ and g is continuous at $f(\boldsymbol{x}_0)$, then $g \circ f$ is continuous at $\boldsymbol{x}_0$.*

Proof If $\boldsymbol{x}_0 \in \boldsymbol{A} \setminus \mathcal{D}(\boldsymbol{A})$ (i.e., if $\boldsymbol{x}_0$ is an isolated point of $\boldsymbol{A}$), the conclusion is trivial from *(i)* of Theorem 1.49 (in this case we do not even need to assume the continuity of g at $f(\boldsymbol{x}_0)$). Let $\boldsymbol{x}_0 \in \boldsymbol{A} \cap \mathcal{D}(\boldsymbol{A})$. We must show that

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} (g \circ f)(\boldsymbol{x}) = (g \circ f)(\boldsymbol{x}_0).$$

This follows immediately from Theorem 1.50.

Theorem 1.52 *Let $(\boldsymbol{X}, \|\cdot\|_X)$, $(\boldsymbol{Y}, \|\cdot\|_Y)$ and $(\boldsymbol{Z}, \|\cdot\|_Z)$ be normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and $\boldsymbol{B} \subseteq \boldsymbol{Y}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$ and $g : \boldsymbol{B} \longrightarrow \boldsymbol{Z}$ be continuous functions satisfying the inclusion $f(\boldsymbol{A}) \subseteq \boldsymbol{B}$. Then, $g \circ f$ is continuous.*

Proof This follows trivially from Theorem 1.51.

The following theorem extends a well-known result for the one-dimensional case, and the proof is essentially the same:

Theorem 1.53 [Sequential Continuity Theorem]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. Then, f is continuous at $\boldsymbol{x}_0$ if and only if, for every sequence $(\boldsymbol{\xi}_k)_{k \in \mathbb{N}}$ with values in $\boldsymbol{A}$ and convergent to $\boldsymbol{x}_0$, the

sequence $(f(\boldsymbol{\xi}_k))_{k\in\mathbb{N}}$ *converges to* $f(\boldsymbol{x}_0)$.

Proof **(If)**. We need to show that

$$\forall\,\varepsilon>0,\ \ \exists\,\delta>0:\ \ \forall\,\boldsymbol{x}\in\boldsymbol{A}\cap\mathcal{B}_\delta(\boldsymbol{x}_0),\qquad \|f(\boldsymbol{x})-f(\boldsymbol{x}_0)\|_Y<\varepsilon, \tag{1.62}$$

where the ball $\mathcal{B}_\delta(\boldsymbol{x}_0)$ is associated with the norm $\|\cdot\|_X$ in $\boldsymbol{X}$. By contradiction, we negate (1.62), and we assume that

$$\exists\,\varepsilon>0,\ \ \forall\,\delta>0:\ \ \exists\,\boldsymbol{x}\in\boldsymbol{A}\cap\mathcal{B}_\delta(\boldsymbol{x}_0),\qquad \|f(\boldsymbol{x})-f(\boldsymbol{x}_0)\|_Y\geq\varepsilon. \tag{1.63}$$

Let ε be the number provided by the statement (1.63). We choose also $\delta=\frac{1}{k}$ with $k\in\mathbb{N}^*$. For this choice, there exists a point $\boldsymbol{\xi}_k\in\boldsymbol{A}$ with $\boldsymbol{\xi}_k\in\mathcal{B}_{\frac{1}{k}}(\boldsymbol{x}_0)$ (so that $\|\boldsymbol{\xi}_k-\boldsymbol{x}_0\|_X<\frac{1}{k}$), and $\|f(\boldsymbol{\xi}_k)-f(\boldsymbol{x}_0)\|_Y\geq\varepsilon$, i.e., $f(\boldsymbol{\xi}_k)\notin\mathcal{B}_\varepsilon(f(\boldsymbol{x}_0))$. It follows that the sequence $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}^*}$ thus constructed converges to $\boldsymbol{x}_0$, but, since $f(\boldsymbol{\xi}_k)\notin\mathcal{B}_\varepsilon(f(\boldsymbol{x}_0))$ for all $k\in\mathbb{N}^*$, then $f(\boldsymbol{\xi}_k)$ does not tend to $f(\boldsymbol{x}_0)$, leading to a contradiction.

(Only if). Let $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ be a sequence with values in $\boldsymbol{A}$ and convergent to $\boldsymbol{x}_0$, i.e.,

$$\forall\,\varepsilon>0,\ \exists\,p\in\mathbb{R}:\ \ \forall\,k>p,\quad \|\boldsymbol{\xi}_k-\boldsymbol{x}_0\|_X<\varepsilon. \tag{1.64}$$

We assume that f is continuous at $\boldsymbol{x}_0$, so that

$$\forall\,\varepsilon>0,\ \exists\,\delta>0:\ \ \forall\,\boldsymbol{x}\in\mathcal{B}_\delta(\boldsymbol{x}_0)\cap\boldsymbol{A},\quad \|f(\boldsymbol{x})-f(\boldsymbol{x}_0)\|_Y<\varepsilon. \tag{1.65}$$

We have to show that the sequence $(f(\boldsymbol{\xi}_k))_{k\in\mathbb{N}}$ converges to $f(\boldsymbol{x}_0)$:

$$\forall\,\varepsilon>0,\ \exists\,p\in\mathbb{R}:\ \ \forall\,k>p,\quad \|f(\boldsymbol{\xi}_k)-f(\boldsymbol{x}_0)\|_Y<\varepsilon. \tag{1.66}$$

Let $\varepsilon>0$. From (1.65),

$$\exists\,\delta>0:\ \ \forall\,\boldsymbol{x}\in\mathcal{B}_\delta(\boldsymbol{x}_0)\cap\boldsymbol{A},\quad \|f(\boldsymbol{x})-f(\boldsymbol{x}_0)\|_Y<\varepsilon. \tag{1.67}$$

We now use (1.64) taking the arbitrary ε appearing therein to be equal to the value of δ thus found, so that

$$\exists\,p\in\mathbb{R}:\ \ \forall\,k>p,\quad \|\boldsymbol{\xi}_k-\boldsymbol{x}_0\|_X<\delta.$$

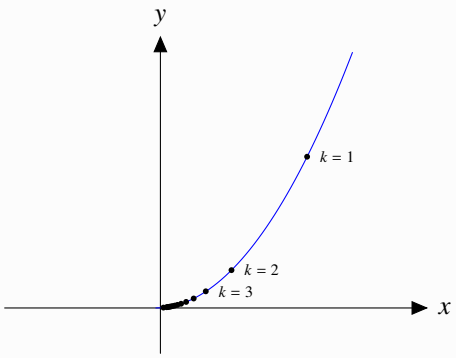
It follows that for all $k>p$ we have $\boldsymbol{\xi}_k\in\mathcal{B}_\delta(\boldsymbol{x}_0)\cap\boldsymbol{A}$, and therefore from (1.67) we conclude that $\|f(\boldsymbol{\xi}_k)-f(\boldsymbol{x}_0)\|_Y<\varepsilon$.

Example *1.67*

The function

$$\begin{aligned} f:\mathbb{R}^2&\longrightarrow\mathbb{R}\\ (x,y)&\mapsto\begin{cases}1 & \text{if } x>0 \text{ and } y=x^2\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

is 1 on an arc of parabola that has $(0,0)$ as limit point, and it is zero elsewhere. Consider the sequence $\left(\frac{1}{k},\frac{1}{k^2}\right)_{k\in\mathbb{N}^*}$, which obviously converges to $(0,0)$ as $k\to\infty$.



If f is continuous at $(0,0)$, we should have that $f\left(\frac{1}{k}, \frac{1}{k^2}\right)$ tends to $f(0,0)$ as $k \to \infty$. On the other hand, we have $f\left(\frac{1}{k}, \frac{1}{k^2}\right) = 1$ for every $k \in \mathbb{N}^*$, whereas $f(0,0) = 0$. Hence, the function f is discontinuous at $(0,0)$. Notice that f does not have a limit as $(x,y) \to (0,0)$: considering in fact the sequence $\left(\frac{1}{k}, 0\right)_{k\in\mathbb{N}^*}$, the sequence $\left(f\left(\frac{1}{k}, 0\right)\right)_{k\in\mathbb{N}}$ is the zero sequence, which tends to 0.

The sequential continuity theorem can be used to prove that the norm is a continuous function.

Theorem 1.54 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. The function $\|\cdot\| : \boldsymbol{X} \longrightarrow \mathbb{R}$ is continuous.*

Proof For the sake of clarity, we denote the function $\|\cdot\|$ by f:

$$\begin{aligned} f : \boldsymbol{X} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\|. \end{aligned}$$

Let $\boldsymbol{x}_0 \in \boldsymbol{X}$. We show that f is continuous at $\boldsymbol{x}_0$ by proving that, given a sequence $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ of elements of $\boldsymbol{X}$ convergent to $\boldsymbol{x}_0$, the sequence $(f(\boldsymbol{\xi}_k))_{k\in\mathbb{N}}$ converges to $f(\boldsymbol{x}_0)$, i.e. that

$$\forall\, \varepsilon > 0, \quad \exists\, p \in \mathbb{R} : \quad \forall\, k > p, \quad |f(\boldsymbol{\xi}_k) - f(\boldsymbol{x}_0)| = \big|\|\boldsymbol{\xi}_k\| - \|\boldsymbol{x}_0\|\big| < \varepsilon.$$

Let $\varepsilon > 0$. Since $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ tends to $\boldsymbol{x}_0$ as $k \to \infty$, there exists $N > 0$ such that for all $k > N$ we have $\|\boldsymbol{\xi}_k - \boldsymbol{x}_0\| < \varepsilon$. Choosing $p = N$ and using the reverse triangle inequality gives

$$\big|\|\boldsymbol{\xi}_k\| - \|\boldsymbol{x}_0\|\big| \leq \big|\|\boldsymbol{\xi}_k\| - \|\boldsymbol{x}_0\|\big| < \varepsilon,$$

as required.

We can rewrite the previous result as

$$\forall\, \boldsymbol{x}_0 \in \boldsymbol{X},\ \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \ \Rightarrow\ \big|\|\boldsymbol{x}\| - \|\boldsymbol{x}_0\|\big| < \varepsilon.$$

Here, the value of δ that satisfies the latter depends both from the choice of $\boldsymbol{x}_0$ and from the choice of ε. However, it is possible to prove an even stronger result: a value of $\delta > 0$ always exists that only depends on the choice of ε, i.e.,

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \quad \forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \qquad \|\boldsymbol{x} - \boldsymbol{y}\| < \delta \ \Rightarrow\ \big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| < \varepsilon.$$

When a condition of this kind is satisfied, we say that the function (the norm in this case) is *uniformly continuous*. Notice the strong relationship between uniform continuity and uniform convergence.

In the case of finite dimensional spaces, for the convergence of the sequences we can consider the sequences of the coordinates with respect to a basis, or, in the case of $\mathbb{R}^n$, the sequences of the components in the n directions. Indeed, if $(\boldsymbol{Y}, \|\cdot\|_Y)$ is a finite dimensional normed space (over a field $\mathbb{F}$, which is $\mathbb{R}$ or $\mathbb{C}$) and we take a basis

$\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ in $\boldsymbol{Y}$, and a function $f : A \longrightarrow \boldsymbol{Y}$ (where A is an arbitrary set), we have $f(\boldsymbol{x}) \in \boldsymbol{Y}$ for every $x \in A$, so that there exist $y_1, \ldots, y_m \in \mathbb{F}$ such that $f(x) = y_1\, \boldsymbol{b}_1 + \ldots + y_m\, \boldsymbol{b}_m$. The functions

$$\begin{aligned} f_i : A &\longrightarrow \mathbb{F} \\ \boldsymbol{x} &\mapsto y_i \end{aligned}$$

are the components of $f(x)$ with respect to the basis $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$.

Definition 1.31 *Let A be a set. Let $(\boldsymbol{Y}, \|\cdot\|_Y)$ be a finite dimensional normed space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $f : A \longrightarrow \boldsymbol{Y}$. If $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$ is a basis of $\boldsymbol{Y}$, for all $x \in A$, let $f_1(x), f_2(x), \ldots, f_m(x)$ denote the coordinates of $f(x)$ with respect to the basis. The functions $f_i : A \longrightarrow \mathbb{F}$ that map $x \in A$ into $f_i(x)$ for all $i \in \{1, 2, \ldots, m\}$ are called* components *of f with respect to the given basis.*
If $\boldsymbol{Y} = \mathbb{R}^m$, given a unit vector $\boldsymbol{u} \in \partial\mathcal{B}_1(0)$, the function that maps $x \in A$ into $f(x) \cdot \boldsymbol{u}$ is called component of f in the direction $\boldsymbol{u}$.

We have the following result, whose proof can once again be obtained by modifying the one-dimensional one.

Theorem 1.55 *Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A}$. Let $\boldsymbol{Y}$ be finite-dimensional, and let $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m)$ be a basis of $\boldsymbol{Y}$. Then, f is continuous at $\boldsymbol{x}_0$ if and only if the m components $f_1, \ldots, f_m$ are continuous at $\boldsymbol{x}_0$.*
If $\boldsymbol{Y} = \mathbb{R}^m$, f is continuous at $\boldsymbol{x}_0$ if and only if, for any unit vector $\boldsymbol{u}$, the component of f with respect to $\boldsymbol{u}$ is continuous at $\boldsymbol{x}_0$.

Proof Since all norms are equivalent in $\boldsymbol{Y}$ (recall that $\boldsymbol{Y}$ is finite dimensional), we are are free to choose whichever norm to use. We can take, as norm of a vector $\boldsymbol{y} \in \boldsymbol{Y}$, the maximum among the absolute value or the modulus (depending on whether $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, respectively) of the components of $\boldsymbol{y}$ with respect to a given basis of $\boldsymbol{Y}$. In other words, if $y_1, \ldots, y_m$ are the coordinates of $\boldsymbol{y}$ with respect to a basis,

$$\|\boldsymbol{y}\|_Y = \max_{1 \le i \le m} |y_i|.$$

Given $\varepsilon > 0$, we have

$$\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon \quad \Leftrightarrow \quad \forall\, i \in \{1, \ldots, m\},\ \ |f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}_0)| < \varepsilon. \tag{1.68}$$

If f_i is continuous at $\boldsymbol{x}_0$ for all $i \in \{1, \ldots, m\}$, then for every $i \in \{1, \ldots, m\}$ and for all $\varepsilon_i > 0$ there exists $\delta_i > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_{\delta_i}(\boldsymbol{x}_0) \cap \boldsymbol{A}$, we have $|f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}_0)| < \varepsilon_i$. Let $\varepsilon > 0$. Take $\varepsilon_1 = \ldots = \varepsilon_m = \varepsilon$ and $\delta = \min\{\delta_1, \ldots, \delta_m\}$, then for all $i \in \{1, \ldots, m\}$ and for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$ we have $|f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}_0)| < \varepsilon$, so that $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$ from (1.68).

Conversely, let f be continuous at $\boldsymbol{x}_0$, so that for all $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$, we have $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon$. Using (1.68), this is the same as saying that for all $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}$, we have $\|f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}_0)\| < \varepsilon$ for all $i \in \{1, \ldots, m\}$. This statement is stronger than (and therefore it implies that)

$$\forall\, i \in \{1, \ldots, m\},\ \forall\, \varepsilon > 0,\ \ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \mathcal{B}_\delta(\boldsymbol{x}_0) \cap \boldsymbol{A}, \qquad |f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}_0)| < \varepsilon.$$

Example *1.68* Consider the function

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (x, y) &\mapsto (\|(x, y)\|_2^2, x\cos(x\,y^2 + 3\,x^2), \ln(1 + x^2)) \end{aligned}$$

is continuous at every point $(x, y) \in \mathbb{R}^2$, because its three components are continuous at every point $(x, y) \in \mathbb{R}^2$ with respect to the standard basis $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$. In fact, the first component, $\|(x, y)\|$, is a polynomial. The second and the third components can be obtained by multiplying, adding and performing the composition of continuous functions. We observe that the functions that map, for every $\boldsymbol{x} \in \mathbb{R}^n$ into its n-th component x_i for $i \in \{1, 2, \ldots, n\}$ are continuous, since $|x_i - y_i| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$ for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, and that a function described by the expression $g(x, y) = h(x)$ can be obtained via the composition of h with one of these functions.

Example *1.69*

Consider the function

$$\begin{aligned} f : [a, b] &\longrightarrow \mathbb{R}^{m \times n} \\ x &\mapsto A(x). \end{aligned}$$

Assigning this function is equivalent to assigning $m\,n$ functions $x \mapsto a_{i,j}(x)$ for $x \in [a, b]$ with values in $\mathbb{R}$. This function is continuous at x_0 if and only if the functions $a_{i,j}$ are continuous at x_0 for all $i = 1, \ldots, m$ and $j = 1, \ldots, m$.

In the case of functions of a real variable, we can check the continuity and the limit at a point in terms of the left and right limits. In the general case of normed vector spaces, we do not have an order, and we can proceed in a different way.

The following result can be applied to functions that are "pointwise elementary":

Theorem 1.56 *Let $(\boldsymbol{X}, \|\cdot\|_x)$ and $(\boldsymbol{Y}, \|\cdot\|_y)$ be two normed spaces. Let $\boldsymbol{A}, \boldsymbol{B} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \cup \boldsymbol{B} \longrightarrow \boldsymbol{Y}$. Let $\boldsymbol{x}_0 \in \boldsymbol{A} \cap \boldsymbol{B}$. Then, f is continuous at $\boldsymbol{x}_0$ if and only if the two restrictions of f to $\boldsymbol{A}$ and $\boldsymbol{B}$ are continuous at $\boldsymbol{x}_0$.*

Example *1.70*

The function

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto \begin{cases} \|(x, y)\|_2 & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

This function is continuous at $(0, 0)$. Indeed, the function

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto \|(x, y)\|_2 \end{aligned}$$

is continuous at $(0, 0)$ (and not only at $(0, 0)$, but at all points, from the inequality $\big|\|\boldsymbol{u}\| - \|\boldsymbol{v}\|\big| \leq \|\boldsymbol{u} - \boldsymbol{v}\|$), and the same is true for the identically zero function. Thus, all the restrictions of these two functions are continuous, and applying the previous result we find that f is continuous at $(0, 0)$.
Clearly, we could have proceeded in a different way to check the continuity, by using the definition with $\delta = \varepsilon$.

Theorem 1.56 can be extended to the case of a finite number of restrictions. If the restrictions are infinite, we cannot deduce the continuity of f at a point $\boldsymbol{x}_0$ of $\mathbb{R}^2$ on the basis of the continuity of all lines through $\boldsymbol{x}_0$. In fact, if the restrictions of f to the lines through $\boldsymbol{x}_0$ are all continuous and we fix $\varepsilon > 0$, for each of these lines, say R, there exists $\delta > 0$ such that the inequality $\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| < \varepsilon$ holds for every point of R which satisfy $\|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta$. Every line R defines a segment S_R with $\boldsymbol{x}_0$ as end-point and entirely contained in R, where such inequality still holds. Thus, this inequality holds at every point of the union of these segments, but such union may not be a neighborhood of $\boldsymbol{x}_0$ In fact, δ may depend on the line considered.

Example *1.71*

Consider again the function

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto \begin{cases} 1 & \text{if } x > 0 \text{ and } y = x^2 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

We saw that this function is discontinuous at the origin. However, the restrictions to the lines through the origin are all continuous. In fact, if the line R does not intersect the first quadrant, we can choose δ arbitrarily for any ε. If R is given by the equation $y = p\,x$, with $p > 0$ and if $\varepsilon \in (0, 1)$, we can still choose an acceptable value of $\delta > 0$, sufficiently small so that the segment S_R does not contain the point (p, p^2) where f is 1. Thus, all the restrictions are continuous. However, the choice of δ in this second case is such that the union of all the segments S_R thus found is not a neighborhood of $(0, 0)$.

A way of seeing the restrictions to the half-lines with $\boldsymbol{x}_0$ as end-point consists in using polar coordinates. A point $\boldsymbol{x} \in \mathbb{R}^n$, distinct from $\boldsymbol{x}_0$, can always be written as

$$\boldsymbol{x} = \boldsymbol{x}_0 + \rho\,\boldsymbol{u}, \qquad \text{with } \rho > 0 \text{ and } \boldsymbol{u} \in \partial\mathcal{B}_1(0) \tag{1.69}$$

by taking

$$\rho = \|\boldsymbol{x} - \boldsymbol{x}_0\| \quad \text{and} \quad \boldsymbol{u} = \frac{\boldsymbol{x} - \boldsymbol{x}_0}{\|\boldsymbol{x} - \boldsymbol{x}_0\|}.$$

This representation is unique. Conversely, every point in the form (1.69) is distinct from $\boldsymbol{x}_0$. The ball $\mathcal{B}_r(\boldsymbol{x}_0)$ is described by the conditions $0 \leq \rho < r$ and $\boldsymbol{u} \in \partial\mathcal{B}_1(0)$. To verify if the restriction of a real function f to this half-line associated with the vector of unit length $\boldsymbol{u}$ is continuous at $\boldsymbol{x}_0$ means to establish if

$$\lim_{\rho \to 0^+} f(\boldsymbol{x}_0 + \rho\,\boldsymbol{u}) = f(\boldsymbol{x}_0) \tag{1.70}$$

holds, whereas to check if f is continuous at $\boldsymbol{x}_0$ means to check that

$$\forall\,\varepsilon > 0,\ \exists\,\delta > 0 :\ \forall\,\rho \in (0, \delta),\ \forall\,\boldsymbol{u} \in \mathcal{B}_1(0),\ |f(\boldsymbol{x}_0 + \rho\,\boldsymbol{u}) - f(\boldsymbol{x}_0)| \leq \varepsilon. \tag{1.71}$$

Condition (1.71) implies (1.70) for every $\boldsymbol{u} \in \mathcal{B}_1(0)$, but it is not equivalent to the set of such conditions. In fact, (1.71) means that

$$\lim_{\rho \to 0^+} f(\boldsymbol{x}_0 + \rho\,\boldsymbol{u}) = f(\boldsymbol{x}_0) \quad \text{uniformly with respect to } \boldsymbol{u} \in \mathcal{B}_1(0)$$

Exercises

Exercise 1.104

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed vector space. Show that the function

$$\begin{aligned} f : \boldsymbol{X} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\| \end{aligned}$$

is continuous (Hint: use the sequential continuity theorem and the reverse triangle inequality). [Solution: (1.104)]

Exercise 1.105

Let $m, n \in \mathbb{N}^*$, and let $A \in \mathbb{R}^{m \times n}$. Show that the function

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \boldsymbol{x} &\mapsto A\,\boldsymbol{x} \end{aligned}$$

is continuous. [Solution: (1.105)]

Exercise 1.106 Consider the function

$$f : [-1,2] \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} x^2 + 2x + 1 & \text{if } x \in [-1,0] \\ -x + 1 & \text{if } x \in (0,2] \end{cases}$$

- Show that $f \in C([-1,2];\mathbb{R})$;
- Find $\|f\|_1$;
- Find $\|f\|_\infty$.

[Solution: (1.106)]

1.8.1 More on uniform convergence and uniform continuity

We have shown with examples that the pointwise limit of a sequence of continuous functions is not necessarily continuous. We show that this is not true for uniform convergence.

Theorem 1.57 *Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$ and consider the sequence $(f_k)_{k\in\mathbb{N}}$ of functions $f_k : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Let this sequence converge uniformly to a function $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. If f_k is continuous for all $k \in \mathbb{N}$, then f is also continuous.*

Proof We have to show that, given any $\boldsymbol{x}_0 \in \boldsymbol{A}$,

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y < \varepsilon.$$

Let $\boldsymbol{x}_0 \in \boldsymbol{A}$, and let $\varepsilon > 0$. Since $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly,

$$\exists\, p \in \mathbb{N} :\ \forall\, k > p,\ \ \forall\, \boldsymbol{x} \in \boldsymbol{A},\ \|f_k(\boldsymbol{x}) - f(\boldsymbol{x})\|_Y < \frac{\varepsilon}{3}.$$

Take any $s > p$, so that

$$\forall\, \boldsymbol{x} \in \boldsymbol{A},\ \|f_s(\boldsymbol{x}) - f(\boldsymbol{x})\|_Y < \frac{\varepsilon}{3}.$$

Moreover, since f_s is continuous,

$$\exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f_s(\boldsymbol{x}) - f_s(\boldsymbol{x}_0)\|_Y < \frac{\varepsilon}{3}.$$

Thus, for all $\boldsymbol{x} \in \boldsymbol{A}$ which satisfies $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$, using the triangle inequality twice we have

$$\begin{aligned} \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\|_Y &= \|f(\boldsymbol{x}) - f_s(\boldsymbol{x}) + f_s(\boldsymbol{x}) - f_s(\boldsymbol{x}_0) + f_s(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)\|_Y \\ &= \|f(\boldsymbol{x}) - f_s(\boldsymbol{x})\|_Y + \|f_s(\boldsymbol{x}) - f_s(\boldsymbol{x}_0)\|_Y + \|f_s(\boldsymbol{x}_0) - f(\boldsymbol{x}_0)\|_Y < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3}. \end{aligned}$$

We conclude with a result that says that we can interchange the limit with the Riemann integral over a bounded domain of integration if the sequence of functions is uniformly convergent.

Theorem 1.58 *Let $A \subseteq \mathbb{R}$ be bounded, and consider the sequence $(f_k)_{k\in\mathbb{N}}$ of continuous functions $f_k : A \longrightarrow \mathbb{R}$. Let this sequence be uniformly convergent. Then,*

$$\lim_{k\to\infty} \int_A f_k(x)\,\mathrm{d}x = \int_A \lim_{k\to\infty} f_k(x)\,\mathrm{d}x.$$

Proof Let $f : A \longrightarrow \mathbb{R}$ be the function such that $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. From Theorem 1.57, f is continuous. We have to show that

$$\lim_{k\to\infty} \int_A f_k(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x.$$

Notice that the integrals in this expression are well defined as Riemann integrals because the functions are continuous and the domain of integration is bounded. From the linearity of the Riemann integral

$$\int_A f_k(x)\,\mathrm{d}x - \int_A f(x)\,\mathrm{d}x = \int_A (f_k(x) - f(x))\,\mathrm{d}x.$$

Let $a, b \in \mathbb{R}$ be such that $a < b$, and let $I = (a, b)$ be a bounded interval with the property that $A \subseteq I$. The length of the interval I is equal to $b - a$. We have to show that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\quad \forall\, k > p,\ \left| \int_A f_k(x)\,\mathrm{d}x - \int_A f(x)\,\mathrm{d}x \right| < \varepsilon.$$

Let $\varepsilon > 0$. Since $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly,

$$\exists\, p \in \mathbb{N} :\quad \forall\, k > p, \forall\, x \in A,\quad |f_k(x) - f(x)| < \frac{\varepsilon}{b-a}.$$

It follows that for all $k > p$

$$\left| \int_A f_k(x)\,\mathrm{d}x - \int_A f(x)\,\mathrm{d}x \right| = \left| \int_A (f_k(x) - f(x))\,\mathrm{d}x \right| \leq \int_A |f_k(x) - f(x)|\,\mathrm{d}t < \frac{\varepsilon}{b-a}\,(b-a) = \varepsilon,$$

as required.

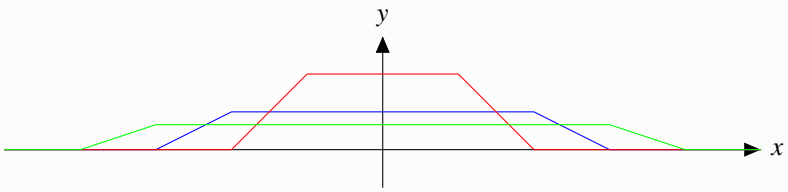
In order to prove an analogous result for the interchange between derivative and limit (which, however, will require not just the uniform convergence of the sequence $(f_k)_{k\in\mathbb{N}}$, but also of the sequence of the derivatives), we need some additional machinery that we will develop in the next chapter.

Example *1.72*

It is possible to show that in Theorem 1.58, the assumption of continuity can be relaxed. However, the assumption that the domain A be bounded is essential (as well as the uniform convergence of the sequence). Consider for example the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ from $\mathbb{R}$ to $\mathbb{R}$ defined for all $t \in \mathbb{R}$ by

$$f_k(t) = \begin{cases} \frac{1}{k} & \text{if } |t| \leq k \\ \frac{(k+1)-|t|}{k} & \text{if } k \leq |t| \leq k+1 \\ 0 & \text{if } |t| \geq k+1 \end{cases}$$

The graphs of f_k for $k = 1$ (red), $k = 2$ (blue) and $k = 3$ (green) are given in the Figure:



For every $k \in \mathbb{N}^*$, the function f_k is continuous. Moreover, the sequence is pointwise convergent to the zero function, since for every $t \in \mathbb{R}$ we have

$$\lim_{k\to\infty} f_k(t) = \lim_{k\to\infty} \frac{1}{k} = 0.$$

From

$$\sup_{x\in\mathbb{R}} |f_k(t)| = \frac{1}{k},$$

we see that the sequence is uniformly convergent to the zero function. However, a simple calculation (or straightforward geometric considerations on the graphs of f_k) shows that

$$\lim_{k\to\infty} \int_{-\infty}^{+\infty} f_k(t)\,\mathrm{d}t = \lim_{k\to\infty} \left(2 + \frac{1}{k}\right) = 2,$$

and therefore the interchange of the limit with the integral is incorrect.

We now formally introduce the notion of uniform continuity.

Definition 1.32 [Uniform Continuity]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. We say that f is uniformly continuous if

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}, \qquad \|\boldsymbol{x} - \boldsymbol{y}\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - f(\boldsymbol{y})\|_Y < \varepsilon$$

Theorem 1.59 [Continuity and Uniform Continuity]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Let $f : \boldsymbol{A} \longrightarrow \boldsymbol{Y}$. Then,

$$f \text{ is uniformly continuous} \quad \Rightarrow \quad f \text{ is continuous.}$$

Proof Let f be uniformly continuous. Let $\epsilon > 0$ and $\boldsymbol{x} \in \boldsymbol{A}$. There exists $\delta > 0$ such that

$$\forall\, \boldsymbol{y}, z \in \boldsymbol{A}, \qquad \|\boldsymbol{y} - z\| < \delta \ \Rightarrow\ \|f(\boldsymbol{y}) - f(z)\| < \epsilon.$$

Then, if $\boldsymbol{y} \in \boldsymbol{A}$ and if $\|\boldsymbol{y} - \boldsymbol{x}\| < \delta$, we obtain $\|f(\boldsymbol{y}) - f(\boldsymbol{x})\| < \epsilon$.

Remark 1.13 The implication in Theorem 1.59 cannot be inverted in general, i.e., a function can be continuous without being uniformly continuous. For example, consider the continuous function

$$\begin{aligned} f : \mathbb{R}_+ &\longrightarrow \mathbb{R} \\ x &\mapsto x^2. \end{aligned}$$

We show that f is not uniformly continuous. Suppose by contradiction that f is uniformly continuous. Then, assuming in the definition of uniformly continuous function $\epsilon = 1$, there exists $\delta > 0$ such that for all $x, y \in \mathbb{R}_+$,

$$|x - y| < \delta \ \Rightarrow\ |x^2 - y^2| < 1.$$

Let $x \geq 1$. For $y \in \mathbb{R}_+$ the inequality $|x^2 - y^2| < 1$ is equivalent to $x^2 - 1 \leq y^2 \leq x^2 + 1$, i.e., $\sqrt{x^2 - 1} \leq y \leq \sqrt{x^2 + 1}$. Since $|x - y| < \delta$ implies $|x^2 - y^2| < 1$, we have

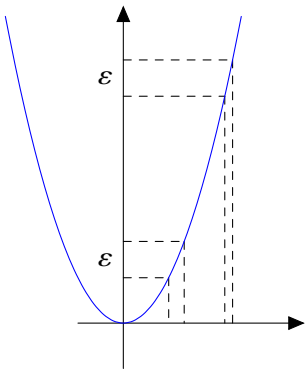
$$(x - \delta, x + \delta) \subseteq (\sqrt{x^2 - 1},\ \sqrt{x^2 + 1}).$$

Thus, $2\,\delta \leq \sqrt{x^2 + 1} - \sqrt{x^2 - 1}$. This inequality holds for $x \geq 1$. One can easily verify that

$$\lim_{\substack{x\to\infty \\ x\geq 1}} \sqrt{x^2 + 1} - \sqrt{x^2 - 1} = 0.$$

Thus, $2\,\delta \leq 0$, which is a contradiction.

Notice that in these considerations we have implicitly used the dependence of x on the neighborhood V of x such that, for all $x \in V \cap \mathbb{R}_+$, there holds $|f(x) - f(y)| < 1$. We can graphically visualize that f is not uniformly continuous. Indeed, if we shift along the y-axis a segment of width ϵ, as the segment moves towards $+\infty$ the points x and y such that $f(x)$ and $f(y)$ lie in the segment need to be chosen progressively closer to each other, and their distance tends to 0.



Theorem 1.60 *Let $a \in \mathbb{R}$, and $f : \mathbb{R} \longrightarrow \mathbb{R}$. Let*

$$\begin{aligned} f_a : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto f(x-a). \end{aligned}$$

Then, f is uniformly continuous if and only if, for every sequence $(a_k)_{k\in\mathbb{N}}$ of real numbers convergent to zero, the sequence $(f_{a_k})_{k\in\mathbb{N}}$ converges uniformly to f.

Proof Let f be uniformly continuous, so that

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, x, y \in A, \qquad |x - y| < \delta \ \Rightarrow\ |f(x) - f(y)| < \varepsilon. \tag{1.72}$$

Let $(a_k)_{k\in\mathbb{N}}$ be an infinitesimal sequence of real numbers, so that

$$\forall\, \varepsilon > 0,\ \exists\, q \in \mathbb{N} :\ \forall\, k > q,\ \ |a_k| < \varepsilon. \tag{1.73}$$

We want to show that the sequence $(f_{a_k})_{k\in\mathbb{N}}$ converges uniformly to f, i.e., that

$$\forall\, \varepsilon > 0,\ \exists\, p > 0 :\ \forall\, k > p,\ \forall\, x \in A,\ |f_{a_k}(x) - f(x)| < \varepsilon.$$

Let $\varepsilon > 0$. Using (1.72) with $x - a_k$ in place of x and with x in place of y, we find that

$$\exists\, \delta > 0 :\ \|x - a_k - x\| < \delta \ \Rightarrow\ |f(x - a_k) - f(x)| < \varepsilon.$$

We can now use this δ as the ε in (1.73), and the corresponding value of q is the required value of p.

We now prove the converse. Suppose that for every infinitesimal sequence $(a_k)_{k\in\mathbb{N}}$ the sequence $(f_{a_k})_{k\in\mathbb{N}}$ converges to f uniformly, but assume by contradiction that f is not uniformly continuous, i.e., that

$$\exists\, \varepsilon > 0,\ \forall\, \delta > 0 :\ \exists\, x, y \in A, \qquad |x - y| < \delta \text{ and } |f(x) - f(y)| \geq \varepsilon. \tag{1.74}$$

Take $k \in \mathbb{N}^*$ and $\delta = 1/k$. From (1.74)

$$\exists\, x_k, y_k \in A : |x_k - y_k| < \delta \text{ and } |f(x_k) - f(y_k)| \geq \varepsilon.$$

For all $k \in \mathbb{N}^*$, let $a_k = x_k - y_k$, so that we obtain a sequence $(a_k)_{k\in\mathbb{N}^*}$. With this definition, we have seen that

$$\exists\, \varepsilon > 0,\ \forall\, k \in \mathbb{N}^* :\ \exists\, a_k \in \mathbb{R} : \qquad |a_k| < \delta \text{ and } |f(x_k) - f_{a_k}(x)| \geq \varepsilon,$$

which says that $(a_k)_{k\in\mathbb{N}^*}$ is infinitesimal and $(f_{a_k})_{k\in\mathbb{N}}$ is not uniformly convergent to f.

Exercises

Exercise 1.107 Consider the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ where, for all $k \in \mathbb{N}^*$, we define

$$\begin{aligned} f_k : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 2k & \text{for } x \in \left[\frac{1}{2k}, \frac{1}{k}\right] \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Show that, denoting by f the pointwise limit of $(f_k)_{k\in\mathbb{N}^*}$, we have

$$\lim_{k\to\infty} \int_0^1 f_k(x)\,\mathrm{d}x \neq \int_0^1 f(x)\,\mathrm{d}x.$$

Show that this result is consistent with the check that shows that $(f_k)_{k\in\mathbb{N}^*}$ is not uniformly convergent to its pointwise limit. Repeat the exercise by considering, instead, the moving bump

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{for } x \in [k, k+1] \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

and by proving the integral inequality

$$\lim_{k\to\infty} \int_{\mathbb{R}} f_k(x)\,\mathrm{d}x \neq \int_{\mathbb{R}} f(x)\,\mathrm{d}x.$$

[Solution: (1.107)]

1.9 Norms of matrices

Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Given $m, n \in \mathbb{N}^*$, we have introduced the 1-norm, the Frobenius norm and the infinity norm for the space $\mathbb{F}^{m\times n}$ of $m \times n$ matrices with entries in $\mathbb{F}$. Our aim is not to add some properties to ensure that the norm of a matrix interacts well with the product of matrices.

Definition 1.33 [Consistent Norms]
Let $n, m, p \in \mathbb{N}^*$. Consider the norms $\|\cdot\|_{m,p}$, $\|\cdot\|_{p,n}$ and $\|\cdot\|_{m,n}$ in the spaces $\mathbb{F}^{m\times p}$, $\mathbb{F}^{p\times n}$ and $\mathbb{F}^{m\times n}$, respectively. We say that $\|\cdot\|_{m,p}$, $\|\cdot\|_{p,n}$ and $\|\cdot\|_{m,n}$ are consistent *if for all $A \in \mathbb{F}^{m\times p}$ and for all $B \in \mathbb{F}^{p\times n}$*

$$\|A\,B\|_{m,n} \leq \|A\|_{m,p}\,\|B\|_{p,n}.$$

For square matrices, we can introduce an interesting type of norm, called *matrix norm*. Consider the vector space $\mathbb{F}^{n\times n}$ of square matrices of order $n \in \mathbb{N}^*$ with entries in $\mathbb{F}$. In this case, the three norms of Definition 1.33 are norms of the same vector space. We can also assume that these three norms are the same norm $\|\cdot\|$ (for example,

all three are 1-norms, or Frobenius norms). Then, a matrix norm is a matrix consistent with itself. This leads to the following definition.

Definition 1.34 [Matrix Norm]
A norm $\|\cdot\|$ *in* $\mathbb{F}^{n\times n}$ *is a* matrix norm *if, for* $A, B \in \mathbb{F}^{n\times n}$*, there holds*

$$\|A\,B\| \le \|A\|\,\|B\|. \tag{1.75}$$

There is no reason why a generic norm should satisfy this condition. For example, if we multiply a norm $\|\cdot\|$ by a real number $\kappa > 0$, we obtain another norm, and (1.75), relative to the norm $\kappa\|\cdot\|$ with $A = B = I$ becomes $\kappa\|I\| \le (\kappa\|I\|)^2$. This is clearly false if $\kappa < 1/\|I\|$.

Example *1.73*

It is easy to see that, given a generic matrix norm $\|\cdot\|$ in $\mathbb{F}^{n\times n}$, we have

- $\|A^k\| \le \|A\|^k$ for all $k \in \mathbb{N}^*$;
- $\|I\| \ge 1$;
- if $A \in \mathbb{F}^{n\times n}$ is invertible, then $\|A^{-1}\| \ge \frac{1}{\|A\|}$.

The first is trivial. By induction, for $k = 1$ it is certainly true, and assuming $\|A^k\| \le \|A\|^k$ for some $k \in \mathbb{N}^*$ gives

$$\|A^{k+1}\| = \|A^k\,A\| \le \|A^k\|\,\|A\| \le \|A\|^k\,\|A\| = \|A\|^{k+1}.$$

To see the second, we apply the sub-multiplicativity to $A = B = I$:

$$\|I\,I\| \le \|I\|\,\|I\| \quad \Leftrightarrow \quad \|I\| \le \|I\|^2.$$

Since $\|I\| \ge 0$, we find $\|I\| \ge 1$. To prove the second, we use the sub-multiplicativity on A and A^{-1}:

$$\|A\,A^{-1}\| = \|I\| \le \|A\|\,\|A^{-1}\| \quad \Rightarrow \quad 1 \le \|A\|\,\|A^{-1}\|.$$

Example *1.74*

We show that the 1-norm is consistent in $\mathbb{F}^{m\times n}$, $\mathbb{F}^{m\times p}$ and $\mathbb{F}^{p\times n}$, i.e., that for every $A \in \mathbb{F}^{m\times p}$ and $B \in \mathbb{F}^{p\times n}$, we have

$$\|A\,B\|_1 \le \|A\|_1\,\|B\|_1,$$

where for the sake of simplicity we have not indicated the spaces where the 1-norms are defined. To see this, using the notations

$$A = (a_{i,j})_{\substack{i=1,\dots,m\\ j=1,\dots,p}}, \quad B = (b_{i,j})_{\substack{i=1,\dots,p\\ j=1,\dots,n}}, \quad C = (c_{i,j})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$$

and $C = A\,B$, we recall that for every $i \in \{1,\dots,m\}$ and every $j \in \{1,\dots,n\}$ we have

$$c_{i,j} = \sum_{k=1}^{p} a_{i,k}\,b_{k,j}.$$

We find

$$\begin{aligned}
\|A\,B\|_1 &= \sum_{i=1}^{m}\sum_{j=1}^{n}|c_{i,j}| = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|\sum_{k=1}^{p}a_{i,k}\,b_{k,j}\right| \\
&\leq \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}|a_{i,k}\,b_{k,j}| = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}|a_{i,k}|\,|b_{k,j}| \\
&\leq \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p}|a_{i,k}|\left(\sum_{\ell=1}^{p}|b_{\ell,j}|\right) \\
&= \sum_{i=1}^{m}\sum_{k=1}^{p}|a_{i,k}|\left(\sum_{j=1}^{n}\sum_{\ell=1}^{p}|b_{\ell,j}|\right) = \|A\|_1\,\|B\|_1,
\end{aligned}$$

where we used the triangle inequality for the absolute value. This shows, in particular, that for square matrices the 1-norm is a matrix norm.

Example *1.75*

The infinity norm in the square case

$$\|A\|_\infty = \max_{i,j\in\{1,\ldots,n\}}|a_{i,j}|$$

is not a matrix norm in $\mathbb{F}^{n\times n}$ if $n > 1$. In fact, taking for example

$$B = A = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$$

gives $A\,B = A^2 = I$, so that $\|A\,B\|_\infty = 1$, but $\|A\|_\infty\,\|B\|_\infty = (\|A\|_\infty)^2 = 4$.

The following result shows that the Frobenius norm is consistent in $\mathbb{F}^{m\times n}$, $\mathbb{F}^{m\times p}\mathbb{F}^{p\times n}$.

Theorem 1.61 *For every $A \in \mathbb{F}^{m\times p}$ and $B \in \mathbb{F}^{p\times n}$, we have*

$$\|A\,B\|_F \leq \|A\|_F\,\|B\|_F.$$

In particular, the Frobenius norm is a matrix norm in $\mathbb{F}^{n\times n}$.

Proof Let $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_m$ be the rows of A and let $\boldsymbol{b}^1,\ldots,\boldsymbol{b}^n$ be the columns of B. Let $C = A\,B$. The Schwarz inequality gives

$$|c_{i,j}| = |\boldsymbol{a}_i\,\boldsymbol{b}^j| \leq |\boldsymbol{a}_i|\,|\boldsymbol{b}^j|, \forall\, i = 1,\ldots,m,\ \forall\, j = 1,\ldots,n.$$

Taking the squares of both sides, computing $|\boldsymbol{a}_i|^2$ and $|\boldsymbol{b}^j|^2$ and adding, gives

$$\sum_{i=1}^{m}\sum_{j=1}^{n}c_{i,j}^2 \leq \sum_{i=1}^{m}\sum_{k=1}^{p}a_{i,k}^2\sum_{k=1}^{p}\sum_{j=1}^{n}b_{k,j}^2.$$

1.9.1 Operator norm

An important class of matrix norms is obtained by those norms induced by a norm of vectors. These are known as *operator norms*. First, we recall that if we have a matrix $A \in \mathbb{F}^{m\times n}$, its multiplication by a vector $\boldsymbol{x} \in \mathbb{F}^n$ can be seen as a function

$$\begin{aligned}
\mathbb{F}^n &\longrightarrow \mathbb{F}^m \\
\boldsymbol{x} &\mapsto A\,\boldsymbol{x}.
\end{aligned}$$

Hence, an $m \times n$ matrix A transforms a vector in $\mathbb{F}^n$ into a vector in $\mathbb{F}^m$.

Example *1.76*

Given $p, q \in [1, +\infty]$, we denote by $\|\boldsymbol{x}\|_{n,q}$ the q-norm in $\mathbb{F}^n$ and by $\|A\,\boldsymbol{x}\|_{m,p}$ the p-norm in $\mathbb{F}^m$, so that if $q \neq +\infty$

$$\|\boldsymbol{x}\|_{n,q} = \Big(\sum_{i=1}^{n} |x_i|^q\Big)^{\frac{1}{q}}$$

and $\|\boldsymbol{x}\|_{n,\infty} = \max\limits_{i=1,\dots,n} |x_i|$ if $q = +\infty$. Likewise, if $p \neq +\infty$

$$\|A\,\boldsymbol{x}\|_{m,p} = \left\|\begin{bmatrix} \sum_{j=1}^n a_{1,j}\,x_j \\ \sum_{j=1}^n a_{2,j}\,x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j}\,x_j \end{bmatrix}\right\|_{m,p} = \Big(\sum_{i=1}^{m} \Big|\sum_{j=1}^{n} a_{i,j}\,x_j\Big|^p\Big)^{\frac{1}{p}}$$

and $\|A\,\boldsymbol{x}\|_{m,\infty} = \max\limits_{i\in\{1,\dots,m\}} \Big|\sum\limits_{j=1}^{n} a_{i,j}\,x_j\Big|$ if $p = +\infty$.

Now, we want to show that

$$\exists\, c \geq 0 : \ \forall\, \boldsymbol{x} \in \mathbb{F}^n, \quad \|A\,\boldsymbol{x}\|_{m,p} \leq c\,\|\boldsymbol{x}\|_{n,q}.$$

We recall that $\|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}}\,\|\boldsymbol{x}\|_\infty$ for $p \geq 1$. Let $p \neq +\infty$. Thus,

$$\begin{aligned}
\|A\,\boldsymbol{x}\|_{m,p} &\leq m^{\frac{1}{p}}\,\|A\,\boldsymbol{x}\|_{m,\infty} = m^{\frac{1}{p}} \left\|\begin{bmatrix} a_{1,1}\,x_1 + a_{1,2}\,x_2 + \ldots + a_{1,n}\,x_n \\ a_{2,1}\,x_1 + a_{2,2}\,x_2 + \ldots + a_{2,n}\,x_n \\ \vdots \\ a_{m,1}\,x_1 + a_{m,2}\,x_2 + \ldots + a_{m,n}\,x_n \end{bmatrix}\right\|_{m,\infty} \\
&= m^{\frac{1}{p}} \max_{i\in\{1,\dots,m\}} \Big|\sum_{j=1}^{n} a_{i,j}\,x_j\Big| \\
&\leq m^{\frac{1}{p}} \max_{i\in\{1,\dots,n\}} \sum_{j=1}^{n} |a_{i,j}|\,|x_j| \\
&\leq m^{\frac{1}{p}} \max_{i\in\{1,\dots,n\}} \sum_{j=1}^{n} |a_{i,j}|\,\|\boldsymbol{x}\|_{q,\infty} \\
&\leq \underbrace{n\,m^{\frac{1}{p}} \Big(\max_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} |a_{i,j}|\Big)}_{c}\,\|\boldsymbol{x}\|_{q,\infty}
\end{aligned}$$

The constant c only depends on the dimensions n and m and A. The case $p = +\infty$ can be addressed in the same way:

$$\|A\,\boldsymbol{x}\|_{m,\infty} = \max_{i\in\{1,\dots,m\}} \Big|\sum_{j=1}^{n} a_{i,j}\,x_j\Big| \leq \max_{i\in\{1,\dots,n\}} \sum_{j=1}^{n} |a_{i,j}|\,\|\boldsymbol{x}\|_{q,\infty} \leq n \Big(\max_{\substack{i\in\{1,\dots,m\}\\ j\in\{1,\dots,n\}}} |a_{i,j}|\Big)\,\|\boldsymbol{x}\|_{q,\infty}.$$

We can generalize the result of the previous example to the case where the norms of $\mathbb{F}^m$ and $\mathbb{F}^n$ are arbitrary (and denoted by $\|\cdot\|_m$ and $\|\cdot\|_n$, respectively), by showing that, given any matrix $A \in \mathbb{F}^{m\times n}$, there exists $c \geq 0$ such that for all $\boldsymbol{x} \in \mathbb{F}^n$ there holds $\|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n$. The function $\mathbb{F}^n \longrightarrow \mathbb{F}^m$ that maps $\boldsymbol{x}$ into $A\,\boldsymbol{x}$ is continuous, because each of its m components is linear in its arguments, and in particular it is continuous at $\boldsymbol{0}$. Choosing $\varepsilon = 1$,

$$\exists\, \delta > 0 : \ \forall\, \boldsymbol{x} \in \mathbb{F}^n, \qquad \|\boldsymbol{x}\|_n \leq \delta \ \Rightarrow\ \|A\,\boldsymbol{x}\|_m \leq 1.$$

If $\boldsymbol{x} = \boldsymbol{0}$, we can take any $c \geq 0$ and the inequality $\|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n$ will be satisfied. Assume $\boldsymbol{x} \neq \boldsymbol{0}$. We have

$$\left\|\delta \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right\|_n = \delta \leq \delta \quad \Rightarrow \quad \left\|A\left(\delta \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right)\right\| \leq 1 \quad \Rightarrow \quad \frac{\delta}{\|\boldsymbol{x}\|_n}\|A\,\boldsymbol{x}\|_m \leq 1 \quad \Rightarrow \quad \|A\,\boldsymbol{x}\|_m \leq \delta^{-1}\,\|\boldsymbol{x}\|_n.$$

We can therefore take $M = \delta^{-1}$. Thus an $m \times n$ matrix changes the vectors in $\mathbb{F}^n$ into vectors in $\mathbb{F}^m$ by a uniformly bounded amount that depends on n, m and the matrix. The constant c of the previous example may not be the smallest. In fact, it is possible, in general, to find $c' < c$ such that $\|A\,\boldsymbol{x}\|_m \leq c'\,\|\boldsymbol{x}\|_q$ for all $\boldsymbol{x} \in \mathbb{F}^n$. We will define the operator norm by assigning to each $A \in \mathbb{F}^{m\times n}$ the least $c \geq 0$ such that $\|A\,\boldsymbol{x}\|_m \leq c'\,\|\boldsymbol{x}\|_n$ for all $\boldsymbol{x} \in \mathbb{F}^n$.

Lemma 1.3 *Consider the normed spaces $(\mathbb{F}^m, \|\cdot\|_m)$ and $(\mathbb{F}^n, \|\cdot\|_n)$. Let $A \in \mathbb{F}^{m\times n}$. Let $c \geq 0$. The following statements are equivalent:*

(A) $\|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n$ *for all* $\boldsymbol{x} \in \mathbb{F}^n$*;*

(B) $\|A\,\boldsymbol{x}\|_m \leq c$ *for all* $\boldsymbol{x} \in \mathbb{F}^n$ *satisfying* $\|\boldsymbol{x}\|_n = 1$.

Proof The implication **(A)** $\Rightarrow$ **(B)** is trivial: if $\|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n$ holds for all $\boldsymbol{x} \in \mathbb{F}^n$, it holds in particular for all $\boldsymbol{x} \in \mathbb{F}^n$ satisfying also $\|\boldsymbol{x}\|_n = 1$, and in this case **(A)** reads exactly as $\|A\,\boldsymbol{x}\|_m \leq c$. We now prove **(B)** $\Rightarrow$ **(A)**. The implication is vacuously true if $\boldsymbol{x} = \boldsymbol{0}$, because in this case $\|\boldsymbol{x}\|_n = 1$ is false. Let $\boldsymbol{x} \neq \boldsymbol{0}$, and notice that

$$\left\|\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n}\right\|_n = \left|\frac{1}{\|\boldsymbol{x}\|_n}\right| \|\boldsymbol{x}\|_n = 1.$$

Thus, we can use the vector $\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n}$ for **(B)**, and we have

$$\|A\,\boldsymbol{x}\|_m = \|\boldsymbol{x}\|_n \left\|A\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n}\right\|_m \leq \|\boldsymbol{x}\|_n\, c.$$

Lemma 1.4 *Consider the normed spaces $(\mathbb{F}^m, \|\cdot\|_m)$ and $(\mathbb{F}^n, \|\cdot\|_n)$. Let $A \in \mathbb{F}^{m\times n}$. Then:*

(i) the set $\{\|A\,\boldsymbol{x}\|_m \,|\, \boldsymbol{x} \in \mathbb{F}^n,\ \ \|\boldsymbol{x}\|_n = 1\}$ *admits maximum;*

(ii) the set $\{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n, \|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n\}$ *admits minimum;*

(iii) the maximum of the first set coincides with the minimum of the second.

Proof The function

$$\begin{aligned} \mathbb{F}^n &\longrightarrow \mathbb{F}^m \\ \boldsymbol{x} &\mapsto A\,\boldsymbol{x} \end{aligned}$$

is continuous, because each component of $A\,\boldsymbol{x}$ is linear in the components of $\boldsymbol{x}$. Moreover, we proved that norms are continuous functions. The composition of these two continuous functions is the continuous function

$$\begin{aligned} f : \mathbb{F}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|A\,\boldsymbol{x}\|_m. \end{aligned}$$

The set $\{\|A\,\boldsymbol{x}\|_m \,|\, \boldsymbol{x} \in \mathbb{F}^n,\ \ \|\boldsymbol{x}\|_n = 1\}$ is therefore the image of f restricted to the unit n-dimensional hypersphere $S = \{\boldsymbol{x} \in \mathbb{F}^n \ \|\boldsymbol{x}\|_n = 1\}$, which is compact (i.e., it is bounded and closed, why?). From Weierstrass extreme value theorem, the restriction of f to S admits maximum, and this proves the first point. To prove the second and third points, let μ denote max $\{\|A\,\boldsymbol{x}\|_m \,|\, \boldsymbol{x} \in \mathbb{F}^n,\ \ \|\boldsymbol{x}\|_n = 1\}$. We prove that $\mu \in \{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n, \|A\,\boldsymbol{x}\|_m \leq c\}$.

Let $\boldsymbol{x} \in \mathbb{F}^n$. We have to show that $\|A\,\boldsymbol{x}\|_m \leq \mu$. This is equivalent to

$$\left\| A \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n} \right\| \leq \mu,$$

which is true since the norm of $\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n}$ is 1. To prove that μ is the minimum of $\{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n, \|A\,\boldsymbol{x}\|_m \leq c\}$, let $\nu \in \{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n, \|A\,\boldsymbol{x}\|_m \leq c\}$. We show that $\nu \geq \mu$. Since $\nu \in \{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n, \|A\,\boldsymbol{x}\|_m \leq c\}$, for all $\boldsymbol{x} \in \mathbb{F}^n$ we have $\|A\,\boldsymbol{x}\|_m \leq \nu\,\|\boldsymbol{x}\|_n$. Let $\boldsymbol{x} \neq \boldsymbol{0}$, so that

$$\left\| A \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_n} \right\| \leq \nu.$$

This implies $\nu \geq \mu$.

Lemma 1.4 allows us to define the operator norm of a real square matrix.

Definition 1.35 [Operator Norm]
Consider the normed spaces $(\mathbb{F}^m, \|\cdot\|_m)$ *and* $(\mathbb{F}^n, \|\cdot\|_n)$. *Let* $A \in \mathbb{F}^{m\times n}$. *We define the* operator norm *of* A *induced by the norms* $\|\cdot\|_n$ *and* $\|\cdot\|_m$ *as*

$$\|A\| \stackrel{def}{=} \min\{c \geq 0 \,|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ \|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n\}.$$

From the previous discussion, we can also write

$$\|A\| = \max_{\|\boldsymbol{x}\|_n = 1} \|A\,\boldsymbol{x}\|_m,$$

or, alternatively,

$$\|A\| = \max_{\|\boldsymbol{x}\|_n \leq 1} \|A\,\boldsymbol{x}\|_m = \max_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|A\,\boldsymbol{x}\|_m}{\|\boldsymbol{x}\|_n}.$$

Notice also that, by definition of operator norm, we have

$$\forall\, \boldsymbol{x} \in \mathbb{F}^n, \quad \|A\,\boldsymbol{x}\|_m \leq \|A\|\,\|\boldsymbol{x}\|_n.$$

Intuitively, the operator norm $\|A\|$ is the largest amount by which A stretches a vector on the unit n-hypersphere in $\mathbb{F}^n$.

Theorem 1.62 *Consider the normed spaces* $(\mathbb{F}^m, \|\cdot\|_m)$ *and* $(\mathbb{F}^n, \|\cdot\|_n)$. *The operator norm induced by the norms* $\|\cdot\|_n$ *and* $\|\cdot\|_m$ *is a norm. Moreover, for square matrix, the operator norm is a matrix norm.*

Proof We have to prove the following properties:

(i) homogeneity: $\forall\, A \in \mathbb{F}^{m\times n}, \forall\, k \in \mathbb{R}, \|k\,A\| = |k|\,\|A\|$;

(ii) sub-additivity: $\forall\, A, B \in \mathbb{F}^{m\times n}, \|A + B\| \leq \|A\| + \|B\|$;

(iii) positivity: $\forall\, A \in \mathbb{F}^{m\times n},\ \|A\| = 0 \ \Rightarrow A = \boldsymbol{0}$;

(iv) if $n = m$, we have the sub-multiplicativity: $\forall\, A, B \in \mathbb{F}^{n\times n}, \|A\,B\| \leq \|A\|\,\|B\|$.

The homogeneity is trivial. Let $k \in \mathbb{R}$. If $k = 0$, then $\|0\,A\| = 0 = |0|\,\|A\|$. Let $k \neq 0$. Then,

$$\begin{aligned}
\|k\,A\| &= \min\{c \geq 0 \mid \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ \|(k\,A)\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n\} \\
&= \min\{c \geq 0 \mid \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ \|k\,(A\,\boldsymbol{x})\|_m \leq c\,\|\boldsymbol{x}\|_n\} \\
&= \min\{c \geq 0 \mid \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ |k|\,\|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n\} \\
&= \min\left\{c \geq 0 \,\middle|\, \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ \|A\,\boldsymbol{x}\|_m \leq \frac{c}{|k|}\,\|\boldsymbol{x}\|_n\right\} \\
&= |k|\,\min\{c \geq 0 \mid \forall\, \boldsymbol{x} \in \mathbb{F}^n,\ \|A\,\boldsymbol{x}\|_m \leq c\,\|\boldsymbol{x}\|_n\} = |k|\,\|A\|.
\end{aligned}$$

We prove the sub-additivity. Given any $\boldsymbol{x} \in \mathbb{F}^n$, from

$$\begin{aligned}
\|(A + B)\,\boldsymbol{x}\|_m &= \|A\,\boldsymbol{x} + B\,\boldsymbol{x}\|_m \\
&\leq \|A\,\boldsymbol{x}\|_m + \|B\,\boldsymbol{x}\|_m \\
&\leq \|A\|\,\|\boldsymbol{x}\|_n + \|B\|\,\|\boldsymbol{x}\|_n \\
&= (\|A\| + \|B\|)\,\|\boldsymbol{x}\|_n.
\end{aligned}$$

It follows immediately that

$$\max_{\|\boldsymbol{x}\|_n = 1} \|(A + B)\,\boldsymbol{x}\|_m \leq \max_{\|\boldsymbol{x}\|_n = 1} (\|A\| + \|B\|)\,\|\boldsymbol{x}\|_n,$$

which is the same as $\|A + B\| \leq \|A\| + \|B\|$. We prove the positivity. By definition, if $\|A\| = 0$, then for all $\boldsymbol{x} \in \mathbb{F}^n$ we have

$$\|A\,\boldsymbol{x}\|_m \leq 0 \quad \Leftrightarrow \quad \|A\,\boldsymbol{x}\|_m = 0 \quad \Leftrightarrow \quad A\,\boldsymbol{x} = \boldsymbol{0}.$$

Since $A\,\boldsymbol{x} = \boldsymbol{0}$ for all $\boldsymbol{x} \in \mathbb{F}^n$, then A must be the zero matrix. Now we prove the sub-multiplicativity. Let $n = m$, and let $\boldsymbol{x} \in \mathbb{F}^n$. Then,

$$\|(A\,B)\,\boldsymbol{x}\|_n = \|A\,(B\,\boldsymbol{x})\|_n \leq \|A\|\,\|B\,\boldsymbol{x}\| \leq \|A\|\,\|B\|\,\|\boldsymbol{x}\|_n,$$

so that

$$\|A\,B\| = \max_{\|\boldsymbol{x}\|_n = 1} \|(A\,B)\,\boldsymbol{x}\|_n \leq \max_{\|\boldsymbol{x}\|_n = 1} \|A\|\,\|B\|\,\|\boldsymbol{x}\|_n = \|A\|\,\|B\|.$$

Consider the normed spaces $(\mathbb{F}^m, \|\cdot\|_m)$ and $(\mathbb{F}^n, \|\cdot\|_n)$, with some abuse of notation, given a matrix $A \in \mathbb{F}^{m\times n}$, we denote by $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ the norms induced by the 1-norm, the 2-norm and the infinity norm of $\mathbb{F}^m$ and $\mathbb{F}^n$, respectively. The following result gives a method to efficiently compute the values of these norms in these three cases. We recall to this end that, given a square real or complex matrix A, the symbol $\rho(A)$ denotes the spectral radius of A, i.e., the largest modulus of its eigenvalues.

Theorem 1.63 *Let $A \in \mathbb{F}^{m\times n}$. Then,*

$$\|A\|_1 = \max_{j\in\{1,\dots,n\}} \sum_{i=1}^{m} |a_{i,j}| \tag{1.76}$$

$$\|A\|_2 = \sqrt{\rho(A^H\,A)} \tag{1.77}$$

$$\|A\|_\infty = \max_{i\in\{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}|. \tag{1.78}$$

Proof Consider the first. With abuse of notation, we denote by $\|\cdot\|_1$ the 1-norm both in $\mathbb{F}^m$ and in $\mathbb{F}^n$. Let

$\boldsymbol{x} \in \mathbb{F}^n$ satisfy $\|\boldsymbol{x}\|_1 = 1$. Then,

$$\begin{aligned}\|A\,\boldsymbol{x}\|_1 &= \sum_{i=1}^{m} \Big| \sum_{j=1}^{n} a_{i,j}\, x_j \Big| \\ &\leq \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{i,j}|\, |x_j| = \sum_{j=1}^{n} \Big(|x_j| \sum_{i=1}^{m} |a_{i,j}| \Big) \\ &\leq \sum_{j=1}^{n} \Big(|x_j| \max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}| \Big) \\ &= \Big(\max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}| \Big) \sum_{j=1}^{n} |x_j| = \max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}|.\end{aligned}$$

This shows that

$$\|A\|_1 \leq \max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}|. \tag{1.79}$$

To conclude the proof of (1.76), it is sufficient to show that there exists $\boldsymbol{x} \in \mathbb{F}^n$ with $\|\boldsymbol{x}\|_1 = 1$ for which the equality holds. Let r be the value of k for which the maximum in the expression (1.79) is attained, so that

$$\max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}| = \sum_{i=1}^{m} |a_{i,r}|.$$

Take $\boldsymbol{x} = \boldsymbol{e}_r$, where $\boldsymbol{e}_r$ is the r-th standard basis vector in $\mathbb{F}^n$. Then,

$$\|A\,\boldsymbol{x}\|_1 = \sum_{i=1}^{m} \Big| \sum_{j=1}^{n} a_{i,j}\, x_j \Big| = \sum_{i=1}^{m} |a_{i,r}\, x_r| = \sum_{i=1}^{m} |a_{i,r}|\, |x_r| = \sum_{i=1}^{m} |a_{i,r}| = \max_{k \in \{1,\ldots,n\}} \sum_{i=1}^{m} |a_{i,k}| = \|A\|_1.$$

Consider (1.77). Let $\boldsymbol{x} \in \mathbb{F}^n$ satisfy $\|\boldsymbol{x}\|_2 = 1$. Consider the complex case generality. Recalling that for a complex number z we have the identity $|z|^2 = \bar{z}\, z$, for a complex vector $\boldsymbol{z} \in \mathbb{C}^n$ we have

$$\|\boldsymbol{z}\|_2^2 = \boldsymbol{z}^*\, \boldsymbol{z},$$

where $\boldsymbol{z}^\top$ is the conjugate transpose of $\boldsymbol{z}$, and in fact the left hand-side is $|z_1|^2 + |z_2|^2 + \ldots + |z_n|^2$, and the right hand-side

$$\boldsymbol{z}^*\, \boldsymbol{z} = \begin{bmatrix} \bar{z}_1 & \bar{z}_2 & \ldots & \bar{z}_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \bar{z}_1\, z_1 + \bar{z}_2\, z_2 + \ldots + \bar{z}_n\, z_n$$

coincides with it. Therefore, we can write

$$\|A\,\boldsymbol{x}\|_2^2 = \boldsymbol{x}^*\, A^*\, A\, \boldsymbol{x}.$$

The matrix $A^*\, A$ is Hermitian, and therefore it is unitarily diagonalizable. In other words, there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ (i.e., a matrix satisfying the equality $U^{-1} = U^*$) such that $U^*\, (A^*\, A)\, U = \Delta$, where Δ is a diagonal matrix with the (real and non-negative) eigenvalues of $A^*\, A$ on its diagonal. We can define $\boldsymbol{y} = U^*\, \boldsymbol{x}$ (which is the same as $\boldsymbol{x} = U\, \boldsymbol{y}$), and

$$\|A\,\boldsymbol{x}\|_2^2 = \boldsymbol{y}^*\, U^*\, A^*\, A\, U\, \boldsymbol{y} = \boldsymbol{y}^*\, \Delta\, \boldsymbol{y} = \sum_{i=1}^{n} \lambda_i\, |y_i|^2 \leq \rho(A^*\, A) \sum_{i=1}^{n} \lambda_i\, |y_i|^2 = \rho(A^*\, A).$$

Indeed, $\|\boldsymbol{y}\|_2 = \|U^* \boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^* U^* U \boldsymbol{x}} = \sqrt{\boldsymbol{x}^* \boldsymbol{x}} = \|\boldsymbol{x}\|_2 = 1$. Notice that this is equivalent to saying that $\|A \boldsymbol{x}\|_2$ is the largest singular value σ_1 of A. We have proved that $\|A \boldsymbol{x}\|_2 \leq \sigma_1$. We now show that there exists $\boldsymbol{x} \in \mathbb{C}^n$ for which such inequality becomes an equality. It suffices to choose $\boldsymbol{x}$ to be the eigenvector of A^*A which corresponds to σ_1:

$$\|A \boldsymbol{x}\|_2 = \boldsymbol{x}^* A^* A \boldsymbol{x} = \boldsymbol{x}^* \sigma_1 \boldsymbol{x} = \sigma_1 (\boldsymbol{x}^* \boldsymbol{x}) = \sigma_1 \|\boldsymbol{x}\|_2 = \sigma_1.$$

This concludes the proof of the second point. We now focus on (1.78). Assume $\|\boldsymbol{x}\|_\infty = 1$, and so in particular $|x_j| \leq 1$ for all $j \in \{1, \dots, n\}$. We have

$$\begin{aligned} \|A \boldsymbol{x}\|_\infty &= \max_{i \in \{1,\dots,m\}} \left| \sum_{j=1}^{n} a_{i,j} x_j \right| \\ &\leq \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}| |x_j| \leq \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}|. \end{aligned}$$

This shows that

$$\|A\|_\infty \leq \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}|. \tag{1.80}$$

To show that the equality holds, we construct $\boldsymbol{x}$ satisfying $\|\boldsymbol{x}\|_\infty = 1$ for which the equality holds. Suppose the maximum in (1.80) is attained at $i = k$, so that

$$\max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}| = \sum_{j=1}^{n} |a_{k,j}|.$$

We define the j-th component of $\boldsymbol{x}$ as

$$x_j = \begin{cases} \frac{\bar{a}_{k,j}}{|a_{k,j}|} & \text{if } a_{k,j} \neq 0 \\ 1 & \text{otherwise,} \end{cases}$$

so that the k-th component of $A \boldsymbol{x}$ is

$$\sum_{j=1}^{n} a_{k,j} x_j = \sum_{\substack{j=1 \\ a_{k,j} \neq 0}}^{n} a_{k,j} \frac{\bar{a}_{k,j}}{|a_{k,j}|} = \sum_{\substack{j=1 \\ a_{k,j} \neq 0}}^{n} |a_{k,j}| = \sum_{j=1}^{n} |a_{k,j}| = \max_{i \in \{1,\dots,m\}} \sum_{j=1}^{n} |a_{i,j}|,$$

as required.

Example *1.77*

We want to find the induced norms $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$ of the matrix

$$A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & -1 & 1 \end{bmatrix}.$$

The 1-norm is

$$\begin{aligned} \|A\|_1 &= \max_{j \in \{1,2,3\}} \sum_{i=1}^{2} |a_{i,j}| = \max\{|a_{1,j}| + |a_{2,j}| \mid j = 1, 2, 3\} \\ &= \max\{|a_{1,1}| + |a_{2,1}|, |a_{1,2}| + |a_{2,2}|, |a_{1,3}| + |a_{2,3}|\} = \max\{1, 1, 3\} = 3. \end{aligned}$$

To find the induced 2-norm, we find the eigenvalues of $A^\top A$:

$$\det(A^\top A - \lambda I) = \begin{vmatrix} 1-\lambda & 0 & -2 \\ 0 & 1-\lambda & -1 \\ -2 & -1 & 5-\lambda \end{vmatrix} = \lambda(1-\lambda)(\lambda-6).$$

The largest singular value of A is therefore $\sigma_1 = \sqrt{6}$, so that $\|A\|_2 = \sqrt{6}$.

Finally,

$$\begin{aligned} \|A\|_\infty &= \max_{i=1,2} \sum_{j=1}^{3} |a_{i,j}| = \max\{|a_{i,1}| + |a_{i,2}| + |a_{i,3}| \mid i = 1, 2\} \\ &= \max\{|a_{1,1}| + |a_{1,2}| + |a_{1,3}|, |a_{2,1}| + |a_{2,2}| + |a_{2,3}|\} = \max\{3, 2\} = 3. \end{aligned}$$

2 Banach and Hilbert spaces

Hilbert spaces arise frequently in mathematics, engineering and physics, typically as spaces of functions. We can say that the theory of Hilbert spaces is concerned with the generalization of the methods of linear algebra and calculus from finite-dimensional vector spaces to possibly infinite-dimensional vector spaces. In fact, a Hilbert space can be seen as a generalization of the notion of Euclidean space. Hilbert spaces are vector spaces for which some fundamental properties of Euclidean spaces are preserved. For example, analogs of the Pythagorean theorem and parallelogram law hold in a Hilbert space. The notion of orthogonal projection onto a subspace can also be defined for Hilbert spaces: this plays a crucial role in the use of Hilbert spaces as the framework to define and solve optimization problems. Moreover, a vector of a Hilbert space can be uniquely specified by its coordinates with respect to an orthonormal basis or an orthonormal system of vectors, in analogy with Cartesian coordinates for Euclidean spaces. We will see that a Hilbert space is a vector space over which an inner product is defined, so that we have notions of norm, distance, angles and orthogonality, and where we have a property called *completeness*: every Cauchy sequence is convergent.

2.1 Cauchy sequences

A sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in a normed space is a Cauchy sequence if the distance between the vectors $\boldsymbol{x}_n$ and $\boldsymbol{x}_m$ is arbitrarily small, provided that the indices n and m are sufficiently large.

Definition 2.1 [Cauchy Sequence]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence in $\boldsymbol{X}$. We say that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence *if*

$$\boxed{\forall\,\varepsilon>0,\ \exists\, p\in\mathbb{N}:\ \forall\, n,m\in\mathbb{N},\qquad n,m>p\ \Rightarrow\ \|\boldsymbol{x}_n-\boldsymbol{x}_m\|<\varepsilon} \tag{2.1}$$

It is clear that there is a connection between Cauchy sequences and convergent sequences: intuitively speaking, if, by taking k to be sufficiently large, the term $\boldsymbol{x}_k$ can be made arbitrarily close to the limit $\hat{\boldsymbol{x}}$, by taking n and m sufficiently large two terms $\boldsymbol{x}_n$ and $\boldsymbol{x}_m$ can also be made arbitrarily close to the same limit, and therefore they can be made arbitrarily close to each other. However, the relationship between convergent sequences and Cauchy sequences is a very delicate point, especially in the context of infinite dimensional vector spaces, and requires a closer investigation. The following results will show that in any normed vector space, every convergent sequence is a Cauchy sequence. Moreover, the Cauchy's condition does not change if we change the norm to an equivalent norm.

Theorem 2.1 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence in $\boldsymbol{X}$. If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent, it is a Cauchy sequence.*

Proof Let $\hat{\boldsymbol{x}}\in\boldsymbol{X}$ be such that $\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} 0$. Let $\varepsilon>0$ and $p\in\mathbb{N}$ be such that $\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\|<\frac{\varepsilon}{2}$ if $k>p$. If we

take $n, m \in \mathbb{N}$ both strictly greater than p, we have

$$\begin{aligned} \|\boldsymbol{x}_n - \boldsymbol{x}_m\| &= \|(\boldsymbol{x}_n - \hat{\boldsymbol{x}}) + (\hat{\boldsymbol{x}} - \boldsymbol{x}_m)\| \\ &\leq \|\boldsymbol{x}_n - \hat{\boldsymbol{x}}\| + \|\hat{\boldsymbol{x}} - \boldsymbol{x}_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

The converse of Theorem 2.1 does not hold in general: in a normed space, there may exist Cauchy sequences which are not convergent. More specifically, the Cauchy's condition is sufficient for the convergence in the case of a finite-dimensional normed vector space.

The first step in this analysis is to show that the converse of Theorem 2.1 is always true for a sequence of real numbers. In this case, a Cauchy sequence is always convergent. One advantage of using Cauchy's condition (2.1) to test for convergence is the fact that it does not involve the knowledge of the limit of the sequence, unlike the definition of convergence sequence.

Theorem 2.2 *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathbb{R}$. If $(x_k)_{k\in\mathbb{N}}$ a Cauchy sequence, then it is convergent.*

Proof Let $(x_k)_{k\in\mathbb{N}}$ be a (real-valued) Cauchy sequence. First, we prove that $(x_k)_{k\in\mathbb{N}}$ is bounded. Since $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, there exists $p \in \mathbb{N}$ such that for all $n, m \in \mathbb{N}$ with $n, m > p$ there holds $|x_n - x_m| < 1$. Choosing $m = p + 1$, we have $|x_n - x_{p+1}| < 1$ for all $n > p$. Thus, for all $n > p$ we have

$$\begin{aligned} |x_n| &\leq |x_n - x_{p+1}| + |x_{p+1}| \\ &< 1 + |x_{p+1}|. \end{aligned}$$

Let

$$M = \max\{|x_0|, |x_1|, \ldots, |x_p|, 1 + |x_{p+1}|\},$$

so that $|x_k| \leq M$ for all $k \in \mathbb{N}$. Thus, $(x_k)_{k\in\mathbb{N}}$ is bounded, and therefore from Bolzano-Weierstrass theorem (Theorem 1.41) there exists a subsequence $(x_{k_h})_{h\in\mathbb{N}}$ of $(x_k)_{k\in\mathbb{N}}$ which is convergent. Let $\ell = \lim_{h\to\infty} x_{k_h}$. We prove that $(x_k)_{k\in\mathbb{N}}$ converges to ℓ as $k \to \infty$. Let $\varepsilon > 0$. We prove that there exists $p \in \mathbb{N}$ such that for all $k \in \mathbb{N}$ with $k > p$ we have $|x_k - \ell| < \varepsilon$. Since $(x_k)_{k\in\mathbb{N}}$ is a Cauchy sequence,

$$\exists\, p \in \mathbb{N} : \quad \forall\, n, m \in \mathbb{N}, \qquad n, m > p \;\Rightarrow |x_n - x_m| < \frac{\varepsilon}{2}. \tag{2.2}$$

Let $k > p$. We prove that $|x_k - \ell| < \varepsilon$. Since $(k_h)_{h\in\mathbb{N}}$ is strictly increasing, k_h tends to $+\infty$ as $h \to \infty$, so that

$$\exists\, q \in \mathbb{N} : \quad \forall\, h \in \mathbb{N}, \qquad h > q \;\Rightarrow k_h > p. \tag{2.3}$$

Since $(x_{k_h})_{h\in\mathbb{N}}$ converges to ℓ as $h \to \infty$,

$$\exists\, r \in \mathbb{N} : \quad \forall\, h \in \mathbb{N}, \qquad h > r \;\Rightarrow |x_{k_h} - \ell| < \frac{\varepsilon}{2}. \tag{2.4}$$

Let us choose $h > \max\{q, r\}$. From (2.3) we obtain $k_h > p$ and from (2.4) we obtain $|x_{k_h} - \ell| < \frac{\varepsilon}{2}$. We now can use (2.2) with $n = k$ and $m = k_h$: for $k > p$ and this choice of h (which guarantees that $k_h > p$) we have $|x_k - x_{k_h}| < \frac{\varepsilon}{2}$. Hence, using the triangle inequality we find

$$|x_k - \ell| \leq |x_k - x_{k_h}| + |x_{k_h} - \ell| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus, for $k > p$, we obtain $|x_k - \ell| < \varepsilon$, which ensures that $(x_k)_{k\in\mathbb{N}}$ converges to ℓ as $k \to \infty$. Hence, the Cauchy sequence $(x_k)_{k\in\mathbb{N}}$ is convergent.

The proof of Theorem 2.2 can be straightforwardly adapted to the case of a sequence of $\mathbb{R}^n$ paired with an arbitrary norm. However, we now show that this result is even more general: it applies to any finite-dimensional

normed space.

Theorem 2.3 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a finite-dimensional normed space over a field $\mathbb{F}$ (either $\mathbb{R}$ or $\mathbb{C}$). Then, every Cauchy sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of $\boldsymbol{X}$ is convergent.*

Proof Let q denote the dimension of $\boldsymbol{X}$. Let $\mathcal{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_q)$ be a basis of $\boldsymbol{X}$. We express every vector $\boldsymbol{\xi}$ in $\boldsymbol{X}$ through its coordinates $\xi_1, \xi_2, \ldots, \xi_q$ with respect to this basis, i.e, we write $\boldsymbol{\xi}$ as $\boldsymbol{\xi} = \xi_1\,\boldsymbol{b}_1 + \xi_2\,\boldsymbol{b}_2 + \ldots + \xi_q\,\boldsymbol{b}_q$, where $\xi_1, \ldots, \xi_q \in \mathbb{F}$ are uniquely determined. We can define the norm

$$\|\boldsymbol{\xi}\|_0 = |\xi_1| + |\xi_2| + \ldots + |\xi_q|. \tag{2.5}$$

Notice that for all $i \in \{1, \ldots, q\}$ we have $|\xi_i| \leq \|\boldsymbol{\xi}\|_0$. Moreover, we know that in a finite-dimensional normed space all norms are equivalent, so that the given norm $\|\cdot\|$ is equivalent to $\|\cdot\|_0$ introduced in (2.5). Thus, there exists $c > 0$ such that for all $\boldsymbol{\xi} \in \boldsymbol{X}$ we have $\|\boldsymbol{\xi}\|_0 \leq c\,\|\boldsymbol{\xi}\|$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a Cauchy sequence of elements of $\boldsymbol{X}$. For all $k \in \mathbb{N}$, we can write

$$\boldsymbol{x}_k = x_{k,1}\,\boldsymbol{b}_1 + x_{k,2}\,\boldsymbol{b}_2 + \ldots + x_{k,q}\,\boldsymbol{b}_q,$$

where, for all $k \in \mathbb{N}$ and $i \in \{1, 2, \ldots, q\}$, the coefficients $x_{k,i}$ are elements of $\mathbb{F}$. This equation defines q sequences $(x_{k,1})_{k\in\mathbb{N}}$, $(x_{k,2})_{k\in\mathbb{N}}$, …, $(x_{k,q})_{k\in\mathbb{N}}$. We begin by showing that each one of these sequences is a Cauchy sequence. Let $i \in \{1, 2, \ldots, q\}$. We have to show that, for all $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that, for $n, m > p$ we have $|x_{n,i} - x_{m,i}| < \varepsilon$. Let $\varepsilon > 0$. Since $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, there exists $p \in \mathbb{N}$ such that for $n, m > p$ we have $\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \frac{\varepsilon}{c}$. On the other hand, for all $i \in \{1, \ldots, q\}$ and for all $n, m > p$ we have

$$|x_{n,i} - x_{m,i}| \leq |x_{n,1} - x_{m,1}| + |x_{n,2} - x_{m,2}| + \ldots + |x_{n,q} - x_m, q| = \|\boldsymbol{x}_n - \boldsymbol{x}_m\|_0 \leq c\,\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < c\,\frac{\varepsilon}{c} = \varepsilon,$$

which shows that for all $i \in \{1, \ldots, q\}$ the sequence $(x_{k,i})_{k\in\mathbb{N}}$ is a Cauchy sequence. Since $(x_{k,i})_{k\in\mathbb{N}}$ is a sequence of elements in $\mathbb{F}$, and $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$, the sequence $(x_{k,i})_{k\in\mathbb{N}}$ is therefore convergent (see Theorem 2.2 and the previous exercise). Let for all $i \in \{1, \ldots, q\}$ denote by ℓ_i its limit, i.e., for all $i \in \{1, 2, \ldots, q\}$

$$\lim_{k\to\infty} x_{k,i} = \ell_i \in \mathbb{F}.$$

Let us now define $\hat{\boldsymbol{x}} = \ell_1\,\boldsymbol{b}_1 + \ell_2\,\boldsymbol{b}_2 + \ldots + \ell_q\,\boldsymbol{b}_q$, so that, using the norm $\|\cdot\|_0$,

$$\begin{aligned}\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|_0 &= \|(x_{k,1}\,\boldsymbol{b}_1 + x_{k,2}\,\boldsymbol{b}_2 + \ldots + x_{k,q}\,\boldsymbol{b}_q) - (\ell_1\,\boldsymbol{b}_1 + \ell_2\,\boldsymbol{b}_2 + \ldots + \ell_q\,\boldsymbol{b}_q)\|_0 \\ &= \|(x_{k,1} - \ell_1)\,\boldsymbol{b}_1 + (x_{k,2} - \ell_2)\,\boldsymbol{b}_2 + \ldots + (x_{k,q} - \ell_q)\,\boldsymbol{b}_q\|_0 = |x_{k,1} - \ell_1| + |x_{k,2} - \ell_2| + \ldots + |x_{k,q} - \ell_q| \overset{k\to\infty}{\longrightarrow} 0.\end{aligned}$$

Since the convergence does not change from a norm to an equivalent one, we also have convergence with respect to the norm $\|\cdot\|$, and we can write

$$\lim_{k\to\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}$$

in the normed space $(\boldsymbol{X}, \|\cdot\|)$. Thus, $(x_k)_{k\in\mathbb{N}}$ is convergent in $(\boldsymbol{X}, \|\cdot\|)$.

Some properties of sets can also be characterized in terms of sequences.

Theorem 2.4 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed vector space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{A}$ is closed if and only if, given any convergent sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ with values in $\boldsymbol{A}$, its limit belongs to $\boldsymbol{A}$.*

Proof **(Only if)**. Let $\boldsymbol{A}$ be a closed set and let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence taking values in $\boldsymbol{A}$. Let us denote by $\hat{\boldsymbol{x}}$ its limit. We need to prove that $\hat{\boldsymbol{x}} \in \boldsymbol{A}$. Suppose by contradiction that $\hat{\boldsymbol{x}} \notin \boldsymbol{A}$, i.e., $\hat{\boldsymbol{x}} \in \boldsymbol{X} \setminus \boldsymbol{A}$. Since $\boldsymbol{X} \setminus \boldsymbol{A}$ is open, there exists $r > 0$ such that $\mathcal{B}_r(\hat{\boldsymbol{x}}) \subseteq \boldsymbol{X} \setminus \boldsymbol{A}$, i.e., it does not intersect A. Then, none of the $\boldsymbol{x}_k \in \boldsymbol{A}$ belongs

to $\mathcal{B}_r(\hat{\boldsymbol{x}})$, and this implies that the sequence does not converge to $\hat{\boldsymbol{x}}$.

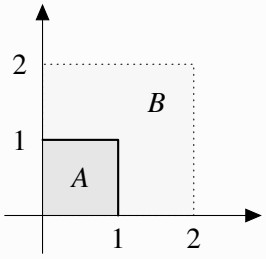
(If). In this part of the proof, we reason using a *contrapositive* argument. If we have two propositions p and q, recall that $p \Rightarrow q$ is logically equivalent to $\neg q \Rightarrow \neg p$ (where $\neg$ denotes the logical negation); this property of the implication is referred to as *modus tollens*. If p is the proposition "any convergent sequence of $\boldsymbol{A}$ tends to an element of $\boldsymbol{A}$" and q is the proposition "$\boldsymbol{A}$ is closed", proving $p \Rightarrow q$ is the same as proving that $\neg q \Rightarrow \neg p$. Hence, suppose that $\boldsymbol{A}$ is not closed, and we construct $\hat{\boldsymbol{x}} \notin \boldsymbol{A}$ which is the limit of a sequence with values in $\boldsymbol{A}$. We choose, as $\hat{\boldsymbol{x}}$, a point not in $\boldsymbol{A}$, but which is an accumulation point for $\boldsymbol{A}$: this is always possible because $\boldsymbol{A}$ is not closed, and therefore does not contain all its limit points (see Theorem 1.12). For every $k \in \mathbb{N}^*$, consider the ball $\mathcal{B}_{\frac{1}{k}}(\hat{\boldsymbol{x}})$: this ball contains at least a point $\boldsymbol{x}_k \in \boldsymbol{A}$ (obviously distinct from $\hat{\boldsymbol{x}}$ since $\hat{\boldsymbol{x}} \notin \boldsymbol{A}$ whereas $\boldsymbol{x}_k \in \boldsymbol{A}$). In this way, we have constructed a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points of $\boldsymbol{A}$ which converges to $\hat{\boldsymbol{x}}$ since $\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| < 1/k$, and $\hat{\boldsymbol{x}} \notin \boldsymbol{A}$.

Definition 2.2 *Given a normed space $(\boldsymbol{X}, \|\cdot\|)$, and $\boldsymbol{A} \subseteq \mathbf{B} \subseteq \boldsymbol{X}$, we say that $\boldsymbol{A}$ is closed in $\mathbf{B}$ if*

$$\overline{\boldsymbol{A}} \cap \mathbf{B} = \boldsymbol{A}.$$

Example *2.1*

Let $B = (0,2) \times (0,2)$ and let $A = (0,1] \times (0,1]$ in $\mathbb{R}^2$. The set A is not closed, but it is closed in B. In fact, $\overline{A} = [0,1] \times [0,1]$ and $B \cap \overline{A} = (0,1] \times (0,1] = A$.



Exercise 2.1

Adapting the proof of Theorem 2.4, show that $\boldsymbol{A}$ is closed in $\mathbf{B}$ if and only if given any sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points of $\boldsymbol{A}$ and convergent to a point of $\mathbf{B}$, its limit belongs to $\boldsymbol{A}$. [Solution: (2.1)]

It is possible to show that $\boldsymbol{A}$ is closed in $\mathbf{B}$ if and only if $\boldsymbol{A}$ contains all the adherent points of $\boldsymbol{A}$ that also belong to $\mathbf{B}$. We can also define $\boldsymbol{A}$ to be open in $\mathbf{B}$ if $\mathring{\boldsymbol{A}} \cap \mathbf{B} = \boldsymbol{A}$. It is easy to show that this condition is equivalent to saying that for all $\boldsymbol{x} \in \boldsymbol{A}$ there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \cap \mathbf{B} \subseteq \boldsymbol{A}$. These notions of relative closure and relative openness are central in the theory of relative topology.

Theorem 2.5 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed vector space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{A}$ is dense in $\boldsymbol{X}$ if and only if every element of $\boldsymbol{X}$ is the limit of a sequence with values in $\boldsymbol{A}$.*

Proof We recall that $\boldsymbol{A}$ is dense in $\boldsymbol{X}$ if the closure of $\boldsymbol{A}$ coincides with $\boldsymbol{X}$. Therefore, the result follows from Theorem 2.4.

In Euclidean spaces, every subspace (and affine subspace) is a closed set: elementary examples are the lines and planes in $\mathbb{R}^3$. This property is **false** is infinite-dimensional spaces:

Theorem 2.6 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{U}$ be a subspace of $\boldsymbol{X}$. If $\boldsymbol{U}$ is finite dimensional, then it is closed.*

Proof To prove that $\boldsymbol{U}$ is closed, we use Theorem 2.4: we take a convergent sequence $(\boldsymbol{u}_k)_{k\in\mathbb{N}}$ of elements in $\boldsymbol{U}$, and we show that its limit $\hat{\boldsymbol{u}} \in \boldsymbol{X}$ is also an element of $\boldsymbol{U}$. Let q be the dimension of $\boldsymbol{U}$, and let $\mathcal{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_q)$ be a basis of $\boldsymbol{U}$. For every $\xi \in \boldsymbol{U}$, we define the usual norm

$$\|\boldsymbol{\xi}\|_0 = |\xi_1| + |\xi_2| + \ldots + |\xi_q|$$

in $\boldsymbol{U}$, where ξ_i is the i-th coordinate of $\boldsymbol{\xi} \in \boldsymbol{U}$ with respect to $\mathcal{B}$. Since the given norm $\|\cdot\|$ is also a norm in $\boldsymbol{U}$ (or, more precisely, a norm in the restriction of $\boldsymbol{X}$ to $\boldsymbol{U}$) and $\boldsymbol{U}$ is finite-dimensional, the two norms $\|\cdot\|$ and $\|\cdot\|_0$ are equivalent, and therefore there exists a constant $c > 0$ such that

$$\forall\, \boldsymbol{\xi} \in \boldsymbol{U}, \quad \|\boldsymbol{\xi}\|_0 \le c\, \|\boldsymbol{\xi}\|.$$

We deduce that

$$\forall\, \boldsymbol{\xi} \in \boldsymbol{U},\ \forall\, i \in \{1, \ldots, q\}, \qquad |\xi_i| \le c\, \|\boldsymbol{\xi}\|.$$

We use this result to show that, if we denote by $u_{k,i}$ the i-th coordinate of $\boldsymbol{u}_k$ with respect to $\mathcal{B}$ (so that $\boldsymbol{u}_k = u_{k,1}\, \boldsymbol{b}_1 + \ldots + u_{k,q}\, \boldsymbol{b}_q$), for all $i \in \{1, \ldots, q\}$, the sequence $(u_{k,i})_{k\in\mathbb{N}}$ is a Cauchy sequence. Let $i \in \{1, \ldots, q\}$. We have to show that for all $\varepsilon > 0$, there exists $p \in \mathbb{R}$ such that, for all $n, m > p$ we have $|u_{n,i} - u_{m,i}| < \varepsilon$. Let $\varepsilon > 0$. Since $(\boldsymbol{u}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{u}}$, there exists $p \in \mathbb{N}$ such that for every $n > p$ we have $\|\boldsymbol{u}_n - \hat{\boldsymbol{u}}\| < \frac{\varepsilon}{2c}$, and, likewise, for every $m > p$ we have $\|\hat{\boldsymbol{u}} - \boldsymbol{u}_m\| < \frac{\varepsilon}{2c}$. It follows that for every $n, m > p$ we have

$$\begin{aligned} |u_{n,i} - u_{m,i}| &\le |u_{n,1} - u_{m,1}| + \ldots + |u_{n,q} - u_{m,q}| = \|\boldsymbol{u}_n - \boldsymbol{u}_m\|_0 \le c\, \|\boldsymbol{u}_n - \boldsymbol{u}_m\| \\ &\le c\, (\|\boldsymbol{u}_n - \hat{\boldsymbol{u}}\| + \|\hat{\boldsymbol{u}} - \boldsymbol{u}_m\|) < c \left(\frac{\varepsilon}{2c} + \frac{\varepsilon}{2c}\right) = \varepsilon. \end{aligned}$$

This shows that, for all $i \in \{1, \ldots, q\}$, the sequence of scalars $(u_{k,i})_{k\in\mathbb{N}}$ is a Cauchy sequence, and therefore it is convergent. Let ℓ_i be its limit, and consider the vector $\hat{\boldsymbol{u}}' = \ell_1\, \boldsymbol{b}_1 + \ell_2\, \boldsymbol{b}_2 + \ldots + \ell_q\, \boldsymbol{b}_q$, which is a vector of $\boldsymbol{U}$ because it is given by a linear combination of the vectors of a basis of $\boldsymbol{U}$. It is easy to see that $\hat{\boldsymbol{u}}'$ is the limit of the given sequence. To see this, we can use the norm $\|\cdot\|_0$:

$$\begin{aligned} \|\boldsymbol{u}_k - \hat{\boldsymbol{u}}'\|_0 &= \|(u_{k,1} - \ell_1)\, \boldsymbol{b}_1 + (u_{k,2} - \ell_2)\, \boldsymbol{b}_2 + \ldots + (u_{k,q} - \ell_q)\, \boldsymbol{b}_q\|_0 \\ &= |u_{k,1} - \ell_1| + |u_{k,2} - \ell_2| + \ldots + |u_{k,q} - \ell_q| \xrightarrow{k\to\infty} 0. \end{aligned}$$

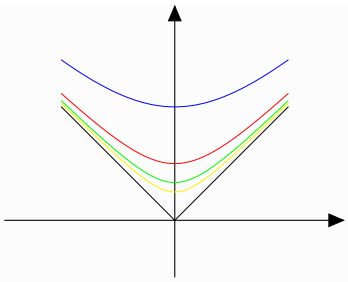
Since the limit does not change by considering equivalent norms, we have also $\|\boldsymbol{u}_k - \hat{\boldsymbol{u}}'\| \xrightarrow{k\to\infty} 0$. On the other hand, the limit is unique, so that $\hat{\boldsymbol{u}} = \hat{\boldsymbol{u}}'$. Since $\hat{\boldsymbol{u}}' \in \boldsymbol{U}$, then we have also $\hat{\boldsymbol{u}} \in \boldsymbol{U}$, and the statement is proved.

The assumptions of Theorem 2.6 are automatically satisfied if the vector space $\boldsymbol{X}$ is finite dimensional. If $\boldsymbol{X}$ is infinite-dimensional, its finite-dimensional subspaces are closed, and its infinite-dimensional subspaces can be closed or not. For example, $\boldsymbol{X}$ is a subspace of itself, and it is therefore closed. However, the following example shows that a generic subspace may not be closed.

Example 2.2 Consider the infinite dimensional normed space $(C([-1, 1]; \mathbb{R}), \|\cdot\|_\infty)$. Consider the subspace $\boldsymbol{U} = C^1([-1, 1]; \mathbb{R})$ of functions of class C^1 with domain $[-1, 1]$ taking real values. We show that $\boldsymbol{U}$ is not closed, by constructing an element $f \in \boldsymbol{X}$ that does not belong to $\boldsymbol{U}$, but is the limit of elements of $\boldsymbol{U}$. Consider

$$f(x) = |x| \quad \text{and} \quad f_k(x) = \sqrt{\frac{1}{k^2} + x^2}, \qquad x \in [-1, 1].$$

Clearly, $f \notin \boldsymbol{U}$, while $f_k \in \boldsymbol{U}$ for all $k \in \mathbb{N}^*$, so that $(f_k)_{k\in\mathbb{N}^*}$ is a sequence taking values in $\boldsymbol{U}$.



The maximum of $f_k - f$ is attained at $x = 0$ (to see this, let $\delta(x) = f_k(x) - f(x)$, and notice that for $x \geq 0$ we have $\delta'(x) = \frac{2x}{2\sqrt{\frac{1}{k^2}+x^2}} - 1 < 0$. Hence, $\delta(x)$ is decreasing, and it reaches its maximum at 0). Since $f_k(x) \geq f(x)$ for all $x \in [-1, 1]$, we have

$$\|f_k - f\|_\infty = \max_{x\in[-1,1]} |f_k(x) - f(x)| = f_k(0) - f(0) = \frac{1}{k},$$

so that f_k converges to f in $\boldsymbol{X}$. It is also possible to prove that $\boldsymbol{U}$ is dense in $\boldsymbol{X}$, i.e., every continuous function in $[-1, 1]$ is the uniform limit of a sequence of functions of class C^1.

Exercises

Exercise 2.2 Use the sequence $(x_n)_{n\in\mathbb{N}^*}$ given by $x_n = \ln(n)$ to show that the definition of Cauchy sequence is **not** equivalent to:

$$\forall\, \varepsilon > 0, \exists N > 0 : \quad \forall\, n > N,\ |x_{n+1} - x_n| < \varepsilon.$$

[Solution: (2.2)]

Exercise 2.3 Use the result on real-valued sequences (Theorem 2.2) to show that, given a sequence $(z_k)_{k\in\mathbb{N}}$ of elements of $\mathbb{C}$, if $(z_k)_{k\in\mathbb{N}}$ a Cauchy sequence, then it is convergent. [Solution: (2.3)]

Exercise 2.4 Using the Cauchy's condition, show that the sequences $(k^2)_{k\in\mathbb{N}}$ and $((-1)^k)_{k\in\mathbb{N}}$ are not Cauchy sequences. [Solution: (2.4)]

Exercise 2.5 Consider the normed space $(C([0, 1], \mathbb{R}), \|\cdot\|_\infty)$. Check whether the following subsets of the vector space $C([0, 1], \mathbb{R})$ are closed:

(i) the set of all functions f that are zero at 0;

(ii) the set of all non-negative functions f;

(iii) the set of all functions f such that $\int_0^1 f(x)\,\mathrm{d}x = 0$;

(iv) the set of all functions f such that $\int_0^1 f(x)\,\mathrm{d}x \geq 0$;

(v) the set of all functions f that are differentiable at $1/2$ and $f'(1/2) = 0$;

(vi) the set of all functions f that are differentiable in $(0, 1)$ and such that f' is identically zero;

(vii) the set of all functions f which are polynomials of degree ≤ 2;

(viii) the set of all functions f which are polynomials of degree equal to 2;

(ix) the set of all polynomial functions f;

(x) the set of all functions f such that $f(0) = f(1)$.

Hint: remember that the convergence in the normed space $(C([0,1],\mathbb{R}),\|\cdot\|_\infty)$ is the uniform convergence. Remember also that the assumption of uniform convergence allows the interchange of the limit with the integral, i.e., if $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f, then

$$\lim_{k\to\infty}\int_a^b f_k(x)\,\mathrm{d}x = \int_a^b \lim_{k\to\infty} f_k(x)\,\mathrm{d}x.$$

[Solution: (2.5)]

Exercise 2.6 Consider the normed space $(C([0,1],\mathbb{R}),\|\cdot\|_\infty)$. Consider the subset of $C([0,1],\mathbb{R})$ of decreasing functions. Is it a subspace? Is it a closed set? [Solution: (2.6)]

Exercise 2.7 Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two equivalent norms for the vector space $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of $\boldsymbol{X}$. Show that if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_1$, it is a Cauchy sequence also with respect to $\|\cdot\|_2$. [Solution: (2.7)]

2.2 Compact sets

Definition 2.3 [Sequentially Compact Set]
Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space. Let $\boldsymbol{K}\subseteq\boldsymbol{X}$. We say that $\boldsymbol{K}$ is sequentially compact *if every sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in $\boldsymbol{K}$ has a convergent subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ which converges to a point in $\boldsymbol{K}$.*

Example 2.3 The space $\mathbb{R}$ is not compact. In fact, the sequence $(x_k)_{k\in\mathbb{N}}$ given by $x_k = k$ is divergent, and every subsequence is divergent, so that the limit is not in $\mathbb{R}$.

Example 2.4 Let $a,b\in\mathbb{R}$ with $a<b$. The closed interval $[a,b]$ is compact. Indeed, let $(x_k)_{k\in\mathbb{N}}$ be a sequence of points in $[a,b]$. Bolzano-Weierstrass Theorem (Theorem 1.41) guarantees that there exists a convergent subsequence $(x_{k_j})_{j\in\mathbb{N}}$. Its limit is also in $[a,b]$, because $[a,b]$, being closed, contains all its limit points.

Exercise 2.8 Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space. Let $\boldsymbol{x}\in\boldsymbol{X}$. Show that $\{\boldsymbol{x}\}$ is sequentially compact. [Solution: (2.8)]

Exercise 2.9 Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space. Let $\boldsymbol{K},\boldsymbol{H}\subseteq\boldsymbol{X}$ be sequentially compact. Show that $\boldsymbol{K}\cap\boldsymbol{H}$ and $\boldsymbol{K}\cup\boldsymbol{H}$ are sequentially compact. [Solution: (2.9)]

Definition 2.4 [Cover and Subcover]
If $(\boldsymbol{X},\|\cdot\|)$ is a normed vector space, $\mathcal{I}$ is an index set and $\boldsymbol{W}\subseteq\boldsymbol{X}$, we say that the sequence $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{I}}$ of open subsets of $\boldsymbol{X}$ is a **cover** *for $\boldsymbol{W}$ if*

$$\boldsymbol{W}\subseteq\bigcup_{\alpha\in\mathcal{I}}\boldsymbol{U}_\alpha.$$

Let $J\subseteq\mathcal{I}$. If

$$\boldsymbol{W}\subseteq\bigcup_{\alpha\in\mathcal{J}}\boldsymbol{U}_\alpha,$$

we say that $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{J}}$ is a **subcover** *for $\boldsymbol{W}$. In particular, we say that $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{J}}$ is a* **proper subcover** *for $\boldsymbol{W}$ if*

$$\boldsymbol{W}\subset\bigcup_{\alpha\in\mathcal{J}}\boldsymbol{U}_\alpha,$$

and that $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{J}}$ *is a* **finite subcover** *for* $\boldsymbol{W}$ *if* $\mathcal{J}$ *is finite.*

Notice that the index set $\mathcal{I}$ can be countable or uncountable.

Example 2.7 The sequence $(U_n)_{n\in\mathbb{N}}$, where $U_n = \{(-n, n)\,|\,n \in \mathbb{N}\}$, covers $\mathbb{R}$.

Example 2.6 The sequence $(U_n)_{n\in\mathbb{N}}$, where $U_n = \{(n, n+2)\,|\,n \in \mathbb{Z}\}$, is a cover of $[0, 2]$, and $\{(-1, 1), (0, 2), (1, 3)\}$ is a finite subcover.

Definition 2.5 [Compact Set]
Let $(\boldsymbol{X}, \|\cdot\|)$ *be a normed vector space. Let* $\boldsymbol{K} \subseteq \boldsymbol{X}$*. Let* I *be an index set. Let* $(\boldsymbol{U}_\alpha)_{\alpha\in I}$ *be a sequence of open subsets of* $\boldsymbol{X}$ *such that*

$$\boldsymbol{K} \subseteq \bigcup_{\alpha\in I} \boldsymbol{U}_\alpha.$$

The set $\boldsymbol{K}$ *is said to be* compact *if there exists a finite subset* J *of* I *such that*

$$\boldsymbol{K} \subseteq \bigcup_{\alpha\in J} \boldsymbol{U}_\alpha.$$

Example 2.7 The set $\mathbb{R}$ is not compact since $\{(n, n+2)\,|\,n \in \mathbb{Z}\}$ is a cover for $\mathbb{R}$ with no finite subcover.

These definitions can be extended to more general contexts than the one of normed spaces. In particular:

(i) If a set is compact, it is sequentially compact;

(ii) For metric spaces (and therefore in particular for normed spaces), a set is compact if and only if it is sequentially compact;

(iii) If a set is compact, it is closed and bounded;

We focus on the relationship between compact and sequentially compact sets in normed spaces (even if all the results that follow can be re-written, without substantial modifications, to metric spaces). Imagine you're trying to cover a set of points on a line with segments of a fixed length. If the set is totally bounded, you can always do this with a finite number of segments, no matter how small the segments are. If the set weren't totally bounded, you'd need infinitely many segments to cover the entire set.

We introduce the notion of totally bounded set. Loosely speaking, a totally bounded set is a set that can be covered with a finite number of arbitrarily small balls.

Definition 2.6 [Totally Bounded Set]
Let $(\boldsymbol{X}, \|\cdot\|)$ *be a normed space. Let* $\boldsymbol{A} \subseteq \boldsymbol{X}$*. We say that* $\boldsymbol{A}$ *is* **totally bounded** *if*

$$\forall\, r > 0,\ \exists\, n \in \mathbb{N}^*,\ \exists\, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \boldsymbol{X}:\quad \boldsymbol{A} \subseteq \bigcup_{i=1}^{n} \mathcal{B}_r(\boldsymbol{x}_i). \tag{2.6}$$

Lemma 2.1 *Let* $(\boldsymbol{X}, \|\cdot\|)$ *be a normed space. Let* $\boldsymbol{K} \subseteq \boldsymbol{X}$ *be sequentially compact. Then,* $\boldsymbol{K}$ *is totally bounded.*

Proof By contradiction, suppose that $\boldsymbol{K}$ is not totally bounded, so that

$$\exists\, r > 0 : \quad \forall\, n \in \mathbb{N}^*,\ \forall\, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \boldsymbol{X} : \quad \boldsymbol{K} \not\subseteq \bigcup_{i=1}^{n} \mathcal{B}_r(\boldsymbol{x}_i).$$

Let r be such value. Let $\boldsymbol{x}_0 \in \boldsymbol{K}$. We define a sequence recursively in this way: for all $n \in \mathbb{N}^*$, since for any $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n \in \boldsymbol{X}$ we have $\boldsymbol{K} \not\subseteq \bigcup_{i=1}^{n} \mathcal{B}_r(\boldsymbol{x}_i)$, we can define

$$\boldsymbol{x}_{n+1} \notin \bigcup_{i=0}^{n} \mathcal{B}_r(\boldsymbol{x}_i).$$

Thus, for all $n, m \in \mathbb{N}$, if $n \neq m$ we have $\|\boldsymbol{x}_n - \boldsymbol{x}_m\| \geq r$. Since $\boldsymbol{K}$ is sequentially bounded, there exists $\hat{\boldsymbol{x}} \in \boldsymbol{K}$ and a subsequence $(\boldsymbol{x}_{n_k})_{k \in \mathbb{N}}$ such that

$$\boldsymbol{x}_{n_k} \xrightarrow{k \to \infty} \hat{\boldsymbol{x}}.$$

Since this subsequence is convergent, it is a Cauchy sequence, and therefore there exists $p \in \mathbb{N}$ such that, for all $h, k \geq p$, we have $\|\boldsymbol{x}_{n_k} - \boldsymbol{x}_{n_h}\| < r$. This contradicts the fact that $\|\boldsymbol{x}_{n_k} - \boldsymbol{x}_{n_h}\| \geq r$ if $h \neq k$.

Lemma 2.2 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{K} \subseteq \boldsymbol{X}$ be sequentially compact. Then, $\boldsymbol{K}$ has a countable dense subset $\boldsymbol{A}$.*

Proof Since $\boldsymbol{K}$ is sequentially compact, it is totally bounded. We can re-write the condition (2.6) as follows:

$$\forall\, n \in \mathbb{N}^* > 0,\ \exists\, m_n \in \mathbb{N}^*,\ \exists\, \boldsymbol{x}_1^{(n)}, \boldsymbol{x}_2^{(n)}, \ldots, \boldsymbol{x}_{m_n}^{(n)} \in \boldsymbol{X} : \quad \boldsymbol{K} \subseteq \bigcup_{i=1}^{m_n} \mathcal{B}_{\frac{1}{n}}(\boldsymbol{x}_i^{(n)}). \tag{2.7}$$

For example, for $n = 1$ there exist $\boldsymbol{x}_1^{(1)}, \ldots, \boldsymbol{x}_{m_1}^{(1)} \in \boldsymbol{X}$ such that $\boldsymbol{K} \subseteq \bigcup_{i=1}^{m_1} \mathcal{B}_1(\boldsymbol{x}_i^{(1)})$. For $n = 2$ there exist $\boldsymbol{x}_1^{(2)}, \ldots, \boldsymbol{x}_{m_2}^{(2)} \in \boldsymbol{X}$ such that $\boldsymbol{K} \subseteq \bigcup_{i=1}^{m_2} \mathcal{B}_{\frac{1}{2}}(\boldsymbol{x}_i^{(2)})$. We can continue in this way, and for all $n \in \mathbb{N}^*$, we define a collection C_n of m_n open balls with radius $1/n$ that covers $\boldsymbol{K}$. Let $\boldsymbol{C}$ denote the set of all the centres of these balls:

$$\boldsymbol{C} = \{\boldsymbol{x} \in \boldsymbol{X} \,|\, \exists\, n \in \mathbb{N}^* : \quad \mathcal{B}_{\frac{1}{n}}(\boldsymbol{x}) \in C_n\}.$$

This set is countable. We show that it is dense in $\boldsymbol{K}$. Let $\boldsymbol{y} \in \boldsymbol{K}$. Then, there exists $\boldsymbol{x}_1 \in C$ such that $\boldsymbol{y} \in \mathcal{B}_1(\boldsymbol{x}_1)$, because the balls of radius 1 cover $\boldsymbol{K}$. There exists $\boldsymbol{x}_2 \in C$ such that $\boldsymbol{y} \in \mathcal{B}_{\frac{1}{2}}(\boldsymbol{x}_2)$, and so on. We construct, in this way, a sequence $(\boldsymbol{x}_k)_{k \in \mathbb{N}^*}$ in C such that $\boldsymbol{y} \in \mathcal{B}_{\frac{1}{k}}(\boldsymbol{x}_k)$. It follows that $\boldsymbol{y}$ is an adherent point for C, and therefore the closure of C will contain $\boldsymbol{y}$.

Theorem 2.7 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{K} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{K}$ is compact if and only if it is sequentially compact.*

Proof We show that if $\boldsymbol{K}$ is compact, it is also sequentially compact. Suppose by contradiction that $\boldsymbol{K}$ is compact but not sequentially compact. Then, there exists a sequence $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ of $\boldsymbol{K}$ which has no convergent subsequence to a point of $\boldsymbol{K}$.
Let $\boldsymbol{x} \in \boldsymbol{K}$. We claim that there exists $\varepsilon_{\boldsymbol{x}} > 0$ such that the ball $\mathcal{B}_{\varepsilon_{\boldsymbol{x}}}(\boldsymbol{x})$ contains no points of the sequence $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ other than $\boldsymbol{x}$ (in case $\boldsymbol{x}$ is also a point of the sequence). By contradiction, assume that for all $\varepsilon > 0$, the ball $\mathcal{B}_\varepsilon(\boldsymbol{x})$ contains points of $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ other than $\boldsymbol{x}$. Taking $\varepsilon = 1/j$ with $j \in \mathbb{N}^*$, we are constructing a subsequence $(\boldsymbol{x}_{k_j})_{j \in \mathbb{N}^*}$ that converges to $\boldsymbol{x}$. We found a subsequence of points of $\boldsymbol{K}$ which converges to a point $\boldsymbol{x}$ of $\boldsymbol{K}$, leading to a contradiction. Therefore, the claim is true: there exists $\varepsilon_{\boldsymbol{x}} > 0$ such that the ball $\mathcal{B}_{\varepsilon_{\boldsymbol{x}}}(\boldsymbol{x})$ contains no points of the sequence $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ other than $\boldsymbol{x}$. Let $\mathcal{B} = \{\mathcal{B}_{\varepsilon_{\boldsymbol{x}}}(\boldsymbol{x}) \,|\, \boldsymbol{x} \in \boldsymbol{K}\}$, which is an open cover

of $\boldsymbol{K}$. Since $\boldsymbol{K}$ is compact, there must exist a finite subcover $\tilde{\mathcal{B}}$. This subcover contains all the points of the sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$. However, each ball in $\tilde{\mathcal{B}}$ can contain at most one point of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$, and since $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ has infinitely distinct points, we get a contradiction.

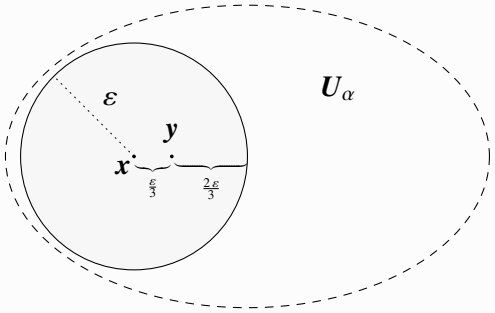
Now, we show that if $\boldsymbol{K}$ is sequentially compact, then it is compact. Let $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{I}}$ be an open cover of $\boldsymbol{K}$, where $\mathcal{I}$ is an index set. We want to show that there exists a finite subcover. From Lemma 2.2, the sequentially compact set $\boldsymbol{K}$ has a countable dense subset $\boldsymbol{A}$. We denote by $\mathcal{B}$ the collection of balls of rational radii around points of $\boldsymbol{A}$:

$$\mathcal{B} = \{\mathcal{B}_r(\boldsymbol{x}) \,|\, \boldsymbol{x}\in\boldsymbol{A} \;\text{ and }\; r\in\mathbb{Q}\cap\mathbb{R}_+\}.$$

The set $\mathcal{B}$ is countable since both $\boldsymbol{A}$ and $\mathbb{Q}$ are countable. Let

$$\boldsymbol{C} = \{\mathcal{B}_r(\boldsymbol{x}) \,|\, \boldsymbol{x}\in\boldsymbol{A} \;\text{ and }\; r\in\mathbb{Q}\cap\mathbb{R}_+ \;\text{ and }\; \exists\,\alpha\in\mathcal{I}:\; \mathcal{B}_r(\boldsymbol{x})\subseteq\boldsymbol{U}_\alpha\}.$$

The set $\boldsymbol{C}$ is contained in $\mathcal{B}$, and it is therefore at most countable. Since $(\boldsymbol{U}_\alpha)_{\alpha\in\mathcal{I}}$ covers $\boldsymbol{K}$, for all $\boldsymbol{x}\in\boldsymbol{K}$ there exists $\alpha\in\mathcal{I}$ such that $\boldsymbol{x}\in\boldsymbol{U}_\alpha$. Moreover, since $\boldsymbol{U}_\alpha$ is open, there exists $\varepsilon>0$ such that $\mathcal{B}_\varepsilon(\boldsymbol{x})\subseteq\boldsymbol{U}_\alpha$. Since $\boldsymbol{A}$ is dense, there exists $\boldsymbol{y}\in\boldsymbol{A}$ such that $\boldsymbol{x}\in\mathcal{B}_{\frac{\varepsilon}{3}}(\boldsymbol{y})$ and $\mathcal{B}_{\frac{2\varepsilon}{3}}(\boldsymbol{y})\subseteq\boldsymbol{U}_\alpha$.



Now, there exists $q\in\mathbb{Q}$ such that $\frac{\varepsilon}{3}<q<\frac{2\varepsilon}{3}$. Thus,

$$\boldsymbol{x}\in\mathcal{B}_q(\boldsymbol{y})\subseteq\mathcal{B}_{\frac{2\varepsilon}{3}}(\boldsymbol{y})\subseteq\boldsymbol{U}_\alpha.$$

The ball $\mathcal{B}_q(\boldsymbol{y})$ is in $\mathcal{B}$, and since $\mathcal{B}_q(\boldsymbol{y})\subseteq\boldsymbol{U}_\alpha$, it is also in $\boldsymbol{C}$. We have shown that $\boldsymbol{C}$ is a countable open cover of $\boldsymbol{K}$. Now, for all $B\in\boldsymbol{C}$ there exists $\alpha_B\in\mathcal{I}$ such that $B\subseteq\boldsymbol{U}_{\alpha_B}$. Consider $\{\boldsymbol{U}_{\alpha_B}\,|\,B\in\boldsymbol{C}\}$: this set is countable since $\boldsymbol{C}$ is countable. Moreover, this set covers $\boldsymbol{K}$, since $B\subseteq\boldsymbol{U}_{\alpha_B}$ and $\{B\,|\,B\in\boldsymbol{C}\}$ covers $\boldsymbol{K}$.

Now, we show that every countable cover of $\boldsymbol{K}$ has a finite subcover. Let $(\boldsymbol{U}_n)_{n\in\mathbb{N}}$ be an open cover of $\boldsymbol{K}$. By contradiction, suppose that there is no finite subcover, i.e., suppose that

$$\forall\, n\in\mathbb{N},\; \bigcup_{k=0}^{n}\boldsymbol{U}_k \not\supseteq \boldsymbol{K}.$$

Let us choose $\boldsymbol{x}_1\in\boldsymbol{K}$. Then, there exists $n_1\in\mathbb{N}$ such that $\boldsymbol{x}_1\in\boldsymbol{U}_{n_1}$. Let us choose $\boldsymbol{x}_2\in\boldsymbol{K}$ such that

$$\boldsymbol{x}_2\notin\bigcup_{k=0}^{n_1}\boldsymbol{U}_k.$$

This is always possible because we are assuming that $\boldsymbol{K}$ does not have a finite subcover. Using the same argument as before, there exists $n_2\in\mathbb{N}$ such that $\boldsymbol{x}_2\in\boldsymbol{U}_{n_2}$. We continue in this way, recursively. At the h-th

step, we choose $\boldsymbol{x}_h$ such that

$$\boldsymbol{x}_h \notin \bigcup_{k=0}^{n_1} \boldsymbol{U}_k \quad \text{and} \quad \exists\, n_h \in \mathbb{N} : \quad \boldsymbol{x}_h \in \boldsymbol{U}_{n_h}.$$

In this way, we have constructed the sequence $(n_h)_{h\in\mathbb{N}}$, which is strictly increasing. Since $\boldsymbol{K}$ is sequentially compact, there exists a sequence $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ which is convergent to a point $\hat{\boldsymbol{x}} \in \boldsymbol{K}$. Since $\boldsymbol{U}_n$ covers $\boldsymbol{K}$, there exists $h \in \mathbb{N}$ such that $\hat{\boldsymbol{x}} \in \boldsymbol{U}_h$. By construction,

$$\boldsymbol{x}_h \notin \bigcup_{k=0}^{n_{h-1}} \boldsymbol{U}_k.$$

and

$$\forall\, h \in \mathbb{N},\ \exists\, k_h \in \mathbb{N} :\ \forall\, m > k_h,\ \boldsymbol{x}_m \notin \boldsymbol{U}_h.$$

It follows that any subsequence cannot converge to a point in $\boldsymbol{U}_h$. The contradiction arises from the fact that we assumed that $\hat{\boldsymbol{x}}$ is a point in $\boldsymbol{U}_h$.

Remark 2.1 In Theorem 2.7 it was proved that in a compact space space every open cover has a countable subcover. Spaces that enjoy this property are referred to as *Lindelöf spaces*.

Remark 2.2 The result of Theorem 2.7 can be extended to metric spaces. The proof is exactly the same, *mutatis mutandis*.

Lemma 2.3 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{A} \subseteq \boldsymbol{X}$. Then, $\boldsymbol{A}$ is bounded if and only if, for every sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in $\boldsymbol{A}$, there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ which is convergent in $\boldsymbol{X}$.*

Proof Suppose that $\boldsymbol{A}$ is bounded, and let a be a sequence of points in $\boldsymbol{A}$. The sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is bounded. From Bolzano-Weierstrass theorem (Theorem 1.41), there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ which is convergent.

Conversely, suppose that every sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in $\boldsymbol{A}$ has a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ which is convergent in $\boldsymbol{X}$. We prove that A is bounded. By contradiction, suppose that $\boldsymbol{A}$ is not bounded. For every $k \in \mathbb{N}$, there exists $\boldsymbol{x}_k \in \boldsymbol{A}$ such that $\|\boldsymbol{x}_k\| \geq k$. Thus, $\|\boldsymbol{x}_k\| \overset{k\to\infty}{\longrightarrow} +\infty$. By assumption, there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ which is convergent in $\boldsymbol{X}$. Let $\ell = \lim_{j\to\infty} \boldsymbol{x}_{k_j}$. Thus, we have simultaneously $\|\boldsymbol{x}_{k_j}\| \overset{j\to\infty}{\longrightarrow} +\infty$ and $\|\boldsymbol{x}_{k_j}\| \overset{j\to\infty}{\longrightarrow} \|\ell\|$, which is a contradiction.

Theorem 2.8 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $\boldsymbol{K}$ be sequentially compact subset of $\boldsymbol{X}$. Then, $\boldsymbol{K}$ is closed and bounded.*

Proof Suppose that $\boldsymbol{K}$ is compact. From Lemma 2.3, the set $\boldsymbol{K}$ is bounded. We need to show that $\boldsymbol{K}$ is closed. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}$ which is convergent in $\boldsymbol{X}$. Since $\boldsymbol{K}$ is sequentially compact, there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ which is convergent in K. Then,

$$\lim_{j\to\infty} \boldsymbol{x}_{k_j} = \lim_{k\to\infty} \boldsymbol{x}_k.$$

and so $\lim_{k\to\infty} \boldsymbol{x}_k \in \boldsymbol{K}$. Thus, from Theorem 2.4 we obtain that $\boldsymbol{K}$ is closed.

If the normed vector space considered is Euclidean, the converse is also true: this is the generalization of the

Heine-Borel theorem for $\mathbb{R}^n$.

Theorem 2.9 [Heine-Borel Theorem]
Consider the normed vector space $(\mathbb{R}^n, \|\cdot\|)$, where $\|\cdot\|$ denotes the Euclidean 2-norm. Let $\boldsymbol{K} \subseteq \mathbb{R}^n$. Then, $\boldsymbol{K}$ is sequentially compact if and only if it is bounded and closed.

Proof From Bolzano-Weierstrass theorem (Theorem 1.41), if $(a_k)_{k\in\mathbb{N}}$ is a bounded sequence of $\mathbb{R}^n$, there exists a subsequence which is convergent in $\mathbb{R}^n$. A simple consequence is that a subset $\boldsymbol{A}$ of $\mathbb{R}^n$ is bounded if and only if for every sequence $(a_k)_{k\in\mathbb{N}}$ in $\boldsymbol{A}$ there exists a subsequence which is convergent. The *only if* part follows immediately from Bolzano-Weierstrass theorem (Theorem 1.41): if $(a_k)_{k\in\mathbb{N}}$ is a sequence of points of $\boldsymbol{A}$, there exists a subsequence which is convergent. To prove the *if* part, we proceed by contradiction. Suppose $\boldsymbol{A}$ is not bounded, so that for every $k \in \mathbb{N}$ there exists $a_k \in \boldsymbol{A}$ such that $\|a_k\| \geq k$. Thus, $\|a_k\| \overset{k\to\infty}{\longrightarrow} +\infty$. However, by assumption there exists a subsequence $(a_{k_j})_{j\in\mathbb{N}}$ of $(a_k)_{k\in\mathbb{N}}$ which is convergent in $\mathbb{R}^n$. Let $\ell = \lim_{j\to\infty} a_{k_j}$. We reach a contradiction since we have $\|a_{k_j}\| \overset{j\to\infty}{\longrightarrow} +\infty$ and $\|a_{k_j}\| \overset{j\to\infty}{\longrightarrow} \|\ell\|$.

Now, suppose that $\boldsymbol{K}$ is compact. From this discussion, since by definition $\boldsymbol{K}$ has a convergent subsequence, the set $\boldsymbol{K}$ is bounded. We need to show that $\boldsymbol{K}$ is closed. Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}$ which is convergent with respect to $\mathbb{R}^n$. Since $\boldsymbol{K}$ is compact, there exists a subsequence $(b_k)_{k\in\mathbb{N}}$ of $(a_k)_{k\in\mathbb{N}}$ which is convergent with respect to $\boldsymbol{K}$. From $\lim_{k\to\infty} b_k = \lim_{k\to\infty} a_k$, we obtain $\lim_{k\to\infty} a_k \in \boldsymbol{K}$. Thus, we obtain that $\boldsymbol{K}$ is closed (recall that a subset $\boldsymbol{U}$ of a normed vector space $(\boldsymbol{X}, \|\cdot\|)$ is closed if and only if, for every convergent sequence of points of $\boldsymbol{U}$, the limit is in $\boldsymbol{U}$).

Conversely, suppose that $\boldsymbol{K}$ is bounded and closed. Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}$. From Bolzano-Weierstrass theorem (Theorem 1.41), there exists a subsequence $(b_k)_{k\in\mathbb{N}}$ of $(a_k)_{k\in\mathbb{N}}$ which is convergent with respect to $\mathbb{R}^n$. Therefore, using again the fact that a subset is closed if and only if every convergent sequence has a limit in it, we have $\lim_{k\to\infty} b_k \in \boldsymbol{K}$. Thus, $(b_k)_{k\in\mathbb{N}}$ is a subsequence of $(a_k)_{k\in\mathbb{N}}$ which is convergent with respect to $\boldsymbol{K}$. Therefore, $\boldsymbol{K}$ is compact.

The Heine-Borel theorem is not true in more general settings and for general metric spaces with a metric different from the Euclidean one.

The following result, known as Heine-Cantor theorem, partially inverts the implication in Theorem 1.59.

Theorem 2.10 [Heine-Cantor Theorem]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed spaces. Let $\boldsymbol{K} \subseteq \boldsymbol{X}$ be sequentially compact. Let $f : \boldsymbol{K} \longrightarrow \boldsymbol{Y}$ be continuous. Then, f is uniformly continuous.

Proof Suppose by contradiction that f is not uniformly continuous. Then, there exists $\varepsilon > 0$ such that for all $\delta > 0$ we can find $a, b \in \boldsymbol{K}$ such that $\|a - b\|_X < \delta$ and $\|f(a) - f(b)\|_Y \geq \varepsilon$. From the arbitrariness of δ, for every $k \in \mathbb{N}^*$ we can choose $\delta = \frac{1}{k}$. Then, there exist $a_k, b_k \in \boldsymbol{K}$ such that $\|a_k - b_k\| < \frac{1}{k}$ and $\|f(a_k) - f(b_k)\|_Y \geq \varepsilon$. In this way, we have defined two sequences $(a_k)_{k\in\mathbb{N}^*}$ and $(b_k)_{k\in\mathbb{N}^*}$ of points in $\boldsymbol{K}$. Since $\boldsymbol{K}$ is compact, there exist a subsequence $(a_{k_j})_{j\in\mathbb{N}^*}$ convergent to a point of $\boldsymbol{K}$. Let $\ell = \lim_{j\to\infty} a_{k_j}$. For the same reason, there exists a subsequence $(b_{k_{h_i}})_{i\in\mathbb{N}^*}$ extracted from $(b_{k_i})_{i\in\mathbb{N}^*}$ which is convergent to a point of $\boldsymbol{K}$. Let $\ell' = \lim_{i\to\infty} b_{k_{h_i}}$ Since $(a_{k_{j_i}})_{i\in\mathbb{N}^*}$ is a subsequence of $(a_{k_j})_{j\in\mathbb{N}^*}$, we also have $\lim_{i\to\infty} a_{k_{j_i}} = \ell$. Since

$$0 \leq \|a_k - b_k\|_X \leq \frac{1}{k} \quad \forall\, k \in \mathbb{N}^*,$$

the sequence $(\|a_k - b_k\|)_{k\in\mathbb{N}^*}$ is convergent and

$$\lim_{k\to\infty} \|a_k - b_k\|_X = 0.$$

Therefore, also $(a_k - b_k)_{k\in\mathbb{N}^*}$ is convergent and

$$\lim_{k\to\infty} (a_k - b_k) = 0.$$

It follows that $(a_{k_{j_i}} - b_{k_{j_i}})_{i\in\mathbb{N}^*}$ is convergent and

$$\lim_{i\to\infty} (a_{k_{j_i}} - b_{k_{j_i}}) = 0.$$

Since

$$\lim_{i\to\infty} (a_{k_{j_i}} - b_{k_{j_i}}) = \lim_{i\to\infty} a_{k_{j_i}} - \lim_{i\to\infty} b_{k_{j_i}} = \ell - \ell',$$

it follows that $\ell = \ell'$. Since f is continuous, the sequences $(f(a_{k_{j_i}}))_{i\in\mathbb{N}^*}$ and $(f(b_{k_{j_i}}))_{i\in\mathbb{N}^*}$ are convergent, and

$$\lim_{i\to\infty} f(a_{k_{j_i}}) = f(\ell) \quad \text{and} \quad \lim_{i\to\infty} f(b_{k_{j_i}}) = f(\ell').$$

Thus, $(f(a_{k_{j_i}}) - f(b_{k_{j_i}}))_{i\in\mathbb{N}^*}$ is convergent and

$$\lim_{i\to\infty} f(a_{k_{j_i}}) - f(b_{k_{j_i}}) = \lim_{i\to\infty} f(a_{k_{k_i}}) - \lim_{i\to\infty} f(b_{k_{j_i}}) = f(\ell) - f(\ell') = 0.$$

There exists $p \in \mathbb{N}$ such that for all $i \geq p$ we have $\|f(a_{k_{j_i}}) - f(b_{k_{j_i}})\|_Y < \varepsilon$. This is a contradiction, because since $\|f(a_k) - f(b_k)\|_Y \geq \varepsilon$ for all $k \in \mathbb{N}^*$, we also must have $\|f(a_{k_{j_i}}) - f(b_{k_{j_i}})\|_Y \geq \varepsilon$.

2.3 Banach spaces

Banach spaces play a central role in functional analysis, and they constitute the natural initial step towards the introduction of Hilbert spaces.

Definition 2.7 [Banach Space]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. We say that $(\boldsymbol{X}, \|\cdot\|)$ is complete*, or that $(\boldsymbol{X}, \|\cdot\|)$ is a* Banach space*, if every Cauchy sequence in $\boldsymbol{X}$ is convergent.*

With this definition, Theorem 2.3 can be re-phrased by saying that every finite-dimensional normed space is complete. In particular, $\mathbb{R}^n$ is complete. We also observe that a normed vector space is complete with respect to a certain norm if and only if it is complete with respect to any norm equivalent to the given one.

Example 2.8 Let $a, b \in \mathbb{R}$ be such that $a < b$. We want to show that $C([a,b],\mathbb{R})$ with the norm

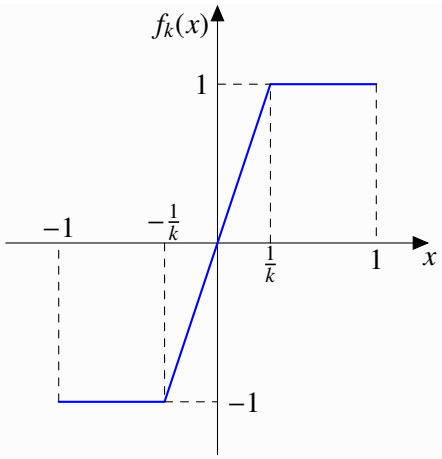
$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x$$

is not complete. To this end, consider $a = -1$ and $b = 1$. We want to show that $C([-1,1],\mathbb{R})$ with the norm

$$\|f\|_1 = \int_{-1}^1 |f(x)|\,\mathrm{d}x$$

is not complete. Consider the sequence of functions $(f_k)_{k\in\mathbb{N}}$ defined by

$$f_k(x) = \begin{cases} k\,x & \text{if } x \in \left(-\dfrac{1}{k}, \dfrac{1}{k}\right) \\ 1 & \text{if } x \in \left[\dfrac{1}{k}, 1\right] \\ -1 & \text{if } x \in \left[-1, -\dfrac{1}{k}\right] \end{cases}$$



Consider also the function

$$\begin{aligned} f : [-1, 1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \in (0, 1] \\ -1 & \text{if } x \in [-1, 0). \end{cases} \end{aligned}$$

We have

$$\begin{aligned} \int_{-1}^{1} |f_k(x) - f(x)|\,\mathrm{d}x &= \int_{-1}^{-\frac{1}{k}} |-1+1|\,\mathrm{d}x + \int_{-\frac{1}{k}}^{0} |k\,x+1|\,\mathrm{d}x + \int_{0}^{\frac{1}{k}} |k\,x-1|\,\mathrm{d}x + \int_{\frac{1}{k}}^{1} |1-1|\,\mathrm{d}x \\ &= 0 + \int_{-\frac{1}{k}}^{0} (k\,x+1)\,\mathrm{d}x + \int_{0}^{\frac{1}{k}} (1-k\,x)\,\mathrm{d}x + 0 \\ &= \left[k\,\frac{x^2}{2} + x\right]_{-\frac{1}{k}}^{0} + \left[x - k\,\frac{x^2}{2}\right]_{0}^{\frac{1}{k}} \\ &= -\left(\frac{1}{2\,k} - \frac{1}{k}\right) + \left(\frac{1}{k} - \frac{1}{2\,k}\right) = \frac{1}{k}. \end{aligned}$$

From this, we deduce that $(f_k)_{k\in\mathbb{N}}$ is a Cauchy sequence. In fact,

$$\begin{aligned} \|f_n - f_m\|_1 &= \int_{-1}^{1} |f_n(x) - f_m(x)|\,\mathrm{d}x \\ &\leq \int_{-1}^{1} |f_m(x) - f(x)|\,\mathrm{d}x + \int_{-1}^{1} |f(x) - f_n(x)|\,\mathrm{d}x = \frac{1}{m} + \frac{1}{n}. \end{aligned}$$

Given an arbitrary $\varepsilon > 0$, taking $n, m > 2/\varepsilon$ ensures that $\frac{1}{m} < \frac{\varepsilon}{2}$ and $\frac{1}{n} < \frac{\varepsilon}{2}$, so that

$$\|f_n - f_m\|_1 \leq \frac{1}{m} + \frac{1}{n} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

We now verify that this sequence does not converge. By contradiction, suppose that this sequence admits limit, i.e., there exists $g \in C([-1, 1], \mathbb{R})$ such that

$$\lim_{k\to\infty} f_k = g,$$

or, equivalently,

$$\lim_{k\to\infty} \|f_k - g\|_1 = 0.$$

We have

$$\int_{-1}^{1} |g(x) - f(x)|\,\mathrm{d}x = \int_{-1}^{1} |g(x) - f_k(x) + f_k(x) - f(x)|\,\mathrm{d}x \leq \int_{-1}^{1} |g(x) - f_k(x)|\,\mathrm{d}x + \int_{-1}^{1} |f_k(x) - f(x)|\,\mathrm{d}x.$$

Consider the terms in the right hand-side: $\int_{-1}^{1} |g(x) - f_k(x)|\,\mathrm{d}x$ goes to zero as $k \to \infty$ by assumption (we are assuming that g is the limit), and $\int_{-1}^{1} |f_k(x) - f(x)|\,\mathrm{d}x = 1/k$ as proved before, and this term also tends to zero as $k \to \infty$. On the other hand, since the left hand-side does not depend on k, we deduce that

$$\int_{-1}^{1} |g(x) - f(x)|\,\mathrm{d}x = 0.$$

Since g and f are continuous at every $x \neq 0$, then $g(x) = f(x)$ for all $x \neq 0$. It follows that g is discontinuous at 0 like f, and this is a contradiction.

We can re-write the conditions for pointwise and uniform convergence of a sequence of functions by using Cauchy's condition. For example, if $(\boldsymbol{Y}, \|\cdot\|)$ is a Banach space, A is a non-empty set and $(f_k)_{k\in\mathbb{N}}$ is a sequence of functions from A to $\boldsymbol{Y}$, the sequence $(f_k)_{k\in\mathbb{N}}$ converges pointwise if and only if

$$\forall\, x \in A,\ \ \forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow \qquad \|f_n(x) - f_m(x)\| < \varepsilon.$$

It is more interesting to see how we can use Cauchy's condition to give an equivalent characterization of uniform convergence.

Theorem 2.11 *Let $(\boldsymbol{Y}, \|\cdot\|)$ be a Banach space. Let A be a non-empty set. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions from A to $\boldsymbol{Y}$. The sequence of functions $(f_k)_{k\in\mathbb{N}}$ converges uniformly if and only if*

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow\ \forall\, x \in A, \qquad \|f_n(x) - f_m(x)\| < \varepsilon. \tag{2.8}$$

Proof We can re-write (2.8) as

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow \qquad \sup_{x\in A} \|f_n(x) - f_m(x)\| < \varepsilon. \tag{2.9}$$

The necessity is left as an exercise (see Exercise 2.10). Consider the sufficiency. Condition (2.9) implies that, for every $x \in A$, the sequence $(f_k(x))_{k\in\mathbb{N}}$ – which is a sequence of elements of $\boldsymbol{Y}$ – is a Cauchy sequence. Since $(\boldsymbol{Y}, \|\cdot\|)$ is a Banach space, $(f_k(x))_{k\in\mathbb{N}}$ is convergent. Hence, $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent, and we need to show that it is also uniformly convergent. Let $f : A \longrightarrow \boldsymbol{Y}$ be the pointwise limit. We have to show that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \in \mathbb{N}, \qquad k > p \ \Rightarrow\ \forall\, x \in A, \qquad \|f_k(x) - f(x)\| < \varepsilon.$$

Let $\varepsilon > 0$. Using (2.8), there exists p_1 such that, for all $n, k > p_1$ and for all $x \in A$, we have

$$\|f_n(x) - f_k(x)\| < \frac{\varepsilon}{2}.$$

Since $(f_k)_{k\in\mathbb{N}}$ is pointwise convergent to f, for every $x \in A$ there exists p_2 such that for all $k > p_2$ we have

$$\|f_k(x) - f(x)\| < \frac{\varepsilon}{2}.$$

Thus, for all $x \in A$, taking $n > p_1$ and $k > \max\{p_1, p_2\}$ we find that

$$\|f_n(x) - f(x)\| \leq \|f_n(x) - f_k(x)\| + \|f_k(x) - f(x)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

It follows that for every $n > p_1$ we have $\sup_{x\in A} \|f_n(x) - f(x)\| < \varepsilon$, which shows that we have the uniform convergence.

Exercise 2.10 Let $(\boldsymbol{Y}, \|\cdot\|)$ be a Banach space. Let A be a non-empty set. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions from A to $\boldsymbol{Y}$ that converges uniformly to a function $f : A \longrightarrow \boldsymbol{Y}$. Show that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow\ \forall\, x \in A, \qquad \|f_n(x) - f_m(x)\| < \varepsilon. \tag{2.10}$$

[Solution: (2.10)]

In Theorem 1.57, we have shown that the uniform limit of a sequence of continuous functions is itself a continuous function. We now take into account the corresponding result for the differentiability. However, here we need to be careful. Consider, for every $k \in \mathbb{N}^*$, the functions $f_k : \mathbb{R} \longrightarrow \mathbb{R}$ and $g_k : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f_k(x) = \frac{\sin(k\,x)}{k} \quad \text{and} \quad g_k(x) = \sqrt{\frac{1}{k} + x^2}.$$

The sequences $(f_k)_{k\in\mathbb{N}^*}$ is uniformly convergent to the zero function. Indeed, for all $x \in \mathbb{R}$ we have

$$\lim_{k\to\infty} \frac{\sin(k\,x)}{k} = 0,$$

so that the pointwise limit is the zero function. Moreover,

$$\sup_{x\in\mathbb{R}} \left|\frac{\sin(k\,x)}{k}\right| = \frac{1}{k},$$

which means that the sequence $\left(\sup_{x\in\mathbb{R}} |f_k - 0|\right)_{k\in\mathbb{N}^*}$ is infinitesimal.

The sequence of functions $(g_k)_{k\in\mathbb{N}^*}$ can be seen to uniformly converge to the zero function and to the function $g(x) = |x|$, respectively. The identically zero function is differentiable and its derivative is zero at each point, but $f_k'(0) = 1$ for every $k \in \mathbb{N}^*$, and the absolute value function is not differentiable at the origin. Therefore, the uniform convergence is not sufficient to ensure that the formula

$$\lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_k(x) = \frac{\mathrm{d}}{\mathrm{d}x} \lim_{k\to\infty} f_k(x) \tag{2.11}$$

holds. A sufficient condition is the uniform convergence of the sequence of the derivatives. The following result addresses this issue.

Theorem 2.12 *Let I be a bounded interval of $\mathbb{R}$. Consider a field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of differentiable functions $f_k : I \longrightarrow \mathbb{F}$. For every $k \in \mathbb{N}$, assume that the sequence $(f_k')_{k\in\mathbb{N}}$ of the derivatives is uniformly convergent. If there exists $x_0 \in I$ such that the sequence $(f_k(x_0))_{k\in\mathbb{N}}$ converges, then $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent in I and (9.1) holds for every $x \in I$.*

Proof We prove this result in the case $\mathbb{F} = \mathbb{R}$. We recall the Lagrange mean value theorem: given a closed interval $[a, b]$, with $a < b$, if we have a function $\varphi : [a, b] \longrightarrow \mathbb{R}$ continuous in $[a, b]$ and differentiable in (a, b), there exists $\xi \in (a, b)$ such that

$$\varphi'(\xi) = \frac{\varphi(b) - \varphi(a)}{b - a}.$$

Consider $a = x_0$ and $b = x$, and, given $n, m \in \mathbb{N}$, let $\varphi = f_n - f_m$, so that the latter says that there exists $\xi \in (x_0, x)$ such that

$$f_n'(\xi) - f_m'(\xi) = \frac{f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)}{x - x_0},$$

i.e.,

$$f_n(x) - f_m(x) = f_n(x_0) - f_m(x_0) + (x - x_0)\,(f_n'(\xi) - f_m'(\xi)).$$

Denoting by $\ell(I)$ the length of I (i.e., $\ell(I) = \sup(I) - \inf(I)$) and using the triangle inequality,

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x_0) - f_m(x_0)| + |x - x_0|\,\left|f_n'(\xi) - f_m'(\xi)\right| \\ &\leq |f_n(x_0) - f_m(x_0)| + |x - x_0| \sup_{x\in I}\left|f_n'(x) - f_m'(x)\right| \\ &\leq |f_n(x_0) - f_m(x_0)| + \ell(I) \sup_{x\in I}\left|f_n'(x) - f_m'(x)\right|. \end{aligned}$$

This inequality says that $|f_n(x_0)-f_m(x_0)|+\ell(I)\ \sup_I\left|f_n'-f_m'\right|$ is an upper bound for the set $\{|f_n(x)-f_m(x)| \mid x \in I\}$, and $\sup_{x\in I}|f_n(x) - f_m(x)|$ is the least upper bound of the same set, so that

$$\sup_{x\in I}|f_n(x) - f_m(x)| \leq |f_n(x_0) - f_m(x_0)| + \ell(I) \sup_{x\in I}\left|f_n'(x) - f_m'(x)\right|. \tag{2.12}$$

Consider the condition for uniform convergence (2.9). We need to show that the sequence $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent, i.e., that if n and m are sufficiently large, we can make $\sup_{x\in I}|f_n(x) - f_m(x)|$ arbitrarily small. Clearly, (2.12) ensures exactly this fact. Indeed, we know in fact that $(f_k(x_0))_{k\in\mathbb{N}}$ converges, and the sequence $(f_k(x_0))_{k\in\mathbb{N}}$ is also a Cauchy sequence (we are working in $(\mathbb{F}, |\cdot|)$) so that by choosing n and m sufficiently large we can make $|f_n(x_0) - f_m(x_0)|$ arbitrarily small. On the other hand, we know also that the sequence $(f_k')_{k\in\mathbb{N}}$ of the derivatives is uniformly convergent, and therefore using (2.9) again we see that we can make $\left|f_n' - f_m'\right|$ (and therefore also $\ell(I) \sup_{x\in I}\left|f_n'(x) - f_m'(x)\right|$) arbitrarily small, by taking n and m to be sufficiently large. This proves the uniform convergence of $(f_k)_{k\in\mathbb{N}}$ to a function $f : I \longrightarrow \mathbb{R}$. Now, we define $\Omega = \{(x, y) \in I \times I \,|\, x \neq y\}$, and we define, for all $(x, y) \in \Omega$, the sequence $(\varphi_k)_{k\in\mathbb{N}}$ where, for all $k \in \mathbb{N}$, $\psi_k : \Omega \longrightarrow \mathbb{R}$ and the function $\Psi : \Omega \longrightarrow \mathbb{R}$ by

$$\psi_k(x, y) = \frac{f_k(x) - f_k(y)}{x - y} \quad \text{and} \quad \Psi(x, y) = \frac{f(x) - f(y)}{x - y}.$$

We show that the sequence $(\psi_k)_{k\in\mathbb{N}}$ is uniformly convergent to Ψ. Using again the mean value theorem, for every $n, m \in \mathbb{N}$ and for every $x \in \Omega$, there exists $\xi \in (x, y)$ such that

$$\begin{aligned} |\psi_n(x, y) - \psi_m(x, y)| &= \left|\frac{(f_n(x) - f_n(y)) - (f_m(x) - f_m(y))}{x - y}\right| \\ &= \left|\frac{(f_n(x) - f_m(x)) - (f_n(y) - f_m(y))}{x - y}\right| \\ &= |f_n'(\xi) - f_m'(\xi)| \leq \sup_{x\in I}|f_n'(x) - f_m'(x)|. \end{aligned}$$

From the arbitrariness of (x, y) in Ω, it follows that $\sup_{x\in I}|f_n'(x) - f_m'(x)|$ is an upper bound for the set $\{|\psi_n(x, y) - \psi_m(x, y)| \mid (x, y) \in \Omega\}$. The least upper bound of this same set is $\sup_{(x,y)\in\Omega}|\psi_n(x, y) - \psi_m(x, y)|$. It follows that for every $n, m \in \mathbb{N}$

$$\sup_{(x,y)\in\Omega}|\psi_n(x, y) - \psi_m(x, y)| \leq \sup_{x\in I}|f_n'(x) - f_m'(x)|.$$

Using (2.9) again, we see that the sequence $(\psi_k)_{k\in\mathbb{N}}$ is uniformly convergent to a function $\gamma : \Omega \longrightarrow \mathbb{R}$. On the other hand, the same sequence is pointwise convergent to Ψ. Thus, γ coincides with Ψ and $(\psi_k)_{k\in\mathbb{N}}$ is uniformly convergent to Ψ. Let now g denote $\lim_{k\to\infty} f_k'$. Let $\varepsilon > 0$. Let m be such that

$$\sup_{x\in I}|f_m'(x) - g(x)| < \frac{\varepsilon}{3} \quad \text{and} \quad \sup_{(x,y)\in\Omega}|\psi_m(x, y) - \Psi(x, y)| < \frac{\varepsilon}{3}.$$

For every $(x, y) \in \Omega$, we find

$$\begin{aligned} |\Psi(x,y) - g(x)| &\leq |\Psi(x,y) - \psi_m(x,y)| + |\psi_m(x,y) - f_m'(x)| + |f_m'(x) - g(x)| \\ &\leq \frac{\varepsilon}{3} + |\psi_m(x,y) - f_m'(x)| + \frac{\varepsilon}{3}. \end{aligned}$$

Let $x \in I$. From the definition of $f_m'(x)$, there exists $\delta > 0$ such that, for all $y \in I$ different from x and such that $|y - x| < \delta$, there holds $|\psi_m(x,y) - f_m'(x)| < \frac{\varepsilon}{3}$. It follows that for such values of y we have $|\Psi(x,y) - g(x)| < \varepsilon$. This shows that $f'(x) = g(x)$ for every $x \in I$.

Example *2.9*

Let $a, b \in \mathbb{R}$ with $a < b$. Consider the vector space $C([a,b], \mathbb{R})$ with the norm

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|.$$

We show that $(C([a,b], \mathbb{R}), \|\cdot\|_\infty)$ is complete. To this end, take a Cauchy sequence $(f_k)_{k \in \mathbb{N}}$ of $C([a,b], \mathbb{R})$, so that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{R} :\quad \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow \|f_n - f_m\|_\infty = \max_{x \in [a,b]} |f_n(x) - f_m(x)| < \varepsilon,$$

which can be re-written as

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{R} :\quad \forall\, n, m \in \mathbb{N}, \qquad n, m > p \ \Rightarrow\ \forall\, x \in [a,b],\ |f_n(x) - f_m(x)| < \varepsilon.$$

This is exactly Cauchy's criterion for the uniform convergence of a sequence of functions. Therefore, there exists $f : [a,b] \longrightarrow \mathbb{R}$ such that

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

uniformly in x, and f is continuous since such are all the functions f_k, and therefore $f \in C([a,b], \mathbb{R})$. Hence

$$\|f_k - f\|_\infty = \max_{x \in [a,b]} |f_k(x) - f(x)| \overset{k \to \infty}{\longrightarrow} 0,$$

so that the limit in $C([a,b], \mathbb{R})$ of the sequence $(f_k)_{k \in \mathbb{N}}$ exists, and this limit coincides with f.

Normed spaces that fail to be complete are "too small": they lack some of the limits of the Cauchy sequences. In the example of $(C([a,b], \mathbb{R}), \|\cdot\|_1)$ the continuous functions are not sufficient, and the sequence considered would not have been a counterexample if the normed space also contained the sign function.

We can therefore consider the idea of taking the space of all integrable functions and introduce the norm $\|\cdot\|_1$ therein. Unfortunately, in doing so, $\|\cdot\|_1$ is no longer a norm: in fact, a function which is zero everywhere except at a single point has zero integral. A way to overcome this difficulty is to identify any two integrable functions f and g such that

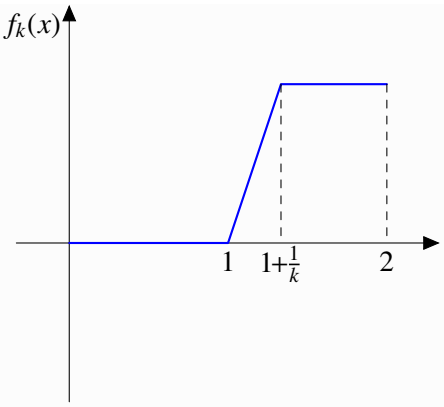
$$\int_a^b |f - g| = 0.$$

In this way $\|\cdot\|_1$ becomes a norm. However, it is possible to show that the resulting normed space is still incomplete: the Riemann integrable functions are still too few. This problem has been solved via the construction of a more sophisticated integration theory, the *Lebesgue integration theory*. In the space of functions which are Lebesgue integrable (with the identification mentioned above), $\|\cdot\|_1$ is a norm that renders the space complete.

Example *2.11*

We show that $(C([a,b], \mathbb{R}), \|\cdot\|_2)$ is not complete. Consider for example $a = 0$ and $b = 2$, and consider the sequence of functions $(f_k)_{k \in \mathbb{N}^*}$ defined by

$$f_k(x) = \begin{cases} 0 & \text{if } x \in [0,1] \\ k(x-1) & \text{if } 1 < x < 1 + \dfrac{1}{k} \\ 1 & \text{if } x \in \left[1 + \dfrac{1}{k}, 2\right] \end{cases}$$



For each $k \in \mathbb{N}^*$, the function f_k is continuous. The sequence $(f_k)_{k\in\mathbb{N}^*}$ is Cauchy in the $\|\cdot\|_2$ norm. Indeed, if $n < m$,

$$\begin{aligned}
\|f_n - f_m\|_2 &= \int_0^2 |f_n(x) - f_m(x)|^2 \,\mathrm{d}x \\
&= \int_1^{1+\frac{1}{n}} |f_n(x) - f_m(x)|^2 \,\mathrm{d}x \\
&= \int_1^{1+\frac{1}{m}} \left(n(x-1) - m(x-1)\right)^2 \mathrm{d}x + \int_{1+\frac{1}{m}}^{1+\frac{1}{n}} \left(n(x-1) - 1\right)^2 \mathrm{d}x \\
&= \int_1^{1+\frac{1}{m}} (n-m)^2 (x-1)^2 \,\mathrm{d}x + \int_{1+\frac{1}{m}}^{1+\frac{1}{n}} \left(n(x-1) - 1\right) \mathrm{d}x \qquad (y = x-1) \\
&= (n-m)^2 \int_0^{\frac{1}{m}} y^2 \,\mathrm{d}y + \int_{\frac{1}{m}}^{\frac{1}{n}} (ny - 1)^2 \,\mathrm{d}y \qquad (\xi = ny - 1) \\
&= (n-m)^2 \left[\frac{y^3}{3}\right]_0^{\frac{1}{m}} + \frac{1}{n} \int_{\frac{n}{m}-1}^{0} \xi^2 \,\mathrm{d}\xi \\
&= (n-m)^2 \frac{1}{3m^3} + \frac{1}{n} \left[\frac{\xi^3}{3}\right]_{\frac{n}{m}-1}^{0} \\
&= \frac{(n-m)^2}{3m^3} - \frac{1}{3n} \frac{(n-m)^3}{m^3} \\
&= \frac{n(n-m)^2 - (n-m)^3}{3m^3 n} \\
&= \frac{(n-m)^2 (n-n+m)}{3m^3 n} \\
&= \frac{(n-m)^2}{3m^2 n} < \frac{1}{3n} \xrightarrow{n\to\infty} 0.
\end{aligned}$$

However, the sequence is not convergent in $C([0,2],\mathbb{R})$ with respect to the norm $\|\cdot\|_2$. In fact, $(f_k)_{k\in\mathbb{N}^*}$ converges to the (shifted) step function

$$f(x) = \begin{cases} 0 & \text{if } x \in [0,1) \\ 1 & \text{if } x \in [1,2] \end{cases}$$

since

$$\begin{aligned}
\|f_k - f\|^2 &= \int_0^2 |f_k(x) - f(x)|^2 \,\mathrm{d}x \\
&= \int_1^{1+\frac{1}{k}} \left(k\,(x-1) - 1\right)^2 \mathrm{d}x \qquad (y = x - 1) \\
&= \int_0^{\frac{1}{k}} (k\,y - 1)^2 \,\mathrm{d}y \qquad (\xi = n\,y - 1) \\
&= \frac{1}{k} \int_{-1}^{0} \xi^2 \,\mathrm{d}\xi \\
&= \frac{1}{k} \left[\frac{\xi^3}{3}\right]_{-1}^{0} \\
&= \frac{1}{3\,k} \overset{k \to \infty}{\longrightarrow} 0.
\end{aligned}$$

Hence, the limit of $(f_k)_{k \in \mathbb{N}^*}$ is f, which does not belong to $C([0,2], \mathbb{R})$.

2.3.1 Banach spaces and series

The definition of series of real or complex numbers uses only the operation of addition and the notion of limit. Thus, the definition of series in a normed vector space is formally identical to the real/complex case.

Definition 2.8 [Sum of the Series and Partial Sum]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ be a sequence of vectors of $\boldsymbol{X}$. The sequence $(\boldsymbol{s}_k)_{k \in \mathbb{N}}$ defined by

$$\boldsymbol{s}_k = \boldsymbol{x}_0 + \boldsymbol{x}_1 + \ldots + \boldsymbol{x}_k$$

is called partial sum. *If the sequence $(\boldsymbol{s}_k)_{k \in \mathbb{N}}$ is convergent, we say that*

$$\sum_{k=0}^{\infty} \boldsymbol{x}_k$$

is convergent, and its limit $\lim_{k \to \infty} \boldsymbol{s}_k$ is called sum of the series.

Notice that we write also

$$\sum_{k=0}^{\infty} \boldsymbol{x}_k = \lim_{k \to \infty} \boldsymbol{s}_k.$$

The terminology used in the case of a normed space is the same used for real/complex series. It is possible to generalize Definition 2.8 to take into account sums where the index starts from a value different from zero. Hence, given $p \in \mathbb{N}^*$, we can define the series

$$\sum_{k=p}^{\infty} \boldsymbol{x}_k$$

by introducing the partial sums $\boldsymbol{s}_k = \boldsymbol{x}_p + \boldsymbol{x}_{p+1} + \ldots + \boldsymbol{x}_k$. This definition is well posed and consistent with the definition of sequence that we gave. Indeed, if $(\boldsymbol{x}_k)_{k \in A}$ is a sequence of vectors of $\boldsymbol{X}$, where A is a subset of $\mathbb{N}$ obtained by removing from $\mathbb{N}$ a finite set (so that $A = \mathbb{N} \setminus F$, where F is finite), then there exists $p \in \mathbb{N}$ such that $p + j \in A$ for all $j \in \mathbb{N}$: any $p > \max F$ will do the job.

Theorem 2.13 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a finite-dimensional normed vector space. Let $n \in \mathbb{N}^*$ be the dimension of $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of vectors of $\boldsymbol{X}$. Let $\mathcal{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n)$ be a basis of $\boldsymbol{X}$, and let $x_{k,i}$ denote the i-th component of $\boldsymbol{x}_k$ with respect to $\mathcal{B}$ (i.e., for all $k \in \mathbb{N}$ we have $\boldsymbol{x}_k = x_{k,1}\,\boldsymbol{b}_1 + x_{k,2}\,\boldsymbol{b}_2 + \ldots + x_{k,n}\,\boldsymbol{b}_n$). Then, the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ converges to $\hat{\boldsymbol{x}}$ if and only if, for every $i \in \{1, 2, \ldots, n\}$, the series $\sum_{k=0}^{\infty} x_{k,i}$ converges to the i-th coordinate $\hat{x}_i$ of $\hat{\boldsymbol{x}}$.*

The proof is left as an exercise (Exercise 2.11).

We can also introduce the notion of absolute convergence.

Definition 2.9 [Absolutely Convergent Series]

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of $\boldsymbol{X}$. We say that the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is absolutely convergent if the series $\sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|$ is convergent.

A series can be convergent without being absolutely convergent (this is what is often referred to as "conditional convergence"): for example, the alternating series $\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$ is convergent (as one can show using the Leibniz test[1]) and the sum of the series is $\ln 2$, but it is not absolutely convergent, because $\sum_{k=1}^{\infty} \left|(-1)^k \frac{1}{k}\right| = \sum_{k=1}^{\infty} \frac{1}{k}$ is the harmonic series. However, in the set of real and complex numbers, the opposite is true: if a series is absolutely convergent, it is also convergent. This result is true, in fact, for all Banach spaces, as the following theorem shows.

Theorem 2.14 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a Banach space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of $\boldsymbol{X}$. If the series $\sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|$ is convergent, then the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ convergent, and*

$$\left\| \sum_{k=0}^{\infty} \boldsymbol{x}_k \right\| \leq \sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|. \tag{2.13}$$

Proof We define the sequences of the partials sums $s_k = \sum_{i=0}^{k} \|\boldsymbol{x}_i\|$ and $\boldsymbol{\sigma}_k = \sum_{i=0}^{k} \boldsymbol{x}_i$. We need to show that, if $(s_k)_{k\in\mathbb{N}}$ is convergent, then $(\boldsymbol{\sigma}_k)_{k\in\mathbb{N}}$ is convergent. Since $(\boldsymbol{X}, \|\cdot\|)$ is assumed to be complete, it is sufficient to show that $(\boldsymbol{\sigma}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence. Hence, we must show that

$$\forall\, \varepsilon > 0, \quad \exists\, p \in \mathbb{N} : \quad \forall n, m > p, \quad \|\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_m\| < \varepsilon.$$

Let $\varepsilon > 0$. From Theorem 2.1, since $(s_k)_{k\in\mathbb{N}}$ is convergent, it is also a Cauchy sequence. Therefore, there exists $p \in \mathbb{N}$ such that, for every $n, m > p$, we have $|s_n - s_m| < \varepsilon$. Now, assuming with no loss of generality

[1] The Leibniz test applies to alternating series in the form $\sum_{k=1}^{\infty} (-1)^k a_k$. It states that if $(a_k)_{k\in\mathbb{N}}$ is non-increasing and it converges to 0, then the series is convergent.

that $n \geq m$ (what happens if $n < m$?), we obtain

$$|s_n - s_m| = \left| \sum_{k=0}^{n} \|\boldsymbol{x}_k\| - \sum_{k=0}^{m} \|\boldsymbol{x}_k\| \right| = \sum_{k=m+1}^{n} \|\boldsymbol{x}_k\| < \varepsilon.$$

Using the triangle inequality, we have

$$\|\boldsymbol{\sigma}_n - \boldsymbol{\sigma}_m\| = \left\| \sum_{k=0}^{n} \boldsymbol{x}_k - \sum_{k=0}^{m} \boldsymbol{x}_k \right\| = \left\| \sum_{k=m+1}^{n} \boldsymbol{x}_k \right\| \leq \sum_{k=m+1}^{n} \|\boldsymbol{x}_k\|,$$

which is $< \varepsilon$ from the former. It follows that $(\boldsymbol{\sigma}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, and therefore, since $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space, $(\boldsymbol{\sigma}_k)_{k\in\mathbb{N}}$ it is also convergent. It follows that the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is convergent.

In order to prove (2.13), we recall that if a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$, the numerical sequence $(\|\boldsymbol{x}_k\|)_{k\in\mathbb{N}}$ converges to $\|\hat{\boldsymbol{x}}\|$ (see Exercise 1.65). In symbols

$$\lim_{k\to\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}} \quad \Rightarrow \quad \lim_{k\to\infty} \|\boldsymbol{x}_k\| = \|\hat{\boldsymbol{x}}\| = \left\| \lim_{k\to\infty} \boldsymbol{x}_k \right\|. \tag{2.14}$$

This means that, in essence, we can interchange the limit with the norm. Now, we observe that

$$\left\| \sum_{k=0}^{n} \boldsymbol{x}_k \right\| \leq \sum_{k=0}^{n} \|\boldsymbol{x}_k\| \leq \sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|.$$

Using the comparison theorem (Theorem 1.25) by defining

$$\begin{aligned} f : A &\longrightarrow \mathbb{R} \\ n &\mapsto \left\| \sum_{k=0}^{n} \boldsymbol{x}_k \right\| \end{aligned} \qquad \text{and} \qquad \begin{aligned} g : A &\longrightarrow \mathbb{R} \\ n &\mapsto \sum_{k=0}^{\infty} \|\boldsymbol{x}_k\| \end{aligned}$$

where $A = \mathbb{N}$ and $\boldsymbol{x}_0 = \infty$ (notice that the second function is constant), we find

$$\lim_{n\to\infty} \left\| \sum_{k=0}^{n} \boldsymbol{x}_k \right\| \leq \sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|.$$

On the other hand, $\sum_{k=0}^{n} \boldsymbol{x}_k = \boldsymbol{\sigma}_n$, and the sequence of the partial sums $(\boldsymbol{\sigma}_k)_{k\in\mathbb{N}}$ is convergent by assumption, so that from (2.14) we find

$$\lim_{n\to\infty} \left\| \sum_{k=0}^{n} \boldsymbol{x}_k \right\| = \lim_{n\to\infty} \|\boldsymbol{\sigma}_n\| = \left\| \lim_{n\to\infty} \boldsymbol{\sigma}_n \right\| = \left\| \sum_{k=0}^{\infty} \boldsymbol{x}_k \right\|.$$

Example 2.11 Consider $(C([a,b];\mathbb{R}), \|\cdot\|_\infty)$. Theorem 2.14 ensures that, if $(f_k)_{k\in\mathbb{N}}$ is a sequence of functions in $C([a,b];\mathbb{R})$ such that

$$\sum_{k=0}^{\infty} \|f_k\|_\infty = \sum_{k=0}^{\infty} \max_{a\leq x\leq b} |f_k(x)| < +\infty,$$

then the series $\sum_{k=0}^{\infty} f_k$ is uniformly convergent to a continuous function.

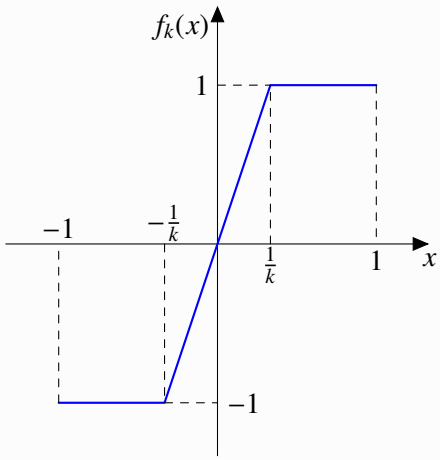
Example 2.12

We now show that Theorem 2.14 does not hold if the normed space is not complete: it is possible to have a series of the norms which is convergent, even if the series of vectors is not convergent. Consider again $C([-1,1],\mathbb{R})$ with the norm

$$\|f\|_1 = \int_{-1}^{1} |f(x)|\,\mathrm{d}x.$$

Consider again the sequence of functions $(f_k)_{k\in\mathbb{N}}$ defined by

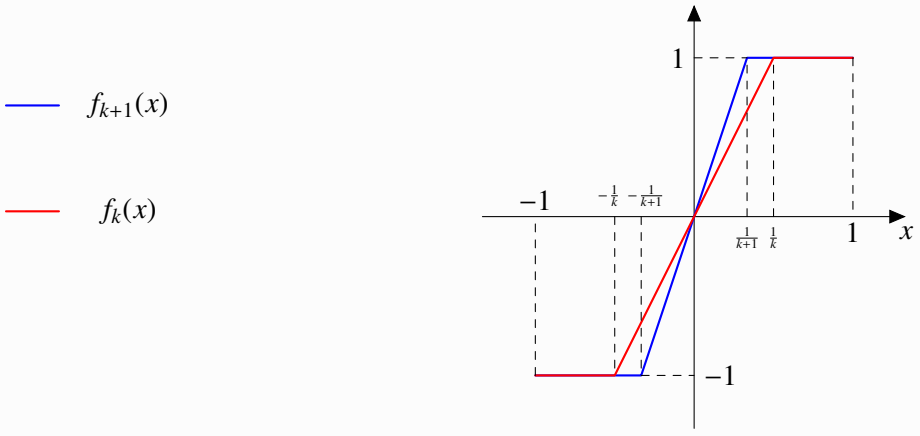
$$f_k(x) = \begin{cases} k\,x & \text{if } x \in \left(-\frac{1}{k}, \frac{1}{k}\right) \\ 1 & \text{if } x \in \left[\frac{1}{k}, 1\right] \\ -1 & \text{if } x \in \left[-1, -\frac{1}{k}\right] \end{cases}$$



We have shown that this sequence does not converge in the space $C([-1,1],\mathbb{R})$ with respect to the norm $\|\cdot\|_1$. In other words, the normed space $(C([-1,1],\mathbb{R}),\|\cdot\|_1)$ is not complete. If we define $u_k = f_{k+1} - f_k$, the sequence of partial sums of $\sum_{k=1}^{\infty} u_k$ is given, for all $k \in \mathbb{N}^*$, by

$$s_k = u_1 + u_2 + \ldots + u_k = \cancel{f_2} - f_1 + \cancel{f_3} - \cancel{f_2} + \ldots + f_{k+1} - \cancel{f_k} = f_{k+1} - f_1,$$

so that $\sum_{k=1}^{\infty} u_k$ does not converge because the sequence of partial sums does not converge. On the other hand, we show that the series of the norms converges. Indeed,



$$
\begin{aligned}
\|u_k\|_1 &= \|f_{k+1} - f_k\|_1 = \int_{-1}^{1} |f_{k+1}(x) - f_k(x)|\,\mathrm{d}x \\
&= \cancel{\int_{-1}^{-\frac{1}{k}} |(-1) - (-1)|\,\mathrm{d}x} + \int_{-\frac{1}{k}}^{-\frac{1}{k+1}} |-1 - k\,x|\,\mathrm{d}x + \int_{-\frac{1}{k+1}}^{\frac{1}{k+1}} \underbrace{|(k+1)\,x - k\,x|}_{=|x|}\,\mathrm{d}x \\
&\quad + \int_{\frac{1}{k+1}}^{\frac{1}{k}} |1 - k\,x|\,\mathrm{d}x + \cancel{\int_{\frac{1}{k}}^{1} |1 - 1|\,\mathrm{d}x} \\
&= \int_{-\frac{1}{k}}^{-\frac{1}{k+1}} (1 + k\,x)\,\mathrm{d}x + \int_{-\frac{1}{k+1}}^{0} -x\,\mathrm{d}x + \int_{0}^{\frac{1}{k+1}} x\,\mathrm{d}x + \int_{\frac{1}{k+1}}^{\frac{1}{k}} (1 - k\,x)\,\mathrm{d}x \\
&= 2\int_{0}^{\frac{1}{k+1}} x\,\mathrm{d}x + 2\int_{\frac{1}{k+1}}^{\frac{1}{k}} (1 - k\,x)\,\mathrm{d}x \\
&= 2\left[\frac{x^2}{2}\right]_0^{\frac{1}{k+1}} + 2\left[x - k\,\frac{x^2}{2}\right]_{\frac{1}{k+1}}^{\frac{1}{k}} \\
&= \frac{1}{(k+1)^2} + 2\,\frac{1}{k} - k\,\frac{1}{k^2} - 2\,\frac{1}{k+1} + k\,\frac{1}{(k+1)^2} = \frac{1}{k\,(k+1)}.
\end{aligned}
$$

Thus, the series $\displaystyle\sum_{k=1}^{\infty} \|u_k\|_1$ converges.

The converse of the statement of Theorem 2.14 holds.

Theorem 2.15 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Then, the following statements are equivalent:*

- **(A)** *$(\boldsymbol{X}, \|\cdot\|)$ is complete;*
- **(B)** *Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\boldsymbol{X}$ such that $\displaystyle\sum_{k=0}^{\infty} \|\boldsymbol{x}_k\|$ is convergent. Then, $\displaystyle\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is convergent.*

Proof The implication **(A)** $\Rightarrow$ **(B)** has already been proved in Theorem 2.14. We now prove that **(B)** $\Rightarrow$ **(A)**. Suppose that any sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ in $\boldsymbol{X}$ such that $\displaystyle\sum_{k\in\mathbb{N}} \|\boldsymbol{x}_k\|$ is convergent is such that $\displaystyle\sum_{k\in\mathbb{N}} \boldsymbol{x}_k$ is also convergent. We must show that $(\boldsymbol{X}, \|\cdot\|)$ is complete. To this end, consider a Cauchy sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$. We must show that this sequence is convergent. From the definition of Cauchy sequence, there exists p_0 such that, for all $n, m > p_0$, we have $\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < 1$. For the same reason, there exists p_1 such that, for all $n, m > p_1$, there holds

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < 2^{-1}.$$

and since $2^{-1} < 1$, it is not restrictive to assume $p_1 > p_0$ (why?). Likewise, there exists p_2 such that, for all $n, m > p_2$, there holds

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\| < 2^{-2},$$

and it is not restrictive to assume $p_2 > p_1$. Continuing in this way, we build a strictly increasing sequence $(p_i)_{i\in\mathbb{N}}$ of $\mathbb{N}$ such that

$$\forall\, i \in \mathbb{N},\ \forall\, n, m > p_i,\ \ \|\boldsymbol{x}_n - \boldsymbol{x}_m\| < 2^{-i}.$$

Since $p_{i+2} > p_{i+1} > p_i$, we deduce in particular

$$\forall\, i \in \mathbb{N},\ \|\boldsymbol{x}_{p_{i+2}} - \boldsymbol{x}_{p_{i+1}}\| < 2^{-i}.$$

Since $\sum_{i=0}^{\infty} 2^{-i} = \sum_{i=0}^{\infty} \left(\frac{1}{2}\right)^i$ is a geometric series with common ratio $1/2 < 1$, the series $\sum_{i=0}^{\infty} \|\boldsymbol{x}_{p_{i+2}} - \boldsymbol{x}_{p_{i+1}}\|$ is convergent (both series have non-negative terms, and we can apply the comparison test). Taking into account the assumption, the series

$$\sum_{i=0}^{\infty} (\boldsymbol{x}_{p_{i+2}} - \boldsymbol{x}_{p_{i+1}})$$

is convergent in $\boldsymbol{X}$. It follows that the sequence of the partial sums $(\boldsymbol{s}_k)_{k\in\mathbb{N}}$, where $\boldsymbol{s}_k = \sum_{i=0}^{k} (\boldsymbol{x}_{p_{i+2}} - \boldsymbol{x}_{p_{i+1}}) = \boldsymbol{x}_{p_{k+2}} - \boldsymbol{x}_{p_1}$, is convergent, and therefore the sequence $(\boldsymbol{x}_{p_i})_{i\in\mathbb{N}}$ is also convergent. Let $\hat{\boldsymbol{x}}$ denote its limit. We show that the given sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$, as well. Let $\varepsilon > 0$. Since $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence,

$$\exists\, p' \in \mathbb{N} : \quad \forall\, n, m > p', \quad \|\boldsymbol{x}_n - \boldsymbol{x}_m\| < \frac{\varepsilon}{2}. \tag{2.15}$$

Moreover, since $(\boldsymbol{x}_{p_i})_{i\in\mathbb{N}}$ converges to $\hat{\boldsymbol{x}}$,

$$\exists\, i' \in \mathbb{N} : \quad \forall\, i > i', \quad \|\boldsymbol{x}_{p_i} - \hat{\boldsymbol{x}}\| < \frac{\varepsilon}{2}. \tag{2.16}$$

Let $j > i'$ be such that $p_j > p'$ (and this is possible because $(p_i)_{i\in\mathbb{N}}$ is strictly increasing, and therefore divergent). We can choose $m = p_j$ and $n = k$ in (2.15) and $i = j$ in (2.16), so that for all $k > p'$

$$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \le \|\boldsymbol{x}_k - \boldsymbol{x}_{p_j}\| + \|\boldsymbol{x}_{p_j} - \hat{\boldsymbol{x}}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Example *2.13* Let $a, b \in \mathbb{R}$ be such that $a < b$. We show that the space

$$C^1([a,b],\mathbb{R}) = \{f : [a,b] \longrightarrow \mathbb{R} \,|\, f \text{ is differentiable and } f' \text{ is continuous}\}$$

is complete with respect to the norm

$$\|f\| = \max_{a\le x\le b} |f(x)| + \max_{a\le x\le b} |f'(x)|.$$

(The fact that $\|\cdot\|$ is a norm is in Exercise 1.59). From the convergence of the series $\sum_{k=0}^{\infty} \|f_k\|$ it follows that the two series $\sum_{k=0}^{\infty} \max_{a\le x\le b} |f_k(x)|$ and $\sum_{k=0}^{\infty} \max_{a\le x\le b} |f_k'(x)|$ are convergent. Thus, the series $\sum_{k=0}^{\infty} f_k$ and $\sum_{k=0}^{\infty} f_k'$ are uniformly convergent. We notice at this point that a result analogous to Theorem 2.12 holds for the differentiation of series

$$\sum_{k=0}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_k(x) = \frac{\mathrm{d}}{\mathrm{d}x} \sum_{k=0}^{\infty} f_k(x),$$

by applying Theorem 2.12 to the partial sums. Hence, the sum f of the first series is differentiable, and its derivative is $\sum_{k=0}^{\infty} f_k'$. It follows that f is of class C^1, and that f is the sum of the considered series in the sense of the given norm.

Exercises

Exercise 2.11 Let $(\boldsymbol{X}, \|\cdot\|)$ be a finite-dimensional normed vector space. Let $n \in \mathbb{N}^*$ be the dimension of $\boldsymbol{X}$. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of vectors of $\boldsymbol{X}$. Let $\mathcal{B} = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n)$ be a basis of $\boldsymbol{X}$, and let $x_{k,i}$ denote the i-th component of $\boldsymbol{x}_k$

with respect to $\mathcal{B}$ (i.e., for all $k \in \mathbb{N}$ we have $\boldsymbol{x}_k = x_{k,1}\,\boldsymbol{b}_1 + x_{k,2}\,\boldsymbol{b}_2 + \ldots + x_{k,n}\,\boldsymbol{b}_n$). Show that the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ converges to $\hat{\boldsymbol{x}}$ if and only if, for every $i \in \{1, 2, \ldots, n\}$, the series $\sum_{k=0}^{\infty} x_{k,i}$ converges to the i-th coordinate $\hat{x}_i$ of $\hat{\boldsymbol{x}}$ (Hint: use Theorem 1.45). [Solution: (2.11)]

Exercise 2.12 Consider the set c_0 of all real sequences $(x_k)_{k\in\mathbb{N}}$ which are convergent and whose limit is 0, i.e., $\lim_{k\to\infty} x_k = 0$. Given $x \in c_0$, we define the norm

$$\|x\|_\infty \stackrel{\text{def}}{=} \sup_{k\in\mathbb{N}} |x_k|.$$

Let c_{00} be the set of sequences with a finite number of non-zero terms.

(i) Show that c_0 is a vector space with respect to the usual addition and scalar multiplication of functions. What is the zero vector of this vector space?

(ii) Show that $(c_0, \|\cdot\|_\infty)$ is a normed space;

(iii) Prove that $(c_0, \|\cdot\|_\infty)$ is a Banach space;

(iv) Show that c_{00} is a subspace of c_0 but it is not closed in $(c_0, \|\cdot\|_\infty)$.

(v) Show that $(c_{00}, \|\cdot\|_\infty)$ is not a Banach space.

[Solution: (2.12)]

Exercise 2.13 Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. Let $p \in \mathbb{R}$. We define the ℓ^p space as the space

$$\ell^p(\mathbb{N}, \mathbb{F}) = \left\{ x : \mathbb{N} \longrightarrow \mathbb{F} \,\middle|\, \sum_{k=0}^{\infty} |x_k|^p < +\infty \right\}. \tag{2.17}$$

(i) Show that $\ell^p(\mathbb{N}, \mathbb{F})$ is a subspace of c_0, where c_0 is defined in Exercise 2.12;

(ii) Show that if $p > 1$, the space $\ell^p(\mathbb{N}, \mathbb{F})$ is a normed space if equipped with the norm

$$\|x\|_p = \Big(\sum_{k=0}^{\infty} |x_k|^p \Big)^{\frac{1}{p}}. \tag{2.18}$$

(Hint: you can adapt the proof of Theorem 1.4, generalizing Hölder's inequality by using Young's inequality, and then obtaining Minkowski's inequality for this norm). Show that if $p < 1$, (2.18) is not a norm.

(iii) Show that $(\ell^p(\mathbb{N}, \mathbb{F}), \|\cdot\|_p)$ is a Banach space.

[Solution: (2.13)]

Exercise 2.14 Let $\mathbb{F}$ be $\mathbb{R}$ or $\mathbb{C}$. We define the ℓ^∞ space as the space

$$\ell^\infty(\mathbb{N}, \mathbb{F}) = \left\{ x : \mathbb{N} \longrightarrow \mathbb{F} \,\middle|\, x \text{ is bounded} \right\}. \tag{2.19}$$

(i) Show that $\ell^\infty(\mathbb{N}, \mathbb{F})$ is a normed space if equipped with the norm

$$\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|.$$

(ii) Show that $(\ell^\infty(\mathbb{N},\mathbb{F}), \|\cdot\|_\infty)$ is a Banach space;

(iii) Show that c_0 defined in Exercise 2.12 is a closed subspace of $(\ell^\infty(\mathbb{N},\mathbb{F}), \|\cdot\|_\infty)$;

(iv) Show that the closure of c_{00}, defined in Exercise 2.12, in $(\ell^\infty(\mathbb{N},\mathbb{F}), \|\cdot\|_\infty)$ is c_0, i.e., $\overline{c_{00}} = c_0$.

[Solution: (2.14)]

Exercise 2.15 Let $p \in [1, +\infty)$. Show that $\ell^p(\mathbb{N};\mathbb{F}) \subset \ell^\infty(\mathbb{N};\mathbb{F})$. [Solution: (2.15)]

Exercise 2.16 Let $p_1, p_2 \in [1, +\infty]$, with $p_1 < p_2$. Let $x \in \ell^{p_1}(\mathbb{N};\mathbb{F}) \cap \ell^{p_2}(\mathbb{N};\mathbb{F})$. Show that $x \in \ell^p(\mathbb{N};\mathbb{F})$ for every $p \in [p_1, p_2]$. [Solution: (2.16)]

Exercise 2.17 Let $p, q \in [1, +\infty]$, with $p < q$. Show that $\ell^p(\mathbb{N};\mathbb{F}) \subset \ell^q(\mathbb{N};\mathbb{F})$. [Solution: (2.17)]

We conclude this section with the notion of unconditionally and conditionally convergent series. Loosely speaking, a series is unconditionally convergent if the series obtained by arbitrarily rearranging the terms of the original series is convergent.

Definition 2.10 [Unconditionally Convergent Series]

Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of $\boldsymbol{X}$. We say that the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is **unconditionally convergent** *if, for any bijection $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$, the series $\sum_{k=0}^{\infty} \boldsymbol{x}_{\sigma(k)}$ is convergent. If the series $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is convergent but it is not unconditionally convergent, we say that it is* **conditionally convergent**.

Example 2.14 We have already recalled that the alternating harmonic series

$$\sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$$

is convergent (in view of Leibniz's test). It is possible to show that the sum of this series is $\ln 2$. By changing the order of the terms of the series by anticipating the terms with even index, we obtain a positively divergent series, whereas by anticipating the terms with odd index the sum tends to $-\infty$. Moreover, the rearrangement of the terms of the series $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ in which the repeating pattern of signs is two non-negative terms followed by one negative term $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \ldots$ converges to $\frac{3}{2} \log 2$.

It is possible to show that an absolutely convergent series in a Banach space is unconditionally convergent. Riemann showed that for series in $\mathbb{R}^n$, the converse is true: an unconditionally convergent is absolutely convergent.

Hence, for finite-dimensional Banach spaces, absolute convergence and unconditional convergence are equivalent notions. On the contrary, it is possible to show that every infinite-dimensional Banach space contains an unconditionally convergent series that does not converge absolutely.

Example 2.15

In the Banach space $(\ell^\infty(\mathbb{N},\mathbb{F}), \|\cdot\|_2)$ consider the sequence $(x^{(j)})_{j\in\mathbb{N}^*}$ defined , for all $i \in \mathbb{N}^*$, by

$$x_i^{(j)} = \begin{cases} \frac{1}{i} & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

Hence, loosely speaking, we have

$$\begin{aligned} x^1 &= (1,0,0,0,\ldots) \\ x^2 &= (0,1/2,0,0,\ldots) \\ x^3 &= (0,0,1/3,0,\ldots) \\ &\vdots \end{aligned}$$

It is easy to see that

$$\sum_{j=1}^{\infty} x^{(j)} \tag{2.20}$$

is not absolutely convergent, because

$$\sum_{j=1}^{\infty} \|x^{(j)}\|_2 = \sum_{j=1}^{\infty} \frac{1}{j}$$

is the harmonic series, which is divergent. However, we show that (2.20) is unconditionally convergent to the sequence $\hat{x} \in \ell^\infty(\mathbb{N},\mathbb{F})$, where for all $i \in \mathbb{N}^*$

$$\hat{x}_i = 1/i,$$

i.e., $\hat{x} = (1, 1/2, 1/3, 1/4, \ldots)$. Consider a rearrangement given by a bijection $\sigma : \mathbb{N}^* \longrightarrow \mathbb{N}^*$. We need to show that

$$\forall\, \varepsilon > 0,\ \exists\, M \in \mathbb{N} :\quad \forall\, M' > M,\ \left\| x - \sum_{k=1}^{M'} x_{\sigma(k)} \right\| < \varepsilon.$$

Let $\varepsilon > 0$. Since

$$\sum_{k=1}^{\infty} \frac{1}{k^2}$$

is convergent,

$$\exists\, N \in \mathbb{N} :\quad \sum_{k=N+1}^{\infty} \frac{1}{k^2} < \varepsilon^2.$$

Let $M = \max\{\sigma^{-1}(1), \sigma^{-1}(2), \ldots, \sigma^{-1}(N)\}$. Let the rearrangement $y^i = x^{(\sigma(i))}$ for all $i \in \mathbb{N}^*$, and we notice that for all $i \leq N$ we have $\sigma^{-1}(i) \leq M$, which implies $y^{\sigma^{-1}(i)} \in \{y^1, y^2, \ldots, y^M\}$, and so $x^{\sigma(\sigma^{-1}(i))} = x^i \in \{y^1, y^2, \ldots, y^M\}$. This means that $x^1, x^2, \ldots, x^N \in \{y^1, y^2, \ldots, y^M\}$, and for all $M' \geq M$ we have $x^1, x^2, \ldots, x^N \in \{y^1, y^2, \ldots, y^{M'}\}$, so that

$$\begin{aligned} \left\| x - \sum_{k=1}^{M'} x^{\sigma(k)} \right\| &\leq \left\| x - \sum_{k=1}^{N} x^k \right\| \\ &= \left\| \sum_{k=N+1}^{\infty} x^k \right\| = \Big(\sum_{k=N+1}^{\infty} \frac{1}{k^2} \Big)^{1/2} < \varepsilon. \end{aligned}$$

2.4 Linear operators

We now consider the case of linear operators, which are essentially linear transformations. However, when using "linear operator" instead of "linear map", we put our emphasis on the fact that these linear transformations can act on vector spaces of functions, which are in general infinite dimensional.

Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be two normed vector spaces over the same field $\mathbb{F}$. We recall that an $\mathbb{F}$-linear transformation between $\boldsymbol{X}$ and $\boldsymbol{Y}$ is a function $T : \boldsymbol{X} \longrightarrow \boldsymbol{Y}$ which preserves the addition and the scalar multiplication:

(A) $T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2)$ for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$;

(B) $T(k\,\boldsymbol{x}) = k\,T(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{X}$ and for all $k \in \mathbb{F}$.

In particular, we notice that an $\mathbb{F}$-linear transformation T transforms the zero vector of $\boldsymbol{X}$ into the zero vector in $\boldsymbol{Y}$, i.e., $T(\mathbf{0}_X) = \mathbf{0}_Y$: for example use **(B)** with $k = 0$. We say that T is an $\mathbb{F}$-linear form if $\boldsymbol{Y}$ is the field $\mathbb{F}$ (recall that a generic field $\mathbb{F}$ can always be viewed as a vector space over itself).

We use the following notation:

$$\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y}) = \{T : \boldsymbol{X} \longrightarrow \boldsymbol{Y} \,|\, T \text{ is an } \mathbb{F}\text{-linear transformation}\}.$$

This set is, itself, a vector space over $\mathbb{F}$ with the standard sum and multiplication of functions. Indeed, consider two linear transformations $T, S \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$, which by definition satisfy

(i) $\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$, $T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2)$ and $S(\boldsymbol{x}_1 + \boldsymbol{x}_2) = S(\boldsymbol{x}_1) + S(\boldsymbol{x}_2)$;

(ii) $\forall\, \boldsymbol{x} \in \boldsymbol{X}$, $\forall\, k \in \mathbb{F}$, $T(k\,\boldsymbol{x}) = k\,T(\boldsymbol{x})$ and $S(k\,\boldsymbol{x}) = k\,S(\boldsymbol{x})$,

We can define the sum of linear transformations T and S and the multiplication of a linear transformation T by a scalar $\alpha \in \mathbb{F}$ as

$$\boxed{\begin{aligned}(T + S)(\boldsymbol{x}) &\stackrel{\text{def}}{=} T(\boldsymbol{x}) + S(\boldsymbol{x}) \\ (\alpha\, T)(\boldsymbol{x}) &\stackrel{\text{def}}{=} \alpha\, T(\boldsymbol{x})\end{aligned}}$$

To show that $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ is a vector space over $\mathbb{F}$, we start showing that it is closed under addition and scalar multiplication. Closure under addition requires that $T + S$ be still an element of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$, and to that end $T + S$ must preserve the addition and the scalar multiplication. Indeed,

$$\begin{aligned}\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}, \qquad (T + S)(\boldsymbol{x}_1 + \boldsymbol{x}_2) &= T(\boldsymbol{x}_1 + \boldsymbol{x}_2) + S(\boldsymbol{x}_1 + \boldsymbol{x}_2) \\ &= T(\boldsymbol{x}_1) + T(\boldsymbol{x}_2) + S(\boldsymbol{x}_1) + S(\boldsymbol{x}_2) \\ &= (T + S)(\boldsymbol{x}_1) + (T + S)(\boldsymbol{x}_2),\end{aligned}$$

thus $T + S$ preserves the addition, and

$$\begin{aligned}\forall\, \boldsymbol{x} \in \boldsymbol{X},\ \forall\, k \in \mathbb{F}, \qquad (T + S)(k\,\boldsymbol{x}) &= T(k\,\boldsymbol{x}) + S(k\,\boldsymbol{x}) = k\,T(\boldsymbol{x}) + k\,S(\boldsymbol{x}) \\ &= k\,(T(\boldsymbol{x}) + S(\boldsymbol{x})) = k\,(T + S)(\boldsymbol{x}),\end{aligned}$$

so that $T + S$ preserves the scalar multiplication.

Closure of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ under scalar multiplication requires $\alpha\, T$ to be an element of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$, and so $\alpha\, T$ must preserve the addition and the scalar multiplication as well. Indeed,

$$\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}, \qquad (\alpha\, T)(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \alpha\, T(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \alpha\, T(\boldsymbol{x}_1) + \alpha\, T(\boldsymbol{x}_2) = (\alpha\, T)(\boldsymbol{x}_1) + (\alpha\, T)(\boldsymbol{x}_2),$$

thus αT preserves the addition, and to see that it preserves the scalar multiplication we only need to observe that

$$\forall \boldsymbol{x} \in \boldsymbol{X},\ \forall k \in \mathbb{F}, \qquad (\alpha T)(k\boldsymbol{x}) = \alpha T(k\boldsymbol{x}) = \alpha k T(\boldsymbol{x}) = k\alpha T(\boldsymbol{x}) = k(\alpha T)(\boldsymbol{x}).$$

All the other axioms in the definition of vector space are easily shown to be verified: they are automatically inherited from the corresponding properties of the standard sum and multiplication. For example, the associative property of the addition requires that

$$\forall T_1, T_2, T_3 \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y}),\quad T_1 + (T_2 + T_3) = (T_1 + T_2) + T_3.$$

Recalling that two functions (over the same domain and codomain) are equal if and only if they take the same values for every value of the independent variable, the latter is satisfied if and only if

$$\forall \boldsymbol{x} \in \boldsymbol{X}, \quad (T_1 + (T_2 + T_3))(\boldsymbol{x}) = ((T_1 + T_2) + T_3)(\boldsymbol{x}).$$

This equation clearly holds: if $\boldsymbol{x} \in \boldsymbol{X}$ we have

$$\begin{aligned}(T_1 + (T_2 + T_3))(\boldsymbol{x}) &= T_1(\boldsymbol{x}) + (T_2 + T_3)(\boldsymbol{x})\\ &= T_1(\boldsymbol{x}) + (T_2(\boldsymbol{x}) + T_3(\boldsymbol{x}))\\ &= (T_1(\boldsymbol{x}) + T_2(\boldsymbol{x})) + T_3(\boldsymbol{x})\\ &= (T_1 + T_2)(\boldsymbol{x}) + T_3(\boldsymbol{x}) = ((T_1 + T_2) + T_3)(\boldsymbol{x}),\end{aligned}$$

where we have exploited the associative property of the vector space $\boldsymbol{Y}$. All the other axioms can be proved in the same way. Notice that the zero element of the vector space $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ is the zero linear transformation:

$$\begin{aligned}0 : \boldsymbol{X} &\longrightarrow \boldsymbol{Y}\\ \boldsymbol{x} &\mapsto \mathbf{0}_{\boldsymbol{Y}},\end{aligned}$$

where $\mathbf{0}_{\boldsymbol{Y}}$ denotes the origin in $\boldsymbol{Y}$. We now ask ourselves what the dimension of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ is. We focus on the case of finite-dimensional vector spaces. We recall that there is an isomorphism between the vector space of $\mathbb{F}$-linear transformations from $\boldsymbol{X}$ to $\boldsymbol{Y}$ and the vector space of $m \times n$ matrices over $\mathbb{F}$: once a basis for $\boldsymbol{X}$ and a basis for $\boldsymbol{Y}$ have been assigned, every element of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ can be written in one and only one way as a matrix in $\mathbb{F}^{m\times n}$ and, conversely, every matrix in $\mathbb{F}^{m\times n}$ can be thought of as a matrix that represents a linear transformation of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ with respect to the given ordered bases. Since the dimension of $\mathbb{F}^{m\times n}$ is $n \cdot m$, we expect the dimension of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ to be equal to $n \cdot m$ as well. The following result proves this result in a more rigorous fashion, and does so in a way to also produce a basis for $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$.

Theorem 2.16 *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two finite-dimensional vector spaces over the field $\mathbb{F}$, of dimensions n and m, respectively. Then,*

$$\dim_{\mathbb{F}}(\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})) = n \cdot m.$$

Proof Let $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a basis for $\boldsymbol{X}$ and let $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m)$ be a basis for $\boldsymbol{Y}$. We construct a basis $\{T_{i,j}\}_{\substack{i\in\{1,\ldots,m\}\\ j\in\{1,\ldots,n\}}}$ for $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$, and we show that such basis consists of $n \cdot m$ linear transformations. A linear transformation can be assigned in a unique way by specifying how it operates on the vectors of a basis of $\boldsymbol{X}$. Hence, for all $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$ we define $T_{i,j}$ by assigning $T_{i,j}(\boldsymbol{x}_1), T_{i,j}(\boldsymbol{x}_2), \ldots, T_{i,j}(\boldsymbol{x}_n)$. Specifically, for all

$i \in \{1, \dots, m\}$ and $j, k \in \{1, \dots, n\}$ let us define $T_{i,j} \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ to be such that

$$T_{i,j}(\boldsymbol{x}_k) = \delta_{k,j}\, \boldsymbol{y}_i, \qquad \delta_{k,j} \overset{\text{def}}{=} \begin{cases} 1 & k = j \\ 0 & k \neq j, \end{cases}$$

where the symbol $\delta_{k,j}$ is the so-called *Kronecker delta*. This is equivalent to choosing, for $k = 1$,

$$\begin{array}{llll} T_{1,1}(\boldsymbol{x}_1) = \boldsymbol{y}_1 & T_{2,1}(\boldsymbol{x}_1) = \boldsymbol{y}_2 & & T_{m,1}(\boldsymbol{x}_1) = \boldsymbol{y}_m \\ T_{1,2}(\boldsymbol{x}_1) = \boldsymbol{0} & T_{2,2}(\boldsymbol{x}_1) = \boldsymbol{0} & & T_{m,2}(\boldsymbol{x}_1) = \boldsymbol{0} \\ \vdots & \vdots & \cdots & \vdots \\ T_{1,n}(\boldsymbol{x}_1) = \boldsymbol{0} & T_{2,n}(\boldsymbol{x}_1) = \boldsymbol{0} & & T_{m,n}(\boldsymbol{x}_1) = \boldsymbol{0}, \end{array}$$

where, for readability, we denoted by $\boldsymbol{0}$ the zero vector of $\boldsymbol{Y}$. For $k = 2$

$$\begin{array}{llll} T_{1,1}(\boldsymbol{x}_2) = \boldsymbol{0} & T_{2,1}(\boldsymbol{x}_2) = \boldsymbol{0} & & T_{m,1}(\boldsymbol{x}_2) = \boldsymbol{0} \\ T_{1,2}(\boldsymbol{x}_2) = \boldsymbol{y}_1 & T_{2,2}(\boldsymbol{x}_2) = \boldsymbol{y}_2 & & T_{m,2}(\boldsymbol{x}_2) = \boldsymbol{y}_m \\ \vdots & \vdots & \cdots & \vdots \\ T_{1,n}(\boldsymbol{x}_2) = \boldsymbol{0} & T_{2,n}(\boldsymbol{x}_2) = \boldsymbol{0} & & T_{m,n}(\boldsymbol{x}_2) = \boldsymbol{0}, \end{array}$$

and so on, up to $k = n$:

$$\begin{array}{llll} T_{1,1}(\boldsymbol{x}_n) = \boldsymbol{0} & T_{2,1}(\boldsymbol{x}_n) = \boldsymbol{0} & & T_{m,1}(\boldsymbol{x}_n) = \boldsymbol{0} \\ T_{1,2}(\boldsymbol{x}_n) = \boldsymbol{0} & T_{2,2}(\boldsymbol{x}_n) = \boldsymbol{0} & & T_{m,2}(\boldsymbol{x}_n) = \boldsymbol{0} \\ \vdots & \vdots & \cdots & \vdots \\ T_{1,n}(\boldsymbol{x}_n) = \boldsymbol{y}_1 & T_{2,n}(\boldsymbol{x}_n) = \boldsymbol{y}_2 & & T_{m,n}(\boldsymbol{x}_n) = \boldsymbol{y}_m. \end{array}$$

We show that $\{T_{i,j}\}_{\substack{i \in \{1,\dots,m\} \\ j \in \{1,\dots,n\}}}$ are linearly independent. Consider a zero linear combination

$$\sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_{i,j}\, T_{i,j} = \boldsymbol{0}_{\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X};\boldsymbol{Y})},$$

where $\alpha_{i,j} \in \mathbb{F}$ for all $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Thus, for every $k \in \{1, 2, \dots, n\}$ we have

$$\sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_{i,j}\, T_{i,j}(\boldsymbol{x}_k) = \sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_{i,j}\, \delta_{k,j}\, \boldsymbol{y}_i = \sum_{i=1}^{m} \alpha_{i,k}\, \boldsymbol{y}_i = \boldsymbol{0}_{\boldsymbol{Y}}.$$

Since $(\boldsymbol{y}_1, \dots, \boldsymbol{y}_m)$ is a basis for $\boldsymbol{Y}$, and since the latter holds true for every $k \in \{1, 2, \dots, n\}$, then $\alpha_{i,j} = 0$ for all $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$, which shows that $\{T_{i,j}\}_{\substack{i \in \{1,\dots,m\} \\ j \in \{1,\dots,n\}}}$ is linearly independent. We show that $\{T_{i,j}\}_{\substack{i \in \{1,\dots,m\} \\ j \in \{1,\dots,n\}}}$ spans $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$. Let $T \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$. We prove that we can write T as

$$T = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}, \tag{2.21}$$

or, which is the same,

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad T(\boldsymbol{x}) = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(\boldsymbol{x}).$$

Let $\boldsymbol{x} = k_1\, \boldsymbol{x}_1 + \ldots + k_n\, \boldsymbol{x}_n \in \boldsymbol{X}$, so that

$$T(\boldsymbol{x}) = T\,(k_1\, \boldsymbol{x}_1 + \ldots + k_n\, \boldsymbol{x}_n) = k_1\, T(\boldsymbol{x}_1) + \ldots + k_n\, T(\boldsymbol{x}_n).$$

Each term $T(\boldsymbol{x}_k)$ (for $k \in \{1, 2, \ldots, n\}$) in the sum is a vector of $\boldsymbol{Y}$, and can therefore be written as

$$T(\boldsymbol{x}_k) = \sum_{i=1}^{m} c_{i,k}\, \boldsymbol{y}_i = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, \delta_{k,j}\, \boldsymbol{y}_i = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(\boldsymbol{x}_k).$$

It follows that

$$\begin{aligned} T(\boldsymbol{x}) &= k_1\, T(\boldsymbol{x}_1) + \ldots + k_n\, T(\boldsymbol{x}_n) \\ &= k_1 \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(\boldsymbol{x}_1) + \ldots + k_n \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(\boldsymbol{x}_n) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(k_1\, \boldsymbol{x}_1 + \ldots + k_n\, \boldsymbol{x}_n) = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{i,j}\, T_{i,j}(\boldsymbol{x}), \end{aligned}$$

as required. Thus, $\{T_{i,j}\}_{\substack{i \in \{1,\ldots,m\} \\ j \in \{1,\ldots,n\}}}$ is a basis for $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$, which in turn implies that the dimension of $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ is $n \cdot m$.

In the proof of Theorem 2.16, we have constructed a basis of linear transformations that corresponds to the standard basis of $\mathbb{F}^{m \times n}$. In fact, consider $T_{1,1}$: we have obtained

$$T_{1,1}(\boldsymbol{x}_1) = \boldsymbol{y}_1, \qquad T_{1,1}(\boldsymbol{x}_2) = T_{1,1}(\boldsymbol{x}_3) = \ldots = T_{1,1}(\boldsymbol{x}_n) = \boldsymbol{0},$$

so that the corresponding matrix $A_{1,1}$ is the matrix

$$A_{1,1} = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Likewise,

$$T_{2,1}(\boldsymbol{x}_1) = \boldsymbol{y}_2, \qquad T_{2,1}(\boldsymbol{x}_2) = T_{2,1}(\boldsymbol{x}_3) = \ldots = T_{2,1}(\boldsymbol{x}_n) = \boldsymbol{0},$$

so that the corresponding matrix $A_{2,1}$ is the matrix

$$A_{2,1} = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Similarly,

$$T_{1,2}(\boldsymbol{x}_1) = \boldsymbol{0}, \qquad T_{1,2}(\boldsymbol{x}_2) = \boldsymbol{y}_1, \qquad T_{1,2}(\boldsymbol{x}_3) = \ldots = T_{1,2}(\boldsymbol{x}_n) = \boldsymbol{0},$$

so that the corresponding matrix $A_{1,2}$ is the matrix

$$A_{1,2} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \end{bmatrix}.$$

Continuing in this way, it is clear that there is a correspondence between the linear transformations $T_{i,j}$ with the matrices $A_{i,j}$ for $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$.

Now we abandon the restriction of finite dimensionality for the vector spaces $\boldsymbol{X}$ and $\boldsymbol{Y}$. We introduce the notion of *bounded operator*.

Definition 2.11 [Bounded Operator]
Let $(\boldsymbol{X}, \|\cdot\|_X)$ *and* $(\boldsymbol{Y}, \|\cdot\|_Y)$ *be two normed vector spaces over the same field* $\mathbb{F}$ *(*$\mathbb{R}$ *or* $\mathbb{C}$*). Let* $T \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$*. We say that* T *is* bounded *if*

$$\exists\, M \geq 0 : \ \forall\, \boldsymbol{x} \in \boldsymbol{X},\ \|T(\boldsymbol{x})\|_Y \leq M \|\boldsymbol{x}\|_X. \tag{2.22}$$

Bounded operators play an important role in many areas of science and engineering. We now study the relation between boundedness and continuity of linear operators. To this end, we define a new set $\mathcal{L}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ of all elements in $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ which are also continuous:

$$\mathcal{L}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y}) = \{T : \boldsymbol{X} \longrightarrow \boldsymbol{Y} \,|\, T \text{ is an } \mathbb{F}\text{-linear transformation and continuous}\}.$$

The following important result holds.

Theorem 2.17 *Let* $(\boldsymbol{X}, \|\cdot\|_X)$ *and* $(\boldsymbol{Y}, \|\cdot\|_Y)$ *be two normed vector spaces over the same field* $\mathbb{F}$ *(*$\mathbb{R}$ *or* $\mathbb{C}$*). Let* $T \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$*. The following statements are equivalent:*

(i) $T \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$*;*

(ii) T *is continuous at* $\mathbf{0}_X$*;*

(iii) T *is bounded.*

Proof The fact that *(i)* implies *(ii)* is obvious. We prove that *(ii)* implies *(iii)*. Let T be continuous at $\mathbf{0}_X$, so that

$$\exists\, \delta > 0 : \ \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|\boldsymbol{x}\|_X \leq \delta \ \Rightarrow \|T(\boldsymbol{x})\|_Y \leq 1,$$

where we used the fact that $T(\mathbf{0}_X) = \mathbf{0}_Y$. We are ready to prove *(iii)*, which is expressed by the inequality (2.22). If $\boldsymbol{x} = \mathbf{0}_X$, both sides of (2.22) are zero, and therefore the statement is true. Let $\boldsymbol{x} \in \boldsymbol{X}$ with $\boldsymbol{x} \neq \mathbf{0}_X$. Clearly,

$$\left\| \frac{\delta}{\|\boldsymbol{x}\|_X} \boldsymbol{x} \right\|_X = \delta \leq \delta,$$

from which

$$\left\| T\left(\frac{\delta}{\|\boldsymbol{x}\|_X} \boldsymbol{x} \right) \right\|_Y \leq 1,$$

and from the linearity of T

$$\left\| T\left(\frac{\delta}{\|\boldsymbol{x}\|_X} \boldsymbol{x} \right) \right\|_Y = \left\| \frac{\delta}{\|\boldsymbol{x}\|_X} T(\boldsymbol{x}) \right\|_Y = \frac{\delta}{\|\boldsymbol{x}\|_X} \|T(\boldsymbol{x})\|_Y,$$

so that $\|T(\boldsymbol{x})\|_Y \leq \delta^{-1} \|\boldsymbol{x}\|_X$. Hence, taking $M = \delta^{-1}$ proves (2.22), and this shows that *(iii)* is true.

We prove that *(iii)* implies *(i)*. Let $\boldsymbol{x}_0 \in \boldsymbol{X}$. We prove that T is continuous at $\boldsymbol{x}_0$. In other words, we show that

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \ \forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \delta \ \Rightarrow \|T(\boldsymbol{x}) - T(\boldsymbol{x}_0)\|_Y \leq \varepsilon.$$

By assumption, there exists $M \geq 0$ such that, for any element $\boldsymbol{x}$ of $\boldsymbol{X}$, satisfies $\|T(\boldsymbol{x})\|_Y \leq M \|\boldsymbol{x}\|_X$. Using the linearity of T

$$\|T(\boldsymbol{x}) - T(\boldsymbol{x}_0)\|_Y = \|T(\boldsymbol{x} - \boldsymbol{x}_0)\|_Y \leq M \|\boldsymbol{x} - \boldsymbol{x}_0\|_X.$$

If $M = 0$, then $\|T(\boldsymbol{x}) - T(\boldsymbol{x}_0)\|_Y \leq \varepsilon$ for every $\boldsymbol{x} \in \boldsymbol{X}$. If $M > 0$, we have $\|T(\boldsymbol{x}) - T(\boldsymbol{x}_0)\|_Y \leq \varepsilon$ if $M \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \varepsilon$, i.e., if

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \frac{\varepsilon}{M}.$$

The statement follows by choosing $\delta = \varepsilon/M$.

If the vector space $\boldsymbol{X}$ is finite dimensional (whether $\boldsymbol{Y}$ is finite dimensional or infinite dimensional), $\mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ coincides with $\mathcal{L}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$. In fact, if n is the dimension of $\boldsymbol{X}$, let $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ be a basis of $\boldsymbol{X}$. If $\boldsymbol{x} \in \boldsymbol{X}$, there exist $c_1, c_2, \ldots, c_n \in \mathbb{F}$ that allows us to write $\boldsymbol{x} = c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 + \ldots + c_n \boldsymbol{x}_n$. If we define $\|\boldsymbol{x}\|_1 = |c_1| + |c_2| + \ldots + |c_n|$, for all $T \in \mathrm{Hom}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ we have

$$\begin{aligned} \|T(\boldsymbol{x})\|_Y &= \|T(c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 + \ldots + c_n \boldsymbol{x}_n)\|_Y \\ &= \|c_1 T(\boldsymbol{x}_1) + c_2 T(\boldsymbol{x}_2) + \ldots + c_n T(\boldsymbol{x}_n)\|_Y \\ &\leq |c_1| \|T(\boldsymbol{x}_1)\|_Y + |c_2| \|T(\boldsymbol{x}_2)\|_Y + \ldots + |c_n| \|T(\boldsymbol{x}_n)\|_Y. \end{aligned}$$

Let

$$M = \max_{1 \leq j \leq n} \|T(\boldsymbol{x}_j)\|_Y.$$

We obtain

$$\|T(\boldsymbol{x})\|_Y \leq M (|c_1| + |c_2| + \ldots + |c_n|) = M \|\boldsymbol{x}\|_1.$$

Thus, $T \in \mathcal{L}_{\mathbb{F}}(\boldsymbol{X}; \boldsymbol{Y})$ if we consider the norm $\|\cdot\|_1$ in $\boldsymbol{X}$. However, since $\boldsymbol{X}$ is finite dimensional, a generic norm $\|\cdot\|_X$ will be equivalent to $\|\cdot\|_1$, i.e., there exists $c > 0$ such that $\|\boldsymbol{x}\|_1 \leq c \|\boldsymbol{x}\|_X$ for all $\boldsymbol{x} \in \boldsymbol{X}$, from which we deduce

$$\|T(\boldsymbol{x})\|_Y \leq M c \|\boldsymbol{x}\|_X$$

which proves the point using the result of the previous theorem.

Example *2.16* Consider the vector space $C([0,1]; \mathbb{R})$ with the norm $\|\cdot\|_1$. Let T be defined by

$$\begin{aligned} T : C([0,1]; \mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto f(1). \end{aligned}$$

The function T is a linear operator with respect to the standard addition and scalar multiplication of functions in $C([0,1]; \mathbb{R})$:

$$\begin{aligned} f + g : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto f(x) + g(x) \end{aligned} \quad \text{and} \quad \begin{aligned} k f : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto k f(x). \end{aligned}$$

Indeed, T preserves the addition since for all $f, g \in C([0,1]; \mathbb{R})$

$$T(f + g) = (f + g)(1) = f(1) + g(1) = T(f) + T(g),$$

and it preserves the scalar multiplication, because for all $k \in \mathbb{R}$ and $f \in C([0,1]; \mathbb{R})$

$$T(k f) = (k f)(1) = k f(1) = k T(f).$$

In particular, in this case, T is a linear form. We show that T is not continuous. For T to be continuous, there

must exist $M \geq 0$ such that for all $f \in C([0,1];\mathbb{R})$

$$|T(f)| = |f(1)| \leq M \int_0^1 |f(x)|\,\mathrm{d}x.$$

Let $k \in \mathbb{N}$. If we choose as f the function $f_k(x) = x^k$, we obtain

$$|T(f_k)| = 1 \leq M \int_0^1 |x^k|\,\mathrm{d}x = M \int_0^1 x^k\,\mathrm{d}x = M \left[\frac{1}{k+1}\,x^{k+1}\right]_0^1 = M\,\frac{1}{k+1}.$$

Thus, $M \geq k+1$. Since this has to be true for every $k \in \mathbb{N}$, the positive constant M does not exist. Hence, T is not continuous.

By taking, instead, $f_k(x) = \sqrt{k}\,x^k$, we still have $\|f_k\|_1 \xrightarrow{k\to\infty} 0$, since

$$\|f_k\|_1 = \int_0^1 \sqrt{k}\,x^k\,\mathrm{d}x = \sqrt{k}\left[\frac{1}{k+1}\,x^{k+1}\right]_0^1 = \frac{\sqrt{k}}{k+1},$$

but $\lim_{k\to\infty} |T(f_k)| = \lim_{k\to\infty} |f_k(1)| = \lim_{k\to\infty} \sqrt{k} = \infty$, since so that the sequence $(T(f_k))_{k\in\mathbb{N}}$ is not even bounded. Hence, in the case of an infinite dimensional vector space, a linear operator may not be bounded at each neighborhood of the origin.

Exercises

Exercise 2.18 Consider a vector space $\boldsymbol{X}$ with two equivalent norms $\|\cdot\|_{X,1}$ and $\|\cdot\|_{X,2}$, and a vector space $\boldsymbol{Y}$ over the same field with two equivalent norms $\|\cdot\|_{Y,1}$ and $\|\cdot\|_{Y,2}$. Let $T : \boldsymbol{X} \longrightarrow \boldsymbol{Y}$ be a linear operator. Show that if T is bounded with respect to the normed spaces $(\boldsymbol{X}, \|\cdot\|_{X,1})$ and $(\boldsymbol{Y}, \|\cdot\|_{Y,1})$, then it is bounded with respect to the normed spaces $(\boldsymbol{X}, \|\cdot\|_{X,2})$ and $(\boldsymbol{Y}, \|\cdot\|_{Y,2})$. [Solution: (2.18)]

Exercise 2.19 Let $\boldsymbol{X} = C^1([0,1],\mathbb{R})$ with the norm $\|\cdot\|_\infty$. Show that the function

$$\begin{aligned} T : C^1([0,1],\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto f'(0) \end{aligned}$$

is a linear form, and it is not continuous. Hint: apply the Sequential Continuity Theorem to the sequence of functions $\left(\frac{\sin(k\pi x)}{\sqrt{k}}\right)_{k\in\mathbb{N}^*}$. [Solution: (2.19)]

Exercise 2.20 Consider the normed vector space $(C([0,1],\mathbb{R}), \|\cdot\|_\infty)$. Let $x_0 \in [0,1]$. Prove that

$$\begin{aligned} T : C([0,1],\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto \int_0^1 f(x)\,\mathrm{d}x \end{aligned}$$

is a linear transformation between the vector spaces $C([0,1],\mathbb{R})$ and $\mathbb{R}$, and that it is continuous. [Solution: (2.20)]

Exercise 2.21 Consider the normed vector space $(C([-2,2],\mathbb{R}), \|\cdot\|_\infty)$. Prove that the functional

$$\begin{aligned} T : C([-2,2],\mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\mapsto f(0) + f(1) \end{aligned}$$

is a linear transformation between the vector spaces $C([-2,2],\mathbb{R})$ and $\mathbb{R}$ and it is continuous. To this end, recall the properties of the $\|\cdot\|_\infty$, the uniform convergence, and the sequential continuity theorem. [Solution: (2.21)]

Exercise 2.22

Let $(\boldsymbol{V},\|\cdot\|_{V,1})$ and $(\boldsymbol{W},\|\cdot\|_{W,1})$ be normed vector spaces. Let $T\in\mathcal{L}_{\mathbb{F}}(\boldsymbol{V},\boldsymbol{W})$, i.e., T is a continuous linear transformation from $(\boldsymbol{V},\|\cdot\|_{V,1})$ to $(\boldsymbol{W},\|\cdot\|_{W,1})$. Let $\|\cdot\|_{V,2}$ be a norm on $\boldsymbol{V}$ equivalent to $\|\cdot\|_{V,1}$, and let $\|\cdot\|_{W,2}$ be a norm on $\boldsymbol{W}$ equivalent to $\|\cdot\|_{W,1}$. Show that T is a continuous linear transformation from $(\boldsymbol{V},\|\cdot\|_{V,2})$ to $(\boldsymbol{W},\|\cdot\|_{W,2})$. [Solution: (2.22)]

Exercise 1.12

Consider the normed vector spaces $(\boldsymbol{V},\|\cdot\|_\infty)$ and $(\boldsymbol{W},\|\cdot\|_\infty)$, where $\boldsymbol{V}=C^1([0,1],\mathbb{R})$ and $\boldsymbol{W}=C([0,1],\mathbb{R})$. Prove that the linear operator

$$\begin{aligned}\varphi:\boldsymbol{V}&\longrightarrow\boldsymbol{W}\\ f&\mapsto f'\end{aligned}$$

is not continuous. (Hint: use the sequence $f_k(x)=\frac{\sin(k\pi x)}{\sqrt{k}}$). [Solution: (2.23)]

Exercise 2.24

Consider the space $\ell^\infty(\mathbb{N},\mathbb{F})$ defined in Exercise 2.14. Let $T:\ell^\infty(\mathbb{N},\mathbb{F})\longrightarrow\ell^\infty(\mathbb{N},\mathbb{F})$ be a function defined, for $x\in\ell^\infty(\mathbb{N},\mathbb{F})$, by

$$T(x_0,x_1,x_2,x_3,\ldots)=(x_1,x_2,x_3,x_4,\ldots).$$

(i) Show that T is a continuous linear operator;

(ii) Show that $S=\{x-T(x)\,\big|\,x\in\ell^\infty(\mathbb{N},\mathbb{F})\}$ is a subspace of $\ell^\infty(\mathbb{N},\mathbb{F})$;

(iii) Show that S is not closed. (Hint: consider the sequence $(z_k)_{k\in\mathbb{N}}$ given by $z_k=\frac{1}{k+1}$).

[Solution: (2.24)]

2.5 Hilbert spaces

In the previous sections, we introduced the notion of completeness of a normed space. We have seen that:

(i) in any normed space, every convergent sequence is a Cauchy sequence;

(ii) in the case of finite-dimensional normed spaces, a Cauchy sequence condition is also convergent;

(iii) in the case of infinite-dimensional vector spaces, the sufficiency of the Cauchy's condition for the convergence depends on the particular space considered: in some spaces, all the Cauchy sequences are convergent, in others there exist Cauchy sequences that are not convergent. We called the first ones *complete* normed spaces, or *Banach spaces*: a normed space is said to be complete, or Banach space, if every Cauchy sequence is convergent. It follows that every finite-dimensional normed space is complete.

In this section we introduce Hilbert spaces, which arise frequently in physics, applied mathematics and engineering: these spaces allow us to generalize results of linear algebra and calculus from Euclidean vector spaces to infinite-dimensional vector spaces.

2.5.1 Sesquilinear forms, quadratic forms and inner products

Before introducing the fundamental notion of Hilbert space, we recall the notion of inner product and inner product space. We start with the notion of *sesquilinear form*.[2]

Definition 2.12 [Sesquilinear Form]
Let $\boldsymbol{X}$ be a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). A sesquilinear form *is a function*

$$\begin{aligned} \mathfrak{b} : \boldsymbol{X} \times \boldsymbol{X} &\longrightarrow \mathbb{F} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) \end{aligned}$$

which satisfies the following two properties: for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{X}$ and for all $\alpha, \beta \in \mathbb{F}$,

- $\mathfrak{b}(\alpha\,\boldsymbol{x} + \beta\,\boldsymbol{y}, \boldsymbol{z}) = \alpha\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \beta\,\mathfrak{b}(\boldsymbol{y}, \boldsymbol{z})$; (Linearity in the First Argument)
- $\mathfrak{b}(\boldsymbol{x}, \alpha\,\boldsymbol{y} + \beta\,\boldsymbol{z}) = \bar{\alpha}\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) + \bar{\beta}\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{z})$. (Anti-Linearity in the Second Argument)

In other words, for a complex vector space, a sesquilinear form satisfies the linearity in the first argument and the anti-linearity in the second argument, while for a real vector space a sesquilinear form is *bilinear*.[3]

Example *2.17*

Let $\boldsymbol{X} = \mathbb{R}^n$, and consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, with $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$. The function $\mathfrak{b} : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = x_1\,y_1 + x_2\,y_2 + \ldots + x_n\,y_n$$

(i.e., the standard dot product in $\mathbb{R}^n$) is a sesquilinear form. Indeed, for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ we have

$$\begin{aligned} \mathfrak{b}(\alpha\,\boldsymbol{x} + \beta\,\boldsymbol{y}, \boldsymbol{z}) &= (\alpha\,x_1 + \beta\,y_1)\,z_1 + \ldots + (\alpha\,x_n + \beta\,y_n)\,z_n \\ &= \alpha\,(x_1\,z_1 + \ldots + x_n\,z_n) + \beta\,(y_1\,z_1 + \ldots + y_n\,z_n) = \alpha\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \beta\,\mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}), \end{aligned}$$

and

$$\begin{aligned} \mathfrak{b}(\boldsymbol{x}, \alpha\,\boldsymbol{y} + \beta\,\boldsymbol{z}) &= x_1\,(\alpha\,y_1 + \beta\,z_1) + \ldots + x_n\,(\alpha\,y_n + \beta\,z_n) \\ &= \alpha\,(x_1\,y_1 + \ldots + x_n\,y_n) + \beta\,(x_1\,z_1 + \ldots + x_n\,z_n) = \alpha\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) + \beta\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}). \end{aligned}$$

Exercise *2.25*

Show that if $\mathfrak{b} : \boldsymbol{X} \times \boldsymbol{X} \longrightarrow \mathbb{C}$ is a sesquilinear form, for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{w} \in \boldsymbol{X}$ and $a, b \in \mathbb{C}$

- $\mathfrak{b}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} + \boldsymbol{w}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \mathfrak{b}(\boldsymbol{x}, \boldsymbol{w}) + \mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}) + \mathfrak{b}(\boldsymbol{y}, \boldsymbol{w})$;
- $\mathfrak{b}(a\,\boldsymbol{x}, b\,\boldsymbol{y}) = a\,\bar{b}\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y})$.

[Solution: (2.25)]

Definition 2.13 [Hermitian Sesquilinear Form]
Let $\boldsymbol{X}$ be a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). A sesquilinear form $\mathfrak{b}$ is said to be Hermitian *if*

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \quad \mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = \overline{\mathfrak{b}(\boldsymbol{y}, \boldsymbol{x})}.$$

[2] The term "sesquilinear" originates from the Latin numerical prefix "sesqui-" meaning "one and a half".

[3] In some areas, and in particular in physics, there is a tendency to consider, *mutatis mutandis*, anti-linearity in the first argument and linearity in the second argument, in accordance with Dirac's bra-ket notation of quantum physics.

In the case where $\mathbb{F} = \mathbb{R}$, a Hermitian sesquilinear form is said to be symmetric.

Notice that for a symmetric sesquilinear form we have

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \quad \mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = \mathfrak{b}(\boldsymbol{y}, \boldsymbol{x}).$$

Definition 2.14 [Quadratic Form]
Let $\boldsymbol{X}$ be a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). A function $\mathfrak{q} : \boldsymbol{X} \longrightarrow \mathbb{R}$ is said to be a quadratic form *if there exists a Hermitian sesquilinear form $\mathfrak{b}$ on $\boldsymbol{X}$ such that*

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x}).$$

If $\mathfrak{q} : \boldsymbol{X} \longrightarrow \mathbb{R}$ is a quadratic form, then:

- *if $\mathbb{F} = \mathbb{R}$, we say that $\mathfrak{q}$ is a* real quadratic form*;*
- *if $\mathbb{F} = \mathbb{C}$, we say that $\mathfrak{q}$ is a* complex quadratic form*.*

We notice that the codomain of a quadratic form is $\mathbb{R}$ even if the vector space $\boldsymbol{X}$ is complex. In fact, since $\mathfrak{b}$ is assumed to be a Hermitian sesquilinear form, for all $\boldsymbol{x} \in \boldsymbol{X}$ we have

$$\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x}) = \overline{\mathfrak{b}(\boldsymbol{x}, \boldsymbol{x})} = \overline{\mathfrak{q}(\boldsymbol{x})},$$

and a complex number is equal to its complex conjugate if and only if it is a real number (why?).

Example 2.18 If $\boldsymbol{X} = \mathbb{R}^n$ and we consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, with $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$, then $\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = x_1\, y_1 + x_2\, y_2 + \ldots + x_n\, y_n$ (i.e., the standard dot product in $\mathbb{R}^n$) is a symmetric sesquilinear form. The corresponding quadratic form is $\mathfrak{q}(\boldsymbol{x}) = x_1^2 + x_2^2 + \ldots + x_n^2$.

Example 2.19 If $\boldsymbol{X} = \mathbb{R}^2$ and we consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$, with $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, then $\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = x_1\, y_1 - x_2\, y_2$ is a symmetric sesquilinear form (why? See Exercise 2.30). The corresponding quadratic form is $\mathfrak{q}(\boldsymbol{x}) = x_1^2 - x_2^2$.

Example 2.20 If $\boldsymbol{X} = \mathbb{C}^n$ over the field $\mathbb{F}$, and we consider $\boldsymbol{z}, \boldsymbol{w} \in \mathbb{C}^n$, with $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$, then $\mathfrak{b}(\boldsymbol{z}, \boldsymbol{w}) = \boldsymbol{w}^*\, \boldsymbol{z} = \bar{w}_1\, z_1 + \bar{w}_2\, z_2 + \ldots + \bar{w}_n\, z_n$ (i.e., the standard Euclidean product in $\mathbb{C}^n$) is a Hermitian sesquilinear form. The corresponding quadratic form is

$$\mathfrak{q}(\boldsymbol{z}) = \bar{z}_1\, z_1 + \bar{z}_2\, z_2 + \ldots + \bar{z}_n\, z_n = |z_1|^2 + |z_2|^2 + \ldots + |z_n|^2.$$

Exercise 2.26 Let $\boldsymbol{X} = C([0, 1], \mathbb{R}) = \{f : [0, 1] \longrightarrow \mathbb{R} \,|\, f \text{ is continuous}\}$ and consider the function

$$\mathfrak{b}(f, g) = \int_0^1 f(t)\, g(t)\, \mathrm{d}t,$$

where $f, g \in \boldsymbol{X}$. Show that this is a symmetric sesquilinear form. [Solution: (2.26)]

Exercise 2.27 If $\boldsymbol{X} = C([0, 1], \mathbb{C}) = \{f : [0, 1] \longrightarrow \mathbb{C} \,|\, f \text{ is continuous}\}$ and $f, g \in \boldsymbol{X}$, show that

$$\mathfrak{b}(f, g) = \int_0^1 \overline{g(t)}\, f(t)\, \mathrm{d}t$$

is a Hermitian sesquilinear form. [Solution: (2.27)]

Since a quadratic form is real valued, we can define its sign as follows.

Definition 2.15 [Sign of a Quadratic Form]
Let $\boldsymbol{X}$ be a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $\mathfrak{q} : \boldsymbol{X} \longrightarrow \mathbb{R}$ be a quadratic form. We say that $\mathfrak{q}$ is:

- positive semidefinite *if $\mathfrak{q}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \boldsymbol{X}$;*
- positive definite *if $\mathfrak{q}(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \boldsymbol{X} \setminus \{\mathbf{0}\}$;*
- negative semidefinite *if $\mathfrak{q}(\boldsymbol{x}) \leq 0$ for all $\boldsymbol{x} \in \boldsymbol{X}$;*
- negative definite *if $\mathfrak{q}(\boldsymbol{x}) < 0$ for all $\boldsymbol{x} \in \boldsymbol{X} \setminus \{\mathbf{0}\}$;*
- indefinite *if $\mathfrak{q}$ is not positive or negative semidefinite.*

Notice that a positive definite quadratic form is also positive semidefinite (which is consistent with the fact that a strictly positive real number is also non-negative), and similarly for negative (semi)-definite quadratic forms.

We observe also that for a quadratic form $\mathfrak{q}$ to be indefinite, there must exists $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{X}$ such that $\mathfrak{q}(\boldsymbol{x}_1) > 0$ and $\mathfrak{q}(\boldsymbol{x}_2) < 0$.

Example *2.21*

Consider the vector space $\mathbb{R}^n$ and the sesquilinear form given by the dot product: $\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$. The corresponding quadratic form

$$\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x}) = x_1^2 + x_2^2 + \ldots + x_n^2$$

is positive definite. Indeed, $\mathfrak{q}(\boldsymbol{x}) \geq 0$ for every $\boldsymbol{x} \in \mathbb{R}^n$, and for $\mathfrak{q}(\boldsymbol{x})$ to be equal zero we need $x_i = 0$ for all $i \in \{1, 2, \ldots, n\}$, i.e., we need $\boldsymbol{x} = \mathbf{0}$.

Example *2.22*

Consider the vector space $\mathbb{C}^n$ and the sesquilinear form given by the Hermitian product: $\mathfrak{b}(\boldsymbol{z}, \boldsymbol{w}) = \boldsymbol{w}^* \boldsymbol{z} = \bar{w}_1 z_1 + \bar{w}_2 z_2 + \ldots + \bar{w}_n z_n$. The corresponding quadratic form

$$\mathfrak{q}(\boldsymbol{z}) = \mathfrak{b}(\boldsymbol{z}, \boldsymbol{z}) = \bar{z}_1 z_1 + \bar{z}_2 z_2 + \ldots + \bar{z}_n z_n = |z_1|^2 + |z_2|^2 + \ldots + |z_1|^n$$

is positive definite. Indeed, $\mathfrak{q}(\boldsymbol{z}) \geq 0$ for every $\boldsymbol{z} \in \mathbb{C}^n$, and for $\mathfrak{q}(\boldsymbol{z})$ to be equal zero we need $z_i = 0$ for all $i \in \{1, 2, \ldots, n\}$, i.e., we need $\boldsymbol{z} = \mathbf{0}$.

Example *2.23*

If $\boldsymbol{X} = \mathbb{R}^2$ and we consider $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$, with $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$, then we have seen that $\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 - x_2 y_2$ is a symmetric sesquilinear form, and it is associated with the corresponding quadratic form is $\mathfrak{q}(\boldsymbol{x}) = x_1^2 - x_2^2$. This quadratic form is not positive semi-definite, since $\mathfrak{q}(1, 0) = 1$ and $\mathfrak{q}(0, 1) = -1$.

Exercise 2.28

Consider the vector space $\mathbb{R}^3$, two vectors $\boldsymbol{x} = (x_1, x_2, x_3)$ and $\boldsymbol{y} = (y_1, y_2, y_3)$ and the form

$$\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}) = x_1 y_1 + x_2 y_2.$$

Show that $\mathfrak{b}$ is a sesquilinear form, and that the associated quadratic form is positive semi-definite. [Solution: (2.28)]

Exercise 2.29

Consider the complex vector space $\mathbb{C}^3$, two vectors $\boldsymbol{z} = (z_1, z_2, z_3)$ and $\boldsymbol{w} = (w_1, w_2, w_3)$ and the function

$$\mathfrak{b}(\boldsymbol{z}, \boldsymbol{w}) = -\bar{w}_1 z_1 - \bar{w}_3 z_3.$$

Show that $\mathfrak{b}$ is a Hermitian semidefinite sesquilinear form, and that the associated quadratic form is negative semi-definite. [Solution: (2.28)]

We are now ready to present the definition of inner product.

Definition 2.16 [Inner Product]
An inner product on a vector space $\boldsymbol{X}$ over the field $\mathbb{F}$ is a function

$$\begin{aligned} \langle \cdot, \cdot \rangle : \boldsymbol{X} \times \boldsymbol{X} &\longrightarrow \mathbb{F} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \langle \boldsymbol{x}, \boldsymbol{y} \rangle, \end{aligned}$$

and satisfies the following axioms:

(i) *Anti-symmetry: $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$ for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$*

(ii) *Linearity in the first argument: for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{X}$ and all $\alpha \in \mathbb{F}$*

- $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$
- $\langle \alpha\, \boldsymbol{x}, \boldsymbol{y} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle$

(iii) *Positive-definiteness: for all $\boldsymbol{x} \in \boldsymbol{X}$*

- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$
- $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0 \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{0}$.

The vector space $\boldsymbol{X}$ with an inner product $\langle \cdot, \cdot \rangle$ is called an inner product space or pre-Hilbert space, and it is denoted by $(\boldsymbol{X}, \langle \cdot, \cdot \rangle)$.
If $\mathbb{F} = \mathbb{C}$, the vector space $\boldsymbol{X}$ with an inner product is also called Hermitian inner product space.

As a consequence of this definition, we have the so-called anti-linearity in the second argument: indeed, recalling that the conjugate of the sum of complex numbers is the sum of the complex conjugates, for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{X}$, we have

$$\langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle = \overline{\langle \boldsymbol{y} + \boldsymbol{z}, \boldsymbol{x} \rangle} = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle + \langle \boldsymbol{z}, \boldsymbol{x} \rangle} = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle} + \overline{\langle \boldsymbol{z}, \boldsymbol{x} \rangle} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle.$$

Moreover, recalling that the complex conjugate of the product of complex numbers is the product of the conjugates, for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ and for all $k \in \mathbb{C}$, we have

$$\langle \boldsymbol{x}, k\, \boldsymbol{y} \rangle = \overline{\langle k\, \boldsymbol{y}, \boldsymbol{x} \rangle} = \overline{k \langle \boldsymbol{y}, \boldsymbol{x} \rangle} = \bar{k}\, \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle} = \bar{k} \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

In the real case, i.e. when $\mathbb{F} = \mathbb{R}$, the anti-symmetry property becomes a symmetry property. In fact, if $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ is real, then it is equal to its conjugate $\overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$ (if $\alpha \in \mathbb{R}$, then $\bar{\alpha} = \alpha$). It follows that in the real case the symmetry and the linearity in the first argument also imply the linearity in the second argument. Indeed, given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \boldsymbol{X}$ and $\alpha \in \mathbb{R}$

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z} \rangle &= \langle \boldsymbol{y} + \boldsymbol{z}, \boldsymbol{x} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle + \langle \boldsymbol{z}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{x}, \boldsymbol{z} \rangle \\ \langle \boldsymbol{x}, \alpha\, \boldsymbol{y} \rangle &= \langle \alpha\, \boldsymbol{y}, \boldsymbol{x} \rangle = \alpha \langle \boldsymbol{y}, \boldsymbol{x} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle. \end{aligned}$$

For this reason, the properties of linearity (only in the real case) are sometimes written as follows:

$$\langle \alpha\, \boldsymbol{x} + \beta\, \boldsymbol{y}, \boldsymbol{z} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta \langle \boldsymbol{y}, \boldsymbol{z} \rangle \qquad \text{and} \qquad \langle \boldsymbol{x}, \alpha\, \boldsymbol{y} + \beta\, \boldsymbol{z} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \beta \langle \boldsymbol{x}, \boldsymbol{z} \rangle.$$

These properties may be summarized by saying that the inner product is *bilinear*. By contrast, the Hermitian inner product, when $\mathbb{F} = \mathbb{C}$, satisfies

$$\langle \alpha\, \boldsymbol{x} + \beta\, \boldsymbol{y}, \boldsymbol{z} \rangle = \alpha \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \beta \langle \boldsymbol{y}, \boldsymbol{z} \rangle \qquad \text{and} \qquad \langle \boldsymbol{x}, \alpha\, \boldsymbol{y} + \beta\, \boldsymbol{z} \rangle = \bar{\alpha} \langle \boldsymbol{x}, \boldsymbol{y} \rangle + \bar{\beta} \langle \boldsymbol{x}, \boldsymbol{z} \rangle,$$

which is anti-linear in the second argument. Axiom **(ii)** in Definition 2.16 says that an inner product is a sesquilinear form. Axiom **(i)** says that this sesquilinear form is Hermitian. Axiom **(iii)** says that the quadratic form associated with this Hermitian sesquilinear form is positive definite. Hence, **an inner product is a Hermitian sesquilinear form whose associated quadratic form is positive definite**.

Example *2.24*

Consider $\boldsymbol{X} = \mathbb{R}^n$ with the dot product. The dot product, defined for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ as

$$\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{y} = x_1\, y_1 + x_2\, y_2 + \ldots + x_n\, y_n,$$

is an inner product $\langle \cdot, \cdot \rangle$, since it is a symmetric sesquilinear form and its associated quadratic form is positive definite. However, we want to check the axioms of inner products to confirm this fact. For any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, using the property of the transposes,

$$\begin{aligned} \boldsymbol{x} \cdot \boldsymbol{y} &= \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{y}^\top \boldsymbol{x} = \boldsymbol{y} \cdot \boldsymbol{x} \\ (\boldsymbol{x} + \boldsymbol{y}) \cdot \boldsymbol{z} &= (\boldsymbol{x} + \boldsymbol{y})^\top \boldsymbol{z} = (\boldsymbol{x}^\top + \boldsymbol{y}^\top)\, \boldsymbol{z} = \boldsymbol{x}^\top \boldsymbol{z} + \boldsymbol{y}^\top \boldsymbol{z} = \boldsymbol{x} \cdot \boldsymbol{z} + \boldsymbol{y} \cdot \boldsymbol{z} \\ (\alpha\, \boldsymbol{x}) \cdot \boldsymbol{y} &= (\alpha\, \boldsymbol{x})^\top \boldsymbol{y} = \alpha\, \boldsymbol{x}^\top \boldsymbol{y} = \alpha\, (\boldsymbol{x} \cdot \boldsymbol{y}) \\ \boldsymbol{x} \cdot \boldsymbol{x} &= x_1^2 + x_2^2 + \ldots + x_n^2 \geq 0, \end{aligned}$$

and, finally,

$$\boldsymbol{x} \cdot \boldsymbol{x} = 0 \quad \Leftrightarrow \quad x_1^2 + x_2^2 + \ldots + x_n^2 = 0 \quad \Leftrightarrow \quad x_1 = x_2 = \ldots = x_n = 0 \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0}.$$

Example *2.25*

Consider the vector space $\mathbb{C}^n$ with the product

$$\boldsymbol{z} \cdot \boldsymbol{w} = \boldsymbol{w}^* \boldsymbol{z} = \bar{w}_1\, z_1 + \bar{w}_2\, z_2 + \ldots + \bar{w}_n\, z_n,$$

where as usual $\boldsymbol{w}^*$ denotes the conjugate transpose. This is a Hermitian inner product, because as we saw before it is a Hermitian sesquilinear form with positive definite quadratic form. However, we want to check the axioms. For any $\boldsymbol{z}, \boldsymbol{w}, \boldsymbol{u} \in \mathbb{C}^n$ and $\alpha \in \mathbb{C}$, using the property of the conjugates, we find

$$\begin{aligned} \boldsymbol{z} \cdot \boldsymbol{w} &= \boldsymbol{w}^* \boldsymbol{z} = \overline{\boldsymbol{z}^* \boldsymbol{w}} = \overline{\boldsymbol{w} \cdot \boldsymbol{z}} \\ (\boldsymbol{z} + \boldsymbol{w}) \cdot \boldsymbol{u} &= \boldsymbol{u}^* (\boldsymbol{z} + \boldsymbol{w}) = \boldsymbol{u}^* \boldsymbol{z} + \boldsymbol{u}^* \boldsymbol{w} = \boldsymbol{z} \cdot \boldsymbol{u} + \boldsymbol{w} \cdot \boldsymbol{u} \\ (\alpha\, \boldsymbol{z}) \cdot \boldsymbol{w} &= \boldsymbol{w}^* (\alpha\, \boldsymbol{z}) = \alpha\, \boldsymbol{w}^* \boldsymbol{z} = \alpha\, (\boldsymbol{z} \cdot \boldsymbol{w}) \\ \boldsymbol{z} \cdot \boldsymbol{z} &= \boldsymbol{z}^* \boldsymbol{z} = \bar{z}_1\, z_1 + \bar{z}_2\, z_2 + \ldots + \bar{z}_n\, z_n = |z_1|^2 + |z_2|^2 + \ldots + |z_n|^2 \geq 0, \end{aligned}$$

and, finally,

$$\boldsymbol{z} \cdot \boldsymbol{z} = 0 \quad \Leftrightarrow \quad |z_1|^2 + |z_2|^2 + \ldots + |z_n|^2 = 0 \quad \Leftrightarrow \quad |z_1| = |z_2| = \ldots = |z_n| = 0 \quad \Leftrightarrow \quad z_1 = z_2 = \ldots = z_n = 0 \quad \Leftrightarrow \quad \boldsymbol{z} = \boldsymbol{0}.$$

This confirms that this product is a Hermitian inner product. We also verify the antisymmetry in the second

argument:

$$z \cdot (w + u) = (w + u)^* z = (w^* + u^*) z = w^* z + u^* z$$
$$z \cdot (\alpha w) = (\alpha w)^* z = \bar{\alpha} w^* z = \bar{\alpha} (z \cdot w).$$

Example *2.26* The vector space $\mathbb{R}^2$ with the product

$$\begin{aligned} p : \mathbb{R}^2 \times \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto x_1 y_1 - x_2 y_2 \end{aligned}$$

is not an inner product. In fact, we saw earlier than even if it is a symmetric sesquilinear form, its quadratic form is not positive definite (we found that, indeed, it is indefinite). As further evidence of this, we see immediately that the axiom of positive definiteness is not satisfied, since

$$p((0, 1), (1, 1)) = -1.$$

Example *2.27* The vector space $\mathbb{C}^3$ with the product

$$\begin{aligned} p : \mathbb{C}^3 \times \mathbb{C}^3 &\longrightarrow \mathbb{C} \\ (z, w) &\mapsto \bar{w}_1 z_1 + \bar{w}_3 z_3 \end{aligned}$$

is not a Hermitian inner product. Indeed,

$$p((0, 1, 0), (0, 1, 0)) = 0,$$

and so the last axiom is not satisfied.

The first notion that we can introduce, once an inner product has been defined, is that of orthogonality between vectors.

Definition 2.17 [Orthogonal Vectors and Orthogonal Complement]
Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space. Let $x, y \in X$. We say that x and y are **orthogonal** *if $\langle x, y \rangle = 0$. The expression $\langle x, y \rangle = 0$ is often written as $x \perp y$.*

Given a subset $W \subseteq X$, we define the **orthogonal complement** *of W in X as the set*

$$W^{\perp} \stackrel{def}{=} \{x \in X \mid \forall\, w \in W,\ \langle x, w \rangle = 0\}.$$

Exercises

Exercise 2.30 If $X = \mathbb{R}^2$ and we consider $x, y \in \mathbb{R}^2$, with $x = (x_1, x_2)$ and $y = (y_1, y_2)$, show that $\mathfrak{b}(x, y) = x_1 y_1 - x_2 y_2$ is a symmetric sesquilinear form. Show also that the associated form is sign indefinite. [Solution: (2.30)]

Exercise 2.31 Verify that

$$\begin{aligned} \mathfrak{b} : \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (u, v) &\mapsto \bar{v}_1 u_1 + \bar{v}_2 u_2 + \ldots + \bar{v}_n u_n, \end{aligned}$$

where $u = (u_1, u_2, \ldots, u_n)$ and $v = (v_1, v_2, \ldots, v_n)$, is a Hermitian sesquilinear form. [Solution: (2.31)]

Exercise 2.32

Let $a, b \in \mathbb{R}$ be such that $a < b$. Verify that

$$\begin{aligned} \mathfrak{b} : C([a,b],\mathbb{C}) \times C([a,b],\mathbb{C}) &\longrightarrow \mathbb{C} \\ (f,g) &\mapsto \int_a^b \overline{g(x)}\, f(x)\, \mathrm{d}x \end{aligned}$$

is a Hermitian sesquilinear form. [Solution: (2.32)]

We now examine how we can introduce a norm in an inner product vector space which is intimately connected with the inner product itself.

Definition 2.18 [Induced Norm]
Let $(\boldsymbol{X}, \langle\cdot,\cdot\rangle)$ be an inner product vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). For every $\boldsymbol{x} \in \boldsymbol{X}$, we define

$$\|\boldsymbol{x}\| \stackrel{def}{=} \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}.$$

We shall name the map

$$\begin{aligned} \boldsymbol{X} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\| \end{aligned}$$

the norm induced by the inner product $\langle\cdot,\cdot\rangle$.

Notice that this definition is well posed because of the positive semi-definiteness condition of the inner product. We still need to convince ourselves that $\|\cdot\|$ is really a norm. To this end, we need to prove the Cauchy-Schwarz inequality for inner product spaces.

Theorem 2.18 [Cauchy-Schwarz Inequality]
Let $(\boldsymbol{X}, \langle\cdot,\cdot\rangle)$ be an inner product vector space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $\|\cdot\|$ be the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Then,

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \quad |\langle \boldsymbol{x}, \boldsymbol{y}\rangle| \leq \|\boldsymbol{x}\|\, \|\boldsymbol{y}\|.$$

Proof Let $\alpha \in \mathbb{F}$. Using the properties of the inner product, for all $t \in \mathbb{R}$

$$\begin{aligned} 0 &\leq \langle \boldsymbol{x} + t\,\alpha\,\boldsymbol{y}, \boldsymbol{x} + t\,\alpha\,\boldsymbol{y}\rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x} + t\,\alpha\,\boldsymbol{y}\rangle + \langle t\,\alpha\,\boldsymbol{y}, \boldsymbol{x} + t\,\alpha\,\boldsymbol{y}\rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x}\rangle + \langle \boldsymbol{x}, t\,\alpha\,\boldsymbol{y}\rangle + \langle t\,\alpha\,\boldsymbol{y}, \boldsymbol{x}\rangle + \langle t\,\alpha\,\boldsymbol{y}, t\,\alpha\,\boldsymbol{y}\rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x}\rangle + t\,\bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{y}\rangle + t\,\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle + t^2\,\alpha\,\bar{\alpha}\,\langle \boldsymbol{y}, \boldsymbol{y}\rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{x}\rangle + t\,(\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{y}\rangle) + t^2\,|\alpha|^2\,\langle \boldsymbol{y}, \boldsymbol{y}\rangle \\ &= \|\boldsymbol{x}\|^2 + 2\,t\,\Re\mathrm{e}\{\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle\} + t^2\,|\alpha|^2\,\|\boldsymbol{y}\|^2 \end{aligned}$$

as

$$\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle + \bar{\alpha}\,\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle + \overline{\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle} = 2\,\Re\mathrm{e}\{\alpha\,\langle \boldsymbol{y}, \boldsymbol{x}\rangle\}.$$

Let us choose

$$\alpha = \begin{cases} \frac{\langle x,y\rangle}{|\langle x,y\rangle|} & \text{if } \langle x,y\rangle \neq 0 \\ 1 & \text{if } \langle x,y\rangle = 0 \end{cases}$$

We observe that in any case $|\alpha| = 1$. Moreover, if $\langle x,y\rangle \neq 0$ we have

$$\alpha\,\langle y,x\rangle = \frac{\langle x,y\rangle}{|\langle x,y\rangle|}\,\langle y,x\rangle = \frac{\langle x,y\rangle\,\overline{\langle x,y\rangle}}{|\langle x,y\rangle|} = \frac{|\langle x,y\rangle|^2}{|\langle x,y\rangle|} = |\langle x,y\rangle|,$$

and if $\langle x,y\rangle = 0$ we have $\alpha\,\langle y,x\rangle = 0$, so that it is still true that $\alpha\,\langle y,x\rangle = |\langle x,y\rangle|$. It follows that in any case

$$\forall\, t \in \mathbb{R}, \quad p(t) = \langle x + t\,\alpha\,y, x + t\,\alpha\,y\rangle = t^2\,\|y\|^2 + 2\,t\left|\langle x,y\rangle\right| + \|x\|^2.$$

Hence, $p(t)$ is a polynomial of degree smaller or equal to 2, with real coefficients, and non-negative for all $t \in \mathbb{R}$. This implies that its discriminant Δ cannot positive, i.e.,

$$\frac{\Delta}{4} = \left|\langle x,y\rangle\right|^2 - \|y\|^2\,\|x\|^2 \leq 0,$$

which leads to the Cauchy-Schwarz inequality by taking the square roots of both sides.

We are now ready to show that the norm induced by an inner product in an inner product vector space is indeed a norm.

Theorem 2.19 *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). The induced norm is a norm.*

Proof We denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Let $\lambda \in \mathbb{F}$. Let $x \in X$. Then, the homogeneity of the norm follows from

$$\|\lambda\,x\| = \sqrt{\langle \lambda\,x, \lambda\,x\rangle} = \sqrt{\lambda\,\bar{\lambda}\,\langle x,x\rangle} = \sqrt{|\lambda|^2\,\langle x,x\rangle} = |\lambda|\,\|x\|.$$

Next, if $x, y \in X$, by virtue of the Cauchy-Schwartz inequality we have

$$\begin{aligned} \|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \langle x,x\rangle + \langle x,y\rangle + \langle y,x\rangle + \langle y,y\rangle \\ &= \|x\|^2 + \langle x,y\rangle + \overline{\langle x,y\rangle} + \|y\|^2 \\ &= \|x\|^2 + 2\,\Re\mathrm{e}\,\{\langle x,y\rangle\} + \|y\|^2 \\ &\leq \|x\|^2 + 2\left|\Re\mathrm{e}\,\{\langle x,y\rangle\}\right| + \|y\|^2 \\ &\leq \|x\|^2 + 2\left|\langle x,y\rangle\right| + \|y\|^2 \\ &\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = \left(\|x\| + \|y\|\right)^2, \end{aligned}$$

where we used the fact that for every $z \in \mathbb{C}$ we have $|\Re\mathrm{e}\{z\}| \leq |z|$ (why?). Taking the square roots on both sides, we obtain $\|x+y\| \leq \|x\| + \|y\|$. Finally, if $\|x\| = 0$, then $\langle x,x\rangle = 0$, which implies $x = \mathbf{0}$.

Example 2.28

Let $z = (z_1, z_2, \ldots, z_n)$ and $w = (w_1, w_2, \ldots, w_n)$ be two vectors $\mathbb{C}^n$. We saw that the Hermitian product is the inner product in $\mathbb{C}^n$ defined as

$$\langle z, w\rangle = \bar{w}_1\,z_1 + \bar{w}_2\,z_2 + \ldots + \bar{w}_n\,z_n$$

Let $z_k = a_k + i\,b_k$, where $a_k, b_k \in \mathbb{R}$ for all $k \in \{1,\ldots,n\}$. Then, $\bar{z}_k\,z_k = (a_k - i\,b_k)(a_k + i\,b_k) = a_k^2 + b_k^2 = |z_k|^2$.

It follows that the induced norm obtained from the Hermitian product is

$$\|z\| = \sqrt{\langle z, z\rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \ldots + |z_n|^2},$$

and it is indeed a norm. Obviously, we can write also

$$\|z\| = \sqrt{a_1^2 + b_1^2 + \ldots + a_n^2 + b_n^2},$$

which coincides with the Euclidean norm in $\mathbb{R}^{2n}$ of the vector $(a_1, b_1, a_2, b_2, \ldots, a_n, b_n)$.

2.5.2 Hilbert spaces: definition and properties

We are now ready to give a precise definition of Hilbert space. A normed space $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ is a Hilbert space if $(\boldsymbol{H}, \|\cdot\|)$ is a Banach space, where $\|\cdot\|$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$ (so that $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$). In other words, a Hilbert space is a vector space equipped with an inner product which induces a norm for which it is a complete normed space.

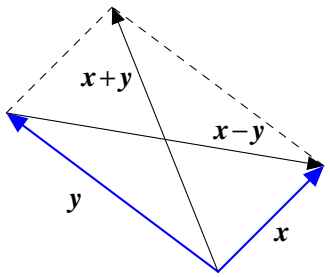
Definition 2.19 [Hilbert Space]
Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be an inner product space. Let us denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. We say that $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ is a Hilbert space *if the normed space $(\boldsymbol{H}, \|\cdot\|)$ is complete.*

The vector spaces $\mathbb{R}^n$ (respectively, $\mathbb{C}^n$), with the standard dot product (respectively, Hermitian product) are Hilbert spaces. The normed space $(C([a,b],\mathbb{R}), \|\cdot\|_2)$ is not a Hilbert space (since it is not a Banach space).

We now introduce the *parallelogram law*, which can be written as

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2\,\|\boldsymbol{x}\|^2 + 2\,\|\boldsymbol{y}\|^2.$$

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.



To prove the so-called parallelogram lemma, we present the following simple properties, which follow straightforwardly from results on sequences of elements of a normed space.

Theorem 2.20 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space where $\boldsymbol{X}$ is a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Then:*

(i) if $(a_k)_{k\in\mathbb{N}}$ is a convergent sequence of elements of $\mathbb{F}$ with limit $\hat{a} \in \mathbb{F}$ and $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a

convergent sequence in $\boldsymbol{X}$ with limit $\hat{\boldsymbol{x}} \in \boldsymbol{X}$, then $(a_k \boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent and its limit is $\hat{a}\,\hat{\boldsymbol{x}}$;

(ii) if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ and $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ are sequences in $\boldsymbol{X}$ with limits $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ and $\hat{\boldsymbol{y}} \in \boldsymbol{X}$, respectively, then $(\boldsymbol{x}_k + \boldsymbol{y}_k)_{k\in\mathbb{N}}$ is a convergent sequence in $\boldsymbol{X}$ and its limit is $\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}}$.

Proof *(i)* We use the triangle inequality:

$$\begin{aligned} 0 \;&\le\; \|a_k\,\boldsymbol{x}_k - \hat{a}\,\hat{\boldsymbol{x}}\| = \|a_k\,(\boldsymbol{x}_k - \hat{\boldsymbol{x}}) + (a_k - \hat{a})\,\hat{\boldsymbol{x}}\| \\ &\le\; |a_k|\,\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| + |a_k - \hat{a}|\,\|\hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} |\hat{a}|\cdot 0 + 0\cdot\|\hat{\boldsymbol{x}}\| = 0. \end{aligned}$$

It follows that $\|a_k\,\boldsymbol{x}_k - \hat{a}\,\hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} 0$, which implies $a_k\,\boldsymbol{x}_k \overset{k\to\infty}{\longrightarrow} \hat{a}\,\hat{\boldsymbol{x}}$.

(ii) We find immediately

$$0 \le \|(\boldsymbol{x}_k + \boldsymbol{y}_k) - (\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}})\| \le \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| + \|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\| \overset{k\to\infty}{\longrightarrow} 0.$$

Theorem 2.21 [Parallelogram Lemma]
Let $(\boldsymbol{X}, \langle\cdot,\cdot\rangle)$ be an inner product space over $\mathbb{F}$, and let $\|\cdot\|$ denote the induced norm. Then,

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}, \quad \|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2\,(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2) \qquad \text{(Parallelogram Law)}. \tag{2.23}$$

Conversely, if $(\boldsymbol{X}, \|\cdot\|)$ is a normed space such that the parallelogram law (2.23) holds, there exists an inner product $\langle\cdot,\cdot\rangle : \boldsymbol{X}\times\boldsymbol{X} \longrightarrow \mathbb{F}$ which induces the norm $\|\cdot\|$, i.e., for all $\boldsymbol{x} \in \boldsymbol{X}$, there holds $\langle \boldsymbol{x}, \boldsymbol{x}\rangle^{\frac{1}{2}} = \|\boldsymbol{x}\|$.

Proof We prove the first part. Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. Expanding the left-hand side of (2.23) yields

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 \;&=\; \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y}\rangle + \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle \\ &=\; \langle \boldsymbol{x}, \boldsymbol{x}\rangle + \langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{y}, \boldsymbol{x}\rangle + \langle \boldsymbol{y}, \boldsymbol{y}\rangle + \langle \boldsymbol{x}, \boldsymbol{x}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle - \langle \boldsymbol{y}, \boldsymbol{x}\rangle + \langle \boldsymbol{y}, \boldsymbol{y}\rangle \\ &=\; 2\,\|\boldsymbol{x}\|^2 + 2\,\|\boldsymbol{y}\|^2. \end{aligned}$$

We prove the converse statement. Consider for the sake of simplicity $\mathbb{F} = \mathbb{R}$. We define

$$\begin{aligned} \langle\cdot,\cdot\rangle : \boldsymbol{X}\times\boldsymbol{X} &\longrightarrow \mathbb{R} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \frac{1}{4}\,(\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2). \end{aligned} \tag{2.24}$$

From the homogeneity of the norm, for all $\boldsymbol{x} \in \boldsymbol{X}$

$$\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \frac{1}{4}\,(\|2\,\boldsymbol{x}\|^2 - 0) = \|\boldsymbol{x}\|^2, \tag{2.25}$$

which says that, if we can show that $\langle\cdot,\cdot\rangle$ is an inner product, it induces the norm $\|\cdot\|$. We now then prove that $\langle\cdot,\cdot\rangle$ satisfies all the axioms of an inner product. First, we see immediately that from (2.25) the positive definiteness (property **(iii)** in Definition 2.16) is satisfied. The symmetry $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{y}, \boldsymbol{x}\rangle$ (property **(i)** in Definition 2.16) is also satisfied for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$ (why? use the properties of the norm). We now need to prove **(ii)** in Definition 2.16, which in the real case is simply the linearity in the first argument (we showed that if $\mathbb{F} = \mathbb{R}$ the linearity in the first argument and the symmetry imply the linearity in the second argument). We go through the following steps:

(a) $2\left\langle \frac{\boldsymbol{x}+\boldsymbol{y}}{2}, z\right\rangle = \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle$;

(b) $\langle \boldsymbol{x}+\boldsymbol{y}, z\rangle = \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle$;

(c) $\forall\, q \in \mathbb{Q}, \quad \langle q\,\boldsymbol{x}, z\rangle = q\,\langle \boldsymbol{x}, z\rangle$;

(d) $\forall\, a \in \mathbb{R}, \quad \langle a\,\boldsymbol{x}, z\rangle = a\,\langle \boldsymbol{x}, z\rangle$.

We note that

$$\begin{aligned}
2\left\langle \frac{\boldsymbol{x}+\boldsymbol{y}}{2}, z\right\rangle &= \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z\right\|^2 - \left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z\right\|^2\right)\\
&= \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right) - \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right)\\
&= \frac{1}{4}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z+\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2 + \left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z-\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2 - \left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z+\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2 - \left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z-\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right)\\
&= \frac{1}{4}\left(\|\boldsymbol{x}+z\|^2 + \|\boldsymbol{y}+z\|^2 - \|\boldsymbol{x}-z\|^2 - \|\boldsymbol{y}-z\|^2\right)\\
&= \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle,
\end{aligned}$$

where the parallelogram law has been used in the third equality (with $\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z$ and $\frac{\boldsymbol{x}+\boldsymbol{y}}{2}$ in place of the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ of (2.23) for the first term, and with $\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z$ and $\frac{\boldsymbol{x}+\boldsymbol{y}}{2}$ in place of $\boldsymbol{x}$ and $\boldsymbol{y}$ for the second term) and in the last equality. This proves **(a)**. Choosing $\boldsymbol{y} = \boldsymbol{0}$ in **(a)** gives

$$\forall\, \boldsymbol{x}, z \in \boldsymbol{X}, \quad 2\left\langle \frac{1}{2}\,\boldsymbol{x}, z\right\rangle = \langle \boldsymbol{x}, z\rangle. \tag{2.26}$$

Using **(a)** again, and using also (2.26) with $\boldsymbol{x}+\boldsymbol{y}$ in place of $\boldsymbol{x}$, yields

$$\forall\, \boldsymbol{x}, \boldsymbol{y}, z \in \boldsymbol{X}, \quad \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle = 2\left\langle \frac{\boldsymbol{x}+\boldsymbol{y}}{2}, z\right\rangle = \langle \boldsymbol{x}+\boldsymbol{y}, z\rangle. \tag{2.27}$$

This proves **(b)**. Notice that **(b)** with $\boldsymbol{x} = \boldsymbol{y}$ gives $2\,\langle \boldsymbol{x}, z\rangle = \langle 2\,\boldsymbol{x}, z\rangle$ for all $\boldsymbol{x}, z \in \boldsymbol{X}$, and using induction we show that

$$\forall\, \boldsymbol{x}, z \in \boldsymbol{X},\ \forall\, m \in \mathbb{N}, \quad m\,\langle \boldsymbol{x}, z\rangle = \langle m\,\boldsymbol{x}, z\rangle. \tag{2.28}$$

Indeed, suppose that for $m \in \mathbb{N}$ there holds $m\,\langle \boldsymbol{x}, z\rangle = \langle m\,\boldsymbol{x}, z\rangle$. Adding $\langle \boldsymbol{x}, z\rangle$ in both sides gives $m\,\langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{x}, z\rangle = \langle m\,\boldsymbol{x}, z\rangle + \langle \boldsymbol{x}, z\rangle$. The right hand-side becomes $(m+1)\,\langle \boldsymbol{x}, z\rangle$ and the left hand-side, using **(b)**, becomes $\langle m\,\boldsymbol{x} + \boldsymbol{x}, z\rangle = \langle (m+1)\,\boldsymbol{x}, z\rangle$, as required. Hence, we proved (2.28).

Re-writing the latter with $\boldsymbol{x}/n$ in place of $\boldsymbol{x}$ and n in place of m, we obtain

$$\forall\, \boldsymbol{x}, z \in \boldsymbol{X},\ \forall\, n \in \mathbb{N}^*, \quad n\,\langle \boldsymbol{x}/n, z\rangle = \langle \boldsymbol{x}, z\rangle.$$

We obtain

$$\forall\, \boldsymbol{x}, z \in \boldsymbol{X},\ \forall\, m \in \mathbb{N},\ \forall\, n \in \mathbb{N}^*, \quad \frac{m}{n}\,\langle \boldsymbol{x}, z\rangle = \frac{m}{n}\,n\left\langle \frac{\boldsymbol{x}}{n}, z\right\rangle = m\left\langle \frac{\boldsymbol{x}}{n}, z\right\rangle = \left\langle \frac{m}{n}\,\boldsymbol{x}, z\right\rangle.$$

This proves **(c)** in the case where the rational number q is non-negative. We want to show that **(c)** holds for any $q \in \mathbb{Q}$. Writing **(b)** with $-\boldsymbol{y}$ in place of $\boldsymbol{y}$ gives

$$\forall\, \boldsymbol{x}, \boldsymbol{y}, z \in \boldsymbol{X}, \quad \langle \boldsymbol{x}, z\rangle - \langle \boldsymbol{y}, z\rangle = \langle \boldsymbol{x}-\boldsymbol{y}, z\rangle. \tag{2.29}$$

From this identity, it follows that

$$\langle -\boldsymbol{x}, z\rangle = \langle \boldsymbol{0}-\boldsymbol{x}, z\rangle = \langle \boldsymbol{0}, z\rangle - \langle \boldsymbol{x}, z\rangle = 0 - \langle \boldsymbol{x}, z\rangle = -\langle \boldsymbol{x}, z\rangle.$$

We now use this identity by considering a negative rational number q:

$$q\,\langle \boldsymbol{x}, z\rangle = -(-q)\,\langle \boldsymbol{x}, z\rangle = -\langle (-q)\,\boldsymbol{x}, z\rangle = -\langle -(q\,\boldsymbol{x}), z\rangle = -(-\langle q\,\boldsymbol{x}, z\rangle) = \langle q\,\boldsymbol{x}, z\rangle.$$

This completes the proof of **(c)**. Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of real numbers which converges to $\hat{a}\in\mathbb{R}$. Let $\boldsymbol{x}, z\in\boldsymbol{X}$. Then, using Theorem 2.20, we have $a_k\,\boldsymbol{x}+z$ and $a_k\,\boldsymbol{x}-z$ converge to $\hat{a}\,\boldsymbol{x}+z$ and $\hat{a}\,\boldsymbol{x}-z$ respectively, so that $\|a_k\,\boldsymbol{x}+z\|$ and $\|a_k\,\boldsymbol{x}-z\|$ converge to $\|\hat{a}\,\boldsymbol{x}+z\|$ and $\|\hat{a}\,\boldsymbol{x}-z\|$ respectively, and thus

$$\langle a_k\,\boldsymbol{x}, z\rangle = \frac{1}{4}\left(\|a_k\,\boldsymbol{x}+z\|^2 - \|a_k\,\boldsymbol{x}-z\|^2\right) \overset{k\to\infty}{\longrightarrow} \frac{1}{4}\left(\|\hat{a}\,\boldsymbol{x}+z\|^2 - \|\hat{a}\,\boldsymbol{x}-z\|^2\right) = \langle \hat{a}\,\boldsymbol{x}, z\rangle.$$

In other words, we can write

$$\lim_{k\to\infty}\langle a_k\,\boldsymbol{x}, z\rangle = \left\langle \lim_{k\to\infty} a_k\,\boldsymbol{x}, z\right\rangle.$$

From the density of $\mathbb{Q}$ in $\mathbb{R}$, we know that every real number can be written as the limit of a sequence of rational numbers. Let a be a real number, and let $(q_k)_{k\in\mathbb{N}}$ be a sequence of rational numbers which converges to a. Then,

$$\langle a\,\boldsymbol{x}, z\rangle = \left\langle \lim_{k\to\infty} q_k\,\boldsymbol{x}, z\right\rangle = \lim_{k\to\infty}\langle q_k\,\boldsymbol{x}, z\rangle = \lim_{k\to\infty} q_k\,\langle \boldsymbol{x}, z\rangle = a\,\langle \boldsymbol{x}, z\rangle.$$

This concludes the proof of **(d)**.

Exercise 2.33

In the second part of the proof of Theorem 2.21 we assumed that the vector space $\boldsymbol{X}$ is real. If $\boldsymbol{X}$ is complex (i.e., if $\mathbb{F}=\mathbb{C}$), adapt the proof of Theorem 2.21 by substituting (2.24) with

$$\begin{aligned}\langle\cdot,\cdot\rangle : \boldsymbol{X}\times\boldsymbol{X} &\longrightarrow \mathbb{C}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto \langle \boldsymbol{x},\boldsymbol{y}\rangle = \frac{1}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\,\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\,\boldsymbol{y}\|^2\right).\end{aligned} \tag{2.30}$$

[Solution: (2.33)]

The expressions (2.24) and (2.30) are called *polarization identities*. More generally, in linear algebra a polarization identity is formula that express the inner product of two vectors in terms of the norm of a normed space.

We recall that if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are normed vector spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, respectively, and $A\subseteq\boldsymbol{X}$ and $f : A\longrightarrow\boldsymbol{Y}$, given $\hat{\boldsymbol{x}}\in A$, the following facts are equivalent:

(A) f is continuous at $\hat{\boldsymbol{x}}$;

(B) for every sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ with values in A such that $\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\|_X \overset{k\to\infty}{\longrightarrow} 0$, we have $\|f(\boldsymbol{x}_k)-f(\hat{\boldsymbol{x}})\|_Y \overset{k\to\infty}{\longrightarrow} 0$.

Let $(\boldsymbol{H},\langle\cdot,\cdot\rangle)$ be a Hilbert space, where $\boldsymbol{H}$ is a vector space over a field $\mathbb{F}$ (as usual, $\mathbb{R}$ or $\mathbb{C}$) and let

$$\begin{aligned} f : \boldsymbol{H}\times\boldsymbol{H} &\longrightarrow \mathbb{F}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto \langle \boldsymbol{x},\boldsymbol{y}\rangle.\end{aligned}$$

The space $\boldsymbol{H}\times\boldsymbol{H}$ is a normed space if we introduce as a norm

$$\|(\boldsymbol{x},\boldsymbol{y})\|_{\boldsymbol{H}\times\boldsymbol{H}} = \sqrt{\|\boldsymbol{x}\|^2+\|\boldsymbol{y}\|^2},$$

where $\|\cdot\|$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$. A sequence $((\boldsymbol{x}_k,\boldsymbol{y}_k))_{k\in\mathbb{N}}$ of $\boldsymbol{H}\times\boldsymbol{H}$ converges to a certain $(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}})$ if and only if

$$\|\boldsymbol{x}_k-\hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} 0 \quad\text{and}\quad \|\boldsymbol{y}_k-\hat{\boldsymbol{y}}\| \overset{k\to\infty}{\longrightarrow} 0.$$

Indeed,

$$\|(\boldsymbol{x}_k\, \boldsymbol{y}_k) - (\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\|_{\boldsymbol{H}\times\boldsymbol{H}} = \|(\boldsymbol{x}_k - \hat{\boldsymbol{x}}, \boldsymbol{y}_k - \hat{\boldsymbol{y}})\|_{\boldsymbol{H}\times\boldsymbol{H}} = \sqrt{\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|^2 + \|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\|^2} \xrightarrow{k\to\infty} 0$$

if and only if both $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ and $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ converge to $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, respectively.

The following result is a generalization of Theorem 1.47 to Hilbert spaces.

Theorem 2.22 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then, the inner product $\langle\cdot,\cdot\rangle$ is continuous as a function of both its arguments.*

Proof We show that the inner product function f is continuous over $\boldsymbol{H}\times\boldsymbol{H}$. Let $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})$ be a generic element of $\boldsymbol{H}\times\boldsymbol{H}$. To prove the continuity of the inner product, we must show that, given a sequence $((\boldsymbol{x}_k, \boldsymbol{y}_k))_{k\in\mathbb{N}}$ of $\boldsymbol{H}\times\boldsymbol{H}$,

$$\|(\boldsymbol{x}_k, \boldsymbol{y}_k) - (\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})\|_{\boldsymbol{H}\times\boldsymbol{H}} \xrightarrow{k\to\infty} 0 \qquad \Rightarrow \qquad |f(\boldsymbol{x}_k, \boldsymbol{y}_k) - f(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}})| \xrightarrow{k\to\infty} 0,$$

or, using the observations above,

$$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \xrightarrow{k\to\infty} 0 \quad \text{and} \quad \|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\| \xrightarrow{k\to\infty} 0 \qquad \Rightarrow \qquad |\langle\boldsymbol{x}_k, \boldsymbol{y}_k\rangle - \langle\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}\rangle| \xrightarrow{k\to\infty} 0.$$

Indeed, using the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\langle\boldsymbol{x}_k, \boldsymbol{y}_k\rangle - \langle\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}\rangle| &= |\langle\boldsymbol{x}_k - \hat{\boldsymbol{x}}, \boldsymbol{y}_k\rangle + \langle\hat{\boldsymbol{x}}, \boldsymbol{y}_k - \hat{\boldsymbol{y}}\rangle| \\ &\le |\langle\boldsymbol{x}_k - \hat{\boldsymbol{x}}, \boldsymbol{y}_k\rangle| + |\langle\hat{\boldsymbol{x}}, \boldsymbol{y}_k - \hat{\boldsymbol{y}}\rangle| \\ &\le \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\|\,\|\boldsymbol{y}_k\| + \|\hat{\boldsymbol{x}}\|\,\|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\| \xrightarrow{k\to\infty} 0, \end{aligned}$$

since $\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \xrightarrow{k\to\infty} 0$ and $\|\boldsymbol{y}_k - \hat{\boldsymbol{y}}\| \xrightarrow{k\to\infty} 0$.

Exercises

Exercise 2.34 Consider the Frobenius norm

$$\|A\|_F = \sqrt{\text{trace}(A^*A)},$$

where A^* is the conjugate transpose of A. Show that $\mathbb{C}^{m\times n}$ with this norm is a Hilbert space (Hint: try to find what Hermitian inner product induces the Frobenius norm). [Solution: (2.34)]

Exercise 2.35 Consider the normed space $(\mathbb{R}^2, \|\cdot\|)$, where, for $\boldsymbol{x}\in\mathbb{R}^2$,

$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + 2\,x_2^2}.$$

Prove that this norm is induced by an inner product $\langle\cdot,\cdot\rangle$ and that $(\mathbb{R}^2, \langle\cdot,\cdot\rangle)$ is a Hilbert space. Given a subspace $A = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$ and a vector $\boldsymbol{x} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$, find the vector $\hat{\boldsymbol{y}}\in A$ that minimizes the distance $\|\hat{\boldsymbol{y}} - \boldsymbol{x}\|$. [Solution: (2.35)]

Exercise 2.36 Show that the norm $\|\cdot\|_1$ in $C([0,1];\mathbb{R})$ is not induced by an inner product. [Solution: (2.36)]

Exercise 2.37 Show that the norm $\|\cdot\|_\infty$ in $C([0,1];\mathbb{R})$ is not induced by an inner product. [Solution: (2.37)]

Exercise 2.38

We have seen in Theorem 1.47 that if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ and $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ are two sequences of $\mathbb{R}^n$ convergent to the limits $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$, respectively, then

$$\lim_{k\to\infty}(\boldsymbol{x}_k\cdot\boldsymbol{y}_k)=\hat{\boldsymbol{x}}\cdot\hat{\boldsymbol{y}}.$$

Adapt the proof of this result to show that if $(\boldsymbol{X},\|\cdot\|_x)$ is a normed space and $(\boldsymbol{Y},\langle\cdot,\cdot\rangle)$ is an inner product space (and we denote by $\|\cdot\|_y$ the norm induced by $\langle\cdot,\cdot\rangle$), if $A\subseteq\boldsymbol{X}$, if $\boldsymbol{x}_0\in\mathcal{D}(A)$ and f and g converge to $\boldsymbol{\ell}_1,\boldsymbol{\ell}_2\in\boldsymbol{Y}$, respectively, as $\boldsymbol{x}\to\boldsymbol{x}_0$, then $\langle f(\boldsymbol{x}),g(\boldsymbol{x})\rangle$ is convergent as $\boldsymbol{x}\to\boldsymbol{x}_0$ and

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\langle f(\boldsymbol{x}),g(\boldsymbol{x})\rangle=\langle\boldsymbol{\ell}_1,\boldsymbol{\ell}_2\rangle.$$

[Solution: (2.38)]

Exercise 2.39

If $\boldsymbol{X}=\mathbb{R}^2$ and we consider $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^2$, with $\boldsymbol{x}=(x_1,x_2)$ and $\boldsymbol{y}=(y_1,y_2)$, show that $\mathfrak{b}(\boldsymbol{x},\boldsymbol{y})=x_1\,y_1-x_2\,y_2$ is a symmetric sesquilinear form. Show also that the associated form is sign indefinite. [Solution: (2.39)]

Exercise 2.40

Consider the vector space

$$\boldsymbol{X}=\left\{\begin{bmatrix}z_1&z_2\\0&z_3\end{bmatrix}\in\mathbb{C}^{2\times2}\,\middle|\,z_1,z_2,z_3\in\mathbb{C}\right\}$$

with the norm

$$\left\|\begin{bmatrix}z_1&z_2\\0&z_3\end{bmatrix}\right\|=\sqrt{|z_1|^2+5\,|z_2|^2+|z_3|^2}.$$

Decide if this norm is induced by an inner product $\langle\cdot,\cdot\rangle$, and establish if $(\boldsymbol{X},\langle\cdot,\cdot\rangle)$ is a Hilbert space. [Solution: (2.40)]

Exercise 2.41

Let $(\boldsymbol{X},\|\cdot\|)$ be a normed space. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\boldsymbol{X}$. Tick the boxes that correspond to statements that are true in general:

- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent, then it is a Cauchy sequence;
- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and it is also convergent, then $(\boldsymbol{X},\|\cdot\|)$ is a Banach space;
- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and it is also convergent, then $(\boldsymbol{X},\|\cdot\|)$ is a Hilbert space;
- ☐ If $(\boldsymbol{X},\|\cdot\|)$ is a Banach space and if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☐ If $(\boldsymbol{X},\|\cdot\|)$ is a Hilbert space and if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☐ If $\boldsymbol{X}$ is finite dimensional, $(\boldsymbol{X},\|\cdot\|)$ is a Banach space;
- ☐ If $\boldsymbol{X}$ is finite dimensional, $(\boldsymbol{X},\|\cdot\|)$ is a Hilbert space;
- ☐ If the norm is induced by an inner product, $(\boldsymbol{X},\|\cdot\|)$ is a Hilbert space;
- ☐ If the norm is induced by an inner product, $(\boldsymbol{X},\|\cdot\|)$ is a Banach space.

[Solution: (2.41)]

Exercise 2.42

Show that the normed spaces $(\ell^p(\mathbb{N};\mathbb{C}),\|\cdot\|_p)$, where $p\in[1,+\infty]$ defined in Exercise 2.13 and 2.14 is a Hilbert space if and only if $p=2$. What is the corresponding inner product in that case? [Solution: (2.42)]

2.6 Orthogonal projections in Hilbert spaces

Consider a normed space $(\boldsymbol{X}, \|\cdot\|)$. Recall that we defined the segment

$$[\boldsymbol{x}, \boldsymbol{y}] \stackrel{\text{def}}{=} \{\boldsymbol{x} + t\,(\boldsymbol{y} - \boldsymbol{x}) \,\big|\, t \in [0,1]\}.$$

with end-points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. Moreover, given $\boldsymbol{A} \subseteq \boldsymbol{X}$, we recall that $\boldsymbol{A}$ is convex if, for any pair of points $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}$, the segment $[\boldsymbol{x}, \boldsymbol{y}]$ is in $\boldsymbol{A}$, i.e., for all $t \in [0,1]$, we have $\boldsymbol{x} + t\,(\boldsymbol{y} - \boldsymbol{x}) \in \boldsymbol{A}$.

The following fundamental result states that, if a subset A of a Hilbert space $\boldsymbol{H}$ is convex and closed, for each point $\boldsymbol{x} \in \boldsymbol{H}$ there exists a unique point $P_A(\boldsymbol{x})$ in $\boldsymbol{A}$ which has minimum distance from $\boldsymbol{x}$. In the case where $\boldsymbol{A}$ is a closed subspace of $\boldsymbol{H}$, then $P_A(\boldsymbol{x})$ may be thought of as the orthogonal projection of $\boldsymbol{x}$ onto $\boldsymbol{A}$. This result is especially significant in numerical analysis, where it forms the basis of least squares methods.

Theorem 2.23 [Hilbert Projection Theorem]
Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space, and let us denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Let $\boldsymbol{A} \subseteq \boldsymbol{H}$ be non-empty, convex and closed. Let $\boldsymbol{x} \in \boldsymbol{H}$. Then, there is a unique point of minimum for the function

$$\begin{aligned} f : \boldsymbol{A} &\longrightarrow \mathbb{R}_+ \\ \boldsymbol{y} &\mapsto \|\boldsymbol{y} - \boldsymbol{x}\|. \end{aligned}$$

Proof We have to show that

$$\exists!\, \hat{\boldsymbol{y}} \in A : \quad \inf_{\boldsymbol{y} \in \boldsymbol{A}} \|\boldsymbol{y} - \boldsymbol{x}\| = \min_{\boldsymbol{y} \in \boldsymbol{A}} \|\boldsymbol{y} - \boldsymbol{x}\| = \|\hat{\boldsymbol{y}} - \boldsymbol{x}\|.$$

We denote by γ the infimum of f , i.e., $\gamma = \inf_{\boldsymbol{y} \in \boldsymbol{A}} \|\boldsymbol{y} - \boldsymbol{x}\|$. Obviously, $\gamma \geq 0$. We show that a sequence $(\boldsymbol{y}_k)_{k \in \mathbb{N}}$ of elements of $\boldsymbol{A}$ exists such that

$$\lim_{k \to \infty} \|\boldsymbol{y}_k - \boldsymbol{x}\| = \gamma. \tag{2.31}$$

Indeed, we recall that if $\Omega \subseteq \mathbb{R}$ and $\alpha \in \mathbb{R}$, then $\alpha = \inf \Omega$ if and only if

(i) $\forall\, x \in \Omega, \quad \alpha \leq x$;

(ii) $\forall\, \varepsilon > 0, \ \exists\, x \in \Omega : x < \alpha + \varepsilon$.

We apply this result to $\Omega = \{\|\boldsymbol{y} - \boldsymbol{x}\| \mid \boldsymbol{y} \in \boldsymbol{A}\}$. For every $k \in \mathbb{N}$, there exists $\boldsymbol{y}_k \in \boldsymbol{A}$ such that

$$\gamma \leq \|\boldsymbol{y}_k - \boldsymbol{x}\| < \gamma + 2^{-k}.$$

Hence, the squeeze theorem allows us to conclude that (2.31) holds.

Now, we show that the sequence $(\boldsymbol{y}_k)_{k \in \mathbb{N}}$ is convergent. Since $\boldsymbol{H}$ is complete, it suffices to show that it is a Cauchy sequence. Let $\varepsilon > 0$. By the parallelogram law, for all $n, m \in \mathbb{N}$,

$$\begin{aligned} \|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 &= \|(\boldsymbol{y}_n - \boldsymbol{x}) - (\boldsymbol{y}_m - \boldsymbol{x})\|^2 \\ &= 2\,(\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + \|\boldsymbol{y}_m - \boldsymbol{x}\|^2) - \|(\boldsymbol{y}_n - \boldsymbol{x}) + (\boldsymbol{y}_m - \boldsymbol{x})\|^2 \\ &= 2\,(\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + \|\boldsymbol{y}_m - \boldsymbol{x}\|^2) - \|\boldsymbol{y}_n + \boldsymbol{y}_m - 2\,\boldsymbol{x}\|^2 \end{aligned}$$

Since $\boldsymbol{A}$ is convex, $\frac{1}{2}\,(\boldsymbol{y}_n + \boldsymbol{y}_m) \in \boldsymbol{A}$. It follows that

$$\|\boldsymbol{y}_n + \boldsymbol{y}_m - 2\,\boldsymbol{x}\|^2 = 4 \left\| \frac{1}{2}\,(\boldsymbol{y}_n + \boldsymbol{y}_m) - \boldsymbol{x} \right\|^2 \geq 4\,\gamma^2.$$

We deduce that

$$\begin{aligned} \|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 &\leq 2\left(\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 + \|\boldsymbol{y}_m - \boldsymbol{x}\|^2\right) - 4\gamma^2 \\ &= 2\left(\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 - \gamma^2\right) + 2\left(\|\boldsymbol{y}_m - \boldsymbol{x}\|^2 - \gamma^2\right). \end{aligned}$$

From (2.31), we have also $\lim_{k\to\infty} \|\boldsymbol{y}_k - \boldsymbol{x}\|^2 = \gamma^2$, so that there exists $p \in \mathbb{N}$ such that for all $k > p$ we have $\|\boldsymbol{y}_k - \boldsymbol{x}\|^2 - \gamma^2 < \frac{\varepsilon^2}{4}$. Hence, for all $n, m > p$ we have

$$\|\boldsymbol{y}_n - \boldsymbol{y}_m\|^2 \leq 2\left(\|\boldsymbol{y}_n - \boldsymbol{x}\|^2 - \gamma^2\right) + 2\left(\|\boldsymbol{y}_m - \boldsymbol{x}\|^2 - \gamma^2\right) < 2\,\frac{\varepsilon^2}{4} + 2\,\frac{\varepsilon^2}{4} = \varepsilon^2,$$

i.e., for all $n, m > p$,

$$\|\boldsymbol{y}_n - \boldsymbol{y}_m\| < \varepsilon.$$

It follows that the sequence $(\boldsymbol{y}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, and therefore it is convergent to some element $\hat{\boldsymbol{y}} \in \boldsymbol{H}$. On the other hand, since $\boldsymbol{A}$ is closed, the limit of every convergent sequence of elements of $\boldsymbol{A}$ is also in $\boldsymbol{A}$, and therefore $\hat{\boldsymbol{y}} \in \boldsymbol{A}$.

From *(iii)* in Theorem 2.20,

$$\|\hat{\boldsymbol{y}} - \boldsymbol{x}\| = \lim_{k\to\infty} \|\boldsymbol{y}_k - \boldsymbol{x}\| = \gamma.$$

Thus, $\gamma = f(\hat{\boldsymbol{y}})$, which means that $\gamma = \min\{\|\boldsymbol{y} - \boldsymbol{x}\| \mid \boldsymbol{y} \in \boldsymbol{A}\}$. In other words, f has a minimum in $\boldsymbol{A}$, which is γ, and $\hat{\boldsymbol{y}}$ is a point of minimum.

To conclude the proof, we need to show the uniqueness of the point of minimum. More explicitly, we need to show that, if $\hat{\boldsymbol{y}}_1, \hat{\boldsymbol{y}}_2$ are elements of $\boldsymbol{A}$ such that

$$\|\hat{\boldsymbol{y}}_1 - \boldsymbol{x}\| = \|\hat{\boldsymbol{y}}_2 - \boldsymbol{x}\| = \gamma \tag{2.32}$$

(i.e., $\hat{\boldsymbol{y}}_1$ and $\hat{\boldsymbol{y}}_2$ are both points of minimum), then $\hat{\boldsymbol{y}}_1 = \hat{\boldsymbol{y}}_2$. To this end, we consider the following function

$$\begin{aligned} \varphi : [0,1] &\longrightarrow \mathbb{R} \\ t &\mapsto \|(1-t)\,\hat{\boldsymbol{y}}_1 + t\,\hat{\boldsymbol{y}}_2 - \boldsymbol{x}\|^2. \end{aligned}$$

Since $\boldsymbol{A}$ is convex, for all $t \in [0,1]$ we have

$$(1-t)\,\hat{\boldsymbol{y}}_1 + t\,\hat{\boldsymbol{y}}_2 = \hat{\boldsymbol{y}}_1 + t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1) \in \boldsymbol{A}.$$

It follows that for all $t \in [0,1]$, we have $\varphi(t) \geq \gamma^2$. On the other hand,

$$\begin{aligned} \|(1-t)\,\hat{\boldsymbol{y}}_1 + t\,\hat{\boldsymbol{y}}_2 - \boldsymbol{x}\| &= \|(1-t)(\hat{\boldsymbol{y}}_1 - \boldsymbol{x}) + t\,(\hat{\boldsymbol{y}}_2 - \boldsymbol{x})\| \\ &\leq (1-t)\,\|\hat{\boldsymbol{y}}_1 - \boldsymbol{x}\| + t\,\|\hat{\boldsymbol{y}}_2 - \boldsymbol{x}\| = \gamma, \end{aligned}$$

where we have used (2.32), so that for all $t \in [0,1]$, we have $\varphi(t) = \gamma^2$, i.e., φ is constant over the interval $[0,1]$. We also observe that

$$\begin{aligned} \varphi(t) &= \|t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1) + \hat{\boldsymbol{y}}_1 - \boldsymbol{x}\|^2 \\ &= \langle t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1) + \hat{\boldsymbol{y}}_1 - \boldsymbol{x},\, t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1) + \hat{\boldsymbol{y}}_1 - \boldsymbol{x}\rangle \\ &= \langle t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1), t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1)\rangle + \langle t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1), \hat{\boldsymbol{y}}_1 - \boldsymbol{x}\rangle + \langle \hat{\boldsymbol{y}}_1 - \boldsymbol{x}, t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1)\rangle + \langle \hat{\boldsymbol{y}}_1 - \boldsymbol{x}, \hat{\boldsymbol{y}}_1 - \boldsymbol{x}\rangle \\ &= \|t\,(\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1)\|^2 + t\,\langle \hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1, \hat{\boldsymbol{y}}_1 - \boldsymbol{x}\rangle + t\,\langle \hat{\boldsymbol{y}}_1 - \boldsymbol{x}, \hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\rangle + \|\hat{\boldsymbol{y}}_1 - \boldsymbol{x}\|^2 \\ &= t^2\,\|\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\|^2 + t\,\overline{\langle \hat{\boldsymbol{y}}_1 - \hat{\boldsymbol{y}}_1, \hat{\boldsymbol{y}}_2 - \boldsymbol{x}\rangle} + t\,\langle \hat{\boldsymbol{y}}_1 - \boldsymbol{x}, \hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\rangle + \|\hat{\boldsymbol{y}}_1 - \boldsymbol{x}\|^2 \\ &= t^2\,\|\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\|^2 + 2\,t\,\mathfrak{Re}\left\{\langle \hat{\boldsymbol{y}}_1 - \boldsymbol{x}, \hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\rangle\right\} + \|\hat{\boldsymbol{y}}_1 - \boldsymbol{x}\|^2. \end{aligned}$$

The second derivative is $\varphi''(t) = 2\|\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\|^2$, and since φ is constant in $[0,1]$, we have $\varphi''(t) = 0$ for all $t \in [0,1]$. Thus, for all $t \in [0,1]$ we have

$$2\|\hat{\boldsymbol{y}}_2 - \hat{\boldsymbol{y}}_1\|^2 = 0.$$

It follows that $\hat{\boldsymbol{y}}_1 = \hat{\boldsymbol{y}}_2$, and the proof is complete.

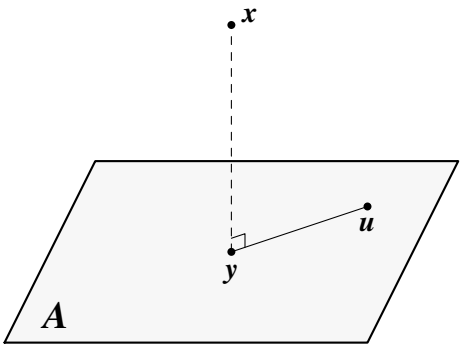
Definition 2.20 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space, and let us denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Let $\boldsymbol{A} \subseteq \boldsymbol{H}$ be non-empty, convex and closed. Let $\boldsymbol{x} \in \boldsymbol{H}$. Then, the unique minimum point in $\boldsymbol{A}$ of the function*

$$\begin{aligned} f : \boldsymbol{A} &\longrightarrow \mathbb{R} \\ \boldsymbol{y} &\mapsto \|\boldsymbol{y} - \boldsymbol{x}\| \end{aligned}$$

is denoted by $P_{\boldsymbol{A}}(\boldsymbol{x})$.

The following result requires, in addition to the previous assumptions, that $\boldsymbol{A}$ be a subspace. A subspace is non-empty and it is closed under linear combinations of its vectors, so that it is always convex. However, as we have seen in Theorem 2.6 and Example 2.2, if the vector space is infinite-dimensional, a subspace needs not be closed.

We recall that two elements $\boldsymbol{x}$ and $\boldsymbol{y}$ of an inner product space $(\boldsymbol{X}, \langle\cdot,\cdot\rangle)$ are orthogonal if $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$. We show that if $\boldsymbol{A}$ is a closed subspace, the point $\boldsymbol{y}$ of $\boldsymbol{A}$ with minimum distance from $\boldsymbol{x}$ enjoys an important property: the vector $\boldsymbol{x} - \boldsymbol{y}$ is orthogonal to $\boldsymbol{A}$, i.e., it is orthogonal to every vector $\boldsymbol{u} \in \boldsymbol{A}$. This is the generalization of a property that you used in Linear Algebra to characterize orthogonal projections.



Theorem 2.24 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space, and let us denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Let $\boldsymbol{A} \subseteq \boldsymbol{H}$ be a closed subspace of $\boldsymbol{H}$. Let $\boldsymbol{x} \in \boldsymbol{H}$. Then, given $\boldsymbol{y} \in \boldsymbol{A}$, the following conditions are equivalent:*

(A) $\boldsymbol{y} = P_{\boldsymbol{A}}(\boldsymbol{x})$*;*

(B) $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle = 0$ *for all* $\boldsymbol{u} \in \boldsymbol{A}$.

Proof We prove that **(A)** implies **(B)**. We check that, if $\boldsymbol{y} = P_{\boldsymbol{A}}(\boldsymbol{x})$, then for all $\boldsymbol{u} \in \boldsymbol{A}$ we have $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle = 0$. Let $\boldsymbol{u} \in \boldsymbol{A}$. Since $\boldsymbol{y} + t\,\boldsymbol{u} \in \boldsymbol{A}$ for all $t \in \mathbb{F}$, the function

$$\begin{aligned} P : \mathbb{F} &\longrightarrow \mathbb{R} \\ t &\mapsto \|\boldsymbol{x} - (\boldsymbol{y} + t\,\boldsymbol{u})\|^2 \end{aligned}$$

has minimum at $t = 0$. We have

$$P(t) = \|\boldsymbol{x} - \boldsymbol{y}\|^2 - 2\,t\,\Re\mathrm{e}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle\} + t^2\,\|\boldsymbol{u}\|^2.$$

Hence, P is a polynomial function in t. Since 0 is a point of minimum for P, we have $P'(0) = 0$, which implies

$$\Re\mathrm{e}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle\} = 0.$$

If $\mathbb{F} = \mathbb{R}$, this part of the proof is complete. If $\mathbb{F} = \mathbb{C}$, we also need to show that $\Im\mathrm{m}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle\} = 0$. To this end, we notice that we have $i\,\boldsymbol{u} \in \boldsymbol{A}$, and therefore also $\Re\mathrm{e}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, i\,\boldsymbol{u}\rangle\} = 0$. Since for every $z \in \mathbb{C}$ we have $\Im\mathrm{m}\{z\} = \Re\mathrm{e}\{-i\,z\}$ (why?), we find

$$\Im\mathrm{m}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle\} = \Re\mathrm{e}\,\{-i\,\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle\} = \Re\mathrm{e}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, i\,\boldsymbol{u}\rangle\} = 0.$$

This shows that **(A)** implies **(B)**. Conversely, let $\boldsymbol{y} \in \boldsymbol{A}$ be such that $\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{u}\rangle = 0$ for all $\boldsymbol{u} \in \boldsymbol{A}$. Then, given $\boldsymbol{u} \in \boldsymbol{A}$, we have

$$\begin{aligned}\|\boldsymbol{x} - \boldsymbol{u}\|^2 &= \|(\boldsymbol{x} - \boldsymbol{y}) + (\boldsymbol{y} - \boldsymbol{u})\|^2\\ &= \|\boldsymbol{x} - \boldsymbol{y}\|^2 + 2\,\Re\mathrm{e}\,\{\langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{y} - \boldsymbol{u}\rangle\} + \|\boldsymbol{y} - \boldsymbol{u}\|^2\\ &= \|\boldsymbol{x} - \boldsymbol{y}\|^2 + \|\boldsymbol{y} - \boldsymbol{u}\|^2 \geq \|\boldsymbol{x} - \boldsymbol{y}\|^2,\end{aligned}$$

where the last equality is a consequence of the fact that $\boldsymbol{y} - \boldsymbol{u} \in \boldsymbol{A}$. Thus, $\boldsymbol{y} = P_{\boldsymbol{A}}(\boldsymbol{x})$.

In the case where $\boldsymbol{A}$ is a finite dimensional subspace of a Hilbert space, it is possible to give an explicit formula for $P_{\boldsymbol{A}}$.

Definition 2.21 [Orthonormal System of Vectors]
Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Let I be an arbitrary set. Let $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ be a family of elements of $\boldsymbol{H}$. We say that $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ is an orthonormal system *if, for all $\alpha, \beta \in I$,*

$$\langle \boldsymbol{e}_\alpha, \boldsymbol{e}_\beta\rangle = \delta_{\alpha,\beta}, \tag{2.33}$$

where $\delta_{\alpha,\beta}$ is the Kronecker delta.

Remark 2.3 In Definition 2.21, we introduced an orthonormal system as a family of vectors, rather than a set of vectors. The family $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ should not be confused with the set $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$. In fact, $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ is the function $\boldsymbol{e}$ with domain I such that $\boldsymbol{e}(\alpha) = \boldsymbol{e}_\alpha$ for any $\alpha \in I$. Thus, if $\alpha_1, \alpha_2 \in I$ and $\alpha_1 \neq \alpha_2$, it may happen that $\boldsymbol{e}_{\alpha_1} = \boldsymbol{e}_{\alpha_2}$. A sequence $(\boldsymbol{e}_n)_{n\in\mathbb{N}}$ is an example of family. We could have defined an orthonormal system of vectors defining it as a set $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$, because the condition (2.33) is never satisfied if we have repetitions (why?). However, in what follows it will be important to refer to families and not sets of vectors.

In essence, in Definition 2.21 we require that the vectors $\boldsymbol{e}_\alpha$ be pairwise orthogonal and have unitary norm. In a completely equivalent way, we can say that $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ is an orthonormal system of vectors in the Hilbert space $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ if

(i) $\forall\, \alpha, \beta \in I, \qquad \alpha \neq \beta \;\Rightarrow\; \boldsymbol{e}_\alpha \perp \boldsymbol{e}_\beta$

(ii) $\forall\, \alpha \in I, \qquad \|\boldsymbol{e}_\alpha\| = 1,$

where $\|\cdot\|$ denotes, as usual, the norm induced by the inner product $\langle\cdot,\cdot\rangle$.

For example, the standard basis of $\mathbb{R}^n$ is an orthonormal system with respect to the standard dot product. The same is true in $\mathbb{C}^n$ with respect to the Hermitian product.

Example 2.29

Consider the Hilbert space $\ell^2(\mathbb{N};\mathbb{F})$. An orthonormal system is given by $(\boldsymbol{e}_i)_{i\in\mathbb{N}}$, where

$$\boldsymbol{e}_1 = (1,0,0,\ldots), \qquad \boldsymbol{e}_2 = (0,1,0,\ldots), \qquad \boldsymbol{e}_3 = (0,0,1,\ldots), \qquad \ldots$$

for all $j \in \mathbb{N}^*$. Indeed, $\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \delta_{i,j}$.

The following result is a generalization of the Gram-Schmidt method that you have seen in Linear Algebra.

Theorem 2.25 [Gram-Schmidt Method]
Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Let $\boldsymbol{A}$ be a finite-dimensional subspace of $\boldsymbol{H}$. Then, $\boldsymbol{A}$ has an orthonormal basis with respect to the inner product of $\boldsymbol{H}$.

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 $\boldsymbol{A}$ starting from an arbitrary basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r$ of $\boldsymbol{A}$, where r is the dimension of $\boldsymbol{A}$. Let us denote by $\|\cdot\|$ the norm induced by the inner product $\langle\cdot,\cdot\rangle$. We define $\boldsymbol{q}_1 = \boldsymbol{u}_1$ and compute

$$\boldsymbol{e}_1 = \frac{1}{\|\boldsymbol{q}_1\|}\,\boldsymbol{q}_1.$$

This is always possible because $\boldsymbol{u}_1 \neq 0$ (it is a basis vector). Thus, $\|\boldsymbol{e}_1\| = 1$. We compute

$$\boldsymbol{q}_2 = \boldsymbol{u}_2 - \langle \boldsymbol{u}_2, \boldsymbol{e}_1\rangle\,\boldsymbol{e}_1.$$

We have $\boldsymbol{q}_2 \neq 0$, because $\boldsymbol{e}_1 \in \operatorname{span}\{\boldsymbol{q}_1\} = \operatorname{span}\{\boldsymbol{u}_1\}$ and $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are linearly independent. We define

$$\boldsymbol{e}_2 = \frac{1}{\|\boldsymbol{q}_2\|}\,\boldsymbol{q}_2.$$

We observe that $\boldsymbol{e}_2$ is orthogonal to $\boldsymbol{e}_1$, because such is $\boldsymbol{q}_2$. Thus, we have constructed $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ with unit norm and such that $\langle \boldsymbol{e}_1, \boldsymbol{e}_2\rangle = 0$. We observe also that, since $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are linearly independent, we have

$$\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}.$$

We assume that we have constructed, for some $j < r$, the set of vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j\}$ such that, for all $i, j \in \{1, 2, \ldots, r-1\}$,

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \delta_{i,j}$$

and

$$\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_j\}. \tag{2.34}$$

We define

$$\boldsymbol{q}_{j+1} = \boldsymbol{u}_{j+1} - \sum_{k=1}^{j} c_k\,\boldsymbol{e}_k$$

in such a way that $\langle \boldsymbol{q}_{j+1}, \boldsymbol{e}_k\rangle = 0$ for $k \in \{1, 2, \ldots, j\}$. If $h \in \{1, 2, \ldots, j\}$, we have

$$\langle \boldsymbol{q}_{j+1}, \boldsymbol{e}_h\rangle = \left\langle \boldsymbol{u}_{j+1} - \sum_{k=1}^{j} c_k\,\boldsymbol{e}_k, \boldsymbol{e}_h \right\rangle = \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_h\rangle - \sum_{k=1}^{j} c_k\,\langle \boldsymbol{e}_k, \boldsymbol{e}_h\rangle = \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_h\rangle - c_h = 0.$$

Defining $c_h = \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_h\rangle$ for all $h \in \{1, 2, \ldots, j\}$, we find

$$\boldsymbol{q}_{j+1} = \boldsymbol{u}_{j+1} - \sum_{k=1}^{j} \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_k\rangle\,\boldsymbol{e}_k. \tag{2.35}$$

We notice that $\boldsymbol{q}_{j+1} \neq \boldsymbol{0}$. In fact, if this was not the case, we would have

$$\boldsymbol{u}_{j+1} = \sum_{k=1}^{j} \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_k \rangle \, \boldsymbol{e}_k,$$

which implies $\boldsymbol{u}_{j+1} \in \operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_j\}$, thus contradicting the linear independence of the set of vectors $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$. We are therefore allowed to define

$$\boldsymbol{e}_{j+1} = \frac{1}{\|\boldsymbol{q}_{j+1}\|} \, \boldsymbol{q}_{j+1}. \tag{2.36}$$

Clearly, $\|\boldsymbol{e}_{j+1}\| = 1$ and $\langle \boldsymbol{e}_{j+1}, \boldsymbol{e}_k \rangle = 0$ for all $k \in \{1, 2, \ldots, j\}$. Moreover,

$$\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j, \boldsymbol{e}_{j+1}\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_j, \boldsymbol{u}_{j+1}\}.$$

The inclusion $\subseteq$ from the left to the right is obvious: every vector $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j, \boldsymbol{e}_{j+1}$ has been defined as a linear combination of $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_j, \boldsymbol{u}_{j+1}$. We only need to justify the inclusion $\supseteq$. If $\boldsymbol{\xi} \in \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_j, \boldsymbol{u}_{j+1}\}$, there exist $a_k \in \mathbb{F}$ for all $k \in \{1, \ldots, j+1\}$ such that

$$\boldsymbol{\xi} = \sum_{k=1}^{j+1} a_k \, \boldsymbol{u}_k.$$

From (2.34), there exist $b_1, b_2, \ldots, b_j \in \mathbb{F}$ such that

$$\sum_{k=1}^{j} a_k \, \boldsymbol{u}_k = \sum_{k=1}^{j} b_k \, \boldsymbol{e}_k,$$

so that

$$\begin{aligned} \boldsymbol{\xi} &= \sum_{k=1}^{j} b_k \, \boldsymbol{e}_k + a_{j+1} \, \boldsymbol{u}_{j+1} \\ &= \sum_{k=1}^{j} b_k \, \boldsymbol{e}_k + a_{j+1} \left(\boldsymbol{q}_{j+1} + \sum_{k=1}^{j} \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_k \rangle \, \boldsymbol{e}_k \right) \\ &= \sum_{k=1}^{j} (b_k + a_{j+1} \langle \boldsymbol{u}_{j+1}, \boldsymbol{e}_k \rangle) \, \boldsymbol{e}_k + a_{j+1} \, \|\boldsymbol{q}_{j+1}\| \, \boldsymbol{e}_{j+1}, \end{aligned}$$

where we have used (2.35) and (2.36). This shows that $\boldsymbol{\xi} \in \operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_j, \boldsymbol{e}_{j+1}\}$. Applying this method r times, we can construct the desired orthonormal basis.

We are going to give a formula for P_A in the case where $\boldsymbol{A}$ is a finite dimensional subspace of $\boldsymbol{H}$: this is another generalization of a well-known result from Linear Algebra.

Theorem 2.26 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space, and let us denote by $\| \cdot \|$ the norm induced by the inner product $\langle \cdot, \cdot \rangle$. Let $\boldsymbol{A}$ be a finite-dimensional subspace of $\boldsymbol{H}$. Let $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_r)$ be an orthonormal basis of $\boldsymbol{A}$. Then,*

$$\forall \, \boldsymbol{x} \in A, \qquad P_A(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle \, \boldsymbol{e}_1 + \langle \boldsymbol{x}, \boldsymbol{e}_2 \rangle \, \boldsymbol{e}_2 + \ldots + \langle \boldsymbol{x}, \boldsymbol{e}_r \rangle \, \boldsymbol{e}_r.$$

Proof We can write

$$P_A(\boldsymbol{x}) = c_1 \, \boldsymbol{e}_1 + c_2 \, \boldsymbol{e}_2 + \ldots + c_r \, \boldsymbol{e}_r$$

for some elements $c_1, c_2, \ldots, c_r \in \mathbb{F}$. Then, for every $j \in \{1, \ldots, r\}$ we must have

$$0 = \langle \boldsymbol{x} - (c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2 + \ldots + c_r \boldsymbol{e}_r), \boldsymbol{e}_j \rangle = \langle \boldsymbol{x}, \boldsymbol{e}_j \rangle - c_j.$$

Using the same notation of Theorem 2.26, we can write

$$\begin{aligned} \|\boldsymbol{x}\|^2 &= \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \langle (\boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x})) + P_{\boldsymbol{A}}(\boldsymbol{x}), (\boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x})) + P_{\boldsymbol{A}}(\boldsymbol{x}) \rangle \\ &= \langle \boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x}), \boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x}) \rangle + \cancel{\langle \boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x}), P_{\boldsymbol{A}}(\boldsymbol{x}) \rangle} + \cancel{\langle P_{\boldsymbol{A}}(\boldsymbol{x}), \boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x}) \rangle} + \langle P_{\boldsymbol{A}}(\boldsymbol{x}), P_{\boldsymbol{A}}(\boldsymbol{x}) \rangle \\ &= \|\boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x})\|^2 + \|P_{\boldsymbol{A}}(\boldsymbol{x})\|^2, \end{aligned}$$

where we have used the fact that, since $\boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x})$ is orthogonal to every vector of $\boldsymbol{A}$, it is in particular orthogonal to $P_{\boldsymbol{A}}(\boldsymbol{x})$. We obtain

$$\|\boldsymbol{x} - P_{\boldsymbol{A}}(\boldsymbol{x})\|^2 = \|\boldsymbol{x}\|^2 - \|P_{\boldsymbol{A}}(\boldsymbol{x})\|^2. \tag{2.37}$$

Exercises

Exercise 2.43

Consider a Hilbert space $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ where $\boldsymbol{H}$ is a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Using Theorem 2.24, show that, if $\boldsymbol{A}$ is a closed subspace in a Hilbert space $\boldsymbol{X}$, the operator $P_{\boldsymbol{A}} : \boldsymbol{H} \longrightarrow \boldsymbol{A}$ is linear. [Solution: (2.43)]

Exercise 2.44

Consider a Hilbert space $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ where $\boldsymbol{H}$ is a vector space over the field $\mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$). Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be pairwise orthogonal elements of $\boldsymbol{H}$. Show that

$$\|\boldsymbol{x}_1 + \ldots + \boldsymbol{x}_n\|^2 = \|\boldsymbol{x}_1\|^2 + \ldots + \|\boldsymbol{x}_n\|^2.$$

This formula can be considered a generalization of the Pythagorean theorem. [Solution: (2.44)]

Exercise 2.45

Let $A \subseteq \mathbb{R}$. Let $\alpha \in \mathbb{R}$. Show that:

- α is the supremum of A if and only if
  - *(i)* $\forall\, x \in A, \quad x \leq \alpha$;
  - *(ii)* $\forall\, \epsilon > 0, \quad \exists\, x \in A : \quad \alpha - \epsilon < x$.
- α is the infimum of A if and only if
  - *(i)* $\forall\, x \in A, \quad \alpha \leq x$;
  - *(ii)* $\forall\, \epsilon > 0, \quad \exists\, x \in A : \quad x < \alpha + \epsilon$.

[Solution: (2.45)]

2.7 Separable Hilbert spaces

The problem that we want to study is the following: if $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ is a Hilbert space and we have an orthonormal system $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ in $\boldsymbol{H}$, where I is an index set, we want to find under which conditions, for every $\boldsymbol{x} \in \boldsymbol{H}$, there exists a set $\{c_\alpha \,|\, \alpha \in I\}$ of scalars such that

$$\boldsymbol{x} = \sum_{\alpha \in I} c_\alpha \, \boldsymbol{e}_\alpha. \tag{2.38}$$

The first task is to give a meaning to the summation in (2.38) in the case where I is possibly not finite (and possibly uncountable). This is accomplished with the following definition.

Definition 2.22 [Unordered Sum]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let I be an arbitrary set. Let $(\boldsymbol{x}_\alpha)_{\alpha\in I}$ be a family of elements of $\boldsymbol{X}$. Let $\boldsymbol{s}\in\boldsymbol{X}$. We shall write

$$\sum_{\alpha\in I}\boldsymbol{x}_\alpha=\boldsymbol{s} \tag{2.39}$$

if, for all $\varepsilon>0$, there exists a finite subset F_ε of I such that, for any finite subset G of I,

$$F_\varepsilon\subseteq G \quad\Rightarrow\quad \Big\|\sum_{\alpha\in G}\boldsymbol{x}_\alpha-\boldsymbol{s}\Big\|<\varepsilon.$$

If (2.39) holds for some $s\in\boldsymbol{X}$, we will say that the unordered sum $\sum_{\alpha\in I}\boldsymbol{x}_\alpha$ *is convergent to $\boldsymbol{s}$, or that its sum is $\boldsymbol{s}$.*

Loosely speaking, this definition says that an unordered sum converges to $\boldsymbol{s}$ if, for every $\varepsilon>0$, there exists a finite subset of its terms $\{\boldsymbol{x}_\alpha\,|\,\alpha\in F_\varepsilon\}$ whose sum constutes a large portion of $\boldsymbol{s}$, in the sense that the terms outside this set are insignificant: no matter what other set of terms we choose, as long as we include $\{\boldsymbol{x}_\alpha\,|\,\alpha\in F_\varepsilon\}$, their corresponding sum is always close to $\boldsymbol{s}$.

Differently from an ordinary sum, an unordered sum is not constrained with a particular order of addition of its terms. It is easy to see that, when I is finite, the unordered sum (2.39) coincides with the usual sum.

Exercise 2.46

Given a family $(\boldsymbol{x}_\alpha)_{\alpha\in I}$ of elements in the normed space $\boldsymbol{X}$, show that there is at most one element $\boldsymbol{s}$ in $\boldsymbol{X}$ for which (2.39) holds true. [Solution: (2.46)]

Exercise 2.47

If $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space and there exists $M\geq 0$ such that, for every F finite subset of I there holds

$$\sum_{\alpha\in F}\|\boldsymbol{x}_\alpha\|\leq M, \tag{2.40}$$

show that:

(a) the unordered sum $\sum_{\alpha\in I}\boldsymbol{x}_\alpha$ is convergent to some $\boldsymbol{s}$, and $\|\boldsymbol{s}\|\leq M$;

(b) for every $\delta>0$, the set $\{\alpha\in I\,|\,\|\boldsymbol{x}_\alpha\|\geq\delta\}$ is finite. Moreover, $\{\alpha\in I\,|\,\|\boldsymbol{x}_\alpha\|\neq 0\}$ is at most countable (i.e., it is finite or countable).

[Solution: (2.47)]

Exercise 2.48

Let $(\boldsymbol{X}, \|\cdot\|_X)$ and $(\boldsymbol{Y}, \|\cdot\|_Y)$ be normed spaces and let $T\in\mathcal{L}(\boldsymbol{X};\boldsymbol{Y})$ (i.e., T is a linear and continuous operator from $\boldsymbol{X}$ to $\boldsymbol{Y}$). Let $(\boldsymbol{x}_\alpha)_{\alpha\in I}$ be a family of elements in $\boldsymbol{X}$ such that $\sum_{\alpha\in I}\boldsymbol{x}_\alpha=\boldsymbol{s}$ (where $\boldsymbol{s}\in\boldsymbol{X}$). Show that

$$\sum_{\alpha\in I}T(\boldsymbol{x}_\alpha)=T(\boldsymbol{s}).$$

[Solution: (2.48)]

Remark 2.4 Moreover:

(III) If $\boldsymbol{X} = \mathbb{C}$ and $I = \mathbb{N}$, if the series $\sum_{i=0}^{\infty} x_i$ is absolutely convergent, then

$$\sum_{i=0}^{\infty} x_i = \sum_{i \in \mathbb{N}} x_i.$$

Definition 2.23 [Cauchy Unordered Sum]
Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let I be an arbitrary set. Let $(\boldsymbol{x}_\alpha)_{\alpha \in I}$ be a family of elements of $\boldsymbol{X}$. The unordered sum

$$\sum_{\alpha \in I} \boldsymbol{x}_\alpha$$

is said to be Cauchy *if, for every $\varepsilon > 0$, there exists a finite subset F_ε of I such that, for all finite subsets G and H of I,*

$$F_\varepsilon \subseteq G \cap H \quad \Rightarrow \quad \left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| < \varepsilon. \tag{2.41}$$

Theorem 2.27 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let I be an arbitrary set. Let $(\boldsymbol{x}_\alpha)_{\alpha \in I}$ be a family of elements of $\boldsymbol{X}$. The unordered sum $\sum_{\alpha \in I} \boldsymbol{x}_\alpha$ is Cauchy if and only if, for every $\varepsilon > 0$, there exists a finite subset F_ε of I such that, for all finite subsets L of I satisfying $F_\varepsilon \cap L = \varnothing$,*

$$\left\| \sum_{\alpha \in L} \boldsymbol{x}_\alpha \right\| < \varepsilon. \tag{2.42}$$

Proof Suppose that the unordered sum $\sum_{\alpha \in I} \boldsymbol{x}_\alpha$ is Cauchy. Let $\varepsilon > 0$. Let F_ε be the set of indices satisfying the condition in Definition 2.23, i.e., such that, for all finite subsets G and H of I, if $F_\varepsilon \subseteq G \cap H$ we have

$$\left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| < \varepsilon. \tag{2.43}$$

Now, we show that for every finite subset L of I satisfying $F_\varepsilon \cap L = \varnothing$, the inequality (2.42) holds. Then, letting $G = F_\varepsilon \cup L$ and $H = F_\varepsilon$, both G and H contain F_ε, and therefore (2.43) gives

$$\left\| \sum_{\alpha \in L} \boldsymbol{x}_\alpha \right\| = \left\| \sum_{\alpha \in F_\varepsilon \cup L} \boldsymbol{x}_\alpha - \sum_{\alpha \in F_\varepsilon} \boldsymbol{x}_\alpha \right\| < \varepsilon.$$

Now, we show that if (2.42) holds, the unordered sum is Cauchy. Let $\varepsilon > 0$, and let F_ε be a finite subset of I such that for every other finite subset L of I we have

$$\left\| \sum_{\alpha \in L} \boldsymbol{x}_\alpha \right\| < \frac{\varepsilon}{2}.$$

For every finite subsets G and H of I containing F_ε, we have

$$\left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| = \left\| \sum_{\alpha \in G \setminus F_\varepsilon} \boldsymbol{x}_\alpha - \sum_{\alpha \in H \setminus F_\varepsilon} \boldsymbol{x}_\alpha \right\|$$
$$\leq \left\| \sum_{\alpha \in G \setminus F_\varepsilon} \boldsymbol{x}_\alpha \right\| + \left\| \sum_{\alpha \in H \setminus F_\varepsilon} \boldsymbol{x}_\alpha \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Theorem 2.28 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. Let I be an arbitrary set. Let $(\boldsymbol{x}_\alpha)_{\alpha \in I}$ be a family of elements of $\boldsymbol{X}$. Let the unordered sum $\sum_{\alpha \in I} \boldsymbol{x}_\alpha$ be convergent. Then, $\sum_{\alpha \in I} \boldsymbol{x}_\alpha$ is Cauchy.*

Exercise 2.49 Prove Theorem 2.28. [Solution: (2.49)]

We now show that if $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space, the converse of Theorem 2.28 holds.

Theorem 2.29 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a Banach space. Let I be an arbitrary set. Let $(\boldsymbol{x}_\alpha)_{\alpha \in I}$ be a family of elements of $\boldsymbol{X}$. The unordered sum $\sum_{\alpha \in I} \boldsymbol{x}_\alpha$ is convergent if and only if it is Cauchy.*

Proof The *only if* part is dealt with in Exercise 2.49. We now show that if the unordered sum is Cauchy, it is convergent. Let $\varepsilon = 1/k$, where $k \in \mathbb{N}^*$. Using the definition of Cauchy unordered sum, there exists a finite subset $R_{1/k}$ of I such that, for all finite subsets G and H of I,

$$R_{1/k} \subseteq G \cap H \quad \Rightarrow \quad \left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| < \varepsilon. \tag{2.44}$$

Now, we define, for all $k \in \mathbb{N}^*$,

$$F_{1/k} = \bigcup_{i=1}^{k} R_{1/i}.$$

The sequence $(F_{1/k})_{k \in \mathbb{N}^*}$ still satisfies the condition of Definition 2.23, but it is monotonically increasing (with respect to the partial ordering $\subseteq$). Now, we define the sequence $(\boldsymbol{s}_k)_{k \in \mathbb{N}^*}$, where

$$\boldsymbol{s}_k = \sum_{\alpha \in F_{1/k}} \boldsymbol{x}_\alpha.$$

We show that this is a Cauchy sequence. Let $k \in \mathbb{N}^*$. Let $n, m > k$. Since $F_{1/n} \supseteq F_{1/k}$ and $F_{1/m} \supseteq F_{1/k}$, we have $F_{1/n} \cap F_{1/m} \supseteq F_{1/k}$. In Definition 2.23 we take $G = F_{1/n}$ and $H = F_{1/m}$, so the condition of Definition 2.23 reads exactly as $F_{1/n} \cap F_{1/m} \supseteq F_{1/k}$. We obtain

$$\|\boldsymbol{s}_n - \boldsymbol{s}_m\| = \left\| \sum_{\alpha \in F_{1/n}} \boldsymbol{x}_\alpha - \sum_{\alpha \in F_{1/m}} \boldsymbol{x}_\alpha \right\| < \frac{1}{k}.$$

It follows that $(\boldsymbol{s}_k)_{k \in \mathbb{N}^*}$ is a Cauchy sequence, and therefore it admits limit $\boldsymbol{s}$ because $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space. We only need to prove that the unordered sum converges to $\boldsymbol{s}$. Let $\varepsilon > 0$. We must show that there exists a finite subset F_ε of I such that, for any finite subset G of I,

$$F_\varepsilon \subseteq G \quad \Rightarrow \quad \left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \right\| < \varepsilon.$$

Let $k_1 \in \mathbb{N}^*$ be such that $\|\boldsymbol{s} - \boldsymbol{s}_k\| < \varepsilon/2$, and let $k_2 \in \mathbb{N}^*$ be such that $k_2 > 2/\varepsilon$, and consider $k \geq \max\{k_1, k_2\}$. Let $G \supseteq F_{1/k}$. Then,

$$\left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \right\| \leq \left\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \sum_{\alpha \in F_{1/k}} \boldsymbol{x}_\alpha \right\| + \left\| \sum_{\alpha \in F_{1/k}} \boldsymbol{x}_\alpha - \boldsymbol{s} \right\| < \frac{1}{k} + \frac{\varepsilon}{2} < \varepsilon.$$

This concludes the proof.

With the notion of unordered sums in mind, we can now better understand the meaning of (2.38). The following result shows what expression the coefficients in (2.38) should have.

Theorem 2.30 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let I be an arbitrary set. Let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. Let $\boldsymbol{x} \in \boldsymbol{H}$, and let*

$$\boldsymbol{x} = \sum_{\alpha \in I} c_\alpha \, \boldsymbol{e}_\alpha. \tag{2.45}$$

Then,

$$\forall\, \alpha \in I, \quad c_\alpha = \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle.$$

Proof Let $\alpha_0 \in I$. Let

$$\begin{aligned} T : \boldsymbol{H} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \langle \boldsymbol{x}, \boldsymbol{e}_{\alpha_0} \rangle. \end{aligned}$$

It is easy to see that $T \in \mathcal{L}(\boldsymbol{H}, \mathbb{C})$. Indeed, the linearity is a consequence of the linearity in the first argument of the inner product. From Cauchy-Schwarz's inequality we have $|\langle \boldsymbol{x}, \boldsymbol{e}_{\alpha_0} \rangle| \leq \|\boldsymbol{x}\| \, \|\boldsymbol{e}_{\alpha_0}\|$, and so by taking $M = \|\boldsymbol{e}_{\alpha_0}\|$ we find $|T(\boldsymbol{x})| \leq M \, \|\boldsymbol{x}\|$, which says that T is bounded, and from Theorem 2.17 we conclude that T is continuous, and so $T \in \mathcal{L}(\boldsymbol{H}, \mathbb{C})$. Then, from Exercise 2.48

$$\langle \boldsymbol{x}, \boldsymbol{e}_{\alpha_0} \rangle = \sum_{\alpha \in I} c_\alpha \, \langle \boldsymbol{e}_\alpha, \boldsymbol{e}_{\alpha_0} \rangle = c_{\alpha_0}.$$

We now recall an important definition in Linear Algebra (but generalized in this context to deal with possibly infinite-dimensional vector spaces).

Definition 2.24 *Let $\boldsymbol{X}$ be a vector space over the field $\mathbb{F}$ and let $\boldsymbol{A} \subseteq \boldsymbol{X}$. We denote by $\operatorname{span}_{\mathbb{F}} \boldsymbol{A}$ the intersection of all subspaces of $\boldsymbol{X}$ containing $\boldsymbol{A}$.*

You will remember that $\operatorname{span}_{\mathbb{F}} \boldsymbol{A}$ is a subspace of $\boldsymbol{X}$. More precisely, it is the smallest subspace of $\boldsymbol{X}$ containing $\boldsymbol{A}$. Moreover, if $\boldsymbol{A} \neq \varnothing$,

$$\operatorname{span}_{\mathbb{F}} \boldsymbol{A} = \left\{ \sum_{\boldsymbol{v} \in \mathcal{F}} c_{\boldsymbol{v}} \, \boldsymbol{v} \;\middle|\; \mathcal{F} \subseteq \boldsymbol{A} \text{ finite and } c_{\boldsymbol{v}} \in \mathbb{F} \right\}. \tag{2.46}$$

We also say that $\operatorname{span} \boldsymbol{A}$ is the subspace of $\boldsymbol{X}$ "generated" by $\boldsymbol{A}$.

Theorem 2.31 [Bessel's Inequality]
Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let $\|\cdot\|$ be the norm induced by the inner product $\langle \cdot, \cdot \rangle$. Let I be an arbitrary set and let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. Let $\boldsymbol{x} \in \boldsymbol{H}$. Then, the unordered sum $\sum_{\alpha \in I} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2$ is

convergent. Moreover, Bessel's inequality

$$\sum_{\alpha \in I} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \leq \|\boldsymbol{x}\|^2 \tag{2.47}$$

holds.

Proof Let F be a finite subset of I. Let

$$\boldsymbol{\Omega} = \text{span}\{\boldsymbol{e}_\alpha \,|\, \alpha \in F\}.$$

Then, $P_{\boldsymbol{\Omega}}(\boldsymbol{x}) = \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha$. From the properties of linearity and anti-linearity of the inner product and from the fact that $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ is an orthonormal system, we find

$$\begin{aligned} \|P_{\boldsymbol{\Omega}}(\boldsymbol{x})\|^2 &= \langle P_{\boldsymbol{\Omega}}(\boldsymbol{x}), P_{\boldsymbol{\Omega}}(\boldsymbol{x}) \rangle \\ &= \left\langle \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha, \sum_{\beta \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\beta \rangle \boldsymbol{e}_\beta \right\rangle \\ &= \sum_{\alpha \in F} \left\langle \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha, \sum_{\beta \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\beta \rangle \boldsymbol{e}_\beta \right\rangle \\ &= \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \overline{\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle} \|\boldsymbol{e}_\alpha\|^2 = \sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2. \end{aligned}$$

From (2.37), we obtain

$$\sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 = \|P_{\boldsymbol{\Omega}}(\boldsymbol{x})\|^2 \leq \|\boldsymbol{x}\|^2.$$

We define also

$$\sigma = \sup \Big\{ \sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \;\Big|\; F \subseteq I \text{ finite} \Big\}.$$

Then, $\sigma \leq \|\boldsymbol{x}\|^2$. We show that

$$\sum_{\alpha \in I} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 = \sigma.$$

Let $\varepsilon > 0$. Then, there exists a finite subset F_ε of I such that

$$\sigma \geq \sum_{\alpha \in F_\varepsilon} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \geq \sigma - \varepsilon.$$

If $F_\varepsilon \subseteq F \subseteq I$ and F is finite,

$$\sigma - \varepsilon \leq \sum_{\alpha \in F_\varepsilon} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \leq \sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \leq \sigma.$$

It follows that

$$\Big| \sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 - \sigma \Big| \leq \varepsilon.$$

Lemma 2.4 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let $\boldsymbol{W}$ and $\boldsymbol{V}$ be closed subspaces of $\boldsymbol{H}$, with $\boldsymbol{W} \subseteq \boldsymbol{V}$. Then,*

$$\forall\, \boldsymbol{x} \in \boldsymbol{H}, \qquad P_W(\boldsymbol{x}) = P_W(P_V(\boldsymbol{x})).$$

Proof From the characterization of the projection in Theorem 2.24, if suffices to show that, for all $\boldsymbol{w} \in \boldsymbol{W}$,

$$\langle \boldsymbol{x} - P_{\boldsymbol{W}}(P_{\boldsymbol{V}}(\boldsymbol{x})), \boldsymbol{w} \rangle = 0.$$

In fact,

$$\langle \boldsymbol{x} - P_{\boldsymbol{W}}(P_{\boldsymbol{V}}(\boldsymbol{x})), \boldsymbol{w} \rangle = \langle \boldsymbol{x} - P_{\boldsymbol{V}}(\boldsymbol{x}), \boldsymbol{w} \rangle + \langle P_{\boldsymbol{V}}(\boldsymbol{x}) - P_{\boldsymbol{W}}(P_{\boldsymbol{V}}(\boldsymbol{x})), \boldsymbol{w} \rangle.$$

We have $\langle \boldsymbol{x} - P_{\boldsymbol{V}}(\boldsymbol{x}), \boldsymbol{w} \rangle = 0$ because $\boldsymbol{W} \subseteq \boldsymbol{V}$. Moreover, $\langle P_{\boldsymbol{V}}(\boldsymbol{x}) - P_{\boldsymbol{W}}(P_{\boldsymbol{V}}(\boldsymbol{x})), \boldsymbol{w} \rangle = 0$ from Theorem 2.24.

Exercise 2.50 Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space and let $\boldsymbol{V}$ be a subspace of $\boldsymbol{X}$. Show that the closure $\overline{\boldsymbol{V}}$ of $\boldsymbol{V}$ in $\boldsymbol{X}$ is a subspace of $\boldsymbol{X}$. [Solution: (2.50)]

Theorem 2.32 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Let I be an arbitrary set. Let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. Let*

$$\boldsymbol{V} = \operatorname{span}(\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}). \tag{2.48}$$

Let $\boldsymbol{x} \in \boldsymbol{H}$. Then, the unordered sum $\sum_{\alpha \in I} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha$ is convergent and

$$\sum_{\alpha \in I} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha = P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}),$$

where $P_{\overline{\boldsymbol{V}}}(\boldsymbol{x})$ denotes the projection of $\boldsymbol{x}$ onto the closure of $\boldsymbol{V}$.

Proof First, we notice that from Exercise 2.50 we know that $\overline{\boldsymbol{V}}$ is a closed subspace of $\boldsymbol{H}$, and therefore $P_{\overline{\boldsymbol{V}}}(\boldsymbol{x})$ is well defined. Let $\varepsilon > 0$. Since $P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}) \in \overline{\boldsymbol{V}}$, there exists $\boldsymbol{w} \in \boldsymbol{V}$ such that $\|\boldsymbol{w} - P_{\overline{\boldsymbol{V}}}(\boldsymbol{x})\| < \varepsilon$. Let F_ε be a finite subset of I such that

$$\boldsymbol{w} \in \operatorname{span}\{\boldsymbol{e}_\alpha \,|\, \alpha \in F_\varepsilon\},$$

and let $F_\varepsilon \subseteq F \subseteq I$, where F is finite. We define

$$\boldsymbol{W} = \operatorname{span}\{\boldsymbol{e}_\alpha \,|\, \alpha \in F\}.$$

From Lemma 2.4,

$$P_{\boldsymbol{W}}(P_{\overline{\boldsymbol{V}}}(\boldsymbol{x})) = P_{\boldsymbol{W}}(\boldsymbol{x}) = \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha.$$

We conclude that

$$\left\| P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}) - \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha \right\| \leq \|\boldsymbol{w} - P_{\overline{\boldsymbol{V}}}(\boldsymbol{x})\| < \varepsilon.$$

Theorem 2.33 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Let $\|\cdot\|$ be the norm induced by the inner product $\langle\cdot,\cdot\rangle$. Let I be an arbitrary set, and let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. Let $\boldsymbol{x} \in \boldsymbol{H}$. The following conditions are equivalent:*

(a) $\sum_{\alpha \in I} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 = \|\boldsymbol{x}\|^2$;

(b) $\boldsymbol{x} \in \overline{\boldsymbol{V}}$, *where* $\boldsymbol{V} = \operatorname{span}(\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\})$;

(c) $\sum_{\alpha \in I} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha = \boldsymbol{x}$.

Proof We show that **(a)** implies **(b)**. Suppose that **(a)** holds. Then, for all $n \in \mathbb{N}^*$ there exists $F_n \subseteq I$ finite such that

$$\sum_{\alpha \in F_n} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \geq \|\boldsymbol{x}\|^2 - \frac{1}{n}.$$

We define

$$\boldsymbol{w}_n = \sum_{\alpha \in F_n} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha.$$

Clearly, $\boldsymbol{w}_n \in \boldsymbol{V}$ and $\boldsymbol{w}_n = P_{\Omega_n}(\boldsymbol{x})$, with $\Omega_n = \text{span}\{\boldsymbol{e}_\alpha \,|\, \alpha \in F_n\}$. From (2.37) we have

$$0 \leq \|\boldsymbol{x} - \boldsymbol{w}_n\|^2 = \|\boldsymbol{x}\|^2 - \|\boldsymbol{w}_n\|^2 = \|\boldsymbol{x}\|^2 - \sum_{\alpha \in F_n} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2 \leq \frac{1}{n},$$

and therefore $\lim_{n \to \infty} \|\boldsymbol{w}_n - \boldsymbol{x}\| = 0$. We deduce that $\boldsymbol{x} \in \overline{\boldsymbol{V}}$. (XXX)
The implication **(b)** implies **(c)** follows from Theorem 2.32, since $\boldsymbol{x} \in \overline{\boldsymbol{V}}$ implies $P_{\overline{V}}(\boldsymbol{x}) = \boldsymbol{x}$.
We show that **(c)** implies **(a)**. Let $\varepsilon > 0$. There exists a finite subset F_ε of I such that, if $F_\varepsilon \subseteq F \subseteq I$ and F is finite

$$\left\| \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha - \boldsymbol{x} \right\|^2 \leq \varepsilon.$$

But we also have

$$\left\| \sum_{\alpha \in F} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha - \boldsymbol{x} \right\|^2 = \|\boldsymbol{x}\|^2 - \sum_{\alpha \in F} |\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle|^2.$$

Using Bessel's inequality, we deduce **(a)**.

Theorem 2.33 shows that the condition

$$\forall\, \boldsymbol{x} \in \boldsymbol{H}, \quad \boldsymbol{v} = \sum_{\alpha \in I} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \boldsymbol{e}_\alpha$$

is equivalent to requiring $\boldsymbol{V} = \text{span}(\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\})$ to be dense in $\boldsymbol{H}$. An equivalent condition on $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ will be given shortly.

Definition 2.25 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let I be an arbitrary set, and let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. We say that $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ is* **maximal** *if, whenever $(\boldsymbol{w}_\beta)_{\beta \in J}$ is an orthonormal system for $\boldsymbol{H}$ satisfying*

$$\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\} \subseteq \{\boldsymbol{w}_\beta \,|\, \beta \in J\}, \tag{2.49}$$

we always have $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\} = \{\boldsymbol{w}_\beta \,|\, \beta \in J\}$.

The following result establishes that the conditions **(a-c)** in Theorem 2.33 are equivalent to saying that $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ is maximal.

Theorem 2.34 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let I be an arbitrary set, and let $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ be an orthonormal system in $\boldsymbol{H}$. The following statements are equivalent:*

(i) $\boldsymbol{V} = \text{span}(\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\})$ is dense in $\boldsymbol{H}$;

(ii) $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ is maximal.

Proof Assume that **(i)** holds. Let $\{\boldsymbol{w}_\beta \,|\, \beta \in J\}$ be an orthonormal system in $\boldsymbol{H}$ such that (2.49) holds. Assume by contradiction that there exists some $\beta_0 \in J$ such that $\boldsymbol{w}_{\beta_0} \notin \{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$. Since **(i)** is assumed to hold, there

exists $F \subseteq I$ finite and $c_\alpha \in \mathbb{F}$ for all $\alpha \in I$ such that

$$\left\| \sum_{\alpha \in F} c_\alpha \, \boldsymbol{e}_\alpha - \boldsymbol{w}_{\beta_0} \right\| < 1. \tag{2.50}$$

We can assume with no loss of generality that the elements $\boldsymbol{e}_\alpha$ in (2.50) are distinct. Therefore, since $\{\boldsymbol{e}_\alpha \,|\, \alpha \in F\} \cup \{\boldsymbol{w}_{\beta_0}\} \subseteq \{\boldsymbol{w}_\beta \,|\, \beta \in J\}$, we find

$$\left\| \sum_{\alpha \in F} c_\alpha \, \boldsymbol{e}_\alpha - \boldsymbol{w}_\beta \right\|^2 = \sum_{\alpha \in F} |c_\alpha|^2 + 1 \geq 1,$$

which contradicts (2.50).
We now show that **(ii)** implies **(i)**. Suppose that $\boldsymbol{V}$ is not dense in $\boldsymbol{H}$. Thus, $\overline{\boldsymbol{V}}$ is a closed proper subspace of $\boldsymbol{H}$. Let $\boldsymbol{x}_0 \in \boldsymbol{H} \setminus \overline{\boldsymbol{V}}$. Then, $\boldsymbol{x}_0 - P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}_0) \neq \boldsymbol{0}$. We define

$$\boldsymbol{w} = \left\| \boldsymbol{x}_0 - P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}_0) \right\|^{-1} (P_{\overline{\boldsymbol{V}}}(\boldsymbol{x}_0)).$$

Then, $\|\boldsymbol{w}\| = 1$ and $\langle \boldsymbol{e}_\alpha, \boldsymbol{w} \rangle = 0$ for all $\alpha \in I$. We take some element β_0 such that $\beta_0 \notin I$, and define $J = I \cup \{\beta_0\}$, and define, for $\beta \in J$,

$$\boldsymbol{w}_\beta = \begin{cases} \boldsymbol{e}_\alpha & \text{if } \beta = \alpha \in I \\ \boldsymbol{w} & \text{if } \beta = \beta_0. \end{cases}$$

It is easily seen that $\{\boldsymbol{w}_\beta \,|\, \beta \in J\}$ is an orthonormal system in $\boldsymbol{H}$, (2.49) holds and $\boldsymbol{w} \in \{\boldsymbol{w}_\beta \,|\, \beta \in J\} \setminus \{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$, and this implies that $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ is not maximal.

We therefore have the following consequence.

Corollary 2.1 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let I be an arbitrary set, and let $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ be an orthonormal system in $\boldsymbol{H}$. The following statements are equivalent:*

(i) $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ is maximal;

(ii) $\forall \, \boldsymbol{x} \in \boldsymbol{H}, \quad \boldsymbol{x} = \sum_{\alpha \in I} \langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle \, \boldsymbol{e}_\alpha$.

Proof The result follows combining Theorem 2.33 and Theorem 2.34.

A criterion which is useful to establish whether an orthonormal system is maximal is the following.

Theorem 2.35 *Let $(\boldsymbol{H}, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let I be an arbitrary set, and let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be an orthonormal system in $\boldsymbol{H}$. Then, the following conditions are equivalent:*

(i) $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ is maximal;

(ii) if $\boldsymbol{x} \in \boldsymbol{H}$ and $\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle = 0$ for all $\alpha \in I$, then $\boldsymbol{x} = \boldsymbol{0}$.

Proof We show that **(i)** implies **(ii)**. Let $(\boldsymbol{e}_\alpha)_{\alpha \in I}$ be a maximal orthonormal system, and let $\boldsymbol{x} \in \boldsymbol{H}$ be such that $\langle \boldsymbol{x}, \boldsymbol{e}_\alpha \rangle = 0$ for all $i \in I$. Then, $\langle \boldsymbol{x}, \boldsymbol{w} \rangle = 0$ for any $\boldsymbol{w} \in \boldsymbol{V}$, where $\boldsymbol{V} = \operatorname{span}\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$. From the continuity of the inner product, we deduce that $\langle \boldsymbol{x}, \boldsymbol{w} \rangle = 0$ for any $\boldsymbol{w} \in \overline{\boldsymbol{V}} = \boldsymbol{H}$, because $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ is maximal. In particular,

$$\|\boldsymbol{x}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0,$$

which implies $\boldsymbol{x} = \boldsymbol{0}$.
We now show that **(ii)** implies **(i)**. If $\{\boldsymbol{e}_\alpha \,|\, \alpha \in I\}$ is not maximal, from Theorem 2.34 we find that $\overline{\boldsymbol{V}}$ is a proper

closed subspace of $\boldsymbol{H}$. Taking $\boldsymbol{x}_0 \in \boldsymbol{H} \setminus \overline{\boldsymbol{V}}$ and defining $\boldsymbol{w} = \boldsymbol{x}_0 - P_{\overline{V}}(\boldsymbol{x}_0)$, we have

$$\forall\, \alpha \in I, \quad \langle \boldsymbol{x}_0 - P_{\overline{V}}(\boldsymbol{x}_0), \boldsymbol{e}_\alpha \rangle = 0,$$

because $\boldsymbol{x}_i \in \overline{\boldsymbol{V}}$, but $\boldsymbol{x}_0 - P_{\overline{V}}(\boldsymbol{x}_0) \neq \boldsymbol{0}$.

Now we consider the problem of the existence of maximal orthonormal systems. The simplest situation is the one in which $\boldsymbol{H}$ is finite dimensional. In this case, a maximal orthonormal system is simply an orthonormal basis. This can be constructed from an arbitrary basis, using the Gram-Schmidt method. A more complex situation is the one of a separable Hilbert space.

Definition 2.26 *Let $(\boldsymbol{X}, \|\cdot\|)$ be a normed space. We say that $\boldsymbol{X}$ is* **separable** *if there exists a countable, dense subset of $\boldsymbol{X}$.*

Theorem 2.36 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be an infinite-dimensional separable Hilbert space. Then, there exists an orthonormal system $(\boldsymbol{e}_\alpha)_{\alpha\in I}$ in $\boldsymbol{H}$ with $I = \mathbb{N}$.*

Proof Let $\{\boldsymbol{w}_k \,|\, k \in \mathbb{N}\}$ be a countable dense subset of $\boldsymbol{H}$. We can construct a sequence $(\boldsymbol{y}_j)_{j\in J}$ such that:

(1) $\boldsymbol{y}_0 \neq \boldsymbol{0}$;

(2) for all $j \in \mathbb{N}$, there holds $\{\boldsymbol{w}_0, \ldots, \boldsymbol{w}_j\} \subseteq \{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j\}$;

(3) $\boldsymbol{y}_{j+1} \notin \operatorname{span}\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j\}$.

It suffices to take $\boldsymbol{y}_0 = \boldsymbol{w}_{k_0}$ with $k_0 = \min\{k \in \mathbb{N} \,|\, \boldsymbol{w}_k \neq \boldsymbol{0}\}$. If for some $j \in \mathbb{N}$ we have $\{\boldsymbol{w}_0, \ldots, \boldsymbol{w}_j\}$, we take $\boldsymbol{y}_{j+1} = \boldsymbol{x}_{k_{j+1}}$, with $k_{j+1} = \min\{k \in \mathbb{N} \,|\, k > j, \quad \boldsymbol{w}_k \notin \operatorname{span}\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j\}$. The index k_{j+1} exists because $\boldsymbol{H}$ is infinite dimensional. Next, we define

$$Y_j = \operatorname{span}\{\boldsymbol{y}_0, \ldots, y_j\}.$$

Clearly, for all $j \in \mathbb{N}$ the set Y_j is a proper subspace of Y_{j+1}. Then, we employ the Gram-Schmidt method to construct an orthonormal system $(\boldsymbol{x}_j)_{j\in\mathbb{N}}$ such that for any j

$$\operatorname{span}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_j\} = \operatorname{span}\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j\}. \tag{2.51}$$

Given $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_j\}$ satisfying (2.51), we take $\boldsymbol{u}_{j+1}$ in the form

$$\boldsymbol{u}_{j+1} = \sum_{k=1}^{j} c_k \boldsymbol{x}_k + \boldsymbol{y}_{j+1},$$

such that $\langle \boldsymbol{u}_{k+1}, \boldsymbol{x}_k \rangle = 0$ for all $k \in \{0, \ldots, j\}$. This can be achieved by taking $c_k = -\langle \boldsymbol{y}_{j+1}, \boldsymbol{x}_k \rangle$, so that

$$\boldsymbol{u}_{j+1} = -\sum_{k=1}^{j} \langle \boldsymbol{y}_{j+1}, \boldsymbol{x}_k \rangle \boldsymbol{x}_k + \boldsymbol{y}_{j+1}.$$

Observe that $\boldsymbol{u}_{j+1} \neq \boldsymbol{0}$, because $\boldsymbol{y}_{j+1} \notin \operatorname{span}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_j\} = \operatorname{span}\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j\}$. We define

$$\boldsymbol{x}_{j+1} = \frac{1}{\|\boldsymbol{u}_{j+1}\|} \boldsymbol{u}_{j+1}.$$

It is easily seen that

$$\operatorname{span}\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_j, \boldsymbol{y}_{j+1}\} = \operatorname{span}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_j, \boldsymbol{x}_{j+1}\}.$$

Clearly, if $\boldsymbol{V} = \text{span}\{\boldsymbol{x}_j \,|\, j \in \mathbb{N}\}$, then $\{\boldsymbol{w}_k \,|\, k \in \mathbb{N}\} \subseteq \boldsymbol{V}$, so that $\boldsymbol{V}$ is dense in $\boldsymbol{H}$ and $(\boldsymbol{x}_j)_{j\in\mathbb{N}}$ is maximal by Theorem 2.34.

We conclude this section showing that in every Hilbert space there exists a maximal orthonormal system. This is a consequence of a deep result of set theory, called Zorn's Lemma, which can be shown to be equivalent to the axiom of choice. We begin with some elementary definitions.

An **ordered set** is a set X, equipped with a relation $\leq$ between the elements of X, which satisfies the following properties:

(i) $\forall\, x \in X, \quad x \leq x$ **(reflexive property)**;

(ii) $\forall\, x, y \in X, \quad$ if $x \leq y$ and $y \leq x$, then $x = y$ **(anti-symmetric property)**;

(iii) $\forall\, x, y, z \in X, \quad$ if $x \leq y$ and $y \leq z$, then $x \leq z$ **(transitive property).**

If X is an ordered set and $A \subseteq X$, then A is ordered by the restriction of the relation $\leq$ to the elements of A. We say that A is totally ordered if

$$\forall\, x, y \in A, \quad x \leq y \text{ or } y \leq x.$$

A maximal element in X is an element M of X such that, for some $x \in X$, if $M \leq x$, then $M = x$. If $A \subseteq X$, an upper bound of A is an element m of X such that $a \leq m$ for all $a \in A$.

Theorem 2.37 [Zorn's Lemma]
Let X be an ordered set. Suppose that every totally ordered subset A of X has an upper bound. Then, there exists a maximal element in X.

Using Zorn's Lemma, we can now prove the following important result.

Theorem 2.38 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be an Hilbert space. Then, there exists a maximal orthonormal system in $\boldsymbol{H}$.*

Proof Let Φ be the collection of subsets $\boldsymbol{A}$ of $\boldsymbol{H}$ satisfying the following conditions:

- if $A \neq \varnothing$ and $\boldsymbol{x} \in \boldsymbol{A}$, then $\|\boldsymbol{x}\| = 1$;
- if $\boldsymbol{x}$ and $\boldsymbol{y}$ are distinct elements of $\boldsymbol{A}$, then $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$.

Hence, the elements $\boldsymbol{A}$ of Φ are the subsets of $\boldsymbol{H}$ in the form

$$\boldsymbol{A} = \{\boldsymbol{x}_i \,|\, i \in I\},$$

where $(\boldsymbol{x}_i)_{i\in\mathbb{N}}$ is an orthonormal system in $\boldsymbol{H}$. We now order Φ by inclusion: if $\boldsymbol{A}, \mathbf{B} \in \Phi$, then $\boldsymbol{A} \leq \mathbf{B}$ means that $\boldsymbol{A} \subseteq \mathbf{B}$. It is easy to see that this is an order relation in Φ. Let Ψ be a totally ordered subset of Φ. Then, if $\boldsymbol{A}, \mathbf{B} \in \Psi$, at least one of the conditions $\boldsymbol{A} \subseteq \mathbf{B}$ or $\mathbf{B} \subseteq \boldsymbol{A}$ holds. We check that Ψ is an upper bound in Φ. In fact, we define $M = \bigcup_{\boldsymbol{A}\in\Psi} \boldsymbol{A}$. Then, $M \in \Phi$. In fact, if $\boldsymbol{x} \in M$, then $\|\boldsymbol{x}\| = 1$. Moreover, if $\boldsymbol{x}, \boldsymbol{y} \in M$ and $\boldsymbol{x} \neq \boldsymbol{y}$, there exist $\boldsymbol{A}$ and $\mathbf{B}$ in Ψ such that $\boldsymbol{x} \in \boldsymbol{A}$ and $\boldsymbol{y} \in \mathbf{B}$. Since Ψ is totally ordered, at least one of the conditions $\boldsymbol{A} \subseteq \mathbf{B}$ or $\mathbf{B} \subseteq \boldsymbol{A}$ holds. Assume, for example, that $\boldsymbol{A} \subseteq \mathbf{B}$. Then, $\boldsymbol{x}$ and $\boldsymbol{y}$ belong to $\mathbf{B}$, $\mathbf{B}$ belongs to Φ, so that $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$. We proved that M is an upper bound for Ψ. From Zorn's Lemma, there exists a maximal element in Φ. Let N be one of them. We construct an orthonormal system $(\boldsymbol{x}_i)_{i\in I}$ such that

$$N = \{\boldsymbol{x}_i \,|\, i \in I\}.$$

Let $I = N$ and define $\boldsymbol{x}_i = i$ for all $i \in I$. Then, $(\boldsymbol{x}_i)_{i\in I}$ is an orthonormal system. We only need to show that it is maximal. We define $\boldsymbol{V} = \text{span}\{\boldsymbol{x}_i \,|\, i \in I\}$. Suppose that $\overline{\boldsymbol{V}}$ is a proper subspace of $\boldsymbol{H}$. Then, we can construct an element $\boldsymbol{x}$ of $\boldsymbol{H}$ such that $\|\boldsymbol{x}\| = 1$ and $\langle \boldsymbol{x}, \boldsymbol{w}\rangle = 0$ for all $\boldsymbol{w} \in \overline{\boldsymbol{V}}$: to this end, we can repeat the argument in the second part of the proof of Theorem 2.34. Now we define $N' = N \cup \{\boldsymbol{x}\}$. Clearly, $N' \in \Phi$, and

$N \subset N'$, contradicting the fact that N is maximal. Thus, necessarily, $\overline{V} = \boldsymbol{H}$, and the conclusion follows from Theorem 2.34.

We conclude this part with the following result.

Theorem 2.39 *Let $(\boldsymbol{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Let $(\boldsymbol{X}_k)_{k\in\mathbb{N}}$ be a sequence of closed subspaces of $\boldsymbol{H}$ such that:*

- $\boldsymbol{X}_k \subseteq \boldsymbol{X}_{k+1}$ *for all* $k \in \mathbb{N}$;
- $\boldsymbol{X} = \bigcup_{k\in\mathbb{N}} \boldsymbol{X}_k$ *is dense in* $\boldsymbol{H}$.

Given $n \in \mathbb{N}$, we denote by P_n the orthogonal projection onto $\boldsymbol{X}_n$. Then, for all $\boldsymbol{x} \in \boldsymbol{H}$,

$$\lim_{k\to\infty} P_n(\boldsymbol{x}) = \boldsymbol{x}.$$

Proof Let $\varepsilon > 0$. There exists $\boldsymbol{x}_0 \in \boldsymbol{X}$ such that $\|\boldsymbol{x} - \boldsymbol{x}_0\| < \varepsilon$. Let $\boldsymbol{x}_0 \in \boldsymbol{X}_{n_0}$ with $n_0 \in \mathbb{N}$. If $k > n_0$, we obtain

$$\|\boldsymbol{x} - P_k(\boldsymbol{x})\| \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| < \varepsilon,$$

where we have used the fact that $\boldsymbol{x}_0 \in \boldsymbol{X}_k$.

Using Theorem 2.24, Show that, if A is a closed subspace in a Hilbert space $\boldsymbol{X}$, the operator $P_A : \boldsymbol{H} \longrightarrow A$ is linear.

Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ be pairwise orthogonal elements of $\boldsymbol{H}$. Show that

$$\|\boldsymbol{x}_1 + \ldots + \boldsymbol{x}_n\|^2 = \|\boldsymbol{x}_1\|^2 + \ldots + \|\boldsymbol{x}_n\|^2.$$

This formula can be considered a generalization of the Pythagorean theorem.

3 Measure theory and Lebesgue integral

A fundamental geometric problem that has been studied since antiquity is that of the measure of a solid body. In one, two, and three dimensions, this measure is usually referred to as the length, area, and volume, respectively. The pivotal idea, which dates back from Archimedes, is to measure a body by partitioning it into elementary components, moving around each component by rigid motions (translations and rotations), and then reassembling those components to form a body which has the same measure of the original one, but whose measure is easier to determine. The assumption that underpins this idea is that we are able to assign a measure to any solid body. This does not seem like an unreasonable assumption if we rely on our geometric intuition. However, as Euclidean geometry was embedded into analytic geometry, it became clear that assigning a measure to any possible subset of $\mathbb{R}^n$ (which obeys axioms that reflect our geometric perception of measure) is not as easy as it seems, and it was proved to potentially lead to paradoxes.

To avoid paradoxes, we need to abandon the idea of assigning a measure to every possible subset of $\mathbb{R}^n$, and only restrict to a class (fortunately very large) of well-behaved subsets of $\mathbb{R}^n$ for which a measure can be successfully defined.

The second part of this chapter deals with the so-called Lebesgue integration and the Lebesgue integral. In your previous courses of Calculus, you studied the so called *Riemann integral*

$$\int_I f(x)\,\mathrm{d}x,$$

defined over an interval I, and where f was assumed to be *integrable*. This notion can be generalized to the integral of functions of more than one variable using the so-called *Peano-Jordan measure theory*. The corresponding integral is still called *Riemann integral*. The drawback of the Riemann integral is that not all functions are Riemann integrable. Luckily, every piecewise continuous function happens to be Riemann integrable, and therefore for some simple applications the Riemann theory of integration appears to be satisfactory. However, in several areas of science and engineering this is not sufficient. Another important issue is that Riemann theory of integration, despite being simple, is not flexible enough to guarantee that some operations, that appear natural, hold true. For example,

$$\lim_{n\to\infty}\int_A f_n(x)\,\mathrm{d}x = \int_A f(x)\,\mathrm{d}x$$

holds only if A is measurable and bounded (for example, when A is a bounded interval), and if there is uniform convergence of the sequence $(f_n)_{n\in\mathbb{N}}$ to the function f, i.e., when

$$\lim_{n\to\infty}\sup_{x\in A}|f_n(x)-f(x)| = 0.$$

A more satisfactory theory of integration hinges on the so-called Lebesgue integral, which can handle all the Riemann integrable functions, plus functions which are not Riemann integrable.

We will present the Lebesgue measure theory in $\mathbb{R}^n$, when $n \geq 1$, using the approach introduced by Carathéodory.

3.1 Arithmetics in $[0, +\infty]$

We are now interested in studying the subset $[0, +\infty]$ of the extended real line. This set is particularly important in the context of measure theory, because the measure of a set is a non-negative quantity that can take the value $+\infty$. In $[0, +\infty]$, the total ordering is inherited from the one defined for $\overline{\mathbb{R}}$, so that if $a, b \in [0, +\infty]$, we say that $a \leq b$ (or, equivalently, $b \geq a$) if one of the following statements is true:

(i) if $a, b \in [0, +\infty)$, then $a \leq b$ in the usual sense in $\mathbb{R}$;

(ii) $b = +\infty$.

As usual, we write $a < b$ (or, equivalently, that $b > a$) if $a \leq b$ and $a \neq b$.

Given a set $A \subseteq [0, +\infty]$:

- if A is non-empty and there exists $u \in \mathbb{R}$ such that $x \leq u$ for all $x \in A$, then $\sup A \in [0, +\infty)$. In this way, the supremum defined here generalizes the standard one defined for subsets of $\mathbb{R}$;
- if A is non-empty and there does not exist $u \in \mathbb{R}$ such that $x \leq u$ for all $x \in A$, then $\sup A = +\infty$;
- if A is empty, then $\sup A = 0$. Indeed, in this case $[0, +\infty]$ is (vacuously) the set of upper bounds of A, and its minimum is 0.

Similarly, the infimum of $A \subseteq [0, +\infty]$ is always finite if A is non-empty. If $A = \varnothing$, every element of $[0, +\infty]$ is a lower bound of A, and therefore $\inf A = +\infty$.

The operations of addition and multiplication in $[0, +\infty]$ are inherited from those defined in $\overline{\mathbb{R}}$. In particular:

- given $a, b \in [0, +\infty]$, then $a + b$ is the usual addition in $\mathbb{R}$ if $a, b \in [0, +\infty)$, and $a + b = +\infty$ if at least one between a and b is equal to $+\infty$;
- multiplication: given $a, b \in [0, +\infty]$, then $a \cdot b$ is the usual multiplication if $a, b \in [0, +\infty)$; if $a \in [0, +\infty]$ and $a \neq 0$ we have
$$a \cdot (+\infty) = (+\infty) \cdot a = +\infty$$
and, finally, we have
$$0 \cdot (+\infty) = (+\infty) \cdot 0 = 0.$$

We saw that the addition is not associative in $\overline{\mathbb{R}}$. However, it is associative in $[0, +\infty]$. Hence, addition and multiplication are still commutative and associative, and the multiplication is still distributive over the addition. Moreover, the order relation $a \leq b$ is preserved under addition or multiplication of both sides by the same quantity. However, as for $\overline{\mathbb{R}}$, the cancellation laws do not apply if some of the variables involved are infinite. For example, from $a + (+\infty) = b + (+\infty)$ or from $a \cdot (+\infty) = b \cdot (+\infty)$ we cannot deduce $a = b$. The use of cancellation laws is possible only when all the quantities involved are finite (and in the case of the multiplication, we obviously also need the quantity being cancelled to be non-zero).

Exercises

Exercise 3.1

Show that the ordering $\leq$ in $[0, +\infty]$ is a total order. In other words, prove that for all $a, b, c \in [0, +\infty]$, the following properties hold:

(i) $a \leq a$ (REFLEXIVE PROPERTY)

(ii) $a \leq b$ and $b \leq c$ imply $a \leq c$ (TRANSITIVE PROPERTY)

(iii) $a \leq b$ and $b \leq a$ imply $a = b$ (ANTI-SYMMETRIC PROPERTY)

(iv) $a \leq b$ or $b \leq a$ (TOTAL ORDERING).

[Solution: (3.1)]

Exercise 3.2 Prove the commutative and the associative property of the addition in $[0, +\infty]$. [Solution: (3.2)]

Exercise 3.3 Prove the commutative and the associative property of the multiplication in $[0, +\infty]$. [Solution: (3.3)]

3.1.1 Sequences in $[0, +\infty]$

We now consider the converges of sequences of elements of $[0, +\infty]$, which still follows from Definition 1.26.

Not all the properties that hold for standard sequences continue to hold for sequences $(x_k)_{k\in\mathbb{N}}$ with values in $[0, +\infty]$. The first property that we present concerns the interaction between Definition 1.26 (but considering sequences of elements in the set $[0, +\infty]$) and the operation of sum, and is a property that holds exactly as in the case of sequences in $[0, +\infty)$.

Theorem 3.1 *Let $\hat{x}, \hat{y} \in [0, +\infty]$. Let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be sequences of elements in the set $[0, +\infty]$ which converge to $\hat{x}$ and $\hat{y}$ respectively. Then,*

$$\lim_{k\to\infty} (x_k + y_k) = \hat{x} + \hat{y}. \tag{3.1}$$

Proof If $\hat{x}$ and $\hat{y}$ are $\neq +\infty$, the statement can be proved in the same way as in the case of real-valued sequences. Let us now assume $\hat{x} = +\infty$. Let $M \in [0, +\infty)$. By assumption, there exists $p \in \mathbb{N}$ such that $x_k > M$ for all $k > p$, and so $x_k + y_k \geq x_k > M$. We deduce that the left hand-side of (3.1) is equal to $+\infty$, and for any $\hat{y} \in [0, +\infty]$ the right hand-side $\hat{x} + \hat{y} = +\infty + \hat{y}$ is equal to $+\infty$, as required.

A similar property does not hold true, in general, for the multiplication. Indeed, if $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ are both sequences of elements in the set $[0, +\infty]$ such that

$$\hat{x} = \lim_{k\to\infty} x_k \quad \text{and} \quad \hat{y} = \lim_{k\to\infty} y_k,$$

it is **not** true, in general, that

$$\lim_{k\to\infty} (x_k\, y_k) = \hat{x}\,\hat{y}.$$

For example, if $x_k = k$ and $y = \frac{1}{k}$ for every $k \in \mathbb{N}^*$, then

$$\lim_{k\to\infty} x_k = +\infty, \quad \lim_{k\to\infty} y_k = 0,$$

but

$$\lim_{k\to\infty} (x_k\, y_k) = 1,$$

and clearly $1 \neq 0 \cdot (+\infty) = 0$. However, the following result shows that $\lim_{k\to\infty} (x_k\, y_k) = \hat{x}\,\hat{y}$ holds true if both sequences are non-decreasing.

Theorem 3.2 *Let $\hat{x}, \hat{y} \in [0, +\infty]$. Let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be sequences of elements in the set $[0, +\infty]$ which converge to $\hat{x}$ and $\hat{y}$, respectively, and such that $x_k \leq x_{k+1}$ and $y_k \leq y_{k+1}$ for all $k \in \mathbb{N}$. Then,*

$$\lim_{k\to\infty} x_k\, y_k = \hat{x}\,\hat{y}.$$

Proof If $\max\{\hat{x}, \hat{y}\} < +\infty$, the statement can be proved in the same way as in the case of real-valued sequences. Let us now assume $\hat{x} = +\infty$. If $\hat{y} = 0$, then $y_k = 0$ for all $k \in \mathbb{N}$ (why? see Exercise 1.66), so that $x_k\, y_k = 0$, and therefore $x_k\, y_k$ tends to 0 as $k \to \infty$. Let us now consider the case $\hat{y} > 0$. Let us fix $0 < c < \hat{y}$. There exists k_1 such that $y_k > c$ for every $k > k_1$ (if such value k_1 did not exist, choosing $\varepsilon \in (0, c)$ we would not be able to find a value $p \in \mathbb{N}$ such that for all $k > p$ we have $|y_k - \hat{y}| < \varepsilon$). Let $M \in [0, +\infty)$. Moreover, since $\hat{x} = +\infty$,

there exists $k_2 \in \mathbb{N}$ such that $x_k > \frac{M}{c}$ for $k > k_2$. It follows that $x_k\, y_k > \frac{M}{c}\, c = M$ for $k > \max\{k_1, k_2\}$. Hence,

$$\lim_{k \to \infty} x_k\, y_k = +\infty = \hat{x}\, \hat{y}.$$

All the other cases can be brought back to these ones.

The result of Theorem 3.2 can be expressed by saying that the multiplication in $[0, +\infty]$ is *upper continuous*, in the sense that if $x_k \in [0, +\infty]$ increases to $\hat{x} \in [0, +\infty]$ and $y_k \in [0, +\infty]$ increases to $\hat{y} \in [0, +\infty]$, then $x_k\, y_k$ increases to $\hat{x}\, \hat{y}$, but it is not *downward continuous*. This asymmetry is in line with the definition that we will give of integration "from below" rather than "from above". This will lead to some fundamental results, like the monotone convergence theorem, to be applicable only to monotonic non-decreasing functions but not necessarily to monotonic non-increasing functions.

3.1.2 Finite and infinite summations in $[0, +\infty]$

We can now introduce summations of elements in $[0, +\infty]$. Let $\mathcal{I} = \{i_1, \ldots, i_n\}$ be a non-empty finite set of elements in $\mathbb{N}$, and, for all $i \in \mathcal{I}$, let $x_i \in [0, +\infty]$. We can define, by recursion on n, the finite sum

$$\sum_{i \in \mathcal{I}} x_i.$$

The commutative and associative properties of the sum ensure that the result is independent from the order of the addends. We define

$$\sum_{i \in \emptyset} x_i = 0.$$

Before we generalize this notion by allowing $\mathcal{I}$ to be an arbitrary (i.e., possibly infinite) set and $x_i \in [0, +\infty]$ for every $i \in \mathcal{I}$, we define a series of elements of the sequence $(x_k)_{k \in \mathbb{N}}$ with terms in $[0, +\infty]$, i.e.,

$$\sum_{k=0}^{\infty} x_i.$$

We recall that if $(x_k)_{k \in \mathbb{N}}$ is a sequence of values in $[0, +\infty)$, the sequence of partial sums $(s_k)_{k \in \mathbb{N}}$ defined by

$$s_k = x_0 + x_1 + \ldots + x_k$$

is monotonically non-decreasing, and therefore it always has limit in $[0, +\infty]$. The limit $\lim_{k \to \infty} s_k$ coincides with $\sup_{k \in \mathbb{N}} s_k$ (where we use the convention earlier, which considers $\sup_{k \in \mathbb{N}} s_k$ to be finite if $\{s_k \,|\, k \in \mathbb{N}\}$ is bounded from above, and $\sup_{k \in \mathbb{N}} s_k = +\infty$ if $\{s_k \,|\, k \in \mathbb{N}\}$ is not bounded from above).

We generalize this result to the case of sequences of terms in $[0, +\infty]$. In this case, we can still define the sequence $(s_k)_{k \in \mathbb{N}}$ of the partial sums, which is still monotonically non-decreasing: if i_0 is the smallest index for which $x_{i_0} = +\infty$, then $(s_k)_{k \in \mathbb{N}}$ will be monotonically non-decreasing and finite for all $k < i_0$, and $s_k = +\infty$ for all $k \geq i_0$. As in the standard case, we wish to define, for a sequence $(x_k)_{k \in \mathbb{N}}$ with values in $[0, +\infty]$, the sum of the series

$$\sum_{k=0}^{\infty} x_k \stackrel{\text{def}}{=} \lim_{k \to \infty} s_k. \tag{3.2}$$

The following theorem shows that, exactly as in the standard case, this limit always exists in $[0, +\infty]$.

Theorem 3.3 *Let $x_i \in [0,+\infty]$ for all $i \in \mathbb{N}$. For all $k \in \mathbb{N}$, let*

$$s_k = \sum_{i=0}^{k} x_i.$$

Then, the limit $\lim_{k\to\infty} s_k$ exists in $[0,+\infty]$, and it is equal to $\sup\{s_k \,|\, k \in \mathbb{N}\}$.

Proof We consider the following three cases:

(1) there exists $i_0 \in \mathbb{N}$ such that $x_{i_0} = +\infty$. It follows that for $k \geq i_0$ we have $s_k = +\infty$, so that $\lim_{k\to\infty} s_k = +\infty = \max\{s_k \,|\, k \in \mathbb{N}\}$;

(2) $x_i \neq +\infty$ for all $i \in \mathbb{N}$ and the sequence $(s_k)_{k\in\mathbb{N}}$ is bounded from above in $\mathbb{R}$. From the corresponding result of series with non-negative real terms we know that $(s_k)_{k\in\mathbb{N}}$ is convergent in $\mathbb{R}$ to $\sup\{s_k \,|\, k \in \mathbb{N}\}$. From Definition 1.26, this is true also in $[0,+\infty]$;

(3) $x_i \neq +\infty$ for all $i \in \mathbb{N}$ and the sequence $(s_k)_{k\in\mathbb{N}}$ is not bounded from above in $\mathbb{R}$. We can use, again, the theory of series with non-negative real terms, which guarantees that $\lim_{k\to\infty} s_k = +\infty = \sup\{s_k \,|\, k \in \mathbb{N}\}$.

Defining $\sum_{k=0}^{\infty} x_k$ as in (3.2), we find that the sum of the series is an element of $[0,+\infty]$.

We can now give the fundamental definition of generalized sum.

Definition 3.1 *Let $\mathcal{I}$ be a set. Let $x_i \in [0,+\infty]$ for all $i \in \mathcal{I}$. We define*

$$\sum_{i\in\mathcal{I}} x_i \overset{def}{=} \sup\Big\{ \sum_{i\in F} x_i \;\Big|\; F \subseteq \mathcal{I} \text{ and } F \text{ is finite} \Big\}. \tag{3.3}$$

If $\mathcal{I}$ is finite, it is straightforward to verify that Definition 3.1 coincides with the notion of sum over a non-empty finite index set (in this case, the supremum in (3.3) is a maximum, and it is attained taking $F = \mathcal{I}$).

We now relate Definition 3.1 with the notion of sum of a series.

Theorem 3.4 *Let $x_i \in [0,+\infty]$ for all $i \in \mathbb{N}$. Then,*

$$\sum_{i\in\mathbb{N}} x_i = \sum_{i=0}^{\infty} x_i.$$

Proof Let $F \subset \mathbb{N}$ be finite. Let $k \in \mathbb{N}$ be such that $i \leq k$ for every $i \in F$. Then, from Theorem 3.3 we find

$$\sum_{i\in F} x_i \leq \sum_{i=0}^{k} x_i \leq \sum_{i=0}^{\infty} x_i.$$

From the arbitrariness of F, we find that $\sum_{i=0}^{\infty} x_i$ is an upper bound of the set $\Big\{ \sum_{i\in F} x_i \;\Big|\; F \subseteq \mathbb{N} \text{ and } F \text{ is finite} \Big\}$. On the other hand, $\sum_{i\in\mathbb{N}} x_i$ is, by Definition 3.1, the least upper bound of $\Big\{ \sum_{i\in F} x_i \;\Big|\; F \subseteq \mathbb{N} \text{ and } F \text{ is finite} \Big\}$, and

therefore

$$\sum_{i\in\mathbb{N}} x_i \le \sum_{i=0}^{\infty} x_i.$$

We now prove the opposite inequality. Let us consider $F = \{0, 1, \dots, k\}$, where $k \in \mathbb{N}$ is arbitrary. Hence, $\sum_{i\in\{0,1,\dots,k\}} x_i$ is an element of $\Big\{ \sum_{i\in F} x_i \;\Big|\; F \subseteq \mathbb{N} \text{ and } F \text{ is finite} \Big\}$. On the other hand, $\sum_{i\in\{0,1,\dots,k\}} x_i = s_k$, and therefore for all $k \in \mathbb{N}$, we have $s_k \le \sum_{i\in\mathbb{N}} x_i$. Taking the limit in both sides yields

$$\sum_{i=0}^{\infty} x_i \le \sum_{i\in\mathbb{N}} x_i.$$

We now introduce some useful properties of generalized sums.

Theorem 3.5 *Let $\mathcal{I}$ and $\mathcal{J}$ be arbitrary sets, and let $x_{i,j} \in [0, +\infty]$ for all $(i, j) \in \mathcal{I} \times \mathcal{J}$. Then,*

$$\sum_{i\in\mathcal{I}} \Big(\sum_{j\in\mathcal{J}} x_{i,j} \Big) = \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} = \sum_{j\in\mathcal{J}} \Big(\sum_{i\in\mathcal{I}} x_{i,j} \Big).$$

Proof We prove the first identity. We start showing that

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} \le \sum_{i\in\mathcal{I}} \Big(\sum_{j\in\mathcal{J}} x_{i,j} \Big).$$

Let $F \subseteq \mathcal{I} \times \mathcal{J}$ be finite. Let G and H be finite subsets of $\mathcal{I}$ and $\mathcal{J}$, respectively, and such that $F \subseteq G \times H$. These sets always exist: in fact, it is always possible to choose for example $G = \pi_{\mathcal{I}}(F)$ and $H = \pi_{\mathcal{J}}(F)$, where $\pi_{\mathcal{I}}$ and $\pi_{\mathcal{J}}$ represent the natural projections of $\mathcal{I} \times \mathcal{J}$ onto $\mathcal{I}$ and $\mathcal{J}$, respectively (so that $\pi_{\mathcal{I}}(F) = \{i \,|\, (i, j) \in F\}$ and $\pi_{\mathcal{J}}(F) = \{j \,|\, (i, j) \in F\}$). Then,

$$\sum_{(i,j)\in F} x_{i,j} \le \sum_{i\in G} \Big(\sum_{j\in H} x_{i,j} \Big) \le \sum_{i\in G} \Big(\sum_{j\in J} x_{i,j} \Big) \le \sum_{i\in\mathcal{I}} \Big(\sum_{j\in\mathcal{J}} x_{i,j} \Big).$$

We find that $\sum_{i\in\mathcal{I}} \Big(\sum_{j\in\mathcal{J}} x_{i,j} \Big)$ is an upper bound for the set $\Big\{ \sum_{(i,j)\in F} x_{i,j} \;\Big|\; F \subseteq \mathcal{I} \times \mathcal{J} \text{ and } F \text{ is finite} \Big\}$ since F is arbitrary. Since $\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j}$ is the least upper bound of the same set, we obtain

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} \le \sum_{i\in\mathcal{I}} \Big(\sum_{j\in\mathcal{J}} x_{i,j} \Big).$$

We now need to show the opposite inequality. This inequality trivially holds if $\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} = +\infty$. Suppose $\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} < +\infty$. Then, $\sum_{j\in\mathcal{J}} x_{i,j} < +\infty$ for all $i \in \mathcal{I}$. Indeed, let $H \subseteq \mathcal{J}$ be finite, and let $i \in \mathcal{I}$. Let $F = \{(i, j) \,|\, j \in H\}$. This set F is finite, and therefore

$$\sum_{j\in H} x_{i,j} = \sum_{(i,j)\in F} x_{i,j} \le \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j}.$$

Hence, $\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j}$ is an upper bound for $\Big\{ \sum_{j\in S} x_{i,j} \;\Big|\; S \subseteq \mathcal{J} \text{ and } S \text{ is finite} \Big\}$ in view of the arbitrariness of H,

and $\sum_{j\in\mathcal{J}} x_{i,j}$ is the supremum of the same set, so that $\sum_{j\in\mathcal{J}} x_{i,j} \leq \sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} < +\infty$. Let $G \subseteq \mathcal{I}$ be finite, and let $\varepsilon > 0$. Let n denote the cardinality of G. For every $i \in G$, from the property of the supremum[a] there exists a finite subset H_i of $\mathcal{J}$ such that

$$\sum_{j\in H_i} x_{i,j} \geq \sum_{j\in\mathcal{J}} x_{i,j} - \frac{\varepsilon}{n}.$$

Let $F = \{(i,j) \,|\, i \in G \text{ and } j \in H_i\}$. Then, F is a finite subset of $\mathcal{I}\times\mathcal{J}$. It follows that

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} \geq \sum_{(i,j)\in F} x_{i,j} = \sum_{i\in G}\Big(\sum_{j\in H_i} x_{i,j}\Big) \geq \sum_{i\in G}\Big(\sum_{j\in\mathcal{J}} x_{i,j}\Big) - \varepsilon.$$

Since ε is arbitrary, we obtain

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} \geq \sum_{i\in G}\Big(\sum_{j\in\mathcal{J}} x_{i,j}\Big)$$

for every finite $G \subseteq \mathcal{I}$. Thus, $\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j}$ is an upper bound for the set $\Big\{\sum_{i\in S}\Big(\sum_{j\in\mathcal{J}} x_{i,j}\Big) \;\Big|\; S \subseteq \mathcal{I} \text{ and } S \text{ is finite}\Big\}$, and the supremum of this set is $\sum_{i\in\mathcal{I}}\Big(\sum_{j\in\mathcal{J}} x_{i,j}\Big)$, and therefore we conclude that

$$\sum_{(i,j)\in\mathcal{I}\times\mathcal{J}} x_{i,j} \geq \sum_{i\in\mathcal{I}}\Big(\sum_{j\in\mathcal{J}} x_{i,j}\Big),$$

as required. The second identity can be obtained in a similar way.

[a] If $\Omega \subseteq \mathbb{R}$ and $\alpha \in \mathbb{R}$, then $\alpha = \sup\Omega$ if and only if

(i) $\forall\, x \in \Omega,\ \ \alpha \geq x$;

(ii) $\forall\, \varepsilon > 0,\ \exists\, a \in \Omega : a > \alpha - \varepsilon$.

In this case, $\Omega = \Big\{\sum_{j\in S} x_{i,j} \;\Big|\; S \subseteq \mathcal{J} \text{ and } S \text{ is finite}\Big\}$.

Theorem 3.6 *Let $\mathcal{I}$ and $\mathcal{J}$ be non-empty sets. Let $\varphi : \mathcal{I} \longrightarrow \mathcal{J}$ be a bijection. For every $j \in \mathcal{J}$, let $x_j \in [0, +\infty]$. Then,*

$$\sum_{j\in\mathcal{J}} x_j = \sum_{i\in\mathcal{I}} x_{\varphi(i)}.$$

Proof Let $F \subseteq \mathcal{J}$ be finite. Since φ is bijective, $\varphi^{-1}(F)$ is a finite subset of $\mathcal{I}$. Then,

$$\sum_{j\in F} x_j = \sum_{i\in\varphi^{-1}(F)} x_{\varphi(i)} \leq \sum_{i\in\mathcal{I}} x_{\varphi(i)}.$$

We can resort to the usual argument: $\sum_{i\in\mathcal{I}} x_{\varphi(i)}$ is an upper bound for the set $\Big\{\sum_{j\in F} x_j \;\Big|\; F \subseteq \mathcal{J} \text{ and } F \text{ is finite}\Big\}$ since F is arbitrary, and the supremum of this set is $\sum_{j\in\mathcal{J}} x_j$, so that

$$\sum_{j\in\mathcal{J}} x_j \leq \sum_{i\in\mathcal{I}} x_{\varphi(i)}.$$

Conversely, let $G \subseteq \mathcal{I}$ be finite. Then, $\varphi(G)$ is a finite subset of $\mathcal{J}$. It follows that

$$\sum_{i \in G} x_{\varphi(i)} = \sum_{j \in \varphi(G)} x_j \leq \sum_{j \in \mathcal{J}} x_j.$$

Hence,

$$\sum_{i \in \mathcal{I}} x_{\varphi(i)} \leq \sum_{j \in \mathcal{J}} x_j.$$

Theorem 3.7 *Let $\varphi : \mathbb{N} \longrightarrow \mathbb{N}$ be a bijection. For every $i \in \mathbb{N}$, let $x_i \in [0, +\infty]$. Then,*

$$\sum_{i=0}^{\infty} x_{\varphi(i)} = \sum_{i=0}^{\infty} x_i.$$

Proof Applying Theorems 3.4 and 3.6 we obtain

$$\sum_{i=0}^{\infty} x_{\varphi(i)} = \sum_{i \in \mathbb{N}} x_{\varphi(i)} = \sum_{i \in \mathbb{N}} x_i = \sum_{i=0}^{\infty} x_i.$$

Remark 3.1 In general, the result of Theorem 3.7 is false for a convergent series $\sum_{i=0}^{\infty} x_i$ with real terms. An example is the series

$$\sum_{k=3}^{\infty} (-1)^{k+1} \frac{1}{k},$$

which, informally, can be written as the infinite sum

$$\frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots$$

We note that the quantities $\frac{1}{3} - \frac{1}{4}$, $\frac{1}{5} - \frac{1}{6}$, $\frac{1}{7} - \frac{1}{8}$, etc, are all positive. More in general, for any $k \geq 1$ we have

$$\frac{1}{2k+1} - \frac{1}{2k+2} = \frac{1}{(2k+1)(2k+2)} > 0,$$

and therefore we can write the original series as

$$\sum_{k=1}^{\infty} \frac{1}{(2k+1)(2k+2)},$$

and there is no doubt that this series is convergent to a positive quantity. One can even show that the sum of the series is $\ln 2 - \frac{1}{2} > 0$. However, if we rearrange the terms of the series so that two negative addends follow each positive one, we obtain

$$\left(\frac{1}{3} - \frac{1}{4} - \frac{1}{6}\right) + \left(\frac{1}{5} - \frac{1}{8} - \frac{1}{10}\right) + \left(\frac{1}{7} - \frac{1}{12} - \frac{1}{14}\right) + \dots$$

and each term within parenthesis is negative. Indeed, each of these terms is equal, for some $k \geq 1$, to the generic term

$$\frac{1}{2k+1} - \frac{1}{4k} - \frac{1}{4k+2} = -\frac{1}{2k(4k+2)},$$

and therefore the original series can be written as

$$-\sum_{k=1}^{\infty} \frac{1}{2\,k\,(4\,k+2)},$$

which converges to a negative number (which can be shown to be equal to $\frac{1}{2}(\ln 2 - 1)$).

This behavior is brought to its extreme consequences with a result (due to Riemann and Dini) that states that the terms of a series that is convergent (like this one, in view of the Leibniz test for alternating series) but not absolutely convergent (like this one, because we know that the harmonic series is not convergent) can be rearranged so that the resulting series can be divergent, or oscillating, or convergent to a different sum of the series.

Theorem 3.8 [Fubini's Theorem for Series]
Let $x_{i,j} \in [0,+\infty]$ for every $i, j \in [0,+\infty]$. Then,

$$\sum_{i=0}^{\infty}\Big(\sum_{j=0}^{\infty} x_{i,j}\Big) = \sum_{j=0}^{\infty}\Big(\sum_{i=0}^{\infty} x_{i,j}\Big).$$

Proof The proof follows immediately from Theorems 3.4 and 3.5.

Exercises

Exercise 3.4 Let $p \in \mathbb{N}^*$. Let $\alpha_k, \beta_k \in [0,+\infty]$ for all $k \in \{1,\dots,p\}$. Using induction on p, show that

$$\sum_{k=1}^{p}(\alpha_k + \beta_k) = \sum_{k=1}^{p}\alpha_k + \sum_{k=1}^{p}\beta_k. \tag{3.4}$$

[Solution: (3.4)]

Exercise 3.5 Let $p, q \in \mathbb{N}^*$. Let $\alpha_{k,\ell} \in [0,+\infty]$ for all $k \in \{1,\dots,p\}$ and $\ell \in \{1,\dots,q\}$. Using induction on q, show that

$$\sum_{k=1}^{p}\left(\sum_{\ell=1}^{q} a_{k,\ell}\right) = \sum_{\ell=1}^{q}\left(\sum_{k=1}^{p} a_{k,\ell}\right). \tag{3.5}$$

[Solution: (3.5)]

Exercise 3.6 Let A, B be sets. Prove that

$$A \setminus B = A \cap B^{\complement}$$

[Solution: (3.6)]

3.2 General considerations about measure

As described in the introduction to this chapter, the main idea that underpins the measure of a solid body is to partition it into elementary components, and then by means of rigid motions reassemble those components into a body that is easy to measure. A classical example is the idea to find the area of a parallelogram by decomposing it

into a triangle and a trapezium, and moving the triangle to form a rectangle:

One could however, take this process to its extreme consequences: in fact, one can view any solid object as the union of an infinite and uncountable number of points, and there is no doubt that the measure of a point is 0. This means that if we view any subset $\boldsymbol{E}$ of $\mathbb{R}^n$ as the union of its points, we end up with the product $\infty \cdot 0$ as its measure. A related consideration, which shows how disassembling and reassembling the object is a very delicate operation in general, is the following: the two intervals $I = [0, 1]$ and $J = [0, 2]$ of $\mathbb{R}$ can be considered to have the same number of points, because between these sets there is a bijection

$$\begin{aligned} \varphi : I &\longrightarrow J \\ x &\mapsto 2\,x. \end{aligned}$$

However, our intuition says (and correctly so) that J is twice as long as I. This leads to the disturbing conclusion that we can disassemble I into an infinite number of points, and then reassemble them to find an interval of twice the length.

The issue seems to be related with the notion of uncountably infinity of points that characterize the disassembly. However, even restricting our attention to countable, or even finite partitions of subsets of $\mathbb{R}^3$ leads to paradoxes. The most famous one is the so-called *Banach-Tarski paradox*: a sphere in $\mathbb{R}^3$ can be disassembled in 5 pieces, which can be reassembled to form two spheres which are identical from the initial one.

It has been proved conclusively only in recent times that the issue lies in the fact that the construction of these pathological decompositions inherently requires the *axiom of choice*.[1] In mathematical frameworks where the axiom of choice is not accepted, these phenomena do not occur. This is good news, because the pathologies that exclusively arise from the axiom of choice are, almost by definition, artificial. On the other hand, this implies also that we have to abandon the idea of being able to measure every subset of $\mathbb{R}^n$, and settle for measuring only the well-behaved subsets of $\mathbb{R}^n$ for which such pathologies cannot occur. These will be referred to as the *Lebesgue measurable sets*, or simply *measurable sets*.

It must become clear that the problem of measuring a set and defining an integral of a function are two related concepts. Indeed, in order to compute the integral

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

we need to "measure" the set $\boldsymbol{A}$: if we wish to integrate the function 1 over the set $\boldsymbol{A}$, then the integral $\int_{\boldsymbol{A}} f$ becomes precisely the measure of $\boldsymbol{A}$, which is a length if $\boldsymbol{A}$ is 1-dimensional, an area if $\boldsymbol{A}$ is 2-dimensional, or a volume if $\boldsymbol{A}$ is 3-dimensional. We use the general term "measure" to refer to lengths, areas and volumes, and their generalizations to higher dimensions.

[1] The **axiom of choice** postulates that, given a family of non-empty sets A_t, we are able to choose a single element a_t from each one, so as to form a set $\{a_t \mid \forall\, t\}$ from all the choices we made. There is a controversial aspect in this axiom: the choices are arbitrary and infinite (if the set of non-empty sets A_t is infinite), and they are made without a rule that shows how the choices have to be made. A large portion of mathematics can be constructed without using the axiom of choice. However, implicitly or explicitly, you have seen results that are valid only once one accepts the axiom of choice: an example is the theorem that says that every vector space has a basis. This result that you saw in Linear Algebra 2 (without proof) hinges on Zorn's Lemma, which in turn can be shown to be equivalent to the axiom of choice. Despite often leading to counterintuitive results, the axiom of choice is, nowadays, generally accepted by mathematicians. An interesting aspect of this axiom is the fact that it is independent (or, to use a more accurate term is *undecidable*) from the other axioms of set theory, in the sense that it can neither be proved nor disproved by the other axioms of set theory.

We also need to point out that the Lebesgue measure is not the only measure theory that can be introduced, and that leads to a notion of integral. As already mentioned, another important measure is the *Peano-Jordan measure*, which leads to the so-called *Riemann integral* (the one which, implicitly or explicitly, you have seen when dealing with integrals of functions of a single variable). However, the class of sets which are Peano-Jordan measurable is smaller than the class of Lebesgue measurable sets: there are sets which are Lebesgue measurable, but they are not Peano-Jordan measurable. Moreover, the definition of integral that arises from the Lebesgue measure behaves better than the Riemann integral under limiting processes.

Given a subset $\boldsymbol{E} \subseteq \mathbb{R}^n$, our aim is to associate a number $\mu(\boldsymbol{E})$, that we call *the measure of* $\boldsymbol{E}$. Ideally, this measure should be chosen in such a way that, given any $\boldsymbol{E} \subseteq \mathbb{R}^n$:

(P1) $\mu(\emptyset) = 0$;

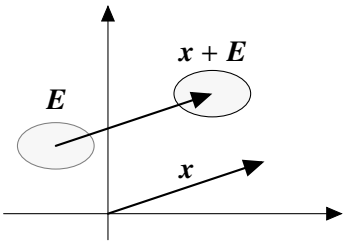
(P2) $\mu(\boldsymbol{E}) \in [0, +\infty]$;

(P3) $\mu(\boldsymbol{E} \cup \boldsymbol{F}) = \mu(\boldsymbol{E}) + \mu(\boldsymbol{F})$ for any two disjoint[2] subsets $\boldsymbol{E}$ and $\boldsymbol{F}$ of $\mathbb{R}^n$; if the sets $\boldsymbol{E}$ and $\boldsymbol{F}$ are not disjoint, we should have $\mu(\boldsymbol{E} \cup \boldsymbol{F}) \leq \mu(\boldsymbol{E}) + \mu(\boldsymbol{F})$;

(P4) $\boldsymbol{E} \subseteq \boldsymbol{F}$ implies $\mu(\boldsymbol{E}) \leq \mu(\boldsymbol{F})$;

(P5) the measure of the n-dimensional hypercube $\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,|\, \forall\, i \in \{1, 2, \ldots, n\}, \quad 0 \leq x_i \leq 1\}$ is equal to 1. This property is called **normalization**;

(P6) if we shift a set $\boldsymbol{E}$ by any vector $\boldsymbol{x}$, the measure should not change, i.e., for all $\boldsymbol{x} \in \mathbb{R}^n$, we should have $\mu(\boldsymbol{x} + \boldsymbol{E}) = \mu(\boldsymbol{E})$. This property is called **translation invariance**. We recall that the translation of the set $\boldsymbol{E}$ by the vector $\boldsymbol{x}$ is the set $\boldsymbol{x} + \boldsymbol{E} = \{\boldsymbol{x} + \boldsymbol{a} \,|\, \boldsymbol{a} \in \boldsymbol{E}\}$.[3]



Unfortunately, it can be proven that a measure function defined for all subsets of $\mathbb{R}^n$, i.e.

$$\mu : \mathcal{P}(\mathbb{R}^n) \longrightarrow [0, +\infty],$$

where $\mathcal{P}(\mathbb{R}^n)$ denotes the power set[4] of $\mathbb{R}^n$, satisfying all these properties does not exist.[5]

Hence, we have to abandon the ambition of assigning a measure to every subset of $\mathbb{R}^n$ in such a way that all those properties (which we perceive as natural for a notion of measure) hold. The problem can be solved by identifying

[2] We recall that two sets are disjoint if they do not have common elements, i.e., if $A \cap B = \emptyset$.

[3] An example of this idea were the affine subspaces of Linear Algebra: given a vector space $\boldsymbol{V}$ and a subspace $\boldsymbol{W}$, we denoted by $\boldsymbol{v} + \boldsymbol{W}$ the set $\{\boldsymbol{v} + \boldsymbol{w} \,|\, \boldsymbol{w} \in \boldsymbol{W}\}$.

[4] The power-set of a set A is the set of all possible subsets of A. It is also denoted by 2^A, because if A is finite and contains n elements, the power-set $\mathcal{P}(A)$ contains 2^n elements (including A and $\emptyset$, which are the trivial subsets of A).

[5] Hausdorff in 1914 showed that for any dimension $n \geq 1$, there is no countably additive measure defined on all subsets of $\mathbb{R}^n$ that is invariant under isometries (translations and rotations) and assigns measure one to the unit hypercube. Even more astonishingly, as already mentioned, Banach and Tarski in 1924 showed that if $n \geq 3$, one can cut up the unit ball $\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,|\, x_1^2 + x_2^2 + \ldots + x_n^2 \leq 1\}$ in $\mathbb{R}^n$ into a finite number of pieces and use isometries (translations and rotations) to reassemble the pieces into a ball of any desired volume: this is the so-called "Banach-Tarski paradox". The construction of these pieces requires the axiom of choice. It was shown in 1970 by Solovay that the the axiom of choice is indispensable to obtain non-Lebesgue measurable sets.

a certain class of subsets of $\mathbb{R}^n$ where these properties are satisfied: this leads to the notion of measurable set. In other words, we will have to define our measure as a function

$$\mu : \mathcal{M}_n \longrightarrow [0, +\infty],$$

where $\mathcal{M}_n$ will denote the set of (Lebesgue) measurable subsets of $\mathbb{R}^n$. Fortunately, the very vast majority of the sets that one has to deal with in mathematics, statistics, physics and science/engineering in general, are Lebesgue measurable.

Our objective is to define a notion of measurable set in such a way that $\mathcal{M}_n$ is a σ-algebra in $\mathbb{R}^n$, which, loosely speaking, means that $\mathcal{M}_n$ is closed under the fundamental set-theoretic operations.

The characterisation of measurable sets that arises from the Lebesgue measure ensures that an additional property is satisfied, in addition to the σ-algebra properties:

- Borel Property: every open subset and every closed subset of $\mathbb{R}^n$ is measurable.

These properties ensure that every set one typically encounters in practice is Lebesgue measurable. However, as aforementioned, it is impossible to exclude the presence of non-measurable sets.

We will see that, if $\boldsymbol{E}$ is measurable, we can define its *Lebesgue measure* $\mu(\boldsymbol{E})$, which satisfies the properties **(P1)-(P6)**. In particular, **(P3)** is satisfied not just for finite unions of measurable sets, but also for countable unions of measurable sets. More precisely, if $(\boldsymbol{E}_j)_{j\in J}$ is an at most countable sequence of **pairwise disjoint**[6] measurable sets, then

$$\mu\Big(\bigcup_{j\in J} \boldsymbol{E}_j\Big) = \sum_{j\in J} \mu(\boldsymbol{E}_j);$$

This property is called **countable additivity**. If the sets of the sequence $(\boldsymbol{E}_j)_{j\in J}$ are not pairwise disjoint, the measure of their union is not greater than the sum of their measures:

$$\mu\Big(\bigcup_{j\in J} \boldsymbol{E}_j\Big) \leq \sum_{j\in J} \mu(\boldsymbol{E}_j);$$

This property is called **countable subadditivity**. As already mentioned, requiring uncountable additivity would lead to the unwanted conclusion that the measure of every subset of $\mathbb{R}^n$ is zero.

There are many alternative ways to construct the Lebesgue measure. Here, we follow the most common approach due to Carathéodory: this method consists in constructing an outer measure on *all* subsets of $\mathbb{R}^n$ by approximating them from the outside by countable unions of "boxes". In doing so, we will observe that the countable additivity property is not satisfied for every subset of $\mathbb{R}^n$. We identify a notion of Lebesgue measurable set that characterizes those subsets of $\mathbb{R}^n$ for which the outer measure is countably additive. We then restrict the outer measure to the subsets of $\mathbb{R}^n$ on which the outer measure is countably additive.

3.3 σ-algebras and measure spaces

As already mentioned, the set of Lebesgue measurable sets that will be introduced in the sequel constitute a so-called σ-algebra. We also want to mention that the importance of the notion of σ-algebra is not limited to the Lebesgue measure; for example, σ-algebras play a very important role in the theory of probability. A σ-algebra is formally defined as follows.

Definition 3.2 [σ-ALGEBRA]
Let X be a set and let $\mathcal{A} \subseteq \mathcal{P}(X)$. We say that $\mathcal{A}$ is a σ-algebra (in X) if it satisfies the following conditions:

(A) $\varnothing \in \mathcal{A}$;

(B) *$A \in \mathcal{A}$ implies $A^{\complement} \in \mathcal{A}$;*

[6] A sequence of sets $(\boldsymbol{E}_j)_{j\in J}$ is said to be pairwise disjoint if $\boldsymbol{E}_i \cap \boldsymbol{E}_j = \varnothing$ whenever $i \neq j$.

(C) *if* $(A_k)_{k\in\mathbb{N}}$ *is a sequence of subsets of* $\mathcal{A}$*, then* $\bigcup_{k=0}^{\infty} A_k \in \mathcal{A}$.

Observe that, for every set X, the sets $\mathcal{P}(X)$ and $\{\varnothing, X\}$ are σ-algebras (why?).

Theorem 3.9 *Let X be a set and let $\mathcal{A}$ be a σ-algebra in X. Then:*

(a) $X \in \mathcal{A}$;

(b) *for every* $m \in \mathbb{N}$*, if* $A_k \in \mathcal{A}$ *for all* $k \in \{0, \dots, m\}$*, then* $\bigcup_{k=0}^{m} A_k \in \mathcal{A}$;

(c) *if* $A_k \in \mathcal{A}$ *for all* $k \in \mathbb{N}$*, then* $\bigcap_{k=0}^{\infty} A_k \in \mathcal{A}$;

(d) *for every* $m \in \mathbb{N}$*, if* $A_k \in \mathcal{A}$ *for all* $k \in \{0, \dots, m\}$*, then* $\bigcap_{k=0}^{m} A_k \in \mathcal{A}$;

(e) *for all* $A, B \in \mathcal{A}$*, we have* $A \setminus B \in \mathcal{A}$.

Proof Statement **(a)** follows from the fact that $X = \varnothing^{\mathbb{C}}$, using **(A)** and **(B)** in Definition 3.2.

We prove **(b)**. We set $A_k = \varnothing$ for $k \in \mathbb{N}$ such that $k > m$. Then, $\bigcup_{k=0}^{m} A_k = \bigcup_{k=0}^{\infty} A_k$, so that **(b)** follows from **(C)** in Definition 3.2.

Statement **(c)** follows from De Morgan's law

$$\bigcap_{k=0}^{\infty} A_k = \left(\bigcup_{k=0}^{\infty} A_k^{\mathbb{C}}\right)^{\mathbb{C}}.$$

We leave the proof of **(d)** as an exercise (see Exercise 3.7), while **(e)** follows from the identity $A \setminus B = A \cap B^{\mathbb{C}}$ and **(d)**.

Exercise 3.7 Prove point **(d)** in the proof of Theorem 3.9. [Solution: (3.7)]

We now define a measure.

Definition 3.3 *Let X be a set and let $\mathcal{A}$ be a σ-algebra in X. Let $\mu : \mathcal{A} \longrightarrow [0, +\infty]$. We say that μ is a* measure *for $(X, \mathcal{A})$ if it satisfies the following conditions:*

(i) $\mu(\varnothing) = 0$;

(ii) if $(A_k)_{k\in\mathbb{N}}$ *is a sequence of pairwise disjoint elements of* $\mathcal{A}$*, then*

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = \sum_{k=0}^{\infty} \mu(A_k).$$

If μ is a measure for $\mathcal{A}$ and X, the triplet $(X, \mathcal{A}, \mu)$ is called measure space.

Exercise 3.8

Let X be a non-empty set. Let $x_0 \in X$. Show that the function

$$\begin{aligned} \mu : \mathcal{P}(X) &\longrightarrow [0, +\infty] \\ A &\mapsto \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is a measure. [Solution: (3.8)]

For the summations and the series in this and in the following sections, we assume to use the conventions established for $[0, +\infty]$.

Theorem 3.10 *Let X be a set and let $\mathcal{A}$ be a σ-algebra in X. Let $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ be a measure. Then:*

(a) *Let $m \in \mathbb{N}$ and let $A_0, \dots, A_m \in \mathcal{A}$ be pairwise disjoint. Then,*

$$\mu\Big(\bigcup_{k=0}^{m} A_k\Big) = \sum_{k=0}^{m} \mu(A_k);$$

(b) *for every $A, B \in \mathcal{A}$ such that $A \subseteq B$, we have $\mu(A) \leq \mu(B)$;*

(c) *for every $A, B \in \mathcal{A}$ such that $A \subseteq B$ and $\mu(B) < +\infty$, then $\mu(A) < +\infty$ and $\mu(B \setminus A) = \mu(B) - \mu(A)$;*

(d) *if $(A_k)_{k\in\mathbb{N}}$ is a sequence of (not necessarily pairwise disjoint) elements of $\mathcal{A}$, then*

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) \leq \sum_{k=0}^{\infty} \mu(A_k);$$

(e) *Let $m \in \mathbb{N}$ and let $A_0, \dots, A_m \in \mathcal{A}$. Then,*

$$\mu\Big(\bigcup_{k=0}^{m} A_k\Big) \leq \sum_{k=0}^{m} \mu(A_k).$$

Proof To prove **(a)**, we set $A_k = \varnothing$ for every $k \in \mathbb{N}$ such that $k > m$. Clearly, in this way all the sets A_k are still pairwise disjoint. Thus,

$$\mu\Big(\bigcup_{k=0}^{m} A_k\Big) = \mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = \sum_{k=0}^{\infty} \mu(A_k) = \sum_{k=0}^{m} \mu(A_k),$$

since $\mu(\varnothing) = 0$.
We prove **(b)**. From **(e)** in Theorem 3.9, we obtain $B \setminus A \in \mathcal{A}$. From **(a)**, we find

$$\mu(B) = \mu(A \cup (B \setminus A)) = \mu(A) + \mu(B \setminus A), \tag{3.6}$$

where we used the fact that $A \cap (B \setminus A) = \varnothing$. This proves **(b)**. Moreover, from **(b)** we conclude that, if $\mu(B) < +\infty$, we also have $\mu(A) < +\infty$. Thus, **(c)** is a consequence of (3.6). We now prove **(d)**. For every $k \in \mathbb{N}$, we define

$$B_k = \begin{cases} A_0 & \text{if } k = 0 \\ A_k \setminus \bigcup_{j=0}^{k-1} A_j & \text{if } k \geq 1. \end{cases}$$

From **(b)** and **(e)** of Theorem 3.9, for all $k \in \mathbb{N}$ we have $B_k \in \mathcal{A}$. Moreover,

$$\bigcup_{k=0}^{\infty} B_k = \bigcup_{k=0}^{\infty} A_k. \tag{3.7}$$

To see this, we first prove the inclusion $\bigcup_{k=0}^{\infty} B_k \subseteq \bigcup_{k=0}^{\infty} A_k$. For every $k \in \mathbb{N}$, we have $B_k \subseteq A_k$, and therefore $\bigcup_{k=0}^{\infty} B_k \subseteq \bigcup_{k=0}^{\infty} A_k$.

Let $x \in \bigcup_{k=0}^{\infty} A_k$, and let k_0 be the smallest index k for which $x \in A_k$. By definition, $x \notin A_j$ for every $j < k_0$. It follows that $x \in B_{k_0}$, so that $x \in \bigcup_{k=0}^{\infty} B_k$, which proves that $\bigcup_{k=0}^{\infty} B_k \supseteq \bigcup_{k=0}^{\infty} A_k$. We have therefore proved (3.7).

We also observe that B_k are pairwise disjoint. Indeed, if $1 \leq i < j$, we have

$$B_i \cap B_j \subseteq A_i \cap B_j = \varnothing,$$

because $y \in A_i$ implies $y \notin B_j$. Applying **(b)**, we finally obtain

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = \mu\Big(\bigcup_{k=0}^{\infty} B_k\Big) = \sum_{k=0}^{\infty} \mu(B_k) \leq \sum_{k=0}^{\infty} \mu(A_k),$$

We leave the proof of **(e)** as an exercise.

Consider a sequence $(A_k)_{k\in\mathbb{N}}$ of subsets of X satisfying $A_k \subseteq A_{k+1}$ or $A_{k+1} \subseteq A_k$ for all $k \in \mathbb{N}$. Intuitively, we can think of $\bigcup_{k=0}^{\infty} A_k$ in the first case, and of $\bigcap_{k=0}^{\infty} A_k$ in the second case as "limits" of the sequence $(A_k)_{k\in\mathbb{N}}$. We now show that a measure is continuous with respect to this notion of limit (with a small exception).

Theorem 3.11 [Continuity of the Measure]
Let X be a set and let $\mathcal{A}$ be a σ-algebra in X. Let $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ be a measure and let $(A_k)_{k\in\mathbb{N}}$ be a sequence of sets in $\mathcal{A}$. Then:

(a) *If $A_k \subseteq A_{k+1}$ for all $k \in \mathbb{N}$,*

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = \lim_{k\to\infty} \mu(A_k);$$

(b) *If $A_{k+1} \subseteq A_k$ for all $k \in \mathbb{N}$ and $\mu(A_0) < +\infty$,*

$$\mu\Big(\bigcap_{k=0}^{\infty} A_k\Big) = \lim_{k\to\infty} \mu(A_k).$$

Proof We start proving **(a)**. First, observe that the limit in the right hand-side exists because, from **(b)** in Theorem 3.10, the sequence $(\mu(A_k))_{k\in\mathbb{N}}$ is monotonically non-decreasing. For $k \in \mathbb{N}$, we define B_k as in the

proof of **(d)** in Theorem 3.10. In this case, we have $B_k = A_k \setminus A_{k-1}$. Applying **(a)** in Theorem 3.10 we find

$$\begin{aligned}\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) &= \mu\Big(\bigcup_{k=0}^{\infty} B_k\Big)\\ &= \sum_{k=0}^{\infty} \mu(B_k)\\ &= \lim_{\ell\to\infty} \sum_{k=0}^{\ell} \mu(B_k)\\ &= \lim_{\ell\to\infty} \mu\Big(\bigcup_{k=0}^{\ell} B_k\Big).\end{aligned}$$

We now show that

$$\bigcup_{k=0}^{\ell} B_k = A_\ell. \tag{3.8}$$

We prove first that the left hand-side set of (3.8) is contained in the right hand-side. Indeed, if $k \le \ell$, we have $B_k \subseteq A_k \subseteq A_\ell$, which proves $\bigcup_{k=0}^{\ell} B_k \subseteq A_\ell$. We now prove the opposite inclusion. If $x \in A_\ell$, let k_0 be the smallest index k such that $x \in A_k$. Clearly, $k_0 \le \ell$, and as in the proof of point **(d)** in Theorem 3.10, we find $x \notin A_r$ for every $r < k_0$, so that $x \in B_{k_0}$, which shows that $x \in \bigcup_{k=0}^{\ell} B_k$. Thus, $\bigcup_{k=0}^{\ell} B_k \supseteq A_\ell$. This proves (3.8), which in turn proves **(a)**.

We now prove **(b)**. Again, the existence of the limit in the right hand-side follows from the monotonicity of the sequence $(\mu(A_k))_{k\in\mathbb{N}}$, which in this case is non-increasing. Moreover, from **(c)** in Theorem 3.10, we have $\bigcap_{k=0}^{\infty} A_k \in \mathcal{A}$. Let us define $A = \bigcap_{k=0}^{\infty} A_k$, and for every $k \in \mathbb{N}$ we define

$$B_k = A_k \setminus A_{k+1}.$$

We prove that

$$A \cup \bigcup_{k=0}^{\infty} B_k = A_0. \tag{3.9}$$

We first show that the left hand-side of (3.9) is contained in the right hand-side of (3.9). If $x \in A$, clearly $x \in A_0$. Moreover, for all $k \in \mathbb{N}$ we have $B_k \subseteq A_k \subseteq A_0$. Hence, $\bigcup_{k=0}^{\infty} B_k \subseteq A_0$, and therefore also $A \cup \bigcup_{k=0}^{\infty} B_k \subseteq A_0$.

Now we prove the opposite inclusion. Let $x \in A_0$. We show that $x \in A$ or $x \in \bigcup_{k=0}^{\infty} B_k$. If x belongs to every A_k, then $x \in A$. If, on the contrary, there exists $k_0 \ge 1$ such that $x \notin A_{k_0}$, from the monotonicity we have $x \notin A_j$ for every $j \ge k_0$, and therefore the set $\{k \in \mathbb{N} \,|\, x \in A_k\}$ is non-empty and finite. Denoting by k_1 the maximum of this set, we find $x \in A_{k_1} \setminus A_{k_1+1} = B_{k_1}$, which shows that $x \in \bigcup_{k=0}^{\infty} B_k$. This proves (3.9). We now prove that $A, B_0, B_1, B_2, \ldots$ are pairwise disjoint. Let $x \in A$. Then, $x \in A_{k+1}$ for all $k \in \mathbb{N}$, and therefore $x \notin B_k$ for all $k \in \mathbb{N}$. This shows that A is disjoint from B_k for every $k \in \mathbb{N}$. We now prove that the sets of the sequence

$(B_k)_{k\in\mathbb{N}}$ are all pairwise disjoint. If $0 \le i < j$, from the monotonicity we find $A_j \subseteq A_i$ and $A_i^{\complement} \subseteq A_j^{\complement}$, so that

$$\begin{aligned} B_i \cap B_j &= (A_i \setminus A_{i+1}) \cap (A_j \setminus A_{j+1}) \\ &= (A_i \cap A_{i+1}^{\complement}) \cap (A_j \cap A_{j+1}^{\complement}) \\ &\subseteq A_{i+1}^{\complement} \cap A_j = \varnothing, \end{aligned}$$

since, from $j \ge i+1$, we have $A_j \subseteq A_{i+1}$. We have proved in this way that $A, B_0, B_1, B_2, \ldots$ are pairwise disjoint. Using the definition of measure we therefore have

$$\mu(A) + \sum_{k=0}^{\infty} \mu(B_k) = \mu(A_0).$$

Since $\mu(A_0) < +\infty$, the series is convergent in $\mathbb{R}$, the sum of the series is a real number, and therefore

$$\begin{aligned} \mu(A) &= \mu(A_0) - \sum_{k=0}^{\infty} \mu(B_k) \\ &= \lim_{\ell\to\infty} \left(\mu(A_0) - \sum_{k=0}^{\ell} \mu(B_k) \right) = \lim_{\ell\to\infty} \left(\mu(A_0) - \mu\Big(\bigcup_{k=0}^{\ell} B_k \Big) \right). \end{aligned}$$

We now verify by induction on ℓ that $\bigcup_{k=0}^{\ell} B_k = A_0 \setminus A_{\ell+1}$. If $\ell = 0$, this is obvious (the left hand-side is B_0 and the right hand-side is $A_0 \setminus A_1$). We now assume that $\bigcup_{k=0}^{\ell} B_k = A_0 \setminus A_{\ell+1}$, and we prove that $\bigcup_{k=0}^{\ell+1} B_k = A_0 \setminus A_{\ell+2}$. We have

$$\begin{aligned} \bigcup_{k=0}^{\ell+1} B_k &= \Big(\bigcup_{k=0}^{\ell} B_k \Big) \cup B_{\ell+1} = (A_0 \setminus A_{\ell+1}) \cup (A_{\ell+1} \setminus A_{\ell+2}) \\ &= (A_0 \cap A_{\ell+1}^{\complement}) \cup (A_{\ell+1} \cap A_{\ell+2}^{\complement}) \\ &= (A_0 \cup A_{\ell+1}) \cap (A_0 \cup A_{\ell+2}^{\complement}) \cap (A_{\ell+1}^{\complement} \cup A_{\ell+1}) \cap (A_{\ell+1}^{\complement} \cup A_{\ell+2}^{\complement}) \\ &= A_0 \cap (A_0 \cup A_{\ell+2}^{\complement}) \cap X \cap A_{\ell+2}^{\complement} \\ &= A_0 \cap A_{\ell+2}^{\complement} = A_0 \setminus A_{\ell+2}, \end{aligned}$$

where we have used the distributive property of the union over the intersection. From point **(b)** of Theorem 3.10 we find

$$\begin{aligned} \mu(A) &= \lim_{\ell\to\infty} \left(\mu(A_0) - \mu\Big(\bigcup_{k=0}^{\ell} B_k \Big) \right) \\ &= \lim_{\ell\to\infty} \left(\mu(A_0) - \mu(A_0 \setminus A_{\ell+1}) \right) \\ &= \lim_{\ell\to\infty} \mu(A_{\ell+1}) = \lim_{\ell\to\infty} \mu(A_\ell). \end{aligned}$$

The assumption $\mu(A_0) < +\infty$ in **(b)** of Theorem 3.11 can be replaced with the weaker assumption $\mu(A_k) < +\infty$ for some sufficiently large k. On the contrary, if $\mu(A_k) = +\infty$ for every $k \in \mathbb{N}$, the conclusion may be false. We will see an example in the sequel.

Exercises

Exercise 3.9

Let X be a non-empty set. Let $x_0 \in X$. Show that the function

$$\begin{aligned} \mu : \mathcal{P}(X) &\longrightarrow [0, +\infty] \\ A &\mapsto \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is a measure. [Solution: (3.9)]

3.3.1 n-dimensional intervals and their volume

Riemann integration of functions of a single variable is usually performed over bounded intervals. In the case of unbounded intervals, we use the theory of improper integrals. For integrals of functions of several variables, the choice of sets over which to integrate — which we will call *measurable sets* — is much broader: already in the case of two variables, it is quite natural to require that the sets of integration include, for example, circles, triangles, etc. Before defining the integral, then, we will identify a class — as broad as possible — of measurable sets, and to assign them a *measure*: the so-called *Lebesgue measure*. This measure must satisfy certain basic properties. As we will see, it is not possible to assign a measure to all subsets of $\mathbb{R}^n$, but the class of subsets measurable in the sense of Lebesgue will still be very rich. The "building blocks" used to construct Lebesgue measure are the n-dimensional intervals, which are, in essence, parallelepipeds with edges parallel to the coordinate axes.

We recall that an interval is a subset I of $\mathbb{R}$ if, for all $a, b, c \in \mathbb{R}$ which satisfy $a \leq b \leq c$ and such that $a, c \in I$, then also $b \in I$. In particular, the empty set is an interval (it vacuously satisfies the definition of interval), and a singleton $[a, a] = \{a\}$, where $a \in \mathbb{R}$, is an interval. A non-degenerate interval of $\mathbb{R}$ is also said to be *proper*. It is not difficult to see that the following statements on an interval I of the real line are equivalent:

- I is non-degenerate;
- I has an infinite number of elements;
- I has at least two distinct elements;
- the interior of I is non-empty.

We begin by defining the length of a bounded interval (which does not need to be open or closed) in an obvious way.

Definition 3.4 [Length of an Interval of $\mathbb{R}$]
If I is a bounded interval in $\mathbb{R}$, we define its **length** *by*

$$\ell(I) \stackrel{def}{=} \begin{cases} \sup I - \inf I & \text{if } I \neq \varnothing \\ 0 & \text{if } I = \varnothing. \end{cases}$$

For example, $\ell([0, 1]) = \sup([0, 1]) - \inf([0, 1]) = \max([0, 1]) - \min([0, 1]) = 1 - 0 = 1$. Likewise, $\ell([0, 1)) = \sup([0, 1)) - \inf([0, 1)) = \sup([0, 1)) - \min([0, 1)) = 1 - 0 = 1$. Similarly, $\ell((0, 1]) = \ell((0, 1)) = 1$.

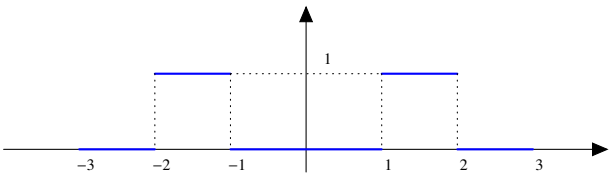
If I consists of a single point, i.e., $I = [a, a]$ with $a \in \mathbb{R}$, then $\ell(I) = \sup I - \inf I = a - a = 0$. It follows that we can have $\ell(I) = 0$ both when I is empty and when the interval I is a single point (i.e., when the interval is degenerate). If $a, b \in I$ and $a \neq b$, we have $\ell(I) \geq |b - a| > 0$.

Remark 3.2 We could extend Definition 3.4 to include intervals which are not bounded. In this case, we would define the length to be equal to $+\infty$. However, for our purposes there is nothing to gain in considering unbounded intervals. In fact, every subset of $\mathbb{R}$, bounded or unbounded, can be "covered" with a countable union of bounded intervals: this is the idea that underpins Lebesgue's measure theory.

Given two sets A and B satisfying $A \subseteq B$, we recall that the function

$$\begin{aligned} \chi_A : B &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is referred to as the *characteristic function* or *indicator function* of A.



Example: characteristic function of $A = (-2,-1) \cup (1,2)$ in $B = [-3,3]$

Moreover, if a function $f : \mathbb{R} \longrightarrow \mathbb{C}$ is integrable as an improper integral over $(-\infty, +\infty)$, we denote the integral with the usual notation

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$$

We now present some properties of the length ℓ that will be useful in the sequel.

Theorem 3.12 *Let I and J be bounded intervals of $\mathbb{R}$. Then:*

(a) $\ell(I) = \displaystyle\int_{-\infty}^{+\infty} \chi_I(x)\,\mathrm{d}x$;

(b) *$I \cap J$ is a bounded interval;*

(c) *there exist two disjoint intervals I_1 and I_2 (not necessarily non-empty) such that $I \setminus J = I_1 \cup I_2$;*

(d) *Let $p \in \mathbb{N}$. Let $I_1, \ldots, I_p$ be bounded and pairwise disjoint (i.e., for all $i, j \in \{1, \ldots, p\}$, then $i \neq j$ implies $I_i \cap I_j = \varnothing$). Let $I = \displaystyle\bigcup_{j=1}^{p} I_j$. Then,*

$$\ell(I) = \sum_{j=1}^{p} \ell(I_j).$$

Proof The first statement is trivially true if $I = \varnothing$. If I is not empty, let $a = \inf I$ and $b = \sup I$. If $y > b$, we have

$$\int_a^y \chi_I(x)\,\mathrm{d}x = \int_a^b \chi_I(x)\,\mathrm{d}x = [x]_a^b = b - a,$$

so that the function

$$\begin{aligned} (b, +\infty) &\longrightarrow \mathbb{R} \\ y &\mapsto \int_a^y \chi_I(x)\,\mathrm{d}x \end{aligned}$$

is convergent as $y \to +\infty$, and taking the limit as $y \to +\infty$, we find

$$\int_a^{+\infty} \chi_I(x)\,\mathrm{d}x = b - a.$$

Moreover, for every $y < a$ we have

$$\int_y^a \chi_I(x)\,\mathrm{d}x = \int_y^a 0\,\mathrm{d}x = 0,$$

so that the function

$$\begin{aligned}(-\infty, a) &\longrightarrow \mathbb{R}\\ y &\mapsto \int_y^a \chi_I(x)\,\mathrm{d}x\end{aligned}$$

is convergent as $y \to -\infty$, and taking the limit as $y \to -\infty$, we find

$$\int_{-\infty}^a \chi_I(x)\,\mathrm{d}x = 0.$$

It follows that χ_I is integrable as an improper Riemann integral over $(-\infty, +\infty)$ and

$$\int_{-\infty}^{+\infty} \chi_I(x)\,\mathrm{d}x = \int_{-\infty}^a \chi_I(x)\,\mathrm{d}x + \int_a^{+\infty} \chi_I(x)\,\mathrm{d}x = 0 + (b-a) = \ell(I).$$

We prove **(b)**. First, note that $I \cap J$ is bounded because it is a subset of bounded sets. Now, we show that $I \cap J$ is an interval. Let $a, b, c \in \mathbb{R}$, where $a, c \in I \cap J$. Then, since $a, c \in I$, then $b \in I$ from the definition of interval. Likewise, $a, c \in J$, so that $b \in J$. Therefore $b \in I \cap J$, so that $I \cap J$ is an interval.

We now prove **(c)**. We have $J^{\complement} = J_1 \cup J_2$, where J_1 and J_2 are disjoint unbounded intervals. Hence,

$$I \setminus J = I \cap J^{\complement} = I \cap (J_1 \cup J_2) = (I \cap J_1) \cup (I \cap J_2).$$

We define $I_1 = I \cap J_1$ and $I_2 = I \cap J_2$. Then, I_1 and I_2 are bounded because they are subsets of I, and they are disjoint because such are J_1 and J_2. As in the proof of **(b)**, we can verify that I_1 and I_2 are intervals (try!), and this concludes the proof of **(c)**.

We prove **(d)**. For all $x \in \mathbb{R}$, we have

$$\chi_I(x) = \sum_{j=1}^p \chi_{I_j}(x).$$

Indeed, if $x \notin I$, then $\chi_{I_j}(x) = 0$ for all $j \in \{1, \ldots, p\}$. If $x \in I$, then x belongs to one and only one of $I_1, \ldots, I_p$. From **(a)**, it follows that

$$\ell(I) = \int_{-\infty}^{+\infty} \chi_I(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \sum_{j=1}^p \chi_{I_j}(x)\,\mathrm{d}x = \sum_{j=1}^p \int_{-\infty}^{+\infty} \chi_{I_j}(x)\,\mathrm{d}x = \sum_{j=1}^p \ell(I_j).$$

We now generalize the notion of interval to $\mathbb{R}^n$: a bounded interval (or n-box, or n-dimensional rectangle) in $\mathbb{R}^n$ is the Cartesian product of n bounded intervals in $\mathbb{R}$.

Definition 3.5 [n-Dimensional Interval]
Let $n \in \mathbb{N}^*$. Let $\boldsymbol{I} \subseteq \mathbb{R}^n$. We say that $\boldsymbol{I}$ is an n-dimensional interval (or n-box, or n-dimensional rectangle) if there exist bounded intervals $I_1, I_2, \ldots, I_n$ of $\mathbb{R}$ such that

$$\boldsymbol{I} = I_1 \times I_2 \times \ldots \times I_n = \prod_{j=1}^n I_j. \tag{3.10}$$

According to this definition, a 1-dimensional interval is a bounded interval of $\mathbb{R}$.

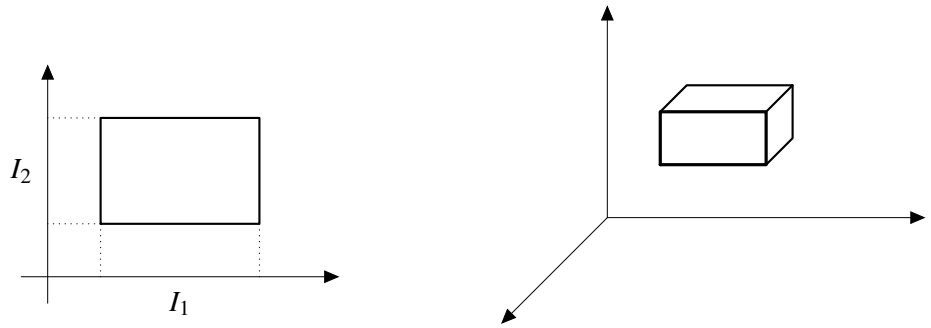
For every $i \in \{1, 2, \ldots, n\}$, let $a_i, b_i \in \mathbb{R}$ be such that $b_i \geq a_i$. An open box in $\mathbb{R}^n$ takes the form

$$\boldsymbol{I} = (a_1, b_1) \times \ldots \times (a_n, b_n) = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid \forall\, i = 1, 2, \ldots, n, \quad x_i \in (a_i, b_i)\}, \tag{3.11}$$

and a closed box in $\mathbb{R}^n$ takes the form

$$\boldsymbol{I} = [a_1, b_1] \times \ldots \times [a_n, b_n] = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid \forall\, i = 1, 2, \ldots, n, \quad x_i \in [a_i, b_i]\}. \tag{3.12}$$

Geometrically, an n-dimensional interval can be viewed as a segment in $\mathbb{R}$, a rectangle in $\mathbb{R}^2$, a parallelepiped in $\mathbb{R}^3$, and so on, with sides/edges oriented parallel to the coordinate axes.



Definition 3.6 [n-Dimensional Volume]
Let $n \in \mathbb{N}^*$. Let $\boldsymbol{I}$ be an n-dimensional interval in $\mathbb{R}^n$. Let $I_1, \ldots, I_n$ be bounded intervals of $\mathbb{R}$ such that $\boldsymbol{I} = I_1 \times I_2 \times \ldots \times I_n$. We define the volume of I as

$$\mathrm{vol}_n(\boldsymbol{I}) \overset{def}{=} \ell(I_1) \cdot \ell(I_2) \cdot \ldots \cdot \ell(I_n) = \prod_{j=1}^{n} \ell(I_j).$$

Notice that there is some abuse of notation in the use of the symbol $\prod$ in Definition 3.5 and in Definition 3.6. Indeed, in Definition 3.5 it is used to denote Cartesian product, whereas in Definition 3.6 it represents an ordinary product of real numbers.

Importantly, Definition 3.6 is well posed: the volume of an n-dimensional interval does not depend on the choice of the bounded intervals $I_1, I_2, \ldots, I_n$ of $\mathbb{R}$ such that $\boldsymbol{I} = I_1 \times I_2 \times \ldots \times I_n$. Indeed, if $\boldsymbol{I} \neq \varnothing$, the intervals $I_1, \ldots, I_n$ are uniquely determined: we have in fact $I_j = \pi_j(\boldsymbol{I})$ for all $j \in \{1, \ldots, n\}$, where π_j is the projection onto the j-th coordinate:

$$\begin{aligned} \pi_j : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x_1, \ldots, x_j, \ldots, x_n) &\mapsto x_j. \end{aligned}$$

If $\boldsymbol{I} = \varnothing$, the intervals $I_1, I_2, \ldots, I_n$ are not uniquely determined: in all cases where at least one of these intervals is empty, we have $\boldsymbol{I} = \varnothing$. However, in any case the product is 0, and therefore $\mathrm{vol}_n(\boldsymbol{I})$ is always 0.

Notice that if $n = 1$, i.e., if I is a bounded interval of $\mathbb{R}$, we have

$$\mathrm{vol}_1(I) = \ell(I).$$

If the n-box is in the form given by (3.11) or (3.12), we have

$$\mathrm{vol}_n(\boldsymbol{I}) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

Remark 3.3 It is possible to extend the definition of n-dimensional interval by relaxing the assumptions that the intervals $I_1, I_2, \ldots, I_n$ be bounded. In this way, we included unbounded n-dimensional intervals of $\mathbb{R}^n$. If I is an unbounded interval of $\mathbb{R}$, we can define its length as $\ell(I) \overset{\text{def}}{=} +\infty$. We can also generalize the notion of volume to these unbounded intervals, but we have to use the convention $0 \cdot (+\infty) = 0$, which is necessary in the case where one of the intervals is unbounded and another one is a singleton. For example, in $\mathbb{R}^2$, the area of the line $\{0\} \times \mathbb{R}$ must be zero. With this definition, if one or more of the intervals $I_1, I_2, \ldots, I_n$ is unbounded, the volume $\mathrm{vol}_n(\boldsymbol{I})$ is $+\infty$, unless one other interval among $I_1, I_2, \ldots, I_n$ has length equal to 0. In that case, $\mathrm{vol}_n(\boldsymbol{I}) = 0$.

However, as explained in Remark 3.2, this generalization is not essential for the treatment of the Lebesgue

measure.

Example *3.1*

Let $\boldsymbol{I} = (0,2) \times [1,2]$ be a 2-dimensional interval. Then

$$\mathrm{vol}_2(\boldsymbol{I}) = \ell((0,2)) \cdot \ell([1,2]) = (2-0) \cdot (2-1) = 2.$$

Example *3.3*

Let $\boldsymbol{I} = [2,3) \times [-1,1] \times [1,6]$ be a 3-dimensional interval. Then

$$\mathrm{vol}_3(\boldsymbol{I}) = \ell([2,3)) \cdot \ell([-1,1]) \cdot \ell([1,6]) = (3-2)(1+1)(6-1) = 10.$$

Example *3.3*

Generalizing what we saw in the previous examples, it is easy to see that, given an n-dimensional interval $\boldsymbol{I}$, there holds

$$\mathrm{vol}_n(\boldsymbol{I}) = \mathrm{vol}_n(\bar{\boldsymbol{I}}) = \mathrm{vol}_n(\mathring{\boldsymbol{I}}).$$

Example *3.4*

The open unit hyper-cube

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,|\, 0 < x_i < 1, \;\; i \in \{1,2,\ldots,n\}\}$$

can be seen as an (open) n-dimensional interval of $\mathbb{R}^n$. Its volume is 1.

Exercise 3.10

Let $\boldsymbol{I} = \prod_{i=1}^{n} I_i$ be an n-dimensional interval. Let $\ell \in \{1,2,\ldots,n\}$. We denote by a_i and b_i the end-points of the interval I_ℓ. If ξ satisfies $a_\ell < \xi < b_\ell$, and we define

$$\boldsymbol{I}_1 = \boldsymbol{I} \cap \{\boldsymbol{x} \in \mathbb{R}^n \,|\, x_\ell < \xi\} \quad \text{and} \quad \boldsymbol{I}_2 = \boldsymbol{I} \cap \{\boldsymbol{x} \in \mathbb{R}^n \,|\, x_\ell > \xi\},$$

show that $\mathrm{vol}_n(\boldsymbol{I}) = \mathrm{vol}_n(\boldsymbol{I}_1) + \mathrm{vol}_n(\boldsymbol{I}_2)$.

Indeed,

$$\mathrm{vol}_n(\boldsymbol{I}_1) = \Big(\prod_{\substack{j=1 \\ j\neq\ell}}^{n} (b_i - a_i)\Big)(\xi - a_\ell)$$

$$\mathrm{vol}_n(\boldsymbol{I}_2) = \Big(\prod_{\substack{j=1 \\ j\neq\ell}}^{n} (b_i - a_i)\Big)(b_\ell - \xi),$$

and so

$$\mathrm{vol}_n(\boldsymbol{I}_1) + \mathrm{vol}_n(\boldsymbol{I}_2) = \Big(\prod_{\substack{j=1 \\ j\neq\ell}}^{n} (b_i - a_i)\Big)(b_\ell - \xi + \xi - a_\ell) = \prod_{j=1}^{n} (b_i - a_i) = \mathrm{vol}_n(\boldsymbol{I}).$$

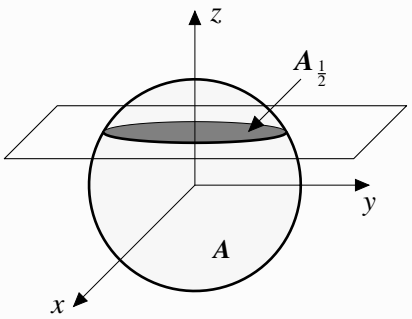
We introduce a notation that we will use extensively in the sequel. Let $n \in \mathbb{N}^*$ be such that $n \geq 2$. Let $\boldsymbol{A} \subseteq \mathbb{R}^n$. Let $x_n \in \mathbb{R}$. We define

$$\boldsymbol{A}_{x_n} \stackrel{\text{def}}{=} \{\boldsymbol{x}' \in \mathbb{R}^{n-1} \,|\, (\boldsymbol{x}', x_n) \in \boldsymbol{A}\}. \tag{3.13}$$

Example 3.5

Let $\boldsymbol{A} = \{(x,y,z) \in \mathbb{R}^3 \,|\, x^2+y^2+z^2 \leq 1\}$. Geometrically, $\boldsymbol{A}$ is the closed unit sphere in $\mathbb{R}^3$. We have

$$\boldsymbol{A}_{\frac{1}{2}} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x^2+y^2 \leq 1-\left(\frac{1}{2}\right)^2\right\} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x^2+y^2 \leq \frac{3}{4}\right\}$$
$$\boldsymbol{A}_{\frac{3}{2}} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x^2+y^2 \leq 1-\left(\frac{3}{2}\right)^2\right\} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x^2+y^2 \leq -\frac{5}{4}\right\} = \varnothing.$$



Exercise 3.11

Let $\boldsymbol{I} = I_1 \times \ldots \times I_n$ be an n-dimensional interval of $\mathbb{R}^n$, where $I_1, \ldots, I_n$ are bounded intervals of $\mathbb{R}$. Show that for all $x_n \in \mathbb{R}$

$$\boldsymbol{I}_{x_n} = \begin{cases} I_1 \times I_2 \times \ldots \times I_{n-1} & \text{if } x_n \in I_n \\ \varnothing & \text{if } x_n \notin I_n. \end{cases}$$

[Solution: (3.10)]

We also need the following elementary result of set theory.

Exercise 3.12

Let X, Y, A, B, C, D be sets satisfying $A, C \subseteq X$ and $B, D \subseteq Y$. Show that

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D).$$

[Solution: (3.11)]

Lemma 3.1 *Let X, Y, A, B, C, D be sets satisfying $A, C \subseteq X$ and $B, D \subseteq Y$. Then,*

$$(A \times B) \setminus (C \times D) = \big((A \setminus C) \times B\big) \cup \big((A \cap C) \times (B \setminus D)\big). \tag{3.14}$$

Moreover, the sets $(A \setminus C) \times B$ and $(A \cap C) \times (B \setminus D)$ in the right hand-side of (3.14) are disjoint.

Proof First, we notice that

$$(A \times B) \setminus (C \times D) = (A \times B) \cap (C \times D)^{\complement}. \tag{3.15}$$

Now, we prove that

$$(C \times D)^{\complement} = (C^{\complement} \times Y) \cup (C \times D^{\complement}). \tag{3.16}$$

We first prove the inclusion $\subseteq$. Let $(x,y) \in (C \times D)^{\complement}$. There are two possibilities: $x \in C^{\complement}$ or $x \in C$. In the first case, $(x,y) \in C^{\complement} \times Y$, and in the second $y \in D^{\complement}$ (in fact, if $y \in D$, then $(x,y) \in C \times D$, against the assumption that $(x,y) \in (C \times D)^{\complement}$). Thus, we proved $\subseteq$. Vice-versa, if $x \in C^{\complement}$, then $(x,y) \notin C \times D$, so that $(x,y) \in (C \times D)^{\complement}$. The same happens when $y \in D^{\complement}$, and this proves the inclusion $\supseteq$ in (3.16).

From (3.15) and (3.16) we can write

$$\begin{aligned}(A\times B)\setminus(C\times D) &= (A\times B)\cap((C^{\complement}\times Y)\cup(C\times D^{\complement}))\\ &= ((A\times B)\cap(C^{\complement}\times Y))\cup((A\times B)\cap(C\times D^{\complement}))\\ &= ((A\cap C^{\complement})\times B)\cup((A\cap C)\times(B\cap D^{\complement}))\\ &= ((A\setminus C)\times B)\cup((A\cap C)\times(B\setminus D)).\end{aligned}$$

We now show that the two sets in this union are disjoint. Indeed, if $(x,y)\in(A\cap C)\times(B\setminus D)$, then $x\in C$, and therefore $x\notin A\setminus C$, and we conclude that $(x,y)\notin(A\setminus C)\times B$.

We are now ready to present some of the fundamental properties of the volume. There is a significant similarity with the properties of intervals of $\mathbb{R}$ of Theorem 3.12. We warn the reader that the intervals I_k of $\mathbb{R}$ in the statement of Theorem 3.12 are not to be confused with the n-dimensional intervals $\boldsymbol{I}^k$ of Theorem 3.13.

Theorem 3.13 *Let $n,p\in\mathbb{N}^*$. Let $\boldsymbol{I}$ and $\boldsymbol{J}$ be n-dimensional intervals. Then,*

(a) *if $n\geq 2$, for every $x_n\in\mathbb{R}$, the set $\boldsymbol{I}_{x_n}$ is an $(n-1)$-dimensional interval; moreover, the function*

$$\begin{aligned}\mathbb{R} &\longrightarrow [0,+\infty)\\ x_n &\mapsto \operatorname{vol}_n(\boldsymbol{I}_{x_n})\end{aligned}$$

is integrable as an improper Riemann integral over $(-\infty,+\infty)$, and

$$\operatorname{vol}_n(\boldsymbol{I})=\int_{-\infty}^{+\infty}\operatorname{vol}_{n-1}(\boldsymbol{I}_{x_n})\,\mathrm{d}x_n; \tag{3.17}$$

(b) *$\boldsymbol{I}\cap\boldsymbol{J}$ is an n-dimensional interval;*

(c) *$\boldsymbol{I}\setminus\boldsymbol{J}$ can be written as the disjoint union of $2n$ (not necessarily all non-empty) n-dimensional intervals;*

(d) *If $\boldsymbol{I}=\bigcup_{j=1}^{p}\boldsymbol{I}^j$ where $\boldsymbol{I}^1,\ldots,\boldsymbol{I}^p$ are pairwise disjoint n-dimensional intervals, then,*

$$\operatorname{vol}_n(\boldsymbol{I})=\sum_{j=1}^{p}\operatorname{vol}_n(\boldsymbol{I}^j).$$

Proof We start proving **(a)**. Let $\boldsymbol{I}=I_1\times\ldots\times I_n$, where $I_1,\ldots,I_n$ are bounded intervals of $\mathbb{R}$ (again, do not confuse these with $I^1,\ldots,I^k$). From Exercise 3.11, for all $x_n\in\mathbb{R}$

$$\boldsymbol{I}_{x_n}=\begin{cases}I_1\times I_2\times\ldots\times I_{n-1} & \text{if } x_n\in I_n\\ \varnothing & \text{if } x_n\notin I_n.\end{cases}$$

It immediately follows that

$$\operatorname{vol}_{n-1}(\boldsymbol{I}_{x_n})=\begin{cases}\displaystyle\prod_{j=1}^{n-1}\ell(I_j) & \text{if } x_n\in I_n\\ 0 & \text{if } x_n\notin I_n,\end{cases}$$

so that

$$\int_{-\infty}^{+\infty} \mathrm{vol}_{n-1}(\boldsymbol{I}_{x_n})\,\mathrm{d}x_n = \int_{-\infty}^{+\infty} \prod_{j=1}^{n-1} \ell(I_j)\chi_{I_n}(x_n)\,\mathrm{d}x_n$$
$$= \prod_{j=1}^{n-1} \ell(I_j) \int_{-\infty}^{+\infty} \chi_{I_n}(x_n)\,\mathrm{d}x_n = \prod_{j=1}^{n} \ell(I_j) = \mathrm{vol}_n(\boldsymbol{I}),$$

which proves **(a)**. We now prove **(b)**. If we write $\boldsymbol{J} = J_1 \times J_2 \times \ldots \times J_n$ and $\boldsymbol{I} = I_1 \times I_2 \times \ldots \times I_n$, where $J_1, \ldots, J_n, I_1, \ldots, I_n$ are bounded intervals of $\mathbb{R}$, then

$$\boldsymbol{I} \cap \boldsymbol{J} = (I_1 \cap J_1) \times (I_2 \cap J_2) \times \ldots \times (I_n \cap J_n),$$

and, for every $k \in \{1, \ldots, n\}$, the set $I_k \cap J_k$ is a bounded interval of $\mathbb{R}$ from **(b)** in Theorem 3.12.

We prove **(c)**, using induction on n. The case $n = 1$ is covered by **(c)** in Theorem 3.12. Assume that the result holds for n. Consider the $(n+1)$-dimensional intervals

$$\boldsymbol{I} = \underbrace{I_1 \times I_2 \times \ldots \times I_n}_{\stackrel{\text{def}}{=} I'} \times I_{n+1} \quad \text{and} \quad \boldsymbol{J} = \underbrace{J_1 \times J_2 \times \ldots \times J_n}_{\stackrel{\text{def}}{=} J'} \times J_{n+1}.$$

From Lemma 3.1, we have

$$\boldsymbol{I} \setminus \boldsymbol{J} = (\boldsymbol{I}' \times I_{n+1}) \setminus (\boldsymbol{J}' \times J_{n+1})$$
$$= ((\boldsymbol{I}' \setminus \boldsymbol{J}') \times I_{n+1}) \cup ((\boldsymbol{I}' \cap \boldsymbol{J}') \times (I_{n+1} \setminus J_{n+1})),$$

where the sets $(\boldsymbol{I}' \setminus \boldsymbol{J}') \times I_{n+1}$ and $(\boldsymbol{I}' \cap \boldsymbol{J}') \times (I_{n+1} \setminus J_{n+1})$ are disjoint. We show that the first term $(\boldsymbol{I}' \setminus \boldsymbol{J}') \times I_{n+1}$ is the union of $2\,n$ disjoint $(n+1)$-dimensional intervals, and the second term $(\boldsymbol{I}' \cap \boldsymbol{J}') \times (I_{n+1} \setminus J_{n+1})$ is the union of 2 disjoint $(n+1)$-dimensional intervals, so that their union $\boldsymbol{I} \setminus \boldsymbol{J}$ is the union of $2\,(n+1)$ disjoint $(n+1)$-dimensional intervals as required.

Consider the second term $(\boldsymbol{I}' \cap \boldsymbol{J}') \times (I_{n+1} \setminus J_{n+1})$. In view of **(b)**, $\boldsymbol{I}' \cap \boldsymbol{J}'$ is an n-dimensional interval. Moreover, from **(c)** in Theorem 3.12 there exist two bounded and disjoint intervals H_1 and H_2 of $\mathbb{R}$ satisfying

$$I_{n+1} \setminus J_{n+1} = H_1 \cup H_2.$$

Hence,

$$(\boldsymbol{I}' \cap \boldsymbol{J}') \times (I_{n+1} \setminus J_{n+1}) = ((\boldsymbol{I}' \cap \boldsymbol{J}') \times H_1) \cup ((\boldsymbol{I}' \cap \boldsymbol{J}') \times H_2),$$

which shows that indeed the second term is the union of 2 disjoint $(n+1)$-dimensional intervals. Now let us consider the first term of the union. From the inductive assumption, we have

$$\boldsymbol{I}' \setminus \boldsymbol{J}' = \bigcup_{r=1}^{2\,n} L_r,$$

where $L_1, \ldots, L_{2\,n}$ are pairwise disjoint n-dimensional intervals. It follows that

$$(\boldsymbol{I}' \setminus \boldsymbol{J}') \times I_{n+1} = \bigcup_{r=1}^{2\,n} (L_r \times I_{n+1}),$$

which says that $(\boldsymbol{I}' \setminus \boldsymbol{J}') \times I_{n+1}$ is the union of $2\,n$ disjoint $(n+1)$-dimensional intervals. This concludes the proof of this statement.
We prove **(d)**, using again induction on n. The case $n = 1$ has been proved Theorem 3.12, point **(d)**. Assume

that the result holds for n. We have already seen in the proof of **(a)** that for every $x_{n+1} \in \mathbb{R}$, the set $\boldsymbol{I}_{x_{n+1}}$ is an n-dimensional interval. The same is true for $\boldsymbol{I}^1_{x_{n+1}}, \dots, \boldsymbol{I}^p_{x_{n+1}}$. We also observe that

$$\bigcup_{k=1}^{p} \boldsymbol{I}^k_{x_{n+1}} = \boldsymbol{I}_{x_{n+1}}, \tag{3.18}$$

and the sets in this union are disjoint. We prove the inclusion $\subseteq$. If $\boldsymbol{x}' \in \bigcup_{k=1}^{p} \boldsymbol{I}^k_{x_{n+1}}$, there exists $k \in \{1, \dots, p\}$ such that $\boldsymbol{x}' \in \boldsymbol{I}^k_{x_{n+1}}$, so that $(\boldsymbol{x}', x_{n+1}) \in \boldsymbol{I}^k$, which implies $(\boldsymbol{x}', x_{n+1}) \in \boldsymbol{I}$ (remember that $\boldsymbol{I}$ is the union of the $\boldsymbol{I}^k$'s), and therefore $\boldsymbol{x}' \in \boldsymbol{I}_{x_{n+1}}$.

We prove the opposite inclusion. If $\boldsymbol{x}' \in \boldsymbol{I}_{x_{n+1}}$, then $(\boldsymbol{x}', x_{n+1}) \in \boldsymbol{I}$, so that there exists $k \in \{1, \dots, p\}$ such that $(\boldsymbol{x}', x_{n+1}) \in \boldsymbol{I}^k$, which proves that $\boldsymbol{x}' \in \boldsymbol{I}^k_{x_{n+1}}$, and therefore $\boldsymbol{x}'$ belongs to the union, as required.

Finally, we prove that the sets involved in the union in the left hand-side of (3.18) are disjoint. If by contradiction $\boldsymbol{x}' \in \boldsymbol{I}^k_{x_{n+1}} \cap \boldsymbol{I}^h_{x_{n+1}}$ for some distinct $h, k \in \{1, \dots, p\}$, then $(\boldsymbol{x}', x_{n+1}) \in \boldsymbol{I}^k \cap I^h$. It follows that $\boldsymbol{I}^1, \dots, \boldsymbol{I}^p$ are not pairwise disjoint. We conclude, using the inductive assumption, that for all $x_{n+1} \in \mathbb{R}$

$$\mathrm{vol}_n(I_{x_{n+1}}) = \sum_{k=1}^{p} \mathrm{vol}_n(I^k_{x_{n+1}}),$$

and therefore, using **(a)**,

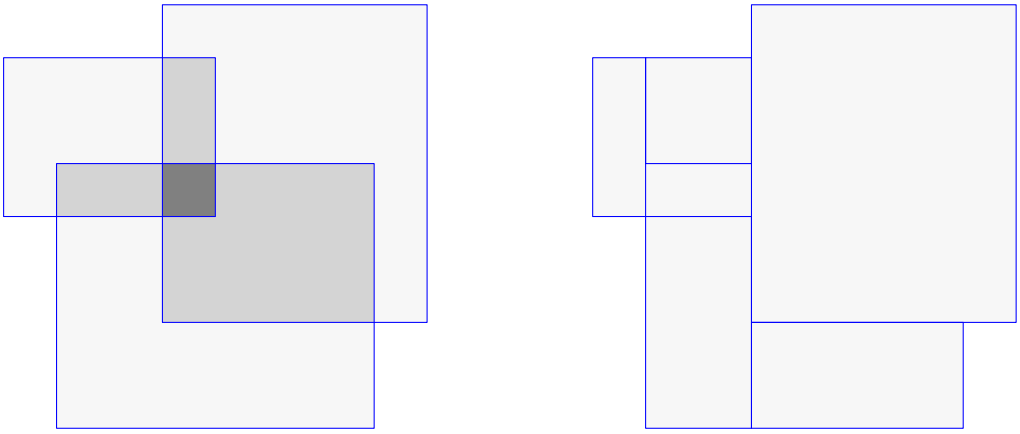
$$\begin{aligned}
\mathrm{vol}_{n+1}(\boldsymbol{I}) &= \int_{-\infty}^{+\infty} \mathrm{vol}_n(\boldsymbol{I}_{x_{n+1}})\, \mathrm{d}x_{n+1} \\
&= \int_{-\infty}^{+\infty} \sum_{k=1}^{p} \mathrm{vol}_n(\boldsymbol{I}^k_{x_{n+1}})\, \mathrm{d}x_{n+1} \\
&= \sum_{k=1}^{p} \int_{-\infty}^{+\infty} \mathrm{vol}_n(\boldsymbol{I}^k_{x_{n+1}})\, \mathrm{d}x_{n+1} = \sum_{k=1}^{p} \mathrm{vol}_{n+1}(\boldsymbol{I}^k).
\end{aligned}$$

Exercises

Exercise 3.13 Show that if a bounded interval I of $\mathbb{R}$ is not degenerate, it has strictly positive length and it contains an infinite number of points. Conclude that an interval of $\mathbb{R}$ can contain 0 points, 1 point, or an infinite number of points. [Solution: (3.12)]

3.3.2 Hints on the Peano-Jordan measure

Peano-Jordan measure is a different approach at measuring subsets of $\mathbb{R}^n$. Consider a subset $\boldsymbol{E}$ of $\mathbb{R}^n$. The idea is to approximate $\boldsymbol{E}$ from the outside and from the inside using n-dimensional intervals of $\mathbb{R}^n$. These n-dimensional intervals may or may not overlap:



The figure shows two possible configurations of 2-dimensional intervals that cover the entire area; the first presents overlaps, whereas the second does not. Intuitively, it is easy to realise that it is always possible to choose the n-dimensional intervals to be "essentially pairwise disjoint": this means that given two of these n-dimensional intervals $\boldsymbol{I}^1$ and $\boldsymbol{I}^2$, their interiors are disjoint, i.e., we will have $\mathring{\boldsymbol{I}}^1 \cap \mathring{\boldsymbol{I}}^2 = \varnothing$. Hence, only their boundary can potentially overlap. This consideration can be made more rigorous as follows: given the union of a finite number of n-dimensional intervals

$$\boldsymbol{P} = \boldsymbol{I}^1 \cup \boldsymbol{I}^2 \cup \ldots \cup \boldsymbol{I}^q,$$

we can define new n-dimensional intervals J^1, J^2, …, J^p such that

$$\boldsymbol{P} = \boldsymbol{J}^1 \cup \boldsymbol{J}^2 \cup \ldots \cup \boldsymbol{J}^p,$$

and for $i, k \in \{1, 2, \ldots, p\}$ we have

$$i \neq k \quad \Rightarrow \quad \mathring{\boldsymbol{J}}^i \cap \mathring{\boldsymbol{J}}^k = \varnothing.$$

Having eliminated the overlaps (or at least having confined them to the boundary), now we can determine the measure of $\boldsymbol{P}$ as the sum of the volumes of each of these rectangles. We define such measure of $\boldsymbol{P}$ as $\tau(\boldsymbol{P})$:

$$\tau(\boldsymbol{P}) = \mathrm{vol}(\boldsymbol{J}^1) + \mathrm{vol}(\boldsymbol{J}^2) + \ldots + \mathrm{vol}(\boldsymbol{J}^p).$$

Given the subset $\boldsymbol{E}$, we can approximate it from the outside, by considering all the unions of n-dimensional intervals that contain $\boldsymbol{E}$. We define

$$m^+(\boldsymbol{E}) = \inf_{\boldsymbol{P}} \{\tau(\boldsymbol{P}) \,|\, \boldsymbol{E} \subseteq \boldsymbol{P}\}.$$

We can approximate it from within, by considering all the unions of n-dimensional intervals that are contained in $\boldsymbol{E}$, and we define

$$m^-(\boldsymbol{E}) = \sup_{\boldsymbol{P}} \{\tau(\boldsymbol{P}) \,|\, \boldsymbol{E} \supseteq \boldsymbol{P}\}.$$

We say that $\boldsymbol{E}$ is Peano-Jordan measurable if $m^+(\boldsymbol{E}) = m^-(\boldsymbol{E})$. In this case, we write $m(\boldsymbol{E}) \stackrel{\text{def}}{=} m^+(\boldsymbol{E}) = m^-(\boldsymbol{E})$.

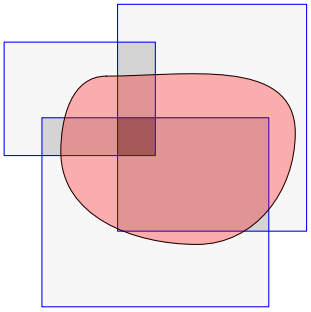
How can the Lebesgue measure do better than this? In essence, it removes the assumption that the number of intervals be finite. However, as we have seen, we cannot use an uncountable number of boxes.

3.3.3 Outer measure

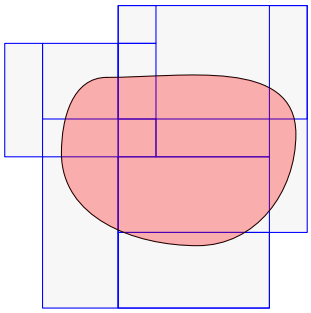
A first approach to obtaining a measure of a set is the notion of n-dimensional outer measure, which consists in covering the set by n-dimensional intervals and then adding up the measure of each n-dimensional interval. Given

for example a set $\boldsymbol{E}$ of $\mathbb{R}^2$ enclosed within the black line in the figure, we cover it by n-dimensional intervals if the n-dimensional intervals $\boldsymbol{I}^k$ are chosen so that

$$\boldsymbol{E} \subseteq \bigcup_k \boldsymbol{I}^k.$$



Obviously, there are infinite ways in which we can accomplish this task: loosely speaking, increasing the number of boxes, we can first of all remove the overlaps.



Increasing the number of boxes potentially leads to a better approximation of the surface, and therefore to a better estimate of the measure: the outer measure is the infimum measure of the covering that we can achieve by considering a countable number of boxes.

We now develop this idea by considering an at-most-countable set J. We say that the family $(\boldsymbol{I}^k)_{k\in J}$ of n-dimensional intervals of $\mathbb{R}^n$ covers $\boldsymbol{E}$ if

$$\boldsymbol{E} \subseteq \bigcup_{k\in J} \boldsymbol{I}^k.$$

If $\boldsymbol{E} \subseteq \mathbb{R}^n$ can be covered by an at-most-countable family of n-dimensional intervals $(\boldsymbol{I}_k)_{k\in J}$, a measure μ_n should satisfy

$$\mu_n(\boldsymbol{E}) \leq \mu_n\Big(\bigcup_{k\in J} \boldsymbol{I}_k\Big) \leq \sum_{k\in J} \mu_n(\boldsymbol{I}^k) = \sum_{k\in J} \mathrm{vol}_n(\boldsymbol{I}^k),$$

where the first inequality follows from the requirement of monotonicity and the second follows from the sub-additivity of μ, and where the last equality expresses our desire to have $\mu_n(\boldsymbol{I}^k) = \mathrm{vol}_n(\boldsymbol{I}^k)$. It follows that for such a function μ we would have

$$\mu_n(\boldsymbol{E}) \leq \inf\Big\{ \sum_{k\in J} \mathrm{vol}_n(\boldsymbol{I}^k) \,\Big|\, J \text{ is at most countable, } (\boldsymbol{I}^k)_{k\in J} \text{ is a family of } n\text{-dimensional intervals and } \boldsymbol{E} \subseteq \bigcup_{k\in J} \boldsymbol{I}^k \Big\}$$

This consideration leads us to defining the outer measure of E as

$$\mu_n^*(\boldsymbol{E}) = \inf\Big\{ \sum_{k\in J} \mathrm{vol}_n(\boldsymbol{I}^k) \,\Big|\, J \text{ is at most countable, } (\boldsymbol{I}^k)_{k\in J} \text{ is a family of } n\text{-dimensional intervals and } \boldsymbol{E} \subseteq \bigcup_{k\in J} \boldsymbol{I}^k \Big\} \tag{3.19}$$

A few observations are in order:

- First, notice that this definition is well posed: indeed, for every $\boldsymbol{E} \subseteq \mathbb{R}^n$, a family $(\boldsymbol{I}^k)_{k\in J}$ of n-dimensional intervals exists which covers E. It is sufficient to consider the case $E = \mathbb{R}^n$, and in this case we can consider $J = \mathbb{N}$, and for all $k \in \mathbb{N}$ we define $\boldsymbol{I}^k = [-k, k]^n$. Hence, the n-dimensional outer measure can be defined for any set $\boldsymbol{E}$ of $\mathbb{R}^n$, because the infimum of a set is well defined if the set is non-empty;[7]
- The infimum in (8.16) and the sums $\sum_{k\in J} \mathrm{vol}_n(\boldsymbol{I}^k)$ are intended in $[0, +\infty]$;
- It is not restrictive to assume that J be countable instead of at most countable: in the case where J is finite, say $J = \{0, 1, \dots, q\}$ with $q \in \mathbb{N}^*$, we can define $\boldsymbol{I}^k = \varnothing$ for all $k > q$ (indeed, n-dimensional intervals are allowed to be empty), and in this case

$$\boldsymbol{E} \subseteq \bigcup_{k=0}^{q} \boldsymbol{I}^k = \bigcup_{k=0}^{\infty} \boldsymbol{I}^k \quad \text{and} \quad \sum_{k=0}^{q} \mathrm{vol}_n(\boldsymbol{I}^k) = \sum_{k\in J} \mathrm{vol}_n(\boldsymbol{I}^k),$$

  because the volume of the empty n-dimensional interval is 0;
- No generality is lost by assuming J to be $\mathbb{N}$: it is only a matter of renumbering the n-dimensional intervals $\boldsymbol{I}^k$ in the family $(\boldsymbol{I}^k)_{k\in J}$, see Theorem 3.6.

From these considerations, we can define the outer measure as follows.

Definition 3.7 [n-Dimensional Outer Measure]
Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. We define the n-dimensional outer measure $\mu_n^*(\boldsymbol{E})$ *of $\boldsymbol{E}$ as the quantity*

$$\mu_n^*(\boldsymbol{E}) = \inf\left\{ \sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k) \,\middle|\, (\boldsymbol{I}^k)_{k\in\mathbb{N}} \text{ is a family of } n\text{-dimensional intervals and } \boldsymbol{E} \subseteq \bigcup_{k\in\mathbb{N}} \boldsymbol{I}^k \right\}.$$

Remark 3.4 We have $\mu^*(\boldsymbol{E}) = +\infty$ if and only if, for any covering $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ of $\boldsymbol{E}$, we have $\sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k) = +\infty$. Vice-versa, we always have $\mu^*(\boldsymbol{E}) \leq \sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k)$ and for every $\varepsilon > 0$, there exists a family $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ of coverings of E such that $\sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k) \leq \mu^*(\boldsymbol{E}) + \varepsilon$ (from the properties of the infimum).

We note that we do not require the n-dimensional intervals in $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ to be pairwise disjoint: thus, the same volume may contribute to multiple terms in the sum $\sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k)$, but this does not affect the value of the infimum.

Example *3.6*

Consider the subset

$$\boldsymbol{E} = \left\{ (x, y) \in \mathbb{R}^2 \,\middle|\, x \geq 1 \quad \text{and} \quad 0 \leq y \leq \frac{1}{x^2} \right\}.$$

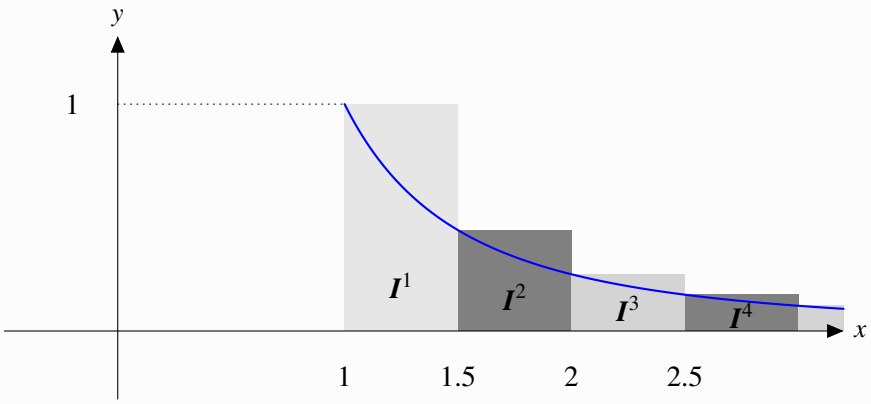
[7] Another covering of the Euclidean space $\mathbb{R}^n$ can be obtained by countably many translates of the open unit hypercube: if we denote by C the unit hypercube, then

$$\mathbb{R}^n = \bigcup_{(i_1,i_2,\dots,i_n)\in\mathbb{Q}\times\mathbb{Q}\times\dots\times\mathbb{Q}} \{(i_1, i_2, \dots, i_n)\} + C.$$

For $k \in \mathbb{N}^*$, consider the 2-dimensional intervals

$$\boldsymbol{I}^k = \left[\frac{k+1}{2}, \frac{k+2}{2}\right] \times \left[0, \frac{4}{(k+1)^2}\right].$$

We have $\boldsymbol{E} \subseteq \bigcup_{k \in \mathbb{N}^*} \boldsymbol{I}^k$, see the figure below:



It follows that

$$\mu_2^*(\boldsymbol{E}) \leq \sum_{k=1}^{\infty} \mathrm{vol}_2(\boldsymbol{I}^k) = \sum_{k=1}^{\infty} \left(\frac{k+2}{2} - \frac{k+1}{2}\right) \cdot \frac{4}{(k+1)^2} = \sum_{k=1}^{\infty} \frac{2}{(k+1)^2}.$$

Since the series is convergent, we can say that $\mu_2^*(\boldsymbol{E}) < +\infty$.

Example *3.7*

Let $n = 1$. We show that $\mu_1^*(\mathbb{N}) = 0$. To this end, it suffices to observe that, for all $k \in \mathbb{N}$, the interval $\{k\}$ is bounded and has length 0. Thus,

$$\mathbb{N} = \bigcup_{k \in \mathbb{N}} \{k\} \qquad \Rightarrow \qquad 0 \leq \mu_1^*(\mathbb{N}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_1(\{k\}) = \sum_{k=0}^{\infty} \ell(\{k\}) = 0.$$

Example *3.8*

We want to prove that the outer measure of $\mathbb{Q} \cap [0,1]$ is zero. Since $\mathbb{Q}$ is countable, so is $\mathbb{Q} \cap [0,1]$, and therefore we can define a sequence $(q_i)_{i \in \mathbb{N}}$ of points of $\mathbb{Q} \cap [0,1]$ which constitutes an enumeration of the elements of this set, and each $\{q_i\}$ is an interval of length 0. Hence, we have the covering

$$\mathbb{Q} \cap [0,1] = \bigcup_{i \in \mathbb{N}} \{q_i\},$$

which can be used to evaluate the outer measure:

$$0 \leq \mu_1^*(\mathbb{Q} \cap [0,1]) \leq \sum_{k=0}^{\infty} \mathrm{vol}_1(\{q_k\}) = \sum_{k=0}^{\infty} \ell(\{q_k\}) = 0.$$

This argument can be used to prove a stronger result: every countable set has 1-dimensional outer measure equal to zero.

The following result presents the fundamental properties of the n-dimensional outer measure.

Theorem 3.14 *We have:*

(a) $\forall\, \boldsymbol{E} \subseteq \mathbb{R}^n$, *we have* $\mu_n^*(\boldsymbol{E}) \in [0, +\infty]$;

(b) $\mu_n^*(\varnothing) = 0$;

(c) $\boldsymbol{E} \subseteq \boldsymbol{F} \subseteq \mathbb{R}^n$ *implies* $\mu_n^*(\boldsymbol{E}) \leq \mu_n^*(\boldsymbol{F})$ (Monotonicity)

(d) *Let* $(\boldsymbol{E}^k)_{k\in\mathbb{N}}$ *be a family of subsets of* $\mathbb{R}^n$ *and let* $\boldsymbol{E} = \bigcup_{k=0}^{\infty} \boldsymbol{E}^k$.

There holds $\mu_n^*(\boldsymbol{E}) \leq \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{E}^k)$ (Sub-Additivity)

(e) *Let* $m \in \mathbb{N}$, *and* $\boldsymbol{E}^0, \dots, \boldsymbol{E}^m \subseteq \mathbb{R}^n$. *Let* $\boldsymbol{E} = \bigcup_{k=0}^{m} \boldsymbol{E}^k$. *Then,*

$\mu_n^*(\boldsymbol{E}) \leq \sum_{k=0}^{m} \mu_n^*(\boldsymbol{E}^k)$ (Finite Sub-Additivity)

(f) *For all* $\boldsymbol{E} \subseteq \mathbb{R}^n$ *and* $\boldsymbol{x} \in \mathbb{R}^n$, *we have* $\mu_n^*(\boldsymbol{x} + \boldsymbol{E}) = \mu_n^*(\boldsymbol{E})$ (Translation Invariance).

Proof To prove **(a)**, note that $\mathrm{vol}_n(\boldsymbol{I}^k)$ is a non-negative quantity for every n-dimensional interval of $\mathbb{R}^n$, and therefore such is also $\sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k)$.

We prove **(b)**. The empty set $\varnothing$ has n-dimensional outer measure equal to 0. In fact, the empty set is a subset of every set, and the measure of every set is always non-negative. Take $\boldsymbol{I}^k = \varnothing$ for every $k \in \mathbb{N}$, so that

$$0 \leq \mu_n^*(\varnothing) \leq \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) = 0,$$

from which we find $\mu_n^*(\varnothing) = 0$.

We prove **(c)**. This property is obvious if $\mu_n^*(\boldsymbol{F}) = +\infty$. Suppose $\mu_n^*(\boldsymbol{F}) < +\infty$. Then, there exists a family $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ of n-dimensional intervals of $\mathbb{R}^n$ for which $\boldsymbol{F} \subseteq \bigcup_{k\in\mathbb{N}} \boldsymbol{I}^k$. Therefore, we have also $\boldsymbol{E} \subseteq \bigcup_{k\in\mathbb{N}} \boldsymbol{I}^k$. It follows that

$$\mu_n^*(\boldsymbol{E}) \leq \sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k).$$

Hence, $\mu_n^*(\boldsymbol{E})$ is a lower bound of the set

$$\Big\{ \sum_{k\in\mathbb{N}} \mathrm{vol}_n(\boldsymbol{I}^k) \,\Big|\, (\boldsymbol{I}^k)_{k\in\mathbb{N}} \text{ is a family of } n\text{-dimensional intervals and } \boldsymbol{F} \subseteq \bigcup_{k\in\mathbb{N}} \boldsymbol{I}^k \Big\}. \tag{3.20}$$

Since $\mu_n^*(\boldsymbol{F})$ is, by definition of infimum, the maximum of the set of lower bounds of (3.20), we obtain $\mu_n^*(\boldsymbol{E}) \leq \mu_n^*(\boldsymbol{F})$.

We now prove **(d)**. This is trivial if $\sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{E}^k) = +\infty$. Suppose $\sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{E}^k) < +\infty$. Then, $\mu_n^*(\boldsymbol{E}^k) < +\infty$ for every $k \in \mathbb{N}$. Let $\varepsilon > 0$. For every $k \in \mathbb{N}$, there exists a family $(\boldsymbol{I}^{(k,j)})_{j\in\mathbb{N}}$ of n-dimensional intervals such that (using Remark 3.4)

$$\boldsymbol{E}^k \subseteq \bigcup_{j=0}^{\infty} \boldsymbol{I}^{(k,j)} \quad \text{and} \quad \sum_{j=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{(k,j)}) \leq \mu^*(\boldsymbol{E}^k) + \frac{\varepsilon}{2^{k+1}}.$$

We recall that $\mathbb{N} \times \mathbb{N}$ is countable, and therefore there exists a bijection $\varphi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ which allows us to

write

$$\boldsymbol{E} = \bigcup_{k=0}^{\infty} \boldsymbol{E}^k \subseteq \bigcup_{k=0}^{\infty} \bigcup_{j=0}^{\infty} \boldsymbol{I}^{(k,j)} = \bigcup_{\ell=0}^{\infty} \boldsymbol{I}^{\varphi(\ell)}.$$

Applying Theorems 3.4, 3.5 and 3.6 we find

$$\begin{aligned}
\mu_n^*(\boldsymbol{E}) &\leq \sum_{\ell=0}^{\infty} \mathrm{vol}_n(I^{\varphi(\ell)}) = \sum_{\ell\in\mathbb{N}} \mathrm{vol}_n(I^{\varphi(\ell)}) \\
&= \sum_{(k,j)\in\mathbb{N}\times\mathbb{N}} \mathrm{vol}_n(I^{(k,j)}) = \sum_{k\in\mathbb{N}} \Big(\sum_{j\in\mathbb{N}} \mathrm{vol}_n(I^{(k,j)}) \Big) \\
&= \sum_{k=0}^{\infty} \Big(\sum_{j=0}^{\infty} \mathrm{vol}_n(I^{(k,j)}) \Big) \\
&\leq \sum_{k=0}^{\infty} \Big(\mu^*(\boldsymbol{E}^k) + \frac{\varepsilon}{2^{k+1}} \Big) = \sum_{k=0}^{\infty} \mu^*(\boldsymbol{E}^k) + \frac{\varepsilon}{2} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \\
&= \sum_{k=0}^{\infty} \mu^*(\boldsymbol{E}^k) + \frac{\varepsilon}{2} \frac{1}{1-\frac{1}{2}} = \sum_{k=0}^{\infty} \mu^*(\boldsymbol{E}^k) + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we obtain

$$\mu^*(\boldsymbol{E}) \leq \sum_{k=0}^{\infty} \mu^*(\boldsymbol{E}^k).$$

We prove **(e)**, which follows from the sub-additivity. Indeed, by setting $\boldsymbol{E}^k = \varnothing$ for all $k \in \mathbb{N}$ such that $k > m$ we find

$$\mu_n^*(\boldsymbol{E}) \leq \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{E}^k) = \sum_{k=0}^{m} \mu_n^*(\boldsymbol{E}^k).$$

We conclude by proving **(f)**. For any n-dimensional interval $\boldsymbol{I}$ we have $\mathrm{vol}_n(\boldsymbol{I}) = \mathrm{vol}_n(\boldsymbol{I} + \{\boldsymbol{x}\})$ for every $\boldsymbol{x} \in \mathbb{R}^n$. For example, if $\boldsymbol{I} = [a_1, b_1) \times \ldots \times [a_n, b_n)$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, then $\boldsymbol{x} + \boldsymbol{I} = [a_1 + x_1, b_1 + x_1) \times \ldots \times [a_n + x_n, b_n + x_n)$, and

$$\mathrm{vol}_n(\boldsymbol{x} + \boldsymbol{I}) = (b_1 + x_1 - a_1 - x_1) \cdot \ldots \cdot (b_n + x_n - a_n - x_n) = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n) = \mathrm{vol}_n(\boldsymbol{I}).$$

Moreover, $(\boldsymbol{I}^k + \boldsymbol{x})_{k\in\mathbb{N}}$ is a cover of $\boldsymbol{x} + \boldsymbol{E}$ if and only if $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ is a cover of $\boldsymbol{E}$ (see Exercise 3.15).

Exercise 3.14 Show that for all $\boldsymbol{x} \in \mathbb{R}^n$ there holds $\mu_n^*(\{\boldsymbol{x}\}) = 0$.

For ecery $\varepsilon > 0$, we have $\{\boldsymbol{x}\} \supseteq \prod_{i=1}^{m} (x_i - \varepsilon, x_i + \varepsilon)$. The n-dimensional cube in the right hand-side has measure $(2\,\varepsilon)^n$, and so

$$\mu_n^*(\{\boldsymbol{x}\}) \leq (2\,\varepsilon)^n,$$

and $(2\,\varepsilon)^n$ can be made arbitrarily small.

With the definition of outer measure given in this section, we have constructed a function

$$\begin{aligned}
\mu^* : \mathcal{P}(\mathbb{R}^n) &\longrightarrow [0, +\infty] \\
\boldsymbol{E} &\mapsto \mu^*(\boldsymbol{E}),
\end{aligned}$$

which is also called, by abuse of nomenclature, *Lebesgue outer measure*.

We now present a result that seems geometrically intuitive at this point, but whose proof is not entirely trivial: the n-dimensional outer measure of an interval of $\mathbb{R}^n$ is equal to its volume.

Theorem 3.15 [Outer Measure of an n-Dimensional Interval]
Let $\boldsymbol{I}$ be an n-dimensional interval of $\mathbb{R}^n$. Then,

$$\mu_n^*(\boldsymbol{I}) = \mathrm{vol}_n(\boldsymbol{I}).$$

Proof Choosing $\boldsymbol{I}^0 = \boldsymbol{I}$ and $\boldsymbol{I}^k = \varnothing$ for all $k > 0$, we have

$$\mu_n^*(\boldsymbol{I}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) = \mathrm{vol}_n(\boldsymbol{I}^0) + \sum_{k=1}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) = \mathrm{vol}_n(\boldsymbol{I}) + 0 = \mathrm{vol}_n(\boldsymbol{I}).$$

We prove the opposite inequality, $\mathrm{vol}_n(\boldsymbol{I}) \leq \mu_n^*(\boldsymbol{I})$, by induction over n. We start from $n = 1$: we have to show that $\ell(I) \leq \mu_1^*(I)$. This is obvious if $\ell(I) = 0$. Assume $\ell(I) > 0$, and let $(I^k)_{k\in\mathbb{N}}$ be a family of bounded intervals such that $I \subseteq \bigcup_{k=0}^{\infty} I^k$. If we can show that

$$\ell(I) \leq \sum_{k=0}^{\infty} \ell(I^k),$$

then $\ell(I)$ is a lower bound for the set $\Big\{ \sum_{k\in\mathbb{N}} \ell(I^k) \,\Big|\, (I^k)_{k\in\mathbb{N}} \text{ is a family of 1-dimensional intervals and } I \subseteq \bigcup_{k\in\mathbb{N}} I^k \Big\}$. Since $\mu_1^*(I)$ is the maximum lower bound of the same set, we can conclude that $\ell(I) \leq \mu_1^*(I)$ as required.

Let $\varepsilon > 0$ be such that $2\,\varepsilon < \ell(I)$. Let $I^\varepsilon = [\inf I + \varepsilon, \sup I - \varepsilon]$ and, for all $k \in \mathbb{N}$, let

$$I^{k,\varepsilon} = \left(\inf I^k - \frac{\varepsilon}{2^{k+2}}, \sup I^k + \frac{\varepsilon}{2^{k+2}}\right).$$

We observe that I^ε is a compact set contained in I, while for all $k \in \mathbb{N}$ the set $I^{k,\varepsilon}$ is an open interval and it contains I^k. Thus, $I^\varepsilon \subseteq I \subseteq \bigcup_{k=0}^{\infty} I^k \subseteq \bigcup_{k=0}^{\infty} I^{k,\varepsilon}$, and this implies $I^\varepsilon \subseteq \bigcup_{k=0}^{\infty} I^{k,\varepsilon}$. We recall that if we have a family $(U^k)_{k\in\mathbb{N}}$ of open sets which is a covering for a set K, if there exists a finite subset H of $\mathbb{N}$ such that $(U^k)_{k\in H}$ is also a covering for K, then K is compact, see Definition 2.5; since we are working in $\mathbb{R}^n$, a compact set is also sequentially compact, and by the Heine-Borel theorem, it is bounded and closed. From these considerations, there exists $m \in \mathbb{N}^*$ such that $I^\varepsilon \subseteq \bigcup_{k=0}^{m} I^{k,\varepsilon}$. Hence, for all $x \in \mathbb{R}$ there holds $\chi_{I^\varepsilon}(x) \leq \sum_{k=0}^{m} \chi_{I^{k,\varepsilon}}(x)$. Applying **(a)**

in Theorem 3.12 we find

$$\begin{aligned}
\ell(I^\varepsilon) &= \int_{-\infty}^{+\infty} \chi_{I^\varepsilon}(x)\,\mathrm{d}x \\
&\leq \int_{-\infty}^{+\infty} \sum_{k=0}^{m} \chi_{I^{k,\varepsilon}}(x)\,\mathrm{d}x \\
&= \sum_{k=0}^{m} \int_{-\infty}^{+\infty} \chi_{I^{k,\varepsilon}}(x)\,\mathrm{d}x \\
&= \sum_{k=0}^{m} \ell(I^{k,\varepsilon}) = \sum_{k=0}^{m} \left(\sup I^k + \frac{\varepsilon}{2^{k+2}} - \inf I^k + \frac{\varepsilon}{2^{k+2}}\right) \\
&= \sum_{k=0}^{m} \left(\ell(I^k) + \frac{\varepsilon}{2^{k+1}}\right) = \sum_{k=0}^{m} \ell(I^k) + \frac{\varepsilon}{2} \sum_{k=0}^{m} \left(\frac{1}{2}\right)^k \\
&= \sum_{k=0}^{m} \ell(I^k) + \frac{\varepsilon}{2} \frac{1 - \left(\frac{1}{2}\right)^{m+1}}{1 - \frac{1}{2}} = \sum_{k=0}^{m} \ell(I^k) + \varepsilon \left(1 - \frac{1}{2^{m+1}}\right) \\
&\leq \sum_{k=0}^{m} \ell(I^k) + \varepsilon \leq \sum_{k=0}^{\infty} \ell(I^k) + \varepsilon,
\end{aligned}$$

where we used the fact that, given $q \in \mathbb{C}$ satisfying $q \neq 1$ and $n \in \mathbb{N}$, there holds $\displaystyle\sum_{i=0}^{n} q^i = \frac{1 - q^{n+1}}{1 - q}$. Since $\ell(I^\varepsilon) = \ell(I) - 2\,\varepsilon$, we obtain

$$\ell(I) \leq \sum_{k=0}^{\infty} \ell(I^k) + 3\,\varepsilon.$$

Since ε can be chosen to be arbitrarily small, we must have

$$\ell(I) \leq \sum_{k=0}^{\infty} \ell(I^k).$$

We conclude that $\ell(I) \leq \mu_1^*(I)$, and this completes the case $n = 1$.

Now, we assume that the result is true for $n \geq 1$. Let $\boldsymbol{I} = \displaystyle\prod_{j=1}^{n+1} I_j$, where $I_1, \ldots, I_{n+1}$ are bounded intervals of $\mathbb{R}$. If $\ell(I_j) = 0$ for some j, then $\mathrm{vol}_{n+1}(\boldsymbol{I}) = 0$, and therefore $\mathrm{vol}_{n+1}(\boldsymbol{I}) \leq \mu_{n+1}^*(\boldsymbol{I})$. Let us assume $\ell(I_j) > 0$ for every $j \in \{1, \ldots, n+1\}$. Let $\varepsilon > 0$ be such that $2\,\varepsilon < \min\limits_{1 \leq j \leq n+1} \ell(I_j)$ and we define

$$\boldsymbol{I}^\varepsilon = \prod_{j=1}^{n+1} \left[\inf I_j + \varepsilon, \sup I_j - \varepsilon\right].$$

The set $\boldsymbol{I}^\varepsilon$ is compact and it is contained in $\boldsymbol{I}$. Moreover,

$$\lim_{\varepsilon \to 0} \mathrm{vol}_{n+1}(\boldsymbol{I}^\varepsilon) = \mathrm{vol}_{n+1}(\boldsymbol{I}).$$

Let $(\boldsymbol{I}^k)_{k\in\mathbb{N}}$ be a family of $(n+1)$-dimensional intervals such that $\boldsymbol{I} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$. We have to show that

$$\mathrm{vol}_{n+1}(\boldsymbol{I}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_{n+1}(\boldsymbol{I}^k),$$

and then we can conclude the proof as in the one-dimensional case. For every $k \in \mathbb{N}$, if $\boldsymbol{I}^k = \prod_{j=1}^{n+1} I_j^k$, we define

$$\boldsymbol{I}^{k,\varepsilon} = \prod_{j=1}^{n+1} \left(\inf I_j^k - \varepsilon_k, \sup I_j^k + \varepsilon_k\right),$$

where $\varepsilon_k > 0$ is such that

$$\mathrm{vol}_{n+1}(\boldsymbol{I}^{k,\varepsilon}) \leq \mathrm{vol}_{n+1}(\boldsymbol{I}^k) + \frac{\varepsilon}{2^{k+1}}.$$

The $(n+1)$-dimensional intervals $\boldsymbol{I}^{k,\varepsilon}$ are open sets, so that, from $\boldsymbol{I}^{\varepsilon} \subseteq \boldsymbol{I} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^{k,\varepsilon}$, it follows that there exists $m \in \mathbb{N}$ such that $\boldsymbol{I}^{\varepsilon} \subseteq \bigcup_{k=0}^{m} \boldsymbol{I}^{k,\varepsilon}$.

We now adopt the notation in (3.13), with $n+1$ instead of n. For every $x_{n+1} \in \mathbb{R}$, we have $\boldsymbol{I}^{\varepsilon}_{x_{n+1}} \subseteq \bigcup_{k=0}^{m} \boldsymbol{I}^{k,\varepsilon}_{x_{n+1}}$, and from the inductive assumption, since $\boldsymbol{I}^{k,\varepsilon}_{x_{n+1}}$ is an n-dimensional interval (from **(a)** in Theorem 3.13), we find

$$\mathrm{vol}_n(\boldsymbol{I}^{\varepsilon}_{x_{n+1}}) = \mu_n^*(\boldsymbol{I}^{\varepsilon}_{x_{n+1}}) \leq \sum_{k=0}^{m} \mu_n^*(\boldsymbol{I}^{k,\varepsilon}_{x_{n+1}}) = \sum_{k=0}^{m} \mathrm{vol}_n(\boldsymbol{I}^{\varepsilon}_{x_{n+1}}).$$

From (3.17) we find

$$\begin{aligned}
\mathrm{vol}_{n+1}(\boldsymbol{I}^{\varepsilon}) &= \int_{-\infty}^{+\infty} \mathrm{vol}_n(\boldsymbol{I}^{\varepsilon}_{x_{n+1}})\,\mathrm{d}x_{n+1} \\
&\leq \sum_{k=0}^{m} \int_{-\infty}^{+\infty} \mathrm{vol}_n(\boldsymbol{I}^{k,\varepsilon}_{x_{n+1}})\,\mathrm{d}x_{n+1} \\
&= \sum_{k=0}^{m} \mathrm{vol}_{n+1}(\boldsymbol{I}^{k,\varepsilon}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_{n+1}(\boldsymbol{I}^{k,\varepsilon}) \\
&\leq \sum_{k=0}^{\infty} \left(\mathrm{vol}_{n+1}(\boldsymbol{I}^k) + \frac{\varepsilon}{2^{k+1}}\right) \\
&= \sum_{k=0}^{\infty} \mathrm{vol}_{n+1}(\boldsymbol{I}^k) + \frac{\varepsilon}{2} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k \\
&= \sum_{k=0}^{\infty} \mathrm{vol}_{n+1}(\boldsymbol{I}^k) + \frac{\varepsilon}{2}\,\frac{1}{1-\frac{1}{2}} \\
&= \sum_{k=0}^{\infty} \mathrm{vol}_{n+1}(\boldsymbol{I}^k) + \varepsilon.
\end{aligned}$$

Letting ε go to 0, we obtain the result.

Example *3.9*

Notice that, in view of Theorem 3.15, the 1-dimensional outer measure of the interval $[0,1]$ in $\mathbb{R}$ is 1. Indeed,

$$\mu_1^*([0,1]) = \mathrm{vol}_1([0,1]) = \ell([0,1]) = 1.$$

The outer measure of the set $\boldsymbol{I} = \{(x,0) \in \mathbb{R}^2 \,|\, 0 \le x \le 1\}$ in $\mathbb{R}^2$ is 0. In fact, we can write this set as $\boldsymbol{I} = [0,1] \times \{0\}$, and therefore

$$\mu_2^*([0,1]\times\{0\}) = \mathrm{vol}_2([0,1]\times\{0\}) = \ell([0,1]) \cdot \ell(\{0\}) = 1 \cdot 0 = 0.$$

From the previous example, we can say that, intuitively, the outer measure is not a property of the sole geometric nature of the set, but also of the space where the set lives: a segment in $\mathbb{R}$ of strictly positive length is always a zero-measure set in $\mathbb{R}^2$, $\mathbb{R}^3$, and so on. A square in $\mathbb{R}^2$ is a zero-measure set in $\mathbb{R}^3$.

Example *3.10*

We want to compute the 1-dimensional outer measure of $\mathbb{R}$. Since $(-M, M) \subset \mathbb{R}$ for all $M > 0$, from the monotonicity of the outer measure and from Theorem 3.15 we have

$$\mu_1^*(\mathbb{R}) \ge \mu_1^*((-M,M)) = \ell((-M,M)) = 2\,M.$$

Since M is arbitrary, we obtain $\mu_1^*(\mathbb{R}) = +\infty$.

However, the 2-dimensional outer measure of the x-axis is 0. Indeed, the x-axis can be written as $\mathbb{R}\times\{0\}$, and

$$\mathbb{R}\times\{0\} = \bigcup_{k=0}^{\infty} [-k,k]\times\{0\}.$$

For each $k \in \mathbb{N}$, the set $[-k,k]$ is a bounded interval of $\mathbb{R}$, and so is $\{0\}$. We can therefore use the countable sub-additivity:

$$\begin{aligned}
\mu_2^*(\mathbb{R}\times\{0\}) &\le \sum_{k=0}^{\infty} \mu_2^*([-k,k]\times\{0\}) \\
&= \sum_{k=0}^{\infty} \mathrm{vol}_2([-k,k]\times\{0\}) \\
&= \sum_{k=0}^{\infty} \ell([-k,k]) \cdot \ell(\{0\}) = \sum_{k=0}^{\infty} (2\,k)\cdot 0 = 0.
\end{aligned}$$

Example *3.11*

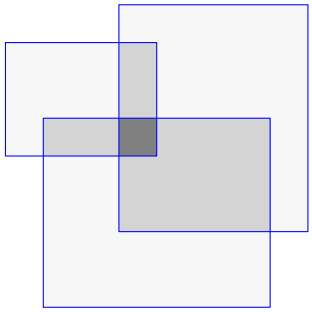
We want to compute the 1-dimensional outer measure of the set of irrational numbers $\mathbb{R} \setminus \mathbb{Q}$. From the sub-additivity, and recalling that the outer measure of a countable set is zero, we find

$$\underbrace{\mu_1^*(\mathbb{R})}_{=+\infty} \le \mu_1^*(\mathbb{R}\setminus\mathbb{Q}) + \underbrace{\mu_1^*(\mathbb{Q})}_{=0},$$

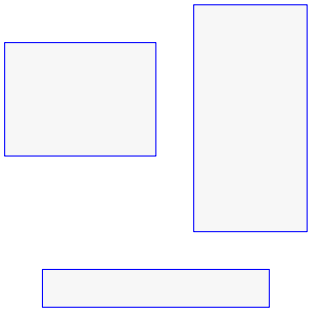
which implies $\mu_1^*(\mathbb{R}\setminus\mathbb{Q}) = +\infty$. Similarly, to find the outer measure of $[0,1]\setminus\mathbb{Q}$, we compute

$$\underbrace{\mu_1^*([0,1])}_{=1} \le \mu_1^*([0,1]\setminus\mathbb{Q}) + \underbrace{\mu_1^*([0,1]\cap\mathbb{Q})}_{=0} \quad\Rightarrow\quad \mu_1^*([0,1]\setminus\mathbb{Q}) = 1.$$

The outer measure μ^* is, to some extent, unsatisfactory. In fact, it does not satisfy the so-called additivity property, but only the subadditivity. We recall that the subadditivity property says that $\boldsymbol{E} = \bigcup_{k=1}^{\infty} \boldsymbol{E}_k$ implies $\mu^*(\boldsymbol{E}) \le \sum_{k=1}^{\infty} \mu^*(\boldsymbol{E}_k)$. The reason why we have $\le$ is because there could be overlaps in the subsets $\boldsymbol{E}_k$:



$\boldsymbol{E}_k$ overlapping



pairwise disjoint $\boldsymbol{E}_k$

However, if the sets $\boldsymbol{E}_k$ are all pairwise disjoint, and $\boldsymbol{E} = \bigcup_{k=1}^{\infty} \boldsymbol{E}_k$, we would like the equality $\mu^*(\boldsymbol{E}) = \sum_{k=1}^{\infty} \mu^*(\boldsymbol{E}_k)$ to hold. Unfortunately, the following results show that the equality does not hold in general. To construct counterexamples, one must use the so-called *axiom of choice*. Loosely speaking, the axiom of choice postulates that given any collection of sets (finite or infinite), each containing at least one element, it is possible to make a selection of exactly one element from each set. Even if most mathematicians nowadays accept the axiom of choice, it is possible to construct mathematical systems without this axiom. In the present case, the axiom of choice is essential in the proof of this result, in the sense that for a mathematical system where such axiom is not assumed, the outer measure is additive.

Theorem 3.16 [Vitali Covering Theorem]
Let J be a countable set. There exists a family $(E_k)_{k\in J}$ of disjoint subsets of $\mathbb{R}$ such that

$$\mu_1^*\Big(\bigcup_{k\in J} E_k\Big) \neq \sum_{k\in J} \mu_1^*(E_k).$$

Proof A set $A \subseteq \mathbb{R}$ is said to be a *coset* of $\mathbb{Q}$ if it is in the form $A = x+\mathbb{Q}$ for some $x \in \mathbb{R}$. For example, $\sqrt{2}+\mathbb{Q}$ is a coset of $\mathbb{Q}$. The same coset A can correspond to the same values of x; for example, $2 + \mathbb{Q}$ is the same as $0 + \mathbb{Q} = \mathbb{Q}$. Moreover, cosets are either identical or disjoint: if $x + \mathbb{Q}$ intersects $y + \mathbb{Q}$ at a point z, then $x - y$ is rational (if $z \in x+\mathbb{Q}$, we can write $z = x+q_1$ with $q_1 \in \mathbb{Q}$, and if $z \in y+\mathbb{Q}$, we can write $z = y+q_2$ with $q_2 \in \mathbb{Q}$, and subtracting we find $x - y = q_1 - q_2 \in \mathbb{Q}$). This implies that $x + \mathbb{Q} = y + \mathbb{Q}$. In fact, given $\xi \in x + \mathbb{Q}$, we can write $\xi = x+q$ with $q \in \mathbb{Q}$, so that $\xi - x = q$, i.e., $\xi - x + x - y = q + x - y$, i.e., $\xi - y = q + x - y \in \mathbb{Q}$, i.e., $\xi \in y+\mathbb{Q}$.

Every coset A of $\mathbb{Q}$ has a non-empty intersection with $[0, 1]$. In fact, if A is a coset of $\mathbb{Q}$, then there exists $x \in \mathbb{R}$ such that $A = x + \mathbb{Q}$. Choosing a rational number q in $[-x, 1 - x]$, we obtain $x + q \in [0, 1]$, and therefore $A \cap [0, 1]$ contains $x + q$.

The set of all cosets of $\mathbb{Q}$ is denoted by $\mathbb{R}/\mathbb{Q}$, and it is referred to as *quotient*. For each coset A in $\mathbb{R}/\mathbb{Q}$, we use the axiom of choice to choose an element ξ_A of $A \cap [0, 1]$, and we denote by E the set of all these ξ_A:

$$E = \{\xi_A \,|\, A \in \mathbb{R}/\mathbb{Q}\},$$

and clearly $E \subseteq [0, 1]$. Let us define the set (also known as *Vitali set*, named after the mathematician Giuseppe Vitali)

$$X = \bigcup_{q\in\mathbb{Q}\cap[-1,1]} (q + E).$$

Since $q + \xi \in [-1, 2]$ for $q \in [-1, 1]$ and $\xi \in E \subseteq [0, 1]$, we obtain $X \subseteq [-1, 2]$. We prove that $X \supseteq [0, 1]$. For any $y \in [0, 1]$, y must belong to a coset A (for example, $y \in y + \mathbb{Q}$), and ξ_A belongs to the same coset, so that

$y - \xi_A$ is equal to a rational number q. Since $y, \xi_A \in [0,1]$, we have $q \in [-1,1]$. Moreover, since $y = q + \xi_A$, we find $y \in q + E$, and therefore $y \in X$.

To conclude the proof, we show that

$$\mu_1^*(X) \neq \sum_{q \in \mathbb{Q} \cap [-1,1]} \mu_1^*(q + E). \tag{3.21}$$

From $[0,1] \subseteq X \subseteq [-1,2]$, we obtain $1 \leq \mu_1^*(X) \leq 3$ by the monotonicity of the outer measure and from Theorem 3.14. Moreover, from the translation invariance

$$\sum_{q \in \mathbb{Q} \cap [-1,1]} \mu_1^*(q + E) = \sum_{q \in \mathbb{Q} \cap [-1,1]} \mu_1^*(E).$$

The set $\mathbb{Q} \cap [-1,1]$ is countable and not finite, so that the right-hand side of (3.21) is zero if $\mu_1^*(E) = 0$ or $+\infty$ if $\mu_1^*(E) > 0$. In any of these two situations, it cannot be between 1 and 3.

It is possible to strengthen the result of Theorem 3.16, by showing that the same result holds for finite collections of sets.

Theorem 3.17 *Let J be finite. There exists a family $(E_k)_{k \in J}$ of disjoint subsets of $\mathbb{R}$ such that*

$$\mu_1^*\Big(\bigcup_{k \in J} E_k\Big) \neq \sum_{k \in J} \mu_1^*(E_k).$$

Hence, while it is desirable that for two arbitrary disjoint sets $\boldsymbol{E}$ and $\boldsymbol{F}$ in $\mathbb{R}^n$ we had

$$\mu_n^*(\boldsymbol{E} \cup \boldsymbol{F}) = \mu_n^*(\boldsymbol{E}) + \mu_n^*(\boldsymbol{F}),$$

unfortunately, this is not true in general. We can only say that the inequality

$$\mu_n^*(\boldsymbol{E} \cup \boldsymbol{F}) \leq \mu_n^*(\boldsymbol{E}) + \mu_n^*(\boldsymbol{F})$$

is true, even when $\boldsymbol{E}$ and $\boldsymbol{F}$ are pairwise disjoint.

Exercises

Exercise 3.15 Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. Show that $(\boldsymbol{x} + \boldsymbol{I}^k)_{k \in \mathbb{N}}$ is a cover of $\boldsymbol{x} + \boldsymbol{E}$ if and only if $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ is a cover of $\boldsymbol{E}$. [Solution: (3.13)]

Exercise 3.16 Show that the 2-dimensional outer measure of $\mathbb{R}^2$ is $+\infty$. [Solution: (3.14)]

Exercise 3.17 Show that the 3-dimensional outer measure of the xy-plane is 0. [Solution: (3.15)]

3.4 Lebesgue measure

To remedy the problem described in the previous section, whereby it is possible to construct (rather artificial) sets for which the countable additivity does not hold, we 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 3.8 [Lebesgue Measurable Set]
Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. We say that $\boldsymbol{E}$ is measurable *or* Lebesgue measurable *if*

$$\forall\, \boldsymbol{A} \subseteq \mathbb{R}^n, \qquad \mu_n^*(\boldsymbol{A}) = \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement}). \tag{3.22}$$

We denote by $\mathcal{M}_n$ the set of subsets of $\mathbb{R}^n$ which are Lebesgue measurable.

The identity (3.22), known as the *Carathéodory condition*, says that we can use $\boldsymbol{E}$ to divide an arbitrary set $\boldsymbol{A}$ into two disjoint parts, $\boldsymbol{A} \cap \boldsymbol{E}$ and $\boldsymbol{A} \cap \boldsymbol{E}^{\complement}$, whose outer measures add up to the outer measure of $\boldsymbol{A}$. In other words, the Carathéodory condition means that a set is measurable if it divides all other sets in a way that the additive property is maintained. Clearly, if the outer measure μ^* were finitely additive, then (3.22) would automatically hold for any subset $\boldsymbol{A}$ of $\mathbb{R}^n$, and therefore any set would be Lebesgue measurable. In this sense, we can think of Lebesgue measurable sets as those sets for which finite additivity holds.

Since μ^* is sub-additive, using **(d)** of Theorem 3.14 we always have

$$\forall\, \boldsymbol{A} \subseteq \mathbb{R}^n, \qquad \mu_n^*(\boldsymbol{A}) \leq \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement}).$$

Therefore, to prove that $\boldsymbol{E} \subseteq \mathbb{R}^n$ is measurable, it is sufficient to show that the opposite inequality holds, i.e.,

$$\mu_n^*(\boldsymbol{A}) \geq \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement}) \tag{3.23}$$

for every $\boldsymbol{A} \subseteq \mathbb{R}^n$.

The class $\mathcal{M}_n$ is very flexible, with respect to the usual operations on sets, as the following result shows.

Theorem 3.18 [Carathéodory's Theorem]
For every $n \in \mathbb{N}^*$, the set $\mathcal{M}_n$ is a σ-algebra in $\mathbb{R}^n$ and μ_n^* is a measure.

Proof We prove this theorem through a number of intermediate steps.

(a) $\varnothing \in \mathcal{M}_n$. The proof is left as an exercise (see Exercise 3.18);

(b) If $\boldsymbol{E} \in \mathcal{M}_n$, then $\boldsymbol{E}^{\complement} \in \mathcal{M}_n$. The proof is left as an exercise (see Exercise 3.19);

(c) If $\boldsymbol{E}_1, \boldsymbol{E}_2 \in \mathcal{M}_n$, then $\boldsymbol{E}_1 \cup \boldsymbol{E}_2 \in \mathcal{M}_n$. Indeed, let $\boldsymbol{A} \subseteq \mathbb{R}^n$. Then

$$\begin{aligned}
\mu_n^*(\boldsymbol{A} \cap (\boldsymbol{E}_1 \cup \boldsymbol{E}_2)) + \mu_n^*(\boldsymbol{A} \cap (\boldsymbol{E}_1 \cup \boldsymbol{E}_2)^{\complement}) &= \mu_n^*\Big(\boldsymbol{A} \cap (\boldsymbol{E}_1 \cup (\boldsymbol{E}_2 \setminus \boldsymbol{E}_1))\Big) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2^{\complement}) \\
&= \mu_n^*\Big((\boldsymbol{A} \cap \boldsymbol{E}_1) \cup (\boldsymbol{A} \cap (\boldsymbol{E}_2 \cap \boldsymbol{E}_1^{\complement}))\Big) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2^{\complement}) \\
&= \mu_n^*((\boldsymbol{A} \cap \boldsymbol{E}_1) \cup (\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2)) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2^{\complement}) \\
&\leq \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement} \cap \boldsymbol{E}_2^{\complement}) \\
&= \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1^{\complement}) = \mu_n^*(\boldsymbol{A}),
\end{aligned}$$

where, in the last line, we used the fact that $\boldsymbol{E}_2 \in \mathcal{M}_n$;

(d) for every $m \in \mathbb{N}$, if $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_m \in \mathcal{M}_n$, then $\bigcup_{k=0}^{m} \boldsymbol{E}_k \in \mathcal{M}_n$; this follows immediately from **(c)** using an induction argument;

(e) If $\boldsymbol{E}, \boldsymbol{F} \in \mathcal{M}_n$, then $\boldsymbol{E} \cap \boldsymbol{F}$ and $\boldsymbol{E} \setminus \boldsymbol{F}$ are elements of $\mathcal{M}_n$. Indeed, $\boldsymbol{E} \cap \boldsymbol{F} = (\boldsymbol{E}^{\complement} \cup \boldsymbol{F}^{\complement})^{\complement}$, and $\boldsymbol{E} \cap \boldsymbol{F} \in \mathcal{M}_n$ follows from **(b)** and **(c)**. We also have $\boldsymbol{E} \setminus \boldsymbol{F} \in \mathcal{M}_n$, since $\boldsymbol{E} \setminus \boldsymbol{F} = \boldsymbol{E} \cap \boldsymbol{F}^{\complement}$;

(f) For every $\boldsymbol{A} \subseteq \mathbb{R}^n$ and for all $m \in \mathbb{N}$, if $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_m$ are pairwise disjoint elements of $\mathcal{M}_n$, then

$$\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m} \boldsymbol{E}_k\Big) = \sum_{k=0}^{m} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k).$$

We prove this statement by induction starting from $m = 1$. Since $\boldsymbol{E}_0 \in \mathcal{M}_n$, then

$$\begin{aligned}\mu_n^*(\boldsymbol{A} \cap (\boldsymbol{E}_0 \cup \boldsymbol{E}_1)) &= \mu_n^*(\boldsymbol{A} \cap (\boldsymbol{E}_0 \cup \boldsymbol{E}_1) \cap \boldsymbol{E}_0) + \mu_n^*(\boldsymbol{A} \cap (\boldsymbol{E}_0 \cup \boldsymbol{E}_1) \cap \boldsymbol{E}_0^{\complement}) \\ &= \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_0) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_1),\end{aligned}$$

since $(\boldsymbol{E}_0 \cup \boldsymbol{E}_1) \cap \boldsymbol{E}_0 = \boldsymbol{E}_0$ and $(\boldsymbol{E}_0 \cup \boldsymbol{E}_1) \cap \boldsymbol{E}_0^{\complement} = (\boldsymbol{E}_0 \cup \boldsymbol{E}_1) \setminus \boldsymbol{E}_0 = \boldsymbol{E}_1$, which takes into account the fact that the sets $\boldsymbol{E}_0$ and $\boldsymbol{E}_1$ are disjoint. Now, suppose that the result is true for a certain m. Then,

$$\begin{aligned}\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m+1} \boldsymbol{E}_k\Big) &= \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m+1} \boldsymbol{E}_k \cap \boldsymbol{E}_{m+1}\Big) + \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m+1} \boldsymbol{E}_k \cap \boldsymbol{E}_{m+1}^{\complement}\Big) \\ &= \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_{m+1}) + \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m} \boldsymbol{E}_k\Big) = \sum_{k=0}^{m+1} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k),\end{aligned}$$

where we used the fact that

$$\bigcup_{k=0}^{m+1} \boldsymbol{E}_k \cap \boldsymbol{E}_{m+1} = \boldsymbol{E}_{m+1} \quad \text{and} \quad \bigcup_{k=0}^{m+1} \boldsymbol{E}_k \cap \boldsymbol{E}_{m+1}^{\complement} = \bigcup_{k=0}^{m} \boldsymbol{E}_k$$

as well as the inductive assumption;

(g) For every $\boldsymbol{A} \subseteq \mathbb{R}^n$, if $(\boldsymbol{E}_k)_{k\in\mathbb{N}}$ is a family of pairwise disjoint subsets belonging to $\mathcal{M}_n$, we want to show that

$$\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) = \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k).$$

We first show the inequality $\leq$. Using **(c)** in Theorem 3.14 we find

$$\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) = \mu_n^*\Big(\bigcup_{k=0}^{\infty} (\boldsymbol{A} \cap \boldsymbol{E}_k)\Big) \leq \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k).$$

To show the opposite inequality, we observe that for every $m \in \mathbb{N}$, from **(b)** in Theorem 3.14 and from **(f)** we find

$$\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) \geq \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m} \boldsymbol{E}_k\Big) = \sum_{k=0}^{m} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k).$$

Taking the limit as $m \to \infty$, we obtain

$$\mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) \geq \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k).$$

(h) if $(\boldsymbol{E}_k)_{k\in\mathbb{N}}$ is a family of pairwise disjoint subsets belonging to $\mathcal{M}_n$, then

$$\mu_n^*\Big(\bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) = \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{E}_k).$$

This follows from **(g)** taking $A = \mathbb{R}^n$.

(i) if $(\boldsymbol{E}_k)_{k\in\mathbb{N}}$ is a family of pairwise disjoint subsets belonging to $\mathcal{M}_n$, then $\bigcup_{k=0}^{\infty} \boldsymbol{E}_k \in \mathcal{M}_n$. To prove this, let $\boldsymbol{A} \subseteq \mathbb{R}^n$ and $m \in \mathbb{N}$. We have

$$\mu_n^*\left(\boldsymbol{A} \cap \Big(\bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big)^{\complement}\right) = \mu_n^*\Big(\boldsymbol{A} \cap \bigcap_{k=0}^{\infty} \boldsymbol{E}_k^{\complement}\Big) \leq \mu_n^*\Big(\boldsymbol{A} \cap \bigcap_{k=0}^{m} \boldsymbol{E}_k^{\complement}\Big).$$

From **(d)** and **(f)** we have

$$\begin{aligned}
\mu_n^*(\boldsymbol{A}) &= \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{m} \boldsymbol{E}_k\Big) + \mu_n^*\left(\boldsymbol{A} \cap \Big(\bigcup_{k=0}^{m} \boldsymbol{E}_k\Big)^{\complement}\right) \\
&= \sum_{k=0}^{m} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k) + \mu_n^*\Big(\boldsymbol{A} \cap \bigcap_{k=0}^{m} \boldsymbol{E}_k^{\complement}\Big) \\
&\geq \sum_{k=0}^{m} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k) + \mu_n^*\Big(\boldsymbol{A} \cap \bigcap_{k=0}^{\infty} \boldsymbol{E}_k^{\complement}\Big).
\end{aligned}$$

Taking the limit as $m \to \infty$ and using **(g)**, we get

$$\begin{aligned}
\mu_n^*(\boldsymbol{A}) &\geq \sum_{k=0}^{\infty} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}_k) + \mu_n^*\Big(\boldsymbol{A} \cap \bigcap_{k=0}^{\infty} \boldsymbol{E}_k^{\complement}\Big) \\
&= \mu_n^*\Big(\boldsymbol{A} \cap \bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big) + \mu_n^*\left(\boldsymbol{A} \cap \Big(\bigcup_{k=0}^{\infty} \boldsymbol{E}_k\Big)^{\complement}\right).
\end{aligned}$$

(l) Let $(\boldsymbol{E}_k)_{k\in\mathbb{N}}$ be a family of $\mathcal{M}_n$. Then, $\bigcup_{k=0}^{\infty} \boldsymbol{E}_k \in \mathcal{M}_n$. As in the proof of **(d)** in Theorem 3.10, we define

$$\boldsymbol{F}_k = \begin{cases} \boldsymbol{E}_0 & \text{if } k = 0 \\ \boldsymbol{E}_k \setminus \bigcup_{j=0}^{k-1} \boldsymbol{E}_j & \text{if } k \geq 1. \end{cases}$$

For every $k \in \mathbb{N}$, from **(d)** and **(e)** we have $\boldsymbol{F}_k \in \mathcal{M}_n$. Moreover, $\boldsymbol{F}_k$ are pairwise disjoint. Applying **(i)** we find

$$\bigcup_{k=0}^{\infty} \boldsymbol{E}_k = \bigcup_{k=0}^{\infty} \boldsymbol{F}_k \in \mathcal{M}_n.$$

This completes the proof.

Exercise 3.18 Prove point **(a)** in the proof of Theorem 3.18. [Solution: (3.16)]

Exercise 3.19 Prove point **(b)** in the proof of Theorem 3.18. [Solution: (3.17)]

Exercise 3.20 Let $\boldsymbol{x}_0 \in \mathbb{R}^n$. Show that $\{\boldsymbol{x}_0\}$ is a Lebesgue measurable set. [Solution: (3.18)]

The problem now is to determine sufficient conditions for a subset $\boldsymbol{E}$ of $\mathbb{R}^n$ to be in $\mathcal{M}_n$. We start with the following result, which is simple but useful.

Theorem 3.19 *Let $\boldsymbol{E} \subseteq \mathbb{R}^n$ be such that $\mu_n^*(\boldsymbol{E}) = 0$. Then, $\boldsymbol{E} \in \mathcal{M}_n$.*

Proof For every $\boldsymbol{A} \subseteq \mathbb{R}^n$, we have

$$0 \leq \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) \leq \mu_n^*(\boldsymbol{E}) = 0.$$

It follows that $\mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) = 0$. On the other hand, $\mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^\complement) \leq \mu_n^*(\boldsymbol{A})$. Thus,

$$\mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^\complement) \leq \mu_n^*(\boldsymbol{A}).$$

If $\boldsymbol{E}$ is measurable, we define the *Lebesgue measure* of $\boldsymbol{E}$ to be $\mu_n(\boldsymbol{E}) \stackrel{\text{def}}{=} \mu_n^*(\boldsymbol{E})$.

Definition 3.9 [Lebesgue Measure]
The Lebesgue measure μ is the restriction of the Lebesgue outer measure $\mu_n^* : \mathcal{P}(\mathbb{R}^n) \longrightarrow [0, +\infty]$ to $\mathcal{M}_n$, i.e., $\mu = \mu^*|_{\mathcal{M}_n}$.

As a result of this definition, we can write

$$\begin{aligned} \mu : \mathcal{M}_n &\longrightarrow [0, +\infty] \\ \boldsymbol{E} &\mapsto \mu^*(\boldsymbol{E}). \end{aligned}$$

Theorem 3.20 *Let $\boldsymbol{I}$ be an n-dimensional interval of $\mathbb{R}^n$. Then, $\boldsymbol{I} \in \mathcal{M}_n$ and*

$$\mu_n(\boldsymbol{I}) = \mathrm{vol}_n(\boldsymbol{I}).$$

Proof The second claim follows from the first and from Theorem 3.15. We prove that $\boldsymbol{I} \in \mathcal{M}_n$, i.e., we show that for every $\boldsymbol{A} \subseteq \mathbb{R}^n$ there holds

$$\mu_n^*(\boldsymbol{A} \cap \boldsymbol{I}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{I}^\complement) \leq \mu_n^*(\boldsymbol{A}).$$

This is obviously trivial if $\mu_n^*(\boldsymbol{A}) = +\infty$. Let us assume that $\mu_n^*(\boldsymbol{A}) < +\infty$. Let $\varepsilon > 0$, and let $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ be a family of n-dimensional intervals such that $\boldsymbol{A} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$ and $\sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \leq \mu_n^*(\boldsymbol{A}) + \varepsilon$. We have $\boldsymbol{A} \cap \boldsymbol{I} \subseteq \bigcup_{k=0}^{\infty} (\boldsymbol{I}^k \cap \boldsymbol{I})$. For every $k \in \mathbb{N}$, the set $\boldsymbol{I}^k \cap \boldsymbol{I}$ is an n-dimensional interval (see **(d)** in Theorem 3.13). Thus,

$$\mu_n^*(\boldsymbol{A} \cap \boldsymbol{I}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k \cap \boldsymbol{I}). \tag{3.24}$$

Moreover,

$$\boldsymbol{A} \cap \boldsymbol{I}^\complement \subseteq \bigcup_{k=0}^{\infty} (\boldsymbol{I}^k \setminus \boldsymbol{I}) = \bigcup_{k=0}^{\infty} \bigcup_{\ell=1}^{2n} \boldsymbol{I}^{k,\ell},$$

where $\boldsymbol{I}^{k,\ell}$ are n-dimensional pairwise disjoint intervals, applying **(c)** in Theorem 3.13. The set $\mathcal{I} = \mathbb{N} \times \{\ell \in \mathbb{N}^* \mid \ell \leq 2n\}$ is countable. Let $\varphi : \mathbb{N} \longrightarrow \mathcal{I}$ be a bijection. Then, $A \cap \boldsymbol{I}^\complement \subseteq \bigcup_{j=0}^{\infty} \boldsymbol{I}^{\varphi(j)}$. Applying Theorems 3.6, 3.5, 3.4, we find

$$\begin{aligned} \mu_n^*(\boldsymbol{A} \cap \boldsymbol{I}^\complement) &\leq \sum_{j=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{\varphi(r)}) \\ &= \sum_{(k,\ell) \in \mathcal{I}} \mathrm{vol}_n(\boldsymbol{I}^{k,\ell}) \\ &= \sum_{k \in \mathbb{N}} \sum_{\ell=1}^{2n} \mathrm{vol}_n(\boldsymbol{I}^{k,\ell}) = \sum_{k=0}^{\infty} \sum_{\ell=1}^{2n} \mathrm{vol}_n(\boldsymbol{I}^{k,\ell}). \end{aligned}$$

Hence,

$$\begin{aligned}\mu_n^*(\boldsymbol{A}\cap\boldsymbol{I})+\mu_n^*(\boldsymbol{A}\cap\boldsymbol{I}^\complement) &\leq \sum_{k=0}^{\infty}\mathrm{vol}_n(\boldsymbol{I}^k\cap\boldsymbol{I})+\sum_{k=0}^{\infty}\sum_{\ell=1}^{2n}\mathrm{vol}_n(\boldsymbol{I}^{k,\ell})\\ &= \sum_{k=0}^{\infty}\Big(\mathrm{vol}_n(\boldsymbol{I}^k\cap\boldsymbol{I})+\sum_{\ell=1}^{2n}\mathrm{vol}_n(\boldsymbol{I}^{k,\ell})\Big).\end{aligned}$$

We recall that $\bigcup_{\ell=1}^{2n}\boldsymbol{I}^{k\ell}=\boldsymbol{I}^k\setminus\boldsymbol{I}=\boldsymbol{I}^k\cap\boldsymbol{I}^\complement$, and the terms of the finite union are pairwise disjoint, so that

$$(\boldsymbol{I}^k\cap\boldsymbol{I})\cup\bigcup_{\ell=1}^{2n}\boldsymbol{I}^{k\ell}=(\boldsymbol{I}^k\cap\boldsymbol{I})\cup(\boldsymbol{I}^k\cap\boldsymbol{I}^\complement)=\boldsymbol{I}^k.$$

Moreover, the terms of the union are pairwise disjoint. It follows from Theorem 3.13 that

$$\mathrm{vol}_n(\boldsymbol{I}^k\cap\boldsymbol{I})+\sum_{\ell=1}^{2n}\mathrm{vol}_n(\boldsymbol{I}^{k,\ell})=\mathrm{vol}_n(\boldsymbol{I}^k).$$

Hence,

$$\mu^*(\boldsymbol{A}\cap\boldsymbol{I})+\mu_n^*(\boldsymbol{A}\cap\boldsymbol{I}^\complement)\leq\sum_{k=0}^{\infty}\mathrm{vol}_n(\boldsymbol{I}^k)\leq\mu_n^*(\boldsymbol{A})+\varepsilon.$$

The result follows from the arbitrariness of ε.

Theorem 3.21 *Let $\boldsymbol{A}$ be an open subset of $\mathbb{R}^n$ or a closed subset of $\mathbb{R}^n$. Then, $\boldsymbol{A}\in\mathcal{M}_n$.*

Proof Let $\boldsymbol{A}$ be open. Let $\mathcal{F}$ be the set of n-dimensional intervals in the form

$$\boldsymbol{I}=\prod_{j=1}^{n}(a_j,b_j),$$

where $a_1,\ldots,a_n,b_1,\ldots,b_n\in\mathbb{Q}$. Since $\mathbb{Q}$ is countable and the Cartesian product of $2n$ countable sets is countable, $\mathcal{F}$ is countable. If we show that

$$\boldsymbol{A}=\bigcup_{\substack{\boldsymbol{I}\in\mathcal{F}\\ \boldsymbol{I}\subseteq\boldsymbol{A}}}\boldsymbol{I}, \tag{3.25}$$

we obtain the result by applying Theorem 3.20 and the fact that $\mathcal{M}_n$ is a σ-algebra. The inclusion $\supseteq$ in (3.25) is trivial. We prove the opposite inclusion. Let $\boldsymbol{x}\in\boldsymbol{A}$. We must show that there exists $\boldsymbol{I}\in\mathcal{F}$ such that $\boldsymbol{x}\in\boldsymbol{I}$ and $\boldsymbol{I}\subseteq\boldsymbol{A}$. Let $\varepsilon>0$. The density of $\mathbb{Q}$ in $\mathbb{R}$ guarantees, for $j\in\{1,\ldots,n\}$, the existence of $a_j,b_j\in\mathbb{Q}$ such that

$$x_j-\varepsilon<a_j<x_j<b_j<x_j+\varepsilon.$$

Let $I=\prod_{j=1}^{n}(a_j,b_j)$. Let $I\subseteq\prod_{j=1}^{n}(x_j-\varepsilon,x_j+\varepsilon)$. We show that it is possible to choose ε in such a way that

$$\prod_{j=1}^{n}(x_j-\varepsilon,x_j+\varepsilon)\subseteq\boldsymbol{A}. \tag{3.26}$$

Let $r > 0$ be such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{A}$, and let us denote by $\|\cdot\|$ the Euclidean norm. We can guarantee that (3.26) is satisfied if the left hand-side of (3.26) is contained in $\mathcal{B}_r(\boldsymbol{x})$. If $\boldsymbol{y} \in \prod_{j=1}^{n}(x_j - \varepsilon, x_j + \varepsilon)$, we have

$$\|\boldsymbol{y} - \boldsymbol{x}\| = \left(\sum_{j=1}^{n}(y_j - x_j)^2\right)^{\frac{1}{2}} \leq \sqrt{\sum_{j=1}^{n}\varepsilon^2} = \sqrt{n}\,\varepsilon.$$

Choosing $\varepsilon > 0$ such that $\sqrt{n}\,\varepsilon < r$, we obtain $\boldsymbol{y} \in \mathcal{B}_r(\boldsymbol{x})$, so that

$$I \subseteq \prod_{j=1}^{n}(x_j - \varepsilon, x_j + \varepsilon) \subseteq \boldsymbol{A}.$$

This proves (3.25).

Given two disjoint sets $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{M}_n$, we have seen that $\mu_n(\boldsymbol{A} \cup \boldsymbol{B}) = \mu_n(\boldsymbol{A}) + \mu_n(\boldsymbol{B})$. We have also seen an analogous formula for a countable union of pairwise disjoint sets:

$$\mu_n\Big(\bigcup_{k=0}^{\infty}\boldsymbol{A}_k\Big) = \sum_{k=0}^{\infty}\mu_n(\boldsymbol{A}_k).$$

The set $\mathcal{M}_n$ contains also all open and closed sets in $\mathbb{R}^n$. Applying the properties of σ-algebras, all subsets of $\mathbb{R}^n$ that one usually encounters in practice are in $\mathcal{M}_n$. As we mentioned, the construction of a subset of $\mathbb{R}^n$ that does not belong to $\mathcal{M}_n$ is complex, and requires the use of the axiom of choice.

Example *3.12*

In this example we now show that the assumption $\mu(A_0) < +\infty$ in **(b)** of Theorem 3.11 is necessary.
Let $X = \mathbb{R}$. Moreover, let $A = \mathcal{M}_1$ and $\mu = \mu_1$. For all $k \in \mathbb{N}$, we define $A_k = [k, +\infty)$, which is closed. Then, $A_k \in \mathcal{M}_1$ from Theorem 3.21. Moreover, for all $j \in \mathbb{N}$

$$\mu_1(A_k) \geq \mu_1([k, k + j]) = j,$$

from Theorem 3.20. It follows that $\mu_1(A_k) = +\infty$ for every $k \in \mathbb{N}$. On the other hand, $\bigcap_{k=0}^{\infty} A_k = \varnothing$. Therefore, the conclusion of Theorem 3.11 is, in this case, false.

Exercise 3.21

Show that:

- a singleton $\{\boldsymbol{x}_0\}$ is Lebesgue measurable and its Lebesgue measure is zero;
- every countable subset of $\mathbb{R}^n$ is Lebesgue measurable and its measure is 0.

[Solution: (3.19)]

3.4.1 Null measure sets

We recall that Theorem 3.19 we showed that $\mu_n^*(\boldsymbol{E}) = 0$ is equivalent to $\boldsymbol{E} \in \mathcal{M}_n$ and $\mu_n(\boldsymbol{E}) = 0$. The following result can be considered as a corollary of Theorem 3.19.

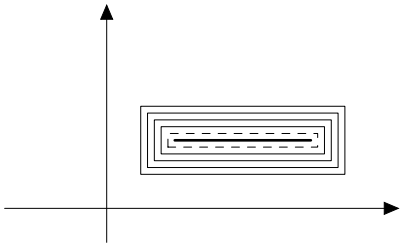
Theorem 3.22 *Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. Then, $\boldsymbol{E} \in \mathcal{M}_n$ and $\boldsymbol{E}$ has zero Lebesgue measure if and only if, for every $\varepsilon > 0$,*

there exists a countable family $(\boldsymbol{I}_k)_{k\in J}$ of n-dimensional intervals such that

$$\boldsymbol{E} \subseteq \bigcup_{k\in\mathbb{N}} \boldsymbol{I}_k \qquad \textit{and} \qquad \sum_{k\in\mathbb{N}} \mathrm{vol}(\boldsymbol{I}_k) \leq \varepsilon. \tag{3.27}$$

Proof By definition of outer measure, $\mu_n^*(\boldsymbol{E}) = 0$ is equivalent to saying that for every $\varepsilon > 0$, there exists a countable family $(\boldsymbol{I}_k)_{k\in J}$ of n-dimensional intervals such that (3.27) holds.

The intuition is the following: consider for example a segment in the plane, that we assume to be parallel to the x-axis for simplicity. We can cover it with boxes (i.e., rectangles) whose area can be made as small as we like. This shows that a segment in $\mathbb{R}^2$ has measure zero. In a similar way, a point on a line can be covered by a box (i.e., an interval) whose length can be made as small as we want, and therefore a point on a line has measure zero:



The same can be said for a rectangle in a 3-dimensional space. This begs the question: what if, instead, we have two points in the plane? or 3? or maybe as many as the natural numbers $\mathbb{N}$? Is it still true that the set has measure zero? The following example shows that the answer is positive: any countable set has measure zero.

Example *3.13*

We can use the characterization in Theorem 3.22 to produce another solution to Exercise 3.21. Hence, we want to show that any countable set $\boldsymbol{E}$ has measure zero. We can view $\boldsymbol{E}$ as the image of a family $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points of $\mathbb{R}^n$. Let $\varepsilon > 0$. For every $k \in \mathbb{N}$, let $\boldsymbol{I}_k$ be an n-dimensional interval containing $\boldsymbol{x}_k$ of measure $\leq 2^{-k-1}\,\varepsilon$. Then, $\boldsymbol{E} \subset \bigcup_{k\in\mathbb{N}} \boldsymbol{I}_k$ and

$$\sum_{k=0}^{\infty} \mathrm{vol}(\boldsymbol{I}_k) \leq \sum_{k=0}^{\infty} \frac{1}{2^{k+1}}\,\varepsilon = \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k \varepsilon = \frac{1}{2}\,\frac{1}{1-\frac{1}{2}}\,\varepsilon = \varepsilon$$

from the properties of the geometric series.

We learned that $\mathbb{N}$ has measure zero. But also, from what we said about countable sets, $\mathbb{Z}$ has measure zero, and $\mathbb{Q}$ has measure zero as well. We can even say that the countable union of null measure sets has measure zero: hence, if we take the union of an infinitely countable number of rectangles, segments and points in $\mathbb{R}^3$, we still obtain a set of measure zero.

3.4.2 Almost everywhere and for almost all

Let p be a proposition which depends on a vector $\boldsymbol{x} \in \mathbb{R}^n$ (also called, in logical mathematics, a *predicate*). We say that $p(\boldsymbol{x})$ holds *almost everywhere* (or *a.e.* in short), or that $p(\boldsymbol{x})$ holds *for almost every* $\boldsymbol{x}$, when the set of points for which $p(\boldsymbol{x})$ is false is a null Lebesgue measure set. Stated differently, for every $\boldsymbol{x} \in \mathbb{R}^n$, the proposition $p(\boldsymbol{x})$ expresses a property of $\boldsymbol{x}$ that can true or false, and we require that the points for which $p(\boldsymbol{x})$ is false has null Lebesgue measure.

Definition 3.10 *Let $\boldsymbol{A} \in \mathcal{M}_n$. We say that a certain property holds* almost everywhere *in $\boldsymbol{A}$, or that it holds* for almost every $\boldsymbol{x} \in \boldsymbol{A}$*, if the set of $\boldsymbol{x}$ for which $\boldsymbol{A}$ for which the property is not true has Lebesgue measure*

zero.

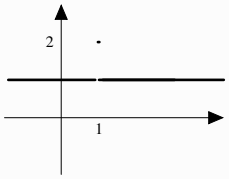
The expression "almost everywhere" is usually abbreviated as "a.e.". For example "$x \in \mathbb{R}$ is not integer" is a property that is true almost everywhere in $\mathbb{R}$, or for almost all $x \in \mathbb{R}$. Indeed, it is false when x is an integer, and the set of integers is a null Lebesgue measure set in $\mathbb{R}$. Likewise, the proposition "x is irrational" is true for almost all $x \in \mathbb{R}$. Hence, it is formally correct to state that "almost every real number is irrational", or that $x \notin \mathbb{Q}$ a.e.

Example *3.14*

Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x \neq 1 \\ 2 & \text{if } x = 1 \end{cases}$$

The graph is of f is:



We can say that $f(x)$ is equal to 1 almost everywhere (or for almost all $x \in \mathbb{R}$), because the set of points where $f(x)$ is different from one is $\{1\}$, which has measure zero in $\mathbb{R}$. Notice that also the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \\ 2 & \text{if } x \in \mathbb{Q} \end{cases}$$

is equal to 1 almost everywhere (or for almost all $x \in \mathbb{R}$), because $\mathbb{Q}$ has measure zero in $\mathbb{R}$ (as it is countable).

In probability theory, the notion of null set is particularly important. For example, if we randomly generate a real number, the probability of obtaining a natural number is equal to 0: this is a consequence of the fact that $\mathbb{N}$ has zero measure in $\mathbb{R}$.

Exercises

Exercise 3.22

Let $\boldsymbol{A}, \boldsymbol{B} \subseteq \mathbb{R}^n$ be Lebesgue measurable. Let $\boldsymbol{A}$ be a zero Lebesgue measure set. Let $\boldsymbol{B} \subseteq \boldsymbol{A}$. Show that the Lebesgue measure of $\boldsymbol{B}$ is also zero. [Solution: (3.20)]

Exercise 3.23

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ satisfy $\mu_n^*(\partial \boldsymbol{A}) = 0$. Show that $\boldsymbol{A} \in \mathcal{M}_n$. [Solution: (3.21)]

3.5 Introduction on Lebesgue integral

In Riemann integration theory, the integral of a continuous function over a closed interval $[a, b]$ is always well defined. However, continuity is not the only criterion for Riemann integrability. In fact, certain types of discontinuous functions are also Riemann integrable. For example, piecewise continuous functions, which have only a finite number of discontinuities, are Riemann integrable over a closed interval. More generally, we can say that for a function to be Riemann integrable, the set of discontinuities must be "small" in a certain sense. A classic example of a function that fails this criterion is the Dirichlet function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases} \end{aligned}$$

which is discontinuous at too many points to be Riemann integrable. Indeed, it can be proved that this function is not Riemann integrable in any proper interval of the real line.

The theory of improper integrals generalizes the theory of ordinary integrals in some directions. For example, it enables us to take into account situations where the integrand function is still regular, but it is not necessarily defined over a closed interval. These include the case where the integrand function has a point at infinity (think for example of $\int_0^2 \frac{1}{\sqrt{x}}\,\mathrm{d}x$), or the case where the interval of integration is not bounded (e.g. $\int_3^{+\infty} \frac{1}{x^2}\,\mathrm{d}x$). This extension, however, does not solve the problem of integrating functions which have too many points of discontinuity, such as the Dirichlet function.

A reasonable question that one might ask at this point is why it is not sufficient to consider continuous or piecewise continuous functions, and to what purpose one might be interested in considering bizarre functions/graphs such as the Dirichlet function.

The first reason is the fact that the space of continuous functions over an interval $[a, b]$ is not complete with respect to norms defined via integrals. Specifically, we have seen that this is true for the norms

$$\|f\|_1 = \int_a^b |f(x)|\,\mathrm{d}x \qquad \text{and} \qquad \|f\|_2 = \left(\int_a^b |f(x)|^2\,\mathrm{d}x\right)^{1/2}.$$

To overcome this difficulty, we can define these norms over the vector space of all Riemann integrable functions. Unfortunately, it is possible to show that, even by doing so, the resulting space would still not be complete. We can only hope to achieve completeness by extending the concept of integral so as that the class of integrable functions becomes wider.

There are, obviously, other reasons behind trying to extend the theory of Riemann integration to include more exotic functions. Indeed, while it is true that the integral allows us to define and then compute e.g. the area of a planar figure or the work of a force, it is also true that the integral is a crucial tool in the solution of several other mathematical problems, and in several contexts the functions to be integrated are very irregular.

3.6 Simple functions

In the classical Riemann integration theory, one first starts to integrate a simplified class of functions: the piecewise constant functions. These functions attain a finite number of values, and they are therefore very special. We use a similar approach for the Lebesgue integral: we start by considering the concept of *simple functions*, which are functions whose range is a finite subset of $\mathbb{R}$.

Definition 3.11 [Simple Function]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$. We say that f is a simple *function if there exist:*

(i) $p \in \mathbb{N}^*$,

(ii) $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p \in \mathcal{M}_n$ *pairwise disjoint, and such that* $\boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \ldots \cup \boldsymbol{A}_p = \boldsymbol{A}$,

(iii) $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$,

such that for all $i \in \{1, 2, \ldots, p\}$

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}_i, \qquad f(\boldsymbol{x}) = \alpha_i.$$

We say that f is a non-negative simple *function if, in addition, the values $\alpha_1, \alpha_2, \ldots, \alpha_p$ are non-negative, i.e., if $\alpha_i \in [0, +\infty)$ for all $i \in \{1, \ldots, p\}$.*

Definition 3.11 says, in essence, that a simple function $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ is a function whose image $f(\boldsymbol{A})$ is finite

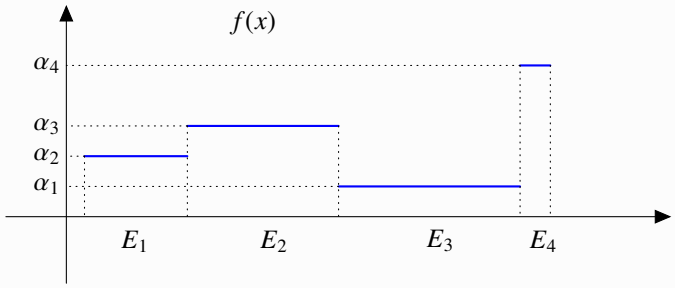
and is such that the domain can be partitioned as the finite union of measurable subsets on every point of which f is constant.

Exercise 3.24 Under the assumptions of Definition 3.11, show that $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ is simple if and only if there exist $p \in \mathbb{N}^*$, there exist $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ and there exist pairwise disjoint and measurable subsets $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$ of $\mathbb{R}^n$ such that $\boldsymbol{A} = \boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \ldots \cup \boldsymbol{A}_p$ and for all $\boldsymbol{x} \in \boldsymbol{A}$

$$f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{A_k}(\boldsymbol{x}). \tag{3.28}$$

[Solution: (3.22)]

Example *3.15* A simple function is, essentially, a piecewise constant function. An example of graph of a non-negative simple function is given in the figure:



Example *3.16* Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $\boldsymbol{E}$ be a measurable subset of $\boldsymbol{A}$. The indicator function of $\boldsymbol{E}$

$$\begin{aligned} \chi_{\boldsymbol{E}} : \boldsymbol{A} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} 1 & \text{if } \boldsymbol{x} \in \boldsymbol{E} \\ 0 & \text{if } \boldsymbol{x} \notin \boldsymbol{E} \end{cases} \end{aligned}$$

is a simple function, because its range $\chi_{\boldsymbol{E}}(\boldsymbol{A})$ is contained in $\{0, 1\}$, which is finite.

Remark 3.5 A given simple function does not uniquely define p, the sets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p \in \mathcal{M}_n$ and the values $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$. In fact, Definition 3.11 allows a simple function to take the same value α over more than one of the sets $\boldsymbol{A}_i$. However, simple functions have a canonical representation. In fact, if we define by $\beta_1, \beta_2, \ldots, \beta_r$ all the distinct values that a simple function f can take, we can also define for all $j \in \{1, 2, \ldots, r\}$ the set

$$\boldsymbol{A}_j = \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) = \beta_j\}.$$

The sets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_r$ are pairwise disjoint, and we can write

$$f(\boldsymbol{x}) = \sum_{j=1}^{r} \beta_j \chi_{A_j}(\boldsymbol{x})$$

for all $\boldsymbol{x} \in \boldsymbol{A}$. The set $\boldsymbol{A}_j$ is the largest set where f takes the value β_j, and it is measurable because it is obtained with a finite number of unions, intersections and set-differences.

We now want to show that the set of simple functions is a real vector space. To this end, we introduce a useful lemma, which says that if two functions over the same domain are simple, we can partition the domain into subsets in such a way that both functions are constant on each subset of the partition.

Lemma 3.2 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f, g : \boldsymbol{A} \longrightarrow \mathbb{R}$ be simple functions. Then, there exists $p \in \mathbb{N}^*$ and there exist measurable and pairwise disjoint subsets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p$ of $\boldsymbol{A}$, whose union is $\boldsymbol{A}$, and real numbers $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_p \in \mathbb{R}$ such that*

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}) \quad \text{and} \quad g(\boldsymbol{x}) = \sum_{k=1}^{p} \beta_k \chi_{g_k}(\boldsymbol{x}). \tag{3.29}$$

Proof Let $f(\boldsymbol{x}) = \sum_{\ell=1}^{q} \gamma_\ell \chi_{\boldsymbol{C}_\ell}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$, where $\gamma_\ell \in \mathbb{R}$ for $\ell \in \{1, \ldots, q\}$ and $\boldsymbol{C}_\ell$ satisfy the second condition in Definition 3.11 (the sets $\boldsymbol{C}_\ell$ are measurable and pairwise disjoint, and their union is $\boldsymbol{A}$). Let $g(\boldsymbol{x}) = \sum_{m=1}^{r} \delta_m \chi_{\boldsymbol{D}_m}(\boldsymbol{x})$ for every $\boldsymbol{x} \in A$, where $\delta_m \in \mathbb{R}$ and $\boldsymbol{D}_m$ satisfying analogous properties. Consider the sets in the form $\boldsymbol{C}_\ell \cap \boldsymbol{D}_m$ with $1 \leq \ell \leq q$ and $1 \leq m \leq r$, that we re-number as a finite family $(\boldsymbol{A}_k)_{1 \leq k \leq qr}$. If (ℓ_1, m_1) and (ℓ_2, m_2) are distinct elements of $\{1, \ldots, q\} \times \{1, \ldots, r\}$, the sets $\boldsymbol{C}_{\ell_1} \cap \boldsymbol{D}_{m_1}$ and $\boldsymbol{C}_{\ell_2} \cap \boldsymbol{D}_{m_2}$ are disjoint. It follows that the sets $\boldsymbol{A}_k$ are pairwise disjoint. Moreover,

$$\begin{aligned}
\bigcup_{k=1}^{qr} \boldsymbol{A}_k &= \bigcup_{\ell=1}^{q} \bigcup_{m=1}^{r} (\boldsymbol{C}_\ell \cap \boldsymbol{D}_m) \\
&= \bigcup_{\ell=1}^{q} \Big(\boldsymbol{C}_\ell \cap \bigcup_{m=1}^{r} \boldsymbol{D}_m \Big) \\
&= \bigcup_{\ell=1}^{q} (\boldsymbol{C}_\ell \cap \boldsymbol{A}) = \bigcup_{\ell=1}^{q} \boldsymbol{C}_\ell = \boldsymbol{A}.
\end{aligned}$$

We conclude by taking $\boldsymbol{A}_k = \boldsymbol{C}_\ell \cap \boldsymbol{D}_m$, $\alpha_k = \gamma_\ell$ and $\beta_k = \delta_m$.

Theorem 3.23 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f, g : \boldsymbol{A} \longrightarrow \mathbb{R}$ be simple functions. Let $\lambda \in \mathbb{R}$. Then,*

(a) *$f + g$, $f\,g$, λf, $\max\{f, g\}$, $\min\{f, g\}$ are simple;*

(b) *if $g(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \boldsymbol{A}$, then $\frac{f}{g}$ is simple;*

(c) *if $h : \mathbb{R} \longrightarrow \mathbb{R}$, then $h \circ f$ is simple.*

Proof We start proving **(a)**. From Lemma 3.2, there exists $p \in \mathbb{N}^*$ and there exist measurable and pairwise disjoint subsets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p$ of $\boldsymbol{A}$, whose union is $\boldsymbol{A}$, and real numbers $\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_p \in \mathbb{R}$ such that, for all $\boldsymbol{x} \in \boldsymbol{A}$,

$$f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}) \quad \text{and} \quad g(\boldsymbol{x}) = \sum_{k=1}^{p} \beta_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}),$$

so that, for all $\boldsymbol{x} \in \boldsymbol{A}$,

$$(f + g)(\boldsymbol{x}) = f(\boldsymbol{x}) + g(\boldsymbol{x}) = \sum_{k=1}^{p} (\alpha_k + \beta_k) \chi_{\boldsymbol{A}_k}(\boldsymbol{x}).$$

This allows us to conclude that $f + g$ is simple. Using the same partition, we see immediately that for all $\boldsymbol{x} \in \boldsymbol{A}$

$$(f\,g)(\boldsymbol{x}) = f(\boldsymbol{x})\,g(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \beta_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}),$$

which shows that $f\,g$ is simple. The case of $\lambda\,f$ is left as an exercise (see Exercise 3.25). Consider $\max\{f,g\}$. For every $\boldsymbol{x}\in\boldsymbol{A}$,

$$\max\{f,g\}(\boldsymbol{x})=\sum_{k=1}^{p}\max\{\alpha_k,\beta_k\}\chi_{A_k}(\boldsymbol{x}).$$

The proof of $\min\{f,g\}$ is analogous.
We prove **(b)**. Under the assumption $g(\boldsymbol{x})\neq 0$ for all $\boldsymbol{x}\in\boldsymbol{A}$, we have $\beta_k\neq 0$ for all $k\in\{1,2,\ldots,p\}$, and therefore for all $\boldsymbol{x}\in\boldsymbol{A}$

$$\frac{f}{g}(\boldsymbol{x})=\sum_{k=1}^{m}\frac{\alpha_k}{\beta_k}\chi_{A_k}(\boldsymbol{x}),$$

which shows that $\frac{f}{g}$ is simple.
We prove **(c)**. Given $\boldsymbol{x}\in\boldsymbol{A}_k$, we have $(h\circ f)(\boldsymbol{x})=h(f(\boldsymbol{x}))=h(\alpha_k)$, which proves the claim.

Exercise 3.25 Prove that if $\lambda\in\mathbb{R}$ and $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ is simple, the product $\lambda\,f$ is simple. [Solution: (3.23)]

We now define the Lebesgue integral for simple functions which, additionally, are non-negative: in other words, we are not only requiring that the image of the function f be a finite subset of real numbers, but we are also requiring f to take non-negative values.

Definition 3.12 [Lebesgue Integral of a Non-Negative Simple Function]
Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ be a simple function satisfying $f(\boldsymbol{x})=\sum_{k=1}^{p}\alpha_k\chi_{A_k}$ for some $p\in\mathbb{N}^*$, pairwise disjoint sets $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_p\in\mathcal{M}_n$ satisfying $\boldsymbol{A}=\boldsymbol{A}_1\cup\ldots\cup\boldsymbol{A}_p$ and $\alpha_1,\ldots,\alpha_p\in[0,+\infty)$. We define

$$\int_{\boldsymbol{A}}f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\overset{def}{=}\sum_{k=1}^{p}\alpha_k\,\mu_n(\boldsymbol{A}_k). \tag{3.30}$$

The integral in (3.30) is also denoted by

$$\int_{\boldsymbol{A}}f\qquad\text{or}\qquad\int_{\boldsymbol{A}}f\,\mathrm{d}\mu.$$

This definition of integral corresponds to the intuitive notion of integration of non-negative functions as the area under the graph of the function (or volume in higher dimensions).

The sum in (3.30) is intended as a sum in $[0,+\infty]$. In fact, by definition, $\alpha_1,\ldots,\alpha_p$ are real numbers, but $\mu_n(\boldsymbol{A}_k)$, for one or more values of $k\in\{1,2,\ldots,p\}$, may be equal to $+\infty$. Thus, the summation in (3.30) is intended as one following the rules of the algebra of $[0,+\infty]$. Hence, in case $\alpha_i=0$, the product $\alpha_i\,\mu_n(\boldsymbol{E}_i)$ gives 0, even when $\mu_n(\boldsymbol{E}_i)=+\infty$. As a consequence, the integral of the identically zero function over an arbitrary measurable set, even of infinite measure, is equal to 0.

Remark 3.6 We have already observed that, from Lemma 3.2, the choice of the partition of $\boldsymbol{A}$ is not unique. Thus, to ensure the integral in Definition 3.12 makes sense, we must show that the integral does not depend on the way we decompose $\boldsymbol{A}$, provided we preserve the property that in each of the parts f is constant. Indeed, let $f(\boldsymbol{x})=\sum_{\ell=1}^{q}\beta_\ell\chi_{B_\ell}(\boldsymbol{x})$, where $\beta_1,\ldots,\beta_q$ are non-negative real numbers and $\boldsymbol{B}_1,\ldots,\boldsymbol{B}_q$ satisfy the same

assumptions of $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p$. Then, for $k \in \{1, \ldots, p\}$,

$$\begin{aligned}\mu_n(\boldsymbol{A}_k) &= \mu_n(\boldsymbol{A}_k \cap \boldsymbol{A}) = \mu_n\left(\boldsymbol{A}_k \cap \Big(\bigcup_{\ell=1}^{q} \boldsymbol{B}_\ell\Big)\right) \\ &= \mu_n\Big(\bigcup_{\ell=1}^{q} (\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\Big) = \sum_{\ell=1}^{q} \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell),\end{aligned}$$

where we have used the subadditivity (notice that since $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_q$ are pairwise disjoint, the intersections $\boldsymbol{A}_k \cap \boldsymbol{B}_1, \ldots, \boldsymbol{A}_k \cap \boldsymbol{B}_q$ are also pairwise disjoint). Applying the operations of addition and multiplication in $[0, +\infty]$ we get

$$\sum_{k=1}^{p} \alpha_k \mu_n(\boldsymbol{A}_k) = \sum_{k=1}^{p} \left(\sum_{\ell=1}^{q} \alpha_k \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\right) = \sum_{\ell=1}^{q} \left(\sum_{k=1}^{p} \alpha_k \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\right).$$

Now, we observe that for all k and ℓ we have

$$\alpha_k \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell) = \beta_\ell \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell).$$

This is obvious if $\mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell) = 0$. If $\mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell) > 0$, we have $\boldsymbol{A}_k \cap \boldsymbol{B}_\ell \neq \emptyset$. Let $\boldsymbol{x} \in \boldsymbol{A}_k \cap \boldsymbol{B}_\ell$. Then, $\alpha_k = f(\boldsymbol{x}) = \beta_\ell$. We conclude that

$$\begin{aligned}\sum_{k=1}^{p} \alpha_k \mu_n(\boldsymbol{A}_k) &= \sum_{\ell=1}^{q} \left(\sum_{k=1}^{p} \beta_\ell \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\right) \\ &= \sum_{\ell=1}^{q} \beta_\ell \left(\sum_{k=1}^{p} \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\right) \\ &= \sum_{\ell=1}^{q} \beta_\ell \mu_n\left(\bigcup_{k=1}^{p} (\boldsymbol{A}_k \cap \boldsymbol{B}_\ell)\right) \\ &= \sum_{\ell=1}^{q} \beta_\ell \mu_n\left(\Big(\bigcup_{k=1}^{p} \boldsymbol{A}_k\Big) \cap \boldsymbol{B}_\ell\right) \\ &= \sum_{\ell=1}^{q} \beta_\ell \mu_n(\boldsymbol{A} \cap \boldsymbol{B}_\ell) = \sum_{\ell=1}^{q} \beta_\ell \mu_n(\boldsymbol{B}_\ell).\end{aligned}$$

Remark 3.7 If $\boldsymbol{A} \in \mathcal{M}_n$ and $f(\boldsymbol{x}) = \alpha$ for all $\boldsymbol{x} \in \boldsymbol{A}$ and with $\alpha \in \mathbb{R}$, then f is simple over $\boldsymbol{A}$. If $\alpha \geq 0$, we have

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \alpha\, \mu_n(\boldsymbol{A}).$$

In particular, if $\alpha = 0$, the integral is zero regardless of $\mu_n(\boldsymbol{A})$. We observe also that if $\boldsymbol{B}$ is a measurable subset of $\boldsymbol{A}$, we can define

$$\chi_{\boldsymbol{B}} : \mathbb{R}^n \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \mapsto \begin{cases} 1 & \text{if } \boldsymbol{x} \in \boldsymbol{B} \\ 0 & \text{if } \boldsymbol{x} \in \boldsymbol{B}^{\complement} \end{cases} \qquad \text{and} \qquad \chi_{\boldsymbol{B}}|_{\boldsymbol{A}} : \boldsymbol{A} \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \mapsto \begin{cases} 1 & \text{if } \boldsymbol{x} \in \boldsymbol{B} \\ 0 & \text{if } \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{B}, \end{cases}$$

where $\chi_{\boldsymbol{B}}|_{\boldsymbol{A}}$ denotes the restriction of $\chi_{\boldsymbol{B}}$ to $\boldsymbol{A}$. This restriction is a simple function over $\boldsymbol{A}$, since for every $\boldsymbol{x} \in \boldsymbol{A}$

$$\chi_{\boldsymbol{B}}|_{\boldsymbol{A}}(\boldsymbol{x}) = 1 \cdot \chi_{\boldsymbol{B}}(\boldsymbol{x}) + 0 \cdot \chi_{\boldsymbol{A} \setminus \boldsymbol{B}}(\boldsymbol{x}).$$

Hence, the integral of the restriction is

$$\int_A \chi_B|_A(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1\cdot\mu_n(\boldsymbol{B}) + 0\cdot\mu_n(\boldsymbol{A}\setminus\boldsymbol{B}) = \mu_n(\boldsymbol{B}).$$

Example *3.17* Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ be such that $f(\boldsymbol{x})=1$ for all $\boldsymbol{x}\in\boldsymbol{A}$. Then,

$$\int_A f(\boldsymbol{x})\,\mathrm{d}x = \mu_n(\boldsymbol{A}).$$

Exercise 3.26 Show that the function

$$\begin{aligned} f:\mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x\in[0,3] \\ 3 & \text{if } x\in(3,5] \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

is simple and compute the integral $\int_A f$. [Solution: (3.24)]

Exercise 3.27 Show that the function

$$\begin{aligned} f:\mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x_1,x_2) &\mapsto \begin{cases} 1 & \text{if } x_1\in[0,1] \text{ and } x_2\in[0,1] \\ 2 & \text{if } x\in(1,2] \text{ and } x_2\in[0,1] \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is simple and compute the integral $\int_A f$. [Solution: (3.25)]

Exercise 3.28 Show that the function

$$\begin{aligned} f:\mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x\in(0,+\infty) \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is simple and compute the integral $\int_A f$. [Solution: (3.26)]

We now examine the fundamental properties of the integral of simple functions.

Theorem 3.24 *Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f,g:\boldsymbol{A}\longrightarrow\mathbb{R}$ be simple and non-negative. Let $c\in[0,+\infty)$. Then:*

(a) $\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\in[0,+\infty]$;

(b) $\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=0$ *if and only if* $\mu_n(\{\boldsymbol{x}\in\boldsymbol{A}\mid f(\boldsymbol{x})>0\})=0$;

(c) $f\le g$ *implies* $\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\le\int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$;

(d) $\displaystyle\int_A (f(\boldsymbol{x}) + g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = \int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$;

(e) $\displaystyle\int_A c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = c\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

Proof **(a)** follows immediately from the definition of integral of a simple function. We prove **(b)**. Let f as in (3.28), where $\alpha_1, \dots, \alpha_p \in [0, +\infty)$. Let $\displaystyle\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$. Then, $\alpha_k\,\mu_n(\boldsymbol{A}_k) = 0$ for all $k \in \{1, \dots, p\}$. This implies that, if for a certain $k \in \mathbb{N}$ we have $\alpha_k > 0$, then necessarily $\mu_n(\boldsymbol{A}_k) = 0$. Thus,

$$\mu_n(\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\}) = \mu_n\Big(\bigcup_{\alpha_k > 0} \boldsymbol{A}_k\Big) = \sum_{\alpha_k > 0} \mu_n(\boldsymbol{A}_k) = 0.$$

Conversely, let $\mu_n(\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\}) = 0$. If $\alpha_k > 0$ for some $k \in \{1, \dots, p\}$, then $\boldsymbol{A}_k \subseteq \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\}$ and therefore $\mu_n(\boldsymbol{A}_k) = 0$. Thus,

$$\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=1}^{p} \alpha_k\,\mu_n(\boldsymbol{A}_k) = 0.$$

We prove **(c)**. From Lemma 3.2, we can assume that for all $\boldsymbol{x} \in \boldsymbol{A}$ we have $g(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{p} \beta_k \chi_{A_k}(\boldsymbol{x})$ with $\alpha_k \le \beta_k$ for all $k \in \{1, \dots, p\}$. It follows that

$$\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=1}^{p} \alpha_k\,\mu_n(\boldsymbol{A}_k) \le \sum_{k=1}^{p} \beta_k\,\mu_n(\boldsymbol{A}_k) = \int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We verify **(d)**. Assuming again $g(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{p} \beta_k \chi_{A_k}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$, we have $f(\boldsymbol{x}) + g(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{p} (\alpha_k + \beta_k)\chi_{A_k}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$. Hence

$$\begin{aligned}\int_A (f(\boldsymbol{x}) + g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} &= \sum_{k=1}^{p} (\alpha_k + \beta_k)\,\mu_n(\boldsymbol{A}_k)\\ &= \sum_{k=1}^{p} \alpha_k\,\mu_n(\boldsymbol{A}_k) + \sum_{k=1}^{p} \beta_k\,\mu_n(\boldsymbol{A}_k)\\ &= \int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Finally, we prove **(e)**. We have

$$c\,f(\boldsymbol{x}) = \sum_{k=0}^{p} c\,\alpha_k \chi_{A_k}(\boldsymbol{x}) \quad \Rightarrow \quad \int_A c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{p} c\,\alpha_k\,\mu(\boldsymbol{A}_k) = c\sum_{k=0}^{p} \alpha_k\,\mu(\boldsymbol{A}_k) = c\int_A f(\boldsymbol{x}).$$

We now consider the integral of a non-negative simple function on a subset of its domain.

Theorem 3.25 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ and let $\boldsymbol{B}$ be a Lebesgue measurable subset of $\boldsymbol{A}$. Then, $f|_{\boldsymbol{B}}$ is simple over $\boldsymbol{B}$. Moreover, if f takes values in $[0, +\infty)$, then*

$$\int_B f|_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_A f(\boldsymbol{x})\,\chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof Let $f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$ with $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$, and $\boldsymbol{A}_k$ satisfying the conditions of Definition 3.11. We note that the sets $\boldsymbol{A}_k \cap \boldsymbol{B}$ are measurable, pairwise disjoint and their union is $\boldsymbol{B}$. If $\boldsymbol{x} \in \boldsymbol{A}_k \cap \boldsymbol{B}$, we have $f(\boldsymbol{x}) = \alpha_k$. Thus, $f|_{\boldsymbol{B}}(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k \cap \boldsymbol{B}}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{B}$. It follows $f|_{\boldsymbol{B}}$ is simple. Moreover, if $\alpha_1, \ldots, \alpha_p > 0$, we have

$$\int_{\boldsymbol{B}} f|_{\boldsymbol{B}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \sum_{k=1}^{p} \alpha_k \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}).$$

In this last case, we observe that $\chi_{\boldsymbol{B}}$ is simple and non-negative, so that $f\chi_{\boldsymbol{B}}$ is simple and non-negative from **(a)** of Theorem 3.23. Moreover,

$$f(\boldsymbol{x})\chi_{\boldsymbol{B}}(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x})\chi_{\boldsymbol{B}}(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k \cap \boldsymbol{B}}(\boldsymbol{x}).$$

Applying **(d)** and **(e)** of Theorem 3.24 and Remark 3.7, we find

$$\begin{aligned} \int_{\boldsymbol{A}} f(\boldsymbol{x})\chi_{\boldsymbol{B}}\, \mathrm{d}\boldsymbol{x} &= \sum_{k=1}^{p} \alpha_k \int_{\boldsymbol{A}} \chi_{\boldsymbol{A}_k \cap \boldsymbol{B}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\ &= \sum_{k=1}^{p} \alpha_k \mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}). \end{aligned}$$

Example *3.18* Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ be a simple non-negative function. Let $q \in \mathbb{N}^*$, and let $\boldsymbol{A} = \boldsymbol{B}_1 \cup \boldsymbol{B}_2 \cup \ldots \cup \boldsymbol{B}_q$, where $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_q$ are measurable and pairwise disjoint. Then, we show that

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{B}_1} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{B}_2} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} + \ldots + \int_{\boldsymbol{B}_q} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x},$$

where, for every $j \in \{1, \ldots, q\}$,

$$\int_{\boldsymbol{B}_j} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \overset{\text{def}}{=} \int_{\boldsymbol{B}_j} f|_{\boldsymbol{B}_j}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

As usual, since f is a simple non-negative function, there exists $p \in \mathbb{N}^*$, there exist measurable and pairwise disjoint subsets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p$ of $\boldsymbol{A}$, whose union is $\boldsymbol{A}$, and real numbers $\alpha_1, \ldots, \alpha_p \geq 0$ such that, for all $\boldsymbol{x} \in \boldsymbol{A}$,

$$f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}).$$

For every $j \in \{1, \ldots, q\}$, we can write

$$\begin{aligned} \boldsymbol{B}_j &= \boldsymbol{A} \cap \boldsymbol{B}_j = \Big(\bigcup_{k=1}^{p} \boldsymbol{A}_k\Big) \cap \boldsymbol{B}_j \\ &= \bigcup_{k=1}^{p} (\boldsymbol{A}_k \cap \boldsymbol{B}_j) = (\boldsymbol{A}_1 \cap \boldsymbol{B}_j) \cup (\boldsymbol{A}_2 \cap \boldsymbol{B}_j) \cup \ldots \cup (\boldsymbol{A}_p \cap \boldsymbol{B}_j), \end{aligned}$$

where the intersections in the right hand-side are pairwise disjoint (why?). We have

$$\begin{aligned}\sum_{j=1}^{q}\int_{\boldsymbol{B}_j} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \alpha_1\,\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_1)+\alpha_2\,\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_1)+\ldots+\alpha_p\,\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_1)\\ &\quad+\alpha_1\,\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_2)+\alpha_2\,\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_2)+\ldots+\alpha_p\,\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_2)\\ &\quad+\ldots+\alpha_1\,\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_q)+\alpha_2\,\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_q)+\ldots+\alpha_p\,\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_q)\\ &= \alpha_1\left(\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_1)+\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_2)+\ldots+\mu_n(\boldsymbol{A}_1\cap\boldsymbol{B}_q)\right)\\ &\quad+\alpha_2\left(\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_1)+\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_2)+\ldots+\mu_n(\boldsymbol{A}_2\cap\boldsymbol{B}_q)\right)\\ &\quad+\ldots+\alpha_p\left(\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_1)+\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_2)+\ldots+\mu_n(\boldsymbol{A}_p\cap\boldsymbol{B}_q)\right)\\ &= \alpha_1\,\mu_n(\boldsymbol{A}_1)+\alpha_2\,\mu_n(\boldsymbol{A}_2)+\ldots+\alpha_p\,\mu_n(\boldsymbol{A}_p)=\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Exercises

Exercise 3.29 Let $\boldsymbol{A}\in\mathcal{M}_n$ and let $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ be a simple function. Let $\lambda\in\mathbb{R}$. Show that λf is a simple function. [Solution: (3.27)]

3.7 Measurable functions

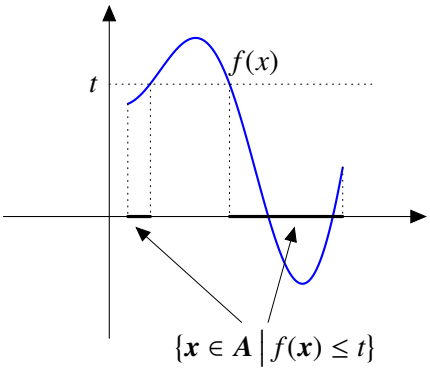
The elementary theory of integration is strictly related to the notion of continuity. Indeed, considering the case of bounded functions defined over a bounded set, we can integrate those functions whose discontinuities form a "small" set, for example a finite set of points in the one-dimensional case, and a finite number of regular arcs in the bi-dimensional case.

By contrast, we will see that the Lebesgue integrability is mostly related to how big the integrand function is rather than its regularity. However, some level of regularity is still required: this will lead to the notion of *measurable function* (not to be confused with the concept of measurable set introduced earlier). This restriction is very mild, so that in practice assumptions of measurability can often be ignored because they are usually satisfied. It is even possible to show that, in order to construct non-measurable functions one needs the axiom of choice.

Definition 3.13 [Measurable Function]
Let $n\in\mathbb{N}^*$. Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f:\boldsymbol{A}\longrightarrow[-\infty,+\infty]$. We say that f is measurable *if, for every $t\in\mathbb{R}$,*

$$\{\boldsymbol{x}\in A\,|\,f(\boldsymbol{x})\le t\}\in\mathcal{M}_n.$$

The graph below shows the geometric meaning of the set appearing in the condition of Definition 3.13.



Example 3.19

Let $E \subseteq \mathbb{R}^n$ be measurable. The characteristic function $\chi_E : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined for all $x \in \mathbb{R}^n$ by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \in \mathbb{R}^n \setminus E \end{cases}$$

is measurable, because for every $t \in \mathbb{R}$ we have

$$\{x \in \mathbb{R}^n \mid \chi_E(x) \leq t\} = \begin{cases} \varnothing & \text{if } t < 0 \\ E & \text{if } 0 \leq t < 1 \\ \mathbb{R}^n & \text{if } t \geq 1. \end{cases}$$

Remark 3.8 In Example 3.19 we saw that the characteristic function of a measurable set is measurable. Now we show something more general: simple functions are measurable. Let $n \in \mathbb{N}^*$, $A \in \mathcal{M}_n$ and $f : A \longrightarrow \mathbb{R}$ be a simple function. For every $x \in A$ we have

$$f(x) = \sum_{k=1}^{p} \alpha_k \chi_{A_k}(x),$$

where $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$ and $A_1, \ldots, A_p$ are pairwise disjoint sets whose union is equal to A. Then, for every $t \in \mathbb{R}$

$$\{x \in A \mid f(x) \leq t\} = \bigcup_{\alpha_k \leq t} A_k,$$

and this set is measurable because it is the union of a finite set of measurable sets.

The following result gives some equivalent conditions for a function to be measurable.

Theorem 3.26 *Let $A \in \mathcal{M}_n$. Let $f : A \longrightarrow [-\infty, +\infty]$. The following statements are equivalent:*

(i) *f is measurable;*

(ii) $\forall\, t \in \mathbb{R}$, $\{x \in A \,|\, f(x) > t\} \in \mathcal{M}_n$;

(iii) $\forall\, t \in \mathbb{R}$, $\{x \in A \,|\, f(x) \geq t\} \in \mathcal{M}_n$;

(iv) $\forall\, t \in \mathbb{R}$, $\{x \in A \,|\, f(x) < t\} \in \mathcal{M}_n$.

Proof We show that **(i)** implies **(ii)**. This follows from **(e)** in Theorem 3.9 and the fact that

$$\{x \in A \,|\, f(x) > t\} = A \setminus \{x \in A \,|\, f(x) \leq t\}.$$

We show that **(ii)** implies **(iii)**. This follows from **(c)** in Theorem 3.9 if we show that

$$\{\boldsymbol{x} \in \boldsymbol{A} \,|\, f(\boldsymbol{x}) \geq t\} = \bigcap_{k=1}^{\infty} \left\{\boldsymbol{x} \in \boldsymbol{A} \,\middle|\, f(\boldsymbol{x}) > t - \frac{1}{k}\right\} \in \mathcal{M}_n.$$

Indeed, if $\boldsymbol{x}$ is an element of $\boldsymbol{A}$ such that $f(\boldsymbol{x}) \geq t$, we have also $f(\boldsymbol{x}) > t - \frac{1}{k}$ for every $k \in \mathbb{N}^*$. Conversely, if $f(\boldsymbol{x}) > t - \frac{1}{k}$ for every $k \in \mathbb{N}^*$, taking the limit for $k \to \infty$ we obtain $f(\boldsymbol{x}) \geq t$ (see also Exercise 1.27). To prove that **(iii)** implies **(iv)**, we observe that

$$\{\boldsymbol{x} \in \boldsymbol{A} \,|\, f(\boldsymbol{x}) < t\} = \boldsymbol{A} \setminus \{\boldsymbol{x} \in \boldsymbol{A} \,|\, f(\boldsymbol{x}) \geq t\} \in \mathcal{M}_n,$$

as the difference of measurable sets is measurable. We now prove that **(iv)** implies **(i)**. This follows again from **(c)** in Theorem 3.9 if we show that

$$\{\boldsymbol{x} \in \boldsymbol{A} \,|\, f(\boldsymbol{x}) \leq t\} = \bigcap_{k=1}^{\infty} \left\{\boldsymbol{x} \in \boldsymbol{A} \,\middle|\, f(\boldsymbol{x}) < t + \frac{1}{k}\right\}.$$

Indeed, if $\boldsymbol{x}$ is an element of A such that $f(\boldsymbol{x}) \leq t$, we have also $f(\boldsymbol{x}) < t + \frac{1}{k}$ for every $k \in \mathbb{N}^*$. Conversely, if $f(\boldsymbol{x}) < t + \frac{1}{k}$ for every $k \in \mathbb{N}^*$, taking the limit for $k \to \infty$ we obtain $f(\boldsymbol{x}) \leq t$.

Another equivalent characterization of measurable function is presented in the following result.

Corollary 3.1 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$. Then, f is measurable if and only if for all $a \in \mathbb{R}$ the set $f^{-1}([a, +\infty])$ is measurable.*

Proof We recall that, given a function $f : X \longrightarrow Y$ and a subset $H \subseteq Y$, the inverse image $f^{-1}(H)$ is defined as $\{x \in X \,|\, f(x) \in H\}$. It follows that, given $a \in \mathbb{R}$,

$$f^{-1}([a, +\infty]) = \{\boldsymbol{x} \in A \,|\, f(\boldsymbol{x}) \in [a, +\infty]\} = \{\boldsymbol{x} \in A \,|\, f(\boldsymbol{x}) \geq a\}.$$

Thus, the condition of the Corollary is equivalent to **(iii)** in Theorem 3.26.

We want to use the result in Corollary 3.1 to show that a function $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ is measurable if and only if the inverse image of every open set of $\mathbb{R}$ is measurable. We first introduce a lemma.

Lemma 3.3 *Let $\Omega \subseteq \mathbb{R}$ be open and non-empty. Then, Ω is the at-most-countable union of pairwise disjoint open intervals.*

Proof Let Ω be a non-empty open subset of $\mathbb{R}$. Let $x \in \Omega$. Since Ω is open, there exists $r > 0$ such that $(x - r, x + r) \subseteq \Omega$. Therefore, there exists $z \in (x - r, x)$ which therefore satisfies $z < x$ and $(z, x) \subseteq \Omega$ (we also know that $[z, x] \subseteq \Omega$, but we do not need to take into account the end-points). Likewise, there exists $y \in (x, x + r)$, which then satisfies $y > x$ and $(x, y) \subseteq \Omega$. Thus, we can define $a_x, b_x \in [-\infty, +\infty]$ as

$$\begin{aligned} a_x &= \inf\{z \in \mathbb{R} \,|\, (z, x) \subseteq \Omega\} \\ b_x &= \sup\{y \in \mathbb{R} \,|\, (x, y) \subseteq \Omega\}. \end{aligned}$$

Since Ω is non-empty, we have $a_x \leq b_x$. Since Ω is not a singleton (because, if that was the case, the set Ω would be closed), we have $a_x \neq b_x$. Hence, $a_x < b_x$, and this implies that we can define $I_x = (a_x, b_x)$, which is an open interval, which contains x. We prove that:

- $I_x \subseteq \Omega$. Let $\xi \in I_x$, and we show that $\xi \in \Omega$. Since $\xi \in I_x$, we have three possibilities:

  - $\xi = x$. In this case $\xi \in \Omega$, because x was chosen to be a point of Ω;
  - $\xi \in (x, b_x)$. Since $\xi < b_x$, by definition of b_x there exists $y > \xi$ such that $(x, y) \subseteq \Omega$, and since $\xi \in (x, y)$ we conclude that $\xi \in \Omega$;
  - $\xi \in (x, a_x)$. This case is the dual of the previous done and can be dealt with using a similar argument (try!).

- $b_x \notin \Omega$. Indeed, if by contradiction we assume that $b_x \in \Omega$, since Ω is open, there exists $\rho > 0$ such that $(b_x - \rho, b_x + \rho) \subseteq \Omega$, and therefore $(x, b_x + \rho) \subseteq \Omega$, thus contradicting the fact that b_x is the supremum of the set $\{y \in \mathbb{R} \,|\, (x, y) \subseteq \Omega\}$.
- $a_x \notin \Omega$. The argument can be adapted from the previous point, *mutatis mutandis*.

We now prove that Ω is the union $\{I_x \,|\, x \in \Omega\}$ of the set of all open intervals I_x constructed in this way, i.e., we show that

$$\Omega = \bigcup_{x \in \Omega} I_x.$$

It is obvious that $\bigcup_{x \in \Omega} I_x \subseteq \Omega$, because for every $x \in \Omega$ the interval I_x is a subset of Ω (and the union of subsets of Ω is still a subset of Ω). It is equally easy to how that $\bigcup_{x \in \Omega} I_x \supseteq \Omega$: given $\xi \in \Omega$, we have $\xi \in I_\xi$, and I_ξ is an element of $\bigcup_{x \in \Omega} I_x$. We show that the set $\{I_x \,|\, x \in \Omega\}$ is pairwise disjoint. Suppose that $\alpha \in I_x$ and $\alpha \in I_\xi$. There are two cases: the first is $I_x = I_\xi$. The other case is the one where $I_x \neq I_\xi$. In this second case, we need to show that $I_x \cap I_\xi = \varnothing$. Suppose by contradiction that there exists $\alpha \in I_x \cap I_\xi$. Then, since $(a_x, b_x) \subseteq \Omega$, we also have $(a_x, \alpha) \subseteq \Omega$. Since $(a_\xi, b_\xi) \subseteq \Omega$, we also have $(\alpha, b_\xi) \subseteq \Omega$. On the other hand, $\alpha \in \Omega$, so that we obtain $(a_x, b_\xi) \subseteq \Omega$. This is a contradiction because $b_x \in (a_x, b_\xi)$ but we showed that $b_x \notin \Omega$.

Thus, we have shown that Ω is the union of a pairwise disjoint set of open intervals. We need to show that this union can be chosen to be at most countable. Since each interval I_x is proper (it is not empty, and it is not a singleton) and it is open, from Theorem 3.81 it also contains a rational number. We can therefore define a bijection between the set of open intervals $\{I_x \,|\, x \in \Omega\}$ and a subset of $\mathbb{Q}$. This subset of $\mathbb{Q}$ is at most countable. Thus, Ω is the at-most-countable union of pairwise disjoint open intervals.

Theorem 3.27 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$. Then, f is measurable if and only if for all open sets Ω in $\mathbb{R}$, the inverse image $f^{-1}(\Omega)$ is measurable.*

Proof First, we show that if for any open set Ω of $\mathbb{R}$ the set $f^{-1}(\Omega)$ is measurable, the function f is measurable. Choosing Ω to be in the form $(c, +\infty)$, using the same idea of Corollary 3.1, given $a \in \mathbb{R}$,

$$f^{-1}((c, +\infty)) = \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \in (c, +\infty)\} = \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > c\}.$$

Using **(ii)** in Theorem 3.26, we conclude that f is measurable.

Conversely, assume that f is measurable. Let Ω be an open subset of $\mathbb{R}$. From Lemma 3.3, the open set Ω can be written as an at-most-countable union of pairwise disjoint open intervals. We assume that this set of collection (as usual, if it is finite, we can use the empty set to complete this finite set to a countable one), so that we can write these open intervals as a sequence $(I_k)_{k \in \mathbb{N}}$. Each I_k can be written as $Y_k \cap X_k$, where $Y_k = (-\infty, y_k)$ and $X_k = (x_k, +\infty)$. Since f is measurable, the sets $f^{-1}(Y_k)$ and $f^{-1}(X_k)$ are measurable. From

Theorem 3.18 and the properties of σ-algebras[a]

$$f^{-1}(\Omega) = f^{-1}\Big(\bigcup_{k=0}^{\infty}(X_k \cap Y_k)\Big) = \bigcup_{k=0}^{\infty}(f^{-1}(X_k) \cap f^{-1}(Y_k)),$$

which shows that $f^{-1}(\Omega)$ is measurable.

[a] We are using a simple property of the inverse image. If H, K are subsets of the codomain of f, then $f^{-1}(H \cap K) = f^{-1}(H) \cap f^{-1}(K)$. Indeed,

$$x \in f^{-1}(H \cap K) \iff f(x) \in H \cap K \iff \begin{cases} f(x) \in H \\ f(x) \in K \end{cases} \iff \begin{cases} x \in f^{-1}(H) \\ x \in f^{-1}(K) \end{cases} \iff x \in f^{-1}(H) \cap f^{-1}(K).$$

Measurability behaves well with respect to limits. In particular, the following result shows that limits of measurable functions are measurable.

Theorem 3.28 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of measurable functions $f_k : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$. Let*

$$\varphi : \boldsymbol{A} \longrightarrow [-\infty,+\infty] \quad \boldsymbol{x} \mapsto \sup_{k\in\mathbb{N}} f_k(\boldsymbol{x}), \qquad \psi : \boldsymbol{A} \longrightarrow [-\infty,+\infty] \quad \boldsymbol{x} \mapsto \inf_{k\in\mathbb{N}} f_k(\boldsymbol{x}),$$

$$\gamma : \boldsymbol{A} \longrightarrow [-\infty,+\infty] \quad \boldsymbol{x} \mapsto \liminf_{k\to\infty} f_k(\boldsymbol{x}), \qquad \rho : \boldsymbol{A} \longrightarrow [-\infty,+\infty] \quad \boldsymbol{x} \mapsto \limsup_{k\to\infty} f_k(\boldsymbol{x}).$$

Then,

(a) *φ and ψ are measurable;*

(b) *γ and ρ are measurable;*

(c) *if there exists a function $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ such that*

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

i.e., such that f_k is pointwise convergent to f (in the sense of the convergence in $\overline{\mathbb{R}}$), then f is measurable.

Proof We prove **(a)**. First, we must show that for all $t \in \mathbb{R}$, the set $\{\boldsymbol{x} \in \boldsymbol{A} \,|\, \varphi(\boldsymbol{x}) \le t\}$ is measurable. Clearly, $\varphi(\boldsymbol{x}) \le t$ is equivalent to the condition $\forall\, k \in \mathbb{N}$, $f_k(\boldsymbol{x}) \le t$, from which we find

$$\{\boldsymbol{x} \in \boldsymbol{A} \,|\, \varphi(\boldsymbol{x}) \le t\} = \bigcap_{k=0}^{\infty}\{\boldsymbol{x} \in \boldsymbol{A} \,|\, f_k(\boldsymbol{x}) \le t\}.$$

Indeed, if $\varphi(\boldsymbol{x}) \le t$, for every $k \in \mathbb{N}$ we have $f_k(\boldsymbol{x}) \le \varphi(\boldsymbol{x}) \le t$. Conversely, if $f_k(\boldsymbol{x}) \le t$ for every $k \in \mathbb{N}$, then t is an upper bound of $\{f_k(\boldsymbol{x}) \,|\, k \in \mathbb{N}\}$, so that $\varphi(\boldsymbol{x}) \le t$. Then, **(a)** follows from **(c)** in Theorem 3.9. To see that ψ is measurable, from Theorem 3.26, to prove that ψ is measurable, we can prove that for every $t \in \mathbb{R}$ the set $\{\boldsymbol{x} \in \boldsymbol{A} \,|\psi(\boldsymbol{x}) \ge t\}$ is measurable. This follows from the fact that

$$\{\boldsymbol{x} \in \boldsymbol{A} \,|\, \psi(\boldsymbol{x}) \ge t\} = \bigcap_{k=0}^{\infty}\{\boldsymbol{x} \in \boldsymbol{A} \,|\, f_k(\boldsymbol{x}) \ge t\},$$

adapting the argument used in the proof of **(a)**. We now prove **(b)**. We must prove that γ is measurable.

Recalling the definition of limit inferior, we have

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \gamma(\boldsymbol{x}) = \liminf_{k\to\infty} f_k(\boldsymbol{x}) = \sup_{k\in\mathbb{N}} \inf_{h\geq k} f_k(\boldsymbol{x}).$$

Given $k \in \mathbb{N}$, using the second of **(a)**, the function

$$\begin{aligned} g_k : \boldsymbol{A} &\longrightarrow [-\infty, +\infty] \\ \boldsymbol{x} &\mapsto \inf_{h\geq k} f_h(\boldsymbol{x}) \end{aligned}$$

is measurable. On the other hand, using the first of **(a)** we see immediately that $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x})$ (which is exactly γ) is also measurable. The proof that ρ is also measurable follows the same argument.

We now prove **(c)**. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ satisfy such

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

We show that

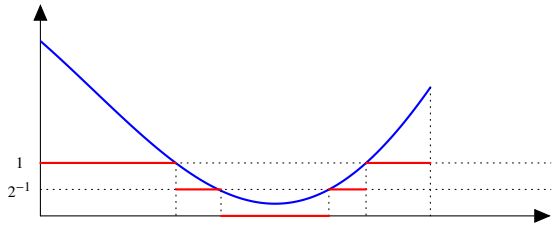
$$f(\boldsymbol{x}) = \liminf_{k\to\infty} f_k(\boldsymbol{x}) = \sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}), \qquad \boldsymbol{x} \in \boldsymbol{A}. \tag{3.31}$$

This is enough to complete the proof, because we have already shown that the function γ is measurable. Consider first the case $f(\boldsymbol{x}) = +\infty$. Then, given $M \in \mathbb{R}$, there exists $p \in \mathbb{N}$ such that, for $k > p$, we have $f_k(\boldsymbol{x}) \geq M$. It follows that, for $k > p$, there holds $g_k(\boldsymbol{x}) \geq M$, and therefore $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) \geq M$ for every $M \in \mathbb{R}$. This implies $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) = +\infty$. Now, consider the case $f(\boldsymbol{x}) = -\infty$. Then, $M \in \mathbb{R}$, there exists $p \in \mathbb{N}$ such that, for $k > p$, we have $f_k(\boldsymbol{x}) \leq M$. It follows that, for $k > p$, there holds $g_k(\boldsymbol{x}) \leq M$, and therefore $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) \leq M$ for every $M \in \mathbb{R}$. This implies $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) = -\infty$. Finally, consider the case $f(\boldsymbol{x}) \in \mathbb{R}$. Let $\varepsilon > 0$. Then, there exists $p \in \mathbb{N}$ such that, for $k > p$, we have $f(\boldsymbol{x}) - \varepsilon < f_k(\boldsymbol{x}) < f(\boldsymbol{x}) + \varepsilon$. It follows that, for $k > p$, we have $f(\boldsymbol{x}) - \varepsilon < g_k(\boldsymbol{x}) < f(\boldsymbol{x}) + \varepsilon$, so that $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) > f(\boldsymbol{x}) - \varepsilon$. Moreover, for every $k \in \mathbb{N}$ we have $g_k(\boldsymbol{x}) \leq g_{k+1}(\boldsymbol{x})$, and this implies $g_k(\boldsymbol{x}) < f(\boldsymbol{x}) + \varepsilon$ for every $k \in \mathbb{N}$, so that $\sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) < f(\boldsymbol{x}) + \varepsilon$. We have obtained

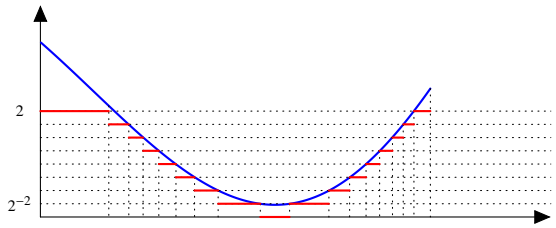
$$f(\boldsymbol{x}) - \varepsilon < \sup_{k\in\mathbb{N}} g_k(\boldsymbol{x}) < f(\boldsymbol{x}) + \varepsilon$$

for every $\varepsilon > 0$. Therefore, in every case (3.31) holds.

In order to define the Lebesgue integral for a class of functions wider than the set of simple functions, we investigate the approximation of a measurable function f with a non-decreasing sequence of non-negative simple functions $(\varphi_k)_{k\in\mathbb{N}}$. An example of construction of this sequence is given in Theorem 3.29. With this construction, when $k = 0$, the function φ_0 is zero. For $k = 1$ and $k = 2$, the functions φ_1 and φ_2 are, respectively, shown in the figure below.



$$\varphi_1(x) = \begin{cases} 0 & \text{if } 0 \leq f(x) < 2^{-1} \\ 2^{-1} & \text{if } 2^{-1} \leq f(x) < 1 \\ 1 & \text{if } f(x) \geq 1 \end{cases}$$



$$\varphi_2(x) = \begin{cases} 0 & \text{if } 0 \le f(\boldsymbol{x}) < 2^{-2} \\ 2^{-2} & \text{if } 2^{-2} \le f(x) < 2 \cdot 2^{-2} \\ 2 \cdot 2^{-2} & \text{if } 2 \cdot 2^{-2} \le f(x) < 3 \cdot 2^{-2} \\ 3 \cdot 2^{-2} & \text{if } 3 \cdot 2^{-2} \le f(x) < 4 \cdot 2^{-2} \\ 4 \cdot 2^{-2} & \text{if } 4 \cdot 2^{-2} \le f(x) < 5 \cdot 2^{-2} \\ 5 \cdot 2^{-2} & \text{if } 5 \cdot 2^{-2} \le f(x) < 6 \cdot 2^{-2} \\ 6 \cdot 2^{-2} & \text{if } 6 \cdot 2^{-2} \le f(x) < 7 \cdot 2^{-2} \\ 7 \cdot 2^{-2} & \text{if } 7 \cdot 2^{-2} \le f(x) < 2 \\ 2 & \text{if } f(x) \ge 2 \end{cases}$$

With this idea in mind, we can better understand what we observed in our preliminary description of the Lebesgue integration: while to compute the Riemann integral of a function, one partitions the domain into subintervals, in the Lebesgue integral, we are in effect partitioning the range of the function.

Theorem 3.29 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$. The following statements are equivalent:*

(I) f is measurable;

(II) there exists a sequence of simple functions $(\varphi_k)_{k\in\mathbb{N}}$ where, for all $k \in \mathbb{N}$, we have $\varphi_k : \boldsymbol{A} \longrightarrow \mathbb{R}$ and such that $0 \le \varphi_k(\boldsymbol{x}) \le \varphi_{k+1}(\boldsymbol{x})$ for all $k \in \mathbb{N}$, and

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x}).$$

Proof From **(c)** of Theorem 3.28 and Remark 3.8, we see immediately that (II) implies (I). We now show that (I) implies (II). Let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable and let $k \in \mathbb{N}$. We define, for $\boldsymbol{x} \in A$,

$$\varphi_k(\boldsymbol{x}) = \begin{cases} 0 & \text{if } 0 \le f(\boldsymbol{x}) < 2^{-k} \\ 2^{-k} & \text{if } 2^{-k} \le f(\boldsymbol{x}) < 2 \cdot 2^{-k} \\ 2 \cdot 2^{-k} & \text{if } 2 \cdot 2^{-k} \le f(\boldsymbol{x}) < 3 \cdot 2^{-k} \\ 3 \cdot 2^{-k} & \text{if } 3 \cdot 2^{-k} \le f(\boldsymbol{x}) < 4 \cdot 2^{-k} \\ \vdots & \\ (k\,2^k - 1) \cdot 2^{-k} & \text{if } (k\,2^k - 1) \cdot 2^{-k} \le f(\boldsymbol{x}) < k \\ k & \text{if } f(\boldsymbol{x}) \ge k \end{cases}$$

or, more succinctly,

$$\varphi_k(\boldsymbol{x}) = \begin{cases} j\,2^{-k} & \text{if } j \in \{0, 1, \ldots, k\,2^k - 1\} \text{ and } j\,2^{-k} \le f(\boldsymbol{x}) < (j+1)\,2^{-k} \\ k & \text{if } f(\boldsymbol{x}) \ge k \end{cases}$$

We observe that, for every $k \in \mathbb{N}$, the function φ_k is a simple function over $\boldsymbol{A}$. Indeed, let us define for $j \in \{0, \ldots, k\,2^k - 1\}$ the set $\boldsymbol{A}_{k,j} = \{\boldsymbol{x} \in \boldsymbol{A} \mid j\,2^{-k} \le f(\boldsymbol{x}) < (j+1)\,2^{-k}\}$ and $\boldsymbol{A}_{k,k\,2^k} = \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \ge k\}$. For $j \in \{0, \ldots, k\,2^k - 1\}$, the set $\boldsymbol{A}_{k,j}$ is measurable because it coincides with the intersection

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid j\,2^{-k} \le f(\boldsymbol{x})\} \cap \{\boldsymbol{x} \in A \mid f(\boldsymbol{x}) < (j+1)\,2^{-k}\},$$

from Theorem 3.26. Moreover, by construction, for all $j \in \{0, \ldots, k\,2^k\}$, the sets $\boldsymbol{A}_{k,j}$ are pairwise disjoint and their union is $\boldsymbol{A}$.

We prove that for every $\boldsymbol{x} \in \boldsymbol{A}$ and for every $k \in \mathbb{N}$, we have $\varphi_k(\boldsymbol{x}) \le \varphi_{k+1}(\boldsymbol{x})$. We distinguish different cases:

(i) We start assuming $f(\boldsymbol{x}) \ge k + 1$. In this case, $\varphi_k(\boldsymbol{x}) = k$ and $\varphi_{k+1}(\boldsymbol{x}) = k + 1$.

(ii) Suppose $k \leq f(\boldsymbol{x}) < k+1$. Then, $\varphi_k(\boldsymbol{x}) = k$. By definition of φ_k we have

$$\varphi_{k+1}(\boldsymbol{x}) = \begin{cases} j\,2^{-k-1} & \text{if } j \in \{0,1,\ldots,(k+1)\,2^{k+1}-1\} \text{ and } j\,2^{-k-1} \leq f(\boldsymbol{x}) < (j+1)\,2^{-k-1} \\ k+1 & \text{if } f(\boldsymbol{x}) \geq k+1. \end{cases}$$

Since $f(\boldsymbol{x}) < k+1$, we will have $\varphi_{k+1}(\boldsymbol{x}) = j\,2^{-k-1}$ for some $j \in \{0,1,\ldots,(k+1)\,2^{k+1}-1\}$. On the other hand, we know that $f(\boldsymbol{x}) \geq k$, so that $k \leq j\,2^{-k-1}$, which implies $j \geq k\,2^{k+1}$. It follows also that $\varphi_{k+1}(\boldsymbol{x}) = j\,2^{-k-1} \geq k\,2^{k+1}\,2^{-k-1} = k = \varphi_k(\boldsymbol{x})$.

(iii) Finally, suppose $f(\boldsymbol{x}) < k$. If $j \in \{0,\ldots,k\,2^k-1\}$ is such that $j\,2^{-k} \leq f(\boldsymbol{x}) < (j+1)\,2^{-k}$, we have $\varphi_k(\boldsymbol{x}) = j\,2^{-k}$. Two cases are possible: $j\,2^{-k} \leq f(\boldsymbol{x}) < j\,2^{-k}+2^{-k-1}$ or $j\,2^{-k}+2^{-k-1} \leq f(\boldsymbol{x}) < (j+1)\,2^{-k}$. In the first case, $(2\,j)\,2^{-k-1} \leq f(\boldsymbol{x}) < (2\,j+1)\,2^{-k-1}$, so that $\varphi_{k+1}(\boldsymbol{x}) = (2\,j)\,2^{-k-1} = j\,2^{-k} = \varphi_k(\boldsymbol{x})$. In the second case, $(2\,j+1)\,2^{-k-1} \leq f(\boldsymbol{x}) < (2\,j+2)\,2^{-k-1}$, so that $\varphi_{k+1}(\boldsymbol{x}) = (2\,j+1)\,2^{-k-1} = \varphi_k(\boldsymbol{x}) + 2^{-k-1}$. Hence, in either case $\varphi_k(\boldsymbol{x}) \leq \varphi_{k+1}(\boldsymbol{x})$.

We finally prove that for all $\boldsymbol{x} \in \boldsymbol{A}$ we have $\lim\limits_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x})$. We observe that the existence of the limit is ensured by the fact that the sequence $(\varphi_k(\boldsymbol{x}))_{k\in\mathbb{N}}$ is non-decreasing. First, suppose that $f(\boldsymbol{x}) = +\infty$. Then, $\varphi_k(\boldsymbol{x}) = k$ for all $k \in \mathbb{N}$, which proves the claim. Now, suppose $f(\boldsymbol{x}) < +\infty$. Let $k \in \mathbb{N}$ be such that $f(\boldsymbol{x}) < k$. Then, $\varphi_k(\boldsymbol{x}) = j\,2^{-k}$ for some $j \in \mathbb{N}$ such that $j\,2^{-k} \leq f(\boldsymbol{x}) < (j+1)\,2^{-k}$. Subtracting $\varphi_k(\boldsymbol{x}) = j\,2^{-k}$ from this inequality yields $j\,2^{-k} - j\,2^{-k} \leq f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x}) < (j+1)\,2^{-k} - j\,2^{-k}$, i.e.,

$$0 \leq f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x}) < 2^{-k} \overset{k\to\infty}{\longrightarrow} 0.$$

The sequence of functions $(\varphi_k)_{k\in\mathbb{N}}$ constructed in Theorem 3.29 is dominated by f, i.e.,

$$\forall\, k \in \mathbb{N},\ \forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad |\varphi_k(\boldsymbol{x})| \leq |\varphi_{k+1}(\boldsymbol{x})| \leq |f(\boldsymbol{x})|.$$

Remark 3.9 In the proof of Theorem 3.29 we have seen that, if f is real-valued and upper bounded, the sequence $(\varphi_k)_{k\in\mathbb{N}}$ can be also chosen to be uniformly convergent.

Example *3.20*

Consider the function

$$\begin{aligned} f : [-1,1] &\longrightarrow \mathbb{R}_+ \\ x &\mapsto x^2. \end{aligned}$$

We know that f is measurable by using the result of Theorem 3.29, by constructing a sequence of non-negative simple functions which is pointwise convergent to f. First, notice that the image of f is $[0,1]$. We also observe that, for $k \in \mathbb{N}^*$

$$\begin{aligned} [0,1] &= \left[0,\frac{1}{k}\right) \cup \left[\frac{1}{k},\frac{2}{k}\right) \cup \ldots \cup \left[\frac{j-1}{k},\frac{j}{k}\right) \cup \ldots \cup \left[\frac{k-1}{k},1\right) \\ &= \left(\bigcup_{j=1}^{k-1} \left[\frac{j-1}{k},\frac{j}{k}\right)\right) \cup \left[\frac{k-1}{k},1\right) \end{aligned}$$

For $j \in \{1, 2, \ldots, k\}$, we define

$$A_{k,j} = \begin{cases} \left\{x \in [-1,1] \,|\, x^2 \in \left[\frac{j-1}{k}, \frac{j}{k}\right)\right\} & \text{if } j \in \{1,2,\ldots,k-1\} \\ \left\{x \in [-1,1] \,|\, x^2 \in \left[\frac{k-1}{k}, 1\right)\right\} & \text{if } j = k \end{cases}$$

$$= \begin{cases} \left(-\sqrt{\frac{j}{k}}, -\sqrt{\frac{j-1}{k}}\right] \cup \left[\sqrt{\frac{j-1}{k}}, \sqrt{\frac{j}{k}}\right) & \text{if } j \in \{1,2,\ldots,k-1\} \\ \left[-1, -\sqrt{\frac{k-1}{k}}\right] \cup \left[\sqrt{\frac{k-1}{k}}, 1\right] & \text{if } j = k \end{cases}$$

It follows that the sets $A_{k,j}$ are pairwise disjoint, their union is $[-1,1]$ and they are measurable. We define, for every $k \in \mathbb{N}^*$, the simple function $\varphi_k : [-1,1] \longrightarrow \mathbb{R}$ such that $\varphi_k(x) = \frac{j-1}{k}$ if $x \in A_{k,j}$

$$\begin{aligned} \varphi_k : [-1,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \frac{j-1}{k} & \text{if } x \in A_{k,j} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Let $x \in [0,1]$, and let $k \in \mathbb{N}^*$. Then, for some $j \in \{1, \ldots, k\}$, we have $x \in A_{k,j}$. This implies that

$$f(x) \in \left[\frac{j-1}{k}, \frac{j}{k}\right) \quad \text{if} \quad j \leq k-1$$

and

$$f(x) \in \left[\frac{k-1}{k}, 1\right) \quad \text{if} \quad j \leq k.$$

In both cases, we have

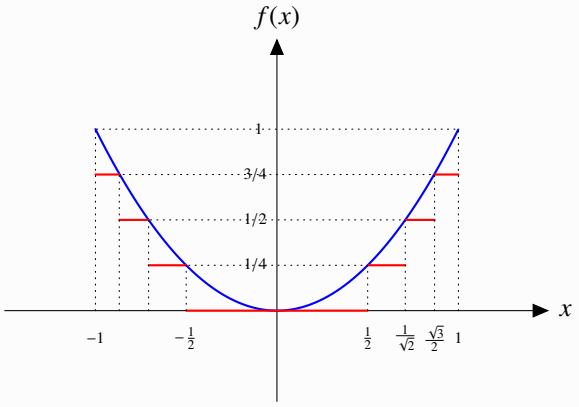
$$0 \leq f(x) - \varphi_k(x) \leq \frac{1}{k},$$

so that for every $x \in [-1,1]$

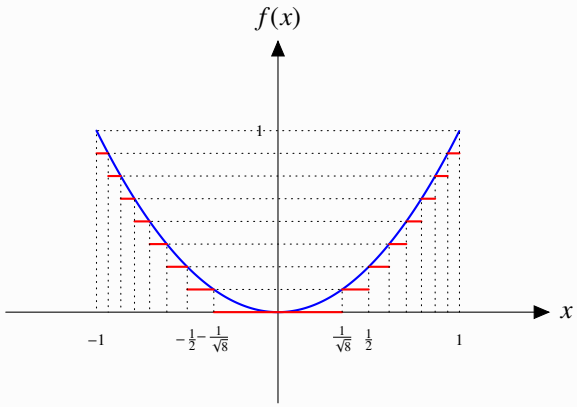
$$\lim_{k \to \infty} \varphi_k(x) = f(x).$$

For example, for $k = 4$ we have drawn the graph of φ_k.

$$\begin{aligned} A_{4,1} &= \left(-\tfrac{1}{2}, 0\right] \cup \left[0, \tfrac{1}{2}\right) \\ A_{4,2} &= \left(-\tfrac{\sqrt{2}}{2}, -\tfrac{1}{2}\right] \cup \left[\tfrac{1}{2}, \tfrac{\sqrt{2}}{2}\right) \\ A_{4,3} &= \left(-\tfrac{\sqrt{3}}{2}, -\tfrac{\sqrt{2}}{2}\right] \cup \left[\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{3}}{2}\right) \\ A_{4,4} &= \left(-1, -\tfrac{\sqrt{3}}{2}\right] \cup \left[\tfrac{\sqrt{3}}{2}, 1\right) \end{aligned}$$



For $k = 8$, we have the following graph of φ_k.



It is clear that this approach to test if a function is measurable is not very efficient in practice. We will have at our disposal more flexible conditions. For example, we will be able to simply use the continuity of f to conclude that f is measurable.

We now want to study the possibility of approximating a measurable function f with simple functions, without assuming f to be non-negative. Before we do it, we introduce the following conventions:

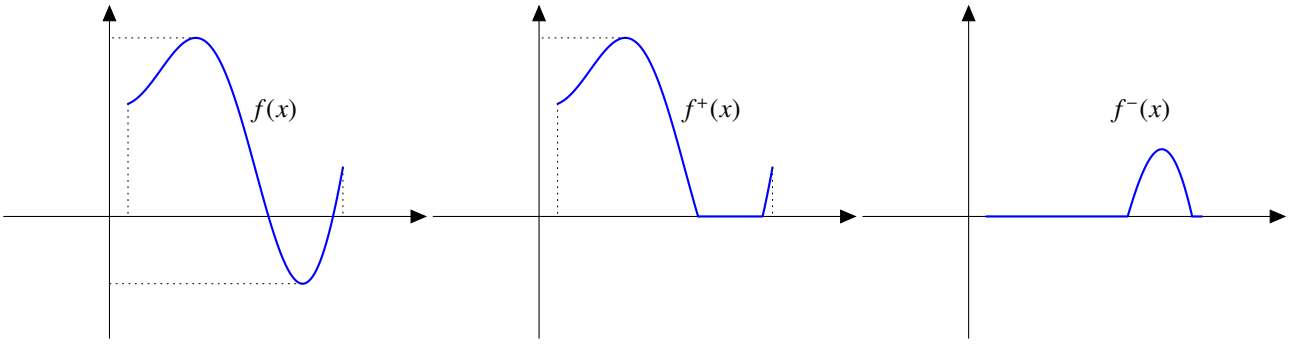
$$|-\infty| = +\infty.$$

Definition 3.14 [Positive and Negative Parts]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ be measurable. We define the positive part *and the* negative part *of f as the functions*

$$\begin{array}{rcl} f^+ : \boldsymbol{A} & \longrightarrow & [-\infty, +\infty] \\ \boldsymbol{x} & \mapsto & \max\{f(\boldsymbol{x}), 0\} \end{array} \quad \textit{and} \quad \begin{array}{rcl} f^- : \boldsymbol{A} & \longrightarrow & [-\infty, +\infty] \\ \boldsymbol{x} & \mapsto & \max\{-f(\boldsymbol{x}), 0\}, \end{array}$$

respectively.

Graphically, the positive and the negative parts of a function can be visualized as in the figure:



The following result holds.

Lemma 3.4 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ be measurable. Then:*

(a) *f^+ and f^- take values in $[0, +\infty]$;*

(b) *$\forall\, \boldsymbol{x} \in \boldsymbol{A}$, $f^+(\boldsymbol{x}) + f^-(\boldsymbol{x}) = |f(\boldsymbol{x})|$;*

(c) $\forall\, \boldsymbol{x} \in \boldsymbol{A},\ f^+(\boldsymbol{x}) - f^-(\boldsymbol{x}) = f(\boldsymbol{x})$;

(d) *f^+ and f^- are measurable;*

Proof The proof of **(a)** is a direct consequence of Definition 3.14.
We prove **(b)**. If $f(\boldsymbol{x}) \geq 0$, we have $f^+(\boldsymbol{x}) + f^-(\boldsymbol{x}) = f(\boldsymbol{x}) + 0 = f(\boldsymbol{x}) = |f(\boldsymbol{x})|$. If $f(\boldsymbol{x}) < 0$, then $f^+(\boldsymbol{x}) + f^-(\boldsymbol{x}) = 0 + (-f(\boldsymbol{x})) = -f(\boldsymbol{x}) = |f(\boldsymbol{x})|$.
We prove **(c)**. First, we verify for every $\boldsymbol{x} \in \boldsymbol{A}$, the quantity $f^+(\boldsymbol{x}) - f^-(\boldsymbol{x})$ is defined. We have to prove that $f^+(\boldsymbol{x})$ and $f^-(\boldsymbol{x})$ cannot be $+\infty$ simultaneously. Indeed, if $f^+(\boldsymbol{x}) = +\infty$, necessarily $f(\boldsymbol{x}) = +\infty$, and therefore $f^-(\boldsymbol{x}) = 0$. Let $f(\boldsymbol{x}) \geq 0$. Then, $f^+(\boldsymbol{x}) - f^-(\boldsymbol{x}) = f(\boldsymbol{x}) - 0 = f(\boldsymbol{x})$. If, on the contrary, $f(\boldsymbol{x}) \leq 0$, then $f^+(\boldsymbol{x}) - f^-(\boldsymbol{x}) = 0 - (-f(\boldsymbol{x})) - f(\boldsymbol{x})$.
We prove **(d)**. For every $t \in \mathbb{R}$,

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid f^+(\boldsymbol{x}) \leq t\} = \begin{cases} \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \leq t\} & \text{if } t \geq 0 \\ \varnothing & \text{if } t < 0 \end{cases} \tag{3.32}$$

and

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid f^-(\boldsymbol{x}) \leq t\} = \begin{cases} \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \geq -t\} & \text{if } t \geq 0 \\ \varnothing & \text{if } t < 0 \end{cases} \tag{3.33}$$

Now, **(d)** is a consequence of Theorem 3.26.

The following result can be seen as a generalization of Theorem 3.29 to functions that do not necessarily take values in $[0, +\infty]$, but in $[-\infty, +\infty]$. In this case, the sequence of simple functions $(\varphi_k)_{k\in\mathbb{N}}$ is not monotonic in general, and its terms are not necessarily non-negative functions.

Theorem 3.30 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$. The following statements are equivalent:*

(I) f is measurable;

(II) there exists a sequence of simple functions $(\varphi_k)_{k\in\mathbb{N}}$, with $\varphi_k : \boldsymbol{A} \longrightarrow \mathbb{R}$ for all $k \in \mathbb{N}$, satisfying

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x}).$$

The sequence $(\varphi_k)_{k\in\mathbb{N}}$ can be chosen in such a way that $|\varphi_k(\boldsymbol{x})| \leq |f(\boldsymbol{x})|$ for every $k \in \mathbb{N}$ and for every $\boldsymbol{x} \in \boldsymbol{A}$. Moreover, if f is bounded and real-valued, $(\varphi_k)_{k\in\mathbb{N}}$ can be chosen to be uniformly convergent.

Proof From **(c)** of Theorem 3.28, we see immediately that (II) implies (I). We now show that (I) implies (II).
We apply Theorem 3.29 and Lemma 3.4 to conclude that there exist two sequences $(\psi_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ of simple functions with domain $\boldsymbol{A}$ such that, for every $\boldsymbol{x} \in \boldsymbol{A}$ and for every $k \in \mathbb{N}$ we have $0 \leq \psi_k(\boldsymbol{x}) \leq \psi_{k+1}(\boldsymbol{x})$, $0 \leq g_k(\boldsymbol{x}) \leq g_{k+1}(\boldsymbol{x})$, $\lim_{k\to\infty} \psi_k(\boldsymbol{x}) = f^+(\boldsymbol{x})$, $\lim_{k\to\infty} g_k(\boldsymbol{x}) = f^-(\boldsymbol{x})$. We define $\varphi_k(\boldsymbol{x}) = \psi_k(\boldsymbol{x}) - g_k(\boldsymbol{x})$. The functions φ_k are simple. For every $\boldsymbol{x} \in \boldsymbol{A}$, $\lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f^+(\boldsymbol{x}) - f^-(\boldsymbol{x}) = f(\boldsymbol{x})$, $|\varphi_k(\boldsymbol{x})| \leq \psi_k(\boldsymbol{x}) + g_k(\boldsymbol{x}) \leq f^+(\boldsymbol{x}) + f^-(\boldsymbol{x}) = |f(\boldsymbol{x})|$.
Finally, if f is real-valued and bounded, so are f^+ and f^-. From Remark 3.9, the convergence of $(\varphi_k)_{k\in\mathbb{N}}$ and $(\psi_k)_{k\in\mathbb{N}}$ to $f^+(\boldsymbol{x})$ and $f^-(\boldsymbol{x})$ is uniform.

The following result presents some properties of measurable functions. For example, the sum of measurable functions is a measurable function. Some care needs to be taken to ensure that the operations in this result are well defined. Indeed, the definition of measurability of a function allowed the codomain of that function to be $[-\infty, +\infty]$. Clearly, if for a certain $\boldsymbol{x} \in \boldsymbol{A}$ we have $f(\boldsymbol{x}) = +\infty$ and $g(\boldsymbol{x}) = -\infty$, we encounter a problem: we want to avoid the operation $+\infty - \infty$. Hence, the following result is presented under some conservative assumptions which ensure that the operations involved there are well defined.

Corollary 3.2 *Let $A \in \mathcal{M}_n$. Let $f, g : A \longrightarrow \mathbb{R}$ be measurable. Let $h : \mathbb{R} \longrightarrow \mathbb{R}$ be continuous, $c \in \mathbb{R}$. Then:*

(a) *$h \circ f$ is measurable;*

(b) *$f + g$, $f\,g$, $c\,f$ are measurable;*

(c) *if $g(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \boldsymbol{A}$, the function $\frac{f}{g}$ is measurable.*

Proof From Theorem 3.30, there exist two sequences of simple functions $(\varphi_k)_{k\in\mathbb{N}}$ and $(\psi_k)_{k\in\mathbb{N}}$ over $\boldsymbol{A}$ such that $\lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x})$ and $\lim_{k\to\infty} \psi_k(\boldsymbol{x}) = g(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$. From **(c)** in Theorem 3.23, for every $k \in \mathbb{N}$, the function $h \circ \varphi_k$ is simple. Since h is continuous, for every $\boldsymbol{x} \in \boldsymbol{A}$ we have

$$(h \circ f)(\boldsymbol{x}) = \lim_{k\to\infty} (h \circ \varphi_k)(\boldsymbol{x}).$$

We found a sequence $(h \circ \varphi_k)_{k\in\mathbb{N}}$ of simple functions which is pointwise convergent to $h \circ f$. Hence, from Theorem 3.30, the function $h \circ f$ is measurable. This proves **(a)**.
Moreover, for every $k \in \mathbb{N}$, the sum $\varphi_k + \psi_k$ is simple from **(a)** in Theorem 3.23, and for every $\boldsymbol{x} \in \boldsymbol{A}$

$$f(\boldsymbol{x}) + g(\boldsymbol{x}) = \lim_{k\to\infty} (\varphi_k(\boldsymbol{x}) + \psi_k(\boldsymbol{x})).$$

Using again Theorem 3.30, we find that $f + g$ is measurable. The remaining part of **(b)** can be proved in the same way.
We prove **(c)**. The issue here is that the assumption $g(\boldsymbol{x}) \neq 0$ for all $\boldsymbol{x} \in \boldsymbol{A}$ does not imply that $\psi_k(\boldsymbol{x})$ is non-zero for every $k \in \mathbb{N}$ and for every $\boldsymbol{x} \in \boldsymbol{A}$. We therefore define, for every $k \in \mathbb{N}^*$, the function

$$\begin{aligned} \gamma_k : \boldsymbol{A} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \psi_k^2(\boldsymbol{x}) + \frac{1}{k}. \end{aligned}$$

For each $k \in \mathbb{N}^*$, the function γ_k is simple (because such is ψ_k) and also non-zero. Moreover, for every $\boldsymbol{x} \in \boldsymbol{A}$,

$$\frac{f(\boldsymbol{x})}{g(\boldsymbol{x})} = \lim_{k\to\infty} \frac{\varphi_k(\boldsymbol{x})\,\psi_k(\boldsymbol{x})}{\psi_k^2(\boldsymbol{x}) + \frac{1}{k}} \lim_{k\to\infty} \frac{\varphi_k(\boldsymbol{x})\,\psi_k(\boldsymbol{x})}{\gamma_k(\boldsymbol{x})},$$

which concludes the proof if we invoke Theorem 3.30, because we wrote f/g as the pointwise limit of a sequence of simple functions $\left(\frac{\varphi_k(\boldsymbol{x})\,\psi_k(\boldsymbol{x})}{\gamma_k(\boldsymbol{x})}\right)_{k\in\mathbb{N}^*}$.

In particular, the previous result has shown that the set of measurable functions is a subspace of the set of functions from $\boldsymbol{A}$ to $\mathbb{R}$.

Remark 3.10 The measurability of $f + g$ as a consequence of the measurability of f and g holds under assumptions that are weaker than those of Corollary 3.2. We can require f and g to assume the values $\pm\infty$, provided that

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) = +\infty\} \cap \{\boldsymbol{x} \in \boldsymbol{A} \mid g(\boldsymbol{x}) = -\infty\} = \varnothing$$
$$\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) = -\infty\} \cap \{\boldsymbol{x} \in \boldsymbol{A} \mid g(\boldsymbol{x}) = +\infty\} = \varnothing$$

The proof is the same as that of Corollary 3.2. On the contrary, the product $f\,g$ is always defined if we assume $0 \cdot (\pm\infty) = 0$.

Another way to obtain measurable functions is to "glue" measurable functions together in the following sense.

Theorem 3.31 *For all $k \in \mathbb{N}$, let $\boldsymbol{A}_k \in \mathcal{M}_n$. Let $\boldsymbol{A} = \bigcup_{k=0}^{\infty} A_k$ and $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ such that for every $k \in \mathbb{N}$, the restriction $f|_{A_k}$ is measurable over $\boldsymbol{A}_k$. Then, f is measurable over $\boldsymbol{A}$.*

Proof We have $\boldsymbol{A} \in \mathcal{M}_n$, because it is the union of a countable sequence of elements of $\mathcal{M}_n$. Let $t \in \mathbb{R}$. Clearly,

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \leq t\} = \bigcup_{k=0}^{\infty} \{\boldsymbol{x} \in \boldsymbol{A}_k \mid f(\boldsymbol{x}) \leq t\},$$

and therefore it belongs to $\mathcal{M}_n$. Hence, f is measurable.

Remark 3.11 In Theorem 3.31, we did not assume $\boldsymbol{A}_k$ to be pairwise disjoint. The result can be used to prove a similar result for the case $\boldsymbol{A} = \boldsymbol{A}_1 \cup \ldots \cup \boldsymbol{A}_m$, where $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathcal{M}_n$. It is sufficient to set, for example, $\boldsymbol{A}_k = \boldsymbol{A}_m$ for every $k > m$.

We conclude this section with the following simple condition for measurability.

Theorem 3.32 *Let $\boldsymbol{A} \in \mathcal{M}_n$ and let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ be continuous. Then, f is measurable.*

Proof We recall that, from Theorem 3.21, open sets are measurable. To prove the result, we show that for every $t \in \mathbb{R}$, the set $\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) < t\}$ is the intersection of $\boldsymbol{A}$ with a certain open set Ω. We conclude that f is measurable from Theorem 3.26.

Let $\boldsymbol{x} \in \boldsymbol{A}$ be such that $f(\boldsymbol{x}) < t$. From the continuity of f, there exists $\delta_x > 0$ such that, if $\boldsymbol{y} \in \boldsymbol{A} \cap \mathcal{B}_{\delta_x}(\boldsymbol{x})$, it is still true that $f(\boldsymbol{y}) < t$. We define

$$\boldsymbol{\Omega} = \bigcup_{\substack{\boldsymbol{x} \in \boldsymbol{A} \\ f(\boldsymbol{x}) < t}} \mathcal{B}_{\delta_x}(\boldsymbol{x}).$$

The set $\boldsymbol{\Omega}$ is open as the union of open sets, and therefore it is measurable. Moreover, by construction, $\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) < t\} = \boldsymbol{A} \cap \boldsymbol{\Omega}$ (why?), which is measurable because $\boldsymbol{A}$ and $\boldsymbol{\Omega}$ are both measurable.

Example *3.21*

Let $\boldsymbol{B} \subseteq \boldsymbol{A} \subseteq \mathbb{R}^n$ and $\boldsymbol{A} \in \mathcal{M}_n$. Verify that $f : \boldsymbol{A} \longrightarrow \mathbb{R}$, $f(\boldsymbol{x}) = \chi_{\boldsymbol{B}}(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$ is measurable if and only if $\boldsymbol{B} \in \mathcal{M}_n$.

Exercises

Exercise 3.30

Let $a, t \in \mathbb{R}$. Show that

$$a \geq t \qquad \Leftrightarrow \qquad \forall\, k \in \mathbb{N}^*, \;\; a > t - \frac{1}{k}.$$

[Solution: (3.28)]

3.8 Integration of non-negative measurable functions

We define the integral of a non-negative measurable function. The idea is to approximate the function from below using simple functions.

Definition 3.15 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable. We define the* Lebesgue integral *to be*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \stackrel{def}{=} \sup\left\{\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \,\middle|\, \varphi : \boldsymbol{A} \longrightarrow \mathbb{R} \text{ simple and } \forall\, \boldsymbol{x} \in \boldsymbol{A},\ \ 0 \le \varphi(\boldsymbol{x}) \le f(\boldsymbol{x})\right\}.$$

Remark 3.12 The integral of φ in Definition 3.15 has to be intended in the sense of Definition 3.12. We observe that in the case where f is simple and non-negative, Definition 3.15 coincides with Definition 3.12. To show this, we denote for the moment by I_1 the integral in the sense of Definition 3.12, and by I_2 the integral in the sense of Definition 3.15. Since, obviously, $0 \le \varphi \le \varphi$, we have $I_1 \le I_2$. On the other hand, from **(a)** of Theorem 3.24, if $\psi : \boldsymbol{A} \longrightarrow \mathbb{R}$ is simple and $0 \le \psi \le \varphi$, we have

$$\int_{\boldsymbol{A}} \psi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le I_1,$$

so that $I_2 \le I_1$.

We now consider the first properties of the integral.

Theorem 3.33 *Let $\boldsymbol{A} \in \mathcal{M}_n$, and let $f, g : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable. Then:*

(a) $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \in [0, +\infty]$;

(b) $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$ *if and only if* $\mu_n(\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\}) = 0$;

(c) *if $f \le g$, then* $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

Proof The proof of **(a)** is straightforward. Consider **(b)**. Let $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$, and let us denote the set $\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\}$ by $\boldsymbol{B}$. From Theorem 3.26, $\boldsymbol{B}$ is measurable. Moreover,

$$\boldsymbol{B} = \bigcup_{k=1}^{\infty} \boldsymbol{B}_k, \quad \text{where} \quad \boldsymbol{B}_k = \left\{\boldsymbol{x} \in \boldsymbol{A} \,\middle|\, f(\boldsymbol{x}) > \frac{1}{k}\right\}.$$

Indeed, if $\boldsymbol{x} \in \boldsymbol{B}$, we have $f(\boldsymbol{x}) > 0$, and taking $\ell < \frac{1}{f(\boldsymbol{x})}$ ensures that $f(\boldsymbol{x}) > 1/\ell$, so that $\boldsymbol{x} \in \boldsymbol{B}_\ell$, and therefore it belongs to the union of the $\boldsymbol{B}_k$. Conversely, if $\boldsymbol{x} \in \boldsymbol{B}_\ell$ for some $\ell \in \mathbb{N}^*$, then $f(\boldsymbol{x}) > 1/\ell > 0$, so that $\boldsymbol{x} \in \boldsymbol{B}$. Moreover, $\boldsymbol{B}_k \subseteq \boldsymbol{B}_{k+1}$ for every $k \in \mathbb{N}^*$. Suppose by contradiction that $\mu_n(\boldsymbol{B}) > 0$. From **(a)** of Theorem 3.11, we find

$$\mu_n(\boldsymbol{B}) = \lim_{k \to \infty} \mu_n(\boldsymbol{B}_k).$$

Therefore, there exists $k \in \mathbb{N}^*$ such that $\mu_n(\boldsymbol{B}_k) > 0$. Consider the function $\varphi : \boldsymbol{A} \longrightarrow \mathbb{R}$ defined by $\varphi(\boldsymbol{x}) = \frac{1}{k}\chi_{\boldsymbol{B}_k}(\boldsymbol{x})$ for $\boldsymbol{x} \in \boldsymbol{A}$. Clearly, φ is simple and $0 \le \varphi \le f$ (recall Remark 3.7). It follows that

$$0 = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \ge \int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \frac{1}{k}\,\mu_n(\boldsymbol{B}_k) > 0.$$

Hence, $\mu_n(\boldsymbol{B})$ must be 0.
Conversely, let $\mu_n(\boldsymbol{B}) = 0$. Let $\varphi : \boldsymbol{A} \longrightarrow \mathbb{R}$ be a simple function such that $0 \le \varphi \le f$. Clearly, $\{\boldsymbol{x} \in \boldsymbol{A} \mid \varphi(\boldsymbol{x}) >$

$0\} \subseteq \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > 0\} = \boldsymbol{B}$, and therefore its measure is 0. From **(b)** in Theorem 3.24, we obtain

$$\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0.$$

From the arbitrariness of φ, we conclude that $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$.

To prove **(c)**, we observe that

$$\left\{\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,\middle|\, \varphi : \boldsymbol{A} \longrightarrow \mathbb{R}, \varphi \text{ is simple and } 0 \le \varphi \le f\right\} \subseteq \left\{\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\,\middle|\, \varphi : \boldsymbol{A} \longrightarrow \mathbb{R}, \varphi \text{ is simple and } 0 \le \varphi \le g\right\}.$$

The supremum of the second set is at least equal to the supremum of the first set.

To find more properties of the Lebesgue integral, we will utilize an important result on the interchange of limit and integral: the so-called *monotone convergence theorem*.

3.8.1 Beppo Levi's monotone convergence theorem

Now we state some basic results of interchange of limits with integrals. Let $\boldsymbol{A} \in \mathcal{M}_n$. For $k \in \mathbb{N}$, let $f_k : \boldsymbol{A} \longrightarrow [0, +\infty]$ and $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be such that, for almost every $\boldsymbol{x} \in \boldsymbol{A}$,

$$\lim_{k \to \infty} f_k(\boldsymbol{x}) = f(\boldsymbol{x}).$$

Assuming that the functions f_k can be integrated according to Definition 3.15, we wonder when it is possible to conclude that f can be integrated, and in that case

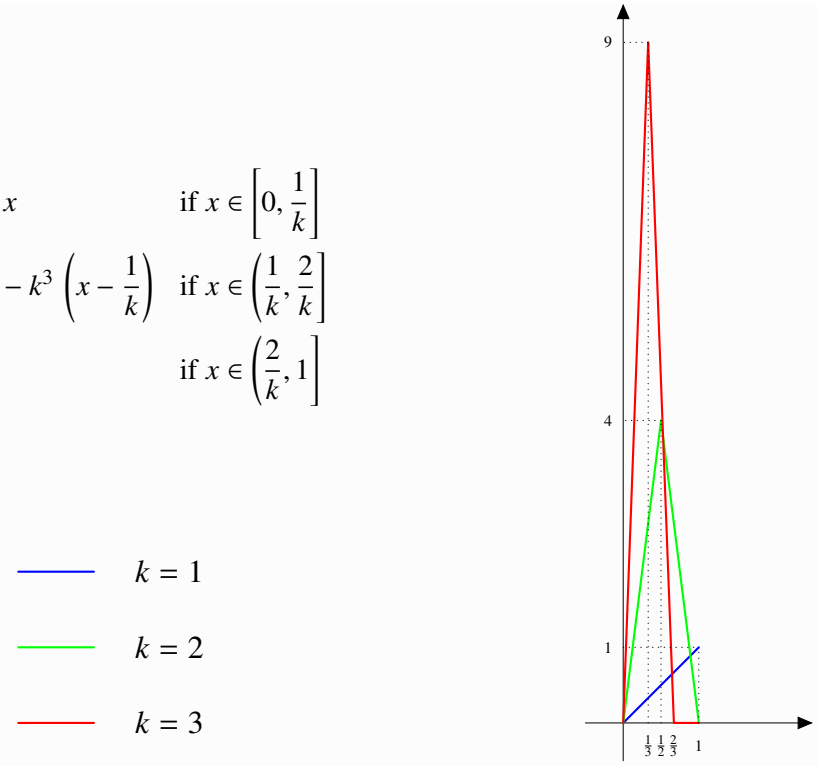
$$\lim_{k \to \infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

The problem is not trivial, as the following example shows.

Example *3.23* Let, for $k \in \mathbb{N}^*$, consider the sequence of functions

$$f_k : [0,1] \longrightarrow \mathbb{R}$$

$$x \mapsto \begin{cases} k^3\, x & \text{if } x \in \left[0, \frac{1}{k}\right] \\ k^2 - k^3 \left(x - \frac{1}{k}\right) & \text{if } x \in \left(\frac{1}{k}, \frac{2}{k}\right] \\ 0 & \text{if } x \in \left(\frac{2}{k}, 1\right] \end{cases}$$



In the figure we have the blue line for $k = 1$, green line for $k = 2$ and red for $k = 3$. It is easy to see that

$$\forall\, x \in [0,1], \quad \lim_{k \to \infty} f_k(x) = 0.$$

This is obvious if $x = 0$. If $x > 0$, if suffices to observe that, if k is large enough, we have $x > \frac{2}{k}$, so that $f_k(x) = 0$. However, $\forall\, k \in \mathbb{N}^*$ we find

$$\begin{aligned} \int_{[0,1]} f_k(x)\, \mathrm{d}x &= \int_0^{\frac{1}{k}} k^3\, x\, \mathrm{d}x + \int_{\frac{1}{k}}^{\frac{2}{k}} (k^2 - k^3\, x + k^2)\, \mathrm{d}x \\ &= \left[k^3\, \frac{x^2}{2}\right]_0^{\frac{1}{k}} + \left[2\, k^2\, x - \frac{1}{2}\, k^3\, x^2\right]_{\frac{1}{k}}^{\frac{2}{k}} = \frac{k}{2} + 4\, k - 2\, k - 2\, k + \frac{k}{2} = k. \end{aligned}$$

Hence,

$$\lim_{k \to \infty} \int_{[0,1]} f_k(x)\, \mathrm{d}x = +\infty.$$

We are going to see two classical theorems of interchange of limit with the integral: Beppo Levi's monotone convergence theorem and Lebesgue dominated convergence theorem.

The monotone convergence theorem says that if $(f_k)_{k \in \mathbb{N}}$ is a monotonically non-decreasing sequence, i.e.,

$$0 \leq f_1 \leq f_2 \leq \ldots \leq f_k \leq f_{k+1} \leq \ldots$$

of functions $f_k : \boldsymbol{A} \longrightarrow [0, +\infty]$ and we define the pointwise limit of $(f_k)_{k \in \mathbb{N}}$ as the function $f : \boldsymbol{A} \longrightarrow [0, +\infty]$, which therefore satisfies

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad f(\boldsymbol{x}) = \lim_{k \to \infty} f_k(\boldsymbol{x})$$

(notice that the limit exists for every $\boldsymbol{x} \in \boldsymbol{A}$ since the sequence $(f_k(\boldsymbol{x}))_{k \in \mathbb{N}}$ is non-decreasing), we can interchange

the limit with the integral sign:

$$\lim_{k\to\infty}\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} \lim_{k\to\infty} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Theorem 3.34 [Monotone Convergence Theorem – Beppo Levi's Theorem]
Let $\boldsymbol{A}\in\mathcal{M}_n$. For all $k\in\mathbb{N}$, let $f_k:\boldsymbol{A}\longrightarrow[0,+\infty]$ be measurable. Let $f_k(\boldsymbol{x})\le f_{k+1}(\boldsymbol{x})$ for all $k\in\mathbb{N}$ and for all $\boldsymbol{x}\in\boldsymbol{A}$. Let $f:\boldsymbol{A}\longrightarrow[0,+\infty]$ be such that

$$\forall\,\boldsymbol{x}\in\boldsymbol{A},\quad f(\boldsymbol{x}) = \lim_{k\to\infty} f_k(\boldsymbol{x}). \tag{3.34}$$

Then:

(a) *f is measurable;*

(b) $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty}\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$

Proof The existence of the limit

$$\lim_{k\to\infty}\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

is a consequence of the monotonicity of $(f_k)_{k\in\mathbb{N}}$. Indeed, for every $k\in\mathbb{N}$, the integral $\displaystyle\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ exists, since for every $k\in\mathbb{N}$ the function f_k is measurable. Moreover, from **(c)** in Theorem 3.33, we have

$$\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \int_{\boldsymbol{A}} f_{k+1}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

so that the sequence $\left(\displaystyle\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)_{k\in\mathbb{N}}$ is also monotonic non-decreasing. Thus, the limit exists.

From **(c)** in Theorem 3.28, we immediately see that **(a)** holds. We now prove **(b)**. For every $k\in\mathbb{N}$, we have $0\le f_k\le f$. From **(c)** in Theorem 3.33, for every $k\in\mathbb{N}$ we have

$$\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.35}$$

We define $\ell = \displaystyle\lim_{k\to\infty}\int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ (which exists as we proved above). Then, from (3.35) we find

$$\ell \le \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.36}$$

We want to prove the reverse inequality. To this end, it is sufficient to show that, given a simple function $\varphi:\boldsymbol{A}\longrightarrow\mathbb{R}$ such that $0\le\varphi\le f$, there holds

$$\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \le \ell. \tag{3.37}$$

Let $\varphi(\boldsymbol{x}) = \displaystyle\sum_{j=1}^{p}\alpha_j\chi_{\boldsymbol{B}_j}(\boldsymbol{x})$, where $\alpha_1,\dots,\alpha_p\in[0,+\infty)$, where $\boldsymbol{B}_1,\dots,\boldsymbol{B}_p$ are pairwise disjoint and their union is $\boldsymbol{A}$. Let $t\in(0,1)$. Consider, for all $k\in\mathbb{N}$, the set

$$\boldsymbol{A}_k = \{\boldsymbol{x}\in\boldsymbol{A}\mid f_k(\boldsymbol{x})\ge t\,\varphi(\boldsymbol{x})\}.$$

For all $k \in \mathbb{N}$, we have $\boldsymbol{A}_k \in \mathcal{M}_n$, because $\boldsymbol{A}_k = \{\boldsymbol{x} \in \boldsymbol{A} \mid f_k(\boldsymbol{x}) - t\,\varphi(\boldsymbol{x}) \geq 0\}$ (the product of a measurable function by a constant is measurable, and the sum of measurable functions is measurable, see **(b)** in Corollary 3.2, and also Remark 3.10 and Theorem 3.26). Let $\psi_k = t\,\varphi\chi_{\boldsymbol{A}_k}$ and $g_k = f_k\chi_{\boldsymbol{A}_k}$ for all $k \in \mathbb{N}$, so that $0 \leq \psi_k \leq g_k \leq f_k$ for all $k \in \mathbb{N}$. Since $\psi_k = \displaystyle\sum_{j=1}^{p} \alpha_j \chi_{\boldsymbol{B}_j}\chi_{\boldsymbol{A}_k} = \sum_{j=1}^{p} t\,\alpha_j \chi_{\boldsymbol{B}_j \cap \boldsymbol{A}_k}$ for all $k \in \mathbb{N}$, we obtain

$$\forall\, k \in \mathbb{N}, \qquad t \sum_{j=1}^{p} \alpha_j\, \mu_n(\boldsymbol{B}_j \cap \boldsymbol{A}_k) \leq \int_{\boldsymbol{A}} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \ell. \tag{3.38}$$

We now observe that for every $k \in \mathbb{N}$, we have $\boldsymbol{A}_k \subseteq \boldsymbol{A}_{k+1}$, since $f_k(\boldsymbol{x}) \leq f_{k+1}(\boldsymbol{x})$. Moreover

$$\bigcup_{k \in \mathbb{N}} \boldsymbol{A}_k = \boldsymbol{A}.$$

In fact, the fact that the countable union of the sets $\boldsymbol{A}_k$ is contained in A is obvious. We show that $\boldsymbol{A} \subseteq \displaystyle\bigcup_{k \in \mathbb{N}} \boldsymbol{A}_k$. Let $\boldsymbol{x} \in \boldsymbol{A}$, and we show that there exists $k \in \mathbb{N}$ such that $\boldsymbol{x} \in \boldsymbol{A}_k$. If $\varphi(\boldsymbol{x}) = 0$, then $\boldsymbol{x} \in \boldsymbol{A}_0$. If $\varphi(\boldsymbol{x}) > 0$, we have $t\,\varphi(\boldsymbol{x}) < \varphi(\boldsymbol{x}) \leq f(\boldsymbol{x})$. Hence, for a sufficiently large value of k, we have $t\,\varphi(\boldsymbol{x}) < f_k(\boldsymbol{x})$. For any such value of k, we have $\boldsymbol{x} \in \boldsymbol{A}_k$.

It follows that for every k and for every $j \in \{1, \ldots, p\}$, we have

$$\boldsymbol{B}_j \cap \boldsymbol{A}_k \subseteq \boldsymbol{B}_j \cap \boldsymbol{A}_{k+1} \quad \text{and} \quad \bigcup_{k=0}^{\infty} (\boldsymbol{B}_j \cap \boldsymbol{A}_k) = \boldsymbol{B}_j \cap \boldsymbol{A} = \boldsymbol{B}_j.$$

From **(a)** in Theorem 3.11,

$$\forall\, j \in \{1, \ldots, p\}, \qquad \lim_{k \to \infty} \mu_n(\boldsymbol{B}_j \cap \boldsymbol{A}_k) = \mu_n\Big(\bigcup_{k=0}^{\infty} (\boldsymbol{B}_j \cap \boldsymbol{A}_k)\Big) = \mu_n(\boldsymbol{B}_j).$$

We can now use Theorem 3.1-3.2 to find

$$\lim_{k \to \infty} t \sum_{j=1}^{p} \alpha_j\, \mu_n(\boldsymbol{B}_j \cap \boldsymbol{A}_k) = t \sum_{j=1}^{p} \alpha_j\, \mu_n(\boldsymbol{B}_j) = t \int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

and from (3.38)

$$t \int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \ell. \tag{3.39}$$

This is true for all $t \in (0, 1)$. Taking the limit for $t \to 1$, we obtain (3.37).

Remark 3.13 Beppo Levi's monotone convergence theorem does not hold in general for Riemann integrals. In general, the pointwise limit of a monotonically increasing, bounded sequence of Riemann integrable functions is not even Riemann integrable. For example, consider a sequence $(q_k)_{k \in \mathbb{N}}$ consisting of an enumeration of the rational numbers in $[0, 1]$ (which is countable), and let

$$\begin{aligned} f_k : [0,1] &\longrightarrow [0, +\infty) \\ x &\mapsto \begin{cases} 1 & \text{if } x = q_i \text{ for some } i \in \{1, 2, \ldots, k\} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

$(f_k)_{k \in \mathbb{N}}$ is a monotonically increasing sequence of bounded positive Riemann integrable functions. The inte-

gral of each of these functions over $[0,1]$ is zero. However, the sequence converges pointwise to $\chi_{\mathbb{Q}\cap[0,1]}$, which is not Riemann integrable.

Example *3.23*

Let $\alpha \in \mathbb{R}$. Consider the function

$$\begin{aligned} f : [0,1] &\longrightarrow [0,+\infty] \\ x &\mapsto \begin{cases} x^{-\alpha} & \text{if } 0 < x \leq 1 \\ +\infty & \text{if } x = 0. \end{cases} \end{aligned}$$

Let, for $k \in \mathbb{N}$,

$$\begin{aligned} f_k : [0,1] &\longrightarrow [0,+\infty] \\ x &\mapsto \begin{cases} x^{-\alpha} & \text{if } 1/k < x \leq 1 \\ k^{\alpha} & \text{if } 0 \leq x < 1/k. \end{cases} \end{aligned}$$

Since for every $k \in \mathbb{N}$ the function f_k is continuous, $(f_k)_{k\in\mathbb{N}}$ is an increasing sequence of Lebesgue measurable functions, and it is pointwise convergent to f. Then, from the monotone convergence theorem we find

$$\int_{[0,1]} f(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_{[0,1]} f_k(x)\,\mathrm{d}x.$$

We obtain

$$\begin{aligned} \int_{[0,1]} f_k(x)\,\mathrm{d}x &= \int_0^{1/k} k^{\alpha}\,\mathrm{d}x + \int_{1/k}^{1} x^{-\alpha}\,\mathrm{d}x = k^{\alpha}\,\frac{1}{k} + \left[\frac{x^{-\alpha+1}}{1-\alpha}\right]_{1/k}^{1} \\ &= \frac{1}{1-\alpha} - \frac{\alpha}{1-\alpha}\,k^{\alpha-1} \overset{k\to\infty}{\longrightarrow} \begin{cases} \frac{1}{1-\alpha} & \text{if } \alpha < 1 \\ +\infty & \text{if } \alpha \geq 1 \end{cases} \end{aligned}$$

It follows that f is integrable on $[0,1]$ if and only if $\alpha < 1$. Hence, the monotone convergence theorem can help us determine whether a function is integrable.

The following theorem shows that the Lebesgue integral is linear. The first claim was anticipated before, but not proved. The last claim is often referred to as the "corollary" of the monotone convergence theorem.

Theorem 3.35 [Linearity of the Lebesgue Integral]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [0,+\infty]$ be measurable. Let $f_k : \boldsymbol{A} \longrightarrow [0,+\infty]$ be measurable for all $k \in \mathbb{N}$. Let $c \in [0,+\infty)$. Then:

(a) *$c\,f$ is measurable, it takes values in $[0,+\infty]$ and*

$$\int_{\boldsymbol{A}} c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x};$$

(b) *for every $m \in \mathbb{N}$, the function $\sum_{k=0}^{m} f_k$ is measurable and*

$$\int_{\boldsymbol{A}} \sum_{k=0}^{m} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{m} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x};$$

(c) *the function* $s(\boldsymbol{x}) = \sum_{k=0}^{\infty} f_k(\boldsymbol{x})$ *(for* $\boldsymbol{x} \in A$*) is measurable and*

$$\int_A s(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty} \int_A f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x};$$

Proof We prove **(a)**. From Theorem 3.29, there exists a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of simple functions with domain $\boldsymbol{A}$ such that for all k there holds $0 \le \varphi_k \le \varphi_{k+1} \le f$ and $\lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$. Then, for all $k \in \mathbb{N}$, we have $0 \le c\,\varphi_k \le c\,\varphi_{k+1} \le c\,f$, and from Theorem 3.2 we obtain

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \lim_{k\to\infty} c\,\varphi_k(\boldsymbol{x}) = c\,f(\boldsymbol{x}).$$

From **(c)** in Theorem 3.28, we find that $c\,f$ is measurable. Using Theorem 3.33, **(e)** of Theorem 3.24 and Theorem 3.2,

$$\int_A c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \int_A c\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} c \int_A \varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = c \int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We prove **(b)**. If $m = 0$, the statement is true: both sides of equation are equal to 0. Let $m = 1$. We have to show that

$$\int_A A(f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = \int_A f_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_A f_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We introduce, for all $k \in \mathbb{N}$, the simple functions $\varphi_{0,k}$ over $\boldsymbol{A}$ such that $0 \le \varphi_{0,k} \le \varphi_{0,k+1} \le f_0$ for all $k \in \mathbb{N}$ and $\lim_{k\to\infty} \varphi_{0,k}(\boldsymbol{x}) = f_0(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$. Similarly, we introduce, for all $k \in \mathbb{N}$, the simple functions $\varphi_{1,k}$ such that $0 \le \varphi_{1,k} \le \varphi_{1,k+1} \le f_1$ for all $k \in \mathbb{N}$ and $\lim_{k\to\infty} \varphi_{1,k}(\boldsymbol{x}) = f_1(\boldsymbol{x})$ for all $\boldsymbol{x} \in \boldsymbol{A}$. It follows that for all k we have $0 \le \varphi_{0,k} + \varphi_{1,k} \le \varphi_{0,k+1} + \varphi_{1,k+1} \le f_0 + f_1$ and

$$\lim_{k\to\infty} (\varphi_{0,k}(\boldsymbol{x}) + \varphi_{1,k}(\boldsymbol{x})) = f_0(\boldsymbol{x}) + f_1(\boldsymbol{x})$$

for all $\boldsymbol{x} \in \boldsymbol{A}$. Applying the monotone convergence theorem (Theorem 3.34), **(d)** in Theorem 3.24 and Theorem 3.2, we find

$$\begin{aligned}
\int_A (f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} &= \lim_{k\to\infty} \int_A (\varphi_{0,k}(\boldsymbol{x}) + \varphi_{1,k}(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \left(\int_A \varphi_{0,k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_A \varphi_{1,k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right) \\
&= \lim_{k\to\infty} \int_A \varphi_{0,k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \lim_{k\to\infty} \int_A \varphi_{1,k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \int_A f_0(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_A f_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

The case $m \ge 2$ can be proved by induction. We leave the details to the reader.
We conclude by proving **(c)**. Let, for every $m \in \mathbb{N}$,

$$s_m(\boldsymbol{x}) = \sum_{k=0}^{m} f_k(\boldsymbol{x}).$$

Then, **(b)** ensures that s_m is measurable for every m, and

$$\int_A s_m(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{m} \int_A f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Since for every m we have $0 \leq s_m \leq s_{m+1} \leq s$ and $\lim_{m\to\infty} s_m(\boldsymbol{x}) = s(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$, from **(c)** in Theorem 3.28 we find that s is measurable. Moreover, from Theorem 3.33,

$$\int_{\boldsymbol{A}} s(\boldsymbol{x}) = \lim_{m\to\infty} \int_{\boldsymbol{A}} s_m(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{m\to\infty} \sum_{k=0}^{m} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Regarding the last statement of the previous theorem: notice that we did not need to introduce any assumptions on the convergence of the two series involved: it can indeed happen that both sides are equal to $+\infty$.

Theorem 3.36 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable. Let $\boldsymbol{A} = \boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \ldots \cup \boldsymbol{A}_m$, with $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_m$ measurable and pairwise disjoint. Then for each $j \in \{1, \ldots, m\}$ the restriction $f|_{\boldsymbol{A}_j}$ is measurable and*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{j=1}^{m} \int_{\boldsymbol{A}_j} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad \textit{with} \quad \int_{\boldsymbol{A}_j} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}_j} f|_{\boldsymbol{A}_j}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof The proof is left as an exercise. First, show that if $A, B \in \mathcal{M}_n$ with $A \subseteq B$ and $f : B \longrightarrow [-\infty, +\infty]$ is measurable, then $f|_A$ is measurable in A. The solution will be given in the next worksheet.

Theorem 3.37 *Let $\boldsymbol{A} \in \mathcal{M}_n$ with $\mu_n(\boldsymbol{A}) = 0$. Let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$. Then, f is measurable and*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0.$$

Proof The proof can be carried out by first considering the case where f is simple and non-negative, recalling that if $\mu_n^*(\boldsymbol{A}) = 0$, then $\boldsymbol{A}$ is Lebesgue measurable. Then, we can apply Definition 3.15.

3.8.2 Fatou's lemma

Before introducing the Lebesgue dominated convergence theorem, we introduce a preliminary result that can be seen as a generalisation of the Beppo Levi's monotone convergence theorem: it goes under the name of Fatou's Lemma. This result is also used in the proof of the Lebesgue dominated convergence theorem.

Theorem 3.38 [Fatou's Lemma]
Let $\boldsymbol{A} \in \mathcal{M}_n$. For all $k \in \mathbb{N}$, let $f_k : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable functions. Then,

(a) *The function*

$$\begin{aligned} A &\longrightarrow [-\infty, +\infty] \\ \boldsymbol{x} &\mapsto \liminf_{k\to\infty} f_k(\boldsymbol{x}) \end{aligned}$$

is measurable and it takes values in $[0, +\infty]$;

(b) *There holds* $\displaystyle\int_{\boldsymbol{A}} \liminf_{k\to\infty} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \liminf_{k\to\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$

Proof First, we observe that **(a)** is a consequence of Theorem 3.28.

We prove **(b)**. Let us define, for all $k \in \mathbb{N}$, the function $g_k(\boldsymbol{x}) = \inf_{h \geq k} f_h(\boldsymbol{x})$. Then, for every $\boldsymbol{x} \in A$ and for every $k \in \mathbb{N}$, we have $g_k(\boldsymbol{x}) \leq g_{k+1}(\boldsymbol{x})$ and $\lim_{k\to\infty} g_k(\boldsymbol{x}) = \liminf_{k\to\infty} f_k(\boldsymbol{x})$. From the monotone convergence theorem (Theorem 3.34) we find

$$\int_{\boldsymbol{A}} \liminf_{k\to\infty} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \int_{\boldsymbol{A}} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.40}$$

Since $g_k \leq f_k$, we then have

$$\begin{aligned}\lim_{k\to\infty} \int_{\boldsymbol{A}} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \liminf_{k\to\infty} \int_{\boldsymbol{A}} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &\leq \liminf_{k\to\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

which leads to the conclusion.

Remark 3.14 In **(b)** of Fatou's Lemma, we can have the strict inequality. We will see an example in Remark 3.23.

3.9 Signed functions

We now want to define the Lebesgue integral for a class of measurable functions that do not necessarily take values in $[0, +\infty)$. We also want to be able to deal with complex-valued functions. Consider a function $f : \boldsymbol{A} \longrightarrow \mathbb{R}$. Recall that we have defined the positive and the negative parts f^+ and f^- of f in Definition 3.14. From **(d)** in Lemma 3.4, if $\boldsymbol{A} \in \mathcal{M}_n$ and $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ is measurable, f^+ and f^- are measurable and they take non-negative values. Hence, the functions f^+ and f^- satisfy the requirements of Definition 3.15, and therefore the integrals $\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and $\int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ are well defined.

A first definition of Lebesgue integral can be given by requiring at least one of the integrals $\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and $\int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ to be finite, or, equivalently, by imposing

$$\min\left\{\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right\} < +\infty. \tag{3.41}$$

For the moment, a function satisfying this condition will be referred to as "Lebesgue integrable". If f is Lebesgue integrable, we can define the Lebesgue integral as

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \stackrel{\text{def}}{=} \int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

and the Lebesgue integrability condition ensures that the form $+\infty - \infty$ is avoided. The drawback of this definition is that the sum of two Lebesgue integrable functions is not, in general, Lebesgue integrable.

Example *3.24*

Consider the functions

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \end{aligned} \qquad \text{and} \qquad \begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} -1 & \text{if } x \leq 0 \\ 0 & \text{if } x > 0. \end{cases} \end{aligned}$$

We have that f^- and g^+ are identically zero functions, and therefore both functions satisfy the condition (3.41). However, their sum is the function

$$f+g:\mathbb{R}\longrightarrow\mathbb{R}$$
$$x\mapsto\begin{cases}1 & \text{if } x>0\\ -1 & \text{if } x<0\\ 0 & \text{if } x=0,\end{cases}$$

and therefore

$$(f+g)^+:\mathbb{R}\longrightarrow\mathbb{R}\qquad x\mapsto\begin{cases}1 & \text{if } x>0\\ 0 & \text{if } x\le 0\end{cases}\qquad\text{and}\qquad (f+g)^-:\mathbb{R}\longrightarrow\mathbb{R}\qquad x\mapsto\begin{cases}1 & \text{if } x<0\\ 0 & \text{if } x\ge 0.\end{cases}$$

Clearly $\displaystyle\int_{\mathbb{R}}(f+g)^+(x)\,\mathrm{d}x=\int_{\mathbb{R}}(f+g)^-(x)\,\mathrm{d}x=+\infty.$

A stronger condition that the one in (3.41) is the so-called *summability*: this condition requires both integrals of the positive and negative part of the function to be finite. We will see that the class of summable functions is a vector space, and this avoids the issue that we described above: the sum of summable functions is summable, and we will have the additivity of the Lebesgue integral. We also warn the reader that the term "integrable" is sometimes used as a synonym of summable. However, sometimes this term is used to describe functions satisfying (3.41). For this reason, we will refrain from using the term "integrable", to avoid confusion.

Definition 3.16 [Real-Valued Summable Functions]
Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ be measurable. We say that f is summable *or* absolutely integrable *in $\boldsymbol{A}$ if*

$$\int_{\boldsymbol{A}}f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}<+\infty\quad\textit{and}\quad\int_{\boldsymbol{A}}f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}<+\infty. \tag{3.42}$$

In this case, we define

$$\int_{\boldsymbol{A}}f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\overset{def}{=}\int_{\boldsymbol{A}}f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}-\int_{\boldsymbol{A}}f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We denote by $\mathcal{L}(\boldsymbol{A};\mathbb{R})$ the class of (real-valued) summable functions.

We notice that condition (3.42) is equivalent to the condition

$$\max\left\{\int_{\boldsymbol{A}}f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},\int_{\boldsymbol{A}}f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right\}<+\infty.$$

Remark 3.15 The integrals $\displaystyle\int_{\boldsymbol{A}}f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ and $\displaystyle\int_{\boldsymbol{A}}f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ are as defined in Definition 3.15. If f takes non-negative values, then $f^+=f$ and $f^-=0$, and in this case Definition 3.16 reduces to Definition 3.15.

Theorem 3.39 *Let $\boldsymbol{A}\in\mathcal{M}_n$. Let $f:\boldsymbol{A}\longrightarrow\mathbb{R}$ be measurable. Then, $f\in\mathcal{L}(A;\mathbb{R})$ if and only if*

$$\int_{\boldsymbol{A}}|f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}<+\infty. \tag{3.43}$$

Proof From **(b)** in Lemma 3.4, we have $|f|=f^++f^-$. It follows that $|f|$ is measurable and it takes values in

$[0, +\infty)$. Moreover, from **(b)** in Theorem 3.35, if $f \in \mathcal{L}(\boldsymbol{A}; \mathbb{R})$, we have

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f^+(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} f^-(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} < +\infty.$$

Conversely, if (3.43) holds true, from **(c)** in Theorem 3.33

$$\max\left\{\int_{\boldsymbol{A}} f^+(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}, \int_{\boldsymbol{A}} f^-(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}\right\} \leq \int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < +\infty.$$

Example *3.26*

Consider the Dirichlet integral

$$\int_0^{+\infty} \frac{\sin x}{x} \, \mathrm{d}x.$$

We show that the integrand is not summable. In fact, suppose by contradiction that

$$\int_0^{+\infty} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x < +\infty.$$

Then, we would have

$$\begin{aligned}
\int_0^{+\infty} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x &= \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \int_{2\pi}^{+\infty} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x \\
&= \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \int_{2k\pi}^{2(k+1)\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x \\
&\geq \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \int_{2k\pi}^{2(k+1)\pi} \left|\frac{\sin x}{2(k+1)\pi}\right| \, \mathrm{d}x \\
&= \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \frac{1}{2(k+1)\pi} \int_{2k\pi}^{2(k+1)\pi} |\sin x| \, \mathrm{d}x \\
&= \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \frac{1}{2(k+1)\pi} \int_0^{2\pi} |\sin x| \, \mathrm{d}x \\
&= \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \frac{1}{2(k+1)\pi} \cdot 4 = \int_0^{2\pi} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x + \sum_{k=1}^{\infty} \frac{2}{(k+1)\pi}.
\end{aligned}$$

The series thus obtained diverges, which contradicts the claim $\int_0^{+\infty} \left|\frac{\sin x}{x}\right| \, \mathrm{d}x < +\infty$. This integral still exists as an improper Riemann integral, because, using contour integration, we saw that both limits

$$\lim_{\delta \to 0^+} \int_{\delta}^{1} \frac{\sin x}{x} \, \mathrm{d}x \quad \text{and} \quad \lim_{M \to +\infty} \int_1^M \frac{\sin x}{x} \, \mathrm{d}x$$

exist in $\mathbb{R}$ (and, indeed, the value of the improper integral was found to be $\pi/2$).

Remark 3.16 If $\boldsymbol{A}$ is measurable, $\mu_n(\boldsymbol{A}) < +\infty$, and if $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ is measurable and bounded, then it is easy to see that f is summable. In fact, if f is bounded, there exists $M > 0$ such that $|f(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in \boldsymbol{A}$, and therefore

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} M \, \mathrm{d}\boldsymbol{x} = M \int_{\boldsymbol{A}} \mathrm{d}\boldsymbol{x} = M \, \mu_n(\boldsymbol{A}) < +\infty.$$

Lemma 3.5 *Let $A \in \mathcal{M}_n$. Let $f : A \longrightarrow \mathbb{R}$. Let $f = h - k$, where $h, k : A \longrightarrow [0, +\infty)$ be measurable and such that* $\max\left\{\int_A h(x)\,dx, \int_A k(x)\,dx\right\} < +\infty$. *Then, $f \in \mathcal{L}(A;\mathbb{R})$ and*

$$\int_A f(x)\,dx = \int_A h(x)\,dx - \int_A k(x)\,dx.$$

Proof First, we notice that f is measurable in view of **(b)** of Corollary 3.2. Moreover, for all $x \in A$ we have also $f^+(x) \le h(x)$ and $f^-(x) \le k(x)$. In fact, if $f(x) \ge 0$,

$$f^+(x) = f(x) = h(x) - k(x) \le h(x).$$

If $f(x) < 0$, then $f^+(x) = 0 \le h(x)$. Moreover, for all $x \in A$ we have

$$f^-(x) = f^+(x) - f(x) \le h(x) - f(x) = k(x).$$

From **(c)** in Theorem 3.33 we obtain

$$\int_A f^+(x)\,dx \le \int_A h(x)\,dx < +\infty \quad \text{and} \quad \int_A f^-(x)\,dx \le \int_A k(x)\,dx < +\infty.$$

It follows that $f \in \mathcal{L}(A;\mathbb{R})$. Moreover, from $f^+ - f^- = h - k$, we obtain $f^+ + k = f^- + h$, and from **(b)** in Theorem 3.35

$$\int_A f^+(x)\,dx + \int_A k(x)\,dx = \int_A f^-(x)\,dx + \int_A h(x)\,dx,$$

so that

$$\int_A f(x)\,dx = \int_A f^+(x)\,dx - \int_A f^-(x)\,dx = \int_A h(x)\,dx - \int_A k(x)\,dx.$$

Theorem 3.40 *Let $A \in \mathcal{M}_n$. Let $c \in \mathbb{R}$. Let $f, g \in \mathcal{L}(A;\mathbb{R})$. Let $c \in \mathbb{R}$. Then:*

(a) *$f + g \in \mathcal{L}(A;\mathbb{R})$ and* $\int_A (f+g)(x)\,dx = \int_A f(x)\,dx + \int_A g(x)\,dx$;

(b) *$cf \in \mathcal{L}(A;\mathbb{R})$ and* $\int_A (cf)(x)\,dx = c\int_A f(x)\,dx$;

(c) *if $f \le g$, then* $\int_A f(x)\,dx \le \int_A g(x)\,dx$.

Proof We prove **(a)**. We have $f + g = (f^+ + g^+) - (f^- + g^-)$. Let $h = f^+ + g^+$ and $k = f^- + g^-$. From **(b)** in Theorem 3.35 we have

$$\begin{aligned}\int_A h(x)\,dx &= \int_A f^+(x)\,dx + \int_A g^+(x)\,dx < +\infty\\ \int_A k(x)\,dx &= \int_A f^-(x)\,dx + \int_A g^-(x)\,dx < +\infty.\end{aligned}$$

Applying Lemma 3.5, we conclude that $f + g \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$ and

$$\begin{aligned}\int_{\boldsymbol{A}} (f(\boldsymbol{x}) + g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \left(\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right) + \left(\int_{\boldsymbol{A}} g^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} g^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\\ &= \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

We prove **(b)**. Suppose $c \geq 0$. Hence, $c\,f = (c\,f^+) - (c\,f^-)$. We have

$$\int_{\boldsymbol{A}} (c\,f^+)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = c\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty \quad \text{and} \quad \int_{\boldsymbol{A}} (c\,f^-)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = c\int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty.$$

Hence, $c\,f \in \mathcal{L}(A;\mathbb{R})$. From Lemma 3.5 and from **(a)** in Theorem 3.35 we have

$$\begin{aligned}\int_{\boldsymbol{A}} c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} c\,f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} c\,f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= c\left(\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right) = c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

If $c < 0$, we repeat the previous argument with $c\,f = (-c)\,f^- - (-c)\,f^+$:

$$\int_{\boldsymbol{A}} (-c\,f^-)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (-c)\int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty \quad \text{and} \quad \int_{\boldsymbol{A}} (-c\,f^+)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (-c)\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty,$$

and

$$\begin{aligned}\int_{\boldsymbol{A}} (c\,f)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} (-c\,f^-)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} (-c\,f^+)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (-c)\left(\int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\\ &= c\left(\int_{\boldsymbol{A}} f^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right) = c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

We now prove **(c)**. We have $g = f + (g - f)$. From **(a)** and **(b)** we find that $g - f \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$. From **(b)** we find

$$\int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} (g - f)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

since $\displaystyle\int_{\boldsymbol{A}} (g - f)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq 0$.

We can now consider complex-valued functions.

Definition 3.17 [Complex-Valued Measurable and Summable Functions]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. We say that f is measurable if $\mathfrak{Re}\,f$ and $\mathfrak{Im}\,f$ are measurable (according to Definition 3.13).
We say that f is summable *or* absolutely integrable *if $\mathfrak{Re}\,f$ and $\mathfrak{Im}\,f$ belong to $\mathcal{L}(\boldsymbol{A};\mathbb{R})$. In this case, we define*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \stackrel{def}{=} \int_{\boldsymbol{A}} \mathfrak{Re}\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i\int_{\boldsymbol{A}} \mathfrak{Im}\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We denote by $\mathcal{L}(\boldsymbol{A};\mathbb{C})$ the class of (complex-valued) summable functions.

Remark 3.17 If f is real-valued, the measurability of f coincides with the one of Definition 3.13, since in this case $\mathfrak{Re} f = f$ and $\mathfrak{Im} f = 0$.

We extend some of the results introduced before to the case of complex-valued functions.

Theorem 3.41 *Let $\boldsymbol{A} \in \mathcal{M}_n$ and let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ be measurable. Then, $|f|$ is measurable and $f \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$ if and only if*

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < +\infty.$$

Proof The measurability of $|f|$ can be written, for $\boldsymbol{x} \in \boldsymbol{A}$, as

$$|f(\boldsymbol{x})| = h((\mathfrak{Re} f(\boldsymbol{x}))^2 + (\mathfrak{Im} f(\boldsymbol{x}))^2),$$

where

$$\begin{aligned} h : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\mapsto |y|^{\frac{1}{2}}, \end{aligned}$$

and applying **(a)** and **(b)** of Corollary 3.2. Moreover,

$$|\mathfrak{Re} f| \leq |f| \leq |\mathfrak{Re} f| + |\mathfrak{Im} f| \quad \text{and} \quad |\mathfrak{Im} f| \leq |f| \leq |\mathfrak{Re} f| + |\mathfrak{Im} f|. \tag{3.44}$$

Hence, if $f \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$, from Theorem 3.39

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} |\mathfrak{Re} f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} |\mathfrak{Im} f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < +\infty.$$

Conversely, if $\displaystyle\int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < +\infty$, using again (3.44), we find

$$\max \left\{ \int_{\boldsymbol{A}} |\mathfrak{Re} f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}, \int_{\boldsymbol{A}} |\mathfrak{Im} f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \right\} \leq \int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x},$$

so that $f \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$.

Theorem 3.42 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $c \in \mathbb{R}$. Let $f, g \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$. Let $c \in \mathbb{C}$. Then:*

(a) *$f + g \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$ and $\displaystyle\int_{\boldsymbol{A}} (f+g)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$;*

(b) *$c\, f \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$ and $\displaystyle\int_{\boldsymbol{A}} (c\, f)(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = c \int_{\boldsymbol{A}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}$;*

(c) $\displaystyle\left| \int_{\boldsymbol{A}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \leq \int_{\boldsymbol{A}} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}.$

Proof We prove **(a)**. We have $\mathfrak{Re}\{f+g\} = \mathfrak{Re} f + \mathfrak{Re} g$ and $\mathfrak{Im}\{f+g\} = \mathfrak{Im} f + \mathfrak{Im} g$. From **(a)** in Theorem 3.40

we find $f+g \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$ and

$$\begin{aligned}\int_{\boldsymbol{A}} (f(\boldsymbol{x})+g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} (\Re f(\boldsymbol{x})+\Re g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} + i\int_{\boldsymbol{A}} (\Im f(\boldsymbol{x})+\Im g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{A}} \Re f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} \Re g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i\left(\int_{\boldsymbol{A}} \Im f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} \Im g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\\ &= \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Let $c = a + i\,b$, where $a, b \in \mathbb{R}$. Thus, $\Re\{c\,f\} = a\,\Re f - b\,\Im f$ and $\Im\{c\,f\} = a\,\Im f + b\,\Re f$, so that $c\,f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$, and

$$\begin{aligned}\int_{\boldsymbol{A}} c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} (a\,\Re f(\boldsymbol{x}) - b\,\Im f(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} + i\int_{\boldsymbol{A}} (a\,\Im f(\boldsymbol{x}) + b\,\Re f(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\\ &= a\int_{\boldsymbol{A}} \Re f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - b\,\Im f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i\,a\,\Im f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i\,b\,\Re f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= (a+i\,b)\left(\int_{\boldsymbol{A}} \Re f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i\int_{\boldsymbol{A}} \Im f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right) = c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

We now prove **(c)**. First, notice that, from Theorem 3.41, $|f| \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$. Let $c \in \mathbb{C}$ be such that $|c| = 1$ and

$$c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|.$$

Such value of c always exists: it suffices to choose

$$c = \frac{\left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|}{\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}$$

if $\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \neq 0$ and $c = 1$ otherwise. From **(b)** we find

$$\begin{aligned}\left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| &= c\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} c\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{A}} \Re\{c\,f(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} |c\,f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

using **(c)** in Theorem 3.40.

Theorem 3.43 *Let $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{M}_n$ with $\boldsymbol{B} \subseteq \boldsymbol{A}$. Let $f : \boldsymbol{A} \longrightarrow [-\infty, +\infty]$ or $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. Then:*

(a) *if f is measurable on $\boldsymbol{A}$, the restriction $f|_{\boldsymbol{B}}$ is measurable on B;*

(b) *if f is measurable on $\boldsymbol{A}$ and takes values in $[0, +\infty]$, then*

$$\int_{\boldsymbol{B}} f|_{\boldsymbol{B}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\chi_{\boldsymbol{B}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.45}$$

(c) *if $f \in \mathcal{L}(\boldsymbol{A},\mathbb{R})$ (resp. $f \in \mathcal{L}(\boldsymbol{A},\mathbb{C})$), then $f|_{\boldsymbol{B}} \in \mathcal{L}(\boldsymbol{B};\mathbb{R})$ (resp. $f|_{\boldsymbol{B}} \in \mathcal{L}(\boldsymbol{B};\mathbb{C})$), and (3.45) holds.*

Proof We first recall that the case where f is simple has already been dealt with in Theorem 3.24. Let f be measurable and assume that it takes values in the set $[0, +\infty]$. Consider a sequence of simple non-negative functions $(\varphi_k)_{k\in\mathbb{N}}$ as in Theorem 3.29. From Theorem 3.25, the restrictions $\varphi_k|_B$ of φ_k to $\boldsymbol{B}$ are simple functions. From **(c)** in Theorem 3.28, we conclude that $f|_B$ is measurable over $\boldsymbol{B}$. Moreover, from the monotone convergence theorem (Theorem 3.34), Theorem 3.25 and in Theorem 3.2, the function $f\chi_B$ is measurable as it is the pointwise limit of $\varphi_k\chi_B$, and

$$\begin{aligned}\int_{\boldsymbol{B}} f|_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \lim_{k\to\infty}\int_B \varphi_k|_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \lim_{k\to\infty}\int_{\boldsymbol{A}} \varphi_k(\boldsymbol{x})\chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

Let now f be measurable and taking values in the set $[-\infty, +\infty]$. The measurability of $f|_B$ can be deduced by observing that $f|_B = (f^+)|_B - (f^-)|_B$, and applying the previous case to f^+ and f^- (see also Remark 3.10). Moreover, if $f \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$ we have

$$\begin{aligned}\int_{\boldsymbol{B}} f|_{\boldsymbol{B}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_B (f^+)|_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_B (f^-)|_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{A}} f^+(\boldsymbol{x})\chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} f^-(\boldsymbol{x})\chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

so that $f\chi_B = f^+\chi_B - f^-\chi_B \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$, and (3.45) holds. Finally, the case where $f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$ can be treated adapting these arguments to $\mathfrak{Re}f$ and $\mathfrak{Im}f$. The details are left to the reader.

Theorem 3.44 *Let $m \in \mathbb{N}^*$. Let $\boldsymbol{A}_1, \dots, \boldsymbol{A}_m \in \mathcal{M}_n$ be pairwise disjoint. Let $\boldsymbol{A} = \bigcup_{k=1}^{m} A_k$ and let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$.*

Then, the following statements are equivalent:

(i) *$f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$;*

(ii) *for every $k \in \{1, \dots, m\}$, defining $f_k = f|_{A_k}$, we have $f_k \in \mathcal{L}(\boldsymbol{A}_k;\mathbb{C})$.*

If **(i)** *holds (or, equivalently, if* **(ii)** *holds), we have*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=1}^{m}\int_{\boldsymbol{A}_k} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof **(i)** implies **(ii)** from Theorem 3.43. Conversely, suppose that **(ii)** holds. To prove that **(i)** holds, we can reason by induction on m. Let $m = 1$, and assume that f is real-valued. Then, f is measurable by Theorem 3.31. Moreover, since

$$|f| = \sum_{k=1}^{m} |f|\chi_{A_k},$$

applying **(b)** in Theorem 3.35 and **(b)** in Theorem 3.43, we find

$$\begin{aligned}\int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &= \sum_{k=1}^{m}\int_{\boldsymbol{A}} |f(\boldsymbol{x})|\chi_{A_k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{k=1}^{m}\int_{\boldsymbol{A}_k} |f_k(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty.\end{aligned}$$

It follows that $f \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$. From **(a)** in Theorem 3.40 and **(c)** in Theorem 3.43,

$$\begin{aligned}\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \sum_{k=1}^{m}\int_{\boldsymbol{A}} f(\boldsymbol{x})\chi_{\boldsymbol{A}_k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \sum_{k=1}^{m}\int_{\boldsymbol{A}_k} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

The case where f is complex-valued can be tackled by writing f as $\mathfrak{Re}f + i\,\mathfrak{Im}f$. We leave the rest of the induction argument (which from this point proceeds straightforwardly) to the reader.

Remark 3.18 Theorem 3.44 can be adapted to the case where f is measurable and takes values in $[0,+\infty]$. We leave the details as an exercise.

We will see in the sequel that the result of Theorem 3.41 can be generalized to the case where $\boldsymbol{A}$ is the union of countable pairwise disjoint measurable sets.

In the sequel, we will use the following notation: if

- $\boldsymbol{A},\boldsymbol{B} \in \mathcal{M}_n$ with $\boldsymbol{B} \subseteq \boldsymbol{A}$
- f is measurable and non-negative on $\boldsymbol{A}$, or if $f \in \mathcal{L}(\boldsymbol{A};\mathbb{R})$ or $f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$,

we define

$$\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \stackrel{\text{def}}{=} \int_{\boldsymbol{B}} f|_{\boldsymbol{B}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We now consider a list of useful properties of measurable functions that takes values in $[0,+\infty]$ with finite integral. The first says that, if f takes values in $[0,+\infty]$, we can make the integral $\displaystyle\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ arbitrarily small by choosing $\boldsymbol{B}$ to have a sufficiently small measure.

Theorem 3.45 [Absolute Continuity of the Lebesgue Integral]
Let $\boldsymbol{A} \in \mathcal{M}_n$ and $f : \boldsymbol{A} \longrightarrow [0,+\infty]$ be measurable and such that $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$. Then, for every $\varepsilon > 0$, there exists $\delta > 0$ such that, for all $\boldsymbol{B} \in \mathcal{M}_n$ with $\boldsymbol{B} \subseteq \boldsymbol{A}$ and $\mu_n(\boldsymbol{B}) < \delta$, there holds

$$\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \varepsilon.$$

Proof First, assume that f is simple, and we write it as $f = \displaystyle\sum_{k=1}^{m}\alpha_k\chi_{\boldsymbol{A}_k}$, where $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_m$ are pairwise disjoint and their union is A; we assume also that $\alpha_k \geq 0$ for every $k \in \{1,\ldots,m\}$. If $f = 0$, the result is obvious. Otherwise, from Theorem 3.25 we se that, for a measurable subset $\boldsymbol{B}$ of $\boldsymbol{A}$, we have

$$\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=1}^{m}\alpha_k\,\mu_n(\boldsymbol{A}_k \cap \boldsymbol{B}) \leq \Big(\sum_{k=1}^{m}\alpha_k\Big)\mu_n(\boldsymbol{B}) < \varepsilon$$

if

$$\mu_n(\boldsymbol{B}) < \delta = \frac{\varepsilon}{\displaystyle\sum_{k=1}^{m}\alpha_k}.$$

Let us now consider the general case. Let $\varepsilon > 0$. From Definition 3.15, there exists a simple function $\varphi : \boldsymbol{A} \longrightarrow [0, +\infty)$ satisfying $\varphi \leq f$ and

$$\int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{\varepsilon}{2}.$$

From Corollary 3.2 and Remark 3.11, we find that $f - \varphi$ is measurable and non-negative. From $f = \varphi + (f - \varphi)$ and from **(b)** in Theorem 3.36, we get

$$\int_{\boldsymbol{A}} (f(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{A}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \frac{\varepsilon}{2}.$$

Let $\boldsymbol{B} \subseteq \boldsymbol{A}$ be measurable and such that $\displaystyle\int_{\boldsymbol{B}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \frac{\varepsilon}{2}$. It follows that

$$\begin{aligned}\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{B}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{B}} (f(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\boldsymbol{B}} \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} (f(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\chi_{\boldsymbol{B}}\,\mathrm{d}\boldsymbol{x} \\ &< \frac{\varepsilon}{2} + \int_{\boldsymbol{A}} (f(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Theorem 3.46 *Let $\boldsymbol{A} \in \mathcal{M}_n$ and $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable and such that $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$. Let $\boldsymbol{A}_\infty \overset{def}{=} \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) = +\infty\}$. Then, $\boldsymbol{A}_\infty$ is measurable and $\mu_n(\boldsymbol{A}_\infty) = 0$.*

Proof The set $\boldsymbol{A}_\infty$ is measurable because it coincides with $\displaystyle\bigcap_{k=0}^{\infty} \{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) > k\}$. We prove that $\mu_n(\boldsymbol{A}_\infty) = 0$. Let $\varphi_k = k\chi_{\boldsymbol{A}_\infty}$ for every $k \in \mathbb{N}$. The function φ_k is simple and $0 \leq \varphi_k \leq f$ for every $k \in \mathbb{N}$. Hence,

$$\forall\, k \in \mathbb{N}, \qquad k\,\mu_n(\boldsymbol{A}_\infty) = \int_{\boldsymbol{A}} \varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.46}$$

If by contradiction $\mu_n(\boldsymbol{A}_\infty) > 0$, we would have $\lim\limits_{k \to \infty} k\,\mu_n(\boldsymbol{A}_\infty) = +\infty$, thus contradicting (3.46).

Lemma 3.6 *Let $\boldsymbol{A} \in \mathcal{M}_n$ and $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable and such that $\displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$. Then, for every $\varepsilon > 0$ there exists a measurable set $\boldsymbol{B} \subseteq \boldsymbol{A}$ with $\mu_n(\boldsymbol{B}) < +\infty$ and*

$$\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \varepsilon.$$

Proof For $k \in \mathbb{N}^*$, let $\boldsymbol{A}_k \overset{\text{def}}{=} \left\{\boldsymbol{x} \in \boldsymbol{A} \mid f(\boldsymbol{x}) \geq \frac{1}{k}\right\}$. We have $\mu_n(\boldsymbol{A}_k) < +\infty$ for every $k \in \mathbb{N}^*$. Indeed, if this was not the case, given

$$\begin{aligned}\varphi : \boldsymbol{A} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \frac{1}{k}\chi_{\boldsymbol{A}_k}(\boldsymbol{x}),\end{aligned}$$

we would have

$$\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \int_A \varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \frac{1}{k}\,\mu_n(\boldsymbol{A}_k) = +\infty.$$

Let $f_k = f\,\chi_{A_k}$. For every $k \in \mathbb{N}^*$, the function f_k is measurable, it satisfies $0 \leq f_k \leq f_{k+1}$ and for every $\boldsymbol{x} \in \boldsymbol{A}$ it satisfies $\lim_{k\to\infty} f_k(\boldsymbol{x}) = f(\boldsymbol{x})$. This is obvious if $f(\boldsymbol{x}) = 0$. If $f(\boldsymbol{x}) > 0$, for a sufficiently large value of k, we have $f(\boldsymbol{x}) \geq \frac{1}{k}$, from which $f_k(\boldsymbol{x}) = f(\boldsymbol{x})$. Applying the monotone convergence theorem (Theorem 3.34) and Theorem 3.43, we obtain

$$\lim_{k\to\infty} \int_{\boldsymbol{A}_k} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \int_{\boldsymbol{A}_k} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Letting $\boldsymbol{B} = \boldsymbol{A}_k$, for a sufficiently large value of k, we obtain the result.

We go back to the problem of finding conditions which guarantee that

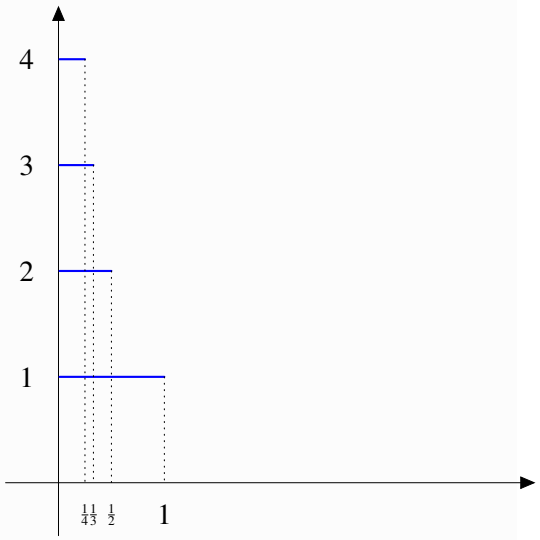
$$\int_A f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \overset{k\to\infty}{\longrightarrow} \int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

for a sequence $(f_k)_{k\in\mathbb{N}}$ which is pointwise convergent to f. This problem amounts to interchanging the limit with the integral, as we have seen. We have already observed that the Riemann integral does not allow us to give a satisfactory answer to this question. A condition that ensures that the interchange of limit and integral is legitimate is the uniform convergence of $(f_k)_{k\in\mathbb{N}}$ to f. However, the uniform convergence assumption is too strong in many contexts. The Lebesgue integral solves this issue: in this case, the pointwise convergence is sufficient, at the expense of some (mild) conditions on the functions f_k of the sequence.

Example *3.26*

Let, for $k \in \mathbb{N}^*$,

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} k & \text{if } 0 < x < \frac{1}{k} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$



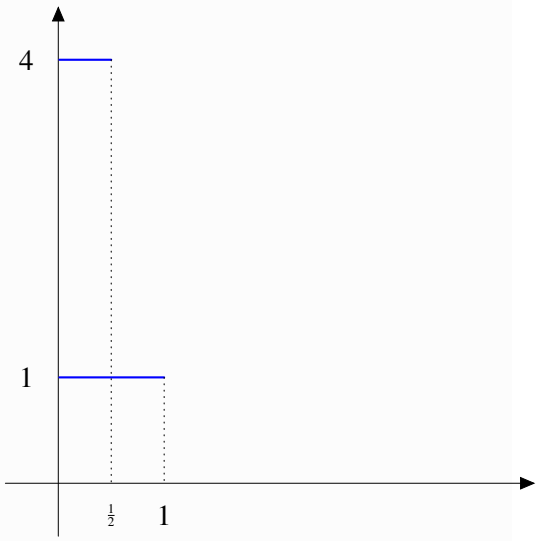
The pointwise limit of $(f_k)_{k\in\mathbb{N}}$ is the zero function, but

$$\forall\, k \in \mathbb{N}^*, \quad \int_{\mathbb{R}} f_k(x)\,\mathrm{d}x = \int_0^{1/k} k\,\mathrm{d}x = \Big[k\,x\Big]_0^{1/k} = 1.$$

Example *3.27*

For $k \in \mathbb{N}^*$, consider the function

$$\begin{array}{rccl} f_k : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \begin{cases} k^2 & \text{if } 0 < x < \frac{1}{k} \\ 0 & \text{otherwise} \end{cases} \end{array}$$



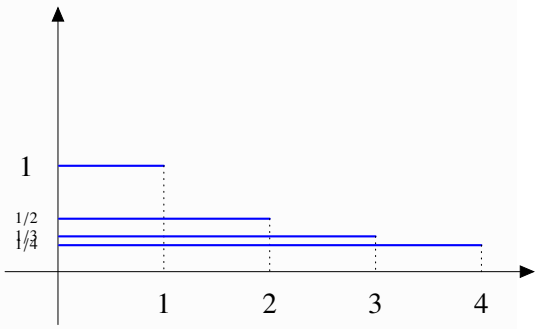
In this case, the pointwise limit of $(f_k)_{k\in\mathbb{N}}$ is still the zero function, but

$$\int_{\mathbb{R}} f_k(x)\,\mathrm{d}x = \int_0^{1/k} k^2\,\mathrm{d}x = \left[k^2\,x\right]_0^{1/k} = k \overset{k\to\infty}{\longrightarrow} +\infty.$$

Example *3.28*

Let, for $k \in \mathbb{N}^*$,

$$\begin{array}{rccl} f_k : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \begin{cases} \frac{1}{k} & \text{if } 0 < x < k \\ 0 & \text{otherwise} \end{cases} \end{array}$$



The sequence $(f_k)_{k\in\mathbb{N}}$ converges uniformly to the zero function, but again

$$\forall\, k \in \mathbb{N}^*, \quad \int_{\mathbb{R}} f_k(x)\,\mathrm{d}x = \int_0^k \frac{1}{k}\,\mathrm{d}x = \left[\frac{1}{k}\,x\right]_0^k = 1.$$

Example *3.29*

Let, for $k \in \mathbb{N}$,

$$\begin{array}{rccl} f_k : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \begin{cases} 1 & \text{if } k < x < k+1 \\ 0 & \text{otherwise} \end{cases} \end{array}$$

Again, the sequence $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the zero function, but

$$\forall\, k \in \mathbb{N}, \quad \int_{\mathbb{R}} f_k(x)\,\mathrm{d}x = \int_k^{k+1} 1\,\mathrm{d}x = \Big[x\Big]_k^{k+1} = k+1-k = 1.$$

Notice that Beppo Levi's monotone convergence theorem, which does not hold in general for Riemann integrals as we have seen, rules out examples like these.

Theorem 3.47 [Dominated Convergence Theorem – Lebesgue Theorem]
Let $\boldsymbol{A} \in \mathcal{M}_n$. For all $k \in \mathbb{N}$, let $f_k : \boldsymbol{A} \longrightarrow \mathbb{C}$ be measurable. Let $g : \boldsymbol{A} \longrightarrow [0,+\infty]$ be measurable. We assume that:

(i) *for every $\boldsymbol{x} \in \boldsymbol{A}$, the limit $f(\boldsymbol{x}) = \lim\limits_{k\to\infty} f_k(\boldsymbol{x})$ exists in $\mathbb{C}$;*

(ii) *for every $k \in \mathbb{N}$ and for every $\boldsymbol{x} \in \boldsymbol{A}$, we have $|f_k(\boldsymbol{x})| \leq g(\boldsymbol{x})$;*

(iii) $\displaystyle\int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty.$

Then:

(a) *for every $k \in \mathbb{N}$, we have $f_k \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$;*

(b) *$f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$;*

(c) $\displaystyle\lim_{k\to\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$

Proof From **(ii)** we find that for every $k \in \mathbb{N}$

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty.$$

Thus, **(a)** holds by Theorem 3.41.

We observe that f is measurable, by applying **(c)** in Theorem 3.28 to $\mathfrak{Re}\,f$ and $\mathfrak{Im}\,f$. Moreover, from **(i)** and **(ii)**, taking the limit $k \to \infty$, we find $|f(\boldsymbol{x})| \leq g(\boldsymbol{x})$ for every $\boldsymbol{x} \in \boldsymbol{A}$. This proves **(b)**. We only need to prove **(c)**.

Let $\varepsilon > 0$. From Lemma 3.6, there exists a measurable subset $\boldsymbol{B}$ of $\boldsymbol{A}$ satisfying $\mu_n(\boldsymbol{B}) < +\infty$ and such that

$$\int_{\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} > \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \varepsilon.$$

Using **(ii)** in Theorem 3.44 and Remark 3.18 we obtain

$$\int_{\boldsymbol{A}\setminus\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \varepsilon.$$

Let, for every $k \in \mathbb{N}$,

$$\boldsymbol{B}_k \overset{\text{def}}{=} \bigcap_{j=k}^{\infty} \left\{ \boldsymbol{x} \in \boldsymbol{B} \;\middle|\; |f_j(\boldsymbol{x}) - f(\boldsymbol{x})| < \frac{\varepsilon}{\max\{1, \mu_n(\boldsymbol{B})\}} \right\}.$$

From Theorem 3.41 and **(c)** in Theorem 3.6, we see that for every $k \in \mathbb{N}$ the set $\boldsymbol{B}_k$ is measurable and

$\boldsymbol{B}_k \subseteq \boldsymbol{B}_{k+1}$. From **(i)** we have $\bigcup_{k=0}^{\infty} \boldsymbol{B}_k = \boldsymbol{B}$. From **(a)** in Theorem 3.9 we obtain

$$\lim_{k\to\infty} \mu_n(\boldsymbol{B}_k) = \mu_n(\boldsymbol{B}).$$

Thus, $\mu_n(\boldsymbol{B} \setminus \boldsymbol{B}_k) \xrightarrow{k\to\infty} 0$; applying Theorem 3.45 there exists $h \in \mathbb{N}$ such that $\int_{\boldsymbol{B}\setminus\boldsymbol{B}_h} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < \varepsilon$. With this choice of h, applying Theorem 3.44, we have

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{B}_h} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{B}\setminus\boldsymbol{B}_h} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}\setminus\boldsymbol{B}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}. \quad (3.47)$$

If $k \geq h$, we can bound the first integral with

$$\frac{\varepsilon}{\max\{1, \mu_n(\boldsymbol{B})\}}\,\mu_n(\boldsymbol{B}_h) < \varepsilon.$$

The second integral can be bounded by $2\int_{\boldsymbol{B}\setminus\boldsymbol{B}_h} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < 2\,\varepsilon$. Finally, the third integral can be bounded by $2\int_{\boldsymbol{A}\setminus\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < 2\,\varepsilon$. We conclude that for $k \geq h$ we have

$$\left|\int_{\boldsymbol{A}} f_k(\boldsymbol{x}) - \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \leq \int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < 5\,\varepsilon,$$

from **(c)** in Theorem 3.42.

Remark 3.19 In the proof of Lebesgue dominated convergence theorem (Theorem 3.47) we have also shown that

$$\lim_{k\to\infty} \int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = 0.$$

Remark 3.20 It is easy to adapt Lebesgue dominated convergence theorem (Theorem 3.47) to a case where a function depends on a continuous variable ε. Consider a measurable function $f_\varepsilon : \boldsymbol{A} \longrightarrow \mathbb{C}$ and a summable function $g : \boldsymbol{A} \longrightarrow [0, +\infty]$ (so that $\int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$). Let f be the pointwise limit of f_ε, i.e.,

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad f(\boldsymbol{x}) = \lim_{\varepsilon\to 0^+} f_\varepsilon(\boldsymbol{x}).$$

Finally, we assume that for every $\varepsilon > 0$ and for every $\boldsymbol{x} \in \boldsymbol{A}$ the inequality $|f_\varepsilon(\boldsymbol{x})| \leq g(\boldsymbol{x})$ holds. Then, f_ε is summable for every $\varepsilon > 0$, f is summable, and

$$\lim_{\varepsilon\to 0^+} \int_{\boldsymbol{A}} f_\varepsilon(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \quad (3.48)$$

The sequential continuity theorem ensures that this statement is equivalent to Lebesgue dominated convergence theorem (Theorem 3.47). Indeed, $\lim_{x\to a} f(x) = \ell$ is equivalent to saying that for every sequence $(x_k)_{k\in\mathbb{N}}$ that converges to a, we have $f(x_k) \xrightarrow{k\to\infty} \ell$. This ensures that, considering the sequence $\left(\frac{1}{k}\right)_{k\in\mathbb{N}^*}$, which converges to 0, (3.48) is equivalent to

$$\lim_{\frac{1}{k}\to 0^+} \int_{\boldsymbol{A}} f_{\frac{1}{k}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \quad (3.49)$$

which is in turn equivalent to

$$\lim_{k\to\infty}\int_A f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{3.50}$$

We now start considering functions defined over a null set (i.e., over a set with zero Lebesgue measure).

Theorem 3.48 *Let $\boldsymbol{A}\in\mathcal{M}_n$ satisfy $\mu_n(\boldsymbol{A})=0$. Let $f:\boldsymbol{A}\longrightarrow[-\infty,+\infty]$ or $f:\boldsymbol{A}\longrightarrow\mathbb{C}$. Then:*

(a) *f is measurable;*

(b) *if f takes values in $[0,+\infty]$, then*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=0;$$

(c) *if f takes values in $\mathbb{C}$, then $f\in\mathcal{L}(\boldsymbol{A};\mathbb{C})$ and*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=0.$$

Proof In view of Theorem 3.19, every subset of $\boldsymbol{A}$ is measurable, and, obviously, its measure is 0. Therefore, recalling Definition 3.13, we can say that every function $f:\boldsymbol{A}\longrightarrow[-\infty,+\infty]$ is measurable. Moreover, from Definition 3.17 we can say that every function with domain $\boldsymbol{A}$ taking complex values is measurable. Since every subset of $\boldsymbol{A}$ has zero measure, the integral of every non-negative simple function with domain $\boldsymbol{A}$ is equal to 0. Hence, the integral of every function taking values in $[0,+\infty]$ is equal to 0. If f is complex-valued, we have

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}=0.$$

From Theorem 3.41, we obtain $f\in\mathcal{L}(\boldsymbol{A};\mathbb{C})$, and applying **(c)** in Theorem 3.42 we find

$$\left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|\leq\int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}=0,$$

from which we obtain **(c)**.

The following result will be useful in the sequel.

Corollary 3.3 *Let $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\in\mathcal{M}_n$ satisfy $\boldsymbol{C}\subseteq\boldsymbol{B}\subseteq\boldsymbol{A}$ and $\mu_n(\boldsymbol{A}\setminus\boldsymbol{C})=0$. Let $f:\boldsymbol{A}\longrightarrow[0,+\infty]$, or $f:\boldsymbol{A}\longrightarrow\mathbb{C}$. Let $g:\boldsymbol{B}\longrightarrow[0,+\infty]$ be either measurable or $g\in\mathcal{L}(\boldsymbol{B};[0,+\infty])$. Assume also that $f(\boldsymbol{x})=g(\boldsymbol{x})$ for all $\boldsymbol{x}\in\boldsymbol{C}$. Then:*

(a) *f is measurable;*

(b) *if f and g take values in $[0,+\infty]$, we have*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=\int_{\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}; \tag{3.51}$$

(c) *if f is complex-valued and $g\in\mathcal{L}(\boldsymbol{B};[0,+\infty])$, then $f\in\mathcal{L}(\boldsymbol{A};\mathbb{C})$ and (3.51) holds.*

Proof Since f and g coincide over $\boldsymbol{C}$, the restriction $f|_{\boldsymbol{C}}$ is measurable. The restriction $f|_{\boldsymbol{A}\setminus\boldsymbol{C}}$ is measurable by Theorem 3.48. Therefore, f is measurable by Theorem 3.31 (and if f is complex-valued, the same theorem must be applied to $\mathfrak{Re}f$ and $\mathfrak{Im}f$). This proves **(a)**.

Let us now consider **(b)**. Applying **(b)** in Theorem 3.43,

$$\begin{aligned}\int_A f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_A f(\boldsymbol{x})\chi_C\,\mathrm{d}\boldsymbol{x} + \int_A f(\boldsymbol{x})\chi_{A\setminus C}\,\mathrm{d}\boldsymbol{x}\\ &= \int_C f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{A\setminus C} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_C f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_C g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},\end{aligned}$$

from **(b)** in Theorem 3.48. Moreover,

$$\begin{aligned}\int_C g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_C g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{B\setminus C} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_A g(\boldsymbol{x})\chi_C\,\mathrm{d}\boldsymbol{x} + \int_A g(\boldsymbol{x})\chi_{B\setminus C}\,\mathrm{d}\boldsymbol{x} = \int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

We can use the same argument to prove **(c)**.

We conclude this part with an application of Theorem 3.39.

Theorem 3.49 *Let $\boldsymbol{A} \in \mathcal{M}_n$, and for all $k \in \mathbb{N}$ let $f_k \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$ be such that*

$$\sum_{k=0}^{\infty}\int_A |f_k(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty.$$

Then:

(a) *the series $\displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x})$ is absolutely convergent almost everywhere in $\boldsymbol{A}$;*

(b) *defining*

$$\begin{aligned} s : \boldsymbol{A} &\longrightarrow \mathbb{C}\\ \boldsymbol{x} &\mapsto \begin{cases} \displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x}) & \text{if the series is convergent}\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

then $s \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$ and

$$\int_A s(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty}\int_A f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof For every $\boldsymbol{x} \in \boldsymbol{A}$, let us define the function $g(\boldsymbol{x}) = \displaystyle\sum_{k=0}^{\infty} |f_k(\boldsymbol{x})|$, which is convergent in $[0,+\infty]$. From **(c)** in Theorem 3.35, the function g is measurable and $\displaystyle\int_A g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} < +\infty$. Thus, applying Theorem 3.46, we can

say that $g(\boldsymbol{x}) < +\infty$ almost everywhere, and we obtain **(a)**. Let us define the sets

$$\begin{aligned}
\boldsymbol{B} &= \{\boldsymbol{x} \in \boldsymbol{A} \,|\, g(\boldsymbol{x}) < +\infty\} \\
\boldsymbol{N} &= \{\boldsymbol{x} \in \boldsymbol{A} \,|\, g(\boldsymbol{x}) = +\infty\} \\
\boldsymbol{N}_1 &= \left\{\boldsymbol{x} \in \boldsymbol{A} \,\middle|\, \sum_{k=0}^{\infty} f_k(\boldsymbol{x}) \text{ is convergent but not absolutely convergent}\right\} \\
\boldsymbol{N}_2 &= \boldsymbol{N} \setminus \boldsymbol{N}_1.
\end{aligned}$$

We have $\boldsymbol{B} = \boldsymbol{A} \setminus \boldsymbol{N}$ and $\boldsymbol{N}_1 \cup \boldsymbol{N}_2 = \boldsymbol{N}$. The set $\boldsymbol{N}$ is measurable and has measure zero. Therefore, $\boldsymbol{B}$, $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are measurable, and $\mu_n(\boldsymbol{N}_1) = \mu_n(\boldsymbol{N}_2) = 0$ (applying Theorem 3.19). We find

$$s(\boldsymbol{x}) = \begin{cases} \displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \boldsymbol{B} \cup \boldsymbol{N}_1 \\ 0 & \text{if } \boldsymbol{x} \in \boldsymbol{N}_2. \end{cases}$$

If $\boldsymbol{x} \in \boldsymbol{B} \cup \boldsymbol{N}_1$, we have $s(\boldsymbol{x}) = \lim_{j\to\infty} s_j(\boldsymbol{x})$ with $s_j(\boldsymbol{x}) = \displaystyle\sum_{k=0}^{j} f_k(\boldsymbol{x})$. Moreover, for every $j \in \mathbb{N}$, we have $|s_j(\boldsymbol{x})| \leq \displaystyle\sum_{k=0}^{j} |f_k(\boldsymbol{x})| \leq g(\boldsymbol{x})$. Applying Lebesgue dominated convergence theorem (Theorem 3.47) and **(c)** in Theorem 3.43, we find that s is summable over $\boldsymbol{B} \cup \boldsymbol{N}_1$ and

$$\begin{aligned}
\int_{\boldsymbol{B}\cup\boldsymbol{N}_1} s(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \lim_{j\to\infty} \int_{\boldsymbol{B}\cup\boldsymbol{N}_1} s_j(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{j\to\infty} \sum_{k=0}^{j} \int_{\boldsymbol{B}\cup\boldsymbol{N}_1} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{k=0}^{\infty} \int_{\boldsymbol{B}\cup\boldsymbol{N}_1} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \sum_{k=0}^{\infty} \left(\int_{\boldsymbol{B}\cup\boldsymbol{N}_1} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{N}_2} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right) = \sum_{k=0}^{\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},
\end{aligned}$$

where we have used **(a)** in Theorem 3.42 and Theorem 3.48. Since $\displaystyle\int_{\boldsymbol{N}_2} s(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$, using Theorem 3.44, we obtain the result.

Example *3.30*

Let $\boldsymbol{A} \in \mathcal{M}_n$ satisfy $\mu_n(\boldsymbol{A}) < +\infty$. For every $k \in \mathbb{N}$, let $f_k : \boldsymbol{A} \longrightarrow \mathbb{C}$ be measurable and such that:

(i) there exists $M > 0$ satisfying $|f_k(\boldsymbol{x})| \leq M$ for all $k \in \mathbb{N}$ and for all $\boldsymbol{x} \in \boldsymbol{A}$;

(ii) for every $\boldsymbol{x} \in \boldsymbol{A}$, there exists $\lim_{k\to\infty} f_k(\boldsymbol{x})$, and let $f(\boldsymbol{x}) = \lim_{k\to\infty} f_k(\boldsymbol{x})$. Show that $f_k \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$ for all $k \in \mathbb{N}$, and $f \in \mathcal{L}(\boldsymbol{A};\mathbb{C})$, and

$$\lim_{k\to\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

The following result generalizes Theorem 3.44. This extension will be of importance in the context of obtaining formulae for the derivatives of distributions.

Theorem 3.50 *Let $(A_k)_{k\in\mathbb{N}}$ be a sequence of measurable pairwise disjoint subsets of $\mathbb{R}^n$, and let*

$$\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{A}_k.$$

Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. Then, the following statements are equivalent:

(i) *$f \in \mathcal{L}(\boldsymbol{A}; \mathbb{C})$;*

(ii) *for every $k \in \mathbb{N}$, defining $f_k = f|_{\boldsymbol{A}_k}$, then $f_k \in \mathcal{L}(\boldsymbol{A}_k; \mathbb{C})$.*

If **(i)** *holds (or, equivalently, if* **(ii)** *holds), we have*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty} \int_{\boldsymbol{A}_k} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{3.52}$$

Proof We prove that **(1)** implies **(2)**. Let f be integrable over $\boldsymbol{A}$, so that $|f|$ is integrable over $\boldsymbol{A}$. Since $\boldsymbol{A}_k \subseteq \boldsymbol{A}$, then f is integrable over $\boldsymbol{A}_k$ (and also $|f|$ is integrable over $\boldsymbol{A}_k$), so that the restriction of $|f|$ to $\boldsymbol{A}_k$, that we denote by $|f|_{\boldsymbol{A}_k}$, is integrable over $\boldsymbol{A}$. Let us define the partial sums

$$s_k = \int_{\boldsymbol{A}_0 \cup \boldsymbol{A}_1 \cup \ldots \cup \boldsymbol{A}_k} |f|.$$

Then, $(s_k)_{k\in\mathbb{N}}$ converges because the integral $\int_{\boldsymbol{A}} f$ exists by assumption. Moreover, since the sets $\boldsymbol{A}_k$ are pairwise disjoint, we have

$$s_k = \int_{\boldsymbol{A}_0} |f|_{\boldsymbol{A}_0} + \int_{\boldsymbol{A}_1} |f|_{\boldsymbol{A}_1} + \ldots + \int_{\boldsymbol{A}_k} |f|_{\boldsymbol{A}_k} = \sum_{j=0}^{k} \int_{\boldsymbol{A}_j} |f|_{\boldsymbol{A}_j}.$$

Letting k go to ∞ we find that $\displaystyle\sum_{k=0}^{\infty} \int_{\boldsymbol{A}_k} |f|_{\boldsymbol{A}_k} = \sum_{k=0}^{\infty} \int_{\boldsymbol{A}_k} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}$ also exists. To prove (3.52), we perform the same steps on the partial sums

$$\sigma_k = \int_{\boldsymbol{A}_0 \cup \boldsymbol{A}_1 \cup \ldots \cup \boldsymbol{A}_k} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{j=0}^{k} \int_{\boldsymbol{A}_j} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

and once again letting k go to ∞ we obtain (3.52).

We prove that **(2)** implies **(1)**. We assume that f is integrable over $\boldsymbol{A}_k$, so that $\displaystyle\int_{\boldsymbol{A}_k} f|_{\boldsymbol{A}_k} < +\infty$. Since by assumption

$$\sum_{k=1}^{\infty} \int_{\boldsymbol{A}_k} |f|_{\boldsymbol{A}_k} < +\infty,$$

then

$$s_k = \sum_{i=1}^{k} \int_{\boldsymbol{A}_i} |f|_{\boldsymbol{A}_i}$$

converges. This means that

$$s_k = \int_{\boldsymbol{A}_1} |f|_{\boldsymbol{A}_1} + \ldots + \int_{\boldsymbol{A}_k} |f|_{\boldsymbol{A}_k} = \int_{\boldsymbol{A}_1 \cup \ldots \cup \boldsymbol{A}_k} |f| \xrightarrow{k\to\infty} \int_{\boldsymbol{A}} |f|.$$

Therefore,

$$\int_{\boldsymbol{A}} |f| = \sum_{i=1}^{\infty} \int_{\boldsymbol{A}_i} |f|_{\boldsymbol{A}_i} < +\infty,$$

which means that f is integrable.

Exercises

Exercise 3.31 Use Lebesgue dominated convergence theorem (Theorem 3.47) to show that

$$\lim_{k\to\infty} \frac{1}{k} \int_{1/k}^{+\infty} \frac{\sin x}{x^2}\, \mathrm{d}x = 0.$$

[Solution: (3.29)]

3.10 Relation between Riemann and Lebesgue integral

In this section we want to compare the Lebesgue integration theory with the Riemann integration theory for functions of a single variable. A summary of the definitions and main properties of the one-dimensional Riemann integration theory are given in the Appendix. In what follows, we will denote by

$$\int_a^b f(x)\,\mathrm{d}x \quad \text{and by} \quad \int_{[a,b]} f(x)\,\mathrm{d}x$$

the Riemann and the Lebesgue integrals, respectively. We prove the following result.

Theorem 3.51 *Let $a, b \in \mathbb{R}$ with $a < b$. Let $f : [a,b] \longrightarrow \mathbb{R}$ be Riemann integrable. Then, $f \in \mathcal{L}([a,b];\mathbb{R})$ and*

$$\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f(x)\,\mathrm{d}x.$$

Proof Let us choose a sequence $(\sigma^k)_{k\in\mathbb{N}}$ of partitions of $[a,b]$ such that $\lim_{k\to\infty} (S(f,\sigma^k) - s(f,\sigma^k)) = 0$. This sequence exists in view of Theorem **??**. It is not restrictive, by virtue of (**??**), to assume that for every $k \in \mathbb{N}$ there holds $\sigma^k \subseteq \sigma^{k+1}$. Whenever such condition is not satisfied, we can, for example, substitute σ^k with $\sigma^1 \cup \ldots \cup \sigma^k$. Let $k \in \mathbb{N}$ be assigned. Let $\sigma_k = \{x_0, \ldots, x_{n_k}\}$ with $a = x_0$ and $x_{n_k} = b$. We define, for $x \in [a,b]$,

$$\varphi_k(x) = \begin{cases} \inf\limits_{x\in[x_j,x_{j+1}]} f(x) & \text{if } x_j \le x < x_{j+1} \text{ for } j \in \{0,\ldots,n_k-2\} \\ \inf\limits_{x\in[x_{n_k-1},x_{n_k}]} f(x) & \text{if } x_{n_k-1} \le x \le x_{n_k} \end{cases}$$

and

$$\psi_k(x) = \begin{cases} \sup\limits_{x\in[x_j,x_{j+1}]} f(x) & \text{if } x_j \le x < x_{j+1} \text{ for } j \in \{0,\ldots,n_k-2\} \\ \sup\limits_{x\in[x_{n_k-1},x_{n_k}]} f(x) & \text{if } x_{n_k-1} \le x \le x_{n_k}. \end{cases}$$

It is easy to see that for every $k \in \mathbb{N}$ the functions φ_k and ψ_k are simple over $[a,b]$, since for all $x \in [a,b]$

$$\varphi_k(x) = \sum_{j=0}^{n_k-2} \Big(\inf_{x\in[x_j,x_{j+1}]} f(x) \Big) \chi_{[x_j,x_{j+1}]}(x) + \Big(\inf_{x\in[x_{n_k-1},x_{n_k}]} f(x) \Big) \chi_{[x_{n_k-1},x_{n_k}]}(x)$$

and

$$\psi_k(x) = \sum_{j=0}^{n_k-2} \Big(\sup_{x\in[x_j,x_{j+1}]} f(x)\Big)\chi_{[x_j,x_{j+1}[}(x) + \Big(\sup_{x\in[x_{n_k-1},x_{n_k}]} f(x)\Big)\chi_{[x_{n_k-1},x_{n_k}]}(x).$$

Moreover, for every $x \in [a,b]$,

$$\varphi_k(x) \le f(x) \le \psi_k(x).$$

Since σ^{k+1} is finer than σ^k, for every $k \in \mathbb{N}$ there holds

$$\varphi_k \le \varphi_{k+1} \quad \text{and} \quad \psi_{k+1} \le \psi_k. \tag{3.53}$$

Moreover,

$$\int_{[a,b]} \varphi_k(x)\,\mathrm{d}x = s(f,\sigma^k) \quad \text{and} \quad \int_{[a,b]} \psi_k(x)\,\mathrm{d}x = S(f,\sigma^k).$$

It follows that

$$\lim_{k\to\infty} \int_{[a,b]} \varphi_k(x)\,\mathrm{d}x = \lim_{k\to\infty} \int_{[a,b]} \psi_k(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Indeed,

$$\left| \int_a^b f(x)\,\mathrm{d}x - \int_{[a,b]} \varphi_k(x)\,\mathrm{d}x \right| \le S(f,\sigma^k) - s(f,\sigma^k)$$

and

$$\left| \int_a^b f(x)\,\mathrm{d}x - \int_{[a,b]} \psi_k(x)\,\mathrm{d}x \right| \le S(f,\sigma^k) - s(f,\sigma^k).$$

For all $x \in [a,b]$ we define $\varphi(x) = \lim\limits_{k\to\infty} \varphi_k(x)$ and $\psi(x) = \lim\limits_{k\to\infty} \psi_k(x)$. The existence of these limits is ensured by the monotonicity in (3.53). There holds

$$\varphi(x) \le f(x) \le \psi(x) \tag{3.54}$$

for every $x \in [a,b]$. We also observe that, for all $x \in [a,b]$ and for every $k \in \mathbb{N}$,

$$\inf_{[a,b]} f \le \varphi_k(x) \le \psi_k(x) \le \sup_{[a,b]} f(x).$$

Choosing $M > 0$ so as to satisfy $-M \le \inf\limits_{[a,b]} f$ and $\sup\limits_{[a,b]} f \le M$, we will have $|\varphi_k(x)| \le M$ and $|\psi_k(x)| \le M$ for every $x \in [a,b]$.

We can now apply Lebesgue dominated convergence theorem (Theorem 3.47) to the two sequences $(\varphi_k)_{k\in\mathbb{N}}$ and $(\psi_k)_{k\in\mathbb{N}}$, by setting $g(x) = M$ for all $x \in [a,b]$. We will find that φ and ψ belong to $\mathcal{L}([a,b];\mathbb{R})$ and

$$\int_{[a,b]} \varphi(x)\,\mathrm{d}x = \int_{[a,b]} \psi(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

From this we obtain

$$\int_{[a,b]} (\psi(x) - \varphi(x))\,\mathrm{d}x = 0.$$

Applying **(b)** in Theorem 3.33, we find that $\varphi(x) = \psi(x)$ almost everywhere in $[a,b]$. Now, let $C = \{x \in [a,b] \,|\, \varphi(x) = \psi(x)\}$. If $x \in C$, by (3.54), we deduce that $f(x) = \varphi(x)$. By **(c)** in Corollary 3.3, with $B = A = [a,b]$, we obtain $f \in \mathcal{L}([a,b];\mathbb{R})$. Moreover,

$$\int_{[a,b]} f(x)\,\mathrm{d}x = \int_{[a,b]} \varphi(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Example *3.31*

The converse of Theorem 3.51 does not hold in general. For example, consider the Dirichlet restricted to $[0,1]$, given by

$$\begin{aligned} f : [0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases} \end{aligned}$$

This function is not Riemann integrable over $[0,1]$. In fact, consider a partition $\sigma = \{x_0, \ldots, x_n\}$ of $[0,1]$ such that $0 = x_0 < x_1 < \ldots < x_n = 1$. Every interval $[x_{k-1}, x_k]$ contains both elements in $\mathbb{Q}$ and elements in $\mathbb{R} \setminus \mathbb{Q}$ (why?). Therefore, for $k \in \{1, \ldots, n\}$, we have

$$\inf_{[x_{k-1},x_k]} f = 0 \quad \text{and} \quad \sup_{[x_{k-1},x_k]} f = 1.$$

It follows that $s(f,\sigma) = 0$, whereas

$$S(f,\sigma) = 1 \cdot ((x_1 - x_0) + \ldots + (x_k - x_{k-1}) + \ldots + (x_n - x_{n-1})) = 1.$$

It follows that

$$\underline{\int_0^1} f(x)\,\mathrm{d}x = 0 \quad \text{and} \quad \overline{\int_0^1} f(x)\,\mathrm{d}x = 1.$$

This shows that f is not Riemann integrable. However, f is a simple function: let $E_1 = \mathbb{Q} \cap [0,1]$ and $E_2 = [0,1] \setminus E_1$, $\alpha_1 = 1$ and $\alpha_2 = 0$. Since $\mathbb{Q}$ is countable, we have $\mu^*(\mathbb{Q}) = 0$, so that $\mu^*(E_1) = 0$. It follows that E_1 is measurable, and therefore so is E_2. Moreover, f is non-negative and

$$\int_{[0,1]} f(x)\,\mathrm{d}x = 1 \cdot \mu(E_1) + 0 \cdot \mu(E_2) = 1 \cdot 0 + 0 \cdot 1 = 0.$$

We now consider improper (or generalized) integrals. Importantly, differently from standard Riemann integrals, not all improper integrals are defined as Lebesgue integrals. We begin recalling the following definition. of improper integral where the domain of integration is either $[a, +\infty)$ or $[a, b)$. These two cases can be captured simultaneously by considering a domain $[a, b)$ where $-\infty < a < b \leq +\infty$.

Definition 3.18 *Let $-\infty < a < b \leq +\infty$. Let $f : [a,b) \longrightarrow \mathbb{R}$ be Riemann integrable over $[a,c]$ for every $c \in [a,b)$. We say that f is:*

- *integrable in a generalized sense in $[a,b)$ if the limit*

$$\lim_{y \to b} \int_a^y f(x)\,\mathrm{d}x$$

exists and it is finite. If this limit exists and it is finite, it is denoted by $\displaystyle\int_a^b f(x)\,\mathrm{d}x$

- *absolutely integrable in a generalized sense in $[a,b)$ if $|f|$ is integrable in a generalized sense in $[a,b)$.*

Recall that an absolutely integrable function in a generalized sense is also integrable in a generalized sense, but the converse is not true in general.

First, we address the case of non-negative functions.

Lemma 3.7 *Let $-\infty < a < b \leq +\infty$. Let $f : [a,b) \longrightarrow [0,+\infty]$ be Riemann integrable over $[a,c]$ for every*

$c \in [a,b)$. Then, f is measurable and

$$\int_{[a,b)} f(x)\,\mathrm{d}x = \lim_{y\to b}\int_a^y f(x)\,\mathrm{d}x. \tag{3.55}$$

Proof First, we observe that the limit in (3.55) exits (finite or infinite) because the function $y \mapsto \int_a^y f(x)\,\mathrm{d}x$ is non-increasing. Let $(b_k)_{k\in\mathbb{N}}$ be an increasing sequence of elements in the interval $[a,b)$ such that $\lim_{k\to\infty} b_k = b$. We define $f_k = f\chi_{[a,b_k]}$. From Theorem 3.31 and Theorem 3.51, for every $k \in \mathbb{N}$, the function f_k is measurable over $[a,b]$. Moreover,

$$\begin{aligned}\int_{[a,b)} f_k(x)\,\mathrm{d}x &= \int_{[a,b)} f(x)\chi_{[a,b_k]}(x)\,\mathrm{d}x\\ &= \int_{[a,b_k]} f(x)\,\mathrm{d}x = \int_a^{b_k} f(x)\,\mathrm{d}x.\end{aligned}$$

For every $x \in [a,b)$, we have $f(x) = \lim_{k\to\infty} f_k(x)$ and for every $k \in \mathbb{N}$, there holds $0 \le f_k \le f_{k+1} \le f$. From the monotone convergence theorem (Theorem 3.34), the function f is measurable and

$$\begin{aligned}\int_{[a,b)} f(x)\,\mathrm{d}x &= \lim_{k\to\infty}\int_{[a,b)} f_k(x)\,\mathrm{d}x\\ &= \lim_{k\to\infty}\int_a^{b_k} f(x)\,\mathrm{d}x = \lim_{y\to b}\int_a^y f(x)\,\mathrm{d}x.\end{aligned}$$

Theorem 3.52 *Let $-\infty < a < b \le +\infty$. Let $f : [a,b) \longrightarrow \mathbb{R}$ be absolutely integrable in a generalized sense over $[a,b)$. Then,*

- *the limit $\lim_{y\to b}\int_a^y f(x)\,\mathrm{d}x$ exists in $\mathbb{R}$;*
- *f is summable in $[a,b)$, i.e., $f \in \mathcal{L}([a,b);\mathbb{R})$;*
- *$\int_{[a,b)} f(x)\,\mathrm{d}x = \lim_{y\to b}\int_a^y f(x)\,\mathrm{d}x$.*

Proof Using the same approach of Lemma 3.7, we can see that f is measurable over $[a,b]$. Moreover,

$$\int_{[a,b)} |f(x)|\,\mathrm{d}x = \int_a^b |f(x)|\,\mathrm{d}x < +\infty.$$

Applying Theorem 3.39, we find $f \in \mathcal{L}([a,b);\mathbb{R})$. Finally, if f_k is the same as in the proof of Lemma 3.7, we have $|f_k| \le |f|$ for every $k \in \mathbb{N}$. Using the same argument of the proof of Lemma 3.7, applying Theorem 3.44, we find that for every $k \in \mathbb{N}$ we have $f_k \in \mathcal{L}([a,b);\mathbb{R})$. Moreover,

$$\int_{[a,b)} f_k(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Applying Lebesgue dominated convergence theorem (Theorem 3.47), we conclude that

$$\int_{[a,b)} f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

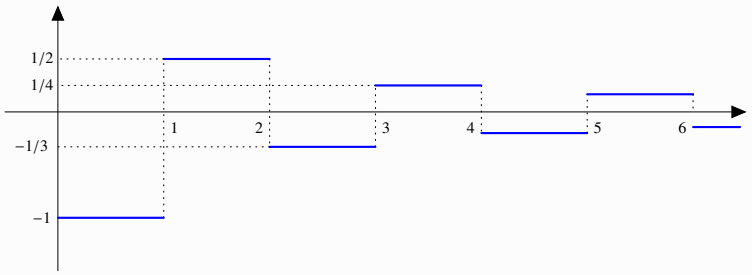
We clearly have analogous results for intervals in the form $(a, b]$ with $-\infty \leq a < b < +\infty$.

Remark 3.21 Theorem 3.52 is not true if f is only integrable in a generalized sense. In fact, we recall that the summability of f implies the summability of $|f|$ (see also Theorem 3.39).

Example *3.32*

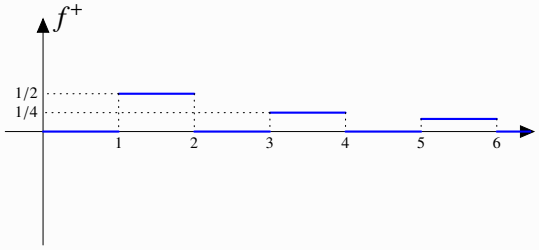
Consider a function $f : (0, +\infty) \longrightarrow \mathbb{R}$ defined in this way. For every $k \in \mathbb{N}^*$ and for every $x \in (k-1, k]$,

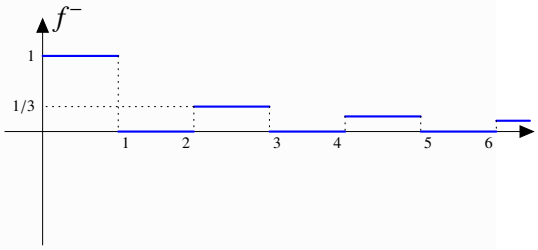
$$f(x) = (-1)^k \frac{1}{k}.$$



For every $k \in \mathbb{N}^*$, for every $x \in (k-1, k]$, we have:

- $f^+(x) = \dfrac{1}{k}$ if k is even, and $f^+(x) = 0$ if k is odd;
- $f^-(x) = \dfrac{1}{k}$ if k is odd, and $f^-(x) = 0$ if k is even.





We find

$$\int_{(0,+\infty)} f^+(x)\,\mathrm{d}x = \sum_{k=1}^{\infty} \frac{1}{2k} = +\infty \quad \text{and} \quad \int_{(0,+\infty)} f^-(x)\,\mathrm{d}x = \sum_{k=0}^{\infty} \frac{1}{2k+1} = +\infty.$$

Hence, in this case the integral $\displaystyle\int_0^{+\infty} f(x)\,\mathrm{d}x$ does not exist as a Lebesgue integral. However, it is easy to see that this integral exists as an improper Riemann integral. Indeed,

$$\lim_{y \to +\infty} \int_0^y f(x)\,\mathrm{d}x = \sum_{k=1}^{\infty} (-1)^k \frac{1}{k}$$

exists in $\mathbb{R}$ (this is due to Leibniz's test, and it can be shown that its value is $-\ln 2$). This is consistent with the fact that f does not satisfy the assumption of Theorem 3.52. In fact, $|f|$ is not integrable in a generalized sense:

$$\lim_{y \to +\infty} \int_0^y |f(x)|\,\mathrm{d}x = \sum_{k=1}^{\infty} \frac{1}{k} = +\infty.$$

With this example, we have, in essence, recast a Lebesgue integral into a series. This series is conditionally convergent: a rearrangement of its elements may lead to convergence to a different (arbitrary) value, or to diverge, or even to oscillate. This seems to violates a fundamental principle: if we decompose a surface into parts, the area should not change. Correspondingly, the value of an integral should not depend on the arrangement of the terms in the sum. In other words, in a case where a function is not absolutely integrable in a generalized sense but whose improper integral exists in $\mathbb{R}$, taking the limit over bounded sections of the domain in a specific way may leads to convergence, but if we permute the positive and negative parts of the function we can make the improper integral converge to different values, or diverge, or oscillate.

Example *3.33*

Consider a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined in this way: for every $k \in \mathbb{N}^*$ and for every $x \in (k-1, k]$,

$$f(x) = (-1)^k \frac{1}{k^2}.$$

In this case, $k \in \mathbb{N}^*$ and for every $x \in (k-1, k]$,

$$f^+(x) = \begin{cases} 1/k^2 & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases} \quad \text{and} \quad f^-(x) = \begin{cases} 1/k^2 & \text{if } k \text{ is odd} \\ 0 & \text{if } k \text{ is even.} \end{cases}$$

The function f is summable, because

$$\int_0^{+\infty} f^+(x)\,\mathrm{d}x = \sum_{k=1}^{\infty} \frac{1}{(2\,k)^2} \quad \text{and} \quad \int_0^{+\infty} f^-(x)\,\mathrm{d}x = \sum_{k=0}^{\infty} \frac{1}{(2\,k+1)^2}$$

are both convergent (and it can be shown that they are convergent, respectively, to $\pi^2/24$ and $\pi^2/8$). Alternatively, one can check the finiteness of the integral of the absolute value:

$$\int_0^{+\infty} |f(x)|\,\mathrm{d}x = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Hence, f is absolutely integrable in a generalized sense, and therefore the improper integral

$$\int_0^{+\infty} f(x)\,\mathrm{d}x$$

is defined also as a Lebesgue integral, and its value is $\pi^2/24 - \pi^2/8 = -\pi^2/12$.

Remark 3.22 As aforementioned, there are counterparts for Lemma 3.7 and Theorem 3.52 considering $-\infty \le a < b < +\infty$. The readers are invited to formulate these for themselves. We have to be more careful in the case where $a = -\infty$ and $b = +\infty$. Indeed, if $f : \mathbb{R} \longrightarrow \mathbb{R}$ and if both integrals

$$\int_{-\infty}^{c} f(x)\,\mathrm{d}x \quad \text{and} \quad \int_c^{+\infty} f(x)\,\mathrm{d}x$$

exist for some $c \in \mathbb{R}$, then we say that the integral $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ exists, and its value is defined to be their sum:

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x \stackrel{\text{def}}{=} \int_{-\infty}^{c} f(x)\,\mathrm{d}x + \int_c^{+\infty} f(x)\,\mathrm{d}x = \lim_{y \to +\infty} \int_{-y}^{c} f(x)\,\mathrm{d}x + \lim_{w \to +\infty} \int_c^{w} f(x)\,\mathrm{d}x. \tag{3.56}$$

It is possible to see that this definition is well-posed, in the sense that the value of (3.56) does not depend on

the value of c. Moreover, it is easy to see that its value is equal to the "symmetric limit"

$$\lim_{y\to+\infty}\int_{-y}^{y} f(x)\,\mathrm{d}x.$$

However, this symmetric limit may exist even when $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ does not exist. An example is the function $f(x)=x$ for $x\in\mathbb{R}$. In this case,

$$\int_{-\infty}^{0} x\,\mathrm{d}x=\lim_{y\to+\infty}\int_{-y}^{0} x\,\mathrm{d}x=-\infty \quad\text{and}\quad \int_{0}^{+\infty} x\,\mathrm{d}x=\lim_{y\to+\infty}\int_{0}^{y} x\,\mathrm{d}x=+\infty,$$

so that $\displaystyle\int_{-\infty}^{+\infty} x\,\mathrm{d}x$ does not exist. However, the symmetric limit is 0:

$$\lim_{y\to+\infty}\int_{-y}^{y} x\,\mathrm{d}x=\lim_{y\to+\infty}\left[\frac{x^2}{2}\right]_{-y}^{y}=\lim_{y\to+\infty}0=0.$$

In this case, the value of the symmetric limit is called the *Cauchy principal value* of the integral $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$. We observe also that the integral $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ does not exist as a Lebesgue integral. In fact, the function is not summable.

Remark 3.23 In **(b)** of Fatou's lemma (Theorem 3.38), the strict inequality may hold. Consider for example

$$\begin{aligned} f_k : [0,+\infty) &\longrightarrow \mathbb{R}\\ x &\mapsto \chi_{[k,+\infty)}(x). \end{aligned}$$

Then, for every $x\in[0,+\infty)$ we have

$$\liminf_{k\to\infty} f_k(x)=\lim_{k\to\infty} f_k(x)=0.$$

On the other hand, from Lemma 3.7, for every $k\in\mathbb{N}$ we have $\displaystyle\int_A f_k(x)\,\mathrm{d}x=+\infty$. It follows that

$$\int_A \liminf_{k\to\infty} f_k(x)=0, \qquad \text{but} \qquad \liminf_{k\to\infty}\int_A f_k(x)=+\infty.$$

Example *3.34*

Let $a\in\mathbb{R}$. We want to find

$$\int_{[1,+\infty)} x^a\,\mathrm{d}x.$$

Let

$$\begin{aligned} f : [1,+\infty) &\longrightarrow \mathbb{R}\\ x &\mapsto x^a. \end{aligned}$$

This function is continuous and non-negative. Let $c\in(1,+\infty)$. Then

$$\int_1^c x^a\,\mathrm{d}x=\begin{cases}\left[\frac{x^{a+1}}{a+1}\right]_1^c=\frac{c^{1+a}-1}{1+a} & \text{if } a\neq-1\\ \ln c & \text{if } a=-1\end{cases}$$

Taking the limit as $c \to +\infty$ we obtain

$$\int_{[1,+\infty)} x^a \, \mathrm{d}x = \begin{cases} \frac{1}{|1+a|} & \text{if } a < -1 \\ +\infty & \text{if } a \geq -1. \end{cases}$$

Example *3.35*

Let $a \in \mathbb{R}$. We want to find

$$\int_{(0,1]} x^a \, \mathrm{d}x.$$

Let

$$\begin{aligned} f : (0,1] &\longrightarrow \mathbb{R} \\ x &\mapsto x^a. \end{aligned}$$

This function is continuous and non-negative. Let $c \in (0,1)$. Then

$$\int_c^1 x^a \, \mathrm{d}x = \begin{cases} \left[\frac{x^{a+1}}{a+1}\right]_c^1 = \frac{1-c^{1+a}}{1+a} & \text{if } a \neq -1 \\ -\ln c & \text{if } a = -1 \end{cases}$$

Taking the limit as $c \to 0^+$ we obtain

$$\int_{(0,1]} x^a \, \mathrm{d}x = \begin{cases} \frac{1}{1+a} & \text{if } a > -1 \\ +\infty & \text{if } a \leq -1. \end{cases}$$

Example *3.36*

Let $a \in \mathbb{R}$. We want to find

$$\int_0^{+\infty} x^a \, \mathrm{d}x.$$

We observe that $(0, +\infty) = (0,1] \cup (1, +\infty)$. From the additivity

$$\int_0^{+\infty} x^a \, \mathrm{d}x = \int_{(0,1]} x^a \, \mathrm{d}x + \int_{[1,+\infty)} x^a \, \mathrm{d}x.$$

From the fact that $\mu(\{1\}) = 0$, we obtain

$$\int_{[1,+\infty)} x^a \, \mathrm{d}x = \int_{\{1\}} x^a \, \mathrm{d}x + \int_{(1,+\infty)} x^a \, \mathrm{d}x = \int_{(1,+\infty)} x^a \, \mathrm{d}x.$$

From the previous two examples, it follows that, for any $a \in \mathbb{R}$, we have

$$\int_0^{+\infty} x^a \, \mathrm{d}x = +\infty.$$

Example *3.37*

Consider the function

$$\begin{aligned} f : [1,+\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\sin \frac{1}{x}}{x^2}. \end{aligned}$$

Since f is continuous, it is Riemann integrable over every interval $[1, c]$ with $1 < c < +\infty$. We verify that f is summable and we compute the integral

$$\int_{[1,+\infty)} f(x) \, \mathrm{d}x.$$

First, we observe that for all $x \geq 1$ we have $|f(x)| \leq \frac{1}{x^2}$. Thus,

$$\lim_{c \to +\infty} \int_1^c |f(x)| \, \mathrm{d}x \leq \lim_{c \to +\infty} \int_1^c \frac{1}{x^2} \, \mathrm{d}x < +\infty,$$

which shows that the assumptions of Theorem 3.52 are satisfied, i.e., f is absolutely integrable in a generalized sense. It follows from Theorem 3.52 that f is summable. Moreover,

$$\int_{[1,+\infty)} f(x) \, \mathrm{d}x = \lim_{c \to +\infty} \int_1^c f(x) \, \mathrm{d}x = \lim_{c \to +\infty} \left[\cos \frac{1}{x} \right]_1^c = \lim_{c \to +\infty} \left(\cos \frac{1}{c} - \cos 1 \right) = 1 - \cos 1.$$

Example *3.38* Let $-\infty \leq a < b < +\infty$ and $f : (a, b] \longrightarrow \mathbb{R}$. State and prove the counterparts of Lemma 3.7 and Theorem 3.52 for the generalized integral

$$\int_a^b f(x) \, \mathrm{d}x = \lim_{y \to a} \int_y^b f(x) \, \mathrm{d}x$$

under the assumption that f is Riemann integrable over $[c, b]$ for every $c \in (a, b]$.

Example *3.39* Writing f as $\mathfrak{Re} f + i \, \mathfrak{Im} f$, generalize the previous definitions and results to the case of complex-valued functions.

3.11 Reduction of integrals

In this section we consider exclusively the multidimensional case, and we present results that allow one, in many concrete situations, to recast multidimensional integrals into integrals of lower dimensions.

Consider the integral

$$\int_{\mathbb{R}^2} f(x, y) \, \mathrm{d}x \, \mathrm{d}y, \tag{3.57}$$

where f is a measurable function. One way to solve this integral is to fix x and compute a one-dimensional integral in y, and then take the resulting quantity (which is now a function of x) and integrate it in x, so that we are effectively computing

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x, y) \, \mathrm{d}y \Big) \, \mathrm{d}x. \tag{3.58}$$

Alternatively, we could swap the roles of x and y and compute

$$\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x, y) \, \mathrm{d}x \Big) \, \mathrm{d}y. \tag{3.59}$$

Roughly speaking, Tonelli's theorem and Fubini's theorem provide conditions under which the three integrals (8.5-3.59) are equal. If these conditions are satisfied.

Tonelli's theorem can be applied when the function f takes values in $[0, +\infty)$, and only requires this function to be measurable. Fubini's theorem does not require f to only take non-negative values, but requires the stronger assumption of summability.

The answer to this question is negative in general, as the following classic example shows.

Example *3.40* Consider the integral

$$\iint_{[0,1] \times [0,1]} \frac{x^2 - y^2}{(x^2 + y^2)^2} \, \mathrm{d}x \, \mathrm{d}y.$$

We have

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2} \, \mathrm{d}y \right) \mathrm{d}x = \int_0^1 \left(\int_0^1 \frac{\partial}{\partial y} \frac{y}{x^2+y^2} \, \mathrm{d}y \right) \mathrm{d}x$$
$$= \int_0^1 \left[\frac{y}{x^2+y^2} \right]_0^1 \mathrm{d}x = \int_0^1 \frac{1}{x^2+1} \, \mathrm{d}x = \arctan 1 = \frac{\pi}{4}$$

and

$$\int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2+y^2)^2} \, \mathrm{d}x \right) \mathrm{d}y = \int_0^1 \left(\int_0^1 \frac{\partial}{\partial x} \frac{-x}{x^2+y^2} \, \mathrm{d}x \right) \mathrm{d}y$$
$$= \int_0^1 \left[\frac{-x}{x^2+y^2} \right]_0^1 \mathrm{d}x = -\int_0^1 \frac{1}{y^2+1} \, \mathrm{d}y = -\frac{\pi}{4}.$$

In this case, we were not authorized to reduce the integral: indeed, reducing it into two different ways gives two different values of the integral. This means that the reduction is only possible under some conditions. The theorems that give us these conditions are two:

- Tonelli's theorem: it requires the function to be non-negative, and only assumes that the integrand be measurable;
- Fubini's theorem: it does not require the function to be non-negative, but it assumes that the integrand be summable.

In this case, neither Tonelli's theorem nor Fubini's theorem can be applied. In fact, the integrand function is certainly measurable, but it is not non-negative in the domain $[0,1]\times[0,1]$. Thus, the assumptions in Tonelli's theorem are not satisfied. One can see that Fubini's theorem cannot be applied either, because the integrand function is not summable. If it was, we would have

$$\int_0^1 \left(\int_0^1 \frac{|x^2 - y^2|}{(x^2+y^2)^2} \, \mathrm{d}x \right) \mathrm{d}y \geq \int_0^1 \left(\int_0^y \frac{|x^2 - y^2|}{(x^2+y^2)^2} \, \mathrm{d}x \right) \mathrm{d}y = \int_0^1 \left(\int_0^y \frac{y^2 - x^2}{(x^2+y^2)^2} \, \mathrm{d}x \right) \mathrm{d}y$$
$$\geq \int_0^1 \left(\int_0^y \frac{y^2 - x^2}{(2\,y^2)^2} \, \mathrm{d}x \right) \mathrm{d}y = \int_0^1 \left(\int_0^y \frac{y^2 - x^2}{4\,y^4} \, \mathrm{d}x \right) \mathrm{d}y$$
$$= \int_0^1 \left[\frac{1}{4\,y^2} x - \frac{x^3}{12\,y^4} \right]_0^y \mathrm{d}y = \int_0^1 \left(\frac{1}{4\,y} - \frac{1}{12\,y} \right) \mathrm{d}y = \frac{1}{6} \int_0^1 \frac{1}{y} \, \mathrm{d}y = +\infty.$$

We start by introducing some important preliminaries.

Definition 3.19 *A* multi-interval *in $\mathbb{R}^n$ is the union of a finite set of n-dimensional intervals.*

We denote by $\mathcal{P}$ the set of multi-intervals in $\mathbb{R}^n$.

Lemma 3.8 *Let $\boldsymbol{A}$ be a multi-interval of $\mathbb{R}^n$. Then, $\boldsymbol{A}$ is the union of a finite set of pairwise disjoint n-dimensional intervals.*

Proof Let $\boldsymbol{A} = \boldsymbol{I}^1 \cup \ldots \boldsymbol{I}^p$, where $\boldsymbol{I}^1, \ldots, \boldsymbol{I}^p$ are n-dimensional intervals. We prove this result by induction over p. If $p = 1$, the result is trivial. Suppose that the result is true for a certain $p \geq 1$. Let $\boldsymbol{A} = \boldsymbol{I}^1 \cup \ldots \boldsymbol{I}^{p+1}$. From the inductive assumption, there exist pairwise disjoint n-dimensional intervals $\boldsymbol{J}^1, \ldots, \boldsymbol{J}^q$ such that

$$\bigcup_{k=1}^{p} \boldsymbol{I}^p = \bigcup_{i=1}^{q} \boldsymbol{J}^i.$$

It follows that

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{I}^{p+1} \cup \bigcup_{i=1}^{q} \boldsymbol{J}^i = \boldsymbol{I}^{p+1} \cup \Big(\bigcup_{i=1}^{q} \boldsymbol{J}^i \setminus \boldsymbol{I}^{p+1} \Big) \\ &= \boldsymbol{I}^{p+1} \cup \bigcup_{i=1}^{q} \boldsymbol{J}^i = \boldsymbol{I}^{p+1} \cup \Big(\bigcup_{i=1}^{q} \boldsymbol{J}^i \cap (\boldsymbol{I}^{p+1})^{\complement} \Big). \end{aligned}$$

In this union, the sets are pairwise disjoint. We now apply **(c)** in Theorem 3.13. For $i \in \{1, \ldots, q\}$ we have

$$\boldsymbol{J}^i \setminus \boldsymbol{I}^{p+1} = \bigcup_{\ell=1}^{2n} \boldsymbol{I}^{i,\ell},$$

where $\boldsymbol{I}^{i,\ell}$ are pairwise disjoint n-dimensional intervals. Thus,

$$\boldsymbol{A} = \boldsymbol{I}^{p+1} \cup \bigcup_{i=1}^{q} \bigcup_{\ell=1}^{2n} \boldsymbol{I}^{i,\ell},$$

and the result is proved because the elements of this union are pairwise disjoint.

We denote by $\mathcal{P}^\infty$ the set of unions of a countable set of n-dimensional intervals.

Lemma 3.9 *Let $m \in \mathbb{N}^*$ and $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathcal{P}^\infty$. Then, $\boldsymbol{A}_1 \cap \ldots \cap \boldsymbol{A}_m \in \mathcal{P}^\infty$.*

Proof We prove this result only in the case $m = 2$. The general case can be treated by induction. If

$$\boldsymbol{A}_1 = \bigcup_{k=0}^{\infty} \boldsymbol{I}^{1,k} \quad \text{and} \quad \boldsymbol{A}_2 = \bigcup_{k=0}^{\infty} \boldsymbol{I}^{2,k},$$

where $(\boldsymbol{I}^{1,k})_{k\in\mathbb{N}}$ and $(\boldsymbol{I}^{2,k})_{k\in\mathbb{N}}$ are sequences of n-dimensional intervals, we have

$$\boldsymbol{A}_1 \cap \boldsymbol{A}_2 = \bigcup_{h=0}^{\infty} \bigcup_{k=0}^{\infty} (\boldsymbol{I}^{1,h} \cap \boldsymbol{I}^{2,k}).$$

From **(b)** in Theorem 3.13, for every $h, k \in \mathbb{N}$, the set $\boldsymbol{I}^{1,h} \cap \boldsymbol{I}^{2,k}$ is an n-dimensional interval. Moreover, $\mathbb{N} \times \mathbb{N}$ is countable, so that $\boldsymbol{A}_1 \cap \boldsymbol{A}_2 \in \mathcal{P}^\infty$.

Definition 3.20 *Let $n, m \in \mathbb{N}^*$. We denote by $(\boldsymbol{x}, \boldsymbol{y})$ the generic element of $\mathbb{R}^{m+n}$ such that $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{y} \in \mathbb{R}^n$. Let $\boldsymbol{A} \subseteq \mathbb{R}^{m+n}$. Let $\boldsymbol{x} \in \mathbb{R}^m$. We denote*

$$\boldsymbol{A}_{\boldsymbol{x}} = \{\boldsymbol{y} \in \mathbb{R}^n \mid (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{A}\}.$$

Theorem 3.53 [Cavalieri's Principle]
Let $n, m \in \mathbb{N}^*$. Let $\boldsymbol{A} \in \mathcal{M}_{m+n}$. Then:

(a) for almost every $\boldsymbol{x} \in \mathbb{R}^m$, we have $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$;

(b) defining

$$g : \mathbb{R}^m \longrightarrow [0, +\infty]$$
$$\boldsymbol{x} \mapsto \begin{cases} \mu_n(\boldsymbol{A}_{\boldsymbol{x}}) & \text{if } \boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n \\ 0 & \text{otherwise} \end{cases}$$

Then, g is measurable;

(b) $\mu_{m+n}(\boldsymbol{A}) = \displaystyle\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.

Proof We denote, for the moment, by $\mathcal{S}$ the set of $\boldsymbol{A} \in \mathcal{M}_{m+n}$ such that, for every $\boldsymbol{x} \in \mathbb{R}^m$, we have $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$ and the function that maps $\boldsymbol{x}$ to $\mu_n(\boldsymbol{A}_{\boldsymbol{x}})$ is measurable over $\mathbb{R}^m$, and

$$\int_{\mathbb{R}^m} \mu_n(\boldsymbol{A}_{\boldsymbol{x}})\,\mathrm{d}x = \mu_{m+n}(\boldsymbol{A}).$$

We carry out the proof in several steps:

(i) Since $\boldsymbol{A}$ is an n-dimensional interval, it is easy to see that $\boldsymbol{A} \in \mathcal{S}$. Indeed, let $\boldsymbol{A} = \boldsymbol{I}_1 \times \boldsymbol{I}_2 \times \ldots \times \boldsymbol{I}_{m+n}$, where $\boldsymbol{I}_1, \ldots, \boldsymbol{I}_{m+n}$ are bounded intervals of $\mathbb{R}$. Then,

$$\boldsymbol{A}_{\boldsymbol{x}} = \begin{cases} \boldsymbol{I}_{m+1} \times \ldots \times \boldsymbol{I}_{m+n} & \text{if } \boldsymbol{x} \in \boldsymbol{I}_1, \ldots, \boldsymbol{I}_m \\ \varnothing & \text{otherwise} \end{cases}$$

From Theorem 3.20,

$$\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = \begin{cases} \ell(\boldsymbol{I}_{m+1}) \cdot \ldots \cdot \ell(\boldsymbol{I}_{m+n}) & \text{if } \boldsymbol{x} \in \boldsymbol{I}_1, \ldots, \boldsymbol{I}_m \\ 0 & \text{otherwise} \end{cases}$$

Hence, in this case g is simple, and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \prod_{j=m+1}^{m+n} \ell(\boldsymbol{I}_j)\,\mu_m\Big(\prod_{j=1}^{m} \boldsymbol{I}_j\Big) = \prod_{j=1}^{m+n} \ell(\boldsymbol{I}_j) = \mu_{m+n}(\boldsymbol{A}).$$

(ii) Let $\boldsymbol{A}^k \in \mathcal{S}$ for all $k \in \mathbb{N}$, and let $\boldsymbol{A}^k$ be pairwise disjoint. We show that if $\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k$, we have $\boldsymbol{A} \in \mathcal{S}$. In fact, for all $\boldsymbol{x} \in \mathbb{R}^m$, there holds $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = \sum_{k=0}^{\infty} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})$. Applying **(c)** of Theorem 3.35, we find that g is measurable and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty} \int_{\mathbb{R}^m} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})\,\mathrm{d}\boldsymbol{x} = \sum_{k=0}^{\infty} \mu_{m+n}(\boldsymbol{A}^k) = \mu_{m+n}(\boldsymbol{A}).$$

(iii) Every n-dimensional interval belongs to $\mathcal{S}$. This follows from Lemma 3.8 and also by **(a-b)**.

(iv) Let $\boldsymbol{A}^k \in \mathcal{S}$ and $\boldsymbol{A}^k \subseteq \boldsymbol{A}^{k+1}$ for all $k \in \mathbb{N}$. Then, if $\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k$, then $\boldsymbol{A} \in \mathcal{S}$. Indeed, for every $\boldsymbol{x} \in \mathbb{R}^m$, we have $\boldsymbol{A}_{\boldsymbol{x}} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k_{\boldsymbol{x}}$ and $\boldsymbol{A}^k_{\boldsymbol{x}} \subseteq \boldsymbol{A}^{k+1}_{\boldsymbol{x}}$ for every $k \in \mathbb{N}$. From **(a)** in Theorem 3.11 it follows that for all $\boldsymbol{x} \in \mathbb{R}^m$ we have

$$\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = \lim_{k \to \infty} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}}).$$

Applying Theorem 3.33, we obtain that g is measurable and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \mu_{m+n}(\boldsymbol{A}^k) = \mu_{m+n}(\boldsymbol{A}),$$

by applying again **(a)** in Theorem 3.11.

(v) $\mathcal{P}^\infty \subseteq \mathcal{S}$. Indeed, let $\boldsymbol{A} = \bigcup_{k=0}^\infty \boldsymbol{I}^k$, where, for all $k \in \mathbb{N}$, the set $\boldsymbol{I}^k$ is an n-dimensional interval. Let $\boldsymbol{A}^k = \bigcup_{j=0}^{k} \boldsymbol{I}^j$. From **(c)**, for every k we have $\boldsymbol{A}^k \in \mathcal{S}$. The result then follows from **(d)**.

(vi) Let $\boldsymbol{A}^k \in \mathcal{S}$, $\boldsymbol{A}^{k+1} \subseteq \boldsymbol{A}^k$ for every $k \in \mathbb{N}$ and $\mu_{m+n}(\boldsymbol{A}^0) < +\infty$. Then, if $\boldsymbol{A} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$, we obtain $\boldsymbol{A} \in \mathcal{S}$. In fact, let $\boldsymbol{x} \in \mathbb{R}^m$, we have $\boldsymbol{A}_{\boldsymbol{x}} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k_{\boldsymbol{x}} \in \mathcal{M}_n$. Let now $\boldsymbol{N} = \{\boldsymbol{x} \in \mathbb{R}^m \,|\, \mu_n(\boldsymbol{A}^0_{\boldsymbol{x}}) = +\infty\}$. From Theorem 3.46, the set $\boldsymbol{N}$ is measurable in $\mathbb{R}^m$ and $\mu_m(\boldsymbol{N}) = 0$. If $\boldsymbol{x} \in \mathbb{R}^m \setminus \boldsymbol{N}$, we obtain, using **(b)** in Theorem 3.11,

$$g(\boldsymbol{x}) = \lim_{k\to+\infty} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}}). \tag{3.60}$$

Thus, g is measurable over $\mathbb{R}^m \setminus \boldsymbol{N}$. It is measurable over N because $\mu_m(\boldsymbol{N}) = 0$ (see **(a)** in Theorem 3.48). Therefore, g is measurable. Taking (3.60) into account, noting that $\mu_n(\boldsymbol{A}^k_{\boldsymbol{x}}) \le \mu_n(\boldsymbol{A}^0_{\boldsymbol{x}})$ for every $k \in \mathbb{N}$ and Theorem 3.39, we conclude that, using **(b)** in Theorem 3.11,

$$\begin{aligned}
\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^m\setminus\boldsymbol{N}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \int_{\mathbb{R}^m\setminus\boldsymbol{N}} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \int_{\mathbb{R}^m} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \mu_{m+n}(\boldsymbol{A}^k) = \mu_{m+n}(\boldsymbol{A}).
\end{aligned}$$

(vii) Every null set $\boldsymbol{A}$ satisfies **(a), (b), (c)**. By definition, for every $k \in \mathbb{N}$ there exists $\boldsymbol{A}^k = \bigcup_{j=0}^{\infty} \boldsymbol{I}^{k,j}$, where $\boldsymbol{I}^{k,j}$ are $(m+n)$-dimensional intervals such that $\boldsymbol{A} \subseteq \boldsymbol{A}^k$ and $\sum_{j=0}^{\infty} \mathrm{vol}_{m+n}(\boldsymbol{I}^{k,j}) \le \frac{1}{k}$. The sets $\boldsymbol{A}^k$ are elements of $\mathcal{P}^\infty$. If needed, we can substitute $\boldsymbol{A}^k$ with $\boldsymbol{A}^0 \cap \ldots \cap \boldsymbol{A}^k$ and use Lemma 3.9, and then assume that $\boldsymbol{A}^{k+1} \subseteq \boldsymbol{A}^k$ for every k. Let $\boldsymbol{B} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$. Let $\boldsymbol{A} \subseteq \boldsymbol{B}$, $\mu_{m+n}(\boldsymbol{B}) = \lim_{k\to\infty} \mu_{m+n}(\boldsymbol{A}^k) = 0$. Applying **(vi)**, $\boldsymbol{B} \in \mathcal{S}$. From the second of Theorem 3.33, we can conclude that $\mu_n(\boldsymbol{B}_{\boldsymbol{x}}) = 0$ almost everywhere in $\mathbb{R}^m$. Then, since $\boldsymbol{A}_{\boldsymbol{x}} \subseteq \boldsymbol{B}_{\boldsymbol{x}}$ for every $\boldsymbol{x} \in \mathbb{R}^m$, applying Theorem 3.19 we obtain that for almost every $\boldsymbol{x} \in \mathbb{R}^m$, we have $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$ and $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = 0$. In other words, $g(\boldsymbol{x})$ is identically zero outside a certain null set N. From Corollary 3.3, we conclude that g is measurable and $\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$.

(viii) Every set $\boldsymbol{A}$ with finite measure satisfies **(a), (b), (c)**. For every $k \in \mathbb{N}$ let $\boldsymbol{A}^k = \bigcup_{j=0}^{\infty} \boldsymbol{I}^{k,j}$, where $\boldsymbol{I}^{k,j}$ are $(m+n)$-dimensional intervals such that $\boldsymbol{A} \subseteq \boldsymbol{A}^k$ and $\sum_{j=0}^{\infty} \mathrm{vol}_{m+n}(\boldsymbol{I}^{k,j}) \le \mu_{m+n}(\boldsymbol{A}) + \frac{1}{k}$. The sets $\boldsymbol{A}^k$ are

elements of $\mathcal{P}^\infty$, and we can assume that $\boldsymbol{A}^{k+1} \subseteq \boldsymbol{A}^k$ for every k. Let $\boldsymbol{B} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$. We have $\boldsymbol{A} \subseteq \boldsymbol{B}$, $\mu_{m+n}(\boldsymbol{B}) = \lim_{k\to\infty} \mu_{m+n}(\boldsymbol{A}^k) = \mu_{m+n}(\boldsymbol{A})$ and, applying **(vi)**, $\boldsymbol{B} \in \mathcal{S}$. Let $\boldsymbol{C} = \boldsymbol{B} \setminus \boldsymbol{A}$. Then, $\mu_{m+n}(\boldsymbol{C}) = 0$. Thus, by **(vii)**, the set $\boldsymbol{C}$ satisfies **(a), (b), (c)**. Then, $\boldsymbol{C}_{\boldsymbol{x}} \in \mathcal{M}_n$ for almost every $\boldsymbol{x}$. It follows that $\boldsymbol{A}_{\boldsymbol{x}} = \boldsymbol{B}_{\boldsymbol{x}} \setminus \boldsymbol{A}_{\boldsymbol{x}}$ belongs to $\mathcal{M}_n$ for almost every $\boldsymbol{x}$. Moreover, $\mu_n(\boldsymbol{C}_{\boldsymbol{x}}) = 0$ almost everywhere in $\mathbb{R}^n$. In this case, from $\boldsymbol{B}_{\boldsymbol{x}} = \boldsymbol{A}_{\boldsymbol{x}} \cup \boldsymbol{C}_{\boldsymbol{x}}$ it follows that $\mu_n(\boldsymbol{B}_{\boldsymbol{x}}) = \mu_n(\boldsymbol{A}_x)$. We conclude that $g(\boldsymbol{x}) = \mu_n(\boldsymbol{B}_{\boldsymbol{x}})$ almost everywhere over $\mathbb{R}^m$. Thus, applying **(a-b)** in Corollary 3.3, we conclude that g is measurable and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^m} \mu_n(\boldsymbol{B}_{\boldsymbol{x}})\,\mathrm{d}\boldsymbol{x} = \mu_{m+n}(\boldsymbol{B}) = \mu_{m+n}(\boldsymbol{A}).$$

(ix) This is the general case. Let $\boldsymbol{A} \in \mathcal{M}_{m+n}$. For every $k \in \mathbb{N}$, we define $\boldsymbol{A}^k = \boldsymbol{A} \cap [-k,k]^{m+n}$. Then, $\boldsymbol{A}^k$ is measurable and

$$\mu_{m+n}(\boldsymbol{A}^k) \leq \mu_{m+n}([-k,k]^{m+n}) = (2\,k)^{m+n}.$$

Then, for all $k \in \mathbb{N}$, the set $\boldsymbol{A}^k$ satisfies **(a), (b), (c)**. We have $\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k$, from which $\boldsymbol{A}_{\boldsymbol{x}} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k_{\boldsymbol{x}}$ for every $\boldsymbol{x} \in \mathbb{R}^m$. Let N_k the set of vectors $\boldsymbol{x} \in \mathbb{R}^m$ such that $\boldsymbol{A}^k_{\boldsymbol{x}}$ does not belong to $\mathcal{M}_n$. Then, $\boldsymbol{N}_k$ is measurable and $\mu_m(\boldsymbol{N}_k) = 0$. We define $\boldsymbol{N} = \bigcup_{k=0}^{\infty} \boldsymbol{N}_k$. Then, $\boldsymbol{N}$ is measurable, and from (d) in Theorem 3.10 we obtain $\mu_m(\boldsymbol{N}) \leq \sum_{k=0}^{\infty} \mu_m(\boldsymbol{N}_k) = 0$. Let $\boldsymbol{x} \in \mathbb{R}^m \setminus \boldsymbol{N}$. Then, for every $k \in \mathbb{N}$, we have $A^k_{\boldsymbol{x}} \in \mathcal{M}_n$. Thus, $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$ and, since $\boldsymbol{A}^k_{\boldsymbol{x}} \subseteq \boldsymbol{A}^{k+1}_{\boldsymbol{x}}$, we have $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = \lim_{k\to\infty} \mu_n(\boldsymbol{A}^k_{\boldsymbol{x}})$. We conclude that **(a)** holds. Let g_k be the function g that corresponds to $\boldsymbol{A}^k$. For every $k \in \mathbb{N}$, the function g_k is measurable and $g(\boldsymbol{x}) = \lim_{k\to\infty} g_k(\boldsymbol{x})$ for every $\boldsymbol{x} \in \mathbb{R}^m \setminus \boldsymbol{N}$. Thus, g is measurable over $\mathbb{R}^m \setminus \boldsymbol{N}$ and, since $\boldsymbol{N}$ has measure zero, g is measurable over $\boldsymbol{N}$. Hence, g is globally measurable. Moreover, from Theorem 3.33

$$\begin{aligned}
\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \int_{\mathbb{R}^m} g_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \lim_{k\to\infty} \mu_{m+n}(\boldsymbol{A}^k) = \mu_{m+n}(\boldsymbol{A}).
\end{aligned}$$

Obviously, an analogous result can be obtained by substituting the first m variables $(x_1, \ldots, x_m)$ with any $(x_{j_1}, \ldots, x_{j_m})$, where $1 \leq j_1 < \ldots < j_m \leq m+n$.

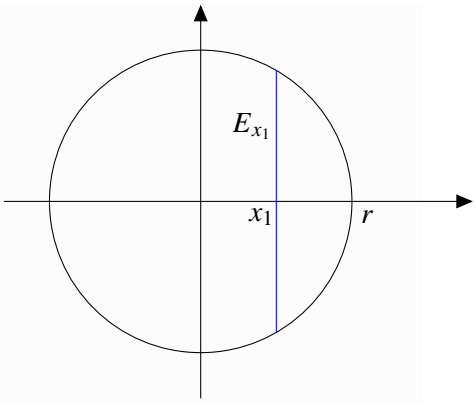
Example *3.41*

Let $r > 0$, and consider the open ball

$$\boldsymbol{E} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1^2 + x_2^2 < r^2\}.$$

Since $\boldsymbol{E}$ is open, it is measurable, and we want to find $\mu_2(\boldsymbol{E})$. Given $x_1 \in \mathbb{R}$, we have

$$E_{x_1} = \{x_2 \in \mathbb{R} \,|\, x_1^2 + x_2^2 < r^2\}$$
$$= \begin{cases} \left(-\sqrt{r^2 - x_1^2},\, \sqrt{r^2 - x_1^2}\right) & \text{if } x_1 \in (-r, r) \\ \varnothing & \text{otherwise} \end{cases}$$



Hence, E_{x_1} is measurable for every $x_1 \in \mathbb{R}$, and

$$\mu_1(E_{x_1}) = \begin{cases} 2\sqrt{r^2 - x_1^2} & \text{if } x_1 \in (-r, r) \\ 0 & \text{otherwise} \end{cases}$$

Therefore,

$$\mu_2(\boldsymbol{E}) = \int_{-r}^{r} 2\sqrt{r^2 - x_1^2}\, \mathrm{d}x_1.$$

Using the change of variable $x_1 = r \sin t$ we find

$$\begin{aligned} \mu_2(E) &= 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{r^2 - r^2 \sin^2 t}\, \cos t \, \mathrm{d}t \\ &= 2r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t \, \mathrm{d}t \\ &= r^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1 + \cos(2t)) \, \mathrm{d}t \\ &= r^2 \left[t + \frac{1}{2}\sin(2t)\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\ &= r^2 \left(\frac{\pi}{2} + \frac{1}{2}\sin\pi - \left[-\frac{\pi}{2} + \frac{1}{2}\sin(-\pi)\right]\right) = \pi r^2. \end{aligned}$$

Obviously, there is nothing special about fixing the first coordinate x_1 and integrate the measure of E_{x_1} when x_1 varies in $\mathbb{R}$: we could have fixed x_2 and integrate the measure of E_{x_2} when x_2 varies in $\mathbb{R}$, and we would have obtained the same result.

Example *3.42*

Let $r > 0$, and consider the open ball

$$\boldsymbol{E} = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \,|\, x_1^2 + x_2^2 + x_3^2 < r^2\}.$$

Since $\boldsymbol{E}$ is open, it is measurable, and we want to find $\mu_3(\boldsymbol{E})$. Given $x_3 \in \mathbb{R}$, we have

$$\boldsymbol{E}_{x_3} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1^2 + x_2^2 + x_3^2 < r^2\} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1^2 + x_2^2 < r^2 - x_3^2\}.$$

If $|x_3| \geq r$, we have $\boldsymbol{E}_{x_3} = \varnothing$. If $|x_3| < r$, then $\boldsymbol{E}_{x_3}$ is the open ball of $\mathbb{R}^2$ centred at $(0, 0)$ with radius $\sqrt{r^2 - x_3^2}$.

From the previous example, we find that for every $\boldsymbol{x}_3 \in \mathbb{R}$

$$\mu_2(\boldsymbol{E}_{x_3}) = \begin{cases} \pi\,(r^2 - x_3^2) & \text{if } x_3 \in (-r, r) \\ 0 & \text{otherwise} \end{cases}$$

Thus, using the change of variable $x_3 = r\,\sin t$,

$$\begin{aligned}
\mu_3(\boldsymbol{E}) &= \int_{-r}^{r} \pi\,(r^2 - x_3^2)\,\mathrm{d}x_3 \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \pi\,(r^2 - r^2\,\sin^2 t)\,r\,\cos t\,\mathrm{d}t \\
&= \pi\,r^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^3 t\,\mathrm{d}t = \pi\,r^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t\,\cos^2 t\,\mathrm{d}t = \pi\,r^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t\,(1 - \sin^2 t)\,\mathrm{d}t \\
&= \pi\,r^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t\,\mathrm{d}t - \pi\,r^3 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t\,\sin^2 t\,\mathrm{d}t \\
&= \pi\,r^3 \left[\sin t - \frac{1}{3}\,\sin^3 x\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \pi\,r^3 \left(1 - \frac{1}{3} + 1 - \frac{1}{3}\right) = \pi\,r^3\,\frac{4}{3} = \frac{4}{3}\,\pi\,r^3.
\end{aligned}$$

Example *3.43*

For every $n \in \mathbb{N}$ and $r > 0$, we find

$$\mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| < r\}).$$

We show also, by induction, that the result is equal to $\omega_n\,r^n$, where $\omega_n = \mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| \leq 1\})$.

If $n = 1$, we have $\mu_1((-r, r)) = 2\,r$, so that $\omega_1 = 2$. We denote by $\|\cdot\|_n$ the Euclidean norm in $\mathbb{R}^n$. Consider the case $n + 1$, assuming that, for n, we have $\mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| < r\}) = \omega_n\,r^n$. We denote by $(x, \boldsymbol{y})$ the generic element of $\mathbb{R}^{n+1}$, where $x \in \mathbb{R}$ and $\boldsymbol{y} \in \mathbb{R}^n$. Let $\boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^{n+1} \mid \|\boldsymbol{x}\|_{n+1} < r\}$. For all $x \in \mathbb{R}$, we have

$$\boldsymbol{A}_x = \begin{cases} \{\boldsymbol{y} \in \mathbb{R}^n \mid \|\boldsymbol{y}\|^2 < r^2 - x^2\} & \text{if } |x| < r \\ \varnothing & \text{otherwise} \end{cases}$$

Using the inductive assumption,

$$\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = \begin{cases} \omega_n(r^2 - x^2)^{\frac{n}{2}} & \text{if } |x| < r \\ 0 & \text{otherwise} \end{cases}$$

Observing that $\mu_1([-r, r] \setminus (-r, r)) = 0$, we obtain

$$\begin{aligned}
\mu_{n+1}(\boldsymbol{A}) &= \int_{\mathbb{R}} \mu_n(\boldsymbol{A}_{\boldsymbol{x}})\,\mathrm{d}x \\
&= \int_{(-r,r)} \omega_n\,(r^2 - x^2)^{\frac{n}{2}}\,\mathrm{d}x \\
&= \int_{[-r,r]} \omega_n\,(r^2 - x^2)^{\frac{n}{2}}\,\mathrm{d}x \\
&= \int_{-r}^{r} \omega_n\,(r^2 - x^2)^{\frac{n}{2}}\,\mathrm{d}x \\
&= \int_{-1}^{1} \omega_n\,r^n\,(1 - \xi^2)^{\frac{n}{2}}\,r\,\mathrm{d}\xi,
\end{aligned}$$

by Theorem 3.51. Thus,

$$\mu_{n+1}(\boldsymbol{A}) = \omega_n\, r^{n+1} \int_{-1}^{1} (1-x^2)^{\frac{n}{2}}\, \mathrm{d}x.$$

We obtain $\omega_{n+1} = \omega_n \displaystyle\int_{-1}^{1} (1-x^2)^{\frac{n}{2}}\, \mathrm{d}x$ and $\mu_{n+1}(\boldsymbol{A}) = \omega_{n+1}\, r^{n+1}$. We now need to compute ω_n. We have

$$\int_{-1}^{1} (1-x^2)^{\frac{n}{2}}\, \mathrm{d}x = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1-\sin^2 t)^{\frac{n}{2}}\, \cos t\, \mathrm{d}t = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n+1} t\, \mathrm{d}t.$$

Let us define

$$I_n = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n t\, \mathrm{d}t$$

for $n \in \mathbb{N}$. We have $I_0 = \pi$ and $I_1 = 2$. If $n \geq 2$,

$$\begin{aligned} I_n &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-1} t\, \cos t\, \mathrm{d}t \\ &= [\cos^{n-1} t\, \sin t]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + (n-1) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2} t\, \sin^2 t\, \mathrm{d}t = 0 + (n-1) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2} t\, (1-\cos^2 t)\, \mathrm{d}t = (n-1)(I_{n-2} - I_n). \end{aligned}$$

It follows that $I_n = \frac{n-1}{n} I_{n-2}$. Hence, if n is even

$$I_n = \frac{n-1}{n} I_{n-2} = \frac{(n-1)(n-3)}{n\,(n-2)} I_{n-4} = \ldots = \frac{(n-1)!!}{n!!} I_0 = \pi \frac{(n-1)!!}{n!!}.$$

If n is odd

$$I_n = \frac{n-1}{n} I_{n-2} = \frac{(n-1)(n-3)}{n\,(n-2)} I_{n-4} = \ldots = \frac{(n-1)!!}{n!!} I_1 = 2 \frac{(n-1)!!}{n!!}.$$

We conclude that, for $n \geq 2$,

$$\omega_n = \omega_{n-1}\, I_n = \ldots = \omega_1\, I_n\, I_{n-1} \cdots I_2,$$

so that

$$\omega_n = \begin{cases} \frac{(2\pi)^{\frac{n}{2}}}{n!!} & \text{if } n \text{ is even} \\ \frac{2\,(2\pi)^{\frac{n-1}{2}}}{n!!} & \text{if } n \text{ is odd} \end{cases}$$

For example, $\omega_2 = \pi$ and $\omega_3 = \frac{4}{3}\pi$.

We are now ready to address the reduction for integrals. We will denote by

$$\int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y}$$

the integral of f over $\mathbb{R}^{m+n}$.

Theorem 3.54 [Tonelli's Theorem]
Let $m, n \in \mathbb{N}^*$. Let $f : \mathbb{R}^{m+n} \longrightarrow [0, +\infty]$ be measurable. Then:

(a) *for almost every $\boldsymbol{x} \in \mathbb{R}^m$, the mapping*

$$\begin{aligned} f(\boldsymbol{x}, \cdot) : \mathbb{R}^n &\longrightarrow [0, +\infty] \\ \boldsymbol{y} &\mapsto f(\boldsymbol{x}, \boldsymbol{y}) \end{aligned} \tag{3.61}$$

is measurable on $\mathbb{R}^n$;

(b) *defining, for* $\boldsymbol{x} \in \mathbb{R}^m$,

$$g(\boldsymbol{x}) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{y} & \text{if } f(\boldsymbol{x}, \cdot) \text{ is measurable} \\ 0 & \text{otherwise} \end{cases}$$

Then, g *is measurable over* $\mathbb{R}^m$*;*

(c) *There holds*

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y}.$$

Proof We start by assuming that $f = \chi_A$, where $\boldsymbol{A} \in \mathcal{M}_{m+n}$. From the first of Theorem 3.53, for almost every $\boldsymbol{x} \in \mathbb{R}^m$, the set $\boldsymbol{A_x}$ is measurable in $\mathbb{R}^n$. Then, since $\chi_A(\boldsymbol{x}, \cdot) = \chi_{A_x}$, then **(a)**. Moreover, function g defined in **(b)** coincides with the function g in Theorem 3.53. It follows that g is measurable and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \mu_{m+n}(A) = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y}.$$

Consider now the case where f is simple and non-negative, so that $f = \displaystyle\sum_{k=1}^{p} \alpha_k \chi_{A^k}$, where $\alpha_k \in [0, +\infty)$ for $k = 1, \ldots, p$ and the sets $\boldsymbol{A}^k$ belonging to $\mathcal{M}_{m+n}$ are pairwise disjoint. We define

$$\boldsymbol{N}_k = \{\boldsymbol{x} \in \mathbb{R}^m \,|\, \boldsymbol{A}^k_{\boldsymbol{x}} \text{ is not measurable}\}.$$

We denote by g_k the function that corresponds to g for χ_{A_k}. Then, for every $k \in \mathbb{N}$, the set $\boldsymbol{N}_k$ is measurable and $\mu_m(\boldsymbol{N}_k) = 0$. It follows that $\mu_m\Big(\displaystyle\bigcup_{k=1}^{p} \boldsymbol{N}_k\Big) = 0$, and if $\boldsymbol{x} \notin \displaystyle\bigcup_{k=1}^{p} \boldsymbol{N}_k$, the function $f(\boldsymbol{x}, \cdot)$ is measurable. Hence, from Theorem 3.35, we have $g(\boldsymbol{x}) = \displaystyle\sum_{k=1}^{p} \alpha_k\, g_k(\boldsymbol{x})$. Then, from Corollary 3.3, g is measurable and

$$\begin{aligned} \int_{\mathbb{R}^m} g(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^m} \sum_{k=1}^{p} \alpha_k\, g_k(\boldsymbol{x}) \\ &= \sum_{k=1}^{p} \alpha_k \int_{\mathbb{R}^m} g_k(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\ &= \sum_{k=1}^{p} \alpha_k \int_{\mathbb{R}^{m+n}} \chi_{A^k}(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y}. \end{aligned}$$

We now consider the general case. From Theorem 3.29, there exists a sequence of simple functions $(\varphi_k)_{k \in \mathbb{N}}$ over $\mathbb{R}^{m+n}$ such that, for every k, we have $0 \leq \varphi_k \leq \varphi_{k+1} \leq f$, and for every $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{m+n}$ we have $f(\boldsymbol{x}, \boldsymbol{y}) = \lim\limits_{k \to \infty} \varphi(\boldsymbol{x}, \boldsymbol{y})$. We define

$$\boldsymbol{M}_k = \{\boldsymbol{x} \in \mathbb{R}^m \,|\, \varphi_k(\boldsymbol{x}, \cdot) \text{ is not measurable}\}.$$

Then, for all k, we have that M_k is measurable and has zero measure. Moreover,

$$\mu_n\Big(\bigcup_{k=0}^{\infty} \boldsymbol{M}_k\Big) \leq \sum_{k=0}^{\infty} \mu_m(\boldsymbol{M}_k) = 0.$$

If $\boldsymbol{x} \notin \bigcup_{k=0}^{\infty} \boldsymbol{M}_k$, we deduce that $f(\boldsymbol{x}, \cdot)$ is measurable, and this proves **(a)**. Moreover, from the monotone convergence theorem (Theorem 3.34), if $\boldsymbol{x} \notin \bigcup_{k=0}^{\infty} \boldsymbol{M}_k$, we have

$$g(\boldsymbol{x}) = \lim_{k \to \infty} g_k(\boldsymbol{x}),$$

where g_k corresponds to g for φ_k. We define, for $k \in \mathbb{N}$,

$$h_k(\boldsymbol{x}) = \begin{cases} g_k(\boldsymbol{x}) & \text{if } \boldsymbol{x} \notin \bigcup_{j=0}^{\infty} \boldsymbol{M}_j \\ 0 & \text{otherwise} \end{cases} \tag{3.62}$$

and

$$h(\boldsymbol{x}) = \begin{cases} g_(\boldsymbol{x}) & \text{if } \boldsymbol{x} \notin \bigcup_{j=0}^{\infty} \boldsymbol{M}_j \\ 0 & \text{otherwise.} \end{cases} \tag{3.63}$$

Then, $h_k(\boldsymbol{x}) = g_k(\boldsymbol{x})$ almost everywhere. Hence, from Corollary 3.3, h_k is measurable and

$$\int_{\mathbb{R}^m} h_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^m} g_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} \varphi_k(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}.$$

Moreover, from the monotone convergence theorem (Theorem 3.34), h is measurable and

$$\begin{aligned} \int_{\mathbb{R}^{m+n}} h(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} &= \lim_{k \to \infty} \int_{\mathbb{R}^{m+n}} h_k(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \\ &= \lim_{k \to \infty} \int_{\mathbb{R}^{m+n}} \varphi_k(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}. \end{aligned}$$

We conclude by observing that $g(\boldsymbol{x}) = h(\boldsymbol{x})$ almost everywhere, and the statement follows from Corollary 3.3.

The statement of Tonelli's theorem required the domain of integration to be the entire space. The following corollary relaxes this assumption.

Corollary 3.4 *Let $m, n \in \mathbb{N}^*$. Let $\boldsymbol{A} \in \mathcal{M}_{m+n}$, let $f : \boldsymbol{A} \longrightarrow [0, +\infty]$ be measurable and let $\boldsymbol{B} \in \mathcal{M}_m$ be such that*

$$\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n, \ \mu_n(\boldsymbol{A}_{\boldsymbol{x}}) > 0, \ f(\boldsymbol{x}, \cdot) \text{ is measurable on } \boldsymbol{A}_{\boldsymbol{x}}\} \subseteq \boldsymbol{B}.$$

Let $g : \boldsymbol{B} \longrightarrow [0, +\infty]$ be defined by

$$g(\boldsymbol{x}) = \begin{cases} \int_{\boldsymbol{A}_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} & \text{if } \boldsymbol{A}_{\boldsymbol{x}} \text{ is measurable and } f(\boldsymbol{x}, \cdot) \text{ is measurable over } \boldsymbol{A}_{\boldsymbol{x}} \\ 0 & \text{otherwise} \end{cases}$$

Then, g is measurable and

$$\int_{\boldsymbol{B}} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}.$$

Proof Let

$$\begin{aligned} \tilde{f} : \mathbb{R}^{m+n} &\longrightarrow [0, +\infty] \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto \begin{cases} f(\boldsymbol{x}, \boldsymbol{y}) & \text{if } (\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{A} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

From Remark 3.44, $\tilde{f}$ is measurable and

$$\int_{\mathbb{R}^{m+n}} \tilde{f}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{A}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}.$$

Let

$$\begin{aligned} \tilde{g} : \mathbb{R}^{m} &\longrightarrow [0, +\infty] \\ \boldsymbol{x} &\mapsto \begin{cases} \displaystyle\int_{\mathbb{R}^n} \tilde{f}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} & \text{if } \tilde{f}(\boldsymbol{x}, \cdot) \text{ is measurable} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

From Theorem 3.54, $\tilde{g}$ is measurable and

$$\int_{\mathbb{R}^m} \tilde{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} \tilde{f}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{A}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}.$$

We now check that $\tilde{g}(\boldsymbol{x}) = 0$ almost everywhere in $\mathbb{R}^m \setminus \boldsymbol{B}$. If $\boldsymbol{x} \notin \boldsymbol{B}$, three cases are possible:

(i) $\boldsymbol{A}_{\boldsymbol{x}} \notin \mathcal{M}_n$

(ii) $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$ and $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = 0$;

(iii) $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$, $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) > 0$ and $f(\boldsymbol{x}, \cdot)$ is not measurable on $\boldsymbol{A}_{\boldsymbol{x}}$.

From Theorem 3.53, the set $\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}_{\boldsymbol{x}} \notin \mathcal{M}_n\}$ is measurable and it has zero measure. If $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$ and $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) = 0$, two cases are possible: $\tilde{f}(\boldsymbol{x}, \cdot)$ is measurable or not. In the second case, $\tilde{g}(\boldsymbol{x}) = 0$ by definition. In the first case, $f(\boldsymbol{x}, \cdot)$ is measurable over $\boldsymbol{A}_{\boldsymbol{x}}$ and we have

$$\tilde{g}(\boldsymbol{x}) = \int_{\mathbb{R}^n} \tilde{f}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{A}_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = 0$$

from Remark 3.44 and the second point of Theorem 3.48. Finally, let $\boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n$, $\mu_n(\boldsymbol{A}_{\boldsymbol{x}}) > 0$ and $f(\boldsymbol{x}, \cdot)$ is not measurable on $\boldsymbol{A}_{\boldsymbol{x}}$. Then, from Theorem 3.43, $\tilde{f}(\boldsymbol{x}, \cdot)$ is not measurable. It follows that $\boldsymbol{x} \in \{\boldsymbol{x} \in \mathbb{R}^m \mid \tilde{f}(\boldsymbol{x}, \cdot) \text{ is not measurable}\}$, which is a zero measure set from Theorem 3.54. It follows that

$$\int_{\mathbb{R}^m} \tilde{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{B}} \tilde{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

Finally, let $\boldsymbol{x} \in B$. Let $\boldsymbol{A}_{\boldsymbol{x}}$ be measurable and let $f(\boldsymbol{x}, \cdot)$ be measurable over $\boldsymbol{A}_{\boldsymbol{x}}$. It follows that $\tilde{f}(\boldsymbol{x}, \cdot)$ is measurable and

$$\tilde{g}(\boldsymbol{x}) = \int_{\mathbb{R}^n} \tilde{f}(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{A}_{\boldsymbol{x}}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = g(\boldsymbol{x}).$$

Moreover, the intersection $\boldsymbol{B} \cap \{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}_{\boldsymbol{x}} \text{ is not measurable}\}$ has measure zero. Finally, $\{\boldsymbol{x} \in \boldsymbol{B} \mid \boldsymbol{A}_{\boldsymbol{x}} \text{ is measurable and } f(\boldsymbol{x}, \cdot) \text{ is}$ is contained in $\{\boldsymbol{x} \in \mathbb{R}^m \mid \tilde{f}(\boldsymbol{x}, \cdot) \text{ is not measurable}\}$, which is measurable and has measure zero. We conclude that $g(\boldsymbol{x}) = \tilde{g}(\boldsymbol{x})$ almost everywhere on B. It follows that g is measurable and

$$\int_{\boldsymbol{B}} g(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{b} \tilde{g}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y}.$$

Fubini's Theorem is essentially the same result for summable functions.

Theorem 3.55 [Fubini's Theorem]
Let $f \in \mathcal{L}(\mathbb{R}^{m+n};\mathbb{C})$, where $m,n \in \mathbb{N}^*$. Then:

(a) *$f(\boldsymbol{x},\cdot) \in \mathcal{L}(\mathbb{R}^n;\mathbb{C})$ for almost every $\boldsymbol{x} \in \mathbb{R}^m$;*

(b) *defining, for $\boldsymbol{x} \in \mathbb{R}^m$,*

$$g(\boldsymbol{x}) = \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} & \text{if } f(\boldsymbol{x},\cdot) \text{ is summable} \\ 0 & \text{otherwise} \end{cases}$$

Then, g is summable over $\mathbb{R}^m$;

(c) *There holds*

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}.$$

Proof We start considering the case where the codomain of f is $\mathbb{R}$. The corresponding functions f^+ and f^- are measurable and have values in $[0,+\infty)$. Let g^+ and g^- be the functions that correspond to g in Theorem 3.54 for f^+ and f^-. We have

$$g^+(\boldsymbol{x}) = \int_{\mathbb{R}^n} f^+(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \quad \text{and} \quad g^-(\boldsymbol{x}) = \int_{\mathbb{R}^n} f^-(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \ \text{ almost everywhere on } \mathbb{R}^m.$$

Since f is summable, from Theorem 3.46, we have $\max\{g^+(\boldsymbol{x}), g^-(\boldsymbol{x})\} < +\infty$ almost everywhere over $\mathbb{R}^m$. Hence,

$$\begin{aligned} \boldsymbol{N} = {} & \{\boldsymbol{x} \in \mathbb{R}^m \mid f^+(\boldsymbol{x},\cdot) \text{ is not measurable}\} \cup \{\boldsymbol{x} \in \mathbb{R}^m \mid f^-(\boldsymbol{x},\cdot) \text{ is not measurable}\} \\ & \cup \{\boldsymbol{x} \in \mathbb{R}^m \mid \max\{g^+(\boldsymbol{x}), g^-(\boldsymbol{x})\} = +\infty\}. \end{aligned}$$

Then, $\boldsymbol{N} \in \mathcal{M}_m$ and $\mu_m(\boldsymbol{N}) = 0$. If $\boldsymbol{x} \notin \boldsymbol{N}$, we have $f(\boldsymbol{x},\cdot) \in \mathcal{L}(\mathbb{R}^n)$ and $g(\boldsymbol{x}) = g^+(\boldsymbol{x}) - g^-(\boldsymbol{x})$. We therefore obtain **(a)**. Moreover, from $g = g^+ - g^-$, it follows that g is summable over $\mathbb{R}^m \setminus \boldsymbol{N}$. Since $\mu_m(\boldsymbol{N}) = 0$, applying Theorem 3.54,

$$\begin{aligned} \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{R}^m} g^+(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\mathbb{R}^m} g^-(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{R}^{m+n}} f^+(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} - \int_{\mathbb{R}^{m+n}} f^-(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} \\ &= \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}. \end{aligned}$$

Finally, N has measure zero. We conclude that, by applying Theorem 3.44, $g \in \mathcal{L}(\mathbb{R}^m)$ and

$$\int_{\mathbb{R}^m} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^m \setminus \boldsymbol{N}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{N}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^{m+n}} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}.$$

The case where the codomain is $\mathbb{C}$ can be dealt with by writing $f = \mathfrak{Re}\{f\} + i\,\mathfrak{Im}\{f\}$.

Corollary 3.5 *Let $m,n \in \mathbb{N}^*$. Let $\boldsymbol{A} \in \mathcal{M}_{m+n}$, let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ be summable and let $\boldsymbol{B} \in \mathcal{M}_m$ be such that*

$$\{\boldsymbol{x} \in \mathbb{R}^m \mid \boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n,\ \mu_n(\boldsymbol{A}_{\boldsymbol{x}}) > 0,\ f(\boldsymbol{x},\cdot) \text{ is measurable on } \boldsymbol{A}_{\boldsymbol{x}}\} \subseteq \boldsymbol{B},$$

where the function $f(\boldsymbol{x},\cdot)$ is as defined in (3.61). Let

$$g : \boldsymbol{B} \longrightarrow \mathbb{C}$$
$$\boldsymbol{x} \mapsto \begin{cases} \displaystyle\int_{\boldsymbol{A}_{\boldsymbol{x}}} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} & \text{if } \boldsymbol{A}_{\boldsymbol{x}} \in \mathcal{M}_n \text{ and } f(\boldsymbol{x},\cdot) \text{ is summable over } \boldsymbol{A}_{\boldsymbol{x}} \\ 0 & \text{otherwise} \end{cases}$$

Then, g is summable over $\boldsymbol{B}$ and

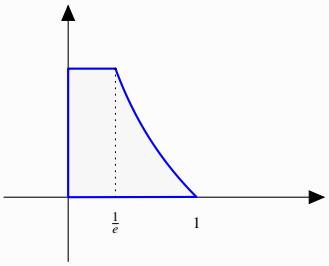
$$\int_{\boldsymbol{B}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}.$$

The proof follows the same lines of Corollary 3.4.

Example *3.44*

Consider the set

$$\boldsymbol{A} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, 0 < x \le 1,\ 0 < y \le \min\left\{1, \ln\tfrac{1}{x}\right\}\right\}$$



and the function

$$f : \boldsymbol{A} \longrightarrow \mathbb{R}$$
$$(x,y) \mapsto xy.$$

We verify that f is summable over $\boldsymbol{A}$ and we find $\displaystyle\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$. Let $\boldsymbol{A} = \boldsymbol{A}_1 \setminus \boldsymbol{A}_2$, where

$$\boldsymbol{A}_1 = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, 0 \le x \le 1,\ 0 \le y \le \min\left\{1, \ln\frac{1}{x}\right\}\right\}$$

obtained by continuation of the function $x \mapsto \min\{1, \ln\frac{1}{x}\}$ to 0 and

$$\boldsymbol{A}_2 = \{(x,y) \in \mathbb{R}^2 \,|\, x = 0\} \cup \{(x,y) \in \mathbb{R}^2 \,|\, y = 0\}.$$

The sets $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are measurable because they are closed. Thus, $\boldsymbol{A}$ is measurable. The function f is measurable over $\boldsymbol{A}$ because it is continuous (see Theorem 3.32). We verify that f is summable over $\boldsymbol{A}$. First, $\boldsymbol{A} \subseteq [0,1] \times [0,1]$, while $|f(x,y)| \le 1$ if $(x,y) \in \boldsymbol{A}$. It follows that

$$\int_{\boldsymbol{A}} |f(x,y)|\,\mathrm{d}x\,\mathrm{d}y \le 1 \cdot \mu_2([0,1] \times [0,1]) = 2.$$

We now compute the integral. Given $x \in \mathbb{R}$, we have

$$A_x = \begin{cases} (0,1] & \text{if } x \in \left(0, \frac{1}{e}\right] \\ \left(0, \ln\frac{1}{x}\right] & \text{if } x \in \left(\frac{1}{e}, 1\right] \\ \varnothing & \text{if } x \notin (0,1] \end{cases}$$

We define $B = (0,1]$. We compute $\int_{B_x} f(x,y)\,\mathrm{d}y$ for every $x \in B$. If $x \in \left(0, \frac{1}{e}\right]$ we have

$$\int_{B_x} f(x,y)\,\mathrm{d}y = \int_{(0,1]} x\,y\,\mathrm{d}y.$$

Since $\mu_1(\{0\}) = 0$, we have

$$\int_{(0,1]} x\,y\,\mathrm{d}y = \int_{[0,1]} x\,y\,\mathrm{d}y = \int_0^1 x\,y\,\mathrm{d}y = \frac{x}{2},$$

applying Theorem 3.51. In the same way, if $x \in \left(\frac{1}{e}, 1\right]$,

$$\int_{B_x} f(x,y)\,\mathrm{d}y = \int_{(0,\ln\frac{1}{x}]} x\,y\,\mathrm{d}y = \int_0^{\ln\frac{1}{x}} x\,y\,\mathrm{d}y = \frac{x\,\ln^2 x}{2},$$

where we used the fact that $\ln^2\frac{1}{x} = \left(\ln\frac{1}{x}\right)^2 = (-\ln x)^2 = \ln^2 x$. We define, for $x \in B$,

$$g(x) = \begin{cases} \frac{1}{2}\,x & \text{if } x \in \left(0, \frac{1}{e}\right] \\ \frac{x\,\ln^2 x}{2} & \text{if } x \in \left(\frac{1}{e}, 1\right]. \end{cases}$$

From Corollary 3.5 we find

$$\begin{aligned} \int_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{(0,1]} g(x)\,\mathrm{d}x = \int_{(0,\frac{1}{e}]} \frac{x}{2}\,\mathrm{d}x + \int_{(\frac{1}{e},1]} \frac{x\,\ln^2 x}{2}\,\mathrm{d}x \\ &= \int_0^{\frac{1}{e}} \frac{x}{2}\,\mathrm{d}x + \int_{\frac{1}{e}}^1 \frac{x\,\ln^2 x}{2}\,\mathrm{d}x \\ &= \frac{1}{4\,e^2} + \left(\frac{1}{8} - \frac{5}{8\,e^2}\right) = \frac{1}{8} - \frac{3}{8\,e^2}. \end{aligned}$$

Example *3.45*

Let

$$\boldsymbol{A} = \left\{(x,y) \in \mathbb{R}^2 \,\middle|\, x \geq 1,\ y \geq x^2\right\}$$

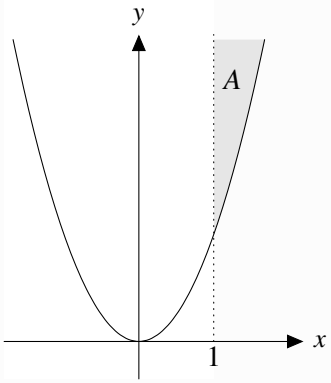
and

$$\begin{aligned} f : \boldsymbol{A} &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto \frac{1}{x^4 + y^2}. \end{aligned}$$

$$\int_{\boldsymbol{A}} \frac{1}{x^4+y^2}\,\mathrm{d}x\,\mathrm{d}y$$

where

$$\boldsymbol{A} = \{(x,y) \in \mathbb{R}^2 \,|\, x \geq 1 \ \text{ and } \ y \geq x^2\}$$



The set $\boldsymbol{A}$ is measurable because it is closed, and f is measurable over $\boldsymbol{A}$ because it is continuous. Moreover, $f(x,y) \geq 0$ for all $(x,y) \in \boldsymbol{A}$, so that we can apply Tonelli's theorem. We find $\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$. For every $x \in \mathbb{R}$

$$A_x = \begin{cases} [x^2, +\infty) & \text{if } x \geq 1 \\ \varnothing & \text{otherwise.} \end{cases}$$

For every $x \geq 1$, applying Lemma 3.7

$$\begin{aligned}
\int_{A_x} f(x,y)\,\mathrm{d}y &= \int_{[x^2,+\infty)} \frac{1}{x^4+y^2}\,\mathrm{d}y \\
&= \lim_{c\to+\infty} \int_{x^2}^{c} \frac{1}{x^4+y^2}\,\mathrm{d}y \\
&= \lim_{c\to+\infty} \frac{1}{x^2} \int_{x^2}^{c} \frac{1}{1+\frac{y^2}{x^4}} \frac{1}{x^2}\,\mathrm{d}y \\
&= \lim_{c\to+\infty} \frac{1}{x^2} \left[\arctan \frac{y}{x^2}\right]_{y=x^2}^{y=c} \\
&= \lim_{c\to+\infty} \frac{1}{x^2} \left(\arctan \frac{c}{x^2} - \arctan 1\right) \\
&= \lim_{c\to+\infty} \frac{1}{x^2} \left(\arctan \frac{c}{x^2} - \frac{\pi}{2}\right) = \frac{1}{x^2}\left(\frac{\pi}{2} - \frac{\pi}{4}\right) = \frac{\pi}{4\,x^2}.
\end{aligned}$$

It follows that

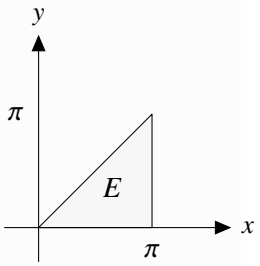
$$\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{[1,+\infty)} \frac{\pi}{4\,x^2}\,\mathrm{d}x = \lim_{c\to+\infty} \int_1^c \frac{\pi}{4\,x^2}\,\mathrm{d}x = \lim_{c\to+\infty} \frac{\pi}{4}\left(1 - \frac{1}{c}\right) = \frac{\pi}{4}.$$

Example 3.46 We want to find

$$\int_{\boldsymbol{E}} x_1 \cos(x_1 + x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2$$

where

$$\boldsymbol{E} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, 0 \le x_1 \le \pi \text{ and } 0 \le x_2 \le x_1\}$$



The integrand is the function

$$\begin{aligned} f : \boldsymbol{E} &\longrightarrow \mathbb{R} \\ (x_1, x_2) &\mapsto x_1 \cos(x_1 + x_2). \end{aligned}$$

It is not a non-negative function, and therefore we cannot apply Tonelli's Theorem. We therefore try to apply Fubini's Theorem. The set $\boldsymbol{E}$ is measurable because it is closed. The function f is measurable because it is continuous. Moreover, $\boldsymbol{E} \subseteq [0, \pi] \times [0, \pi]$, so that $\mu(\boldsymbol{E}) \le \pi^2$. Since $\boldsymbol{E}$ is closed and bounded, from Weierstrass extreme value theorem, f is bounded, and therefore from Remark 3.16 we find that f is summable in $\boldsymbol{E}$. We can therefore apply Fubini's theorem. Let, for all $x_1 \in \mathbb{R}$

$$E_{x_1} = \begin{cases} [0, x_1] & \text{if } 0 \le x_1 \le \pi \\ \varnothing & \text{otherwise} \end{cases}$$

Let $B = [0, \pi]$. By virtue of Theorem 3.51, for all $x_1 \in B$ the function $f(x_1, \cdot)$ is summable over E_{x_1} and

$$\begin{aligned} g(x_1) &= \int_{E_{x_1}} f(x_1, x_2)\, \mathrm{d}x_2 = \int_0^{x_1} x_1 \cos(x_1 + x_2)\, \mathrm{d}x_2 \\ &= x_1 \int_0^{x_1} \cos(x_1 + x_2)\, \mathrm{d}x_2 = x_1 \left[\sin(x_1 + x_2)\right]_{x_2 = 0}^{x_2 = x_1} = x_1 \left(\sin(2\, x_1) - \sin x_1\right). \end{aligned}$$

It follows that

$$\int_{\boldsymbol{E}} f(x_1, x_2)\, \mathrm{d}x_1\, \mathrm{d}x_2 = \int_0^{\pi} x_1 \left(\sin(2\, x_1) - \sin x_1\right) \mathrm{d}x_1.$$

We can use the integration by parts to find a primitive of $x_1 \left(\sin(2\, x_1) - \sin x_1\right)$:

$$\int x_1 \sin x_1\, \mathrm{d}x_1 = -x_1 \cos x_1 - \int 1 \cdot (-\cos x_1)\, \mathrm{d}x_1 = -x_1 \cos x_1 + \int \cos x_1 = -x_1 \cos x_1 + \sin x_1$$

and

$$\begin{aligned} \int x_1 \sin(2\, x_1)\, \mathrm{d}x_1 &= -\frac{1}{2} x_1 \cos x_1 - \int -\frac{1}{2} \cos(2\, x_1)\, \mathrm{d}x_1 \\ &= -\frac{1}{2} x_1 \cos x_1 + \frac{1}{2} \int \cos(2\, x_1)\, \mathrm{d}x_1 = -\frac{1}{2} x_1 \cos x_1 + \frac{1}{4} \sin(2\, x). \end{aligned}$$

It follows that

$$\begin{aligned} \int x_1 \left(\sin(2\, x_1) - \sin x_1\right) \mathrm{d}x_1 &= -\frac{1}{2} x_1 \cos x_1 + \frac{1}{4} \sin(2\, x) + x_1 \cos x_1 - \sin x_1 \\ &= -\frac{1}{2} x_1 \cos x_1 + \frac{1}{2} \sin x_1 \cos x_1 + x_1 \cos x_1 - \sin x_1 \\ &= -\frac{1}{2} x_1 \cos x_1 + x_1 \cos x_1 + \sin x \left(\frac{1}{2} \cos x - 1\right). \end{aligned}$$

Using this expression we find

$$\begin{aligned}\int_{\boldsymbol{E}} f(x_1,x_2)\,\mathrm{d}x_1\,\mathrm{d}x_2 &= \left[-\frac{1}{2}\,x_1\,\cos x_1 + x_1\,\cos x_1 + \sin x\left(\frac{1}{2}\,\cos x - 1\right)\right]_{x_1=0}^{x_1=\pi}\\ &= -\frac{1}{2}\,\pi\cdot 1 + \pi\cdot(-1) + 0 = -\frac{3}{2}\,\pi.\end{aligned}$$

A general strategy could be to verify the summability of f by first applying Tonelli's theorem to $|f|$ (which is non-negative), and then apply Fubini's theorem.

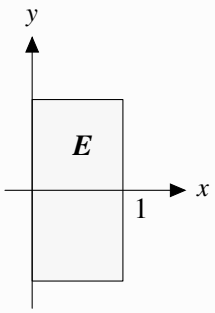
Example *3.47*

Consider the integral

$$\int_{\boldsymbol{E}} \frac{y}{x}\,\mathrm{d}x\,\mathrm{d}y$$

where

$$\boldsymbol{E} = (0,1)\times[-1,1]$$



The function

$$\begin{aligned} f : \boldsymbol{E} &\longrightarrow \mathbb{R}\\ (x,y) &\mapsto \frac{y}{x}\end{aligned}$$

is not non-negative, so we cannot solve the integral using Tonelli's theorem. To use Fubini's theorem, we must first ensure that the function is summable. To this end, we apply Tonelli's theorem to the function $|f|$, which is non-negative, and if it turns out that $\int_{\boldsymbol{E}} |f| < +\infty$, then f is summable and we can apply Fubini's theorem. Let $B = (0,1)$. Then, using Tonelli's theorem

$$\begin{aligned}\int_{\boldsymbol{E}} |f(x,y)|\,\mathrm{d}x\,\mathrm{d}y &= \int_0^1\left(\int_{-1}^1 \left|\frac{y}{x}\right|\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_0^1\left(\int_{-1}^1 \frac{|y|}{x}\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_0^1 \frac{1}{x}\left(\int_{-1}^1 |y|\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_0^1 \frac{1}{x}\underbrace{\left(\int_{-1}^0 -y\,\mathrm{d}y + \int_0^1 y\,\mathrm{d}y\right)}_{=\frac{1}{2}+\frac{1}{2}=1}\mathrm{d}x = \int_0^1 \frac{1}{x}\,\mathrm{d}x = +\infty.\end{aligned}$$

It follows that f is not summable, and therefore we cannot apply Fubini's theorem. The question now is: what happens if we perform the interchange? We find that what we obtain exchanging the order of integration is not the same:

$$\int_{(0,1)\times[-1,1]} \frac{y}{x}\,\mathrm{d}x\,\mathrm{d}y = \int_0^1\left(\int_{-1}^{+1}\frac{y}{x}\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 \frac{1}{x}\left[\frac{y^2}{2}\right]_{-1}^{+1}\mathrm{d}x = \int_0^1 \frac{1}{x}\cdot 0\,\mathrm{d}x = 0.$$

On the other hand,

$$\int_{(0,1)\times[-1,1]} \frac{y}{x}\,\mathrm{d}x\,\mathrm{d}y = \int_{-1}^{+1} y\left(\int_0^1 \frac{1}{x}\,\mathrm{d}x\right)\mathrm{d}y,$$

and the inner integral does not exist.

Theorem 3.56 [Measurability of Cartesian Products of Sets]
Let $m, n \in \mathbb{N}^*$. Let $\boldsymbol{A} \in \mathcal{M}_m$ and $\boldsymbol{B} \in \mathcal{M}_n$. Then, $\boldsymbol{A} \times \boldsymbol{B} \in \mathcal{M}_{m+n}$ and $\mu_{m+n}(\boldsymbol{A} \times \boldsymbol{B}) = \mu_m(\boldsymbol{A})\,\mu_n(\boldsymbol{B})$.

Proof The second statement follows from the first, and from Theorem 3.53. We observe that

$$\boldsymbol{A} \times \boldsymbol{B} = (\boldsymbol{A} \times \mathbb{R}^n) \cap (\mathbb{R}^m \times \boldsymbol{B}).$$

We can therefore limit ourselves to proving the result in the case where $\boldsymbol{A} = \mathbb{R}^m$ or $B = \mathbb{R}^n$. Suppose $\boldsymbol{B} = \mathbb{R}^n$ (the other case being similar). We start assuming that $\boldsymbol{A}$ is an m-dimensional interval. Then

$$\boldsymbol{A} \times \mathbb{R}^n = \bigcup_{k=0}^{\infty} (\boldsymbol{A} \times [-k, k]^n).$$

For every $k \in \mathbb{N}$, the set $\boldsymbol{A} \times [-k, k]^n$ is an $(m + n)$-dimensional interval and therefore it lives in $\mathcal{M}_{m+n}$. It follows that $\boldsymbol{A} \times \mathbb{R}^n \in \mathcal{M}_{m+n}$. Let $\boldsymbol{A} \in \mathcal{P}^\infty$, and $\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$, where $\boldsymbol{I}^k$ is an n-dimensional interval. Then,

$$\boldsymbol{A} \times \mathbb{R}^n = \bigcup_{k=0}^{\infty} (\boldsymbol{I}^k \times \mathbb{R}^n),$$

from which we find $\boldsymbol{A} \times \mathbb{R}^n \in \mathcal{M}_{m+n}$.
Let $\mu_m(\boldsymbol{A}) = 0$. Let $\varepsilon > 0$. There exists a sequence $(\boldsymbol{I}^j)_{j\in\mathbb{N}}$ of m-dimensional intervals such that

$$\boldsymbol{A} \subseteq \bigcup_{j=0}^{\infty} \boldsymbol{I}^j \quad \text{and} \quad \sum_{j=0}^{\infty} \mathrm{vol}_m(\boldsymbol{I}^j) < \varepsilon.$$

Let $k \in \mathbb{N}$. We have

$$\boldsymbol{A} \times [-k, k]^n \subseteq \bigcup_{j=0}^{\infty} (\boldsymbol{I}^j \times [-k, k]^n).$$

For every j, the set $\boldsymbol{I}^j \times [-k, k]^n$ is an $(m + n)$-dimensional interval and $\mathrm{vol}_{m+n}(\boldsymbol{I}^j \times [-k, k]^n) = \mathrm{vol}_m(\boldsymbol{I}^j)\,(2\,k)^n$. It follows that

$$\mu_n^*(\boldsymbol{A} \times [-k, k]^n) \leq \sum_{j=0}^{\infty} \mathrm{vol}_{m+n}(I^j \times [-k, k]^n) \leq \varepsilon\,(2\,k)^n.$$

Since ε is arbitrary, we obtain $\mu_n^*(\boldsymbol{A}\times[-k, k]^n) = 0$. Hence, $\boldsymbol{A}\times[-k, k]^n \in \mathcal{M}_{m+n}$, and therefore $\boldsymbol{A}\times\mathbb{R}^n \in \mathcal{M}_{m+n}$, because it coincides with $\bigcup_{k=0}^{\infty} (\boldsymbol{A} \times [-k, k]^n)$.

Let us now consider the case $\mu_m(\boldsymbol{A}) < +\infty$. In the proof of *(viii)* in Theorem 3.53 we saw that $\boldsymbol{A} = \boldsymbol{B} \setminus \boldsymbol{C}$, with $\boldsymbol{B} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$ and $\boldsymbol{A}^k \in \mathcal{P}^\infty$, $\mu_m(\boldsymbol{C}) = 0$. Hence,

$$\boldsymbol{A} \times \mathbb{R}^n = (\boldsymbol{B} \times \mathbb{R}^n) \setminus (\boldsymbol{C} \times \mathbb{R}^n) = \Big(\bigcap_{k=0}^{\infty} (\boldsymbol{A}^k \times \mathbb{R}^n) \Big) \setminus (\boldsymbol{C} \times \mathbb{R}^n).$$

Applying the previous cases, we obtain $\boldsymbol{A} \times \mathbb{R}^n \in \mathcal{M}_{m+n}$. Finally, given $\boldsymbol{A} \in \mathcal{M}_m$, we find $\boldsymbol{A} = \bigcup_{k=0}^{\infty} \boldsymbol{A}^k$, with $\boldsymbol{A}^k = \boldsymbol{A} \cap [-k, k]^n$. It follows that $\boldsymbol{A} \times \mathbb{R}^n = \bigcup_{k=0}^{\infty} (\boldsymbol{A}^k \times \mathbb{R}^n)$. Applying the previous case, we obtain the conclusion.

3.11.1 Change of variable theorem

Now we consider the change of variable theorem, which is another fundamental tool in the evaluation of integrals. We consider a particular class of n-dimensional intervals: the n-dimensional hypercubes.

Definition 3.21 *Let $I_1, I_2, \ldots, I_n$ be intervals with $\ell(I_1) = \ell(I_2) = \ldots = \ell(I_n)$. The set $\boldsymbol{I} = I_1 \times I_2 \times \ldots \times I_n$ is called n-dimensional hypercube.*

Lemma 3.10 *Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. Then*

$$\mu_n^*(\boldsymbol{E}) = \inf \left\{ \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \,\Big|\, (\boldsymbol{I}^k)_{k \in \mathbb{N}} \text{ sequence of open } n\text{-dimensional hypercubes such that } \boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k \right\}.$$

Proof Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. We let, for now,

$$\lambda(\boldsymbol{E}) = \inf \left\{ \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \,\Big|\, (\boldsymbol{I}^k)_{k \in \mathbb{N}} \text{ sequence of open } n\text{-dimensional hypercubes such that } \boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k \right\}.$$

Since open hypercubes are n-dimensional intervals, we have $\mu_n^*(\boldsymbol{E}) \leq \lambda(\boldsymbol{E})$ for every $\boldsymbol{E} \subseteq \mathbb{R}^n$. The conclusion follows if $\mu_n^*(\boldsymbol{E}) = +\infty$.
Let $\boldsymbol{I}$ be a non-empty n-dimensional interval such that $\boldsymbol{I} = I_1 \times \ldots \times I_n$, where $I_1, \ldots, I_n$ are bounded intervals of $\mathbb{R}$. Let $a_j = \inf I_j$ and $b_j = \sup I_j$ for all $j \in \{1, \ldots, n\}$. Let $k \in \mathbb{N}$. We denote by r_j the largest integer such that $r_j 2^{-k} \leq a_j$ and by s_j the smallest integer such that $s_j 2^{-k} > b_j$. Let $\boldsymbol{L}$ denote the set of elements $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n) \in \mathbb{Z}^m$ such that, for all $j \in \{1, \ldots, n\}$, there holds $r_j \leq \ell_j \leq s_j$. Given $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_n) \in \boldsymbol{L}$, we denote by $\boldsymbol{I}^{\boldsymbol{\ell}}$ the hypercube

$$\prod_{j=1}^{n} \left[\ell_j 2^{-k}, (\ell_j + 1) 2^{-k} \right).$$

Then, the hypercubes $\boldsymbol{I}^{\boldsymbol{\ell}}$ are pairwise disjoint for distinct values of $\boldsymbol{\ell}$ and

$$\bigcup_{\boldsymbol{\ell} \in \boldsymbol{L}} \boldsymbol{I}^{\boldsymbol{\ell}} = \prod_{j=1}^{n} \left[\ell_j 2^{-k}, (\ell_j + 1) 2^{-k} \right).$$

From **(d)** in Theorem 3.13,

$$\sum_{\ell \in \boldsymbol{L}} \mathrm{vol}_n(\boldsymbol{I}^\ell) = \prod_{j=1}^{n} (s_j - r_j)\, 2^{-k}.$$

For every j, we have

$$(s_j - r_j - 2)\, 2^{-k} \leq b_j - a_j \leq (s_j - r_j)\, 2^{-k},$$

so that

$$\mathrm{vol}_n(\boldsymbol{I}) \leq \sum_{\ell \in \boldsymbol{L}} \mathrm{vol}_n(\boldsymbol{I}^\ell) \leq \prod_{j=1}^{n} (b_j - a_j + 2^{1-k}).$$

It follows that, for every $\varepsilon > 0$, there exists a finite set $\boldsymbol{L}$ of n-dimensional hypercubes such that

$$\boldsymbol{I} \subseteq \bigcup_{\ell \in \boldsymbol{L}} \boldsymbol{I}^\ell \tag{3.64}$$

and

$$\sum_{\ell \in \boldsymbol{L}} \mathrm{vol}_n(\boldsymbol{I}^\ell) \leq \mathrm{vol}_n(\boldsymbol{I}) + \varepsilon. \tag{3.65}$$

Given an n-dimensional hypercube $\boldsymbol{J} = J_1 \times \ldots \times J_n$, it is clear that for every $\eta > 0$ it is possible to construct an n-dimensional open hypercube $\boldsymbol{J}'$ containing $\boldsymbol{J}$ and such that $\mathrm{vol}_n(\boldsymbol{J}') \leq \mathrm{vol}_n(\boldsymbol{J}) + \eta$: it is sufficient to define

$$\boldsymbol{J}' = \prod_{j=1}^{n} \left(\inf J_j - \delta, \sup J_j + \delta\right)$$

for a sufficiently small $\delta > 0$. We can therefore substitute each $\boldsymbol{I}^\ell$ in (3.64) with an open hypercube $\boldsymbol{J}^\ell$ with a volume arbitrarily close to $\mathrm{vol}_n(\boldsymbol{I}^\ell)$. In other words, for every $\varepsilon > 0$ there exists a finite set $\boldsymbol{L}$ of open n-dimensional hypercubes such that

$$\boldsymbol{I} \subseteq \bigcup_{\ell \in \boldsymbol{L}} \boldsymbol{I}^\ell \tag{3.66}$$

and

$$\sum_{\ell \in \boldsymbol{L}} \mathrm{vol}_n(\boldsymbol{I}^\ell) \leq \mathrm{vol}_n(\boldsymbol{I}) + \varepsilon. \tag{3.67}$$

We are now ready to conclude the proof. Let $\boldsymbol{E} \subseteq \mathbb{R}^n$ be such that $\mu_n^*(\boldsymbol{E}) < +\infty$. Let $\varepsilon > 0$. There exists a sequence $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ of n-dimensional intervals such that $\boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$ and $\sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \leq \mu_n^*(\boldsymbol{E}) + \frac{\varepsilon}{2}$. For every k, there exists a sequence $(\boldsymbol{I}^{k\ell})_{\ell \in \mathbb{N}}$ of open n-dimensional hypercubes such that

$$\boldsymbol{I}^k \subseteq \bigcup_{\ell=0}^{\infty} \boldsymbol{I}^{k\ell} \quad \text{and} \quad \sum_{\ell=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{k\ell}) \leq \mathrm{vol}_n(\boldsymbol{I}^k) + \frac{\varepsilon}{2^{k+1}}.$$

If follows that

$$\boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \bigcup_{\ell=0}^{\infty} \boldsymbol{I}^{k\ell}.$$

Then, if $\varphi : \mathbb{N} \longrightarrow \mathbb{N}^2$ is bijective, using Theorems 3.4, 3.5 and 3.6,

$$\lambda(\boldsymbol{E}) \leq \sum_{j=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{\varphi(j)}) = \sum_{k=0}^{\infty} \left(\sum_{\ell=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{k\,\ell}) \right)$$
$$\leq \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) + \frac{\varepsilon}{2^{k+1}} \leq \mu_n^*(\boldsymbol{E}) + \varepsilon.$$

and from the arbitrariness of ε we obtain $\lambda(\boldsymbol{E}) \leq \mu^*(\boldsymbol{E})$.

Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. We define

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \max_{j \in \{1,\ldots,n\}} |x_j - y_j|.$$

We have seen that d is a metric in $\mathbb{R}^n$ that induces the Euclidean topology. If $\boldsymbol{I}$ is the open hypercube $(a_1, b_1) \times \ldots \times (a_n, b_n)$, where $a_j, b_j \in \mathbb{R}$ for $j \in \{1, \ldots, n\}$ and $b_j - a_j = 2\,r > 0$, then $\boldsymbol{I}$ coincides with the open ball with respect to this metric of center $\left(\frac{a_1+b_1}{2}, \ldots, \frac{a_n+b_n}{2}\right)$ and radius r.

Let $\boldsymbol{E} \subseteq \mathbb{R}^n$. Let $T : \boldsymbol{E} \longrightarrow \mathbb{R}^n$. We say that T is Lipschitz if there exists $M > 0$ such that for every $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{E}$

$$\mathrm{d}(T(\boldsymbol{x}), T(\boldsymbol{y})) \leq M\,\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}). \tag{3.68}$$

We say that T is locally Lipschitz if, for every compact set $\boldsymbol{K} \subseteq \boldsymbol{E}$, the restriction $T|_{\boldsymbol{K}}$ is Lipschitz.

Theorem 3.57 *Let $\boldsymbol{\Omega}$ be an open subset of $\mathbb{R}^n$. Let $\boldsymbol{A} \subseteq \boldsymbol{\Omega}$ be measurable and let $T : \boldsymbol{A} \longrightarrow \mathbb{R}^n$ be injective, Lipschitz and open. Then, $T(\boldsymbol{A}) \in \mathcal{M}_n$.*

Proof We start assuming $\mu_n(\boldsymbol{A}) < +\infty$ and that T is Lipschitz. Then, there exists a sequence $(\boldsymbol{A}^k)_{k \in \mathbb{N}}$ of open sets such that, for all $k \in \mathbb{N}$,

$$\boldsymbol{A} \subseteq \boldsymbol{A}_{k+1} \subseteq \boldsymbol{A}_k \subseteq \boldsymbol{\Omega} \quad \text{and} \quad \lim_{k \to \infty} \mu_n(\boldsymbol{A}^k) = \boldsymbol{A}.$$

In fact, from Lemma 3.64, for every $k \in \mathbb{N}$ there exists a sequence $(\boldsymbol{I}^{k,j})_{j \in \mathbb{N}}$ of open hypercubes such that $\boldsymbol{A} \subseteq \bigcup_{j=0}^{\infty} \boldsymbol{I}^{k,j}$ and $\sum_{j=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^{k,j}) \leq \mu_n(\boldsymbol{A}) + \frac{1}{k}$. We define $\boldsymbol{\Omega}^k = \bigcup_{j=0}^{\infty} \boldsymbol{I}^{k,j}$. We let $\boldsymbol{A}^k = \boldsymbol{\Omega}^1 \cap \ldots \cap \boldsymbol{\Omega}^k \cap \boldsymbol{\Omega}$. Let $\boldsymbol{B} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$. From the injectivity of T we have

$$T(\boldsymbol{A}) = T(\boldsymbol{B}) \setminus T(\boldsymbol{B} \setminus \boldsymbol{A}).$$

The set $T(\boldsymbol{B})$ is measurable, because it coincides with $\bigcap_{k=0}^{\infty} T(\boldsymbol{A}^k)$. Indeed, T is open, and therefore for every k the set $T(\boldsymbol{A}^k)$ is open. We only need to show that $T(\boldsymbol{B} \setminus \boldsymbol{A})$ is measurable. We notice that $\mu_n(\boldsymbol{B}) = \lim_{k \to \infty} \mu_n(\boldsymbol{A}^k) = \mu_n(\boldsymbol{A})$. Therefore, $\mu_n(\boldsymbol{B} \setminus \boldsymbol{A}) = 0$. Let $\varepsilon > 0$. There exists a sequence $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ of open hypercubes such that $\boldsymbol{B} \setminus \boldsymbol{A} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$ and $\sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \leq \varepsilon$. If $\boldsymbol{I}^k = (a_1, b_1) \times \ldots \times (a_n, b_n)$ with $b_j - a_j = 2\,r$ for $j \in \{1, \ldots, n\}$, if (3.68) holds for every $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{\Omega}$ and if $\boldsymbol{x}^0 \in \boldsymbol{I}^k \cap \boldsymbol{\Omega}$,

$$T(\boldsymbol{\Omega} \cap \boldsymbol{I}^k) \subseteq \prod_{j=1}^{n} (T_j(\boldsymbol{x}^0) - M\,r, T_j(\boldsymbol{x}^0) + M\,r).$$

It follows that

$$T(\boldsymbol{B} \setminus \boldsymbol{A}) \subseteq \bigcup_{k=1}^{\infty} T(\boldsymbol{\Omega} \cap \boldsymbol{I}^k) \subseteq \bigcup_{k=1}^{\infty} \boldsymbol{J}^k,$$

where, for every $k \in \mathbb{N}$, the set $\boldsymbol{J}^k$ is an open hypercube whose volume is not greater than $M^n \ \mathrm{vol}_n(\boldsymbol{I}^k)$. Hence,

$$\mu_n^*(T(\boldsymbol{B} \setminus \boldsymbol{A})) \leq M^n \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k) \leq M^n \, \varepsilon.$$

From the arbitrariness of ε, we get $\mu_n^*(T(\boldsymbol{B} \setminus \boldsymbol{A})) = 0$. Therefore, $T(\boldsymbol{B} \setminus \boldsymbol{A})$ is measurable by Theorem 3.19. We now prove the general result. For every $k \in \mathbb{N}$, we define

$$\boldsymbol{\Omega}_k = \left\{ \boldsymbol{x} \in \boldsymbol{\Omega} \,\middle|\, \|\boldsymbol{x}\| \leq k, \ \ \mathrm{d}(\boldsymbol{x}, \partial\boldsymbol{\Omega})) \geq \frac{1}{k} \right\}.$$

For every $k \in \mathbb{N}$, the set $\boldsymbol{\Omega}_k$ is open, $\overline{\boldsymbol{\Omega}}_k$ is a compact set contained in $\boldsymbol{\Omega}$ and $\bigcup_{k=0}^{\infty} \boldsymbol{\Omega}_k = \boldsymbol{\Omega}$. Thus,

$$T(\boldsymbol{A}) = \bigcup_{k=0}^{\infty} T(\boldsymbol{A} \cap \boldsymbol{\Omega}_k).$$

We can now exploit the first part of the proof and conclude that for every $k \in \mathbb{N}$ the set $T(\boldsymbol{A} \cap \boldsymbol{\Omega}_k)$ is measurable, and therefore $T(\boldsymbol{A})$ is measurable.

Remark 3.24 In the proof of Theorem 3.19, we have implicitly shown that, if $\mu_n(A) = 0$, then $\mu_n(T(A)) = 0$.

We now consider the case where $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is linear and invertible (i.e., it is an *automorphism*). We start with a preliminary result, which needs the following notation: if $\boldsymbol{x} \in \mathbb{R}^n$, $r \in \mathbb{R}$ and $\boldsymbol{A} \subseteq \mathbb{R}^n$, we have defined

$$\boldsymbol{x} + \boldsymbol{A} = \{\boldsymbol{x} + \boldsymbol{a} \,|\, \boldsymbol{a} \in \boldsymbol{A}\}$$

and we now define

$$r\,\boldsymbol{A} = \{r\,\boldsymbol{a} \,|\, \boldsymbol{a} \in \boldsymbol{A}\}.$$

Lemma 3.11 *Let $\boldsymbol{A} \in \mathcal{M}_n$, $\boldsymbol{x} \in \mathbb{R}^n$ and $r > 0$. Then, $\boldsymbol{x} + \boldsymbol{A}$ and $r\,\boldsymbol{A}$ are measurable. Moreover,*

$$\mu_n(\boldsymbol{x} + \boldsymbol{A}) = \mu_n(\boldsymbol{A}) \quad \textit{and} \quad \mu_n(r\,\boldsymbol{A}) = r^n \, \mu_n(\boldsymbol{A}).$$

Proof The sets $\boldsymbol{x} + \boldsymbol{A}$ and $r\,\boldsymbol{A}$ are measurable in view of Theorem 3.57, by setting $T(\boldsymbol{y}) = \boldsymbol{x} + \boldsymbol{y}$ and $T(\boldsymbol{y}) = r\,\boldsymbol{y}$. Let $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ be a sequence of n-dimensional intervals such that $\boldsymbol{A} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$. Then, $\boldsymbol{x} + \boldsymbol{A} \subseteq \bigcup_{k=0}^{\infty} (\boldsymbol{x} + \boldsymbol{I}^k)$. Each $\boldsymbol{x} + \boldsymbol{I}^k$ is an n-dimensional interval and $\mathrm{vol}_n(\boldsymbol{x} + \boldsymbol{I}^k) = \mathrm{vol}_n(\boldsymbol{I}^k)$. Thus, $\mu_n(\boldsymbol{x} + \boldsymbol{A}) \leq \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k)$, so that $\mu_n(\boldsymbol{x} + \boldsymbol{A}) \leq \mu_n(\boldsymbol{A})$. Since $\boldsymbol{A} = (-\boldsymbol{x}) + (\boldsymbol{x} + \boldsymbol{A})$, we also have the opposite inequality.

Moreover, $r\,\boldsymbol{A} \subseteq \bigcup_{k=0}^{\infty} r\,\boldsymbol{I}^k$. If $\boldsymbol{I} = I_1 \times \ldots \times I_n$, where I_j are bounded intervals in $\mathbb{R}$, we have $r\,\boldsymbol{I} = (r\,I_1) \times \ldots \times (r\,I_n)$, and therefore

$$\mathrm{vol}_n(r\,\boldsymbol{I}) = r^n \ \mathrm{vol}_n(\boldsymbol{I}).$$

It follows that $\mu_n(A) \leq r^n \sum_{k=0}^{\infty} \mathrm{vol}_n(\boldsymbol{I}^k)$, which leads to $\mu_n(r\,\boldsymbol{A}) \leq r^n\,\mu_n(\boldsymbol{A})$. Since $\boldsymbol{A} = r^{-1}(r\,\boldsymbol{A})$, we obtain the opposite inequality.

If $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is linear and invertible, we define

$$\Delta(T) = \mu_n(T([0,1)^n)). \tag{3.69}$$

We notice that $\Delta(T)$ is well defined in view of Theorem 3.57. Moreover, $0 < \Delta(T) < +\infty$. Indeed, $[0,1)^n \subset [0,1]^n$, and $[0,1]^n$ is compact. It follows that $T([0,1)^n)$ is bounded, i.e., $\Delta < +\infty$. Moreover, $(0,1)^n \subset [0,1)^n$, and therefore $T([0,1)^n)$ contains $T((0,1)^n)$, which is open and non-empty. Thus, $\Delta(T) > 0$.

Theorem 3.58 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be linear and invertible. Then, $T(\boldsymbol{A}) \in \mathcal{M}_n$ and*

$$\mu_n(T(\boldsymbol{A})) = \Delta(T)\,\mu_n(\boldsymbol{A}).$$

Proof The first claim follows from Theorem 3.57. Let

$$\boldsymbol{A} = \prod_{j=1}^{n} [a_j, a_j + r)$$

where $a_1, \ldots, a_n \in \mathbb{R}$ and $r > 0$. We define $\boldsymbol{a} = (a_1, \ldots, a_n)$. Then, $\boldsymbol{A} = \boldsymbol{a} + r\,[0,1)^n$, and by linearity

$$T(\boldsymbol{A}) = T(\boldsymbol{a}) + r\,T([0,1)^n).$$

From Lemma 3.11 we find

$$\mu_n(T(\boldsymbol{A})) = \mu_n(r\,T([0,1)^n) = r^n\,\Delta(T) = \Delta(T)\,\mu_n(\boldsymbol{A}).$$

We prove the result in the case where $\boldsymbol{A}$ is open. Let $k \in \mathbb{N}$. Let $\boldsymbol{r} = (r_1, \ldots, r_n) \in \mathbb{Z}^n$. We denote by $\boldsymbol{I}^{r\,k}$ the n-dimensional hypercube

$$\prod_{j=1}^{n} [r_j\,2^{-k}, (r_j + 1)\,2^{-k}). \tag{3.70}$$

We define also

$$\boldsymbol{A}^k = \bigcup_{\substack{\boldsymbol{r} \in \mathbb{Z}^n \\ \boldsymbol{I}^{r\,k} \subseteq \boldsymbol{A}}} \boldsymbol{I}^{r\,k}.$$

For every $k \in \mathbb{N}$, we have $\boldsymbol{A}^k \subseteq \boldsymbol{A}^{k+1}$. Indeed, if $\boldsymbol{r} \in \mathbb{Z}^n$, the set in (3.70) is the union of hypercubes

$$\prod_{j=1}^{n} [s_j\,2^{-(k+1)}, (s_j + 1)\,2^{-(k+1)}), \quad \text{with } 2\,r_j \leq s_j \leq 2\,r_j + 1.$$

Moreover,

$$\bigcup_{k=0}^{\infty} \boldsymbol{A}^k = \boldsymbol{A}. \tag{3.71}$$

The inclusion $\subseteq$ is trivial. Conversely, let $\boldsymbol{x} \in \boldsymbol{A}$. Since $\boldsymbol{A}$ is open, there exists $r > 0$ such that

$$\prod_{j=1}^{n}(x_j - r, x_j + r) \subseteq \boldsymbol{A}.$$

Let $k \in \mathbb{N}$ be such that $2^{-k} < r$. Moreover, let $r_j = \max\{m \in \mathbb{Z} \,|\, m\,2^{-k} \leq x_j\}$. Then, $r_j\,2^{-k} \leq x_j < (r_j + 1)\,2^{-k}$, from which we find $r_j\,2^{-k} > x_j - 2^{-k} > x_j - r$ and $(r_j + 1)\,2^{-k} < x_j + 2^{-k} < x_j + r$. It follows that

$$\prod_{j=1}^{n}[r_j\,2^{-k}, (r_j + 1)\,2^{-k}) \subseteq \boldsymbol{A},$$

which proves (3.71). For every $k \in \mathbb{N}$, we have

$$\begin{aligned}\mu_n(T(\boldsymbol{A}^k)) &= \mu_n\Big(\bigcup_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{r\,k}\subseteq\boldsymbol{A}}} T(\boldsymbol{I}^{r\,k})\Big)\\ &= \sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{r\,k}\subseteq\boldsymbol{A}}} \mu_n(T(\boldsymbol{I}^{r\,k})) = \Delta(T)\sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{r\,k}\subseteq\boldsymbol{A}}} \mu_n(\boldsymbol{I}^{r\,k}) = \Delta(T)\,\mu_n(\boldsymbol{A}^k).\end{aligned}$$

Hence,

$$\begin{aligned}\mu_n(T(\boldsymbol{A})) &= \mu_n\Big(\bigcup_{k=0}^{\infty} T(\boldsymbol{A}^k)\Big)\\ &= \lim_{k\to\infty}\mu_n(T(\boldsymbol{A}^k)) = \lim_{k\to\infty}\Delta(T)\,\mu_n(\boldsymbol{A}^k) = \Delta(T)\,\mu_n(\boldsymbol{A}).\end{aligned}$$

We now prove the result in the case $\mu_n(\boldsymbol{A}) < +\infty$. In the proof of Theorem 3.57, we showed that there exists a sequence of open sets $(\boldsymbol{A}^k)_{k\in\mathbb{N}}$ such that, for all $k \in \mathbb{N}$, we have $\boldsymbol{A} \subseteq \boldsymbol{A}_{k+1} \subseteq \boldsymbol{A}_k$ and $\mu_n(\boldsymbol{A}_k) \leq \mu_n(\boldsymbol{A}) + \frac{1}{k}$. We define $\boldsymbol{B} = \bigcap_{k=0}^{\infty} \boldsymbol{A}^k$. Then,

$$T(\boldsymbol{A}) = T(\boldsymbol{B}) \setminus T(\boldsymbol{B} \setminus \boldsymbol{A}),$$

and since $\mu_n(\boldsymbol{B} \setminus \boldsymbol{A}) = 0$, in view of Remark 3.24, we obtain

$$\mu_n(T(\boldsymbol{A})) = \mu_n(T(\boldsymbol{B})) = \mu_n\Big(\bigcap_{k=0}^{\infty} T(\boldsymbol{A}^k)\Big). \tag{3.72}$$

We get $\mu_n(T(\boldsymbol{A}_0)) = \Delta(T)\,\mu_n(\boldsymbol{A}_0) < +\infty$. From the second point of Theorem 3.11 we obtain

$$\mu_n(T(\boldsymbol{A})) = \lim_{k\to\infty}\mu_n(T(\boldsymbol{A}^k)) = \lim_{k\to\infty}\Delta(T)\,\mu_n(\boldsymbol{A}^k) = \Delta(T)\,\mu_n(\boldsymbol{A}).$$

We now consider the general case. For every $k \in \mathbb{N}$, we define $\boldsymbol{A}^k = \boldsymbol{A} \cap [-k, k]^n$. Then,

$$\begin{aligned}\mu_n(T(\boldsymbol{A})) &= \mu_n\Big(\bigcup_{k=0}^{\infty} T(\boldsymbol{A}^k)\Big)\\ &= \lim_{k\to\infty}\mu_n(T(\boldsymbol{A}^k))\\ &= \lim_{k\to\infty}\Delta(T)\,\mu_n(\boldsymbol{A}^k) = \Delta(T)\,\mu_n(\boldsymbol{A}).\end{aligned}$$

We want to find an expression for $\Delta(T)$. To this end, we introduce the following notation. Let i, j be distinct natural numbers between 1 and n. We denote by $C^{i,j}$ be the unique linear transformation from $\mathbb{R}^n$ to itself which,

for $k \in \{1, \dots, n\}$, satisfies

$$C^{i,j}(\boldsymbol{e}_k) = \begin{cases} \boldsymbol{e}_k & \text{if } k \notin \{i, j\} \\ \boldsymbol{e}_j & \text{if } k = i \\ \boldsymbol{e}_i & \text{if } k = j. \end{cases} \tag{3.73}$$

The inverse of this transformation is $C^{i,j}$ itself, i.e., $(C^{i,j})^{-1} = C^{i,j}$. Let $i, j \in \{1, \dots, n\}$ be again distinct, and let $c \in \mathbb{R}$. We denote by $D^{i,j}(c)$ the linear transformation from $\mathbb{R}^n$ in itself which, for $k \in \{1, \dots, n\}$, satisfies

$$D^{i,j}(c)(\boldsymbol{e}_k) = \begin{cases} \boldsymbol{e}_k & \text{if } k \neq j \\ \boldsymbol{e}_j + c\,\boldsymbol{e}^i & \text{if } k = j. \end{cases} \tag{3.74}$$

We observe that $D^{i,j}(c)^{-1} = D^{i,j}(-c)$. We say that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is upper triangular if it is in the form

$$T(x_1, x_2, \dots, x_n) = (a_{1,1}\, x_1 + \dots + a_{1,n}\, x_n, a_{2,2}\, x_2 + \dots + a_{2,n}\, x_n, \dots, a_{n,n}\, x_n). \tag{3.75}$$

Lemma 3.12 *Let $n \geq 2$. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be an invertible linear transformation. Then, T can be expressed as the composition of a finite number of invertible linear transformations in the forms (3.73), (3.74) and (3.75).*

Proof Let

$$T(x_1, x_2, \dots, x_n) = (a_{1,1}\, x_1 + \dots + a_{1,n}\, x_n, a_{2,1}\, x_1 + \dots + a_{2,n}\, x_n, \dots, a_{n,1}\, x_1 + \dots + a_{n,n}\, x_n). \tag{3.76}$$

Since T is invertible, there exists $j \in \{1, \dots, n\}$ such that $a_{1,j} \neq 0$. Then, if $a_{1,1} = 0$, substituting T with $T \circ C^{1,j}$, we are brought back to the case where $a_{1,1} \neq 0$. If this condition is not satisfied, we can consider the composition

$$D^{n,1}\left(-\frac{a_{n,1}}{a_{1,1}}\right) \circ \dots \circ D^{2,1}\left(-\frac{a_{2,1}}{a_{1,1}}\right) \circ T.$$

Considering this transformation in place of T, we have

$$a_{1,1} \neq 0, \quad a_{2,1} = \dots = a_{n,1} = 0. \tag{3.77}$$

Under this last assumption, if we have $a_{2,2} = 0$, we choose $j > 2$ such that $a_{2,j} \neq 0$ and, replacing T with $T \circ C^{2\,j}$ we are brought back to the case $a_{2,2} \neq 0$. Condition (3.77) is maintained. Suppose that now $a_{2,2} \neq 0$. We substitute T with

$$D^{n,2}\left(-\frac{a_{n,2}}{a_{2,2}}\right) \circ \dots \circ D^{3,2}\left(-\frac{a_{3,2}}{a_{2,2}}\right) \circ T.$$

In this way, we are in the case where

$$a_{1,1} \neq 0, \quad a_{2,2} \neq 0, \quad a_{2,1} = \dots = a_{n,1} = a_{3,2} = \dots = a_{n,2} = 0. \tag{3.78}$$

In a finite number of steps, we obtain an upper triangular transformation S. To summarize, we have found linear transformations $C_1, \dots, C_r$ in the form (3.73) and $D_1, \dots, D_s$ in the form (3.74) such that

$$S = D_1 \circ D_2 \circ \dots \circ D_s \circ T \circ C_1 \circ \dots \circ C_r$$

is upper triangular. It follows that

$$T = D_s^{-1} \circ \dots \circ D_1^{-1} \circ S \circ C_1^{-1} \circ \dots \circ C_r^{-1},$$

and as already observed the transformations $D_s^{-1}, \ldots, D_1^{-1}$ are still in the form (3.74), whereas the transformations $C_1^{-1}, \ldots, C_r^{-1}$ are in the form (3.73).

We can associate an $n \times n$ matrix to any linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ (for example, with respect to the standard basis of $\mathbb{R}^n$). If T is in the form (3.76), the matrix of T has $a_{i,j}$ as its entry in position (i, j).

Theorem 3.59 *Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be an invertible linear transformation. Then,*

$$\Delta(T) = |\det T|.$$

Proof First, we notice that the result is true if $n = 1$. In this case, T is in the form $T(x) = c\,x$, with $c \in \mathbb{R}^*$. We have

$$T([0,1)) = \begin{cases} [0, c) & \text{if } c > 0 \\ (c, 0] & \text{if } c < 0, \end{cases}$$

which yields $\Delta(T) = |c| = |\det T|$.
Let $n \geq 2$. From Theorem 3.58, if $T = S\,U$, where S and U are invertible linear transformations from $\mathbb{R}^n$ into itself, there holds

$$\Delta(T) = \Delta(S)\,\Delta(U).$$

From the properties of determinants and from Lemma 3.12, we see that to prove the result it is sufficient to prove it for linear transformations in the form (3.73), (3.74) and (3.75). We define $\boldsymbol{Q}_n = [0,1)^n$. We will write simply $\boldsymbol{Q}$ if n is clear from the context. We start considering linear transformations in the form (3.73). For every (i, j), with $i \neq j$, we have $C^{i,j}(\boldsymbol{Q}) = \boldsymbol{Q}$, so that $\Delta(C^{i,j}) = 1$, whereas $\det(C^{i,j}) = 1$, which implies that the result is true.
Consider a linear transformation $D^{i,j}(c)$ in the form (3.74). We have $\det(D^{i,j}(c)) = 1$. Suppose, for example, $i < j$. We let

$$\boldsymbol{A} = \{(x_1, \ldots, x_i, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) \in \mathbb{R}^{n-1} \,\big|\, 0 \leq x_k < 1, \quad \forall\, k \in \{1, \ldots, j-1, j+1, \ldots, n\}\}.$$

It is easy to see that

$$D^{i,j}(c)(\boldsymbol{Q}) = \{(y_1, \ldots, y_j, \ldots, y_n) \in \mathbb{R}^n \,|\, (y_1, \ldots, y_i, \ldots, y_{j-1}, y_{j+1}, \ldots, y_n) \in \boldsymbol{A},\ c\,y_i \leq y_j < c\,y_i + 1\}.$$

We denote, for the sake of conciseness, $D^{i,j}(c)(\boldsymbol{Q})$ with the symbol B. If $\boldsymbol{y}' = (y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_n)$ and $B_{\boldsymbol{y}'} = \{y_j \in \mathbb{R} \,|\, (y_1, \ldots, y_{j-1}, y_j, y_{j+1}, \ldots, y_n) \in B\}$, we have $\mu_1(B_{\boldsymbol{y}'}) = 1$ for every $\boldsymbol{y}' \in \boldsymbol{A}$. Hence, from Theorem 3.53,

$$\mu_n(D^{i,j}(c)(\boldsymbol{Q})) = \int_{\boldsymbol{A}} 1\, \mathrm{d}\boldsymbol{y}' = \mu_{n-1}(\boldsymbol{A}) = 1.$$

The result in this second particular case is proved. We now have to consider the case where T is upper triangular and invertible. Let

$$T(x_1, x_2, \ldots, x_n) = (a_{1,1}\,x_1 + \ldots + a_{1,n}\,x_n, a_{2,2}\,x_2 + \ldots + a_{2,n}\,x_n, \ldots, a_{n,n}\,x_n).$$

We have $\det T = a_{1,1}\,a_{2,2} \cdots a_{n,n}$, so that T is invertible if and only if $a_{j,j} \neq 0$ for all $j \in \{1, \ldots, n\}$. We show that $\Delta(T) = |a_{1,1}\,a_{2,2} \cdots a_{n,n}|$ by induction over n. The result is obvious if $n = 1$. Suppose that the result is true for n, and we prove it for $n + 1$. Let $T : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^{n+1}$ such that

$$\begin{aligned} T(x_1, x_2, \ldots, x_n, x_{n+1}) = \quad & (a_{1,1}\,x_1 + \ldots + a_{1,n}\,x_n + a_{1,n+1}\,x_{n+1}, \\ & a_{2,2}\,x_2 + \ldots + a_{2,n}\,x_n + a_{2,n+1}\,x_{n+1}, \ldots, a_{n,n}\,x_n + a_{n,n+1}\,x_{n+1}, a_{n+1,n+1}\,x_{n+1}), \end{aligned}$$

where $a_{1,1} \cdots a_{n,n}\, a_{n+1,n+1} \neq 0$. Let $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be defined by

$$S(x_1, \ldots, x_n) = (a_{1,1}\, x_1 + \ldots + a_{1,n}\, x_n, a_{2,2}\, x_2 + \ldots + a_{2\,n}\, x_n, \ldots, a_{n,n}\, x_n)$$

and

$$A = \begin{cases} [0, a_{n+1,n+1}) & \text{if } a_{n+1,n+1} > 0 \\ (a_{n+1,n+1}, 0] & \text{if } a_{n+1,n+1} < 0 \end{cases}$$

and $\boldsymbol{c} = (a_{1,n+1}, a_{2,n+1}, \ldots, a_{n,n+1}) \in \mathbb{R}^n$. It is easy to verify that for every $x_{n+1} \in \mathbb{R}$

$$T(\boldsymbol{Q}_{n+1})x_{n+1} = \begin{cases} S(\boldsymbol{Q}_n) + \frac{x_{n+1}}{a_{n+1,n+1}}\, \boldsymbol{c} & \text{if } x_{n+1} \in A \\ \varnothing & \text{otherwise} \end{cases}$$

Applying Theorem 3.53, Lemma 3.11 and the inductive assumption, we obtain

$$\mu_{n+1}(T(\boldsymbol{Q}_{n+1})) = \mu_n(S(\boldsymbol{Q}_n))\, |a_{n+1,n+1}| = |a_{1,1}\ \ldots\ a_{n,n}\, a_{n+1,n+1}|.$$

We now consider a class of transformations which are not necessarily linear.

Definition 3.22 [Change of Variable]
Let $\boldsymbol{E}$ be an open subset of $\mathbb{R}^n$. Let $T : \boldsymbol{E} \longrightarrow \mathbb{R}^n$. We say that T is a change of variable *if*

(i) T is injective;

(ii) T is of class C^1;

(iii) for every $x \in \boldsymbol{E}$, the determinant $\det J_T(\boldsymbol{x})$ *of the Jacobian matrix of T at $\boldsymbol{x}$ is not zero.*

Remark 3.25 The inverse function theorem ensures that T is open (which means that the image of any open set is open). Let $\boldsymbol{A} = [a_1, b_1] \times \ldots \times [a_n, b_n]$ where $a_j, b_j \in \mathbb{R}$ and $a_j < b_j$ for every $j \in \{1, \ldots, n\}$ and $\boldsymbol{A} \subseteq \boldsymbol{\Omega}$. Then, for every $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n) \in \boldsymbol{A}$, applying the mean value theorem, we find

$$\mathrm{d}(T(\boldsymbol{x}), T(\boldsymbol{y})) = \max_{j \in \{1,\ldots,n\}} |T_j(\boldsymbol{x}) - T_j(\boldsymbol{y})| \leq \max_{j \in \{1,\ldots,n\}} \max_{z \in \boldsymbol{A}} \left| \frac{\partial T_j}{\partial x_i}(z) \right| \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}). \tag{3.79}$$

Another consequence of the local invertibility theorem is the fact that if T is a change of variable, such is also T^{-1}.

Lemma 3.13 *Let T be a change of variable in $\mathbb{R}^n$. Then, T satisfies the assumptions of Theorem 3.57.*

Proof By assumption, T is injective, and we have already observed that T is open. We only need to show that T is locally Lipschitz. Let H be a compact subset of $\boldsymbol{\Omega}$. By contradiction, suppose that for every $k \in \mathbb{N}$ there exist $\boldsymbol{x}^k$ and $\boldsymbol{y}^k$ in H such that

$$\mathrm{d}(T(\boldsymbol{x}^k), T(\boldsymbol{y}^k)) > k\, \mathrm{d}(\boldsymbol{x}^k, y^k). \tag{3.80}$$

Since H is compact, we can assume that the two sequences $(\boldsymbol{x}^k)_{k \in \mathbb{N}}$ and $(\boldsymbol{y}^k)_{k \in \mathbb{N}}$ converge, respectively, to $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$. We observe that $\hat{\boldsymbol{y}} = \hat{\boldsymbol{x}}$. In fact, if this was not the case, the sequence $(\mathrm{d}(T(\boldsymbol{x}^k), T(\boldsymbol{y}^k)))_{k \in \mathbb{N}}$ would tend to $+\infty$. This is not possible, because T is continuous, and therefore bounded over H. Since $\boldsymbol{\Omega}$ is open, there exists $r > 0$ such that

$$\prod_{j=1}^{n} [\hat{x}_j - r, \hat{x}_j + r] \subseteq \boldsymbol{\Omega}.$$

From (3.79), there exists $M > 0$ such that

$$\mathrm{d}(T(\boldsymbol{x}), T(\boldsymbol{y})) \leq M\, \mathrm{d}(\boldsymbol{x}, \boldsymbol{y})$$

for every choice of $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\prod_{j=1}^{n} [\hat{x}_j - r, \hat{x}_j + r] \subseteq \boldsymbol{\Omega}$. Since $\lim_{k\to\infty} \boldsymbol{x}^k = \lim_{k\to\infty} \boldsymbol{y}^k = \hat{\boldsymbol{x}}$, for sufficiently large values of k

$$\mathrm{d}(T(\boldsymbol{x}^k), T(\boldsymbol{y}^k)) \leq M\, \mathrm{d}(\boldsymbol{x}^k, \boldsymbol{y}^k),$$

which contradicts (3.80).

Given a change of variable of domain $\boldsymbol{\Omega}$ and given $\boldsymbol{A} \in \mathcal{M}_n$ contained in $\boldsymbol{\Omega}$, we want an expression for $\mu_n(T(\boldsymbol{A}))$ that generalizes Theorem 3.58. The following lemma will be crucial.

Lemma 3.14 *Let $\boldsymbol{\Omega}$ be an open set of $\boldsymbol{\Omega}$. Let T be a change of variable with domain $\boldsymbol{\Omega}$. Let $\boldsymbol{A}$ be open, and let $\overline{\boldsymbol{A}} \subseteq \boldsymbol{\Omega}$. Let $\varepsilon \in (0, 1]$. Then, there exists $r_0 > 0$ (dependent upon $\boldsymbol{\Omega}$ and ε) such that*

(a) *for every $\boldsymbol{x} \in \boldsymbol{A}$, there holds $\boldsymbol{x} + [-r_0, r_0]^n \subseteq \boldsymbol{\Omega}$;*

(b) *for every $\boldsymbol{x} \in \boldsymbol{A}$ and for every $r \in (0, r_0]$*

$$\begin{aligned} T(\boldsymbol{x}) + \mathrm{d}T(\boldsymbol{x})(-r(1-\varepsilon), r(1-\varepsilon)]^n) \subset\ & T(\boldsymbol{x} + [-r, r]^n) \\ \subseteq\ & T(\boldsymbol{x}) + \mathrm{d}T(\boldsymbol{x})([-r(1+\varepsilon), r(1+\varepsilon)]^n). \end{aligned}$$

Proof The existence of $r_0 > 0$ for which **(a)** holds is a simple consequence of the fact that $\overline{\boldsymbol{A}}$ is compact. This condition also ensures the existence of $M > 0$ such that, for every $\boldsymbol{x} \in \boldsymbol{A}$ and for every $\boldsymbol{h} \in \mathbb{R}^n$,

$$\|\mathrm{d}T(\boldsymbol{x})(\boldsymbol{h})\| \leq M\, \|\boldsymbol{h}\|, \ \ \|\mathrm{d}T(\boldsymbol{x})^{-1}(\boldsymbol{h})\| \leq M\, \|\boldsymbol{h}\|. \tag{3.81}$$

For $\boldsymbol{x}$ and $\boldsymbol{x} + \boldsymbol{h}$ belonging to $\boldsymbol{\Omega}$, we define

$$r(\boldsymbol{x}, \boldsymbol{h}) = T(\boldsymbol{x} + \boldsymbol{h}) - T(\boldsymbol{x}) - \mathrm{d}T(\boldsymbol{x})(\boldsymbol{h}).$$

We denote by $\mathrm{d}r(\boldsymbol{x}, \boldsymbol{h})$ the partial differential of r with respect to the variable $\boldsymbol{h}$: we have $\mathrm{d}r(\boldsymbol{x}, \boldsymbol{h}) = \mathrm{d}T(\boldsymbol{x} + \boldsymbol{h}) - \mathrm{d}T(\boldsymbol{x})$. From the uniform continuity of the partial derivatives of T over the compact subsets of $\boldsymbol{\Omega}$ and from the mean value theorem (applied componentwise), for every $\eta > 0$ there exists $r > 0$ (that we can assume to be non-greater than r_0) such that, for every $\boldsymbol{x} \in \boldsymbol{A}$ and for every $\boldsymbol{h}^1, \boldsymbol{h}^2 \in \mathbb{R}^n$ with $\max\{\|\boldsymbol{h}^1\|, \|\boldsymbol{h}^2\|\} \leq r$, there holds

$$\|r(\boldsymbol{x}, \boldsymbol{h}^1) - r(\boldsymbol{x}, \boldsymbol{h}^2)\| \leq \eta\, \|\boldsymbol{h}^1 - \boldsymbol{h}^2\|.$$

Let $\varepsilon > 0$ and $\boldsymbol{h} \in [-r, r]^n$ with $0 < r \leq r_0$. We want a solution $\boldsymbol{k} \in [-r(1+\varepsilon), r(1+\varepsilon)]^n$ of the equation

$$T(\boldsymbol{x} + \boldsymbol{h}) = T(\boldsymbol{x}) + \mathrm{d}T(\boldsymbol{x})(\boldsymbol{k}). \tag{3.82}$$

This equation has only one solution in $\mathbb{R}^n$, given by

$$\boldsymbol{k} = \mathrm{d}T(\boldsymbol{x})^{-1}(T(\boldsymbol{x} + \boldsymbol{h}) - T(\boldsymbol{x})) = \boldsymbol{h} + \mathrm{d}T(\boldsymbol{x})^{-1} r(\boldsymbol{x}, \boldsymbol{h}).$$

Therefore, for $j \in \{1, \ldots, n\}$, if $\sqrt{n}\, r \leq r$,

$$|k_j| \leq |h_j| + M\, \|r(\boldsymbol{x}, \boldsymbol{h})\| \leq |h_j| + M\, \eta\, \|\boldsymbol{h}\| \leq (1 + M\, \eta\, \sqrt{n})\, r \leq (1 + \varepsilon)\, r$$

if $M\, \eta\, \sqrt{n} \leq \varepsilon$. Let $\boldsymbol{k} \in [-r(1-\varepsilon), r(1-\varepsilon)]^n$. Consider the equation

$$T(\boldsymbol{x} + \boldsymbol{h}) = T(\boldsymbol{x}) + \mathrm{d}T(\boldsymbol{x})(\boldsymbol{k}). \tag{3.83}$$

We want a solution $\boldsymbol{h} \in [-r, r]^n$. Eq. (3.83) is equivalent to

$$\boldsymbol{h} = \boldsymbol{k} - \mathrm{d}T(\boldsymbol{x})^{-1} r(\boldsymbol{x}, \boldsymbol{h}). \tag{3.84}$$

Let

$$\begin{aligned} g : [-r, r]^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{h} &\mapsto \boldsymbol{k} \mathrm{d}T(\boldsymbol{x})^{-1} r(\boldsymbol{x}, \boldsymbol{h}). \end{aligned}$$

We look for a fixed point of g. First, for $j \in \{1, \ldots, n\}$ and $\boldsymbol{h} \in [-r, r]^n$, if $\sqrt{n}\, r \leq r$,

$$\begin{aligned} |g_j(\boldsymbol{h})| &\leq |k_j| + \|\mathrm{d}T(\boldsymbol{x})^{-1} r(\boldsymbol{x}, \boldsymbol{h})\| \\ &\leq |k_j| + M\, \eta\, \|\boldsymbol{h}\| \\ &\leq r\,(1 - \varepsilon) + M\, \eta\, \sqrt{n}\, r \leq r \end{aligned}$$

if $\eta \leq \frac{\varepsilon}{\sqrt{n}\, M}$. It follows that, if r is sufficiently small, $g([-r, r]^n) \subseteq [-r, r]^n$. Moreover, if $\boldsymbol{h}^1, \boldsymbol{h}^2$ belong to $[-r, r]^n$, we have, for $\sqrt{n}\, r \leq r$,

$$\|g(\boldsymbol{h}^1 - g(\boldsymbol{h}^2)\| = \|\mathrm{d}T(\boldsymbol{x})^{-1}(r(\boldsymbol{x}, \boldsymbol{h}^1) - \boldsymbol{r}(\boldsymbol{x}, \boldsymbol{h}^2))\| \leq M\, \eta\, \|\boldsymbol{h}^1 - \boldsymbol{h}^2\|,$$

Choosing η such that $M\, \eta < 1$, the function g is a contraction over $[-r, r]^n$. Therefore, using the Contraction Theorem, there exists a unique $\boldsymbol{h} \in [-r, r]^n$ such that $g(\boldsymbol{h}) = \boldsymbol{h}$. This implies that (3.83) has a unique solution $\boldsymbol{h} \in [-r, r]^n$.

We are now able to prove the following important generalization of Theorem 3.58.

Theorem 3.60 *Let $\boldsymbol{\Omega}$ be an open subset of $\mathbb{R}^n$. Let T be a change of variable of domain $\boldsymbol{\Omega}$. Let $\boldsymbol{A}$ be a measurable subset of $\boldsymbol{\Omega}$. Then, $T(\boldsymbol{A})$ is measurable and*

$$\mu_n(T(\boldsymbol{A})) = \int_{\boldsymbol{A}} |\det J_T(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x}. \tag{3.85}$$

Proof The first statement follows from Lemma 3.13. We also observe that the mapping $\boldsymbol{x} \mapsto |\det J_T(\boldsymbol{x})|$ is continuous to $\boldsymbol{A}$ and has positive values. Hence, by Theorem 3.32, the integral in (3.85) is well defined. We start proving the identity in the case where $\boldsymbol{A}$ is open and its closure $\overline{\boldsymbol{A}}$ is compact in $\boldsymbol{\Omega}$. We define $\boldsymbol{I}^{rk}$ (with $r \in \mathbb{Z}^n$ and $k \in \mathbb{N}$) and $\boldsymbol{A}^k$ as in the proof of Theorem 3.58. Let $\varepsilon \in \left(0, \frac{1}{2}\right)$. We choose k so that $2^{-1-k} \leq r_0$, where r_0 is the value that appears in the statement of Lemma 3.14. Let $r \in \mathbb{Z}^n$ and $k \in \mathbb{N}$ be such that $\boldsymbol{I}^{rk} \subseteq \boldsymbol{A}$. We have $\boldsymbol{I}^{rk} = z^{rk} + [-2^{-k-1}, 2^{-k-1})^n$, where $z^{rk} = \left(\left(r_1 + \frac{1}{2}\right) 2^{-k}, \ldots, \left(r_n + \frac{1}{2}\right) 2^{-k}\right)$. From Lemma 3.14

$$\begin{aligned} T(z^{rk}) + \mathrm{d}T(z^{rk})([-2^{-k-1}(1 - 2\,\varepsilon), 2^{-k-1}(1 - 2\,\varepsilon)]^n) &\subseteq T(\boldsymbol{I}^{rk}) \\ &\subseteq T(z^{rk}) + \mathrm{d}T(z^{rk})([-2^{-k-1}(1 + \varepsilon), 2^{-k-1}(1 + \varepsilon)]^n), \end{aligned}$$

and applying Lemma 3.11 and Theorems 3.58 and 3.59

$$|\det J_T(z^{rk})|\, \mu_n(\boldsymbol{I}^{rk})(1 - 2\,\varepsilon)^n \leq \mu_n(T(\boldsymbol{I}^{rk})) \leq |\det J_T(z^{rk})|\, \mu_n(\boldsymbol{I}^{rk})(1 + \varepsilon)^n.$$

Since the closure of $\boldsymbol{A}$ is compact in $\boldsymbol{\Omega}$, we obtain

$$\mu_n(T(\boldsymbol{A})) \leq +\infty, \qquad \int_{\boldsymbol{A}} |\det J_T(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} < +\infty.$$

Moreover,

$$\begin{aligned}\left|\int_{A^k} |\det J_T(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \mu_n(T(\boldsymbol{A}^k))\right| &= \left|\sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}}\left(\int_{\boldsymbol{I}^{rk}} |\det J_T(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \mu_n(T(\boldsymbol{I}^{rk}))\right)\right| \\ &= \left|\sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}}\left(\int_{\boldsymbol{I}^{rk}} (|\det J_T(\boldsymbol{x})| - |\det J_T(\boldsymbol{z}^{rk})|)\,\mathrm{d}\boldsymbol{x}\right)\right. \\ &\quad \left. + \sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}}\left(|\det J_T(\boldsymbol{z}^{rk})|\,\mu_n(\boldsymbol{I}^{rk}) - \mu_n(T(\boldsymbol{I}^{rk}))\right)\right| \\ &\leq \sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}}\left(\int_{\boldsymbol{I}^{rk}} \left||\det J_T(\boldsymbol{x})| - |\det \boldsymbol{J}_T(\boldsymbol{z}^{rk})|\right|\right)\mathrm{d}\boldsymbol{x} \\ &\quad + \sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}} |\det J_T(\boldsymbol{z}^{rk})|\,\mu_n(\boldsymbol{I}^{rk})\,\max\{(1+\varepsilon)^n - 1, 1 + (1-2\,\varepsilon)^2\}.\end{aligned}$$

The uniform continuity of the function $\boldsymbol{x}\mapsto|\det J_T(\boldsymbol{x})|$ over $\boldsymbol{A}$, taking a sufficiently large value of k, we can assume that $\left||\det J_T(\boldsymbol{x})| - |\det \boldsymbol{J}_T(\boldsymbol{z}^{rk})|\right| \leq \varepsilon$ for every $\boldsymbol{x}\in\boldsymbol{I}^{rk}$. Let $M>0$ be such that $|\det J_T(\boldsymbol{x})|\leq M$ for every $\boldsymbol{x}\in\boldsymbol{A}$. Then,

$$\begin{aligned}\left|\int_{A^k} |\det(\boldsymbol{J}_T(\boldsymbol{x}))|\,\mathrm{d}\boldsymbol{x} - \mu_n(T(\boldsymbol{A}^k))\right| &\leq (\varepsilon + M\,\max\{(1+\varepsilon)^n - 1, 1-(1-2\,\varepsilon)^n\})\sum_{\substack{\boldsymbol{r}\in\mathbb{Z}^n\\ \boldsymbol{I}^{rk}\subseteq \boldsymbol{A}}}\mu_n(\boldsymbol{I}^{rk}) \\ &\leq (\varepsilon + M\,\max\{(1+\varepsilon)^n - 1, 1-(1-2\,\varepsilon)^n\})\,\mu_n(\boldsymbol{A})\end{aligned}$$

if the value of k is sufficiently large. We conclude that

$$\lim_{k\to\infty}\left(\int_{A^k} |\det J_T(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \mu_n(T(\boldsymbol{A}^k))\right) = 0.$$

Moreover,

$$\lim_{k\to\infty}\mu_n(T(\boldsymbol{A}^k)) = \mu_n\Big(\bigcup_{k=0}^{\infty} T(\boldsymbol{A}^k)\Big) = \mu_n(T(\boldsymbol{A}))$$

and

$$\lim_{k\to\infty}\int_{A^k} |\det J_T(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty}\int_{\boldsymbol{A}} |\det J_T(\boldsymbol{x})|\,\chi_{A^k}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} |\det J_T(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}$$

from Theorem 3.33. Hence, we proved the result in the case where $\boldsymbol{A}$ is open and its closure is compact in $\boldsymbol{\Omega}$. We now suppose that A is measurable and $\boldsymbol{A}\subseteq\boldsymbol{O}$, where $\boldsymbol{O}$ is open and its closure is compact in $\boldsymbol{\Omega}$. We can construct a sequence of open sets $(\boldsymbol{A}^k)_{k\in\mathbb{N}}$ such that for every $k\in\mathbb{N}$, $\boldsymbol{A}\subseteq\boldsymbol{A}^{k+1}\subseteq\boldsymbol{A}^k\subseteq\boldsymbol{O}$ and $\lim_{k\to\infty}\mu_n(\boldsymbol{A}^k) = \mu_n(\boldsymbol{A})$ (we notice that in the case that we are considering $\mu_n(\boldsymbol{A})<+\infty$). This can be accomplished by using *(viii)* of the proof of Theorem 3.53, by taking $\boldsymbol{I}^{kr}$ to be all open. Let $C = \bigcap_{k=0}^{\infty}\boldsymbol{A}^k$. Then, $\boldsymbol{A}\subseteq C$ and

$\mu_n(C \setminus A) = 0$. From Remark 3.64, we find $\mu_n(T(C \setminus A)) = 0$. Thus,

$$\begin{aligned}\mu_n(T(A)) &= \mu_n(T(C)) = \mu_n\Big(\bigcap_{k=0}^{\infty} T(A^k)\Big)\\ &= \lim_{k\to\infty} \mu_n(T(A^k)) = \lim_{k\to\infty} \int_{A^k} |\det J_T(x)|\,\mathrm{d}x\\ &= \lim_{k\to\infty} \int_A |\det J_T(x)|\,\chi_{A^k}(x)\,dx = \int_O |\det J_T(x)|\,\chi_C(x)\,dx\end{aligned}$$

by Theorem 3.48. Finally, we consider the general case. Let, for every $k \in \mathbb{N}$,

$$\Omega^k = \left\{ x \in \Omega \,\middle|\, \mathrm{d}(x, \partial\Omega) > \frac{1}{k},\ \ \|x\| < k \right\}.$$

Then, Ω^k is an open set with compact closure in Ω for every $k \in \mathbb{N}$ and $\Omega^k \subseteq \Omega^{k+1}$ and $\bigcup_{k=0}^{\infty} \Omega^k = \Omega$. In view of the previous argument, we have

$$\begin{aligned}\mu_n(T(A)) &= \lim_{k\to\infty} \mu_n(T(A \cap \Omega^k))\\ &= \lim_{k\to\infty} \int_{A\cap\Omega^k} |\det J_T(x)|\,\mathrm{d}x\\ &= \lim_{k\to\infty} \int_A |\det J_T(x)|\,\chi_{A\cap\Omega^k}(x)\,\mathrm{d}x = \int_A |\det J_T(x)|\,\mathrm{d}x\end{aligned}$$

by the monotone convergence theorem (Theorem 3.34).

Theorem 3.61 [Change of Variable]
Let Ω be an open subset of $\mathbb{R}^n$. Let $T : \Omega \longrightarrow \mathbb{R}^n$ be a change of variable. Let A be a measurable subset of $T(\Omega)$. Then:

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

(b) *Let $f : A \longrightarrow [0, +\infty]$ be measurable. Then, $x \mapsto f(T(x))\,|\det J_T(x)|$ is measurable over $T^{-1}(A)$. In such case,*

$$\int_A f(y)\,\mathrm{d}y = \int_{T^{-1}(A)} f(T(x))\,|\det J_T(x)|\,\mathrm{d}x; \tag{3.86}$$

(c) *Let f be summable over A. Then, $x \mapsto f(T(x))\,|\det J_T(x)|$ is summable over $T^{-1}(A)$ and (3.86) holds.*

Proof We have already observed that T^{-1} is a change of variable. Then, **(a)** follows from Lemma 3.13. Consider the case $f = \chi_B$, where $B \subseteq A$ and measurable. Thus, $f(T(x)) = \chi_{T^{-1}(B)}(x)$ for every $x \in T^{-1}(A)$. It follows that $x \mapsto f(T(x))\,|\det J_T(x)|$ is measurable over $T^{-1}(A)$. Moreover, using Theorem 3.60,

$$\begin{aligned}\int_A f(y)\,\mathrm{d}y &= \mu_n(B) = \int_{T^{-1}(B)} |\det J_T(x)|\,\mathrm{d}x\\ &= \int_{T^{-1}(A)} \chi_{T^{-1}(B)}(x)\ \det J_T(x)\,|\mathrm{d}x\\ &= \int_{T^{-1}(A)} f(T(x))\,|\det J_T(x)|\,\mathrm{d}x.\end{aligned}$$

This proves **(b)**. Point **(c)** can be proved from **(b)** by decomposing f in the form $f^+ - f^-$ in the case where f is real-valued, and in the form $\Re\mathrm{e}\, f + i\,\Im\mathrm{m}\, f$ in the case where it is complex-valued.

Remark 3.26 It is easy to see that the first conclusion of point **(b)** in Theorem 3.61 (i.e., the measurability) holds without assuming that f takes values in $[0, +\infty]$.

Example *3.48*

Consider the case $n = 1$. Let Ω be an open interval in $\mathbb{R}$ and $T : \Omega \longrightarrow \mathbb{R}$ be of class C^1, with $T'(x) \neq 0$ for every $x \in \Omega$. Then, T' has constant sign over Ω and T is a change of coordinates. Let $[a, b] \subseteq T(\Omega)$, where $a, b \in \mathbb{R}$ and $a < b$. Let $f : [a, b] \longrightarrow \mathbb{R}$ be Riemann integrable. Then, by Theorem 3.51, using the notation introduced when comparing Lebesgue and Riemann integrals, we can write

$$\int_a^b f(y)\,\mathrm{d}y = \int_{T^{-1}([a,b])} f(T(x))\,|T'(x)|\,\mathrm{d}x. \tag{3.87}$$

If $T'(x) > 0$ for every $x \in \Omega$, then T is increasing, and therefore $T^{-1}([a, b]) = [T^{-1}(a), T^{-1}(b)]$. From (3.87) and Theorem 3.51, it follows that

$$\int_a^b f(y)\,\mathrm{d}y = \int_{[T^{-1}(a),T^{-1}(b)]} f(T(x))\,|T'(x)|\,\mathrm{d}x = \int_{T^{-1}(a)}^{T^{-1}(b)} f(T(x))\,|T'(x)|\,\mathrm{d}x.$$

If $T'(x) < 0$ for every $x \in \Omega$, then T is decreasing, and therefore $T^{-1}([a, b]) = [T^{-1}(b), T^{-1}(a)]$. Hence,

$$\int_a^b f(y)\,\mathrm{d}y = -\int_{[T^{-1}(b),T^{-1}(a)]} f(T(x))\,|T'(x)|\,\mathrm{d}x = \int_{T^{-1}(a)}^{T^{-1}(b)} f(T(x))\,|T'(x)|\,\mathrm{d}x,$$

considered as a Riemann integral over an oriented interval. We have re-obtained in this case the well-known change of variable formula of the one-dimensional Riemann integral.

Example *3.49*

Consider the ball $\mathcal{B}_1(\mathbf{0})$ in $\mathbb{R}^n$. This ball has finite measure, since

$$\mathcal{B}_1(\mathbf{0}) \subseteq (-1, 1)^n \quad \Rightarrow \quad \mu_n(\mathcal{B}_1(\mathbf{0})) \leq 2^n.$$

Let

$$\omega_n = \mu_n(\mathcal{B}_1(\mathbf{0})).$$

Given $\boldsymbol{x}_0 \in \mathbb{R}^n$ and $r > 0$. We want to prove that

$$\mu_n(\mathcal{B}_r(\boldsymbol{x}_0)) = \omega_n\, r^n.$$

Notice that this is consistent with the well-known results for $n = 1, 2, 3$ (why?). We define

$$\begin{aligned} T : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{x} &\mapsto r\,\boldsymbol{x} + \boldsymbol{x}_0. \end{aligned}$$

The function T is of class C^1. It is easy to see that it is injective: if $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n$ satisfy $T(\boldsymbol{x}_1) = T(\boldsymbol{x}_2)$, then $r\,\boldsymbol{x}_1 + \boldsymbol{x}_0 = r\,\boldsymbol{x}_2 + \boldsymbol{x}_0$, which implies $r\,(\boldsymbol{x}_1 - \boldsymbol{x}_2) = 0$, and since $r > 0$ we obtain $\boldsymbol{x}_1 = \boldsymbol{x}_2$. Moreover, the image is $\mathbb{R}^n$. Indeed, given $\boldsymbol{y} \in \mathbb{R}^n$, defining $\boldsymbol{x} = r^{-1}\,(\boldsymbol{y} - \boldsymbol{x}_0)$ gives $T(\boldsymbol{x}) = \boldsymbol{y}$. We now compute the Jacobian matrix of T at a generic point $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. Writing $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2}, \ldots, x_{0,n})$, we have

$$T(\boldsymbol{x}) = T(x_1, x_2, \ldots, x_n) = r\,\boldsymbol{x} + \boldsymbol{x}_0 = r\,(x_1, x_2, \ldots, x_n) + (x_{0,1}, x_{0,2}, \ldots, x_{0,n}) = (r\,x_1 + x_{0,1}, r\,x_2 + x_{0,2}, \ldots, r\,x_n + x_{0,n}),$$

so that

$$\begin{aligned}
\frac{\partial T_1}{\partial x_1}(\boldsymbol{x}) &= r, \quad \frac{\partial T_1}{\partial x_2}(\boldsymbol{x}) = 0, \quad \ldots, \frac{\partial T_1}{\partial x_n}(\boldsymbol{x}) = 0,\\
\frac{\partial T_2}{\partial x_1}(\boldsymbol{x}) &= 0, \quad \frac{\partial T_2}{\partial x_2}(\boldsymbol{x}) = r, \quad \ldots, \frac{\partial T_2}{\partial x_n}(\boldsymbol{x}) = 0,\\
&\vdots\\
\frac{\partial T_n}{\partial x_1}(\boldsymbol{x}) &= 0, \quad \frac{\partial T_n}{\partial x_2}(\boldsymbol{x}) = 0, \quad \ldots, \frac{\partial T_n}{\partial x_n}(\boldsymbol{x}) = r,
\end{aligned}$$

so that the Jacobian matrix is diagonal:

$$\boldsymbol{J}_T(\boldsymbol{x}) = \begin{bmatrix} r & 0 & \ldots & 0 \\ 0 & r & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r \end{bmatrix},$$

and its determinant is $r^n > 0$. We want to compute

$$\mu_n(\mathcal{B}_r(\boldsymbol{x}_0)) = \int_{\|\boldsymbol{y}-\boldsymbol{x}_0\|<r} 1 \, \mathrm{d}\boldsymbol{y},$$

and to do this we use the change of variable theorem:

$$\int_{\boldsymbol{A}} f(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} = \int_{T^{-1}(\boldsymbol{A})} f(T(\boldsymbol{x})) \, |\det J_T(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}. \tag{3.88}$$

In the present case, $\boldsymbol{A} = \mathcal{B}_r(\boldsymbol{x}_0) = \{\boldsymbol{y} \in \mathbb{R}^n \mid \|\boldsymbol{y} - \boldsymbol{x}_0\| < r\}$, and we need to determine $T^{-1}(A)$. We find:

$$\begin{aligned}
T^{-1}(\boldsymbol{A}) &= \{\boldsymbol{x} \in \mathbb{R}^n \mid T(\boldsymbol{x}) \in \boldsymbol{A}\}\\
&= \{\boldsymbol{x} \in \mathbb{R}^n \mid r\,\boldsymbol{x} + \boldsymbol{x}_0 \in \boldsymbol{A}\}\\
&= \{\boldsymbol{x} \in \mathbb{R}^n \mid \|(r\,\boldsymbol{x} + \boldsymbol{x}_0) - \boldsymbol{x}_0\| < r\}\\
&= \{\boldsymbol{x} \in \mathbb{R}^n \mid \|r\,\boldsymbol{x}\| < r\}\\
&= \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| < 1\} = \mathcal{B}_1(0).
\end{aligned}$$

Hence, (3.88) becomes

$$\int_{\|\boldsymbol{y}-\boldsymbol{x}_0\|<r} 1 \, \mathrm{d}y = \int_{\|\boldsymbol{x}\|<1} 1 \, r^n \, \mathrm{d}\boldsymbol{x} = r^n \int_{\|\boldsymbol{x}\|<1} 1 \, \mathrm{d}\boldsymbol{x} = r^n \, \mu_n(\mathcal{B}_1(0)) = r^n \, \omega_n.$$

We have observed in Remark 3.25 that the inverse T^{-1} of a change of variable is itself a change of variable. In fact, $T(\boldsymbol{\Omega})$ is open, T^{-1} is of class C^1 and $\det J_{T^{-1}}(\boldsymbol{y}) \neq 0$ for all $\boldsymbol{y} \in T(\boldsymbol{\Omega})$ using the inverse function theorem.[8]

[8] The inverse function theorem in differential calculus provides sufficient conditions for a function to admit a local inverse, i.e., for a function to be invertible in a suitable neighborhood of a point of its domain. Let $\boldsymbol{A}$ be an open subset of $\mathbb{R}^n$, and let $\boldsymbol{x}_0 \in \boldsymbol{A}$. Let $f : \boldsymbol{\Omega} \longrightarrow \mathbb{R}^n$ is a function of class C^1 such that the determinant of the Jacobian at $\boldsymbol{x}_0$ is non-zero, i.e., $J_f(\boldsymbol{x}_0) \neq 0$, then there exists a neighborhood $\boldsymbol{U}$ of $\boldsymbol{x}_0$ such that the restriction of f to $\boldsymbol{U}$, that we denote by $F : \boldsymbol{U} \longrightarrow F(\boldsymbol{U})$, is invertible. Its inverse F^{-1} is of class C^1. Moreover, for every $\boldsymbol{y} \in F(\boldsymbol{U})$ there holds

$$J_{T^{-1}}(\boldsymbol{y}) = J_F^{-1}(F^{-1}(\boldsymbol{y})).$$

We often encounter integrals in the form

$$\int_{\boldsymbol{A}} f(S(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y},$$

where S is a change of variable. Defining $T = S^{-1}$ (which, as aforementioned, is itself a change of variable), we have

$$\begin{aligned}\int_{\boldsymbol{A}} f(S(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} &= \int_{S(\boldsymbol{A})} f(S(S^{-1}(\boldsymbol{x})))\,|\det J_{S^{-1}}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\\ &= \int_{S(\boldsymbol{A})} f(\boldsymbol{x})\,|\det J_S^{-1}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{3.89}$$

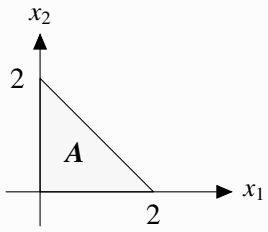
Example *3.50*

We want to evaluate the integral

$$I = \iint_{\boldsymbol{A}} e^{\frac{x_2-x_1}{x_1+x_2}}\,\mathrm{d}x_1\,\mathrm{d}x_2,$$

where

$$\boldsymbol{A} = \{(x_1, x_2) \in \mathbb{R}^2 \,|\, x_1 > 0 \text{ and } x_2 > 0 \text{ and } x_1 + x_2 < 2\}$$



The set $\boldsymbol{A}$ is measurable because it is open. The integrand is measurable over $\boldsymbol{A}$ because it is continuous, and it is non-negative. Therefore, the integral is defined (although it may be equal to $+\infty$).

We would like to write the integral as

$$\int_{\boldsymbol{A}} f(S(x_1, x_2))\,\mathrm{d}x_1\,\mathrm{d}x_2, \quad \text{with} \quad f(y_1, y_2) = e^{\frac{y_1}{y_2}}.$$

To this end, we can use the transformation

$$\begin{aligned} S : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2\\ (x_1, x_2) &\mapsto (x_2 - x_1, x_1 + x_2), \end{aligned}$$

which is linear, and the determinant of its Jacobian matrix is

$$\det J_S(x_1, x_2) = \det \begin{bmatrix} \dfrac{\partial S_1}{\partial x_1}(x_1, x_2) & \dfrac{\partial S_1}{\partial x_2}(x_1, x_2) \\ \dfrac{\partial S_2}{\partial x_1}(x_1, x_2) & \dfrac{\partial S_2}{\partial x_2}(x_1, x_2) \end{bmatrix} = \det \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = -1 - 1 = -2.$$

Applying (3.89) we find

$$I = \int_{\boldsymbol{A}} f(S(x_1, x_2))\,\mathrm{d}x_1\,\mathrm{d}x_2 = \frac{1}{2}\int_{S(\boldsymbol{A})} e^{\frac{y_1}{y_2}}\,\mathrm{d}y_1\,\mathrm{d}y_2.$$

We must determine $S(\boldsymbol{A})$. Take an element $(y_1, y_2) \in S(\boldsymbol{A})$. Hence, there exists $(x_1, x_2) \in \boldsymbol{A}$ such that $(y_1, y_2) = S(x_1, x_2)$. From the definition of S, we find $(y_1, y_2) = (x_2 - x_1, x_1 + x_2)$, and

$$\begin{cases} y_1 = x_2 - x_1 \\ y_2 = x_1 + x_2 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x_1 = \frac{y_2 - y_1}{2} \\ x_2 = \frac{y_1 + y_2}{2}. \end{cases}$$

Since $(x_1, x_2) \in \boldsymbol{A}$, the components x_1 and x_2 satisfy $x_1 > 0$, $x_2 > 0$ and $x_1 + x_2 < 2$, so that y_1 and y_2 must satisfy

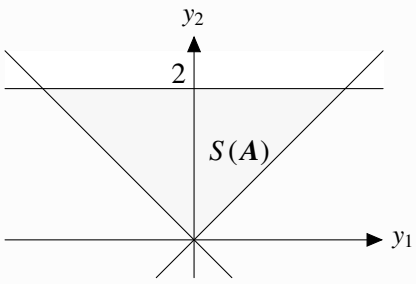
$$\frac{y_2 - y_1}{2} > 0, \quad \frac{y_1 + y_2}{2} > 0, \quad \frac{y_2 - y_1}{2} + \frac{y_1 + y_2}{2} < 2,$$

i.e.,

$$y_2 > y_1, \quad y_2 > -y_1, \quad y_2 < 2.$$

It follows that

$$S(\boldsymbol{A}) = \{(y_1, y_2) \in \mathbb{R}^2 \,|\, 2 > y_2 > |y_1|\}.$$



We can apply Tonelli's theorem, first with respect to y_1 (in other words, we fix y_2 instead of fixing y_1). We get

$$I = \frac{1}{2} \int_{(0,2)} \left(\int_{(-y_2, y_2)} e^{\frac{y_1}{y_2}} \, \mathrm{d}y_1 \right) \mathrm{d}y_2.$$

Since $\mu_1(\{-y_2, y_2\}) = 0$, we have, for all $y_2 \in (0, 2)$,

$$\int_{(-y_2, y_2)} e^{\frac{y_1}{y_2}} \, \mathrm{d}y_1 = \int_{[-y_2, y_2]} e^{\frac{y_1}{y_2}} \, \mathrm{d}y_1 = \int_{-y_2}^{y_2} e^{\frac{y_1}{y_2}} \, \mathrm{d}y_1 = \left[y_2 \, e^{\frac{y_1}{y_2}} \right]_{-y_2}^{y_2} = (e - e^{-1}) \, y_2.$$

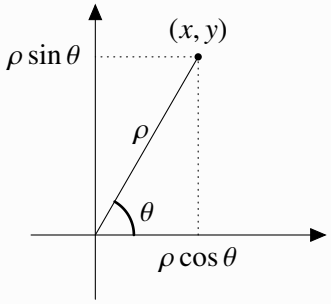
We conclude that

$$I = \frac{1}{2} \int_{(0,2)} (e - e^{-1}) \, y_2 \, \mathrm{d}y_2 = \frac{e - e^{-1}}{2} \int_0^2 y_2 \, \mathrm{d}y_2 = e - e^{-1}.$$

Example *3.51*

We introduce the change of variable into polar coordinates. Consider the equations in $\mathbb{R}^2$

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta \end{cases}$$



where $(\rho, \theta) \in \boldsymbol{\Omega} = (0, +\infty) \times (0, 2\pi)$, which is an open set. Let

$$\begin{aligned} T : \boldsymbol{\Omega} &\longrightarrow \mathbb{R}^2 \\ (\rho, \theta) &\mapsto (\rho \cos\theta, \rho \sin\theta). \end{aligned}$$

Clearly, T is of class C^1. It is also injective and its image is

$$T(\boldsymbol{\Omega}) = \mathbb{R}^2 \setminus \{(x, 0) \in \mathbb{R}^2 \,|\, x \geq 0\}.$$

Moreover, for all $(\rho,\theta)\in\mathbf{\Omega}$,

$$\det J_T(\rho,\theta)=\det\begin{bmatrix}\dfrac{\partial T_1}{\partial\rho}(\rho,\theta) & \dfrac{\partial T_1}{\partial\theta}(\rho,\theta)\\ \dfrac{\partial T_2}{\partial\rho}(\rho,\theta) & \dfrac{\partial T_2}{\partial\theta}(\rho,\theta)\end{bmatrix}=\det\begin{bmatrix}\cos\theta & -\rho\sin\theta\\ \sin\theta & \rho\cos\theta\end{bmatrix}=\rho>0.$$

It follows that T is a change of variable.

We now want to use this change of variable to compute the integral

$$\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

where $\boldsymbol{A}$ is measurable in $\mathbb{R}^2$ and f is measurable over $\boldsymbol{A}$. In order to use directly the result of the change of variable theorem (Theorem 3.61), we must assume $\boldsymbol{A}\subseteq T(\mathbf{\Omega})$ (which means, geometrically, that $\boldsymbol{A}$ does not intersect the positive x-axis). In this case

$$\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_{T^{-1}(\boldsymbol{A})} f(\rho\cos\theta,\rho\sin\theta)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta.$$

However, requiring $\boldsymbol{A}\subseteq T(\mathbf{\Omega})$ can be very restrictive. We show that if this assumption does not hold, we can still use the change of variable. Suppose that $\boldsymbol{A}\not\subseteq T(\mathbf{\Omega})$. We write $\boldsymbol{A}$ as the union of two disjoint sets:

$$\boldsymbol{A}=(\boldsymbol{A}\cap T(\mathbf{\Omega}))\cup(\boldsymbol{A}\cap T(\mathbf{\Omega})^{\complement}).$$

The set $\boldsymbol{A}\cap T(\mathbf{\Omega})^{\complement}$ has measure zero because $T(\mathbf{\Omega})^{\complement}$ has measure zero. It follows that

$$\int_{\boldsymbol{A}\cap T(\mathbf{\Omega})^{\complement}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y=0.$$

Now, we can use the change of variable theorem (Theorem 3.61):

$$\int_{\boldsymbol{A}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\int_{\boldsymbol{A}\cap T(\mathbf{\Omega})} f(x,y)\,\mathrm{d}x\,\mathrm{d}y+\underbrace{\int_{\boldsymbol{A}\cap T(\mathbf{\Omega})^{\complement}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y}_{=0}=\int_{T^{-1}(\boldsymbol{A}\cap T(\mathbf{\Omega}))} f(\rho\,\cos\theta,\rho\,\sin\theta)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta.$$

It is not difficult to see that

$$T^{-1}(\boldsymbol{A}\cap T(\mathbf{\Omega}))=\{(\rho,\theta)\in\mathbf{\Omega}\,|\,T(\rho,\theta)\in\boldsymbol{A}\}. \tag{3.90}$$

In fact, suppose that $(\rho,\theta)\in T^{-1}(\boldsymbol{A}\cap T(\mathbf{\Omega}))$, and we show that $(\rho,\theta)\in\{(\rho,\theta)\in\mathbf{\Omega}\,|\,T(\rho,\theta)\in\boldsymbol{A}\}$. Since $(\rho,\theta)\in T^{-1}(\boldsymbol{A}\cap T(\mathbf{\Omega}))$, by definition of inverse image there exists $(x,y)\in\boldsymbol{A}\cap T(\mathbf{\Omega})$ such that $T(\rho,\theta)=(x,y)$. Since $(x,y)\in\boldsymbol{A}$, this equation says that $T(\rho,\theta)\in\boldsymbol{A}$. Moreover, since $(x,y)\in\mathbf{\Omega}$, the same equation can be written as $T(\rho,\theta)\in T(\mathbf{\Omega})$, which gives $(\rho,\theta)\in\mathbf{\Omega}$. It follows that $(\rho,\theta)\in\{(\rho,\theta)\in\mathbf{\Omega}\,|\,T(\rho,\theta)\in\boldsymbol{A}\}$.
Now we prove the opposite inclusion. Suppose that $(\rho,\theta)\in\{(\rho,\theta)\in\mathbf{\Omega}\,|\,T(\rho,\theta)\in\boldsymbol{A}\}$, and we show that $(\rho,\theta)\in T^{-1}(\boldsymbol{A}\cap T(\mathbf{\Omega}))$, which is the same (by definition of inverse image of a function) to $T(\rho,\theta)\in\boldsymbol{A}\cap T(\mathbf{\Omega})$. This inclusion holds, because $(\rho,\theta)\in\mathbf{\Omega}$, which implies $T(\rho,\theta)\in T(\mathbf{\Omega})$, and $T(\rho,\theta)\in\boldsymbol{A}$. Thus, $T(\rho,\theta)\in\boldsymbol{A}\cap T(\mathbf{\Omega})$.

We now consider the specific problem of finding, again, $\mu_2(\mathcal{B}_r((0,0)))$ with $r>0$. We have

$$\mu_2(\mathcal{B}_r((0,0)))=\int_{\mathcal{B}_r((0,0))}1\,\mathrm{d}x\,\mathrm{d}y=\int_{\|(x,y)\|<r}1\,\mathrm{d}x\,\mathrm{d}y=\int_{\sqrt{x^2+y^2}<r}1\,\mathrm{d}x\,\mathrm{d}y=\int_{x^2+y^2<r^2}1\,\mathrm{d}x\,\mathrm{d}y.$$

Notice that $\boldsymbol{A} = \mathcal{B}_r((0,0))$ is not contained in $T(\boldsymbol{\Omega})$. We have

$$\mathcal{B}_r((0,0)) \cap T(\boldsymbol{\Omega}) = \{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 < r^2 \quad \text{and} \quad (y \neq 0 \text{ if } x \geq 0)\},$$

and by (3.90) we get

$$\begin{aligned} T^{-1}\big(\mathcal{B}_r((0,0)) \cap T(\boldsymbol{\Omega})\big) &= \{(\rho,\theta) \in (0,+\infty) \times (0,2\pi) \,|\, T(\rho,\theta) \in \boldsymbol{A}\} \\ &= \{(\rho,\theta) \in (0,+\infty) \times (0,2\pi) \,|\, (\rho\,\cos\theta, \rho\,\sin\theta) \in \boldsymbol{A}\} \\ &= \{(\rho,\theta) \in (0,+\infty) \times (0,2\pi) \,|\, \rho^2\,\cos^2\theta + \rho^2\,\sin^2\,\theta < r^2\} \\ &= \{(\rho,\theta) \in (0,+\infty) \times (0,2\pi) \,|\, \rho^2 < r^2\} = (0,r) \times (0,2\pi). \end{aligned}$$

It follows that

$$\mu_2(\mathcal{B}_r((0,0))) = \int_{(0,r)\times(0,2\pi)} \rho \,\mathrm{d}\rho\,\mathrm{d}\theta.$$

We can apply Tonelli's theorem:

$$\mu_2(\mathcal{B}_r((0,0))) = \int_{(0,r)} \left(\int_{(0,2\pi)} \rho \,\mathrm{d}\theta \right) \mathrm{d}\rho = \int_{(0,r)} 2\,\pi\,\rho \,\mathrm{d}\rho = \int_0^r 2\,\pi\,\rho \,\mathrm{d}\rho = \left[2\,\pi\,\frac{\rho^2}{2} \right]_0^r = \pi\,r^2.$$

Remark 3.27 In the previous exercise, we claimed that

$$\begin{aligned} T : \boldsymbol{\Omega} &\longrightarrow \mathbb{R}^2 \\ (\rho,\theta) &\mapsto (\rho\,\cos\theta, \rho\,\sin\theta). \end{aligned}$$

is injective, and that its image is $\mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2 \mid x \geq 0\}$. To see that T is injective, suppose for $\rho_1, \theta_1, \rho_2, \theta_2$ we have $T(\rho_1,\theta_1) = T(\rho_2,\theta_2)$. We have to show that $\rho_1 = \rho_2$ and $\theta_1 = \theta_2$. The identity $T(\rho_1,\theta_1) = T(\rho_2,\theta_2)$ leads immediately to $(\rho_1\,\cos\theta_1, \rho_1\,\sin\theta_1) = (\rho_2\,\cos\theta_2, \rho_2\,\sin\theta_2)$, which can be written as the pair of equations

$$\begin{cases} \rho_1\,\cos\theta_1 = \rho_2\,\cos\theta_2 \\ \rho_1\,\sin\theta_1 = \rho_2\,\sin\theta_2 \end{cases}$$

Since $\rho_2 \neq 0$ we find

$$\frac{\rho_1}{\rho_2}\,\cos\theta_1 = \cos\theta_2, \qquad \frac{\rho_1}{\rho_2}\,\sin\theta_1 = \sin\theta_2.$$

Taking the square of both equations and adding the result gives

$$\left(\frac{\rho_1}{\rho_2}\right)^2 = 1,$$

which is satisfied if $\rho_1 = \rho_2$ or $\rho_1 = -\rho_2$. However, since both of them are positive, we must have $\rho_1 = \rho_2$. Thus, the equations become

$$\begin{cases} \cos\theta_1 = \cos\theta_2 \\ \sin\theta_1 = \sin\theta_2 \end{cases}$$

Two angles in a period (e.g. between 0 and $2\,\pi$) which have the same sine and the same cosine are equal, so that $\theta_1 = \theta_2$).

We now show that $T(\boldsymbol{\Omega}) = \mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2 \mid x \geq 0\}$. To this end, we study the solvability in $\rho \in (0,+\infty)$ and $\theta \in (0,2\pi)$ of the equations

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta. \end{cases}$$

We find immediately $x^2+y^2 = \rho^2$, so that $\rho = \sqrt{x^2+y^2}$ (we discard the negative solution because $\rho \in (0,+\infty)$). We must have $(x,y) \neq (0,0)$ to ensure $\rho > 0$. Let us consider the two cases:

- if $x \neq 0$, we can divide the second equation by the first, to obtain

$$\frac{y}{x} = \tan\theta.$$

The tangent is invertible in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. We distinguish the cases:

  - if $x > 0$ and $y > 0$, we have
$$\theta = \arctan\frac{y}{x} \in \left(0, \frac{\pi}{2}\right).$$
  - if $x > 0$ and $y < 0$, we have
$$\theta = 2\pi + \arctan\frac{y}{x} \in \left(\frac{3\pi}{2}, 2\pi\right),$$
since $\arctan\frac{y}{x} \in \left(-\frac{\pi}{2}, 0\right)$;
  - if $x < 0$ and $y < 0$, then
$$\theta = \pi + \arctan\frac{y}{x} \in \left(\pi, \frac{3\pi}{2}\right);$$
  - if $x < 0$ and $y > 0$, then
$$\theta = \pi + \arctan\frac{y}{x} \in \left(\frac{\pi}{2}, \pi\right);$$
  - if $y = 0$, then $\sin\theta = 0$, and therefore $\theta = \pi$ (the solution $\theta = 0$ is not in the range $(0,2\pi)$), so that $x = -\rho < 0$. In this case $\theta = \arctan\frac{0}{x} + \pi = \pi$, and can be included in one of the previous two cases.

- if $x = 0$, we have $\cos\theta = 0$, i.e., $\theta = \frac{\pi}{2}$ or $\theta = \frac{3\pi}{2}$. In the first case, $y = \rho > 0$, and in the second $y = -\rho < 0$.

The case $x > 0$ and $y = 0$ has no solutions in $(\rho,\theta) \in (0,+\infty) \times (0,2\pi)$. In fact, $y = 0$ implies $\theta = \pi$ (because $\theta = 0$ is not in $(0,2\pi)$), and for $\theta = \pi$ we have $x = -\rho < 0$. To summarize, given $(x,y) \in \mathbb{R}^2 \setminus \{(x,0) \in \mathbb{R}^2 \mid x \geq 0\}$, we have

$$T^{-1}(x,y) = \begin{cases} \left(\sqrt{x^2+y^2}, \arctan\frac{y}{x}\right) & \text{if } x > 0 \text{ and } y > 0 \\ \left(\sqrt{x^2+y^2}, 2\pi + \arctan\frac{y}{x}\right) & \text{if } x > 0 \text{ and } y < 0 \\ \left(\sqrt{x^2+y^2}, \pi + \arctan\frac{y}{x}\right) & \text{if } x < 0 \\ \left(\sqrt{x^2+y^2}, \frac{\pi}{2}\right) & \text{if } x = 0 \text{ and } y > 0 \\ \left(\sqrt{x^2+y^2}, \frac{3}{2}\pi\right) & \text{if } x = 0 \text{ and } y < 0. \end{cases}$$

Example *3.52* Let

$$\boldsymbol{A} = \{(x,y) \in \mathbb{R}^2 \mid 9 < x^2+y^2 < 8\,y\}.$$

The set $\boldsymbol{A}$ is measurable because it is open. We compute $\mu_2(\boldsymbol{A})$. Using the polar coordinates in $\mathbb{R}^2$, we have

$$\mu_2(\boldsymbol{A}) = \int_{\boldsymbol{A}} 1 \,\mathrm{d}x\,\mathrm{d}y = \int_{T^{-1}(\boldsymbol{A})} \rho \,\mathrm{d}\rho\,\mathrm{d}\theta,$$

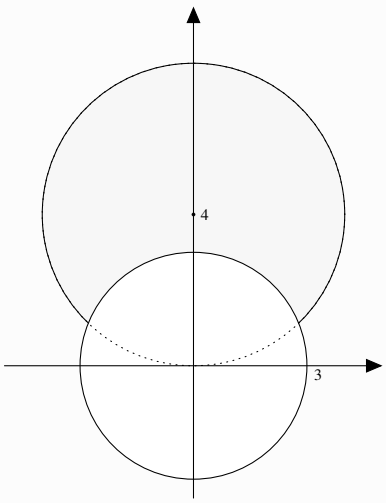
where $T^{-1}(\boldsymbol{A}) = \{(\rho,\theta) \in (0,+\infty)\times(0,2\pi) \,|\, 9 < \rho^2 < 8\rho\,\sin\theta\}$. Let $\boldsymbol{B} = T^{-1}(\boldsymbol{A})$. For every $\theta \in \mathbb{R}$ we have $B_\theta = \{\rho > 0 \,|\, (\rho,\theta) \in \boldsymbol{B}\}$, so that $\rho^2 > 9$ and $\rho\,(\rho - 8\,\sin\theta) < 0$. These two give $\rho > 3$ and $\rho < 8\,\sin\theta$, which lead to

$$B_\theta = \begin{cases} (3,\, 8\,\sin\theta) & \text{if } \theta \in (0,2\pi) \text{ and } 3 < 8\,\sin\theta \\ \varnothing & \text{otherwise} \end{cases}$$

Moreover, $3 < 8\,\sin\theta$ can be written as $\sin\theta > 3/8$, which is satisfied for $\theta \in \left(\arcsin\frac{3}{8}, \pi - \arcsin\frac{3}{8}\right)$. Using Tonelli's theorem, it follows that

$$\begin{aligned}
\mu_2(\boldsymbol{A}) &= \int_{\left(\arcsin\frac{3}{8},\pi-\arcsin\frac{3}{8}\right)} \left(\int_{(3,\,8\,\sin\theta)} \rho\,\mathrm{d}\rho\right)\mathrm{d}\theta \\
&= \int_{\left(\arcsin\frac{3}{8},\pi-\arcsin\frac{3}{8}\right)} \left(\int_{[3,\,8\,\sin\theta]} \rho\,\mathrm{d}\rho\right)\mathrm{d}\theta \\
&= \int_{\left(\arcsin\frac{3}{8},\pi-\arcsin\frac{3}{8}\right)} \left(\int_{3}^{8\,\sin\theta} \rho\,\mathrm{d}\rho\right)\mathrm{d}\theta \\
&= \int_{\left(\arcsin\frac{3}{8},\pi-\arcsin\frac{3}{8}\right)} \frac{64\,\sin^2\theta - 9}{2}\,\mathrm{d}\theta \\
&= \int_{\arcsin\frac{3}{8}}^{\pi-\arcsin\frac{3}{8}} \frac{64\,\sin^2\theta - 9}{2}\,\mathrm{d}\theta = \frac{23}{2}\left(\pi - \arcsin\frac{3}{8}\right) + 3\,\sqrt{55}.
\end{aligned}$$

The equation $x^2+y^2 > 9$ represents all the points outside the circle centred at $(0,0)$ of radius 3. The equation $x^2 + y^2 - 8\,y = 0$ can be written as $x^2 + y^2 - 8\,y + 16 - 16 = 0$, i.e., $x^2 + (y-4)^2 = 16$. Thus, $x^2 + y^2 < 8\,y$ represents the points inside the circle centred at $(0,4)$ of radius 4:



Example *3.53* We compute the integral

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y.$$

First, we observe that the function from $\mathbb{R}^2$ to $\mathbb{R}$ defined by $(x, y) \mapsto e^{-(x^2+y^2)}$ is continuous and non-negative over $\mathbb{R}^2$, and this guarantees that the integral is well defined. We now use polar coordinates as before. This time, $\boldsymbol{A} = \mathbb{R}^2$ intersects the positive x-axis, and therefore

$$\int_{\mathbb{R}^2} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = \int_{\{(\rho,\theta)\in\Omega\,|\,T(\rho,\theta)\in A\}} e^{-\rho^2}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta = \int_{\{(\rho,\theta)\in(0,+\infty)\times(0,2\pi)\,|\,(\rho\cos\theta,\rho\sin\theta)\in\mathbb{R}^2\}} e^{-\rho^2}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta,$$

and since for every $(\rho, \theta) \in (0, +\infty) \times (0, 2\pi)$ we have $(\rho\cos\theta, \rho\sin\theta) \in \mathbb{R}^2$, we can re-write the integral as

$$\begin{aligned}
\int_{(0,+\infty)\times(0,2\pi)} e^{-\rho^2}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta &= \int_{(0,+\infty)} e^{-\rho^2}\,\rho\left(\int_{(0,2\pi)}\mathrm{d}\theta\right)\mathrm{d}\rho \\
&= 2\pi\int_{(0,+\infty)} e^{-\rho^2}\,\rho\,\mathrm{d}\rho \\
&= 2\pi\int_{[0,+\infty)} e^{-\rho^2}\,\rho\,\mathrm{d}\rho \qquad \text{since } \mu_1(\{0\}) = 0 \\
&= 2\pi\lim_{c\to+\infty}\int_0^c e^{-\rho^2}\,\rho\,\mathrm{d}\rho = \pi\lim_{c\to+\infty}(1-e^{-c^2}) = \pi,
\end{aligned}$$

where we have used Theorem 3.54 and Lemma 3.7. From Theorem 3.54 we obtain

$$\pi = \int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x\int_{\mathbb{R}} e^{-y^2}\,\mathrm{d}y.$$

We can conclude with the well-known formula

$$\int_{\mathbb{R}} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

Example *3.54*

The same considerations of the previous example can be applied to the case of cylindrical coordinates in $\mathbb{R}^3$. Let

$$\boldsymbol{\Omega} = (0, +\infty) \times (0, 2\pi) \times \mathbb{R},$$

and the function

$$\begin{aligned}
T : \boldsymbol{\Omega} &\longrightarrow \mathbb{R}^3 \\
(\rho, \theta, \zeta) &\mapsto (\rho\cos\theta, \rho\sin\theta, \zeta).
\end{aligned}$$

Notice that ρ and θ are the polar coordinates of (x, y) in $\mathbb{R}^2$. Again, it is easy to see that T is injective and its image is

$$T(\boldsymbol{\Omega}) = \mathbb{R}^3 \setminus \{(x, 0, z) \in \mathbb{R}^3 \,|\, x \geq 0\}.$$

Clearly, T is of class C^1. Moreover, for all $(\rho, \theta, \zeta) \in \boldsymbol{\Omega}$) we have

$$\det J_T(\rho,\theta) = \det\begin{bmatrix}
\frac{\partial T_1}{\partial\rho}(\rho,\theta,\zeta) & \frac{\partial T_1}{\partial\theta}(\rho,\theta,\zeta) & \frac{\partial T_1}{\partial\zeta}(\rho,\theta,\zeta) \\
\frac{\partial T_2}{\partial\rho}(\rho,\theta,\zeta) & \frac{\partial T_2}{\partial\theta}(\rho,\theta,\zeta) & \frac{\partial T_2}{\partial\zeta}(\rho,\theta,\zeta) \\
\frac{\partial T_3}{\partial\rho}(\rho,\theta,\zeta) & \frac{\partial T_3}{\partial\theta}(\rho,\theta,\zeta) & \frac{\partial T_3}{\partial\zeta}(\rho,\theta,\zeta)
\end{bmatrix} = \rho > 0$$

for every $(\rho, \theta, \zeta) \in \boldsymbol{\Omega}$. It follows that T is a change of variable. Since $T(\boldsymbol{\Omega}) \neq \mathbb{R}^3$, the change of variable T is not surjective, but $\mu_3(\mathbb{R}^3 \setminus T(\boldsymbol{\Omega})) = 0$. Let $\boldsymbol{A} \in \mathcal{M}_3$ and let f be summable. If $\boldsymbol{A} \in \mathcal{M}_3$ and $\boldsymbol{A} \subseteq T(\boldsymbol{\Omega})$,

applying Theorem 3.61 and using similar considerations of the previous example we find

$$\int_{\boldsymbol{A}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{T^{-1}(\boldsymbol{A})} f(\rho\,\cos\theta, \rho\,\sin\theta, \zeta)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\zeta.$$

If the condition $\boldsymbol{A} \subseteq T(\boldsymbol{\Omega})$ does not hold. As in the case of polar coordinates, we have

$$\int_{\boldsymbol{A}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega}))} f(\rho\,\cos\theta, \rho\,\sin\theta, \zeta)\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\zeta,$$

and, again, it is straightforward to show that

$$T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega})) = \{(\rho,\theta,\zeta)\in\boldsymbol{\Omega}\,|\,T(\rho,\theta,\zeta)\in\boldsymbol{A}\}.$$

Cylindrical coordinates are useful for the calculation of integrals involving terms (in the definition of the domain of integration or in the integrand function) in the form x^2+y^2. For example, we want to determine the value of the integral

$$\int_{\boldsymbol{A}} e^{-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \qquad \text{where} \quad A = \{(x,y,z)\in\mathbb{R}^3\,|\,x^2+y^2\le z^2 \ \text{ and } \ z\ge 0\}.$$

The integral is defined because $\boldsymbol{A}$ is closed, and therefore measurable, and the integrand function is continuous (and therefore measurable) and non-negative. Thus,

$$\int_{\boldsymbol{A}} e^{-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\boldsymbol{B}} e^{-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$

where

$$\begin{aligned}\boldsymbol{B} &= \{(\rho,\theta,\zeta)\in(0,+\infty)\times(0,2\pi)\times\mathbb{R}\,\big|\,(\rho\cos\theta)^2+(\rho\sin\theta)^2\le\zeta^2 \ \text{ and } \ \zeta\ge 0\}\\ &= \{(\rho,\theta,\zeta)\in(0,+\infty)\times(0,2\pi)\times\mathbb{R}\,\big|\,\zeta>0 \ \text{ and } \ 0<\rho\le\zeta\}.\end{aligned}$$

We define

$$\boldsymbol{C} = \{(\rho,\zeta)\in\mathbb{R}^2\,\big|\,\zeta>0 \ \text{ and } \ 0<\rho\le\zeta\}.$$

From Tonelli's theorem

$$\begin{aligned}\int_{\boldsymbol{A}} e^{-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_{\boldsymbol{C}}\left(\int_{(0,2\pi)} e^{-\zeta}\,\rho\,\mathrm{d}\theta\right)\mathrm{d}\rho\,\mathrm{d}\zeta = 2\pi\int_{\boldsymbol{C}} e^{-\zeta}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\zeta\\ &= 2\pi\int_{(0,+\infty)}\left(\int_{(0,\zeta]} e^{-\zeta}\,\rho\,\mathrm{d}\rho\right)\mathrm{d}\zeta = 2\pi\int_{(0,+\infty)} e^{-\zeta}\left(\int_0^{\zeta}\rho\,\mathrm{d}\rho\right)\mathrm{d}\zeta\\ &= \pi\int_{(0,+\infty)} e^{-\zeta}\,\zeta^2\,\mathrm{d}\zeta = \pi\int_{[0,+\infty)} e^{-\zeta}\,\zeta^2\,\mathrm{d}\zeta\\ &= \pi\lim_{c\to+\infty}\int_0^c e^{-\zeta}\,\zeta^2\,\mathrm{d}\zeta\\ &= \pi\lim_{c\to+\infty} 2\,(1-e^{-c}\,c-e^{-c}) = 2\pi,\end{aligned}$$

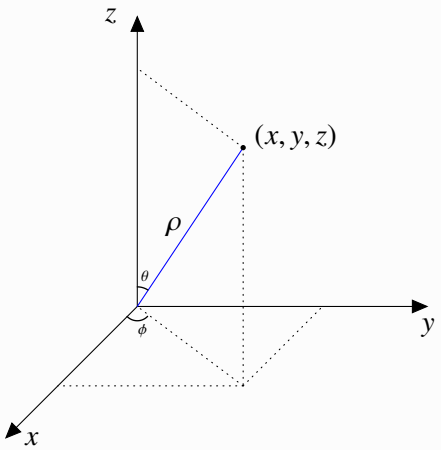
where we have used integration by parts:

$$\begin{aligned}\int x^2\, e^{-x} &= -x^2\, e^{-x} - \int -2\, x\, e^{-x} \\ &= -x^2\, e^{-x} + 2 \int x\, e^{-x} \\ &= -x^2\, e^{-x} + 2\Big(- x\, e^{-x} - \int -e^{-x}\Big) \\ &= -x^2\, e^{-x} - 2\, x\, e^{-x} + 2 \int e^{-x} = -x^2\, e^{-x} - 2\, x\, e^{-x} - 2\, e^{-x}.\end{aligned}$$

Example *3.55*

We now introduce the polar coordinates in $\mathbb{R}^3$ (also known as "spherical coordinates"), which are described by the relations

$$\begin{cases} x = \rho\, \sin\theta\, \cos\phi \\ y = \rho\, \sin\theta\, \sin\phi \\ z = \rho\, \cos\theta \end{cases}$$



In this case,

$$\mathbf{\Omega} = (0, +\infty) \times (0, \pi) \times (0, 2\,\pi).$$

Consider the function

$$\begin{aligned} T : \mathbf{\Omega} &\longrightarrow \mathbb{R}^3 \\ (\rho, \theta, \varphi) &\mapsto (\rho\, \sin\theta\, \cos\phi, \rho\, \sin\theta\, \sin\phi, \rho\, \cos\theta). \end{aligned}$$

One can verify that T is injective and of class C^1. Its image is

$$T(\mathbf{\Omega}) = \mathbb{R}^3 \setminus \{(x, 0, z) \in \mathbb{R}^3 \,|\, x \geq 0\}.$$

Geometrically, this represents a half-plane (it is the part of the xz-plane with $x \geq 0$). For every $(\rho, \theta, \varphi) \in \mathbf{\Omega}$

we have

$$\begin{aligned}
\det J_T(\rho,\theta,\phi) = &= \begin{vmatrix} \sin\theta\cos\phi & \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} \\
&= (\cos\theta)\begin{vmatrix} \rho\cos\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \rho\cos\theta\sin\phi & \rho\sin\theta\cos\phi \end{vmatrix} \\
&\quad +\rho\sin\theta\begin{vmatrix} \sin\theta\cos\phi & -\rho\sin\theta\sin\phi \\ \sin\theta\sin\phi & \rho\sin\theta\cos\phi \end{vmatrix} \\
&= \cos\theta(\rho^2\sin\theta\cos\theta\cos^2\phi + \rho^2\sin\theta\cos\theta\sin^2\phi) \\
&\quad \rho\sin\theta(\rho\sin^2\theta\cos^2\phi + \rho\sin^2\theta\sin^2\phi) \\
&= \rho^2\sin\theta\cos^2\theta + \rho^2\sin^3\theta \\
&= \rho^2\sin\theta(\cos^2\theta + \sin^2\theta) = \rho^2\sin\theta.
\end{aligned}$$

Since $\rho > 0$ and $\theta \in (0,\pi)$ (so that $\sin\theta > 0$), we have $\det J_T(\rho,\theta,\phi) > 0$ for all $(\rho,\theta,\phi) \in \boldsymbol{\Omega}$. If $\boldsymbol{A} \in \mathcal{M}(\mathbb{R}^3)$, if $\boldsymbol{A} \subseteq T(\boldsymbol{\Omega})$ and if f is measurable, we obtain

$$\int_{\boldsymbol{A}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{T^{-1}(\boldsymbol{A})} f(\rho\sin\theta\cos\phi, \rho\sin\theta\sin\phi, \rho\cos\theta)\rho^2\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\phi.$$

If $\boldsymbol{A} \not\subseteq T(\boldsymbol{\Omega})$, we find again

$$\int_{\boldsymbol{A}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega}))} f(\rho\sin\theta\cos\phi, \rho\sin\theta\sin\phi, \rho\cos\theta)\rho^2\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\phi.$$

We have also

$$\begin{aligned}
T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega}) &= \{(\rho,\theta,\phi) \in \boldsymbol{\Omega} \,|\, T(\rho,\theta,\varphi) \in \boldsymbol{A}\} \\
&= \{(\rho,\theta,\phi) \in \boldsymbol{\Omega} \,|\, (\rho\sin\theta\cos\phi, \rho\sin\theta\sin\phi, \rho\cos\theta) \in \boldsymbol{A}\}.
\end{aligned}$$

Spherical coordinates are useful in the calculation of integrals where we have terms in the form $x^2 + y^2 + z^2$. Consider for example $\alpha \in \mathbb{R}$, and the integral

$$I_\alpha = \int_{\boldsymbol{A}} \|\boldsymbol{x}\|^\alpha\,\mathrm{d}\boldsymbol{x}, \qquad \text{where } \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^3 \,|\, \|\boldsymbol{x}\| \geq 1\}.$$

We observe that, for any α, the integral I_α is well defined because $\boldsymbol{A}$ is closed and the integrand function is continuous and non-negative. Then,

$$I_\alpha = \int_{x^2+y^2+z^2\geq 1} (\sqrt{x^2+y^2+z^2})^\alpha\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{x^2+y^2+z^2\geq 1} (x^2+y^2+z^2)^{\frac{\alpha}{2}}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Notice that replacing (x,y,z) with $(\rho\sin\theta\cos\phi, \rho\sin\theta\sin\phi, \rho\cos\theta)$ yields

$$x^2+y^2+z^2 = (\rho^2\sin^2\theta\cos^2\phi + \rho^2\sin^2\theta\sin^2\phi) + \rho^2\cos^2\theta = \rho^2\sin^2\theta + \rho^2\cos^2\theta = \rho^2,$$

so that

$$I_\alpha = \int_{T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega}))} \rho^\alpha\rho^2\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\phi,$$

and

$$T^{-1}(\boldsymbol{A}\cap T(\boldsymbol{\Omega})) = \{(\rho,\theta,\phi) \in \boldsymbol{\Omega} \,|\, (\rho\sin\theta\cos\phi)^2 + (\rho\sin\theta\sin\phi)^2 + (\rho\cos\theta)^2 \geq 1\} = [1,+\infty)\times(0,\pi)\times(0,2\pi).$$

Applying Tonelli's theorem we find

$$\begin{aligned} I_\alpha &= \int_{[1,+\infty)\times(0,\pi)} \left(\int_{(0,2\pi)} \rho^{\alpha+2} \sin\theta \, \mathrm{d}\phi \right) \mathrm{d}\rho \, \mathrm{d}\theta \\ &= 2\pi \int_{[1,+\infty)} \left(\int_{(0,\pi)} \rho^{\alpha+2} \sin\theta \, \mathrm{d}\theta \right) \mathrm{d}\rho \\ &= 2\pi \int_{[1,+\infty)} \rho^{\alpha+2} \left(\int_0^\pi \sin\theta \, \mathrm{d}\theta \right) \mathrm{d}\rho \\ &= 4\pi \int_{[1,+\infty)} \rho^{\alpha+2} \, \mathrm{d}\rho \\ &= 4\pi \lim_{c\to+\infty} \int_1^c \rho^{\alpha+2} \, \mathrm{d}\rho = \begin{cases} 4\pi \lim\limits_{c\to+\infty} \dfrac{c^{\alpha+3}-1}{\alpha+3} & \text{if } \alpha \neq -3 \\ 4\pi \lim\limits_{c\to+\infty} \ln c & \text{if } \alpha = -3. \end{cases} \end{aligned}$$

We obtain

$$I_\alpha = \begin{cases} \dfrac{4\pi}{|\alpha+3|} & \text{if } \alpha < -3 \\ +\infty & \text{if } \alpha \geq -3. \end{cases}$$

Example *3.56*

We now introduce polar coordinates in $\mathbb{R}^n$. Let $n \geq 3$. Let

$$\boldsymbol{\Omega} = (0,+\infty) \times (0,2\pi) \times (0,\pi)^{n-2}.$$

Moreover, let $T : \boldsymbol{\Omega} \longrightarrow \mathbb{R}^n$ where

$$\begin{aligned} T_1(\rho,\theta_1,\ldots,\theta_{n-2},\theta_{n-1}) &= \rho \sin\theta_1 \ldots \sin\theta_{n-2} \sin\theta_{n-1} \\ T_2(\rho,\theta_1,\ldots,\theta_{n-2},\theta_{n-1}) &= \rho \cos\theta_1 \ldots \sin\theta_{n-2} \sin\theta_{n-1} \\ T_j(\rho,\theta_1,\ldots,\theta_{n-2},\theta_{n-1}) &= \rho \cos\theta_{j-1} \ldots \sin\theta_j \sin\theta_{n-1} \\ &\vdots \\ T_n(\rho,\theta_1,\ldots,\theta_{n-2},\theta_{n-1}) &= \rho \cos\theta_{n-1}. \end{aligned}$$

We verify that $T(\boldsymbol{\Omega}) = \mathbb{R}^n \setminus \{(x_1,x_2,\ldots,x_n) \in \mathbb{R}^n \,|\, x_1 = 0, \quad x_2 \geq 0\}$ and T is injective. We see immediately that ρ must coincide with $\sqrt{x_1^2+\ldots+x_n^2}$, which is strictly positive if and only if $(x_1,\ldots,x_n) \neq (0,\ldots,0)$. We have thus brought ourselves to considering the system

$$\begin{cases} \sin\theta_1 \ldots \sin\theta_{n-1} = y_1 \\ \cos\theta_1 \ldots \sin\theta_{n-1} = y_2 \\ \quad\vdots \\ \cos\theta_{j-1} \sin\theta_j \ldots \sin\theta_{n-1} = y_j \\ \quad\vdots \\ \cos\theta_{n-1} = y_n \end{cases} \tag{3.91}$$

where $y_1,\ldots,y_n$ are real and $y_1^2+\ldots+y_n^2 = 1$. We are looking for a unique solution $(\theta_1,\ldots,\theta_{n-1})$ in $(0,2\pi)\times(0,\pi)^{n-2}$. We verify that it exists if and only if $(y_1,y_2) \notin \{0\}\times[0,+\infty)$. We proceed by induction over $n \geq 3$. The case $n = 3$ is left to the reader. Suppose that the result is true for a certain $n \geq 3$, and consider the $(n+1)$-dimensional case. Let $(y_1,\ldots,y_n,y_{n+1}) \in \mathbb{R}^{n+1}$ with $y_1^2+\ldots+y_n^2+y_{n+1}^2 = 1$. Consider the system (3.91) by replacing n with $n+1$. We must have

$$y_{n+1} = \cos\theta_n.$$

Since θ_n must belong to $(0,\pi)$, we must have

$$\theta_n = \arccos y_{n+1}.$$

This value belongs to $(0,\pi)$ if and only if $|y_{n+1}| < 1$, i.e., if and only if $(y_1,\dots,y_n) \neq (0,\dots,0)$. Thus,

$$\sin\theta_n = \sqrt{1-y_{n+1}^2} = \sqrt{y_1^2+\dots+y_n^2} \neq 0,$$

and we are back to system (3.91) with $\left(\frac{y_1}{\sqrt{y_1^2+\dots+y_n^2}},\dots,\frac{y_n}{\sqrt{y_1^2+\dots+y_n^2}}\right)$ in place of $(y_1,\dots,y_n)$. From the inductive assumption, this system has one solution in $(0,2\pi)\times(0,\pi)^{n-2}$ if and only if $\left(\frac{y_1}{\sqrt{y_1^2+\dots+y_n^2}},\frac{y_2}{\sqrt{y_1^2+\dots+y_n^2}}\right) \neq \{0\}\times[0,+\infty)$, and this happens if and only if $(y_1,y_2)\neq\{0\}\times[0,+\infty)$. Now, we verify that for every $(\rho,\theta_1,\dots,\theta_{n-1})\in\Omega$ there holds

$$|\det J_T(\rho,\theta_1,\dots,\theta_{n-1})| = \rho^{n-1}\ \sin\theta_2\cdots(\sin\theta_j)^{j-1}\dots(\sin\theta_{n-1})^{n-2}.$$

This can be done directly for the case $n=3$. Suppose that this result is true for $n-1$ (assuming of course $n\geq 2$), and we prove it for n. Then, for every $(\rho,\theta_1,\dots,\theta_{n-1})\in\boldsymbol{\Omega}$

$$T(\rho,\theta_1,\dots,\theta_{n-1}) = (S(\rho,\theta_1,\dots,\theta_{n-2})\sin\theta_{n-1},\dots,S_{n-1}(\rho,\theta_1,\dots,\theta_{n-2})\ \sin\theta_{n-1},\rho\ \cos\theta_{n-1}),$$

and omitting the argument $(\rho,\theta_1,\dots,\theta_{n-2})$ of the functions, we can write

$$J_T(\rho,\theta_1,\dots,\theta_{n-1}) = \begin{bmatrix} \frac{\partial S_1}{\partial\rho}\sin\theta_{n-1} & \dots & \frac{\partial S_1}{\partial\theta_{n-2}}\sin\theta_{n-1} & S_1\cos\theta_{n-1} \\ \frac{\partial S_2}{\partial\rho}\sin\theta_{n-1} & \dots & \frac{\partial S_2}{\partial\theta_{n-2}}\sin\theta_{n-1} & S_2\cos\theta_{n-1} \\ \vdots & \dots & \ddots & \vdots \\ \frac{\partial S_{n-1}}{\partial\rho}\sin\theta_{n-1} & \dots & \frac{\partial S_{n-1}}{\partial\theta_{n-2}}\sin\theta_{n-1} & S_{n-1}\cos\theta_{n-1} \\ \cos\theta_{n-1} & \dots & 0 & -\rho\sin\theta_{n-1} \end{bmatrix}.$$

Expanding the determinant with respect to the last row and taking into account that $S_j = \rho\dfrac{\partial S_j}{\partial\rho}$ for every $j\in\{1,\dots,n-1\}$, we find

$$\begin{aligned}\det J_T(\rho,\theta_1,\dots,\theta_{n-1}) &= -\rho\left[(\sin\theta_{n-1})^n + \cos^2\theta_{n-1}(\sin\theta_{n-1})^{n-2}\right]\det J_S(\rho,\theta_1,\dots,\theta_{n-2}) \\ &= -\rho(\sin\theta_{n-1})^{n-2}\det J_S(\rho,\theta_1,\dots,\theta_{n-2}),\end{aligned}$$

as expected. We can then conclude that, if $\boldsymbol{A}\in\mathcal{M}_n$ and $f\in\mathcal{L}(\boldsymbol{A})$, the following formula holds:

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{T^{-1}(A)} f(T(\rho,\theta_1,\dots,\theta_{n-1}))\rho^{n-1}\ \sin\theta_2\ \dots(\sin\theta_j)^{j-1}\cdots(\sin\theta_{n-1})^{n-2}\,\mathrm{d}\rho\,\mathrm{d}\theta_1\dots\mathrm{d}\theta_{n-1}.$$

Example 3.57

Let $\boldsymbol{x}_0\in\mathbb{R}^n$. Let $\boldsymbol{A} = \{\boldsymbol{x}\in\mathbb{R}^n \mid 0<\|\boldsymbol{x}-\boldsymbol{x}_0\|<1\}$. We want to determine, given $\alpha\in\mathbb{R}$, the value of

$$\int_{\boldsymbol{A}} \|\boldsymbol{x}-\boldsymbol{x}_0\|^\alpha\,\mathrm{d}x.$$

We observe that the integral is well defined, because $\boldsymbol{A}$ is open and the integrand function is continuous and

non-negative. Let

$$\begin{aligned} S : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{x} &\mapsto \boldsymbol{x} - \boldsymbol{x}_0, \end{aligned}$$

and from Theorem 3.57

$$\int_{\boldsymbol{A}} \|\boldsymbol{x} - \boldsymbol{x}_0\|^\alpha \, \mathrm{d}\boldsymbol{x} = \int_A \|S(\boldsymbol{x})\|^\alpha \, \mathrm{d}\boldsymbol{x} = \int_{S(\boldsymbol{A})} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x},$$

where $S(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, 0 < \|\boldsymbol{x}\| < 1\}$. Using the polar coordinates in $\mathbb{R}^n$, we find

$$\begin{aligned} \int_{\boldsymbol{A}} \|\boldsymbol{x} - \boldsymbol{x}_0\|^\alpha \, \mathrm{d}\boldsymbol{x} &= \int_{T^{-1}(\boldsymbol{A})} \rho^{\alpha+n-1} \, \sin\theta_2 \cdots)(\sin\theta_j)^{j-1} \cdots (\sin\theta_{n-1})^{n-2} \, \mathrm{d}\rho \, \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_{n-1} \\ &= \int_{(0,1)} \rho^{\alpha+n-1} \, \mathrm{d}\rho \int_{(0,2\pi)\times(0,\pi)^{n-2}} \sin\theta_2 \cdots)(\sin\theta_j)^{j-1} \cdots (\sin\theta_{n-1})^{n-2} \, \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_{n-1}, \end{aligned}$$

which is finite if and only if $\alpha + n - 1 > -1$, i.e., $a > -n$. let now $\boldsymbol{B} = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \|\boldsymbol{x} - \boldsymbol{x}_0\| > 1\}$. Given $\alpha \in \mathbb{R}$, we compute

$$\int_{\boldsymbol{B}} \|\boldsymbol{x} - \boldsymbol{x}_0\|^\alpha \, \mathrm{d}\boldsymbol{x}.$$

We have again

$$\int_{\boldsymbol{B}} \|\boldsymbol{x} - \boldsymbol{x}_0\|^\alpha \, \mathrm{d}\boldsymbol{x} = \int_{S(\boldsymbol{B})} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x},$$

where $S(\boldsymbol{B}) = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \|\boldsymbol{x}\| > 1\}$.

Exercises

Exercise 3.32 Compute

$$\lim_{k\to\infty} \int_0^k \frac{\sin x}{x} \, \mathrm{d}x$$

using the integral $\frac{1}{x} = \int_0^{+\infty} e^{-xt} \, \mathrm{d}t$ for $x > 0$ and Fubini/Tonelli's Theorem. [Solution: (3.30)]

Exercise 3.33 Let $f \in L^1(\mathbb{R};\mathbb{R})$. Show that

$$\lim_{k\to\infty} \frac{1}{2k} \int_{-k}^k f(x) \, \mathrm{d}x = 0.$$

[Solution: (3.31)]

Exercise 3.34 Let $a > 0$. Compute

$$\lim_{n\to\infty} \int_0^a \sin\left(1 - e^{-\frac{x}{n}}\right) e^{-\frac{x}{n}} \, \mathrm{d}x.$$

[Solution: (3.32)]

3.12 Integral convergence

Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of summable functions from $\boldsymbol{A}$ to $\mathbb{C}$ and let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ be summable, and

$$\lim_{k\to\infty} \int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = 0. \tag{3.92}$$

We want to approximate f in the sense of (3.92) with continuous functions f_k. We begin with an approximation result of a summable function with continuous functions in $\mathbb{R}^n$.

Theorem 3.62 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be summable, and let $\varepsilon > 0$. Then, there exists a continuous function g, zero outside a compact set, such that*

$$\int_{\mathbb{R}^n} |f(\boldsymbol{x}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \varepsilon.$$

Proof We denote by $\mathcal{A}$ the class of summable functions that enjoy the aforementioned property. We have to prove that $\mathcal{A}$ is equal to the set of summable functions from $\mathbb{R}^n$ to $\mathbb{C}$. It is easy to see that:

(i) linear combinations of elements of $\mathcal{A}$ are elements of $\mathcal{A}$;

(ii) if

- $f_k \in \mathcal{A}$ for every $k \in \mathbb{N}$;
- $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is summable;
- $\lim_{k \to \infty} \int_{\mathbb{R}^n} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = 0$,

then $f \in \mathcal{A}$. Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. We have to show that $f \in \mathcal{A}$. From *(i)*, we can restrict our attention to the case where f takes real values. In this case, writing $f = f^+ - f^-$, we can limit ourselves to the case $f \geq 0$. From Theorem 3.29, there exists a sequence of simple non-negative functions $(\varphi_k)_{k \in \mathbb{N}}$ such that, for all $k \in \mathbb{N}$, we have $0 \leq \varphi_k \leq \varphi_{k+1}$ for all $k \in \mathbb{N}$, and

$$\forall \, \boldsymbol{x} \in \boldsymbol{A}, \quad \lim_{k \to \infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x}).$$

From $|f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})| \leq 2 \, f(\boldsymbol{x})$ and the Dominated Convergence Theorem (Theorem 3.47), it follows that

$$\lim_{k \to \infty} \int_{\mathbb{R}^n} |f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = 0,$$

and from *(ii)*, we only need to prove that every simple non-negative function belongs to $\mathcal{A}$. From *(i)*, this follows from the fact that characteristic functions of measurable sets with finite measure belong to $\mathcal{A}$. Let $f = \chi_{\boldsymbol{A}}$, with $\boldsymbol{A} \in \mathcal{M}_n$, with $\mu_n(A) < +\infty$. By definition, given $\eta > 0$, a family of n-dimensional intervals $(I_k)_{k \in \mathbb{N}}$ exists such that $\boldsymbol{A} \subseteq \boldsymbol{B}$, with

$$\boldsymbol{B} = \bigcup_{k=0}^{\infty} I^k,$$

and

$$\mu_n(\boldsymbol{A}) \subseteq \mu_n(\boldsymbol{B}) \subseteq \sum_{k=0}^{\infty} \mathrm{vol}_n(I_k) \leq \mu_n(\boldsymbol{A}) + \eta.$$

Thus,

$$\int_{\mathbb{R}^n} |\chi_{\boldsymbol{A}}(\boldsymbol{x}) - \chi_{\boldsymbol{B}}(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = \mu_n(\boldsymbol{B} \setminus \boldsymbol{A}) < \eta.$$

From *(ii)*, we can limit ourselves to the case $f = \chi_{\boldsymbol{A}}$, where

$$\boldsymbol{A} = \bigcup_{k=0}^{\infty} I^k,$$

where I^k are n-dimensional intervals and

$$\sum_{k=0}^{\infty} \mathrm{vol}_n(I^k) < +\infty.$$

Define

$$\boldsymbol{A}_k = \sum_{r=0}^{k} I^k.$$

We have

$$\boldsymbol{A} \setminus \boldsymbol{A}_k \subseteq \bigcup_{r=k+1}^{\infty} I^r.$$

Hence

$$\begin{aligned}\int_{\mathbb{R}^n} |\chi_{\boldsymbol{A}}(\boldsymbol{x}) - \chi_{\boldsymbol{A}_k}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^n} |\chi_{\boldsymbol{A}\setminus\boldsymbol{A}_k}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\ &\leq \int_{\mathbb{R}^n} \sum_{k=r+1}^{\infty} \chi_{I^r}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \chi_{I^r}(\boldsymbol{x})\,\mu_n(I^r),\end{aligned}$$

applying **(c)** in Theorem 3.35, and so

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} |\chi_{\boldsymbol{A}}(\boldsymbol{x}) - \chi_{\boldsymbol{A}_k}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = 0.$$

From *(ii)*, it is sufficient to consider the case where $f = \chi_{\boldsymbol{A}}$, where $\boldsymbol{A}$ is a multi-interval. By Lemma 3.8,

$$\boldsymbol{A} = \bigcup_{k=0}^{p} I^k,$$

where I^k are pairwise disjoint n-dimensional intervals, and so

$$\chi_{\boldsymbol{A}} = \sum_{k=0}^{p} \chi_{I^k}.$$

From *(i)*, this allows us to just consider $f = \chi_{\boldsymbol{A}}$, where $\boldsymbol{A}$ is an n-dimensional interval. If $\boldsymbol{A} = \varnothing$, the result is trivial. Let $\boldsymbol{A}$ be non-empty, and

$$\boldsymbol{A} = \prod_{j=1}^{n} I_j,$$

where I_j is a non-empty interval of $\mathbb{R}$ for each $j \in \{1, 2, \ldots, n\}$. Let $\eta > 0$. For $j \in \{1, 2, \ldots, n\}$, let $a_j = \inf I_j$ and $b_j = \sup I_j$. We define

$$\begin{aligned} f_{\eta,j} : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} 1 & \text{if } a_j \leq t \leq b_j \\ \eta^{-1}(t - a_j + \eta) & \text{if } a_j - \eta \leq t \leq a_j \\ \eta^{-1}(b_j + \eta - t) & \text{if } b_j \leq t \leq a_j + \eta \\ 0 & \text{if } t \notin [a_j - \eta, b_j + \eta] \end{cases}\end{aligned}$$

and

$$f_\eta : \mathbb{R}^n \longrightarrow \mathbb{R}$$
$$(x_1, x_2, \ldots, x_n) \mapsto \prod_{j=1}^{n} f_{\eta,j}(x_j).$$

Function f_η is continuous, and it is zero outside $\prod_{j=1}^{n}[a_j - \eta, b_j + \eta]$. Moreover,

$$\chi_A \leq f_\eta \leq \chi_{\prod_{j=1}^{n}[a_j-\eta, b_j+\eta]}.$$

It follows that

$$\int_{\mathbb{R}^n} |\chi_A(\boldsymbol{x}) - f_\eta(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \prod_{j=1}^{n}(b_j - a_j + \eta) - \prod_{j=1}^{n}(b_j - a_j),$$

which tends to zero as η tends to zero.

A simple but important consequence of Theorem 3.62 is the following result.

Theorem 3.63 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be summable. Then*

$$\lim_{\boldsymbol{h}\to 0} \int_{\mathbb{R}^n} |f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = 0. \tag{3.93}$$

Proof The integral in (3.93) is well defined in view of Theorem 3.61. If f is continuous and with compact support, the result is obvious. In general, let $\varepsilon > 0$. From Theorem 3.62, there exists a continuous function g, with compact support, such that

$$\int_{\mathbb{R}^n} |f(\boldsymbol{x}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \frac{\varepsilon}{3}.$$

Given $\boldsymbol{h} \in \mathbb{R}^n$, applying Theorem 3.61,

$$\begin{aligned}\int_{\mathbb{R}^n} |f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} &\leq \int_{\mathbb{R}^n} |f(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x} + \boldsymbol{h})| \, \mathrm{d}\boldsymbol{x} + \int_{\mathbb{R}^n} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} + \int_{\mathbb{R}^n} |g(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \\ &= 2 \int_{\mathbb{R}^n} |g(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} + \int_{\mathbb{R}^n} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \\ &\leq \frac{2\,\varepsilon}{3} + \int_{\mathbb{R}^n} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}.\end{aligned}$$

Let $\delta > 0$ be such that, if $\|\boldsymbol{h}\| < \delta$, then

$$\int_{\mathbb{R}^n} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \frac{\varepsilon}{3}.$$

Then,

$$\int_{\mathbb{R}^n} |f(\boldsymbol{x} + \boldsymbol{h}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \varepsilon.$$

The following result establishes a link between the convergence in the sense of (3.92) and the pointwise convergence.

Theorem 3.64 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of summable functions from $\boldsymbol{A}$ to $\mathbb{C}$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ be a summable function such that (3.92) holds. Then, there exists a subsequence of $(f_k)_{k\in\mathbb{N}}$ that converges*

pointwise to f almost everywhere in $\boldsymbol{A}$.

Proof Let $\varepsilon > 0$. Let $k_\varepsilon \in \mathbb{N}$ be such that, for $k > k_\varepsilon$,

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \frac{\varepsilon}{2}.$$

Then, if j and k are natural numbers and $\min\{j, k\} > k_\varepsilon$, we have

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f_j(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} |f_j(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \varepsilon.$$

Therefore, for every $\varepsilon > 0$, there exists $k_\varepsilon \in \mathbb{N}$ such that, if $\min\{j, k\} > k_\varepsilon$, we have

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f_j(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < \varepsilon.$$

Let $(k_r)_{r \in \mathbb{N}}$ be a strictly increasing sequence of natural numbers such that, for every $r \in \mathbb{N}$, if k and j are natural numbers and $\min\{j, k\} > k_r$, then

$$\int_{\boldsymbol{A}} |f_k(\boldsymbol{x}) - f_j(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} < 2^{-r}.$$

We show that for almost all $\boldsymbol{x} \in \boldsymbol{A}$ there holds

$$\lim_{r \to \infty} f_{k_r}(\boldsymbol{x}) = f(\boldsymbol{x}).$$

Let $g_r(\boldsymbol{x}) = f_{k_r}(\boldsymbol{x})$ for every $r \in \mathbb{N}$. Consider the series of functions

$$g_1(\boldsymbol{x}) + \sum_{r=1}^{\infty} (g_{r+1}(\boldsymbol{x}) - g_r(\boldsymbol{x})). \tag{3.94}$$

We have

$$\sum_{r=1}^{\infty} \int_{\boldsymbol{A}} |g_{r+1}(\boldsymbol{x}) - g_r(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \sum_{r=1}^{\infty} 2^{-r} < +\infty.$$

From Theorem **??**, the series of functions (3.94) is absolutely convergent almost everywhere in $\boldsymbol{A}$. Moreover, for every $k \in \mathbb{N}$

$$g_1(\boldsymbol{x}) + \sum_{r=1}^{\infty} (g_{r+1}(\boldsymbol{x}) - g_r(\boldsymbol{x})) = g_{k+1}(\boldsymbol{x}),$$

and so the sequence $(g_k(\boldsymbol{x}))_{k \in \mathbb{N}}$ is pointwise convergent almost everywhere in $\boldsymbol{A}$. We denote by $\boldsymbol{A}'$ the set of $\boldsymbol{x}$ in $\boldsymbol{A}$ such that the series of functions (3.94) is convergent in $\boldsymbol{x}$. This set coincides with

$$\{\boldsymbol{x} \in \boldsymbol{A} \mid (g_r(\boldsymbol{x}))_{r \in \mathbb{N}} \text{ is convergent}\}.$$

For $\boldsymbol{x} \in \boldsymbol{A}'$, let $s(\boldsymbol{x}) = \lim_{r \to \infty} g_r(\boldsymbol{x})$. By Theorem 3.49, s is summable over $\boldsymbol{A}'$. From Theorem 3.38 (Fatou's Lemma),

$$\int_{\boldsymbol{A}'} |s(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \liminf_{r \to \infty} \int_{\boldsymbol{A}'} |g_r(\boldsymbol{x}) - f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} = 0.$$

We conclude that $s(\boldsymbol{x}) = f(\boldsymbol{x})$ almost everywhere on $\boldsymbol{A}'$, and this concludes the proof.

In general, from the assumptions of Theorem 3.64, we cannot deduce that $\lim_{k \to \infty} f_k(\boldsymbol{x}) = f(\boldsymbol{x})$ almost everywhere, as the following example shows.

Example *3.58*

Let

$$a_n = \left\lfloor \sum_{k=1}^{n} \frac{1}{k} \right\rfloor,$$

i.e., a_n is the fractional part of $\displaystyle\sum_{k=1}^{n} \frac{1}{k}$, which belongs to $[0,1)$. We denote by A_n the interval $[a_n, a_{n+1}]$ if $a_n \leq a_{n+1}$, or $[0, a_{n+1}] \cup [a_n, 1)$ if $a_{n+1} < a_n$. For every $n \in \mathbb{N}$,

$$\mu_1(A_n) = \frac{1}{n+1}.$$

It follows that

$$\lim_{n\to\infty} \int_{[0,1)} \chi_{A_n}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 0.$$

On the other hand, since $\displaystyle\sum_{k=1}^{n} \frac{1}{k}$ is divergent, every element of $[0,1)$ belongs to an infinite number of A_n. Therefore, no $x \in [0,1)$ exists such that $\lim_{n\to\infty} \chi_{A_n}(x) = 0$.

3.13 Differentiation of integral functions

We recall that the Fundamental Theorem of Calculus says that, given a continuous function $f : \mathbb{R} \longrightarrow \mathbb{C}$ and given $x_0 \in \mathbb{R}$, the function

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{C} \\ x &\mapsto \int_{x_0}^{x} f(y)\, \mathrm{d}y \end{aligned}$$

is differentiable and $F'(x) = f(x)$ for all $x \in \mathbb{R}$. In this section, we present some important extensions of this result.

Lemma 3.15 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be summable. Let $r > 0$ and $x \in \mathbb{R}^n$. Let*

$$\begin{aligned} f_r : \mathbb{R}^n &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} f(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}. \end{aligned} \tag{3.95}$$

Then:

(a) f_r is uniformly continuous;

(b) f_r is summable;

(c) $\displaystyle\lim_{r\to 0} \int_{\mathbb{R}^n} |f_r(\boldsymbol{x}) - f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = 0.$

Proof (a) We observe that, by virtue of Lemma 3.65 and Example 3.43 , for all $x \in \mathbb{R}^n$

$$\mu_n(\mathcal{B}_r(\boldsymbol{x})) = \omega_n\, r^n.$$

Let $\boldsymbol{x}$ and $\boldsymbol{x}+\boldsymbol{h}$ be in $\mathbb{R}^n$. Then,

$$\begin{aligned}|f_r(\boldsymbol{x}+\boldsymbol{h})-f_r(\boldsymbol{x})| &= \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))}\left|\int_{\mathcal{B}_r(\boldsymbol{x})}(f(\boldsymbol{y}+\boldsymbol{h})-f_(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\right| \\ &\leq \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\left|\int_{\mathcal{B}_r(\boldsymbol{x})}(f(\boldsymbol{y}+\boldsymbol{h})-f_(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\right|,\end{aligned}$$

and the result follows from Theorem 3.63.

(b) Note that the function $(\boldsymbol{x},\boldsymbol{y})\mapsto f(\boldsymbol{x}+\boldsymbol{y})$ is measurable in $\mathbb{R}^{2n}$.[a] Taking the composition of this function with the diffeomorphism

$$\begin{aligned}\mathbb{R}^{2n} &\longrightarrow \mathbb{R}^{2n}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto (\boldsymbol{x}+\boldsymbol{y},\boldsymbol{y})\end{aligned}$$

and applying **(b)** of Theorem 3.61 and Remark 3.26,

$$\begin{aligned}\int_{\mathbb{R}^n}|f_r(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &\leq \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\int_{\mathbb{R}^n}\left(\int_{\mathcal{B}_r(\boldsymbol{0})}|f(\boldsymbol{x}+\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x}\\ &= \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\int_{\mathcal{B}_r(\boldsymbol{0})}\left(\int_{\mathbb{R}^n}|f(\boldsymbol{x}+\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x}=\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y},\end{aligned}$$

where the last equality is obtained by applying Corollary 3.4.

(c) For every $\boldsymbol{x}\in\mathbb{R}^n$

$$\begin{aligned}|f_r(\boldsymbol{x})-f(\boldsymbol{x})| &= \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\left|\int_{\mathcal{B}_r(\boldsymbol{0})}(f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{y}\right|\\ &= \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\int_{\mathcal{B}_r(\boldsymbol{0})}\left|f(\boldsymbol{x}+\boldsymbol{y})-f_(\boldsymbol{x})\right|\mathrm{d}\boldsymbol{x}.\end{aligned}$$

From previous considerations, the function

$$\begin{aligned}\mathbb{R}^{2n} &\longrightarrow \mathbb{R}^{2n}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto |f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})|\end{aligned}$$

is measurable. Thus,

$$\begin{aligned}\mathbb{R}^{n} &\longrightarrow \mathbb{R}^{n}\\ \boldsymbol{y} &\mapsto |f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})|\end{aligned}$$

is measurable for every $\boldsymbol{x}\in\mathbb{R}^n$, and so

$$\begin{aligned}\mathbb{R}^{n} &\longrightarrow \mathbb{R}^{n}\\ \boldsymbol{x} &\mapsto \int_{\mathcal{B}_r(\boldsymbol{0})}|f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{y}\end{aligned}$$

is measurable. By Corollary 3.4

$$\begin{aligned}\int_{\mathbb{R}^n}|f_r(\boldsymbol{x})-f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &\leq \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\int_{\mathbb{R}^n}\left(\int_{\mathcal{B}_r(\boldsymbol{0})}|f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x}\\ &= \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{0}))}\int_{\mathcal{B}_r(\boldsymbol{0})}\left(\int_{\mathbb{R}^n}|f(\boldsymbol{x}+\boldsymbol{y})-f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\right)\mathrm{d}\boldsymbol{y}<\varepsilon,\end{aligned}$$

if $r < r_\varepsilon$ for some $r_\varepsilon > 0$, see Theorem 3.63.

[a]Given the measurable functions $f : \mathbb{R}^m \longrightarrow \mathbb{C}$ and $g : \mathbb{R}^n \longrightarrow \mathbb{C}$, the tensor product

$$\begin{aligned} f \otimes g : \mathbb{R}^{m+n} &\longrightarrow \mathbb{C} \\ (\boldsymbol{x}, \boldsymbol{y}) &\mapsto f(\boldsymbol{x})\, g(\boldsymbol{y}) \end{aligned}$$

is measurable. To see this, you can start with simple functions, applying Theorem **??**, and then using Theorem 3.30.

Theorem 3.65 [Vitali's Lemma]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $\mathcal{I}$ be an arbitrary set and, for all $i \in \mathcal{I}$, let $\mathcal{B}_i$ be a ball with positive diameter in $\mathbb{R}^n$, such that

(i) $\boldsymbol{A} \subseteq \bigcup_{i \in \mathcal{I}} \mathcal{B}_i$;

(ii) $\sup_{i \in \mathcal{I}} \operatorname{diam}(\mathcal{B}_i) < +\infty$.

Then, there exists a finite or countable subfamily $(C_k)_{k \in \mathcal{F}}$ of $(\mathcal{B}_i)_{i \in \mathcal{I}}$ such that C_k are pairwise disjoint and

$$\sum_{k \in \mathcal{F}} \mu_n(C_k) \geq 5^{-n}\, \mu_n(\boldsymbol{A}).$$

Proof Choose $C_1 = \mathcal{B}_{i_1}$, so that

$$\operatorname{diam}(C_1) \geq \frac{1}{2} \sup_{i \in \mathcal{I}} \operatorname{diam}(\mathcal{B}_i).$$

Assume that the non-empty sets $C_1, C_2, \ldots, C_k$ have been chosen. Let $\mathcal{F} = \{1, 2, \ldots, k\}$ if, for every $i \in \mathcal{I}$, we have

$$\Big(\bigcup_{r=1}^{k} C_r\Big) \cap \mathcal{B}_i \neq \varnothing.$$

Vice-versa, choose $C_{k+1} = \mathcal{B}_{i_{k+1}}$, so that

$$\operatorname{diam}(\mathcal{B}_{i_{k+1}}) \geq \sup_{\mathcal{B}_i \cap \left(\bigcup_{r=1}^{k} C_r\right) = \varnothing} \operatorname{diam}(\mathcal{B}_i). \tag{3.96}$$

In this way, we have $\mathcal{F} = \{1, 2, \ldots, k\}$ for some $k \in \mathbb{N}$, or $\mathcal{F} = \mathbb{N}^*$ if, for every $k \in \mathbb{N}^*$, there exists $i \in \mathcal{I}$ such that

$$\bigcup_{r \leq k} C_r \cap \mathcal{B}_i = \varnothing.$$

If $\sum_{k \in \mathcal{F}} \mu_n(C_k) = +\infty$, the conclusion is trivial. Suppose

$$\sum_{k \in \mathcal{F}} \mu_n(C_k) < +\infty. \tag{3.97}$$

Given $k \in \mathcal{F}$, we denote by $\tilde{C}_k$ the closed ball with the same center as C_k and whose diameter is 5 times the diameter of C_k. We verify that

$$\boldsymbol{A} \subseteq \bigcup_{k \in \mathcal{F}} \tilde{C}_k.$$

Let $\boldsymbol{x} \in \boldsymbol{A}$. Then, there exists $i \in \mathcal{I}$ such that $\boldsymbol{x} \in \mathcal{B}_i$. There exists k such that $\mathcal{B}_i \cap C_k \neq \emptyset$. This is certainly true if $\mathcal{F}$ is finite. Conversely, if $\mathcal{F} = \mathbb{N}^*$, (3.97) implies

$$\lim_{k \to \infty} \operatorname{diam}(C_k) = 0.$$

From (3.96), we must have $\operatorname{diam}(\mathcal{B}_i) \leq 2 \operatorname{diam}(C_k)$ for every k, and this implies $\operatorname{diam}(\mathcal{B}_i) = 0$. Therefore, there exists k such that $\mathcal{B}_I \cap C_j \neq \emptyset$. Let k be the minimal index in $\mathcal{F}$ with this property. Therefore, $\mathcal{B}_i \cap C_j = \emptyset$ if $j < k$, from which, by (3.96), we have $\operatorname{diam}(\mathcal{B}_i) \leq 2 \operatorname{diam}(C_k)$. From these considerations, we obtain $\mathcal{B}_i \subseteq \tilde{C}_k$. Therefore,

$$\mu_n(\boldsymbol{A}) \leq \sum_{k \in \mathcal{F}} \mu_n(\tilde{C}_k) = 5^n \sum_{k \in \mathcal{F}} \mu_n(C_k).$$

Definition 3.23 [Locally Summability]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ be measurable. We say that f is **locally summable** *if its restriction to every compact subset of $\boldsymbol{A}$ is summable.*

Definition 3.24 [Hardy-Littlewood Maximal Function]
Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be measurable. We define

$$\begin{aligned} f^{\#} : \mathbb{R}^n &\longrightarrow [0, +\infty] \\ \boldsymbol{x} &\mapsto \sup_{r>0} \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y}. \end{aligned}$$

The function $f^{\#}$ is called Hardy-Littlewood maximal function.

Theorem 3.66 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be measurable. Then, $f^{\#}$ is measurable.*

Proof We verify that, for all $t \in \mathbb{R}$, the set $\{\boldsymbol{x} \in \mathbb{R}^n \mid f^{\#}(\boldsymbol{x}) > t\}$ is open. Let $\boldsymbol{x} \in \mathbb{R}^n$ be such that $f^{\#}(\boldsymbol{x}) > t$ for some $t \in \mathbb{R}$. Then, there exists $r > 0$ such that

$$\int_{\mathcal{B}_r(\boldsymbol{x})} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} > t \, \mu_n(\mathcal{B}_r(\boldsymbol{x})).$$

Applying Theorem 3.34,

$$\begin{aligned} \int_{\mathcal{B}_r(\boldsymbol{x})} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} &= \lim_{k \to \infty} \int_{\mathcal{B}_r(\boldsymbol{x})} |f(\boldsymbol{y})| \chi_{\mathcal{B}_{r-\frac{1}{k}}}(\boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} \\ &= \lim_{k \to \infty} \int_{\mathcal{B}_{r-\frac{1}{k}}(\boldsymbol{x})} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} \end{aligned}$$

It follows that there exists $k \in \mathbb{N}$ such that

$$\int_{\mathcal{B}_{r-\frac{1}{k}}(\boldsymbol{x})} |f(\boldsymbol{y})| \, \mathrm{d}\boldsymbol{y} > t \, \mu_n(\mathcal{B}_r(\boldsymbol{x})).$$

Let $\boldsymbol{h} \in \mathbb{R}^n$ be such that $\|\boldsymbol{h}\| < 1/k$. Then, $\mathcal{B}_{r-\frac{1}{k}}(\boldsymbol{x}) \subseteq \mathcal{B}_r(\boldsymbol{x}+\boldsymbol{h})$. Then

$$\begin{aligned} f^{\#}(\boldsymbol{x}+\boldsymbol{h}) &\geq \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}+\boldsymbol{h}))} \int_{\mathcal{B}_r(\boldsymbol{x}+\boldsymbol{h})} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \\ &\geq \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_{r-\frac{1}{k}}(\boldsymbol{x})} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} > t. \end{aligned}$$

Theorem 3.67 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be summable. Then, for all $t \in \mathbb{R}_+^*$,*

$$\mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \mid f^{\#}(\boldsymbol{x}) > t\}) \leq \frac{\displaystyle\int_{\mathbb{R}^n} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}}{t}.$$

Proof Let $\boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{R}^n \mid f^{\#}(\boldsymbol{x}) > t\}$, and let $\boldsymbol{x} \in \boldsymbol{A}$. Then, there exists $r(\boldsymbol{x}) > 0$ such that

$$\int_{\mathcal{B}_r(\boldsymbol{x})} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} > t\,\mu_n(\mathcal{B}_r(\boldsymbol{x})).$$

In this way, we have constructed a family of balls with positive radius that covers $\boldsymbol{A}$. By Theorem **??**, this family has a finite or countable subfamily $\{C_k = \mathcal{B}_{r_k}(\boldsymbol{x}_k) \,|k \in \mathcal{F}\}$ such that the C_k are pairwise disjoint and

$$\sum_{k\in\mathcal{F}} \mu_n(C_k) \geq 5^{-n}\,\mu_n(\boldsymbol{A}).$$

Thus,

$$\begin{aligned} \mu_n(\boldsymbol{A}) &\leq 5^n \sum_{k\in\mathcal{F}} \mu_n(C_k) \\ &\leq \frac{5^n}{t} \sum_{k\in\mathcal{F}} \int_{C_k} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \\ &\leq \frac{5^n}{t} \int_{\bigcup_{k\in\mathcal{F}} C_k} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \\ &\leq \frac{5^n \displaystyle\int_{\mathbb{R}^n} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}}{t}. \end{aligned}$$

The following result is the most important of this section.

Theorem 3.68 *Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ be locally summable. Then,*

$$\lim_{r\to 0} \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = f(\boldsymbol{x}) \tag{3.98}$$

almost everywhere in $\mathbb{R}^n$.

Proof First, let us consider the case where f is summable. We can assume that f takes real values. For $\boldsymbol{x} \in \mathbb{R}^n$, let

$$\Omega_f(\boldsymbol{x}) = \lim_{r\to 0} \Big(\sup_{0<\rho\leq r} f_\rho(\boldsymbol{x}) - \inf_{0<\rho\leq r} f_\rho(\boldsymbol{x}) \Big), \tag{3.99}$$

where f_r is defined in (3.95). Notice that the limit exists in $[0, +\infty]$ because the function

$$\begin{aligned} \mathbb{R}_+^* &\longrightarrow \mathbb{R} \\ r &\mapsto \sup_{0<\rho\le r} f_\rho(\boldsymbol{x}) - \inf_{0<\rho\le r} f_\rho(\boldsymbol{x}) \end{aligned}$$

is non-decreasing. We also notice that $\Omega_f(\boldsymbol{x}) = 0$ if and only if the limit

$$\lim_{r\to 0} \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \tag{3.100}$$

exists in $\mathbb{R}$, as one can easily see. Moreover, for every $\boldsymbol{x} \in \mathbb{R}^n$

$$\Omega_f(\boldsymbol{x}) \le 2\, f^{\#}(\boldsymbol{x}). \tag{3.101}$$

Given two summable functions $g, h : \mathbb{R}^n \longrightarrow \mathbb{R}$, we have

$$\Omega_{g+h} \le \Omega_g + \Omega_h. \tag{3.102}$$

Let $\varepsilon > 0$. From Theorem 3.62, there exists a continuous and summable function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that

$$\int_{\mathbb{R}^n} |f(\boldsymbol{x}) - g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < \varepsilon.$$

Let $h = f - g$. From (3.102), since Ω_g is identically zero, we obtain $\Omega_f \le \Omega_h$. Applying (3.101) and Theorem 3.67, for every $t > 0$

$$\begin{aligned} \mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \,|\, \Omega_f(\boldsymbol{x}) > t\}) &\le \mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \,|\, \Omega_h(\boldsymbol{x}) > t\}) \\ &\le \mu_n\left(\left\{\boldsymbol{x} \in \mathbb{R}^n \,\middle|\, h^{\#}(\boldsymbol{x}) > \frac{t}{2}\right\}\right) \\ &\le 2\,\frac{5^n \int_{\mathbb{R}^n} |h(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}}{t} \le 2\,\frac{5^n\,\varepsilon}{t}. \end{aligned}$$

Since ε is arbitrary, we obtain $\mu_n(\{\boldsymbol{x} \in \mathbb{R}^n \,|\, \Omega_f(\boldsymbol{x}) > t\}) = 0$ for every $t > 0$, and so $\Omega_f(\boldsymbol{x}) = 0$ almost everywhere in $\mathbb{R}^n$. As already mentioned, this means that for almost all $\boldsymbol{x} \in \mathbb{R}^n$ the limit (3.100) exists in $\mathbb{R}$. We only need to show that this limit coincides with $f(x)$. Using (c) in Lemma 3.15 and Theorem 3.64, there exists a sequence $(r_k)_{k\in\mathbb{N}}$ of positive real number converging to zero such that $\lim_{k\to\infty} f_{r_k}(\boldsymbol{x}) = f(\boldsymbol{x})$ almost everywhere in $\mathbb{R}^n$. This concludes the proof in the case where f is summable.

Let us now consider the case where f is locally summable. For $k \in \mathbb{N}^*$, let $f_k = f\,\chi_{[-k,k]^n}$. For every k, the function f_k is summable. Let N_k the set of $\boldsymbol{x} \in (-k, k)^n$ for which the limit

$$\lim_{r\to 0} \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} f_k(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

does not exist in $\mathbb{R}$, or it exists in $\mathbb{R}$ but does not coincide with $f_k(\boldsymbol{x})$. From the first part of the proof, N_k is measurable and with Lebesgue measure zero. Then,

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n \setminus \bigcup_{k=1}^{\infty} N_k, \qquad \lim_{r\to 0} \frac{1}{\mu_n(\mathcal{B}_r(\boldsymbol{x}))} \int_{\mathcal{B}_r(\boldsymbol{x})} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = f(\boldsymbol{x}),$$

which allows us to conclude.

A result that is useful in the theory of differential equations is the following.

Corollary 3.6 *Let $a, b \in \mathbb{R}$ be such that $a < b$. Let $f : [a, b] \longrightarrow \mathbb{R}$ be summable. Then, f is zero almost everywhere if and only if for every subinterval I of $[a, b]$*

$$\int_I f(t)\,\mathrm{d}t = 0. \tag{3.103}$$

Proof **(Only if)** is obvious: if f is zero almost everywhere, the integral over any interval contained in $[a, b]$ is zero.
(If). When $n = 1$, (3.98) can be written as

$$\lim_{h \to 0^+} \frac{1}{2\,h} \int_{(x-h,x+h)} f(y)\,\mathrm{d}y = f(x).$$

If for every subinterval I of $[a, b]$ the integral is zero, the left hand-side of the latter is zero, and therefore f is zero almost everywhere.

From the proof of Theorem 3.68, it emerges that several variations of this result can be formulated. For example, we can replace the balls $\mathcal{B}_r(\boldsymbol{x})$ with cubes

$$\prod_{j=1}^{n} (x_j - r, x_j + r),$$

and the 1-dimensional case we can consider, instead of $(x - r, x + r)$, the intervals $(x, x + r)$ or $(x - r, x)$. In this way, we obtain the following generalization of the Fundamental Theorem of Calculus:

Corollary 3.7 *Let $f : \mathbb{R} \longrightarrow \mathbb{C}$ be locally summable. Let $x_0 \in \mathbb{R}$. Let*

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{C} \\ x &\mapsto \begin{cases} \displaystyle\int_{(x_0,x)} f(y)\,\mathrm{d}y & \text{if } x_0 \le x \\ \displaystyle -\int_{(x,x_0)} f(y)\,\mathrm{d}y & \text{if } x \le x_0. \end{cases} \end{aligned}$$

Then, F is differentiable almost everywhere in $\mathbb{R}$ and $F'(x) = f(x)$.

3.14 Complements A: Countable sets

Intuitively speaking, a countable set is characterized by the property that its elements can be counted one at a time (although the counting may never finish). Every element of the set is associated with a unique natural number.

First, we recall that given two sets A and B, we say that A is equipotent to B if there exists a bijection $f : A \longrightarrow B$. The existence of a bijection between A and B corresponds to the intuitive idea that A and B have the same number of elements. To say that equipotent sets have the same number of elements is in accordance with our intuition for finite sets. However, for infinite sets it can happen that $B \subset A$ (i.e., B is a proper subset of A) and yet B is equipotent to A. For example, the set of even numbers $\{2\,n \,|\, n \in \mathbb{N}\}$ is a proper subset of $\mathbb{N}$. Since the function

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \{2\,n \,|\, n \in \mathbb{N}\} \\ n &\mapsto 2\,n \end{aligned}$$

is bijective (and in fact the inverse is the function $f^{-1} : \{2\,n \,|\, n \in \mathbb{N}\} \longrightarrow \mathbb{N}$ that maps m to $\frac{m}{2}$), we find that these two sets are equipotent.

An infinite set is defined to be a set that is equipotent to at least one of its proper subsets. Hence, if A is a set, we say that A is infinite if there exists $B \subset A$ such that B is equipotent to A. We say that A is finite if it not infinite.

Theorem 3.69 *The set of natural numbers $\mathbb{N}$ is infinite.*

Proof The set $\{2n \mid n \in \mathbb{N}\}$ is a proper subset of $\mathbb{N}$ (i.e., $\{2n \mid n \in \mathbb{N}\} \subset \mathbb{N}$). Since the function

$$\begin{aligned} f : \mathbb{N} &\longrightarrow \{2n \mid n \in \mathbb{N}\} \\ n &\mapsto 2n \end{aligned}$$

is bijective (and in fact the inverse is the function $f^{-1} : \{2n \mid n \in \mathbb{N}\} \longrightarrow \mathbb{N}$ that maps m to $\frac{m}{2}$), we can conclude that $\mathbb{N}$ is infinite.

A set A is said to be:

countable if A is equipotent to $\mathbb{N}$. Hence, if A is countable, it is infinite (why?).

at most countable if A is finite or countable

uncountable if A is not finite and it is not countable.

Theorem 3.70 *Let $B \subset \mathbb{N}$. Then, B is at most countable.*

Proof If B is finite, the statement is true. Suppose that B is infinite. We define by induction a function $a : \mathbb{N} \longrightarrow B$. For $n = 0$, let $a_0 = \min B$. Suppose that a_i are defined for $i \in \{1, 2, \ldots, n\}$ and let $a_i \in B$ for $i \in \{1, 2, \ldots, n\}$. Since B is infinite, we have $B \neq \{a_0, a_1, \ldots, a_n\}$. Thus, $B \setminus \{a_0, a_1, \ldots, a_n\} \neq \emptyset$. We can therefore define $a_{n+1} = \min(B \setminus \{a_0, a_1, \ldots, a_n\})$. We show that a is strictly increasing. To this end, we show that $\forall\, n \in \mathbb{N}$ we have $a_{n+1} > a_n$. Let $n \in \mathbb{N}$. We have $a_{n+1} \in B \setminus \{a_0, a_1, \ldots, a_n\} \subseteq B \setminus \{a_0, a_1, \ldots, a_{n-1}\}$. Thus, $a_{n+1} \in B \setminus \{a_0, a_1, \ldots, a_{n-1}\}$. Since $a_n = \min(B \setminus \{a_0, a_1, \ldots, a_{n-1}\})$, we have $a_n \leq a_{n+1}$. Moreover, $a_{n+1} \neq a_n$. Thus, $a_n < a_{n+1}$. We have proved that a is strictly increasing, and therefore it is also injective. We prove that a is surjective. Let $k \in B$, and we prove that there exists $n \in \mathbb{N}$ such that $a_n = k$. Suppose by contradiction that for all $n \in \mathbb{N}$ we have $a_n \neq k$. We show by induction that $\forall\, n \in \mathbb{N}$ we have $a_n < k$. Indeed, by definition of a_0 we have $a_0 \leq k$. Since $a_0 \neq k$, we have also $a_0 < k$. Suppose $a_i < k$ for $i \in \{1, 2, \ldots, n\}$, and we show that $a_{n+1} < k$. Since $a_i \neq k$ for $i \in \{0, 1, 2, \ldots, n\}$, we find $k \in B \setminus \{a_0, a_1, \ldots, a_n\}$, so that $a_{n+1} \leq k$; since $a_{n+1} \neq k$, we obtain $a_{n+1} < k$. We have shown that for all $n \in \mathbb{N}$ the strict inequality $a_n < k$ holds. In particular, $a_k < k$. Since a is strictly increasing, for all $n \in \mathbb{N}$ we have $a_n \geq n$, so that $a_k \geq k$, which leads to a contradiction. Since a is injective and surjective, a is bijective. Thus, B is equipotent to $\mathbb{N}$.

The following result shows that a set with less elements than a countable set is at most countable. Intuitively, there are no infinite sets with a number of elements that is strictly smaller than the number of elements of a countable set.

Theorem 3.71 *Let A be a set. Let A be at most countable. Let B be a subset of A. Then, B is at most countable.*

Proof If A is countable, then this result follows directly from Theorem 3.70. If A is finite, the statement follows from Theorem 3.70.

Theorem 3.72 *Let A and B be sets. Then:*

(i) if A is finite and B is countable, then $A \cup B$ is countable;

(ii) if A and B are countable, then $A \cup B$ is countable;

(iii) if A and B are at most countable, then $A \cup B$ is at most countable.

Proof Let A be finite and let B be countable. We show that we can assume with no loss of generality that A and B are disjoint. Indeed, we have

$$A \cup B = (A \setminus B) \cup B, \quad (A \setminus B) \cap B = \varnothing, \quad \text{and} \quad A \setminus B \text{ is finite.}$$

Hence, if A and B are not disjoint, we can replace A with $A \setminus B$. Therefore, there exists $p \in \mathbb{N}$, a bijective function $f : \{1, 2, \dots, p\} \longrightarrow A$ and a bijective function $g : \mathbb{N} \longrightarrow B$. We define the function

$$\begin{array}{rcl} h : \mathbb{N} & \longrightarrow & A \cup B \\ n & \mapsto & \begin{cases} f(n+1) & \text{for } n \in \{0, 1, \dots, p-1\} \\ g(n-p) & \text{for } n \geq p, \end{cases} \end{array}$$

which is bijective. The injectivity is a consequence of the fact that f and g are injective and from the fact that A and B are disjoint. We prove the surjectivity. Let $k \in A \cup B$. We have $k \in A$ or $k \in B$. If $k \in A$, since f is surjective, there exists $\ell \in \{1, 2, \dots, p\}$ such that $f(\ell) = k$, so that $h(\ell - 1) = k$. If $k \in B$, since g is surjective, there exists $\mu \in \mathbb{N}$ such that $g(\mu) = k$, so that $h(\mu + p) = k$. Since h is bijective, $A \cup B$ is countable. This proves the first statement.

Let A and B be countable. Then

$$A \cup B = (A \setminus B) \cup B \quad \text{and} \quad (A \setminus B) \cap B = \varnothing.$$

Obviously $A \setminus B \subseteq A$. Hence, $A \setminus B$ is finite or countable. If it is finite, we are back to the first point of the theorem. Suppose that $A \setminus B$ is countable. Replacing A with $A \setminus B$, we can assume that A and B are disjoint. Therefore, there exist two bijective functions $f : \mathbb{N} \longrightarrow A$ and $g : \mathbb{N} \longrightarrow B$; we can define the function

$$\begin{array}{rcl} h : \mathbb{N} & \longrightarrow & A \cup B \\ n & \mapsto & \begin{cases} f(\frac{n}{2}) & \text{for even } n \\ g(\frac{n-1}{2}) & \text{for odd } n, \end{cases} \end{array}$$

so that $h(0) = f(0)$, $h(1) = g(0)$, $h(2) = f(1)$, $h(3) = g(1)$, $h(4) = f(2)$, $h(5) = g(2)$, ... From the bijectivity of f and g, and the fact that A and B have been assumed with no loss of generality to be disjoint, h is bijective. Therefore, $A \cup B$ is countable. This proves the second point.

The third point is a consequence of the fact that the cardinality of the union of disjoint finite sets is the sum of the cardinalities.

Corollary 3.8 *Let X and Λ be sets. Let $a : \Lambda \longrightarrow \mathcal{P}(X)$. Then*

- *if Λ is finite and, $\forall\, \lambda \in \Lambda$, $a(\lambda)$ is finite, then $\bigcup_{\lambda \in \Lambda} a(\lambda)$ is finite;*
- *if Λ is finite, $\Lambda \neq \varnothing$, and, $\forall\, \lambda \in \Lambda$, $a(\lambda)$ is countable, then $\bigcup_{\lambda \in \Lambda} a(\lambda)$ is countable;*
- *if Λ is finite and, $\forall\, \lambda \in \Lambda$, $a(\lambda)$ is at most countable, then $\bigcup_{\lambda \in \Lambda} a(\lambda)$ is at most countable.*

Proof If follows from Theorem 3.72, using induction on $n = \text{card}(\Lambda)$.

The following theorem shows that an infinite set has at least as many elements as a countable set. In other words, we can say that countable sets are the smallest among infinite sets. If we order the sets according to the number of their elements, the sets with fewer number of elements are the finite sets (ordered in turn according to their cardinality); the next sets that we find immediately after the finite sets are the countable sets. ⋆ We will see

in the sequel that the sets that are equipotent to $\mathbb{R}$ are infinite sets with a strictly larger number of elements than the number of elements of a countable set.

Theorem 3.73 *Let A be a set. Let A be infinite. Then, there exists a subset $B \subseteq A$ such that B is countable.*

Proof Since A is infinite, there exists $C \subset A$ such that C is equipotent to A. Therefore, there exists a bijection $f : A \longrightarrow C$. Since $C \neq A$, then $A \setminus C \neq \emptyset$. Thus, there exists $b \in A \setminus C$. Let a be the sequence of elements of A defined by induction from $a_0 = b$ and $a_{n+1} = f(a_n)$. We have $a_0 \in A \setminus C$, and for all $n \in \mathbb{N}$ such that $n \geq 1$, we have $a_n \in C$. It follows that for all $n \in \mathbb{N}$ such that $n \geq 1$ we have $a_n \neq a_0$. We show that a is injective. Suppose by contradiction that a is not injective. Let

$$M = \{n \in \mathbb{N} \,|\, \exists k \in \mathbb{N} : \; k < n \text{ and } a_k = a_n\}.$$

Since a is not injective, then $M \neq \emptyset$. We can define $m = \min M$. Thus, $m \geq 1$. Since $m \in M$, there exists $k \in \mathbb{N}$ such that $k < m$ and $a_k = a_m$. Since $m \geq 1$, we have $a_m \in C$. Therefore, $k \neq 0$. It follows that $k \geq 1$. From $a_k = f(a_{k-1})$ and $a_m = f(a_{m-1})$, since f is injective, we obtain $a_{k-1} = a_{m-1}$. Moreover, $k - 1 < m - 1$. Thus, $m - 1 \in M$, which implies that m is not the minimum of M; this leads to a contradiction.

We proved that a is injective. Let $B = a(\mathbb{N})$. The function

$$\begin{aligned} \mathbb{N} &\longrightarrow B \\ n &\mapsto a_n \end{aligned}$$

is bijective (it is injective because such is a as proved above, and it is surjective because the codomain coincides with the image). Then B is a countable set of A.

Theorem 3.74 *Let A be a set. Let $B \subseteq A$. Let B be at most countable. Let $A \setminus B$ be infinite. Then, A and $A \setminus B$ are equipotent.*

Proof Since $A \setminus B$ is infinite, from Theorem 3.73 there exists $N \subseteq A \setminus B$ such that N is countable. Since B is at most countable, $N \cup B$ is countable. Hence, $N \cup B$ and N are equipotent, and there exists a bijection $f : N \cup B \longrightarrow N$. We find

$$\begin{aligned} A &= (A \setminus B) \cup B \\ &= (((A \setminus B) \setminus N) \cup N) \cup B \\ &= ((A \setminus (N \cup B)) \cup N) \cup B \\ &= (A \setminus (N \cup B)) \cup (N \cup B) \end{aligned}$$

and $(A \setminus (N \cup B)) \cap (N \cup B) = \emptyset$. Since $A \setminus (N \cup B) \subseteq A \setminus B$,

$$x \in A \setminus (N \cup B) \;\Rightarrow\; x \in A \setminus B.$$

Since $N \subseteq A \setminus B$,

$$x \in N \cup B \;\Rightarrow\; f(x) \in A \setminus B.$$

Thus, we can define

$$\begin{aligned} g : A &\longrightarrow A \setminus B \\ x &\mapsto \begin{cases} x & \text{for } x \in A \setminus (N \cup B) \\ f(x) & \text{for } x \in N \cup B \end{cases} \end{aligned}$$

To prove the statement of the theorem, it is sufficient to show that g is bijective. We start proving that g is injective. Let $x, y \in A$ such that $x \neq y$. We prove that $g(x) \neq g(y)$. If $x, y \in A \setminus (N \cup B)$, we have $g(x) = x \neq y = g(y)$, so that $g(x) \neq g(y)$. If $x, y \in N \cup B$, we have $g(x) = f(x)$ and $g(y) = f(y)$. Since f is injective and $x \neq y$, we have $f(x) \neq f(y)$, so that $g(x) \neq g(y)$. Finally, let $x \in A \setminus (N \cup B)$ and $y \in N \cup B$. Then, $g(x) = x \in A \setminus (N \cup B)$ and $g(y) = f(y) \in N \cup B$. Thus, $g(x) \neq g(y)$. It follows that g is injective. We prove that it is surjective. Let $a \in A \setminus B$. We prove that there exists $x \in A$ such that $g(x) = a$. We have $A \setminus B = (A \setminus (N \cup B)) \cup N$, so that $a \in A \setminus (N \cup B)$ or $a \in N$. If $a \in A \setminus (N \cup B)$, taking $x = a$ gives $g(x) = a$. If $a \in N$, taking $x = f^{-1}(a)$ gives again $g(x) = a$. This shows that g is surjective.

The following result shows that the number of elements of an infinite set does not change if we add to that set another countable set.

Theorem 3.75 *Let A and B be sets. Let A be infinite. Let B be at most countable. Then, A and $A \cup B$ are equipotent.*

Proof We have $A \cup B = A \cup (B \setminus A)$, We have $B \setminus A \subseteq A \cup (B \setminus A)$ and $B \setminus A$ is at most countable. Moreover, since $(A \cup (B \setminus A)) \setminus (B \setminus A) = A$, it follows that $(A \cup (B \setminus A)) \setminus (B \setminus A)$ is infinite. Then, by Theorem 3.74, $A \cup (B \setminus A)$ and $(A \cup (B \setminus A)) \setminus (B \setminus A)$ are equipotent. Thus, $A \cup B$ and A are equipotent.

Theorem 3.76 *Let $r \in \mathbb{N}$. Then*

(i) $r \neq 0 \Rightarrow \{1, 2, \ldots, r\} \times \mathbb{N}$ *is countable;*

(ii) $\mathbb{N} \times \mathbb{N}$ *is countable.*

Proof Let $r \neq 0$. We have

$$\{1, 2, \ldots, r\} \times \mathbb{N} = \bigcup_{i \in \{1,2,\ldots,r\}} (\{i\} \times \mathbb{N}).$$

For every $i \in \{1, 2, \ldots, r\}$, the set $\{i\} \times \mathbb{N}$ is countable. Then, $\{1, 2, \ldots, r\} \times \mathbb{N}$ is countable from Corollary 3.8. This proves the first statement.

Now we order $\mathbb{N} \times \mathbb{N}$ according to the following scheme:

$$\begin{array}{ccccc}
(0,0) & (0,1) & (0,2) & (0,3) & \cdot \\
(1,0) & (1,1) & (1,2) & (1,3) & \cdot \\
(2,0) & (2,1) & (2,2) & (2,3) & \cdot \\
(3,0) & (3,1) & (3,2) & (3,3) & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot
\end{array}$$

We define the function from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ that maps the pair $(0, 0)$ to the natural number 0, the pair $(1, 0)$ to the natural number 1, the pair $(0, 1)$ to the natural number 2, the pair $(2, 0)$ to the natural number 3, the pair $(1, 1)$ to the natural number 4, the pair $(0, 2)$ to the natural number 5 and so forth. More precisely, since for

every $n, m \in \mathbb{N}$ the natural number $(n+m)(n+m+1)$ is even, we define

$$\begin{aligned} f : \mathbb{N} \times \mathbb{N} &\longrightarrow \mathbb{N} \\ (n, m) &\mapsto \frac{(n+m)(n+m+1)}{2} + m. \end{aligned}$$

From the scheme shown above it is clear that f is bijective. We now prove this point more rigorously. Let $(n_1, m_1), (n_2, m_2) \in \mathbb{N} \times \mathbb{N}$. Let $(n_1, m_1) \neq (n_2, m_2)$. To prove that f is injective, we show that $f(n_1, m_1) \neq f(n_2, m_2)$. Suppose first that $n_1 + m_1 \neq n_2 + m_2$. For example $n_1 + m_1 < n_2 + m_2$. Then, $n_1 + m_1 + 1 \leq n_2 + m_2$. Thus,

$$\begin{aligned} f(n_1, m_1) &= \frac{(n_1+m_1)(n_1+m_1+1)}{2} + m_1 \\ &< \frac{(n_1+m_1)(n_1+m_1+1)}{2} + n_1 + m_1 + 1 \\ &= \frac{(n_1+m_1+1)(n_1+m_1+2)}{2} \\ &\leq \frac{(n_2+m_2)(n_2+m_2+1)}{2} \\ &\leq \frac{(n_2+m_2)(n_2+m_2+1)}{2} + m_2 = f(n_1, m_2). \end{aligned}$$

Therefore $f(n_1, m_1) < f(n_2, m_2)$; in particular $f(n_1, m_1) \neq f(n_2, m_2)$. Let us now consider the case $n_1 + m_1 = n_2 + m_2$. Since we have assumed that $(n_1, m_1) \neq (n_2, m_2)$, it follows that $m_1 \neq m_2$. Since

$$\frac{(n_1+m_1)(n_1+m_1+1)}{2} = \frac{(n_2+m_2)(n_2+m_2+1)}{2},$$

we obtain

$$\frac{(n_1+m_1)(n_1+m_1+1)}{2} + m_1 \neq \frac{(n_2+m_2)(n_2+m_2+1)}{2} + m_2,$$

which leads to $f(n_1, m_1) \neq f(n_2, m_2)$.

We now prove that f is surjective. Let $p \in \mathbb{N}$. We prove that there exists $(n, m) \in \mathbb{N} \times \mathbb{N}$ such that $f(n, m) = p$. Let

$$A = \left\{ k \in \mathbb{N} \;\middle|\; \frac{k\,(k+1)}{2} \leq p \right\}.$$

Since $0 \in A$, then $A \neq \emptyset$. Let $k \in A$. Then $k + k^2 \leq 2\,p$, so that $k \leq k^2 + k \leq 2\,p$, which gives $A \subseteq \{i \in \mathbb{N} \,|\, i \leq 2\,p\}$. Hence, A is finite; thus, A admits a maximum. Let $q = \max A$. By definition of maximum, we find

$$\frac{q\,(q+1)}{2} \leq p < \frac{(q+1)(q+2)}{2}.$$

Indeed, $\frac{(q+1)(q+2)}{2} \leq p$ would imply that q is not the maximum of A. Thus,

$$\frac{(q+1)(q+2)}{2} - 1 \geq p.$$

It follows that

$$\frac{(q+1)(q+2)}{2} - 1 = \frac{(q+1)(q+2) - 2}{2} = \frac{(q+1)\,q + 2\,q}{2} = q + \frac{q\,(q+1)}{2}.$$

Therefore, $q + \frac{q(q+1)}{2} \geq p$. Since $q + \frac{q(q+1)}{2} \geq p$ and $\frac{q(q+1)}{2} \leq p$, we can define

$$n = q + \frac{q(q+1)}{2} - p \quad \text{and} \quad m = p - \frac{q(q+1)}{2}.$$

We have $n + m = q$, so that

$$f(n,m) = \frac{(n+m)(n+m+1)}{2} + m = \frac{q(q+1)}{2} + p - \frac{q(q+1)}{2} = p.$$

Hence, f is surjective. This proves the second statement.

Lemma 3.16 *Let A and B be finite sets. Then, $A \times B$ is finite and*

$$\operatorname{card}(A \times B) = \operatorname{card}(A)\ \operatorname{card}(B).$$

Theorem 3.77 *Let A and B be sets. Then:*

- *if A is finite and non-empty, and B is countable, then $A \times B$ is countable;*
- *if A and B are countable, then $A \times B$ is countable;*
- *if A and B are at most countable, then $A \times B$ is at most countable.*

Proof The first two points follow from Theorem 3.76. The third follows from Lemma 3.16.

The following theorem shows that the set of integers is countable. This is intuitive, since we can build the following enumeration $\varphi : \mathbb{N} \longrightarrow \mathbb{Q}$:

$$\varphi(0) = 0, \quad \varphi(1) = 1, \quad \varphi(2) = -1, \quad \varphi(3) = 2, \quad \varphi(4) = -2, \quad \varphi(5) = 3, \quad \varphi(6) = -3, \quad \ldots$$

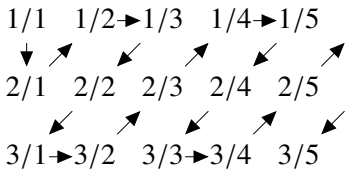
Theorem 3.78 *The set of integers $\mathbb{Z}$ is countable.*

Proof We can construct the function

$$\begin{aligned} \varphi : \mathbb{N} &\longrightarrow \mathbb{Z} \\ n &\mapsto \begin{cases} 0 & \text{if } k = 0 \\ \dfrac{k+1}{2} & \text{if } k \text{ is odd} \\ -\dfrac{k}{2} & \text{if } k \text{ is even.} \end{cases} \end{aligned}$$

It is easy to see that φ is a bijection.

We now show that $\mathbb{Q}$ is also countable. A possible enumeration of positive rational numbers is



Theorem 3.79 *The set of rational numbers $\mathbb{Q}$ is countable.*

Proof Recall that $\mathbb{Q} = \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z} \text{ and } n \in \mathbb{N}^* \right\}$. We can construct the function

$$\begin{aligned} \varphi : \mathbb{Z} \times \mathbb{N}^* &\longrightarrow \mathbb{Q} \\ (m, n) &\mapsto \frac{m}{n}. \end{aligned}$$

This function is bijective by definition of $\mathbb{Q}$.

Theorem 3.80 *The set of real numbers $\mathbb{R}$ is not countable.*

Proof We prove that the open interval $(0, 1)$ is not countable. From this, it is trivial to show that $\mathbb{R}$ is not countable. Suppose by contradiction that there exists a bijection

$$\begin{aligned} \varphi : \mathbb{N} &\longrightarrow (0, 1), \\ k &\mapsto a_{k,0}\, 10^{-1} + a_{k,1}\, 10^{-2} + a_{k,2}\, 10^{-3} + \ldots = \sum_{i=0}^{\infty} a_{k,i}\, 10^{-i-1}. \end{aligned}$$

where $a_{k,i}$ are the decimals of $\varphi(k)$. In other words,

$$\begin{aligned} \varphi(0) &= 0.a_{0,0}\, a_{0,1}\, a_{0,2}\, a_{0,3} \ldots \\ \varphi(1) &= 0.a_{1,0}\, a_{1,1}\, a_{1,2}\, a_{1,3} \ldots \\ \varphi(2) &= 0.a_{2,0}\, a_{2,1}\, a_{2,2}\, a_{2,3} \ldots \\ &\vdots \end{aligned}$$

We show that there exists $b \in (0, 1)$, written in decimal representation as

$$b = b_0\, 10^{-1} + b_1\, 10^{-2} + b_2\, 10^{-3} + \ldots = \sum_{i=0}^{\infty} b_i\, 10^{-i-1},$$

which is not in the image of b, which contradicts the assumption that φ is surjective. Let us choose b_i to be different from $a_{i,i}$ and different from 9. The reason why b_i should be different from 9 is that we want to avoid generating a number of the form $0.44999999\ldots$, because this is equal to 0.05: the number b we might end up generating would have different decimals than any other number in the list, but it would represent the same element of $(0, 1)$. In this way, we obtain a number b such that there exists no n such that $\varphi(n) = b$: the n-th decimal digit of $\varphi(n)$ is different from the n-th decimal digit of b by construction.

Theorem 3.81 *Let I be a proper interval of $\mathbb{R}$. Then, I and $\mathbb{R}$ are equipotent.*

Proof The function

$$\begin{aligned} \mathbb{R} &\longrightarrow (-1, 1) \\ x &\mapsto \frac{x}{1 + |x|} \end{aligned}$$

is bijective, so that $(-1,1)$ is equipotent to $\mathbb{R}$. Let $a, b \in \mathbb{R}$ with $a < b$. The function

$$\begin{aligned}(a,b) &\longrightarrow (-1,1)\\ x &\mapsto \frac{2}{b-a}x - \frac{a+b}{b-a}\end{aligned}$$

is bijective. Thus, (a,b) and $(-1,1)$ are equipotent. It follows also that (a,b) and $\mathbb{R}$ are equipotent. The function

$$\begin{aligned}(0,+\infty) &\longrightarrow \mathbb{R}\\ x &\mapsto \frac{x^2-1}{x}\end{aligned}$$

is bijective, so that $(0,+\infty)$ and $\mathbb{R}$ are equipotent. The function

$$\begin{aligned}(0,+\infty) &\longrightarrow (a,+\infty)\\ x &\mapsto a+x\end{aligned}$$

is bijective, so that $(a,+\infty)$ and $\mathbb{R}$ are equipotent. The function

$$\begin{aligned}(0,+\infty) &\longrightarrow (-\infty,a)\\ x &\mapsto a-x\end{aligned}$$

is bijective, so that $(-\infty,a)$ and $\mathbb{R}$ are equipotent. The statement is then proved for open intervals. If I is not open, the statement follows from Theorem 3.75.

Theorem 3.82 *Let I be a proper interval of $\mathbb{R}$. Then*

(i) $I \cap \mathbb{Q}$ is non-empty and it is countable;

(ii) $I \setminus \mathbb{Q}$ and $\mathbb{R}$ are equipotent.

Proof First, we prove that $I \cap \mathbb{Q}$ is non-empty. Since I is a proper interval, there exist $a, b \in I$ such that $a < b$. Thus, $[a,b] \subseteq I$. It is therefore sufficient to show that $[a,b] \cap \mathbb{Q} \neq \varnothing$. From the density of $\mathbb{Q}$ in $\mathbb{R}$, we know that

$$\forall\, x \in \mathbb{R},\ \forall\, \varepsilon > 0,\ \ \exists\, q \in \mathbb{Q} :\ \ |x-q| < \varepsilon.$$

Choosing $x = \frac{a+b}{2}$ and $\varepsilon = \frac{b-a}{2}$ gives the result.
We now show that $I \cap \mathbb{Q}$ is countable. From Theorem 3.72, $I \cap \mathbb{Q}$ is at most countable. To prove the first statement, it is suffices to show that $I \cap \mathbb{Q}$ is not finite. There exists $a, b \in I$ such that $a < b$. Since $(a,b) \subseteq I$, it is enough to prove that $(a,b) \cap \mathbb{Q}$ is not finite. Suppose by contradiction that $(a,b) \cap \mathbb{Q}$ is finite. Let $m = \min((a,b) \cap \mathbb{Q})$. Thus, $m > a$ and $(a,m) \cap \mathbb{Q} = \varnothing$. This contradicts the fact that $I \cap \mathbb{Q}$ is non-empty.

Let us now prove the second statement. We have $I \setminus \mathbb{Q} = I \setminus (I \cap \mathbb{Q})$. From Theorem 3.74, I and $I \setminus \mathbb{Q}$ are equipotent. From Theorem 3.81, we have that $I \setminus \mathbb{Q}$ and $\mathbb{R}$ are equipotent. This proves the second statement.

4 L^p spaces

L^p spaces are, loosely speaking, spaces of functions defined using a generalization of the p-norm for finite-dimensional vector spaces. These spaces are also called *Lebesgue spaces* (named after Henri Lebesgue). Lebesgue spaces are used also in physics, statistics, and engineering.

4.1 Identification of measurable functions

We introduced the following norms in $\mathbb{R}^n$, for an arbitrary vector $\boldsymbol{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$:

- the $\|\cdot\|_1$ norm: $\|\boldsymbol{x}\|_1 = |x_1| + |x_2| + \dots + |x_n|$;
- the Euclidean norm: $\|\boldsymbol{x}\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_n|^2}$;
- the infinity norm $\|\boldsymbol{x}\|_\infty = \max_{k=1,\dots,n} |x_k|$.

We have also seen that these norms are particular cases of the so-called p-norm, defined by

$$\|\boldsymbol{x}\|_p = \left(|x_1|^p + |x_2|^p + \dots + |x_n|^p\right)^{\frac{1}{p}}.$$

The vector space $\mathbb{R}^n$ is turned into a complete normed space by any of these norms, because $\mathbb{R}^n$ is finite dimensional (and in fact we have seen that in a final dimensional vector normed space all Cauchy sequences are convergent).

We then introduced analogous norms for functions. In particular, for the space of continuous functions $C([a,b];\mathbb{R})$ over the interval $[a,b]$ with $a, b \in \mathbb{R}$ and $a < b$, we defined

- the $\|\cdot\|_1$ norm: $\|f\|_1 = \displaystyle\int_a^b |f(x)|\,\mathrm{d}x$;
- the Euclidean norm: $\|f\|_2 = \sqrt{\displaystyle\int_a^b |f(x)|^2\,\mathrm{d}x}$;
- the infinity norm $\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$.

These norms can be thought of as particular cases of the p-norm

$$\|f\|_p = \Big(\int_a^b |f(x)|^p\,\mathrm{d}x\Big)^{\frac{1}{p}}.$$

We have seen, however, that the normed space $C([a,b];\mathbb{R})$ is not complete with respect to all these norms. For example, it is not complete with respect to the 1-norm and to the 2-norm, but it is complete if endowed with the norm $\|\cdot\|_\infty$. We already hinted at the fact that the incompleteness is due to the fact that the vector space does not contain the limits of all the Cauchy sequences; indeed, for the norms $\|\cdot\|_1$ and $\|\cdot\|_2$, such limits may not be continuous. A possible solution is to replace the set of continuous functions with the set of Lebesgue summable functions (as we mentioned, the set of Riemann integrable functions is not rich enough to lead to a complete normed space). However, now we have a problem: are we sure that, by replacing $C([a,b];\mathbb{R})$ with the set of Lebesgue summable functions the norm $\|\cdot\|_p$ is still a norm? We proved that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms in $C([a,b];\mathbb{R})$, but the proof made explicit use of the continuity. Abandoning this assumption creates an issue that we need to address.

Consider for the sake of argument the norm $\|\cdot\|_1$. A function which is not identically zero only because it has non-zero values on a measure zero set (for example at a single point of the domain) still has zero norm, thus contradicting the axiom of positivity that defines every norm. For example, the two functions

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 0 \end{aligned} \qquad \text{and} \qquad \begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{for } x \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & \text{for } x \in \mathbb{Q} \end{cases} \end{aligned}$$

satisfy

$$\int_{\mathbb{R}} |f(x)|\,\mathrm{d}x = \int_{\mathbb{R}} |g(x)|\,\mathrm{d}x = 0,$$

since $\mathbb{Q}$, being countable, has Lebesgue measure zero in $\mathbb{R}$, and so $\|f\|_1 = \|g\|_1 = 0$. However, only f is identically zero. Therefore, if we remove the assumption of continuity, it is no longer true that that $\|f\|_1 = 0$ implies $f = 0$.

There are two ways to address this issue. The first is to weaken the notion of norm, by removing the positivity axiom "$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|\boldsymbol{x}\| = 0 \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{0}$". This leads in particular to the notion of *pseudo-metric* and *seminorm* (or *pseudo-norm*). Another approach, which is the one we will pursue, is to **identify functions that differ only by the values they take on sets of Lebesgue zero measure**, i.e., functions which coincide almost everywhere. In the example above, the two functions f and g are identified as the same object, being equal almost everywhere. We recall that a property which depends on a variable $\boldsymbol{x}$ holds "almost everywhere" if the set of values of $\boldsymbol{x}$ for which said property does not hold has null Lebesgue measure. More in general, we recall that, if $\boldsymbol{A} \in \mathcal{M}_n$, a proposition $p(\boldsymbol{x})$ that depends on a variable $\boldsymbol{x} \in \mathbb{R}^n$ is true *almost everywhere* (in short *a.e.*) in $\boldsymbol{A}$, or that $p(\boldsymbol{x})$ is true for almost every $\boldsymbol{x} \in \boldsymbol{A}$ if the set $\{\boldsymbol{x} \in \boldsymbol{A} \mid p(\boldsymbol{x}) \text{ is false}\}$ is measurable and has null Lebesgue measure.

The first task is to be able to compare functions which are not defined over the same domain, but on domains that differ by a zero measure set. For example, we want to be able to say that the functions

$$\begin{aligned} f_1 : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned} \qquad \text{and} \qquad \begin{aligned} f_2 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 7 & \text{for } x = 0 \\ \dfrac{1}{x} & \text{for } x \neq 0 \end{cases} \end{aligned}$$

are equivalent even if they are defined over different domains, and, to this end, we need to be able to say that both these functions are defined almost everywhere in $\mathbb{R}$ (even if f_2) is defined for every $x \in \mathbb{R}$.

Definition 4.1 *Let $\boldsymbol{B} \in \mathcal{M}_n$. Let $f : \boldsymbol{B} \longrightarrow \mathbb{C}$. Let $\boldsymbol{N} \in \mathcal{M}_n$ be a zero measure set (i.e., $\mu_n(\boldsymbol{N}) = 0$). We say that f is* defined almost everywhere in $\boldsymbol{B} \cup \boldsymbol{N}$.

We will say that a function f is defined almost everywhere in $\boldsymbol{A}$ if the set $\boldsymbol{A}$ is measurable and if there exists a zero measure subset N of $\boldsymbol{A}$ such that $\boldsymbol{A} \setminus N$ is the domain of f, i.e., such that $f : \boldsymbol{A} \setminus \boldsymbol{N} \longrightarrow \mathbb{C}$. For example, we can say that the function $f : \mathbb{R}^* \longrightarrow \mathbb{R}$ given by $f(x) = 1/x$ is defined almost everywhere in $\mathbb{R}$. Notice that in Definition 4.1 we assumed, to fix ideas, that the codomain of the functions is $\mathbb{C}$. We will maintain this assumption in the sequel. However, no significant differences arise if we consider, instead, $\mathbb{R}$, or $[-\infty, +\infty]$, or even $\mathbb{C} \cup \{\infty\}$ as codomains.

Definition 4.2 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let f, g be measurable functions defined almost everywhere in $\boldsymbol{A}$. We say that f is equivalent to g, and we write $f \sim g$, if $f(\boldsymbol{x}) = g(\boldsymbol{x})$ almost everywhere in $\boldsymbol{A}$.*

First, we note that Definition 4.2 is well posed: if f and g are measurable functions defined almost everywhere in $\boldsymbol{A}$, there exist two zero measure sets $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ such that $f : \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C}$ and $g : \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}$. For every $\boldsymbol{x} \in A \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2)$, both $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are defined, and can be compared, and the set $\boldsymbol{N}_1 \cup \boldsymbol{N}_2$ has measure zero because the countable union of zero measure sets has measure zero.

Exercise 4.1

Show that the "almost everywhere" equivalence $\sim$ is an equivalence relation. In other words, given three arbitrary measurable functions f, g, h defined almost everywhere in $A \in \mathcal{M}_n$, show that

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

(ii) $f \sim g$ implies $g \sim f$ (symmetric property);

(iii) $f \sim g$ and $g \sim h$ imply $f \sim h$ (transitive property).

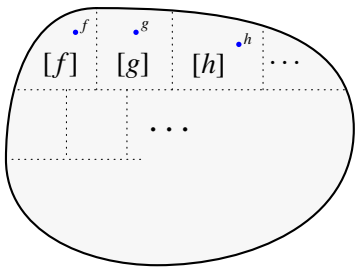
[Solution: (4.1)]

From Exercise 4.1, we learned that $\sim$ is an equivalence relation. Thus, given a measurable function f defined almost everywhere in a measurable subset A of $\mathbb{R}^n$, we can define the equivalence class $[f]$ of f as

$$[f] \stackrel{\text{def}}{=} \{g |\ g \text{ is defined a.e. in } \boldsymbol{A},\ g \text{ is measurable and } \ g \sim f\}.$$

The equivalence classes with respect to this equivalence relation are still called, with some abuse of language, "measurable functions".

An equivalence relation – such as the one we introduced here – allows us to partition the whole set (in this case the set of measurable functions from $\boldsymbol{A} \subseteq \mathbb{R}^n$ to $\mathbb{C}$) into equivalence classes (or *cosets*). A partition means that the union of all equivalence classes is the whole set, and the equivalence classes are pairwise disjoint. Visually, we can imagine the set of measurable functions $\boldsymbol{A} \longrightarrow \mathbb{C}$ to be represented as the regions below:



Remark 4.1 If we denote by $\mathfrak{M}(\boldsymbol{A}; \mathbb{C})$ the set of measurable functions defined almost everywhere in $\boldsymbol{A}$, then $\mathfrak{M}(\boldsymbol{A}; \mathbb{C})/\sim$ represents the quotient space with respect to the equivalent relation $\sim$, i.e., the set of equivalence classes obtained by grouping together functions that are equal almost everywhere. Let us now consider the mapping $f \mapsto [f]$, where $f \in \mathfrak{M}(\boldsymbol{A}; \mathbb{C})$. The image of this mapping is the set of all possible equivalent classes, i.e, $\mathfrak{M}(\boldsymbol{A}; \mathbb{C})/\sim$. It follows that the function

$$\begin{aligned} \pi : \mathfrak{M}(\boldsymbol{A}; \mathbb{C}) &\longrightarrow \mathfrak{M}(\boldsymbol{A}; \mathbb{C})/\sim \\ f &\mapsto [f] \end{aligned}$$

is surjective. This function is called *canonical projection*.

Theorem 4.1 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let f, g, f_1, f_2, g_1, g_2 be measurable functions defined almost everywhere in $\boldsymbol{A}$. Let $c \in \mathbb{C}$. Then:*

(i) $f_1 \sim g_1$ *and* $f_2 \sim g_2 \quad \Rightarrow \quad f_1 + f_2 \sim g_1 + g_2$;

(ii) $f \sim g \quad \Rightarrow \quad c\, f \sim c\, g$;

(iii) *if* $f \sim g$ *and* $h : \mathbb{C} \longrightarrow \mathbb{C}$ *is continuous, then* $h \circ f \sim h \circ g$;

(iv) $f \sim g \quad \Rightarrow \quad \mathfrak{Re}\{f\} \sim \mathfrak{Re}\{g\}$ *and* $\mathfrak{Im}\{f\} \sim \mathfrak{Im}\{g\}$;

(v) *if $f \sim g$, f, g are real-valued and f is measurable, then g is also measurable;*

(vi) *if $f \sim g$ and f is summable, then g is summable, and*

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Proof The proof of *(i-ii)* is left as an exercise (see Exercise 4.2 and Exercise 4.3). We prove *(iii)*. Since f and g are defined almost everywhere in $\boldsymbol{A}$, there exist two null sets $\boldsymbol{N}_1, \boldsymbol{N}_2 \subseteq \mathbb{R}^n$ such that

$$\begin{aligned} f &: \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C} \\ g &: \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}. \end{aligned}$$

The continuity of h ensures that, if f and g are measurable, then $h \circ f$ and $h \circ g$ are measurable (see Corollary 3.2). Moreover,

$$\{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, h(f(\boldsymbol{x})) \neq h(g(\boldsymbol{x}))\} \subseteq \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}. \tag{4.1}$$

Since the set in the right hand-side of (4.1) has measure zero, the set in the left hand-side of (4.1) has also measure zero. This implies that $h \circ f \sim h \circ g$.

We prove *(v)*. Let

$$\begin{aligned} f &: \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C} \\ g &: \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}, \end{aligned}$$

where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ have measure zero, and define

$$\boldsymbol{N} = \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}.$$

Since $f \sim g$, the set $\boldsymbol{N}$ is measurable and has measure 0. Since f is assumed to be measurable in its domain $\boldsymbol{A} \setminus \boldsymbol{N}_1$, it is also measurable in the subset $\boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N})$ of its domain, which is itself measurable, see Theorem 3.43. Since g coincides with f in $\boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N})$, it is measurable in $\boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N})$. Since $\boldsymbol{N}_1 \cup \boldsymbol{N}$ has zero measure, then g is also measurable.

We prove *(vi)*. Let f be defined over $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and g be defined over $\boldsymbol{A} \setminus \boldsymbol{N}_2$, where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are null sets. Now, let $N_3 = \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}$; from our assumption, we have $\mu_n(N_3) = 0$. Let $\boldsymbol{N} = \boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3$, which also satisfies $\mu_n(\boldsymbol{N}) = 0$. We now define the function

$$\begin{aligned} \tilde{f} : \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \begin{cases} f(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_1 \\ \text{arbitrary} & \text{if } \boldsymbol{x} \in \boldsymbol{N}_1. \end{cases} \end{aligned}$$

Since $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and $\boldsymbol{N}_1$ are disjoint and their union gives $\boldsymbol{A}$, using Theorem 3.43 and Theorem 3.48, we find

$$\int_{\boldsymbol{A}} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}_1} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{N}_1} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}_1} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}.$$

Now, we observe that $\boldsymbol{N}_2 \subseteq \boldsymbol{N}$, and, moreover,

$$(\boldsymbol{A} \setminus \boldsymbol{N}) \cup (\boldsymbol{N} \setminus \boldsymbol{N}_2) = \boldsymbol{A} \setminus \boldsymbol{N}_2 \quad \text{and} \quad (\boldsymbol{A} \setminus \boldsymbol{N}) \cap (\boldsymbol{N} \setminus \boldsymbol{N}_2) = \varnothing,$$

and since $\boldsymbol{N} \setminus \boldsymbol{N}_2$ also has Lebesgue measure zero, we find

$$\begin{aligned}\int_{\boldsymbol{A}} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}\setminus\boldsymbol{N}} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{N}} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\boldsymbol{A}\setminus\boldsymbol{N}} |g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}} |g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{N}\setminus\boldsymbol{N}_2} |g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}_2} |g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

It follows that

$$\int_{\boldsymbol{A}\setminus\boldsymbol{N}_1} |\tilde{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}_2} |g(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x},$$

so that, since f is summable, such is also g. Using the same argument, we find that

$$\int_{\boldsymbol{A}\setminus\boldsymbol{N}_1} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}\setminus\boldsymbol{N}_2} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Exercise 4.2 Let f_1, f_2, g_1, g_2 be measurable and defined almost everywhere over a measurable set $A \subseteq \mathbb{R}^n$, and with a codomain $\mathbb{C}$. Let these functions satisfy $f_1 \sim g_1$ and $f_2 \sim g_2$. Show that $f_1 + f_2 \sim g_1 + g_2$. [Solution: (4.2)]

Exercise 4.3 Let $\boldsymbol{A} \in \mathcal{M}_n$. Let f, g be measurable functions defined almost everywhere in $\boldsymbol{A}$, taking values in $\mathbb{C}$, and such that $f \sim g$. Show that for all $c \in \mathbb{C}$ we have $c\,f \sim c\,g$. [Solution: (4.3)]

Thanks to the last point of Theorem 4.1, if $\boldsymbol{A} \in \mathcal{M}_n$ and f is a summable functions defined almost everywhere in $\boldsymbol{A}$, it makes sense to define its Lebesgue integral over A as the Lebesgue integral over an arbitrary continuation of f over $\boldsymbol{A}$. In other words, if $\boldsymbol{N}$ is a null set such that $\boldsymbol{A} \setminus \boldsymbol{N}$ is the domain of f, then we can define an arbitrary continuation

$$\begin{aligned}\tilde{f} : \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \begin{cases} f(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \\ \text{arbitrary} & \text{if } \boldsymbol{x} \in \boldsymbol{N}. \end{cases}\end{aligned}$$

From *(vi)* in Theorem 4.1, since f and $\tilde{f}$ are equivalent, the summability of f over $\boldsymbol{A} \setminus \boldsymbol{N}$ is equivalent to the summability of $\tilde{f}$ over A. Thus, we can define

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \overset{\text{def}}{=} \int_{\boldsymbol{A}} \tilde{f}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{4.2}$$

With some abuse of language, we will say that f is summable over $\boldsymbol{A}$ if $\tilde{f}$ is summable over $\boldsymbol{A}$.

4.1.1 Operations on equivalent classes

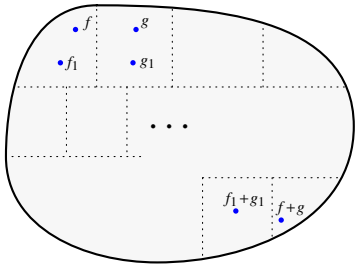
From Theorem 4.1, if f and g are measurable functions defined almost everywhere in $\boldsymbol{A}$, the following expressions are well defined:

$$[f] + [g] \overset{\text{def}}{=} [f + g] \tag{4.3}$$

and

$$\lambda\,[f] \overset{\text{def}}{=} [\lambda\, f]. \tag{4.4}$$

In fact, consider (4.3). Consider $f_1 \sim f$ and $g_1 \sim g$. Hence, f_1 lives in $[f]$ and g_1 lives in $[g]$. However, in general, $f_1 + g_1$ is different from $f + g$. We would be in trouble if, however, $f_1 + g_1$ did not live in the same equivalence class as $f + g$, i.e., if we were in the situation shown below:



If this was the case, the result of adding $[f]$ and $[g]$ would depend on which functions are chosen from the sets $[f]$ and $[g]$. Fortunately, everything works well because $f_1 \sim f$ and $g_1 \sim g$ imply $f_1 + g_1 \sim f + g$, as point *(i)* in Theorem 4.1 has shown. Similar considerations apply to $\lambda\,[f]$: try to elaborate on this case, which follows along the same lines.

Theorem 4.2 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let f, g, f_1, g_1 be functions defined almost everywhere in $\boldsymbol{A}$. Let f and g be real for almost every $\boldsymbol{x} \in \boldsymbol{A}$ and let $f(\boldsymbol{x}) \leq g(\boldsymbol{x})$ almost everywhere. Let $f_1 \sim f$ and $g_1 \sim g$. Then f_1 and g_1 are real for almost every $\boldsymbol{x} \in \boldsymbol{A}$ and $f_1(\boldsymbol{x}) \leq g_1(\boldsymbol{x})$ almost everywhere.*

Proof Let $\boldsymbol{N}_1, \boldsymbol{N}_2, \boldsymbol{N}_3, \boldsymbol{N}_4$ be null sets in $\mathbb{R}^n$ such that

$$\begin{aligned} f &: \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C} \\ g &: \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C} \\ f_1 &: \boldsymbol{A} \setminus \boldsymbol{N}_3 \longrightarrow \mathbb{C} \\ g_1 &: \boldsymbol{A} \setminus \boldsymbol{N}_4 \longrightarrow \mathbb{C}. \end{aligned}$$

We define the sets

$$\begin{aligned} \boldsymbol{N}_a &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_3) \,|\, f_1(\boldsymbol{x}) \neq f(\boldsymbol{x})\} \\ \boldsymbol{N}_b &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_1 \,|\, f(\boldsymbol{x}) \notin \mathbb{R}\} \\ \boldsymbol{N}_c &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_2 \cup \boldsymbol{N}_4) \,|\, g_1(\boldsymbol{x}) \neq g(\boldsymbol{x})\} \\ \boldsymbol{N}_d &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_2 \,|\, g(\boldsymbol{x}) \notin \mathbb{R}\} \\ \boldsymbol{N}_e &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \in \mathbb{R},\, g(\boldsymbol{x}) \in \mathbb{R},\; f(\boldsymbol{x}) \not\leq g(\boldsymbol{x})\}, \end{aligned}$$

which all have zero measure by assumption. For all $\boldsymbol{x} \in (\boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_3)) \setminus (\boldsymbol{N}_a \cup \boldsymbol{N}_b) = \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_a \cup \boldsymbol{N}_b)$, we have $f(\boldsymbol{x}) = f_1(\boldsymbol{x})$ and $f(\boldsymbol{x}) \in \mathbb{R}$. Since the finite union of null sets $\boldsymbol{N}_1 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_a \cup \boldsymbol{N}_b$ is a null set, this implies that $f_1(\boldsymbol{x}) \in \mathbb{R}$ almost everywhere in $\boldsymbol{A}$. Similarly, for all $\boldsymbol{x} \in (\boldsymbol{A} \setminus (\boldsymbol{N}_2 \cup \boldsymbol{N}_4)) \setminus (\boldsymbol{N}_c \cup \boldsymbol{N}_d) = \boldsymbol{A} \setminus (\boldsymbol{N}_2 \cup \boldsymbol{N}_4 \cup \boldsymbol{N}_c \cup \boldsymbol{N}_d)$, we have $g(\boldsymbol{x}) = g_1(\boldsymbol{x})$ and $g(\boldsymbol{x}) \in \mathbb{R}$, and since $\boldsymbol{N}_2 \cup \boldsymbol{N}_4 \cup \boldsymbol{N}_c \cup \boldsymbol{N}_d$ has measure zero we conclude that $g_1(\boldsymbol{x}) \in \mathbb{R}$ almost everywhere in $\boldsymbol{A}$.

Finally, for all $\boldsymbol{x} \in (\boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4)) \setminus (\boldsymbol{N}_a \cup \boldsymbol{N}_b \cup \boldsymbol{N}_c \cup \boldsymbol{N}_d \cup \boldsymbol{N}_e) = \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4 \cup \boldsymbol{N}_a \cup \boldsymbol{N}_b \cup \boldsymbol{N}_c \cup \boldsymbol{N}_d \cup \boldsymbol{N}_e)$ we have

$$\mathbb{R} \ni f_1(\boldsymbol{x}) = f(\boldsymbol{x}) \leq g(\boldsymbol{x}) = g_1(\boldsymbol{x}) \in \mathbb{R}.$$

Since $\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4 \cup \boldsymbol{N}_a \cup \boldsymbol{N}_b \cup \boldsymbol{N}_c \cup \boldsymbol{N}_d \cup \boldsymbol{N}_e$ is the finite union of null sets, it is a null set. Thus, $f_1 \leq g_1$ almost everywhere.

Up to this point, we have seen that some definitions and properties, originally introduced for functions, are

invariant under changes over null measure sets. For example, we saw in Theorem 4.1 that if f is summable and g is equivalent to f, then g is also summable and their integrals over $\boldsymbol{A}$ are equal. A list of operations and properties, some of which we have already proved to be invariant under the identification of equivalent functions, are:

- the usual algebraic operations of addition and multiplication by a scalar;
- the composition $h \circ f$ with a continuous function h;
- the real and imaginary parts (this is a particular case of the former);
- the positive and negative parts in the real case;
- the powers with natural exponent, the modulus and its powers with real exponent;
- the integral;
- the restriction to a set with positive measure;
- the comparison of functions in the form $f(\boldsymbol{x}) \leq g(\boldsymbol{x})$ a.e. (or similar ones, but not an inequality involving every point of the domain). In particular, the notions of upper/lower bound, where we say that for example an upper bound of f is a constant M such that $f(\boldsymbol{x}) \leq M$ a.e.;
- the convergence almost everywhere (but the pointwise convergence is not invariant).

A concept that fails to be consistent with the identification of equivalent functions is the value of a function at a specific point $\boldsymbol{x}_0$. If we change f just at $\boldsymbol{x}_0$, we obtain another function g that can be identified with f (since every point is a null measure set), but $f(\boldsymbol{x}_0) \neq g(\boldsymbol{x}_0)$. It follows that with this identification we cannot talk about the value of a function at a specific point.
We can assign, to each equivalence class, the value that the equivalence class takes for a specific, privileged element contained in it, which is called *representative* of the equivalence class. Let us see now if an equivalence class $[f]$ contains a special element, that we can use as representative for the entire class. If the equivalence class $[f]$ contains a continuous function, we can consider as the value of the function f at a point $\boldsymbol{x}_0 \in \boldsymbol{A}$ the value of that continuous function at that point, and in this way we have solved the problem of the lack of pointwise values of the elements of each equivalence class. Obviously, for this to be possible we have to be able to prove that if the equivalence class $[f]$ contains a continuous function, it must be unique, as the following result shows.

Theorem 4.3 *Let $\boldsymbol{A} \in \mathcal{M}_n$ be open. Let f be defined almost everywhere in $\boldsymbol{A}$. If there exists a continuous function $g : \boldsymbol{A} \longrightarrow \mathbb{C}$ equal to f almost everywhere in $\boldsymbol{A}$, it is unique.*

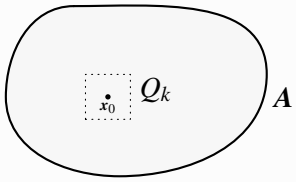
Proof Let $\boldsymbol{N}$ to be a null set, such that $f : \boldsymbol{A} \setminus \boldsymbol{N} \longrightarrow \mathbb{C}$. Let $g : \boldsymbol{A} \longrightarrow \mathbb{C}$ and $h : \boldsymbol{A} \longrightarrow \mathbb{C}$ be continuous functions equal to f almost everywhere in $\boldsymbol{A}$. Let

$$\begin{aligned} \boldsymbol{N}_1 &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \,|\, g(\boldsymbol{x}) \neq f(\boldsymbol{x})\} \\ \boldsymbol{N}_2 &= \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \,|\, h(\boldsymbol{x}) \neq f(\boldsymbol{x})\}. \end{aligned}$$

The sets $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ have measure zero, by assumption. We show that $g = h$. This is the same as showing that, given a generic point $\boldsymbol{x}_0 \in \boldsymbol{A}$, we have $g(\boldsymbol{x}_0) = h(\boldsymbol{x}_0)$. We denote by $x_{0,j}$ the j-th component of $\boldsymbol{x}_0$, so that $\boldsymbol{x}_0 = (x_{0,1}, \ldots, x_{0,n})$. Since $\boldsymbol{A}$ is open, there exists a sufficiently large $k \in \mathbb{N}^*$ such that the hypercube

$$\boldsymbol{Q}_k = \prod_{j=1}^{n} \left[x_{0,j} - \frac{1}{k}, x_{0,j} + \frac{1}{k} \right]$$

is entirely contained in $\boldsymbol{A}$ (see Exercise 4.4).



The hypercube $\boldsymbol{Q}_k$ has non-zero measure for any $k \in \mathbb{N}^*$. In fact,

$$\mu_n(\boldsymbol{Q}_k) = \left(\frac{2}{k}\right)^n > 0.$$

Since $\boldsymbol{N} \cup \boldsymbol{N}_1 \cup \boldsymbol{N}_2$ is a null set, being the finite union of null sets, the measure of $\boldsymbol{Q}_k \setminus (\boldsymbol{N} \cup \boldsymbol{N}_1 \cup \boldsymbol{N}_2)$ is equal to the measure of $\boldsymbol{Q}_k$, which is positive. Therefore, the set $\boldsymbol{Q}_k \setminus (\boldsymbol{N} \cup \boldsymbol{N}_1 \cup \boldsymbol{N}_2)$ is not empty (the measure of the empty set is 0). Hence, for every $k \in \mathbb{N}^*$, there exists $\boldsymbol{x}_k \in \boldsymbol{Q}_k \setminus (\boldsymbol{N} \cup \boldsymbol{N}_1 \cup \boldsymbol{N}_2)$. We therefore have $f(\boldsymbol{x}_k) = g(\boldsymbol{x}_k)$ and $f(\boldsymbol{x}_k) = h(\boldsymbol{x}_k)$, so that also $g(\boldsymbol{x}_k) = h(\boldsymbol{x}_k)$. Clearly,

$$\lim_{k \to \infty} \boldsymbol{x}_k = \boldsymbol{x}_0,$$

and, since g and h are continuous, using the sequential continuity theorem we obtain

$$g(\boldsymbol{x}_0) = \lim_{k \to \infty} g(\boldsymbol{x}_k) = \lim_{k \to \infty} h(\boldsymbol{x}_k) = h(\boldsymbol{x}_0).$$

Exercise 4.4

Let A be an open set in $\mathbb{R}^n$. Let $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2}, \ldots, x_{0,n}) \in A$. Show that there exists a sufficiently large value of $k \in \mathbb{N}^*$ such that the hyper-cube

$$Q_k = \prod_{j=1}^{n} \left[x_{0,j} - \frac{1}{k}, x_{0,j} + \frac{1}{k} \right]$$

is entirely contained in A. [Solution: (4.4)]

Example *4.1*

Consider the function

$$\begin{aligned} f : \mathbb{R} \setminus \{0\} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\sin x}{x}. \end{aligned}$$

This function is defined almost everywhere in $\mathbb{R}$. Moreover, f is equal almost everywhere to a continuous function in $\mathbb{R}$, because there exists a continuation of f to $\mathbb{R}$:

$$\begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases} \end{aligned}$$

Hence, we can say that the value of $[f]$ at 0 is 1.

Example *4.2*

Consider the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x = \frac{1}{k} \text{ for some } k \in \mathbb{N}^* \\ 2 & \text{otherwise} \end{cases} \end{aligned}$$

This function is equal almost everywhere to the continuous function

$$\begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 2. \end{aligned}$$

In fact, every countable set (such as $\mathbb{N}^*$) has Lebesgue measure zero. Hence, the value of $[f]$ at 0 is 2.

Unfortunately, not every equivalence class in the set of measurable functions contains a continuous function, so this is only a partial solution to the impossibility of assigning a value to an equivalence class at a point $\boldsymbol{x} \in \boldsymbol{A}$. Can you construct an example of an equivalence class which does not contain a continuous function?

Exercises

Exercise 4.5 Let f and g be two complex-valued functions defined almost everywhere on a measurable set $\boldsymbol{A} \subseteq \mathbb{R}^n$. Show that if f is equal almost everywhere to g, then the real part of f is equal almost everywhere to the real part of g. [Solution: (4.5)]

Exercise 4.6 Let $p \in [1, +\infty]$. We know that it is possible to define the sum and the scalar product in $L^p(\boldsymbol{A}; \mathbb{C})$, where $\boldsymbol{A} \in \mathcal{M}_n$: indeed, we showed that $L^p(\boldsymbol{A}; \mathbb{C})$ is a vector space with respect to these operations, by proving that these operations are well defined. Establish if it possible to define also the product of elements in $L^p(\boldsymbol{A}; \mathbb{C})$ as follows: for every $[f], [g] \in L^p(\boldsymbol{A}; \mathbb{C})$,

$$[f] \cdot [g] \stackrel{\text{def}}{=} [f\, g],$$

where the product $f\, g$ is defined in the obvious way. [Solution: (4.6)]

4.2 L^p spaces

We are now ready to introduce the Lebesgue spaces L^p. In this notation, p is a real number greater or equal to 1.[1]

Definition 4.3 [L^p SPACES]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $p \geq 1$. We define

$$L^p(\boldsymbol{A}; \mathbb{C}) = \left\{ [f] \,\middle|\, f \text{ is defined a.e. in } \boldsymbol{A}, \text{ measurable and such that } \int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty \right\}.$$

The set $L^p(\boldsymbol{A}; \mathbb{C})$ is often denoted simply as $L^p(\boldsymbol{A})$, without distinguishing the real from the complex case. If the domain is clear from the context, it is frequent to find also the notation L^p.

We should make sure that the set $L^p(\boldsymbol{A}; \mathbb{C})$ is well defined. Indeed, if for two equivalent functions f and f_1 we had $\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty$ and $\int_{\boldsymbol{A}} |f_1(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} = +\infty$ we would be in trouble: whether $[f]$ is an element or not of $L^p(\boldsymbol{A}; \mathbb{C})$ would depend on the particular element chosen from $[f]$, and the definition of $L^p(\boldsymbol{A}; \mathbb{C})$ would become ambiguous. To make sure that this definition is well posed, we need to show that if $\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty$ and $g \sim f$, then $\int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty$, or, stated differently, that $f \sim g$ implies

[1] In the case $p < 1$, some fundamental property of L^p spaces cease to be true; an example is the triangle inequality of the norm, as we have seen also for normed spaces.

$|f|^p \sim |g|^p$. This is the case because we can write $|f|^p$ as $h \circ f$ and $|g|^p$ as $h \circ g$, where the function

$$\begin{aligned} h : \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\mapsto |z|^p \end{aligned}$$

is continuous, and we have already proved (point *(iii)* in Theorem 4.1) that $h \circ f \sim h \circ g$, i.e., $|f|^p \sim |g|^p$, as required. The same argument shows that if $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$ and $f \sim g$, the value of the integrals $\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x}$ and $\int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x}$ is the same.

Remark 4.2 With some abuse of notation and language, we can say that the set $L^1(\boldsymbol{A}; \mathbb{C})$ is the space of Lebesgue summable functions over $\boldsymbol{A}$. The abuse consists in the fact that $L^1(\boldsymbol{A}; \mathbb{C})$ is not a set of functions, but a set of equivalent classes, and, moreover, if $[f] \in L^1(\boldsymbol{A}; \mathbb{C})$, a function g may not be defined over A, but over $\boldsymbol{A} \setminus \boldsymbol{N}$ where $\boldsymbol{N}$ is a null set. However, from *(vi)* in Theorem 4.1, we know that if $[f] \in L^1(\boldsymbol{A}; \mathbb{C})$, all the functions in the equivalence class $[f]$ are summable in the sense of (4.2), i.e., any arbitrary extension is summable over $\boldsymbol{A}$.

Remark 4.3 In what follows, when we write $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$, we shall mean that f is defined almost everywhere in A and its equivalence class $[f]$ is an element of $L^p(\boldsymbol{A}; \mathbb{C})$ (instead of writing $\Phi \in L^p(\boldsymbol{A}; \mathbb{C})$ such that there exists f defined almost everywhere in $\boldsymbol{A}$ satisfying $[f] = \Phi$). This is legitimate because the function that maps an element to its equivalence class is surjective.

We now consider the space L^p when $p = +\infty$. To deal with this case, we need to introduce the class of essentially bounded functions.

Definition 4.4 [Essentially Bounded Functions]
Let $\boldsymbol{A} \in \mathcal{M}_n$. We say that the function f defined almost everywhere in $\boldsymbol{A}$ and taking values in $\mathbb{C}$ is essentially bounded *if it is bounded almost everywhere in $\boldsymbol{A}$, i.e., if there exists $M \geq 0$ such that*

$$|f(\boldsymbol{x})| \leq M \quad \textit{for almost all } \boldsymbol{x} \in \boldsymbol{A}.$$

We can also define

$$\operatorname{ess\,sup} f = \inf\{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ for almost all } \boldsymbol{x} \in \boldsymbol{A}\}. \tag{4.5}$$

If the set in the right hand-side of the latter is empty, we have $\operatorname{ess\,sup} f = +\infty$. If f takes real values, we can also define

$$\operatorname{ess\,inf} f = \sup\{M \geq 0 \,|\, f(\boldsymbol{x}) \geq M \text{ for almost all } \boldsymbol{x} \in \boldsymbol{A}\}, \tag{4.6}$$

and $\operatorname{ess\,inf} f = -\infty$ when the set in the right hand-side of (4.6) is empty.

Example 4.3 The function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} k & \text{if } x = k \text{ for } k \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is not bounded, but it is essentially bounded: the set of points at which f is not zero has zero measure, being the countable union of pairwise disjoint sets with zero Lebesgue measure.

Definition 4.5 [L^∞ SPACES]
Let $\boldsymbol{A} \in \mathcal{M}_n$. We define

$$L^\infty(\boldsymbol{A};\mathbb{C}) = \left\{ [f] \,\middle|\, f \text{ is defined a.e. in } \boldsymbol{A}, \text{ measurable and essentially bounded} \right\}.$$

We now show that the set $L^\infty(\boldsymbol{A};\mathbb{C})$ is well defined, i.e., we make sure that the condition of being "essentially bounded" in the set defining $L^\infty(\boldsymbol{A};\mathbb{C})$ does not give different answers for different functions belonging to the same equivalence class (which would render the definition of $L^\infty(\boldsymbol{A};\mathbb{C})$ ill-posed: the same equivalence class $[f]$ could simultaneously belong to $L^\infty(\boldsymbol{A};\mathbb{C})$ or not, depending on which element of $[f]$ we choose). Let $[f] \in L^\infty(\boldsymbol{A};\mathbb{C})$ and take $g \sim f$. We show that g is also essentially bounded. Indeed, if $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are null sets such that $f : \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C}$ and $g : \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}$, since $g \sim f$ the set $\boldsymbol{N}_3 = \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}$ has measure zero. Hence, for all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3)$, we have $|g(\boldsymbol{x})| = |f(\boldsymbol{x})| \leq M$, and since $\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3$ has measure zero, we can conclude that $|g(\boldsymbol{x})| \leq M$ almost everywhere in $\boldsymbol{A}$.

Example *4.4*

Let $a > 0$. We recall that $\frac{1}{x^a}$ is:

- summable in $(0,1)$ $\quad\Leftrightarrow\quad$ $a < 1$;
- summable in $(1,+\infty)$ $\quad\Leftrightarrow\quad$ $a > 1$.

Let $\alpha, \beta \in \mathbb{R}$ be such that $\frac{1}{2} \leq \alpha < 1$ and $\beta > 1$. Consider the function

$$\begin{aligned} u : (0,+\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \frac{1}{x^\alpha} & \text{if } 0 < x \leq 1 \\ \frac{1}{x^\beta} & \text{if } x > 1 \end{cases} \end{aligned}$$

It is easy to see, using the previous considerations, that u is summable, so that $u \in L^1((0,+\infty);\mathbb{R})$. However, $\int_0^{+\infty} |u|^2(x)\,\mathrm{d}x = +\infty$, because $\int_0^1 |u|^2(x)\,\mathrm{d}x = +\infty$, from the fact that $1 \leq 2\alpha < 2$. It follows that $u \notin L^2((0,+\infty);\mathbb{R})$. Finally, since u is not bounded, we have $u \notin L^\infty((0,+\infty);\mathbb{R})$. We can conclude that that

$$u \notin L^2((0,+\infty);\mathbb{R}) \cup L^\infty((0,+\infty);\mathbb{R}).$$

Example *4.5*

Let $\alpha, \beta \in \mathbb{R}$ be parameters satisfying the conditions $0 < \alpha < \frac{1}{2}$ and $\beta > \frac{1}{2}$. Consider the function

$$\begin{aligned} f : (0,+\infty) &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} \frac{1}{t^\alpha} & \text{if } 0 < t \leq 1 \\ \frac{1}{t^\beta} & \text{if } t > 1 \end{cases} \end{aligned}$$

We have $\int_0^{+\infty} |f|^2(t)\,\mathrm{d}t < +\infty$, so that $f \in L^2((0,+\infty);\mathbb{R})$. However, f is not summable (notice that β is not greater than 1), so that $f \notin L^1((0,+\infty);\mathbb{R})$. Finally, since f is not bounded, we have $f \notin L^\infty((0,+\infty);\mathbb{R})$. It follows that

$$f \notin L^1((0,+\infty);\mathbb{R}) \cup L^\infty((0,+\infty);\mathbb{R}).$$

Example 4.6

Consider a constant function $u : (0, +\infty) \longrightarrow \mathbb{R}$. Since $|u|$ is bounded, then $u \in L^\infty((0, +\infty); \mathbb{R})$. However, neither $|u|$ nor $|u|^2$ is integrable, so that

$$u \notin L^1((0, +\infty); \mathbb{R}) \cup L^2((0, +\infty); \mathbb{R}).$$

These examples show that, in general, there is no inclusion relation between L^p spaces for different values of p. In other words, if $1 \leq p < q \leq +\infty$, taking $\alpha \in \left(\frac{1}{q}, \frac{1}{p}\right)$ allows us to define a function

$$\begin{aligned} f : (0, +\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x^\alpha} \chi_{(0,1)} \end{aligned}$$

which is in $L^p((0, +\infty); \mathbb{R})$ but not in $L^q((0, +\infty); \mathbb{R})$, as well as a function

$$\begin{aligned} g : (0, +\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x^\alpha} \chi_{(1,+\infty)} \end{aligned}$$

which is in $L^q((0, +\infty); \mathbb{R})$ but not in $L^p((0, +\infty); \mathbb{R})$.

We now want to establish a result that shows that, for every $p \in [1, +\infty]$, the space $L^p(A; \mathbb{C})$ is a vector space. To this end, we need to define the notion of addition and scalar multiplication for the elements of $L^p(A; \mathbb{C})$, which – we recall – are equivalence classes (i.e., sets of functions).

We have seen in *(i)* of Theorem 4.1 that, if f_1, f_2, g_1, g_2 are defined almost everywhere over $\boldsymbol{A} \in \mathcal{M}_n$ and are such that $f_1 \sim g_1$ and $f_2 \sim g_2$, then $f_1 + f_2 \sim g_1 + g_2$. Therefore, in (4.3) we defined the sum of the corresponding equivalence classes as $[f] + [g] \stackrel{\text{def}}{=} [f + g]$. We say that this definition is independent from the particular representatives f and g used to represent the equivalence classes. In a similar way, using *(ii)* of Theorem 4.1, we can define the product of an equivalence class by a scalar. Let $\lambda \in \mathbb{C}$, then we defined in (4.4) $\lambda\,[f] \stackrel{\text{def}}{=} [\lambda f]$. With these operations, $L^p(\boldsymbol{A}; \mathbb{C})$ is a vector space. Since functions from any set taking values in a vector space form a vector space, we must only verify that:

(i) if $[f], [g] \in L^p(\boldsymbol{A}; \mathbb{C})$, then $[f] + [g] = [f + g] \in L^p(\boldsymbol{A}; \mathbb{C})$;

(ii) if $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$ and $\lambda \in \mathbb{C}$, then $\lambda\,[f] = [\lambda f] \in L^p(\boldsymbol{A}; \mathbb{C})$.

We first consider the case $1 \leq p < +\infty$. Since $[f], [g] \in L^p(\boldsymbol{A}; \mathbb{C})$, we have

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty \quad \text{and} \quad \int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty. \tag{4.7}$$

If f is defined over $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and g over $\boldsymbol{A} \setminus \boldsymbol{N}_2$, where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are null sets, for every $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2)$ we have

$$|f(\boldsymbol{x}) + g(\boldsymbol{x})| \leq |f(\boldsymbol{x})| + |g(\boldsymbol{x})| \leq \max\{2\,|f(\boldsymbol{x})|, 2\,|g(\boldsymbol{x})|\} = 2\,\max\{|f(\boldsymbol{x})|, |g(\boldsymbol{x})|\},$$

we obtain

$$|f(\boldsymbol{x}) + g(\boldsymbol{x})|^p \leq 2^p\,\max\{|f(\boldsymbol{x})|^p, |g(\boldsymbol{x})|^p\} \leq 2^p\,(|f(\boldsymbol{x})|^p + |g(\boldsymbol{x})|^p).$$

Integrating over $\boldsymbol{A}$ (arbitrarily extending f over $\boldsymbol{N}_1$ and g over $\boldsymbol{N}_2$) yields

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x}) + g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} 2^p\,(|f(\boldsymbol{x})|^p + |g(\boldsymbol{x})|^p)\,\mathrm{d}\boldsymbol{x} \leq 2^p \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \right) < +\infty$$

from (4.7), so that indeed $[f + g] \in L^p(\boldsymbol{A}; \mathbb{C})$.

Now we prove that, for any $\lambda \in \mathbb{C}$, there holds $\lambda\,[f] = [\lambda\, f] \in L^p(\boldsymbol{A};\mathbb{C})$. We have

$$\int_{\boldsymbol{A}} |\lambda\, f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x} = |\lambda|^p \int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x} < +\infty,$$

since $[f] \in L^p(\boldsymbol{A};\mathbb{C})$.

Now, consider $p = +\infty$. We have to show that, if $[f], [g] \in L^\infty(\boldsymbol{A};\mathbb{C})$, then $[f] + [g] \in L^\infty(\boldsymbol{A};\mathbb{C})$. Let $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ be null sets such that $f : \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C}$ and $g : \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}$. Since f is essentially bounded, there exists a null set $\boldsymbol{N}_3$ and there exists $M_1 \geq 0$ such that for all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_3)$, there holds $|f(\boldsymbol{x})| \leq M_1$. Likewise, since g is essentially bounded, there exists a null set $\boldsymbol{N}_4$ and there exists $M_2 \geq 0$ such that for all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_2 \cup \boldsymbol{N}_4)$, there holds $|g(\boldsymbol{x})| \leq M_2$. Hence, for all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4)$ we have $|f(\boldsymbol{x}) + g(\boldsymbol{x})| \leq |f(\boldsymbol{x})| + |g(\boldsymbol{x})| \leq M_1 + M_2$, and since $\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4$ has measure zero we conclude that $f + g$ is essentially bounded, i.e., $[f + g] \in L^\infty(\boldsymbol{A};\mathbb{C})$.

Moreover, if $\lambda \in \mathbb{C}$ and $[f] \in L^\infty(\boldsymbol{A};\mathbb{C})$, we can write $f : \boldsymbol{A} \setminus \boldsymbol{N}_1 \longrightarrow \mathbb{C}$ where $\boldsymbol{N}_1$ is a null set, and, since f is essentially bounded, there exists a null set $\boldsymbol{N}_2$ and there exists $M \geq 0$ such that for all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2)$ we have $|f(\boldsymbol{x})| \leq M$. For all $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2)$, we have $|\lambda\, f(\boldsymbol{x})| \leq |\lambda|\, M$, so that $|\lambda\, f(\boldsymbol{x})| \leq |\lambda|\, M$ almost everywhere in $\boldsymbol{A}$, from which we conclude that $[\lambda\, f] \in L^\infty(\boldsymbol{A};\mathbb{C})$.

We are now ready to define the norms in $L^p(\boldsymbol{A};\mathbb{C})$.

Definition 4.6 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $p \in [1, +\infty)$ and $[f] \in L^p(\boldsymbol{A};\mathbb{C})$. We define*

$$\|\,[f]\,\|_{L^p(\boldsymbol{A};\mathbb{C})} \stackrel{def}{=} \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x} \right)^{\frac{1}{p}}.$$

Let $p = +\infty$, and $[f] \in L^\infty(\boldsymbol{A};\mathbb{C})$. We define

$$\|\,[f]\,\|_{L^\infty(\boldsymbol{A};\mathbb{C})} \stackrel{def}{=} \operatorname{ess\,sup} f = \inf\{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ a.e. on } \boldsymbol{A}\}.$$

We will often use the symbols $\|\,[f]\,\|_{\boldsymbol{A},p}$ and $\|\,[f]\,\|_{A,\infty}$, or $\|\,[f]\,\|_p$ and $\|\,[f]\,\|_\infty$, instead of $\|\,[f]\,\|_{L^p(\boldsymbol{A};\mathbb{C})}$ and $\|\,[f]\,\|_{L^\infty(\boldsymbol{A};\mathbb{C})}$, respectively, when there is no possibility of misunderstanding.

We need to make a few considerations for the case $p = +\infty$.

Theorem 4.4 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $[f] \in L^\infty(\boldsymbol{A};\mathbb{C})$. The set*

$$\{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ a.e. in } \boldsymbol{A}\}$$

admits a minimum.

Proof Let $S = \{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ a.e. on } \boldsymbol{A}\}$, and let $m = \inf S$. We have to prove that $m \in S$ (because if the infimum of a set belongs to the set, it is the minimum), i.e., that $|f(\boldsymbol{x})| \leq m$ for almost all $\boldsymbol{x} \in \boldsymbol{A}$. From the properties of the infimum, for every $k \in \mathbb{N}^*$, there exists $M_k \in S$ such that $M_k \leq m + \frac{1}{k}$. By definition of S, for

all $k \in \mathbb{N}^*$ we have $M_k \geq 0$ such that $|f(\boldsymbol{x})| \leq M_k$ a.e. on $\boldsymbol{A}$, and

$$m \leq M_k \leq m + \frac{1}{k}.$$

Let $\boldsymbol{N}_0$ be a null set that allows us to write the domain of f as $\boldsymbol{A} \setminus \boldsymbol{N}_0$. Let us also define, for all $k \in \mathbb{N}^*$, the set $N_k = \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_0 \,|\, |f(\boldsymbol{x})| > M_k\}$. Since for all $k \in \mathbb{N}^*$, $|f(\boldsymbol{x})| \leq M_k$ a.e. on $\boldsymbol{A}$, it follows that $\mu_n(\boldsymbol{N}_k) = 0$ for all $k \in \mathbb{N}^*$. Let

$$\boldsymbol{x} \in \boldsymbol{A} \setminus \bigcup_{k \in \mathbb{N}} \boldsymbol{N}_k.$$

Then, $|f(\boldsymbol{x})| \leq M_k \leq m + \frac{1}{k}$ for all $k \in \mathbb{N}^*$, so that $|f(\boldsymbol{x})| \leq m$. We only need to show that

$$\mu_n\Big(\bigcup_{k \in \mathbb{N}} \boldsymbol{N}_k\Big) = 0.$$

This follows recalling that

$$\mu_n\Big(\bigcup_{k \in \mathbb{N}} \boldsymbol{N}_k\Big) \leq \sum_{k \in \mathbb{N}} \mu_n(\boldsymbol{N}_k) = 0.$$

Hence, m is the minimum of S.

As a result of Theorem 4.4, if $[f] \in L^\infty(\boldsymbol{A}; \mathbb{C})$, we can write

$$\|\,[f]\,\|_{\infty,\boldsymbol{A}} = \min\{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ a.e. on } \boldsymbol{A}\}.$$

We now need to show that the norms introduced for the L^p spaces are indeed norms. In other words, we need to prove that for all $p \in [1, +\infty]$, the norm $\|\cdot\|_p$ satisfies the following properties:

(i) $\|\,[f]\,\|_p = 0$ if and only if $[f] = [0]$. Let $p \in [1, +\infty)$. We know that

$$\|\,[f]\,\|_p = \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} = 0 \quad \Leftrightarrow \quad \int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x} = 0,$$

from which $|f(\boldsymbol{x})|^p = 0$ for almost every $\boldsymbol{x} \in \boldsymbol{A}$, so that $f(\boldsymbol{x}) = 0$ a.e. on $\boldsymbol{A}$, which means $[f] = [0]$. Let $p = +\infty$. In this case $\|\,[f]\,\|_\infty = \min\{M \geq 0 \,|\, |f(\boldsymbol{x})| \leq M \text{ a.e. on } \boldsymbol{A}\} = 0$ if and only if $|f(\boldsymbol{x})| \leq 0$ a.e. on $\boldsymbol{A}$, i.e., $|f(\boldsymbol{x})| = 0$ a.e. on $\boldsymbol{A}$, so that $[f] = [0]$;

(ii) for all $[f] \in L^p(A; \mathbb{C})$ and for all $\lambda \in \mathbb{C}$, we have $\|\lambda\,[f]\,\|_p = |\lambda|\,\|\,[f]\,\|_p$. Again, consider first $p \in [1, +\infty)$. We have

$$\|\lambda\,[f]\,\|_p = \|\,[\lambda\, f]\,\|_p = \left(\int_{\boldsymbol{A}} |\lambda\, f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} = |\lambda| \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} = |\lambda|\,\|\,[f]\,\|_p.$$

Now, consider $p = +\infty$. We have $\|\lambda\,[f]\,\|_\infty = \|\,[\lambda\, f]\,\|_\infty$ and $|f(\boldsymbol{x})| \leq \|\,[f]\,\|_\infty$ a.e. on $\boldsymbol{A}$, from which $|\lambda\, f(\boldsymbol{x})| \leq |\lambda|\,\|\,[f]\,\|_\infty$ a.e. on $\boldsymbol{A}$. It follows that $|\lambda|\,\|\,[f]\,\|_\infty \in \{M \geq 0 \,|\, |\lambda\, f(\boldsymbol{x})| \leq M \text{ a.e. in } \boldsymbol{A}\}$. Since $\|\,[\lambda\, f]\,\|_\infty$ is the minimum of $\{M \geq 0 \,|\, |\lambda\, f(\boldsymbol{x})| \leq M \text{ a.e. in } \boldsymbol{A}\}$, we obtain $\|\,[\lambda\, f]\,\|_\infty \leq |\lambda|\,\|\,[f]\,\|_\infty$. We need to prove the opposite inequality.
If $\lambda = 0$, then clearly $\|\,[0]\,\|_\infty = |0|\,\|\,[f]\,\|_\infty$. Let us now assume $\lambda \neq 0$. We can write $f = \lambda^{-1}\,(\lambda\, f)$. From what we have seen before $\|\,[f]\,\|_\infty \leq |\lambda|^{-1}\,\|\,[\lambda\, f]\,\|_\infty$, so that $|\lambda|\,\|\,[f]\,\|_\infty \leq \|\,[\lambda\, f]\,\|_\infty$. This proves the statement;

(iii) For all $[f], [g] \in L^p(\boldsymbol{A}; \mathbb{C})$, we must have $\|\,[f]+[g]\,\|_p \leq \|\,[f]\,\|_p + \|\,[g]\,\|_p$. We show this inequality for $p = 1$ and $p = +\infty$. If $p = 1$, we have

$$\|\,[f]+[g]\,\|_1 = \|\,[f+g]\,\|_1 = \int_{\boldsymbol{A}} |f(\boldsymbol{x}) + g(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{A}} |f(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{A}} |g(\boldsymbol{x})| \,\mathrm{d}\boldsymbol{x} = \|\,[f]\,\|_1 + \|\,[g]\,\|_1.$$

If $p = +\infty$, we have $\|[f] + [g]\|_\infty = \|[f + g]\|_\infty$, and we also have the inequalities $|f(\boldsymbol{x}) + g(\boldsymbol{x})| \leq |f(\boldsymbol{x})| + |g(\boldsymbol{x})|$, and $|f(\boldsymbol{x})| \leq \|[f]\|_\infty$ a.e. on $\boldsymbol{A}$, and, finally, $|g(\boldsymbol{x})| \leq \|[g]\|_\infty$ a.e. on $\boldsymbol{A}$. Let the domains of f and g be, respectively, $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and $\boldsymbol{A} \setminus \boldsymbol{N}_2$, where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are null sets. Moreover, let $\boldsymbol{N}_3 = \{\boldsymbol{x} \in \boldsymbol{A} \setminus N_1 \,|\, |f(\boldsymbol{x})| > \|[f]\|_\infty\}$ and $\boldsymbol{N}_4 = \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_2 \,|\, |g(\boldsymbol{x})| > \|[g]\|_\infty\}$. We know that $\mu_n(\boldsymbol{N}_3) = \mu_n(\boldsymbol{N}_4) = 0$, and for $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3 \cup \boldsymbol{N}_4)$ we have $|f(\boldsymbol{x}) + g(\boldsymbol{x})| \leq \|[f]\|_\infty + \|[g]\|_\infty$. This is true almost everywhere on $\boldsymbol{A}$ since $\mu_n(N_1 \cup N_2 \cup N_3 \cup N_4) = 0$. Thus, $\|[f]\|_\infty + \|[g]\|_\infty$ is an element of the set $\{M \geq 0 \,|\, |f(\boldsymbol{x}) + g(\boldsymbol{x})| \leq M \text{ a.e.}\}$. Since $\|[f + g]\|_\infty$ is the minimum of the same set, we find $\|[f + g]\|_\infty \leq \|[f]\|_\infty + \|[g]\|_\infty$, which is the same as $\|[f] + [g]\|_\infty \leq \|[f]\|_\infty + \|[g]\|_\infty$.

The verification of the triangular inequality

$$\|[f] + [g]\|_p \leq \|[f]\|_p + \|[g]\|_p$$

in the general case is not trivial: we use the same ideas used to prove that the p-norm in $\mathbb{R}^n$ is a norm: first, we use Young's inequality

$$a\, b \leq \frac{a^p}{p} + \frac{b^q}{q},$$

where $a, b \geq 0$ and $p, q > 1$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$ (see Theorem 1.3), to prove Hölder's inequality for the L^p-norm. Finally, we prove the triangle inequality (also called *Minkowski inequality*).

Theorem 4.5 [Hölder's Inequality]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $p, q \in [1, +\infty]$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ (where we set $\frac{1}{+\infty} = 0$). Let $[f] \in L^p(\boldsymbol{A}, \mathbb{C})$ and $[g] \in L^q(\boldsymbol{A}, \mathbb{C})$. Then, $f\, g \in L^1(\boldsymbol{A}, \mathbb{C})$. Moreover,

$$\|[f\, g]\|_1 \leq \|[f]\|_p \, \|[g]\|_q. \tag{4.8}$$

Proof If $\|[f]\|_p = 0$, then f is zero a.e., and the product $f\, g$ is zero a.e., so that the left-hand side of (4.8) is zero, which means that (4.8) is trivially satisfied. The same is true for $\|[g]\|_q = 0$. Thus, we can assume that both $\|[f]\|_p$ and $\|[g]\|_q$ are strictly positive. It follows that we can assume that $\|[f]\|_p$ and $\|[g]\|_q$ are in $(0, +\infty)$: this will be necessary to apply Young's inequality. Moreover, Young's inequality requires p and q to be > 1, whereas the statement of Hölder's inequality also holds when $p = 1$ or $q = 1$. Clearly, if $p = 1$, then $q = +\infty$, and, vice-versa, when $q = 1$ we have $p = +\infty$. Therefore these cases need to be addressed independently from Young's inequality. If $p = 1$ and $q = +\infty$, (4.8) becomes $\|[f]\,[g]\|_1 \leq \|[f]\|_1 \|[g]\|_\infty$ a.e. (why?), i.e.,

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})\, g(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \leq \int_A |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x} \cdot \operatorname{ess\,sup} g,$$

and the result follows from fact that the Lebesgue integral of a non-negative function is monotonic, taking into account that for almost all $\boldsymbol{x} \in A$ we have $|f(\boldsymbol{x})\, g(\boldsymbol{x})| \leq |f(\boldsymbol{x})| \operatorname{ess\,sup} g$. The same can be said for the case $p = +\infty$ and $q = 1$. Hence, we can focus on the case $p, q \in (1, +\infty)$. We can use Young's inequality with

$$a = \frac{|f(\boldsymbol{x})|}{\|[f]\|_p} \quad \text{and} \quad b = \frac{|g(\boldsymbol{x})|}{\|[g]\|_q},$$

where $\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2)$, assuming that the domain of f is $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and the domain of g is $\boldsymbol{A} \setminus \boldsymbol{N}_2$, where $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ are null sets. We obtain

$$a\, b = \frac{|f(\boldsymbol{x})| \cdot |g(\boldsymbol{x})|}{\|[f]\|_p \, \|[g]\|_q} \leq \frac{|f(\boldsymbol{x})|^p}{p\, \|[f]\|_p} + \frac{|g(\boldsymbol{x})|^q}{q\, \|[g]\|_q}.$$

Integrating both sides yields

$$\frac{1}{\|[f]\|_p \|[g]\|_q} \int_A |f(\boldsymbol{x})\, g(\boldsymbol{x})|\, \mathrm{d}x \leq \frac{1}{p\,\|[f]\|_p^p} \int_A |f(\boldsymbol{x})|^p\, \mathrm{d}x + \frac{1}{q\,\|[g]\|_q^q} \int_A |g(\boldsymbol{x})|^q\, \mathrm{d}x,$$

i.e.,

$$\frac{1}{\|[f]\|_p \|[g]\|_q} \|[f\, g]\|_1 \leq \frac{1}{p\,\|[f]\|_p^p} \|[f]\|_p^p + \frac{1}{q\,\|[g]\|_q^q} \|[g]\|_q^q = \frac{1}{p} + \frac{1}{q} = 1.$$

We are now ready to prove the Minkowski's inequality, which is the way the triangle inequality

$$\|[f] + [g]\|_p \leq \|[f]\|_p + \|[g]\|_p$$

is called for L^p spaces.

Theorem 4.6 [Minkowski's Inequality]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $p \in [1, +\infty]$. Let $[f], [g] \in L^p(\boldsymbol{A}, \mathbb{C})$. Then,

$$\|[f] + [g]\|_p \leq \|[f]\|_p + \|[g]\|_p.$$

Proof First, we have already observed that $L^p(\boldsymbol{A}, \mathbb{C})$ is closed under the addition defined in (4.3). Hence, if $[f], [g] \in L^p(\boldsymbol{A}, \mathbb{C})$, the norms $\|[f]\|_p$ and $\|[g]\|_p$ are finite, and since $[f] + [g] \in L^p(\boldsymbol{A}, \mathbb{C})$, the norm $\|[f] + [g]\|_p$ is also finite.

Now, if $\|[f + g]\|_p = 0$, the inequality is trivially true. We can consider the case $\|[f + g]\|_p > 0$. Then,

$$\begin{aligned}
\|[f + g]\|_p^p &= \int_A |f(\boldsymbol{x}) + g(\boldsymbol{x})|^p\, \mathrm{d}\boldsymbol{x} \\
&= \int_A |f(\boldsymbol{x}) + g(\boldsymbol{x})| \cdot |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x} \\
&\leq \int_A (|f(\boldsymbol{x})| + |g(\boldsymbol{x})|)\, |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x} \\
&= \int_A |f(\boldsymbol{x})| |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x} + \int_A |g(\boldsymbol{x})| |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x}. \qquad (4.9)
\end{aligned}$$

Consider the first integral. We use Hölder's inequality by considering this integral as the 1-norm of the product of the functions $\varphi = f$ and $\psi = (f + g)^{p-1}$, and choosing q so that $\frac{1}{p} + \frac{1}{q} = 1$ holds, i.e., $q = \frac{p}{p-1}$ (which is possible because $p > 1$):

$$\int_A |f(\boldsymbol{x})| |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x} = \|[\varphi\, \psi]\|_1 \leq \|[\varphi]\|_p + \|[\psi]\|_q = \left(\int_A |f(\boldsymbol{x})|^p\, \mathrm{d}x \right)^{\frac{1}{p}} \left(\int_A |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{(p-1)\left(\frac{p}{p-1}\right)}\, \mathrm{d}\boldsymbol{x} \right)^{1-\frac{1}{p}}$$

We can do the same for the second integral in (4.9):

$$\int_A |g(\boldsymbol{x})| |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{p-1}\, \mathrm{d}\boldsymbol{x} \leq \left(\int_A |g(\boldsymbol{x})|^p\, \mathrm{d}x \right)^{\frac{1}{p}} \left(\int_A |f(\boldsymbol{x}) + g(\boldsymbol{x})|^{(p-1)\left(\frac{p}{p-1}\right)}\, \mathrm{d}\boldsymbol{x} \right)^{1-\frac{1}{p}},$$

so that continuing from (4.9) we find

$$\begin{aligned}\|[f+g]\|_p^p &\leq \left[\left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}x\right)^{\frac{1}{p}} + \left(\int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \,\mathrm{d}x\right)^{\frac{1}{p}}\right]\left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})+g(\boldsymbol{x})|^{(p-1)\left(\frac{p}{p-1}\right)} \,\mathrm{d}\boldsymbol{x}\right)^{1-\frac{1}{p}} \\ &= \left[\left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \,\mathrm{d}x\right)^{\frac{1}{p}} + \left(\int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \,\mathrm{d}x\right)^{\frac{1}{p}}\right]\left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})+g(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}\right)^{1-\frac{1}{p}} \\ &= (\|[f]\|_p + \|[g]\|_p)\,\frac{\|[f+g]\|_p^p}{\|[f+g]\|_p}\end{aligned}$$

Multiplying both sides by $\frac{\|[f+g]\|_p}{\|[f+g]\|_p^p}$ proves the claim.

We can conclude this section with the result that uses all the considerations established so far, including the Minkowski's inequality, to state that $L^p(\boldsymbol{A},\mathbb{C})$ is a complex vector space.

Corollary 4.1 *Let $\boldsymbol{A} \in \mathcal{M}_n$. Let $p \in [1,+\infty]$. Then, $L^p(\boldsymbol{A},\mathbb{C})$ is a vector space over the field $\mathbb{C}$, by considering the addition and the scalar multiplication as defined in (4.3) and (4.4), respectively.*

4.3 Completeness of L^p spaces

Consider the case $p = 2$, let $\boldsymbol{A} \in \mathcal{M}_n$ and let $f, g : \boldsymbol{A} \longrightarrow \mathbb{C}$ be two measurable functions, such that $[f], [g] \in L^2(\boldsymbol{A},\mathbb{C})$. Then,

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^2 \,\mathrm{d}\,\boldsymbol{x} < +\infty \quad \text{and} \quad \int_{\boldsymbol{A}} |g(\boldsymbol{x})|^2 \,\mathrm{d}\,\boldsymbol{x} < +\infty.$$

It is easy to see that $\bar{g}\, f$ is summable. Indeed, using Young's inequality with $p = q = 2$, we find that for all $a, b \in \mathbb{R}$

$$a\, b \leq \frac{1}{2}\,(a^2 + b^2).$$

It follows that $|f\,\bar{g}| = |f|\,|g| \leq \frac{1}{2}\,(|f|^2 + |g|^2)$, so that $\bar{g}\, f$ is measurable and its modulus can be bounded by a summable function. This ensures that $\bar{g}\, f$ is summable. We can define the inner product in $L^2(\boldsymbol{A};\mathbb{C})$ between $[f]$ and $[g]$ as

$$\langle [f], [g] \rangle = \int_{\boldsymbol{A}} \overline{g(\boldsymbol{x})}\, f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x},$$

from which we obtain

$$\|[f]\|_2 = (\langle [f], [f] \rangle)^{\frac{1}{2}}.$$

We will find that $L^2(\boldsymbol{A};\mathbb{C})$ is a Hilbert space, because, as we will see shortly, $L^p(\boldsymbol{A};\mathbb{C})$ is a Banach space for all $p \in [1,+\infty]$, and the norm in $L^2(\boldsymbol{A};\mathbb{C})$ is induced by an inner product.

Before we prove the following result, we recall that a normed vector space $(\boldsymbol{X}, \|\cdot\|)$ is complete if and only if every absolutely convergent series is also convergent (Theorem 12 in the set of notes on Banach and Hilbert spaces). This is the property that will be used to prove the completeness of L^p spaces.

Theorem 4.7 [Completeness of L^p Spaces — Riesz-Fischer Theorem]
Let $\boldsymbol{A} \in \mathcal{M}_n$. Then, $(L^p(A;\mathbb{C}), \|\cdot\|_p)$ is a Banach space for all $p \in [1,+\infty]$.

Proof Let first $p \geq 1$. To prove that $(L^p(\boldsymbol{A};\mathbb{C}), \|\cdot\|_p)$ is complete, we show that any absolutely convergent series of elements of $L^p(\boldsymbol{A};\mathbb{C})$ is also convergent with respect to the norm $\|\cdot\|_p$. Let $([f_k])_{k\in\mathbb{N}}$ be a sequence of $L^p(\boldsymbol{A},\mathbb{C})$ such that the associated series is absolutely convergent, i.e., $\displaystyle\sum_{k=0}^{\infty} \|[f_k]\|_p < +\infty$. Let

$$M = \sum_{k=0}^{\infty} \|[f_k]\|_p. \tag{4.10}$$

We have to show that $\displaystyle\sum_{k=0}^{\infty}[f_k]$ is convergent. This means that its sequence of the partial sums is convergent in $L^p(\boldsymbol{A},\mathbb{C})$. To construct the sequence of the partial sums we proceed as follows. For each $k \in \mathbb{N}$, let the domain of f_k be $\boldsymbol{A} \setminus \boldsymbol{N}_k$, where $\boldsymbol{N}_k$ is a null set. Since the countable union of null sets is a null set, we can say that $\boldsymbol{N} = \displaystyle\bigcup_{k=0}^{\infty} \boldsymbol{N}_k$ has measure zero. We define, for every $k \in \mathbb{N}$, the function

$$\begin{aligned} s_k : \boldsymbol{A} \setminus \boldsymbol{N} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \sum_{\ell=0}^{k} f_\ell(\boldsymbol{x}). \end{aligned}$$

Hence, for every $k \in \mathbb{N}$ we have

$$[s_k] = \left[\sum_{\ell=0}^{k} f_\ell\right] = \sum_{\ell=0}^{k}[f_\ell],$$

so that $([s_k])_{k\in\mathbb{N}}$ is the sequence of the partial sums of the series $\displaystyle\sum_{k=0}^{\infty}[f_k]$. To prove that $\displaystyle\sum_{k=0}^{\infty}[f_k]$ is convergent in $L^p(\boldsymbol{A},\mathbb{C})$, we show that the sequence of the partial sums $([s_k])_{k\in\mathbb{N}}$ admits limit in $L^p(\boldsymbol{A},\mathbb{C})$, i.e., that there exists a function s defined almost everywhere in A such that $[s] \in L^p(\boldsymbol{A},\mathbb{C})$ and

$$\lim_{k\to\infty} \|[s_k] - [s]\|_p = 0.$$

To this end, we define for every $k \in \mathbb{N}$ the function

$$\begin{aligned} g_k : \boldsymbol{A} \setminus \boldsymbol{N} &\longrightarrow [0, +\infty] \\ \boldsymbol{x} &\mapsto \sum_{\ell=0}^{k} |f_\ell(\boldsymbol{x})|. \end{aligned}$$

Using Minkowski's inequality, for every $k \in \mathbb{N}$ we find

$$\|[g_k]\|_p = \left\|\left[\sum_{\ell=0}^{k}|f_\ell|\right]\right\|_p = \left\|\sum_{\ell=0}^{k}[\,|f_\ell|\,]\right\|_p \leq \sum_{\ell=0}^{k}\|[\,|f_\ell|\,]\|_p = \sum_{\ell=0}^{k}\|[f_\ell]\|_p \leq \sum_{\ell=0}^{\infty}\|[f_\ell]\|_p = M,$$

so that, for all $k \in \mathbb{N}$,

$$\left(\int_{\boldsymbol{A}} |g_k(\boldsymbol{x})|^p \,\mathrm{d}x\right)^{\frac{1}{p}} \leq M \quad \Rightarrow \quad \int_{\boldsymbol{A}} g_k^p(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq M^p.$$

We notice that the sequence $(g_k)_{k\in\mathbb{N}}$ is monotonically non-decreasing. Therefore, by the monotone convergence theorem, the function g defined pointwise by

$$\forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}, \quad g(\boldsymbol{x}) = \lim_{k\to\infty} g_k(\boldsymbol{x})$$

is measurable, it takes non-negative values, and

$$\int_{\boldsymbol{A}} |g(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} g^p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \lim_{k\to\infty} \int_{\boldsymbol{A}} g_k^p(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \leq \lim_{k\to\infty} M^p = M^p.$$

This shows that $[g] \in L^p(\boldsymbol{A}, \mathbb{C})$ and $g(\boldsymbol{x}) < +\infty$ almost everywhere in $\boldsymbol{A}$. For the values of $\boldsymbol{x}$ for which $g(\boldsymbol{x})$ is finite, the series $\sum_{k=0}^{\infty} |f_k(\boldsymbol{x})|$ is convergent, i.e., the series $\sum_{k=0}^{\infty} f_k(\boldsymbol{x})$ is absolutely convergent. This means that the set

$$\boldsymbol{N}_0 = \left\{ \boldsymbol{x} \in \boldsymbol{A} \setminus N \,\middle|\, \sum_{k=0}^{\infty} f_k(\boldsymbol{x}) \text{ is not absolutely convergent} \right\} = \left\{ \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \,\middle|\, \sum_{k=0}^{\infty} |f_k(\boldsymbol{x})| = +\infty \right\}$$

has zero Lebesgue measure. We can therefore define the function

$$\begin{aligned} s : \boldsymbol{A} \setminus \boldsymbol{N} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \begin{cases} \displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in (\boldsymbol{A} \setminus \boldsymbol{N}) \setminus \boldsymbol{N}_0 \\ 0 & \text{if } \boldsymbol{x} \in (\boldsymbol{A} \setminus \boldsymbol{N}) \cap \boldsymbol{N}_0 \end{cases} \end{aligned}$$

Hence, at those points where the series $\sum_{k=0}^{\infty} f_k(\boldsymbol{x})$ is absolutely convergent, we have $s(\boldsymbol{x}) = \sum_{k=0}^{\infty} f_k(\boldsymbol{x})$, and therefore s coincides a.e. with the sum of the series. The sets $(\boldsymbol{A} \setminus \boldsymbol{N}) \setminus \boldsymbol{N}_0$ and $(\boldsymbol{A} \setminus \boldsymbol{N}) \cap \boldsymbol{N}_0$ are measurable. Moreover, the function s is measurable since $s|_{(\boldsymbol{A}\setminus\boldsymbol{N})\setminus\boldsymbol{N}_0}$ is measurable, being the pointwise limit of measurable functions. Furthermore, $s|_{(\boldsymbol{A}\setminus\boldsymbol{N})\cap\boldsymbol{N}_0}$ is measurable since $\boldsymbol{N}$ has measure zero. Thus, s (or, more precisely, an arbitrary extension of s to $\boldsymbol{A}$) is measurable in $\boldsymbol{A}$. We show that $[s] \in L^p(\boldsymbol{A}, \mathbb{C})$. For every $\boldsymbol{x} \in (\boldsymbol{A} \setminus \boldsymbol{N}) \setminus N_0$ and for all $k \in \mathbb{N}$ we have

$$|s_k(\boldsymbol{x})| = \left| \sum_{\ell=0}^{k} f_\ell(\boldsymbol{x}) \right| \leq \sum_{\ell=0}^{k} |f_\ell(\boldsymbol{x})| = g_k(\boldsymbol{x}) \leq g(\boldsymbol{x}).$$

The inequality is preserved taking the limit as $k \to \infty$, and we obtain $|s(\boldsymbol{x})| \leq g(\boldsymbol{x})$ almost everywhere in $\boldsymbol{A}$. From the monotonicity of the Lebesgue integral, since $[g] \in L^p(\boldsymbol{A}, \mathbb{C})$, so is $[s]$. Moreover, for all $\boldsymbol{x} \in (\boldsymbol{A} \setminus N) \setminus \boldsymbol{N}_0$

$$|s_k(\boldsymbol{x}) - s(\boldsymbol{x})| \leq |s_k(\boldsymbol{x})| + |s(\boldsymbol{x})| \leq g(\boldsymbol{x}) + g(\boldsymbol{x}) = 2\, g(\boldsymbol{x}).$$

Since $[g] \in L^p(\boldsymbol{A}, \mathbb{C})$, we can apply the dominated convergence theorem:

$$\lim_{k\to\infty} \int_{\boldsymbol{A}} |s_k(\boldsymbol{x}) - s(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} \lim_{k\to\infty} |s_k(\boldsymbol{x}) - s(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} = 0.$$

Hence,

$$\lim_{k\to\infty} \| [s_k - s] \|_p^p = 0 \quad \Rightarrow \quad \lim_{k\to\infty} \| [s_k - s] \|_p = 0.$$

This concludes the proof for the case $p \in [1, +\infty)$.
We now prove that $L^\infty(\boldsymbol{A}, \mathbb{C})$ is complete. Let $([f_k])_{k\in\mathbb{N}}$ be a sequence of $L^\infty(\boldsymbol{A}, \mathbb{C})$. For each $k \in \mathbb{N}$, the domain of f_k is $\boldsymbol{A} \setminus \boldsymbol{N}_k$, where $\boldsymbol{N}_k$ is a null set. Assume that

$$\sum_{k=0}^{\infty} \| [f_k] \|_\infty < +\infty.$$

We show that $\sum_{k=0}^{\infty} f_k$ is convergent. For all $k \in \mathbb{N}$, we have $|f_k(\boldsymbol{x})| \leq \| [f_k] \|_\infty$ a.e. on $\boldsymbol{A}$. Then, for all $k \in \mathbb{N}$ the set

$$\boldsymbol{N}'_k = \left\{ \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}_k \,\middle|\, |f_k(\boldsymbol{x})| > \| [f_k] \|_\infty \right\}$$

has measure zero, and we can define

$$\boldsymbol{N} = \Big(\bigcup_{k=0}^{\infty} \boldsymbol{N}_k\Big) \cup \Big(\bigcup_{k=0}^{\infty} \boldsymbol{N}'_k\Big).$$

Clearly, $\boldsymbol{N}$ has also measure zero, being the countable union of null sets. Let $\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}$. For all $k \in \mathbb{N}$ we have $|f_k(\boldsymbol{x})| \leq \|[f_k]\|_\infty$. It follows that the series $\displaystyle\sum_{k=0}^{\infty} |f_k(\boldsymbol{x})|$ is convergent, so that the series $\displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x})$ is absolutely convergent in $\boldsymbol{A} \setminus \boldsymbol{N}$. We can now define the function

$$\begin{aligned} s : \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \begin{cases} \displaystyle\sum_{k=0}^{\infty} f_k(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \\ 0 & \text{if } \boldsymbol{x} \in \boldsymbol{N} \end{cases} \end{aligned}$$

This function is essentially bounded, because if $\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}$, we have

$$|s(\boldsymbol{x})| = \left|\sum_{k=0}^{\infty} f_k(\boldsymbol{x})\right| \leq \sum_{k=0}^{\infty} |f_k(\boldsymbol{x})| \leq \sum_{k=0}^{\infty} \|[f_k]\|_\infty = M,$$

where $M \in [0, +\infty)$ is an appropriate constant. We deduce that $|s(\boldsymbol{x})| \leq M$ a.e. on $\boldsymbol{A}$, from which we find $[s] \in L^\infty(\boldsymbol{A}, \mathbb{C})$. We need to verify that the sequence of partial sums tends to the equivalence class of f, i.e., that $[s] = \displaystyle\sum_{k=0}^{\infty} [f_k]$. Let $\ell \in \mathbb{N}$. We observe that

$$\left\| [s] - \sum_{k=0}^{\ell} [f_k] \right\|_\infty = \left\| \left[s - \sum_{k=0}^{\ell} f_k \right] \right\|_\infty,$$

and if $\boldsymbol{x} \notin \boldsymbol{A} \setminus N$, we have

$$\left| s(\boldsymbol{x}) - \sum_{k=0}^{\ell} f_k(\boldsymbol{x}) \right| = \left| \sum_{k=0}^{\infty} f_k(\boldsymbol{x}) - \sum_{k=0}^{\ell} f_k(\boldsymbol{x}) \right| = \left| \sum_{k=\ell+1}^{\infty} f_k(\boldsymbol{x}) \right| \leq \sum_{k=\ell+1}^{\infty} |f_k(\boldsymbol{x})| \leq \sum_{k=\ell+1}^{\infty} \|[f_k]\|_\infty.$$

It follows that

$$\left\| \left[s - \sum_{k=0}^{\ell} f_k \right] \right\|_\infty \leq \sum_{k=\ell+1}^{\infty} \|[f_k]\|_\infty \overset{\ell \to +\infty}{\longrightarrow} 0.$$

so that the series is convergent as required.

We have already seen that for $p, q \in \mathbb{R}$ satisfying $1 \leq p < q \leq +\infty$, both the inclusion $L^p \subseteq L^q$ and $L^p \supseteq L^q$ are false in general. However, when $\boldsymbol{A}$ has finite measure, it is possible to establish the following inclusions.

Theorem 4.8 *Let $\boldsymbol{A} \in \mathcal{M}_n$ be a set of finite measure. Then,*

$$L^\infty(\boldsymbol{A}, \mathbb{C}) \subseteq L^2(\boldsymbol{A}, \mathbb{C}) \subseteq L^1(\boldsymbol{A}, \mathbb{C}).$$

Proof Let $f \in L^\infty(\boldsymbol{A}, \mathbb{C})$. We need to show that $f \in L^2(\boldsymbol{A}, \mathbb{C})$. Since $f \in L^\infty(\boldsymbol{A}, \mathbb{C})$, there exists $M > 0$ such

that $|f(\boldsymbol{x})| \leq M$ a.e. on $\boldsymbol{A}$. Since $\boldsymbol{A}$ is measurable, then M^2 is integrable on $\boldsymbol{A}$ and

$$\int_{\boldsymbol{A}} M^2 \, \mathrm{d}\boldsymbol{x} = M^2 \int_{\boldsymbol{A}} \mathrm{d}\boldsymbol{x} = M^2 \, \mu_n(\boldsymbol{A}).$$

Since $|f(\boldsymbol{x})| \leq M$ a.e. on $\boldsymbol{A}$, then also $|f(\boldsymbol{x})|^2 \leq M^2$ a.e. on $\boldsymbol{A}$, so that $f \in L^2(\boldsymbol{A}, \mathbb{C})$.

Let us now assume that $f \in L^2(\boldsymbol{A}, \mathbb{C})$, and we show that $f \in L^1(\boldsymbol{A}, \mathbb{C})$. We can use the Cauchy-Schwarz inequality by observing that the constant function 1 is in $L^2(\boldsymbol{A}, \mathbb{C})$:

$$\left| \int_{\boldsymbol{A}} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} \right| \leq \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{2}} \cdot \left(\int_{\boldsymbol{A}} 1 \, \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{2}} = (\mu_n(\boldsymbol{A}))^{\frac{1}{2}} \cdot \left(\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} \right)^{\frac{1}{2}} < +\infty,$$

since $(\mu_n(\boldsymbol{A}))^{\frac{1}{2}}$ is bounded.

Example *4.7*

Consider the space $C([-1, 1]; \mathbb{R})$ with the norm $\| \cdot \|_2$, induced by the inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x) \, g(x) \, \mathrm{d}x.$$

Suppose that we consider the final-dimensional subspace of polynomial functions of degree less or equal to 2. We cannot determine an orthonormal basis for A using Gram-Schmidt because this vector space, together with this inner product, is not complete. However, if we consider instead $\boldsymbol{H} = L^2([-1, 1])$, then we have indeed a complete space, and therefore $\boldsymbol{H}$ with this norm is a Hilbert space. Let A be the subspace of 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 of $\boldsymbol{H}$, and a basis of A is $\{[1], [x], [x^2]\}$. In other words, the elements in A are functions which, almost everywhere in $[-1, 1]$, can be written as $\alpha + \beta \, x + \gamma \, x^2$ for some real coefficients α, β, γ.

In the following, following a standard convention, given $f \in L^2([-1, 1]; \mathbb{R})$, we shall write f instead of $[f]$, for the sake of notational simplicity. In this way, we will not be distinguishing notationally between functions and the corresponding equivalent classes. We will also simply say that the function f is defined in $[-1, 1]$ by only implicitly assuming that this is true almost everywhere.

Let $u_1(x) = 1$, with $x \in [-1, 1]$. We have

$$\|u_1\| = \sqrt{\int_{-1}^{1} 1^2 \, \mathrm{d}t} = \sqrt{2}.$$

We define the function $v_1 : [-1, 1] \longrightarrow \mathbb{R}$ as

$$v_1(x) = \frac{1}{\sqrt{2}}, \qquad x \in [-1, 1].$$

We compute

$$q_2 = u_2 - \langle u_2, v_1 \rangle \, v_1 = t - \int_{-1}^{2} \frac{x}{\sqrt{2}} \, \mathrm{d}x \, \frac{1}{\sqrt{2}} = x.$$

We find the norm of q_2:

$$\|q_2\| = \sqrt{\int_{-1}^{2} x^2 \, \mathrm{d}x} = \sqrt{\frac{2}{3}}.$$

Thus, we define the function $v_2 : [-1,1] \longrightarrow \mathbb{R}$ as

$$v_2(x) = \sqrt{\frac{3}{2}}\, x, \qquad x \in [-1,1].$$

Finally,

$$\begin{aligned} q_3 &= u_3 - \langle u_3, v_1\rangle\, v_1 - \langle u_3, v_2\rangle\, v_2 = x^2 - \int_{-1}^{2} \frac{x^2}{\sqrt{2}}\,\mathrm{d}x\, \frac{1}{\sqrt{2}} - \int_{-1}^{2} x^2 \left(\sqrt{\frac{3}{2}}\, x\right) \mathrm{d}x\, \sqrt{\frac{3}{2}}\, x \\ &= x^2 - \frac{\sqrt{2}}{3}\,\frac{1}{\sqrt{2}} - 0 \cdot \sqrt{\frac{3}{2}}\, x = x^2 - \frac{1}{3}. \end{aligned}$$

Next,

$$\|q_3\| = \frac{2}{3}\,\sqrt{\frac{2}{5}},$$

which leads to

$$v_3(x) = \frac{3}{2}\,\sqrt{\frac{5}{2}}\left(x^2 - \frac{1}{3}\right).$$

The set $\{[v_1], [v_2], [v_3]\}$ is an orthonormal basis for A.

Example *4.8*

Let

$$\boldsymbol{U} = \{[f] \in L^2(\mathbb{R};\mathbb{R}) \,|\, [x\, f] \in L^2(\mathbb{R};\mathbb{R})\},$$

where by $x\, f$ we mean the function defined over the same domain of f, i.e., $\mathbb{R} \setminus N$ where N has measure zero, and

$$\begin{aligned} x\, f : \mathbb{R} \setminus N &\longrightarrow \mathbb{R} \\ x &\mapsto x\, f(x). \end{aligned}$$

Show that:

(i) $\boldsymbol{U}$ is a subset of $L^1(\mathbb{R};\mathbb{R})$;

(ii) the expression

$$\|\,[f]\,\| = \sqrt{\int_{\mathbb{R}} (1 + x^2)\, |f(x)|^2\, \mathrm{d}x}$$

is a norm in $\boldsymbol{U}$;

(iii) $(\boldsymbol{U}, \|\cdot\|)$ is complete.

Solution 4.1 *(i) Let $[f] \in \boldsymbol{U}$. We show that $[f] \in L^1(\mathbb{R};\mathbb{R})$. We have*

$$\int_{\mathbb{R}} |f(x)|\,\mathrm{d}x = \int_{|x|\le 1} |f(x)|\,\mathrm{d}x + \int_{|x|>1} |f(x)|\,\mathrm{d}x.$$

Using the Cauchy-Schwartz inequality, the first integral in the right hand-side gives

$$\int_{|x|\le 1} 1 \cdot |f(x)|\,\mathrm{d}x \le \sqrt{\int_{|x|\le 1} 1\,\mathrm{d}x}\, \sqrt{\int_{|x|\le 1} |f(x)|^2\,\mathrm{d}x} \le 2 \cdot \sqrt{\int_{x\in\mathbb{R}} |f(x)|^2\,\mathrm{d}x} < +\infty,$$

because $[f] \in L^2(\mathbb{R};\mathbb{R})$.

(ii) We prove the homogeneity. Let $k \in \mathbb{R}$ and $[f] \in \boldsymbol{U}$. Then,

$$\begin{aligned} \|\,[k\,f]\,\| &= \sqrt{\int_{\mathbb{R}} (1+x^2)\,|k\,f(x)|^2\,\mathrm{d}x} \\ &= \sqrt{\int_{\mathbb{R}} k^2\,(1+x^2)\,|f(x)|^2\,\mathrm{d}x} \\ &= |k|\,\sqrt{\int_{\mathbb{R}} (1+x^2)\,|f(x)|^2\,\mathrm{d}x} = |k|\,\|\,[f]\,\|. \end{aligned}$$

We prove the subadditivity. Let $[f], [g] \in \boldsymbol{U}$. Then,

$$\begin{aligned} \|\,[f+g]\,\| &= \sqrt{\int_{\mathbb{R}} (1+x^2)\,|f(x)+g(x)|^2\,\mathrm{d}x} \\ &\leq \sqrt{\int_{\mathbb{R}} (1+x^2)\,(|f(x)|+|g(x)|)^2\,\mathrm{d}x} \\ \text{Correggi!}\quad &\leq \sqrt{\int_{\mathbb{R}} (1+x^2)\,|f(x)|^2\,\mathrm{d}x} + \sqrt{\int_{\mathbb{R}} (1+x^2)\,|g(x)|^2\,\mathrm{d}x} = \|\,[f]\,\| + \|\,[g]\,\|. \end{aligned}$$

The positivity is trivial.

(iii) We prove that, given a Cauchy sequence $([f_k])_{k\in\mathbb{N}}$ in $(\boldsymbol{U}, \|\cdot\|)$:

- *$([f_k])_{k\in\mathbb{N}}$ is a Cauchy sequence in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$;*
- *$([x\,f_k])_{k\in\mathbb{N}}$ is a Cauchy sequence in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$.*

For evert $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that, for all $n, m > p$, we have $\|\,[f_n - f_m]\,\| < \varepsilon$. Then

$$\begin{aligned} \|\,[f_n - f_m]\,\|_2^2 &= \int_{\mathbb{R}} |f_n(x) - f_m(x)|^2\,\mathrm{d}x \\ &\leq \int_{\mathbb{R}} (1+x^2)\,|f_n(x) - f_m(x)|^2\,\mathrm{d}x = \|\,[f_n - f_m]\,\|^2. \end{aligned}$$

This proves that $([f_k])_{k\in\mathbb{N}}$ is a Cauchy sequence in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$. Similarly,

$$\begin{aligned} \|\,[x\,f_n - x\,f_m]\,\|_2^2 &= \int_{\mathbb{R}} |x\,f_n(x) - x\,f_m(x)|^2\,\mathrm{d}x \\ &= \int_{\mathbb{R}} x^2\,|f_n(x) - f_m(x)|^2\,\mathrm{d}x \\ &\leq \int_{\mathbb{R}} (1+x^2)\,|f_n(x) - f_m(x)|^2\,\mathrm{d}x = \|\,[f_n - f_m]\,\|^2. \end{aligned}$$

This proves that $([x\,f_k])_{k\in\mathbb{N}}$ is a Cauchy sequence in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$.

Now, let $(f_k)_{k\in\mathbb{N}}$ be a sequence of functions that converges to f in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$, and let $(x\,f_k)_{k\in\mathbb{N}}$ converges to g in $(L^2(\mathbb{R};\mathbb{R}), \|\cdot\|_2)$. We prove that $x\,f = g$ almost everywhere in $\mathbb{R}$. First, we show that $x\,f = g$ almost everywhere in $[-1,1]$. Let $\varepsilon > 0$. There exists $p \in \mathbb{N}$ such that for all $k > p$ we have

$$\|\,[f_k - f]\,\|_{L^2([-1,1];\mathbb{R})} \leq \frac{\varepsilon}{2} \quad \textit{and} \quad \|\,[x\,f_k - g]\,\|_{L^2([-1,1];\mathbb{R})} \leq \frac{\varepsilon}{2}.$$

Then,

$$\begin{aligned}
\| [x f - g] \|^2_{L^2([-1,1];\mathbb{R})} &= \int_{-1}^{1} |x f(x) - g(x)|^2 \, \mathrm{d}x \\
&= \int_{-1}^{1} |x f(x) - x f_k(x) + x f_k(x) - g(x)|^2 \, \mathrm{d}x \\
&\leq \int_{-1}^{1} |x f(x) - x f_k(x)|^2 \, \mathrm{d}x + \int_{-1}^{1} |x f_k(x) - g(x)|^2 \, \mathrm{d}x \\
&\leq \int_{-1}^{1} x^2 |f(x) - f_k(x)|^2 \, \mathrm{d}x + \frac{\varepsilon^2}{4} \\
&\leq \int_{-1}^{1} |f(x) - f_k(x)|^2 \, \mathrm{d}x + \frac{\varepsilon^2}{4} \leq \frac{\varepsilon^2}{2}.
\end{aligned}$$

We conclude that $\| [x f - g] \|^2_{L^2([-1,1];\mathbb{R})} = 0$*, and thus* $x f = g$ *almost everywhere in* $[-1, 1]$*. Let us now consider* $(-\infty, -1) \cup (1, +\infty)$*. Consider the norm*

$$\begin{aligned}
\left\| \left[f - \frac{g}{x} \right] \right\|^2_{L^2((-\infty,-1)\cup(1,+\infty);\mathbb{R})} &= \int_{|x|>1} \left| f(x) - \frac{g(x)}{x} \right|^2 \, \mathrm{d}x \\
&= \int_{|x|>1} \left| f(x) - f_k(x) + f_k(x) - \frac{g(x)}{x} \right|^2 \, \mathrm{d}x \\
&\leq \int_{|x|>1} |f(x) - f_k(x)|^2 \, \mathrm{d}x + \int_{|x|>1} \left| f_k(x) - \frac{g(x)}{x} \right|^2 \, \mathrm{d}x \\
&\leq \frac{\varepsilon^2}{4} + \int_{|x|>1} \frac{1}{x^2} |x f_k(x) - g(x)|^2 \, \mathrm{d}x \\
&\leq \frac{\varepsilon^2}{4} + \int_{|x|>1} |x f_k(x) - g(x)|^2 \, \mathrm{d}x \leq \frac{\varepsilon^2}{4} + \frac{\varepsilon^2}{4} = \frac{\varepsilon^2}{2}.
\end{aligned}$$

This shows that $x f = g$ *almost everywhere in* $\mathbb{R}$*. We have shown that* $x f_k$ *converges to* $x f$ *in* L^2*. Now, we show that* f_k *converges to* f *in* $\boldsymbol{U}$*:*

$$\begin{aligned}
\| [f_k - f] \|^2 &= \int_{\mathbb{R}} (1 + x^2) |f_k(x) - f(x)|^2 \, \mathrm{d}x \\
&= \int_{\mathbb{R}} |f_k(x) - f(x)|^2 \, \mathrm{d}x + \int_{\mathbb{R}} x^2 |f_k(x) - f(x)|^2 \, \mathrm{d}x \\
&= \int_{\mathbb{R}} |f_k(x) - f(x)|^2 \, \mathrm{d}x + \int_{\mathbb{R}} |x f_k(x) - x f(x)|^2 \, \mathrm{d}x.
\end{aligned}$$

The first integral in the right hand-side is 0 because we have convergence in L^2*. The second converges to zero because* $x f_k$ *converges to* $x f$ *in* L^2*.*

Exercises

Exercise 4.7

Consider the function

$$\begin{aligned}
u : (0, +\infty) &\longrightarrow \mathbb{R} \\
x &\mapsto \begin{cases} x^{-3/8} & \text{if } 0 < x < 1 \\ x^{-1} & \text{if } 1 \leq x < +\infty \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Find if u is in $L^1((0, +\infty), \mathbb{R})$, and/or $L^2((0, +\infty), \mathbb{R})$, and/or $L^\infty((0, +\infty), \mathbb{R})$. [Solution: (4.7)]

Exercise 4.8

Find the maximum, the minimum (if they are defined), the supremum, the infimum, the essential supremum and the essential infimum of the following functions:

(i) $f : \mathbb{R} \longrightarrow \mathbb{R}$ described, for all $x \in \mathbb{R}$, by $f(x) = \begin{cases} 3 & \text{if } x \in \mathbb{R} \setminus \{-1, +1\} \\ 4 & \text{if } x = -1 \\ -5 & \text{if } x = 1 \end{cases}$

(ii) $f : \mathbb{R} \longrightarrow \mathbb{R}$ described, for all $x \in \mathbb{R}$, by $f(x) = \begin{cases} x & \text{if } x \in \mathbb{Q} \\ \arctan x & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$

(iii) $f : \mathbb{R} \longrightarrow \mathbb{R}$ described, for all $x \in \mathbb{R}$, by $f(x) = x^3$.

[Solution: (4.8)]

Exercise 4.9

Using the Hölder's inequality in a generic Hilbert space, prove that if $(u_k)_{k\in\mathbb{N}}$ and $(v_k)_{k\in\mathbb{N}}$ are sequences of $L^2(A, \mathbb{C})$ that converge in $L^2(A; \mathbb{C})$ to the two functions u and v, respectively, then $u_k\, v_k \overset{k\to\infty}{\longrightarrow} u\, v$ in $L^1(A; \mathbb{C})$. Moreover, show that if $u_k \overset{k\to\infty}{\longrightarrow} 0$ in $L^2(A, \mathbb{C})$ and $(v_k)_{k\in\mathbb{N}}$ is bounded in $L^2(A; \mathbb{C})$, then $u_k\, v_k \overset{k\to\infty}{\longrightarrow} 0$ in $L^1(A; \mathbb{C})$. [Solution: (4.9)]

Exercise 4.10

Let $\lambda \in \mathbb{R}$ and consider the sequence $(u_k)_{k\in\mathbb{N}^*}$ described by

$$\begin{aligned} u_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto k^\lambda \chi_{[0,\frac{1}{k}]}(x). \end{aligned}$$

For $p = 1, 2, +\infty$, find for which values of λ we have $\|u_k\|_p \overset{k\to\infty}{\longrightarrow} 0$. [Solution: (4.10)]

Exercise 4.11

Verify that the following sequences converge a.e. in $\mathbb{R}$, and find for which values of $p \in \{1, 2, +\infty\}$ they converge in $L^p(\mathbb{R}, \mathbb{R})$:

(i) $e^{-k|x|}$;

(ii) $e^{-|x-k|}$;

(iii) $k^{1/2}\, e^{-k^2|x|}$;

(iv) $e^{-k|x-k|}$;

(v) $\frac{1}{k}\, e^{-\frac{|x|}{k}}$;

(vi) $k\, e^{-k|x|}$.

[Solution: (4.11)]

Exercise 4.12

Verify that the following sequences converge a.e. in $\mathbb{R}$, and find for which values of $p \in \{1, 2, +\infty\}$ they converge in $L^p(\mathbb{R}, \mathbb{R})$:

(i) $k^{-\frac{1}{2}} \chi_{[k,k+1]}$;

(ii) $k^{-\frac{1}{2}} \chi_{[k,2k]}$;

(iii) $k\chi_{[0,\frac{1}{k}]}$;

(iv) $k\chi_{[k,k+\frac{1}{k}]}$;

(v) $k^{-\frac{2}{3}} \chi_{[0,k]}$;

(vi) $k^{-\frac{1}{2}} \chi_{[-k,k^2]}$.

[Solution: (4.12)]

Exercise 4.13 Determine which one(s) of the following functions is in $L^1(\mathbb{R};\mathbb{R})$ (or, in other words, which one is summable):

(i) $f(x) = c$, where $c \in \mathbb{R}$;

(ii) $f(x) = \chi_{[a,b]}(x)$, where $a, b \in \mathbb{R}$ are such that $a < b$;

(iii) $f(x) = e^{-x}$;

(iv) $f(x) = e^{x}$;

(v) $f(x) = e^{-|x|}$;

(vi) $f(x) = e^{-x^2}$;

(vii) $f(x) = \sin x$.

[Solution: (4.13)]

Exercise 4.14 Determine whether the function

(i) Find if the function $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ defined by $f(x) = \arctan \frac{1}{x}$ is in $L^1(\mathbb{R} \setminus \{0\};\mathbb{R})$;

(ii) Find if the function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by $f(\boldsymbol{x}) = e^{-\|\boldsymbol{x}\|^2}$ is in $L^1(\mathbb{R}^n;\mathbb{R})$ (Hint: use Fubini's Theorem).

[Solution: (4.14)]

Exercise 4.15 Find if the function $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ defined by $f(x) = \arctan \frac{1}{x}$ is in $L^2(\mathbb{R} \setminus \{0\};\mathbb{R})$. [Solution: (4.15)]

Exercise 4.16 Establish if the functions

$$f : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} \frac{1}{x} & \text{for } x \geq 1 \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad g : \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto \begin{cases} \frac{1}{\sqrt{x}} & \text{for } x \in (0,1] \\ 0 & \text{otherwise} \end{cases}$$

are in $L^1(\mathbb{R};\mathbb{R})$ and/or in $L^2(\mathbb{R};\mathbb{R})$. [Solution: (4.16)]

5 Distributions

In this part of the unit we introduce the concept of *distribution*, also known as *generalized function*, which is a pillar of modern mathematics, and it is the natural instrument to describe impulsive phenomena and for the study of differential equations. Dirac systematically used the δ distribution in his research into quantum mechanics, but the rigorous mathematical foundations of a theory of distributions was developed by Sobolev and Schwartz. A simple example of a situation where distributions arise is the collision of a point particle P of mass m which is moving with a uniform linear motion on a horizontal plane (without friction) against a wall, perpendicular to the direction of the motion. If the collision is elastic, we would like the fundamental laws of mechanics

$$m\,(v_2 - v_1) = \int_{t_1}^{t_2} F(t)\,\mathrm{d}t \tag{5.1}$$

to continue to hold, where v_1 and v_2 are the velocities of P at time t_1 and t_2, respectively, and $F(t)$ is the overall force acting on P at time t. We denote by t_0 the instant of the collision:

- $t_1 < t_2 < t_0$ or $t_0 < t_1 < t_2$: in these cases $v_1 = v_2$ and F is zero in $[t_1, t_2]$;
- if $t_1 < t_0 < t_2$, the left-hand side of (5.1) is $2\,m\,v_2 \neq 0$ since $v_1 = -v_2 \neq 0$, whereas the right hand-side is zero since $F(t) = 0$ for all $t \neq t_0$.

Things work out if we introduce new objects, which are similar to functions, that can be zero except for a single point and, nevertheless, have a non-zero integral. These objects are indeed new, because the integral of any "ordinary" function equal to zero for $t \neq t_0$ must be zero. In the case of the collision, this new object will be a *distribution*: it will be a multiple of the so-called *Dirac delta* δ at t_0:

$$2\,m\,v_2\,\delta(t - t_0).$$

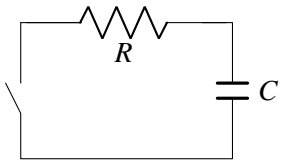
Distributions are also extremely useful in the study of the solutions of partial differential equations. For example, non-linear partial differential equations describing waves necessitate the possibility to accept discontinuous solutions. Classical theories of differential equations are not suitable to that end. For example, consider the characteristic function of $[0, +\infty)$

$$\begin{aligned} \psi : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} 1 & t \in [0, +\infty) \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

and the function $v = 2\,\psi$. These two functions are treated exactly in the same way if we consider their derivative: they both have derivative equal to zero in $\mathbb{R} \setminus \{0\}$ and neither of them is differentiable at $t = 0$. A "good definition" of discontinuous solution of a differential equation needs to also take into account the "jump" of the function at the origin.

Another simple example of a situation where we need a mathematical description of an impulsive behaviour is the following: consider the following circuit with a resistance R and a capacitor C, and a

switch.



The capacitor, initially charged, starts to uncharge from the instant $t_0 = 0$ in which the switch is closed. The potential of the capacitor decays from the initial value V_0 with the law[1]

$$v(t) = V_0\, e^{-\frac{t}{RC}}.$$

Thus, the current is

$$i(t) = \frac{v(t)}{R} = \frac{V_0}{R}\, e^{-\frac{t}{RC}}.$$

We deduce

$$\int_0^{+\infty} i(t)\,\mathrm{d}t = \int_0^{+\infty} \frac{V_0}{R}\, e^{-\frac{t}{RC}}\,\mathrm{d}t = \frac{V_0}{R}\,(-R\,C)\lim_{y\to+\infty}\left[e^{-\frac{t}{RC}}\right]_0^y = -C\,V_0\,(0-1) = C\,V_0.$$

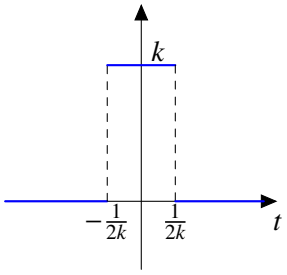
The integral of the current does not depend on the resistance R. When the value R tends to zero, the current $i(t)$ tends to decay more rapidly. At the limit, $i(t)$ tends to be zero for every $t > 0$, and it tends to be ∞ at $t = 0$, while its integral remains the same. This limiting behavior cannot be described by considering $i(t)$ as a function of time, because – as already pointed out – the integral of a function that is zero everywhere except for a single point cannot be $\neq 0$.

A simple but incorrect way of defining an object (we will call it δ) that behaves in the way described above, i.e., such that $\delta(t) = 0$ for every $t \neq 0$ but such that

$$\int_{-\infty}^{+\infty} \delta(t)\,\mathrm{d}t = 1, \tag{5.2}$$

is the following: consider the sequence of functions

$$\Delta_k(t) = \begin{cases} k & \text{if } t \in \left[-\frac{1}{2\,k}, \frac{1}{2\,k}\right]; \\ 0 & \text{otherwise} \end{cases}$$



We clearly have for all $k \in \mathbb{N}^*$

$$\int_{-\infty}^{+\infty} \Delta_k(t)\,\mathrm{d}t = 1,$$

[1] To see this, write Kirchhoff's second law of the closed loop as $v(t) + R\,i(t) = 0$, and since $i(t) = C\,\frac{\mathrm{d}v(t)}{\mathrm{d}t}$, we obtain $v(t) + R\,C\,\frac{\mathrm{d}v(t)}{\mathrm{d}t} = 0$, which is the linear differential equation $\dot{v}(t) = -\frac{1}{RC}\,v(t)$ with initial condition $v(0) = V_0$.

and therefore it seems natural to define δ as the limit

$$\delta = \lim_{k \to +\infty} \Delta_k. \tag{5.3}$$

In fact, intuitively speaking, these functions tend to be zero everywhere except at the point $t = 0$, and at $t = 0$ they have an infinite amplitude. However, their integral over the real line seems to be independent of k and equal to 1, and therefore at the limit we have (5.2).

This way of proceeding is very widespread in many areas of science and engineering, but it is incorrect for several reasons. Loosely speaking, the limit (5.3) is ill-defined: it would give

$$\delta(t) = \begin{cases} +\infty & \text{for } t = 0 \\ 0 & \text{otherwise} \end{cases}$$

and $+\infty$ is not a number, and so this equality is meaningless. More formally, using the appropriate language we would need to ask ourselves if the operation (5.3) is well legitimate, and obviously it is not: the sequence $(\Delta_k)_{k \in \mathbb{N}}$ is not pointwise convergent, because the sequence of numbers $(\Delta_k(0))_{k \in \mathbb{N}}$ is not convergent. If it is not pointwise convergent, it is not uniformly convergent, and therefore bringing the limit outside the integral in

$$\int_{-\infty}^{+\infty} \delta(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} \lim_{k \to +\infty} \Delta_k(t)\,\mathrm{d}t = \lim_{k \to +\infty} \int_{-\infty}^{+\infty} \Delta_k(t)\,\mathrm{d}t = \lim_{k \to +\infty} 1 = 1$$

is incorrect.

From these considerations, it turns out that we need more sophisticated mathematical tools to define this special object δ: this will lead to the notion of distribution. It will not be possible to say what the value of a distribution is at a specific point, but we will describe these objects from the way they act on other functions: we could say that we are not going to define what a distribution is, but rather what it does.

Before continuing, we present another example arising from physics, which is particularly important historically, being closer to what motivated Dirac to introduce distributions. Suppose that we want to determine the density $\delta(\boldsymbol{x})$ of a material point of unit mass, that we assume to be at the origin of our coordinates $\boldsymbol{x}$. We calculate $\delta(\boldsymbol{x})$ as the limit of the average density by first assuming that the mass is *distributed* inside a sphere $\overline{\mathcal{B}_\eta(0)}$. The average density is

$$\begin{aligned} f_\eta : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \frac{3}{4\pi\eta^3} & \text{if } \|\boldsymbol{x}\| \leq \eta \\ 0 & \text{if } \|\boldsymbol{x}\| > \eta. \end{cases} \end{aligned}$$

Defining δ as the limit of f_η, as $\eta \to 0$, would give

$$\delta(\boldsymbol{x}) = \lim_{\eta \to 0} f_\eta(\boldsymbol{x}) = \begin{cases} +\infty & \text{if } \|\boldsymbol{x}\| = \boldsymbol{0} \\ 0 & \text{if } \|\boldsymbol{x}\| \neq \boldsymbol{0}. \end{cases} \tag{5.4}$$

It is also required that the integral of the density δ over any volume $\boldsymbol{V}$ gives the mass:

$$\int_V \delta(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \boldsymbol{V} \\ 0 & \text{if } \boldsymbol{x} \notin \boldsymbol{V}. \end{cases}$$

On the other hand, the integral $\int_V \delta(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ should always be zero from (5.4) if we interpret this integral as a Lebesgue integral.

To solve the problem, we can analyse what happens at the limit as $\eta \to 0$ when f_η acts on a continuous function φ. In particular, we are interested in showing that

$$\lim_{\eta\to 0}\int_{\mathbb{R}^3} f_\eta(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \varphi(0).$$

Since φ is continuous at the origin, for every $\varepsilon > 0$ there exists $\zeta > 0$ such that $\|\boldsymbol{x}\| < \zeta$ implies $|\varphi(\boldsymbol{x}) - \varphi(0)| \le \varepsilon$. It follows that, for $\eta \le \zeta$, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^3} f_\eta(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \varphi(0)\right| &= \left|\int_{\|\boldsymbol{x}\|\le\eta} \frac{3}{4\,\pi\,\eta^3}\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \varphi(0)\right| \\ &= \frac{3}{4\,\pi\,\eta^3}\left|\int_{\|\boldsymbol{x}\|\le\eta} (\varphi(\boldsymbol{x}) - \varphi(0))\,\mathrm{d}x\right| \\ &\le \frac{3}{4\,\pi\,\eta^3}\int_{\|\boldsymbol{x}\|\le\eta} \varepsilon\,\mathrm{d}\boldsymbol{x} = \varepsilon,\end{aligned}$$

as required. Loosely speaking, f_η tends, as $\eta \to 0$, to a functional that maps φ to $\varphi(0)$, and this will be, for us, the definition of the Dirac delta function.

5.1 Support and test functions

The definition of distribution is one of the most delicate points in modern mathematics. As explained in the introduction, distributions are odd objects that cannot be defined pointwise: such attempt would result in the need of impose their value to be infinity at specific points. The standard approach is to specify how a distribution T operates when it is applied to another function φ. Loosely speaking, this action will be expressed by the relation

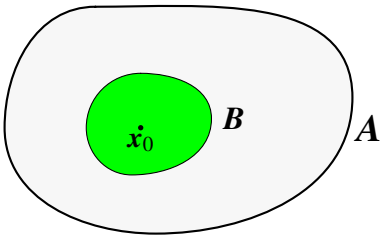
$$T(\varphi) = \int_{\mathbb{R}^n} T\,\varphi. \tag{5.5}$$

Since a distribution has a very irregular type of behaviour (a pulse, intuitively, is certainly not smooth), the first step consists in applying T to a function φ that can be as regular as it can be, so as to ensure that the expression (5.5) is well defined (and, as we shall see, distributions will be defined linear continuous functionals – we will see what we mean by "continuous" later on – over the space of test functions: the narrower the space of test functions, the more linear continuous functionals there are over it). In first instance, we choose φ to be differentiable for all degrees of differentiation (i.e., to be of class C^∞), and also to be equal to zero outside a compact subset of $\mathbb{R}^n$, so as to avoid problems of convergence of the integral in (5.5). Hence, the first step consists in defining the set $\mathcal{D}$ of C^∞ functions with compact support. Once this set of extremely well-behaved and regular functions has been defined, we will define the set of distributions as the dual $\mathcal{D}'$ of $\mathcal{D}$.

In what follows, we consider a number $n \in \mathbb{N}^*$, and we consider functions from an open set $\boldsymbol{A} \subseteq \mathbb{R}^n$ to the complex numbers. All the results presented in these notes do not change for functions from $\boldsymbol{A} \subseteq \mathbb{R}^n$ to real numbers.

Remark 5.1 Consider a function f defined almost everywhere on a measurable set $\boldsymbol{A} \in \mathcal{M}_n$. Consider a subset $\boldsymbol{B}$ of $\boldsymbol{A}$. Then, it is easy to see that f is also defined almost everywhere in $\boldsymbol{B}$. Indeed, since f is defined almost everywhere in $\boldsymbol{A}$, there exists a null set $\boldsymbol{N} \subseteq \mathbb{R}^n$ such that $\boldsymbol{A} \setminus \boldsymbol{N}$ is the domain of f. This implies that $f(\boldsymbol{x})$ is defined for all $\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N}$, and therefore $f(\boldsymbol{x})$ is defined for all $\boldsymbol{x} \in \boldsymbol{B} \setminus \boldsymbol{N}$. Thus, f is defined almost everywhere in $\boldsymbol{B}$. With abuse of notation, we denote by $f|_{\boldsymbol{B}}$ the function $f|_{\boldsymbol{B} \cap \boldsymbol{N}}$.

We now introduce the notion of essential support. Consider a complex-valued function f defined almost everywhere in $\boldsymbol{A}$. The essential support of f is the subset of $\boldsymbol{A}$, denoted by $\operatorname{supp} f$, that satisfies the following property: $x_0 \in \boldsymbol{A}$ does **not** belong to $\operatorname{supp} f$ if there exists an open neighbourhood $\boldsymbol{B} \subset \boldsymbol{A}$ of $\boldsymbol{x}_0$ such that $f(\boldsymbol{x})$ is zero almost everywhere in $\boldsymbol{B}$.



Definition 5.1 [Support of a Measurable Function]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be an open set. Let f be a complex-valued measurable function defined almost everywhere in $\boldsymbol{A}$. Then, we define the essential support *(or simply* support*) of f as the set*

$$\operatorname{supp} f \stackrel{def}{=} \boldsymbol{A} \setminus \{\boldsymbol{x} \in \boldsymbol{A} \mid \exists\, \boldsymbol{B} \subseteq \boldsymbol{A} \text{ open with } \boldsymbol{x} \in \boldsymbol{B} : f|_{\boldsymbol{B}} = 0 \text{ almost everywhere}\}.$$

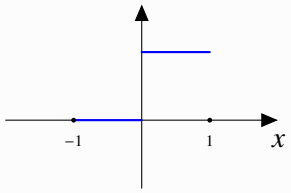
In Definition 5.1, the restriction $f|_B$ is intended in the sense of Remark 5.1. Moreover, $\boldsymbol{A}$ is measurable because it is assumed to be open.

A simple consequence of Definition 5.1 is the fact that the essential support of a function is the complement in $\boldsymbol{A}$ of the largest open set for which $f = 0$ almost everywhere. The following example shows that $\operatorname{supp} f$ is not necessarily a closed set.

Example *5.1*

Consider the function

$$\begin{aligned} f : (-1,1) &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & x > 0 \\ 0 & -1 < x \leq 0 \end{cases} \end{aligned}$$



The points in $(-1,0)$ do not belong to $\operatorname{supp} f$, because for each point $x_0 \in (-1,0)$ there exists a neighbourhood of x_0 for which the function f restricted to it is zero. Consider the point $x = 0$. The interval $(-\varepsilon, \varepsilon)$ centred at $x = 0$ is such that $\{x \in (-\varepsilon,\varepsilon) \mid f(x) \neq 0\} = (0,\varepsilon)$, whose measure is positive for any $\varepsilon > 0$. Then $0 \in \operatorname{supp} f$. Thus, in this case $\operatorname{supp} f = [0,1)$. Therefore, the essential support is not a closed set in general, even if it is the complement of an open set. This is because it is not the complement of an open set with respect

to a closed set, but with respect to A, which in this case is the open interval $(-1, 1)$.[a]

[a] This distinction is at the basis of the so-called *relative topology*. The set $[0, 1)$ is not closed in $\mathbb{R}$, because the point 1 is adherent to $[0, 1)$ but it is not contained in it. On the other hand, when considered as a subset of $A = (-1, 1)$, the set $[0, 1)$ is closed: the point 1 is not an element of A, and so is no longer considered an adherent point of $[0, 1)$, and therefore $[0, 1)$ contains all its adherent points.

It is easy to see that two functions that are coincident almost everywhere have the same essential support, as the following lemma shows.

Lemma 5.1 *Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. Let f, g be complex-valued measurable functions defined almost everywhere in $\boldsymbol{A}$ and such that $f = g$ almost everywhere in $\boldsymbol{A}$. Then,* $\operatorname{supp} f = \operatorname{supp} g$.

Proof Let $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$ be null sets such that $f : \boldsymbol{A} \setminus N_1 \longrightarrow \mathbb{C}$ and $g : \boldsymbol{A} \setminus \boldsymbol{N}_2 \longrightarrow \mathbb{C}$. The set

$$\boldsymbol{N} = \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}$$

has measure zero by assumption. To show that $\operatorname{supp} f = \operatorname{supp} g$, we prove that $(\operatorname{supp} f)^{\complement} = (\operatorname{supp} g)^{\complement}$. This amount to showing that, given $\boldsymbol{x}_0 \in \boldsymbol{A}$, if $\boldsymbol{x}_0 \notin \operatorname{supp} f$, then also $\boldsymbol{x}_0 \notin \operatorname{supp} g$. By definition, $\boldsymbol{x}_0 \notin \operatorname{supp} f$ implies that there exists an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ containing $\boldsymbol{x}_0$ such that $f(\boldsymbol{x}) = 0$ for almost all $\boldsymbol{x} \in \boldsymbol{B}$. Thus, the set $\boldsymbol{N}_3 \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \boldsymbol{B} \setminus \boldsymbol{N}_1 \,|\, f(\boldsymbol{x}) \neq 0\}$ has zero Lebesgue measure. It follows that the set $\boldsymbol{B} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3)$ also has zero Lebesgue measure. For all $\boldsymbol{x} \in \boldsymbol{B} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2 \cup \boldsymbol{N}_3)$, both $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ exist, they are equal, and $f(\boldsymbol{x}) = 0$, so that $g(\boldsymbol{x}) = 0$. This implies that $g(\boldsymbol{x}) = 0$ for almost all $\boldsymbol{x} \in \boldsymbol{B}$. Then, $\boldsymbol{x}_0 \notin \operatorname{supp} g$.

Remark 5.2 The notion of essential support should not be confused with another notion of support used in other branches of mathematics. In particular, the so-called *closed support* is defined for a function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ as

$$\overline{\{\boldsymbol{x} \in A \,|\, f(\boldsymbol{x}) \neq 0\}}. \tag{5.6}$$

Consider for example the Dirichlet function $f : [0, 1] \longrightarrow \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases} \tag{5.7}$$

The essential support is empty, because f is equal almost everywhere to the zero function. The closed support (5.6) is the entire interval $[0, 1]$. This shows that the essential support of a function can be smaller than the closed support.

In this part of the unit, we will only be dealing with the essential support and, for this reason, we accept the abuse of language consisting in referring to it simply as the "support" of a function.

The following result shows that the support of a function f defined almost everywhere over an open set $\boldsymbol{A}$ is closed in $\boldsymbol{A}$, see Definition 2.2. In other words, if $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ is a sequence of points in $\operatorname{supp} f$ that converges to a point $\hat{\boldsymbol{x}}$ of $\boldsymbol{A}$ (or, stated differently, there exists $\hat{\boldsymbol{x}} \in \boldsymbol{A}$ such that $\lim_{k \to \infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}$), then $\hat{\boldsymbol{x}} \in \operatorname{supp} f$.

Lemma 5.2 *Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. Let f be a complex-valued measurable function defined almost everywhere in $\boldsymbol{A}$. Then,* supp f *is closed in $\boldsymbol{A}$.*

Proof Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence in supp f that converges to $\hat{\boldsymbol{x}} \in \boldsymbol{A}$. Suppose by contradiction that $\hat{\boldsymbol{x}} \notin \text{supp}\, f$. Then, there exists an open neighbourhood $\boldsymbol{B}$ of $\hat{\boldsymbol{x}}$ such that $f(\boldsymbol{x}) = 0$ for almost every $\boldsymbol{x} \in \boldsymbol{B}$. For large values of k we have $\boldsymbol{x}_k \in \boldsymbol{B}$, so that $\boldsymbol{x}_k \notin \text{supp}\, f$, which is a contradiction since $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a sequence of points in supp f.

Example *5.2*

The set $[0, 1)$ is closed in $A = (-1, 1)$. In fact, the limit of a sequence in $[0, 1)$ is in $[0, 1]$. Indeed, if by contradiction the limit ℓ is not in $[0, 1]$, there is a ball $\mathcal{B}_r(\ell)$ that has empty intersection with $[0, 1]$, and as $k \to \infty$ the elements of this sequence are not contained in such ball. If we assume that $\ell \in A$, then $\ell \neq 1$, and therefore ℓ must be in $[0, 1]$ but not equal to 1, so that $\ell \in [0, 1)$. Hence, $[0, 1)$ is closed in A.

Exercises

Exercise 5.1

Consider two continuous functions $f, g : A \longrightarrow \mathbb{C}$. Show that

(i) $f = 0 \quad \Leftrightarrow \quad \text{supp}\, f = \varnothing$;

(ii) $\text{supp}(f\, g) \subseteq \text{supp}\, f \cap \text{supp}\, g$.

[Solution: (5.1)]

Definition 5.2 [Measurable Functions with Compact Support]
Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. Let f be a complex-valued measurable function defined almost everywhere in $\boldsymbol{A}$. The function f is said to have compact support *if* supp f *is a compact subset of $\boldsymbol{A}$.*

Example *5.3*

We found that the support of the function

$$f : (-1, 1) \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} 1 & x > 0 \\ 0 & -1 < x \leq 0 \end{cases}$$

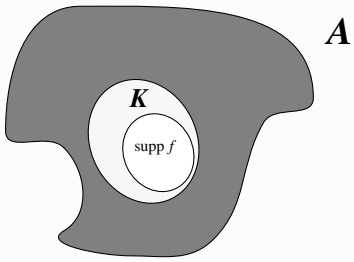
is $[0, 1)$. Therefore, f does not have compact support.

We recall a few notions about compact sets. Given a normed space $(\boldsymbol{X}, \|\cdot\|)$ and $\boldsymbol{K} \subseteq \boldsymbol{X}$, we recall that $\boldsymbol{K}$ is *sequentially compact* if every sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in $\boldsymbol{K}$ has a convergent subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ which converges to a point in $\boldsymbol{K}$. It is not difficult to show that the union and intersection of two compact sets is compact. It is possible to show that, if $\boldsymbol{K}$ is compact, then K is bounded and closed. If the normed space considered is Euclidean and endowed with the 2-norm, the converse is also true: this is the generalization of the Heine-Borel theorem for $\mathbb{R}^n$: consider the normed space $(\mathbb{R}^n, \|\cdot\|_2)$, and let $K \subseteq \mathbb{R}^n$. Then, K is compact if and only if K is bounded and closed.

It is easy to see that a Lebesgue measurable function f defined almost everywhere in $\boldsymbol{A}$ has compact support if and only if there exists a compact set $\boldsymbol{K} \subset \boldsymbol{A}$ such that $f(\boldsymbol{x})$ is zero almost everywhere in $\boldsymbol{A} \setminus \boldsymbol{K}$, as the following result shows. This result is consistent with what we obtained in the previous example, since supp $f = [0, 1)$ is not contained in any compact subset of $A = (-1, 1)$.

Theorem 5.1 *Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. Let f be a complex-valued measurable function defined almost everywhere in $\boldsymbol{A}$. Then, the support of f is contained in a compact set of $\boldsymbol{A}$ if and only if* supp f *is compact.*

Proof **(If)**. It is obvious that if supp f is compact, then the support of f is contained in a compact set of $\boldsymbol{A}$.
(Only if). We now show that if supp $f \subseteq \boldsymbol{K}$, where $\boldsymbol{K}$ is a compact subset of $\boldsymbol{A}$, then supp f is compact. We have to show in particular that supp f is bounded and closed. The fact that supp f is bounded is trivial, since it is contained in the compact set $\boldsymbol{K}$ which is bounded. We now show that supp f is closed, i.e., that the limit of any convergent sequence of points of supp f is in supp f. Consider a convergent sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points in supp f, and let us denote by $\hat{\boldsymbol{x}}$ the limit. Since K is compact, there exists a subsequence of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ that converges to $\hat{\boldsymbol{x}} \in \boldsymbol{K}$ (obviously the subsequence will converge to the same limit of the original sequence, but now we know that the limit is in $\boldsymbol{K}$). Since $\boldsymbol{K} \subset \boldsymbol{A}$, then $\hat{\boldsymbol{x}} \in \boldsymbol{A}$.



This subsequence is itself a sequence of points of supp f which converges to a point of $\boldsymbol{A}$. Therefore, in view of Lemma 5.2, which established that supp f is closed in $\boldsymbol{A}$, we have $\hat{\boldsymbol{x}} \in$ supp f.

A typical application of this result is for a function $f : \mathbb{R}^n \longrightarrow \mathbb{C}$. If this function has bounded support, it also has compact support. In fact, if supp f is bounded, there exists $R > 0$ such that $\overline{\mathcal{B}_R(0)} \supseteq$ supp f, and Theorem 5.1 can be applied because $\overline{\mathcal{B}_R(0)}$ is compact.

It is also obvious that if f is continuous, the fact that f is zero almost everywhere in B also implies that f is zero everywhere in B. We now introduce the concept of test functions. Such functions are differentiable an arbitrary number of times and have compact support. We recall that $C^\infty(\boldsymbol{A};\mathbb{C})$ denotes the set of all continuous function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ whose partial derivatives of all orders with respect to all the variables exist (and they are automatically continuous in $\boldsymbol{A}$).

Definition 5.3 [The Space $\mathcal{D}(\boldsymbol{A};\mathbb{C})$]
Let $\boldsymbol{A}$ be open set of $\mathbb{R}^n$. We denote by $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ the set of infinitely differentiable functions whose support is compact:

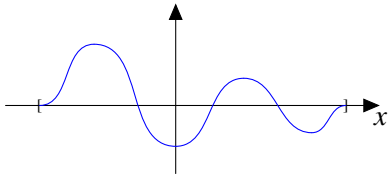
$$\mathcal{D}(\boldsymbol{A};\mathbb{C}) \stackrel{def}{=} \{f \in C^\infty(\boldsymbol{A};\mathbb{C}) \mid \text{supp } f \text{ is compact}\}.$$

The elements of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ are called test functions *on $\boldsymbol{A}$.*

Very often, the codomain is dropped in the symbol $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, i.e., you will often find the notation $\mathcal{D}(\boldsymbol{A})$, as well as $C_0^\infty(\boldsymbol{A})$.

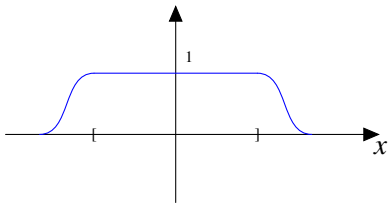
It is easy to see that $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ is a vector space, see Exercise 5.2. In particular, it contains the identically zero function, whose support is contained in the compact set $\boldsymbol{K} = \varnothing$. Loosely speaking, test functions on

A are infinitely differentiable functions that are zero outside a certain compact set contained in $\boldsymbol{A}$.



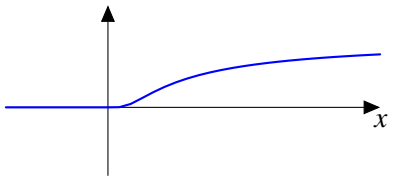
We will also prove that if $\boldsymbol{A}$ is open, given a compact subset of $\boldsymbol{A}$, it is possible to have a test function whose value is equal to 1 over $\boldsymbol{K}$:

$$\exists\, f \in \mathcal{D}(\boldsymbol{A};\mathbb{C}): \qquad \forall\, \boldsymbol{x} \in \boldsymbol{K},\quad f(\boldsymbol{x}) = 1.$$



In order to prove that such functions exist, we first observe that we can define a function ψ in $C^\infty(\mathbb{R};\mathbb{R})$ such that $\psi(x) = 0$ for $x \le 0$ and $\psi(x) > 0$ for $x > 0$. One example is the function

$$\begin{aligned} \psi : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} e^{-\frac{1}{x}} & x > 0 \\ 0 & x \le 0. \end{cases} \end{aligned}$$



Indeed, ψ is continuous, since $\lim_{x\to 0^+} e^{-\frac{1}{x}} = 0$. We now show that ψ is differentiable and its derivative is continuous. First, we observe that the restrictions $\psi|_{(0,+\infty)}$ and $\psi|_{(-\infty,0)}$ are both differentiable, and their derivatives are, respectively,

$$\psi'|_{(0,+\infty)}(x) = \frac{1}{x^2}\, e^{-\frac{1}{x}},\ x \in (0,+\infty) \qquad \text{and} \qquad \psi'|_{(-\infty,0)}(x) = 0,\ x \in (-\infty,0).$$

The only point where ψ is potentially not differentiable is $x = 0$. However, the left and right derivatives coincide at $x = 0$. Indeed, using the change of variable $y = 1/x$, we find

$$\lim_{x\to 0^+} \psi'(x) = \lim_{x\to 0^+} \frac{1}{x^2}\, e^{-\frac{1}{x}} = \lim_{y\to+\infty} y^2\, e^{-y} = 0.$$

It follows that the derivative of ψ is

$$\begin{aligned} \psi' : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \dfrac{1}{x^2}\, e^{-\frac{1}{x}} & x > 0 \\ 0 & x \le 0. \end{cases} \end{aligned}$$

This argument can be iterated for the derivatives of higher order. Indeed, by induction one can easily see that the m-th derivative of ψ is given by

$$\psi^{(m)} : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} \dfrac{p_m(x)}{x^{2m}}\, e^{-\frac{1}{x}} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0, \end{cases}$$

where $p_m(x)$ is a polynomial of degree $m-1$. In fact, by induction we find that for $x > 0$

$$\begin{aligned} \psi^{(m+1)}(x) &= \frac{x^{2m}\, p_m'(x) - 2m\, x^{2m-1}\, p_m(x)}{x^{4m}}\, e^{-\frac{1}{x}} + \frac{p_m(x)}{x^{2m}}\, \frac{1}{x^2}\, e^{-\frac{1}{x}} \\ &= \frac{1}{x^{2m+2}} \left(p_m'(x)\, x^2 - 2m\, x\, p_m(x) + p_m(x) \right) e^{-\frac{1}{x}}, \end{aligned}$$

and the degree of the polynomial $p_m'(x)\, x^2 - 2m\, x\, p_m(x) + p_m(x)$ is equal to the degree of $2m\, x\, p_m(x)$, i.e., m.

All the derivatives are continuous at $x = 0$, since for all $m \in \mathbb{N}$ we have

$$\lim_{x \to 0^+} \frac{e^{-\frac{1}{x}}}{x^m} = 0.$$

This proves that, for any $m \in \mathbb{N}$, the derivative $\psi^{(m)}$ exists and is continuous at $x = 0$. Thus, $\psi^{(m)} \in C(\mathbb{R}; \mathbb{R})$ for any $m \in \mathbb{N}$, which implies that $\psi \in C^\infty(\mathbb{R}; \mathbb{R})$. We now show that, using this function ψ and any $r > 0$, we can construct $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$ such that

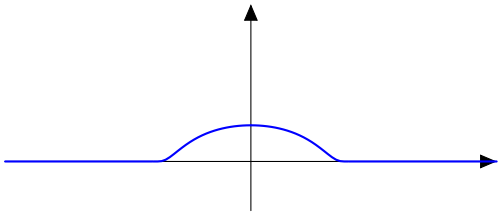
- $\varphi(\boldsymbol{x}) > 0$ for $\|\boldsymbol{x}\| < r$,
- $\varphi(\boldsymbol{x}) = 0$ for $\|\boldsymbol{x}\| \geq r$,
- $\displaystyle\int_{\mathbb{R}^n} \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 1.$

Indeed, we can define a function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ in this way:

$$\varphi(\boldsymbol{x}) = \kappa\, \psi(r^2 - \|\boldsymbol{x}\|^2) = \begin{cases} \kappa\, e^{-\frac{1}{r^2 - \|\boldsymbol{x}\|^2}} & \|\boldsymbol{x}\| < r \\ 0 & \|\boldsymbol{x}\| \geq r \end{cases}$$

for $\kappa > 0$. For example, for $n = 1$, $\kappa = 1$ and $r = 1$ we have

$$\mathbb{R} \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} e^{-\frac{1}{1-x^2}} & |x| < 1 \\ 0 & |x| \geq 1. \end{cases}$$



This function is in $C^\infty(\mathbb{R}^n; \mathbb{R})$ because it is the composition of infinitely differentiable functions. Moreover, from the properties of the function ψ, we have

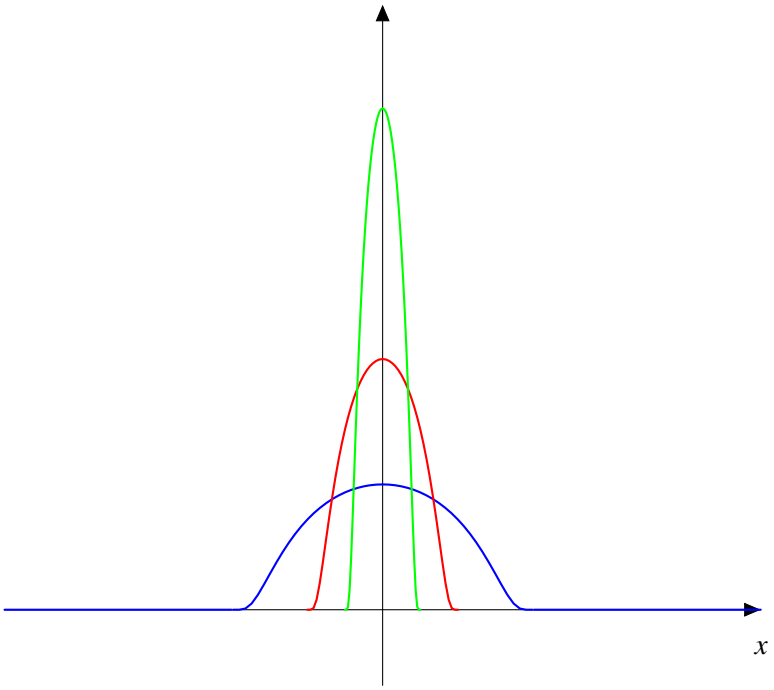
$$\begin{aligned} \varphi(\boldsymbol{x}) > 0 \quad &\Leftrightarrow \quad r^2 - \|\boldsymbol{x}\|^2 > 0 \quad \Leftrightarrow \quad \|\boldsymbol{x}\| < r \\ \varphi(\boldsymbol{x}) = 0 \quad &\Leftrightarrow \quad \|\boldsymbol{x}\|^2 \geq r^2. \end{aligned}$$

Moreover, it is also possible to choose $k > 0$ such that

$$\int_{\mathbb{R}^n} \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \kappa \int_{\|\boldsymbol{x}\| < r} \psi(r^2 - \|\boldsymbol{x}\|^2)\, \mathrm{d}\boldsymbol{x} = 1$$

by choosing $\kappa = \dfrac{1}{\displaystyle\int_{\|\boldsymbol{x}\|<r} \psi(r^2 - \|\boldsymbol{x}\|^2)\,\mathrm{d}\boldsymbol{x}}$: notice that $\kappa > 0$ because the integrand is non-negative.

The figure below shows the function φ obtained with different values of κ and r.



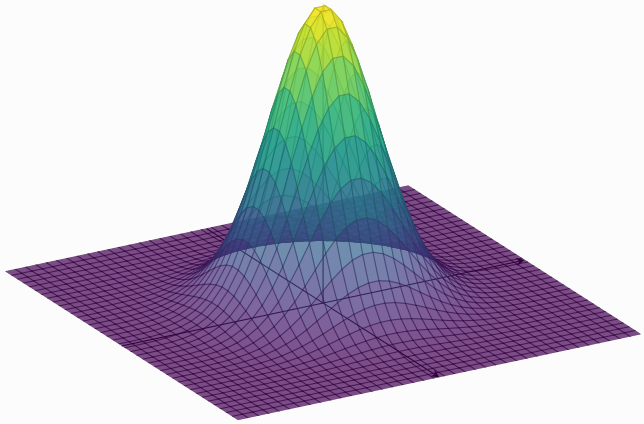
Functions of this type are often referred to also as *bump functions*.

Example *5.4*

The function

$$\begin{aligned} \varphi : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\mapsto \exp\left(-\frac{1}{1 - (x^2 + y^2)}\right) \chi_{B_1(0)}(x, y) \end{aligned}$$

is an element of $\mathcal{D}(\mathbb{R}^2; \mathbb{R})$, i.e., it is a test function on $\mathbb{R}^2$. The support of φ is the closed unit disc $\overline{B_1(0)}$. Moreover, φ is non-zero on $B_1(0)$ and it is equal to zero in $\mathbb{R}^2 \setminus \overline{B_1(0)}$, see figure below:



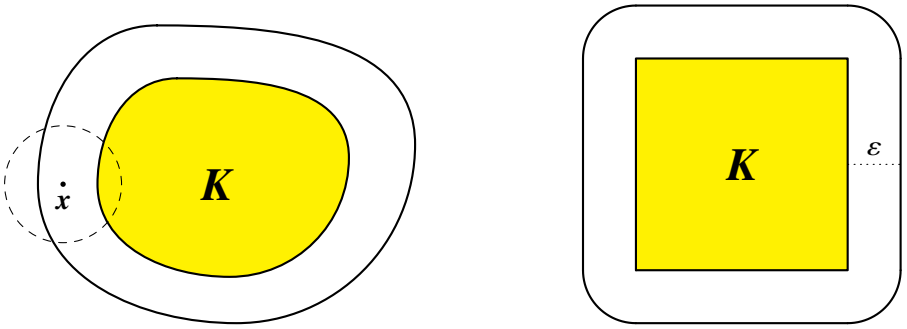
Exercises

Exercise 5.2 Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Show that $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ is a vector space over $\mathbb{C}$. [Solution: (5.2)]

Definition 5.4 *Let $\boldsymbol{K} \subseteq \mathbb{R}^n$ be compact. Let $\varepsilon > 0$. We define*

$$\boldsymbol{K}_\varepsilon \stackrel{def}{=} \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \overline{\mathcal{B}_\varepsilon(\boldsymbol{x})} \cap \boldsymbol{K} \neq \varnothing\}.$$

The set $\boldsymbol{K}_\varepsilon$ is the set of points of $\mathbb{R}^n$ whose distance from $\boldsymbol{K}$ is smaller than ε.



Theorem 5.2 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $\boldsymbol{K} \subseteq \mathbb{R}^n$ be compact. Then:*

(1) $\forall\, \varepsilon > 0, \quad \boldsymbol{K}_\varepsilon$ *is compact;*

(2) $\forall\, \varepsilon_1, \varepsilon_2 > 0, \qquad \varepsilon_1 < \varepsilon_2 \;\Rightarrow \boldsymbol{K}_{\varepsilon_1} \subseteq \boldsymbol{K}_{\varepsilon_2}$;

(3) *if* $\boldsymbol{K} \subseteq \boldsymbol{A}$, *then* $\exists\, \varepsilon > 0 : \quad \boldsymbol{K}_\varepsilon \subseteq \boldsymbol{A}$.

Proof **(1)** Let $\varepsilon > 0$. We show that $\boldsymbol{K}_\varepsilon$ is bounded and closed. We first show that it is bounded. Let $\boldsymbol{x} \in \boldsymbol{K}_\varepsilon$. Then, by definition of $\boldsymbol{K}_\varepsilon$ there exists $\boldsymbol{y} \in \overline{\mathcal{B}_\varepsilon(\boldsymbol{x})} \cap \boldsymbol{K}$. Thus, $\|\boldsymbol{x} - \boldsymbol{y}\| \leq \varepsilon$. Thus,

$$\|\boldsymbol{x}\| = \|(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{y}\| \leq \|\boldsymbol{y}\| + \|\boldsymbol{x} - \boldsymbol{y}\| \leq \max_{\boldsymbol{y} \in \boldsymbol{K}} \|\boldsymbol{y}\| + \varepsilon,$$

and the maximum exists since K is bounded (being compact itself). It follows that $\boldsymbol{K}_\varepsilon$ is bounded. We show that $\boldsymbol{K}_\varepsilon$ is closed: consider a sequence $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ of elements of $\boldsymbol{K}_\varepsilon$ which converges to $\hat{\boldsymbol{x}}$, and we show that $\hat{\boldsymbol{x}} \in \boldsymbol{K}_\varepsilon$. Since for every $k \in \mathbb{N}$ we have $\boldsymbol{x}_k \in \boldsymbol{K}_\varepsilon$, for every $k \in \mathbb{N}$ there exists $\boldsymbol{y}_k \in \overline{\mathcal{B}_\varepsilon(\boldsymbol{x}_k)} \cap \boldsymbol{K}$, and so for every $k \in \mathbb{N}$ there exists $\boldsymbol{y}_k \in \boldsymbol{K}$ such that $\|\boldsymbol{x}_k - \boldsymbol{y}_k\| \leq \varepsilon$, from the definition of $\boldsymbol{K}_\varepsilon$. We have constructed a sequence $(\boldsymbol{y}_k)_{k \in \mathbb{N}}$ of elements of $\boldsymbol{K}$. Since $\boldsymbol{K}$ is compact, there exists a subsequence $(\boldsymbol{y}_{k_\ell})_{\ell \in \mathbb{N}}$ which converges to a point of $\boldsymbol{K}$, say to $\hat{\boldsymbol{y}} \in \boldsymbol{K}$. Thus,

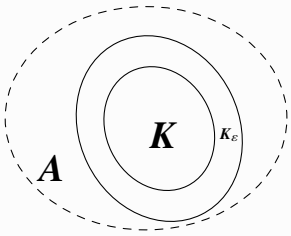
$$\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\| = \lim_{\ell \to +\infty} \|\boldsymbol{x}_{k_\ell} - \boldsymbol{y}_{k_\ell}\| \leq \varepsilon,$$

because the norms are all $\leq \varepsilon$. We conclude that $\hat{\boldsymbol{y}} \in \boldsymbol{K} \cap \overline{\mathcal{B}_\varepsilon(\hat{\boldsymbol{x}})}$, and therefore $\hat{\boldsymbol{x}} \in \boldsymbol{K}_\varepsilon$.

(2) Let $\varepsilon_1, \varepsilon_2 > 0$ be such that $\varepsilon_1 < \varepsilon_2$. We must prove that $\boldsymbol{K}_{\varepsilon_1} \subseteq \boldsymbol{K}_{\varepsilon_2}$. This is a direct consequence of the fact that $\mathcal{B}_{\varepsilon_1}(\boldsymbol{x}) \subseteq \mathcal{B}_{\varepsilon_2}(\boldsymbol{x})$, so that

$$\overline{\mathcal{B}_{\varepsilon_1}(\boldsymbol{x})} \cap \boldsymbol{K} \subseteq \overline{\mathcal{B}_{\varepsilon_2}(\boldsymbol{x})} \cap \boldsymbol{K}.$$

(3) Let $\boldsymbol{A} \subseteq \boldsymbol{K}$ be open. We have to prove that there exists $\varepsilon > 0$ such that $\boldsymbol{K}_\varepsilon \subseteq \boldsymbol{A}$. By contradiction, suppose such ε does not exist. This is equivalent to saying that for all $\varepsilon > 0$, we have $\boldsymbol{K}_\varepsilon \not\subseteq \boldsymbol{A}$, which means that for all $\varepsilon > 0$ there exists $\boldsymbol{x}_\varepsilon \in \boldsymbol{K}_\varepsilon \setminus \boldsymbol{A}$. Let us consider a sequence $(\varepsilon_k)_{k \in \mathbb{N}^*}$ which tends to 0 as $k \to \infty$. For example, we could take $\varepsilon_k = 1/k$ for all $k \in \mathbb{N}^*$.



From what we say before, for every $k \in \mathbb{N}^*$ there exists $\boldsymbol{x}_k \in \boldsymbol{K}_{\varepsilon_k} \setminus \boldsymbol{A}$. We are constructing in this way a sequence of points of $\boldsymbol{K}_{\varepsilon_k} \setminus \boldsymbol{A}$. We show that the limit of this sequence lies in $\boldsymbol{K}$. Indeed, since $\boldsymbol{x}_k \in \boldsymbol{K}_{\varepsilon_k}$, there exists $\boldsymbol{y}_k \in \overline{\mathcal{B}_{\varepsilon_k}(\boldsymbol{x})} \cap \boldsymbol{K}$, so that $\boldsymbol{y}_k \in \boldsymbol{K}$ and $\|\boldsymbol{x}_k - \boldsymbol{y}_k\| \leq \varepsilon_k$. Since $\boldsymbol{K}$ is compact, there exists a subsequence $(\boldsymbol{y}_{k_\ell})_{\ell \in \mathbb{N}}$ of elements of $\boldsymbol{K}$ which converges to an element of $\boldsymbol{K}$, say $\hat{\boldsymbol{y}} \in \boldsymbol{K}$. We obtain

$$\|\boldsymbol{x}_{k_\ell} - \hat{\boldsymbol{y}}\| = \|(\boldsymbol{x}_{k_\ell} - \boldsymbol{y}_{k_\ell}) + (\boldsymbol{y}_{k_\ell} - \hat{\boldsymbol{y}})\| \leq \|\boldsymbol{x}_{k_\ell} - \boldsymbol{y}_{k_\ell}\| + \|\boldsymbol{y}_{k_\ell} - \hat{\boldsymbol{y}}\| \leq \varepsilon_{k_\ell} + \|\boldsymbol{y}_{k_\ell} - \hat{\boldsymbol{y}}\|$$

which goes to 0 as $\ell \to +\infty$. On the other hand, $\boldsymbol{x}_{k_\ell} \notin \boldsymbol{A}$. To summarize what we found, $(\boldsymbol{x}_{k_\ell})_{\ell \in \mathbb{N}}$ converges to $\hat{\boldsymbol{y}} \in \boldsymbol{K} \subseteq \boldsymbol{A}$, but $\boldsymbol{x}_{k_\ell} \notin \boldsymbol{A}$ for every ℓ. This is a contradiction: by definition of limit, there must exist $p \in \mathbb{N}$ such that for every $\ell > p$ we have $\boldsymbol{x}_{k_\ell} \in \boldsymbol{A}$ if the sequence converges to a point of $\boldsymbol{A}$.

Theorem 5.3 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $\boldsymbol{K} \subseteq \boldsymbol{A}$ be compact. Then, there exists $f \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ such that $f(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{K}$.*

Proof From the third point of Theorem 5.2, we know that there exists $\varepsilon > 0$ such that $\boldsymbol{K}_\varepsilon \subseteq \boldsymbol{A}$. We have proved that there exists $\omega \in C^\infty(\mathbb{R}^n; \mathbb{R})$ such that

- $\omega(\boldsymbol{x}) > 0$ for $\|\boldsymbol{x}\| < \frac{\varepsilon}{2}$
- $\omega(\boldsymbol{x}) = 0$ for $\|\boldsymbol{x}\| \geq \frac{\varepsilon}{2}$

- $\int_{\mathbb{R}^n} \omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1$.

The function ω is summable over compact sets (which are Lebesgue measurable), so that for all $\boldsymbol{x} \in \boldsymbol{A}$ we can define

$$f(\boldsymbol{x}) = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \omega(\boldsymbol{x} - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

Indeed, the integral exists, as we can easily see by writing

$$|f(\boldsymbol{x})| \leq \int_{\mathbb{R}^n} |\omega(\boldsymbol{x} - \boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n} |\omega(\boldsymbol{w})|\,\mathrm{d}\boldsymbol{w} = \int_{\mathbb{R}^n} \omega(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w} = 1$$

using the change of variable $\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{w}$, and the absolute value of the Jacobian is 1. Indeed, recall that, given $\boldsymbol{w} = S(\boldsymbol{y})$,

$$\int_{\mathbb{R}^n} g(\boldsymbol{x} - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n} g(S(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} = \int_{S(\mathbb{R}^n)} g(S(S^{-1}(\boldsymbol{w})))\,|\det J_S^{-1}(\boldsymbol{w})|\,\mathrm{d}\boldsymbol{w} = \int_{\mathbb{R}^n} g(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w},$$

because in this case $J_S^{-1}(\boldsymbol{w})$ is $-I$, and its determinant is ± 1.
We show that the function f thus constructed satisfies the requirements of the theorem. First, we show that $f \in C^\infty(\boldsymbol{A}; \mathbb{C})$. We begin by showing that f is continuous. Consider a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of $\boldsymbol{A}$ which converges to $\hat{\boldsymbol{x}} \in \boldsymbol{A}$. To prove that f is continuous, we need to prove that

$$f(\boldsymbol{x}_k) = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \omega(\boldsymbol{x}_k - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

converges to $f(\hat{\boldsymbol{x}})$. To this end, we use Lebesgue's dominated convergence theorem (Theorem 3.47). Since ω is of class C^∞, it is continuous, and for every $\boldsymbol{y} \in \boldsymbol{g}\boldsymbol{K}_{\frac{\varepsilon}{2}}$, we have

$$\omega(\boldsymbol{x}_k - \boldsymbol{y}) \overset{k\to+\infty}{\longrightarrow} \omega(\hat{\boldsymbol{x}} - \boldsymbol{y}).$$

Now, if we can interchange the limit with the integral in the following chain of equalities

$$\lim_{k\to\infty} f(\boldsymbol{x}_k) = \lim_{k\to\infty} \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \omega(\boldsymbol{x}_k - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \lim_{k\to\infty} \omega(\boldsymbol{x}_k - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \omega(\hat{\boldsymbol{x}} - \boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = f(\hat{\boldsymbol{x}}) \tag{5.8}$$

we have proved the statement (and $\omega(\hat{\boldsymbol{x}} - \boldsymbol{y})$ is summable over $\boldsymbol{K}_{\frac{\varepsilon}{2}}$). However, to apply Lebesgue's dominated convergence theorem (Theorem 3.47) we have to bound $|\omega(\boldsymbol{x}_k - \boldsymbol{y})|$ with a function that is independent of k and is summable over the compact set $\boldsymbol{K}_{\frac{\varepsilon}{2}}$. We can use the inequality $|\omega(\boldsymbol{x}_k - \boldsymbol{y})| \leq \sup|\omega|$, and remember that constant functions are integrable over compact sets, so that Lebesgue's dominated convergence theorem (Theorem 3.47) can be applied, and it gives indeed the chain of equalities in (5.8).
Now we have to verify that the first-order partial derivatives exist, i.e., that for every $j \in \{1, \dots, n\}$, denoting by $\boldsymbol{e}_j$ the j-th vector of the standard basis of $\mathbb{R}^n$, the limit of the difference quotient

$$\lim_{t\to 0} \frac{f(\boldsymbol{x} + t\,\boldsymbol{e}_j) - f(\boldsymbol{x})}{t} = \lim_{t\to 0} \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \frac{\omega(\boldsymbol{x} + t\,\boldsymbol{e}_j - \boldsymbol{y}) - \omega(\boldsymbol{x} - \boldsymbol{y})}{t}\,\mathrm{d}\boldsymbol{y}$$

exists and it is finite: this limit will be then equal to the j-th partial derivative of f, that we denote by $\partial_j f(\boldsymbol{x})$. The limit as $t \to 0$ of the integrand exists, and it is equal to $\partial_j \omega(\boldsymbol{x} - \boldsymbol{y})$ (remember that ω is of class C^∞). If we can bound the integrand with a summable function over $\boldsymbol{K}_{\frac{\varepsilon}{2}}$, we can conclude that $\partial_j \omega(\boldsymbol{x} - \boldsymbol{y})$ is summable

over $\boldsymbol{K}_{\frac{\varepsilon}{2}}$ and we can interchange the limit with the integral to obtain

$$\begin{aligned}\partial_j f(\boldsymbol{x}) &= \lim_{t\to 0}\frac{f(\boldsymbol{x}+t\,\boldsymbol{e}_j)-f(\boldsymbol{x})}{t} = \lim_{t\to 0}\int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\frac{\omega(\boldsymbol{x}+t\,\boldsymbol{e}_j-\boldsymbol{y})-\omega(\boldsymbol{x}-\boldsymbol{y})}{t}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\lim_{t\to 0}\frac{\omega(\boldsymbol{x}+t\,\boldsymbol{e}_j-\boldsymbol{y})-\omega(\boldsymbol{x}-\boldsymbol{y})}{t}\,\mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\partial_j\,\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}. \end{aligned} \tag{5.9}$$

From the mean value theorem, there exists $\theta\in(0,1)$ such that

$$\partial_j\,\omega(\boldsymbol{x}+\theta\,t\,\boldsymbol{e}_j-\boldsymbol{y}) = \frac{\omega(\boldsymbol{x}+t\,\boldsymbol{e}_j-\boldsymbol{y})-\omega(\boldsymbol{x}-\boldsymbol{y})}{t},$$

so that

$$\left|\frac{\omega(\boldsymbol{x}+t\,\boldsymbol{e}_j-\boldsymbol{y})-\omega(\boldsymbol{x}-\boldsymbol{y})}{t}\right| = \left|\partial_j\,\omega(\boldsymbol{x}+\theta\,t\,\boldsymbol{e}_j-\boldsymbol{y})\right| \leq \sup_{\mathbb{R}^n}|\partial_j\,\omega|,$$

where $0<\theta<1$. We can therefore conclude that (5.9) holds, leading to the result. This argument can be iterated for higher order partial derivatives. This shows that the partial derivatives of every order exist.

Now, we have to prove that the function defined for all $\boldsymbol{x}\in\boldsymbol{A}$ by

$$f(\boldsymbol{x}) = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

is equal to 1 for all $\boldsymbol{x}\in\boldsymbol{K}$. Let $\boldsymbol{x}\in\boldsymbol{K}$. We know by definition of ω that

$$\int_{\mathbb{R}^n}\omega(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1,$$

and we have already seen that a change of variable leads to $\displaystyle\int_{\mathbb{R}^n}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = 1$. Since

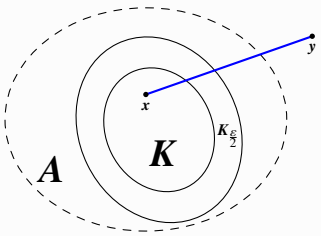
$$\forall\,\boldsymbol{x}\in\boldsymbol{A},\quad 1 = \int_{\mathbb{R}^n}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} + \int_{\mathbb{R}^n\setminus\boldsymbol{K}_{\frac{\varepsilon}{2}}}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y},$$

if we prove that

$$\int_{\mathbb{R}^n\setminus\boldsymbol{K}_{\frac{\varepsilon}{2}}}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = 0 \qquad \forall\,\boldsymbol{x}\in\boldsymbol{K}, \tag{5.10}$$

then we will have succeeded in showing that

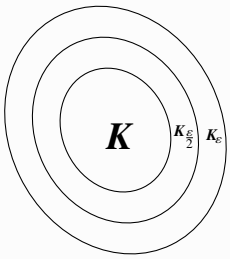
$$f(\boldsymbol{x}) = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}}\omega(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = 1 \qquad \forall\,\boldsymbol{x}\in\boldsymbol{K}.$$



Since $\boldsymbol{x} \in \boldsymbol{K}$, and the $\boldsymbol{y}$ in the integral in (5.10) varies in $\mathbb{R}^n \setminus \boldsymbol{K}_{\frac{\varepsilon}{2}}$, we have $\|\boldsymbol{x} - \boldsymbol{y}\| > \frac{\varepsilon}{2}$, so that $\omega(\boldsymbol{x} - \boldsymbol{y}) = 0$ by definition of ω. The integrand in (5.10) is therefore the identically zero function, which shows that (5.10) holds as required.

Now we have to show that the support of f is compact. We start proving that if $\boldsymbol{x} \notin \boldsymbol{K}_\varepsilon$, we have $f(\boldsymbol{x}) = 0$. We recall that

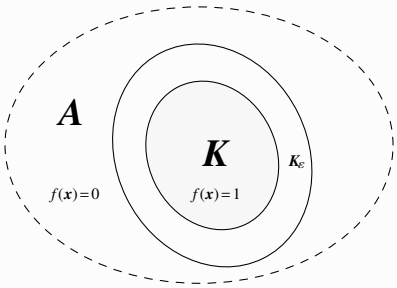
$$f(\boldsymbol{x}) = \int_{\boldsymbol{K}_{\frac{\varepsilon}{2}}} \omega(\boldsymbol{x} - \boldsymbol{y})\, \mathrm{d}\boldsymbol{y}.$$



We need to show that $\omega(\boldsymbol{x} - \boldsymbol{y}) = 0$ if $\boldsymbol{x} \notin \boldsymbol{K}_\varepsilon$ and $\boldsymbol{y} \in \boldsymbol{K}_{\frac{\varepsilon}{2}}$. To this end, it suffices to show that if $\boldsymbol{x} \notin \boldsymbol{K}_\varepsilon$ and $\boldsymbol{y} \in \boldsymbol{K}_{\frac{\varepsilon}{2}}$ we have $\|\boldsymbol{x} - \boldsymbol{y}\| \geq \frac{\varepsilon}{2}$, again by definition of ω. Suppose by contradiction that $\boldsymbol{x} \notin \boldsymbol{K}_\varepsilon$ and $\boldsymbol{y} \in \boldsymbol{K}_{\frac{\varepsilon}{2}}$, but $\|\boldsymbol{x} - \boldsymbol{y}\| < \frac{\varepsilon}{2}$. Since $\boldsymbol{y} \in \boldsymbol{K}_{\frac{\varepsilon}{2}}$, there exists $z \in \boldsymbol{K}$ such that $\|\boldsymbol{y} - z\| \leq \frac{\varepsilon}{2}$. Thus,

$$\|\boldsymbol{x} - z\| = \|(\boldsymbol{x} - \boldsymbol{y}) + (\boldsymbol{y} - z)\| \leq \|\boldsymbol{x} - \boldsymbol{y}\| + \|\boldsymbol{y} - z\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

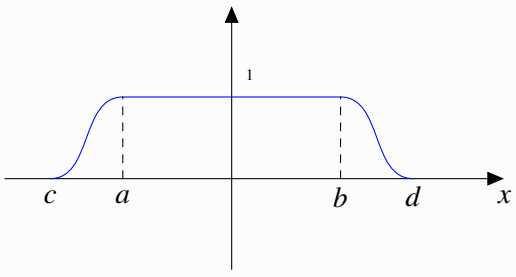
which implies $\boldsymbol{x} \in \boldsymbol{K}_\varepsilon$. This is a contradiction. We have therefore proved that $f(\boldsymbol{x})$ is zero outside $\boldsymbol{K}_\varepsilon$, so that its support is contained in $\boldsymbol{K}_\varepsilon$, i.e., $\mathrm{supp}(f) \subseteq \boldsymbol{K}_\varepsilon$. Since $A \setminus \boldsymbol{K}_\varepsilon$ is open, given $\boldsymbol{x} \in A \setminus \boldsymbol{K}_\varepsilon$ there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{A} \setminus \boldsymbol{K}_\varepsilon$, and for every $\boldsymbol{y} \in \mathcal{B}_r(\boldsymbol{x})$ we have $f(\boldsymbol{y}) = 0$.



We now examine some relevant particular cases of this result.

Example 5.5

Let $a, b, c, d \in \mathbb{R}$ such that $c < a < b < d$. Considering $K = [a, b]$ and $A = (c, d)$, there exists a function $f \in \mathcal{D}((c, d); \mathbb{R})$ which is equal to 1 for all $x \in [a, b]$:

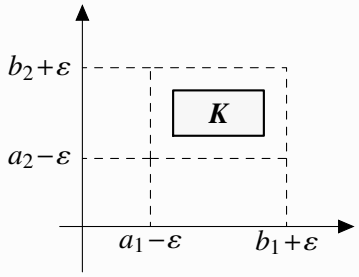


Example 5.6

Let $\boldsymbol{K} = \prod_{j=1}^{n}[a_j, b_j]$, where for all $j \in \{1, \ldots, n\}$ we have $a_j, b_j \in \mathbb{R}$ with $a_j < b_j$. Take $\varepsilon > 0$ and

$$\boldsymbol{A} = \prod_{j=1}^{n}(a_j - \varepsilon, b_j + \varepsilon).$$

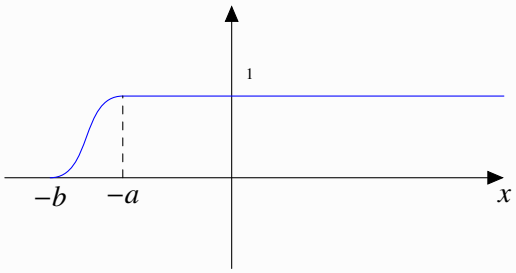
Then, there exists $f \in \mathcal{D}(\boldsymbol{A}; \mathbb{R})$ such that $f(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{K}$.



Example 5.7

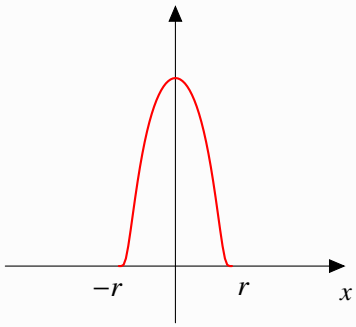
Let $a, b \in \mathbb{R}$ such that $0 < a < b$. We want to construct a function $\psi \in C^\infty(\mathbb{R}; \mathbb{R})$ such that

$$\psi(x) = \begin{cases} 1 & \text{for } x \geq -a \\ 0 & \text{for } x \leq -b \end{cases}$$



Let $r = \dfrac{b-a}{2}$. We have seen that we can construct functions $\varphi \in C^\infty(\mathbb{R}; \mathbb{R})$ such that

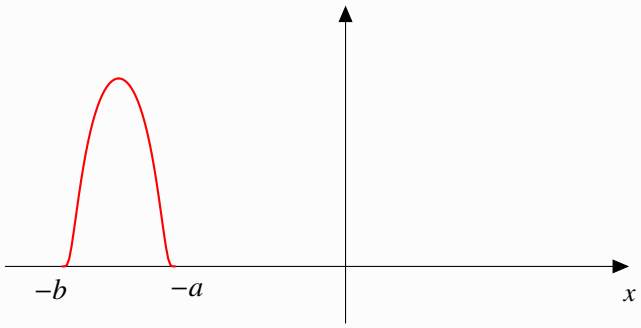
- $\varphi(x) > 0$ for $x \in (-r, r)$;
- $\varphi(x) = 0$ for $|x| > r$;
- $\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x = 1$.



We can shift this function to the left, by defining a new function $\theta(x) = \varphi(x + \alpha)$, where $\alpha = b - r$, so that

- $\theta(x) > 0$ for $-b < x < -a$;

- $\theta(x) = 0$ for $x \notin (-b, -a)$;
- $\int_{\mathbb{R}} \theta(x)\,\mathrm{d}x = 1$.



Now, we define $\psi(x) = \int_{-\infty}^{x} \theta(y)\,\mathrm{d}y$. From the fundamental theorem of Calculus, this function is in C^∞ and $\psi'(x) = \theta(x)$. If $x < -b$, we have $\psi(x) = 0$, and for $x > -a$ we have

$$\psi(x) = \int_{-\infty}^{x} \theta(y)\,\mathrm{d}y = \int_{-\infty}^{+\infty} \theta(y)\,\mathrm{d}y = 1.$$

5.1.1 Density of $\mathcal{D}$ in L^p

Given a normed vector space $(\boldsymbol{X}, \|\cdot\|)$ and $\boldsymbol{W} \subseteq \boldsymbol{X}$, we recall that $\boldsymbol{W}$ is dense in $\boldsymbol{X}$ if $\overline{\boldsymbol{W}} = \boldsymbol{X}$, i.e., for all $\boldsymbol{x} \in \boldsymbol{X}$, there exists a sequence $(\boldsymbol{w}_k)_{k\in\mathbb{N}}$ of elements of $\boldsymbol{W}$ that converges to $\boldsymbol{x}$, i.e., $\lim_{k\to\infty} \boldsymbol{w}_k = \boldsymbol{x}$.

Given an open set $\boldsymbol{A}$ of $\mathbb{R}^n$, we want to study the density of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ in $L^p(\boldsymbol{A};\mathbb{C})$. We need to clarify immediately one point: the set $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ contains functions, while the set $L^p(\boldsymbol{A};\mathbb{C})$ contains equivalence classes, which are sets of functions. Thus, the inclusion $\mathcal{D}(\boldsymbol{A};\mathbb{C}) \subseteq L^p(\boldsymbol{A};\mathbb{C})$ is meaningless. However, if we consider the normed space $(L^p(\boldsymbol{A};\mathbb{C}), \|\cdot\|_p)$, we will show that $\boldsymbol{W} = \{[f] \,|\, f \in \mathcal{D}(\boldsymbol{A};\mathbb{C})\}$ is a subset of $L^p(\boldsymbol{A};\mathbb{C})$, and it is dense in $L^p(\boldsymbol{A};\mathbb{C})$, provided $p \in [1, +\infty)$ (we will show that this result is false if $p = +\infty$).

Before we prove this result, we show that the set of simple functions is dense in $L^p(\boldsymbol{A};\mathbb{C})$.

Theorem 5.4 *The set*

$$S = \left\{ [g] \,\middle|\, g : \mathbb{R}^n \longrightarrow \mathbb{C} \text{ is simple and } \int_{\mathbb{R}^n} |g(\boldsymbol{x})|^p\,\mathrm{d}\boldsymbol{x} < +\infty \right\}$$

is dense in $L^p(\mathbb{R}^n;\mathbb{C})$ for all $1 \le p < +\infty$.

Proof Let $[f] \in L^p(\mathbb{R}^n;\mathbb{C})$. Our objective is to determine a sequence of functions in S that converges to f. We begin assuming that f is real and non-negative, i.e., $f(\boldsymbol{x}) \ge 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Given a non-negative measurable function, we can construct a sequence of simple non-negative functions $(\varphi_k)_{k\in\mathbb{N}}$ such that $\varphi_k(\boldsymbol{x}) \le \varphi_{k+1}(\boldsymbol{x}) \le f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$ and

$$\lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = f(\boldsymbol{x}) \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^n,$$

see Theorem 3.30. We show that the sequence $(\varphi_k)_{k\in\mathbb{N}}$ is a good candidate as a sequence in S that converges

to f. First we have to show that $(\varphi_k)_{k\in\mathbb{N}}$ is indeed a sequence of functions in S. Since for all $k \in \mathbb{N}$ we have

$$\int_{\mathbb{R}^n} |\varphi_k(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \leq \int_{\mathbb{R}^n} |f(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < +\infty,$$

then $[\varphi_k] \in S$ for all $k \in \mathbb{N}$. Now we show that $(\varphi_k)_{k\in\mathbb{N}}$ converges to f in $L^p(\mathbb{R}^n;\mathbb{R})$, i.e., that $\left\|[f] - [\varphi_k]\right\|_p \xrightarrow{k\to\infty} 0$. By definition of p-norm and using the continuity of the power function

$$\left\|[f] - [\varphi_k]\right\|_p = \left(\int_{\mathbb{R}^n} |f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \qquad \Rightarrow \qquad \lim_{k\to+\infty} \left\|[f] - [\varphi_k]\right\|_p = \left(\lim_{k\to+\infty} \int_{\mathbb{R}^n} |f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}}.$$

If we can apply Lebesgue's dominated convergence theorem (Theorem 3.47), we can immediately conclude that the limit is indeed equal to 0. Pointwise, the integrand converges to 0 as $k \to +\infty$. We bound the integrand with a summable function that does not depend on k. From

$$|f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})| \leq |f(\boldsymbol{x})| + |\varphi_k(\boldsymbol{x})| \leq 2\,|f(\boldsymbol{x})|,$$

we find

$$|f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})|^p \leq 2^p\,|f(\boldsymbol{x})|^p,$$

which is integrable in $\mathbb{R}^n$ because, by assumption, $[f] \in L^p(\mathbb{R}^n;\mathbb{C})$, so that

$$\left(\int_{\mathbb{R}^n} |f(\boldsymbol{x}) - \varphi_k(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \xrightarrow{k\to+\infty} 0.$$

Recall that we have assumed that f is real-valued and non-negative. If f is real, but not necessarily non-negative, then we can write it as $f = f_+ - f_-$ where f_+ and f_- are non-negative, and $|f| = f_+ + f_-$. We have

$$f_+^p \leq |f|^p \quad \text{and} \quad f_-^p \leq |f|^p,$$

so that $[f_+] \in L^p(\mathbb{R}^n;\mathbb{R})$ and $[f_-] \in L^p(\mathbb{R}^n;\mathbb{R})$. Since f_+ and f_- are non-negative, there exist two non-decreasing sequences of simple functions $(\varphi_k)_{k\in\mathbb{N}}$ and $(\psi_k)_{k\in\mathbb{N}}$ that converge to $[f_+]$ and $[f_-]$, respectively, and these are in $L^p(\mathbb{R}^n;\mathbb{R})$ as shown before, i.e.,

$$[\varphi_k] \xrightarrow{k\to+\infty} [f_+] \in L^p(\mathbb{R}^n;\mathbb{R}) \qquad \text{and} \qquad [\psi_k] \xrightarrow{k\to+\infty} [f_-] \in L^p(\mathbb{R}^n;\mathbb{R}).$$

We consider the difference

$$[\varphi_k - \psi_k] \xrightarrow{k\to+\infty} [f_+ - f_-] = [f] \in L^p(\mathbb{R}^n;\mathbb{R}).$$

Finally, we can consider the complex case $[f] \in L^p(\mathbb{R}^n;\mathbb{C})$. The functions $\mathfrak{Re}\,f$ and $\mathfrak{Im}\,f$ are now verifying the previous conditions, so that there exist two sequences of simple functions $(\varphi_k)_{k\in\mathbb{N}}$ and $(\psi_k)_{k\in\mathbb{N}}$ that converge to $[\mathfrak{Re}\,f]$ and $[\mathfrak{Im}\,f]$, respectively, and these are in $L^p(\mathbb{R}^n;\mathbb{R})$, i.e.,

$$[\varphi_k] \xrightarrow{k\to+\infty} [\mathfrak{Re}\,f] \in L^p(\mathbb{R}^n;\mathbb{R}) \qquad \text{and} \qquad [\psi_k] \xrightarrow{k\to+\infty} [\mathfrak{Im}\,f] \in L^p(\mathbb{R}^n;\mathbb{R}),$$

so that

$$[\varphi_k + i\,\psi_k] \xrightarrow{k\to+\infty} [\mathfrak{Re}\,f + i\,\mathfrak{Im}\,f] = [f] \in L^p(\mathbb{R}^n;\mathbb{C}).$$

The following result establishes that $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ is dense in $L^p(\boldsymbol{A};\mathbb{C})$. This claim hides an abuse of notation that, from now on, we will adopt consistently: the elements of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ are functions, whereas the elements of $L^p(\boldsymbol{A};\mathbb{C})$ are equivalence classes of functions.

Theorem 5.5 [Density of $\mathcal{D}$ in L^p]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $p \in [1, +\infty)$. Then, $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$ is dense in $L^p(\boldsymbol{A}; \mathbb{C})$.

Partial proof First, we show that $\mathcal{D}(\boldsymbol{A}; \mathbb{C}) \subseteq L^p(\boldsymbol{A}; \mathbb{C})$, in the sense that if $f \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$, then $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$. Let $f \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. We show that for every $p \geq 1$ we have

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}x < +\infty.$$

If $\boldsymbol{x} \notin \operatorname{supp} f$, there exists an open subset $\boldsymbol{B}$ of $\boldsymbol{A}$ such that $f(\boldsymbol{y}) = 0$ for almost every $\boldsymbol{y}$ in $\boldsymbol{B}$. By continuity, $f(\boldsymbol{y}) = 0$ for all $\boldsymbol{y} \in \boldsymbol{B}$. It follows that $f(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \notin \operatorname{supp} f$. Then, using Weierstrass extreme value theorem

$$\int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p \, \mathrm{d}x = \int_{\operatorname{supp} f} |f(\boldsymbol{x})|^p \, \mathrm{d}x \leq \max_{\operatorname{supp} f} |f|^p \, \mu_n(\operatorname{supp} f) < +\infty,$$

because $\operatorname{supp} f$ is compact. We conclude that $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$. Thus, $\{[f] \,|\, f \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})\} \subseteq L^p(\boldsymbol{A}; \mathbb{C})$.

We assume for simplicity that $\boldsymbol{A} = \mathbb{R}^n$. From Theorem 5.4, an equivalence class $[f]$ in $L^p(\mathbb{R}^n; \mathbb{C})$ can be approximated with arbitrary precision with a sequence of simple functions $(f_k)_{k \in \mathbb{N}}$ in $L^p(\mathbb{R}^n; \mathbb{C})$. It is possible to prove that it is not restrictive to assume that these simple functions can be chosen so that their range is constant on a decomposition of their domain into disjoint boxes, i.e., each can be written as

$$f_k(\boldsymbol{x}) = \sum_{j=1}^{m} \alpha_j \chi_{\boldsymbol{B}_j}(\boldsymbol{x}),$$

where $m \in \mathbb{N}^*$, $\alpha_j \in \mathbb{C} \setminus \{0\}$ and $\boldsymbol{B}_j \subseteq \mathbb{R}^n$ are pairwise disjoint compact n-dimensional boxes of $\mathbb{R}^n$, so that we can write for all $j \in \{1, \ldots, m\}$

$$\boldsymbol{B}_j = \prod_{k=1}^{n} [a_k^j, b_k^j], \qquad a_k^j, b_k^j \in \mathbb{R} \text{ and } a_k^j < b_k^j.$$

We now show that each f_k can be approximated with arbitrary precision with a function $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$, i.e., for every $\varepsilon > 0$ we can find $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$ such that $\|[\varphi] - [f_k]\|_p \leq \varepsilon$. Let $\varepsilon > 0$. We construct φ in this way. Let $h > 0$, and consider the boxes

$$S_j = \prod_{k=1}^{n} [a_k^j - h, b_k^j + h].$$

For every $j \in \{1, \ldots, m\}$, from Theorem 5.3 there exists $\varphi_j \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$ satisfying $\varphi_j(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in \boldsymbol{B}_j$ and $\operatorname{supp} \varphi_j \subseteq S_j$, so that $\varphi_j(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \notin S_j$. Then, we define

$$\varphi(\boldsymbol{x}) = \sum_{j=1}^{m} \alpha_j \varphi_j(\boldsymbol{x}).$$

We prove that for every $\eta > 0$, there exists h small enough so that

$$\|[\chi_{\boldsymbol{B}_j}] - [\varphi_j]\|_p \leq \eta.$$

Recalling that for $\boldsymbol{x} \in \boldsymbol{B}_j$ both $\varphi_j(\boldsymbol{x})$ and $\chi_{\boldsymbol{B}_j}(\boldsymbol{x})$ are equal to 1, we have

$$\begin{aligned}\left\|[\chi_{\boldsymbol{B}_j} - \varphi_j]\right\|_p &= \left(\int_{\mathbb{R}^n} \left|\chi_{\boldsymbol{B}_j}(\boldsymbol{x}) - \varphi_j(\boldsymbol{x})\right|^p \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \\ &= \left(\int_{S_j \setminus \boldsymbol{B}_j} \left|\chi_{\boldsymbol{B}_j}(\boldsymbol{x}) - \varphi_j(\boldsymbol{x})\right|^p \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \\ &= \left(\int_{S_j \setminus \boldsymbol{B}_j} \left|\varphi_j(\boldsymbol{x})\right|^p \mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} \leq \left(\int_{S_j \setminus \boldsymbol{B}_j} 1 \,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{p}} = \left(\mu_n(S_j \setminus \boldsymbol{B}_j)\right)^{\frac{1}{p}},\end{aligned}$$

since $|\varphi_j(\boldsymbol{x})| \leq 1$ over $S_j \setminus \boldsymbol{B}_j$, with

$$\mu_n(S_j \setminus \boldsymbol{B}_j) = \prod_{k=1}^{n} [a_k^j - h, b_k^j + h] - \prod_{k=1}^{n} [a_k^j, b_k^j] \xrightarrow{h \to 0} 0.$$

It follows that we can choose h sufficiently small so that

$$\left\|[\chi_{\boldsymbol{B}_j} - \varphi_j]\right\|_p \leq \eta = \frac{\varepsilon}{\sum_{j=1}^{m} |\alpha_j|},$$

and so

$$\begin{aligned}\|[f_k] - [\varphi]\|_p &= \left\|\sum_{j=1}^{m} \alpha_j \left[\chi_{\boldsymbol{B}_j} - \varphi_j\right]\right\|_p \\ &\leq \sum_{j=1}^{m} |\alpha_j| \left\|[\chi_{\boldsymbol{B}_j} - \varphi_j]\right\|_p \leq \eta \sum_{j=1}^{m} |\alpha_j| = \varepsilon.\end{aligned}$$

This allows us to conclude that a function in $L^p(\mathbb{R}^n;\mathbb{C})$ can be approximated with arbitrary precision with a function in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$; in other words, a function in $L^p(\mathbb{R}^n;\mathbb{C})$ can be approximated by a simple function, which is the limit in $L^p(\mathbb{R}^n;\mathbb{C})$ of functions of the type f_k, which, in turn, can be approximated by functions in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, as required.

We observe that the result of Theorem 5.4 is false for $p = +\infty$, for any open set $\boldsymbol{A} \subseteq \mathbb{R}^n$. Consider in fact, for example, the constant function $f = 1$ (i.e., $f(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{A}$), and consider any function $v \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, which has compact support. The Lebesgue measure of the set $\boldsymbol{B} = \boldsymbol{A} \setminus \operatorname{supp} v$ is positive, so that

$$\begin{aligned}\|[v] - [f]\|_{\infty,\boldsymbol{A}} &= \min\{M \geq 0 \mid |v(\boldsymbol{x}) - f(\boldsymbol{x})| \leq M \text{ a.e. in } \boldsymbol{A}\} \\ &\geq \min\{M \geq 0 \mid |v(\boldsymbol{x}) - f(\boldsymbol{x})| \leq M \text{ a.e. in } \boldsymbol{A} \setminus \operatorname{supp} f\} = \|[f]\|_{\infty,\boldsymbol{B}} = 1.\end{aligned}$$

Hence, $\|[v] - [f]\|_{\infty,\boldsymbol{A}} \geq 1$, and this implies that $\|[v] - [f]\|_{\infty,\boldsymbol{A}}$ cannot be rendered arbitrarily small.

Exercises

Exercise 5.3 Let $0 < a < b$. Show that the function $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$\varphi(x) = \begin{cases} \exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right) & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

is in $\mathcal{D}(\mathbb{R};\mathbb{R})$. Define the function

$$\Phi(x) = \frac{\int_x^b \varphi(t)\,\mathrm{d}t}{\int_a^b \varphi(t)\,\mathrm{d}t}.$$

After observing that

$$\Phi(x) = \begin{cases} 1 & \text{if } x \le a \\ 0 & \text{if } x \ge b, \end{cases}$$

show that the function $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}$ given by $\psi(x_1, x_2, \ldots, x_n) = \Phi(x_1^2 + x_2^2 + \ldots + x_n^2)$ is in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. [Solution: (5.3)]

5.2 Locally summable functions

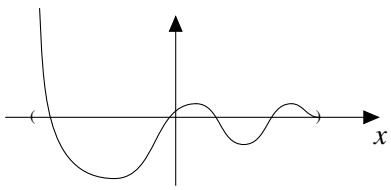
We now give the definition of locally summable function. Loosely, a locally summable function (also called locally integrable function) is a function which is summable (i.e., its integral is finite) on every compact subset of its domain. This concept is particularly important, because we will see that distributions are a generalization of this concept. Indeed, distributions are not precisely an extension of a function but, more precisely, a generalization of locally summable function.

Definition 5.5 [Locally Summable Functions]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be an open set. Let the complex-valued function f be defined almost everywhere in $\boldsymbol{A}$ be Lebesgue measurable. Then, $[f]$ is said to be locally summable *or* locally integrable *if it is summable on every compact set contained in $\boldsymbol{A}$.*

The set of all locally summable (complex-valued) functions defined a.e. in A is denoted by $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$:

$$L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C}) = \{[f] \,|\, f \text{ is defined a.e. on } \boldsymbol{A} \text{ and is summable on every compact subset of } \boldsymbol{A}\}.$$

As usual, with some abuse of notation we will consider $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ as a space of functions, even if, rigorously speaking, its elements are equivalent classes of functions. Loosely speaking, an element of $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ is a function which is "not too large" in the interior of $\boldsymbol{A}$, while near the border (or at infinity if $\boldsymbol{A}$ is unbounded) this function can have an arbitrary behaviour.



Example 5.8 We show that all polynomial functions belong to $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Consider a closed interval $K = [a,b]$ in $\mathbb{R}$, with

$a < b$, and coefficients $a_0, \ldots, a_n \in \mathbb{R}$. We have

$$\int_a^b (a_n\, x^n + a_{n-1}\, x^{n-1} + \ldots + a_1\, x + a_0)\, \mathrm{d}x = \left[a_n \frac{x^{n+1}}{n+1} + a_{n-1} \frac{x^n}{n} + \ldots + a_1 \frac{x^2}{2} + a_0\, x \right]_a^b$$
$$= a_n \frac{b^{n+1} - a^{n+1}}{n+1} + a_{n-1} \frac{b^n - a^n}{n} + \ldots + a_1 \frac{b^2 - a^2}{2} + a_0\,(b-a) < +\infty.$$

Example *5.9*

Consider the function $f(x) = \frac{1}{x}$, defined almost everywhere on $\mathbb{R}$. It is easy to see that $f \notin L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Indeed, taking for example the closed interval $[0,1]$, we have

$$\int_0^1 \frac{1}{x}\, \mathrm{d}x,$$

which is not finite, because of the "strong" singularity at the origin. However, it is easy to see that $f \in L^1_{\mathrm{loc}}(\mathbb{R} \setminus \{0\};\mathbb{R})$. In fact, consider a compact set contained in $\mathbb{R} \setminus \{0\}$ in the form $[-b,-\varepsilon]$ or $[\varepsilon, b]$, where $\varepsilon > 0$. If for example we take $K = [\varepsilon, b]$, we find

$$\int_\varepsilon^b \frac{1}{x}\, \mathrm{d}x = \ln b - \ln \varepsilon < +\infty.$$

Example *5.10*

Consider the function $f(x) = \frac{1}{\sqrt{x}}$, defined almost everywhere on $\mathbb{R}_+ = [0, +\infty)$. It is easy to see that $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R})$. Indeed, taking a generic closed interval $[a,b]$, with $a < b$, we have

$$\int_a^b \frac{1}{\sqrt{x}}\, \mathrm{d}x = [2\,\sqrt{x}]_a^b = 2\,(\sqrt{b} - \sqrt{a}).$$

Theorem 5.6 *Let $A \subseteq \mathbb{R}^n$ be an open set. Let $f : A \longrightarrow \mathbb{C}$ be continuous. Then, $f \in L^1_{\mathrm{loc}}(A;\mathbb{C})$.*

Proof Let $\boldsymbol{K}$ be a compact in $\boldsymbol{A}$. In view of Weierstrass theorem, there exists $M > 0$ such that $f(\boldsymbol{x}) \leq |f(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in \boldsymbol{K}$. Moreover, the constant function equal to M is summable over compact sets because compact sets have finite Lebesgue measure, so that $\int_{\boldsymbol{K}} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} < +\infty$. Thus, f is summable over $\boldsymbol{K}$, which implies $f \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$.

Remark 5.3 Notice that in Theorem 5.6 we are not assuming that f is defined almost everywhere in $\boldsymbol{A}$. We can see easily that in that case the result is false: the function $f(x) = \frac{1}{x}$ defined almost everywhere in $\mathbb{R}$ is continuous, but it is not in $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$.

In view of Theorem 5.6, we can write

$$C^0(\boldsymbol{A};\mathbb{C}) \subseteq L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C}).$$

We now show that also the spaces L^p are contained in L^1_{loc}. We begin introducing the following lemma, which will be used also in the sequel.

Lemma 5.3 *Let $p \in \mathbb{R}$ be such that $p \geq 1$. Then,*

$$\forall\, x \geq 0,\ \forall\, \eta > 0, \qquad x \leq \eta + \eta^{1-p}\, x^p.$$

Proof If $x \leq \eta$, the inequality is obvious. Consider the case $x > \eta$. We have $x = x^{1-p}\, x^p$ with $1 - p \leq 0$ (since it is assumed that $p \geq 1$). Since x^α with $\alpha < 0$ and $x > 0$ is a decreasing function, we have

$$x > \eta \quad \Rightarrow \quad x^{1-p} < \eta^{1-p}.$$

It follows that

$$x = x^{1-p}\, x^p \leq \eta^{1-p}\, x^p < \eta + \eta^{1-p}\, x^p.$$

Hence, in both cases $x \leq \eta$ and $x > \eta$, we have $x \leq \eta + \eta^{1-p}\, x^p$.

Theorem 5.7 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be an open set, and let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. Let $[f] \in L^p(\boldsymbol{A}; \mathbb{C})$ with $1 \leq p \leq +\infty$. Then, $[f] \in L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$.*

Proof First, consider the case $p = \infty$. Let $f : \boldsymbol{A} \setminus \boldsymbol{N} \longrightarrow \mathbb{C}$, where $\boldsymbol{N}$ has zero Lebesgue measure. We recall that $[f] \in L^\infty(\boldsymbol{A}; \mathbb{C})$ means that f is essentially bounded, i.e., that there exists $M \geq 0$ such that $|f(\boldsymbol{x})| \leq M$ for almost every $\boldsymbol{x}$ in $\boldsymbol{A}$. Let $\boldsymbol{N}_1 = \{\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{N} \,|\, |f(\boldsymbol{x})| > M\}$, which obviously has zero Lebesgue measure. Let K be a compact subset of $\boldsymbol{A}$. Then

$$\int_{\boldsymbol{K}} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{K} \cap \boldsymbol{N}_1} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{K} \setminus \boldsymbol{N}_1} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{K} \setminus \boldsymbol{N}_1} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} \leq M\, \mu_n(\boldsymbol{K}) < +\infty,$$

because $\boldsymbol{K} \cap \boldsymbol{N}_1$ also has zero Lebesgue measure, and for the same reason $\mu_n(\boldsymbol{K}) = \mu_n(\boldsymbol{K} \cap \boldsymbol{N}_1)$. We deduce that $[f] \in L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$.

Consider the case $1 \leq p < \infty$. We use the result in Lemma 5.3 with $\eta = 1$ and $|f(\boldsymbol{x})|$ in place of x, and we can write $|f(\boldsymbol{x})| \leq 1 + |f(\boldsymbol{x})|^p$. Thus,

$$\int_K |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{K}} (1 + |f(\boldsymbol{x})|^p)\, \mathrm{d}\boldsymbol{x} \leq \mu_n(\boldsymbol{K}) + \int_{\boldsymbol{A}} |f(\boldsymbol{x})|^p\, \mathrm{d}\boldsymbol{x} < +\infty.$$

As a result of the latter, we can write

$$L^p(\boldsymbol{A}; \mathbb{C}) \subseteq L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$$

for all $1 \leq p \leq +\infty$.

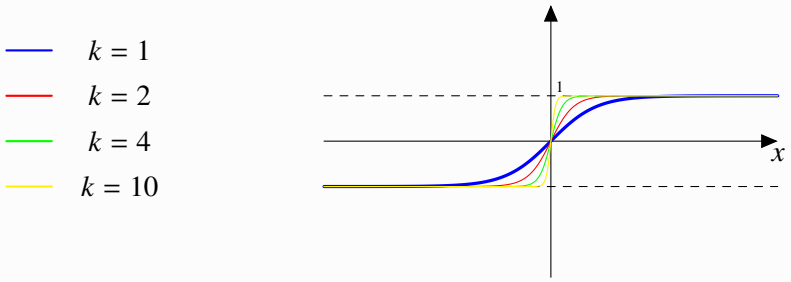
We now introduce a notion of convergence in $L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$.

Definition 5.6 [Convergence in L^1_{loc}]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $[f] \in L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$. Let $([f]_k)_{k \in \mathbb{N}}$ be a sequence of equivalence classes in $L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$. We say that $([f_k])_{k \in \mathbb{N}}$ converges to $[f]$ in $L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$ if for every compact set K contained in $\boldsymbol{A}$ we have

$$\lim_{k \to \infty} \int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = 0.$$

Example 5.11

Let $A = \mathbb{R}$, and let $f_k(x) = \tanh(k\,x)$ for $x \in \mathbb{R}$. Recall that $\tanh y = \frac{\sinh y}{\cosh y} = \frac{e^y - e^{-y}}{e^y + e^{-y}}$. For all $k \in \mathbb{N}$ the function f_k is in $L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$ because it is continuous.



We show that the sequence $(f_k)_{k\in\mathbb{N}}$ converges to the sign function in $L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$, which amounts to showing that for any compact in $\mathbb{R}$ we have

$$\lim_{k\to\infty} \int_K |\tanh(k\,x) - \operatorname{sgn}(x)|\,\mathrm{d}x = 0.$$

To this end, we use Lebesgue's dominated convergence theorem (Theorem 3.47). First, we observe that the sequence $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the sign function since

$$\lim_{k\to\infty} \tanh(k\,x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

It follows that $(f_k - \operatorname{sgn})_{k\in\mathbb{N}}$ is pointwise convergent to the zero function. Moreover, the sign function is in $L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$ because it is measurable and bounded. Moreover, for all $k \in \mathbb{N}$ we have

$$|\tanh(k\,x) - \operatorname{sgn}(x)| \le |\tanh(k\,x)| + 1 \le 2,$$

which is a constant, and therefore it is summable over K. Lebesgue's dominated convergence theorem (Theorem 3.47) ensures that

$$\lim_{k\to\infty} \int_K |\tanh(k\,x) - \operatorname{sgn}(x)|\,\mathrm{d}x = \int_K \lim_{k\to\infty} |\tanh(k\,x) - \operatorname{sgn}(x)|\,\mathrm{d}x = 0.$$

Example 5.12

Consider the sequence of function $(f_k)_{k\in\mathbb{N}}$ defined by

$$f_k(t) = \frac{2}{\sqrt{\pi}} \int_0^{k\,t} e^{-\tau^2}\,\mathrm{d}\tau$$

(also called *error function*). We show that $(f_k)_{k\in\mathbb{N}}$ converges to the sign function in $L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$. From the well-known integral

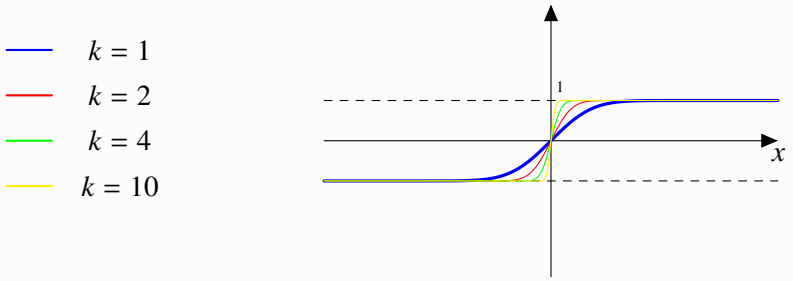
$$\int_{\mathbb{R}} e^{-\tau^2}\,\mathrm{d}\tau = \sqrt{\pi}$$

we find that

$$\lim_{k\to\infty} \frac{2}{\sqrt{\pi}} \int_0^{k\,t} e^{-\tau^2}\,\mathrm{d}\tau = \begin{cases} 1 & t > 0 \\ 0 & t = 0 \\ -1 & t < 0. \end{cases}$$

The result follows by applying Lebesgue's dominated convergence theorem (Theorem 3.47), after noting that

$$\left|\frac{2}{\sqrt{\pi}}\int_0^{kt} e^{-\tau^2}\,\mathrm{d}\tau - \operatorname{sgn} t\right| \le \frac{2}{\sqrt{\pi}}\sup\Big(\int_0^{kt} e^{-\tau^2}\,\mathrm{d}\tau\Big) + 1 = \frac{2}{\sqrt{\pi}}\,\frac{\sqrt{\pi}}{2} + 1 = 2.$$



We now show that the convergence in L^p implies the convergence in L^1_{loc}.

Theorem 5.8 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be an open set. Let $p \in [1, +\infty)$. Let $([f]_k)_{k\in\mathbb{N}}$ be a sequence in $L^p(\boldsymbol{A};\mathbb{C})$. Let $[f] \in L^p(\boldsymbol{A};\mathbb{C})$. Let $([f_k])_{k\in\mathbb{N}}$ converge to $[f]$ in $L^p(\boldsymbol{A};\mathbb{C})$. Then, $([f_k])_{k\in\mathbb{N}}$ converges to $[f]$ in $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$.*

Proof Let $([f_k])_{k\in\mathbb{N}}$ converge to $[f]$ in $L^p(\boldsymbol{A};\mathbb{C})$. To show the convergence in $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, we have to show that for any compact subset $\boldsymbol{K}$ in $\boldsymbol{A}$ we have

$$\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \overset{k\to+\infty}{\longrightarrow} 0.$$

In other words, we show that for any $\varepsilon > 0$ there exists $p \in \mathbb{N}$ such that, for all $k > p$, we have

$$\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < \varepsilon.$$

Let $\varepsilon > 0$. We exploit the inequality of Lemma 5.3

$$\forall\, x \ge 0,\ \forall\, \eta > 0, \qquad x \le \eta + \eta^{1-p}\, x^p,$$

where $p \ge 1$. Let $\eta > 0$. We have

$$|f_k(\boldsymbol{x}) - f(\boldsymbol{x})| \le \eta + \eta^{1-p}\,|f_k(\boldsymbol{x}) - f(\boldsymbol{x})|^p,$$

so that

$$\begin{aligned}\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &\le \eta\,\mu_n(\boldsymbol{K}) + \eta^{1-p}\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|^p\,\mathrm{d}\boldsymbol{x}\\ &\le \eta\,\mu_n(\boldsymbol{K}) + \eta^{1-p}\,\|\,[f_k] - [f]\,\|_p^p.\end{aligned}$$

We choose $\eta > 0$ so that $\eta\,\mu_n(K) < \frac{\varepsilon}{2}$. Moreover, since $\|\,[f_k] - [f]\,\|_p$ tends to zero as $k \to \infty$ in view of the convergence in $L^p(\boldsymbol{A};\mathbb{C})$, there exists $p \in \mathbb{N}$ such that for all $k > p$ we have $\|\,[f_k] - [f]\,\|_p < \left(\frac{\varepsilon}{2\,\eta^{1-p}}\right)^{\frac{1}{p}}$, so that for all $k > p$ we have also $\eta^{1-p}\,\|\,[f_k] - [f]\,\|_p^p < \frac{\varepsilon}{2}$. Thus,

$$\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \le \eta\,\mu_n(\boldsymbol{K}) + \eta^{1-p}\,\|\,[f_k] - [f]\,\|_p^p < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Exercises

Exercise 5.4 Let $\boldsymbol{A} \subseteq \mathbb{R}^2$ be open. The sets $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and $L^\infty(\boldsymbol{A};\mathbb{C})$ contain different types of elements. Explain in which sense we can write the inclusion $\mathcal{D}(\boldsymbol{A};\mathbb{C}) \subseteq L^\infty(\boldsymbol{A};\mathbb{C})$, and provide a simple proof. [Solution: (5.4)]

Exercise 5.5 Establish whether the functions

(i) $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = c$, where $c \neq 0$

(ii) $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = e^x$

(iii) $f : \mathbb{R} \longrightarrow \mathbb{R}$ described by $f(x) = \sin x$

(iv) $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ such that $f(x) = \arctan \frac{1}{x}$

are in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. [Solution: (5.5)]

Exercise 5.6 Show that the sequence of function $(f_k)_{k\in\mathbb{N}}$ defined by

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \frac{2}{\pi} \arctan(k\,t) \end{aligned}$$

converges to the sign function in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. [Solution: (5.6)]

Exercise 5.7 Establish if the function

$$f(x) = \frac{1}{x^2 - 1}$$

is in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. [Solution: (5.7)]

Exercise 5.8 Consider the space

$$\boldsymbol{W} = \{[f] \in L^2(\mathbb{R};\mathbb{R}) \,|\, [x\,f] \in L^2(\mathbb{R};\mathbb{R})\}.$$

Show that $\boldsymbol{W} \subseteq L^1(\mathbb{R};\mathbb{R})$. [Solution: (5.8)]

5.3 Linear functionals associated with locally summable functions

Given an open set $\boldsymbol{A} \subseteq \mathbb{R}^n$ and the function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$, we want to study how the integral of the product of f by the generic function φ depends on the latter. In other words, we want to study the mapping

$$\varphi \mapsto \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{5.11}$$

We need in particular to establish which assumptions we need on f and φ so that the integral in (5.11) converges. Roughly speaking, if f is chosen from an ample functional space, we need to impose more stringent restrictions on φ to ensure that the generality of f is compensated by the special structure of φ in the product. In particular, if $f \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$, we can restrict φ to be of compact support: in this case the integral in (5.11) is well defined, because φ is zero outside a certain compact set, whereas f is summable

over every compact set. More precisely, from $\operatorname{supp}\varphi \subset \boldsymbol{K}$, we find that $f\,\varphi$ is summable over $\boldsymbol{A}$:

$$\begin{aligned}\int_{\boldsymbol{A}} |f(\boldsymbol{x})\,\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,|\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,|\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,(\sup_{\boldsymbol{x}\in\boldsymbol{K}} |\varphi(\boldsymbol{x})|)\,\mathrm{d}\boldsymbol{x}\\ &= \sup_{\boldsymbol{x}\in\boldsymbol{K}} |\varphi(\boldsymbol{x})| \int_{\boldsymbol{g}} K|f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty,\end{aligned}$$

since $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. Thus, a sufficient condition for the convergence of the integral in (5.11) is that φ is bounded and with compact support. Since in the sequel we will want to substitute f with distributions, which are much more irregular objects, we restrict φ to belong to a very regular functional space that can compensate the irregular behaviour of f: we assume that φ belongs to the class of infinitely differentiable functions with compact support. We therefore construct the following functional, which does exactly the job of the mapping (5.11): it takes an infinitely differentiable complex-valued function with compact support (i.e., a function in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$), and returns a number. A function of this kind is also referred to as a *functional*: in more general terms, the term "functional" denotes a mapping from a vector space into its underlying scalar field (in this case the vector space is $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and the field is $\mathbb{C}$). This functional will be denoted by L_f, because it depends on the choice of f, and it allows us to write (5.11) as

$$L_f(\varphi) = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Definition 5.7 [The Functional L_f]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $f \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. Let us define the functional L_f as

$$\begin{aligned} L_f : \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C}\\ \varphi &\mapsto \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \end{aligned} \tag{5.12}$$

An obvious property of the functional L_f is the linearity (try!). We want to show that this functional L_f is also continuous: our aim is to show that if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, then $L_f(\varphi_k)$ tends to $L_f(\varphi)$ as $k \to \infty$. This property requires first to define a notion of convergence on $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. To this end, we introduce the multi-index notation, which generalizes the concept of an integer index to an ordered tuple of indices. A multi-index is an n-tuple of natural numbers: $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{N}^n$. Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open, and let $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ be "sufficiently regular": we can assume for the sake of simplicity that f be of class C^∞ to ensure that Schwarz's theorem holds for the interchange of the partial derivatives. We define

$$\partial^{\boldsymbol{\alpha}} f \stackrel{\text{def}}{=} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f,$$

where $|\boldsymbol{\alpha}| = \alpha_1 + \ldots + \alpha_n$ denotes the absolute value (or length, or weight of the multi-index). We write also $\partial^{\boldsymbol{\alpha}} = \partial_1^{\alpha_1}\,\partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n}$. If $\partial^{\boldsymbol{\alpha}} f$ is continuous in $\boldsymbol{A}$, we say that f is of class $C^{|\boldsymbol{\alpha}|}$. As usual, if $\boldsymbol{\alpha} = (0, 0, \ldots, 0)$, then $\partial^{\boldsymbol{\alpha}} f = f$. Finally, given two multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}^n$, when we write $\boldsymbol{\beta} \le \boldsymbol{\alpha}$ we will mean that $\beta_j \le \alpha_j$ for all $j \in \{1, \ldots, n\}$.

Example 5.13

Consider the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, y, z) &\mapsto x\, y\, \sin z + e^x \cos(y\, z). \end{aligned}$$

Let $\boldsymbol{\alpha} = (2, 2, 1)$. We want to determine $\partial^{\alpha} f$:

$$\partial^{\alpha} f(x, y, z) = \frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2} \frac{\partial}{\partial z} f(x, y, z) = -2\, z\, e^x \cos(y\, z) + e^x\, z^2\, y \sin(y\, z).$$

The following result shows how Leibniz rule can be generalised for multi-indices. We recall that, in the case $n = 1$, if A is an open set of $\mathbb{R}$, then the usual Leibniz rule is

$$(f\, g)^{(\alpha)}(x) = \sum_{\beta=0}^{\alpha} \binom{\alpha}{\beta} f^{(\alpha-\beta)}(x)\, g^{(\beta)}(x), \qquad \binom{\alpha}{\beta} = \frac{\alpha!}{\beta!\,(\alpha-\beta)!}.$$

When $\alpha = 1$ we have

$$(f\, g)' = \binom{1}{0} f^{(1)}\, g^{(0)} + \binom{1}{1} f^{(0)}\, g^{(1)} = f^{(1)}\, g^{(0)} + f^{(0)}\, g^{(1)} = f'\, g + f\, g'.$$

In the case of multi-indices, the binomial coefficient becomes

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} = \binom{\alpha_1}{\beta_1}\binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n} = \frac{\boldsymbol{\alpha}!}{\boldsymbol{\beta}!\,(\boldsymbol{\alpha}-\boldsymbol{\beta})!}, \tag{5.13}$$

where the factorial of a multi-index $\boldsymbol{\alpha}$ is defined by

$$\boldsymbol{\alpha}! = \alpha_1!\, \alpha_2! \cdots \alpha_n!.$$

Lemma 5.4 [Generalized Leibniz Rule]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $f, g \in C^{\infty}(\boldsymbol{A}; \mathbb{C})$. Consider the multi-index $\boldsymbol{\alpha} \in \mathbb{N}^n$. Then,

$$\partial^{\alpha}(f\, g) = \sum_{\boldsymbol{\beta} \leq \boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}} f\, \partial^{\boldsymbol{\beta}} g. \tag{5.14}$$

Proof We proceed by induction. Let $\boldsymbol{\alpha} = \boldsymbol{e}_j$, where $\boldsymbol{e}_j$ is the vector $(0, \ldots, 0, 1, 0, \ldots, 0)$ with n components, where the 1 is in position j. Denoting by ∂_j the first-order partial derivative with respect to x_j, we have, in view of the standard Leibniz's rule,

$$\partial^{\boldsymbol{e}_j}(f\, g) = \partial_j(f\, g) = (\partial_j f)\, g + f\, (\partial_j g).$$

Thus, in this case the summation in (5.14) involves the terms $\boldsymbol{\beta} = (0, 0, \ldots, 0)$ and $\beta = \boldsymbol{e}_j$, so that (5.14) becomes

$$\partial^{\boldsymbol{e}_j}(f\, g) = \binom{\boldsymbol{e}_j}{\boldsymbol{0}} (\partial_j f)\, (\partial^{(0,0,\ldots,0)} g) + \binom{\boldsymbol{e}_j}{\boldsymbol{e}_j} (\partial^{(0,0,\ldots,0)} f)\, (\partial_j g) = \binom{\boldsymbol{e}_j}{\boldsymbol{0}} (\partial_j f)\, g + \binom{\boldsymbol{e}_j}{\boldsymbol{e}_j} f\, (\partial_j g),$$

where $\boldsymbol{0} = (0, 0, \ldots, 0) \in \mathbb{N}^n$. Using (5.13) we see immediately that $\binom{\boldsymbol{e}_j}{\boldsymbol{0}} = \binom{\boldsymbol{e}_j}{\boldsymbol{e}_j} = 1$, since $\binom{1}{0} = \binom{0}{0} = 1$. Let us

now assume that (5.14) holds for $\boldsymbol{\alpha}$, and we prove it for $\boldsymbol{\alpha}+\boldsymbol{e}_k$, with $k \in \{1,\dots,n\}$. We have

$$\begin{aligned}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k}(f\,g) &= \partial_k\,\partial^{\boldsymbol{\alpha}}(f\,g) = \partial_k\left[\sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}}f\,\partial^{\boldsymbol{\beta}}g\right]\\ &= \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}}\left[\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\beta}}f\,\partial^{\boldsymbol{\beta}}g + \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}}f\,\partial^{\boldsymbol{\beta}+\boldsymbol{e}_k}g\right]\\ &= \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\beta}}f\,\partial^{\boldsymbol{\beta}}g + \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}}f\,\partial^{\boldsymbol{\beta}+\boldsymbol{e}_k}g.\end{aligned}$$

We change variable with $\boldsymbol{\gamma}=\boldsymbol{\beta}$ in the first summation, and $\boldsymbol{\gamma}=\boldsymbol{\beta}+\boldsymbol{e}_k$ in the second summation:

$$\begin{aligned}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k}(f\,g) &= \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\gamma}}f\,\partial^{\boldsymbol{\gamma}}g + \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}+\boldsymbol{e}_k}\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}\partial^{\boldsymbol{\alpha}-\boldsymbol{\gamma}+\boldsymbol{e}_k}f\,\partial^{\boldsymbol{\gamma}}g\\ &= \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}}\left[\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}}+\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}\right]\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\gamma}}f\,\partial^{\boldsymbol{\gamma}}g + \sum_{\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{e}_k}\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}\partial^{\boldsymbol{\alpha}-\boldsymbol{\gamma}+\boldsymbol{e}_k}f\,\partial^{\boldsymbol{\gamma}}g \qquad (5.15)\end{aligned}$$

We now want to show that

$$\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}}+\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}=\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}.$$

To this end, we first consider the case $n,k\in\mathbb{N}$ with $k\le n$. In this case

$$\begin{aligned}\binom{n}{k}+\binom{n}{k-1} &= \frac{n!}{k!\,(n-k)!}+\frac{n!}{(k-1)!\,(n-k+1)!} = \frac{n!\,(n-k+1)+n!\,k}{k!\,(n-k+1)!}\\ &= \frac{n!\,(n+1)}{k!\,(n+1-k)!} = \frac{(n+1)!}{k!\,(n+1-k)!} = \binom{n+1}{k}.\end{aligned}$$

Hence,

$$\begin{aligned}\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}}+\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k} &= \binom{\alpha_1}{\gamma_1}\binom{\alpha_2}{\gamma_2}\cdots\binom{\alpha_n}{\gamma_n}+\binom{\alpha_1}{\gamma_1}\binom{\alpha_2}{\gamma_2}\cdots\binom{\alpha_{k-1}}{\gamma_{k-1}}\binom{\alpha_k}{\gamma_k-1}\binom{\alpha_{k+1}}{\gamma_{k+1}}\cdots\binom{\alpha_n}{\gamma_n}\\ &= \binom{\alpha_1}{\gamma_1}\binom{\alpha_2}{\gamma_2}\cdots\binom{\alpha_{k-1}}{\gamma_{k-1}}\left[\binom{\alpha_k}{\gamma_k}+\binom{\alpha_k}{\gamma_k-1}\right]\binom{\alpha_{k+1}}{\gamma_{k+1}}\cdots\binom{\alpha_n}{\gamma_n}\\ &= \binom{\alpha_1}{\gamma_1}\binom{\alpha_2}{\gamma_2}\cdots\binom{\alpha_{k-1}}{\gamma_{k-1}}\binom{\alpha_k+1}{\gamma_k}\binom{\alpha_{k+1}}{\gamma_{k+1}}\cdots\binom{\alpha_n}{\gamma_n}=\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}.\end{aligned}$$

We use this result in (5.15) and we obtain

$$\begin{aligned}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k}(f\,g) &= \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\gamma}}f\,\partial^{\boldsymbol{\gamma}}g + \sum_{\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{e}_k}\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}\partial^{\boldsymbol{\alpha}-\boldsymbol{\gamma}+\boldsymbol{e}_k}f\,\partial^{\boldsymbol{\gamma}}g\\ &= \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\gamma}}f\,\partial^{\boldsymbol{\gamma}}g + \sum_{\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{e}_k}\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}-\boldsymbol{\gamma}+\boldsymbol{e}_k}f\,\partial^{\boldsymbol{\gamma}}g = \sum_{\boldsymbol{\gamma}\le\boldsymbol{\alpha}+\boldsymbol{e}_k}\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}\partial^{\boldsymbol{\alpha}+\boldsymbol{e}_k-\boldsymbol{\gamma}}f\,\partial^{\boldsymbol{\gamma}}g,\end{aligned}$$

where the second equality follows from the fact that, when $\boldsymbol{\gamma}=\boldsymbol{\alpha}+\boldsymbol{e}_k$, we have $\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}=\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}$. Indeed, $\binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}-\boldsymbol{e}_k}=\binom{\boldsymbol{\alpha}}{\boldsymbol{\alpha}}=1$, and $\binom{\boldsymbol{\alpha}+\boldsymbol{e}_k}{\boldsymbol{\gamma}}=\binom{\boldsymbol{\gamma}}{\boldsymbol{\gamma}}=1$. This concludes the proof.

We are now ready to introduce the convergence in the vector space $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. We will consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and one particular function $\varphi\in\mathcal{D}(\boldsymbol{A};\mathbb{C})$, and we will say that $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ if **(1)** there exists a compact set $\boldsymbol{K}$ in $\boldsymbol{A}$ that contains the supports of each φ_k, and **(2)** for every multi-index $\boldsymbol{\alpha}$, the sequence $(\partial^{\boldsymbol{\alpha}}\varphi_k)_{k\in\mathbb{N}}$ is uniformly convergent to $\partial^{\boldsymbol{\alpha}}\varphi$.

Recalling the definition of uniform convergence, this means that for all $\varepsilon > 0$ there exists $p \in \mathbb{N}$ such that for all $k > p$ we have $|\partial^\alpha \varphi_k(\boldsymbol{x}) - \partial^\alpha \varphi(\boldsymbol{x})| \leq \varepsilon$ for all $\boldsymbol{x} \in \boldsymbol{A}$.

Definition 5.8 [Convergence in $\mathcal{D}$]
Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of functions of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, and let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. We say that $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ if the following two conditions are satisfied:

- $\exists\, \boldsymbol{K} \subseteq \boldsymbol{A}$ *compact such that* $\operatorname{supp}\varphi_k \subseteq \boldsymbol{K}$ *for all* $k \in \mathbb{N}$;
- $\forall\, \boldsymbol{\alpha} \in \mathbb{N}^n, \quad \lim_{k\to\infty} \partial^\alpha \varphi_k(\boldsymbol{x}) = \partial^\alpha \varphi(\boldsymbol{x})$ *uniformly.*

To say that $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, we write

$$\lim_{k\to\infty} \varphi_k = \varphi \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}).$$

If $n = 1$, we write the second condition in Definition 5.8 as

$$\forall\, \alpha \in \mathbb{N}, \quad \lim_{k\to\infty} \varphi_k^{(\alpha)}(x) = \varphi^{(\alpha)}(x) \quad \text{uniformly on } \mathcal{D}(\boldsymbol{A};\mathbb{C}).$$

Remark 5.4 If $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, and if $\boldsymbol{K} \subseteq \boldsymbol{A}$ is a compact such that $\operatorname{supp}\varphi_k \subseteq \boldsymbol{K}$ for all $k \in \mathbb{N}$ (which exists by Definition 5.8), then we have also $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$. Indeed, the condition of the uniform convergence holds in particular for $\boldsymbol{\alpha} = (0, 0, \ldots, 0)$, and this implies that $(\varphi_k)_{k\in\mathbb{N}}$ converges uniformly to φ, and therefore it also converges pointwise to φ. Then, for $\boldsymbol{x} \notin \boldsymbol{K}$ we have

$$\varphi(\boldsymbol{x}) = \lim_{k\to\infty} \varphi_k(\boldsymbol{x}) = \lim_{k\to\infty} 0 = 0.$$

Example 5.14 Let $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and let $\varphi_k(x) = 2^{-k}\psi(x)$ for all $x \in \mathbb{R}$, i.e., for all $k \in \mathbb{N}$

$$\begin{aligned} \varphi_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 2^{-k}\psi(x). \end{aligned}$$

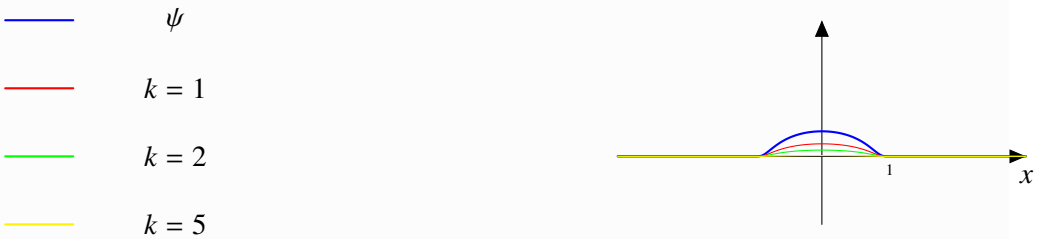
The sequence $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$ (why?). We want to show that the sequence $(\varphi_k)_{k\in\mathbb{N}}$ converges to $\varphi = 0$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$. First, we find the compact set $K \subset \mathbb{R}$ that contains the supports of each φ_k. Since $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, there exists a compact subset K of $\mathbb{R}$ such that $\operatorname{supp}\psi \subseteq K$. Now, if $x \notin K$, then $\varphi_k(x) = 0$. Therefore, the supports of all φ_k are contained in K. Now we have to check the second condition. We first observe that the identically zero function is the pointwise limit of $(\varphi_k)_{k\in\mathbb{N}}$. We need to prove that it is also the uniform limit, not just of $(\varphi_k)_{k\in\mathbb{N}}$, but also of all its derivatives. Let $\alpha \in \mathbb{N}$. To prove the uniform convergence of the derivatives, we need to show that the sequence $\left(\sup_{x\in\mathbb{R}} \left|\varphi_k^{(\alpha)}(x) - \varphi^{(\alpha)}(x)\right|\right)_{k\in\mathbb{N}^*}$ is infinitesimal. On the other hand, we have

$$\begin{aligned} \sup_{x\in\mathbb{R}} \left|\varphi_k^{(\alpha)}(x) - \varphi^{(\alpha)}(x)\right| &= \sup_{x\in\mathbb{R}} \left|\varphi_k^{(\alpha)}(x)\right| \\ &= \sup_{x\in\mathbb{R}} \left|2^{-k}\psi^{(\alpha)}(x)\right| \\ &= 2^{-k} \max_{x\in K} \left|\psi^{(\alpha)}(x)\right| \xrightarrow{k\to\infty} 0 \end{aligned}$$

by Weierstrass extreme value theorem (indeed, $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ implies $\psi^{(\alpha)} \in \mathcal{D}(\mathbb{R};\mathbb{R})$ for all $\alpha \in \mathbb{N}$). Thus,

$\varphi_k^{(\alpha)}(x) \xrightarrow{k\to\infty} 0$ uniformly on $\mathbb{R}$, so that $(\varphi_k)_{k\in\mathbb{N}}$ converges to 0 in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

In the figure below, we considered as ψ the function $\psi(x) = e^{-\frac{1}{1-x^2}}\chi_{|x|<1}$.



Example 5.15

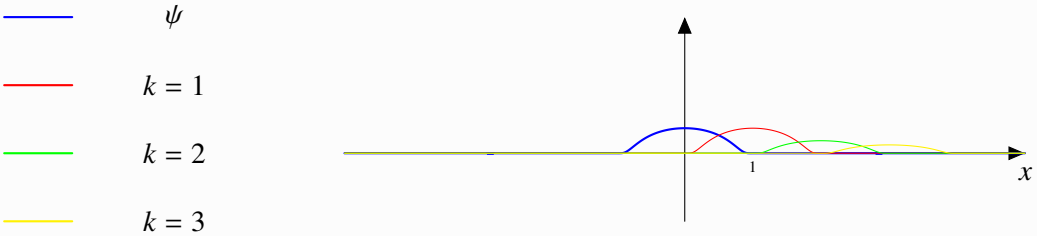
Let $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and let, for all $x \in \mathbb{R}$ and $k \in \mathbb{N}^*$,

$$\varphi_k(x) = \frac{1}{k}\psi(x-k).$$

We want to determine whether or not the sequence $(\varphi_k)_{k\in\mathbb{N}}$ converges in $\mathcal{D}(\mathbb{R};\mathbb{R})$. First, we notice that for every $k \in \mathbb{N}^*$ the function φ_k is indeed in $\mathcal{D}(\mathbb{R};\mathbb{R})$. In fact, φ_k is in $C^\infty(\mathbb{R}\,\mathbb{R})$, because it is the composition of functions in $C^\infty(\mathbb{R}\,\mathbb{R})$. Since ψ has compact support, there exists $R > 0$ such that $\psi(t) = 0$ for all $|t| > R$, which implies $\psi(x-k) = 0$ for all $|x-k| > R$. It follows that $\operatorname{supp}\varphi_k \subseteq [-R+k, R+k]$, and therefore $\varphi_k \in \mathcal{D}(\mathbb{R};\mathbb{R})$. The second condition in Definition 5.8 is satisfied. Indeed, to prove the uniform convergence of the derivatives, we need to show that the sequence $\left(\sup_{x\in\mathbb{R}}\left|\varphi_k^{(\alpha)}(x) - \varphi^{(\alpha)}(x)\right|\right)_{k\in\mathbb{N}^*}$ is infinitesimal, and this is the case because

$$\begin{aligned}\sup_{x\in\mathbb{R}}\left|\varphi_k^{(\alpha)}(x) - \varphi^{(\alpha)}(x)\right| &= \sup_{x\in\mathbb{R}}\left|\frac{1}{k}\psi^{(\alpha)}(x-k)\right| \\ &= \sup_{t\in\mathbb{R}}\left|\frac{1}{k}\psi^{(\alpha)}(t)\right| \\ &= \frac{1}{k}\max_{t\in K}\left|\psi^{(\alpha)}(t)\right| \xrightarrow{k\to\infty} 0.\end{aligned}$$

The first condition in Definition 5.8, however, is not satisfied, unless ψ is identically zero. In fact, let $x_0 \in \mathbb{R}$ be such that $\psi(x_0) \neq 0$. Then, $\varphi_k(x_0+k) \neq 0$, which implies $x_0 + k \in \operatorname{supp}\varphi_k$. Since compact sets are bounded, there cannot be exist a compact set that contains $\{x_0 + k \,|\, k \in \mathbb{N}^*\}$. It follows that the sequence $(\varphi_k)_{k\in\mathbb{N}}$ does not converge in $\mathcal{D}(\mathbb{R};\mathbb{R})$.



Exercise 5.9

Let A be an open set of $\mathbb{R}$. Let $[f] \in L^1_{\text{loc}}(A;\mathbb{R})$ and $\psi \in C(A;\mathbb{R})$. Show that $[\psi f] \in L^1_{\text{loc}}(A;\mathbb{R})$. [Solution: (5.9)]

Exercise 5.10

Let A be an open set of $\mathbb{R}$. Let $\psi \in C^\infty(A;\mathbb{R})$ and $\varphi \in \mathcal{D}(A;\mathbb{R})$. Show that $\psi\varphi \in \mathcal{D}(A;\mathbb{R})$. [Solution: (5.10)]

Now that we have introduced a notion of convergence in $\mathcal{D}$, we are ready to prove that the functional L_f is continuous with respect to this type of convergence.

Theorem 5.9 [Continuity of L_f]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. The functional L_f defined in (5.7) is continuous.

Proof Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then, continuity with respect to the notion of convergence introduced in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ means that if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, then

$$\lim_{k\to\infty} L_f(\varphi_k) = L_f(\varphi),$$

or, equivalently, by $|L_f(\varphi_k) - L_f(\varphi)| \xrightarrow{k\to\infty} 0$. Since $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ that converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, for every $k \in \mathbb{N}$, the function φ_k is zero outside a certain compact set $\boldsymbol{K}$ (stated differently, there exists a compact set $\boldsymbol{K} \subseteq \boldsymbol{A}$ such that $\operatorname{supp}\varphi_k \subseteq \boldsymbol{K}$ for every $k \in \mathbb{N}$), so that

$$\begin{aligned}
|L_f(\varphi_k) - L_f(\varphi)| &= \left| \int_{\boldsymbol{A}} f(\boldsymbol{x})\,(\varphi_k(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \right| \\
&\leq \int_{\boldsymbol{A}} |f(\boldsymbol{x})|\,|\varphi_k(\boldsymbol{x}) - \varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,|\varphi_k(\boldsymbol{x}) - \varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\
&\leq \sup_{\boldsymbol{x}\in\boldsymbol{K}} |\varphi_k(\boldsymbol{x}) - \varphi(\boldsymbol{x})| \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} 0
\end{aligned}$$

because $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ uniformly.

Consider now $[f_1], [f_2] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. If f_1 is equal to f_2 almost everywhere in $\boldsymbol{A}$, then $L_{f_1} = L_{f_2}$, i.e.,

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \int_{\boldsymbol{A}} f_1(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f_2(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

(why?). The following result shows that the converse is also true.

Lemma 5.5 [Fundamental Lemma]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $[f_1], [f_2] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. Then, $L_{f_1} = L_{f_2}$ if and only if $f_1(\boldsymbol{x}) = f_2(\boldsymbol{x})$ for almost every $\boldsymbol{x} \in \boldsymbol{A}$.

Proof The **only if** part, which was anticipated before the statement of this theorem, is straightforward and it is left as an exercise. We show that if $L_{f_1} = L_{f_2}$, then $f_1 = f_2$ almost everywhere. Let $f \stackrel{\text{def}}{=} f_1 - f_2$. We show that there exists a sequence $(\boldsymbol{B}_k)_{k\in\mathbb{N}}$ of open balls such that

$$\boldsymbol{A} = \bigcup_{k\in\mathbb{N}} \boldsymbol{B}_k.$$

Since $\boldsymbol{A}$ is open, for any $\boldsymbol{x} \in A$ there exists $r_{\boldsymbol{x}} > 0$ such that $\mathcal{B}_{r_{\boldsymbol{x}}}(\boldsymbol{x}) \subseteq \boldsymbol{A}$. The set $\boldsymbol{A}$ is the union

$$\boldsymbol{A} = \bigcup_{\substack{\boldsymbol{x}\in\boldsymbol{A} \\ r_{\boldsymbol{x}}>0 \text{ s.t. } \mathcal{B}_{r_{\boldsymbol{x}}}(\boldsymbol{x})\subseteq\boldsymbol{A}}} \mathcal{B}_{r_{\boldsymbol{x}}}(\boldsymbol{x}).$$

Indeed, every point of $\boldsymbol{A}$ belongs to the union, because the union is taken for every $\boldsymbol{x} \in \boldsymbol{A}$, and every point of the union is a point of A from the way the union depends on $r_{\boldsymbol{x}}$. In this union, we can choose $r_{\boldsymbol{x}} \in \mathbb{Q}^*_+$, and we

can consider $\boldsymbol{q}_{\boldsymbol{x}} \in \mathcal{B}_{\frac{r_{\boldsymbol{x}}}{3}}(\boldsymbol{x})$. With this choice, $\boldsymbol{x} \in \mathcal{B}_{\frac{2r_{\boldsymbol{x}}}{3}}(\boldsymbol{q}_{\boldsymbol{x}}) \subseteq \mathcal{B}_{r_{\boldsymbol{x}}}(\boldsymbol{x})$. The inclusion follows from the fact that

$$\boldsymbol{p} \in \mathcal{B}_{\frac{2r_{\boldsymbol{x}}}{3}}(\boldsymbol{q}_{\boldsymbol{x}}) \quad \Rightarrow \quad \|\boldsymbol{p} - \boldsymbol{q}_{\boldsymbol{x}}\| < \frac{2\,r_{\boldsymbol{x}}}{3}$$

and

$$\|\boldsymbol{p} - \boldsymbol{x}\| \leq \|\boldsymbol{p} - \boldsymbol{q}_{\boldsymbol{x}}\| + \|\boldsymbol{q}_{\boldsymbol{x}} - \boldsymbol{x}\| < \frac{2\,r_{\boldsymbol{x}}}{3} + \frac{r_{\boldsymbol{x}}}{3} = r_{\boldsymbol{x}}.$$

Thus,

$$\boldsymbol{A} = \bigcup_{\substack{\boldsymbol{q}_{\boldsymbol{x}} \in \boldsymbol{A} \\ r_{\boldsymbol{x}} \in \mathbb{Q}_+^* \text{ s.t. } \mathcal{B}_{\frac{2r_{\boldsymbol{x}}}{3}}(\boldsymbol{q}_{\boldsymbol{x}}) \subseteq \boldsymbol{A}}} \mathcal{B}_{\frac{2r_{\boldsymbol{x}}}{3}}(\boldsymbol{q}_{\boldsymbol{x}}). \tag{5.16}$$

Now, the set $\{(\boldsymbol{q}, r) \in \mathbb{Q}^n \times \mathbb{Q} \,|\, r > 0\}$ is countable because it is a subset of $\mathbb{Q}^{n+1}$ (and we know that the Cartesian product of countable sets is countable). The set $P = \{\mathcal{B}_r(\boldsymbol{q}) \,|\, \boldsymbol{q} \in \mathbb{Q}^n \text{ and } r \in \mathbb{Q}_+^*\}$ is therefore also countable. The collection

$$\left\{\mathcal{B}_{r_{\boldsymbol{x}}}(\boldsymbol{q}_{\boldsymbol{x}}) \,|\, \boldsymbol{q}_{\boldsymbol{x}} \in \boldsymbol{A},\ r_{\boldsymbol{x}} \in \mathbb{Q}_+^* \text{ such that } \mathcal{B}_{\frac{2r_{\boldsymbol{x}}}{3}}(\boldsymbol{q}_{\boldsymbol{x}}) \subseteq \boldsymbol{A}\right\}$$

is a subset of P, which is countable, and therefore it is itself countable. We proved that the right hand-side of (5.16) is the union of a countable number of balls, as required. Since $\boldsymbol{A}$ can be viewed as a union of a sequence of balls, it is enough to show that given an open ball $\boldsymbol{B}$ such that $\overline{\boldsymbol{B}} \subset \boldsymbol{A}$, we have $f = 0$ almost everywhere in $\boldsymbol{B}$. We assume for simplicity that f is real-valued. Let $\boldsymbol{B}$ be such ball. For the sake of readability, we continue to denote by f the restriction of f to $\boldsymbol{B}$. We have $[f] \in L^1(\boldsymbol{B}; \mathbb{R})$. Given an arbitrary $\varepsilon > 0$, since $\mathcal{D}(\boldsymbol{B}; \mathbb{R})$ is dense in $L^1(\boldsymbol{B}; \mathbb{R})$, we can choose $h \in \mathcal{D}(\boldsymbol{B}; \mathbb{R})$ such that

$$\|[h] - [f]\|_1 \leq \varepsilon. \tag{5.17}$$

For all $k \in \mathbb{N}^*$, we define the functions

$$\begin{aligned} \varphi_k : \boldsymbol{B} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \tanh(k\,h(\boldsymbol{x})). \end{aligned}$$

We have $\varphi_k \in \mathcal{D}(\boldsymbol{B}; \mathbb{R})$ and $|\varphi_k(\boldsymbol{x})| \leq 1$ for all $\boldsymbol{x} \in \boldsymbol{B}$ and for all $k \in \mathbb{N}^*$. Moreover, it is easy to see that $h\,\varphi_k$ is pointwise convergent to $|h|$ in B, since $\tanh(k\,h(\boldsymbol{x}))$ is pointwise convergent to $\operatorname{sgn}(h(\boldsymbol{x}))$. Using Lebesgue's dominated convergence theorem (Theorem 3.47), we obtain

$$\lim_{k \to \infty} \int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{B}} |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}.$$

Notice that this is possible because

$$\begin{aligned} \big|h(\boldsymbol{x})\,\tanh(k\,h(\boldsymbol{x})) - |h(\boldsymbol{x})|\big| &= \big|h(\boldsymbol{x})\,\tanh(k\,h(\boldsymbol{x})) - h(\boldsymbol{x})\,\operatorname{sgn}(h(\boldsymbol{x}))\big| \\ &= |h(\boldsymbol{x})|\,(|\tanh(k\,h(\boldsymbol{x}))| + 1) \leq 2\,|h(\boldsymbol{x})|, \end{aligned}$$

and h is summable. This means that k can be chosen so that

$$\left|\int_{\boldsymbol{B}} |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\right| \mathrm{d}\boldsymbol{x} \leq \varepsilon.$$

We can now use (5.17) in the chain of inequalities

$$\begin{aligned}
\|f\|_1 &= \int_{\boldsymbol{B}} |f(\boldsymbol{x}) - h(\boldsymbol{x}) + h(\boldsymbol{x})|\,\mathrm{d}x \\
&\leq \int_{\boldsymbol{B}} |f(\boldsymbol{x}) - h(\boldsymbol{x})|\,\mathrm{d}x + \int_P |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \leq \int_{\boldsymbol{B}} |f(\boldsymbol{x}) - h(\boldsymbol{x})|\,\mathrm{d}x + \left|\int_{\boldsymbol{B}} |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\right| \\
&= \|[h] - [f]\|_1 + \left|\int_{\boldsymbol{B}} |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&\leq \|[h] - [f]\|_1 + \left|\int_{\boldsymbol{B}} |h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| + \left|\int_{\boldsymbol{B}} h(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| + \left|\int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&\leq \varepsilon + \varepsilon + \|[h] - [f]\|_1 \max_{\boldsymbol{x}\in\overline{\boldsymbol{B}}} |\varphi_k(\boldsymbol{x})| + \left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&\leq 3\,\varepsilon + \left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| = 3\,\varepsilon.
\end{aligned}$$

Since ε is arbitrary, we obtain $\|[f]\|_1 = 0$, so that $f(\boldsymbol{x}) = 0$ almost everywhere in $\boldsymbol{B}$.

Lemma 5.5 has established that the mapping between $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and the set $\{L_f \,|\, [f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})\}$ is injective.

Remark 5.5 In view of the fundamental lemma 5.5, the mapping $f \mapsto L_f$ is a bijection. As a consequence, a function f such that $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ can be identified with the corresponding functional L_f.

Exercises

Exercise 5.11

Let $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, where $\boldsymbol{A} \subseteq \mathbb{R}^n$ is open. Consider the function

$$\begin{aligned}
L_f : \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C} \\
\varphi &\mapsto \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

(i) Show that L_f is well defined, i.e., that $L_f(\varphi) = \displaystyle\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ exists in $\mathbb{C}$ for every $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$.

(ii) Show that L_f is a linear transformation.

[Solution: (5.11)]

Exercise 5.12

Let $[f_1], [f_2] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, where $\boldsymbol{A} \subseteq \mathbb{R}^n$ is open. If f_1 is equal to f_2 almost everywhere in $\boldsymbol{A}$, show that $L_{f_1} = L_{f_2}$. [Solution: (5.12)]

Exercise 5.13

Let $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and let $\varphi_k(x) = 2^{-k}\,\psi(k^5\,x)$ for all $x \in \mathbb{R}$ and for all $k \in \mathbb{N}^*$.

(i) Show that $(\varphi_k)_{k\in\mathbb{N}^*}$ is a sequence of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$;

(ii) Establish $(\varphi_k)_{k\in\mathbb{N}^*}$ converges to the identically zero function $\varphi = 0$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

[Solution: (5.13)]

Exercise 5.14

Let $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and consider the sequence $(\varphi_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$ described, for all $x \in \mathbb{R}$, by

$$\varphi_k(x) = 2^{-k}\,\psi\left(\frac{x}{k}\right).$$

(i) Show that $(\varphi_k)_{k\in\mathbb{N}^*}$ is a sequence of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$;

(ii) Establish whether $(\varphi_k)_{k\in\mathbb{N}^*}$ converges to the identically zero function $\varphi = 0$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

[Solution: (5.14)]

5.4 Definition of distribution

We begin this section giving a definition of distribution.

Definition 5.9 [Distribution]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$ be a linear continuous functional with respect to the convergence of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then, T is said to be a distribution over $\boldsymbol{A}$. We denote by $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ the set of distributions over $\boldsymbol{A}$.

We recall that the continuity of a linear functional $T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$ with respect to the convergence of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ means that, for any sequence $(\varphi_k)_{k\in\mathbb{N}}$ of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and any $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, we have

$$\lim_{k\to\infty} \varphi_k = \varphi \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad \lim_{k\to\infty} T(\varphi_k) = T(\varphi). \tag{5.18}$$

Show that, by linearity, (5.18) is equivalent to the condition

$$\lim_{k\to\infty} \varphi_k = 0 \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad \lim_{k\to\infty} T(\varphi_k) = 0,$$

which is sometimes easier to check.

The only examples of distributions that we know for now are the functionals L_f, where $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. In other words, the set $\mathcal{L}(\boldsymbol{A};\mathbb{C}) = \{L_f \,|\, [f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})\}$ of all linear and continuous functionals L_f defined by

$$\begin{aligned} L_f : \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \end{aligned}$$

where $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, is a subset of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$.

On the other hand, in view of the identification introduced in Remark 5.5 between the functions in $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and the corresponding functionals, the sets $\mathcal{L}(\boldsymbol{A};\mathbb{C})$ and $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ are isomorphic, and for this reason we can also think of $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ as a subset of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$.

In view of this identification, the notations introduced in the following definition are consistently used in different areas of mathematics, physics, engineering and applied mathematics.

Definition 5.10 [Notation]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then, the symbol $T(\varphi)$ is also denoted using one of the following symbols:

- $T[\varphi]$,

- $\langle T, \varphi \rangle$,
- (T, φ),
- $\displaystyle\int_A T(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$,
- $\displaystyle\int_A T\,\varphi$.

The symbol of integration used in this definition is not a Lebesgue integral, but it is usually still referred to as an integral. Indeed, several properties (for example additivity) of the Lebesgue integral obviously do not extend to this operation. This is the reason why sometimes the symbol $\langle T, \varphi \rangle$ is preferred. We also observe that the expression $\displaystyle\int_A T(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ hides an important abuse of notation, that will become more clear in the sequel: we will see that it makes no sense to write $T(\boldsymbol{x})$, i.e., the value of a distribution at a point is an ill-defined concept. However, this notation is still very much in use.

The symbol $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ used to denote the set of distributions $T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$ shows that this set is the *dual* of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$.[2]

An interesting continuity criterion, that will be used a few times in these notes, gives the following condition that can be used to check whether a certain functional over $\mathcal{D}$ is indeed a distribution.

Theorem 5.10 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $L : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$ be linear. Then, L is a distribution (i.e., $L \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$) if and only if, for every compact set $\boldsymbol{K}$ contained in $\boldsymbol{A}$, there exist two constants $M \geq 0$ and $m \in \mathbb{N}$ such that, for every $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$, the inequality*

$$|L(\varphi)| \leq M \max_{|\boldsymbol{\alpha}| \leq m} \|\partial^{\boldsymbol{\alpha}}\varphi\|_\infty \tag{5.19}$$

holds.

Proof **(If)**. We assume that (5.19) holds, and we prove that L is a distribution. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ that converges to the zero function $0 \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. It follows that there exists a compact set K such that the supports of each function φ_k are all contained in $\boldsymbol{K}$, i.e., $\operatorname{supp}\varphi_k \subseteq \boldsymbol{K}$ for all $k \in \mathbb{N}$. Now, by assumption, given $\boldsymbol{K}$ we know that two constants M and m can be found such that (5.19) holds. Let M and m be those two constants. Then,

$$|L(\varphi_k)| \leq M \max_{|\boldsymbol{\alpha}| \leq m} \|\partial^{\boldsymbol{\alpha}}\varphi_k\|_\infty \xrightarrow{k\to\infty} 0,$$

because $(\partial^{\boldsymbol{\alpha}}\varphi_k)_{k\in\mathbb{N}}$ goes to zero uniformly. It follows that $|L(\varphi_k)| \xrightarrow{k\to\infty} 0$.

(Only if). We have to show that, if for every compact set $\boldsymbol{K}$ contained in $\boldsymbol{A}$, there exist two constants $M \geq 0$ and $m \in \mathbb{N}$ such that,

$$\forall\,\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \operatorname{supp}\varphi \subseteq \boldsymbol{K} \;\Rightarrow\; |L(\varphi)| \leq M \max_{|\boldsymbol{\alpha}| \leq m} \|\partial^{\boldsymbol{\alpha}}\varphi\|_\infty,$$

[2] Given an n-dimensional vector space $\boldsymbol{V}$ over a field $\mathbb{F}$, a linear transformation $\phi : \boldsymbol{V} \longrightarrow \mathbb{F}$ is said to be a functional, or a linear form. The set of linear forms $\phi : \boldsymbol{V} \longrightarrow \mathbb{F}$ is itself an n-dimensional vector space over $\mathbb{F}$, it is denoted by $\boldsymbol{V}'$ and it is referred to as the *dual* of $\boldsymbol{V}$.

then L is a distribution, and in particular it is continuous. We proceed by contradiction: we assume that there exists a compact set $\boldsymbol{K}$ contained in $\boldsymbol{A}$ such that for all $M \geq 0$ and for all $m \in \mathbb{N}$, there exists $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ with $\operatorname{supp} \varphi \subseteq \boldsymbol{K}$ such that

$$|L(\varphi)| > M \max_{|\alpha| \leq m} \|\partial^\alpha \varphi\|_\infty.$$

Let K be this compact set. We can now choose M and m arbitrarily. Let us choose $M = m = k$, where k is an arbitrary strictly positive integer. Then, there exists $\psi_k \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ such that

$$|L(\psi_k)| > k \max_{|\alpha| \leq k} \|\partial^\alpha \psi_k\|_\infty. \tag{5.20}$$

Since k is arbitrary in $\mathbb{N}^*$, we are constructing a sequence $(\psi_k)_{k \in \mathbb{N}}$ in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$ whose supports are all contained in $\boldsymbol{K}$. Let us now define

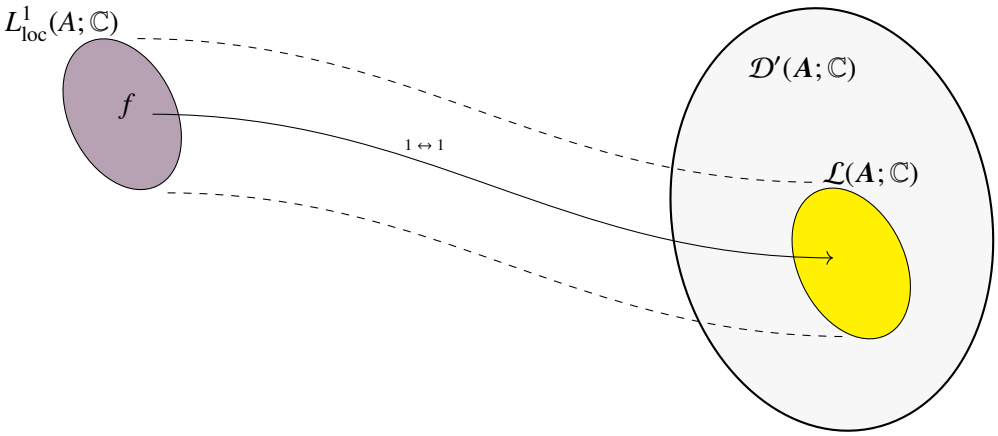
$$\varphi_k = \frac{\psi_k}{|L(\psi_k)|},$$

so that by linearity $|L(\varphi_k)| = \frac{|L(\psi_k)|}{|L(\psi_k)|} = 1$ for all $k \in \mathbb{N}^*$, which implies that the sequence $(L(\varphi_k))_{k \in \mathbb{N}^*}$ is not infinitesimal. Nevertheless, this sequence tends to 0 in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$. Indeed, first, by construction, the supports of each φ_k is contained in $\boldsymbol{K}$. Moreover, dividing both sides of (5.20) by $k\,|L(\psi_k)|$, using the linearity of L and of the derivatives as well as the homogeneity of the norm we obtain

$$\max_{|\alpha| \leq k} \|\partial^\alpha \varphi_k\|_\infty < \frac{1}{k}, \qquad \forall\, k \in \mathbb{N}^*,$$

which means that $\|\partial^\alpha \varphi_k\|_\infty < 1/k$ for every $k \in \mathbb{N}^*$ and for every differentiation of order $\leq k$. This shows that $\varphi_k \xrightarrow{k \to \infty} 0$ in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$, and therefore L is not continuous, which implies that L is not a distribution. This concludes the proof.

We now show that $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ is greater than $\mathcal{L}(\boldsymbol{A}; \mathbb{C})$, i.e., there are distributions that are not functionals obtainable from functions in $L^1_{\text{loc}}(\boldsymbol{A}; \mathbb{C})$.



In other words, distributions can be considered as a generalization of the concept of locally integrable functions. To prove that $\mathcal{D}'(\boldsymbol{A}; \mathbb{C}) \supsetneq \mathcal{L}(\boldsymbol{A}; \mathbb{C})$, it suffices to provide an example of a distribution that is not in $\mathcal{L}(\boldsymbol{A}; \mathbb{C})$. To this end, we introduce the Dirac delta, which is a function $\delta : \mathcal{D}(\mathbb{R}^n; \mathbb{R}) \longrightarrow \mathbb{R}$ defined in this way: for every $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$,

$$\delta(\varphi) = \langle \delta, \varphi \rangle = \varphi(0).$$

Before we can call δ a distribution, we must show that this definition is in line with the requirements of Definition 5.9. The following result takes care of this.

Lemma 5.6 *Let*

$$\begin{aligned} \delta : \mathcal{D}(\mathbb{R}^n;\mathbb{R}) &\longrightarrow \mathbb{R} \\ \varphi &\mapsto \varphi(0). \end{aligned}$$

Then, $\delta \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$.

Proof The linearity is trivial (try!). To prove the continuity, consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ that converges to φ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We have to prove that $\delta(\varphi_k) \xrightarrow{k\to\infty} \delta(\varphi)$. Since $\delta(\varphi_k) = \varphi_k(0)$ and $\delta(\varphi) = \varphi(0)$, this means that we have to prove that $\varphi_k(0) \xrightarrow{k\to\infty} \varphi(0)$; this follows from the convergence of $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$, which implies that $(\varphi_k)_{k\in\mathbb{N}}$ is uniformly convergent to φ, and therefore in particular also pointwise convergent.

An alternative proof of Lemma 5.6 consists in proving the linearity first, and then using the necessary and sufficient condition (5.19) in Theorem 5.10. Indeed, for every compact subset $\boldsymbol{K}$ of $\mathbb{R}^n$, choosing $M = 1$ and $m = 0$, for every test function $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ with support contained in $\boldsymbol{K}$ we have

$$|\delta(\varphi)| = |\varphi(0)| \leq \max_{|\alpha|=0} \left(\sup_{\boldsymbol{K}} |\partial^\alpha \varphi| \right) = \sup_{\boldsymbol{K}} |\varphi|.$$

Definition 5.11 [Dirac δ distribution]
Let $\delta \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ *be defined by*

$$\forall\, \varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R}), \qquad \delta(\varphi) = \varphi(0).$$

Then, δ *is called* Dirac δ *distribution.*

The main point that we want to prove is that $\delta \notin \mathcal{L}(\mathbb{R}^n;\mathbb{R})$.

Theorem 5.11

$$\nexists\, [f] \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R}) : \qquad \forall\, \varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R}), \quad \int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \varphi(0). \tag{5.21}$$

Proof Suppose by contradiction that a function $[f] \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ exists such that $\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \varphi(0)$ holds for every $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Consider in particular a function $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ such that $\operatorname{supp}\varphi \subseteq \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ (this is always possible because we have seen that bump functions can be shifted to obtain other bump functions), so that $\varphi(0) = 0$. Then,

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0.$$

Consider the restriction $f|_{\mathbb{R}^n\setminus\{\boldsymbol{0}\}}$. We have

$$\int_{\mathbb{R}^n\setminus\{\boldsymbol{0}\}} f|_{\mathbb{R}^n\setminus\{\boldsymbol{0}\}}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \qquad \forall\, \varphi \in \mathcal{D}(\mathbb{R}^n \setminus \{\boldsymbol{0}\};\mathbb{R}),$$

because the value of the integral does not change by only removing a point from the domain of integration. As

a direct consequence of Lemma 5.5, if $\boldsymbol{B} \subseteq \mathbb{R}^n$ is open and $[g] \in L^1_{\mathrm{loc}}(\boldsymbol{B};\mathbb{R})$, and

$$\int_{\boldsymbol{B}} g(\boldsymbol{x})\,\psi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$$

for all $\psi \in \mathcal{D}(\boldsymbol{B};\mathbb{R})$, then $g(\boldsymbol{x}) = 0$ almost everywhere in $\boldsymbol{B}$. If we apply this result with $\boldsymbol{B} = \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ and $g = f|_{\mathbb{R}^n\setminus\{\boldsymbol{0}\}}$, we find that $f|_{\mathbb{R}^n\setminus\{\boldsymbol{0}\}}(\boldsymbol{x}) = 0$ for almost every $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$; therefore, also $f(\boldsymbol{x}) = 0$ for almost all $\boldsymbol{x} \in \mathbb{R}^n$. This implies that

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \quad \forall\,\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R}).$$

Then, by choosing $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ such that $\varphi(0) \neq 0$ (and this is possible using a bump function centered at the origin), we obtain a contradiction.

Definition 5.12 [Distributions of Function Type]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We say that T is of function type *or* regular *if there exists $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ such that*

$$T(\varphi) = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \tag{5.22}$$

for every $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$.

When a distribution is of function type, the distribution itself is often identified with the function f, with $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, for which (5.22) holds. Specifically, it is frequent to use an abuse of notation consisting in writing $T(\varphi)$ as $\langle f,\varphi\rangle$ or even $f(\varphi)$, instead of $\langle L_f,\varphi\rangle$ or $L_f(\varphi)$. This point is important, because it allows us to understand what it is meant by saying that distributions generalize functions. A distribution **is** a function from every point of view: $T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$. When T is of function type, we can identify T with a function f such that $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, but when T is not of function type, a distribution can be viewed as an extension of the elements of $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$.

In view of Lemma 5.5, a distribution of function type T can be identified with the function f such that (5.22) holds for all $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. From Theorem 5.11, it turns out that the Dirac δ distribution is **not** a function-type distribution. Distributions that are not of function type are also called *singular*.

Exercises

Exercise 5.15

Let $[f] \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Let us define the functional

$$\begin{aligned} T : \mathcal{D}(\mathbb{R}^2;\mathbb{R}) &\longrightarrow \mathbb{R} \\ \varphi &\mapsto \int_{\mathbb{R}} f(x)\,\varphi(x,0)\,\mathrm{d}x. \end{aligned}$$

Show that T is a distribution and it is not of function type, adapting the arguments in the proof of Theorem 5.11. [Solution: (5.15)]

Exercise 5.16

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T : \mathcal{D}(A;\mathbb{C}) \longrightarrow \mathbb{C}$ be a linear functional with respect to the convergence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Show that the following two statements (involving the linearity of T) are equivalent:

(A) $\lim_{k\to\infty} \varphi_k = \varphi$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C}) \quad \Rightarrow \quad \lim_{k\to\infty} T(\varphi_k) = T(\varphi)$

(B) $\lim_{k\to\infty} \varphi_k = 0$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C}) \quad \Rightarrow \quad \lim_{k\to\infty} T(\varphi_k) = 0.$

[Solution: (5.16)]

Exercise 5.17 Prove the linearity of δ. [Solution: (5.17)]

Exercise 5.18 Consider the functional

$$\begin{aligned} T : \mathcal{D}(\mathbb{R};\mathbb{R}) &\longrightarrow \mathbb{R} \\ \varphi &\mapsto \int_0^{+\infty} x\,\varphi'(x)\,\mathrm{d}x. \end{aligned}$$

Show that T is a distribution of function type, and determine the function f such that $[f] \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ that allows us to write

$$T(\varphi) = \int_{\mathbb{R}} f(x)\,\varphi(x)\,\mathrm{d}x.$$

(Hint: Use integration by parts). [Solution: (5.18)]

Exercise 5.19 Let

$$\begin{aligned} f_k : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \dfrac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))} & \text{if } \boldsymbol{x} \in \mathcal{B}_{\frac{1}{k}}(0) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Show that $\lim_{k\to\infty} L_{f_k} = \delta$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. [Solution: (5.19)]

Exercise 5.20 Consider the function

$$f_k(x) = \begin{cases} \sqrt{k} & x \in \left[0, \frac{1}{k}\right] \\ 0 & \text{otherwise} \end{cases}$$

Study the convergence of L_{f_k} in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. [Solution: (5.20)]

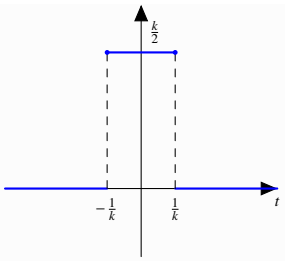
We now introduce a useful notion of convergence in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, which will give an explanation as to why the definition of the Dirac delta described in the introduction, while being mathematically meaningless, works in the most common situations.

Definition 5.13 [Convergence in $\mathcal{D}'$]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, and let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We say that $(T_k)_{k\in\mathbb{N}}$ converges to T in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ (or in the sense of distributions*) if*

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \lim_{k\to\infty} T_k(\varphi) = T(\varphi)$$

Example 5.16 Let $A = \mathbb{R}$. Consider, for every $k \in \mathbb{N}^*$, the function

$$f_k(t) = \begin{cases} \dfrac{k}{2} & t \in \left[-\dfrac{1}{k}, \dfrac{1}{k}\right] \\ 0 & \text{otherwise} \end{cases}$$



For every $k \in \mathbb{N}^*$, the equivalence class $[f_k]$ is in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$, and

$$\forall\, k \in \mathbb{N}, \qquad \int_{\mathbb{R}} f_k(t)\,\mathrm{d}t = 1.$$

Let us now consider $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and compute

$$\langle L_{f_k}, \varphi\rangle = \int_{\mathbb{R}} f_k(t)\,\varphi(t)\,\mathrm{d}t = \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} \varphi(t)\,\mathrm{d}t.$$

We want to find the limit, as $k \to \infty$, of this quantity: we add and subtract the quantity $\varphi(0)$, and obtain

$$\begin{aligned} \langle L_{f_k}, \varphi\rangle &= \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} \varphi(t)\,\mathrm{d}t - \varphi(0) + \varphi(0) \\ &= \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} \varphi(t)\,\mathrm{d}t - \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} \varphi(0)\,\mathrm{d}t + \varphi(0) \\ &= \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} (\varphi(t) - \varphi(0))\,\mathrm{d}t + \varphi(0). \end{aligned}$$

It is easy to see that the first term in the right hand-side of the latter vanishes. Indeed, since φ is continuous, for every $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that $|t| < \delta_\varepsilon$ implies $|\varphi(t) - \varphi(0)| < \varepsilon$. Thus, since in the integral $|t| < \frac{1}{k}$, choosing k so that $\frac{1}{k} < \delta_\varepsilon$, i.e., such that $k > \frac{1}{\delta_\varepsilon}$, we find

$$\begin{aligned} \left|\frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} (\varphi(t) - \varphi(0))\,\mathrm{d}t\right| &\le \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} |\varphi(t) - \varphi(0)|\,\mathrm{d}t \\ &\le \frac{k}{2}\int_{-\frac{1}{k}}^{\frac{1}{k}} \varepsilon\,\mathrm{d}t < \frac{\cancel{k}}{\cancel{2}}\varepsilon\frac{\cancel{2}}{\cancel{k}} = \varepsilon. \end{aligned}$$

We have proved that for any $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ we have

$$\langle L_{f_k}, \varphi\rangle \xrightarrow{k\to\infty} \varphi(0) = \langle \delta, \varphi\rangle.$$

Thus, $\lim_{k\to\infty} L_{f_k} = \delta$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$.

Example *5.17*

The result of the previous example can be generalised to $\boldsymbol{A} = \mathbb{R}^n$ by considering the function

$$\begin{aligned} f_k : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \dfrac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))} & \text{if } \boldsymbol{x} \in \mathcal{B}_{\frac{1}{k}}(0) \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Try, using the same arguments of the previous example, that $\lim_{k\to\infty} L_{f_k} = \delta$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$.

Example 5.18

Let $A = \mathbb{R}$. We want to study the convergence of the functional L_{f_k}, where, for all $k \in \mathbb{N}$, f_k is the function

$$f_k(x) = \begin{cases} k^2 & x \in \left[0, \frac{1}{k}\right] \\ 0 & \text{otherwise.} \end{cases}$$

Note that $[f_k] \in L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned} \langle L_{f_k}, \varphi\rangle &= \int_{\mathbb{R}} f_k(x)\,\varphi(x)\,\mathrm{d}x = k^2 \int_0^{1/k} \varphi(x)\,\mathrm{d}x \\ &= k^2 \int_0^{1/k} (\varphi(x) - \varphi(0))\,\mathrm{d}x + k\,\varphi(0), \end{aligned}$$

which does not converge for any $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Thus, L_{f_k} does not converge in $\mathcal{D}'(\mathbb{R};\mathbb{R})$.

Example 5.19

Let $[g] \in L^1(\mathbb{R}^n;\mathbb{R})$. Let $f_k(\boldsymbol{x}) = k^n\, g(k\,\boldsymbol{x})$. Consider the sequence $(L_{f_k})_{k\in\mathbb{N}}$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. We want to find if the limit $\lim_{k\to\infty} L_{f_k}$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ exists. Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We find

$$\begin{aligned} \langle L_{f_k}, \varphi\rangle &= \int_{\mathbb{R}^n} k^n\, g(k\,\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= k^n \int_{\mathbb{R}^n} g(k\,\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= k^n \int_{\mathbb{R}^n} g(\boldsymbol{y})\,\varphi(k^{-1}\,\boldsymbol{y})\,k^{-n}\,\mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n} g(\boldsymbol{y})\,\varphi(k^{-1}\,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \end{aligned}$$

where we have introduced the new variable $\boldsymbol{y} = k\,\boldsymbol{x}$. Now, $\lim_{k\to\infty} g(\boldsymbol{y})\,\varphi(k^{-1}\,\boldsymbol{y}) = g(\boldsymbol{y})\,\varphi(0)$ pointwise. Moreover, since $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$, there exists $M > 0$ such that $|\varphi(\boldsymbol{x})| \le M$ for all $\boldsymbol{x} \in \mathbb{R}^n$, so that $|g(\boldsymbol{y})\,\varphi(k^{-1}\,\boldsymbol{y})| \le M\,|g(\boldsymbol{y})|$, and the function $M\,g$ is summable because such is g. Thus, using Lebesgue's dominated convergence theorem (Theorem 3.47) we find

$$\lim_{k\to\infty} \langle L_{f_k}, \varphi\rangle = \left(\int_{\mathbb{R}^n} g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\varphi(0).$$

It follows that $\lim_{k\to\infty} L_{f_k} = \left(\int_{\mathbb{R}^n} g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\delta$.

Example 5.20

We want to verify that for all $\lambda \in \mathbb{R}$, there holds $\lim_{k\to\infty} k^\lambda \sin(k\,t) = 0$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. This means that we are considering the sequence $(f_k)_{k\in\mathbb{N}}$ with $f_k(t) = k^\lambda \sin(k\,t)$ for $t \in \mathbb{R}$ in (with $[f_k] \in L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$) and we have to prove that for every $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ we have

$$\lim_{k\to\infty} L_{f_k}(\varphi) = \lim_{k\to\infty} \int_{\mathbb{R}} k^\lambda \sin(k\,t)\,\varphi(t)\,\mathrm{d}t = 0.$$

This is true when $\lambda < 0$, because

$$\left|\int_{\mathbb{R}} k^\lambda \sin(k\,t)\,\varphi(t)\,\mathrm{d}t\right| \le k^\lambda \int_{\mathbb{R}} |\varphi(t)|\,\mathrm{d}t \xrightarrow{k\to\infty} 0.$$

Let us consider now $\lambda \geq 0$. Since $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, there exists $M > 0$ such that $\varphi(t) = 0$ if $|t| > M$, so that

$$\begin{aligned}
\int_{\mathbb{R}} k^\lambda \sin(k\,t)\,\varphi(t)\,\mathrm{d}t &= k^\lambda \int_{-M}^{M} \sin(k\,t)\,\varphi(t)\,\mathrm{d}t \\
&= k^\lambda \left(\cancel{\left[-\frac{\cos(k\,t)}{k}\,\varphi(t) \right]_{-M}^{M}} + \int_{-M}^{M} \frac{\cos(k\,t)}{k}\,\varphi'(t)\,\mathrm{d}t \right) \\
&= k^{\lambda-1} \int_{-M}^{M} \cos(k\,t)\,\varphi'(t)\,\mathrm{d}t \\
&= -k^{\lambda-2} \int_{-M}^{M} \sin(k\,t)\,\varphi''(t)\,\mathrm{d}t \\
&\vdots \\
&= \pm k^{\lambda-m} \int_{-M}^{M} g(k\,t)\,\varphi^{(m)}(t)\,\mathrm{d}t,
\end{aligned}$$

where g is either the sine or the cosine function, and where we can choose m in such a way that $\lambda - m < 0$. For what we observed for $\lambda < 0$, we deduce that $\lim_{k\to\infty} k^\lambda \sin(k\,t) = 0$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ for all $\lambda \in \mathbb{R}$.

We now show that the notion of convergence in $\mathcal{D}'$ is consistent with the notion of convergence in L^1_{loc} introduced earlier.

Theorem 5.12 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $([f_k])_{k\in\mathbb{N}}$ be a sequence in $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. Let $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$. Then,*

$$f_k \xrightarrow{k\to\infty} f \text{ in } L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad L_{f_k} \xrightarrow{k\to\infty} L_f \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}).$$

Proof Since the sequence $([f_k])_{k\in\mathbb{N}}$ converges to $[f]$ in $L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, for every compact set $\boldsymbol{K}$ contained in $\boldsymbol{A}$ we have

$$\int_{\boldsymbol{K}} |f_k(\boldsymbol{x}) - f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} 0. \tag{5.23}$$

Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. We need to show that

$$\lim_{k\to\infty} \int_{\boldsymbol{A}} f_k(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \quad \text{or, equivalently,} \quad \int_{\boldsymbol{A}} (f_k(\boldsymbol{x}) - f(\boldsymbol{x}))\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} 0.$$

Let $\boldsymbol{K}$ be a compact set contained in $\boldsymbol{A}$ such that $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$. Then,

$$\begin{aligned}
\left| \int_{\boldsymbol{A}} (f_k(\boldsymbol{x}) - f(\boldsymbol{x}))\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right| &= \left| \int_{\boldsymbol{K}} (f_k(\boldsymbol{x}) - f(\boldsymbol{x}))\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \right| \\
&\leq \int_{\boldsymbol{K}} \left| f_k(\boldsymbol{x}) - f(\boldsymbol{x}) \right| |\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\
&\leq \max_{\boldsymbol{K}} |\varphi| \int_{\boldsymbol{K}} \left| f_k(\boldsymbol{x}) - f(\boldsymbol{x}) \right| \mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} 0,
\end{aligned}$$

where we used Weierstrass extreme value theorem, and the integral goes to zero in view of (5.23).

Theorem 5.13 [Completeness Theorem]

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. If the limit

$$\lim_{k\to\infty} T_k(\varphi)$$

exists in $\mathbb{C}$, then $L \overset{def}{=} \lim_{k\to\infty} T_k \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$.

We omit the proof of this result.[3] The linearity of L is easy to prove: it follows from the linearity of T_k for all $k \in \mathbb{N}$ (try as an exercise). The more challenging part is to prove the continuity, i.e., that

$$\varphi_k \xrightarrow{k\to\infty} \varphi \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad L(\varphi_k) \xrightarrow{k\to\infty} L(\varphi).$$

5.4.1 Series

Using the notion of convergence in $\mathcal{D}'$, we can give a meaning of a series of a sequence of distributions.

Lemma 5.7 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let*

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \sum_{k=0}^{\infty} T_k(\varphi) \tag{5.24}$$

(which is a series of complex numbers) be convergent. Then, the function

$$\begin{aligned} \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \sum_{k=0}^{\infty} T_k(\varphi) \end{aligned}$$

is a distribution.

Proof We can consider the partial sums

$$\sum_{k=0}^{m} T_k(\varphi) = T_0(\varphi) + T_1(\varphi) + \ldots + T_m(\varphi).$$

Since $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ is a vector space over $\mathbb{C}$, we can define for all $m \in \mathbb{N}$ the sequence of the partial sums $(s_m)_{m\in\mathbb{N}}$ as

$$s_m = T_0 + T_1 + \ldots + T_m,$$

which is itself a distribution, so that

$$s_m(\varphi) = \sum_{k=0}^{m} T_k(\varphi).$$

Therefore, the condition (5.24) can be expressed by saying that the limit

$$\lim_{m\to+\infty} s_m(\varphi)$$

[3] The standard proof of this result uses the so-called Banach-Steinhaus theorem (or uniform boundedness principle), which in turn uses Baire category theorem.

exists in $\mathbb{C}$ for every $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, and the limit is, with some abuse of notation, still denoted by $\sum_{k=0}^{\infty} T_k(\varphi)$. From the completeness theorem, this is a distribution, and therefore we can write

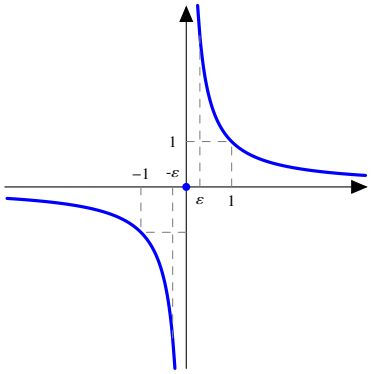
$$T(\varphi) = \sum_{k=0}^{\infty} T_k(\varphi),$$

where T is an element of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$.

5.4.2 Principal value of $1/x$

Consider the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases} \end{aligned}$$



We observe that $[f] \notin L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$, because the integral over a closed interval $[-M, M]$ (with $M > 0$), which is compact, does not exist. It follows that f does not generate a distribution of function type L_f. However, we can still define a distribution over $\mathbb{R}$, that we call *principal value* of $1/x$, and we denote it by $\operatorname{pv} \frac{1}{x}$. This distribution will be defined by the expression

$$\left\langle \operatorname{pv} \frac{1}{x}, \varphi \right\rangle = \lim_{\varepsilon \to 0^+} \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} \, \mathrm{d}x, \tag{5.25}$$

where $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$.[4] For this definition to be well posed, we need to verify that for all $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ the limit in (5.25) exists in $\mathbb{R}$.

We decompose the integral as

$$\begin{aligned} \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x} \, \mathrm{d}x &= \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x)}{x} \, \mathrm{d}x + \int_{|x|>1} \frac{\varphi(x)}{x} \, \mathrm{d}x \\ &= \int_{\varepsilon \leq |x| \leq 1} \frac{\varphi(x) - \varphi(0)}{x} \, \mathrm{d}x + \varphi(0) \underbrace{\int_{\varepsilon \leq |x| \leq 1} \frac{1}{x} \, \mathrm{d}x}_{=0} + \int_{|x|>1} \frac{\varphi(x)}{x} \, \mathrm{d}x. \end{aligned}$$

The second integral is zero because $\frac{1}{x}$ is odd. The function for positive and negative values of x goes to zero in the same way from the left and the right. The last integral exists because $\frac{\varphi(x)}{x}$ is C^∞ in $(0, +\infty)$

[4] Sometimes the definition of principal value is equivalently written as

$$\operatorname{pv} \frac{1}{x} = \lim_{\varepsilon \to 0^+} \frac{1}{x} \chi_{\mathbb{R} \setminus [-\varepsilon, \varepsilon]}.$$

and with compact support. Let us now focus on the first integral. We can write

$$\lim_{\varepsilon\to 0^+}\int_{\varepsilon\le|x|\le 1}\frac{\varphi(x)-\varphi(0)}{x}\,\mathrm{d}x = \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^*}\frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,-\varepsilon]\cup[\varepsilon,1]}(x)\,\mathrm{d}x$$
$$= \lim_{k\to\infty}\int_{\mathbb{R}^*}\frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,-\frac{1}{k}]\cup[\frac{1}{k},1]}(x)\,\mathrm{d}x.$$

We can apply tLebesgue's dominated convergence theorem (Theorem 3.47) on the sequence of functions defined for $k\in\mathbb{N}^*$ by

$$\begin{aligned} f_k : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,-\frac{1}{k}]\cup[\frac{1}{k},1]}(x). \end{aligned}$$

This sequence converges pointwise to the function

$$\begin{aligned} f : \mathbb{R}^* &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,1]\setminus\{0\}}(x). \end{aligned}$$

Moreover, defining $w:\mathbb{R}\longrightarrow\mathbb{R}$ by

$$w(x)=\begin{cases}\dfrac{\varphi(x)-\varphi(0)}{x} & \text{if } x\neq 0\\ \varphi'(0) & \text{if } x=0,\end{cases}$$

which is continuous, we have for all $k\in\mathbb{N}^*$ and for all $x\in\mathbb{R}^*$

$$|f_k(x)| = \left|\frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,-\frac{1}{k}]\cup[\frac{1}{k},1]}(x)\right| \le \max_{x\in[-1,1]}|w(x)|.$$

Interchanging the limit with the integral yields

$$\lim_{k\to\infty}\int_{\mathbb{R}^*}\frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,-\frac{1}{k}]\cup[\frac{1}{k},1]}(x)\,\mathrm{d}x = \int_{\mathbb{R}^*}\frac{\varphi(x)-\varphi(0)}{x}\chi_{[-1,1]\setminus\{0\}}(x)\,\mathrm{d}x$$
$$= \int_{[-1,1]\setminus\{0\}}\frac{\varphi(x)-\varphi(0)}{x}\,\mathrm{d}x = \int_{[-1,1]\setminus\{0\}}w(x)\,\mathrm{d}x = \int_{-1}^{1}w(x)\,\mathrm{d}x,$$

since w is continuous in $[-1,1]$. This integral is well defined because it is the integral of a continuous function over a compact set.

Now, we need to show that the principal value is a distribution. To this end, we consider the sequence $([f_k])_{k\in\mathbb{N}}$ in $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ in this way:

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases}\dfrac{1}{x} & \text{for } |x|\ge\frac{1}{k}\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

and for $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$ we have

$$\lim_{k\to\infty}\langle L_{f_k},\varphi\rangle = \lim_{k\to\infty}\int_{|x|\ge\frac{1}{k}}\frac{\varphi(x)}{x}\,\mathrm{d}x = \left\langle \mathrm{pv}\,\frac{1}{x},\varphi\right\rangle.$$

Now we know that the principal value is a distribution using the completeness theorem: we obtained it as the limit of a convergent sequence of distributions. However, it is not of function type. In fact, suppose by contradiction that there is $[f] \in L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$ such that $\text{pv}\,\frac{1}{x} = L_f$. This means that for every $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ we have

$$L_f(\varphi) = \left\langle \text{pv}\,\frac{1}{x}, \varphi \right\rangle = \lim_{\varepsilon \to 0^+} \int_{|x| \geq \varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x = \int_{\mathbb{R}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}x.$$

We can apply the fundamental lemma 5.5 to the open set $\mathbb{R}^*$, and from

$$\int_{\mathbb{R}^*} f|_{\mathbb{R}^*}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}x = \int_{x \in \mathbb{R}^*} \frac{\varphi(x)}{x}\,\mathrm{d}x = \int_{|x|>0} \frac{\varphi(x)}{x}\,\mathrm{d}x \qquad \forall\, \varphi \in \mathcal{D}(\mathbb{R}^*;\mathbb{R}),$$

we find that $f(x) = 1/x$ almost everywhere in $\mathbb{R}^*$. It follows that $f(x) = 1/x$ almost everywhere in $\mathbb{R}$. On the other hand, f is locally summable in $\mathbb{R}$, whereas $1/x$ is not, leading to a contradiction.

Remark 5.6 If φ is a function defined over $\mathbb{R}$, and regular except for a point x_0 and integrable over the complement of every neighbourhood of x_0, the principal value of the integral of f is defined as

$$\text{pv} \int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x = \lim_{\varepsilon \to 0^+} \int_{|x-x_0| \geq \varepsilon} \varphi(x)\,\mathrm{d}x,$$

when it exists and it is finite. This justifies the name *principal value*. Another case where we can define the principal value of the integral is the one of a locally integrable function in $\mathbb{R}$, which may not be integrable. In this case

$$\text{pv} \int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x = \lim_{R \to +\infty} \int_{-R}^{R} \varphi(x)\,\mathrm{d}x,$$

if it exists and it is finite. The two definitions can be combined and extended in the case where we have more than one singularity. For example, if φ is locally integrable over $\mathbb{R} \setminus \{x_0\}$ we define

$$\text{pv} \int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x = \lim_{\substack{\varepsilon \to 0^+ \\ R \to +\infty}} \int_{\substack{|x| \leq R \\ |x-x_0| \geq \varepsilon}} \varphi(x)\,\mathrm{d}x,$$

if it exists and it is finite.

Exercises

Exercise 5.21 Find the limit of $\sin^2(k\,t)$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. [Solution: (5.21)]

Exercise 5.22 Let $h \in L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. Let us define the functional

$$\begin{aligned} T : \mathcal{D}(\mathbb{R}^2;\mathbb{R}) &\longrightarrow \mathbb{R} \\ \varphi &\mapsto \int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x. \end{aligned}$$

(i) Show that T is a distribution. To this end, prove that:

- the function $\psi(x) = \varphi(x,0)$ is in $\mathcal{D}(\mathbb{R};\mathbb{R})$;
- T is linear and continuous (with respect to the convergence in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

(ii) Show that T is not of function type. To this end, try to use the same argument that was used to prove that δ is not of function type, and proceed by contradiction, assuming that $T = L_f$, where $[f] \in L^1_{\text{loc}}(\mathbb{R}^2;\mathbb{R})$. Hint: consider a restriction of the integral to the set $A = \{(x,y) \in \mathbb{R}^2 \,|\, y \neq 0\}$ to show that T needs to be the zero distribution.

(iii) Show that T can be obtained as the limit as $\epsilon \to 0^+$ of functionals L_{f_ϵ} defined by

$$\begin{aligned} f_\epsilon : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto \begin{cases} \frac{1}{\epsilon} h(x) & \text{for } 0 \le y < \epsilon \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

[Solution: (5.22)]

Exercise 5.23

Let $A = \mathbb{R}$. Consider the sequence of functions $(f_k)_{k\in\mathbb{N}^*}$ from A to $\mathbb{R}$ defined by

$$f_k(t) = \begin{cases} 3\,k & t \in \left[0, \frac{1}{k}\right] \\ 0 & \text{otherwise} \end{cases}$$

Study the convergence of the distribution of function type L_{f_k} as $k \to +\infty$ in $\mathcal{D}'(A;\mathbb{R})$. [Solution: (5.23)]

Exercise 5.24

Given a set $\boldsymbol{A} \subseteq \mathbb{R}^n$, recall that the conjugate of a function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ is the function

$$\begin{aligned} \bar{f} : \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \overline{f(\boldsymbol{x})}. \end{aligned}$$

Recall also that f is summable if $\mathfrak{Re}\, f$ and $\mathfrak{Im}\, f$ are summable, and

$$\int_{\boldsymbol{A}} f = \int_A \mathfrak{Re}\, f + i \int_{\boldsymbol{A}} \mathfrak{Im}\, f.$$

Let $\boldsymbol{A} \subseteq \mathbb{R}^2$ be open. Let $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$.

(a) Show that $[\bar{f}] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$;

(b) Show that for any $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ we have

$$\langle L_{\bar{f}}, \varphi\rangle = \overline{\langle L_f, \bar{\varphi}\rangle};$$

(c) Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Define $\bar{T}$ as

$$\langle \bar{T}, \varphi\rangle \stackrel{\text{def}}{=} \overline{\langle T, \bar{\varphi}\rangle};$$

Show that this definition is well posed, and that $\bar{T} \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$;

(d) Show that, given a sequence $(T_k)_{k\in\mathbb{N}}$ of $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$, then

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \quad \Rightarrow \quad \bar{T}_k \xrightarrow{k\to\infty} \bar{T} \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}).$$

[Solution: (5.24)]

5.5 Operations on distributions

In this section we extend some elementary properties of locally integrable functions to distributions. Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$. The first operations are those that endow $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ with a vector space structure: we have already implicitly used them before. If T_1 and T_2 are two distributions, the sum $T_1 + T_2$ is the functional

$$\begin{aligned} T_1 + T_2 : \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto T_1(\varphi) + T_2(\varphi). \end{aligned}$$

This is clearly a distribution. If T is a distribution and k is a scalar, the functional $k\,T$ is the functional

$$\begin{aligned} k\,T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto k\,T(\varphi), \end{aligned}$$

which is also a distribution. With these two operations, $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ is a vector space and $T(\varphi)$ is a bilinear form.

We now take into account other operations. The usual approach is the following:

- we first examine the operation for locally integrable functions that we wish to extend to distributions; we will isolate a property that we wish to be maintained for distributions. This property will be expressed in terms of a formula whose right hand-side continues to be well posed even in a distributional sense;
- we show that such term in the right hand-side defines a distribution when the test function varies in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$;
- we present a definition according to what is established as a result of the previous step.

5.5.1 Multiplication of a distribution by a function of class C^∞

We know that we can multiply a distribution by a scalar: this follows from the vector space structure of the set $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. A scalar can be thought of as a constant function, and this begs the question: can we multiply a distribution by other type of functions? The answer is yes, but only if we assume that function to be very infinitely differentiable (and the constant function falls within this class, obviously).

Let $\boldsymbol{A}$ be an open set of $\mathbb{R}^n$, and let $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$; let h be a function from $\boldsymbol{A}$ to $\mathbb{C}$. The product $f\,h$ satisfies the identity

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \int_{\boldsymbol{A}} (f(\boldsymbol{x})\,h(\boldsymbol{x}))\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,(h(\boldsymbol{x})\,\varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}. \tag{5.26}$$

This suggests that, interpreting f as its associated distribution, the expression

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \langle T\,h, \varphi\rangle = \langle T, h\,\varphi\rangle$$

can be used as a definition of multiplication of the distribution T by the function h. However, we must ensure that this expression makes sense:

- when T is associated with $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$, we should have $[f\,h] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$, or else the integral $\langle L_f\,h, \varphi\rangle = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,h(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ may not converge;

- $h\,\varphi$ should be in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$.

The second requirement imposes us to choose h from the class C^∞.[5] In other words, we cannot expect to be able to multiply a distribution by a generic function, but only by those functions that preserve the regularity of φ. Thus, now we prove that by taking $h \in C^\infty(\boldsymbol{A};\mathbb{C})$, the equivalence class $[f\,h]$ is still in $L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$ (so that the left hand-side of (5.26) is well defined). Then, we need to show that the right hand-side of (5.26) is also well-defined, and to this end we prove that when $h \in C^\infty(\boldsymbol{A};\mathbb{C})$, the product $h\,\varphi$ is still in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. We establish these results in the following lemma.

Lemma 5.8 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$; let $h \in C^\infty(\boldsymbol{A};\mathbb{C})$ and let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then,*

- $[f\,h] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$;
- $h\,\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$.

Proof We begin with the first statement. Let $\boldsymbol{K} \subset \boldsymbol{A}$ be a compact set. Since h is continuous, $|h|$ admits a maximum over $\boldsymbol{K}$, so that there exists $M > 0$ such that $|h(\boldsymbol{x})| \leq M$ for all $\boldsymbol{x} \in \boldsymbol{K}$ (in view of Weierstrass extreme value theorem). Thus,

$$\int_{\boldsymbol{K}} |f(\boldsymbol{x})\,h(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \leq M \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty$$

since $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$.
We now prove the second statement. We have $h\,\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$, because the product of two infinitely differentiable functions is infinitely differentiable. It is also clear that $h\,\varphi$ has compact support in $\boldsymbol{A}$, i.e., if $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$, where $\boldsymbol{K}$ is a compact set contained in $\boldsymbol{A}$, then also $\operatorname{supp}(h\,\varphi) \subseteq \boldsymbol{K}$. In fact, let $\boldsymbol{x}_0 \in \operatorname{supp}(h\,\varphi)$, and suppose by contradiction that $\boldsymbol{x}_0 \notin \boldsymbol{K}$. Since $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$, then $\boldsymbol{x}_0 \notin \operatorname{supp}\varphi$; this means that there exists an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ containing $\boldsymbol{x}_0$ such that $\varphi(\boldsymbol{x}) = 0$ for every $\boldsymbol{x} \in \boldsymbol{B}$, which also implies $h(\boldsymbol{x})\,\varphi(\boldsymbol{x}) = 0$ for every $\boldsymbol{x} \in \boldsymbol{B}$, so that $\boldsymbol{x}_0 \notin \operatorname{supp}(h\,\varphi)$. In particular, $\operatorname{supp}(h\,\varphi) \subseteq \operatorname{supp}\varphi$.

Lemma 5.8 allows us to introduce the following definition.

Definition 5.14 [Multiplication of a Distribution by a Function of Class C^∞]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $h \in C^\infty(\boldsymbol{A};\mathbb{C})$. We define

$$\forall \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C}), \qquad \langle T\,h, \varphi\rangle \stackrel{def}{=} \langle T, h\,\varphi\rangle.$$

The only fact that we need to prove for this definition to be well-posed is that $T\,h$ as defined above is linear and continuous. Proving linearity is trivial (try!). We now focus on the continuity. We need to show that if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ which converges to $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ in the sense of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, then, given a function h of class C^∞ we have $(T\,h)(\varphi_k) \xrightarrow{k\to\infty} (T\,h)(\varphi)$. If we show that the convergence of $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{D}(A;\mathbb{C})$ implies the convergence of $(h\,\varphi_k)_{k\in\mathbb{N}}$ to $h\,\varphi$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, the solution is trivial:

$$\lim_{k\to\infty} \langle h\,T, \varphi_k\rangle = \lim_{k\to\infty} \langle T, h\,\varphi_k\rangle = \langle T, h\,\varphi\rangle = \langle T\,h, \varphi\rangle.$$

Thus, we only need to prove the following result.

[5] We cannot weaken this requirement to only ask h to be of class C^∞ on the support of φ, because the definition involves the arbitrariness of φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, and a bump function can be shifted to any point of the domain.

Lemma 5.9 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $h \in C^\infty(\boldsymbol{A};\mathbb{C})$. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then*

$$\varphi_k \xrightarrow{k\to\infty} \varphi \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad h\,\varphi_k \xrightarrow{k\to\infty} h\,\varphi \text{ in } \mathcal{D}(\boldsymbol{A};\mathbb{C}).$$

Proof First, for all $k \in \mathbb{N}$, the supports of $h\,\varphi_k$ are contained in a single compact set $\boldsymbol{K}$, because this is true for each φ_k. Indeed, since $(\varphi_k)_{k\in\mathbb{N}}$ converge to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, there exists a compact set $\boldsymbol{K}$ contained in $\boldsymbol{A}$ such that $\operatorname{supp}\varphi_k \subseteq \boldsymbol{K}$ for all $k \in \mathbb{N}$, so that also $\operatorname{supp}(h\,\varphi_k) \subseteq \boldsymbol{K}$ for all $k \in \mathbb{N}$.

We now need to show that $\partial^\alpha(h\,\varphi_k) \xrightarrow{k\to\infty} \partial^\alpha(h\,\varphi)$ uniformly for any $\boldsymbol{\alpha} \in \mathbb{N}^n$. We use Leibniz formula (Lemma 5.4) to see that given $\boldsymbol{\alpha} \in \mathbb{N}^n$

$$\begin{aligned}
|\partial^\alpha(h\,\varphi_k) - \partial^\alpha(h\,\varphi)| &= \left|\sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \partial^{\alpha-\beta}\,h(\boldsymbol{x})\,(\partial^\beta\varphi_k(\boldsymbol{x}) - \partial^\beta\varphi(\boldsymbol{x}))\right| \\
&\le \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \left|\partial^{\alpha-\beta}\,h(\boldsymbol{x})\right| \left|\partial^\beta\varphi_k(\boldsymbol{x}) - \partial^\beta\varphi(\boldsymbol{x})\right| \\
&\le \sum_{\boldsymbol{\beta}\le\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} \max_{\boldsymbol{x}\in\boldsymbol{K}} \left|\partial^{\alpha-\beta}\,h(\boldsymbol{x})\right| \left|\partial^\beta\varphi_k(\boldsymbol{x}) - \partial^\beta\varphi(\boldsymbol{x})\right|,
\end{aligned}$$

where we have used the fact that the supports of the functions φ_k and that of φ are all contained in $\boldsymbol{K}$. Thus, $\binom{\alpha}{\beta} \max_{\boldsymbol{x}\in\boldsymbol{K}} \left|\partial^{\alpha-\beta}\,h(\boldsymbol{x})\right|$ are constant terms, whereas $\left|\partial^\beta\varphi_k(\boldsymbol{x}) - \partial^\beta\varphi(\boldsymbol{x})\right|$ tends to zero uniformly.

Example 5.21 Let $h \in C^\infty(\mathbb{R}^n;\mathbb{R})$. Then, $h\,\delta = h(0)\,\delta$. Indeed, for all $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ we have

$$\langle h\,\delta, \varphi\rangle = \langle \delta, h\,\varphi\rangle = h(0)\,\varphi(0) = \langle h(0)\,\delta, \varphi\rangle.$$

Example 5.22 We recall that for all $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ we have defined

$$\left\langle \operatorname{pv}\frac{1}{x}, \varphi\right\rangle = \lim_{\varepsilon\to 0^+} \int_{|x|\ge\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x.$$

We want to verify that

$$x\,\operatorname{pv}\frac{1}{x} = 1,$$

where by 1 we mean the function-type distribution associated with the constant function equal to 1 (whose equivalence class is obviously in $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ because $\int_K 1\,\mathrm{d}x = \mu_1(K) < +\infty$ for any compact set $K \subset \mathbb{R}$):

$$\langle L_1, \varphi\rangle = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x.$$

Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\left\langle x\,\operatorname{pv}\frac{1}{x}, \varphi\right\rangle = \left\langle \operatorname{pv}\frac{1}{x}, x\,\varphi\right\rangle = \lim_{\varepsilon\to 0^+} \int_{|x|\ge\varepsilon} \frac{x\,\varphi(x)}{x}\,\mathrm{d}x.$$

We want to show that

$$\lim_{\varepsilon\to 0^+} \int_{|x|\ge\varepsilon} \varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}x.$$

To this end, we write

$$\lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon}\varphi(x)\,\mathrm{d}x = \lim_{k\to\infty}\int_{|x|\geq\frac{1}{k}}\varphi(x)\,\mathrm{d}x = \lim_{k\to\infty}\int_{\mathbb{R}}\varphi(x)\,\chi_{(-\infty,-\frac{1}{k}]\cup[\frac{1}{k},+\infty]}(x)\,\mathrm{d}x. \tag{5.27}$$

If we can bring the limit inside the integral we have concluded. To this end, we use Lebesgue's dominated convergence theorem (Theorem 3.47). For all $x\in\mathbb{R}$,

$$\lim_{k\to\infty}\chi_{(-\infty,-\frac{1}{k}]\cup[\frac{1}{k},+\infty)]}(x)\,\varphi(x) = \begin{cases}\varphi(x) & \text{if } x\neq 0\\ 0 & \text{if } x=0.\end{cases}$$

The function is therefore bounded in a neighbourhood of the origin. For all $k\in\mathbb{N}^*$ we have $|\chi_{(-\infty,-\frac{1}{k}]\cup[\frac{1}{k},+\infty)]}(x)\,\varphi(x)|\leq|\varphi(x)|$ for all $x\in\mathbb{R}$, and $|\varphi(x)|$ is continuous and bounded, and therefore summable. This allows the interchange of the limit with the integral, and this leads to the result.

We now observe that the multiplication by functions in C^∞ is continuous with respect to the convergence in $\mathcal{D}'$.

Theorem 5.14 *Let $A\subseteq\mathbb{R}^n$ be open. Let $h\in C^\infty(A;\mathbb{C})$. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(A;\mathbb{C})$ and let $T\in\mathcal{D}'(A;\mathbb{C})$. Then,*

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(A;\mathbb{C}) \quad\Rightarrow\quad h\,T_k \xrightarrow{k\to\infty} h\,T \text{ in } \mathcal{D}'(A;\mathbb{C}).$$

The proof is left as an exercise, see Question **??**.

Example 5.23

We can give another proof of the fact that $\delta\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ is a singular distribution. Suppose again by contradiction that $f\in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ is such that $\langle L_f,\varphi\rangle=\varphi(0)$ for all $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Let us now consider the linear form π_1 that maps the vector $\boldsymbol{x}=(x_1,x_2,\dots,x_n)$ to its first component x_1, i.e.,

$$\begin{aligned}\pi_1:\mathbb{R}^n &\longrightarrow \mathbb{R}\\ \boldsymbol{x} &\mapsto x_1.\end{aligned}$$

This function is C^∞, and therefore

$$\langle f,x_1\,\varphi\rangle = \int_{\mathbb{R}} f(\boldsymbol{x})\,x_1\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = x_1\,\varphi(\boldsymbol{x})|_{x=0} = 0 = \langle\pi_1 f,\varphi\rangle.$$

From the fundamental lemma 5.5, $\pi_1 f$ is zero almost everywhere, and therefore f is zero almost everywhere (why?), so that $\langle L_f,\varphi\rangle=\varphi(0)$ cannot hold for a test function φ with $\varphi(0)\neq 0$.

What we have seen begs an important question: can we define the product of two distributions? The answer is, in general, negative. Even the product of locally summable functions is not, in general, locally summable: think for example about $(|x|^{-\frac{1}{2}})^2=|x|^{-1}$ in $\mathbb{R}$. More is true: it is possible to prove that it is impossible to define a product of distributions which is associative and commutative. Indeed, recalling that $x\ \mathrm{pv}\,\frac{1}{x}$, we would obtain

$$0 = 0\cdot\mathrm{pv}\,\frac{1}{x} = (x\,\delta(x))\cdot\mathrm{pv}\,\frac{1}{x} = (\delta(x)\,x)\cdot\mathrm{pv}\,\frac{1}{x} = \delta(x)\cdot\left(x\ \mathrm{pv}\,\frac{1}{x}\right) = \delta(x).$$

Loosely speaking, the product of distributions T and S can only be defined in such a way that if T is singular in the neighbourhood of any point, then S must be regular in this neighbourhood. For example, it is possible to define the product $\delta(x-1)\,\delta(x-2)$ (and this product is zero).

Exercises

Exercise 5.25 Let $A \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(A;\mathbb{C})$. Let $h \in C^\infty(A;\mathbb{C})$. Show that the functional $T\,h$ is linear. [Solution: (5.25)]

Exercise 5.26 Let $A \subseteq \mathbb{R}^n$ be open. Let $h \in C^\infty(A;\mathbb{C})$. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(A;\mathbb{C})$ and let $T \in \mathcal{D}'(A;\mathbb{C})$. Show that

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(A;\mathbb{C}) \qquad \Rightarrow \qquad h\,T_k \xrightarrow{k\to\infty} h\,T \text{ in } \mathcal{D}'(A;\mathbb{C}).$$

[Solution: (5.26)]

5.5.2 Composition of distributions with diffeomorphisms in C^∞

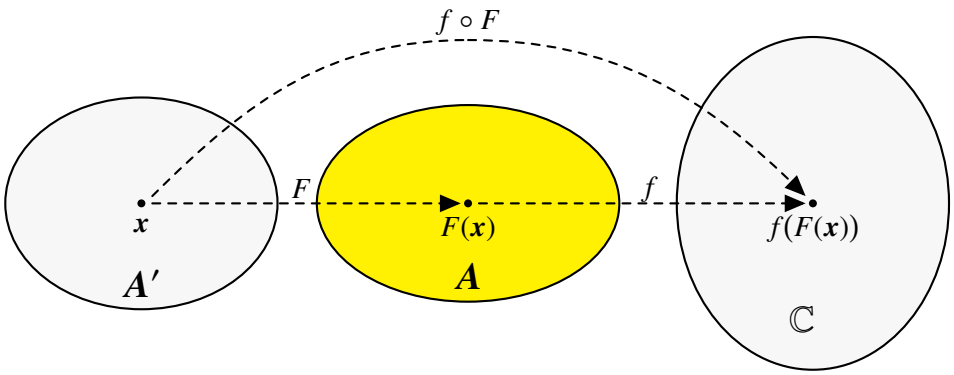
One of the shortcomings of the theory of distributions is the fact that there is no natural definition for the product of two distributions (nor for the composition of two distributions) which enjoy nice properties with respect to the notion of convergence in the sense of distributions.

In this section we show how the composition can be defined under some restrictive assumptions. We will not be able to consider the composition $F \circ T$, where F is a function and T is a distribution. However, we will be able to define the composition $T \circ F$; this gives a meaning, for example, to the expressions $\delta(2\,x)$ or $\delta(x-1)$, which can be indeed viewed as the composition of δ with certain functions: in these cases, these functions are

$$\begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & 2\,x \end{array} \qquad \text{and} \qquad \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & x-1, \end{array}$$

respectively. We begin, as usual, considering what happens if in place of a distribution T we are dealing with a function f. More precisely, we deal with the problem of defining the composition $f \circ F$, when $f : A \longrightarrow \mathbb{C}$ and $F : A' \longrightarrow A$, so that $f \circ F : A' \longrightarrow \mathbb{C}$, and

$$\forall\, \boldsymbol{x} \in A', \qquad (f \circ F)(\boldsymbol{x}) = f(F(\boldsymbol{x})).$$



We recall that if F is differentiable at $\boldsymbol{x}$ and f is differentiable at $F(\boldsymbol{x})$, then $f \circ F$ is differentiable at $\boldsymbol{x}$ and

$$\mathrm{d}(f \circ F)(\boldsymbol{x}) = \mathrm{d}f(F(\boldsymbol{x})) \circ \mathrm{d}F(\boldsymbol{x}).$$

This is the generalization of the chain rule to the multivariable case. This is an identity involving linear transformations; we can re-write it in terms of the corresponding Jacobian matrices:

$$J_{f\circ F}(\boldsymbol{x}) = J_f(F(\boldsymbol{x}))\, J_F(\boldsymbol{x}).$$

We now wish to express the integral

$$\int_{A'} (f \circ F)(\boldsymbol{x})\, \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

as an integral over A. Clearly, the tool that we can use to this purpose is the change of variable, and this will impose some first restrictions on F. First, we need A' to be an open set of the same space $\mathbb{R}^n$. Moreover, it is convenient to assume F to be of class C^1 and invertible, and with its inverse of class C^1. In other words, we require F to be a diffeomorphism.[6] Under these assumptions, and with reasonable assumptions over f, we can use the change of variable to turn the integral over $\boldsymbol{A}'$ into an integral over $\boldsymbol{A}$ as follows:

$$\int_{A'} (f \circ F)(\boldsymbol{x})\, \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{A'} f(F(\boldsymbol{x}))\, \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \overset{\boldsymbol{y}=F(\boldsymbol{x})}{=} \int_{A} f(\boldsymbol{y})\, \varphi(F^{-1}(\boldsymbol{y}))\, |\det J_{F^{-1}}(\boldsymbol{y})|\, \mathrm{d}\boldsymbol{y},$$

where $J_{F^{-1}}$ denotes the Jacobian matrix of F^{-1}. Requiring F^{-1} to be of class C^1 is equivalent to requiring the determinant of the Jacobian matrix J_F to be different from zero at every point. Indeed, if F is a diffeomorphism, then it is invertible, and therefore for all $\boldsymbol{x} \in \boldsymbol{A}'$ we have $(F^{-1} \circ F)(\boldsymbol{x}) = \boldsymbol{x}$. This identity can also be written as $F^{-1} \circ F = I_{A'}$, where $I_{A'}$ is the identity function that maps $\boldsymbol{x} \in \boldsymbol{A}'$ into itself. Since both F and F^{-1} are of class C^1, we can apply the differential of the composition (chain rule) recalled above to the expression $F^{-1} \circ F = I_{A'}$, or, equivalently, to

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}', \qquad (F^{-1} \circ F)(\boldsymbol{x}) = \boldsymbol{x},$$

taking into account that the derivative as a linear transformation of $I_{A'}$ is the identity function in $\mathbb{R}^n$, denoted by $I_{\mathbb{R}^n}$, we obtain

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}', \qquad (F^{-1})'(F(\boldsymbol{x})) \circ F'(\boldsymbol{x}) = I_{\mathbb{R}^n}.$$

In terms of the corresponding Jacobian matrices, we get

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}', \qquad J_{F^{-1}}(F(\boldsymbol{x}))\, J_F(\boldsymbol{x}) = I_n,$$

which says that the Jacobian of F is invertible for all $\boldsymbol{x} \in \boldsymbol{A}'$ (and therefore so is the Jacobian of F^{-1} for all $\boldsymbol{y} \in \boldsymbol{A}$), and, moreover,

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}', \quad J_{F^{-1}}(F(\boldsymbol{x})) = J_F^{-1}(\boldsymbol{x}).$$

[6] Given two open subsets $\boldsymbol{A}$ and $\boldsymbol{B}$ of $\mathbb{R}^n$, the function $f : gA \longrightarrow \boldsymbol{B}$ is a diffeomorphism if

(i) f is injective and surjective, i.e., it is invertible;

(ii) f is of class C^1;

(iii) f^{-1} is of class C^1.

For diffeomorphisms, we have a formula on the derivative of the inverse function which generalises the usual single variable one. In particular, if $f : \boldsymbol{A} \longrightarrow \boldsymbol{B}$ is a diffeomorphism, for every $\boldsymbol{x} \in \boldsymbol{A}$ we have:

- the linear transformation $f'(\boldsymbol{x})$ is invertible (or, equivalently, $J_{f^{-1}}(\boldsymbol{x})$ is invertible);
- $(f^{-1})'(f(\boldsymbol{x})) \circ f'(\boldsymbol{x}) = I_A$, where I_A is the identity linear transformation (or, equivalently, $J_{f^{-1}}(f(\boldsymbol{x})) = (J_f(\boldsymbol{x}))^{-1}$)

To ensure that $f \circ F$ is a distribution, we must ensure that the function which multiplies $f(\boldsymbol{y})$ in the last integral is a function in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, and for this to happen we need other restrictions on F. This request is fulfilled if both F and F^{-1} are of class C^∞. Hence, we will assume that $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ is a diffeomorphism of class C^∞, and this can be summarized by the following requirements:

(i) F is invertible;

(ii) F is of class C^∞;

(iii) F^{-1} is of class C^∞.

Our aim is now to make these considerations more precise. We begin by showing that under these assumptions, if $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$, then $[f \circ F] \in L^1_{\text{loc}}(\boldsymbol{A}';\mathbb{C})$.

We begin with a lemma that you may have seen before (it is a result that is used in the proof of Weierstrass extreme value theorem): the continuous image of a compact set is compact.

Theorem 5.15 *Let $\boldsymbol{A}, \boldsymbol{A}' \subseteq \mathbb{R}^n$ be open. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be continuous and let $\boldsymbol{K}$ be a compact subset of $\boldsymbol{A}'$. Then, $F(\boldsymbol{K})$ is a compact subset of A.*

Proof To prove that $F(\boldsymbol{K})$ is compact, we take a sequence $(\boldsymbol{b}_k)_{k\in\mathbb{N}}$ of $F(\boldsymbol{K})$ and we verify that it admits a subsequence $(\boldsymbol{b}_{k_j})_{j\in\mathbb{N}}$ that converges in $F(\boldsymbol{K})$. Since for all $k \in \mathbb{N}$ we have $\boldsymbol{b}_k \in F(\boldsymbol{K})$, for every $k \in \mathbb{N}$, there exists $\boldsymbol{a}_k \in \boldsymbol{K}$ such that $F(\boldsymbol{a}_k) = \boldsymbol{b}_k$. We have thus constructed a sequence $(\boldsymbol{a}_k)_{k\in\mathbb{N}}$ of points in $\boldsymbol{K}$. Since $\boldsymbol{K}$ is compact, there exists a subsequence $(\boldsymbol{a}_{k_j})_{j\in\mathbb{N}}$ of $(\boldsymbol{a}_k)_{k\in\mathbb{N}}$ which is convergent in $\boldsymbol{K}$. Since F is continuous, the implication

$$\boldsymbol{a}_{k_j} \xrightarrow{j\to\infty} \boldsymbol{\ell} \in \boldsymbol{K} \qquad \Rightarrow \qquad F(\boldsymbol{a}_{k_j}) \xrightarrow{j\to\infty} F(\boldsymbol{\ell}) \in F(\boldsymbol{K})$$

holds. Recalling that $F(\boldsymbol{a}_k) = \boldsymbol{b}_k$ for every $k \in \mathbb{N}$, we also have

$$\boldsymbol{b}_{k_j} \xrightarrow{j\to\infty} F(\boldsymbol{\ell}) \in F(\boldsymbol{K}).$$

This says that the subsequence $(\boldsymbol{b}_{k_j})_{j\in\mathbb{N}}$ converges in $F(\boldsymbol{K})$, which is what we had to prove.

Lemma 5.10 *Let $\boldsymbol{A}, \boldsymbol{A}' \subseteq \mathbb{R}^n$ be open. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{g}A$ be a diffeomorphism of class C^∞. Let $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$. Then,*

$$[f \circ F] \in L^1_{\text{loc}}(\boldsymbol{A}';\mathbb{C})$$

Proof We need to show that $f \circ F$ is integrable in any compact subset of $\boldsymbol{A}'$. Let $\boldsymbol{K}$ be a compact set contained in $\boldsymbol{A}'$. We must show that the integral

$$\int_{\boldsymbol{K}} \left|f(F(\boldsymbol{x}))\right| \mathrm{d}\boldsymbol{x} \overset{\boldsymbol{y}=F(\boldsymbol{x})}{=} \int_{F(\boldsymbol{K})} |f(\boldsymbol{y})|\,|\det J_{F^{-1}}(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}$$

is finite, where we have changed variable with $F(\boldsymbol{x}) = \boldsymbol{y}$. Since continuous functions map compact sets into compact sets, $F(\boldsymbol{K})$ is a compact set of $\boldsymbol{A}$, and the mapping

$$\begin{aligned} \boldsymbol{A} &\longrightarrow \mathbb{R}_+ \\ \boldsymbol{y} &\mapsto |\det J_{F^{-1}}(\boldsymbol{y})| \end{aligned}$$

is continuous because its terms are sums and products of first order partial derivatives. Thus, by the Weierstrass

extreme value theorem there exists $M > 0$ such that $|\det J_{F^{-1}}(\boldsymbol{y})| \leq M$ for every $\boldsymbol{y} \in F(K)$. Hence,

$$\int_{F(K)} |f(\boldsymbol{y})|\,|\det J_{F^{-1}}(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \leq M \int_{F(K)} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} < \infty.$$

Thus, $[f \circ F] \in L^1_{\mathrm{loc}}(\boldsymbol{A}'; \mathbb{C})$.

From the previous theorem, $f \circ F$ can be viewed as a function-type distribution. Let us see how it operates, before we extend its definition to distributions in $\mathcal{D}'$. Let us consider $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$. By applying the theorem of change of variable we find

$$\begin{aligned}\langle L_{f\circ F}, \varphi\rangle &= \int_{\boldsymbol{A}'} f(F(\boldsymbol{x}))\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \overset{\boldsymbol{y}=F(\boldsymbol{x})}{=} \int_{\boldsymbol{A}} f(\boldsymbol{y})\,\varphi(F^{-1}(\boldsymbol{y}))\,|\det J_{F^{-1}}(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} \\ &= \langle L_f, (\varphi \circ F^{-1})\,|\det J_{F^{-1}}|\rangle. \end{aligned} \tag{5.28}$$

If we want the latter integral to define a distribution in $\mathcal{D}'(\boldsymbol{A}'; \mathbb{C})$, we need to show that the function

$$\begin{aligned} \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{y} &\mapsto \varphi(F^{-1}(\boldsymbol{y}))\,|\det \boldsymbol{J}_{F^{-1}}(\boldsymbol{y})| \end{aligned} \tag{5.29}$$

is in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$. Since φ and F^{-1} are both of class C^∞, we have $\varphi \circ F^{-1} \in C^\infty(\boldsymbol{A}; \mathbb{C})$. The function $\det \boldsymbol{J}_{F^{-1}}$, i.e., the mapping

$$\begin{aligned} \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{y} &\mapsto \det J_{F^{-1}}(\boldsymbol{y}) \end{aligned} \tag{5.30}$$

is also of class C^∞. In general, the absolute value of a function of class C^∞ is not of class C^∞ because of the singularity at zero. However, since as aforementioned $\det J_{F^{-1}}$ is never zero under our assumptions, then

$$\boldsymbol{y} \mapsto |\det J_{F^{-1}}(\boldsymbol{y})|$$

is of class C^∞. Thus, (5.29) is of class C^∞. We show that it has compact support. If $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$, where $\boldsymbol{K}$ is a compact set of $\boldsymbol{A}'$, then

$$\boldsymbol{x} \notin \boldsymbol{K} \quad \Rightarrow \quad \varphi(\boldsymbol{x}) = 0,$$

or, with the definition $\boldsymbol{y} = F(\boldsymbol{x})$,

$$F^{-1}(\boldsymbol{y}) \notin \boldsymbol{K} \quad \Rightarrow \quad \varphi(F^{-1}(\boldsymbol{y})) = 0,$$

i.e.,

$$\boldsymbol{y} \notin F(\boldsymbol{K}) \quad \Rightarrow \quad \varphi(F^{-1}(\boldsymbol{y})) = 0.$$

Thus, $\operatorname{supp}(\varphi \circ F^{-1}) \subseteq F(\boldsymbol{K})$. As recalled above, $F(\boldsymbol{K})$ is compact. Hence, the last expression in (5.28) is well defined. This suggests the following definition.

Definition 5.15 [Composition of Distributions with Diffeomorphisms in C^∞]
Let $\boldsymbol{A}, \boldsymbol{A}'$ be open sets in $\mathbb{R}^n$. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ and $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$. We define

$$\langle T \circ F, \varphi\rangle \overset{def}{=} \langle T, (\varphi \circ F^{-1})\,|\det J_{F^{-1}}|\rangle. \tag{5.31}$$

Notice that in the case $n = 1$, (5.31) can be written as

$$\langle T \circ F, \varphi \rangle \stackrel{\text{def}}{=} \langle T, (\varphi \circ F^{-1})\,|(F^{-1})'| \rangle. \tag{5.32}$$

The last fact that we need to prove is that the functional $T \circ F$ that maps the right hand-side of (5.31) to each $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$ is indeed a distribution in $\boldsymbol{A}'$. Before we prove this fact, we have to introduce the following preliminary lemma.

Lemma 5.11 [Faà di Bruno's Formula]
Let $\boldsymbol{A}, \boldsymbol{B} \subseteq \mathbb{R}^n$ be open. Let $g : \boldsymbol{A} \longrightarrow \boldsymbol{B}$ and $f : \boldsymbol{B} \longrightarrow \mathbb{C}$ be of class C^∞. Then, for every $\boldsymbol{\alpha} \in \mathbb{N}^n$

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}}(f \circ g)(\boldsymbol{x}) = \sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})\, \frac{\partial^{|\boldsymbol{\beta}|} f}{\partial \boldsymbol{x}^{\boldsymbol{\beta}}}(g(\boldsymbol{x})), \tag{5.33}$$

where $d_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in C^\infty(\boldsymbol{A}; \mathbb{C})$ depend only on g.

Proof We can reason by induction on $|\boldsymbol{\alpha}|$. If $|\boldsymbol{\alpha}| = 1$, say $\boldsymbol{\alpha} = (0, 0, \ldots, 1, 0, \ldots, 0)$, where the 1 is in position k (so that we can write $\boldsymbol{\alpha} = \boldsymbol{e}_k$). Using the chain rule we find

$$\frac{\partial}{\partial x_k}(f \circ g)(\boldsymbol{x}) = \frac{\partial f}{\partial x_k}(g(\boldsymbol{x}))\, \frac{\partial}{\partial x_k} g(\boldsymbol{x}).$$

Let us compare this expression with (5.33), whose sum in the case $\boldsymbol{\alpha} = \boldsymbol{e}_k$ has only two terms: $\boldsymbol{\beta}$ equal to the zero vector, and $\boldsymbol{\beta} = \boldsymbol{\alpha} = \boldsymbol{e}_k$. Thus,

$$\frac{\partial^{|\boldsymbol{e}_k|}}{\partial \boldsymbol{x}^{\boldsymbol{e}_k}}(f \circ g)(\boldsymbol{x}) = \frac{\partial}{\partial x_k}(f \circ g)(\boldsymbol{x}) = d_{\boldsymbol{\alpha},0}\, f(g(\boldsymbol{x})) + d_{\boldsymbol{\alpha},\boldsymbol{\alpha}}\, \frac{\partial f}{\partial x_k}(g(\boldsymbol{x})).$$

These two expressions are consistent by taking $d_{\boldsymbol{\alpha},0} = 0$ and $d_{\boldsymbol{\alpha},\boldsymbol{\alpha}} = \frac{\partial}{\partial x_k} g(\boldsymbol{x})$. This shows that the statement is true for $|\boldsymbol{\alpha}| = 1$. Let us assume that the statement is true for $\boldsymbol{\alpha}$, and let us prove it for $\boldsymbol{\alpha} + \boldsymbol{e}_j$ (note that $|\boldsymbol{\alpha} + \boldsymbol{e}_j| = |\boldsymbol{\alpha}| + 1$). We have

$$\begin{aligned}
\frac{\partial^{|\boldsymbol{\alpha}|+1}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}+\boldsymbol{e}_j}}(f \circ g)(\boldsymbol{x}) &= \frac{\partial}{\partial x_j}\, \frac{\partial^{|\boldsymbol{\alpha}|}}{\partial \boldsymbol{x}^{\boldsymbol{\alpha}}} f(g(\boldsymbol{x})) \\
&= \frac{\partial}{\partial x_j} \left(\sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})\, \frac{\partial^{|\boldsymbol{\beta}|} f}{\partial \boldsymbol{x}^{\boldsymbol{\beta}}}(g(\boldsymbol{x})) \right) \\
&= \sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} \left(\frac{\partial d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})}{\partial x_j}\, \frac{\partial^{|\boldsymbol{\beta}|} f}{\partial \boldsymbol{x}^{\boldsymbol{\beta}}}(g(\boldsymbol{x})) + d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})\, \frac{\partial^{|\boldsymbol{\beta}|+1} f}{\partial \boldsymbol{x}^{\boldsymbol{\beta}+\boldsymbol{e}_j}}(g(\boldsymbol{x}))\, \frac{\partial g(\boldsymbol{x})}{\partial x_j} \right).
\end{aligned}$$

We can write this last summation as

$$\sum_{\boldsymbol{b} \le \boldsymbol{\alpha}+\boldsymbol{e}_j} \hat{d}_{\boldsymbol{\alpha}+\boldsymbol{e}_j,\boldsymbol{b}}(\boldsymbol{x})\, \frac{\partial^{|\boldsymbol{b}|} f}{\partial \boldsymbol{x}^{\boldsymbol{b}}}(g(\boldsymbol{x})),$$

where

$$\hat{d}_{\boldsymbol{\alpha}+\boldsymbol{e}_j,\boldsymbol{b}}(\boldsymbol{x}) = \begin{cases} \frac{\partial d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})}{\partial x_j} & \text{if } \boldsymbol{b} \le \boldsymbol{\alpha} \\ d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{x})\, \frac{\partial g(\boldsymbol{x})}{\partial x_j} & \text{if } \boldsymbol{b} = \boldsymbol{\alpha} + \boldsymbol{e}_j \end{cases}$$

Thus, we re-obtained (5.33). Notice that the coefficients depend only on g, and from this dependence we find that these functions are all of class C^∞, because such is g.

Theorem 5.16 *Let $\boldsymbol{A}, \boldsymbol{A}' \subseteq \mathbb{R}^n$ be open. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$. The composition $T \circ F$ defined in (5.31) is linear and continuous.*

Proof The proof of linearity is left as a tutorial exercise. As a hint, start by verifying that given three functions f, g, h with domains and codomains defined appropriately and for a constant k we have

$$(f + g) \circ h = (f \circ h) + (g \circ h) \qquad \text{and} \qquad (k\, f) \circ g = k\,(f \circ g).$$

We now prove that $T \circ F : \mathcal{D}(\boldsymbol{A}'; \mathbb{C}) \longrightarrow \mathbb{C}$ as defined in (5.31) is continuous. Consider a sequence $(\varphi_k)_{k \in \mathbb{N}}$ in $\mathcal{D}(\boldsymbol{A}'; \mathbb{C})$ that converges to $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$. We need to prove that

$$\langle T \circ F, \varphi_k \rangle \xrightarrow{k \to \infty} \langle T \circ F, \varphi \rangle,$$

i.e., that

$$\langle T, (\varphi_k \circ F^{-1}) \,|\det J_{F^{-1}}| \rangle \xrightarrow{k \to \infty} \langle T, (\varphi \circ F^{-1}) \,|\det J_{F^{-1}}| \rangle,$$

which follows from the definition of T as an element of $\mathcal{D}'(\boldsymbol{A}'; \mathbb{C})$, if we prove that

$$\varphi_k \xrightarrow{k \to \infty} \varphi \text{ in } \mathcal{D}(\boldsymbol{A}'; \mathbb{C}) \quad \Rightarrow \quad (\varphi_k \circ F^{-1}) \,|\det J_{F^{-1}}| \xrightarrow{k \to \infty} (\varphi \circ F^{-1}) \,|\det J_{F^{-1}}| \text{ in } \mathcal{D}(\boldsymbol{A}; \mathbb{C}),$$

i.e., that

$$\varphi_k \xrightarrow{k \to \infty} \varphi \text{ in } \mathcal{D}(\boldsymbol{A}'; \mathbb{C}) \quad \Rightarrow \quad \varphi_k \circ F^{-1} \xrightarrow{k \to \infty} \varphi \circ F^{-1} \text{ in } \mathcal{D}(\boldsymbol{A}; \mathbb{C}),$$

Since $(\varphi_k)_{k \in \mathbb{N}}$ in $\mathcal{D}(\boldsymbol{A}'; \mathbb{C})$ converges to $\varphi \in \mathcal{D}(\boldsymbol{A}'; \mathbb{C})$, there exists a compact set $\boldsymbol{K}$ of $\boldsymbol{A}'$ such that $\operatorname{supp}(\varphi_k) \subseteq \boldsymbol{K}$ for all $k \in \mathbb{N}$. As we have seen before, this says that $\boldsymbol{x} \notin \boldsymbol{K}$ implies $\varphi(\boldsymbol{x}) = 0$, which is equivalent to saying that $F^{-1}(\boldsymbol{y}) \notin \boldsymbol{K}$ implies $\varphi(F^{-1}(\boldsymbol{y})) = 0$, which is in turn equivalent to saying that $\boldsymbol{y} \notin F(\boldsymbol{K})$ implies $(\varphi \circ F^{-1})(\boldsymbol{y}) = 0$. Therefore, $\operatorname{supp}(\varphi_k \circ F^{-1}) \subseteq F(\boldsymbol{K})$, and $F(\boldsymbol{K})$ is compact in $\boldsymbol{A}$.

Now we have to prove the uniform convergence of the partial derivatives of every order, by using the previous lemma. Given $\boldsymbol{\alpha} \in \mathbb{N}^n$, for all $\boldsymbol{y} \in \boldsymbol{A}$ there holds

$$\partial^{\boldsymbol{\alpha}}(\varphi \circ F^{-1})(\boldsymbol{y}) = \sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{y})\, \partial^{\boldsymbol{\beta}} \varphi(F^{-1}(\boldsymbol{y})),$$

where $d_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ is of class C^∞ and only depends on F^{-1}. Then, using the result of the previous lemma, we find

$$\begin{aligned} |\partial^{\boldsymbol{\alpha}}(\varphi_k \circ F^{-1})(\boldsymbol{y}) - \partial^{\boldsymbol{\alpha}}(\varphi \circ F^{-1})(\boldsymbol{y})| &= \Big| \sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} d_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\boldsymbol{y}) \Big((\partial^{\boldsymbol{\beta}} \varphi_k)(F^{-1}(\boldsymbol{y})) - (\partial^{\boldsymbol{\beta}} \varphi)(F^{-1}(\boldsymbol{y})) \Big) \Big| \\ &\le \sum_{\boldsymbol{\beta} \le \boldsymbol{\alpha}} \sup_{F(\boldsymbol{K})} |d_{\boldsymbol{\alpha},\boldsymbol{\beta}}| \sup_{\boldsymbol{A}'} \big| \partial^{\boldsymbol{\beta}} \varphi_k - \partial^{\boldsymbol{\beta}} \varphi \big|, \end{aligned}$$

which tends to zero because $|\partial^{\boldsymbol{\beta}} \varphi_k - \partial^{\boldsymbol{\beta}} \varphi|$ tends uniformly to zero, since φ_k converges uniformly to φ in $\mathcal{D}(\boldsymbol{A}'; \mathbb{C})$. The number of addends in the summation is finite, and therefore the sum tends to 0 uniformly.

The next result shows that if T_k converges to T in $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$, then $T_k \circ F$ converges to $T \circ F$ in $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$.

Theorem 5.17 *Let $\boldsymbol{A}, \boldsymbol{A}'$ be open sets in $\mathbb{R}^n$. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ and let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Then,*

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A}; \mathbb{C}) \qquad \Rightarrow \qquad T_k \circ F \xrightarrow{k\to\infty} T \circ F \text{ in } \mathcal{D}'(\boldsymbol{A}'; \mathbb{C}).$$

The proof is left as an exercise (see Question **??**).

Example 5.24

Consider $\delta \in \mathcal{D}'(\mathbb{R}; \mathbb{R})$. Let $F : \mathbb{R} \longrightarrow \mathbb{R}$ be invertible, and such that both F and F^{-1} are of class C^∞. Thus, both F and F^{-1} are strictly monotonic, and $F^{-1}(x) \neq 0$ for all $x \in \mathbb{R}$. Let $\varphi \in \mathcal{D}(\mathbb{R}; \mathbb{R})$. Then,

$$\begin{aligned} \langle \delta \circ F, \varphi \rangle &= \langle \delta, (\varphi \circ F^{-1}) \, |(F^{-1})'| \rangle \\ &= \varphi(F^{-1}(0)) \, |(F^{-1})'(0)| \end{aligned}$$

Consider for example $F(x) = 2\,x$, which is clearly a diffeomorphism in C^∞, and $F^{-1}(y) = y/2$, so that $(F^{-1})'(0) = 1/2$. Then,

$$\langle \delta \circ F, \varphi \rangle = \varphi(F^{-1}(0)) \, |(F^{-1})'(0)| = \frac{1}{2}\,\varphi(0),$$

where with some abuse of notation we have denoted the distribution $\delta \circ F$ as $\delta(2\,x)$. With the same abuse, we can write the expression

$$\delta(2\,x) = \frac{1}{2}\delta(x),$$

which often replaces the most precise identity $\delta \circ F = \frac{1}{2}\,\delta$.

Example 5.25

Consider $\delta \in \mathcal{D}'(\mathbb{R}; \mathbb{R})$. Consider the diffeomorphism $F(x) = x - a$ where $a \in \mathbb{R}$, so that $F^{-1}(y) = y + a$ and $(F^{-1})'(0) = 1$, and

$$\begin{aligned} \langle \delta \circ F, \varphi \rangle &= \langle \delta, (\varphi \circ F^{-1}) \, |(F^{-1})'| \rangle \\ &= \varphi(F^{-1}(0)) \, |(F^{-1})'(0)| = \varphi(a). \end{aligned}$$

Example 5.26

The "object" $\delta(|x|)$ cannot be seen as the composition of a distribution with a diffeomorphism, because the function $F(x) = |x|$ (with $x \in \mathbb{R}$) is not invertible (and also not C^∞).

Example 5.27

A lot of care needs to be taken in the way the composition of a distribution with a diffeomorphism of class C^∞ is defined. For example, consider $\lambda \neq 0$, and consider the problem of defining $\delta(e^{\lambda t})$. The function $F(x) = e^{\lambda t}$ is certainly of class C^∞, and defining its codomain appropriately it becomes a C^∞-diffeomorphism:

$$\begin{aligned} F : \mathbb{R} &\longrightarrow (0, +\infty) \\ x &\mapsto e^{\lambda\, x} \end{aligned}$$

Indeed,

$$\begin{aligned} F^{-1} : (0, +\infty) &\longrightarrow \mathbb{R} \\ y &\mapsto \tfrac{1}{\lambda} \ln y \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} (F^{-1})' : (0, +\infty) &\longrightarrow \mathbb{R} \\ y &\mapsto \tfrac{1}{\lambda y}. \end{aligned}$$

When considering Definition 5.15, however, we have $A' = \mathbb{R}$ and $A = (0, +\infty)$. Thus, $\varphi \in \mathcal{D}(\mathbb{R}; \mathbb{C})$ and the distribution must be an element of $\mathcal{D}'((0, +\infty); \mathbb{C})$. In this case, δ does not live in this set, and therefore the

object $\delta(e^{\lambda t})$ is ill-defined if we want to use Definition 5.15. Using an integral representation

$$\int_{\mathbb{R}} \delta(e^{\lambda x})\,\varphi(x)\,\mathrm{d}x = \int_{F(\mathbb{R})} \delta(y)\,\varphi\left(\frac{1}{\lambda}\,\ln y\right)\left|\frac{1}{\lambda\, y}\right|\,\mathrm{d}y = \int_0^{+\infty} \delta(y)\,\varphi\left(\frac{1}{\lambda}\,\ln y\right)\frac{1}{|\lambda|\, y}\,\mathrm{d}y,$$

it becomes even more obvious that we cannot set y to 0 in $\varphi\left(\frac{1}{\lambda}\,\ln y\right)\frac{1}{|\lambda| y}$.

The composition of a distribution with a diffeomorphism of class C^∞ can be used to introduce a number of distributions that extend some notions for functions: we will introduce radial, homogeneous, even/odd, periodic distributions. These are very useful in the context of Fourier analysis.

Exercises

Exercise 5.27

Which of the following functions is a C^∞ diffeomorphism? Can you relate this result to the Jacobian matrix?

$$f_1 : \begin{array}{ccc}(1,2) & \longrightarrow & (1,\sqrt{2}) \\ x & \mapsto & \sqrt{x}\end{array} \qquad \text{and} \qquad f_2 : \begin{array}{ccc}\mathbb{R} & \longrightarrow & \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \\ x & \mapsto & \arctan x\end{array} \qquad \text{and} \qquad f_3 : \begin{array}{ccc}\mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & x^3 - 3\,x.\end{array}$$

[Solution: (5.27)]

Exercise 5.28

Let $\boldsymbol{A}, \boldsymbol{A}' \subseteq \mathbb{R}^n$ be open. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{A}';\mathbb{C})$. Show that the composition $T \circ F$ is linear. To this end, verify that given three functions f, g, h with domains and codomains defined appropriately and for any constant k we have

$$(f+g)\circ h = (f\circ h) + (g\circ h) \qquad \text{and} \qquad (k\,f)\circ g = k\,(f\circ g).$$

[Solution: (5.28)]

Exercise 5.29

Let $\boldsymbol{A}, \boldsymbol{A}'$ be open sets in $\mathbb{R}^n$. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Show that

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \qquad \Rightarrow \qquad T_k \circ F \xrightarrow{k\to\infty} T\circ F \text{ in } \mathcal{D}'(\boldsymbol{A}';\mathbb{C}).$$

[Solution: (5.29)]

Exercise 5.30

Let $T \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. Prove that given an invertible matrix $A \in \mathbb{R}^{n\times n}$ and a vector $\boldsymbol{b} \neq 0$, we have for all $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$

$$\langle T(A\,\boldsymbol{x} - \boldsymbol{b}), \varphi(\boldsymbol{x})\rangle = \frac{1}{|\det A|}\left\langle T(\boldsymbol{y}), \varphi(A^{-1}(\boldsymbol{y}-\boldsymbol{b}))\right\rangle.$$

and use this result to find

$$\langle \delta(x_1 + 1, 2\,x_1 + 2\,x_2 + 2), \varphi(x_1, x_2)\rangle,$$

where δ is the Dirac delta in $\mathbb{R}^2$. [Solution: (5.30)]

Exercise 5.31

Determine the distribution pv $\frac{1}{2x-3}$, given by the composition of the distribution pv $\frac{1}{x}$ with the function

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 2x - 3. \end{aligned}$$

[Solution: (5.31)]

Radial distributions

We recall that a matrix $Q \in \mathbb{R}^{n\times n}$ is orthogonal if its columns are an orthonormal basis for $\mathbb{R}^n$. The following statements on a matrix $Q \in \mathbb{R}^{n\times n}$ are equivalent:

(i) Q is orthogonal;

(ii) Q is invertible and $Q^\top = Q^{-1}$;

(iii) $\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|Q\,\boldsymbol{x}\| = \|\boldsymbol{x}\|$;

(iv) $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \quad (Q\,\boldsymbol{x}) \cdot (Q\,\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$. (why?)

We also recall that the determinant of an orthogonal matrix is either 1 or -1. In fact,

$$\det(Q^{-1}\,Q) = \det I = 1 \quad \text{and} \quad \det(Q^{-1}\,Q) = \det(Q^\top Q) = \det Q^\top \,\det Q = (\det Q)^2,$$

so that $(\det Q)^2 = 1$ implies $\det Q \in \{-1, 1\}$.

Moreover, if $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ are such that $\|\boldsymbol{x}\| = \|\boldsymbol{y}\|$, then there exists an orthogonal matrix Q such that $Q\,\boldsymbol{x} = \boldsymbol{y}$.[7]

We introduce the class of *radial functions*, which are functions whose value only depends on the Euclidean norm of its argument.

Definition 5.16 [Radial Function]
Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. We say that f is radial *if*

$$\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \qquad \|\boldsymbol{x}\| = \|\boldsymbol{y}\| \;\Rightarrow\; f(\boldsymbol{x}) = f(\boldsymbol{y}).$$

It is easy to see (try!) that a function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ is radial if and only if there exists a function $\psi : [0, +\infty) \longrightarrow \mathbb{C}$ or $\psi : (0, +\infty) \longrightarrow \mathbb{C}$ such that

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad f(\boldsymbol{x}) = \psi(\|\boldsymbol{x}\|).$$

The function $\boldsymbol{x} \mapsto \|\boldsymbol{x}\|$ is not a diffeomorphism, and therefore this definition cannot be used directly to define a distributional counterpart of a radial function. We will work towards finding an alternative characterization which is easier to extend to distributions.

[7] If $\boldsymbol{x}$ and $\boldsymbol{y}$ are the zero vector, every orthogonal matrix Q will do. If they are not, we can normalise them: $\boldsymbol{x}_1 = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ and $\boldsymbol{y}_1 = \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|}$, so that $\|\boldsymbol{x}_1\| = \|\boldsymbol{y}_1\| = 1$. Using Gram-Schmidt, we can complete $\boldsymbol{x}_1$ to an orthonormal basis $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ of $\mathbb{R}^n$, and we can do the same for $\boldsymbol{y}_1$: this is always possible using the Gram-Schmidt method. In this way, we obtain two orthonormal bases $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ and $(\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n)$. Now, the matrix Q is defined by $Q\,\boldsymbol{x}_i = \boldsymbol{y}_i$ for all $i \in \{1, \ldots, n\}$, it is orthogonal, and also satisfies $Q\,\boldsymbol{x} = \boldsymbol{y}$. Indeed, defining the orthogonal matrices $X = [\ x_1 \quad x_2 \quad \ldots \quad x_n\ ]$ and $Y = [\ y_1 \quad y_2 \quad \ldots \quad y_n\ ]$, we have $Q = Y\,X^{-1} = Y\,X^\top$, and $Q^{-1} = X\,Y^{-1} = X\,Y^\top = Q^\top$, so that Q is also orthogonal.

The following result expresses the fact that a function is radial if and only if it is invariant under all orthogonal transformations.

Theorem 5.18 [Radial Functions and Orthogonal Matrices]
Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. Let $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. Then, f is radial if and only if for all orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and for all $\boldsymbol{x} \in \mathbb{R}^n$ there holds

$$f(Q\,\boldsymbol{x}) = f(\boldsymbol{x}).$$

Proof If f is radial, given any orthogonal matrix Q, for every $\boldsymbol{x} \in \boldsymbol{A}$ we have $\|Q\,\boldsymbol{x}\| = \|\boldsymbol{x}\|$, and therefore $f(\boldsymbol{x}) = f(Q\,\boldsymbol{x})$.
Vice-versa, let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{A}$ such that $\|\boldsymbol{x}\| = \|\boldsymbol{y}\|$. We have to show that $f(\boldsymbol{x}) = f(\boldsymbol{y})$. There exists an orthogonal matrix Q such that $Q\,\boldsymbol{x} = \boldsymbol{y}$. Thus, $f(\boldsymbol{y}) = f(Q\,\boldsymbol{x}) = f(\boldsymbol{x})$.

This characterization of radial function lends itself to be used as a definition of radial distributions, considering the orthogonal matrix as a C^∞ diffeomorphism.

Definition 5.17 *Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We say that T is radial if, for all orthogonal matrices $Q \in \mathbb{R}^{n\times n}$, there holds $T \circ Q = T$.*

Example 5.28 The distribution $\delta \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ is radial. In fact, given $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ and any $n \times n$ orthogonal matrix Q, we have

$$\langle \delta \circ Q, \varphi\rangle = \left\langle \delta, (\varphi \circ Q^{-1}) \frac{1}{|\det Q|}\right\rangle = \varphi(Q^{-1}(0)) = \varphi(0) = \langle \delta, \varphi\rangle,$$

where we have used the fact that $\det Q \in \{-1, 1\}$.

Homogeneous functions

Definition 5.18 [Homogeneous Function]
Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. We say that $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ is (positively) homogeneous *if there exists $\lambda \in \mathbb{R}$ such that*

$$\forall\, \boldsymbol{x} \in \boldsymbol{A},\ \ \forall\, t > 0, \qquad f(t\,\boldsymbol{x}) = t^\lambda\, f(\boldsymbol{x}). \tag{5.34}$$

The real number λ is called degree of homogeneity.

Example 5.29 The function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ described by $f(x, y) = x^2 + y^2$ is homogeneous of degree 2. In fact

$$f(t\,x, t\,y) = (t\,x)^2 + (t\,y)^2 = t^2\,(x^2 + y^2) = t^2\,(x, y).$$

Example 5.30 Given two vector spaces $\boldsymbol{V}$ and $\boldsymbol{W}$ over the same field, a linear transformation $T : \boldsymbol{V} \longrightarrow \boldsymbol{W}$ is homogeneous of degree 1. Indeed, by definition T preserves the scalar product:

$$\forall\, \boldsymbol{v} \in \boldsymbol{V},\ \forall\, \alpha \in \mathbb{F}, \qquad T(\alpha\,\boldsymbol{v}) = \alpha\, T(\boldsymbol{v}).$$

We want to extend the definition of homogeneous function to distributions. By looking at (5.34), this

seems possible, provided we consider the diffeomorphism

$$\begin{aligned} F : \boldsymbol{A} &\longrightarrow \boldsymbol{A} \\ \boldsymbol{x} &\mapsto t\,\boldsymbol{x}. \end{aligned} \tag{5.35}$$

This function is indeed a diffeomorphism of class C^∞.

Definition 5.19 [Homogeneous Distribution]
Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. We say that T is (positively) homogeneous *if there exists $\lambda \in \mathbb{R}$ such that*

$$\forall\, t > 0, \qquad T(t\,\boldsymbol{x}) = t^\lambda\, T(\boldsymbol{x}). \tag{5.36}$$

The real number λ is called degree of homogeneity.

We have been sloppy in the way we have written $T(t\,\boldsymbol{x})$ instead of $T \circ F$, where F is as in (5.35). We have followed, however, a common trend, which consists in replacing the name of the function/distribution with the value of the function/distribution at a generic point $\boldsymbol{x}$. This notation has the advantage of being more explicit.

Example *5.31*

We want to show that $\delta \in \mathcal{D}'(\mathbb{R}^n; \mathbb{R})$ is homogeneous of degree $-n$. Hence, given $t > 0$, we must show that for every $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$, there holds $\langle \delta \circ F, \varphi \rangle = t^{-n} \langle \delta, \varphi \rangle$, where F is the diffeomorphism

$$\begin{aligned} F : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{x} &\mapsto t\,\boldsymbol{x}. \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} F^{-1} : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{y} &\mapsto t^{-1}\,\boldsymbol{y}. \end{aligned}$$

We have $F^{-1}(\mathbf{0}) = \mathbf{0}$, and

$$\det J_{F^{-1}}(\boldsymbol{y}) = \det \begin{bmatrix} t^{-1} & 0 & \dots & 0 \\ 0 & t^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t^{-1} \end{bmatrix} = t^{-n}.$$

Thus,

$$\begin{aligned} \langle \delta \circ F, \varphi \rangle &= \langle \delta, (\varphi \circ F^{-1}) \,|\det J_{F^{-1}}| \rangle \\ &= \varphi(F^{-1}(\mathbf{0}))\, t^{-n} = t^{-n} \langle \delta, \varphi \rangle. \end{aligned}$$

Odd and even distributions

Consider a function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$, where $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. We recall that f is even if

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad f(-\boldsymbol{x}) = f(\boldsymbol{x}),$$

and it is odd if

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \qquad f(-\boldsymbol{x}) = -f(\boldsymbol{x}).$$

Considering the diffeomorphism

$$\begin{aligned} F : \boldsymbol{A} &\longrightarrow \boldsymbol{A} \\ \boldsymbol{x} &\mapsto -\boldsymbol{x} \end{aligned} \tag{5.37}$$

it is clear how to extend this definition to distributions.

Definition 5.20 [Odd and Even Distributions]
Let $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. We say that:

(i) T is even *if $T(-\boldsymbol{x}) = T(\boldsymbol{x})$;*

(ii) T is odd *if $T(-\boldsymbol{x}) = -T(\boldsymbol{x})$.*

Example 5.32

We want to show that $\delta \in \mathcal{D}'(\mathbb{R}^n; \mathbb{R})$ is an even distribution. For every $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$ we have

$$\langle \delta(-\boldsymbol{x}), \varphi(\boldsymbol{x}) \rangle = \langle \delta(\boldsymbol{y}), \varphi(-\boldsymbol{y}) \rangle = \varphi(0) = \langle \delta, \varphi \rangle,$$

where we have used the fact that $F(\boldsymbol{x}) = -\boldsymbol{x}$, the inverse is $F^{-1}(\boldsymbol{y}) = -\boldsymbol{y}$, and therefore the Jacobian matrix is $-I$. The determinant of $-I$ is 1 or -1, depending on whether n is odd or even (precisely, $\det(-I) = (-1)^{n+1}$), but the absolute value is always 1.

Example 5.33

The principal value of $\frac{1}{x}$ is an odd distribution. Indeed, let $A = \mathbb{R}$ and let $T = \operatorname{pv} \frac{1}{x}$. Let $\varphi \in \mathcal{D}(\mathbb{R}; \mathbb{R})$. Then,

$$\langle T(-x), \varphi(x) \rangle = \langle T(y), \varphi(-y) \rangle = \lim_{\varepsilon \to 0^+} \int_{|y| \geq \varepsilon} \frac{\varphi(-y)}{y} \, \mathrm{d}y = \lim_{\varepsilon \to 0^+} \int_{|z| \geq \varepsilon} \frac{\varphi(z)}{-z} \, \mathrm{d}z = -\langle T, \varphi \rangle.$$

Notice that in this change of variable we used the fact that the Jacobian matrix of the change of variable $z = -y$ in this case is simply a derivative, and its absolute value is 1. The same result can be obtained using the orientation of intervals:

$$\begin{aligned}
\lim_{\varepsilon \to 0^+} \int_{|y| \geq \varepsilon} \frac{\varphi(-y)}{y} \, \mathrm{d}y &= \lim_{\varepsilon \to 0^+} \left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(-y)}{y} \, \mathrm{d}y + \int_{\varepsilon}^{+\infty} \frac{\varphi(-y)}{y} \, \mathrm{d}y \right) \\
&= \lim_{\varepsilon \to 0^+} \left(\int_{+\infty}^{\varepsilon} \frac{\varphi(z)}{-z} (-1) \, \mathrm{d}z + \int_{-\varepsilon}^{-\infty} \frac{\varphi(z)}{-z} (-1) \, \mathrm{d}z \right) \\
&= \lim_{\varepsilon \to 0^+} \left(- \int_{\varepsilon}^{+\infty} \frac{\varphi(z)}{z} \, \mathrm{d}z - \int_{-\infty}^{-\varepsilon} \frac{\varphi(z)}{z} \, \mathrm{d}z \right) \\
&= - \lim_{\varepsilon \to 0^+} \left(\int_{-\infty}^{-\varepsilon} \frac{\varphi(z)}{z} \, \mathrm{d}z + \int_{\varepsilon}^{+\infty} \frac{\varphi(z)}{z} \, \mathrm{d}z \right) = - \lim_{\varepsilon \to 0^+} \int_{|z| \geq \varepsilon} \frac{\varphi(z)}{z} \, \mathrm{d}z.
\end{aligned}$$

Periodic distributions

Given a function $f : \mathbb{R} \longrightarrow \mathbb{C}$, we say that f is periodic with period t_0 (or t_0-periodic) if

$$\forall\, x \in \mathbb{R}, \qquad f(x + t_0) = f(x).$$

We can clearly extend this definition to distributions, by considering the diffeomorphism

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x + t_0. \end{aligned} \tag{5.38}$$

Definition 5.21 [Periodic Distribution]
*Let $t_0 \in \mathbb{R}$. Let $T \in \mathcal{D}'(\mathbb{R}; \mathbb{C})$. We say that T is t_0-*periodic *if*

$$T(x + t_0) = T(x). \tag{5.39}$$

We now know how to interpret the identity (5.39). Indeed, this is just way of rewriting the identity $T \circ F = T$, where F is defined in (5.38).

Example *5.34*

Consider the following distribution

$$T(x) = \sum_{k=-\infty}^{+\infty} \delta(x-k).$$

We have already seen that for $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$

$$\langle \delta(x-k), \varphi(x) \rangle = \varphi(k),$$

so that

$$\langle T(x), \varphi(x) \rangle = \sum_{k=-\infty}^{+\infty} \langle \delta(x-k), \varphi(x) \rangle = \sum_{k=-\infty}^{+\infty} \varphi(k).$$

Notice that there are no issues of convergence because φ has compact support: it is zero outside a certain closed interval $[-M, M]$. It is also clear that T defines a distribution from the theorem of completeness: T can be obtained as the limit of a sequence of distributions:

$$T_n(x) = \sum_{k=-n}^{n} \delta(x-k).$$

Clearly, for every $n \in \mathbb{N}$ we have $T_n(x) \in \mathcal{D}'(\mathbb{R};\mathbb{R})$, and for all $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$

$$\lim_{n \to +\infty} \langle T_n(x), \varphi(x) \rangle = \langle T(x), \varphi(x) \rangle.$$

We now show that T is 1-periodic. We use the diffeomorphism $F(x) = x + 1$, so that $F^{-1}(y) = y - 1$, and

$$\langle T(x+1), \varphi(x) \rangle = \langle T, \varphi \circ F^{-1} \rangle = \sum_{k=-\infty}^{+\infty} \varphi(F^{-1}(k)) = \sum_{k=-\infty}^{+\infty} \varphi(k-1) = \sum_{k=-\infty}^{+\infty} \varphi(k) = \langle T(x), \varphi(x) \rangle.$$

Generalizations of the composition

In some cases, it is possible to define the composition of a distribution with a function which is not a C^∞-diffeomorphism. For example, let $a \in \mathbb{R}^*$. We want to be able to define $\delta(x^2 - a^2)$. However, the function $F(x) = x^2 - a^2$ is not injective, and therefore it is not invertible. We can view $\langle \delta(x^2 - a^2), \varphi \rangle$ as an integral, that we decompose into two integrals where the function F is injective. In other words,

$$\langle \delta(x^2 - a^2), \varphi \rangle = \int_{-\infty}^{0} \delta(x^2 - a^2)\, \varphi(x)\, \mathrm{d}x + \int_{0}^{+\infty} \delta(x^2 - a^2)\, \varphi(x)\, \mathrm{d}x. \tag{5.40}$$

We define the functions

$$\begin{array}{rcl} F_1 : \mathbb{R}_+^* & \longrightarrow & (-a^2, +\infty) \\ x & \mapsto & x^2 - a^2 \end{array} \quad \text{and} \quad \begin{array}{rcl} F_2 : \mathbb{R}_-^* & \longrightarrow & (-a^2, +\infty) \\ x & \mapsto & x^2 - a^2. \end{array}$$

The two functions F_1 and F_2 are C^∞-diffeomorphisms, and their inverses are, respectively,

$$\begin{array}{rcl} F_1^{-1} : (-a^2, +\infty) & \longrightarrow & \mathbb{R}_+^* \\ y & \mapsto & \sqrt{y + a^2} \end{array} \quad \text{and} \quad \begin{array}{rcl} F_2^{-1} : (-a^2, +\infty) & \longrightarrow & \mathbb{R}_-^* \\ y & \mapsto & -\sqrt{y + a^2}. \end{array}$$

The corresponding derivatives are

$$(F_1^{-1})' : (-a^2, +\infty) \longrightarrow \mathbb{R}_+^* \quad y \mapsto \frac{1}{2\sqrt{y+a^2}} \qquad \text{and} \qquad (F_2^{-1})' : (-a^2, +\infty) \longrightarrow \mathbb{R}_-^* \quad y \mapsto -\frac{1}{2\sqrt{y+a^2}}.$$

Using the rule of the composition of a distribution for each of the two integrals in (5.40) we find

$$\begin{aligned}\langle \delta(x^2-a^2), \varphi\rangle &= \int_{-\infty}^{0} \delta(y)\, \varphi(-\sqrt{y+a^2})\, \frac{1}{2\sqrt{y+a^2}}\, \mathrm{d}x + \int_{0}^{+\infty} \delta(y)\, \varphi(\sqrt{y+a^2})\, \frac{1}{2\sqrt{y+a^2}}\, \mathrm{d}x \\ &= \varphi(-|a|)\, \frac{1}{2|a|} + \varphi(|a|)\, \frac{1}{2|a|} = \frac{1}{2|a|}\, (\varphi(|a|) + \varphi(-|a|)) \\ &= \frac{1}{2|a|}\, (\langle \delta(x-|a|), \varphi\rangle + \langle \delta(x+|a|), \varphi\rangle).\end{aligned}$$

Therefore, we can define

$$\delta(x^2-a^2) = \frac{1}{2|a|}\, (\delta(x-a) + \delta(x+a)).$$

5.5.3 Restriction to an open set

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be an open set, and let $\boldsymbol{B} \subseteq \boldsymbol{A}$ be an open set. Consider $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A}; \mathbb{C})$. Then, clearly $[f|_{\boldsymbol{B}}] \in L^1_{\mathrm{loc}}(\boldsymbol{B}; \mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{B}; \mathbb{C})$. Then,

$$\langle L_{f|_{\boldsymbol{B}}}, \varphi\rangle = \int_{\boldsymbol{B}} f(\boldsymbol{x})\, \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

Consider the trivial continuation of φ to $\boldsymbol{A}$:

$$\tilde{\varphi} : \boldsymbol{A} \longrightarrow \mathbb{C}, \quad \boldsymbol{x} \mapsto \begin{cases} \varphi(\boldsymbol{x}) & \boldsymbol{x} \in \boldsymbol{B} \\ 0 & \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{B}. \end{cases}$$

Then, $\tilde{\varphi} \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$, because φ is zero on $\partial \boldsymbol{B}$ (φ has compact support: it is therefore zero on the boundary, because the compact does not intersect the boundary). Thus,

$$\begin{aligned}\langle L_{f|_{\boldsymbol{B}}}, \varphi\rangle &= \int_{\boldsymbol{B}} f(\boldsymbol{x})\, \varphi(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \\ &= \int_{\boldsymbol{A}} f(\boldsymbol{x})\, \tilde{\varphi}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \langle L_f, \tilde{\varphi}\rangle.\end{aligned}$$

This suggests the following definition.

Definition 5.22 [Restriction of Distributions to Open Sets]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $\boldsymbol{B} \subseteq \boldsymbol{A}$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. We define $T|_{\boldsymbol{B}} \in \mathcal{D}'(\boldsymbol{B}; \mathbb{C})$ to be the distribution such that

$$\forall\, \varphi \in \mathcal{D}(\boldsymbol{B}; \mathbb{C}), \qquad \langle T|_{\boldsymbol{B}}, \varphi\rangle = \langle T, \tilde{\varphi}\rangle. \tag{5.41}$$

To ensure that this definition is well-posed, we need to make sure that linearity and continuity hold. The linearity is obvious. We now prove the continuity.

Theorem 5.19 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $\boldsymbol{B} \subseteq A$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of functions in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Then,*

$$\varphi_k \xrightarrow{k\to\infty} \varphi \text{ in } \mathcal{D}(\boldsymbol{B};\mathbb{C}) \qquad \Rightarrow \qquad \langle T|_{\boldsymbol{B}}, \varphi_k\rangle \xrightarrow{k\to\infty} \langle T|_{\boldsymbol{B}}, \varphi\rangle.$$

Proof Let $(\varphi_k)_{k\in\mathbb{N}}$ converge to φ in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$. We show that $\langle T|_{\boldsymbol{B}}, \varphi_k\rangle \xrightarrow{k\to\infty} \langle T|_{\boldsymbol{B}}, \varphi\rangle$, i.e., that $\langle T, \tilde{\varphi}_k\rangle \xrightarrow{k\to\infty} \langle T, \tilde{\varphi}\rangle$, where $\tilde{\varphi}_k$ and $\tilde{\varphi}$ denote the trivial continuations of φ_k and φ to $\boldsymbol{A}$, respectively. It suffices to prove that if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$, then $(\tilde{\varphi}_k)_{k\in\mathbb{N}}$ converges to $\tilde{\varphi}$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. It is easy to see that there exists a compact set in $\boldsymbol{A}$ that contains the supports of all the functions $\tilde{\varphi}_k$. Indeed, we know that there exists a compact set $\boldsymbol{K} \subset \boldsymbol{B}$ that contains the supports of all the functions φ_k. On the other hand, the supports of $\tilde{\varphi}_k$ coincide with the supports of φ_k for any $k \in \mathbb{N}$, and the compact set $\boldsymbol{K}$ is also a compact set in $\boldsymbol{A}$, so that

$$\operatorname{supp}\tilde{\varphi}_k = \operatorname{supp}\varphi_k \subseteq \boldsymbol{K} \subset \boldsymbol{B} \subseteq \boldsymbol{A}.$$

Now, since $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$, for every $\alpha \in \mathbb{N}^n$ we have

$$\partial^\alpha \tilde{\varphi}_k(\boldsymbol{x}) = \begin{cases} \partial^\alpha \varphi_k(\boldsymbol{x}) & \boldsymbol{x} \in \boldsymbol{B} \\ 0 & \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{B} \end{cases}$$

so that

$$\sup_{\boldsymbol{A}} |\partial^\alpha \tilde{\varphi}_k - \partial^\alpha \tilde{\varphi}| = \sup_{\boldsymbol{B}} |\partial^\alpha \varphi_k - \partial^\alpha \varphi| \xrightarrow{k\to\infty} 0.$$

Example *5.35* Let B be an open set in $\mathbb{R}$. Let $\varphi \in \mathcal{D}(B;\mathbb{R})$. Then

$$\langle \delta|_B, \varphi\rangle = \langle \delta, \tilde{\varphi}\rangle = \tilde{\varphi}(0) = \begin{cases} 0 & \text{if } 0 \notin B \\ \varphi(0) & \text{if } 0 \in B \end{cases}$$

Thus,

$$\delta|_B = \begin{cases} 0 & \text{if } 0 \notin B \\ \delta & \text{if } 0 \in B \end{cases}$$

Example *5.36* Consider the distribution $T = \operatorname{pv}\frac{1}{x}$. Consider $B = \mathbb{R}\setminus\{0\}$. We want to determine $T|_B$. First, we observe that the function

$$\begin{aligned} f : \mathbb{R}\setminus\{0\} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x} \end{aligned}$$

is in $L^1_{\text{loc}}(\mathbb{R}\setminus\{0\};\mathbb{R})$, because it is continuous in $\mathbb{R}\setminus\{0\}$. We want to show that

$$T|_B = L_f.$$

Let $\varphi \in \mathcal{D}(B;\mathbb{R})$. We have

$$\langle T|_B, \varphi\rangle = \langle T, \tilde{\varphi}\rangle = \lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} \frac{\tilde{\varphi}(x)}{x}\,\mathrm{d}x, \qquad \text{where} \quad \tilde{\varphi}(x) = \begin{cases} \varphi(x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Since φ has compact support in $\mathbb{R}\setminus\{0\}$, we have $0 \notin \operatorname{supp}\varphi$, which means that there is a neighbourhood of 0 such that φ restricted to this neighbourhood is identically zero. In other words, there exists $\alpha > 0$ such that

$\varphi(x) = 0$ for all $|x| \leq \alpha$, and therefore for every $\varepsilon \leq \alpha$ we have

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x &= \lim_{\varepsilon\to 0^+}\left(\int_{-\infty}^{-\alpha}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{-\alpha}^{-\varepsilon}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{\varepsilon}^{\alpha}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{\alpha}^{+\infty}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x\right)\\ &= \lim_{\varepsilon\to 0^+}\left(\int_{-\infty}^{-\alpha}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{\alpha}^{+\infty}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x\right)\\ &= \lim_{\varepsilon\to 0^+}\left(\int_{-\infty}^{-\alpha}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{-\alpha}^{0}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{0}^{\alpha}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x+\int_{\alpha}^{+\infty}\frac{\tilde\varphi(x)}{x}\,\mathrm{d}x\right)=\int_{\mathbb{R}\setminus\{0\}}\frac{\varphi(x)}{x}\,\mathrm{d}x.\end{aligned}$$

We conclude with the following result.

Theorem 5.20 *Let $\boldsymbol{A}\subseteq\mathbb{R}^n$ be open. Let $\boldsymbol{B}\subseteq\boldsymbol{A}$ be open. Let $T\in\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of distributions in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Then,*

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \qquad\Rightarrow\qquad T_k|_{\boldsymbol{B}} \xrightarrow{k\to\infty} T|_{\boldsymbol{B}} \text{ in } \mathcal{D}'(\boldsymbol{B};\mathbb{C}).$$

The proof is left as an exercise (see Question **??**).

Exercises

Exercise 5.32 Let $\boldsymbol{A}\subseteq\mathbb{R}^n$ be open. Let $\boldsymbol{\Theta}\subseteq\boldsymbol{A}$ be open. Let $T\in\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of distributions in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $T\in\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Show that

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \qquad\Rightarrow\qquad T_k|_{\boldsymbol{\Theta}} \xrightarrow{k\to\infty} T|_{\boldsymbol{\Theta}} \text{ in } \mathcal{D}'(\boldsymbol{\Theta};\mathbb{C}).$$

[Solution: (5.32)]

Exercise 5.33 Let $\boldsymbol{A}\subseteq\mathbb{R}^n$ be open. Let $\boldsymbol{\Theta}\subseteq\boldsymbol{A}$ be open. Let $T\in\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Prove that $T|_{\boldsymbol{\Theta}}$ is a linear functional. [Solution: (5.33)]

5.5.4 Tensor product

The next operation is the so-called tensor product (also known as direct product, or sometimes Dirichlet product). Given two functions of n and m variables, respectively, their tensor product is a function of $n+m$ variables and is defined as follows.

Definition 5.23 [Tensor Product]
Let $\boldsymbol{A},\boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $[f_1]\in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and $[f_2]\in L^1_{\mathrm{loc}}(\boldsymbol{B};\mathbb{C})$. The tensor product of f_1 and f_2, denoted by $f_1\otimes f_2$, is the function $f_1\otimes f_2:\boldsymbol{A}\times\boldsymbol{B}\longrightarrow\mathbb{C}$ defined by

$$(f_1\otimes f_2)(\boldsymbol{x},\boldsymbol{y}) \stackrel{def}{=} f_1(\boldsymbol{x})\,f_2(\boldsymbol{y}),\quad (\boldsymbol{x},\boldsymbol{y})\in\boldsymbol{A}\times\boldsymbol{B}.$$

The next result shows that if two functions are locally summable, such is also their tensor product.

Theorem 5.21 *Let $\boldsymbol{A},\boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $[f_1]\in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and $[f_2]\in$*

$L^1_{\text{loc}}(\boldsymbol{B};\mathbb{C})$. *Then,* $[f_1 \otimes f_2] \in L^1_{\text{loc}}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$.

Proof We prove that $f_1 \otimes f_2$ is measurable. Consider for the sake of simplicity that f_1 and f_2 are real-valued. First, we show that if f_1 and f_2 are simple functions, then $f_1 \otimes f_2$ is simple. In fact, if f_1 is simple, there exist $p \in \mathbb{N}^*$, there exist pairwise disjoint measurable sets $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_p$ whose union is $\boldsymbol{A}$, and $\alpha_1,\ldots,\alpha_p \in \mathbb{R}$ such that

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad f_1(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}).$$

Likewise, if f_2 is simple, there exist $q \in \mathbb{N}^*$, there exist pairwise disjoint measurable sets $\boldsymbol{B}_1,\ldots,\boldsymbol{B}_q$ whose union is $\boldsymbol{B}$, and $\beta_1,\ldots,\beta_q \in \mathbb{R}$ such that

$$\forall\, \boldsymbol{y} \in \boldsymbol{B}, \quad f_2(\boldsymbol{y}) = \sum_{j=1}^{q} \beta_j \chi_{\boldsymbol{B}_j}(\boldsymbol{y}).$$

Hence, for all $(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{A}\times\boldsymbol{B}$

$$\begin{aligned}(f_1 \otimes f_2)(\boldsymbol{x},\boldsymbol{y}) &= f_1(\boldsymbol{x})\, f_2(\boldsymbol{y}) = \left(\sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x})\right)\left(\sum_{j=1}^{q} \beta_j \chi_{\boldsymbol{B}_j}(\boldsymbol{y})\right)\\ &= \sum_{k=1}^{p}\sum_{j=1}^{q} \alpha_k\beta_j \chi_{\boldsymbol{A}_k}(\boldsymbol{x})\chi_{\boldsymbol{B}_j}(\boldsymbol{y}) = \sum_{k=1}^{p}\sum_{j=1}^{q} \alpha_k\beta_j \chi_{\boldsymbol{A}_k\times\boldsymbol{B}_j}(\boldsymbol{x},\boldsymbol{y}),\end{aligned}$$

Thus, $f_1 \otimes f_2$ is simple. Now we apply Theorem 32 in the set of notes on the Lebesgue measure and integral: $f_1 \otimes f_2$ is measurable if and only if there exists a sequence of simple functions $(\varphi_k)_{k\in\mathbb{N}}$, with $\varphi_k : \boldsymbol{A}\times\boldsymbol{B} \longrightarrow \mathbb{R}$ for all $k \in \mathbb{N}$ (which can be chosen in such a way that $|\varphi_k(\boldsymbol{x},\boldsymbol{y})| \le |f_1(\boldsymbol{x})\, f_2(\boldsymbol{y})|$ for every $k \in \mathbb{N}$ and for every $(\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{A}\times\boldsymbol{B}$) satisfying

$$\forall\, (\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{A}\times\boldsymbol{B}, \quad \lim_{k\to\infty} \varphi_k(\boldsymbol{x},\boldsymbol{y}) = f(\boldsymbol{x},\boldsymbol{y}).$$

Since f_1 is measurable, there exists a sequence of simple functions $(\varphi^1_k)_{k\in\mathbb{N}}$, with $\varphi^1_k : \boldsymbol{A} \longrightarrow \mathbb{R}$ for all $k \in \mathbb{N}$, satisfying

$$\forall\, \boldsymbol{x} \in \boldsymbol{A}, \quad \lim_{k\to\infty} \varphi^1_k(\boldsymbol{x}) = f_1(\boldsymbol{x}).$$

Likewise, since f_2 is measurable, there exists a sequence of simple functions $(\varphi^2_k)_{k\in\mathbb{N}}$, with $\varphi^2_k : \boldsymbol{B} \longrightarrow \mathbb{R}$ for all $k \in \mathbb{N}$, satisfying

$$\forall\, \boldsymbol{y} \in \boldsymbol{B}, \quad \lim_{k\to\infty} \varphi^2_k(\boldsymbol{y}) = f_2(\boldsymbol{y}).$$

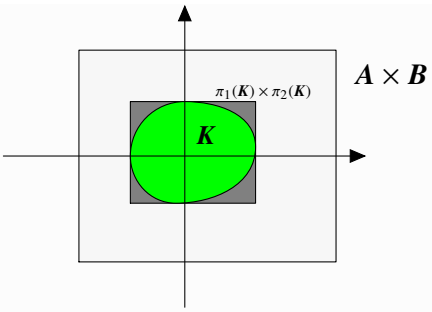
The function $\varphi_k : \boldsymbol{A}\times\boldsymbol{B} \longrightarrow \mathbb{R}$ defined by $\varphi_k(\boldsymbol{x},\boldsymbol{y}) = \varphi^1_k(\boldsymbol{x})\,\varphi^2_k(\boldsymbol{y})$ is simple, and

$$\forall\, (\boldsymbol{x},\boldsymbol{y}) \in \boldsymbol{A}\times\boldsymbol{B}, \quad \lim_{k\to\infty} \varphi_k(\boldsymbol{x},\boldsymbol{y}) = \lim_{k\to\infty} \varphi^1_k(\boldsymbol{x})\varphi^2_k(\boldsymbol{y}) = f_1(\boldsymbol{x})\, f_2(\boldsymbol{y}) = (f_1 \otimes f_2)(\boldsymbol{x},\boldsymbol{y})$$

as required. This shows that $f_1 \otimes f_2$ is measurable. The generalization to the case where f_1 and f_2 take complex values is standard. Now, we show that $f_1 \otimes f_2$ is summable over compact sets of $\boldsymbol{A}\times\boldsymbol{B}$. Consider the projection maps

$$\begin{array}{rcl}\pi_1 : \mathbb{R}^n\times\mathbb{R}^m &\longrightarrow& \mathbb{R}^n\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto& \boldsymbol{x}\end{array} \quad \text{and} \quad \begin{array}{rcl}\pi_2 : \mathbb{R}^n\times\mathbb{R}^m &\longrightarrow& \mathbb{R}^m\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto& \boldsymbol{y}.\end{array}$$

Consider a compact set $\boldsymbol{K}$ of $\boldsymbol{A}\times\boldsymbol{B}$. Since π_1 and π_2 are continuous, $\pi_1(\boldsymbol{K})$ is compact in $\boldsymbol{A}$ and $\pi_2(K)$ is compact in $\boldsymbol{B}$. Thus, $\boldsymbol{K} \subseteq \pi_1(\boldsymbol{K})\times\pi_2(\boldsymbol{K})$.



By using Tonelli's theorem (Theorem 3.54), we can see that the tensor product is measurable in $\pi_1(K) \times \pi_2(K)$. Indeed,

$$\begin{aligned}\int_K |(f_1 \otimes f_2)(x,y)|\,dx\,dy &\le \int_{\pi_1(K)\times\pi_2(K)} |(f_1 \otimes f_2)(x,y)|\,dx\,dy \\ &= \int_{\pi_1(K)\times\pi_2(K)} |f_1(x)|\,|f_2(y)|\,dx\,dy \\ &= \int_{\pi_1(K)} |f_1(x)|\,dx \int_{\pi_2(K)} |f_2(x)|\,dx.\end{aligned}$$

Both integrals $\int_{\pi_1(K)} |f_1(x)|\,dx$ and $\int_{\pi_2(K)} |f_2(x)|\,dx$ are finite because f_1 and f_2 are locally summable.

Remark 5.7 It is easy to see that, if f_1 and f_2 are summable, then $f_1 \otimes f_2$ is also summable. In fact, one can use Tonelli's theorem (Theorem 3.54) by setting $f(x,y) = |(f_1 \otimes f_2)(x,y)| = |f_1(x)||f_2(y)|$ to conclude that

$$\int_{A\times B} |(f_1 \otimes f_2)(x,y)|\,dx\,dy = \int_A \left(\int_B |f_1(x)|\,|f_2(y)|\,dy\right) dx = \int_A |f_1(x)|\,dx \int_B |f_2(y)|\,dy.$$

To determine how the tensor product can be defined for distributions, we consider $\varphi \in \mathcal{D}(A \times B; \mathbb{C})$. Under the assumptions of the previous theorem, we have the two equalities

$$\int_{A\times B} f_1(x)\, f_2(y)\, \varphi(x,y)\,dx\,dy = \int_A f_1(x)\Big(\int_B f_2(y)\,\varphi(x,y)\,dy\Big)\,dx, \tag{5.42}$$

and

$$\int_{A\times B} f_1(x)\, f_2(y)\, \varphi(x,y)\,dx\,dy = \int_B f_2(y)\Big(\int_A f_1(x)\,\varphi(x,y)\,dx\Big)\,dy, \tag{5.43}$$

where Fubini's theorem (Theorem 3.8) has been used, using fact that $f_1(x)\, f_2(y)\, \varphi(x,y)$ is summable. Now we can use one of these two expressions, say (5.42), to see how to define the tensor product for distributions. For the inner integral to be interpreted as the action of a linear functional L_{f_2} on a test function, we need to ensure that, for each $x \in A$, the function $\varphi(x,y)$ (which is now a function of y) is a test function in $\mathcal{D}(B;\mathbb{C})$. Then, to interpret the outer integral as the action of another linear functional L_{f_1} on a test function, the function

$$x \mapsto \int_B f_2(y)\,\varphi(x,y)\,dy$$

must be in $\mathcal{D}(A;\mathbb{C})$.

Lemma 5.12 *Let $\boldsymbol{A}, \boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $[f_1] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$ and $f_2 \in L^1_{\text{loc}}(\boldsymbol{B};\mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$. Then,*

(i) for all $\boldsymbol{x} \in \boldsymbol{A}$, the map

$$\begin{aligned} \varphi(\boldsymbol{x},\cdot) : \boldsymbol{B} &\longrightarrow \mathbb{C} \\ \boldsymbol{y} &\mapsto \varphi(\boldsymbol{x},\boldsymbol{y}), \end{aligned} \tag{5.44}$$

is in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$;

(ii) the map

$$\begin{aligned} w : \boldsymbol{A} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \end{aligned} \tag{5.45}$$

is in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$.

Proof We want to show that, for all $\boldsymbol{x} \in \boldsymbol{A}$, the function $\varphi(\boldsymbol{x},\cdot)$ is in $\mathcal{D}(B;\mathbb{C})$. First, we show that this function has compact support in $\boldsymbol{B}$. Let $\boldsymbol{K} \subset \boldsymbol{A}\times\boldsymbol{B}$ be the compact set such that $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$. Thus, $\varphi(\boldsymbol{x},\boldsymbol{y}) = 0$ if $(\boldsymbol{x},\boldsymbol{y}) \notin \boldsymbol{K}$. Consider the functions π_1 and π_2 defined in the proof of Theorem 5.21. Using the same argument of that proof, we find that $\pi_1(\boldsymbol{K})$ and $\pi_2(\boldsymbol{K})$ are compact sets in $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively, and $\boldsymbol{K} \subseteq \pi_1(\boldsymbol{K})\times\pi_2(\boldsymbol{K})$. Moreover, for all $\boldsymbol{x} \in \boldsymbol{A}$, we have $\operatorname{supp}\varphi(\boldsymbol{x},\cdot) \subseteq \pi_2(\boldsymbol{K})$, which shows that the function in (5.44) has compact support. Moreover, it is in $C^\infty(\boldsymbol{B};\mathbb{C})$, and it is therefore in $\mathcal{D}(\boldsymbol{B};\mathbb{C})$.

Now, we show that $w \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. First, we show that w has compact support. If $\boldsymbol{x} \notin \pi_1(\boldsymbol{K})$, then $\varphi(\boldsymbol{x},\cdot)$ is zero, so that $\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = 0$, and this implies that $\operatorname{supp} w \subseteq \pi_1(\boldsymbol{K})$.

Now, we prove that $w \in C^\infty(\boldsymbol{A};\mathbb{C})$. We start proving that w is continuous, i.e., given a sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points of $\boldsymbol{A}$ converging to $\hat{\boldsymbol{x}} \in \boldsymbol{A}$, we show that

$$w(\boldsymbol{x}_k) \xrightarrow{k\to\infty} w(\hat{\boldsymbol{x}}).$$

Since we have proved that $\varphi(\boldsymbol{x},\cdot) \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$, the sequence $\partial_{\boldsymbol{y}}^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x}_k,\cdot)$ converges uniformly to $\partial_{\boldsymbol{y}}^{\boldsymbol{\alpha}}\varphi(\hat{\boldsymbol{x}},\cdot)$ in $\boldsymbol{B}$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$, and therefore, since the functional L_{f_2} is continuous, we obtain

$$w(\boldsymbol{x}_k) = \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x}_k,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \xrightarrow{k\to\infty} \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\hat{\boldsymbol{x}},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = w(\hat{\boldsymbol{x}}).$$

This shows that w is continuous. For all $\boldsymbol{x} \in \boldsymbol{A}$,

$$\frac{\varphi(\boldsymbol{x}+h\,\boldsymbol{e}_1,\cdot)-\varphi(\boldsymbol{x},\cdot)}{h} \xrightarrow{h\to 0} \frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1} \tag{5.46}$$

pointwise. We want to show that

$$\frac{\partial w(\boldsymbol{x})}{\partial x_1} = \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\frac{\partial\varphi(\boldsymbol{x},\boldsymbol{y})}{\partial x_1}\,\mathrm{d}\boldsymbol{y}.$$

We have

$$\begin{aligned} \frac{w(\boldsymbol{x}+h\,\boldsymbol{e}_1)-w(\boldsymbol{x})}{h} &= \frac{1}{h}\left(\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x}+h\,\boldsymbol{e}_1,\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} - \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right) \\ &= \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\frac{\varphi(\boldsymbol{x}+h\,\boldsymbol{e}_1,\boldsymbol{y})-\varphi(\boldsymbol{x},\boldsymbol{y})}{h}\,\mathrm{d}\boldsymbol{y}. \end{aligned}$$

Taking the limit on both sides as $h \to 0$ yields the result if we are able to apply Lebesgue's dominated convergence theorem (Theorem 3.47) to the latter expression. In other words, replacing the continuous variable h with a discrete variable k, we want to be able to write

$$\lim_{k\to\infty} \frac{w\left(\boldsymbol{x}+\frac{1}{k}\boldsymbol{e}_1\right)-w(\boldsymbol{x})}{\frac{1}{k}} = \lim_{k\to\infty}\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\frac{\varphi\left(\boldsymbol{x}+\frac{1}{k}\boldsymbol{e}_1,\boldsymbol{y}\right)-\varphi(\boldsymbol{x},\boldsymbol{y})}{\frac{1}{k}}\,\mathrm{d}\boldsymbol{y} = \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\lim_{k\to\infty}\frac{\varphi\left(\boldsymbol{x}+\frac{1}{k}\boldsymbol{e}_1,\boldsymbol{y}\right)-\varphi(\boldsymbol{x},\boldsymbol{y})}{\frac{1}{k}}\,\mathrm{d}\boldsymbol{y}.$$

From (5.46) we find

$$\forall\,\varepsilon>0,\ \exists\,\delta>0:\ \forall\,k\in\mathbb{N},\qquad 0<\frac{1}{k}<\delta\ \Rightarrow\ \left|\frac{\varphi(\boldsymbol{x}+h\boldsymbol{e}_1,\cdot)-\varphi(\boldsymbol{x},\cdot)}{h}-\frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1}\right|<\varepsilon.$$

Using the reverse triangle inequality, if k is sufficiently large

$$\left|\frac{\varphi(\boldsymbol{x}+h\boldsymbol{e}_1,\cdot)-\varphi(\boldsymbol{x},\cdot)}{h}\right|-\left|\frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1}\right|\le\left|\frac{\varphi(\boldsymbol{x}+h\boldsymbol{e}_1,\cdot)-\varphi(\boldsymbol{x},\cdot)}{h}-\frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1}\right|<\varepsilon,$$

i.e.,

$$\left|\frac{\varphi(\boldsymbol{x}+h\boldsymbol{e}_1,\cdot)-\varphi(\boldsymbol{x},\cdot)}{h}\right|<\varepsilon+\left|\frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1}\right|.$$

It follows that, as long as k is sufficiently large,

$$\left|f_2(\boldsymbol{y})\,\frac{\varphi\left(\boldsymbol{x}+\frac{1}{k}\boldsymbol{e}_1,\boldsymbol{y}\right)-\varphi(\boldsymbol{x},\boldsymbol{y})}{\frac{1}{k}}\right|\le|f_2(\boldsymbol{y})|\left(\varepsilon+\left|\frac{\partial\varphi(\boldsymbol{x},\cdot)}{\partial x_1}\right|\right),$$

and the left hand-side is summable. This is sufficient to guarantee that Lebesgue's dominated convergence theorem (Theorem 3.47) can be applied, and we therefore have the differentiability of w with respect to the multi-index $\boldsymbol{e}_1$. Repeating the same argument for all $j\in\{1,\dots,n\}$, we find

$$\frac{\partial w(\boldsymbol{x})}{\partial x_j}=\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\frac{\partial\varphi(\boldsymbol{x},\boldsymbol{y})}{\partial x_j}\,\mathrm{d}\boldsymbol{y}.$$

We can apply the same argument to this formula, and we find that we have the same result for the second-order derivatives, and so forth. Since $\partial^\alpha\varphi$ is in $\mathcal{D}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$, we can conclude that w is of class C^∞.

It follows that

$$\langle L_{f_1\otimes f_2},\varphi\rangle=\langle L_{f_1},\langle L_{f_2},\varphi\rangle\rangle.$$

This notation has the disadvantage that it does not explain that the functional L_{f_2} is acting upon the last m components of $\varphi(\boldsymbol{x},\boldsymbol{y})$, and L_{f_1} on the first n components of $\langle L_{f_2},\varphi\rangle$. For this reason, it is often preferred to write it as

$$\langle(f_1\otimes f_2)(\boldsymbol{x},\boldsymbol{y}),\varphi(\boldsymbol{x},\boldsymbol{y})\rangle=\langle f_1(\boldsymbol{x}),\langle f_2(\boldsymbol{y}),\varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle,$$

despite the usual abuse of notation.

Now we can define the tensor product for distributions as follows.

Definition 5.24 [Tensor Product of Distributions]
Let $\boldsymbol{A},\boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1\in\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and $T_2\in\mathcal{D}'(\boldsymbol{B};\mathbb{C})$. Let $\varphi\in\mathcal{D}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$. We define

$$\langle T_1\otimes T_2,\varphi\rangle\overset{def}{=}\langle T_1(\boldsymbol{x}),\langle T_2(\boldsymbol{y}),\varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle. \tag{5.47}$$

The only claim that remains to be proved, to ensure this definition is well posed, is that the functional in (5.47) is indeed a distribution over $\boldsymbol{A} \times \boldsymbol{B}$. To this end, we need to consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{D}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$ that converges to 0. We verify that the corresponding sequence $(w_k)_{k\in\mathbb{N}}$ defined by

$$w_k(\boldsymbol{x}) = \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi_k(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

tends to 0 in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Since we have already proved that the supports of w_k is contained in $\pi_1(\boldsymbol{K})$, for all $k \in \mathbb{N}$, we only need to show that, for every differentiation $\partial_{\boldsymbol{x}}^{\alpha}$ with respect to the variables x_i, the sequence $(\partial_{\boldsymbol{x}}^{\alpha} w_k)_{k\in\mathbb{N}}$ tends uniformly to 0. Let M and m be as in Theorem 5.10 that correspond to f_2 and $\pi_2(\boldsymbol{K})$. Then, for all $\psi \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$ with support contained in $\pi_2(\boldsymbol{K})$ we have

$$\left|\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\psi(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right| \le M \max_{|\boldsymbol{\beta}|\le m} \|\partial_{\boldsymbol{y}}^{\boldsymbol{\beta}}\psi\|_{\infty,\boldsymbol{B}}$$

where $\partial_{\boldsymbol{y}}^{\boldsymbol{\beta}}$ is the differentiation of any order with respect to the variables y_i. Taking in particular $\psi = \partial_{\boldsymbol{x}}^{\alpha}\varphi_k(\boldsymbol{x},\cdot)$, we obtain for all $\boldsymbol{x} \in A$

$$\left|\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\partial_{\boldsymbol{x}}^{\alpha}\varphi_k(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right| \le M \max_{|\boldsymbol{\beta}|\le m} \|\partial_{\boldsymbol{x}}^{\alpha}\partial_{\boldsymbol{y}}^{\boldsymbol{\beta}}\varphi\|_{\infty,\boldsymbol{B}}$$

Now, since

$$\partial_{\boldsymbol{x}}^{\alpha} w(\boldsymbol{x}) = \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\partial_{\boldsymbol{x}}^{\alpha}\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} \qquad \forall\, \boldsymbol{x} \in \boldsymbol{A},$$

we obtain

$$\|\partial_{\boldsymbol{x}}^{\alpha} w_k(\boldsymbol{x})\|_{\infty,\Omega_1} \le M \max_{|\boldsymbol{\beta}|\le m} \|\partial_{\boldsymbol{x}}^{\alpha}\partial_{\boldsymbol{y}}^{\boldsymbol{\beta}}\varphi_k\|_{\infty,\boldsymbol{A}\times\boldsymbol{B}} \xrightarrow{k\to\infty} 0.$$

Since f_1 is by assumption a distribution, the sequence $\int_{\boldsymbol{A}} f_1\, w_k$ is infinitesimal. Therefore, the functional

$$\varphi \mapsto \int_{\boldsymbol{A}} f_1(\boldsymbol{x}) w(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \qquad \varphi \in \mathcal{D}(\boldsymbol{A}\times\boldsymbol{B};\mathbb{C})$$

is a distribution over $\boldsymbol{A} \times \boldsymbol{B}$. Finally (5.42) holds by definition of f.

Remark 5.8 As already mentioned, if we consider again two functions $f_1 \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$ and $f_2 \in L^1_{\mathrm{loc}}(\boldsymbol{B};\mathbb{C})$, we observe that in view of Fubini's theorem (Theorem 3.8)

$$\begin{aligned}\int_{\boldsymbol{A}\times\boldsymbol{B}} f_1(\boldsymbol{x})\,f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x},\boldsymbol{y})\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} &= \int_{\boldsymbol{A}} f_1(\boldsymbol{x})\Big(\int_{\boldsymbol{B}} f_2(\boldsymbol{y})\,\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\Big)\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{B}} f_2(\boldsymbol{y})\Big(\int_{\boldsymbol{A}} f_1(\boldsymbol{x})\,\varphi(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\Big)\,\mathrm{d}\boldsymbol{y},\end{aligned}$$

we can alternatively define the tensor product for the two distributions $T_1 \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and $T_2 \in \mathcal{D}'(\boldsymbol{B};\mathbb{C})$ as

$$\langle T_1 \otimes T_2, \varphi\rangle \stackrel{\text{def}}{=} \langle T_2(\boldsymbol{y}), \langle T_1(\boldsymbol{x}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle.$$

When dealing with the Fourier series, we will have the background material to show that the two definitions are equivalent.

Example 5.37

We consider the tensor product of two Dirac delta distributions. Let δ_1 the Dirac delta in $\mathbb{R}^n$ and let δ_2 the Dirac delta in $\mathbb{R}^m$. Let $\varphi \in \mathcal{D}(\mathbb{R}^{n+m}; \mathbb{R})$. Then

$$\langle \delta_1 \otimes \delta_2, \varphi \rangle = \langle \delta_1(\boldsymbol{x}), \langle \delta_2(\boldsymbol{y}), \varphi(\boldsymbol{x}, \boldsymbol{y}) \rangle \rangle = \langle \delta_1(\boldsymbol{x}), \varphi(\boldsymbol{x}, 0) \rangle = \varphi(0, 0).$$

It follows that, if we denote by δ the Dirac delta distribution in $\mathbb{R}^{n+m}$, we have $\delta = \delta_1 \otimes \delta_2$.

Example 5.38

We consider the tensor product of a Dirac delta with a function-type distribution. Let $[f] \in L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$. Then, $L_f \otimes \delta \in \mathcal{D}'(\mathbb{R}^2; \mathbb{R})$, and for $\varphi \in \mathcal{D}(\mathbb{R}^2; \mathbb{R})$ we have

$$\begin{aligned} \langle (L_f \otimes \delta)(x, y), \varphi(x, y) \rangle &= \langle f(x), \langle \delta(y), \varphi(x, y) \rangle \rangle \\ &= \langle f(x), \varphi(x, 0) \rangle = \int_{\mathbb{R}} f(x)\, \varphi(x, 0)\, \mathrm{d}x. \end{aligned}$$

Since in the case of three functions f_1, f_2, f_3 we have

$$(f_1 \otimes f_2) \otimes f_3 = f_1 \otimes (f_2 \otimes f_3),$$

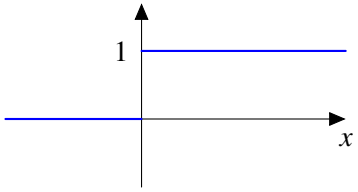
it is now easy to see that this property continues to hold in the case of distributions. Therefore, we can write unambiguously $f_1 \otimes f_2 \otimes f_3$ or, in the case of distributions, $T_1 \otimes T_2 \otimes T_3$.

5.5.5 Derivative of distributions

The derivative is the most important among the operations involving distributions, and it enjoys some remarkable properties: first, any distribution is infinitely differentiable. Second, the result of differentiation does not depend on the order of differentiation. Third, we can extend Leibniz's formula for differentiation for the product of a distribution by a C^∞ function. Before we start with our investigation, we introduce some preliminary observations.

Consider the following "step function", or Heaviside function $H = \chi_{\mathbb{R}_+}$:

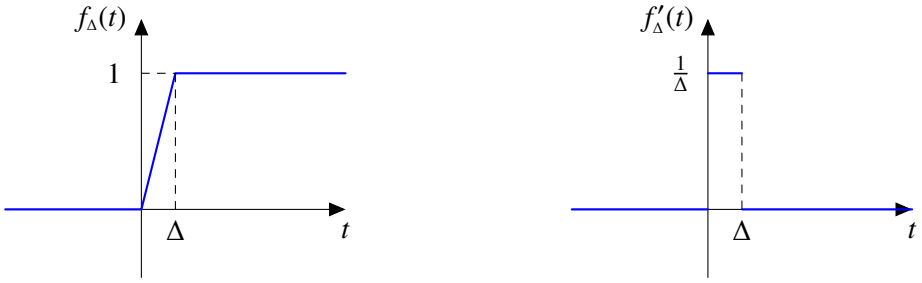
$$\begin{aligned} H : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \end{aligned}$$



Because of the discontinuity at $x = 0$, this function is not differentiable: only its restrictions to $(-\infty, 0)$ and $(0, +\infty)$ are differentiable. However, there is an alternative way of looking at the derivative, which will be developed in this section.
Consider a variable Δ, allowed to vary with continuity in $\mathbb{R}_+$. Consider the function f_Δ defined as

$$\begin{aligned} f_\Delta : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} 0 & \text{if } t < 0 \\ \frac{1}{\Delta} t & \text{if } t \in [0, \Delta] \\ 1 & \text{if } t > \Delta \end{cases} \end{aligned}$$

The graph of the function f_Δ is given below, along with the graph its derivative, where the derivative is defined:



We notice that f_Δ is continuous, but it is not differentiable. It is differentiable in the three restrictions to $(-\infty, 0)$, $(0, \Delta)$ and $(\Delta, +\infty)$. We observe that, for any $\Delta > 0$, the area defined by the graph of f'_Δ is 1. If we imagine to let Δ go to zero, $f_\Delta(t)$ tends to the unit step H, and its derivative continues to have unit area, but it tends to have infinite amplitude and infinitesimal duration: in other words, in intuitive terms it tends to become an impulse.

As usual, this limit process is ill-posed, but it is useful to gain intuition: by developing a theory of differentiation for distributions, we will be able to claim that the derivative of the unit step exists in the sense of distributions, and that derivative is exactly δ:

$$H' = \delta.$$

This will not only be true for the unit step. Other functions that do not have the derivative in the classical sense (absolute value, sign function, ceiling/floor functions, as well as other less obvious functions such as $\ln|x|$) will have a derivative if we interpret them as distributions, and this has tremendous repercussions in the context of differential equations.

As usual, we first examine the derivative when f is locally summable, and we start from a simple case: the case where $n = 1$, and the domain of our function is an open interval I of $\mathbb{R}$. We consider $f \in C^1(I; \mathbb{R})$ and $\varphi \in \mathcal{D}(I; \mathbb{R})$. Since φ has compact support in I, there exist a closed interval $[a, b]$ in I (with $a, b \in I$ satisfying $a < b$) such that, if $x \notin [a, b]$, we have $\varphi(x) = 0$. Since φ is itself of class C^∞, in particular $\varphi(a) = \varphi(b) = 0$. Then, using integration by parts:

$$\langle L_{f'}, \varphi \rangle = \int_I f'(x)\,\varphi(x)\,\mathrm{d}x = \int_a^b f'(x)\,\varphi(x)\,\mathrm{d}x = [f(x)\,\varphi(x)]_a^b - \int_I f(x)\,\varphi'(x)\,\mathrm{d}x = \langle L_f, \varphi' \rangle.$$

We can generalize this argument by considering the domain of f to be a generic open subset of $\mathbb{R}$. We can write it as the union of a countable family of open pairwise disjoint intervals as

$$A = \bigcup_{k=1}^{\infty} I_k.$$

Thus,

$$\begin{aligned}
\int_A f'(x)\,\varphi(x)\,\mathrm{d}x &= \int_{\bigcup_{k=1}^{\infty} I_k} f'(x)\,\varphi(x)\,\mathrm{d}x \\
&= \sum_{k=1}^{\infty} \int_{I_k} f'(x)\,\varphi(x)\,\mathrm{d}x \\
&= \sum_{k=1}^{\infty} \int_{a_k}^{b_k} f'(x)\,\varphi(x)\,\mathrm{d}x \\
&= \sum_{k=1}^{\infty} \left([f(x)\,\varphi(x)]_{x=a_k}^{x=b_k} - \int_{a_k}^{b_k} f(x)\,\varphi'(x)\,\mathrm{d}x \right) \\
&= -\sum_{k=1}^{\infty} \int_{a_k}^{b_k} f(x)\,\varphi'(x)\,\mathrm{d}x \\
&= -\sum_{k=1}^{\infty} \int_{I_k} f(x)\,\varphi'(x)\,\mathrm{d}x \\
&= -\int_A f(x)\,\varphi'(x)\,\mathrm{d}x.
\end{aligned}$$

Notice that the integral over the countable union of $(I_k)_{k\in\mathbb{N}}$ has been written as the countable sum of integrals. This step is justified by Theorem 52 in the set of notes on the Lebesgue integral.

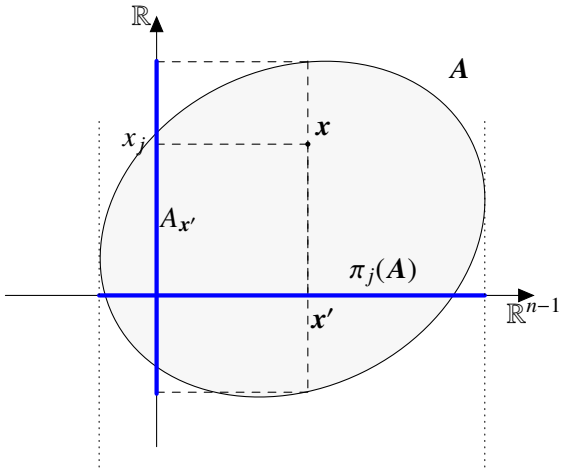
We now want to generalize the previous arguments to the case $n > 1$. Let us assume $f \in C^1(\boldsymbol{A};\mathbb{C})$. Since f is continuous, it is also locally summable (see Theorem 5.6). Thus, f is associated with a distribution of function type L_f; since its derivatives are also continuous, we can associate to them distributions of function type. Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. For $j \in \{1,\dots,n\}$, we now need to evaluate

$$\left\langle L_{\frac{\partial f}{\partial x_j}}, \varphi \right\rangle = \int_{\boldsymbol{A}} \frac{\partial f}{\partial x_j}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$. We define $\boldsymbol{x}' = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$. We also define the function

$$\pi_j(x_1,\dots,x_{j-1},x_j,x_{j+1},\dots,x_n) = (x_1,\dots,x_{j-1},x_{j+1},\dots,x_n).$$

Finally, for $\boldsymbol{x}' \in \pi_j(\boldsymbol{A})$, we define $A_{\boldsymbol{x}'} = \{x_j \in \mathbb{R} \,|\, (x_1,\dots,x_j,\dots,x_n) \in \boldsymbol{A}\}$.



Applying Fubini's theorem (Theorem 3.8) we find

$$\int_{\boldsymbol{A}} \frac{\partial f}{\partial x_j}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\pi_j(\boldsymbol{A})} \Big(\int_{A_{\boldsymbol{x}'}} \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \Big) \mathrm{d}\boldsymbol{x}'. \tag{5.48}$$

It is easy to see that for all $\boldsymbol{x}' \in \pi_j(\boldsymbol{A})$ the set $A_{\boldsymbol{x}'} \subseteq \mathbb{R}$ is open,[8] and therefore it can be written as the union of a finite or countable family of open pairwise disjoint intervals of $\mathbb{R}$, i.e.,

$$A_{\boldsymbol{x}'} = \bigcup_{\ell=0}^{\infty} I_\ell, \tag{5.49}$$

where I_ℓ are open pairwise disjoint intervals (dependent on $\boldsymbol{x}'$) of $\mathbb{R}$. Moreover, the function

$$x_j \mapsto \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)$$

is continuous and with compact support in $A_{\boldsymbol{x}'}$, and it is therefore summable over $A_{\boldsymbol{x}'}$.

Consider the inner integral of (5.48), that we consider as an integral over the union (5.49), and therefore, using Theorem 3.50, we can write it as a sum of integrals. Moreover, using the fact that for each ℓ we have $\varphi(\boldsymbol{x}', x_j) = 0$ if x_j does not belong to a compact $[a_\ell, b_\ell] \subseteq I_\ell$, it is found that

$$\begin{aligned}
\int_{A_{\boldsymbol{x}'}} \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)\,\mathrm{d}x_j &= \int_{\bigcup_{\ell=0}^{\infty} I_\ell} \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \\
&= \sum_{\ell=0}^{\infty} \int_{I_\ell} \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \\
&= \sum_{\ell=0}^{\infty} \int_{a_\ell}^{b_\ell} \frac{\partial f}{\partial x_j}(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \\
&= \sum_{\ell=0}^{\infty} \Big(\cancel{\Big[f(\boldsymbol{x}', x_j)\,\varphi(\boldsymbol{x}', x_j) \Big]_{a_\ell}^{b_\ell}} - \int_{a_\ell}^{b_\ell} f(\boldsymbol{x}', x_j)\,\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \Big) \\
&= -\sum_{\ell=0}^{\infty} \int_{a_\ell}^{b_\ell} f(\boldsymbol{x}', x_j)\,\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \\
&= -\int_{A_{\boldsymbol{x}'}} f(\boldsymbol{x}', x_j)\,\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}', x_j)\,\mathrm{d}x_j.
\end{aligned}$$

Therefore, if $f \in C^1(A;\mathbb{C})$ and $\varphi \in \mathcal{D}(A;\mathbb{C})$, we find

$$\int_A \frac{\partial f}{\partial x_j}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = -\int_{\pi_j(A)} \Big(\int_{A_{\boldsymbol{x}'}} f(\boldsymbol{x}', x_j)\,\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}', x_j)\,\mathrm{d}x_j \Big) \mathrm{d}\boldsymbol{x}' = -\int_A f(\boldsymbol{x})\,\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

[8] Let $\boldsymbol{A} \subseteq \mathbb{R}^{m+n}$ be open, and denote its points as $(\boldsymbol{x}, \boldsymbol{y})$ with $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{y} \in \mathbb{R}^n$. Let $\boldsymbol{\xi} \in \mathbb{R}^m$, and $\boldsymbol{A}_{\boldsymbol{\xi}} = \{y \in \mathbb{R}^n \,|\, (\boldsymbol{\xi}, y) \in \boldsymbol{A}\}$. We show that $\boldsymbol{A}_{\boldsymbol{\xi}}$ is open in $\mathbb{R}^n$. Consider an arbitrary $\boldsymbol{y} \in \boldsymbol{A}_{\boldsymbol{\xi}}$, so that $(\boldsymbol{\xi}, \boldsymbol{y}) \in \boldsymbol{A}$. Since $\boldsymbol{A}$ is open, there exists $r > 0$ such that $\mathcal{B}_r((\boldsymbol{\xi}, \boldsymbol{y})) \subseteq \boldsymbol{A}$. We want to show that $\mathcal{B}_r(y) \subseteq \boldsymbol{A}_{\boldsymbol{\xi}}$. Consider a point $z \in \mathcal{B}_r(y)$. We have

$$\|(\boldsymbol{\xi}, z) - (\boldsymbol{\xi}, \boldsymbol{y})\|_1 = \|(0, z - \boldsymbol{y})\|_1 \leq \|z - \boldsymbol{y}\|_1 < r,$$

where the norm in $\|(0, z - \boldsymbol{y})\|_1$ is the 1-norm in $\mathbb{R}^{m+n}$ and the norm in $\|z - \boldsymbol{y}\|_1$ is the 1-norm in $\mathbb{R}^n$. It follows that $(\boldsymbol{\xi}, z) \in \mathcal{B}_r((\boldsymbol{\xi}, \boldsymbol{y})) \subseteq \boldsymbol{A}$, which implies $(\boldsymbol{\xi}, z) \in \boldsymbol{A}$, which in turn implies $z \in \boldsymbol{A}_{\boldsymbol{\xi}}$.

or, written in terms of distributions of function type,

$$\left\langle L_{\frac{\partial f}{\partial x_j}}, \varphi \right\rangle = -\left\langle L_f, \frac{\partial \varphi}{\partial x_j} \right\rangle,$$

which makes sense because $\frac{\partial \varphi}{\partial x_j}$ is still of class C^∞ with compact support.

These results suggest that a definition of derivative of a distribution can be given as follows.

Definition 5.25 [Derivative of Distributions]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let $j \in \{1, \dots, n\}$. We define $\partial_j T \stackrel{def}{=} \frac{\partial T}{\partial x_j} \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ as the functional such that for all $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$

$$\left\langle \partial_j T, \varphi \right\rangle = -\left\langle T, \partial_j \varphi \right\rangle.$$

Notice that, because of the way distributions have been defined and in view of Definition 5.25, for every distribution T the the derivatives $\partial_1 T, \partial_2 T, \dots, \partial_n T$ are defined. Therefore, we do not need to introduce a concept of differentiability. We will see that even more is true: every distribution is infinitely differentiable.

In the case $n = 1$, Definition 5.25 can be written as

$$\langle L_{f'}, \varphi \rangle \stackrel{\text{def}}{=} -\langle L_f, \varphi' \rangle \quad \forall \varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C}).$$

Notice that the definition is well posed since, when $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$, then also $\partial_j \varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$.

Theorem 5.22 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $f \in C^1(\boldsymbol{A}; \mathbb{C})$. Let $j \in \{1, \dots, n\}$. Then,*

$$\partial_j(L_f) = L_{\frac{\partial f}{\partial x_j}}.$$

Proof Let $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. Then, in view of the calculations shown above, we find

$$\langle \partial_j L_f, \varphi \rangle = -\left\langle L_f, \frac{\partial \varphi}{\partial x_j} \right\rangle = -\int_{\boldsymbol{A}} f(\boldsymbol{x}) \frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} \frac{\partial f}{\partial x_j}(\boldsymbol{x}) \varphi(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} = \left\langle L_{\frac{\partial f}{\partial x_j}}, \varphi \right\rangle.$$

We now verify that $\partial_j T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$.

Theorem 5.23 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let $j \in \{1, \dots, n\}$. Then, $\partial_j T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$.*

Proof It is straightforward to verify that $\partial_j T$ is linear (try!). We prove the continuity. Let $(\varphi_k)_{k \in \mathbb{N}}$ be a sequence in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$ which converges to $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$. We need to prove that

$$\lim_{k \to \infty} \left\langle \partial_j T, \varphi_k \right\rangle = \left\langle \partial_j T, \varphi \right\rangle,$$

which can be written as

$$-\left\langle T, \partial_j \varphi_k \right\rangle \xrightarrow{k \to \infty} -\left\langle T, \partial_j \varphi \right\rangle.$$

To prove this, it suffices to prove that

$$\partial_j \varphi_k \xrightarrow{k\to\infty} \partial_j \varphi$$

in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. In other words, we have to show that if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, then $(\frac{\partial\varphi_k}{\partial x_j})_{k\in\mathbb{N}}$ converges to $\frac{\partial\varphi_k}{\partial x_j}$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Indeed, there exists a compact set $\boldsymbol{K}\subset\boldsymbol{A}$ such that $\operatorname{supp}\varphi_k\subseteq\boldsymbol{K}$ for all $k\in\mathbb{N}$, from which we find $\operatorname{supp}(\frac{\partial\varphi_k}{\partial x_j})\subseteq\operatorname{supp}\varphi_k\subset\boldsymbol{K}$. Consider the multi-index $\boldsymbol{\alpha}\in\mathbb{N}^n$. We have

$$\partial^\alpha\left(\frac{\partial\varphi_k}{\partial x_j}\right)=\partial^{\alpha+\boldsymbol{e}_j}\varphi_k,$$

where $\boldsymbol{e}_j$ is the j-th vector of the standard basis of $\mathbb{R}^n$, so that

$$\partial^{\alpha+\boldsymbol{e}_j}\varphi_k\xrightarrow{k\to\infty}\partial^{\alpha+\boldsymbol{e}_j}\varphi$$

uniformly, in view of the convergence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Thus,

$$-\left\langle T,\frac{\partial\varphi_k}{\partial x_j}\right\rangle\xrightarrow{k\to\infty}-\left\langle T,\frac{\partial\varphi}{\partial x_j}\right\rangle=\langle\partial_j T,\varphi\rangle.$$

Example *5.39*

Consider the absolute value function

$$\begin{aligned}f:\mathbb{R}&\longrightarrow\mathbb{R}\\ t&\mapsto|t|,\end{aligned}$$

which is not differentiable. However, f is continuous and therefore it is locally summable. We can thus view it as a function-type distribution. We want to show that f has the derivative if we consider it as a distribution, and the derivative of f in the sense of distributions is the sign function. Let $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$ satisfy $\operatorname{supp}\varphi\subseteq[-M,M]$ for $M>0$. Then

$$\begin{aligned}\langle L_{f'},\varphi\rangle&=-\langle L_f,\varphi'\rangle\\&=-\int_{\mathbb{R}}|t|\,\varphi'(t)\,\mathrm{d}t\\&=\int_{-\infty}^{0}t\,\varphi'(t)\,\mathrm{d}t-\int_0^{\infty}t\,\varphi'(t)\,\mathrm{d}t.\end{aligned}$$

Then, integrating by parts gives

$$\int_{-\infty}^{0}t\,\varphi'(t)\,\mathrm{d}t=\int_{-M}^{0}t\,\varphi(t)\,\mathrm{d}t=[t\,\varphi(t)]_{-M}^{0}-\int_{-M}^{0}\varphi(t)\,\mathrm{d}t=-\int_{-\infty}^{0}\varphi(t)\,\mathrm{d}t,$$

and

$$\int_0^{+\infty}t\,\varphi(t)\,\mathrm{d}t=\int_0^{M}t\,\varphi(t)\,\mathrm{d}t=[t\,\varphi(t)]_0^{M}-\int_0^{M}\varphi(t)\,\mathrm{d}t=-\int_0^{M}\varphi(t)\,\mathrm{d}t=-\int_0^{+\infty}\varphi(t)\,\mathrm{d}t$$

so that

$$\langle L_{f'},\varphi\rangle=\int_{-\infty}^{0}-\varphi(t)\,\mathrm{d}t+\int_0^{\infty}\varphi(t)\,\mathrm{d}t=\int_{-\infty}^{+\infty}\operatorname{sgn}(t)\,\varphi(t)\,\mathrm{d}t.$$

Thus, $f'=\operatorname{sgn}$, i.e., $\partial|t|=\operatorname{sgn}(t)$ in the sense of distributions. We also observe that in the classical sense we can write $\frac{\mathrm{d}}{\mathrm{d}t}|t|=\operatorname{sgn}(t)$ almost everywhere on $\mathbb{R}$, i.e., everywhere on $\mathbb{R}\setminus\{0\}$.

Example 5.40

We want to show that the derivative of the Heaviside function as a distribution is the Dirac delta distribution. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then,

$$\langle L_{H'}, \varphi\rangle = -\langle L_H, \varphi'\rangle = -\int_0^{+\infty} \varphi'(t)\,\mathrm{d}t = -[\varphi(t)]_0^{+\infty} = \varphi(0) = \langle \delta, \varphi\rangle,$$

so that $H' = \delta$ in the sense of distributions.

Example 5.41

It is easy to prove that the identity $H' = \delta$ can be seen as a limit (in the distributional sense) of a difference quotient:

$$\lim_{h\to 0} \frac{H(t+h) - H(t)}{h} = \delta.$$

Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then, for $h \in \mathbb{R}$, using the linearity of the Lebesgue integral and a change of variable, we find

$$\begin{aligned}
\left\langle L_{\frac{H(t+h)-H(t)}{h}}, \varphi \right\rangle &= \int_{\mathbb{R}} \frac{H(t+h) - H(t)}{h}\,\varphi(t)\,\mathrm{d}t = \frac{1}{h}\left(\int_{\mathbb{R}} H(t+h)\,\varphi(t)\,\mathrm{d}t - \int_{\mathbb{R}} H(t)\,\varphi(t)\,\mathrm{d}t\right)\\
&= \frac{1}{h}\left(\int_{\mathbb{R}} H(s)\,\varphi(s-h)\,\mathrm{d}s - \int_{\mathbb{R}} H(t)\,\varphi(t)\,\mathrm{d}t\right)\\
&= \int_{\mathbb{R}} H(t)\,\frac{\varphi(t-h) - \varphi(t)}{h}\,\mathrm{d}t\\
&= \int_0^{+\infty} \frac{\varphi(t-h) - \varphi(t)}{h}\,\mathrm{d}t \xrightarrow{h\to 0} \int_0^{+\infty} -\varphi'(t)\,\mathrm{d}t = [-\varphi(t)]_0^{+\infty} = \varphi(0) = \langle \delta, \varphi\rangle,
\end{aligned}$$

where one can use Lebesgue's dominated convergence theorem (Theorem 3.47) to justify the interchange of the limit and the integral, writing

$$\lim_{h\to 0} \int_0^{+\infty} \frac{\varphi(t-h) - \varphi(t)}{h}\,\mathrm{d}t = \lim_{k\to\infty} \int_0^{+\infty} \frac{\varphi\left(t - \frac{1}{k}\right) - \varphi(t)}{\frac{1}{k}}\,\mathrm{d}t,$$

noting that the integrand is pointwise convergent to $-\varphi'(t)$ and that the absolute value of the integrand is bounded by a constant function. Indeed, if $\operatorname{supp}\varphi \subseteq [-M, M]$ for $M > 0$, then $\varphi\left(t - \frac{1}{k}\right) = 0$ for $t > M + \frac{1}{k}$, so that the support of the integrand function is in $[0, M+1]$, and for all $k \in \mathbb{N}^*$

$$\left|\frac{\varphi\left(t - \frac{1}{k}\right) - \varphi(t)}{\frac{1}{k}}\right| \leq \max_{t\in[0,M+1]} \left|\frac{\varphi\left(t - \frac{1}{k}\right) - \varphi(t)}{\frac{1}{k}}\right|.$$

Example 5.42

Let $A = (0, 3)$, and consider the floor function

$$\begin{aligned}
f : A &\longrightarrow \mathbb{R}\\
t &\mapsto \lfloor t\rfloor = \max\{n \in \mathbf{Z} \mid n \leq t\}.
\end{aligned}$$

We want to compute f' in the sense of distributions. To this end, let $\varphi \in \mathcal{D}(A;\mathbb{R})$. Then

$$\begin{aligned}
\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi'\rangle = -\int_0^3 \lfloor t\rfloor\,\varphi'(t)\,\mathrm{d}t\\
&= -\left(\int_0^1 0\cdot\varphi'(t)\,\mathrm{d}t + \int_1^2 1\cdot\varphi'(t)\,\mathrm{d}t + \int_2^3 2\cdot\varphi'(t)\,\mathrm{d}t\right)\\
&= -(\varphi(2) - \varphi(1) + \cancel{2\,\varphi(3)} - 2\,\varphi(2))\\
&= \varphi(1) + \varphi(2) = \langle \delta(t-1), \varphi(t)\rangle + \langle \delta(t-2), \varphi(t)\rangle.
\end{aligned}$$

Thus, $f'(t) = \delta(t-1) + \delta(t-2)$.

Example 5.43

We show that $\mathrm{pv}\,\frac{1}{x}$ can be viewed as the distributional derivative of the function $f(x) = \ln|x|$. This function is defined over $\mathbb{R} \setminus \{0\}$, but we define a function that gives f an arbitrary value at $x = 0$, for example 0. In other words, we define the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} \ln|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases} \end{aligned}$$

We show that f is locally summable. Since by considering the integrability over compact sets the only problem of possible divergence is at the origin, it is sufficient to show that

$$\forall\, M > 0, \qquad \int_{-M}^{M} \big|\ln|x|\big|\,\mathrm{d}x < +\infty.$$

We can assume with no loss of generality that $M \leq 1$, so that $0 < |x| \leq 1$. Thus, $\big|\ln|x|\big| = -\ln|x|$, and so

$$\int_{-M}^{M} \big|\ln|x|\big|\,\mathrm{d}x = 2\int_{0}^{M} \big[-\ln|x|\big]\,\mathrm{d}x = 2\lim_{j\to+\infty} \int_{(0,M]} \chi_{\left[\frac{1}{j},M\right]}(x)\big[-\ln|x|\big]\,\mathrm{d}x$$

where the last equality follows from Beppo Levi's monotone convergence theorem, and therefore[a]

$$\int_{-M}^{M} \big|\ln|x|\big|\,\mathrm{d}x = 2\lim_{j\to+\infty} \int_{\frac{1}{j}}^{M} -\ln x\,\mathrm{d}x = 2\lim_{j\to+\infty} [x\ln x - x]_{\frac{1}{j}}^{M} = 2M(1-\ln M) < +\infty.$$

Therefore $[f] \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Now, let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\langle L_{f'}, \varphi\rangle = -\langle L_f, \varphi'\rangle = -\int_{\mathbb{R}} \ln|x|\,\varphi'(x)\,\mathrm{d}x = -\lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} \ln|x|\,\varphi'(x)\,\mathrm{d}x$$

from Lebesgue's dominated convergence theorem (Theorem 3.47). We obtain, breaking down the integral and using integration by parts

$$\begin{aligned} \langle L_{f'}, \varphi\rangle &= -\lim_{\varepsilon\to 0^+} \left(\int_{[\varepsilon,+\infty)} \ln x\,\varphi'(x)\,\mathrm{d}x + \int_{(-\infty,-\varepsilon]} \ln(-x)\,\varphi'(x)\,\mathrm{d}x \right) \\ &= -\lim_{\varepsilon\to 0^+} \left(-\ln\varepsilon\,\varphi(\varepsilon) - \int_{[\varepsilon,+\infty)} \frac{\varphi(x)}{x}\,\mathrm{d}x + \ln\varepsilon\,\varphi(-\varepsilon) - \int_{(-\infty,-\varepsilon]} \frac{\varphi(x)}{x}\,\mathrm{d}x \right) \\ &= -\lim_{\varepsilon\to 0^+} \left(\ln(\varepsilon)\,(\varphi(\varepsilon) - \varphi(-\varepsilon)) - \int_{|x|\geq\varepsilon} \frac{\varphi(x)}{x}\,\mathrm{d}x \right). \end{aligned}$$

Applying Taylor with initial point 0 we find

$$\ln(\varepsilon)\,(\varphi(\varepsilon) - \varphi(-\varepsilon)) = \ln(\varepsilon)\,(\varphi(0) + \varphi'(0)\,\varepsilon + O(\varepsilon^2) - \varphi(0) - \varphi'(0)\,\varepsilon + O(\varepsilon^2)) = \ln\varepsilon\,O(\varepsilon^2),$$

which goes to 0 as $\varepsilon \to 0^+$.

[a] We recall that a primitive of $\ln x$ is $x\ln x - x$, as one can obtain by integrating by parts.

Higher order derivatives

Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. For $i, j \in \{1,\dots,n\}$ we define $\partial_i\partial_j T = \partial_i(\partial_j T)$, or, more explicitly,

$$\frac{\partial^2 T}{\partial x_i\,\partial x_j} \stackrel{\text{def}}{=} \frac{\partial}{\partial x_i}\left(\frac{\partial T}{\partial x_j}\right).$$

We recall that, given a function $f : \boldsymbol{A} \longrightarrow \mathbb{R}$ or $f : \boldsymbol{A} \longrightarrow \mathbb{C}$, if the second partial derivatives exist and they are continuous, then we can exchange the order of integration. For distributions the continuity of the second partial derivatives is not needed, because the second partial derivatives of a distribution are brought back to the second partial derivatives of the function φ, which is of class C^∞, and therefore this condition is automatically satisfied.

Theorem 5.24 [Generalized Schwarz-Clairaut's Theorem]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(A;\mathbb{C})$. For $i, j \in \{1,\dots,n\}$ there holds

$$\partial_i\,\partial_j T = \partial_j\,\partial_i T.$$

Proof Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. We find

$$\begin{aligned}\left\langle \partial_i\,\partial_j\,T,\varphi\right\rangle &= -\left\langle \partial_j T,\partial_i\varphi\right\rangle\\ &= \left\langle T,\partial_j\,\partial_i\,\varphi\right\rangle\\ &= \left\langle T,\partial_i\,\partial_j\,\varphi\right\rangle\\ &= \left\langle \partial_j\,\partial_i\,T,\varphi\right\rangle,\end{aligned}$$

where the classical version of Schwarz-Clairaut's Theorem has been applied to φ: this is allowed because φ is in $C^\infty(\boldsymbol{A};\mathbb{C})$.

In terms of multi-indices, we can write $\partial_i\,\partial_j\,T = \partial^{\boldsymbol{\alpha}}\,T$, where $\boldsymbol{\alpha} = (0,\dots,1,0,\dots,1,\dots,0)$, where the terms equal to 1 are those in positions i and j (i.e., $\boldsymbol{\alpha} = \boldsymbol{e}_i + \boldsymbol{e}_j$). If $i = j$, then $\boldsymbol{\alpha} = (0,\dots,0,2,0,\dots,0)$. We can define derivatives of any order using the multi-index $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\dots,\alpha_n)$:

$$\partial^{\boldsymbol{\alpha}} T = \partial_1^{\alpha_1}\,\partial_2^{\alpha_2}\,\dots\,\partial_n^{\alpha_n}\,T.$$

Thus, for $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, we have

$$\langle \partial^{\boldsymbol{\alpha}} T,\varphi\rangle = \left\langle \partial_1^{\alpha_1}\,\partial_2^{\alpha_2}\,\dots\,\partial_n^{\alpha_n}\,T,\varphi\right\rangle = -\left\langle \partial_1^{\alpha_1-1}\,\partial_2^{\alpha_2}\,\dots\,\partial_n^{\alpha_n}\,T,\partial_1\varphi\right\rangle.$$

One by one, we can shift all the derivatives from the distribution to the function φ. It is now straightforward to see that, continuing in this fashion, for any $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ there holds

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

because every time a derivative is shifted, the sign changes.

Theorem 5.25 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of distributions in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let*

$T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. *Then,*

$$T_k \xrightarrow{k\to\infty} T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \quad \Rightarrow \quad \forall\, \boldsymbol{\alpha} \in \mathbb{N}^n,\ \ \partial^\alpha T_k \xrightarrow{k\to\infty} \partial^\alpha T \text{ in } \mathcal{D}'(\boldsymbol{A};\mathbb{C}).$$

The proof is left as an exercise, see Question **??**.

A similar result holds for series: if $\sum_{k=0}^{\infty} T_k$ is a convergent series in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, then for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ the series $\sum_{k=0}^{\infty} \partial^\alpha T_k$ converges in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, and

$$\sum_{k=0}^{\infty} \partial^\alpha T_k = \partial^\alpha \sum_{k=0}^{\infty} T_k.$$

This result can be proved by applying the previous theorem to the partial sums (see Question **??**).

Derivative of the product

Consider $\psi \in C^\infty(\boldsymbol{A};\mathbb{C})$ and $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We show that

$$\partial_j(\psi\, T) = (\partial_j\psi)\, T + \psi\, (\partial_j T).$$

Indeed, let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then,

$$\begin{aligned}
\langle \partial_j(\psi\, T), \varphi\rangle &= -\langle \psi\, T, \partial_j\varphi\rangle \\
&= -\langle T, \psi\, (\partial_j\varphi)\rangle \\
&= -\langle T, \partial_j(\psi\,\varphi) - (\partial_j\psi)\,\varphi\rangle \\
&= \langle (\partial_j\psi)\, T, \varphi\rangle + \langle \partial_j T, \psi\,\varphi\rangle \\
&= \langle (\partial_j\psi)\, T + \psi\, (\partial_j\, T), \varphi\rangle,
\end{aligned}$$

where we have used the fact that $\psi\,(\partial_j\varphi) = \partial_j(\psi\,\varphi) - (\partial_j\psi)\,\varphi$.

Derivative of the composition

We only consider the one-dimensional case. Let $\boldsymbol{A}, \boldsymbol{A}'$ be open sets in $\mathbb{R}$. Let $F : \boldsymbol{A}' \longrightarrow \boldsymbol{A}$ be a diffeomorphism of class C^∞ (i.e., F is a bijection and $F'(x') \neq 0$ for all $x' \in \boldsymbol{A}'$). Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We expect to find

$$(T \circ F)' = F'(T' \circ F). \tag{5.50}$$

Indeed, let $\varphi \in \mathcal{D}(\boldsymbol{A}';\mathbb{C})$. We notice that

$$(F^{-1})'(x) = \operatorname{sgn}((F^{-1})'(x))\, |(F^{-1})'(x)|$$

for all $x \in \boldsymbol{A}$, which yields $(F^{-1})'(x) = \operatorname{sgn} F'\, |(F^{-1})'(x)|$ (in the sign we do not indicate the point because it is a constant value: to change sign, since F^{-1} is C^∞, we should attain the value 0, and that would contradict our assumption that the Jacobian is invertible). Thus,

$$|(F^{-1})'| = \operatorname{sgn} F'\, (F^{-1})'. \tag{5.51}$$

We notice also that

$$(\varphi \circ F^{-1})' = (F^{-1})'\,(\varphi' \circ F^{-1}). \tag{5.52}$$

Finally, the theorem of the derivative of the inverse function gives

$$(F^{-1})'(x) = \frac{1}{F'(F^{-1}(x))} \quad \Rightarrow \quad (F^{-1})' = \frac{1}{F' \circ F^{-1}}. \tag{5.53}$$

Then

$$\begin{aligned}
\langle (T \circ F)', \varphi \rangle &= -\langle T \circ F, \varphi' \rangle \\
&= -\langle T, (\varphi' \circ F^{-1}) |(F^{-1})'| \rangle \qquad \text{using (5.51)} \\
&= -\langle T, (\varphi' \circ F^{-1})((F^{-1})')\ \operatorname{sgn}(F') \rangle \qquad \text{using (5.52)} \\
&= -\operatorname{sgn}(F') \langle T, (\varphi \circ F^{-1})' \rangle \\
&= \operatorname{sgn}(F') \langle T', \varphi \circ F^{-1} \rangle \\
&= \langle T', (\varphi \circ F^{-1})(F' \circ F^{-1})(F^{-1})' \rangle \operatorname{sgn}(F') \qquad \text{using (5.53)} \\
&= \langle T', ((\varphi\, F') \circ F^{-1})\, |(F^{-1})'| \rangle \\
&= \langle T' \circ F, \varphi\, F' \rangle = \langle F'(T' \circ F), \varphi \rangle,
\end{aligned}$$

where we used the fact that

$$(\varphi \circ F^{-1})(F' \circ F^{-1}) = (\varphi\, F') \circ F^{-1},$$

which is the same as

$$\varphi(F^{-1}(y))\, F'(F^{-1}(y)) = (\varphi\, F')(F^{-1}(y)),$$

and this is simply the definition of product of two functions.

Example *5.44* There holds

$$(\delta(2\,t))' = 2\,\delta'(2\,t).$$

Derivative as limit of the difference quotient

Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{R})$. We can therefore consider the distribution

$$T(x+h) = T \circ F, \qquad F(x) = x + h,$$

We expect

$$\lim_{h \to 0} \frac{T(x+h) - T(x)}{h} = T' \quad \text{in } \mathcal{D}'(\mathbb{R};\mathbb{R}).$$

Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We would like to show that

$$\lim_{h \to 0} \left\langle \frac{T(x+h) - T(x)}{h}, \varphi(x) \right\rangle = \langle T'(x), \varphi(x) \rangle \quad \text{in } \mathcal{D}'(\mathbb{R};\mathbb{R}).$$

Consider a sequence $(h_k)_{k\in\mathbb{N}}$ of elements of $\mathbb{R}\setminus\{0\}$ which converges to 0. We want to show that

$$\lim_{k\to\infty}\left\langle\frac{T(x+h_k)-T(x)}{h_k},\varphi(x)\right\rangle=\langle T'(x),\varphi(x)\rangle\quad\text{in }\mathcal{D}'(\mathbb{R};\mathbb{R}).$$

We have

$$\langle T(x+h_k)-T(x),\varphi\rangle=\langle T(y),\varphi(y-h_k)\cdot 1\rangle,$$

where 1 is the determinant of the Jacobian of the function $F(x)=x+h_k$, which gives $F^{-1}(y)=y-h_k$. It follows that

$$\left\langle\frac{T(x+h_k)-T(x)}{h_k},\varphi(x)\right\rangle=\left\langle T(y),\frac{\varphi(y-h_k)-\varphi(y)}{h_k}\right\rangle.$$

We can define the sequence $(\varphi_k)_{k\in\mathbb{N}}$ such that

$$\varphi_k(y)=\frac{\varphi(y-h_k)-\varphi(y)}{h_k},\tag{5.54}$$

and we show that $\varphi_k\xrightarrow{k\to\infty}-\varphi'$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$. There are two conditions to satisfy: first, let $R>0$ be such that $\operatorname{supp}\varphi\subseteq[-R,R]$. We notice that $\varphi_k(y)\neq 0$ implies $|y-h_k|\leq R$ or $|y|\leq R$. First, suppose $|y-h_k|\leq R$. We have

$$|y|=|(y-h_k)+h_k|\leq|y-h_k|+|h_k|\leq R+|h_k|\leq R+H,$$

where $|h_k|\leq H$ for all $k\in\mathbb{N}$ (in fact, since the sequence $(h_k)_{k\in\mathbb{N}}$ tends to 0, it is bounded). Thus, if $|y|>R+H$ we have that neither condition $|y-h_k|\leq R$ nor $|y|\leq R$ is satisfied, i.e., $\varphi_k(y)=0$. Therefore, we obtain

$$\forall\,k\in\mathbb{N},\qquad\operatorname{supp}\varphi_k\subseteq[-R-H,R+H].$$

The second condition that we must check is the uniform convergence of the derivatives. We start by showing that

$$\lim_{k\to\infty}\varphi_k=-\varphi'\quad\text{uniformly.}$$

We apply the fundamental theorem of calculus to (5.54):

$$\begin{aligned}\varphi_k(y)&=h_k^{-1}\,(\varphi(y-h_k)-\varphi(y))\\&=h_k^{-1}\int_y^{y-h_k}\varphi'(z)\,\mathrm{d}z.\end{aligned}$$

Subtracting $\varphi'(y)$ from both sides yields

$$\varphi_k(y)-\varphi'(y)=h_k^{-1}\int_y^{y-h_k}(\varphi'(z)-\varphi'(y))\,\mathrm{d}z.$$

We have

$$|\varphi_k(y)-\varphi'(y)|=\left|h_k^{-1}\int_y^{y-h_k}(\varphi'(z)-\varphi'(y))\,\mathrm{d}z\right|\leq|h_k|^{-1}\int_y^{y-h_k}\left|\varphi'(z)-\varphi'(y)\right|\mathrm{d}z.$$

From $\varphi'\in\mathcal{D}(\mathbb{R};\mathbb{R})$, we deduce that φ' is uniformly continuous. This means that

$$\forall\,\varepsilon>0,\quad\exists\,\delta>0:\qquad|y-z|\leq\delta\;\Rightarrow\;\left|\varphi'(z)-\varphi'(y)\right|\leq\varepsilon.$$

We can always choose k to be sufficiently large so that $h_k \leq \delta$, and therefore $|y - z| \leq \delta$. Therefore,

$$|\varphi_k(y) - \varphi'(y)| \leq |h_k|^{-1} \int_y^{y-h_k} \left|\varphi'(z) - \varphi'(y)\right| \mathrm{d}z \leq |h_k|^{-1} \varepsilon\, h_k = \varepsilon,$$

where ε is a bound for the integrand and $|h_k|$ is the length of the domain of integration. Since this is all independent of y, we obtain

$$\lim_{k \to \infty} \varphi_k = -\varphi' \quad \text{uniformly.}$$

The same process can be carried out for the derivatives. Given $\alpha \in \mathbb{N}^*$, we find

$$\varphi_k^{(\alpha)}(y) = \frac{\varphi^{(\alpha)}(y - h_k) - \varphi^{(\alpha)}(y)}{h_k} \xrightarrow{k \to \infty} -\varphi^{(\alpha+1)}(y) \quad \text{uniformly.}$$

This allows us to conclude.

Derivative of a homogeneous function

We have seen that if $\boldsymbol{A} = \mathbb{R}^n$ or $\boldsymbol{A} = \mathbb{R}^n \setminus \{0\}$, we said that a distribution $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ is homogeneous with degree of homogeneity $\lambda \in \mathbb{R}$ if, for all $t > 0$, we have

$$T(t\,\boldsymbol{x}) = t^\lambda\, T(\boldsymbol{x}),$$

where the argument is to be intended as the composition with the diffeomorphism

$$\begin{aligned} F : \boldsymbol{A} &\longrightarrow \boldsymbol{A} \\ \boldsymbol{x} &\mapsto t\,\boldsymbol{x}. \end{aligned}$$

We show that in this case the derivative $\partial_j T$ is homogeneous with degree of homogeneity $\lambda - 1$, i.e.,

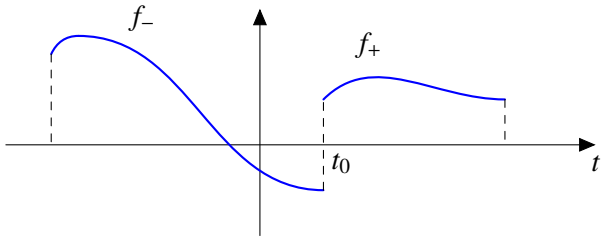
$$\forall\, t > 0, \qquad (\partial_j T)(t\,\boldsymbol{x}) = t^{\lambda-1} \partial_j T(\boldsymbol{x}).$$

Consider $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. The Jacobian of F^{-1} is t^{-n}, so that

$$\begin{aligned} \langle (\partial_j T)(t\,\boldsymbol{x}), \varphi(\boldsymbol{x}) \rangle &= \langle \partial_j T(\boldsymbol{y}), \varphi(t^{-1}\,\boldsymbol{y})\, t^{-n} \rangle \\ &= -\langle T(\boldsymbol{y}), \partial_j \varphi(t^{-1}\,\boldsymbol{y})\, t^{-n-1} \rangle \\ &= -t^{-1}\, \langle T(\boldsymbol{y}), \partial_j \varphi(t^{-1}\,\boldsymbol{y})\, t^{-n} \rangle \\ &= -t^{-1}\, \langle T(t\,\boldsymbol{y}), \partial_j \varphi(\boldsymbol{y}) \rangle \\ &= -t^{\lambda-1}\, \langle T(\boldsymbol{y}), \partial_j \varphi(\boldsymbol{y}) \rangle \\ &= t^{\lambda-1}\, \langle \partial_j T(\boldsymbol{y}), \varphi(\boldsymbol{y}) \rangle = \langle t^{\lambda-1}\, \partial_j T(\boldsymbol{y}), \varphi(\boldsymbol{y}) \rangle. \end{aligned}$$

Derivatives and "δ-jumps"

We have seen that discontinuous functions, such as the sign function or the Heaviside function, can admit a derivative in the sense of distributions. Now we examine this fact more closely.
Let $A = I$ be an open interval. Let $f : A \longrightarrow \mathbb{C}$. Let $t_0 \in I$. We assume that the restrictions of f to $I \cap (-\infty, t_0)$ and to $I \cap (t_0, +\infty)$ are of class C^1.



The left and right limits of f at t_0 exist in $\mathbb{C}$, i.e.,

$$\lim_{t \to t_0^\pm} f(t) = f(t_0^\pm).$$

Since the restrictions are of class C^1, the same is true for the first derivative:

$$\lim_{t \to t_0^\pm} f'(t) = f'(t_0^\pm).$$

From our assumptions, f is locally summable. We want to calculate the derivative in the sense of distributions. We will denote by ∂f the distributional derivative of f, and by $\mathrm{D}f$ the function that is coincident with the pointwise derivative, arbitrarily extended at $t = t_0$. Let $\varphi \in \mathcal{D}(I; \mathbb{C})$. We find

$$\begin{aligned}
\langle L_{\partial f}, \varphi \rangle &= -\langle L_f, \varphi' \rangle \\
&= -\int_I f(t)\,\varphi'(t)\,\mathrm{d}t \\
&= -\lim_{\varepsilon \to 0^+} \left(\int_{I \cap (-\infty, t_0 - \varepsilon]} f(t)\,\varphi'(t)\,\mathrm{d}t + \int_{I \cap [t_0 + \varepsilon, \infty)} f(t)\,\varphi'(t)\,\mathrm{d}t \right).
\end{aligned}$$

Let us consider the first integral in the right hand-side of the latter. Let $a, b \in I$ be such that $\operatorname{supp} \varphi \subseteq [a, b]$, and obviously $a \le t_0 \le b$. We find

$$\begin{aligned}
\int_{I \cap (-\infty, t_0 - \varepsilon]} f(t)\,\varphi'(t)\,\mathrm{d}t &= \int_a^{t_0 - \varepsilon} f(t)\,\varphi'(t)\,\mathrm{d}t \\
&= [f(t)\,\varphi(t)]_a^{t_0 - \varepsilon} - \int_a^{t_0 - \varepsilon} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t \\
&= f(t_0 - \varepsilon)\,\varphi(t_0 - \varepsilon) - \int_{I \cap (-\infty, t_0 - \varepsilon]} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t.
\end{aligned}$$

In the same way we get

$$\begin{aligned}
\int_{I \cap [t_0 + \varepsilon, \infty)} f(t)\,\varphi'(t)\,\mathrm{d}t &= \int_{[t_0 + \varepsilon, b)} f(t)\,\varphi'(t)\,\mathrm{d}t \\
&= [f(t)\,\varphi(t)]_{t_0 + \varepsilon}^b - \int_{[t_0 + \varepsilon, b)} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t \\
&= -f(t_0 + \varepsilon)\,\varphi(t_0 + \varepsilon) - \int_{I \cap [t_0 + \varepsilon, +\infty)} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t.
\end{aligned}$$

Thus, putting everything together and remembering that φ is continuous (so that $\varphi(t_0+\varepsilon)$ and $\varphi(t_0-\varepsilon)$ both tend to $\varphi(t_0)$ as $\varepsilon \to 0^+$), we obtain

$$\begin{aligned}\langle L_{\partial f}, \varphi\rangle &= \lim_{\varepsilon\to 0^+}\Big(\int_{I\cap(-\infty,t_0-\varepsilon]} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t + \int_{I\cap[t_0+\varepsilon,+\infty)} \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t \\ &\qquad + f(t_0+\varepsilon)\,\varphi(t_0+\varepsilon) - f(t_0-\varepsilon)\,\varphi(t_0-\varepsilon)\Big) \\ &= \int_I \mathrm{D}f(t)\,\varphi(t)\,\mathrm{d}t + \big(f(t_0^+) - f(t_0^-)\big)\,\varphi(t_0).\end{aligned}$$

Thus we can write

$$\partial f = \mathrm{D}f + \big(f(t_0^+) - f(t_0^-)\big)\,\delta(\cdot - t_0).$$

It follows that the derivative of f in the distributional sense is equal to the classical pointwise derivative of f extended arbitrarily at $t = t_0$ plus the jump of f at $t = t_0$ multiplied by the Dirac delta shifted by t_0. In the case $f(t) = |t|$, we have in particular $t_0 = 0$ and $f(0^+) - f(0^-) = 0$, so that

$$\partial f = \mathrm{D}f + 0\,\delta(t)$$

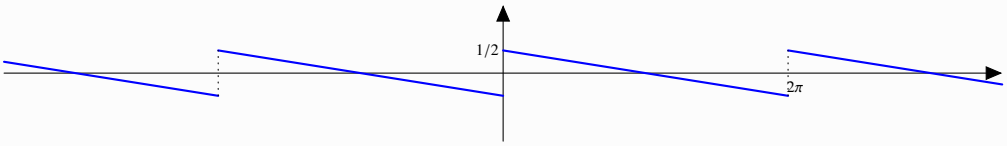
where the point-wise derivative $\mathrm{D}f$ of $|t|$ extended arbitrarily at $t = t_0$, is exactly $\mathrm{sgn}(t)$. In the case $f(t) = H(t)$, we have $t_0 = 0$, and the jump is $f(t_0^+) - f(t_0^-) = 1$, so that

$$\partial f = \mathrm{D}f + \big(f(0^+) - f(0^-)\big)\,\delta(t) = \delta(t).$$

Example *5.45*

Consider a 2π-periodic function f, which is equal to

$$f(x) = \frac{1}{2} - \frac{x}{2\pi}, \qquad \text{for } x \in [0, 2\pi).$$



The derivative in the sense of distributions of f is

$$f'(x) = -\frac{1}{2\pi} + \sum_{k=-\infty}^{+\infty} \delta(x - 2k\pi).$$

Example *5.46*

Let $\alpha \in \mathbb{R}$. Consider the function

$$\begin{aligned} f_\alpha : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^\alpha & \text{if } \boldsymbol{x} \neq 0 \\ 0 & \text{if } \boldsymbol{x} = 0 \end{cases}\end{aligned}$$

We want to find the values of α for which $[f_\alpha] \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$. We will find that the values are $\alpha > -n$. This is consistent with our intuition in the case $n = 1$. In fact, in this case we know that $|x|^\alpha$ is summable over closed and bounded sets if and only if $\alpha > -1$ (the only singularity we need to take care of is the origin). In $\mathbb{R}^n$, we must determine the values of α for which the integral of f_α is finite over every compact set of $\mathbb{R}^n$. The problem of integrability only occurs at the origin, so that we can simply show that for $\alpha > -n$ we have

$$\int_K \|\boldsymbol{x}\|^\alpha\,\mathrm{d}\boldsymbol{x} < +\infty$$

when $\boldsymbol{K} \subseteq \mathbb{R}^n$ is a compact set such that there exists $R > 0$ for which $\boldsymbol{K} \subseteq \mathcal{B}_R(0)$. In other words, it suffices to show that

$$\forall\, R > 0, \quad \int_{\overline{\mathcal{B}_R(0)}} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} < +\infty \qquad \Leftrightarrow \qquad \alpha > -n.$$

It is clear that, if we find a value R_1 for which this is true, then the statement will be true for every $R > 0$. Indeed, suppose that the claim is true for $R = 1$, i.e., suppose that

$$\int_{\overline{\mathcal{B}_1(0)}} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} < +\infty.$$

We want to show that for any $R > 0$ the statement is still true, i.e., $\int_{\overline{\mathcal{B}_R(0)}} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} < +\infty$. A change of variable $\frac{\boldsymbol{x}}{R} = \boldsymbol{y}$ gives

$$\int_{\overline{\mathcal{B}_R(0)}} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} = \int_{\overline{\mathcal{B}_1(0)}} R^\alpha \, \|\boldsymbol{y}\|^\alpha \, R^n \, \mathrm{d}\boldsymbol{y} = R^{\alpha+n} \int_{\|\boldsymbol{y}\| \leq 1} \|\boldsymbol{y}\|^\alpha \, \mathrm{d}\boldsymbol{y} < +\infty.$$

Thus, we are not losing any generality if we only investigate the existence of the integral

$$\int_{\|\boldsymbol{x}\| \leq 1} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x}.$$

Let us define, for all $k \in \mathbb{N}^*$, the set

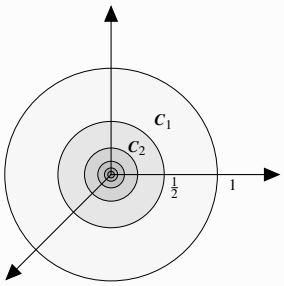
$$\boldsymbol{C}_k = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, 2^{-k} < \|\boldsymbol{x}\| \leq 2^{1-k}\}.$$

More explicitly, we have

$$\boldsymbol{C}_1 = \left\{\boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \frac{1}{2} < \|\boldsymbol{x}\| \leq 1\right\}, \quad \boldsymbol{C}_2 = \left\{\boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \frac{1}{4} < \|\boldsymbol{x}\| \leq \frac{1}{2}\right\}, \quad \boldsymbol{C}_3 = \left\{\boldsymbol{x} \in \mathbb{R}^n \,\middle|\, \frac{1}{8} < \|\boldsymbol{x}\| \leq \frac{1}{4}\right\}, \quad \ldots$$

We have

$$\{\boldsymbol{x} \in \mathbb{R}^n \,|\, 0 < \|\boldsymbol{x}\| \leq 1\} = \bigcup_{k=1}^{\infty} \boldsymbol{C}_k.$$



Since the sets $\boldsymbol{C}_k$ are pairwise disjoint, we have

$$\int_{0<\|\boldsymbol{x}\|\leq 1} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} = \sum_{k=1}^{\infty} \int_{\boldsymbol{C}_k} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x}.$$

On the other hand,

$$\begin{aligned} \int_{\boldsymbol{C}_k} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} &= \int_{2^{-k} < \|\boldsymbol{x}\| \leq 2^{1-k}} \|\boldsymbol{x}\|^\alpha \, \mathrm{d}\boldsymbol{x} \\ &\overset{2^k \boldsymbol{x} = \boldsymbol{y}}{=} \int_{1 < \|\boldsymbol{y}\| \leq 2} 2^{-k\alpha} \|\boldsymbol{y}\|^\alpha \, 2^{-nk} \mathrm{d}\boldsymbol{y} = A \, 2^{-(\alpha+n)k}, \end{aligned}$$

where A is a constant equal to

$$A = \int_{1<\|\boldsymbol{y}\|\leq 2} \|\boldsymbol{y}\|^{\alpha}\,\mathrm{d}\boldsymbol{y} < +\infty.$$

We obtained

$$\int_{0<\|\boldsymbol{x}\|\leq 1} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x} = A\sum_{k=1}^{\infty} 2^{-(\alpha+n)k} = A\sum_{k=1}^{\infty}\left(\frac{1}{2^{n+\alpha}}\right)^{k}.$$

The geometric series converges if and only if the common ratio is < 1, i.e., if and only if $n+\alpha > 0$, i.e. $\alpha > -n$ as required.

Example *5.47*

Consider the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^{-1} & \text{if } \boldsymbol{x}\neq 0 \\ 0 & \text{if } \boldsymbol{x} = 0 \end{cases} \end{aligned}$$

which, from the previous example, is in $L^1_{\mathrm{loc}}(\mathbb{R}^3;\mathbb{R})$. We want to determine the Laplacian of f in the sense of distributions. We recall that the Laplacian operator is defined as

$$\Delta = \nabla^2 = \sum_{j=1}^{n}\frac{\partial^2}{\partial x_j^2}.$$

To this end, we use Green's formula, expressed in the following form: Let $\boldsymbol{A}\subseteq\mathbb{R}^n$ be an open set with "regular" boundary, i.e., a boundary on which it is possible to define a normal unit vector external to $\boldsymbol{A}$ for all $\boldsymbol{x}\in\partial\boldsymbol{A}$. We denote by $\nu(\boldsymbol{x})$ the external normal unit vector. Let $u\in C^1(\overline{\boldsymbol{A}};\mathbb{C})$: this means that $u\in C^1(\boldsymbol{A};\mathbb{C})$, and it admits a continuation, together with its derivatives, to the closure $\overline{\boldsymbol{A}}$ of $\boldsymbol{A}$. Then

$$\int_{\boldsymbol{A}}\frac{\partial u}{\partial x_j}\,\mathrm{d}\boldsymbol{x} = \int_{\partial\boldsymbol{A}} u\,\nu_j\,\mathrm{d}s,$$

where $j\in\{1,\dots,n\}$, and ν_j is the j-th component of the external normal unit vector. The integral in $\mathrm{d}s$ is a curvilinear integral when $n=2$, and a surface integral if $n=3$.

Let $\varphi\in\mathcal{D}(\mathbb{R}^3;\mathbb{R})$. We have

$$\begin{aligned} \langle\Delta f,\varphi\rangle &= \sum_{j=1}^{3}\left\langle\frac{\partial^2 f}{\partial x_j^2},\varphi\right\rangle \\ &= \sum_{j=1}^{3}\left\langle f,\frac{\partial^2\varphi}{\partial x_j^2}\right\rangle \\ &= \sum_{j=1}^{3}\int_{\mathbb{R}^3}\frac{\partial^2\varphi}{\partial x_j^2}\frac{1}{\|\boldsymbol{x}\|}\,\mathrm{d}\boldsymbol{x} \\ &= \lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial^2\varphi}{\partial x_j^2}\frac{1}{\|\boldsymbol{x}\|}\,\mathrm{d}\boldsymbol{x}. \end{aligned}$$

Now, we observe that

$$\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\right) = \frac{\partial^2\varphi}{\partial x_j^2}(\boldsymbol{x})\frac{1}{\|\boldsymbol{x}\|} + \frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{\partial}{\partial x_j}\left(\frac{1}{\|\boldsymbol{x}\|}\right),$$

so that

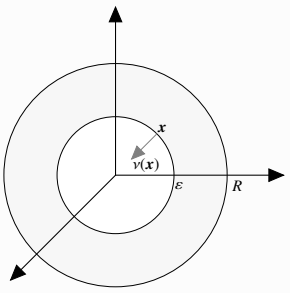
$$\begin{aligned}\langle \Delta f, \varphi\rangle &= \lim_{\varepsilon\to 0^+} \sum_{j=1}^{3} \int_{\|\boldsymbol{x}\|\geq\varepsilon} \frac{\partial^2\varphi}{\partial x_j^2}\frac{1}{\|\boldsymbol{x}\|}\,\mathrm{d}\boldsymbol{x}\\ &= \lim_{\varepsilon\to 0^+}\left(\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x} - \sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{\partial}{\partial x_j}\left(\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x}\right).\end{aligned}$$

Since φ has compact support, there exists $R>0$ such that $\varphi(\boldsymbol{x})=0$ for all $\|\boldsymbol{x}\|\geq R$. Hence,

$$\begin{aligned}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x} &= \int_{\varepsilon\leq\|\boldsymbol{x}\|\leq R}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x}\\ &= \cancel{\int_{\|\boldsymbol{x}\|=R}\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\nu_j(\boldsymbol{x})\,\mathrm{d}s} + \int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\nu_j(\boldsymbol{x})\,\mathrm{d}s\\ &= \int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\nu_j(\boldsymbol{x})\,\mathrm{d}s,\end{aligned}$$

where we applied Green's Theorem in the bounded subset $\{\boldsymbol{x}\in\mathbb{R}^3\,|\,\varepsilon\leq\|\boldsymbol{x}\|\leq R\}$ of $\mathbb{R}^3$ and where we used the fact that $\varphi\in\mathcal{D}(\mathbb{R}^3;\mathbb{R})$ is zero for $\|\boldsymbol{x}\|=R$.

$$\nu(\boldsymbol{x}) = -\frac{1}{\varepsilon}\boldsymbol{x}$$



Using this calculation, together with the fact that $\nu_j(\boldsymbol{x}) = -\frac{x_j}{\varepsilon}$ (the external normal unit vector at $\boldsymbol{x}$, where $\|\boldsymbol{x}\|=\varepsilon$, is $\nu(\boldsymbol{x})=-\boldsymbol{x}/\varepsilon$, and the minus sign is because $\nu(\boldsymbol{x})$ is pointing towards the origin, and therefore in the opposite direction of $\boldsymbol{x}$) and the fact that

$$\frac{\partial}{\partial x_j}(\|\boldsymbol{x}\|^{-1}) = \frac{\partial}{\partial x_j}\left((\|\boldsymbol{x}\|^2)^{-\frac{1}{2}}\right) = -\frac{1}{2}\left((\|\boldsymbol{x}\|^2)^{-\frac{3}{2}}\right)\frac{\partial}{\partial x_j}(x_1^2+x_2^2+\ldots+x_j^2+\ldots+x_n^2) = -\frac{1}{2}\left((\|\boldsymbol{x}\|^2)^{-\frac{3}{2}}\right)2\,x_j = -\frac{x_j}{\|\boldsymbol{x}\|^3},$$

we obtain

$$\begin{aligned}\langle \Delta f, \varphi\rangle &= \lim_{\varepsilon\to 0^+}\left(\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial}{\partial x_j}\left(\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x} - \sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{\partial}{\partial x_j}\left(\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x}\right)\\ &= \lim_{\varepsilon\to 0^+}\left(\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\nu_j(\boldsymbol{x})\,\mathrm{d}s - \sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{\partial}{\partial x_j}\left(\frac{1}{\|\boldsymbol{x}\|}\right)\mathrm{d}\boldsymbol{x}\right)\\ &= \lim_{\varepsilon\to 0^+}\left(-\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\frac{1}{\|\boldsymbol{x}\|}\frac{x_j}{\varepsilon}\,\mathrm{d}s + \sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{x_j}{\|\boldsymbol{x}\|^3}\,\mathrm{d}\boldsymbol{x}\right)\\ &= \lim_{\varepsilon\to 0^+}\left(-\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\frac{1}{\varepsilon}\frac{x_j}{\varepsilon}\,\mathrm{d}s + \sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{x_j}{\|\boldsymbol{x}\|^3}\,\mathrm{d}\boldsymbol{x}\right).\end{aligned}$$

Now, using

$$\frac{\partial}{\partial x_j}\left(\varphi(\boldsymbol{x})\,\frac{x_j}{\|\boldsymbol{x}\|^3}\right)=\frac{\partial\varphi(\boldsymbol{x})}{\partial x_j}\,\frac{x_j}{\|\boldsymbol{x}\|^3}+\varphi(\boldsymbol{x})\,\frac{\partial}{\partial x_j}\left(\frac{x_j}{\|\boldsymbol{x}\|^3}\right),$$

we obtain

$$\begin{aligned}\langle\Delta f,\varphi\rangle &= \lim_{\varepsilon\to 0^+}\left(-\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{1}{\varepsilon}\,\frac{x_j}{\varepsilon}\,\mathrm{d}s+\sum_{j=1}^{3}\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial\varphi}{\partial x_j}(\boldsymbol{x})\frac{x_j}{\|\boldsymbol{x}\|^3}\,\mathrm{d}\boldsymbol{x}\right)\\ &= \lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\left(-\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s+\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial}{\partial x_j}\left(\varphi(\boldsymbol{x})\,\frac{x_j}{\|\boldsymbol{x}\|^3}\right)\mathrm{d}\boldsymbol{x}-\int_{\|\boldsymbol{x}\|\geq\varepsilon}\varphi(\boldsymbol{x})\,\frac{\partial}{\partial x_j}\left(\frac{x_j}{\|\boldsymbol{x}\|^3}\right)\mathrm{d}\boldsymbol{x}\right)\end{aligned}$$

We now show that the last integral gives a zero contribution to the sum, when we add over j. In fact,

$$\begin{aligned}\frac{\partial}{\partial x_j}\left(\frac{x_j}{\|\boldsymbol{x}\|^3}\right) &= \frac{1}{\|\boldsymbol{x}\|^3}+x_j\,\frac{\partial}{\partial x_j}\left(\frac{1}{\|\boldsymbol{x}\|^3}\right)=\frac{1}{\|\boldsymbol{x}\|^3}+x_j\,\frac{\partial}{\partial x_j}\left((\|\boldsymbol{x}\|^2)^{-\frac{3}{2}}\right)\\ &= \frac{1}{\|\boldsymbol{x}\|^3}+x_j\left(-\frac{3}{2}\left(\|\boldsymbol{x}\|^2\right)^{-\frac{5}{2}}\right)2\,x_j\Big)\\ &= \frac{1}{\|\boldsymbol{x}\|^3}-3\,\|\boldsymbol{x}\|^{-5}\,x_j^2=\|\boldsymbol{x}\|^{-6}\left(\|\boldsymbol{x}\|^3-3\,\|\boldsymbol{x}\|\,x_j^2\right).\end{aligned}$$

It follows that

$$\sum_{j=1}^{3}\frac{\partial}{\partial x_j}\left(\frac{x_j}{\|\boldsymbol{x}\|^3}\right)=\sum_{j=1}^{3}\|\boldsymbol{x}\|^{-6}\left(\|\boldsymbol{x}\|^3-3\,\|\boldsymbol{x}\|\,x_j^2\right)=\|\boldsymbol{x}\|^{-6}\left(3\,\|\boldsymbol{x}\|^3-3\,\|\boldsymbol{x}\|^3\right)=0.$$

We obtain

$$\langle\Delta f,\varphi\rangle = \lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\left(-\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s+\int_{\|\boldsymbol{x}\|\geq\varepsilon}\frac{\partial}{\partial x_j}\left(\varphi(\boldsymbol{x})\,\frac{x_j}{\|\boldsymbol{x}\|^3}\right)\mathrm{d}\boldsymbol{x}\right)$$

We can now apply Green's Theorem again to the second integral and we obtain

$$\begin{aligned}\langle\Delta f,\varphi\rangle &= \lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\left(-\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s+\int_{\|\boldsymbol{x}\|=\varepsilon}\varphi(\boldsymbol{x})\,\frac{x_j}{\|\boldsymbol{x}\|^3}\left(-\frac{x_j}{\varepsilon}\right)\mathrm{d}s\right)\\ &= \lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\left(-\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s-\int_{\|\boldsymbol{x}\|=\varepsilon}\varphi(\boldsymbol{x})\,\frac{x_j^2}{\varepsilon^4}\,\mathrm{d}s\right)\\ &= -\lim_{\varepsilon\to 0^+}\sum_{j=1}^{3}\left(\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s+\int_{\|\boldsymbol{x}\|=\varepsilon}\varphi(\boldsymbol{x})\,\frac{x_j^2}{\varepsilon^4}\,\mathrm{d}s\right).\end{aligned}$$

Now we want to show that the first integral is zero. More precisely, for $j=1,2,3$ we have

$$\int_{\|\boldsymbol{x}\|=\varepsilon}\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\,\mathrm{d}s\overset{\varepsilon\to 0^+}{\longrightarrow}0.$$

We bound the integrand by considering that on the surface we have $|x_j|\leq\|\boldsymbol{x}\|=\varepsilon$, so that

$$\left|\frac{\partial\varphi}{\partial x_j}\,\frac{x_j}{\varepsilon^2}\right|\leq\frac{1}{\varepsilon}\,\sup\left|\frac{\partial\varphi}{\partial x_j}\right|.$$

It follows that

$$\int_{\|\boldsymbol{x}\|=\varepsilon} \frac{\partial \varphi}{\partial x_j} \frac{x_j}{\varepsilon^2} \,\mathrm{d}s \leq \frac{1}{\cancel{\varepsilon}} \sup \left| \frac{\partial \varphi}{\partial x_j} \right| 4\pi \cancel{\varepsilon^2} \xrightarrow{\varepsilon \to 0^+} 0.$$

Thus,

$$\begin{aligned}
\langle \Delta f, \varphi \rangle &= -\lim_{\varepsilon \to 0^+} \sum_{j=1}^{3} \int_{\|\boldsymbol{x}\|=\varepsilon} \varphi(\boldsymbol{x}) \frac{x_j^2}{\varepsilon^4} \,\mathrm{d}s \\
&= -\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon^2} \int_{\|\boldsymbol{x}\|=\varepsilon} \varphi(\boldsymbol{x}) \,\mathrm{d}s \\
&= -\lim_{\varepsilon \to 0^+} \left(\frac{1}{\varepsilon^2} \int_{\|\boldsymbol{x}\|=\varepsilon} (\varphi(\boldsymbol{x}) - \varphi(0)) \,\mathrm{d}s + \frac{1}{\varepsilon^2} \int_{\|\boldsymbol{x}\|=\varepsilon} \varphi(0) \,\mathrm{d}s \right) \\
&= -\lim_{\varepsilon \to 0^+} \left(\frac{1}{\varepsilon^2} \int_{\|\boldsymbol{x}\|=\varepsilon} (\varphi(\boldsymbol{x}) - \varphi(0)) \,\mathrm{d}s + \frac{1}{\cancel{\varepsilon^2}} \varphi(0) 4\pi \cancel{\varepsilon^2} \right).
\end{aligned}$$

where we used the fact that $\displaystyle\sum_{j=1}^{3} x_j^2 = \varepsilon^2$. Now we only need to show that the remaining integral goes to 0 as $\varepsilon \to 0^+$. We exploit the continuity of φ. Let $\eta > 0$. Then, there exists $\delta_\eta > 0$ such that if $\|\boldsymbol{x}\| \leq \delta_\eta$, we have $|\varphi(\boldsymbol{x}) - \varphi(0)| \leq \eta$. For $\varepsilon \leq \delta_\eta$ we have

$$\left| \frac{1}{\varepsilon^2} \int_{\|\boldsymbol{x}\|=\varepsilon} (\varphi(\boldsymbol{x}) - \varphi(0)) \,\mathrm{d}s \right| \leq \frac{1}{\varepsilon^2} \eta 4\pi \varepsilon^2,$$

and since η is arbitrarily small the integral goes to 0. Thus,

$$\langle \Delta f, \varphi \rangle = -4\pi \varphi(0),$$

which proves that $\Delta f = -4\pi \delta$.

Exercises

Exercise 5.34

Let $(\varphi_k)_{k \in \mathbb{N}}$ be a sequence in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$ and let $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$, where $A \subseteq \mathbb{R}^n$ is open. Prove that, if $\varphi_k \xrightarrow{k \to \infty} \varphi$ in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$, then for every $j \in \{1, \ldots, n\}$

$$\frac{\partial \varphi_k}{\partial x_j} \xrightarrow{k \to \infty} \frac{\partial \varphi}{\partial x_j} \quad \text{in } \mathcal{D}(\boldsymbol{A}; \mathbb{C}).$$

[Solution: (5.34)]

Exercise 5.35

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k \in \mathbb{N}}$ be a sequence of distributions in $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Then,

$$T_k \xrightarrow{k \to \infty} T \text{ in } \mathcal{D}'(\boldsymbol{A}; \mathbb{C}) \quad \Rightarrow \quad \forall \alpha \in \mathbb{N}^n, \ \partial^\alpha T_k \xrightarrow{k \to \infty} \partial^\alpha T \text{ in } \mathcal{D}'(\boldsymbol{A}; \mathbb{C}).$$

[Solution: (5.35)]

Exercise 5.36

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T_1, T_2 \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let $\alpha \in \mathbb{N}^n$. Show that

$$\partial^\alpha T_1 + \partial^\alpha T_2 = \partial^\alpha (T_1 + T_2).$$

[Solution: (5.36)]

Exercise 5.37 Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(T_k)_{k\in\mathbb{N}}$ be a sequence of distributions in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Show that, if $\sum_{k=0}^{\infty} T_k$ is a convergent series in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, then for all $\alpha \in \mathbb{N}^n$ the series $\sum_{k=0}^{\infty} \partial^\alpha T_k$ converges in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$, and

$$\sum_{k=0}^{\infty} \partial^\alpha T_k = \partial^\alpha \sum_{k=0}^{\infty} T_k.$$

[Solution: (5.37)]

Exercise 5.38 Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Let $j \in \{1,\dots,n\}$. Show that $\partial_j T$ is linear. [Solution: (5.38)]

Exercise 5.39 Show that the distributional derivative of the sign function is $2\,\delta$. [Solution: (5.39)]

Exercise 5.40 Show that the derivative T' of an even distribution $T \in \mathcal{D}'(\mathbb{R};\mathbb{R})$ is odd. [Solution: (5.40)]

Exercise 5.41 Show that the derivative T' of an odd distribution $T \in \mathcal{D}'(\mathbb{R};\mathbb{R})$ is even. [Solution: (5.41)]

Exercise 5.42 Find

$$x^2 \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{pv}\,\frac{1}{x}\right).$$

[Solution: (5.42)]

Exercise 5.43 Let $\alpha \in \mathbb{R}$. Show that the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^\alpha & \text{if } 0 < \|\boldsymbol{x}\| < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is in $L^1(\mathbb{R}^3;\mathbb{R})$ if and only if $\alpha > -3$, using spherical coordinates. [Solution: (5.43)]

Exercise 5.44 Show that the derivative of a t_0-periodic distribution in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ is t_0-periodic. [Solution: (5.44)]

Exercise 5.45 Let $\alpha \in \mathbb{R}$. Show that the function

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\|^\alpha \end{aligned}$$

is summable over $\{\boldsymbol{x} \in \mathbb{R}^n \,|\, 0 < \|\boldsymbol{x}\| \leq 1\}$ if and only if $\alpha > -n$. To this end, verify that for $j \in \mathbb{N}^*$

$$\int_{2^{-j} < \|\boldsymbol{x}\| \leq 2^{1-j}} \|\boldsymbol{x}\|^\alpha \,\mathrm{d}\boldsymbol{x} = 2^{-(n+\alpha)(j-1)} \int_{\frac{1}{2} < \|\boldsymbol{x}\| \leq 1} \|\boldsymbol{x}\|^\alpha \,\mathrm{d}\boldsymbol{x}.$$

[Solution: (5.45)]

Exercise 5.46 Find the first-order partial derivatives of the characteristic function of the first open quadrant of $\mathbb{R}^2$

$$\{(x, y) \in \mathbb{R}^2 \,|\, x > 0 \text{ and } y > 0\}.$$

[Solution: (5.46)]

Exercise 5.47 Let $A = (0, 3)$, and consider the ceiling function

$$\begin{aligned} f : A &\longrightarrow \mathbb{R} \\ t &\mapsto \lceil t \rceil = \min\{n \in \mathbb{Z} \,|\, n \geq t\}. \end{aligned}$$

Compute the distributional derivative f'. [Solution: (5.47)]

Exercise 5.48 Find the first, second and third derivative of the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x\,|x| \end{aligned}$$

in the sense of distributions. [Solution: (5.48)]

Exercise 5.49 Find the derivatives of order ≤ 2

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto e^{-x} H(x), \end{aligned}$$

where H is the Heaviside function. [Solution: (5.49)]

Exercise 5.50 Simplify the distribution

$$e^x \delta''.$$

[Solution: (5.50)]

Exercise 5.51 Find the first and second derivative of the distribution of function type L_f, where

$$f(x) = (H(x) - H(x-2))\,(x^2 - x - 2),$$

and where H represents the Heaviside function. [Solution: (5.51)]

Exercise 5.52 Find the derivative of the distribution of function type $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(x) = H(x)\, \ln|3\,x| + H(-x)\, \ln|2\,x|.$$

[Solution: (5.52)]

5.5.6 Primitives

We consider the case $n = 1$. Let I be an open interval and let $T \in \mathcal{D}'(I; \mathbb{C})$. We ask ourselves if there are distributions whose derivative is T. In other words, we are looking for a distribution $u \in \mathcal{D}'(I; \mathbb{C})$ such that $T = u'$. We will see that this problem always has solutions. Moreover, if there exists $u_0 \in \mathcal{D}'(I; \mathbb{C})$ such that $T = u_0'$, the solutions of $T = u'$ are all those, and only those, that can be obtained by adding to

u_0 constants.

To prove this result, we need a preliminary lemma.

Lemma 5.13 *Let I be an open interval of $\mathbb{R}$. Let $\psi \in \mathcal{D}(I;\mathbb{C})$ be such that*

$$\int_I \psi(x)\,\mathrm{d}x = 1.$$

For all $\varphi \in \mathcal{D}(I;\mathbb{C})$, we define

$$w_\varphi(x) = \int_{I\cap(-\infty,x]} (\varphi(y) - \langle 1,\varphi\rangle\,\psi(y))\,\mathrm{d}y.$$

Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(I;\mathbb{C})$. Then:

(a) $w_\varphi \in \mathcal{D}(I;\mathbb{R})$;

(b) $\forall\, x \in I, \quad w'_\varphi(x) = \varphi(x) - \langle 1,\varphi\rangle\,\psi(x)$;

(c) $\varphi_k \xrightarrow{k\to\infty} \varphi$ *in* $\mathcal{D}(I;\mathbb{C})$ $\quad\Rightarrow\quad$ $w_{\varphi_k} \xrightarrow{k\to\infty} w_\varphi$ *in* $\mathcal{D}(I;\mathbb{C})$.

Proof First, let $I = (a,b)$, where $a, b \in \overline{\mathbb{R}}$ with $a < b$. Then, we can write

$$w_\varphi(x) = \int_a^x (\varphi(y) - \langle 1,\varphi\rangle\,\psi(y))\,\mathrm{d}y.$$

First, we prove that $w_\varphi \in \mathcal{D}(I,\mathbb{R})$. In view of the fundamental theorem of calculus, w_φ is differentiable and

$$w'_\varphi(x) = \varphi(x) - \langle 1,\varphi\rangle\,\psi(x)$$

is of class C^∞, because such is also φ. We need to show that w_φ has compact support. If $\operatorname{supp}\varphi \subseteq [a',b']$ and $\operatorname{supp}\psi \subseteq [a'',b'']$, we will have $a < a' < b' < b$ and $a < a'' < b'' < b$, and we show that

$$\operatorname{supp} w_\varphi \subseteq [\min\{a',a''\}, \max\{b',b''\}].$$

Clearly, for $a < x < \min\{a',a''\}$ we have $w_\varphi(x) = 0$, because both φ and ψ are zero. We show that for $\max\{b',b''\} < x < b$ we also have $w_\varphi(x) = 0$. In fact, recalling that $w'_\varphi(x) = \varphi(x) - \langle 1,\varphi\rangle\,\psi(x)$, and since for $\max\{b',b''\} < x < b$ both $\varphi(x)$ and $\psi(x)$ are zero, we obtain $w'_\varphi(x) = 0$, which implies that $w_\varphi(x)$ is constant. Moreover,

$$w_\varphi(\max\{b',b''\}) = \int_a^{\max\{b',b''\}} \varphi(y)\,\mathrm{d}y - \langle 1,\varphi\rangle \int_a^{\max\{b',b''\}} \psi(y)\,\mathrm{d}y = \langle 1,\varphi\rangle - \langle 1,\varphi\rangle\cdot 1 = 0,$$

where we have used the fact that $\int_I \psi(x)\,\mathrm{d}x = 1$. Now, we need to show that if $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(I,\mathbb{R})$, then $w_{\varphi_k} \xrightarrow{k\to\infty} w_\varphi$ in $\mathcal{D}(I,\mathbb{R})$, which shows that w_φ depends on φ with continuity. We need to check two things: first, that there exists a compact which is independent of k that contains the supports of all functions w_{φ_k}; second, that we have uniform convergence of all the derivatives of w_{φ_k}. Since by assumption $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(I,\mathbb{R})$, there exists a compact set $K \subseteq I$ such that $\operatorname{supp}\varphi_k \subseteq K$ for all $k \in \mathbb{N}$. Let $c = \min K$ and $d = \max K$. Thus, $\operatorname{supp}\varphi_k \subseteq [c,d]$ for all $k \in \mathbb{N}$. From the argument used before, we have $\operatorname{supp} w_{\varphi_k} \subseteq [\min\{c,a''\}, \max\{d,b''\}]$, which is a compact that does not depend on k. Let $\alpha = \min\{c,a''\}$ and $\beta = \max\{d,b''\}$.

For $x \in [\alpha, \beta]$ we find

$$|w_{\varphi_k}(x) - w_\varphi(x)| = \left| \int_a^x (\varphi_k(y) - \langle 1, \varphi_k \rangle \psi(y)) \, dy - \int_a^x (\varphi(y) - \langle 1, \varphi \rangle \psi(y)) \, dy \right|.$$

On the other hand, both $\operatorname{supp} \varphi$ and $\operatorname{supp} \varphi_k$, for all $k \in \mathbb{N}$, are contained in $[\alpha, \beta]$, so that

$$\begin{aligned} |w_{\varphi_k}(x) - w_\varphi(x)| &= \left| \int_\alpha^x (\varphi_k(y) - \varphi(y)) \, dy - \langle 1, \varphi_k - \varphi \rangle \int_\alpha^x \psi(y) \, dy \right| \\ &\leq \int_\alpha^x \left| \varphi_k(y) - \varphi(y) \right| dy - |\langle 1, \varphi_k + \varphi \rangle| \int_\alpha^x |\psi(y)| \, dy \\ &\leq \sup_{[\alpha,\beta]} |\varphi_k - \varphi| (\beta - \alpha) + \sup_{[\alpha,\beta]} |\varphi_k - \varphi| (\beta - \alpha) \int_I |\psi(y)| \, dy, \end{aligned}$$

and both addends go to zero uniformly in view of the uniform convergence to φ of $(\varphi_k)_{k \in \mathbb{N}}$. Finally, we must show that we have uniform convergence of all the derivatives of every order. Consider the first derivative: we have to prove that

$$\lim_{k \to +\infty} w'_{\varphi_k}(x) = w'_\varphi(x)$$

uniformly. Indeed,

$$\begin{aligned} |w'_\varphi(x) - w'_{\varphi_k}(x)| &= |\varphi_k(x) - \varphi(x) - \langle 1, \varphi_k \rangle \psi(x) + \langle 1, \varphi \rangle \psi(x)| \\ &\leq \sup_I |\varphi_k - \varphi| + |\langle 1, \varphi_k - \varphi \rangle| \, |\psi(x)|, \end{aligned}$$

which goes to 0 uniformly because $\varphi_k \xrightarrow{k \to \infty} \varphi$ uniformly and $\langle 1, \varphi_k \rangle \xrightarrow{k \to \infty} \langle 1, \varphi \rangle$. We can iterate the same argument for all the higher order derivatives. Indeed, $w_\varphi^{(\alpha)}(x) = \varphi^{(\alpha)}(x) + \langle 1, \varphi \rangle \psi^{(\alpha)}(x)$, where $\alpha \in \mathbb{N}$, leads to

$$\begin{aligned} |w_\varphi^{(\alpha+1)}(x) - w_{\varphi_k}^{(\alpha+1)}(x)| &= \left| \varphi_k^{(\alpha)}(x) - \varphi^{(\alpha)}(x) - \langle 1, \varphi_k \rangle \psi^{(\alpha)}(x) + \langle 1, \varphi \rangle \psi^{(\alpha)}(x) \right| \\ &\leq \sup_I |\varphi_k^{(\alpha)} - \varphi^{(\alpha)}| + |\langle 1, \varphi_k - \varphi \rangle| \, |\psi^{(\alpha)}(x)|, \end{aligned}$$

which goes to 0 uniformly.

Theorem 5.26 *Let I be an open interval of $\mathbb{R}$. Let $T \in \mathcal{D}'(I; \mathbb{C})$. Then:*

(i) the equation $u' = T$ always admits solutions in $u \in \mathcal{D}'(I, \mathbb{C})$;

(ii) if $u_0 \in \mathcal{D}'(I, \mathbb{C})$ is such that $u_0' = T$, then the set of solutions of $u' = T$ is given by $u = u_0 + c$, with $c \in \mathbb{C}$.

Proof We prove the first point by showing that:

- if $u' = T$ has a solution $u \in \mathcal{D}'(I; \mathbb{C})$, then u satisfies

$$\langle u, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c, \varphi \rangle, \qquad \text{for some } c \in \mathbb{C}, \tag{5.55}$$

  where w_φ is the function introduced in the previous lemma.
- A distribution defined as in (5.55) satisfies the equation $u' = T$ for any $c \in \mathbb{C}$.

Finally, we will prove the second point.

Let $\varphi \in \mathcal{D}(I;\mathbb{C})$. Suppose that there exists $u \in \mathcal{D}'(I;\mathbb{C})$ such that $u' = T$. We have

$$\langle u', w_\varphi \rangle = \langle T, w_\varphi \rangle.$$

Moreover,

$$\begin{aligned} \langle u', w_\varphi \rangle &= -\langle u, w_\varphi' \rangle \\ &= -\langle u, \varphi - \langle 1, \varphi \rangle \psi \rangle \\ &= -\langle u, \varphi \rangle + \langle 1, \varphi \rangle \langle u, \psi \rangle, \end{aligned}$$

so that $\langle T, w_\varphi \rangle = -\langle u, \varphi \rangle + \langle 1, \varphi \rangle \langle u, \psi \rangle$. We obtain

$$\langle u, \varphi \rangle = -\langle T, w_\varphi \rangle + c \langle 1, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c, \varphi \rangle,$$

where $c = \langle u, \psi \rangle \in \mathbb{C}$ (since ψ is a fixed function). If $u : \mathcal{D}(I;\mathbb{C}) \longrightarrow \mathbb{C}$ is a functional that satisfies, for each $\varphi \in \mathcal{D}(I;\mathbb{C})$, the equation $\langle u, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c, \varphi \rangle$, then $u \in \mathcal{D}'(I;\mathbb{C})$. The linearity is obvious, because w_φ is linear (try!). Let us now consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of $\mathcal{D}(I;\mathbb{C})$ such that $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(I;\mathbb{C})$. We have seen in the previous lemma that this implies $w_{\varphi_k} \xrightarrow{k\to\infty} w_\varphi$ in $\mathcal{D}(I;\mathbb{C})$. Then

$$\langle u, \varphi_k \rangle = -\langle T, w_{\varphi_k} \rangle + \langle c, \varphi \rangle \xrightarrow{k\to\infty} -\langle T, w_\varphi \rangle + \langle c, \varphi \rangle = \langle u, \varphi \rangle.$$

This shows that $u \in \mathcal{D}'(I;\mathbb{C})$. This is enough to prove that if $u' = T$ is solvable, we can write the solution u as the distribution satisfying

$$\langle u, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c, \varphi \rangle, \qquad c \in \mathbb{C}.$$

To conclude, we must also show that a distribution u defined in this way is necessarily a solution of $u' = T$. Now,

$$\begin{aligned} \langle u', \varphi \rangle &= -\langle u, \varphi' \rangle \\ &= \langle T, w_{\varphi'} \rangle - \langle c, \varphi' \rangle, \end{aligned}$$

and

$$w_{\varphi'}(x) = \int_a^x (\varphi'(y) - \langle 1, \varphi' \rangle \psi(y))\, \mathrm{d}y,$$

where

$$\langle 1, \varphi' \rangle = \int_I \varphi'(y)\, \mathrm{d}y = \int_c^d \varphi'(y)\, \mathrm{d}y = \varphi(d) - \varphi(c) = 0,$$

since $\operatorname{supp} \varphi \subseteq [c, d]$, so that

$$w_{\varphi'}(x) = \int_a^x \varphi'(y)\, \mathrm{d}y = \int_c^x \varphi'(y)\, \mathrm{d}y = \varphi(x),$$

and therefore $w_{\varphi'}(x) = \varphi(x)$. We obtain

$$\begin{aligned} \langle u', \varphi \rangle &= \langle T, \varphi \rangle - \langle c, \varphi' \rangle \\ &= \langle T, \varphi \rangle + \cancel{\langle c', \varphi \rangle} = \langle T, \varphi \rangle. \end{aligned}$$

Thus, we obtain $u' = T$ as required.

We prove the second point. If we have $u_1, u_2 \in \mathcal{D}'(I;\mathbb{C})$ that satisfy $u_1' = T$ and $u_2' = T$, then u_1 and u_2 differ by a constant. Indeed, subtracting $\langle u_1, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c_1, \varphi \rangle$ from $\langle u_2, \varphi \rangle = -\langle T, w_\varphi \rangle + \langle c_2, \varphi \rangle$, we find $\langle u_1 - u_2, \varphi \rangle = \langle c_1 - c_2, \varphi \rangle$, which shows that $u_1 - u_2 = c_1 - c_2$.

We observe that if $u' = 0$, then u satisfies $\langle u, \varphi\rangle = \langle c, \varphi\rangle$. Then, the solutions of the equation $u' = 0$ are the function-type distributions associated to constants.

5.5.7 Derivative of the tensor product

Let $\boldsymbol{A}, \boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1 \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ and $T_2 \in \mathcal{D}'(\boldsymbol{B}; \mathbb{C})$. Let $\varphi \in \mathcal{D}(\boldsymbol{A} \times \boldsymbol{B}; \mathbb{C})$. We recall that we defined

$$\langle T_1 \otimes T_2, \varphi\rangle \stackrel{\text{def}}{=} \big\langle T_1(\boldsymbol{x}), \langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x}, \boldsymbol{y})\rangle\big\rangle,$$

where $\boldsymbol{x}$ denotes a generic variable in $\boldsymbol{A}$ and $\boldsymbol{y}$ denotes a generic variable in $\boldsymbol{B}$. We have also mentioned that the order is irrelevant: we could have defined the tensor product in an entirely equivalent way as

$$\langle T_1 \otimes T_2, \varphi\rangle = \big\langle T_2(\boldsymbol{y}), \langle T_1(\boldsymbol{x}), \varphi(\boldsymbol{x}, \boldsymbol{y})\rangle\big\rangle.$$

Theorem 5.27 *Let $\boldsymbol{A}, \boldsymbol{B}$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1 \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ and $T_2 \in \mathcal{D}'(\boldsymbol{B}; \mathbb{C})$. Let $\alpha \in \mathbb{N}^n$. Then,*

$$\partial_{\boldsymbol{x}}^{\alpha}(T_1 \otimes T_2) = \partial_{\boldsymbol{x}}^{\alpha} T_1 \otimes T_2.$$

Proof Consider $\varphi \in \mathcal{D}(\boldsymbol{A} \times \boldsymbol{B}; \mathbb{C})$. We have

$$\begin{aligned}\langle \partial_{\boldsymbol{x}}^{\alpha}(T_1 \otimes T_2), \varphi\rangle &= (-1)^{|\alpha|}\langle (T_1 \otimes T_2)(\boldsymbol{x}, \boldsymbol{y}), \partial_{\boldsymbol{x}}^{\alpha}\varphi(\boldsymbol{x}, \boldsymbol{y})\rangle \\ &= (-1)^{|\alpha|}\big\langle T_1(\boldsymbol{x}), \langle T_2(\boldsymbol{y}), \partial_{\boldsymbol{x}}^{\alpha}\varphi(\boldsymbol{x}, \boldsymbol{y})\rangle\big\rangle.\end{aligned}$$

We proved that the function $\boldsymbol{x} \mapsto \langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x}, \boldsymbol{y})\rangle$ is in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$, and therefore we can move the differentiation outside the distribution pairing:

$$\begin{aligned}\langle \partial_{\boldsymbol{x}}^{\alpha}(T_1 \otimes T_2), \varphi\rangle &= (-1)^{|\alpha|}\big\langle T_1(\boldsymbol{x}), \partial_{\boldsymbol{x}}^{\alpha}\langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x}, \boldsymbol{y})\rangle\big\rangle \\ &= \big\langle \partial_{\boldsymbol{x}}^{\alpha} T_1(\boldsymbol{x}), \langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x}, \boldsymbol{y})\rangle\big\rangle = \langle \partial_{\boldsymbol{x}}^{\alpha} T_1 \otimes T_2, \varphi\rangle.\end{aligned}$$

Exercises

Exercise 5.53 Let $\boldsymbol{A}_1, \boldsymbol{A}_2$ be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1 \in \mathcal{D}'(\boldsymbol{A}_1; \mathbb{C})$ and $T_2 \in \mathcal{D}'(\boldsymbol{A}_2; \mathbb{C})$. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $\boldsymbol{\beta} \in \mathbb{N}^m$. Show that

$$\partial_{\boldsymbol{x}}^{\alpha}\partial_{\boldsymbol{y}}^{\beta}(T_1 \otimes T_2) = \partial_{\boldsymbol{x}}^{\alpha} T_1 \otimes \partial_{\boldsymbol{y}}^{\beta} T_2.$$

[Solution: (5.53)]

Exercise 5.54 Show that, if $T, S \in \mathcal{D}'(\mathbb{R}; \mathbb{R})$, there holds

$$\frac{\partial^2(T \otimes S)}{\partial x\, \partial y} = T' \otimes S'.$$

Use this result to show that

$$\frac{\partial^2(H \otimes H)}{\partial x\, \partial y} = \delta,$$

where H is the Heaviside function and δ is the Dirac delta in $\mathcal{D}'(\mathbb{R}^2; \mathbb{R})$. [Solution: (5.54)]

5.5.8 Support of a distribution

Up to this point, we know how a distribution can operate on a function in the space $\mathcal{D}$, which, however, only includes functions that have a very special type of structure: none of the functions that you encountered before are C^∞ with finite support. Thus, if we were to stop now, you would only be able to use distributions on an extremely narrow set of functions. Our aim is to introduce the notion of distribution with compact support, because these can be applied to a much larger class of functions: those which are in C^∞. To start, we begin by introducing the notion of support for a distribution, mimicking what was done for functions.

We know that, by definition, a distribution $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$ is a map $T : \mathcal{D}(\boldsymbol{A};\mathbb{C}) \longrightarrow \mathbb{C}$. Therefore, a distribution cannot be evaluated at a point $\boldsymbol{x} \in \boldsymbol{A}$ as it is usually possible for a function $f : \boldsymbol{A} \longrightarrow \mathbb{C}$. However, we will see that a distribution can be evaluated over open subsets of $\boldsymbol{A}$. This means that we can define the support of a distribution, which plays an important role.

The definition that we provide for the support of a distribution should encompass the one for locally integrable functions. Stated intuitively, given an open set A and a distribution, a point $\boldsymbol{x}$ does not belong to the support of the distribution if there exists an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ such that the restriction of the distribution to $\boldsymbol{B}$ is zero.

Definition 5.26 [Support of a Distribution]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. We define

$$\operatorname{supp} T \stackrel{def}{=} \boldsymbol{A} \setminus \{\boldsymbol{x} \in \boldsymbol{A} \mid \exists\, \boldsymbol{B} \text{ open with } \boldsymbol{x} \in \boldsymbol{B} \subseteq \boldsymbol{A} \ : \ T|_{\boldsymbol{B}} = 0\}$$

The definition given above can be equivalently expressed as follows. We have $\boldsymbol{x}_0 \notin \operatorname{supp} T$ if and only if an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ exists such that $\boldsymbol{x}_0 \in B$ and for all $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ with $\operatorname{supp}\varphi \subset \boldsymbol{B}$ there holds $\langle T, \varphi\rangle = 0$. In other words, we are saying that the two conditions

(A) $T|_{\boldsymbol{B}} = 0$

(B) $\operatorname{supp}\varphi \subseteq \boldsymbol{B} \quad \Rightarrow \quad \langle T, \varphi\rangle = 0$

are equivalent. To see this, we first show that **(A)** implies **(B)**. Indeed, taking $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that $\operatorname{supp}\varphi \subset \boldsymbol{B}$, we have

$$\varphi = \widetilde{\varphi|_{\boldsymbol{B}}},$$

where $\widetilde{\varphi|_{\boldsymbol{B}}}$ denotes the trivial extension of $\varphi|_{\boldsymbol{B}}$; thus, $\langle T, \varphi\rangle = \langle T, \widetilde{\varphi|_{\boldsymbol{B}}}\rangle = \langle T|_B, \varphi|_{\boldsymbol{B}}\rangle = 0$.

Now, we show that **(B)** implies **(A)**. In other words, we have to prove that if $\psi \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$, then $\langle T|_{\boldsymbol{B}}, \psi\rangle = 0$. Denoting by $\tilde{\psi}$ the trivial continuation of ψ to $\boldsymbol{A}$, we find that $\tilde{\psi}$ is in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ and $\operatorname{supp}\tilde{\psi} \subseteq \boldsymbol{B}$, which from **(B)** implies that $\langle T, \tilde{\psi}\rangle = 0$. Since by definition of restriction of a distribution to an open set we have $\langle T|_{\boldsymbol{B}}, \psi\rangle = \langle T, \tilde{\psi}\rangle$, we obtain $\langle T|_{\boldsymbol{B}}, \psi\rangle = 0$ as required.

Hence, we proved the equivalence of the two conditions

(i) $\boldsymbol{x}_0 \notin \operatorname{supp} T \quad \Leftrightarrow \quad \exists\, \boldsymbol{B} \subseteq \boldsymbol{A}$ open such that $\boldsymbol{x}_0 \in B$ and $T|_{\boldsymbol{B}} = 0$;

(ii) $\boldsymbol{x}_0 \notin \operatorname{supp} T \quad \Leftrightarrow$
$\exists\, \boldsymbol{B} \subseteq \boldsymbol{A}$ open such that $\boldsymbol{x}_0 \in \boldsymbol{B}$ and $\forall\, \varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$,
$\operatorname{supp}\varphi \subseteq \boldsymbol{B} \ \Rightarrow \ \langle T, \varphi\rangle = 0.$

Theorem 5.28 [Support of the Dirac δ]
$\operatorname{supp}\delta = \{0\}$.

Proof Let $\boldsymbol{x}_0 \in \mathbb{R}^n \setminus \{0\}$. We show that $\boldsymbol{x}_0 \notin \operatorname{supp}\delta$, i.e., that there exists an open set $\boldsymbol{B} \subseteq \mathbb{R}^n$ such that $\boldsymbol{x}_0 \in \boldsymbol{B}$ for which $\delta|_{\boldsymbol{B}} = 0$. Let us choose $\boldsymbol{B} = \mathbb{R}^n \setminus \{0\}$. Clearly, $\boldsymbol{x}_0 \in \boldsymbol{B}$. Moreover, if we take $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ such that $\operatorname{supp}\varphi \subseteq \boldsymbol{B} = \mathbb{R}^n \setminus \{0\}$, then $\varphi(0) = 0$. Thus, $\langle \delta, \varphi\rangle = 0$, which in turn implies that $\delta|_{\boldsymbol{B}} = 0$. Thus, $\boldsymbol{x}_0 \notin \operatorname{supp}\delta$. Now we show that $\{0\} \in \operatorname{supp}\delta$. Let B be any open subset of $\boldsymbol{A}$ such that $0 \in B$. We show that $\delta|_{\boldsymbol{B}} \neq 0$, i.e., there exists $\varphi \in \mathcal{D}(\boldsymbol{B};\mathbb{R})$ such that $\langle \delta|_{\boldsymbol{B}}, \varphi\rangle \neq 0$. Consider a ball $\mathcal{B}_r(0)$ where $r > 0$, such that $\overline{\mathcal{B}_r(0)} \subset \boldsymbol{B}$. We can define functions in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ whose support is in this ball and that are not zero at the origin. For example, consider the single variable function

$$\psi(x) = \begin{cases} e^{-\frac{1}{x}} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

that we use to construct the function $\varphi(\boldsymbol{x}) = \psi(r^2 - \|\boldsymbol{x}\|^2)$. This function φ is of class C^∞ with compact support inside the ball. There holds $\varphi(0) = \psi(r^2) > 0$ and $\varphi|_{\boldsymbol{B}} \in \mathcal{D}(\boldsymbol{B};\mathbb{R})$, so that $\langle \delta|_{\boldsymbol{B}}, \varphi|_{\boldsymbol{B}}\rangle = \varphi(0) \neq 0$. This shows that the restriction of δ to any open set that contains 0 is never zero.

The next result shows that the definition of support given for a distribution is consistent with the one that we gave for a locally integrable function.

Theorem 5.29 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $f \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$. Then*

$$\operatorname{supp} L_f = \operatorname{supp} f.$$

Proof We show that if $\boldsymbol{x}_0 \notin \operatorname{supp} L_f$, then $\boldsymbol{x}_0 \notin \operatorname{supp} f$. Let $\boldsymbol{x}_0 \notin \operatorname{supp} L_f$, which means that there exists an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ for which $\boldsymbol{x}_0 \in \boldsymbol{B}$ such that $L_f|_{\boldsymbol{B}} = 0$. This means that if we take $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ with $\operatorname{supp}\varphi \subset \boldsymbol{B}$, we have $\langle L_f, \varphi\rangle = 0$. On the other hand, saying that

$$\langle L_f, \varphi\rangle = \int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$$

for all $\varphi \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$ implies, in view of the fundamental lemma 5.5, that $f(\boldsymbol{x}) = 0$ for almost every $\boldsymbol{x} \in \boldsymbol{B}$, which in turn implies that $\boldsymbol{x}_0 \notin \operatorname{supp} f$. Now we prove the converse, i.e., that if $\boldsymbol{x}_0 \notin \operatorname{supp} f$, then $\boldsymbol{x}_0 \notin \operatorname{supp} L_f$. Let $\boldsymbol{x}_0 \notin \operatorname{supp} f$. Thus, there exists an open set $\boldsymbol{B} \subseteq \boldsymbol{A}$ with $\boldsymbol{x}_0 \in \boldsymbol{B}$ such that $f(\boldsymbol{x}) = 0$ almost everywhere on $\boldsymbol{B}$. Thus, if $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ with $\operatorname{supp}\varphi \subset \boldsymbol{B}$, we have

$$\langle L_f, \varphi\rangle = \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{B}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0,$$

which implies $\boldsymbol{x}_0 \notin \operatorname{supp} L_f$.

We now show that the support of the derivative of a distribution is contained in the support of the original distribution.

Theorem 5.30 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$. Then, for every $\alpha \in \mathbb{N}^n$,*

$$\operatorname{supp}(\partial^\alpha T) \subseteq \operatorname{supp} T.$$

Proof Let $\boldsymbol{B}$ be an open set of $\boldsymbol{A}$. We begin by showing that $\partial^\alpha(T|_{\boldsymbol{B}}) = (\partial^\alpha T)|_{\boldsymbol{B}}$. Let $\varphi \in \mathcal{D}(\boldsymbol{B};\mathbb{C})$. We have

$$\begin{aligned}
\langle \partial^\alpha(T|_{\boldsymbol{B}}), \varphi\rangle &= (-1)^{|\alpha|}\langle T|_{\boldsymbol{B}}, \partial^\alpha\varphi\rangle \\
&= (-1)^{|\alpha|}\langle T, \widetilde{\partial^\alpha\varphi}\rangle \\
&= (-1)^{|\alpha|}\langle T, \partial^\alpha\tilde{\varphi}\rangle \qquad \text{why?} \\
&= \langle \partial^\alpha T, \tilde{\varphi}\rangle \\
&= \langle (\partial^\alpha T)|_{\boldsymbol{B}}, \varphi\rangle,
\end{aligned}$$

where the tilde indicates the trivial continuation.

If $\boldsymbol{x}_0 \notin \operatorname{supp} T$, there is an open set B such that $T|_{\boldsymbol{B}} = 0$. Therefore, $\partial^\alpha(T|_{\boldsymbol{B}}) = (\partial^\alpha T)|_{\boldsymbol{B}} = 0$, which shows that $\boldsymbol{x}_0 \notin \operatorname{supp}(\partial^\alpha T)$.

It is easy to see that the support of a distribution $T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C})$ is closed in $\boldsymbol{A}$, i.e., if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a sequence in $\operatorname{supp} T$ and $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}_0 \in \boldsymbol{A}$, then $\boldsymbol{x}_0 \in \operatorname{supp} T$.

The following result is known as *partition of unity*.

Theorem 5.31 [Partition of Unity]
Let $\boldsymbol{K} \subseteq \mathbb{R}^n$ be a compact set. Let $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_\ell$ be open sets of $\mathbb{R}^n$ such that $\boldsymbol{K} \subseteq \boldsymbol{A}_1 \cup \boldsymbol{A}_2 \cup \ldots \cup \boldsymbol{A}_\ell$. Then, for all $j \in \{1, \ldots, \ell\}$, there exist $\psi_j \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ such that:

(i) $\operatorname{supp}\psi_j \subseteq \boldsymbol{A}_j$ *for all* $j \in \{1, \ldots, \ell\}$;

(ii) $\psi_j(\boldsymbol{x}) \geq 0$ *for every* $\boldsymbol{x} \in \mathbb{R}^n$;

(iii) $\psi_1(\boldsymbol{x}) + \psi_2(\boldsymbol{x}) + \ldots + \psi_\ell(\boldsymbol{x}) \leq 1$ *for every* $\boldsymbol{x} \in \mathbb{R}^n$;

(iv) $\psi_1(\boldsymbol{x}) + \psi_2(\boldsymbol{x}) + \ldots + \psi_\ell(\boldsymbol{x}) = 1$ *for every* $\boldsymbol{x} \in \boldsymbol{K}$.

Proof Let $\boldsymbol{K}_i = \overline{\boldsymbol{A}_i \cap \boldsymbol{K}}$ for all $i \in \{1, \ldots, \ell\}$, so that $\boldsymbol{K}_i$ is compact and

$$\boldsymbol{K} \subseteq \bigcup_{i=1}^{\ell} \boldsymbol{K}_i.$$

From Theorem 5.3, for each $i \in \{1, \ldots, \ell\}$ there exists a function $f_i \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that:

- $\operatorname{supp} f_i \subseteq \boldsymbol{A}_i$
- $f_i(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{K}_i$.

Let

$$\begin{cases} \psi_1 = f_1 \\ \psi_i = f_i\,(1 - f_1)\cdot(1 - f_2)\cdots(1 - f_{i-1}), \quad i = 2, \ldots, \ell. \end{cases}$$

First, notice that $\operatorname{supp}\psi_i \subseteq \boldsymbol{A}_i$ for all $i \in \{1, \ldots, \ell\}$. Indeed, this is true for $i = 1$ by construction of f_i. This continues to be true for all $i \in \{1, \ldots, \ell\}$, since ψ_i is the product of f_i by another function. Now, notice that since for all $i \in \{1, \ldots, \ell\}$ we have $0 \leq f_i \leq 1$, we have also $\psi_i \geq 0$. Also, all factors of ψ_i are between 0 and 1, and therefore the product is ≤ 1. We now show that $\psi_1(\boldsymbol{x}) + \psi_2(\boldsymbol{x}) + \ldots + \psi_\ell(\boldsymbol{x}) = 1$ for every $\boldsymbol{x} \in \boldsymbol{K}$. To this end, we show that

$$\sum_{i=1}^{\ell} \psi_i = 1 - (1 - f_1)(1 - f_2)\cdots(1 - f_\ell). \tag{5.56}$$

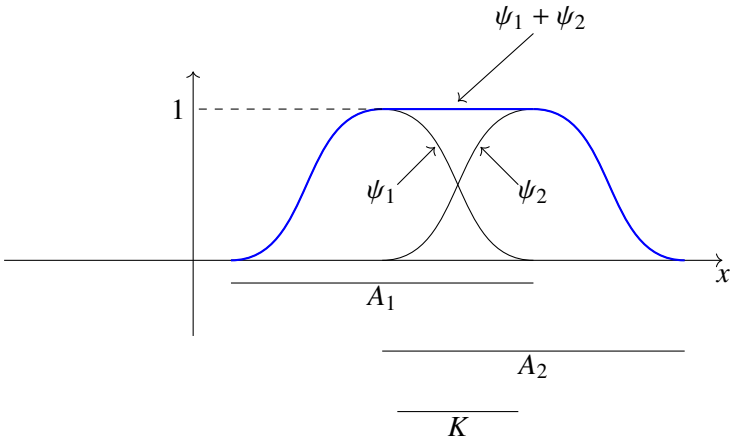
This is sufficient, because for $\boldsymbol{x} \in \boldsymbol{K}$ we have that $\boldsymbol{x} \in \boldsymbol{K}_i$ for at least some $i \in \{1, \ldots, \ell\}$, and therefore at least one f_i will be equal to 1 at $\boldsymbol{x}$, causing $1 - f_i(\boldsymbol{x})$ to be zero. We now prove (5.56). By definition of ψ_i, this means showing that

$$f_1 + f_2\,(1 - f_1) + \cdots + f_\ell\,(1 - f_1)(1 - f_2)\cdots(1 - f_\ell) = 1 - (1 - f_1)(1 - f_2)\cdots(1 - f_\ell). \tag{5.57}$$

For $\ell = 1$, it is trivial to see that the right and the left hand-sides of (5.57) are equal. Let us assume that it is true for $\ell = k$. We prove it for $k+1$. Writing the left hand-side for $\ell = k+1$ and using the inductive assumption on the first k terms give

$$\begin{aligned} &f_1 + f_2\,(1 - f_1) + \cdots + f_k\,(1 - f_1)(1 - f_2)\cdots(1 - f_{k-1}) + f_{k+1}\,(1 - f_1)(1 - f_2)\cdots(1 - f_k) \\ &\quad 1 - (1 - f_1)\cdots(1 - f_k) + f_{k+1}\,(1 - f_1)(1 - f_2)\cdots(1 - f_k) = 1 - (1 - f_{k+1})(1 - f_1)(1 - f_2)\cdots(1 - f_k), \end{aligned}$$

which equals the required right hand-side.



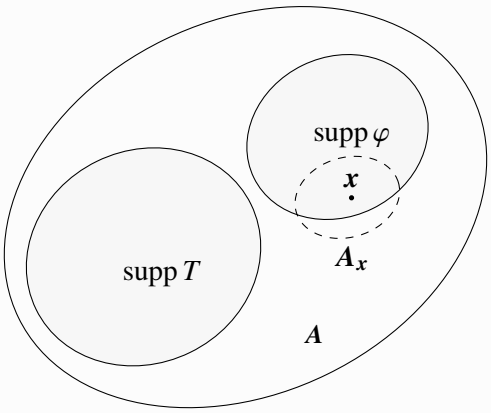
We use the partition of unity to prove the following important result.

Theorem 5.32 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$ and $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. Then,*

$$\operatorname{supp} T \cap \operatorname{supp} \varphi = \varnothing \quad \Rightarrow \quad \langle T, \varphi \rangle = 0.$$

Proof Let $\boldsymbol{x} \in \operatorname{supp} \varphi$. Since $\operatorname{supp} T \cap \operatorname{supp} \varphi = \varnothing$, we have $\boldsymbol{x} \notin \operatorname{supp} T$. This means that there exists an open set $\boldsymbol{A}_{\boldsymbol{x}} \subseteq \boldsymbol{A}$ with $\boldsymbol{x} \in \boldsymbol{A}_{\boldsymbol{x}} \subset \boldsymbol{A}$ such that $T|_{A_x} = 0$ (i.e., such that for all $w \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ with $\operatorname{supp} w \subset \boldsymbol{A}_{\boldsymbol{x}}$, there holds $\langle T, w \rangle = 0$). Since this can be said for all points of $\operatorname{supp} \varphi$, we find

$$\operatorname{supp} \varphi \subseteq \bigcup_{\boldsymbol{x} \in \operatorname{supp} \varphi} \boldsymbol{A}_{\boldsymbol{x}}.$$



Since $\operatorname{supp}\varphi$ is compact, there exists a finite subfamily of $\boldsymbol{A}_{\boldsymbol{x}}$ that continues to cover $\operatorname{supp}\varphi$, i.e., there exist $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^p \in \operatorname{supp}\varphi$ such that $\operatorname{supp}\varphi \subseteq \boldsymbol{A}_{\boldsymbol{x}^1} \cup \boldsymbol{A}_{\boldsymbol{x}^2} \cup \ldots \boldsymbol{A}_{\boldsymbol{x}^p}$. In view of Theorem 5.31, for every $j = 1, \ldots, p$, there exist $\psi_j \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$ with $\operatorname{supp}\psi_j \subseteq \boldsymbol{A}_{\boldsymbol{x}^j}$ such that $\psi_1(\boldsymbol{x}) + \ldots + \psi_p(\boldsymbol{x}) \leq 1$ for every $\boldsymbol{x} \in \mathbb{R}^n$ and $\psi_1(\boldsymbol{x}) + \ldots + \psi_p(\boldsymbol{x}) = 1$ for every $\boldsymbol{x} \in \operatorname{supp}\varphi$. Since φ is zero outside its support, we find

$$\sum_{j=1}^{p} \psi_j(\boldsymbol{x})\,\varphi(\boldsymbol{x}) = \varphi(\boldsymbol{x}),$$

i.e., $\varphi = \displaystyle\sum_{j=1}^{p} \psi_j\,\varphi$. Then, by linearity we find

$$\begin{aligned}\langle T, \varphi\rangle &= \left\langle T, \sum_{j=1}^{p} \psi_j\,\varphi \right\rangle \\ &= \sum_{j=1}^{p} \langle T, \psi_j\,\varphi\rangle.\end{aligned}$$

We now prove that for every $j \in \{1, \ldots, p\}$, there holds $\langle T, \psi_j\,\varphi\rangle = 0$. This is a simple consequence of the fact that $\operatorname{supp}(\psi_j\,\varphi) \subseteq \boldsymbol{A}_{\boldsymbol{x}^j}$, and the restriction of T to each $\boldsymbol{A}_{\boldsymbol{x}^j}$ is the zero distribution, i.e., $T|_{A_{x^j}} = 0$.

We now turn our attention to distributions with compact support. Clearly the Dirac δ distribution has compact support. The following result represents the first step in establishing that if distribution T has compact support, it is possible to choose the test function φ to belong to a wider class of functions, i.e., $C^\infty(\boldsymbol{A}; \mathbb{C})$ instead of $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$.

Theorem 5.33 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. Let U be an open subset of $\boldsymbol{A}$ that contains* $\operatorname{supp} T$. *Let $\psi \in C^\infty(\boldsymbol{A}; \mathbb{C})$ be such that $\psi(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in U$. Then, $\psi\,T = T$.*

Proof We need to prove that $\psi\,T - T = (\psi - 1)\,T = 0$. Let $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$; we show that $\langle (\psi - 1)\,T, \varphi\rangle = 0$. Clearly, $\langle (\psi - 1)\,T, \varphi\rangle = \langle T, (\psi - 1)\,\varphi\rangle$. Therefore, it suffices to prove that $\operatorname{supp} T \cap \operatorname{supp}((\psi - 1)\,\varphi) = \varnothing$, because this allows to use Theorem 5.32 to conclude that $\langle T, (\psi - 1)\,\varphi\rangle = 0$. Let $\boldsymbol{x}_0 \in \operatorname{supp} T$. Hence, $\boldsymbol{x}_0 \in U$, which implies that $\psi(\boldsymbol{x}_0) = 1$. Thus, $(\psi(\boldsymbol{x}_0) - 1)\,\varphi(\boldsymbol{x}_0) = 0$, which in turn implies that $\boldsymbol{x}_0 \notin \operatorname{supp}((\psi - 1)\,\varphi)$.

Remark 5.9 Notice that we cannot simplify the statement of Theorem 5.33 by claiming that given $\psi \in C^\infty(\boldsymbol{A};\mathbb{C})$ such that $\psi(\boldsymbol{x}) = 1$ in $\operatorname{supp} T$, then $\psi\, T = T$. Consider for example $A = \mathbb{R}$, $T = \delta'$, and $\psi(x) = 1 + x$. In this case we have $\operatorname{supp}\delta' \subseteq \operatorname{supp}\delta = \{0\}$, so that indeed ψ is certainly equal to 1 in $\operatorname{supp}\delta'$. Now,

$$\begin{aligned}\langle (1+x)\,\delta'(x), \varphi(x)\rangle &= \langle \delta'(x), (1+x)\,\varphi(x)\rangle \\ &= -\langle \delta(x), \varphi(x) + (1+x)\,\varphi'(x)\rangle = -\varphi(0) - \varphi'(0).\end{aligned}$$

This implies that $(1+x)\,\delta'(x) = -\delta(x) + \delta'(x)$. On the other hand, if it was true that $(1+x)\,\delta' = \delta'$, we would obtain $\delta' = -\delta + \delta'$, which implies $\delta = 0$. This is clearly false.

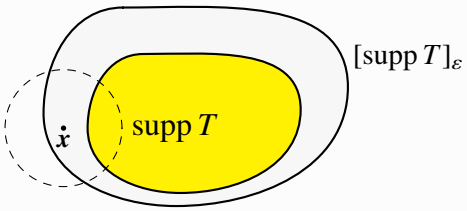
Definition 5.27 [Distributions with Compact Support]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$. We define

$$\mathcal{E}'(\boldsymbol{A};\mathbb{C}) \stackrel{def}{=} \{T \in \mathcal{D}'(\boldsymbol{A};\mathbb{C}) \mid \operatorname{supp} T \text{ is compact on } \boldsymbol{A}\}.$$

The following result shows that, for $T \in \mathcal{E}'(\boldsymbol{A};\mathbb{C})$, there exists $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ that is equal to 1 on an open set U that contains $\operatorname{supp} T$.

Theorem 5.34 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{E}'(\boldsymbol{A};\mathbb{C})$. There exists an open set U of $\boldsymbol{A}$ such that $\operatorname{supp} T \subset U$ and there exists $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that $\zeta(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in U$.*

Proof We have $[\operatorname{supp} T]_\varepsilon \stackrel{\text{def}}{=} \left\{\boldsymbol{x} \in \mathbb{R}^n \mid \overline{\mathcal{B}_\varepsilon(\boldsymbol{x})} \cap \operatorname{supp} T \neq \emptyset\right\}$.



We have shown in Theorem 5.2 that there exists $\varepsilon > 0$ such that $[\operatorname{supp} T]_\varepsilon \subseteq \boldsymbol{A}$ and that $[\operatorname{supp} T]_\varepsilon$ is compact. Let $\varepsilon > 0$ be such that $[\operatorname{supp} T]_\varepsilon \subseteq \boldsymbol{A}$. Then, $[\operatorname{supp} T]_{\varepsilon/2} \subseteq \boldsymbol{A}$ is compact, and $\operatorname{supp} T \subseteq [\operatorname{supp} T]_{\varepsilon/2} \subseteq \boldsymbol{A}$. There exists $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ that is equal to 1 for every $\boldsymbol{x} \in [\operatorname{supp} T]_{\varepsilon/2}$. Let $\boldsymbol{U} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathcal{B}_{\varepsilon/2}(\boldsymbol{x}) \cap \operatorname{supp} T \neq \emptyset\}$. Thus, $\operatorname{supp} T \subseteq \boldsymbol{U} \subseteq [\operatorname{supp} T]_{\varepsilon/2}$. Therefore $\zeta(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{U}$. We need to prove that U is open, i.e., that all points that are close enough to a point $\boldsymbol{x} \in \boldsymbol{U}$ are still in $\boldsymbol{U}$. There exists $\boldsymbol{y} \in \operatorname{supp} T$ such that $\|\boldsymbol{x} - \boldsymbol{y}\| = \eta < \varepsilon/2$. Let $\boldsymbol{z} \in \mathcal{B}_r(\boldsymbol{x})$. We have $\|\boldsymbol{z} - \boldsymbol{y}\| \leq \|\boldsymbol{z} - \boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{y}\| < r + \eta$. If $r < (\varepsilon/2) - \eta$, then $\|\boldsymbol{z} - \boldsymbol{y}\| < \varepsilon/2$. It follows that $\boldsymbol{z} \in \boldsymbol{U}$, because $\boldsymbol{y} \in \mathcal{B}_{\varepsilon/2}(\boldsymbol{z}) \cap \operatorname{supp} T$. From the arbitrariness of $\boldsymbol{z}$ in $\mathcal{B}_r(\boldsymbol{x})$, we obtain $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{U}$ as required.

Combining Theorems 5.33 and 5.34, we find that given $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that $\zeta(\boldsymbol{x}) = 1$ for all $\boldsymbol{x}$ belonging to an open subset $\boldsymbol{U}$ of $\boldsymbol{A}$ containing $\operatorname{supp} T$, we have $\zeta\, T = T$, so that for all $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ we

have

$$\begin{aligned}\langle \zeta\, T, \varphi\rangle &= \langle T, \varphi\rangle \ \text{ because } \zeta\, T = T \\ &= \langle T, \zeta\, \varphi\rangle \ \text{ because } \zeta \in C^\infty(A;\mathbb{C})\end{aligned}$$

As such,

$$\langle T, \varphi\rangle = \langle T, \zeta\, \varphi\rangle. \tag{5.58}$$

On the other hand, the right hand-side of the latter is well defined even when $\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$ instead of $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. This consideration justifies the following definition.

Definition 5.28 [Extension of Distributions to Functions in C^∞]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$. Let $T \in \mathcal{E}'(\boldsymbol{A};\mathbb{C})$. Let $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ be such that $\zeta(\boldsymbol{x}) = 1$ for $\boldsymbol{x} \in \boldsymbol{U}$, where $\boldsymbol{U}$ is an open set of $\boldsymbol{A}$ such that $\operatorname{supp} T \subseteq \boldsymbol{U}$. *Let $\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$. We define*

$$\langle T, \varphi\rangle \overset{def}{=} \langle T, \zeta\, \varphi\rangle.$$

The following result shows that this definition does not depend on the function ζ.

Theorem 5.35 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$. Let $T \in \mathcal{E}'(\boldsymbol{A};\mathbb{C})$. For $j \in \{1,2\}$, let $\zeta_j \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ be such that $\zeta_j(\boldsymbol{x}) = 1$ for $\boldsymbol{x} \in \boldsymbol{U}_j$, where $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ are open sets of $\boldsymbol{A}$ such that* $\operatorname{supp} T \subseteq \boldsymbol{U}_j$. *Then,*

$$\langle T, \zeta_1\, \varphi\rangle = \langle T, \zeta_2\, \varphi\rangle \quad \forall\, \varphi \in C^\infty(\boldsymbol{A};\mathbb{C}).$$

Proof We show that $\langle T, (\zeta_1 - \zeta_2)\,\varphi\rangle = 0$ for all $\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$. We have $(\zeta_1 - \zeta_2)\,\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, and $(\zeta_1 - \zeta_2)\,\varphi$ is identically zero in the open set $\boldsymbol{U}_1 \cap \boldsymbol{U}_2$, which also contains $\operatorname{supp} T$. Then, if $\boldsymbol{x} \in \operatorname{supp} T$, we have $\boldsymbol{x} \in \boldsymbol{U}_1 \cap \boldsymbol{U}_2$ and therefore $\boldsymbol{x} \notin \operatorname{supp}((\zeta_1 - \zeta_2)\,\varphi)$. It follows that $\operatorname{supp} T \cap \operatorname{supp}((\zeta_1 - \zeta_2)\,\varphi) = \varnothing$, from which we immediately find $\langle T, (\zeta_1 - \zeta_2)\,\varphi\rangle = 0$.

We now introduce the following convergence in $C^\infty(\boldsymbol{A};\mathbb{C})$, which requires a notion of uniform convergence on any compact set contained in $\boldsymbol{A}$.

Definition 5.29 [Convergence in C^∞]
Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $C^\infty(\boldsymbol{A};\mathbb{C})$, and let $\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$. Then, we say that

$$\lim_{k\to\infty} \varphi_k = \varphi$$

in $C^\infty(\boldsymbol{A};\mathbb{C})$ if for every compact subset $\boldsymbol{K}$ of $\boldsymbol{A}$ and for every $\alpha \in \mathbb{N}^n$ we have $\lim_{k\to\infty} \partial^\alpha \varphi_k(\boldsymbol{x}) = \partial^\alpha \varphi(\boldsymbol{x})$ uniformly in $\boldsymbol{K}$.

Theorem 5.36 *Let $\boldsymbol{A} \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{E}'(\boldsymbol{A};\mathbb{C})$. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $C^\infty(\boldsymbol{A};\mathbb{C})$ that converges to $\varphi \in C^\infty(\boldsymbol{A};\mathbb{C})$. Then,*

$$\lim_{k\to\infty} \langle T, \varphi_k\rangle = \langle T, \varphi\rangle.$$

Proof Let $\zeta \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$ such that $\zeta(\boldsymbol{x}) = 1$ for every $\boldsymbol{x} \in \boldsymbol{U}$, where $\boldsymbol{U}$ is an open subset of $\boldsymbol{A}$ with $\operatorname{supp} T \subseteq \boldsymbol{U}$. Since $\langle T, \varphi_k\rangle = \langle T, \zeta\,\varphi_k\rangle$, it suffices to verify that $\lim_{k\to\infty}\langle T, \zeta\,\varphi_k\rangle = \langle T, \zeta\,\varphi\rangle$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, i.e., $\lim_{k\to\infty}(\zeta\,\varphi_k) = \zeta\,\varphi$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Therefore, we need to prove that **(i)** there exists a compact set that contains the supports of every $\zeta\,\varphi_k$. This is easy to show, because $\operatorname{supp}(\zeta\,\varphi_k) \subseteq \operatorname{supp}\zeta$ is a compact that is independent from k; **(ii)** The derivatives of $\zeta\,\varphi_k$ uniformly converge to the derivatives of $\zeta\,\varphi$. Let $\boldsymbol{x} \in \operatorname{supp}\zeta$ and $\boldsymbol{\alpha} \in \mathbb{N}^n$. We have

$$\begin{aligned}
|\partial^{\alpha}(\zeta\,\varphi_k)(\boldsymbol{x}) - \partial^{\alpha}(\zeta\,\varphi)(\boldsymbol{x})| &= \Big|\sum_{\beta\le\alpha} c_{\alpha,\beta}\partial^{\alpha-\beta}\zeta(\boldsymbol{x})(\partial^{\beta}\varphi_k(\boldsymbol{x}) - \partial^{\beta}\varphi(\boldsymbol{x}))\Big| \\
&\le \sum_{\beta\le\alpha} c_{\alpha,\beta}|\partial^{\alpha-\beta}\zeta(\boldsymbol{x})|\,|\partial^{\beta}\varphi_k(\boldsymbol{x}) - \partial^{\beta}\varphi(\boldsymbol{x})| \\
&\le \sum_{\beta\le\alpha} c_{\alpha,\beta}\sup_{\boldsymbol{A}}|\partial^{\alpha-\beta}\zeta|\sup_{\boldsymbol{A}}|\partial^{\beta}\varphi_k(\boldsymbol{x}) - \partial^{\beta}\varphi(\boldsymbol{x})| \\
&\le \sum_{\beta\le\alpha} c_{\alpha,\beta}\sup_{\boldsymbol{A}}|\partial^{\alpha-\beta}\zeta|\sup_{\operatorname{supp}\zeta}|\partial^{\beta}\varphi_k - \partial^{\beta}\varphi| \xrightarrow{k\to\infty} 0
\end{aligned}$$

uniformly because $\partial^{\beta}\varphi_k \xrightarrow{k\to\infty} \partial^{\beta}\varphi$ over compact sets.

Example *5.48*

Consider the Dirac delta distribution δ, whose support is compact in $\mathbb{R}^n$ (recall that $\operatorname{supp}\delta = \{0\}$). Let $\zeta \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ be such that $\zeta(\boldsymbol{x}) = 1$ for $\boldsymbol{x} \in U$, where U is an open set of $\mathbb{R}^n$ such that $0 \in U$. Consider $\varphi \in C^{\infty}$. Then,

$$\langle\delta(\boldsymbol{x}), \varphi(\boldsymbol{x})\rangle = \langle\delta, \zeta\,\varphi\rangle = \zeta(0)\,\varphi(0) = \varphi(0).$$

Therefore the extension of the Dirac delta distribution maps every function in C^{∞} to its value at the origin.

Remark 5.10 For a distribution with compact support, we can also define the integral

$$\int_{\boldsymbol{A}} T(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \overset{\text{def}}{=} \langle T, 1\rangle.$$

For example, $\displaystyle\int_{\mathbb{R}} \delta(t)\,\mathrm{d}t = \langle\delta, 1\rangle = 1.$

Example *5.49*

We show that, given $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$, we have $x\,T = 0$ if and only if there exists $c \in \mathbb{C}$ such that $T = c\,\delta$. Clearly, by $x\,T$ we mean the product of the distribution T with the function

$$\begin{aligned}
\mathbb{R} &\longrightarrow \mathbb{R} \\
x &\mapsto x.
\end{aligned}$$

If $T = c\,\delta$, the result is obvious: $\langle x\,(c\,\delta), \varphi\rangle = \langle\delta, c\,x\,\varphi\rangle = 0 \cdot c\,\varphi(0) = 0$ for every $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$. Now, suppose that $x\,T = 0$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$. For any $x \in \mathbb{R}$ we have

$$\varphi(x) = \varphi(x) - \varphi(0) + \varphi(0) = \int_0^x \varphi'(t)\,\mathrm{d}t + \varphi(0) = \int_0^1 \varphi'(s\,x)\,x\,\mathrm{d}s + \varphi(0) = x\int_0^1 \varphi'(s\,x)\,\mathrm{d}s + \varphi(0).$$

We want to show that $\langle T, \varphi\rangle = c\,\varphi(0)$. Taking $\psi \in \mathcal{D}(\mathbb{R};\mathbb{C})$ such that $\psi(x) = 1$ in a neighborhood containing

$\operatorname{supp}\varphi \cup \{0\}$, we find

$$\begin{aligned}\langle T, \varphi\rangle &= \left\langle T, \psi\, x \int_0^1 \varphi'(s\,x)\,\mathrm{d}s + \psi\,\varphi(0)\right\rangle \\ &= \varphi(0)\,\langle T, \psi\rangle + \left\langle x\,T, \psi \int_0^1 \varphi'(s\,x)\,\mathrm{d}s\right\rangle = \varphi(0)\,\langle T, \psi\rangle,\end{aligned}$$

where we used the fact that $x\,T = 0$. We only need to prove that $\langle T, \psi\rangle$ is a constant, which only depends on T. We show that if ψ_1 and ψ_2 are equal in a neighborhood of $\operatorname{supp}\varphi \cup \{0\}$, then $\langle T, \psi_1\rangle = \langle T, \psi_2\rangle$. Using linearity, this means $\langle T, \psi_1 - \psi_2\rangle = 0$. We only need to prove that if g is zero in a neighborhood of 0, then $\langle T, g\rangle = 0$. Let

$$v(x) = \begin{cases} g(x)/x & \text{if } x \in \operatorname{supp} g \\ 0 & \text{if } x \notin \operatorname{supp} g \end{cases}$$

This function is in $\mathcal{D}(\mathbb{R};\mathbb{C})$ and $x\,v = g$, so that

$$\langle T, g\rangle = \langle T, x\,v\rangle = \langle x\,T, v\rangle = 0,$$

as required.

Example *5.50*

We show that, given $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$, we have $x\,T = 1$ if and only if there exists $c \in \mathbb{C}$ such that $T = c\,\delta + \operatorname{pv}\frac{1}{x}$. Recall that in a previous exercise we proved that $x\,\operatorname{pv}\frac{1}{x} = 1$, so that

$$x\,T = 1 = x\,\operatorname{pv}\frac{1}{x} \quad \Rightarrow \quad x\left(T - \operatorname{pv}\frac{1}{x}\right) = 0.$$

From the previous exercise, this is true if and only if there exists $c \in \mathbb{C}$ such that $T - \operatorname{pv}\frac{1}{x} = c\,\delta$.

Exercises

Exercise 5.55

Using the change of variable $y = k\,x$ and Lebesgue's dominated convergence theorem (Theorem 3.47), show that

$$\lim_{k\to\infty} \frac{k}{k^2\,x^2 + 1} = \pi\,\delta$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. [Solution: (5.55)]

Exercise 5.56

Find the limit of the sequence $(T_k)_{k\in\mathbb{N}^*}$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$, where

$$T_k = k^2\,x\,\delta(2\,k\,x - 1).$$

[Solution: (5.56)]

Exercise 5.57

Find

$$\lim_{k\to\infty} k\left(\delta\Big(x - \frac{1}{k}\Big) - \delta\Big(x + \frac{5}{k}\Big)\right)$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. [Solution: (5.57)]

Exercise 5.58

Let $\varepsilon > 0$. Let

$$\begin{aligned} f_\varepsilon : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x,y,z) &\mapsto \begin{cases} \frac{1}{\varepsilon} & \text{if } 1 \le y^2+z^2 \le (1+\varepsilon)^2 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Find

$$\lim_{\varepsilon \to 0} f_\varepsilon$$

in $\mathcal{D}'(\mathbb{R}^3;\mathbb{R})$. [Solution: (5.58)]

Exercise 5.59

Find the limit

$$\lim_{n\to\infty} \frac{n}{1+4\,n^2\,x^2}$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. [Solution: (5.59)]

Exercise 5.60

Consider the distribution defined by

$$\begin{aligned} T : \mathcal{D}(\mathbb{R}^2;\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \int_{\mathbb{R}} \varphi(x,-3\,x)\,\mathrm{d}x. \end{aligned}$$

Show that T is not of function type (i.e., it is a singular distribution), and that

$$\partial_x T - 3\,\partial_y T = 0.$$

[Solution: (5.60)]

Exercise 5.61

Evaluate the limit

$$\lim_{n\to\infty} n^\alpha \left(\delta\left(x+\frac{3}{n}\right) - \delta(x)\right)$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ as a function of $\alpha \in \mathbb{R}$. [Solution: (5.61)]

Exercise 5.62

Find

$$\lim_{\varepsilon \to 0^+} \varepsilon\,|x|^{3\,\varepsilon-1}.$$

[Solution: (5.62)]

Exercise 5.63

Let u denote the distribution p. v. $\frac{1}{x}$. Show that

$$(x^2+3\,x)\,u'(x) + 3\,u(x) = -1.$$

[Solution: (5.63)]

Exercise 5.64

Find

$$\lim_{k\to\infty} k^\alpha\, \chi_{[0,\frac{1}{3k}]}$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ as a function of α. [Solution: (5.64)]

Exercise 5.65

Find the limit

$$\lim_{\varepsilon \to 0^+} \varepsilon^{-5}\, H(x)\, H(\varepsilon - x)\, x^{\alpha}$$

in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ with $\alpha > -1$. [Solution: (5.65)]

Exercise 5.66

Let $\varepsilon > 0$, and let $u_\varepsilon : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be defined by

$$u_\varepsilon(x,y) = \begin{cases} \frac{3}{\varepsilon} & \text{if } 0 \le y \le 4\,\varepsilon \\ 0 & \text{otherwise} \end{cases}$$

Show that

$$\lim_{\varepsilon \to 0} \left(\partial_x u_\varepsilon - \partial_y u_\varepsilon \right) = -12\, \partial_y (1 \otimes \delta).$$

[Solution: (5.66)]

Exercise 5.67

Let u be the characteristic function of $\{(x,y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \le 9\}$. Verify using Green's theorem that, in the sense of distributions,

$$x\, \frac{\partial u}{\partial x} + y\, \frac{\partial u}{\partial y} = c\, T,$$

where, for $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$, $T(\varphi)$ is a curvilinear integral of φ and c is a complex number to determine. [Solution: (5.67)]

5.6 Convolution

Convolution is an operation on two functions that yields a new function which, loosely speaking, expresses how the shape of one is modified by the other. The convolution refers is defined as the integral of the product of the two given functions after one is reversed and shifted. The integral is evaluated for all values of the shift.

The convolution plays a fundamental role in statistics, physics and signal processing.

Given two functions $f, g : \mathbb{R}^n \longrightarrow \mathbb{R}$, we define their convolution as

$$(f * g)(\boldsymbol{x}) \stackrel{\text{def}}{=} \int_{\mathbb{R}^n} f(\boldsymbol{x} - \boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{R}^n. \tag{5.59}$$

The integral in (5.59) is usually called *convolution integral*. We first want to examine conditions on the functions f and g that ensure that this definition is well posed for every $\boldsymbol{x} \in \mathbb{R}^n$, or at least for almost every $\boldsymbol{x} \in \mathbb{R}^n$. If f and g are continuous functions, the convergence of this integral only depends on the behaviour at infinity of f and g. In other words, given $\boldsymbol{x} \in \mathbb{R}^n$, what needs to be integrable is the product of the two maps

$$\boldsymbol{y} \mapsto f(\boldsymbol{x} - \boldsymbol{y}) \quad \text{and} \quad \boldsymbol{y} \mapsto g(\boldsymbol{y})$$

The first map is obtained with an operation of symmetry and a shift. Indeed, $f(\boldsymbol{x} - \boldsymbol{y}) = h(\boldsymbol{y} - \boldsymbol{x})$ and $h(\boldsymbol{x}) = f(-\boldsymbol{x})$. Thus, in loose terms, if $n = 1$ the behaviour of f at $-\infty$ has to be compensated by the behaviour of g at $+\infty$ and *vice-versa*. This means that we will not be able to impose assumptions on each of the two functions f and g independently, but rather pairs of compatibility conditions, such that the convolution integral makes sense.

Definition 5.30 [Convolvable Functions]
Let $f, g \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{R})$. We say that f and g are convolvable *if for every compact set $\boldsymbol{K}$ of $\mathbb{R}^n$ the integral*

$$\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} \tag{5.60}$$

is finite.

First, we want to convince ourselves that the integral in (5.60) is well defined, i.e., that the function

$$\begin{aligned}\mathbb{R}^{2n} &\longrightarrow \mathbb{R}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\end{aligned}$$

is measurable in $\mathbb{R}^{2n}$. To this end, we recall that we have defined the tensor product between two functions $f, g \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{R})$ as the function

$$\begin{aligned}f\otimes g : \mathbb{R}^{2n} &\longrightarrow \mathbb{R}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto (f\otimes g)(\boldsymbol{x},\boldsymbol{y}) = f(\boldsymbol{x})\,g(\boldsymbol{y}),\end{aligned}$$

and we proved that this is an element of $L^1_{\text{loc}}(\mathbb{R}^{2n};\mathbb{R})$: the tensor product of locally integrable functions is locally integrable. On the other hand, the function

$$\begin{aligned}\mathbb{R}^{2n} &\longrightarrow \mathbb{R}\\ (\boldsymbol{x},\boldsymbol{y}) &\mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\end{aligned}$$

can be obtained as the composition of a tensor product and an invertible linear transformation $A : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^{2n}$ such that $A(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})$. In other words, for every $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$,

$$f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y}) = (f\otimes g)(A(\boldsymbol{x},\boldsymbol{y})) = ((f\otimes g)\circ A)(\boldsymbol{x},\boldsymbol{y}).$$

From the change of variable theorem of the Lebesgue integration theory, the function $(\boldsymbol{x},\boldsymbol{y}) \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ is measurable in $\mathbb{R}^{2n}$.

We also notice that in the definition of convolvable functions we require that the function be integrable over every set $\boldsymbol{K}\times\mathbb{R}^n$, where $\boldsymbol{K}\subseteq\mathbb{R}^n$ is compact, and this is a more restrictive condition than requiring that the function $(\boldsymbol{x},\boldsymbol{y}) \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ be in $L^1_{\text{loc}}(\mathbb{R}^{2n};\mathbb{R})$ (which would require assuming the integrability over any compact of $\mathbb{R}^{2n}$). The reason for this is the fact that, as we now show, this assumption ensures that the convolution integral $\displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$ exists for almost every $\boldsymbol{x}\in\mathbb{R}^n$.

If f and g are convolvable, then the function $(\boldsymbol{x},\boldsymbol{y}) \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ is integrable over $\boldsymbol{K}\times\mathbb{R}^n$ for every compact set $\boldsymbol{K}\subseteq\mathbb{R}^n$. Using Fubini's theorem (Theorem 3.8), we can say that for almost every $\boldsymbol{x}\in\boldsymbol{K}$ the function $\boldsymbol{y}\mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ is integrable over $\mathbb{R}^n$, and therefore the integral

$$\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$$

exists for almost every $\boldsymbol{x}\in\boldsymbol{K}$. On the other hand, $\mathbb{R}^n$ can be expressed as the union of a countable family of compact sets, i.e., as

$$\mathbb{R}^n = \bigcup_{k=0}^{\infty} \overline{\mathcal{B}_k(0)}.$$

We can construct, for each $k \in \mathbb{N}$, the sets

$$\boldsymbol{N}_k = \left\{ \boldsymbol{x} \in \overline{\mathcal{B}_k(0)} \,\middle|\, \int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} \ \text{ does not exist} \right\}.$$

For every $k \in \mathbb{N}$, the set $\boldsymbol{N}_k$ has zero Lebesgue measure, and therefore also

$$\mu_n(\boldsymbol{N}) = 0, \quad \text{where } \boldsymbol{N} \stackrel{\text{def}}{=} \bigcup_{k=0}^{\infty} \boldsymbol{N}_k.$$

Thus, for $\boldsymbol{x} \notin \boldsymbol{N}$, the integral $\displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}$ exists. We just proved that if f and g are convolvable, then for almost every $\boldsymbol{x} \in \mathbb{R}^n$ the integral $\displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}$ exists (or, stated differently, for almost every $\boldsymbol{x} \in \mathbb{R}^n$ the function $\boldsymbol{y} \mapsto f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})$ is in $L^1(\mathbb{R}^n;\mathbb{R})$). This gives a meaning to the following definition.

Definition 5.31 [Convolution]
Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R})$ be convolvable. We define

$$(f * g)(\boldsymbol{x}) \stackrel{def}{=} \begin{cases} \displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} & \text{if the integral exists (i.e., for } \boldsymbol{x} \notin \boldsymbol{N}) \\ 0 & \text{otherwise (i.e., for } \boldsymbol{x} \in \boldsymbol{N}) \end{cases} \tag{5.61}$$

First, we notice that the term "convolution" refers both to the operation and the result of that operation. Notice also that the zero in the definition of $f * g$, whenever the integral $\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y}$ does not exist, is simply an arbitrary value. Its value is irrelevant, since we identify functions that coincide almost everywhere.

We observe that the convolution is locally integrable. In other words, if f and g are convolvable, then $f * g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$. Indeed, since f and g are convolvable, for every compact subset $\boldsymbol{K}$ of $\mathbb{R}^n$, in view of Fubini's theorem (Theorem 3.8), we find

$$\begin{aligned} \int_{\boldsymbol{K}} \left|(f*g)(\boldsymbol{x})\right| \mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{K}} \left| \int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\, g(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} \right| \mathrm{d}\boldsymbol{x} \\ &\le \int_{\boldsymbol{K}} \left(\int_{\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\, |g(\boldsymbol{y})|\, \mathrm{d}\boldsymbol{y} \right) \mathrm{d}\boldsymbol{x} \\ &= \int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\, |g(\boldsymbol{y})|\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y} < +\infty. \end{aligned}$$

5.7 Sufficient conditions for convolvability

As we observed, it is impossible to decouple the role of the functions f and g in the convolution integral. This means, in particular, that we have only sufficient conditions on the two functions for them to be convolvable.

Theorem 5.37 *Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$. Let g be with compact support. Then, f and g are convolvable.*

Proof Let $\boldsymbol{K}$ be a compact subset of $\mathbb{R}^n$. We show that

$$\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < +\infty.$$

If $\operatorname{supp} g \subset \boldsymbol{H}$, where $\boldsymbol{H}$ is a compact set, and we consider the mapping $A(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})$, we have

$$\begin{aligned}\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} &= \int_{\boldsymbol{K}\times\boldsymbol{H}} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})\,|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\boldsymbol{K}\times\boldsymbol{H}} |(f\otimes g)(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\boldsymbol{K}\times\boldsymbol{H}} |(f\otimes g)(A(\boldsymbol{x},\boldsymbol{y}))|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{A(\boldsymbol{K}\times\boldsymbol{H})} |(f\otimes g)(\boldsymbol{u},\boldsymbol{v})|\,\mathrm{d}\boldsymbol{u}\,\mathrm{d}\boldsymbol{v} < +\infty,\end{aligned}$$

since $A(\boldsymbol{K}\times\boldsymbol{H})$ is compact in $\mathbb{R}^{2n}$ (continuous functions map compact sets into compact sets). Here, we used the change of variable $A(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{u},\boldsymbol{v})$, so that $A^{-1}(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{u}+\boldsymbol{v},\boldsymbol{v})$.

The following theorem shows that the convolution is commutative.

Theorem 5.38 *Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$. Let f and g be convolvable. Then, g and f are convolvable and*

$$f * g = g * f.$$

Proof First, we prove that if f and g are convolvable, then also g and f are convolvable, since for every compact subset K of $\mathbb{R}^n$ we have

$$\int_K \Big(\int_{\mathbb{R}^n} |g(\boldsymbol{x}-\boldsymbol{y})|\,|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\Big)\,\mathrm{d}\boldsymbol{x} = \int_K \Big(\int_{\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{z})|\,|g(\boldsymbol{z})|\,\mathrm{d}\boldsymbol{z}\Big)\,\mathrm{d}\boldsymbol{x}$$

by changing variable with $\boldsymbol{x}-\boldsymbol{y} = \boldsymbol{z}$ and Fubini's theorem (Theorem 3.8). The first integral is finite if and only if the second is finite, i.e., g and f are convolvable if and only if f and g are convolvable.

To show that the convolution is the same, we recall that for almost every $\boldsymbol{x}\in\mathbb{R}^n$ the integral $\displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}$ exists, and therefore also the integral $\displaystyle\int_{\mathbb{R}^n} g(\boldsymbol{x}-\boldsymbol{z})\,f(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}$ exists for almost every $\boldsymbol{x}\in\mathbb{R}^n$. Since

$$\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n} g(\boldsymbol{x}-\boldsymbol{z})\,f(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z} \tag{5.62}$$

almost everywhere in $\mathbb{R}^n$, we find $f * g = g * f$.

Example *5.51* If we consider the case $n = 1$, the change of variable $s = t-\tau$ in the convolution would lead to a minus sign:

$$\int_{\mathbb{R}} f(t-\tau)\,g(\tau)\,\mathrm{d}\tau = -\int_{\mathbb{R}} f(s)\,g(t-s)\,\mathrm{d}s.$$

What happened to the minus sign in (5.62)? The answer is in the end-points of the interval of integration.

The following result shows that, if $f, g \in L^1(\mathbb{R}^n;\mathbb{R})$ (and not just in $L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{R})$), the convolution is defined and it is also in $L^1(\mathbb{R}^n;\mathbb{R})$. Even more is true: one of the two functions is allowed to live in $L^1(\mathbb{R}^n;\mathbb{R})$, and the other in $L^p(\mathbb{R}^n;\mathbb{R})$ for any $p \in [1,+\infty]$: the functions are still convolvable.

Theorem 5.39 *Let $p \in [1,+\infty]$. Let $f \in L^p(\mathbb{R}^n;\mathbb{R})$ and $g \in L^1(\mathbb{R}^n;\mathbb{R})$. Then:*

(i) f and g are convolvable;

(ii) $f * g \in L^p(\mathbb{R}^n;\mathbb{R})$;

(iii) $\|f * g\|_p \le \|f\|_p \, \|g\|_1$ (Hausdorff-Young Inequality).

Proof We prove this result for $p = 1$ and $p = 2$ (the case $p = +\infty$ is left as an exercise).

- Consider the case $p = 1$. Let $\boldsymbol{K}$ be a compact subset of $\mathbb{R}^n$. We want to show that

$$\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < +\infty.$$

  Then, using Fubini's theorem (Theorem 3.8) we find

$$\begin{aligned}\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} &\le \int_{\mathbb{R}^n\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\right)|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(\boldsymbol{z})|\,\mathrm{d}\boldsymbol{z}\right)|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^n} |f(\boldsymbol{z})|\,\mathrm{d}\boldsymbol{z}\int_{\mathbb{R}^n} |g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} = \|f\|_1\,\|g\|_1 < +\infty.\end{aligned}$$

  It follows that f and g are convolvable.

  From this sequence of steps, we deduce also that the map $(\boldsymbol{x},\boldsymbol{y}) \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ is in $L^1(\mathbb{R}^{2n};\mathbb{R})$, and from Fubini's theorem (Theorem 3.8) for almost every $\boldsymbol{x} \in \mathbb{R}^n$ the map $\boldsymbol{y} \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})$ is in $L^1(\mathbb{R}^n;\mathbb{R})$. Hence,

$$\int_{\mathbb{R}^{2n}} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n}(f*g)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

  from which we find that $f * g$ is in $L^1(\mathbb{R}^n;\mathbb{R})$. Finally,

$$\|f*g\|_1 = \int_{\mathbb{R}^n}|(f*g)(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \le \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x} = \|f\|_1\,\|g\|_1.$$

- Consider the case $p = 2$. Let $\boldsymbol{K} \subseteq \mathbb{R}^n$ be compact. We want to show that

$$\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < +\infty.$$

From Fubini's theorem (Theorem 3.8), we have

$$\int_{K\times\mathbb{R}^n} |f(x-y)|\,|g(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^n}\left(\int_K |f(x-y)|\,\mathrm{d}x\right)|g(y)|\,\mathrm{d}y.$$

We now consider the internal integral and apply the Cauchy-Schwarz's inequality in $L^2(\mathbb{R}^n;\mathbb{R})$:

$$\begin{aligned}\int_K |f(x-y)|\,\mathrm{d}x &= \int_{\mathbb{R}^n} |f(x-y)|\chi_K(x)\,\mathrm{d}x = \langle |f(\cdot-y)|,\chi_K\rangle\\ &\le \left(\int_{\mathbb{R}^n} |f(x-y)|^2\,\mathrm{d}x\right)^{\frac12}\left(\int_{\mathbb{R}^n}\chi_K(x)\,\mathrm{d}x\right)^{\frac12}\\ &\le \left(\int_{\mathbb{R}^n}|f(z)|^2\,\mathrm{d}z\right)^{\frac12}\sqrt{\mu_n(K)} = \|f\|_2\sqrt{\mu_n(K)},\end{aligned}$$

where we denoted the inner product in $L^2(\mathbb{R}^n;\mathbb{R})$ with $\langle\cdot,\cdot\rangle$. It follows that

$$\int_{K\times\mathbb{R}^n}|f(x-y)|\,|g(y)|\,\mathrm{d}x\,\mathrm{d}y \le \|f\|_2\sqrt{\mu_n(K)}\int_{\mathbb{R}^n}|g(y)|\,\mathrm{d}y < +\infty,$$

since $g\in L^1(\mathbb{R}^n;\mathbb{R})$. This proves that f and g are convolvable.

Now we prove that the convolution is a function in $L^2(\mathbb{R}^n;\mathbb{R})$. For almost every $x\in\mathbb{R}^n$ we have

$$\begin{aligned}|(f*g)(x)| &\le \int_{\mathbb{R}^n}|f(x-y)|\,|g(y)|\,\mathrm{d}y\\ &= \int_{\mathbb{R}^n}|f(x-y)|\,|g(y)|^{\frac12}\,|g(y)|^{\frac12}\mathrm{d}y\\ &\le \left(\int_{\mathbb{R}^n}|f(x-y)|^2\,|g(y)|\,\mathrm{d}y\right)^{\frac12}\left(\int_{\mathbb{R}^n}|g(y)|\,\mathrm{d}y\right)^{\frac12}\\ &= \|g\|_1^{\frac12}\left(\int_{\mathbb{R}^n}|f(x-y)|^2\,|g(y)|\,\mathrm{d}y\right)^{\frac12},\end{aligned}$$

where we have used again Cauchy-Schwarz's inequality. Thus,

$$\begin{aligned}\int_{\mathbb{R}^n}|(f*g)(x)|^2\,\mathrm{d}x &\le \int_{\mathbb{R}^n}\|g\|_1\left(\int_{\mathbb{R}^n}|f(x-y)|^2\,|g(y)|\,\mathrm{d}y\right)\mathrm{d}x\\ &= \|g\|_1\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|f(x-y)|^2\,\mathrm{d}x\right)|g(y)|\,\mathrm{d}y = \|f\|_2^2\,\|g\|_1^2.\end{aligned}$$

This implies $f*g\in L^2(\mathbb{R}^n;\mathbb{R})$. We use the same equality to prove the last point:

$$\|f*g\|_2 = \left(\int_{\mathbb{R}^n}|(f*g)(x)|^2\,\mathrm{d}x\right)^{\frac12} \le \|f\|_2\,\|g\|_1.$$

We now present one last condition for a pair of functions to be convolvable. This condition applies to functions over $\mathbb{R}$.

Theorem 5.40 *Let $f,g\in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Let* $\operatorname{supp} f\subseteq[0,+\infty)$ *and* $\operatorname{supp} g\subseteq[0,+\infty)$. *Then:*

(i) f and g are convolvable;

(ii) $\operatorname{supp}(f * g) \in [0, +\infty)$.

Proof Let $K \subseteq \mathbb{R}$ be compact. We want to prove that

$$\int_{K\times\mathbb{R}} |f(x-y)|\,|g(y)|\,\mathrm{d}x\,\mathrm{d}y < +\infty.$$

We know that $g(y)$ is zero for $y < 0$ and $f(x-y)$ is zero for $y > x$. Let $K \subseteq [-r, r]$ for some $r > 0$. If $f(x-y)\,g(y) \neq 0$, then we must have $y \geq 0$ and $y \leq x$. It follows that $0 \leq y \leq x \leq r$ (in other words, the product can be different from zero only if $0 \leq y \leq x \leq r$, and y cannot be outside $[0, r]$). Thus,

$$\int_{K\times\mathbb{R}} |f(x-y)|\,|g(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{K\times[0,r]} |f(x-y)|\,|g(y)|\,\mathrm{d}x\,\mathrm{d}y,$$

and $K \times [0, r]$ is a compact subset of $\mathbb{R}^2$. We have already seen that the function $(x, y) \mapsto f(x-y)\,g(y)$ is locally integrable, and therefore is in $L^1_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{R})$. Thus, the functions f and g are convolvable.

Now we need to show that $\operatorname{supp}(f * g) \subseteq [0, +\infty)$. We have

$$(f * g)(x) = \int_{\mathbb{R}} f(x-y)\,g(y)\,\mathrm{d}y.$$

It is easy to see that, if $x < 0$, we have $f(x-y)\,g(y) = 0$ for all $y \in \mathbb{R}$. Indeed, if $y \geq 0$, we have $x - y < 0$, so that $f(x-y) = 0$ since $\operatorname{supp} f \subseteq [0, +\infty)$. If $y < 0$, then $g(y) = 0$ since $\operatorname{supp} g \subseteq [0, +\infty)$. In both cases, $f(x-y)\,g(y) = 0$. This allows us to conclude that $\operatorname{supp}(f * g) \subseteq [0, +\infty)$.

5.8 Convolution of distributions

Our task now is to extend the convolution to distributions. Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ be convolvable. Then, $f * g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ defines a distribution of function type. Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We can write

$$\langle f * g, \varphi\rangle = \int_{\mathbb{R}^n} (f * g)(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\Big)\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We want to make sure that this integral exists. Recall that by definition of convolvable function, we have

$$\int_{\boldsymbol{K}\times\mathbb{R}^n} |f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} < \infty$$

for every compact set $\boldsymbol{K}$ of $\mathbb{R}^n$. This condition can be alternatively expressed by saying that the function

$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\chi_{\boldsymbol{K}}(\boldsymbol{x})$$

is integrable in $\mathbb{R}^{2n}$ for every compact set $\boldsymbol{K}$ of $\mathbb{R}^n$, where $\chi_{\boldsymbol{K}}$ is the characteristic function of $\boldsymbol{K}$, i.e., such that

$$\chi_{\boldsymbol{K}}(\boldsymbol{x}) = \begin{cases} 1 & \boldsymbol{x} \in \boldsymbol{K} \\ 0 & \text{otherwise.} \end{cases}$$

If $\boldsymbol{K}$ is such that $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$, then $|\varphi(\boldsymbol{x})| \leq \sup_{\boldsymbol{x}\in\boldsymbol{K}} |\varphi(\boldsymbol{x})|\,\chi_{\boldsymbol{K}}(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$, which implies that

$$|f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,|\varphi(\boldsymbol{x})| \leq \sup_{\boldsymbol{x}\in\boldsymbol{K}} |\varphi|\,|f(\boldsymbol{x}-\boldsymbol{y})|\,|g(\boldsymbol{y})|\,\chi_{\boldsymbol{K}}(\boldsymbol{x})$$

is integrable in $\mathbb{R}^n$. Using Fubini's theorem (Theorem 3.8)

$$\begin{aligned}
\int_{\mathbb{R}^n} (f * g)(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\Big)\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} f(\boldsymbol{z})\,g(\boldsymbol{x}-\boldsymbol{z})\,\mathrm{d}\boldsymbol{z}\Big)\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}^n} f(\boldsymbol{z}) \Big(\int_{\mathbb{R}^n} g(\boldsymbol{x}-\boldsymbol{z})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\Big)\mathrm{d}\boldsymbol{z} \\
&= \int_{\mathbb{R}^n} f(\boldsymbol{z}) \Big(\int_{\mathbb{R}^n} g(\boldsymbol{y})\,\varphi(\boldsymbol{z}+\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\Big)\mathrm{d}\boldsymbol{z} \\
&= \int_{\mathbb{R}^n} f(\boldsymbol{x}) \Big(\int_{\mathbb{R}^n} g(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\Big)\mathrm{d}\boldsymbol{x} \\
&= \langle (f\otimes g)(\boldsymbol{x},\boldsymbol{y}), (\varphi\circ s)(\boldsymbol{x},\boldsymbol{y})\rangle,
\end{aligned}$$

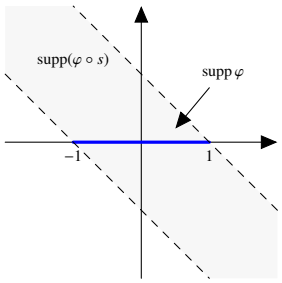
where

$$\begin{aligned}
s : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\
(\boldsymbol{x},\boldsymbol{y}) &\mapsto \boldsymbol{x}+\boldsymbol{y}.
\end{aligned}$$

This seems to suggest that a valid definition for convolution for distributions could be

$$\langle f * g, \varphi\rangle = \langle f\otimes g, \varphi\circ s\rangle,$$

Unfortunately, however, this is not a correct definition for the convolution of distributions. The reason is that $\varphi\circ s$ does not have compact support in $\mathbb{R}^{2n}$. Indeed, consider for example the case $n=1$, and take any $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$ with $\varphi(x)\neq 0$ for $|x|<1$.



Then $\varphi(x+y)\neq 0$ for every pair (x,y) such that $|x+y|<1$, and such pairs live in the strip of $\mathbb{R}^2$ within the lines $x+y=\pm 1$. Thus, the support of $\varphi\circ s$ is not bounded, and $\varphi\circ s\notin\mathcal{D}(\mathbb{R}^2;\mathbb{R})$. A solution to this problem is to construct a sequence that approximates $\varphi\circ s$. To this end, we introduce the notion of sequence of functions in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ converging to unity.

Definition 5.32 [Sequences of Functions Converging to Unity]
Consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ of elements of $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We say that $(\eta_k)_{k\in\mathbb{N}}$ converges to unity if

(i) for every compact subset $\boldsymbol{K}$ of $\mathbb{R}^n$ there exists $p_{\boldsymbol{K}}$ such that for all $k\geq p_K$ we have $\eta_k(\boldsymbol{x})=1$ for all $\boldsymbol{x}\in\boldsymbol{K}$;

(ii) for all $\alpha\in\mathbb{N}^n$ there exists $c(\alpha)>0$ such that for all $k\in\mathbb{N}$ and for all $\boldsymbol{x}\in\mathbb{R}^n$ there holds $|\partial^\alpha_{\boldsymbol{x}}\eta_k(\boldsymbol{x})|\leq c(\alpha)$.

We express the second condition in Definition 5.32 by saying that the sequence $(\eta_k)_{k\in\mathbb{N}}$ is uniformly bounded, together with its derivatives of any order, in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$.

Example 5.52

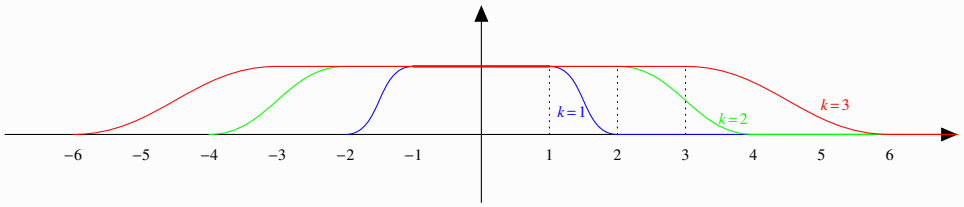
Consider a function $\eta \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ such that

$$\eta(\boldsymbol{x}) = \begin{cases} 1 & \|\boldsymbol{x}\| \leq 1 \\ 0 & \|\boldsymbol{x}\| \geq 2. \end{cases}$$

Let us define, for $k \in \mathbb{N}^*$, the function $\eta_k(\boldsymbol{x}) = \eta\left(\frac{\boldsymbol{x}}{k}\right)$, so that

$$\eta_k(\boldsymbol{x}) = \begin{cases} 1 & \|\boldsymbol{x}\| \leq k \\ 0 & \|\boldsymbol{x}\| \geq 2\,k. \end{cases}$$

From the way it was constructed, it is clear the this sequence satisfies the first condition of Definition 5.32. Indeed, given any compact set $\boldsymbol{K}$ of $\mathbb{R}^n$, we have $K \subseteq \overline{\mathcal{B}_R(0)}$, with a sufficiently large value of $R > 0$. Since $\eta_k(\boldsymbol{x}) = 1$ if $\|\boldsymbol{x}\| \leq k$, if $k > R$, we have $\eta_k(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{K}$.



It also satisfies the second condition since

$$\partial^\alpha_{\boldsymbol{x}} \eta_k(\boldsymbol{x}) = \partial^\alpha_{\boldsymbol{x}} \eta\left(\frac{\boldsymbol{x}}{k}\right) = k^{-|\alpha|} \partial^\alpha_{\frac{\boldsymbol{x}}{k}} \eta\left(\frac{\boldsymbol{x}}{k}\right),$$

so that $|\partial^\alpha_{\boldsymbol{x}} \eta_k(\boldsymbol{x})| \leq \sup |\partial^\alpha_{\boldsymbol{x}} \eta| = c(\alpha)$, since $\partial^\alpha_{\boldsymbol{x}} \eta$ is bounded in $\mathbb{R}^n$ (as η is of class C^∞).

It follows that for every $k \in \mathbb{N}$ we can apply the tensor product to η_k

$$\langle f \otimes g, (\varphi \circ s)\, \eta_k \rangle,$$

and require that the limit of this expression exists. We also need to impose that the value of the limit does not depend on η_k.

Lemma 5.14 *Let $(\eta_k)_{k\in\mathbb{N}}$ be convergent to unity in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Then, for all $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$, the sequence $(\eta_k\, \varphi)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$.*

The proof is left as an exercise (see Exercise 5.68).

Clearly, now for any $k \in \mathbb{N}$ the object

$$\langle f \otimes g, (\varphi \circ s)\, \eta_k \rangle$$

is well-defined. However, it depends on the choice of η_k.

Definition 5.33 [Convolvable Distributions]
Let $T, S \in \mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. We say that T and S are convolvable if for every $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ and for every sequence $(\eta_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{C})$ converging to unity in $\mathbb{R}^{2n}$ the limit

$$\lim_{k\to\infty} \langle (T \otimes S)(\boldsymbol{x},\boldsymbol{y}), \varphi(\boldsymbol{x}+\boldsymbol{y})\, \eta_k(\boldsymbol{x},\boldsymbol{y}) \rangle$$

exists (in $\mathbb{C}$) and is independent from $(\eta_k)_{k\in\mathbb{N}}$.

We now need to define the convolution of two convolvable distributions T and S. First, for every $k \in \mathbb{N}$ the functional

$$\begin{aligned} \mathcal{D}(\mathbb{R}^n;\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \langle (T\otimes S)(\boldsymbol{x},\boldsymbol{y}), \varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle \end{aligned}$$

can be shown to be linear and continuous over $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$. At this point, using the Completeness Theorem, the function that maps φ to the limit is a distribution:

$$\begin{aligned} T * S : \mathcal{D}(\mathbb{R}^n;\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \lim_{k\to\infty} \langle (T\otimes S)(\boldsymbol{x},\boldsymbol{y}), \varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle \end{aligned}$$

The following theorem shows that every distribution is convolvable with the Dirac delta distribution, and that the Dirac delta is the identity element of the operation of convolution.

Theorem 5.41 *Let $T \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. Then, T and δ are convolvable, and $T * \delta = T$.*

Proof Consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity in $\mathbb{R}^{2n}$. We have

$$\begin{aligned} \langle (T\otimes\delta), (\varphi\circ s)\,\eta_k\rangle &= \langle T(\boldsymbol{x}), \langle \delta(\boldsymbol{y}), \varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \\ &= \langle T(\boldsymbol{x}), \varphi(\boldsymbol{x})\,\eta_k(\boldsymbol{x},0)\rangle. \end{aligned}$$

Consider the compact set $K = \operatorname{supp}\varphi \times \{0\}$ of $\mathbb{R}^{2n}$. For large values of k we have $\eta_k(\boldsymbol{x},\boldsymbol{y}) = 1$ if $(\boldsymbol{x},\boldsymbol{y}) \in K$, i.e., $\eta_k(\boldsymbol{x},0) = 1$ if $\boldsymbol{x} \in \operatorname{supp}\varphi$. It follows that for large values of k we have

$$\varphi(\boldsymbol{x})\,\eta_k(\boldsymbol{x},0) = \varphi(\boldsymbol{x}),$$

for every $\boldsymbol{x} \in \mathbb{R}^n$, so that if k is sufficiently large

$$\langle (T\otimes\delta), (\varphi\circ s)\,\eta_k\rangle = \langle T(\boldsymbol{x}), \varphi(\boldsymbol{x})\,\eta_k(\boldsymbol{x},0)\rangle = \langle T,\varphi\rangle,$$

and therefore $T * \delta = T$.

The following result shows that the definition of convolution for distributions that we have given is consistent with the convolution of functions in $L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$.

Theorem 5.42 *Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$ be convolvable in $L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$. Then, f and g are convolvable as distributions, and the two convolutions coincide.*

Proof Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$, and consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity in $\mathbb{R}^{2n}$. Then,

$$\begin{aligned} \langle f\otimes g, (\varphi\circ s)\,\eta_k\rangle &= \int_{\mathbb{R}^n} f(\boldsymbol{\xi}) \left(\int_{\mathbb{R}^n} g(\boldsymbol{\tau})\,\varphi(\boldsymbol{\xi}+\boldsymbol{\tau})\,\eta_k(\boldsymbol{\xi},\boldsymbol{\tau})\,\mathrm{d}\boldsymbol{\tau}\right) \mathrm{d}\boldsymbol{\xi} \\ &= \int_{\mathbb{R}^{2n}} f(\boldsymbol{\xi})\,g(\boldsymbol{\tau})\,\varphi(\boldsymbol{\xi}+\boldsymbol{\tau})\,\eta_k(\boldsymbol{\xi},\boldsymbol{\tau})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\tau} \\ &= \int_{\mathbb{R}^{2n}} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\varphi(\boldsymbol{x})\,\eta_k(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}, \end{aligned}$$

where the substitutions $\boldsymbol{x}-\boldsymbol{y}=\boldsymbol{\xi}$ and $\boldsymbol{y}=\boldsymbol{\tau}$ have been used (and the absolute value of the determinant of the Jacobian of the mapping $(\boldsymbol{x},\boldsymbol{y})\mapsto(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})$ is 1). Since f and g are convolvable, the function $(\boldsymbol{x},\boldsymbol{y})\mapsto f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\varphi(\boldsymbol{x})$ is in $L^1(\mathbb{R}^{2n};\mathbb{R})$. Using Lebesgue's dominated convergence theorem (Theorem 3.47), that we can use since $|\eta_k(\boldsymbol{x}-\boldsymbol{y},\boldsymbol{y})|\leq c_0$ where c_0 is a constant, we have

$$\begin{aligned}\lim_{k\to\infty}\langle f\otimes g,(\varphi\circ s)\,\eta_k\rangle &= \int_{\mathbb{R}^{2n}} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\varphi(\boldsymbol{x})\;\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\varphi(\boldsymbol{x})\;\mathrm{d}\boldsymbol{x}=\langle f*g,\varphi\rangle,\end{aligned}$$

where Fubini's theorem (Theorem 3.8) has been used again. This is exactly the definition of convolution in $L^1_{\mathrm{loc}}(\mathbb{R}^n;\mathbb{R})$.

We now show that the commutative property also holds for distributions.

Theorem 5.43 *Let $T,S\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ be convolvable. Then, S and T are also convolvable, and*

$$T*S=S*T.$$

Proof From the properties of the tensor product we have that for $\psi\in\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ there holds

$$\langle T\otimes S,\psi\rangle=\Big\langle T(\boldsymbol{x}),\langle S(\boldsymbol{y}),\psi(\boldsymbol{x},\boldsymbol{y})\rangle\Big\rangle=\Big\langle S(\boldsymbol{y}),\langle T(\boldsymbol{x}),\psi(\boldsymbol{x},\boldsymbol{y})\rangle\Big\rangle.$$

Thus, given $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ and a sequence $(\eta_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity in $\mathbb{R}^{2n}$, we have

$$\begin{aligned}\langle T*S,\varphi\rangle=\lim_{k\to\infty}\langle T\otimes S,(\varphi\circ s)\,\eta_k\rangle &= \lim_{k\to\infty}\Big\langle T(\boldsymbol{x}),\langle S(\boldsymbol{y}),\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle\Big\rangle\\ &= \lim_{k\to\infty}\Big\langle S(\boldsymbol{y}),\langle T(\boldsymbol{x}),\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle\Big\rangle=\langle S*T,\varphi\rangle.\end{aligned}$$

The convolution is a linear operation in the first and in the second argument: if $T,T_1,T_2,S,S_1,S_2\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ and $\lambda,\mu\in\mathbb{R}$, then:

- If T_1*S and T_2*S exist, then $(\lambda T_1+\mu T_2)*S$ exists and

$$(\lambda T_1+\mu T_2)*S=\lambda(T_1*S)+\mu(T_2*S).$$

- If $T*S_1$ and $T*S_2$ exist, then $T*(\lambda S_1+\mu S_2)$ exists and

$$T*(\lambda S_1+\mu S_2)=\lambda(T*S_1)+\mu(T*S_2).$$

Consider for example the second. Since $T*S_1$ and $T*S_2$ exist, then also $T*(\lambda S_1)$ and $T*(\mu S_2)$ exist, which implies that $T*(\lambda S_1+\mu S_2)$ exists. Let $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$, and let $(\eta_k)_{k\in\mathbb{N}}$ be a sequence in

$\mathcal{D}'(\mathbb{R}^{2n};\mathbb{R})$ converging to unity in $\mathbb{R}^{2n}$. Then

$$\begin{aligned}
\langle T*(\lambda S_1+\mu S_2),\varphi\rangle &= \lim_{k\to\infty}\Big\langle (T\otimes(\lambda S_1+\mu S_2))(\boldsymbol{x},\boldsymbol{y}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\Big\rangle\\
&= \lim_{k\to\infty}\Big\langle T(\boldsymbol{x}),\langle \lambda S_1(\boldsymbol{x})+\mu S_2(\boldsymbol{x}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle\\
&= \lim_{k\to\infty}\Big(\Big\langle T(\boldsymbol{x}),\langle \lambda S_1(\boldsymbol{x}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle+\Big\langle T(\boldsymbol{x}),\langle \mu S_2(\boldsymbol{y}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle\Big)\\
&= \lim_{k\to\infty}\Big(\lambda\Big\langle T(\boldsymbol{x}),\langle S_1(\boldsymbol{x}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle+\mu\Big\langle T(\boldsymbol{x}),\langle S_2(\boldsymbol{y}),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle\Big)\\
&= \lim_{k\to\infty}\Big(\lambda\langle (T\otimes S_1),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle+\mu\langle (T\otimes S_2),\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big)\\
&= \langle\lambda(T*S_1),\varphi\rangle+\mu(T*S_2),\varphi\rangle = \langle\lambda(T*S_1)+\mu(T*S_2),\varphi\rangle.
\end{aligned}$$

The following two theorems give conditions for two distributions to be convolvable.

Theorem 5.44 *Let $T,S\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. If either T or S has compact support, then, T and S are convolvable.*

Proof Let $\operatorname{supp} S\subseteq \boldsymbol{K}$, where $\boldsymbol{K}$ is compact. Let $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ and let $(\eta_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity in $\mathbb{R}^{2n}$. We have to prove that

$$\langle T(\boldsymbol{x}),\langle S(\boldsymbol{y}),\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle\rangle$$

admits limit as $k\to\infty$, and the value of the limit does not depend on η_k. We show in particular that from a certain value of k onward the limit is constant. Let $\zeta\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ be such that $\zeta(\boldsymbol{y})=1$ if $\boldsymbol{y}\in\boldsymbol{U}$, where $\boldsymbol{U}$ is an open set containing $\boldsymbol{K}$. Thus, $\zeta S=S$. It follows that

$$\langle S(\boldsymbol{y}),\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle=\langle S(\boldsymbol{y}),\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle.$$

The function $(\boldsymbol{x},\boldsymbol{y})\mapsto\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})$ is in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$. To see this, we only need to show that it has compact support in $\mathbb{R}^{2n}$. Indeed, there exist $r,s>0$ such that $\operatorname{supp}\zeta\subseteq\overline{\mathcal{B}_r(0)}$ and $\operatorname{supp}\varphi\subseteq\overline{\mathcal{B}_s(0)}$. For $\|\boldsymbol{y}\|\le r$ and $\|\boldsymbol{x}+\boldsymbol{y}\|\le s$, we have

$$\|\boldsymbol{x}\|=\|\boldsymbol{x}+\boldsymbol{y}-\boldsymbol{y}\|\le\|\boldsymbol{x}+\boldsymbol{y}\|+\|\boldsymbol{y}\|\le s+r.$$

It follows that the support of the function $(\boldsymbol{x},\boldsymbol{y})\mapsto\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})$ is contained in $\overline{\mathcal{B}_{r+s}(0)}\times\overline{\mathcal{B}_r(0)}$. Hence, it has compact support. Since for a sufficiently large value of k we have $\eta_k(\boldsymbol{x},\boldsymbol{y})=1$ over the support, we have that for sufficiently large values of k

$$\forall\,(\boldsymbol{x},\boldsymbol{y})\in\mathbb{R}^{2n},\quad \zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})=\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y}).$$

It follows that for sufficiently large values of k

$$\langle S(\boldsymbol{y}),\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\,\eta_k(\boldsymbol{x},\boldsymbol{y})\rangle=\langle S(\boldsymbol{y}),\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle,$$

and therefore, again for a sufficiently large k,

$$\langle T\otimes S,(\varphi\circ s)\,\eta_k\rangle=\langle T(\boldsymbol{x}),\langle S(\boldsymbol{y}),\zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\rangle.$$

The following result focuses on the case $n=1$.

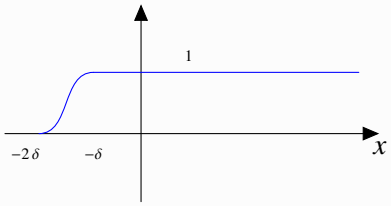
Theorem 5.45 *Let $T,S\in\mathcal{D}'(\mathbb{R};\mathbb{R})$. If $\operatorname{supp}T\cup\operatorname{supp}S\subseteq[0,+\infty)$, then, T and S are convolvable, and $\operatorname{supp}(T*S)\subseteq[0,+\infty)$.*

Proof Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ and let $(\eta_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}'(\mathbb{R}^2;\mathbb{R})$ converging to unity in $\mathbb{R}^2$. As in the proof of the previous theorem, we show that the limit as $k \to \infty$ of

$$\langle T(x), \langle S(y), \varphi(x+y)\,\eta_k(x,y)\rangle\rangle$$

becomes constant from a certain value of k onward. Let $\delta > 0$. Let $\zeta \in C^\infty(\mathbb{R};\mathbb{R})$ be such that

$$\zeta(x) = \begin{cases} 1 & \text{if } x \geq -\delta \\ 0 & \text{if } x \leq -2\,\delta. \end{cases}$$



In other words, ζ is equal to 1 on an open set that contains $[0,+\infty)$, so that $\zeta\,T = T$ and $\zeta\,S = S$. It follows that

$$\begin{aligned} \langle T(x), \langle S(y), \varphi(x+y)\,\eta_k(x,y)\rangle\rangle &= \langle \zeta(x)\,T(x), \langle \zeta(y)\,S(y), \varphi(x+y)\,\eta_k(x,y)\rangle\rangle \\ &= \langle T(x), \langle \zeta(x)\,\zeta(y)\,S(y), \varphi(x+y)\,\eta_k(x,y)\rangle\rangle \\ &= \langle T(x), \langle S(y), \zeta(x)\,\zeta(y)\,\varphi(x+y)\,\eta_k(x,y)\rangle\rangle \end{aligned}$$

We now show that the function $(x,y) \mapsto \zeta(x)\,\zeta(y)\,\varphi(x+y)\,\eta_k(x,y)$ has compact support, and this will be enough to prove that it is in $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$. For $\zeta(x)\,\zeta(y)\,\varphi(x+y)\,\eta_k(x,y)$, we need $x \geq -2\,\delta$ and $y \geq -2\,\delta$. Moreover, φ has compact support, and if $\operatorname{supp}\varphi \subseteq [-r,r]$, then $-r \leq x+y \leq r$. Thus, we have three conditions:

$$\begin{cases} x \geq -2\,\delta \\ y \geq -2\,\delta \\ -r \leq x+y \leq r \end{cases}$$

The last one says that $x \leq r - y$, and using the second written as $-y \leq 2\,\delta$ we obtain $x \leq r + 2\,\delta$. Thus, for the product to be $\neq 0$ we must have

$$\begin{cases} -2\,\delta \leq x \leq r + 2\,\delta \\ -2\,\delta \leq y \leq r + 2\,\delta \end{cases}$$

It follows that the support of $(x,y) \mapsto \zeta(x)\,\zeta(y)\varphi(x+y)\,\eta_k(x,y)$ is contained in the square $[-2\,\delta, r+2\,\delta] \times [-2\,\delta, r+2\,\delta]$. For a sufficiently large value of k, the value of $\eta_k(x,y)$ is equal to 1 on this square, so that

$$\zeta(x)\,\zeta(y)\,\varphi(x+y)\,\eta_k(x,y) = \zeta(x)\,\zeta(y)\,\varphi(x+y),$$

where $\zeta(x) = 1$ in $(-\delta,+\infty)$ for some $\delta > 0$. Now, we show that $\operatorname{supp}(T * S) \subseteq [0,+\infty)$. Let $x_0 < 0$. We show that $x_0 \notin \operatorname{supp}(T * S)$, i.e., that there is a neighbourhood of x_0 such that, if we have $\operatorname{supp}\varphi$ in this neighbourhood, the convolution is equal to 0. Let $\delta' > 0$ be such that $x_0 + \delta' < 0$, and we verify that

$$\operatorname{supp}\varphi \subseteq (x_0 - \delta, x_0 + \delta) \quad \Rightarrow \quad \langle T * S, \varphi\rangle = 0.$$

The function ζ can be chosen with an arbitrarily small δ. For an arbitrarily small $\delta > 0$, the function $\zeta(x)\,\zeta(y)\,\varphi(x+y)$ is identically zero, i.e.,

$$\zeta(x)\,\zeta(y)\,\varphi(x+y) = 0 \quad \forall\,(x,y) \in \mathbb{R}^2.$$

We have $\varphi(x+y) \neq 0$ when $x + y \leq x_0 + \delta'$. Moreover, $\zeta(x) \neq 0$ when $x \geq -2\,\delta$ and $\zeta(y) \neq 0$ when $y \geq -2\,\delta$. Thus, to have $\zeta(x)\,\zeta(y)\,\varphi(x+y) \neq 0$ we must have

$$\begin{cases} x + y \leq x_0 + \delta' \\ x \geq -2\,\delta \\ y \geq -2\,\delta \end{cases}$$

The last two say that $x + y \geq -4\,\delta$, and choosing $\delta > 0$ so that $-4\,\delta > x_0 + \delta'$ we find

$$\zeta(x)\,\zeta(y)\,\varphi(x+y) = 0 \quad \forall\,(x,y) \in \mathbb{R}^2.$$

It follows that $\langle T * S, \varphi\rangle = 0$. To summarise, we took a point $x_0 < 0$, and we found an interval of x such that, choosing φ with support in this neighbourhood, the convolution is zero. It follows that x_0 does not belong to the support.

5.8.1 Derivative of the convolution

We begin with the following result.

Lemma 5.15 *Let $T, S \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ be convolvable. Consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ convergent to unity in $\mathcal{D}(\mathbb{R}^{2\,n};\mathbb{R})$. Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Then,*

$$\lim_{k\to\infty} \langle (T\otimes S)(\boldsymbol{x},\boldsymbol{y}), \partial^{\alpha}_{\boldsymbol{x},\boldsymbol{y}}\eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle = 0$$

for every multi-index $\alpha \in \mathbb{N}^{2\,n}$ such that $|\alpha| \geq 1$.

Proof Consider the sequence $(\eta_k + \partial^{\alpha}_{\boldsymbol{x},\boldsymbol{y}}\eta_k)_{k\in\mathbb{N}}$. We show that this sequence is also convergent to unity. Consider a compact subset $\boldsymbol{K}$ of $\mathbb{R}^{2\,n}$, and let $R > 0$ be such that $\boldsymbol{K} \subseteq \overline{\mathcal{B}_R(0)}$. For a sufficiently large k (i.e., say, for k greater or equal to a certain k_0), the function η_k is equal to 1 in $\overline{\mathcal{B}_R(0)}$, and

$$\partial^{\alpha}_{\boldsymbol{x},\boldsymbol{y}}\eta_k(\boldsymbol{x},\boldsymbol{y}) = 0 \quad \forall\,(\boldsymbol{x},\boldsymbol{y}) \in \overline{\mathcal{B}_R(0)} \text{ and } k \geq k_0,$$

because this amounts to differentiate a constant function. Therefore, the first condition in Definition 5.32 is satisfied. For the second condition on the uniform boundedness, we have to prove that for every $\gamma \in \mathbb{N}^{2\,n}$, there must exist $c(\gamma) > 0$ such that for every $k \in \mathbb{N}$ and for every $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{2\,n}$ there holds $|\partial^{\gamma}(\eta_k(\boldsymbol{x},\boldsymbol{y}) + \partial^{\alpha}\,\eta_k(\boldsymbol{x},\boldsymbol{y})| \leq c(\gamma)$. This inequality can be rewritten as $|\partial^{\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y}) + \partial^{\alpha+\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})| \leq c(\gamma)$. We have $|\partial^{\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y}) + \partial^{\alpha+\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})| \leq |\partial^{\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})| + |\partial^{\alpha+\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})|$, and the statement follows from the fact that $(\eta_k)_{k\in\mathbb{N}}$ is convergent to unity: there exists $c_1 > 0$ such that $|\partial^{\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})| \leq c_1$ and there exists $c_2 > 0$ such that $|\partial^{\alpha+\gamma}\eta_k(\boldsymbol{x},\boldsymbol{y})| \leq c_2$ for every $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{2\,n}$, so that it suffices to choose $c(\gamma) = c_1 + c_2$.
Since both $\eta_k)_{k\in\mathbb{N}}$ and $(\eta_k + \partial^{\alpha}_{\boldsymbol{x},\boldsymbol{y}}\eta_k)_{k\in\mathbb{N}}$ are sequences of $\mathcal{D}(\mathbb{R}^{2\,n},\mathbb{R})$ converging to unity, we obtain

$$\langle T * S, \varphi\rangle = \lim_{k\to\infty}\langle T\otimes S, (\varphi\circ s)\,\eta_k\rangle = \lim_{k\to\infty}\langle T\otimes S, (\varphi\circ s)(\eta_k + \partial^{\alpha}_{\boldsymbol{x},\boldsymbol{y}}\eta_k)\rangle,$$

and the result follows by linearity.

Theorem 5.46 *Let $T, S \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ be convolvable. Let $j \in \{1,\dots,n\}$. Then, $\partial_j T$ and S are convolvable, and T and $\partial_j S$ are convolvable, and we have*

$$\partial_j(T * S) = \partial_j T * S = T * \partial_j S.$$

Proof Consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ convergent to unity in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$. Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We have

$$\begin{aligned}\langle T\otimes\partial_j S, \eta_k\,(\varphi\circ s)\rangle &= \Big\langle T(\boldsymbol{x}), \langle \partial_j S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle \\ &= -\left\langle T(\boldsymbol{x}), \left\langle S(\boldsymbol{y}), \frac{\partial\eta_k}{\partial y_j}(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\right\rangle\right\rangle - \Big\langle T(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\partial_j\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle.\end{aligned}$$

Now we take the limit:

$$\begin{aligned}\lim_{k\to\infty}\langle T\otimes\partial_j S, \eta_k\,(\varphi\circ s)\rangle &= \lim_{k\to\infty}\left[-\left\langle T\otimes S, \frac{\partial\eta_k}{\partial y_j}\,(\varphi\circ s)\right\rangle - \Big\langle T\otimes S, \eta_k\,(\partial_j\varphi)\circ s\Big\rangle\right] \\ &= 0 - \langle T * S, \partial_j\varphi\rangle = \langle \partial_j(T*S),\varphi\rangle,\end{aligned}$$

where we have applied Lemma 5.15. Thus, T and $\partial_j S$ are convolvable and $T * \partial_j S = \partial_j(T*S)$. The identity $\partial_j T * S = \partial_j(T*S)$ follows from the commutative property.

This result can iterated. In general, given $\alpha\in\mathbb{N}^n$, and $T,S\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ convolvable, we have $\partial^\alpha_{\boldsymbol{x}} T$ and S are convolvable, and

$$\partial^\alpha_{\boldsymbol{x}}(T*S) = \partial^\alpha_{\boldsymbol{x}} T * S = T * \partial^\alpha_{\boldsymbol{x}} S.$$

We can also generalize even further, by writing

$$\partial^{\alpha+\beta}_{\boldsymbol{x}}(T*S) = \partial^\alpha_{\boldsymbol{x}} T * \partial^\beta_{\boldsymbol{x}} S.$$

We also focus on "negative" results about the convolution. For example if $T,S\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ and $(S_k)_{k\in\mathbb{N}}$ is a sequence of distributions in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$, if

- S_k is convolvable with T for any $k\in\mathbb{N}$ and S is convolvable with T
- $S_k \xrightarrow{k\to\infty} S$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$

it is not true that

$$T * S_k \xrightarrow{k\to\infty} T * S.$$

Example 5.53

Let $T\in\mathcal{D}'(\mathbb{R};\mathbb{R})$ be the constant function 1. Let $S_k(x) = \delta(x-k)$ for $k\in\mathbb{N}$. This distribution has compact support (the only point k), so the convolution is well defined. We know that

$$\langle \delta(x-k),\varphi(x)\rangle = \varphi(k) \qquad \forall\,\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R}).$$

We now want to determine $T * S_k$. Let $(\eta_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$ that converges to unity. Let $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$. Since S_k has compact support, we can introduce a function $\zeta\in\mathcal{D}(\mathbb{R};\mathbb{R})$ which is equal to 1 on an open set containing the support of S_k. In other words, $\zeta(y) = 1$ in a neighbourhood of k. We find

$$\begin{aligned}\langle T*S,\varphi\rangle &= \langle T(x), \langle S_k(y), \zeta(y)\,\varphi(x+y)\rangle\rangle \\ &= \langle 1(x), \langle \delta(y-k), \zeta(y)\,\varphi(x+y)\rangle\rangle \\ &= \langle 1(x), \langle \zeta(k)\,\varphi(x+k)\rangle\rangle \\ &= \langle 1(x), \varphi(x+k)\rangle = \int_{\mathbb{R}}\varphi(x+k)\,\mathrm{d}x = \int_{\mathbb{R}}\varphi(x)\,\mathrm{d}x = \langle 1,\varphi\rangle.\end{aligned}$$

Thus, $T * S_k = T = 1$ for every $k\in\mathbb{N}$. We now study the behaviour of S_k as $k\to\infty$. Let $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$:

$$\langle S_k,\varphi\rangle = \langle \delta(x-k),\varphi(x)\rangle = \varphi(k) \xrightarrow{k\to\infty} 0.$$

It follows that

$$\lim_{k\to\infty} S_k(x) = 0 \quad \text{in } \mathcal{D}'(\mathbb{R};\mathbb{R}).$$

We conclude that S is the zero distribution, and therefore $T * S = 0$. We have found that

$$0 = T * S \neq \lim_{k\to\infty} T * S_k = 1.$$

Remark 5.11 In general, the associative property does not hold for the convolution. In other words, given $T, S, R \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$ such that $(T * S) * R$ and $T * (S * R)$ are defined, we have

$$(T * S) * R \neq T * (S * R).$$

Consider for example $n = 1$, and let T to be the Heaviside function (i.e., $T = H$), $S = \delta'$ and $R = 1$. Then

$$T * S = H * \delta' = (H * \delta)' = H' = \delta,$$

where we notice that $H * \delta'$ makes sense because δ' has compact support. Thus,

$$(T * S) * R = \delta * R = R = 1.$$

However, $S * R = \delta' * 1 = (\delta * 1)' = 1' = 0$, and therefore

$$T * (S * R) = T * 0 = 0.$$

Exercises

Exercise 5.68 Let $(\eta_k)_{k\in\mathbb{N}}$ be convergent to unity in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Show that, for all $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$, the sequence $(\eta_k \varphi)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. [Solution: (5.68)]

Exercise 5.69 Let $T_1, T_2, S \in \mathcal{D}'(\mathbb{R}^n;\mathbb{R})$, and let $\lambda, \mu \in \mathbb{R}$. Show that, if $T_1 * S$ and $T_2 * S$ exist, then $(\lambda T_1 + \mu T_2) * S$ exists and

$$(\lambda T_1 + \mu T_2) * S = \lambda (T_1 * S) + \mu (T_2 * S).$$

[Solution: (5.69)]

Exercise 5.70 Let $(\eta_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity. For all $k \in \mathbb{N}$ and for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^n$, let

$$\tilde{\eta}_k(\boldsymbol{x}, \boldsymbol{y}) = \eta_k(\boldsymbol{y}, \boldsymbol{x}).$$

Show that $(\eta_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{D}(\mathbb{R}^{2n};\mathbb{R})$ converging to unity. [Solution: (5.70)]

6 Fourier transform

In this chapter, we introduce the Fourier transform and we present its most important properties, first in the context of integrable functions, and later on in the context of tempered distributions. We will be able to treat large classes of functions, and in particular $L^1(\mathbb{R}^n, \mathbb{C})$. We will also see how to apply the Fourier transform to study the solutions of equations arising in mathematical physics.

6.1 Fourier transform in $L^1(\mathbb{R}^n, \mathbb{C})$

We recall that $L^1(\mathbb{R}^n, \mathbb{C})$ is the vector space of equivalence classes of integrable functions $f : \mathbb{R}^n \longrightarrow \mathbb{C}$. Moreover, $L^1(\mathbb{R}^n, \mathbb{C})$ is a Banach space with the norm

$$\| [f] \|_1 = \int_{\mathbb{R}^n} |f(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{x}.$$

For the sake of simplicity, given $f \in L^1(\mathbb{R}^n, \mathbb{C})$, we shall write f instead of $[f]$, and we shall think of the elements of $L^1(\mathbb{R}^n, \mathbb{C})$ as summable functions, identifying the functions which coincide almost everywhere.

We anticipate that the formal definition of Fourier transform reads as

$$\boxed{\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \, f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}} \qquad \text{where } \boldsymbol{\xi} \in \mathbb{R}^n. \tag{6.1}$$

The first problem is to find under which conditions on f the new function $\hat{f}$ is well defined. Interpreting the integral in (6.1) as a Lebesgue integral, we need to impose that the function be summable, and since, as we will see, $\left| e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \, f(\boldsymbol{x}) \right| = |f(\boldsymbol{x})|$, because $\boldsymbol{\xi}$ and $\boldsymbol{x}$ are real, the natural assumption is $f \in L^1(\mathbb{R}^n; \mathbb{C})$.

Definition 6.1 [Fourier Transform]
Let $n \in \mathbb{N}^*$. Let $f \in L^1(\mathbb{R}^n, \mathbb{C})$. The function

$$\begin{aligned} \hat{f} : \mathbb{R}^n &\longrightarrow \mathbb{C} \\ \boldsymbol{\xi} &\mapsto \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \, f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \end{aligned}$$

is called Fourier transform of f. *We also denote the function $\hat{f}$ with the symbol $\mathcal{F}\{f\}$.*

The exponent $\boldsymbol{x} \cdot \boldsymbol{\xi}$ represents the standard dot product in $\mathbb{R}^n$, so that, denoting by $x_1, \ldots, x_n$ the components of x and by $\xi_1, \ldots, \xi_n$ the components of $\boldsymbol{\xi}$, we have

$$\boldsymbol{x} \cdot \boldsymbol{\xi} = x_1 \, \xi_1 + x_2 \, \xi_2 + \ldots + x_n \, \xi_n.$$

The integral in Definition 6.1 exists for every $\boldsymbol{\xi} \in \mathbb{R}^n$ because $i\,\boldsymbol{x} \cdot \boldsymbol{\xi}$ is a purely imaginary number, and therefore the function $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})$ is summable.

In particular, from Definition 6.1 we see immediately that

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \hat{f}(0).$$

Notice that in the literature there are alternative but entirely equivalent definitions of Fourier transform, which are obtained by changing the *Fourier kernel* $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$ with $e^{-2\pi i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$ or $(2\pi)^{-n/2}\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$. These choices do not alter the essence, and they are motivated by the desire to simplify the constants that appear in the various formulae, ideally to reduce them to 1. However, this simplification is always partial, as it cannot be achieved for all formulae.

Example *6.1*

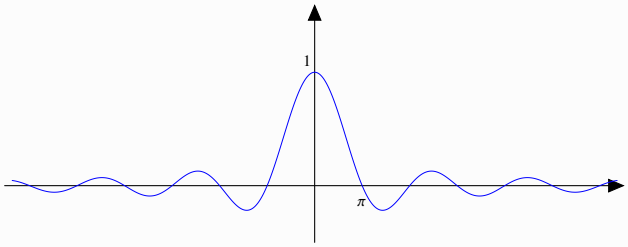
We want to determine the Fourier transform of the characteristic function of the closed interval $[-1, 1]$

$$\begin{aligned} \chi_{[-1,1]} : \mathbb{R} &\longrightarrow \mathbb{C} \\ x &\mapsto \begin{cases} 1 & \text{if } x \in [-1,1] \\ 0 & \text{if } x \notin [-1,1]. \end{cases} \end{aligned}$$

Let $\xi \in \mathbb{R}$. First, observe that $\chi_{[-1,1]}$ is a real even function. Thus,

$$\mathcal{F}\{\chi_{[-1,1]}\}(\xi) = \int_{-1}^{1} e^{-i\,x\,\xi}\,\mathrm{d}x = \begin{cases} 2 & \text{if } \xi = 0 \\ \left[\frac{e^{-i\,x\,\xi}}{-i\,\xi}\right]_{-1}^{1} = 2\,\frac{e^{i\,x\,\xi} - e^{-i\,x\,\xi}}{2\,i\,\xi} = 2\,\frac{\sin\xi}{\xi} & \text{if } \xi \neq 0. \end{cases}$$

This function is also called *sinc* (which stands for sine cardinal)



We observe in particular that the Fourier transform is:

- bounded
- continuous
- infinitesimal at infinity.

These properties are true for all integrable functions (i.e., for functions in $\mathcal{L}^1(\mathbb{R}^n; \mathbb{C})$). However, we also observe that the Fourier transform is not in $\mathcal{L}^1(\mathbb{R}^n; \mathbb{C})$. Indeed, if $f(x) = \chi_{[-1,1]}(x)$, we have

$$\int_0^{+\infty} |\hat{f}(\xi)|\,\mathrm{d}\xi = +\infty.$$

Hence, the properties of f are not mapped into corresponding properties of $\hat{f}$.

Now we study the fundamental properties of the Fourier transform $\hat{f}$ in terms of the properties of f. We will see that some formulae will be valid under some cumbersome assumptions. This is one of the aspects that will push us towards extending the Fourier transform for distributions: in this setting everything works without assumptions.

The first set of properties is given in the following theorem: it shows that the Fourier transform is bounded, uniformly continuous and it vanishes at ∞: this last property is known as *Riemann-Lebesgue Lemma*, and, unlike the other properties, its proof requires more care.

Theorem 6.1 *Let $n \in \mathbb{N}^*$. Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. Then:*

(i) $\hat{f}$ *is bounded;*

(ii) $\hat{f}$ *is uniformly continuous;*

(iii) $\hat{f} \overset{\|\xi\| \to \infty}{\longrightarrow} 0$ (Riemann-Lebesgue Lemma).

Proof First, we show that $\hat{f}$ is bounded. In particular, we prove that $\hat{f}$ is bounded by the 1-norm. Indeed, for every $\boldsymbol{\xi} \in \mathbb{R}^n$ we have

$$|\hat{f}(\boldsymbol{\xi})| = \left| \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \right| \leq \int_{\mathbb{R}^n} \left| e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x}) \right| \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = \|f\|_1. \tag{6.2}$$

Now we show that $\hat{f}$ is continuous. Let $(\boldsymbol{\xi}_k)_{k\in\mathbb{N}}$ be a sequence of vectors in $\mathbb{R}^n$ that converges to $\boldsymbol{\xi}^0 \in \mathbb{R}^n$. We have to show that $\hat{f}(\boldsymbol{\xi}_k) \xrightarrow{k\to\infty} \hat{f}(\boldsymbol{\xi}^0)$, i.e., that

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}_k}\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}^0}\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

Since there is convergence of the integrands almost everywhere, it is sufficient to find a dominating integrable function independent from k and apply Lebesgue's dominated convergence theorem (Theorem 3.47). For all $k \in \mathbb{N}$ we have

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \qquad \left| e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}_k}\, f(\boldsymbol{x}) \right| = |f(\boldsymbol{x})|,$$

which means that we can take $|\boldsymbol{f}|$ as the dominating function. Now, for all $\boldsymbol{x} \in \mathbb{R}^n$ we can apply Lebesgue's dominated convergence theorem (Theorem 3.47) to conclude that

$$\lim_{k\to\infty} \hat{f}(\boldsymbol{\xi}_k) = \lim_{k\to\infty} \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}_k}\, f(\boldsymbol{x})\, \mathrm{d}x = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}^0}\, f(\boldsymbol{x})\, \mathrm{d}x = \hat{f}(\boldsymbol{\xi}^0).$$

Thus, $\hat{u}$ is continuous. However, uniform continuity of $\hat{f}$ is a stronger property, which by definition requires that

$$\forall\, \varepsilon > 0, \quad \exists\, \delta > 0 : \quad \forall\, \boldsymbol{\xi}, \boldsymbol{h} \in \mathbb{R}^n, \qquad |\boldsymbol{h}| < \delta \;\Rightarrow\; |\hat{f}(\boldsymbol{\xi}+\boldsymbol{h}) - \hat{f}(\boldsymbol{\xi})| \leq \varepsilon.$$

We have

$$\begin{aligned} |\hat{f}(\boldsymbol{\xi}+\boldsymbol{h}) - \hat{f}(\boldsymbol{\xi})| &= \left| \int_{\mathbb{R}^n} \left(e^{-i\,\boldsymbol{x}\cdot(\boldsymbol{\xi}+\boldsymbol{h})} - e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\right) f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \right| \\ &\leq \int_{\mathbb{R}^n} \left| e^{-i\,\boldsymbol{x}\cdot\boldsymbol{h}} - 1 \right| |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x}. \end{aligned}$$

In this expression $\boldsymbol{\xi}$ has disappeared. Thus, we can show the uniform continuity by showing that this expression vanishes as $\boldsymbol{h} \to 0$, i.e., we must show that

$$\lim_{\boldsymbol{h}\to 0} \int_{\mathbb{R}^n} \left| e^{-i\,\boldsymbol{x}\cdot\boldsymbol{h}} - 1 \right| |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} = 0.$$

This follows directly from Lebesgue's dominated convergence theorem (Theorem 3.47):

$$\left|e^{-i\,\boldsymbol{x}\cdot\boldsymbol{h}} - 1\right| |f(\boldsymbol{x})| \leq \left(\left|e^{-i\,\boldsymbol{x}\cdot\boldsymbol{h}}\right| + 1\right) |f(\boldsymbol{x})| = 2\,|f(\boldsymbol{x})|,$$

and $2\,|f(\boldsymbol{x})|$ is integrable over $\mathbb{R}^n$ by assumption, so that we can bring the limit inside the integral, thus proving the statement.

We prove the last point. We begin considering $f \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$. Multiplying both sides of the Fourier transform by $1 + \|\boldsymbol{\xi}\|^2$ we find

$$(1 + \|\boldsymbol{\xi}\|^2)\,|\hat{f}(\boldsymbol{\xi})| = \left|\int_{\mathbb{R}^n} (1 + \|\boldsymbol{\xi}\|^2)\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|.$$

We have

$$\begin{aligned}
\frac{\partial}{\partial x_j} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} &= \frac{\partial}{\partial x_j} e^{-i\,(\xi_1\,x_1 + \xi_2\,x_2 + \ldots + \xi_n\,x_n)} = -i\,\xi_j\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \\
\frac{\partial^2}{\partial x_j^2} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} &= \frac{\partial}{\partial x_j}\left(-i\,\xi_j\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\right) = -\xi_j^2\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\nabla_{\boldsymbol{x}}^2 e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} &= \sum_{j=1}^{n} \left(\frac{\partial^2}{\partial x_j^2} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\right) \\
&= -(\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2)\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} = -\|\boldsymbol{\xi}\|^2\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}.
\end{aligned}$$

Subtracting $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$ and changing sign on both sides yields

$$(1 + \|\boldsymbol{\xi}\|^2)\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} = (1 - \nabla_{\boldsymbol{x}}^2)\,(e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}).$$

It follows that

$$\left|\int_{\mathbb{R}^n} (1 + \|\boldsymbol{\xi}\|^2)\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| = \left|\int_{\mathbb{R}^n} (1 - \nabla_{\boldsymbol{x}}^2)\,(e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}})\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right|,$$

and by viewing the derivative as the derivative in the sense of distributions (which is allowed since f is assumed to live in $\mathcal{D}(\mathbb{R}^n; \mathbb{C})$) we obtain

$$\begin{aligned}
(1 + \|\boldsymbol{\xi}\|^2)\,|\hat{f}(\boldsymbol{\xi})| &= \left|\int_{\mathbb{R}^n} (1 - \nabla_{\boldsymbol{x}}^2)\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&= \left|\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,(1 - \nabla_{\boldsymbol{x}}^2)\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \\
&= \left|\mathcal{F}\left\{(1 - \nabla_{\boldsymbol{x}}^2)\, f\right\}(\boldsymbol{\xi})\right|,
\end{aligned}$$

which is bounded by $\|(1 - \nabla_{\boldsymbol{x}}^2)\, f\|_1$. Thus,

$$(1 + \|\boldsymbol{\xi}\|^2)\,|\hat{f}(\boldsymbol{\xi})| \leq \|(1 - \nabla_{\boldsymbol{x}}^2)\, f\|_1 \quad \Rightarrow \quad |\hat{f}(\boldsymbol{\xi})| \leq \frac{\|(1 - \nabla_{\boldsymbol{x}}^2)\, f\|_1}{(1 + \|\boldsymbol{\xi}\|^2)} \xrightarrow{\|\boldsymbol{\xi}\| \to \infty} 0,$$

as required. We now generalise, by using the density of functions in $\mathcal{D}(\mathbb{R}^n; \mathbb{C})$ in $\mathcal{L}^1(\mathbb{R}^n; \mathbb{C})$. Let $f \in \mathcal{L}^1(\mathbb{R}^n; \mathbb{C})$. Then, there exists a sequence $(f_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{D}(\mathbb{R}^n; \mathbb{C})$ such that

$$\|f - f_k\|_1 \xrightarrow{k \to \infty} 0.$$

On the other hand, for every $\boldsymbol{\xi} \in \mathbb{R}^n$, we have

$$\begin{aligned} |\hat{f}(\boldsymbol{\xi})| &\leq \left|\hat{f}(\boldsymbol{\xi}) - \hat{f}_k(\boldsymbol{\xi})\right| + \left|\hat{f}_k(\boldsymbol{\xi})\right| \\ &\leq \|f - f_k\|_1 + \left|\hat{f}_k(\boldsymbol{\xi})\right|, \end{aligned}$$

where $\|f - f_k\|_1 \xrightarrow{k \to \infty}$ by assumption. Thus, let $\varepsilon > 0$. Let k be sufficiently large so that $\|f - f_k\|_1 \leq \frac{\varepsilon}{2}$, and we can write that for all $\boldsymbol{\xi} \in \mathbb{R}^n$ we have $|\hat{f}(\boldsymbol{\xi})| \leq \frac{\varepsilon}{2} + |\hat{f}_k(\boldsymbol{\xi})|$. However, we know that $f_k \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$, and from what we proved before its Fourier transform $\hat{f}_k$ vanishes, which means that there exists $R_\varepsilon > 0$ such that for all $\boldsymbol{\xi} \in \mathbb{R}^n$ such that $\|\boldsymbol{\xi}\| > R_\varepsilon$ we have $|\hat{f}_k(\boldsymbol{\xi})| \leq \frac{\varepsilon}{2}$. Hence, for such vectors $\boldsymbol{\xi}$ we have $|\hat{f}(\boldsymbol{\xi})| \leq \varepsilon$, which is enough to conclude.

Remark 6.1 To prove

$$\lim_{\|\boldsymbol{\xi}\| \to \infty} \hat{f}(\boldsymbol{\xi}) = \lim_{\|\boldsymbol{\xi}\| \to \infty} \int_{\mathbb{R}^n} e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 0,$$

one is tempted to use Lebesgue's dominated convergence theorem (Theorem 3.47) to bring the limit inside the integral. However, it is not true that the integrand is pointwise convergent to the zero function: indeed, the exponential has purely imaginary exponent, and therefore as $\|\boldsymbol{\xi}\|$ increases the term $e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}}$ represents a complex number that moves around the unit circle. The exponential does not even converge, in general: the only value for which we have convergence is $\boldsymbol{x} = \mathbf{0}$.

The first property of Theorem 6.1 can be re-written as

$$\|\hat{f}\|_\infty \leq \|f\|_1.$$

This says, in particular, that the mapping $f \mapsto \hat{f}$ is linear and continuous from $L^1(\mathbb{R}^n; \mathbb{C})$ to $L^\infty(\mathbb{R}^n; \mathbb{C})$. In other words, if $f_k \xrightarrow{k \to \infty} f$ in $L^1(\mathbb{R}^n; \mathbb{C})$, then $\hat{f}_k \xrightarrow{k \to \infty} \hat{f}$ uniformly. This statement will be refined later on.

As a result of this discussion, we have

$$\mathcal{F}(\mathcal{L}^1(\mathbb{R}^n; \mathbb{C})) \subset \mathcal{L}^\infty(\mathbb{R}^n; \mathbb{C}).$$

Corollary 6.1 *Let $n \in \mathbb{N}^*$. Let $f \in L^1(\mathbb{R}^n; \mathbb{C})$. Then,*

$$\lim_{\|\boldsymbol{\xi}\| \to \infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \lim_{\|\boldsymbol{\xi}\| \to \infty} \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = 0.$$

Proof If $f(\boldsymbol{x})$ is real almost everywhere, we have $\cos(\boldsymbol{x}\cdot\boldsymbol{\xi}) = \Re\mathrm{e}(e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}})$ and $\sin(\boldsymbol{x}\cdot\boldsymbol{\xi}) = \Im\mathrm{m}(e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}})$, so that

$$\int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = \Re\mathrm{e}\, \hat{f}(\boldsymbol{\xi}) \quad \text{and} \quad \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} = -\Im\mathrm{m}\, \hat{f}(\boldsymbol{\xi}).$$

The statement then follows from Riemann-Lebesgue's Lemma. If $f(\boldsymbol{x})$ is not real-valued, the argument applies to $\Re\mathrm{e}\, f$ and $\Im\mathrm{m}\, f$, which are integrable and real-valued.

Given a set $\boldsymbol{A} \subseteq \mathbb{R}^n$, we denote by $BC(\boldsymbol{A}; \mathbb{C})$ the vector space of functions $f : \boldsymbol{A} \longrightarrow \mathbb{C}$ which are

continuous and bounded. In this vector space we have the norm $\|\cdot\|_\infty$, defined by

$$\|f\|_\infty = \sup\{|f(w)| \in \mathbb{R}_+ \,|\, w \in \boldsymbol{A}\}.$$

Arguing as in the case of $BC(\boldsymbol{A};\mathbb{R})$, the vector space $BC(\boldsymbol{A};\mathbb{C})$ equipped with this norm is a Banach space. We present another corollary of the Riemann-Lebesgue Lemma.

Corollary 6.2 *Let*

$$\begin{aligned} \mathcal{F} : L^1(\mathbb{R}^n,\mathbb{C}) &\longrightarrow BC(\mathbb{R}^n;\mathbb{C}) \\ f &\mapsto \mathcal{F}\{f\} \end{aligned}$$

Then, $\mathcal{F}$ is linear and continuous from $L^1(\mathbb{R}^n,\mathbb{C})$ to $BC(\mathbb{R}^n;\mathbb{C})$.

Proof The proof of the linearity is left as an exercise. We show that $\mathcal{F}$ is continuous. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^1(\mathbb{R}^n,\mathbb{C})$ such that $\lim_{k\to\infty} f_k = f$ in $L^1(\mathbb{R}^n,\mathbb{C})$. We have to show that $\lim_{k\to\infty} \mathcal{F}f_k = \mathcal{F}f$ in $BC(\mathbb{R}^n;\mathbb{C})$. Since as already observed in the proof of the Riemann-Lebesgue Lemma there holds

$$|\hat{f}(\boldsymbol{\xi})| \le \int_{\mathbb{R}^n} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \|f\|_1,$$

it follows that for every $g \in L^1(\mathbb{R}^n,\mathbb{C})$ we have $\|\mathcal{F}g\|_\infty \le \|g\|_1$. This implies that

$$\|\mathcal{F}f - \mathcal{F}f_k\|_\infty = \|\mathcal{F}(f-f_k)\|_\infty \le \|f-f_k\| \xrightarrow{k\to\infty} 0.$$

6.2 Properties of the Fourier transform

We now present some important properties of the Fourier transform.

Theorem 6.2 *Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$, $\boldsymbol{y} \in \mathbb{R}^n$ and $A \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$ invertible (i.e., A is an automorphism). Then:*

(P1) $g(\boldsymbol{x}) = f(\boldsymbol{x}-\boldsymbol{y}) \quad \Rightarrow \quad \hat{g}(\boldsymbol{\xi}) = e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi})$;

(P2) $g(\boldsymbol{x}) = e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,f(\boldsymbol{x}) \quad \Rightarrow \quad \hat{g}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}-\boldsymbol{y})$;

(P3) $g(\boldsymbol{x}) = f(A\,\boldsymbol{x}) \quad \Rightarrow \quad \hat{g}(\boldsymbol{\xi}) = \frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi})$;

(P4) $g = \bar{f} \quad \Rightarrow \quad \hat{g}(\boldsymbol{\xi}) = \overline{\hat{f}(-\boldsymbol{\xi})}$.

Proof Recall that we can express the change of variable theorem as

$$\int_A f(S(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y} = \int_{S(A)} f(\boldsymbol{x})\,|\det J_S^{-1}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}, \tag{6.3}$$

where S is a change of variable, i.e., it is injective, of class C^1, and its Jacobian matrix $J_S(\boldsymbol{x})$ is invertible for

every $\boldsymbol{x}$. We prove **(P1)**. We have

$$\hat{g}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} e^{-i\,(z+y)\cdot\boldsymbol{\xi}} f(z)\,\mathrm{d}z = e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi}),$$

where we used the change of variable $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{y}$ (so that $\boldsymbol{z} = S(\boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{y}$ and therefore $J_S(\boldsymbol{x})$ is the identity matrix for every $\boldsymbol{x} \in \mathbb{R}^n$). Now we prove **(P2)**:

$$\hat{g}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, e^{i\,\boldsymbol{x}\cdot\boldsymbol{y}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot(\boldsymbol{\xi}-\boldsymbol{y})} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \hat{f}(\boldsymbol{\xi}-\boldsymbol{y}).$$

Consider **(P3)**:

$$\hat{g}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f(A\,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} e^{-i\,A^{-1}z\cdot\boldsymbol{\xi}} f(z)\,\frac{1}{|\det A|}\,\mathrm{d}z = \frac{1}{|\det A|}\int_{\mathbb{R}^n} e^{-i\,z\cdot A^{-\top}\boldsymbol{\xi}} f(z)\,\mathrm{d}z = \frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}),$$

where we used the change of variable $\boldsymbol{z} = A\,\boldsymbol{x}$ (so that $S(\boldsymbol{x}) = A\,\boldsymbol{x}$ and $J_S(\boldsymbol{x}) = A$ for every $\boldsymbol{x} \in \mathbb{R}^n$). Finally, we prove **(P4)**:

$$\hat{g}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\overline{f(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} \overline{e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \overline{\int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}} = \overline{\hat{f}(-\boldsymbol{\xi})}.$$

Remark 6.2 It is very common to see Properties **(P1)** written in the following way:

$$\widehat{f(\boldsymbol{x}-\boldsymbol{y})}(\boldsymbol{\xi}) = e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi}) \qquad \text{or} \qquad \mathcal{F}\{f(\boldsymbol{x}-\boldsymbol{y})\}(\boldsymbol{\xi}) = e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi}). \tag{6.4}$$

When we use the notation (6.4), we are implicitly assuming that $f(\boldsymbol{x}-\boldsymbol{y})$ is the name that we are giving to the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto f(\boldsymbol{x}-\boldsymbol{y}). \end{aligned}$$

In some instances, it is possible to avoid this abuse by denoting the independent variable with a dot, e.g., in the case of (6.4) we would have

$$\mathcal{F}\{f(\cdot-\boldsymbol{y})\}(\boldsymbol{\xi}) = e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi}).$$

This last notation is not always clear, especially when the function is more complex. For example, using this notation on **(P2)** would lead to $\mathcal{F}\left\{e^{i\,\cdot\cdot\,\boldsymbol{y}}\,f(\cdot)\right\}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}-\boldsymbol{y})$, which is notationally cumbersome.

Using the conventions in (6.4), the properties **(P2)-(P4)** can be written as

$$\begin{aligned} \widehat{e^{i\,\boldsymbol{x}\cdot\boldsymbol{y}}\,f(\boldsymbol{x})}(\boldsymbol{\xi}) &= \hat{f}(\boldsymbol{\xi}-\boldsymbol{y}) && \text{or} && \mathcal{F}\left\{e^{i\,\boldsymbol{x}\cdot\boldsymbol{y}}\,f(\boldsymbol{x})\right\}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}-\boldsymbol{y}) \\ \widehat{f(A\,\boldsymbol{x})}(\boldsymbol{\xi}) &= \tfrac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) && \text{or} && \mathcal{F}\{f(A\,\boldsymbol{x})\}(\boldsymbol{\xi}) = \tfrac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) \\ \widehat{\bar{f}(\boldsymbol{x})}(\boldsymbol{\xi}) &= \overline{\hat{f}(-\boldsymbol{\xi})} && \text{or} && \mathcal{F}\left\{\bar{f}(\boldsymbol{x})\right\}(\boldsymbol{\xi}) = \overline{\hat{f}(-\boldsymbol{\xi})}. \end{aligned}$$

Before we proceed, we notice that the statements of Theorem 6.2 lead us to ask ourselves the following question: if a certain operation allows us to go from a function f to another function g, what kind of operation on the Fourier transforms allows us to go directly from $\hat{f}$ to $\hat{g}$?

Loosely speaking, properties **(P1)** and **(P2)** say that shifts are mapped into products by exponentials and vice-versa. **(P3)** says that a linear change of coordinate corresponds to another linear change of coordinates. Finally, **(P4)** expresses the fact that the complex conjugation corresponds to itself.

As a particular case of **(P3)**, we have a property for a so-called homothety:

(P5) $g(\boldsymbol{x}) = f(\lambda\,\boldsymbol{x}) \qquad \Rightarrow \qquad \hat{g}(\boldsymbol{\xi}) = \frac{1}{|\lambda|^n}\,\hat{f}\left(\frac{\boldsymbol{\xi}}{\lambda}\right),$

see Exercise 6.3. This says that a change of scale corresponds to another change of scale: one contracts, the other dilates.

Corollary 6.3 *Let $n \in \mathbb{N}^*$. Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. Then:*

(Q1) *f is even* $\Rightarrow$ *$\hat{f}$ is even*

(Q2) *f is odd* $\Rightarrow$ *$\hat{f}$ is odd*

(Q3) *f is radial* $\Rightarrow$ *$\hat{f}$ is radial*

Proof If f is even, then $f(-\boldsymbol{x}) = f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Defining $A \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$ as $A\,\boldsymbol{x} = -\boldsymbol{x}$, and the identity $f(-\boldsymbol{x}) = f(\boldsymbol{x})$ can be rewritten as $f \circ A = f$. Taking the Fourier transform of both sides gives $\widehat{f \circ A} = \hat{f}$. We can use property **(P3)** in Theorem 6.2, and we find

$$\frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}), \quad \forall\,\boldsymbol{\xi} \in \mathbb{R}^n.$$

On the other hand, $A\,\boldsymbol{x} = -\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$, which implies that A is a diagonal matrix with -1 on the diagonal. This means that A is orthogonal and symmetric, and therefore $A = A^{-1} = A^{\top}$ and $|\det A| = 1$. It follows that

$$\frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) = (\hat{f} \circ A)(\boldsymbol{\xi}) \quad \Rightarrow \quad \hat{f} \circ A = \hat{f},$$

which implies that $\hat{f}$ is even.

The second statement follows the same argument and is left as an exercise.

In Chapter 5, we saw that f is radial if $f \circ Q = f$ for any orthogonal matrix Q: the function in this case depends only on the Euclidean norm of the independent variable. We need to show that $\hat{f} \circ Q = \hat{f}$. Taking the Fourier transform of both sides gives $\widehat{f \circ Q} = \hat{f}$, and from property **(P3)** in Theorem 6.2 we obtain

$$\frac{1}{|\det Q|}\,\hat{f} \circ Q^{-\top} = \hat{f}.$$

Since Q is orthogonal, $Q^{-1} = Q^{\top}$. Moreover, $\det Q \in \{1, -1\}$. This is enough to conclude that $\hat{f} \circ Q = \hat{f}$.

If the domain of f is $\mathbb{R}$, and if the image of f is real we can express the Fourier transform in a simpler way.

Theorem 6.3 [Fourier Transform of a Real Function]

Let $f \in L^1(\mathbb{R}, \mathbb{C})$. Let $\xi \in \mathbb{R}$. Then,

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x - i \int_{-\infty}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x.$$

Moreover, if $f(x) \in \mathbb{R}$ for every $x \in \mathbb{R}$, we have

$$\mathfrak{Re}\{\hat{f}(\xi)\} = \int_{-\infty}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x \quad \textit{and} \quad \mathfrak{Im}\{\hat{f}(\xi)\} = -\int_{-\infty}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x.$$

Proof The result follows from Euler's formulae: since $x\,\xi \in \mathbb{R}$ we find immediately

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\,x\,\xi}\, f(x)\,\mathrm{d}x = \int_{\mathbb{R}} (\cos(-x\,\xi) + i\,\sin(-x\,\xi))\, f(x)\,\mathrm{d}x = \int_{\mathbb{R}} (\cos(x\,\xi) - i\,\sin(x\,\xi))\, f(x)\,\mathrm{d}x,$$

where we used the fact that the cosine is an even function and the sine is an odd function.

The expression in Theorem 6.3 can be further simplified in the case where f is an odd or an even function.

Theorem 6.4 [Fourier Transform of a Real Odd/Even Function]
Let $n \in \mathbb{N}^*$. Let $f \in L^1(\mathbb{R}^n, \mathbb{C})$. For almost every $x \in \mathbb{R}$, let $f(\boldsymbol{x}) \in \mathbb{R}$. Let $\xi \in \mathbb{R}$. Then:

- *if f is an even function, we have*

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x = 2 \int_{0}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x.$$

- *if f is an odd function, we have*

$$\hat{f}(\xi) = -i \int_{-\infty}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x = -2\,i \int_{0}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x.$$

Proof The statements follow immediately from Theorem 6.3, and by recalling that the product of two even functions is an even function, and so is the product of two odd functions (see Exercise 6.1). On the contrary, the product of an even function with an odd function is an odd function. The details are left as an exercise.

Example 6.2 We denote by f the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \frac{1}{1+x^2}, \end{aligned}$$

which is in $L^1(\mathbb{R};\mathbb{R})$, since $\displaystyle\int_{\mathbb{R}} \frac{1}{1+x^2}\,\mathrm{d}x = [\arctan x]_{-\infty}^{+\infty} = \pi$. We want to determine the Fourier transform of f. Let $\xi \in \mathbb{R}$. Clearly, $\frac{1}{1+x^2}$ is an even function, so that

$$\mathcal{F}\left\{\frac{1}{1+x^2}\right\}(\xi) = \int_{-\infty}^{+\infty} \frac{1}{1+x^2} \cos(x\,\xi)\,\mathrm{d}x.$$

Let $\xi \geq 0$. Consider the complex function

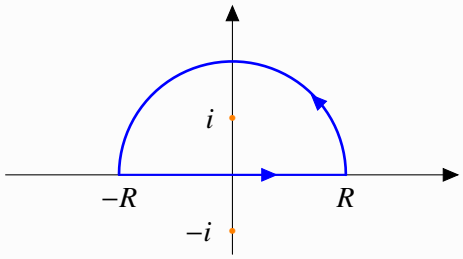
$$\begin{aligned} f : \mathbb{C} \setminus \{\pm i\} &\longrightarrow \mathbb{C} \\ z &\mapsto \frac{1}{1+z^2}\, e^{iz\xi}, \end{aligned}$$

and the complex integral

$$\oint_\Gamma \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z$$

over the contour Γ represented in the figure:

$$\Gamma = [-R, R] + C_R^+(0)$$



Here, we assume $R > 1$ so that the pole $z = i$ (the only pole with positive imaginary part) is encircled by the contour. From the Residue Theorem we have

$$\begin{aligned} \oint_\Gamma \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z &= 2\pi i\, \mathrm{Res}\left(\frac{1}{1+z^2}\, e^{iz\xi}; i\right) \\ &= 2\pi i \lim_{z\to i}(z-i)\, \frac{1}{1+z^2}\, e^{iz\xi} = 2\pi i \lim_{z\to i} \frac{e^{iz\xi}}{z+i} = 2\pi i\, \frac{e^{-\xi}}{2i} = \pi e^{-\xi}. \end{aligned}$$

On the other hand

$$\oint_\Gamma \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z = \int_{[-R,R]} \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z + \int_{C_R^+(0)} \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z.$$

Using the parameterizations

$$\begin{aligned} \varphi : [-R, R] &\longrightarrow \mathbb{C} \\ t &\mapsto t \end{aligned} \qquad \text{and} \qquad \begin{aligned} \psi : [0, \pi] &\longrightarrow \mathbb{C} \\ t &\mapsto R\, e^{it} \end{aligned}$$

for $[-R, R]$ and $C_R^+(0)$, respectively, we obtain

$$\oint_\Gamma \frac{1}{1+z^2}\, e^{iz\xi}\, \mathrm{d}z = \int_{-R}^{R} \frac{1}{1+t^2}\, e^{it\xi}\, \mathrm{d}t + \int_0^\pi \frac{e^{iR(\cos t + i\,\sin t)\xi}}{1+R^2\, e^{2it}}\, R\, i\, e^{it}\, \mathrm{d}t.$$

The first integral tends, as $R \to +\infty$, to $\int_{-\infty}^{\infty} \frac{1}{1+x^2}\, e^{ix\xi}\, \mathrm{d}x$. We need to show that the second integral vanishes. We have

$$\begin{aligned} \left| \int_0^\pi \frac{e^{iR(\cos t + i\,\sin t)\xi}}{1+R^2\, e^{2it}}\, R\, i\, e^{it}\, \mathrm{d}t \right| &\leq \int_0^\pi \left| \frac{e^{iR(\cos t + i\,\sin t)\xi}}{1+R^2\, e^{2it}}\, R\, i\, e^{it} \right| \mathrm{d}t \\ &= \int_0^\pi \frac{|e^{-R(\sin t)\xi}|}{|1+R^2\, e^{2it}|}\, |R\, i\, e^{it}|\, \mathrm{d}t \\ &\leq \int_0^\pi \frac{1}{R^2-1}\, R\, \mathrm{d}t = \frac{R\pi}{R^2-1} \xrightarrow{R\to+\infty} 0, \end{aligned}$$

where we used the fact that $e^{-R(\sin t)\xi} < 1$ (note that the exponent is non-positive when $t \in [0,\pi]$) and the reversed triangle inequality for the denominator (taking also into account that $R > 1$):

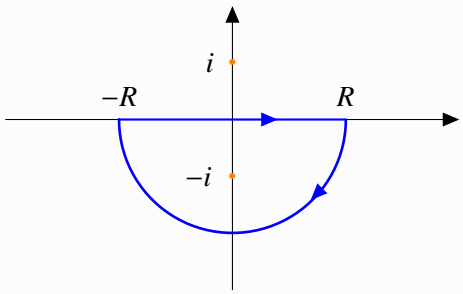
$$|1 + R^2\, e^{2\,i\,t}| = |R^2\, e^{2\,i\,t} - (-1)| \geq \left||R^2\, e^{2\,i\,t}| - |-1|\right| = |R^2 - 1| = R^2 - 1$$

Finally,

$$\int_{-\infty}^{+\infty} \frac{1}{1+x^2}\, e^{i\,x\,\xi}\, \mathrm{d}x = \Re\mathrm{e}\left\{\oint_\Gamma \frac{1}{1+z^2}\, e^{i\,z\,\xi}\, \mathrm{d}z\right\} = \Re\mathrm{e}\left\{\pi\, e^{-\xi}\right\} = \pi\, e^{-\xi}.$$

In the case $\xi < 0$ we proceed in a similar way, using the following contour

$$\Gamma = [-R, R] - C_R^-(0)$$



Finally, we obtain

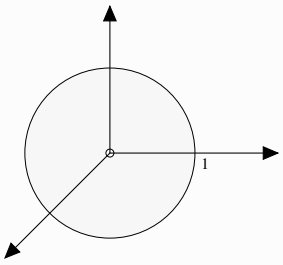
$$\mathcal{F}\left\{\frac{1}{1+x^2}\right\}(\xi) = \pi\, e^{-|\xi|}.$$

Example *6.3*

Let $\alpha \in \mathbb{R}$. Consider the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^\alpha & \text{if } 0 < \|\boldsymbol{x}\| \leq 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

In Example 5.46, we found that this function is integrable if $\alpha > -n = -3$. This, choosing $\alpha > -3$ ensures that $f \in L^1(\mathbb{R}^3;\mathbb{R})$. We want to find the Fourier transform of this function, exploiting the fact that it is radial since f is radial.



Indeed, this means that $\hat{f}(\boldsymbol{\xi})$ is equal to $\hat{f}$ computed at any point with Euclidean norm equal to $\|\boldsymbol{\xi}\|$, if we assume that $\boldsymbol{\xi} \neq \boldsymbol{0}$. For example, considering the vector $\boldsymbol{e}_3 = (0,0,1)$ of the standard basis of $\mathbb{R}^3$, we find

$$\hat{f}(\boldsymbol{\xi}) = \hat{f}(\|\boldsymbol{\xi}\|\,\boldsymbol{e}_3) = \int_{0<\|\boldsymbol{x}\|\leq 1} e^{-i\,x_3\,\|\boldsymbol{\xi}\|}\, \|\boldsymbol{x}\|^\alpha\, \mathrm{d}\boldsymbol{x}.$$

We use the spherical coordinates

$$\begin{cases} x_1 = \rho \sin\theta \cos\varphi \\ x_2 = \rho \sin\theta \sin\varphi \\ x_3 = \rho \cos\theta, \end{cases} \qquad \text{with } \rho \in (0,1], \quad \theta \in (0,\pi), \quad \varphi \in (0, 2\pi).$$

The determinant of the Jacobian is

$$\begin{aligned} \begin{vmatrix} \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\cos\theta\sin\varphi & \rho\sin\theta\cos\varphi \\ \cos\theta & -\rho\sin\theta & 0 \end{vmatrix} &= (\cos\theta) \begin{vmatrix} \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \rho\cos\theta\sin\varphi & \rho\sin\theta\cos\varphi \end{vmatrix} \\ &\quad + \rho\sin\theta \begin{vmatrix} \sin\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & \rho\sin\theta\cos\varphi \end{vmatrix} \\ &= \cos\theta(\rho^2 \sin\theta\cos\theta\cos^2\varphi + \rho^2\sin\theta\cos\theta\sin^2\varphi) \\ &\quad \rho\sin\theta(\rho\sin^2\theta\cos^2\varphi + \rho\sin^2\theta\sin^2\varphi) \\ &= \rho^2\sin\theta\cos^2\theta + \rho^2\sin^3\theta \\ &= \rho^2\sin\theta(\cos^2\theta + \sin^2\theta) = \rho^2\sin\theta. \end{aligned}$$

Using Fubini's theorem (Theorem 3.8), we find

$$\begin{aligned} \hat{f}(\boldsymbol{\xi}) &= \int_{(0,1]\times(0,\pi)\times(0,2\pi)} e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|} \rho^\alpha \rho^2 \sin\theta \,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= 2\pi \int_{(0,1]\times(0,\pi)} e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|} \rho^{\alpha+2} \sin\theta \,\mathrm{d}\rho\,\mathrm{d}\theta \\ &= 2\pi \int_{(0,1]} \rho^{\alpha+2} \left(\int_{(0,\pi)} e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|} \sin\theta \,\mathrm{d}\theta \right) \mathrm{d}\rho \\ &= 2\pi \int_{(0,1]} \rho^{\alpha+2} \left[\frac{e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|}}{i\rho\|\boldsymbol{\xi}\|} \right]_{\theta=0}^{\theta=\pi} \mathrm{d}\rho \\ &= 2\pi \int_{(0,1]} \rho^{\alpha+2} \frac{e^{i\rho\|\boldsymbol{\xi}\|} - e^{-i\rho\|\boldsymbol{\xi}\|}}{i\rho\|\boldsymbol{\xi}\|} \mathrm{d}\rho \\ &= 4\pi \int_{(0,1]} \rho^{\alpha+2} \frac{\sin(\rho\|\boldsymbol{\xi}\|)}{\rho\|\boldsymbol{\xi}\|} \mathrm{d}\rho \\ &= 4\pi \int_0^{\|\boldsymbol{\xi}\|} \|\boldsymbol{\xi}\|^{-\alpha-2} r^{2+\alpha} \frac{\sin r}{r} \|\boldsymbol{\xi}\|^{-1} \mathrm{d}r \qquad (r = \rho\|\boldsymbol{\xi}\|) \\ &= 4\pi \|\boldsymbol{\xi}\|^{-\alpha-3} \int_0^{\|\boldsymbol{\xi}\|} r^{1+\alpha} \sin r \,\mathrm{d}r. \end{aligned}$$

Example 6.4 Let $t \in \mathbb{R}_+$. Let $n \in \mathbb{N}^*$. Consider the fundamental solution of the heat equation

$$u_t(\boldsymbol{x}) = (4\pi t)^{-\frac{n}{2}} \exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right)$$

Applying the Fourier transform gives

$$\hat{u}_t(\boldsymbol{\xi}) = (4\pi t)^{-\frac{n}{2}} \int_{\mathbb{R}^n} \exp(-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}) \exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right) \mathrm{d}\boldsymbol{x}.$$

We can write

$$\exp(-i\,\boldsymbol{x}\cdot\boldsymbol{\xi})\,\exp\left(-\frac{\|\boldsymbol{x}\|^2}{4\,t}\right)=\exp\left(-i\,x_1\,\xi_1-\frac{x_1^2}{4\,t}\right)\cdot\ldots\cdot\exp\left(-i\,x_n\,\xi_n-\frac{x_n^2}{4\,t}\right)=\prod_{j=1}^{n}\exp\left(-i\,x_j\,\xi_j-\frac{x_j^2}{4\,t}\right),$$

and use this result to apply Tonelli's theorem (Theorem 3.54) by writing the integral as the product of n integrals in a single variable:

$$\begin{aligned}\hat{u}_t(\boldsymbol{\xi}) &= (4\,\pi\,t)^{-\frac{n}{2}}\prod_{j=1}^{n}\int_{\mathbb{R}}\exp\left(-i\,x_j\,\xi_j-\frac{x_j^2}{4\,t}\right)\mathrm{d}x_j\\ &= (4\,\pi\,t)^{-\frac{n}{2}}\prod_{j=1}^{n}\int_{\mathbb{R}}\exp\left(-i\,\sqrt{4\,t}\,\xi_j\,y-y^2\right)\sqrt{4\,t}\,\mathrm{d}y\qquad\left(y=\frac{x_j}{\sqrt{4\,t}}\right).\end{aligned}$$

Let $\eta=\sqrt{4\,t}\,\xi_j$. We show that

$$\int_{\mathbb{R}}\exp\left(-i\,\eta\,x-x^2\right)\mathrm{d}x=\sqrt{\pi}\,e^{-\frac{\eta^2}{4}},\tag{6.5}$$

which implies that

$$\hat{u}_t(\boldsymbol{\xi})=(4\,\pi\,t)^{-\frac{n}{2}}\prod_{j=1}^{n}\sqrt{4\,t}\,\sqrt{\pi}\,e^{-\frac{4\,t\,\xi_j^2}{4}}=e^{-t\,\|\boldsymbol{\xi}\|^2}.$$

Let us now prove (6.5). We start from the case $\eta=0$. In this case, we have to show that

$$\int_{\mathbb{R}}e^{-x^2}\,\mathrm{d}x=\sqrt{\pi}.\tag{6.6}$$

The integral in the left hand-side of (6.6) is the so-called Gaussian integral. Consider the integral

$$\int_{\mathbb{R}^2}e^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y,$$

and the polar coordinates $x=\rho\cos\theta$ and $y=\rho\sin\theta$, with $\rho\in\mathbb{R}_+$ and $\theta\in(0,2\,\pi)$. We obtain, applying Tonelli's theorem (Theorem 3.54)

$$\int_{\mathbb{R}^2}e^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y=\int_{\mathbb{R}_+\times(0,2\,\pi)}e^{-\rho^2}\,\rho\,\mathrm{d}\rho\,\mathrm{d}\theta=2\,\pi\int_0^{+\infty}e^{-\rho^2}\,\rho\,\mathrm{d}\rho=\frac{2\,\pi}{-2}\int_0^{+\infty}e^{-\rho^2}\,(-2\,\rho)\,\mathrm{d}\rho=-\pi\left[e^{-\rho^2}\right]_0^{+\infty}=\pi.$$

We now apply again Tonelli's theorem (Theorem 3.54):

$$\pi=\int_{\mathbb{R}^2}e^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y=\int_{\mathbb{R}}\left(\int_{\mathbb{R}}e^{-y^2}\,\mathrm{d}y\right)e^{-x^2}\,\mathrm{d}x=\left(\int_{\mathbb{R}}e^{-x^2}\,\mathrm{d}x\right)^2.$$

The integral cannot take the value $-\sqrt{\pi}$ because the integrand is non-negative, and therefore (6.6) is proved. We now consider the general case where $\eta\in\mathbb{R}$. We can write

$$x^2+i\,\eta\,x=x^2+2\,x\,\frac{i\,\eta}{2}-\frac{\eta^2}{4}+\frac{\eta^2}{4}=\left(x+\frac{i\,\eta}{2}\right)^2+\frac{\eta^2}{4}.$$

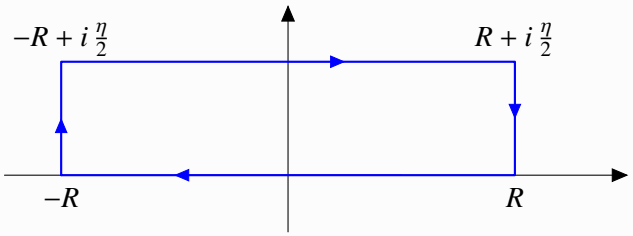
The integral that we have to evaluate in (6.5) can therefore be written as

$$I(\eta)=e^{-\frac{\eta^2}{4}}\int_{\mathbb{R}}\exp\left(-\left(x+\frac{i\,\eta}{2}\right)^2\right)\mathrm{d}x.$$

We show that for every $\eta \in \mathbb{R}$

$$\int_{\mathbb{R}} \exp\left(-\left(x + \frac{i\eta}{2}\right)^2\right) \mathrm{d}x = \sqrt{\pi}. \tag{6.7}$$

Suppose that $\eta > 0$. Then, we compute the integral of e^{z^2} over the contour Γ_R in the complex plane:



The contour Γ_R is given by

$$\Gamma_R = \left[-R + i\frac{\eta}{2}, R + i\frac{\eta}{2}\right] + \left[R + i\frac{\eta}{2}, R\right] - [-R, R] + \left[-R, -R + i\frac{\eta}{2}\right].$$

The integral

$$\int_{\Gamma_R} e^{-z^2}\,\mathrm{d}z$$

is zero, because the integrand is an entire function. We write the same integral as the sum of four integrals, taking into account the parameterisations of the four segments:

$$\gamma_1 : [-R, R] \longrightarrow \mathbb{C} \quad t \mapsto t + \tfrac{i\eta}{2}, \qquad \gamma_2 : \left[0, \tfrac{\eta}{2}\right] \longrightarrow \mathbb{C} \quad t \mapsto R + it, \qquad \gamma_3 : [-R, R] \longrightarrow \mathbb{C} \quad t \mapsto t, \qquad \gamma_4 : \left[0, \tfrac{\eta}{2}\right] \longrightarrow \mathbb{C} \quad t \mapsto -R + it,$$

as

$$0 = \int_{\Gamma_R} e^{-z^2}\,\mathrm{d}z = \int_{-R}^{R} \exp\left(-\left(t + \frac{i\eta}{2}\right)^2\right) \mathrm{d}t - \int_0^{\frac{\eta}{2}} \exp(-(R + it)^2)\, i\,\mathrm{d}t - \int_{-R}^{R} e^{-t^2}\,\mathrm{d}t + \int_0^{\frac{\eta}{2}} \exp(-(-R + it)^2)\, i\,\mathrm{d}t.$$

As $R \to +\infty$, the first integral in the right hand-side tends to the integral that we have to compute. The third integral tends to the Gaussian integral, and its value as $R \to +\infty$ is $\sqrt{\pi}$. The integrals on the segments on the sides go to zero. In fact

$$\begin{aligned}
\left|\int_0^{\frac{\eta}{2}} \exp(-(R + it)^2)\, i\,\mathrm{d}t\right| &\leq \int_0^{\frac{\eta}{2}} |\exp(-(R + it)^2)|\,\mathrm{d}t \\
&= \int_0^{\frac{\eta}{2}} \left|e^{-(R^2 - t^2 + 2iRt)}\right| \mathrm{d}t \\
&= \int_0^{\frac{\eta}{2}} |e^{-(R^2 - t^2)}|\,\mathrm{d}t \\
&= e^{-R^2} \int_0^{\frac{\eta}{2}} e^{t^2}\,\mathrm{d}t,
\end{aligned}$$

which tends to 0 as $R \to +\infty$. A similar calculation shows that the last integral (the one over γ_4) also tends to 0 as $R \to +\infty$, and this completes the exercise.

Exercises

Exercise 6.1

Prove that:

- the product of two even functions is an even function
- the product of two odd functions is an even function
- the product of an even function and an odd functions is an odd function

Moreover, prove that, if the improper integral $\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x$ exists, show that

- if f is even, then
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 2\int_{0}^{\infty} f(x)\,\mathrm{d}x.$$
- if f is odd, then
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 0.$$

Use these results to show that:

- if f is an even function, we have
$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x = 2\int_{0}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x. \tag{6.8}$$
- if f is an odd function, we have
$$\hat{f}(\xi) = -i\int_{-\infty}^{+\infty} \sin(\xi\, x)\, u(x)\,\mathrm{d}x = -2\,i\int_{0}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x. \tag{6.9}$$

[Solution: (6.1)]

Exercise 6.2

Find the Fourier transform of the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by $f(x) = \chi_+(x)\, e^{-x}$, where χ_+ is the characteristic function of $\mathbb{R}_+$. Verify that this is an example where $f \in L^1(\mathbb{R}^n;\mathbb{R})$, but $\hat{f} \notin L^1(\mathbb{R}^n;\mathbb{R})$. [Solution: (6.2)]

Exercise 6.3

Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. Show that
$$\mathcal{F}\{f(\lambda\,\boldsymbol{x})\}(\boldsymbol{\xi}) = \frac{1}{|\lambda|^n}\,\mathcal{F}\{f\}\left(\frac{\boldsymbol{\xi}}{\lambda}\right).$$

[Solution: (6.3)]

Exercise 6.4

Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. Justify the following statements:

- If f is real and even, then $\hat{f}$ is real and even;
- If f is real and odd, then $\hat{f}$ is purely imaginary and odd.

[Solution: (6.4)]

Exercise 6.5

Recalling that the Fourier transform of the function $\frac{1}{1+x^2}$ is $\pi\, e^{-|\xi|}$, determine the Fourier transform of the following functions, using Properties **(P1-P4)** in Theorem 6.2:

(i) $\dfrac{\cos x}{1+x^2}$

(ii) $\dfrac{1}{1+(x+3)^2}$

(iii) $\dfrac{1}{4+x^2}$

[Solution: (6.5)]

Exercise 6.6 Let $f \in L^1(\mathbb{R}^n;\mathbb{C})$. Show that

$$\lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0$$

in the general case where f may not be real a.e. [Solution: (6.6)]

Exercise 6.7 Let H denote the Heaviside function (i.e., $H(x) = 1$ for all $x \geq 0$ and $H(x) = 0$ for all $x < 0$). Check that the following functions are in $L^1(\mathbb{R};\mathbb{R})$ and then determine their Fourier transform:

(i) $f(x) = e^{-x}\,H(x)$

(ii) $f(x) = e^{x}\,H(-x)$

(iii) $f(x) = e^{-|x|}$.

[Solution: (6.7)]

Exercise 6.8 Let $f \in L^1(\mathbb{R};\mathbb{C})$. For all $k \in \mathbb{N}$, consider the function $f_k = f\,\chi_{[-k,k]}$. Show that $\hat{f}_k \xrightarrow{k\to\infty} \hat{f}$ uniformly. [Solution: (6.8)]

Exercise 6.9 Let $A \in \mathbb{R}^{n\times n}$ be invertible. Let $f \in L^1(\mathbb{R}^n;\mathbb{C})$. Show that $f \circ A \in L^1(\mathbb{R}^n;\mathbb{C})$. [Solution: (6.9)]

Exercise 6.10 Let $f \in L^1(\mathbb{R};\mathbb{R})$ be odd. Show that $\hat{f}$ is odd and purely imaginary. [Solution: (6.10)]

Exercise 6.11 Let $\alpha \in \mathbb{R}$. Show that the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^{\alpha} & \text{if } 0 < \|\boldsymbol{x}\| < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

is in $L^1(\mathbb{R}^3;\mathbb{R})$ if and only if $\alpha > -3$. [Solution: (6.11)]

Exercise 6.12 Find the Fourier transform of the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ described by

$$f(x) = e^{-|x|}\ \operatorname{sgn} x,$$

where sgn is the sign function. [Solution: (6.12)]

Exercise 6.13 Find the Fourier transform of the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto e^{-x^2}. \end{aligned}$$

[Solution: (6.13)]

Exercise 6.14 Find the Fourier transform of the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ described by

$$f(x) = x\, e^{-|x|}\ \operatorname{sgn} x,$$

where sgn is the sign function. [Solution: (6.14)]

Exercise 6.15 Let $a, b \in \mathbb{R}$ be such that $a < b$. Show that $\chi_{[a,b]}$ is in $L^1(\mathbb{R}; \mathbb{R})$ and then find its Fourier transform. Verify that the Fourier transform is continuous and that it tends to 0 as $|\xi| \to +\infty$. [Solution: (6.15)]

Exercise 6.16 Find the Fourier transform of the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto e^{-\|x\|^2}. \end{aligned}$$

[Solution: (6.16)]

6.3 Fourier transform and derivatives

Now we examine the interaction between the Fourier transform and the differentiation. We anticipate that, under suitable assumptions, in the one-dimensional case we will find that

$$\mathcal{F}\{f'(x)\} = i\,\xi\,\hat{f}(\xi) \quad \text{and} \quad \mathcal{F}\{x\, f(x)\} = i\,\hat{f}\,'(\xi),$$

where we used the notation described in Remark 6.2. These properties highlight an important fact: the operation of differentiation is mapped into the multiplication of the function by its independent variable. This means that the Fourier transform transforms the differentiation into an algebraic operation. This fact has great relevance in the solution of partial differential equations.

We begin showing that, if $f \in L^1(\mathbb{R}^n; \mathbb{C})$, and for $j \in \{1, \dots, n\}$ we have also $\partial_j f \in L^1(\mathbb{R}^n; \mathbb{C})$, then

$$\widehat{\partial_j f}(\boldsymbol{\xi}) = (i\,\xi_j)\,\hat{f}(\boldsymbol{\xi}),$$

where the derivative $\partial_j f$ is to be intended in the sense of distributions. Intuitively, this result can be seen to be the result of the following steps:

$$\widehat{\partial_j f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = -\int_{\mathbb{R}^n} \partial_j(e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}})\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} i\,\xi_j\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = (i\,\xi_j)\,\hat{f}(\boldsymbol{\xi}).$$

However, we see immediately why we have been sloppy in this chain of identities: the function $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$ is in C^∞, but not of compact support. We now prove this result more rigorously.

Theorem 6.5 *Let $f \in L^1(\mathbb{R}^n; \mathbb{C})$ be of class C^1. Let $j \in \{1, \dots, n\}$, and let $\partial_j f \in L^1(\mathbb{R}^n; \mathbb{C})$. Then,*

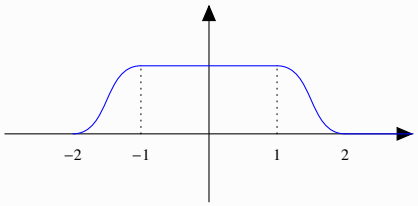
$$\widehat{\partial_j f}(\boldsymbol{\xi}) = (i\,\xi_j)\,\hat{f}(\boldsymbol{\xi}).$$

Proof Consider

$$\widehat{\partial_j f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

The function $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$ does not have compact support. To solve this problem, we use the same idea that we used

to extend the convolution to distributions: we consider $\eta \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ to be such that $\eta(\boldsymbol{x}) = 1$ for $\|\boldsymbol{x}\| \leq 1$ and $\operatorname{supp}\eta \subseteq \overline{\mathcal{B}_2(0)}$.



Consider the sequence $(\eta_k)_{k\in\mathbb{N}^*}$, where $\eta_k = \eta\left(\frac{\boldsymbol{x}}{k}\right)$. We want to show that

$$\widehat{\partial_j f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \lim_{k\to\infty}\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x})\,\partial_j f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

where obviously $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x}) \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$. To this end, we use Lebesgue's dominated convergence theorem (Theorem 3.47), since $\eta_k(\boldsymbol{x}) \xrightarrow[k\to\infty]{} 1$ and

$$\left|e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x})\,\partial_j f(\boldsymbol{x})\right| \leq \sup|\eta|\,\left|\partial_j f(\boldsymbol{x})\right|.$$

Now,

$$\begin{aligned}
\widehat{\partial_j f}(\boldsymbol{\xi}) &= \lim_{k\to\infty}\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x})\,\partial_j f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\
&= \lim_{k\to\infty}\left\langle \partial_j f(\boldsymbol{x}), e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x})\right\rangle\\
&= -\lim_{k\to\infty}\left\langle f(\boldsymbol{x}), -i\,\xi_j\,e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x}) + e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j\eta_k(\boldsymbol{x})\right\rangle\\
&= \lim_{k\to\infty}\int_{\mathbb{R}^n} f(\boldsymbol{x})\left(i\,\xi_j\,e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x}) - e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j\eta_k(\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}.
\end{aligned}$$

We use again Lebesgue's dominated convergence theorem (Theorem 3.47) to take the limit inside the integral. We observe in fact that all the derivatives of $\eta_k(x)$ are equal to zero for $\|\boldsymbol{x}\| \leq 1$, and therefore $\partial_j\eta_k(\boldsymbol{x}) \xrightarrow[k\to\infty]{} 0$. It follows that

$$i\,\xi_j\,e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x}) - e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j\eta_k(\boldsymbol{x}) \xrightarrow[k\to\infty]{} i\,\xi_j\,e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}.$$

Moreover, since

$$\partial_j\eta_k(\boldsymbol{x}) = \frac{1}{k}\frac{\partial\eta}{\partial x_j}\left(\frac{\boldsymbol{x}}{k}\right) \xrightarrow[k\to\infty]{} 0 \qquad (\text{as } \frac{\partial\eta}{\partial x_j}(0) = 0),$$

we have

$$\begin{aligned}
\left|f(\boldsymbol{x})\left(i\,\xi_j\,e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\eta_k(\boldsymbol{x}) - e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\partial_j\eta_k(\boldsymbol{x})\right)\right| &\leq |f(\boldsymbol{x})|\left(|\xi_j|\,|\eta_k(\boldsymbol{x})| + |\partial_j\eta_k(\boldsymbol{x})|\right)\\
&\leq |f(\boldsymbol{x})|\left(|\xi_j|\sup_{\mathbb{R}^n}|\eta| + \frac{1}{k}\sup_{\mathbb{R}^n}|\partial_j\eta|\right)\\
&\leq |f(\boldsymbol{x})|\left(|\xi_j|\sup_{\mathbb{R}^n}|\eta| + \sup_{\mathbb{R}^n}|\partial_j\eta|\right),
\end{aligned}$$

which is a constant that does not depend on k. Taking the limit inside the integral gives

$$\widehat{\partial_j f}(\boldsymbol{\xi}) = i\,\xi_j\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = i\,\xi_j\,\hat{f}(\boldsymbol{\xi})$$

as required.

We now show how, iterating this result, we obtain a generalization of the previous expression. To this end, we extend the multi-index notation to the power. In other words, given $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ and $\alpha \in \mathbb{N}^n$, then

$$z^\alpha = z_1^{\alpha_1} z_2^{\alpha_2} \dots z_n^{\alpha_n}.$$

With this notation, we use the convention that $0^0 = 1$.

Example 6.5 Let $z = (0, 2\,i)$ and $\alpha = (0, 3)$. Then

$$z^\alpha = 0^0 \cdot (2\,i)^3 = 1 \cdot (-8\,i) = -8\,i.$$

Assume that, given $m \in \mathbb{N}$, we have $\partial_x^\alpha f \in L^1(\mathbb{R}^n; \mathbb{C})$ for any multi-index α satisfying $|\alpha| \le m$, and assume also that f is of class C^m. Then, the repeated application of Theorem 6.5 yields

$$\begin{aligned}
\widehat{\partial_x^\alpha f}(\xi) &= \widehat{(\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} f)}(\xi) \\
&= i\,\xi_1\, \widehat{(\partial_{x_1}^{\alpha_1 - 1} \dots \partial_{x_n}^{\alpha_n} f)}(\xi) = \dots = (i\,\xi_1)^{\alpha_1}\, \widehat{(\partial_{x_2}^{\alpha_2} \dots \partial_{x_n}^{\alpha_n} f)}(\xi) \\
&= i\,\xi_n\, \widehat{(\partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n - 1} f)}(\xi) = \dots = (i\,\xi_n)^{\alpha_n}\, \widehat{(\partial_{x_2}^{\alpha_2} \dots \partial_{x_{n-1}}^{\alpha_{n-1}} f)}(\xi).
\end{aligned}$$

The multi-index notation will render this notation more compact in the following corollary.

Corollary 6.4 *Let $m \in \mathbb{N}$. Let $f \in L^1(\mathbb{R}^n; \mathbb{C}) \cap C^m(\mathbb{R}^n; \mathbb{C})$. Let $\alpha \in \mathbb{N}^n$. Let $\partial_x^\alpha f \in L^1(\mathbb{R}^n; \mathbb{C})$ for $|\alpha| \le m$. Then,*

$$\forall\, \xi \in \mathbb{R}^n, \qquad \widehat{\partial_x^\alpha f}(\xi) = (i\,\xi)^\alpha \hat{f}(\xi).$$

We now want to make a very important comment on what we have obtained. Considering for simplicity the case $n = 1$, we have found that

$$\widehat{f'}(\xi) = i\,\xi\,\hat{f}(\xi).$$

This equation says that $\xi\,\hat{f}(\xi)$ is the Fourier transform of a function, and therefore we can apply Riemann-Lebesgue's Lemma to it. This says that $\xi\,\hat{f}(\xi)$ is infinitesimal at infinity, i.e.,

$$\lim_{|\xi| \to +\infty} \xi\,\hat{f}(\xi) = 0.$$

It follows that $\hat{f}(\xi) = o(1/\xi)$ as $|\xi| \to \infty$. In other words, the higher is the order derivatives of f are in L^1, the more pronounced is the decay at infinity of the Fourier transform.

We now want to study the second property mentioned in the beginning of this section, involving the identity $\mathcal{F}\{x\,f(x)\} = i\,\hat{f}'(\xi)$. First, we state this result for $n \ge 1$. If $f \in L^1(\mathbb{R}^n; \mathbb{C})$ and we assume that for some $j \in \{1, \dots, n\}$ we also have $x_j f \in L^1(\mathbb{R}^n; \mathbb{C})$, then $\hat{f}$ admits partial derivative $\partial_j \hat{f}(\xi)$ at every point $\xi \in \mathbb{R}^n$. Moreover, for all $\xi \in \mathbb{R}^n$

$$\partial_j \hat{f}(\xi) = -i\,\widehat{x_j\,f}(\xi).$$

Differentiating formally the Fourier transform with respect to ξ_j, we have

$$\partial_j \hat{f}(\xi) = \frac{\partial}{\partial \xi_j} \int_{\mathbb{R}^n} e^{-i\,x\cdot\xi}\, f(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} e^{-i\,x\cdot\xi}\,(-i\,x_j)\,f(x)\,\mathrm{d}x = -i \int_{\mathbb{R}^n} e^{-i\,x\cdot\xi}\,x_j\,f(x)\,\mathrm{d}x = -i\,\widehat{x_j\,f}(\xi).$$

We make this argument more precise with the following result.

Theorem 6.6 *Let $f \in L^1(\mathbb{R}^n;\mathbb{C})$. Let $j \in \{1,\dots,n\}$, and let $x_j f \in L^1(\mathbb{R}^n;\mathbb{C})$. Then, $\hat{f}$ is differentiable with respect to ξ_j (in the classical sense) and*

$$\frac{\partial \hat{f}}{\partial \xi_j}(\boldsymbol{\xi}) = \widehat{-i\,x_j\,f}\,(\boldsymbol{\xi}).$$

Proof Let $\boldsymbol{e}_j$ be the j-th standard basis vector of $\mathbb{R}^n$, and consider the difference quotient

$$\frac{\hat{f}(\xi + t\,\boldsymbol{e}_j) - \hat{f}(\boldsymbol{\xi})}{t} = \int_{\mathbb{R}^n} \frac{e^{-i\,\boldsymbol{x}\cdot(\xi+t\boldsymbol{e}_j)} - e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}}{t}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\frac{e^{-i\,t\,x_j}-1}{t}\, f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

We want to take the limit of this quantity as $t \to 0$, and to this end we need to ensure that we can exchange the limit with the integral. We use Lebesgue's dominated convergence theorem (Theorem 3.47) by noting that, in view of de L'Hôpital's rule (or, alternatively, the power series expansion of the exponential), we have

$$\lim_{t\to 0} \frac{e^{-i\,t\,x_j}-1}{t} = -i\,x_j,$$

so that the limit of the integrand is $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,(-i\,x_j)\,f(\boldsymbol{x})$. Next, we bound the integrand with a summable function that does not depend on t: from the fundamental theorem of integral calculus, since

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{-i\,t\,x_j}) = -i\,x_j\,e^{-i\,t\,x_j},$$

we find

$$\left|\frac{e^{-i\,t\,x_j}-1}{t}\right| = \left|\frac{1}{t}\int_0^t (-i\,x_j)\,e^{-i\,s\,x_j}\,\mathrm{d}s\right| \le \frac{1}{|t|}\,|t|\sup_{s\in\mathbb{R}}|(-i\,x_j)\,e^{-i\,x_j\,s}| = |x_j|$$

(notice that we need the absolute value of t because t could be negative). It follows that

$$\left|e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\frac{e^{-i\,t\,x_j}-1}{t}\,f(\boldsymbol{x})\right| \le |x_j|\,|f(\boldsymbol{x})|,$$

and the right hand-side is in $L^1(\mathbb{R}^n;\mathbb{C})$ by assumption. Bringing the limit inside the integral gives

$$\lim_{t\to 0}\frac{\hat{f}(\xi+t\,\boldsymbol{e}_j)-\hat{f}(\boldsymbol{\xi})}{t} = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,(-i\,x_j)\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

as required.

Once again, we can iterate this result as follows. Suppose that $f \in L^1(\mathbb{R}^n;\mathbb{C})$, and $\boldsymbol{\alpha} \in \mathbb{N}^n$ with $|\boldsymbol{\alpha}| \le m$, where m is a natural number. We observe that

$$\boldsymbol{x}^{\boldsymbol{\alpha}} = x_1^{\alpha_1}\cdots x_n^{\alpha_n},$$

and this implies that $|\boldsymbol{x}^{\boldsymbol{\alpha}}| \le \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|}$ since $|x_j| \le \|\boldsymbol{x}\|$ for all $j \in \{1,\dots,n\}$. Indeed

$$|\boldsymbol{x}^{\boldsymbol{\alpha}}| = |x_1^{\alpha_1}\cdots x_n^{\alpha_n}| = |x_1|^{\alpha_1}\cdots|x_n|^{\alpha_n} \le \|\boldsymbol{x}\|^{\alpha_1}\cdots\|\boldsymbol{x}\|^{\alpha_n} = \|\boldsymbol{x}\|^{\alpha_1+\ldots+\alpha_n} = \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|}.$$

If m is such that $(1+\|\boldsymbol{x}\|)^m f(\boldsymbol{x}) \in L^1(\mathbb{R}^n;\mathbb{C})$, then $\boldsymbol{x}^{\boldsymbol{\alpha}}\,f(\boldsymbol{x}) \in L^1(\mathbb{R}^n;\mathbb{C})$ for all $\alpha \in \mathbb{N}^n$ such that $|\boldsymbol{\alpha}| \le m$, since

$$|\boldsymbol{x}^{\boldsymbol{\alpha}}\,f(\boldsymbol{x})| \le \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|}\,|f(\boldsymbol{x})| \le (1+\|\boldsymbol{x}\|)^m\,|f(\boldsymbol{x})|.$$

Corollary 6.5 *Let $m \in \mathbb{N}$. Let $(1 + \|\boldsymbol{x}\|)^m f \in L^1(\mathbb{R}^n; \mathbb{C})$. Then, $\hat{f} \in C^m(\mathbb{R}^n; \mathbb{C})$. Moreover, for every $\boldsymbol{\alpha} \in \mathbb{N}^n$ such that $|\boldsymbol{\alpha}| \leq m$, we have*

- $(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x}) \in L^1(\mathbb{R}^n; \mathbb{C})$;
- $\forall\, \boldsymbol{\xi} \in \mathbb{R}^n, \quad \partial^{\alpha}_{\boldsymbol{\xi}} \hat{f}(\boldsymbol{\xi}) = \widehat{(-i\,\boldsymbol{x})^{\alpha} f}\,(\boldsymbol{\xi})$.

Proof We first observe that, if $(1 + \|\boldsymbol{x}\|)^m f \in L^1(\mathbb{R}^n; \mathbb{C})$, we have $f \in L^1(\mathbb{R}^n; \mathbb{C})$, since for all $\boldsymbol{x} \in \mathbb{R}^n$

$$|f(\boldsymbol{x})| \leq (1 + \|\boldsymbol{x}\|)^m |f(\boldsymbol{x})|.$$

We now show that, if $|\boldsymbol{\alpha}| \leq m$, we have $(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x}) \in L^1(\mathbb{R}^n; \mathbb{C})$. We use the same argument presented before, noting that

$$(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x}) = (-i\,x_1)^{\alpha_1} \cdots (-i\,x_n)^{\alpha_n} f(\boldsymbol{x}),$$

so that

$$\left|(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x})\right| = |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} |f(\boldsymbol{x})| \leq \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} |f(\boldsymbol{x})|, \tag{6.10}$$

where we used the fact that $|x_j| \leq \|\boldsymbol{x}\|$ for all $j \in \{1, \ldots, n\}$. We now show that $\|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} \leq (1 + \|\boldsymbol{x}\|)^m$. We can distinguish two cases:

- if $\|\boldsymbol{x}\| \leq 1$, then $\|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} \leq 1 \leq (1 + \|\boldsymbol{x}\|)^m$, and the statement is true;
- if $\|\boldsymbol{x}\| > 1$, then $\|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} \leq \|\boldsymbol{x}\|^m \leq (1 + \|\boldsymbol{x}\|)^m$, and the statement is, again, true.

Therefore, getting back to (6.10), we obtain

$$\left|(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x})\right| \leq \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} |f(\boldsymbol{x})| \leq (1 + \|\boldsymbol{x}\|)^m |f(\boldsymbol{x})|.$$

Since $(1 + \|\boldsymbol{x}\|)^m f \in L^1(\mathbb{R}^n; \mathbb{C})$ by assumption, we can conclude that $(-i\,\boldsymbol{x})^{\alpha} f(\boldsymbol{x}) \in L^1(\mathbb{R}^n; \mathbb{C})$. Moreover, by virtue of Riemann-Lebesgue's Theorem, we have $\partial^{\alpha} \hat{f} \in BC(\mathbb{R}^n; \mathbb{C})$, so that $\hat{f} \in C^m(\mathbb{R}^n; \mathbb{C})$, and $\partial^{\alpha} \hat{f}$ is continuous, and

$$\lim_{\|\boldsymbol{\xi}\| \to +\infty} \partial^{\alpha} \hat{f} = 0.$$

Now, we find

$$\begin{aligned}
\frac{\partial \hat{f}}{\partial \xi_j}(\boldsymbol{\xi}) &= \mathcal{F}((-i\,x_1)\, f)(\boldsymbol{\xi}) \\
\frac{\partial^2 \hat{f}}{\partial \xi_j^2}(\boldsymbol{\xi}) &= \mathcal{F}((-i\,x_1)(-i\,x_1)\, f)(\boldsymbol{\xi}) = \mathcal{F}((-i\,x_1)^2 f)(\boldsymbol{\xi}) \\
&\vdots \\
\partial^{\alpha_1}_{\xi_1} \hat{f}(\boldsymbol{\xi}) &= \mathcal{F}((-i\,x_1)^{\alpha_1} f)(\boldsymbol{\xi}) \\
&\vdots \\
\partial^{\alpha_1}_{\xi_1} \partial^{\alpha_2}_{\xi_2} \ldots \partial^{\alpha_n}_{\xi_n} \hat{f}(\boldsymbol{\xi}) &= \mathcal{F}((-i\,x_1)^{\alpha_1} (-i\,x_2)^{\alpha_2} \ldots (-i\,x_n)^{\alpha_n} f)(\boldsymbol{\xi}) = \mathcal{F}((-i\,\boldsymbol{x})^{\alpha} f).
\end{aligned}$$

Example 6.6 Let $t \in \mathbb{R}_+$. Let $n \in \mathbb{N}^*$. The fundamental solution of the heat equation

$$u_t(\boldsymbol{x}) = (4\,\pi\,t)^{-\frac{n}{2}} \exp\left(-\frac{\|\boldsymbol{x}\|^2}{4\,t}\right)$$

vanishes rapidly as $\|\boldsymbol{x}\| \to +\infty$. The function $\exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right)$ is bounded, and if we multiply it by any polynomial, it remains bounded. Thus, given $p \in \mathbb{N}$, the function $\exp\left(-\frac{\|\boldsymbol{x}\|^2}{4t}\right)(1+\|\boldsymbol{x}\|)^p$ is bounded for every $\boldsymbol{x} \in \mathbb{R}^n$, so that

$$|u_t(\boldsymbol{x})|(1+\|\boldsymbol{x}\|)^p \leq c_{t,p}, \tag{6.11}$$

where $c_{t,p}$ is a positive constant. We want to prove that $(1+\|\boldsymbol{x}\|)^m |u_t(\boldsymbol{x})|$ is integrable for all $m \in \mathbb{N}$. In fact, from (6.11)

$$(1+\|\boldsymbol{x}\|)^{m+n+1} |u_t(\boldsymbol{x})| \leq c_{t,m+n+1},$$

where $c_{t,n+m+1}$ is a positive constant, and therefore

$$|u_t(\boldsymbol{x})|\,(1+\|\boldsymbol{x}\|)^m \leq c_{t,m+n+1}\,(1+\|\boldsymbol{x}\|)^{-n-1}.$$

As $\|\boldsymbol{x}\| \to \infty$, the right hand-side goes to infinity in the same way as $\|\boldsymbol{x}\|^{-n-1}$. In Chapter 5, we saw that the condition that guarantees the convergence at 0 of the integral of $\|\boldsymbol{x}\|^\alpha$ is $\alpha > -n$, and this implies that the condition that guarantees the convergence at ∞ of the integral of $\|\boldsymbol{x}\|^\alpha$ is $\alpha < -n$. In this case, this condition is satisfied because $-n-1 < -n$.

We can apply Corollary 6.5, which shows that $\hat{u}_t$ is of class C^∞, and

$$\forall\, \boldsymbol{\alpha} \in \mathbb{N}^n, \quad \partial^\alpha_{\boldsymbol{\xi}} \hat{u}_t(\boldsymbol{\xi}) = \mathcal{F}((-i\,\boldsymbol{x})^\alpha\, u_t)(\boldsymbol{\xi}).$$

In particular, all the derivatives $\partial^\alpha_{\boldsymbol{\xi}} \hat{u}_t$ of the Fourier transform are bounded and tend to 0 as $\|\boldsymbol{\xi}\|$ tends to ∞.

We conclude this section with the fundamental inversion formula.

Theorem 6.7 *Let $f \in L^1(\mathbb{R}^n;\mathbb{C})$ and assume that $\hat{f} \in L^1(\mathbb{R}^n;\mathbb{C})$. Then,*

$$f(\boldsymbol{x}) = \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, \hat{f}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} = \frac{1}{(2\,\pi)^n}\, \hat{\hat{f}}(-\boldsymbol{x}) \qquad \textit{almost everywhere in } \mathbb{R}^n. \tag{6.12}$$

Proof We prove the result under the further assumption that f is bounded and continuous, i.e., $f \in BC(\mathbb{R}^n;\mathbb{C})$. By definition of Fourier transform, we have

$$\frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, \hat{f}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} = \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \left(\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} \right) \mathrm{d}\boldsymbol{\xi}.$$

We cannot interchange the two integrals using Fubini's theorem (Theorem 3.8), because the function $e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi} - i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{y})$ of the two variables $\boldsymbol{y}$ and $\boldsymbol{\xi}$ would need to be summable in $\boldsymbol{y}$ and $\boldsymbol{\xi}$, but

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi} - i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{y})|\, \mathrm{d}\boldsymbol{y}\, \mathrm{d}\boldsymbol{\xi} = \int_{\mathbb{R}^n \times \mathbb{R}^n} |f(\boldsymbol{y})|\, \mathrm{d}\boldsymbol{y}\, \mathrm{d}\,\boldsymbol{\xi} = +\infty.$$

We overcome this obstacle by introducing an integrating factor and then applying Lebesgue's dominated convergence theorem (Theorem 3.47):

$$\begin{aligned} \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, \hat{f}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} &= \lim_{k\to\infty} \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi} - \frac{\|\boldsymbol{\xi}\|^2}{k}}\, \hat{f}(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \\ &= \lim_{k\to\infty} \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi} - \frac{\|\boldsymbol{\xi}\|^2}{k}} \left(\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{y})\, \mathrm{d}\boldsymbol{y} \right) \mathrm{d}\boldsymbol{\xi} \end{aligned}$$

The first equality holds in view of Lebesgue's dominated convergence theorem (Theorem 3.47), since $e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi} - \frac{\|\boldsymbol{\xi}\|^2}{k}}\, \hat{f}(\boldsymbol{\xi})$

converges to $e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi})$ as $k\to\infty$, and since

$$\left|e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\hat{f}(\boldsymbol{\xi})\right| = e^{-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,|\hat{f}(\boldsymbol{\xi})| \le |\hat{f}(\boldsymbol{\xi})|$$

where the right hand-side is integrable by assumption. We now want to apply Fubini's theorem (Theorem 3.8). To this end, we evaluate

$$\left|e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,f(\boldsymbol{y})\right| = e^{-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,|f(\boldsymbol{y})|$$

and we show that it is integrable in $(\boldsymbol{y},\boldsymbol{\xi})$. Indeed, using Tonelli's theorem (Theorem 3.54) (for which only the measurability is required since the functions involved are non-negative)

$$\begin{aligned}\int_{\mathbb{R}^n\times\mathbb{R}^n} e^{-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{y} &= \int_{\mathbb{R}^n} e^{-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\mathrm{d}\boldsymbol{\xi}\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} = \left(\int_{\mathbb{R}} e^{-\frac{t^2}{k}}\,\mathrm{d}t\right)^n\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\\ &= \left(\int_{\mathbb{R}} e^{-\tau^2}\,\sqrt{k}\,\mathrm{d}\tau\right)^n\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} = k^{\frac{n}{2}}\left(\int_{\mathbb{R}} e^{-\tau^2}\,\mathrm{d}\tau\right)^n\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y}\\ &= k^{\frac{n}{2}}\,(\sqrt{\pi})^n\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} = (k\,\pi)^{\frac{n}{2}}\int_{\mathbb{R}^n}|f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} < +\infty.\end{aligned}$$

We can therefore apply Fubini's theorem (Theorem 3.8) and we find

$$\frac{1}{(2\,\pi)^n}\int_{\mathbb{R}^n} e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\left(\int_{\mathbb{R}^n} e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{\xi} = \frac{1}{(2\,\pi)^n}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\mathrm{d}\boldsymbol{\xi}\right)f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

We now apply Fubini's theorem (Theorem 3.8) in the internal integral and use the change of variable $t_j = \xi_j/\sqrt{k}$:

$$\begin{aligned}\int_{\mathbb{R}^n} e^{i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\mathrm{d}\boldsymbol{\xi} &= \prod_{j=1}^n\int_{\mathbb{R}} e^{i(x_j-y_j)\,\xi_j-\frac{\xi_j^2}{k}}\,\mathrm{d}\xi_j\\ &= \prod_{j=1}^n\int_{\mathbb{R}} e^{i(x_j-y_j)\,\sqrt{k}\,t_j-t_j^2}\,k^{\frac{1}{2}}\,\mathrm{d}t_j\\ &= k^{\frac{n}{2}}\prod_{j=1}^n\int_{\mathbb{R}} e^{i(x_j-y_j)\,\sqrt{k}\,t_j-t_j^2}\,\mathrm{d}t_j\\ &= k^{\frac{n}{2}}\prod_{j=1}^n\int_{\mathbb{R}} e^{-i\,\eta_j\,t_j-t_j^2}\,\mathrm{d}t_j,\end{aligned}$$

where $\eta_j = k^{\frac{1}{2}}\,(y_j-x_j)$. Now, recalling (6.5), which says that $\displaystyle\int_{\mathbb{R}} e^{-i\,x\,\eta-x^2}\,\mathrm{d}x = \sqrt{\pi}\,e^{-\frac{\eta^2}{4}}$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} e^{i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\mathrm{d}\boldsymbol{\xi} &= k^{\frac{n}{2}}\prod_{j=1}^n\sqrt{\pi}\,e^{-\frac{\eta_j^2}{4}}\\ &= (k\,\pi)^{\frac{n}{2}}\prod_{j=1}^n e^{-\frac{k\,(y_j-x_j)^2}{4}}\\ &= (k\,\pi)^{\frac{n}{2}}\,e^{-\frac{k\,(y_1-x_1)^2}{4}-\frac{k\,(y_2-x_2)^2}{4}-\ldots-\frac{k\,(y_n-x_n)^2}{4}} = (k\,\pi)^{\frac{n}{2}}\,e^{-\frac{k\,\|\boldsymbol{x}-\boldsymbol{y}\|^2}{4}}.\end{aligned}$$

We use this result in the original integral

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} e^{i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{\xi}-\frac{\|\boldsymbol{\xi}\|^2}{k}}\,\mathrm{d}\boldsymbol{\xi}\right) f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \left(\frac{k}{4\pi}\right)^{\frac{n}{2}}\int_{\mathbb{R}^n} e^{-\frac{k\|\boldsymbol{x}-\boldsymbol{y}\|^2}{4}} f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \pi^{-\frac{n}{2}}\int_{\mathbb{R}^n} e^{-\|\boldsymbol{z}\|^2} f\left(\boldsymbol{x}-\frac{2\,\boldsymbol{z}}{\sqrt{k}}\right)\mathrm{d}\boldsymbol{z},$$

where we changed variable using $\boldsymbol{z} = \sqrt{k}\,\frac{\boldsymbol{x}-\boldsymbol{y}}{2}$, and the inverse of the determinant of the Jacobian is $\left(\frac{2}{\sqrt{k}}\right)^n = \left(\frac{4}{k}\right)^{\frac{n}{2}}$. Since f is assumed to be continuous, applying the Sequential Continuity Theorem we have for all $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$

$$\lim_{k\to\infty} e^{-\|\boldsymbol{z}\|^2} f\left(\boldsymbol{x}-\frac{2\,\boldsymbol{z}}{\sqrt{k}}\right) = e^{-\|\boldsymbol{z}\|^2} f(\boldsymbol{x}).$$

Moreover, for every $k \in \mathbb{N}$, for every $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$, we have

$$\left|e^{-\|\boldsymbol{z}\|^2} f\left(\boldsymbol{x}-\frac{2\,\boldsymbol{z}}{\sqrt{k}}\right)\right| \le \sup_{\mathbb{R}^n}|f|\, e^{-\|\boldsymbol{z}\|^2},$$

which is integrable. We can therefore apply Lebesgue's dominated convergence theorem (Theorem 3.47) to bring the limit inside the integral, to conclude that

$$\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \pi^{-\frac{n}{2}}\int_{\mathbb{R}^n} e^{-\|\boldsymbol{z}\|^2}\,\mathrm{d}\boldsymbol{z}\, f(\boldsymbol{x}) = f(\boldsymbol{x}),$$

from the value of the Gaussian integral. In fact, using again Fubini's theorem (Theorem 3.8)

$$\int_{\mathbb{R}^n} e^{-\|\boldsymbol{z}\|^2}\,\mathrm{d}\boldsymbol{z} = \int_{\mathbb{R}^n} e^{-z_1^2-z_2^2-\dots-z_n^2}\,\mathrm{d}z_1\,\mathrm{d}z_2\,\dots\,\mathrm{d}z_n = \left(\int_{\mathbb{R}} e^{-t^2}\,\mathrm{d}t\right)^n = (\sqrt{\pi})^n = \pi^{\frac{n}{2}}.$$

We observe that the equality in the inversion formula (6.12) holds for almost every $\boldsymbol{x} \in \mathbb{R}^n$. If f is continuous (as it was assumed in the previous proof), then the equality holds for all $\boldsymbol{x} \in \mathbb{R}^n$.

A direct consequence of the inversion formula is that, assuming that $f, \hat{f} \in L^1(\mathbb{R}^n;\mathbb{C})$ and that f is continuous, if $\hat{f}(\boldsymbol{\xi}) = 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^n$, then $f = 0$.

Corollary 6.6 *The linear operator*

$$\begin{aligned}\mathcal{F} : L^1(\mathbb{R}^n,\mathbb{C}) &\longrightarrow BC(\mathbb{R}^n;\mathbb{C})\\ f &\mapsto \mathcal{F}\{f\},\end{aligned}$$

as defined in Corollary 6.2, is injective.

Proof Since $\mathcal{F}$ is linear, it suffices to show that its kernel has only one element 0, which is the zero of the vector space $L^1(\mathbb{R}^n;\mathbb{C})$. This is a direct consequence of the previous discussion and of the inversion formula (6.12).

The following result will be useful in the sequel, to extend the Fourier transform to distributions.

Theorem 6.8 *Let $f, g \in L^1(\mathbb{R}^n;\mathbb{C})$. Then,*

$$\int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \int_{\mathbb{R}^n} f(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}. \tag{6.13}$$

Proof First, we notice that the integrals in (6.13) exist. Indeed, for example, from Riemann-Lebesgue's Lemma we know that $\hat{f}$ is bounded, and

$$|\hat{f}(\boldsymbol{\xi})\, g(\boldsymbol{\xi})| \leq \sup_{\mathbb{R}^n} |\hat{f}|\, |g(\boldsymbol{\xi})| = \|\hat{f}\|_\infty\, |g(\boldsymbol{\xi})| \in L^1(\mathbb{R}^n;\mathbb{C}).$$

The same argument applies to the integral in the right hand-side. Using the definition of Fourier transform, we write the integral in the left hand-side of (6.13) as

$$\int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} \right) g(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}. \tag{6.14}$$

We want to apply Fubini's theorem (Theorem 3.8) to the right hand-side of (6.14). To this end, we observe that the mapping $(\boldsymbol{x},\boldsymbol{\xi}) \mapsto e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\, g(\boldsymbol{\xi})$ (which is, in essence, given by the product of the the tensor product $f(\boldsymbol{x})\, g(\boldsymbol{\xi})$, which is measurable, with the measurable function $e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$) is integrable over $\mathbb{R}^n \times \mathbb{R}^n$, since $|e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, f(\boldsymbol{x})\, g(\boldsymbol{\xi})| = |f(\boldsymbol{x})|\, |g(\boldsymbol{\xi})|$, and using Tonelli's theorem (Theorem 3.54)

$$\int_{\mathbb{R}^{2n}} |f(\boldsymbol{x})|\, |g(\boldsymbol{\xi})|\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{\xi} = \int_{\mathbb{R}^n} |f(\boldsymbol{x})|\, \mathrm{d}\boldsymbol{x} \cdot \int_{\mathbb{R}^n} |g(\boldsymbol{\xi})|\, \mathrm{d}\boldsymbol{\xi} = \|f\|_1 \cdot \|g\|_1 < +\infty.$$

Therefore, using Fubini's theorem (Theorem 3.8) to the right hand-side of (6.14), we find

$$\int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi})\, g(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} = \int_{\mathbb{R}^n} f(\boldsymbol{x}) \left(\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, g(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi} \right) \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} f(\boldsymbol{x})\, \hat{g}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

This result suggests how to define the Fourier transform for a distribution. If $f \in L^1(\mathbb{R}^n;\mathbb{C})$ and $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, in accordance with the identity (6.13), we could define

$$\langle \hat{f}, \varphi \rangle = \langle f, \hat{\varphi} \rangle.$$

The difficulty is the fact that, in general, $\hat{\varphi} \notin \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, because in general it does not have compact support. On the contrary, it is possible to prove that the only function $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ such that $\hat{\varphi} \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ is the identically zero function. Consider for example the case $n = 1$. We have

$$\hat{\varphi}(\xi) = \int_{\mathbb{R}} e^{-i\,x\,\xi}\, \varphi(x)\, \mathrm{d}x.$$

We can consider $\xi \in \mathbb{C}$, and what we obtain is an entire function, i.e., holomorphic in the entire complex plane $\mathbb{C}$. On the other hand, $\hat{\varphi}$ should be zero outside a certain closed interval of the real line, in order to belong to $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$. From the principle of identity of analytic functions, if the set of points where a holomorphic function over an open set is zero has accumulation points in that set, then the function is identically zero. It follows that $\hat{\varphi}(\xi) = 0$ for all $\xi \in \mathbb{C}$, and by the inversion formula we obtain $\varphi = 0$.

Exercises

Exercise 6.17 Recalling that the Fourier transform of the function $\frac{1}{1+x^2}$ is $\pi\, e^{-|\xi|}$, determine the Fourier transform of the following functions, using the properties of the derivatives:

(i) $\dfrac{x}{(1+x^2)^2}$

(ii) $\dfrac{x^2}{(1+x^2)^2}$

(iii) $\dfrac{1}{(1+x^2)^2}$

[Solution: (6.17)]

6.4 Regularity at infinity of the Fourier transform

First, consider the case $n = 1$. In this case, considering a function $f \in \mathcal{L}^1(\mathbb{R};\mathbb{C}) \cap \boldsymbol{C}^1(\mathbb{R};\mathbb{C})$ such that $f' \in \mathcal{L}^1(\mathbb{R};\mathbb{C})$, we have seen that $\mathcal{F}\{f'\}(\xi) = i\,\xi\,\hat{f}(\xi)$. In other words, $i\,\xi\,\hat{f}(\xi)$ is the Fourier transform of the function f', and we can apply the Riemann-Lebesgue Lemma to it:

$$\xi\,\hat{f}(\xi) \overset{\xi\to\infty}{\longrightarrow} 0.$$

This implies that

$$\frac{\hat{f}(\xi)}{\frac{1}{\xi}} \overset{\xi\to\infty}{\longrightarrow} 0 \qquad \Rightarrow \qquad \hat{f}(\xi) \overset{\xi\to\infty}{\ll} \frac{1}{|\xi|}.$$

This means, in particular, that as $|\xi|$ goes to $+\infty$, the Fourier transform $\hat{f}(\xi)$ not only goes to zero (we knew that from Riemann-Lebesgue), but it goes to zero faster that $|\xi|^{-1}$.

The same result applies for the case $n \neq 1$. Consider $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C}) \cap \boldsymbol{C}^1(\mathbb{R}^n;\mathbb{C})$. For all $j \in \{1,2,\ldots,n\}$, let $\partial_j f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. We have seen that for all $j \in \{1,2,\ldots,n\}$ we have $\mathcal{F}\{\partial_j f\}(\boldsymbol{\xi}) = i\,\xi_j\,\hat{f}(\boldsymbol{\xi})$. Hence, $\xi_k\,\hat{f}(\boldsymbol{\xi}) \overset{\|\boldsymbol{\xi}\|\to+\infty}{\longrightarrow} 0$. Since $\|\boldsymbol{\xi}\| \leq |\xi_1| + |\xi_2| + \ldots + |\xi_n|$, we have also $\|\boldsymbol{\xi}\|\,\hat{f}(\boldsymbol{\xi}) \overset{\xi\to+\infty}{\longrightarrow} 0$. Thus

$$\frac{\hat{f}(\boldsymbol{\xi})}{\frac{1}{\|\boldsymbol{\xi}\|}} \overset{\xi\to\infty}{\longrightarrow} 0 \quad \Rightarrow \quad \hat{f}(\boldsymbol{\xi}) \overset{x\to\infty}{\ll} \frac{1}{\|\boldsymbol{\xi}\|}.$$

We can generalise this result to higher-order derivatives: Let $m \in \mathbb{N}$. Let $f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C}) \cap \boldsymbol{C}^m(\mathbb{R}^n;\mathbb{C})$. For every $\boldsymbol{\alpha} \in \mathbb{N}^n$ such that $|\boldsymbol{\alpha}| \leq m$, let $\partial^\alpha f \in \mathcal{L}^1(\mathbb{R}^n;\mathbb{C})$. Then,

$$\hat{f}(\boldsymbol{\xi}) \overset{x\to+\infty}{\ll} \frac{1}{\|\boldsymbol{\xi}\|^m}.$$

Loosely speaking, the higher is the order of differentiability, the faster the Fourier transform will converge to zero as $\|\boldsymbol{\xi}\|$ tends to ∞. Stated differently, f is smooth if and only if $\hat{f}$ quickly goes to 0 as $\|\boldsymbol{\xi}\| \to \infty$. Since, as we say, the inverse Fourier transform follows analogous rules of the Fourier transform, the converse of this rule of thumb holds true, and therefore we can say that f quickly vanishes as $\|\boldsymbol{x}\| \to \infty$ if and only if $\hat{f}$ is smooth.

6.4.1 Fourier transform of the convolution

If the functions f and g are in $L^1(\mathbb{R}^n;\mathbb{C})$, the Fourier transform of their convolution is the product of the Fourier transforms of f and g:

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\}\mathcal{F}\{g\}.$$

This result makes sense, because we know that the convolution of functions that live in $L^1(\mathbb{R}^n;\mathbb{C})$ is a function that also lives in $L^1(\mathbb{R}^n;\mathbb{C})$, and the Fourier transform $\mathcal{F}\{f * g\}$ is well defined. Moreover, the product in the right hand-side involves two continuous functions which tend to 0 as $\|\boldsymbol{\xi}\| \to \infty$.

Theorem 6.9 [Fourier Transform and Convolution]
Let $f, g \in L^1(\mathbb{R}^n;\mathbb{C})$. Then, for all $\boldsymbol{\xi} \in \mathbb{R}^n$

$$\widehat{f * g}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi}).$$

Proof We use the definition of convolution and of Fourier transform:

$$\begin{aligned}\widehat{f * g}(\boldsymbol{\xi}) &= \int_{\mathbb{R}^n} e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}}\,(f * g)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathbb{R}^n} e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}} \left(\int_{\mathbb{R}^n} f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-i(\boldsymbol{x}-\boldsymbol{y})\boldsymbol{\xi}}\,f(\boldsymbol{x}-\boldsymbol{y})\,e^{-i\boldsymbol{y}\boldsymbol{\xi}}\,g(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}\boldsymbol{x},\end{aligned}$$

since $e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}} = e^{-i(\boldsymbol{x}-\boldsymbol{y})\boldsymbol{\xi}}\,e^{-i\boldsymbol{y}\boldsymbol{\xi}}$. We want to change the order of integration. To this end, we need to show that the integrand is integrable in $\mathbb{R}^{2n}$. However, we have already established, when dealing with the convolution, that

$$|e^{-i(\boldsymbol{x}-\boldsymbol{y})\boldsymbol{\xi}}\,f(\boldsymbol{x}-\boldsymbol{y})\,e^{-i\boldsymbol{y}\boldsymbol{\xi}}\,g(\boldsymbol{y})| = |f(\boldsymbol{x}-\boldsymbol{y})\,g(\boldsymbol{y})|$$

is integrable in $\mathbb{R}^{2n}$. We can therefore apply Fubini's theorem (Theorem 3.8):

$$\begin{aligned}\widehat{f * g}(\boldsymbol{\xi}) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-i(\boldsymbol{x}-\boldsymbol{y})\boldsymbol{\xi}}\,f(\boldsymbol{x}-\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\right) g(\boldsymbol{y})\,e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{y}\\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-iz\boldsymbol{\xi}}\,f(z)\,\mathrm{d}z\right) g(\boldsymbol{y})\,e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{y}\\ &= \hat{f}(\boldsymbol{\xi}) \int_{\mathbb{R}^n} g(\boldsymbol{y})\,e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{y} = \hat{f}(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi}).\end{aligned}$$

Exercises

Exercise 6.18

Using a converse of Fubini's theorem (Theorem 3.8), verify that, if $f_1, f_2, \ldots, f_n \in L^1(\mathbb{R};\mathbb{C})$, then

(i) $f_1 \otimes \ldots \otimes f_n \in L^1(\mathbb{R}^n;\mathbb{C})$

(ii) $\widehat{f_1 \otimes \ldots \otimes f_n} = \hat{f}_1 \otimes \ldots \otimes \hat{f}_n$.

Use this result to find the Fourier transform of an n-dimensional cube $B = [-1,1] \times \ldots \times [-1,1]$. [Solution: (6.18)]

6.5 Schwartz space

The previous discussion showed a duality between the operations on the function and the operations on the Fourier transform:

- translations correspond to products by exponentials with purely imaginary exponent and *vice-versa*:

$$\widehat{f(\cdot - \boldsymbol{y})}(\boldsymbol{\xi}) = e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{f}(\boldsymbol{\xi}) \qquad \text{and} \qquad \widehat{e^{i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,f}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi}-\boldsymbol{y});$$

- changes of variables correspond to other changes of variables:

$$\widehat{f \circ A}(\boldsymbol{\xi}) = \frac{1}{|\det A|} (\hat{f} \circ A^{-\top})(\boldsymbol{\xi});$$

- differentiation with respect to a variable x_j and multiplication by the dual variable ξ_j corresponds to the multiplication by x_j and differentiation with respect to ξ_j:

$$\widehat{\frac{\partial f}{\partial x_j}}(\boldsymbol{\xi}) = (i\,\xi_j)\,\hat{f}(\boldsymbol{\xi}) \qquad \text{and} \qquad \frac{\partial \hat{f}}{\partial \xi_j}(\boldsymbol{\xi}) = \widehat{-i\,x_j\,f}\,(\boldsymbol{\xi}).$$

Similar considerations apply to the properties of the original function and its Fourier transform: the smoothness of f corresponds to how fast $\hat{f}$ decays at ∞, and, conversely, the speed of decay of f depends on the regularity of $\hat{f}$.

We have also observed that these analogies are not perfect. The two fundamental results on the relation of the Fourier transform with the operation of differentiation are not entirely symmetric: the hypothesis of Theorem 6.5 is not the thesis of Theorem 6.6 and *vice-versa*. Finally, the formula in Theorem 6.7 was referred to as *inversion formula*, but, rigorously, it should not be regarded as a formula for the inverse Fourier transform, because $\mathcal{F}^{-1}\mathcal{F} f = f$ does not hold for every function in $L^1(\mathbb{R}^n;\mathbb{C})$: we need to restrict ourselves to the class of functions which are continuous and bounded. These considerations motivate our desire to look for functional spaces different from $L^1(\mathbb{R}^n;\mathbb{C})$, where each property holds with the corresponding dual property. There are two possibilities at this point. One consists in seeking to extend the Fourier transform to functions in $L^2(\mathbb{R}^n;\mathbb{C})$. This goal is certainly within reach, but it does not solve all the aforementioned issues. The second possibility is more general, and useful for our purposes: extend the notion of Fourier transform for distributions. This will naturally encompass the case of functions in $L^2(\mathbb{R}^n;\mathbb{C})$. This is the choice that we make in these notes.

We begin introducing the space of rapidly decreasing functions of class C^∞, also known as the *Schwartz space* (from Laurent Schwartz): it is the space comprising smooth functions φ whose derivatives of every order are rapidly decreasing. This request is accomplished by requiring that the derivatives of any order, multiplied by any power of x, remains bounded.

Definition 6.2 [Schwartz Space]
We call space of rapidly decreasing functions *or* Schwartz space *the set* $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ *defined by*

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \stackrel{def}{=} \{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n, \quad \forall m \in \mathbb{N}, \quad \sup_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{x}\|^m \,|\partial^\alpha \varphi(\boldsymbol{x})| < +\infty\}$$

To understand the meaning of this definition, we first write $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ as

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \exists\, c > 0 : \;\; \forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|^m \,|\partial^\alpha \varphi(\boldsymbol{x})| \le c\}.$$

We are interested in what happens as $\|\boldsymbol{x}\| \to \infty$. In a neighbourhood of ∞, the quantity $\|\boldsymbol{x}\|^m$ is invertible, and therefore the condition $\|\boldsymbol{x}\|^m \,|\partial^\alpha \varphi(\boldsymbol{x})| \le c$ can be written as

$$|\partial^\alpha \varphi(\boldsymbol{x})| \le \frac{c}{\|\boldsymbol{x}\|^m}.$$

As $\|\boldsymbol{x}\| \to \infty$, the right hand-side goes to 0, and therefore the left hand-side goes to zero as well, and it goes to zero faster than $\frac{c}{\|\boldsymbol{x}\|^m}$. Since m is arbitrary, we can say that $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is the set of all functions of class $C^\infty(\mathbb{R}^n;\mathbb{C})$ which decrease, together with all their derivatives, faster than any power of $\|\boldsymbol{x}\|^{-1}$, as $\|\boldsymbol{x}\| \to \infty$.

Sometimes, the alternative definition for $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ (which is useful in making the above argument independent from a neighbourhood of ∞) is given:

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \exists\, c > 0 : \;\; \forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad (1 + \|\boldsymbol{x}\|^m)\, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \le c\},$$

see Question **??**. This last form is convenient because for every $\boldsymbol{x} \in \mathbb{R}^n$ the quantity $1 + \|\boldsymbol{x}\|^m$ is different from zero, and we can write

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \left\{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \exists\, c > 0 : \;\; \forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \le \frac{c}{(1 + \|\boldsymbol{x}\|^m)}\right\}.$$

This expression is saying that the function $\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})$ is negligible[1] with respect to $\frac{1}{(1+\|\boldsymbol{x}\|^m)}$ at ∞. In symbols:

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \left\{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,\middle|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}) \overset{x \to +\infty}{\ll} \frac{1}{1 + \|\boldsymbol{x}\|^m}\right\},$$

or, also,

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) = \left\{\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,\middle|\, \forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}) = \mathrm{o}_{x \to +\infty}\left(\frac{1}{1 + \|\boldsymbol{x}\|^m}\right)\right\}.$$

We notice that $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is a vector space (why?). Thus, in particular, it contains the identically zero function $\varphi(\boldsymbol{x}) = 0$ (which is of class C^∞, and for every $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$ satisfies $\|\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| = 0 < +\infty$).

An immediate consequence of the definition is the fact that, if $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then, $\partial^{\boldsymbol{\alpha}} \varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ for every $\boldsymbol{\alpha} \in \mathbb{N}^n$.

We can say even more. We can use the multi-index notation to define multivariate polynomials in the indeterminate $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ of degree $\le m$ by considering a sum over all multi-indices whose length is $\le m$:

$$p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, \boldsymbol{x}^{\boldsymbol{\alpha}}, \tag{6.15}$$

where, for each $\boldsymbol{\alpha} \in \mathbb{N}^n$ with $|\boldsymbol{\alpha}| \le m$, the coefficient of the monomial $\boldsymbol{x}^{\boldsymbol{\alpha}}$ is $a_{\boldsymbol{\alpha}} \in \mathbb{C}$. The set of polynomials defined in this way is denoted, in accordance with the usual notation, by $\mathbb{C}_m[\boldsymbol{x}]$. For example, if $n = 3$ and we write $\boldsymbol{x} = (x, y, z)$, a generic multivariate polynomial of degree ≤ 2 can be written as

$$\begin{aligned} p(\boldsymbol{x}) &= \sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, (x, y, z)^{(\alpha_1, \alpha_2, \alpha_3)} = \sum_{|\boldsymbol{\alpha}| \le 2} a_{\boldsymbol{\alpha}}\, x^{\alpha_1}\, y^{\alpha_2}\, z^{\alpha_3} \\ &= a_{0,0,0} + a_{1,0,0}\, x + a_{0,1,0}\, y + a_{0,0,1}\, z \\ &\quad + a_{1,1,0}\, x\, y + a_{1,0,1}\, x\, z + a_{0,1,1}\, y\, z + a_{2,0,0}\, x^2 + a_{0,2,0}\, y^2 + a_{0,0,2}\, z^2. \end{aligned}$$

[1] We recall that if $\Omega \subseteq \mathbb{R}^n$ and we have two functions $f, g : \Omega \longrightarrow \mathbb{R}^m$, and $\boldsymbol{x}_0 \in \overline{\Omega}$, then f is asymptotically negligible with respect to f at $\boldsymbol{x}_0$ (or, f is a "little o" of g as $\boldsymbol{x} \to \boldsymbol{x}_0$), and we write $f \overset{x \to x_0}{\ll} g$, if there exists a neighbourhood $\boldsymbol{U}$ of $\boldsymbol{x}_0$ such that for all $\boldsymbol{x} \in \boldsymbol{U} \cap \Omega$ we have $\|f(\boldsymbol{x})\| \le \epsilon\, \|g(\boldsymbol{x})\|$. If in addition $g(\boldsymbol{x}) \neq 0$ for every $\boldsymbol{x} \in \boldsymbol{U}$, then $f \overset{x \to x_0}{\ll} g$ if and only if $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{\|f(\boldsymbol{x})\|}{\|g(\boldsymbol{x})\|} = 0$.

With this notation, we can define a linear differential operator with constant coefficients as a polynomial defined as in (6.15) where the indeterminate $\boldsymbol{x}$ is replaced by the symbol of differentiation ∂, and where the constant term $a_{0,0,0}$ is replaced by the operator $a_{0,0,0}\, I$, which represents the multiplication by $a_{0,0,0}$:

$$Q(\partial) = \sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, \partial^{\boldsymbol{\alpha}} \in \mathbb{C}_m[\partial].$$

Example *6.7*

The Laplacian operator is obtained by the multivariate polynomial $p(\boldsymbol{x}) = p(x_1, x_2, \ldots, x_n) = x_1^2 + x_2^2 + \ldots + x_n^2$, where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, and in fact the corresponding differential operator is written as

$$Q(\partial) = Q(\partial_1, \partial_2, \ldots, \partial_n) = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \ldots + \frac{\partial^2}{\partial x_n^2} = \nabla^2.$$

Example *6.8*

The differential operator associated with the polynomial $p(\boldsymbol{x}) = 1 - \|\boldsymbol{x}\|^2 = 1 - \sum_{j=1}^{n} x_j$ is $I - \sum_{j=1}^{n} \partial_j^2 = I - \nabla^2 = I - \Delta$, which maps a sufficiently regular function φ into $\varphi - \Delta\varphi$.

With this notation, we can generalize the previous discussion with the following theorem.

Theorem 6.10 *Let $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. Let $m \in \mathbb{N}$ and $Q \in \mathbb{C}_m[\partial]$. Then,*

$$Q(\partial)\, \varphi = \sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, \partial^{\boldsymbol{\alpha}} \varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C}).$$

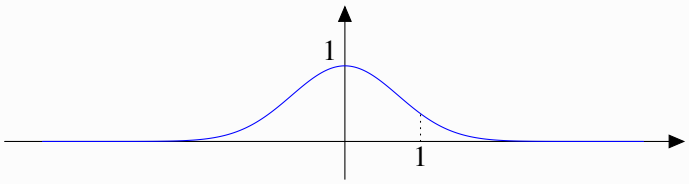
Example *6.9*

The prototypical example of a function in $\mathcal{S}(\mathbb{R}^n; \mathbb{R})$ is the function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ described by

$$\varphi(\boldsymbol{x}) = e^{-\|\boldsymbol{x}\|^2}, \qquad \boldsymbol{x} \in \mathbb{R}^n.$$

This function is not in $\mathcal{D}(\mathbb{R}^n; \mathbb{R})$ because it does not vanish, and therefore its support is not compact.

$n = 1$



However, it is in $\mathcal{S}(\mathbb{R}^n; \mathbb{R})$. Indeed, $\|\boldsymbol{x}\|^m \left| e^{-\|\boldsymbol{x}\|^2} \right|$ is bounded for all $m \in \mathbb{N}$, since $\varphi(\boldsymbol{x}) \to 0$ as $\|\boldsymbol{x}\| \to 0$. Moreover, given a multi-index $\boldsymbol{\alpha} \in \mathbb{N}^n$,

$$\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}) = p(\boldsymbol{x})\, e^{-\|\boldsymbol{x}\|^2}$$

where p is a polynomial of degree $|\boldsymbol{\alpha}|$. For example, if $n = 3$, $\boldsymbol{x} = (x, y, z)$ and $\boldsymbol{\alpha} = (1, 2, 0)$, we find

$$\varphi(\boldsymbol{x}) = e^{-x^2 - y^2 - z^2}$$

and

$$\begin{aligned}
\partial_x \varphi(\boldsymbol{x}) &= (-2\,x)\, e^{-x^2 - y^2 - z^2}, \\
\partial_x \partial_y \varphi(\boldsymbol{x}) &= (-2\,x)\,(-2\,y)\, e^{-x^2 - y^2 - z^2} = 4\,x\,y\, e^{-x^2 - y^2 - z^2}, \\
\partial_x \partial_y^2 \varphi(\boldsymbol{x}) &= (4\,x - 8\,x\,y^2)\, e^{-x^2 - y^2 - z^2},
\end{aligned}$$

so that

$$\partial^{\alpha}\varphi(\boldsymbol{x}) = (4\,x - 8\,x\,y^2)\,e^{-x^2-y^2-z^2}.$$

A more precise proof of this fact requires induction over $\boldsymbol{\alpha}$ (try!).

Recalling that, for a multi-index $\boldsymbol{\alpha} \in \mathbb{N}^n$ we have

$$|\boldsymbol{x}^{\boldsymbol{\alpha}}| = |x_1^{\alpha_1} \cdots x_n^{\alpha_n}| = |x_1|^{\alpha_1} \cdots |x_n|^{\alpha_n} \leq \|\boldsymbol{x}\|^{\alpha_1} \cdots \|\boldsymbol{x}\|^{\alpha_n} = \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|},$$

we find

$$|p(\boldsymbol{x})| = \Big|\sum_{|\boldsymbol{\beta}|\leq m} a_{\boldsymbol{\beta}}\,\boldsymbol{x}^{\boldsymbol{\beta}}\Big| \leq \sum_{|\boldsymbol{\beta}|\leq m} |a_{\boldsymbol{\beta}}|\,|\boldsymbol{x}^{\boldsymbol{\beta}}| \leq \sum_{|\boldsymbol{\beta}|\leq p} |a_{\boldsymbol{\beta}}|\,\|\boldsymbol{x}\|^{|\boldsymbol{\beta}|},$$

so that

$$\|\boldsymbol{x}\|^m\,|\partial^{\boldsymbol{\alpha}}\varphi| = \|\boldsymbol{x}\|^m\,|p(\boldsymbol{x})\,e^{-\|\boldsymbol{x}\|^2}| \leq \sum_{|\boldsymbol{\beta}|\leq m} |a_{\boldsymbol{\beta}}|\,\|\boldsymbol{x}\|^{|\boldsymbol{\beta}|+m}\,e^{-\|\boldsymbol{x}\|^2},$$

which is bounded.

Example *6.10*

The function

$$\begin{aligned}\varphi : \mathbb{R} &\longrightarrow \mathbb{R}\\ x &\mapsto e^{-x^2}\,\sin(e^{x^2})\end{aligned}$$

is not in $\mathcal{S}(\mathbb{R};\mathbb{R})$. In fact, its order of infinitesimal as $x \to +\infty$ is greater than any real number (or, in other words,

$$|x|^m\,|\varphi(x)| = |x|^m\,e^{-x^2}\,\sin(e^{x^2})$$

is bounded for every $m \in \mathbb{N}$ since $|\sin(e^{x^2})| \leq 1$. However, this is not true for the derivatives of every order. For example, the first derivative

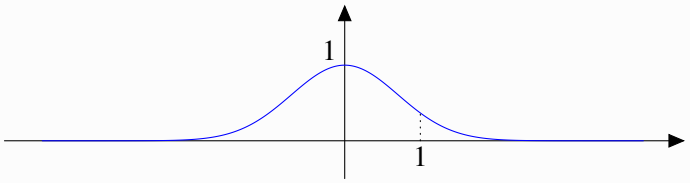
$$\varphi'(x) = -2\,x\,e^{-x^2}\,\sin(e^{x^2}) + e^{-x^2}\,\cos(e^{x^2})\,2\,x\,e^{x^2} = 2\,x\,\cos(e^{x^2}) - 2\,x\,e^{-x^2}\,\sin(e^{x^2}),$$

is not even infinitesimal.

Example *6.11*

Consider the function

$$\begin{aligned}\varphi : \mathbb{R} &\longrightarrow \mathbb{R}\\ x &\mapsto e^{-|x|}\end{aligned}$$



Function φ satisfies the condition on the rate of decrease of the function with all its derivatives, but it is not of class C^{∞}: it not differentiable at the origin.

Example *6.12*

We notice that no polynomial functions can be in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, except for the zero polynomial function. Consider for example $n = 1$ and the polynomial function

$$\varphi(x) = a_0 + a_1\,x + \ldots + a_q\,x^q,$$

where $q \in \mathbb{N}^*$ (which is the degree of $\varphi(x)$). Then, using $m = 0$ and $\alpha = 1$ yields the condition

$$\sup_{x\in\mathbb{R}} |\varphi'| = \sup_{x\in\mathbb{R}} |a_1 + 2\,a_2\,x + \ldots + q\,a_q\,x^{q-1}|.$$

The supremum in this case is $+\infty$.

The next theorem shows that a function in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ (i.e., in C^∞ and with compact support) is also in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.11 *There holds*

$$\mathcal{D}(\mathbb{R}^n;\mathbb{C}) \subset \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

Proof Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$. Therefore, there exists $R > 0$ such that $\varphi(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$ such that $\|\boldsymbol{x}\| > R$. From the Weierstrass Extreme Value Theorem, for all $\boldsymbol{x} \in \mathbb{R}^n$ such that $\|\boldsymbol{x}\| \leq R$ we have

$$\|\boldsymbol{x}\|^m \, |\varphi(\boldsymbol{x})| \leq R^m \max_{\|\boldsymbol{x}\| \leq R} |\varphi(\boldsymbol{x})|.$$

On the other hand, for $\|\boldsymbol{x}\| > R$ we have $\varphi(\boldsymbol{x}) = 0$. Thus, $\|\boldsymbol{x}\|^m \, |\varphi(\boldsymbol{x})|$ is bounded over $\mathbb{R}^n$. The same argument can be used for the derivatives of any order of φ. Thus, $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. The inclusion is strict because we have seen an example of a function ($\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$, with $\varphi(\boldsymbol{x}) = e^{-\|\boldsymbol{x}\|^2}$) in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ whose support is not compact.

We now observe that, if $f \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, then $[f] \in L^p(\mathbb{R}^n;\mathbb{C})$ for any $p \in [1,+\infty]$ (why?). Thus, modulo the usual identification of a function f with its equivalence class $[f]$, we can say that $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ is contained in $L^p(\mathbb{R}^n;\mathbb{C})$ for any $p \in [1,+\infty]$. We established also that $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ is dense in $L^p(\mathbb{R}^n;\mathbb{C})$ for any $p \in [1,+\infty)$. We now show that the same is true for the Schwartz space $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, even if in this case this result is slightly less obvious. In other words, $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is contained in every space $L^p(\mathbb{R}^n;\mathbb{C})$ in the same sense as for $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$: the set

$$\boldsymbol{W} = \{[\varphi] \,|\, \varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})\}$$

is contained in $L^p(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.12 *For every $p \in [1,+\infty]$ there holds*

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subseteq L^p(\mathbb{R}^n;\mathbb{C}).$$

Proof If $p = \infty$ the statement is obvious, since in $L^\infty(\mathbb{R}^n;\mathbb{C})$ we have boundedness almost everywhere, while in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ we have boundedness of the derivatives of every order. Suppose $p \neq \infty$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We want to show that

$$\int_{\mathbb{R}^n} |\varphi(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} < \infty.$$

Since $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, given a natural number $m \in \mathbb{N}$, there exists $c_m > 0$ such that $\|\boldsymbol{x}\|^m |\varphi(\boldsymbol{x})| \leq c_m$; let us choose m to be satisfying the inequality $m > n/p$. Let c_0 be such that $|\varphi(\boldsymbol{x})| \leq c_0$. Combining these inequalities and defining $C_m = c_0 + c_m$ yields

$$|\varphi(\boldsymbol{x})| \leq C_m \, (1 + \|\boldsymbol{x}\|^m)^{-1},$$

so that

$$\int_{\mathbb{R}^n} |\varphi(\boldsymbol{x})|^p \, \mathrm{d}\boldsymbol{x} \leq C_m^p \int_{\mathbb{R}^n} (1 + \|\boldsymbol{x}\|^m)^{-p} \, \mathrm{d}\boldsymbol{x}.$$

The integral in the right-hand side is finite because as $\|\boldsymbol{x}\|$ goes to infinity the function $(1 + \|\boldsymbol{x}\|^m)^{-p}$ behaves like the function $\|\boldsymbol{x}\|^{-m\,p}$, and $-m\,p < -n$ since $m > n/p$.

Our task will be, eventually, to define distributions for test functions in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Since distributions were defined as linear functionals on $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ which are continuous with respect to the notion of convergence introduced in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, we introduce an analogous notion of convergence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, which is inevitably different from the one given for $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$.

Definition 6.3 [Convergence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$]
Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We say that $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ if for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$ we have

$$\lim_{k\to\infty} \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}\varphi_k(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x})| = 0$$

uniformly in $\mathbb{R}^n$.

This notion of convergence can sometimes be expressed in an apparently more general way:

$$p(\boldsymbol{x})\, q(\partial)\, \varphi_k(\boldsymbol{x}) \xrightarrow{k\to\infty} p(\boldsymbol{x})\, q(\partial)\, \varphi(\boldsymbol{x}) \quad \text{uniformly in } \mathbb{R}^n$$

for every pair of polynomials p and q. However, since a polynomial is a finite linear combination of monomials, to check this condition it is sufficient to consider the case where p and q are monomials. For example, if $n = 1$ it is sufficient to check the uniform convergence with $p(x) = x^m$ and $q(x) = x^r$, with arbitrary $m, r \in \mathbb{N}$. In this case, $q(\partial)$ is the r-th derivative.

Remark 6.3 Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. If $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, for all multivariate polynomials p

$$p(\partial)\,\varphi_k \xrightarrow{k\to\infty} p(\partial)\,\varphi$$

in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ (prove this result: the solution is in the Worksheet). In particular, if $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ we also have

$$\partial^{\boldsymbol{\alpha}}\varphi_k \xrightarrow{k\to\infty} \partial^{\boldsymbol{\alpha}}\varphi \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

We will need this result for the well posedness of the definition of derivative of tempered distributions.

We now show that if a sequence $(\varphi_k)_{k\in\mathbb{N}}$ converges in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, it converges also in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. In other words, the notion of convergence $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ introduced in this section is weaker than the one for $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.13 *Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, and let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$. Then,*

$$\lim_{k\to\infty} \varphi_k = \varphi \ \text{ in } \mathcal{D}(\mathbb{R}^n;\mathbb{C}) \quad \Rightarrow \quad \lim_{k\to\infty} \varphi_k = \varphi \ \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C})$$

Proof We assume that $\lim_{k\to\infty} \varphi_k = \varphi$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$. Thus, the supports of all the φ_k are contained in a compact set that does not depend on k, so that there exists $R > 0$ such that $\operatorname{supp}\varphi_k \subseteq \overline{\mathcal{B}_R(0)}$ for all $k \in \mathbb{N}$. Outside

$\mathcal{B}_R(0)$, all the φ_k are zero. For all $\boldsymbol{x} \in \mathcal{B}_R(0)$, for every $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $m \in \mathbb{N}$ we find

$$\begin{aligned} \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi_k(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| &\leq \sup_{\mathbb{R}^n} \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi_k(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \\ &\leq \sup_{\|\boldsymbol{x}\| \leq R} \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi_k(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \\ &\leq R^m \sup_{\|\boldsymbol{x}\| \leq R} |\partial^{\boldsymbol{\alpha}} \varphi_k(\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})|, \end{aligned}$$

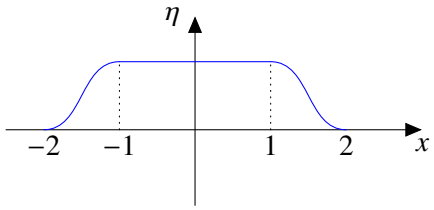
which converges to zero uniformly as $k \to \infty$ because the derivatives of φ_k uniformly converge to the derivatives of φ.

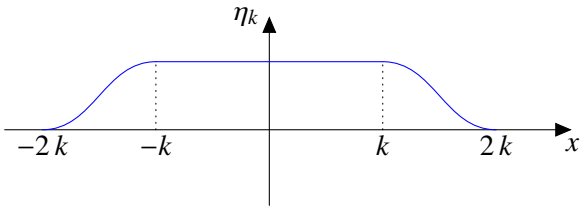
Remark 6.4 The converse of Theorem 6.13 does not hold in general: there are sequences in $\mathcal{D}(\mathbb{R}^n; \mathbb{C})$ that converge in the sense of $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$ but not in the sense of $\mathcal{D}(\mathbb{R}^n; \mathbb{C})$. We propose an example. Let $\eta \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$ be a non-negative function such that $\eta(\boldsymbol{x}) = 1$ when $\|\boldsymbol{x}\| \leq 1$ and $\eta(\boldsymbol{x}) = 0$ when $\|\boldsymbol{x}\| \geq 2$.

Let us now define the sequence of non-negative functions $(\eta_k)_{k \in \mathbb{N}}$ of $\mathcal{D}(\mathbb{R}^n; \mathbb{R})$ in this way: $\eta_k(t) \overset{\text{def}}{=} \eta(\frac{\boldsymbol{x}}{k})$, for $k \in \mathbb{N}^*$, so that

$$\eta : \mathbb{R} \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \mapsto \begin{cases} 1 & \text{if } \|\boldsymbol{x}\| \leq 1 \\ 0 & \text{if } \|\boldsymbol{x}\| \geq 2 \end{cases} \qquad \text{and} \qquad \eta_k : \mathbb{R} \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \mapsto \begin{cases} 1 & \text{if } \|\boldsymbol{x}\| \leq k \\ 0 & \text{if } \|\boldsymbol{x}\| \geq 2\,k \end{cases}$$

and in the case $n = 1$ we have the following examples of graphical representations of η and η_k (for some $k \in \mathbb{N}^*$):





Let us define, for all $k \in \mathbb{N}^*$, the functions $\varphi_k \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$ in the following way: for all, $\boldsymbol{x} \in \mathbb{R}^n$, we define $\varphi_k(\boldsymbol{x}) = 2^{-k} \, \eta_k(\boldsymbol{x})$, so that

$$\varphi_k : \mathbb{R} \longrightarrow \mathbb{R}, \quad \boldsymbol{x} \mapsto \begin{cases} 2^{-k} & \text{if } \|\boldsymbol{x}\| \leq k \\ 0 & \text{if } \|\boldsymbol{x}\| \geq 2\,k. \end{cases}$$

The sequence $(\varphi_k)_{k \in \mathbb{N}^*}$ does not converge in $\mathcal{D}(\mathbb{R}^n; \mathbb{R})$, because there does not exist a compact set K in $\mathbb{R}^n$ which contains $\operatorname{supp} \varphi_k = [-2\,k, 2\,k]$ for every $k \in \mathbb{N}^*$. We show, nevertheless, that this sequence converges in $\mathcal{S}(\mathbb{R}^n; \mathbb{R})$ to the zero function. Indeed, for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $m \in \mathbb{N}$

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi_k(\boldsymbol{x})| = \|\boldsymbol{x}\|^m \left| 2^{-k} \, \partial^{\boldsymbol{\alpha}} \eta\left(\frac{\boldsymbol{x}}{k}\right) \right|$$

If $\|\boldsymbol{x}\| \geq 2\,k$, this quantity is zero because $\operatorname{supp} \eta_k = \overline{\mathcal{B}_{2\,k}(0)}$, and therefore the partial derivatives of

every order of η_k are also zero. Thus, we can bound this quantity assuming $\|\boldsymbol{x}\| \leq 2\,k$:

$$\|\boldsymbol{x}\|^m\,|\partial^{\boldsymbol{\alpha}}\varphi_k(\boldsymbol{x})| \;=\; \|\boldsymbol{x}\|^m\,\left|2^{-k}\,\partial^{\boldsymbol{\alpha}}\eta\left(\frac{\boldsymbol{x}}{k}\right)\right| \leq (2\,k)^m\,2^{-k}\sup_{[-2\,k,2\,k]}|\partial^{\boldsymbol{\alpha}}\eta|$$
$$= 2^m\,(k^m\,2^{-k})\sup_{[-2\,k,2\,k]}|\partial^{\boldsymbol{\alpha}}\eta| \xrightarrow{k\to\infty} 0,$$

since $k^m\,2^{-k} \xrightarrow{k\to\infty} 0$ for any $m \in \mathbb{N}$.

Exercises

Exercise 6.19

Let $\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C})$. Prove that the following two conditions are equivalent:

(A) $\forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \exists\, c > 0 : \;\; \forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|^m\,|\partial^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x})| \leq c$

(B) $\forall \boldsymbol{\alpha} \in \mathbb{N}^n \;\; \forall m \in \mathbb{N}, \;\; \exists\, c > 0 : \;\; \forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad (1 + \|\boldsymbol{x}\|^m)\,|\partial^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x})| \leq c$

[Solution: (6.19)]

Exercise 6.20

Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. If $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, show that for all multivariate polynomials p

$$p(\partial)\,\varphi_k \xrightarrow{k\to\infty} p(\partial)\,\varphi$$

in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. [Solution: (6.20)]

6.5.1 Slowly increasing functions

To define the space of tempered distributions $\mathcal{S}'$ by using the set of test functions $\mathcal{S}$ introduced above, we will need to ask ourselves what happens to the operations that we have introduced for $\mathcal{D}'$. We will see that the operations of differentiation and of composition with a C^∞-diffeomorphism are continuous from $\mathcal{S}$ to $\mathcal{S}$. By contrast, the set $\mathcal{S}$ is not closed under multiplication by a function of class C^∞. For example, $\varphi = e^{-\|\boldsymbol{x}\|^2}$ is a function in $\mathcal{S}(\mathbb{R}^n;\mathbb{R})$, but its product by the function $e^{\|\boldsymbol{x}\|^2}$, which is of class C^∞, gives the constant function 1 (i.e., the function $\mathbb{R}^n \longrightarrow \mathbb{R}$ that maps $\boldsymbol{x} \in \mathbb{R}^n$ into 1), which is not an element of $\mathcal{S}(\mathbb{R}^n;\mathbb{R})$.

Hence, the product of a tempered distribution by a function of class C^∞ needs to be defined in a different way. In particular, we define the set of functions of class C^∞ which grow at infinity, together with all their derivatives, no faster than a polynomial (and these functions are therefore referred to as *slowly increasing functions* or *polynomially increasing functions*).

Definition 6.4 [Slowly Increasing Functions]
A function $\psi \in C^\infty(\mathbb{R}^n;\mathbb{C})$ is said to be a slowly increasing function *if for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ there exists $c_{\boldsymbol{\alpha}} \geq 0$ and $m_{\boldsymbol{\alpha}} \in \mathbb{N}$ such that*

$$|\partial^{\boldsymbol{\alpha}}\psi(\boldsymbol{x})| \leq c_{\boldsymbol{\alpha}}\,(1 + \|\boldsymbol{x}\|)^{m_\alpha} \tag{6.16}$$

for all $\boldsymbol{x} \in \mathbb{R}^n$. We denote by $\mathcal{O}_M$ the set of slowly increasing functions, i.e.,

$$\mathcal{O}_M(\mathbb{R}^n;\mathbb{C}) \stackrel{def}{=} \{\psi \in C^\infty(\mathbb{R}^n;\mathbb{C}) \,|\, \forall\, \boldsymbol{\alpha} \in \mathbb{N}^n,\ \exists\, c_{\boldsymbol{\alpha}} > 0,\ \exists\, m_{\boldsymbol{\alpha}} \in \mathbb{N}\ :\ \ \forall\, \boldsymbol{x} \in \mathbb{R}^n\ \ |\partial^{\boldsymbol{\alpha}}\psi(\boldsymbol{x})| \leq c_{\boldsymbol{\alpha}}\,(1+\|\boldsymbol{x}\|)^{m_{\boldsymbol{\alpha}}}\}.$$

The elements of $\mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$ are the functions of class C^∞ that do not increase faster than polynomials. The reason why, in the right hand-side of the inequality (6.16), we have the factor $(1+\|\boldsymbol{x}\|)^{m_\alpha}$ instead of just $\|\boldsymbol{x}\|^{m_\alpha}$ is due to the fact that we do not want this term to vanish for $\boldsymbol{x} = 0$. Indeed, if we replaced $(1+\|\boldsymbol{x}\|)^{m_\alpha}$ with $\|\boldsymbol{x}\|^{m_\alpha}$ in (6.16), we would end up requiring ψ, with all its partial derivatives of every order, to be zero at the origin.

Example *6.13*

Polynomials in $\mathbb{C}[\boldsymbol{x}]$, where $\boldsymbol{x} \in \mathbb{R}^n$, are in $\mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$. Indeed, we know we can write an element of $\mathbb{C}[\boldsymbol{x}]$ as

$$p(\boldsymbol{x}) = \sum_{|\boldsymbol{\alpha}|\leq m} a_{\boldsymbol{\alpha}}\,\boldsymbol{x}^{\boldsymbol{\alpha}},$$

and, moreover,

$$|\boldsymbol{x}^{\boldsymbol{\alpha}}| \leq \|\boldsymbol{x}\|^{|\boldsymbol{\alpha}|} \leq (1+\|\boldsymbol{x}\|)^{|\boldsymbol{\alpha}|} \leq (1+\|\boldsymbol{x}\|)^m.$$

Hence,

$$|p(\boldsymbol{x})| \leq \sum_{|\boldsymbol{\alpha}|\leq m} |a_{\boldsymbol{\alpha}}|\,(1+\|\boldsymbol{x}\|)^m.$$

A similar argument holds for the derivatives of any order.

We now show that $\mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$ is a space of multipliers for $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ that does not cause us to go beyond the limits of the set $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. In other words, the product of a rapidly decreasing function by a slowly increasing function is rapidly decreasing.

Theorem 6.14 *Let $\psi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$, and $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then, $\psi\,\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$.*

Proof We need to show that for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$ there exists a constant C such that $\|\boldsymbol{x}\|^m\,|\partial^{\boldsymbol{\alpha}}(\psi\,\varphi)(\boldsymbol{x})| \leq C$. We recall the generalization of Leibniz rule

$$\partial^{\boldsymbol{\alpha}}(\psi\,\varphi) = \sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} c_{\boldsymbol{\alpha},\boldsymbol{\beta}}\,\partial^{\boldsymbol{\alpha}-\boldsymbol{\beta}}\psi\,\partial^{\boldsymbol{\beta}}\,\varphi,$$

where $c_{\boldsymbol{\alpha},\boldsymbol{\beta}} \geq 0$. Since in the function ψ and φ can be interchanged in the product (as they are both scalar functions), and using the fact that $\psi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$, which implies that for every $\boldsymbol{\beta} \in \mathbb{N}^n$ we can write $|\partial^{\boldsymbol{\beta}}\psi(\boldsymbol{x})| \leq$

$c_\beta\,(1+\|\boldsymbol{x}\|)^{m_\beta}$ for some $c_\beta \geq 0$ and $m_\beta \in \mathbb{N}$, then

$$\begin{aligned}
\|\boldsymbol{x}\|^m\,|\partial^\alpha(\psi\,\varphi)(\boldsymbol{x})| &= \|\boldsymbol{x}\|^m \left|\sum_{\beta\leq\alpha} c_{\alpha,\beta}\,\partial^\beta\psi\,\partial^{\alpha-\beta}\varphi\right| \\
&\leq \|\boldsymbol{x}\|^m \sum_{\beta\leq\alpha} c_{\alpha,\beta}\left|\partial^\beta\psi\right|\left|\partial^{\alpha-\beta}\varphi\right| \\
&\leq \|\boldsymbol{x}\|^m \sum_{\beta\leq\alpha} c_{\alpha,\beta}\,c_\beta\,(1+\|\boldsymbol{x}\|)^{m_\beta}\,\left|\partial^{\alpha-\beta}\varphi\right| \\
&\leq \|\boldsymbol{x}\|^m \sum_{\beta\leq\alpha} c_{\alpha,\beta}\,c_\beta \sum_{j=0}^{m_\beta}\binom{m_\beta}{j}\|\boldsymbol{x}\|^j\,\left|\partial^{\alpha-\beta}\varphi\right| \\
&= \sum_{\beta\leq\alpha}\sum_{j=0}^{m_\beta} c_{\alpha,\beta}\,c_\beta\binom{m_\beta}{j}\|\boldsymbol{x}\|^{m+j}\,\left|\partial^{\alpha-\beta}\varphi\right|
\end{aligned}$$

where we used the Newton's formula.[a]

The function is the right hand-side is bounded, since since the functions $\|\boldsymbol{x}\|^{m+j}\,|\partial^\beta\,\varphi(\boldsymbol{x})|$ are bounded from the assumption that $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

[a] Recall Newton's formula of the binomial expansion theorem:

$$(a+b)^n = \sum_{j=0}^{n}\binom{n}{k}a^{n-k}\,b^k.$$

The following result shows that the multiplication by functions of $\mathcal{O}_M$ is a continuous operator.

Theorem 6.15 *Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Let $\psi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$. Then,*

$$\varphi_k \xrightarrow{k\to\infty} \varphi \;\; \text{in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}) \quad\Rightarrow\quad \psi\,\varphi_k \xrightarrow{k\to\infty} \psi\,\varphi \;\; \text{in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

Proof We have to show that, for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$,

$$\|\boldsymbol{x}\|^m\,|\partial^\alpha(\varphi_k\,\psi)(\boldsymbol{x}) - \partial^\alpha(\varphi\,\psi)(\boldsymbol{x})| \xrightarrow{k\to\infty} 0$$

uniformly in $\mathbb{R}^n$. We use Leibniz rule, and the fact that ψ is in $O_M(\mathbb{R}^n;\mathbb{C})$:

$$\begin{aligned}
\|\boldsymbol{x}\|^m \left|\partial^\alpha(\varphi_k\,\psi)(\boldsymbol{x}) - \partial^\alpha(\varphi\,\psi)(\boldsymbol{x})\right| &= \|\boldsymbol{x}\|^m \left|\partial^\alpha(\psi\,(\varphi_k - \varphi))\right| \\
&= \|\boldsymbol{x}\|^m \left|\sum_{\beta\le\alpha} c_{\alpha,\beta}\,\partial^\beta\psi(\boldsymbol{x})\,\partial^{\alpha-\beta}(\varphi_k(\boldsymbol{x}) - \varphi(\boldsymbol{x}))\right| \\
&\le \|\boldsymbol{x}\|^m \sum_{\beta\le\alpha} c_{\alpha,\beta}\left|\partial^\beta\psi(\boldsymbol{x})\right|\left|\partial^{\alpha-\beta}\varphi_k(\boldsymbol{x}) - \partial^{\alpha-\beta}\varphi(\boldsymbol{x})\right| \\
&\le \|\boldsymbol{x}\|^m \sum_{\beta\le\alpha} c_{\alpha,\beta}\,c_\beta\,(1+\|\boldsymbol{x}\|)^{m_\beta}\left|\partial^{\alpha-\beta}\varphi_k(\boldsymbol{x}) - \partial^{\alpha-\beta}\varphi(\boldsymbol{x})\right| \\
&= \sum_{\beta\le\alpha} c_{\alpha,\beta}\,c_\beta\,\|\boldsymbol{x}\|^m \sum_{j=0}^{m_\beta}\binom{m_\beta}{j}\|\boldsymbol{x}\|^j\left|\partial^{\alpha-\beta}\varphi_k(\boldsymbol{x}) - \partial^{\alpha-\beta}\varphi(\boldsymbol{x})\right| \\
&= \sum_{\beta\le\alpha} c_{\alpha,\beta}\,c_\beta\sum_{j=0}^{m_\beta}\binom{m_\beta}{j}\|\boldsymbol{x}\|^{m+j}\left|\partial^{\alpha-\beta}\varphi_k(\boldsymbol{x}) - \partial^{\alpha-\beta}\varphi(\boldsymbol{x})\right| \xrightarrow{k\to\infty} 0
\end{aligned}$$

uniformly, because the sum is finite and all the functions are bounded, and therefore we have a linear combination of terms $\|\boldsymbol{x}\|^{m+j}\left|\partial^{\alpha-\beta}\varphi_k(\boldsymbol{x}) - \partial^{\alpha-\beta}\varphi(\boldsymbol{x})\right|$ which go to zero uniformly from the convergence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ of the sequence $(\varphi_k)_{k\in\mathbb{N}}$.

Exercises

Exercise 6.21 Prove that $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is a vector space. [Solution: (6.21)]

Exercise 6.22 Consider the function

$$\begin{aligned}
\varphi : \mathbb{R}^n &\longrightarrow \mathbb{R} \\
\boldsymbol{x} &\mapsto e^{-\|\boldsymbol{x}\|^2}.
\end{aligned}$$

Show that for all $\boldsymbol{\alpha}\in\mathbb{N}^n$ there exists a polynomial $p(\boldsymbol{x})$ of degree $|\boldsymbol{\alpha}|$ satisfying

$$\partial^\alpha\varphi(\boldsymbol{x}) = p(\boldsymbol{x})\,\varphi(\boldsymbol{x}).$$

[Solution: (6.22)]

Exercise 6.23 Consider the function $\phi : \mathbb{R}\longrightarrow\mathbb{R}$ defined by

$$\phi(x) = e^{-\frac{1}{x^2}-x^2}\,\chi_{(0,1)}(x).$$

(i) Prove that $\phi(x)\in\mathcal{S}(\mathbb{R};\mathbb{R})$;

(ii) Show that, given $a,b\in\mathbb{R}$ with $a<b$, the function $\varphi:\mathbb{R}\longrightarrow\mathbb{R}$ defined by

$$\varphi(x) = \phi(x-a)\,\phi(b-x)$$

is in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

[Solution: (6.23)]

6.6 Tempered distributions

The space of tempered distributions is the dual of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$: it is a set of distributions for which we can define an object in the form

$$\int_{\mathbb{R}^n} T\,\varphi,$$

for every $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Its elements are also called *slowly increasing distributions*. These is a particular class of distributions which, at infinity, has a more controlled behaviour than the generic distributions: this allows more generality in the choice of test functions. Not all functionals arising from functions in $L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{C})$ are tempered. For example, in the case $n = 1$ the exponential function is not tempered, because its order of infinity, as $x \to +\infty$, is greater than any other power of x.

Definition 6.5 [Tempered Distribution]
A tempered distribution is a linear functional $T : \mathcal{S}(\mathbb{R}^n;\mathbb{C}) \longrightarrow \mathbb{C}$ such that, if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then

$$\lim_{k\to\infty} \langle T, \varphi_k\rangle = \langle T, \varphi\rangle.$$

We denote by $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ the set of tempered distributions $T : \mathcal{S}(\mathbb{R}^n;\mathbb{C}) \longrightarrow \mathbb{C}$.

Hence, a tempered distribution is a linear functional on $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ which is continuous with respect to the convergence on $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ defined in the previous section.

Theorem 6.16 *Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Then,*

$$T|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})} \in \mathcal{D}'(\mathbb{R}^n;\mathbb{C}).$$

Proof Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ and let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$. If $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, then $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Moreover, since $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, we have $\langle T, \varphi_k\rangle \xrightarrow{k\to\infty} \langle T, \varphi\rangle$, and this can be written also as

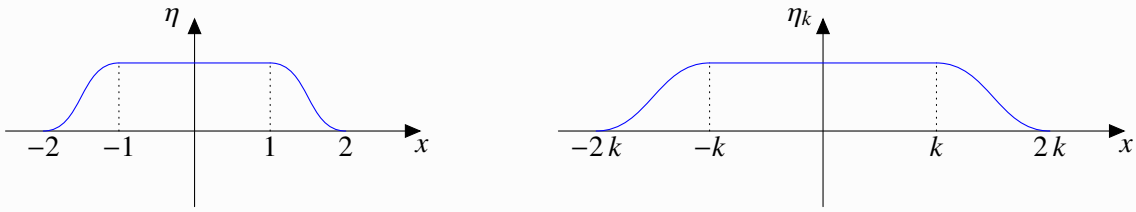
$$\langle T|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})}, \varphi_k\rangle \xrightarrow{k\to\infty} \langle T|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})}, \varphi\rangle.$$

We now want to show that, stated loosely, the set of tempered distributions can be considered as a subset of the set of distributions. This is not surprising: after all, distributions are linear continuous functionals over $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, and tempered distributions are linear continuous functionals over $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Since $\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subset \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, a tempered distribution can also be considered a linear continuous functional over $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$. However, to make this claim more precise, we must ensure that there are no different tempered distributions that act in the same way over all the elements of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$. Stated more precisely, if two tempered distributions have the same restriction to $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, they must coincide. To prove this fact, we first present the following important result.

Theorem 6.17 [Density of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$]
Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. There exists a sequence $(\varphi_k)_{k\in\mathbb{N}}$ of elements in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ that converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

Proof Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Let $\eta \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ be such that $\eta(\boldsymbol{x}) = 1$ when $\|\boldsymbol{x}\| \leq 1$ and $\eta(\boldsymbol{x}) = 0$ when $\|\boldsymbol{x}\| \geq 2$, so that $\operatorname{supp}\eta \subseteq \mathcal{B}_2(0)$.

Let us now define the sequence $(\eta_k)_{k\in\mathbb{N}}$ of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ in this way: $\eta_k(t) \stackrel{\text{def}}{=} \eta(\frac{\boldsymbol{x}}{k})$, for $k \in \mathbb{N}^*$, and therefore $\operatorname{supp}\eta_k \subseteq \overline{\mathcal{B}_{2\,k}(0)}$.



As k goes to infinity, $\eta_k(\boldsymbol{x})$ tends to the constant function 1. Now consider $\varphi_k(\boldsymbol{x}) \stackrel{\text{def}}{=} \varphi(\boldsymbol{x})\,\eta_k(\boldsymbol{x})$; the sequence $(\varphi_k)_{k\in\mathbb{N}}$ is in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ for all $k \in \mathbb{N}$. We want to prove that the limit of this sequence is φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, i.e., that for all $m \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{N}^n$, we have $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}(\varphi_k - \varphi)(\boldsymbol{x})| \xrightarrow[k\to\infty]{} 0$ uniformly in $\mathbb{R}^n$. If $\|\boldsymbol{x}\| < k$, we have $\varphi_k(\boldsymbol{x}) = \varphi(\boldsymbol{x})$, so that $\partial^{\alpha}\varphi_k = \partial^{\alpha}\varphi$ for any $\boldsymbol{\alpha}$. It follows that $\partial^{\alpha}(\varphi_k - \varphi)(\boldsymbol{x}) = 0$ for $\|\boldsymbol{x}\| < k$. Therefore we only need to prove the convergence to zero of the elements $\boldsymbol{x} \in \mathbb{R}^n$ for which $\|\boldsymbol{x}\| \geq k$. Let $\|\boldsymbol{x}\| \geq k$. We find

$$\|\boldsymbol{x}\|^m\,|\partial^{\alpha}(\varphi_k - \varphi)(\boldsymbol{x})| \leq \|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})| + \|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi(\boldsymbol{x})|.$$

We now show that the right-hand side of the latter goes to zero uniformly as $k \to \infty$.

Consider the term $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi(\boldsymbol{x})|$: we know that $\partial^{\alpha}\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, so that, by definition, there exists $c_{m+1} > 0$ such that $\|\boldsymbol{x}\|^{m+1}\,|\partial^{\alpha}\varphi(\boldsymbol{x})| \leq c_{m+1}$ for all $\boldsymbol{x} \in \mathbb{R}^n$, and in particular for all $\boldsymbol{x} \in \mathbb{R}^n$ such that $\|\boldsymbol{x}\| \geq k$, from which we find

$$\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi(\boldsymbol{x})| \;=\; \|\boldsymbol{x}\|^{-1}\,\|\boldsymbol{x}\|^{m+1}|\partial^{\alpha}\varphi(\boldsymbol{x})| \leq \frac{1}{k}c_{m+1}.$$

Therefore, given $\varepsilon > 0$, there exists $k_{1,\varepsilon} \in \mathbb{N}$ such that if $k > k_{1,\varepsilon}$, we have $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi(t)| \leq \varepsilon/2$: take $k_{1,\varepsilon} \geq 2\,c_{m+1}/\varepsilon$.

Let us now consider the term $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})|$. We recall that $\eta_k(\boldsymbol{x}) = \eta\left(\frac{\boldsymbol{x}}{k}\right)$ and $\eta_k(\boldsymbol{x}) = 0$ if $\|\boldsymbol{x}\| \geq 2\,k$. For all $\boldsymbol{\gamma} \in \mathbb{N}$, by virtue of the derivative of the composition $\partial^{\gamma}\eta_k(\boldsymbol{x}) = k^{-|\gamma|}(\partial^{\gamma}\eta)\left(\frac{\boldsymbol{x}}{k}\right)$ for any multi-index $\boldsymbol{\gamma} \in \mathbb{N}^n$; therefore, it suffices to consider $k \leq \|\boldsymbol{x}\| \leq 2\,k$. We use these considerations on the term $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})|$:

$$\begin{aligned}
\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})| \;&=\; \|\boldsymbol{x}\|^m\,|\partial^{\alpha}(\eta_k\,\varphi)(\boldsymbol{x})| \\
&=\; \|\boldsymbol{x}\|^m\,\Big|\sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} c_{\alpha,\beta}\,\partial^{\alpha-\beta}\eta_k(\boldsymbol{x})\,\partial^{\beta}\varphi(\boldsymbol{x})\Big| \\
&=\; \|\boldsymbol{x}\|^m\,\Big|\sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} c_{\alpha,\beta}\,k^{-|\alpha-\beta|}(\partial^{\alpha-\beta}\eta)\left(\frac{\boldsymbol{x}}{k}\right)\partial^{\beta}\varphi(\boldsymbol{x})\Big| \\
&\leq\; \sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} c_{\alpha,\beta}\,\sup_{\overline{\mathcal{B}_2(0)}}|\partial^{\alpha-\beta}\eta|\,\|\boldsymbol{x}\|^m\,|\partial^{\beta}\varphi(\boldsymbol{x})|,
\end{aligned}$$

where we used the fact that $k^{-|\alpha-\beta|} \leq 1$ since the power is negative. Now, for all $\boldsymbol{\beta} \in \mathbb{N}^n$, we know that $\partial^{\beta}\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, from which we find $\|\boldsymbol{x}\|^{m+1}\,|\partial^{\beta}\varphi(\boldsymbol{x})| \leq c_{m+1,\beta}$. Thus, $\|\boldsymbol{x}\|^m\,|\partial^{\beta}\varphi(\boldsymbol{x})| \leq c_{m+1,\beta}\,\|\boldsymbol{x}\|^{-1} \leq (1/k)\,c_{m+1,\beta}$ since it is enough to consider only $\|\boldsymbol{x}\| \in \mathbb{R}^n$ such that $k \leq \|\boldsymbol{x}\| \leq 2\,k$. We obtain

$$\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})| \;\leq\; \sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}} c_{\alpha,\beta}\,\sup_{\overline{\mathcal{B}_2(0)}}|\partial^{\alpha-\beta}\eta|\,c_{m+1,\beta}\,\frac{1}{k}.$$

This implies that, given $\varepsilon > 0$, there exists $k_{2,\varepsilon} \in \mathbb{N}$ such that if $k > k_{2,\varepsilon}$ we have $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}\varphi_k(\boldsymbol{x})| \leq \varepsilon/2$ for all $\boldsymbol{x} \in \mathbb{R}^n$. By choosing $\tilde{k} = \max\{k_{1,\varepsilon}\,k_{2,\varepsilon}\}$, for all $\varepsilon > 0$ if $k > \tilde{k}$ we have $\|\boldsymbol{x}\|^m\,|\partial^{\alpha}(\varphi_k - \varphi)(\boldsymbol{x})| \leq \varepsilon$ for all $\boldsymbol{x} \in \mathbb{R}^n$.

We are now ready to prove the following important result, which shows that the function that maps a tempered distribution to its restriction to $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ is injective. Hence, as already explained, the restriction of a tempered distribution to $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ uniquely determines the tempered distribution itself, and this is what allows us to identify the set of tempered distributions with a subset of the set of distributions.

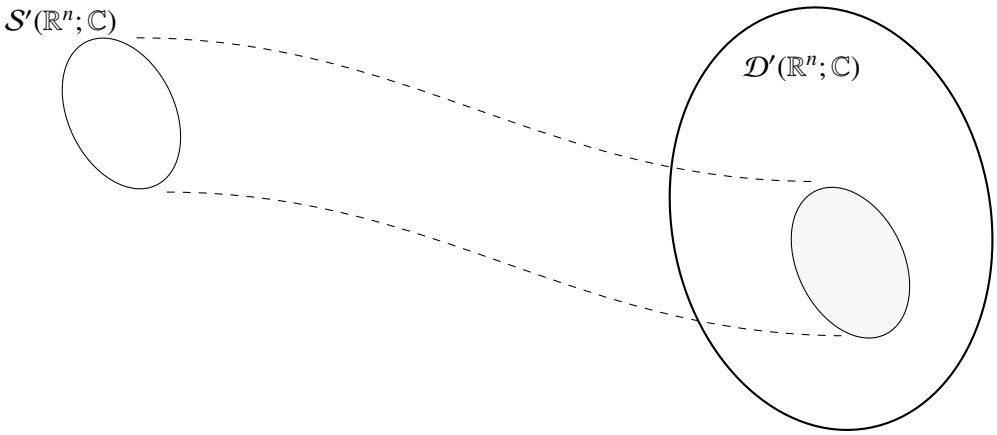
Theorem 6.18 *Let $T_1, T_2 \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ such that $T_1|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})} = T_2|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})}$. Then, $T_1 = T_2$.*

Proof To prove that $T_1 = T_2$, we have to show that for every $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ we have $\langle T_1, \varphi\rangle = \langle T_2, \varphi\rangle$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. In view of Theorem 6.17, there exists a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ that converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. On the other hand, since $T_1|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})} = T_2|_{\mathcal{D}(\mathbb{R}^n;\mathbb{C})}$, for all $k \in \mathbb{N}$ we have $\langle T_1, \varphi_k\rangle = \langle T_2, \varphi_k\rangle$. Since T_1 and T_2 are tempered distributions, we have

$$\lim_{k\to\infty} \langle T_1,\, \varphi_k\rangle = \langle T_1,\, \varphi\rangle \qquad \text{and} \qquad \lim_{k\to\infty} \langle T_2,\, \varphi_k\rangle = \langle T_2,\, \varphi\rangle.$$

Thus, we have $\langle T_1,\, \varphi\rangle = \langle T_2,\, \varphi\rangle$ from the uniqueness of the limit of a sequence of real or complex numbers.

Tempered distributions are functionals over $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and since $\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subset \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, they are also functionals over $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, just like the elements of $\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. Moreover, each tempered distribution is unambiguously identified by the way it acts on the elements of $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, and the mapping it induces over these elements is a distribution that lives in $\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. This says that the set of tempered distributions is not larger than $\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. It remains to prove that there are distributions that are not tempered distributions, and that will show that $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ can be considered a subset of $\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$.



In general, the functional L_f induced by a function $f \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{C})$ is not a tempered distribution. The following theorem shows that if $f \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{C})$, and there exists $m \in \mathbb{N}$ such that $\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m} \in L^1(\mathbb{R}^n;\mathbb{C})$, then f is tempered.

Theorem 6.19 *Let $f \in L^1_{\text{loc}}(\mathbb{R}^n;\mathbb{C})$. Let $m \in \mathbb{N}$ be such that $\int_{\mathbb{R}^n} \frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}\,\mathrm{d}\boldsymbol{x} < +\infty$. Then:*

(i) $f\varphi \in L^1(\mathbb{R}^n;\mathbb{C})$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$;

(ii) the functional

$$\begin{aligned} \mathcal{S}(\mathbb{R}^n;\mathbb{C}) &\longrightarrow \mathbb{C} \\ \varphi &\mapsto \int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \end{aligned}$$

is in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$.

Proof We begin proving that $f\,\varphi$ is integrable. We have

$$\begin{aligned} |f(\boldsymbol{x})\,\varphi(\boldsymbol{x})| &= \frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}(1+\|\boldsymbol{x}\|)^m\,|\varphi(\boldsymbol{x})| \\ &= \frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}\sum_{j=0}^{m}\binom{m}{j}\|\boldsymbol{x}\|^j\,|\varphi(\boldsymbol{x})| \le c\,\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m} \end{aligned}$$

for some non-negative constant c, since $\|\boldsymbol{x}\|^j\,|\varphi(\boldsymbol{x})|$ is bounded for each $j \in \{1,\dots,m\}$ from the assumption $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Moreover, $\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}$ is integrable by assumption. This shows that $f\,\varphi \in L^1(\mathbb{R}^n;\mathbb{C})$, and therefore the integral

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$$

is well defined.
We now we prove the second point. Let us now consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then

$$\begin{aligned} \left|\int_{\mathbb{R}^n} f(\boldsymbol{x})\,(\varphi_k(\boldsymbol{x})-\varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}\right| &\le \int_{\mathbb{R}^n}\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}(1+\|\boldsymbol{x}\|)^m\,|\varphi_k(\boldsymbol{x})-\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \\ &\le \sup_{\mathbb{R}^n}\left((1+\|\boldsymbol{x}\|)^m\,|\varphi_k(\boldsymbol{x})-\varphi(\boldsymbol{x})|\right)\int_{\mathbb{R}^n}\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}\,\mathrm{d}\boldsymbol{x} \xrightarrow{k\to\infty} 0 \end{aligned}$$

because $(1+\|\boldsymbol{x}\|)^m\,|\varphi_k(\boldsymbol{x})-\varphi(\boldsymbol{x})|$ tends to 0 uniformly in view of the definition of convergence in $\mathcal{S}$. Therefore, the integral $\int_{\mathbb{R}^n} f(\boldsymbol{x})\,(\varphi_k(\boldsymbol{x})-\varphi(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x}$ goes to zero as $k\to\infty$.

We now show that if $f \in L^p(\mathbb{R}^n;\mathbb{C})$, then f can be extended to a tempered distribution.

Theorem 6.20 *For all* $p \in [1,\infty]$

$$L^p(\mathbb{R}^n;\mathbb{C}) \subseteq \mathcal{S}'(\mathbb{R}^n;\mathbb{C}).$$

Proof Let $p<\infty$. Let $f \in L^p(\mathbb{R}^n;\mathbb{C})$. From the previous result, the functional associated with f is a tempered distribution if there exists $m\in\mathbb{N}$ such that $\frac{f(\boldsymbol{x})}{(1+\|\boldsymbol{x}\|)^m}$ is summable. From the inequality $|f(\boldsymbol{x})| \le 1+|f(\boldsymbol{x})|^p$ (which comes from the inequality $x \le \eta + \eta^{1-p}\,x^p$ for all $x\ge 0$ and $\eta>0$ proved in Lemma 5.3, by considering $\eta=1$ and $|f(\boldsymbol{x})|$ in place of x), we find

$$\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m} \le \frac{1}{(1+\|\boldsymbol{x}\|)^m} + \frac{|f(\boldsymbol{x})|^p}{(1+\|\boldsymbol{x}\|)^m} \le \frac{1}{(1+\|\boldsymbol{x}\|)^m} + |f(\boldsymbol{x})|^p,$$

where $|f(\boldsymbol{x})|^p$ is integrable and $(1+\|\boldsymbol{x}\|)^{-m}$ is integrable if $m>n$. Thus, we found a condition for m which ensures that $\int_{\mathbb{R}^n}\frac{|f(\boldsymbol{x})|}{(1+\|\boldsymbol{x}\|)^m}\,\mathrm{d}\boldsymbol{x} < +\infty$, which is what we needed to ensure that the corresponding functional is a tempered distribution.

Example 6.14

We now introduce an example of a distribution that is not tempered. Consider $f : \mathbb{R} \longrightarrow \mathbb{R}$ be the exponential function $f(x) = e^x$. We show that $L_f \notin \mathcal{S}'(\mathbb{R};\mathbb{R})$. Indeed, if by contradiction we assume that L_f is tempered, there should exists a functional $U : \mathcal{S}(\mathbb{R};\mathbb{R}) \longrightarrow \mathbb{R}$ which is linear and continuous with respect to the notion of convergence in $\mathcal{S}$, and such that $L_f = U|_{\mathcal{D}(\mathbb{R};\mathbb{R})}$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Let

$$\langle L_f, \varphi \rangle = \int_{\mathbb{R}} e^x \, \varphi(x) \, \mathrm{d}x.$$

Analogously to what was done in in Remark 6.4, but now in the scalar case, we construct a sequence $(\varphi_k)_{k\in\mathbb{N}^*}$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$ which converges in $\mathcal{S}(\mathbb{R};\mathbb{R})$ but not in $\mathcal{D}(\mathbb{R};\mathbb{R})$. Let $\eta \in \mathcal{D}(\mathbb{R};\mathbb{R})$ be defined by

$$\eta(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 & |x| \geq 2 \end{cases}$$

and $\eta(x) \geq 0$ for all $x \in \mathbb{R}$. Let $\varphi_k(x) = 2^{-k}\,\eta\left(\frac{x}{k}\right)$. We have already noticed that the sequence $(\varphi_k)_{k\in\mathbb{N}^*}$ does not converge in $\mathcal{D}(\mathbb{R};\mathbb{R})$ because its support increases monotonically with k. However, we have already shown in Remark 6.4 that this sequence converges in $\mathcal{S}(\mathbb{R};\mathbb{R})$. Then, by definition of tempered distributions, L_f is tempered if

$$\lim_{k\to\infty} \langle L_f, \varphi_k \rangle = \langle L_f, \varphi \rangle = 0.$$

Using the fact that $\varphi_k(x) = 2^{-k}$ for every x in the interval $[-k, k]$, we find

$$\begin{aligned} \langle L_f, \varphi_k \rangle &= \int_{\mathbb{R}} e^x \, \varphi_k(x) \, \mathrm{d}x \\ &\geq \int_{-k}^{k} e^x \, \varphi_k(x) \, \mathrm{d}x \\ &= 2^{-k} \int_{-k}^{k} e^x \, \mathrm{d}x = 2^{-k}(e^k - e^{-k}), \end{aligned}$$

which does not converge to zero for $k \to \infty$ since $e > 2$. Therefore, L_f is not tempered.

We now want to show that the distributions with compact support are tempered. Since δ has been shown to have compact support, in particular this fact will imply that δ, along with its derivatives, is tempered. We recall that we denoted by $\mathcal{E}'(\Omega;\mathbb{C})$ the set of distributions with compact support on an open set $\Omega \subseteq \mathbb{R}^n$. Take $T \in \mathcal{E}'(\mathbb{R}^n;\mathbb{C})$ and a function $\zeta \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ satisfying $\zeta(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \boldsymbol{U}$ where $\boldsymbol{U}$ is an open set containing the support of T, i.e., $\operatorname{supp} T \subseteq \boldsymbol{U}$. Given $\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C})$, we have defined

$$\langle T, \varphi \rangle = \langle T, \zeta\,\varphi \rangle.$$

Moreover, if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence of functions in $C^\infty(\mathbb{R}^n;\mathbb{C})$ which converges to $\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C})$, then $\langle T, \varphi_k \rangle \xrightarrow{k\to\infty} \langle T, \varphi \rangle$. Since $\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subset C^\infty(\mathbb{R}^n;\mathbb{C})$, this extension is also valid when $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Therefore, to verify that T is tempered it is necessary to prove that the convergence of the sequence $(\varphi_k)_{k\in\mathbb{N}}$ of elements of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ implies the convergence in $C^\infty(\mathbb{R}^n;\mathbb{C})$, i.e., that for every compact set $K \subseteq \mathbb{R}^n$ and for every $\boldsymbol{\alpha} \in \mathbb{N}^n$ we have $\partial^{\boldsymbol{\alpha}}\varphi_k \xrightarrow{k\to\infty} \partial^{\boldsymbol{\alpha}}\varphi$ uniformly on K. This is obvious to see, because the convergence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ ensures that $\partial^{\boldsymbol{\alpha}}\varphi_k \xrightarrow{k\to\infty} \partial^{\boldsymbol{\alpha}}\varphi$ uniformly on $\mathbb{R}^n$, and therefore also on every compact subset of $\mathbb{R}^n$.

Therefore, distributions with compact support are tempered. We now briefly examine the operations on tempered distributions. As aforementioned, in particular $\delta \in \mathcal{E}'(\mathbb{R}^n;\mathbb{R})$ is tempered, i.e., $\delta \in \mathcal{S}'(\mathbb{R}^n;\mathbb{R})$.

Definition 6.6 [Multiplication with functions in $\mathcal{O}_M$]
Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ and $\psi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We define

$$\langle \psi\, T, \varphi\rangle \stackrel{def}{=} \langle T, \psi\,\varphi\rangle.$$

This definition is well posed because $\psi\, T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Indeed, if $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, we have defined $\langle \psi\, T, \varphi\rangle = \langle T, \psi\,\varphi\rangle$. This definition continues to hold even if $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ because $\psi\,\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then

$$\lim_{k\to\infty} \langle \psi\, T, \varphi_k\rangle = \lim_{k\to\infty} \langle T, \psi\,\varphi_k\rangle = \langle T, \psi\,\varphi\rangle = \langle \psi\, T, \varphi\rangle,$$

where we used the fact that $\psi\,\varphi_k \xrightarrow{k\to\infty} \psi\,\varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, see Theorem 6.15. This shows that $\psi\, T$ is indeed tempered.

We now consider the composition of a tempered distribution with C^∞-diffeomorphisms in $\mathbb{R}^n$. In this context, we have restrictions on the types of diffeomorphisms that we can take into account. In particular, we can consider the following cases.

- Let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be an invertible linear transformation, and let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Then $T \circ A$ is tempered. In the case $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, we already know that $\langle T\circ A, \varphi\rangle = \langle T, \varphi\circ A^{-1}\rangle |\det A|^{-1}$. In order to extend this operation to $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, we need to verify two properties:

  (i) if $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $\varphi\circ A^{-1} \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Indeed, given $m \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{N}^n$, we need to show that $\|\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}}(\varphi\circ A^{-1})(\boldsymbol{x})| = \|\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})|$ is bounded. Recalling that an invertible linear transformation are continuous and are bounded by a so-called Lipschitz constant $L \geq 0$ (i.e., for all $\boldsymbol{y} \in \mathbb{R}^n$, we have $\|A\,\boldsymbol{y}\| \leq L\,\boldsymbol{y}$, see Theorem 2.17), we have $\|\boldsymbol{x}\| = \|A\,(A^{-1}\,\boldsymbol{x})\| \leq L\,\|A^{-1}\,\boldsymbol{x}\|$, so that

  $$\|\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})| \leq L^m\, \|A^{-1}\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})|,$$

  and since $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ the quantity $\|A^{-1}\boldsymbol{x}\|^m\, |\partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})|$ is bounded (simply write $\boldsymbol{y}$ in place of $A^{-1}\boldsymbol{x}$).

  (ii) If $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ such that $\lim_{k\to\infty} \varphi_k = \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $\lim_{k\to\infty} \varphi_k\circ A^{-1} = \varphi\circ A^{-1}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Indeed, given $m \in \mathbb{N}$ and $\boldsymbol{\alpha} \in \mathbb{N}^n$, using the same argument we find

  $$\|\boldsymbol{x}\|^m \left|\partial^{\boldsymbol{\alpha}}\varphi_k(A^{-1}\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})\right| \leq L^m\, \|A^{-1}\,\boldsymbol{x}\|^m \left|\partial^{\boldsymbol{\alpha}}\varphi_k(A^{-1}\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})\right|,$$

  which converges to zero uniformly in $\mathbb{R}^n$ since $\|A^{-1}\,\boldsymbol{x}\|^m \left|\partial^{\boldsymbol{\alpha}}\varphi_k(A^{-1}\boldsymbol{x}) - \partial^{\boldsymbol{\alpha}}\varphi(A^{-1}\boldsymbol{x})\right|$ goes to zero uniformly in $\mathbb{R}^n$ (and this is due to the convergence of the sequence $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$).

  Thus, it makes sense to define $\langle T\circ A, \varphi\rangle = |\det A|^{-1}\langle T, \varphi\circ A^{-1}\rangle$, and indeed $T\circ A \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ since $\langle T\circ A, \varphi_k\rangle = |\det A|^{-1}\langle T, \varphi_k\circ A^{-1}\rangle \xrightarrow{k\to\infty} |\det A|^{-1}\langle T, \varphi\circ A^{-1}\rangle = \langle T\circ A, \varphi\rangle$.

  Now, given $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ and an invertible linear transformation $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, we define $T\circ A$ by defining, for all $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$

  $$\langle T\circ A, \varphi\rangle \stackrel{\text{def}}{=} \langle T, \varphi\circ A^{-1}\rangle |\det A|^{-1}.$$

- The second case is the case of translations. Let $\boldsymbol{y} \in \mathbb{R}^n$. We define

$$\begin{aligned} \Phi : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ \boldsymbol{x} &\mapsto \boldsymbol{x} - \boldsymbol{y}. \end{aligned}$$

The function Φ is a diffeomorphism. Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ and $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We define

$$\langle T(\boldsymbol{x}-\boldsymbol{y}), \varphi(\boldsymbol{x})\rangle \stackrel{\text{def}}{=} \langle T(\boldsymbol{x}), \varphi(\boldsymbol{x}+\boldsymbol{y})\rangle.$$

To make sure that this definition is well posed, we need to check that the function

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{C} \\ \boldsymbol{x} &\mapsto \varphi(\boldsymbol{x}+\boldsymbol{y}) \end{aligned}$$

is in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. In other words, we need to check that for all $m \in \mathbb{N}$ and for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ the quantity $\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}+\boldsymbol{y})|$ is bounded. First,

$$\|\boldsymbol{x}\|^m = \|(\boldsymbol{x}+\boldsymbol{y})-\boldsymbol{y}\|^m \leq (\|\boldsymbol{x}+\boldsymbol{y}\| + \|\boldsymbol{y}\|)^m \leq \sum_{j=0}^{m} \binom{m}{j} \|\boldsymbol{y}\|^{m-j} \, \|\boldsymbol{x}+\boldsymbol{y}\|^j,$$

which can be used in

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}+\boldsymbol{y})| \leq \sum_{j=0}^{m} \binom{m}{j} \|\boldsymbol{y}\|^{m-j} \, \|\boldsymbol{x}+\boldsymbol{y}\|^j \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}+\boldsymbol{y})| = \sum_{j=0}^{m} \binom{m}{j} \|\boldsymbol{y}\|^{m-j} \, \|\boldsymbol{x}+\boldsymbol{y}\|^j \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x}+\boldsymbol{y})|,$$

which is bounded because $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

It follows that

$$\langle T(\boldsymbol{x}-\boldsymbol{y}), \varphi_k(\boldsymbol{x})\rangle = \langle T(\boldsymbol{x}), \varphi_k(\boldsymbol{x}+\boldsymbol{y})\rangle \xrightarrow{k\to\infty} \langle T(\boldsymbol{x}), \varphi(\boldsymbol{x}+\boldsymbol{y})\rangle = \langle T(\boldsymbol{x}-\boldsymbol{y}), \varphi(\boldsymbol{x})\rangle,$$

which guarantees that $T(\cdot - \boldsymbol{y}) \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$.

- Consider $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Then, for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ we have $\partial^{\boldsymbol{\alpha}} T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. In the case $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$, we defined $\langle \partial^{\boldsymbol{\alpha}} T, \varphi\rangle = (-1)^{|\boldsymbol{\alpha}|}\langle T, \partial^{\boldsymbol{\alpha}}\varphi\rangle$. The extension to the case $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is straightforward, because $\partial^{\boldsymbol{\alpha}}\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and if $\lim_{k\to\infty} \varphi_k = \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ we have $\lim_{k\to\infty} \partial^{\boldsymbol{\alpha}}\varphi_k = \partial^{\boldsymbol{\alpha}}\varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

6.7 Fourier transform of tempered distributions

First, we study how the Fourier transform interacts with the Schwartz space $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. First, we recall that, from Theorem 6.12, for all $p \in [1,+\infty]$ the inclusion

$$\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subseteq L^p(\mathbb{R}^n;\mathbb{C})$$

holds true. Thus, in particular, $\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subseteq L^1(\mathbb{R}^n;\mathbb{C})$. This means that we can compute the Fourier transform of Schwartz functions: if $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $\hat{\varphi}$ is defined. The next important result, Theorem 6.21, will show that the Fourier transform $\hat{\varphi}$ is still in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Before we present this result, however, we

want to present two formulae for the interaction of linear differential operators with the Fourier transform.

Given a multivariate polynomial

$$Q(\xi) = \sum_{|\alpha| \leq m} a_\alpha \, \xi^\alpha,$$

where $m \in \mathbb{N}$ and where, for every multi-index $\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq m$ the coefficient a_α is complex, we have seen that $Q(\partial)\,\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ (or, in other words, the Schwartz space $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is closed under the action of linear differential operators). It follows that we can also take the Fourier transform of $Q(\partial)\,\varphi$, and using the linearity of the Fourier transform

$$\widehat{Q(\partial)\,\varphi}\,(\boldsymbol{\xi}) = \sum_{|\alpha| \leq m} a_\alpha \, \widehat{\partial_x^\alpha \varphi}\,(\boldsymbol{\xi}) = \sum_{|\alpha| \leq m} a_\alpha \, (i\,\boldsymbol{\xi})^\alpha \, \hat{\varphi}(\boldsymbol{\xi}) = Q(i\,\boldsymbol{\xi})\,\hat{\varphi}(\boldsymbol{\xi}). \tag{6.17}$$

The second formula consists in applying $Q(\partial)$ to $\hat{\varphi}$:

$$Q(\partial)\,\hat{\varphi}(\boldsymbol{\xi}) = \sum_{|\alpha| \leq m} a_\alpha \, \partial_\xi^\alpha \hat{\varphi} = \sum_{|\alpha| \leq m} a_\alpha \, \mathcal{F}\left\{(-i\,\boldsymbol{x})^\alpha \, \varphi\right\} = \mathcal{F}\left\{\sum_{|\alpha| \leq m} a_\alpha \, (-i\,\boldsymbol{x})^\alpha \, \varphi\right\} = \mathcal{F}\left\{Q(-i\,x)\,\varphi\right\}, \tag{6.18}$$

where we used the linearity of the Fourier transform. To summarise, we obtained the formulae

$$\boxed{\begin{aligned} \widehat{Q(\partial)\,\varphi} &= Q(i\,\boldsymbol{\xi})\,\hat{\varphi} \\ Q(\partial)\,\hat{\varphi} &= \widehat{Q(-i\,\boldsymbol{x})\,\varphi} \end{aligned}}$$

Example *6.15*

We want to find the Fourier transform of $\nabla^2 \varphi$ using the previous result. We recall that the polynomial associated with the Laplacian operator is

$$Q(\xi_1, \xi_2, \ldots, \xi_n) = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2.$$

Using (6.17) gives

$$\widehat{\nabla^2 \varphi}(\boldsymbol{\xi}) = ((i\,\xi_1)^2 + (i\,\xi_2)^2 + \ldots + (i\,\xi_n)^2)\,\hat{\varphi}(\boldsymbol{\xi}) = -\|\boldsymbol{\xi}\|^2 \, \hat{\varphi}(\boldsymbol{\xi}).$$

Example *6.16*

Let $m \in \mathbb{N}$. We want to show that the Fourier transform of $\Delta^m \varphi = (\nabla^2)^m \, \varphi$ is $(-\|\boldsymbol{\xi}\|^2)^m \, \hat{\varphi}$. First, the polynomial associated with this differential operator is

$$Q(\xi_1, \xi_2, \ldots, \xi_n) = (\xi_1^2 + \xi_2^2 + \ldots + \xi_n^2)^m.$$

From the multinomial theorem[a] and using the linearity of the Fourier transform

$$\begin{aligned} \widehat{\Delta^m \varphi} &= \mathcal{F}\left\{(\partial_1^2 + \ldots + \partial_n^2)^m \varphi\right\} \\ &= \mathcal{F}\left\{\sum_{k_1+k_2+\ldots+k_n \leq m} \binom{m}{k_1, k_2, \ldots, k_n} \partial_1^{2\,k_1} \, \partial_2^{2\,k_2} \, \cdots \, \partial_n^{2\,k_n} \, \varphi\right\} \\ &= \sum_{k_1+k_2+\ldots+k_n \leq m} \binom{m}{k_1, k_2, \ldots, k_n} \mathcal{F}\left\{\partial_1^{2\,k_1} \, \partial_2^{2\,k_2} \, \cdots \, \partial_n^{2\,k_n} \, \varphi\right\} \\ &= \sum_{k_1+k_2+\ldots+k_n \leq m} \binom{m}{k_1, k_2, \ldots, k_n} \left((i\,\xi_1)^{2\,k_1} \, \cdots \, (i\,\xi_n)^{2\,k_n} \, \hat{\varphi}\right) \\ &= \sum_{k_1+k_2+\ldots+k_n \leq m} \binom{m}{k_1, k_2, \ldots, k_n} \left((-\xi_1^2)^{k_1} \, \cdots \, (-\xi_n^2)^{k_n}\right) \hat{\varphi} \\ &= (-\xi_1^2 - \ldots - \xi_n^2)^m \, \hat{\varphi} = (-\|\boldsymbol{\xi}\|^2)^m \, \hat{\varphi} \end{aligned}$$

[a] The multinomial theorem is a generalization of Newton's binomial formula:

$$(x_1+x_2+\ldots+x_n)^m = \sum_{k_1+k_2+\ldots+k_n \leq m} \binom{m}{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} \quad \text{where} \quad \binom{m}{k_1, k_2, \ldots, k_n} = \frac{n!}{k_1! \cdot k_2! \cdot \ldots \cdot k_n!}.$$

Example 6.17

Let $m \in \mathbb{N}$. We want to show that the Fourier transform of $(1-\Delta)^m \varphi$ is $(1+\|\boldsymbol{\xi}\|^2)^m \hat{\varphi}$. First, the polynomial associated with this differential operator is

$$Q(\xi_1, \xi_2, \ldots, \xi_n) = (1 - \xi_1^2 - \xi_2^2 - \ldots - \xi_n^2)^m.$$

Using (6.17) gives

$$\begin{aligned}
\widehat{(1-\Delta)^m \varphi}(\boldsymbol{\xi}) &= \mathcal{F}\left\{(1-\Delta)^m \varphi\right\} \\
&= \mathcal{F}\left\{\sum_{j=0}^{m} \binom{m}{j} (-\Delta^j) \varphi\right\} \\
&= \sum_{j=0}^{m} \binom{m}{j} \mathcal{F}\left\{(-\Delta^j) \varphi\right\} \\
&= \sum_{j=0}^{m} \binom{m}{j} (\|\boldsymbol{\xi}\|^2)^j \hat{\varphi} = (1+\|\boldsymbol{\xi}\|^2)^m \hat{\varphi}.
\end{aligned}$$

We show that the Fourier transform of an element in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is still in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.21 *Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then, $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$.*

Proof First, we show that $\hat{\varphi}$ is of class C^∞. From Corollary 6.5, if $f \in L^1(\mathbb{R}^n;\mathbb{C})$ and if $(1+\|\boldsymbol{x}\|)^m f \in L^1(\mathbb{R}^n;\mathbb{C})$, then $\hat{f} \in C^m(\mathbb{R}^n;\mathbb{C})$. Therefore, it is sufficient to verify that for all $m \in \mathbb{N}$ we have $(1+\|\boldsymbol{x}\|)^m \varphi \in L^1(\mathbb{R}^n;\mathbb{C})$. Since $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, for all $p \in \mathbb{N}$ we have $\|\boldsymbol{x}\|^p |\varphi(\boldsymbol{x})| \leq c_{p,1}$ for all $\boldsymbol{x} \in \mathbb{R}^n$, and the same is true for $p=0$, so that $|\varphi(\boldsymbol{x})| \leq c_{p,0}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. It follows that

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad (1+\|\boldsymbol{x}\|^p)\,|\varphi(\boldsymbol{x})| \leq c_p, \quad \text{where} \quad c_p = c_{p,0} + c_{p,1}.$$

This implies $|\varphi(\boldsymbol{x})| \leq c_p\,(1+\|\boldsymbol{x}\|^p)^{-1}$. Hence, we can write

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad (1+\|\boldsymbol{x}\|)^m\,|\varphi(\boldsymbol{x})| \leq c_p\,(1+\|\boldsymbol{x}\|)^m\,(1+\|\boldsymbol{x}\|^p)^{-1}.$$

The left hand-side, as $\|\boldsymbol{x}\| \to \infty$, is asymptotically equivalent to $\|\boldsymbol{x}\|^{m-p}$, and therefore we the function is integrable if and only if $m-p<-n$, which means $p>m+n$. In other words, for any $m \in \mathbb{N}$ it is possible to find p for which inequality holds, and this shows that indeed $(1+\|\boldsymbol{x}\|)^m \varphi \in L^1(\mathbb{R}^n;\mathbb{C})$, which in turn implies that $\varphi \in C^m(\mathbb{R}^n;\mathbb{C})$ for every $m \in \mathbb{N}$, i.e., $\varphi \in C^\infty(\mathbb{R}^n;\mathbb{C})$.

We need to show that $\|\boldsymbol{\xi}\|^m\,|\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})|$ is bounded over $\mathbb{R}^n$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$. Again from Corollary 6.5, since we showed that for all $m \in \mathbb{N}$ we have $(1+\|\boldsymbol{x}\|)^m \varphi \in L^1(\mathbb{R}^n;\mathbb{C})$, for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ with $|\boldsymbol{\alpha}| \leq m$

$$\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi}) = \mathcal{F}\big((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi\big)(\boldsymbol{\xi}).$$

From Riemann-Lebesgue Lemma, $\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})$ is bounded over $\mathbb{R}^n$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$. We now prove that $\|\boldsymbol{\xi}\|^m\,|\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})|$

is bounded over $\mathbb{R}^n$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$. From the previous example, $\widehat{\nabla^2 \varphi} = -\|\boldsymbol{\xi}\|^2 \hat{\varphi}$, so that

$$\begin{aligned} \|\boldsymbol{\xi}\|^{2m} |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})| &= \|\boldsymbol{\xi}\|^{2m} |\mathcal{F}((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi)(\boldsymbol{\xi})| \\ &= |\mathcal{F}((\nabla^2)^m ((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi)(\boldsymbol{\xi})| \end{aligned}$$

We observe that $(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \in \mathcal{O}_M(\mathbb{R}^n; \mathbb{C})$, because $(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}$ is a polynomial. Thus, $(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. Therefore, $(\nabla^2)^m ((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi) \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, and this implies $(\nabla^2)^m ((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi) \in L^1(\mathbb{R}^n; \mathbb{C})$. From Riemann-Lebesgue Lemma, the Fourier transform of a function in L^1 is bounded, so that a non-negative constant c exists such that

$$\forall\, \boldsymbol{\xi} \in \mathbb{R}^n, \quad |\mathcal{F}((\nabla^2)^m ((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \varphi)(\boldsymbol{\xi})| \leq c,$$

which implies that $\|\boldsymbol{\xi}\|^{2m} |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})|$ is bounded over $\mathbb{R}^n$ for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$. On the other hand, we needed to prove the boundedness for $\|\boldsymbol{\xi}\|^m |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})|$. We notice that

$$\|\boldsymbol{\xi}\|^m \leq \begin{cases} 1 & \text{if } \|\boldsymbol{\xi}\| \leq 1 \\ \|\boldsymbol{\xi}\|^{2m} & \text{if } \|\boldsymbol{\xi}\| > 1. \end{cases}$$

This allows us to conclude, because

$$\|\boldsymbol{\xi}\|^m |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})| \leq \begin{cases} |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})| & \text{if } \|\boldsymbol{\xi}\| \leq 1 \\ \|\boldsymbol{\xi}\|^{2m} |\partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{\varphi}(\boldsymbol{\xi})| & \text{if } \|\boldsymbol{\xi}\| > 1. \end{cases}$$

Theorem 6.22 *The Fourier transform*

$$\mathcal{F} : \mathcal{S}(\mathbb{R}^n; \mathbb{C}) \longrightarrow \mathcal{S}(\mathbb{R}^n; \mathbb{C}) \tag{6.19}$$

is a bijection.

Proof First, we notice that the linear operator $\mathcal{F}$ can be written as in (6.19) because, if $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, then $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ in view of Theorem 6.21. Since $\mathcal{F}$ is linear, to prove the injectivity we only need to show that its null-space is the origin. In other words, if $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ and $\mathcal{F}\{\varphi\} = 0$, then we must show that $\varphi = 0$. We know that $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ implies $\varphi \in L^1(\mathbb{R}^n; \mathbb{C})$, since, from Theorem 6.12, for every $p \in [1, +\infty]$ the Schwartz space $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$ is contained in $L^p(\mathbb{R}^n; \mathbb{C})$, and therefore it is contained, in particular, in $L^1(\mathbb{R}^n; \mathbb{C})$. Thus, if $\mathcal{F}\{\varphi\} = 0$, we can use the inversion formula in Theorem 6.7; since φ is in $L^1(\mathbb{R}^n; \mathbb{C})$, it is continuous and bounded (because it is also in $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$), and we obtain $\varphi = 0$.

Now, we need to prove that $\mathcal{F}$ is surjective. Let $w \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. We have to prove that there exists $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ such that $\hat{\varphi} = w$. Let us define φ in the following way:

$$\varphi(\boldsymbol{x}) = \frac{1}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, w(\boldsymbol{\xi})\, \mathrm{d}\boldsymbol{\xi}, \qquad \boldsymbol{x} \in \mathbb{R}^n.$$

We want to show that this function satisfies the requirements: $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ and $\hat{\varphi} = w$. The first one is straightforward: we can write

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \varphi(\boldsymbol{x}) = \frac{1}{(2\,\pi)^n}\, \hat{w}(-\boldsymbol{x}). \tag{6.20}$$

Clearly, since $w \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, we have $\hat{w} \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, and therefore also $\hat{w}(-\boldsymbol{x}) \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$ (why? Think about the definition of $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$). It follows that $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. Now, we show that $\hat{\varphi} = w$. First, we introduce the

following notation for a function $g : \mathbb{R}^n \longrightarrow \mathbb{C}$:

$$\check{g}(\boldsymbol{\xi}) = g(-\boldsymbol{\xi}),$$

which is another way of writing $\check{g} = g \circ A$, where $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the linear transformation $A\,\boldsymbol{x} = -\boldsymbol{x}$. With this notation, we can re-write (6.20) as

$$\varphi = \frac{1}{(2\,\pi)^n}\,\check{\hat{w}} \quad \Rightarrow \quad \hat{\varphi} = \frac{1}{(2\,\pi)^n}\,\hat{\check{w}}.$$

We show that we can interchange the operations $\check{}$ and $\hat{}$ using the rule of the Fourier transform of the composition:

$$\hat{\check{g}} = \widehat{g \circ A} = \frac{1}{|\det A|}\,\hat{g} \circ A^{-\top} = \hat{g} \circ A = \check{\hat{g}},$$

since $|\det A| = 1$ and $A^{-\top} = A$. It follows that

$$\hat{\varphi} = \frac{1}{(2\,\pi)^n}\,\check{\hat{w}} = w,$$

where we used the inversion formula (6.12).[a]

[a] We recall that in view of (6.12) we have precisely

$$w(\boldsymbol{x}) = \frac{1}{(2\,\pi)^n}\int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\hat{w}(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \frac{1}{(2\,\pi)^n}\,\hat{\hat{w}}(-\boldsymbol{x}) \quad \forall\,\boldsymbol{x} \in \mathbb{R}^n.$$

The remaining step is to show that the Fourier transform is a continuous operator on $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.23 *Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then,*

$$\varphi_k \xrightarrow{k\to\infty} \varphi \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}) \quad \Leftrightarrow \quad \hat{\varphi}_k \xrightarrow{k\to\infty} \hat{\varphi} \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

Proof We first prove $\Rightarrow$. We need to show that if $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then for all $\boldsymbol{\alpha} \in \mathbb{N}^n$ and for all $m \in \mathbb{N}$,

$$\|\boldsymbol{\xi}\|^m\,|\partial^{\boldsymbol{\alpha}}_{\boldsymbol{\xi}}\hat{\varphi}_k(\boldsymbol{\xi}) - \partial^{\boldsymbol{\alpha}}_{\boldsymbol{\xi}}\hat{\varphi}(\boldsymbol{\xi})| \xrightarrow{k\to\infty} 0$$

uniformly in $\mathbb{R}^n$. First, we observe that for all $\boldsymbol{\xi} \in \mathbb{R}^n$, we have $\|\boldsymbol{\xi}\|^m \le (1 + \|\boldsymbol{\xi}\|^2)^m$. Indeed, if $\|\boldsymbol{\xi}\| \le 1$ we have $\|\boldsymbol{\xi}\|^m \le 1 \le (1 + \|\boldsymbol{\xi}\|^2)^m$, and if $\|\boldsymbol{\xi}\| > 1$ we have $\|\boldsymbol{\xi}\|^m \le \|\boldsymbol{\xi}\|^{2\,m} \le (1 + \|\boldsymbol{\xi}\|^2)^m$. Thus,

$$\begin{aligned}
\|\boldsymbol{\xi}\|^m\,|\partial^{\boldsymbol{\alpha}}_{\boldsymbol{\xi}}\hat{\varphi}_k(\boldsymbol{\xi}) - \partial^{\boldsymbol{\alpha}}_{\boldsymbol{\xi}}\hat{\varphi}(\boldsymbol{\xi})| &= \|\boldsymbol{\xi}\|^m\,\left|\mathcal{F}\left\{(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}\varphi_k\right\} - \mathcal{F}\left\{(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}\varphi\right\}\right| \\
&\le \left|(1 + \|\boldsymbol{\xi}\|^2)^m\,\mathcal{F}\left\{(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}(\varphi_k - \varphi)\right\}\right| \\
&= \left|\mathcal{F}\left\{(1 - \nabla^2_{\boldsymbol{x}})^m\,((-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}(\varphi_k - \varphi))\right\}\right|,
\end{aligned}$$

where we have used the formula $\widehat{Q(\partial)\,\varphi} = Q(i\,\boldsymbol{\xi})\,\hat{\varphi}$ with $Q(\boldsymbol{\xi}) = (1 + \|\boldsymbol{\xi}\|^2)^m$ in (6.17), which gives $Q(i\,\boldsymbol{\xi}) = Q(\boldsymbol{\xi}) = (1 + \|\boldsymbol{\xi}\|^2)^m$ and $Q(\partial) = (1 + \partial^2)^m = (1 + \nabla^2)^m$. Since $(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}} \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$ for every $\boldsymbol{\alpha} \in \mathbb{N}^n$, the multiplication of $(\varphi_k - \varphi)$ by $(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}$ is a continuous operation, and therefore

$$(-i\,\boldsymbol{x})^{\boldsymbol{\alpha}}(\varphi_k - \varphi) \xrightarrow{k\to\infty} 0 \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

Therefore, we have also, from the observation in Remark 6.3

$$(1-\nabla_{\boldsymbol{x}}^2)^m\left((-i\,\boldsymbol{x})^\alpha(\varphi_k-\varphi)\right)\xrightarrow{k\to\infty}0\ \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

Let $w_k(\boldsymbol{x}) = (1-\nabla_{\boldsymbol{x}}^2)^m\left((-i\,\boldsymbol{x})^\alpha(\varphi_k-\varphi)\right)$. We only have to show that if $w_k \xrightarrow{k\to\infty} 0$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $|\hat{w}_k(\boldsymbol{\xi})|\xrightarrow{k\to\infty}0$ uniformly in $\mathbb{R}^n$. From Riemann-Lebesgue Lemma and, in particular, from (6.2), we know that $|\hat{w}_k(\boldsymbol{\xi})|\le\|w_k\|_1$, and therefore it is sufficient to prove that the convergence to zero of w_k in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ implies the convergence to zero of w_k in $L^1(\mathbb{R}^n;\mathbb{C})$. Indeed, since w_k converges to 0 in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $(1+\|\boldsymbol{x}\|^m)\,|w_k(\boldsymbol{x})|$ tends to 0 uniformly, so that

$$\begin{aligned}\|w_k\|_1 &= \int_{\mathbb{R}^n}|w_k(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}=\int_{\mathbb{R}^n}\frac{1}{1+\|\boldsymbol{x}\|^m}(1+\|\boldsymbol{x}\|^m)\,|w_k(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\\ &\le \sup_{\mathbb{R}^n}(1+\|\boldsymbol{x}\|^m)\,|w_k(\boldsymbol{x})|\int_{\mathbb{R}^n}\frac{1}{1+\|\boldsymbol{x}\|^m}\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

If $m>n$, the integral converges and we have the convergence of $\|w_k\|_1$ as $k\to\infty$ as required.

Now we prove the converse: if $(\hat{\varphi}_k)_{k\in\mathbb{N}}$ converges to $\hat{\varphi}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $(\varphi_k)_{k\in\mathbb{N}}$ converges to φ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. This follows by

$$\varphi_k=\frac{1}{(2\,\pi)^n}\,\mathcal{F}\left\{\check{\hat{\varphi}}_k\right\}$$

Since $(\hat{\varphi}_k)_{k\in\mathbb{N}}$ converges to $\hat{\varphi}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, in view of the previous implication $(\hat{\hat{\varphi}}_k)_{k\in\mathbb{N}}$ converges to $\hat{\hat{\varphi}}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and obviously $(\check{\hat{\varphi}}_k)_{k\in\mathbb{N}}$ converges to $\check{\hat{\varphi}}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$.

Definition 6.7 *Let $T\in\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $\varphi\in\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We define*

$$\langle\hat{T},\varphi\rangle\stackrel{def}{=}\langle T,\hat{\varphi}\rangle.$$

The distribution $\hat{T}$ is tempered. Indeed, we know that if $\varphi_k\xrightarrow{k\to\infty}\varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $\hat{\varphi}_k\xrightarrow{k\to\infty}\hat{\varphi}$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, so that

$$\langle\hat{T},\varphi_k\rangle=\langle T,\hat{\varphi}_k\rangle\xrightarrow{k\to\infty}\langle T,\hat{\varphi}\rangle=\langle\hat{T},\varphi\rangle.$$

We observe that this definition is consistent with the definition of Fourier transform for a function in $L^1(\mathbb{R}^n;\mathbb{C})$. Indeed, if $f\in L^1(\mathbb{R}^n;\mathbb{C})$,

$$\hat{f}(\boldsymbol{\xi})=\int_{\mathbb{R}^n}e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Now, if $\varphi\in\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, we find

$$\langle\hat{f},\varphi\rangle=\int_{\mathbb{R}^n}\hat{f}(\boldsymbol{\xi})\,\varphi(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}=\int_{\mathbb{R}^n}f(\boldsymbol{x})\,\hat{\varphi}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=\langle f,\hat{\varphi}\rangle=\langle\hat{f},\varphi\rangle.$$

Example *6.18* We find the Fourier transform of the Dirac delta $\delta\in\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$:

$$\langle\hat{\delta},\varphi\rangle=\langle\delta,\hat{\varphi}\rangle=\hat{\varphi}(0)=\int_{\mathbb{R}^n}\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}=\langle 1,\varphi\rangle.$$

Thus, $\hat{\delta} = 1$.

The properties of the Fourier transform continue to hold for tempered distributions, and, importantly, they hold without any assumption.

Theorem 6.24 *Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, $\boldsymbol{y} \in \mathbb{R}^n$ and $A \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$ invertible. Then:*

(P1) $S(\boldsymbol{x}) = T(\boldsymbol{x}-\boldsymbol{y}) \quad \Rightarrow \quad \hat{S}(\boldsymbol{\xi}) = e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{T}(\boldsymbol{\xi})$;

(P2) $S(\boldsymbol{x}) = e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,T(\boldsymbol{x}) \quad \Rightarrow \quad \hat{S}(\boldsymbol{\xi}) = \hat{T}(\boldsymbol{\xi}-\boldsymbol{y})$;

(P3) $S(\boldsymbol{x}) = T(A\,\boldsymbol{x}) \quad \Rightarrow \quad \hat{S}(\boldsymbol{\xi}) = \frac{1}{|\det A|}\,\hat{T}(A^{-\top}\boldsymbol{\xi})$;

(P4) $S(\boldsymbol{x}) = \bar{T}(\boldsymbol{x}) \quad \Rightarrow \quad \hat{S}(\boldsymbol{\xi}) = \overline{\hat{T}(-\boldsymbol{\xi})}$;

(P5) *T is even* $\Rightarrow$ *$\hat{T}$ is even;*

(P6) *T is odd* $\Rightarrow$ *$\hat{T}$ is odd;*

(P7) *T is radial* $\Rightarrow$ *$\hat{T}$ is radial.*

Proof We prove **(P1)**, which can be rewritten as

$$\widehat{T(\cdot-\boldsymbol{y})} = e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\hat{T}.$$

First, we notice that $\boldsymbol{\xi} \mapsto e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}$ is a function in $O_M(\mathbb{R}^n;\mathbb{C})$, because it is bounded (indeed, $|e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}| = 1$), and the derivatives are polynomials multiplied by the function itself:

$$\partial^{\alpha}_{\boldsymbol{\xi}}(e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}) = p_{\alpha}(\boldsymbol{y},\boldsymbol{\xi})\,e^{-i\boldsymbol{y}\cdot\boldsymbol{\xi}}$$

Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then

$$\begin{aligned}
\langle\widehat{T(\cdot-\boldsymbol{y})},\varphi\rangle &= \langle T(\boldsymbol{x}-\boldsymbol{y}),\hat{\varphi}(\boldsymbol{x})\rangle\\
&= \langle T(\boldsymbol{x}),\hat{\varphi}(\boldsymbol{x}+\boldsymbol{y})\rangle\\
&= \langle T(\boldsymbol{x}),\widehat{e^{-i\boldsymbol{y}\cdot\boldsymbol{z}}\,\varphi(\boldsymbol{z})}\,(\boldsymbol{x})\rangle\\
&= \langle\hat{T},e^{-i\boldsymbol{y}\cdot\boldsymbol{z}}\,\varphi(\boldsymbol{z})\rangle\\
&= \langle e^{-i\boldsymbol{y}\cdot\boldsymbol{z}}\,\hat{T}(\boldsymbol{z}),\varphi(\boldsymbol{z})\rangle,
\end{aligned}$$

which proves the first point. We prove **(P2)**, which can be re-written as

$$\widehat{e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,T} = \hat{T}(\cdot-\boldsymbol{y}).$$

Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then

$$\begin{aligned}
\langle\widehat{e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,T},\varphi\rangle &= \langle e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,T,\hat{\varphi}\rangle\\
&= \langle T,e^{i\boldsymbol{x}\cdot\boldsymbol{y}}\,\hat{\varphi}\rangle\\
&= \langle T,\hat{\varphi}(\boldsymbol{x}+\boldsymbol{y})\rangle\\
&= \langle\hat{T},\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle = \langle\hat{T}(\boldsymbol{x}-\boldsymbol{y}),\varphi\rangle.
\end{aligned}$$

Now, we prove **(P3)**, which can be re-written as

$$\widehat{T\circ A} = \frac{1}{|\det A|}\hat{T}\circ A^{-\top}.$$

Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then

$$\begin{aligned}\langle \widehat{T\circ A},\varphi\rangle &= \langle T\circ A,\hat{\varphi}\rangle\\ &= \frac{1}{|\det A|}\langle T(\boldsymbol{\xi}),\hat{\varphi}(A^{-1}\boldsymbol{\xi})\rangle\end{aligned}$$

On the other hand, from the properties of inner products

$$\begin{aligned}\hat{\varphi}(A^{-1}\boldsymbol{\xi}) &= \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot A^{-1}\boldsymbol{\xi}}\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathbb{R}^n} e^{-i\,(A^{-1})^{\top}\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{y}\cdot\boldsymbol{\xi}}\,\varphi(A^{\top}\boldsymbol{y})\,|\det A|\,\mathrm{d}\boldsymbol{y} \qquad (\boldsymbol{y}=A^{-\top}\boldsymbol{x})\\ &= |\det A|\,\widehat{\varphi\circ A^{\top}}.\end{aligned}$$

We use this expression to find

$$\begin{aligned}\langle \widehat{T\circ A},\varphi\rangle &= \frac{1}{|\det A|}\langle T(\boldsymbol{\xi}),\hat{\varphi}(A^{-1}\boldsymbol{\xi})\rangle\\ &= \langle T(\boldsymbol{\xi}),\widehat{\varphi\circ A^{\top}}(\boldsymbol{\xi})\rangle\\ &= \langle \hat{T},\varphi\circ A^{\top}\rangle = \langle \hat{T}\circ A^{-\top},\varphi\rangle\,\frac{1}{|\det A|}.\end{aligned}$$

We prove **(P4)**, which can be written as

$$\hat{\hat{T}} = \overline{\hat{T}(-\cdot)}.$$

Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then

$$\begin{aligned}\langle \hat{\hat{T}},\varphi\rangle &= \langle \bar{T},\hat{\varphi}\rangle\\ &= \overline{\langle T,\bar{\hat{\varphi}}\rangle}\\ &= \overline{\langle T,\check{\hat{\varphi}}\rangle}.\end{aligned}$$

From **(P3)** we find $\widehat{T(-\boldsymbol{x})} = \hat{T}(-\cdot)$, which implies $\hat{\check{T}} = \check{\hat{T}}$. It follows that

$$\langle \hat{\check{T}},\varphi\rangle = \overline{\langle \hat{T},\check{\bar{\varphi}}\rangle} = \overline{\langle \check{\hat{T}},\bar{\varphi}\rangle} = \langle \bar{\hat{T}},\varphi\rangle,$$

which yields $\hat{\hat{T}} = \overline{\hat{T}(-\cdot)}$ as required.
We prove **(P5)**. We must prove that $\hat{T}(-\boldsymbol{\xi}) = \hat{T}(\boldsymbol{\xi})$. Since T is even, $\check{T} = T$, so that given $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ we find

$$\langle \hat{\check{T}},\varphi\rangle = \langle \check{T},\hat{\varphi}\rangle = \langle T,\hat{\varphi}\rangle = \langle \hat{T},\varphi\rangle.$$

The proof of **(P6)** is similar, taking into account that, since T is odd, there holds $\check{T} = -T$.

We prove **(P7)**. Since T is radial, for every orthogonal matrix Q we have $T\circ Q = T$. Using **(P3)** we find

$$\hat{T} = \widehat{T\circ Q} = \frac{1}{|\det Q|}\hat{T}\circ Q^{-\top} = \hat{T}\circ Q,$$

since $Q^{-\top} = Q$. Thus, $\hat{T} = \hat{T}\circ Q$, which shows that $\hat{T}$ is radial.

6.7.1 Derivatives of tempered distributions

The definition of derivative of a tempered distribution is formally equal to the one for general distributions.

Definition 6.8 *Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then,*

$$\langle \partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} T, \varphi \rangle \stackrel{def}{=} (-1)^{|\alpha|} \langle T, \partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \varphi \rangle.$$

We observe that this definition is well posed because $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ implies $\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, see Theorem 6.10. Moreover, in Remark 6.3 we have seen that if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ converging to $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then

$$\partial^{\alpha} \varphi_k \xrightarrow{k\to\infty} \partial^{\alpha} \varphi \text{ in } \mathcal{S}(\mathbb{R}^n;\mathbb{C}).$$

This ensures that $\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} T$ is a tempered distribution.

We now extend the properties that relate the Fourier transform of an $L^1(\mathbb{R}^n;\mathbb{C})$ function with the notion of derivative to the class of tempered distributions. It is worth noting the absence of the "unnatural" assumptions that we saw in the $L^1(\mathbb{R}^n;\mathbb{C})$ case.

Theorem 6.25 *Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$. Then*

$$\boxed{\widehat{\partial^{\alpha} T} = (i\boldsymbol{\xi})^{\alpha} \hat{T}}$$

Proof Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then,

$$\begin{aligned}
\langle \widehat{\partial^{\alpha} T}, \varphi \rangle &= \langle \partial^{\alpha} T, \hat{\varphi} \rangle \\
&= (-1)^{|\alpha|} \langle T, \partial^{\alpha} \hat{\varphi} \rangle \\
&= (-1)^{|\alpha|} \langle T, \mathcal{F}\{(-i\boldsymbol{\xi})^{\alpha} \varphi\} \rangle \\
&= (-1)^{|\alpha|} \left\langle \hat{T}, (-i\boldsymbol{\xi})^{\alpha} \varphi \right\rangle \\
&= \left\langle (i\boldsymbol{\xi})^{\alpha} \hat{T}, \varphi \right\rangle.
\end{aligned}$$

Theorem 6.26 *Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$. Then*

$$\boxed{\widehat{(i\boldsymbol{x})^{\alpha} T} = \partial_{\boldsymbol{\xi}}^{\boldsymbol{\alpha}} \hat{T}}$$

Proof Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then,

$$\begin{aligned}
\langle \widehat{(i\boldsymbol{x})^{\alpha} T}, \varphi \rangle &= \langle (i\boldsymbol{x})^{\alpha} T, \hat{\varphi} \rangle \\
&= \langle T, (i\boldsymbol{x})^{\alpha} \hat{\varphi} \rangle \\
&= (-1)^{|\alpha|} \langle T, \widehat{\partial^{\alpha} \varphi} \rangle \\
&= (-1)^{|\alpha|} \langle \hat{T}, \partial^{\alpha} \varphi \rangle = \langle \partial^{\alpha} \hat{T}, \varphi \rangle.
\end{aligned}$$

6.7.2 Inversion formula

We now consider the problem of the antitransformation. To present the main result, we recall that given a function $g : \mathbb{R}^n \longrightarrow \mathbb{C}$ we have defined the unary operation $\check{g}$ as the composition with an invertible linear transformation $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ governed by $A\,\boldsymbol{x} = -\boldsymbol{x}$:

$$\check{g} = g \circ A \quad \Leftrightarrow \quad \check{g}(\boldsymbol{x}) = g(-\boldsymbol{x}).$$

Theorem 6.27 [Inversion Formula]
The linear operator

$$\begin{aligned} \mathcal{F} : \mathcal{S}'(\mathbb{R}^n;\mathbb{C}) &\longrightarrow \mathcal{S}'(\mathbb{R}^n;\mathbb{C}) \\ T &\mapsto \hat{T} \end{aligned}$$

is bijective. Moreover, if $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, then

$$\mathcal{F}^{-1}\{T\} = \frac{1}{(2\pi)^n}\,\check{\hat{T}}. \tag{6.21}$$

Proof First, we show that $\mathcal{F}$ is injective. To this end, we show that if $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ is such that $\hat{T} = 0$, then $T = 0$, or, in other words, the kernel of $\mathcal{F}$ is zero. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Since we have shown in Theorem 6.22 that $\mathcal{F}|_{\mathcal{S}(\mathbb{R}^n;\mathbb{C})}$ (which is a function from $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ to itself) is bijective, there exists $w \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ such that $\hat{w} = \varphi$. It follows that

$$\langle T, \varphi\rangle = \langle T, \hat{w}\rangle = \langle \hat{T}, w\rangle = 0,$$

which shows that $T = 0$.
We now show that $\mathcal{F}$ is surjective, i.e., that every tempered distribution is the Fourier transform of another tempered distribution, and at the same time we prove (6.21). Let $S \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. We want to find $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ such that $\hat{T} = S$. Consider the formula

$$T = \frac{1}{(2\pi)^n}\,\check{\hat{S}}.$$

Clearly, T is tempered. We need to show that indeed $\hat{T} = S$. Taking the Fourier transform of both sides yields

$$\hat{T} = \frac{1}{(2\pi)^n}\,\hat{\check{\hat{S}}}.$$

We notice that for all $R \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ we have $\hat{\check{R}} = \check{\hat{R}}$. Indeed, given the linear transformation $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ ruled by $A\boldsymbol{x} = -\boldsymbol{x}$, we have $\check{R} = R \circ A$, so that

$$\widehat{R \circ A} = \frac{1}{|\det A|}\hat{R} \circ A^{-\top} = \hat{R} \circ A.$$

We use this result in the expression above:

$$\begin{aligned} \hat{T} &= \frac{1}{(2\pi)^n}\,\hat{\check{\hat{S}}} \\ &= \frac{1}{(2\pi)^n}\,\check{\hat{\hat{S}}}. \end{aligned}$$

To prove that the right hand-side is the tempered distribution S, we consider $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$:

$$\begin{aligned}\left\langle \frac{1}{(2\pi)^n}\check{\hat{\hat{S}}},\varphi\right\rangle &= \frac{1}{(2\pi)^n}\left\langle \hat{\hat{S}},\check{\varphi}\right\rangle \\ &= \left\langle S,\frac{1}{(2\pi)^n}\hat{\hat{\varphi}}\right\rangle \\ &= \left\langle S,\frac{1}{(2\pi)^n}\check{\hat{\hat{\varphi}}}\right\rangle = \langle S,\varphi\rangle,\end{aligned}$$

because we have seen that $\varphi = \frac{1}{(2\pi)^n}\check{\hat{\hat{\varphi}}}$, see proof of Theorem 6.22.

Example *6.19*

Consider the constant function

$$\begin{aligned} f : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto 1. \end{aligned}$$

The functional L_f is a tempered distribution because f is bounded. Sometimes, the Fourier transform of this function is denoted by

$$\hat{1} = \int_{\mathbb{R}^n} e^{-i\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{x}.$$

However, the integral in the right hand-side is not a Lebesgue integral: it is part of a symbolic notation to denote the Fourier transform of f. We have seen that $\hat{\delta} = 1$. We use this fact to find $\hat{1}$. Clearly, $\hat{1} = \hat{\hat{\delta}}$. Since for any function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ we have $\check{\check{g}} = g$ (why?), using the inversion formula (6.21) we get

$$\delta = \frac{1}{(2\pi)^n}\check{\hat{\hat{\delta}}} \qquad \Leftrightarrow \qquad \check{\delta} = \frac{1}{(2\pi)^n}\hat{\hat{\delta}} \qquad \Leftrightarrow \qquad \delta = \frac{1}{(2\pi)^n}\hat{\hat{\delta}},$$

where the last equivalence is due to the fact that δ is an even distribution, so that $\check{\delta} = \delta$. The last equality, together with $\hat{1} = \hat{\hat{\delta}}$ established above, yields

$$\hat{1} = (2\pi)^n\,\delta.$$

Example *6.20*

We want to find the Fourier transform of the sine function. We use Euler's formulae:

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \qquad \Rightarrow \qquad \widehat{\sin x} = \frac{\widehat{e^{ix}} - \widehat{e^{-ix}}}{2i}.$$

The Fourier transform of e^{ix} can be computed as

$$\widehat{e^{ix}} = \widehat{\widetilde{\delta(\cdot+1)}} = 2\pi\,\widetilde{\delta(\cdot+1)}.$$

On the other hand, given $\varphi \in \mathcal{S}(\mathbb{R};\mathbb{R})$, we find

$$\langle\widetilde{\delta(\cdot+1)},\varphi\rangle = \langle\delta(x+1),\varphi(-x)\rangle = \langle\delta(y),\varphi(1-y)\rangle = \varphi(1) = \langle\delta(\cdot-1),\varphi\rangle.$$

Thus,

$$\widehat{e^{ix}} = 2\pi\,\widetilde{\delta(\cdot+1)} = 2\pi\,\delta(\xi-1).$$

Likewise, we get

$$\widehat{e^{-ix}} = 2\pi\,\delta(\xi+1).$$

We conclude that

$$\widehat{\sin x} = -i\,\pi\,(\delta(\xi - 1) - \delta(\xi + 1)).$$

Using the same calculations, it is easy to find

$$\widehat{\cos x} = \pi\,(\delta(\xi - 1) + \delta(\xi + 1)).$$

Example *6.21*

Consider the identity function in $\mathbb{R}$:

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x. \end{aligned}$$

First, we want to show that the corresponding distribution is tempered. To do this, we look for a value of $m \in \mathbb{N}$ for which

$$\int_{\mathbb{R}} \frac{|x|}{(1 + |x|)^m}\,\mathrm{d}x < +\infty.$$

The integrand, as $x \to \pm\infty$, is asymptotically equivalent to $1/|x|^{m-1}$, so that the integral converges if and only if $m - 1 > 1$, i.e., if and only if $m > 2$. Thus, L_f is a tempered distribution. Recalling that $\widehat{g'} = (i\,x)\,\hat{g}$, we find

$$x = \frac{1}{i}\,(i\,x)\cdot 1 = -i\,(i\,x)\,\hat{\delta} = -i\,\widehat{\delta'}.$$

Taking the Fourier transform on both sides gives

$$\hat{x} = -i\,\widehat{\widehat{\delta'}}.$$

On the other hand, from the inversion formula we have

$$\delta' = \frac{1}{2\,\pi}\,\overset{\vee}{\widehat{\widehat{\delta'}}} \qquad \Leftrightarrow \qquad \widehat{\widehat{\delta'}} = 2\,\pi\,\overset{\vee}{\delta'} = -2\,\pi\,\delta'$$

because the derivative of δ is odd (why?). Thus,

$$\hat{x} = 2\,\pi\,i\,\delta'.$$

Example *6.22*

Let us consider $T(x) = \operatorname{pv}\frac{1}{x}$. We show that T is a tempered distribution. We recall that $T = Z'$, where $S(x) = \ln|x|$. If we show that $Z \in \mathcal{S}'(\mathbb{R};\mathbb{R})$, we will have $T \in \mathcal{S}'(\mathbb{R};\mathbb{R})$. We can proceed by asking ourselves if there exists $m \in \mathbb{N}$ such that

$$\frac{\ln|x|}{(1 + |x|)^m} \in L^1(\mathbb{R};\mathbb{R}).$$

As $x \to 0$ we have no issues of integrability because $\ln|x| = \mathrm{o}_{x\to 0}(|x|^{-\varepsilon})$ for all $\varepsilon > 0$. For $x \to +\infty$, we recall that $\ln|x| = \mathrm{o}_{x\to\infty}(|x|^{\varepsilon})$. Hence, once ε has been fixed, we can choose m so that $\varepsilon - m < -1$, i.e., $m > 1 + \varepsilon$. In this way, we deduce that Z is tempered, and therefore so is T. We recall that we have found

$$x\,\operatorname{pv}\frac{1}{x} = 1.$$

Taking the Fourier transform of both sides yields

$$\widehat{x\,\operatorname{pv}\tfrac{1}{x}} = \hat{1} = 2\,\pi\,\delta.$$

Hence,

$$\widehat{(-i\,x)\,\operatorname{pv}\tfrac{1}{x}} = -2\,\pi\,i\,\delta.$$

On the other hand, we recall that $\widehat{(-i\,x)\,T} = \partial\hat{T}$, so that

$$\partial\,\widehat{\mathrm{pv}\,\tfrac{1}{x}} = -2\,\pi\,i\,\delta.$$

Clearly, this problem has been reduced to the problem of finding a primitive of $-2\,\pi\,i\,\delta$. We know that if two distributions have the same derivative, they differ by a constant, and a primitive of δ is the Heaviside function. It follows that

$$\widehat{\mathrm{pv}\,\tfrac{1}{x}} = -2\,\pi\,i\,H + C,$$

where $C \in \mathbb{R}$. The problem is now to determine the value of the constant C. To this end, we recall that $\mathrm{pv}\,\frac{1}{x}$ is odd, and therefore its Fourier transform is also odd. This implies that C must be such that $-2\,\pi\,i\,H + C$ is odd. More explicitly, the function at hand is

$$-2\,\pi\,i\,H(\xi) + C = \begin{cases} C - 2\,\pi\,i & \text{if } \xi \geq 0 \\ C & \text{if } \xi < 0. \end{cases}$$

Imposing $\hat{T}(-\xi) = -\hat{T}(\xi)$ gives

$$C = -(C - 2\,\pi\,i) \qquad \Rightarrow \qquad C = \pi\,i.$$

Hence,

$$\widehat{\mathrm{pv}\,\tfrac{1}{x}} = \pi\,i - 2\,\pi\,i\,H.$$

On the other hand

$$\pi\,i - 2\,\pi\,i\,H(\xi) = \begin{cases} -\pi\,i & \text{if } \xi \geq 0 \\ \pi\,i & \text{if } \xi < 0. \end{cases}$$

We conclude that

$$\widehat{\mathrm{pv}\,\tfrac{1}{x}} = -\pi\,i\,\mathrm{sgn}(\xi).$$

Example *6.23* Consider $n = 1$. Find the Fourier transform of the distributions $\delta^{(k)}$ and x^k, where $k \in \mathbb{N}$.

6.7.3 Convergence in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$

We introduce a notion of convergence in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$.

Definition 6.9 *Let $(T_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Then, we say that*

$$T_k \xrightarrow{k\to\infty} T \quad \text{in } \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$$

if for all $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$

$$\lim_{k\to\infty} \langle T_k, \varphi\rangle = \langle T, \varphi\rangle.$$

You can verify that, given a sequence $(T_k)_{k\in\mathbb{N}}$ in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ and $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, if $(T_k)_{k\in\mathbb{N}}$ converges to T in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, the following operations are continuous:

(i) $\forall\,\psi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C}), \quad \lim_{k\to+\infty} \psi\,T_k = \psi\,T;$

(ii) $\forall\,\alpha \in \mathbb{N}^n, \quad \lim_{k\to+\infty} \partial^\alpha\,T_k = \partial^\alpha\,T;$

(iii) if $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is linear and invertible, $\quad \lim_{k\to+\infty} T_k \circ A = T \circ A;$

(iv) $\forall\, \boldsymbol{y} \in \mathbb{R}^n, \quad \lim_{k \to +\infty} T_k(-\boldsymbol{y}) = T(-\boldsymbol{y});$

(v) $\lim_{k \to +\infty} \hat{T}_k = \hat{T}.$

Example *6.24*

We want to find the Fourier transform of the function

$$\begin{aligned} f : \mathbb{R}^3 \setminus \{0\} &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|\boldsymbol{x}\|^{-1}. \end{aligned}$$

First, we notice that f is locally integrable, i.e., $f \in L^1_{\text{loc}}(\mathbb{R}^3;\mathbb{R})$, since $-1 > -3$. However, $f \notin L^1(\mathbb{R}^3;\mathbb{R})$, and therefore to find its Fourier transform we need to first check that L_f is a tempered distribution, by looking for $m \in \mathbb{N}$ such that

$$\int_{\mathbb{R}^3} \frac{\|\boldsymbol{x}\|^{-1}}{(1+\|\boldsymbol{x}\|)^m}\, \mathrm{d}\boldsymbol{x} < +\infty.$$

At the origin we do not have problems of convergence, because once again $-1 > -3$. At ∞ we have $-1 - m < -3$, which gives $m > 2$. Therefore, $L_f \in \mathcal{S}'(\mathbb{R}^3;\mathbb{R})$. We find the Fourier transform, by noting that f can be obtained as the limit in $\mathcal{S}'(\mathbb{R}^3;\mathbb{R})$ of the sequence $(f_k)_{k\in\mathbb{N}}$ defined by

$$f_k(\boldsymbol{x}) = \begin{cases} \|\boldsymbol{x}\|^{-1} & \text{if } \|\boldsymbol{x}\| \le k \\ 0 & \text{if } \|\boldsymbol{x}\| > k. \end{cases}$$

Since $f_k \xrightarrow{k\to\infty} f$ in $\mathcal{S}'(\mathbb{R}^3;\mathbb{R})$, we have also $\hat{f}_k \xrightarrow{k\to\infty} \hat{f}$ in $\mathcal{S}'(\mathbb{R}^3;\mathbb{R})$. Since for every $k \in \mathbb{N}$ the function f_k is in $L^1(\mathbb{R}^3;\mathbb{R})$, we find immediately

$$\hat{f}_k(\boldsymbol{\xi}) = \int_{\|\boldsymbol{x}\| \le k} \|\boldsymbol{x}\|^{-1}\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, \mathrm{d}\boldsymbol{x}.$$

The function f_k is radial, and therefore $\hat{f}_k$ is also radial: we can replace $\boldsymbol{\xi}$ with a vector which has the same norm: we choose $\|\boldsymbol{\xi}\|\,\boldsymbol{e}_3$, where $\boldsymbol{e}_3 = (0,0,1)$, so that

$$\hat{f}_k(\boldsymbol{\xi}) = \hat{f}_k(0,0,\|\boldsymbol{\xi}\|) = \int_{\|\boldsymbol{x}\| \le k} \|\boldsymbol{x}\|^{-1}\, e^{-i\,x_3\,\|\boldsymbol{\xi}\|}\, \mathrm{d}\boldsymbol{x}.$$

Now, we can use the spherical coordinates

$$\begin{cases} x_1 = \rho\, \sin\theta\, \cos\varphi \\ x_2 = \rho\, \sin\theta\, \sin\varphi \\ x_3 = \rho\, \cos\theta, \end{cases} \qquad \text{with } \rho \in (0,1],\ \ \theta \in (0,\pi),\ \ \varphi \in (0, 2\,\pi).$$

The determinant of the Jacobian is $\rho^2 \sin\theta$. We obtain, using Fubini's theorem (Theorem 3.8),

$$\begin{aligned}
\hat{f}_k(\boldsymbol{\xi}) &= \int_{(0,k]\times(0,\pi)\times(0,2\pi)} e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|}\rho^{-1}\rho^2\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&= 2\pi\int_0^k \rho\left(\int_0^\pi e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|}\sin\theta\,\mathrm{d}\theta\right)\mathrm{d}\rho \\
&= 2\pi\int_0^k \rho\left[\frac{e^{-i\rho\cos\theta\|\boldsymbol{\xi}\|}}{i\rho\|\boldsymbol{\xi}\|}\right]_{\theta=0}^{\theta=\pi}\mathrm{d}\rho \\
&= 2\pi\int_0^k \not{\rho}\,\frac{e^{i\rho\|\boldsymbol{\xi}\|}-e^{-i\rho\|\boldsymbol{\xi}\|}}{i\not{\rho}\|\boldsymbol{\xi}\|}\mathrm{d}\rho \\
&= 2\pi\int_0^k 2\,\frac{\sin(\rho\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\mathrm{d}\rho \\
&= 4\pi\left[-\cos(\rho\|\boldsymbol{\xi}\|)\frac{1}{\|\boldsymbol{\xi}\|^2}\right]_{\rho=0}^{\rho=k} \\
&= 4\pi\frac{1}{\|\boldsymbol{\xi}\|^2}\left(1-\cos(k\|\boldsymbol{\xi}\|)\right).
\end{aligned}$$

We prove that

$$\lim_{k\to\infty}\hat{f}_k(\boldsymbol{\xi}) = \frac{4\pi}{\|\boldsymbol{\xi}\|^2} \quad \text{in } \mathcal{S}'(\mathbb{R}^3;\mathbb{R}).$$

To this end, we only have to show that

$$\lim_{k\to\infty}\frac{\cos(k\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|^2} = 0 \quad \text{in } \mathcal{S}'(\mathbb{R}^3;\mathbb{R}).$$

Let $\varphi\in\mathcal{S}(\mathbb{R}^3;\mathbb{R})$. We have, using again the spherical coordinates and Fubini's theorem (Theorem 3.8)

$$\begin{aligned}
\lim_{k\to\infty}\left\langle\frac{\cos(k\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|^2},\varphi(\boldsymbol{\xi})\right\rangle &= \lim_{k\to\infty}\int_{\mathbb{R}^3}\frac{\cos(k\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|^2}\varphi(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \\
&= \lim_{k\to\infty}\int_{(0,+\infty)\times(0,\pi)\times(0,2\pi)}\frac{\cos(k\rho)}{\rho^2}\varphi(\rho\sin\theta\cos\phi,\rho\sin\theta\sin\phi,\rho\cos\theta)\rho^2\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\phi \\
&= \lim_{k\to\infty}\int_{(0,+\infty)}\cos(k\rho)\left(\int_{(0,\pi)\times(0,2\pi)}\varphi(\rho\sin\theta\cos\phi,\rho\sin\theta\sin\phi,\rho\cos\theta)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\right)\mathrm{d}\rho \\
&= \lim_{k\to\infty}\int_{\mathbb{R}}\cos(k\rho)\,g(\rho)\,\mathrm{d}\rho,
\end{aligned}$$

where

$$\begin{aligned}
g:\mathbb{R} &\longrightarrow \mathbb{R} \\
\rho &\mapsto \begin{cases}\displaystyle\int_{(0,\pi)\times(0,2\pi)}\varphi(\rho\sin\theta\cos\phi,\rho\sin\theta\sin\phi,\rho\cos\theta)\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi & \text{if } \rho>0 \\ 0 & \text{if } \rho\le 0.\end{cases}
\end{aligned}$$

From the change of variable theorem we see that $g\in L^1(\mathbb{R};\mathbb{R})$. Moreover, g is continuous, and it is rapidly increasing, since $\rho^m\,|\varphi(\rho\sin\theta\cos\phi,\rho\sin\theta\sin\phi,\rho\cos\theta)|$ is bounded, which implies that $\rho^m\,|\varphi(\rho)|$ is bounded for every $m\in\mathbb{N}$. Therefore, we have integrability at ∞. Since $g\in L^1(\mathbb{R};\mathbb{R})$, Riemann-Lebesgue's Corollary ensures that

$$\int_{\mathbb{R}}\cos(k\rho)\,g(\rho)\,\mathrm{d}\rho = 0.$$

It follows that

$$\widehat{\|\boldsymbol{x}\|^{-1}} = \frac{4\pi}{\|\boldsymbol{\xi}\|^2}.$$

6.7.4 Fourier transform of distributions with compact support

We have already seen that distributions with compact support are tempered. For these special types of tempered distributions, there is an alternative way to compute the Fourier transform.

Theorem 6.28 *Let $T \in \mathcal{E}'(\mathbb{R}^n;\mathbb{C})$. Then, $\hat{T} \in C^\infty(\mathbb{R}^n;\mathbb{C})$, and*

$$\forall\, \boldsymbol{\xi} \in \mathbb{R}^n, \qquad \hat{T}(\boldsymbol{\xi}) = \langle T(\boldsymbol{x}), e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\rangle.$$

We do not present a proof for this result.[2] It is also possible to prove that:

- $\hat{T} \in O_M(\mathbb{R}^n;\mathbb{C})$;
- $\forall\, \boldsymbol{\alpha} \in \mathbb{N}^n$,

$$\partial^\alpha_{\boldsymbol{\xi}} \hat{T}(\boldsymbol{\xi}) = \langle T(\boldsymbol{x}), (-i\,\boldsymbol{x})^\alpha\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\rangle,$$

where $(-i\,\boldsymbol{x})^\alpha\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} = \partial^\alpha_{\boldsymbol{\xi}} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}$.

Example 6.25 We have already seen that $\hat{\delta} = 1$. We obtain the same result using Theorem 6.28. Let $\zeta \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$ be such that $\zeta(\boldsymbol{x}) = 1$ in a neighbourhood of 0. Then,

$$\hat{\delta}(\boldsymbol{\xi}) = \langle \delta(\boldsymbol{x}), e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\rangle = \langle \delta(\boldsymbol{x}), \zeta(\boldsymbol{x})\, e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\rangle = 1.$$

Example 6.26 Let $R > 0$. We want to find the Fourier transform of a distributed mass over a spherical surface

$$S_R(0) = \{\boldsymbol{x} \in \mathbb{R}^3 \mid \|\boldsymbol{x}\| = R\}$$

in $\mathbb{R}^3$. This result is not just important *per se*, but it is also instrumental in the solution of the wave equation. Consider the distribution T described by

$$\langle T, \varphi\rangle = \int_{\|\boldsymbol{x}\|=R} \varphi(\boldsymbol{x})\, \mathrm{d}S(\boldsymbol{x}).$$

We notice that $T \in \mathcal{E}'(\mathbb{R}^3;\mathbb{R})$. In fact, if $\boldsymbol{x} \notin S_R(0)$, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \cap S_R(0) = \varnothing$. Therefore, if $\operatorname{supp}\varphi \subseteq \mathcal{B}_r(\boldsymbol{x})$, we have $\langle T, \varphi = 0$, which implies $\operatorname{supp} T \subseteq S_R(0)$.
We prove that the distribution T is radial. In fact, let Q be orthogonal. We show that $T \circ Q = T$. Let $\varphi \in \mathcal{D}(\mathbb{R}^3;\mathbb{R})$. Since $|\det Q^{-1}| = 1$, we find

$$\langle T \circ Q, \varphi\rangle = \langle T, \varphi \circ Q^{-1}\rangle = \int_{\|\boldsymbol{x}\|=R} \varphi(Q^{-1}\,\boldsymbol{x})\, \mathrm{d}S(\boldsymbol{x}).$$

We parameterize the sphere $S_R(0)$ with $\Phi : D \longrightarrow \mathbb{R}^3$. Since Q is linear, we find

$$\frac{\partial(Q\circ\Phi)}{\partial u}(u,v) = Q\,\frac{\partial\Phi}{\partial u}(u,v) \quad \text{and} \quad \frac{\partial(Q\circ\Phi)}{\partial v}(u,v) = Q\,\frac{\partial\Phi}{\partial v}(u,v).$$

[2] A proof can be found in the textbook "Equations of Mathematical Physics", by V.S. Vladimorov, pag 127.

Hence,

$$\int_{\|\boldsymbol{x}\|=R} \varphi(Q^{-1}\,\boldsymbol{x})\,\mathrm{d}S\,(\boldsymbol{x}) \;=\; \int_D \varphi(\Phi(u,v))\,\left|Q\,\frac{\partial\Phi}{\partial u}(u,v)\wedge Q\,\frac{\partial\Phi}{\partial v}(u,v)\right|\,\mathrm{d}u\,\mathrm{d}v.$$

By definition of vector product

$$\begin{aligned}\left\|Q\,\frac{\partial\Phi}{\partial u}\wedge Q\,\frac{\partial\Phi}{\partial v}\right\| &= \left\|Q\,\frac{\partial\Phi}{\partial u}(u,v)\right\|\,\left\|Q\,\frac{\partial\Phi}{\partial v}(u,v)\right\|\,\sin\left(\arccos\frac{\left\langle Q\,\frac{\partial\Phi}{\partial u},Q\,\frac{\partial\Phi}{\partial v}\right\rangle}{\left|Q\,\frac{\partial\Phi}{\partial u}(u,v)\right|\,\left|Q\,\frac{\partial\Phi}{\partial v}(u,v)\right|}\right)\\ &= \left|\frac{\partial\Phi}{\partial u}\wedge\frac{\partial\Phi}{\partial v}\right|,\end{aligned}$$

because Q is orthogonal and Φ preserves the norms and the inner product. It follows that

$$\begin{aligned}\langle T\circ Q,\varphi\rangle &= \int_D \varphi(\Phi(u,v))\,\left|\frac{\partial\Phi}{\partial u}(u,v)\wedge\frac{\partial\Phi}{\partial v}(u,v)\right|\,\mathrm{d}u\,\mathrm{d}v\\ &= \int_{\|\boldsymbol{x}\|=R}\varphi(\boldsymbol{x})\,\mathrm{d}S\,(\boldsymbol{x}) = \langle T,\varphi\rangle,\end{aligned}$$

which shows that T is radial.

Using Theorem 6.28 we can write

$$\hat{T}(\boldsymbol{\xi}) = \langle T(\boldsymbol{x}), e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\rangle.$$

Using the fact that T is radial we find

$$\hat{T}(\boldsymbol{\xi}) = \hat{T}(0,0,\|\boldsymbol{\xi}\|) = \langle T(\boldsymbol{x}), e^{-i\,x_3\,\|\boldsymbol{\xi}\|}\rangle = \int_{S_R(0)} e^{-i\,x_3\,\|\boldsymbol{\xi}\|}\,\mathrm{d}S\,(\boldsymbol{x}).$$

Consider the parameterisation

$$\begin{aligned}\Phi : [0,\pi]\times[0,2\,\pi] &\longrightarrow \mathbb{R}^3\\ (\theta,\phi) &\mapsto (R\,\sin\theta\,\cos\phi, R\,\sin\theta\,\sin\phi, R\,\cos\theta)\end{aligned}$$

of $S_R(0)$. We have

$$\begin{aligned}\frac{\partial\Phi}{\partial\theta}(\theta,\phi) &= (R\,\cos\theta\,\cos\phi, R\,\cos\theta\,\sin\phi, -R\,\sin\theta)\\ \frac{\partial\Phi}{\partial\phi}(\theta,\phi) &= (-R\,\sin\theta\,\sin\phi, R\,\sin\theta\cos\phi, 0).\end{aligned}$$

Using the expansion of the cross product[a]

$$\frac{\partial\Phi}{\partial\theta}(\theta,\phi)\wedge\frac{\partial\Phi}{\partial\phi}(\theta,\phi) = R^2\,(\sin^2\theta\,\cos\phi, \sin^2\theta\,\sin\phi, \sin\theta\,\cos\phi).$$

It follows that

$$\begin{aligned}\left\|\frac{\partial\Phi}{\partial\theta}(\theta,\phi)\wedge\frac{\partial\Phi}{\partial\phi}(\theta,\phi)\right\| &= R^2\,\sqrt{\sin^4\,\cos^2\phi+\sin^4\theta\,\sin^2\phi+\sin^2\theta\,\cos^2\theta}\\ &= R^2\,\sqrt{\sin^4\theta+\sin^2\theta\,\cos^2\theta}\\ &= R^2\,\sqrt{\sin^2\theta\,(\sin^2\theta+\cos^2\theta)}\\ &= R^2\,\sin\theta.\end{aligned}$$

We obtain

$$\begin{aligned}\int_{S_R(0)} e^{-i\,x_3\,\|\xi\|}\,\mathrm{d}S(\boldsymbol{x}) &= \int_{[0,\pi]\times[0,2\pi]} e^{-i\,R\,\cos\theta\,\|\xi\|}\,R^2\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi\\ &= 2\,\pi\,R^2\int_0^{\pi} e^{-i\,R\,\cos\theta\,\|\xi\|}\,\sin\theta\,\mathrm{d}\theta\\ &= \frac{2\,\pi\,R^2}{i\,R\,\|\xi\|}\left[e^{-i\,R\,\cos\theta\,\|\xi\|}\right]_0^{\pi}\\ &= 2\,\frac{2\,\pi\,R}{i\,\|\xi\|}\,\frac{e^{i\,R\,\|\xi\|}-e^{-i\,R\,\|\xi\|}}{2} = 4\,\pi\,R\,\frac{\sin(R\,\|\xi\|)}{\|\xi\|}.\end{aligned}$$

[a] If $\boldsymbol{v}=a_1\,\boldsymbol{i}+a_2\,\boldsymbol{j}+a_3\,\boldsymbol{k}$ and $\boldsymbol{w}=b_1\,\boldsymbol{i}+b_2\,\boldsymbol{j}+b_3\,\boldsymbol{k}$, the cross product is

$$\boldsymbol{v}\wedge\boldsymbol{w}=(a_2\,b_3-a_3\,b_2)\,\boldsymbol{i}+(a_3\,b_1-a_1\,b_3)\,\boldsymbol{j}+(a_1\,b_2-a_2\,b_1)\,\boldsymbol{k}.$$

Exercises

Exercise 6.24 Let $T\in\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. Let us define, for $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{C})$,

$$\langle\bar{T},\varphi\rangle\stackrel{\text{def}}{=}\overline{\langle T,\bar{\varphi}\rangle}.$$

Show that $\bar{T}\in\mathcal{D}'(\mathbb{R}^n;\mathbb{C})$. Verify that a similar result holds for $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. [Solution: (6.24)]

Exercise 6.25 Let $T\in\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ be odd. Show that $\hat{T}$ is odd. [Solution: (6.25)]

Exercise 6.26 If $T\in\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ is λ-homogeneous, show that its Fourier transform is $(-\lambda-n)$-homogeneous. [Solution: (6.26)]

Exercise 6.27 Consider $n=1$. Find the Fourier transform of the distributions $\delta^{(k)}$ and x^k, where $k\in\mathbb{N}$. [Solution: (6.27)]

Exercise 6.28 Find $\widehat{\cos^2 x}$ and $\widehat{\sin^2 x}$. [Solution: (6.28)]

Exercise 6.29 Find the Fourier transform of H and $H(-\cdot)$ (Hint: recall that $\widehat{\mathrm{sgn}}(\xi)=\frac{2}{i}\,\mathrm{pv}\,\frac{1}{\xi}$). [Solution: (6.29)]

Exercise 6.30 Let $T\in\mathcal{D}'(\mathbb{R};\mathbb{R})$. Show that, if $x\,T=0$, then:

(i) $\mathrm{supp}\,T\subseteq\{0\}$;

(ii) $T\in\mathcal{S}'(\mathbb{R};\mathbb{R})$.

To this end, we use a result proved in a previous Tutorial.[a] Moreover, find all the distributions $T\in\mathcal{S}'(\mathbb{R};\mathbb{C})$ such that $x\,T=0$. [Solution: (6.30)]

[a] Let $\Omega\subseteq\mathbb{R}^n$ be open. Let $T\in\mathcal{D}'(\Omega;\mathbb{C})$ and let $\psi\in C^\infty(\Omega;\mathbb{C})$. We showed that, if $\psi\,T=0$, then

$$\mathrm{supp}\,T\subseteq\{\boldsymbol{x}\in\Omega\,|\,\psi(\boldsymbol{x})=0\}.$$

6.8 Fourier transform in $L^2(\mathbb{R}^n;\mathbb{C})$

Since $L^2(\mathbb{R}^n;\mathbb{C})$ is contained in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, every function in $L^2(\mathbb{R}^n;\mathbb{C})$ has Fourier transform. The following result is of crucial importance. It establishes that the Fourier transform of an element of $L^2(\mathbb{R}^n;\mathbb{C})$ is in $L^2(\mathbb{R}^n;\mathbb{C})$. Thus, the restriction $\mathcal{F}|_{L^2(\mathbb{R}^n;\mathbb{C})}$ gives rise to a bijection.

Theorem 6.29 [Plancherel Theorem]
Let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Then,

$$T \in L^2(\mathbb{R}^n;\mathbb{C}) \qquad \Leftrightarrow \qquad \hat{T} \in L^2(\mathbb{R}^n;\mathbb{C}).$$

Moreover, given $f, g \in L^2(\mathbb{R}^n;\mathbb{C})$, we have

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\overline{g(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi})\,\overline{\hat{g}(\boldsymbol{\xi})}\,\mathrm{d}\boldsymbol{\xi} \qquad \text{[Parseval's Identity]}$$

and

$$\|f\|_2 = (2\pi)^{-\frac{n}{2}}\,\|\hat{f}\|_2. \tag{6.22}$$

Proof We observe that $\mathcal{S}(\mathbb{R}^n;\mathbb{C}) \subseteq L^2(\mathbb{R}^n;\mathbb{C})$. In fact, given $f \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, we have

$$\int_{\mathbb{R}^n} |f(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} (1+\|\boldsymbol{x}\|)^{-2m}\,\left((1+\|\boldsymbol{x}\|)^m\,|f(\boldsymbol{x})|\right)^2\,\mathrm{d}\boldsymbol{x} \le c_m \int_{\mathbb{R}^n} (1+\|\boldsymbol{x}\|)^{-2m}\,\mathrm{d}\boldsymbol{x},$$

where c_m is a positive constant satisfying $\left((1+\|\boldsymbol{x}\|)^m\,|f(\boldsymbol{x})|\right)^2 \le c_m$. The integral converges if $-2m < -n$, i.e., if $m > n/2$.

Let $f, g \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then, using the inversion formula $f = \frac{1}{(2\pi)^n}\,\check{\hat{f}}$, changing variable and bringing the Fourier transform from one function to the other we get

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\overline{g(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \check{\hat{f}}\,\bar{g}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}\,\check{\bar{g}}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}\,\hat{\bar{g}}\,\mathrm{d}\boldsymbol{x} \tag{6.23}$$

On the other hand,

$$\begin{aligned}
\hat{\bar{g}}(\boldsymbol{\xi}) &= \check{\bar{g}}(\boldsymbol{\xi}) = \hat{\bar{g}}(-\boldsymbol{\xi}) \\
&= \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\overline{g(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}^n} \overline{e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,g(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} \\
&= \overline{\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}} = \overline{\hat{g}(\boldsymbol{\xi})}.
\end{aligned}$$

Substituting this back into (6.23) gives

$$\int_{\mathbb{R}^n} f(\boldsymbol{x})\,\overline{g(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}\,\bar{\hat{g}}\,\mathrm{d}\boldsymbol{\xi},$$

which is exactly Parseval's identity. The second formula is obtained simply by writing Parseval's identity with $f = g$:

$$\|f\|_2^2 = \int_{\mathbb{R}^n} f(\boldsymbol{x})\,\overline{f(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}(\boldsymbol{\xi})\,\overline{\hat{f}(\boldsymbol{\xi})}\,\mathrm{d}\boldsymbol{\xi} = (2\pi)^{-n}\,\|\hat{f}\|_2^2$$

immediately gives the required expression.

Now, we consider $f \in L^2(\mathbb{R}^n;\mathbb{C})$. We have seen that, for $p \in [1,+\infty)$, the space $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$ is dense in $L^p(\mathbb{R}^n;\mathbb{C})$. Therefore, $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ is dense in $L^2(\mathbb{R}^n;\mathbb{C})$. It follows that there exists a sequence $(f_k)_{k\in\mathbb{N}}$ of functions in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ that converges to f in $L^2(\mathbb{R}^n;\mathbb{C})$, i.e., such that

$$\|f_k - f\|_2 \xrightarrow{k\to\infty} 0.$$

Let $k, j \in \mathbb{N}$. Using (6.22) we find

$$\|\hat{f}_k - \hat{f}_j\|_2 = \|\widehat{f_k - f_j}\|_2 = (2\pi)^{\frac{n}{2}}\,\|f_k - f_j\|_2$$

Since $L^2(\mathbb{R}^n;\mathbb{C})$ is complete, the sequence $(f_k)_{k\in\mathbb{N}}$ (which is assumed to converge to f in $L^2(\mathbb{R}^n;\mathbb{C})$ is a Cauchy sequence; thus, given $\epsilon > 0$, there exists $p \in \mathbb{N}$ such that, if $k, j \geq p$, we have

$$\|\hat{f}_k - \hat{f}_j\|_2 = (2\pi)^{\frac{n}{2}}\,\|f_k - f_j\|_2 \leq \epsilon.$$

This shows that $(\hat{f}_k)_{k\in\mathbb{N}}$ is also a Cauchy sequence. Therefore, it will have a limit, that we call v:

$$v = \lim_{k\to\infty} \hat{f}_k \quad \text{in } L^2(\mathbb{R}^n;\mathbb{C})$$

The convergence of $(f_k)_{k\in\mathbb{N}}$ to f in $L^2(\mathbb{R}^n;\mathbb{C})$ implies the convergence in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$:

$$v = \lim_{k\to\infty} \hat{f}_k \quad \text{in } \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$$

and therefore we have convergence in $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ of the Fourier transforms:

$$\lim_{k\to\infty} f_k = f \quad \text{in } \mathcal{S}'(\mathbb{R}^n;\mathbb{C}) \qquad \Rightarrow \qquad \lim_{k\to\infty} \hat{f}_k = \hat{f} \quad \text{in } \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$$

It follows that $v = \hat{f}$. Since $v \in L^2(\mathbb{R}^n;\mathbb{C})$, then also $\hat{f}$, which proves that the Fourier transform of an element of $L^2(\mathbb{R}^n;\mathbb{C})$ is in $L^2(\mathbb{R}^n;\mathbb{C})$.

Conversely, let $T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$, and $\hat{T} \in L^2(\mathbb{R}^n;\mathbb{C})$. Since we proved above that the Fourier transform of an element of $L^2(\mathbb{R}^n;\mathbb{C})$ is in $L^2(\mathbb{R}^n;\mathbb{C})$, and therefore $\hat{\hat{T}} \in L^2(\mathbb{R}^n;\mathbb{C})$. Thus, $\check{\hat{T}} \in L^2(\mathbb{R}^n;\mathbb{C})$. It follows from the inversion formula

$$T = \frac{1}{(2\pi)^n}\,\check{\hat{T}}$$

that $T \in L^2(\mathbb{R}^n;\mathbb{C})$.

The last point to prove is Parsevel's identity for $f, g \in L^2(\mathbb{R}^n;\mathbb{C})$ (we have proved its validity in the case where f and g were in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$). Consider the sequences $(f_k)_{k\in\mathbb{N}}$ and $(g_k)_{k\in\mathbb{N}}$ of elements of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ which converge to f and g in $L^2(\mathbb{R}^n;\mathbb{C})$, respectively, in view of the density of $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ in $L^2(\mathbb{R}^n;\mathbb{C})$. From the continuity of the inner product

$$\int_{\mathbb{R}^n} f\,\bar{g} = \lim_{k\to\infty}\int_{\mathbb{R}^n} f_k\,\bar{g}_k = \lim_{k\to\infty}\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}_k\,\bar{\hat{g}}_k = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \hat{f}\,\bar{\hat{g}}.$$

Again, (6.22) follows by writing Parseval's identity with $g = f$.

Example 6.27

We want to find the Fourier transform of the function

$$f(x) = \frac{\sin x}{x}, \qquad x \in \mathbb{R}.$$

First, we want to convince ourselves that this function is in $L^2(\mathbb{R};\mathbb{R})$. There are no problems of integrability at the origin because the limit of f, as x goes to 0, is 1. Moreover, $|f(x)|^2 \leq \frac{1}{x^2}$, which is integrable, and therefore $f \in L^2(\mathbb{R};\mathbb{R})$. It follows that its Fourier transform must also be in $L^2(\mathbb{R};\mathbb{R})$. First, we write

$$x\,f(x) = \sin x \quad \Rightarrow \quad \widehat{x\,f} = \widehat{\sin x}. \tag{6.24}$$

We know that $\hat{f}' = \widehat{-i\,x\,f}$, which implies $\widehat{x\,f} = i\,\hat{f}'$. Moreover, we have found in a previous example that

$$\widehat{\sin}(\xi) = -i\,\pi\,(\delta(\xi-1) - \delta(\xi+1)).$$

Using these facts in (6.24) yields

$$i\,\hat{f}' = i\,\pi\,(\delta(\xi+1) - \delta(\xi-1)).$$

It follows that the required Fourier transform is a primitive of the distribution $\pi\,(\delta(\xi+1) - \delta(\xi-1))$:

$$\hat{f}(\xi) = \pi\,(H(\xi+1) - H(\xi-1)) + C.$$

We can find the constant C by imposing $\hat{f} \in L^2(\mathbb{R};\mathbb{R})$. We have

$$H(\xi+1) - H(\xi-1) = \begin{cases} 0 & \text{if } |\xi| > 1 \\ 1 & \text{if } |\xi| \leq 1 \end{cases}$$

or, written equivalently, $H(\xi+1) - H(\xi-1) = \chi_{[-1,1)}(\xi)$. Thus,

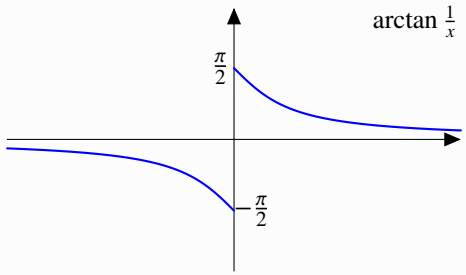
$$\int_{\mathbb{R}} |\hat{f}(\xi)|^2\,\mathrm{d}\xi \geq \int_{(1,+\infty)} |c|^2\,\mathrm{d}\xi,$$

and the latter integral converges if and only if $c = 0$. We conclude that

$$\hat{f}(\xi) = \pi\,\chi_{[-1,1)}(\xi).$$

Example 6.28

We want to find the Fourier transform of the function $f : \mathbb{R}\setminus\{0\} \longrightarrow \mathbb{R}$ defined by $f(x) = \arctan\frac{1}{x}$. First, we want to show that f is in $L^2(\mathbb{R}\setminus\{0\};\mathbb{R})$.



The function is integrable at the origin since

$$\lim_{x\to 0^{\pm}} f(x) = \pm\frac{\pi}{2}.$$

Hence, the function remains bounded in a punctured neighbourhood of the origin. We study the behavior at infinity of $f(x)$. Since as $y \to 0$ the function $\arctan y$ is asymptotically equivalent to y, we find that as $x \to +\infty$

the function $\arctan\frac{1}{x}$ is asymptotically equivalent to $\frac{1}{x}$. Therefore $|\arctan\frac{1}{x}^2|$ is asymptotically equivalent to $\frac{1}{x^2}$, which is integrable. We conclude that $\arctan\frac{1}{x} \in L^2(\mathbb{R}\setminus\{0\};\mathbb{R})$.

We compute the derivative of f in the sense of distributions: we have seen that

$$f' = -\frac{1}{x^2+1} + \pi\,\delta(x).$$

Recall that the Fourier transform of $\frac{1}{x^2+1}$ is $\pi\,e^{-|\xi|}$, so that

$$\widehat{f'} = -\pi\,e^{-|\xi|} + \pi,$$

which leads to

$$i\,\xi\,\hat{f}(\xi) = \pi\,(1 - e^{-|\xi|}).$$

We can divide through by $i\,\xi$ if and only if $\xi \neq 0$, and in that case

$$\hat{f}(\xi) = \frac{\pi\,(1-e^{-|\xi|})}{i\,\xi}.$$

We can give $\hat{f}(\xi)$ an arbitrary value for $\xi = 0$.

6.8.1 An application to partial differential equations

Consider the Laplace equation

$$(\lambda - \nabla^2)\,u = f$$

where $\lambda \in \mathbb{R}_+^*$. We show that, for every $f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{R})$, there exists one and only one tempered distribution u that solves this equation. We can see it by taking the Fourier transform of both sides. Indeed, from

$$\widehat{\nabla^2 u} = \widehat{\frac{\partial^2 u}{\partial x_1^2}\;\frac{\partial^2 u}{\partial x_2^2}\;\cdots\;\frac{\partial^2 u}{\partial x_n^2}} = -\|\xi\|^2\,\hat{u}(\xi).$$

Thus, $\hat{u}$ solves

$$(\lambda + \|\xi\|^2)\,\hat{u}(\xi) = \hat{f}(\xi).$$

Since $\lambda + \|\xi\|^2 \neq$ for all $\xi \in \mathbb{R}^n$. Therefore, we can define

$$\phi(\xi) = \frac{1}{\lambda + \|\xi\|^2},$$

and it is easy to see that $\phi \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{R})$. We find

$$\hat{u}(\xi) = \frac{1}{\lambda + \|\xi\|^2}\,\hat{f}(\xi) \in \mathcal{S}'(\mathbb{R}^n;\mathbb{R}).$$

This shows that there is at most one solution:

$$u = \mathcal{F}^{-1}\left\{\frac{1}{\lambda + \|\xi\|^2}\,\hat{f}\right\}. \tag{6.25}$$

It is possible to show that it is a solution by repeating the same steps backward, to find

$$(\lambda + \|\xi\|^2)\,\hat{u}(\xi) = \hat{f}(\xi),$$

and $\lambda + \|\xi\|^2 = \mathcal{F}\{\lambda - \nabla^2 u\}$, and since $\mathcal{F}$ is injective $(\lambda - \nabla^2 u) = f$. It follows that there is only one solution, and that solution is (6.25).

6.9 Sobolev spaces

We introduce a class of distributions of great importance in modern analysis, and in particular in the solution of certain classes of partial differential equations.

Definition 6.10 *Let $m \in \mathbb{N}$. We define*

$$\mathcal{H}^m(\mathbb{R}^n;\mathbb{C}) \stackrel{def}{=} \{f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C}) \,|\, \forall\, \boldsymbol{\alpha} \in \mathbb{N}^n, \qquad |\boldsymbol{\alpha}| \le m \;\Rightarrow\; \partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} f \in L^2(\mathbb{R}^n;\mathbb{C})\}.$$

Notice that we wrote the element of $\mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ as f, instead of the usual symbol T that we employ for distributions, because we can replace $f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ in $\mathcal{H}^m(\mathbb{R}^n;\mathbb{C})$ with $f \in L^2(\mathbb{R}^n;\mathbb{C})$. In fact, for $\boldsymbol{\alpha} = 0$, the condition $\partial_{\boldsymbol{x}}^{\boldsymbol{\alpha}} f \in L^2(\mathbb{R}^n;\mathbb{C})$ becomes $f \in L^2(\mathbb{R}^n;\mathbb{C})$.

Theorem 6.30 [Sobolev's Theorem]
Let $m \in \mathbb{N}$. Then,

$$\mathcal{H}^m(\mathbb{R}^n;\mathbb{C}) = \{f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C}) \,|\, (1 + \|\boldsymbol{\xi}\|^2)^{\frac{m}{2}}\, \hat{f} \in L^2(\mathbb{R}^n;\mathbb{C})\}.$$

Proof We have to show that

$$f \in \mathcal{H}^m(\mathbb{R}^n;\mathbb{C}) \qquad \Leftrightarrow \qquad (1 + \|\boldsymbol{\xi}\|^2)^{\frac{m}{2}}\, \hat{f} \in L^2(\mathbb{R}^n;\mathbb{C}).$$

We start proving $\Rightarrow$. Let $f \in \mathcal{H}^m(\mathbb{R}^n;\mathbb{C})$. We start proving that there exists $c > 0$ such that

$$\forall\, \boldsymbol{\xi} \in \mathbb{R}^n, \qquad (1 + \|\boldsymbol{\xi}\|^2)^{\frac{m}{2}} \le c \sum_{|\boldsymbol{\alpha}| \le m} \|\boldsymbol{\xi}^{\boldsymbol{\alpha}}\|.$$

Let $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}^n$ and let $j_0 \in \{1, \ldots, n\}$ be such that

$$|\xi_{j_0}| = \max_{1 \le j \le n} |\xi_j|.$$

It follows that $\|\boldsymbol{\xi}\|^2 = \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 \le n\,\xi_{j_0}^2$. Hence,

$$1 + \|\boldsymbol{\xi}\|^2 \le 1 + n\,\xi_{j_0}^2 \quad \Rightarrow \quad (1 + \|\boldsymbol{\xi}\|^2)^{\frac{m}{2}} \le (1 + n\,\xi_{j_0}^2)^{\frac{m}{2}}. \tag{6.26}$$

Let us now consider the function

$$\begin{aligned} [0, +\infty) &\longrightarrow \mathbb{R} \\ t &\mapsto \frac{(1 + n\,t^2)^{\frac{m}{2}}}{1 + t^m}. \end{aligned}$$

This function is bounded, because it is asymptotically equivalent to $n^{\frac{m}{2}}$ (the dominating term t^m appears both in the numerator and in the denominator). It follows that there exists $c > 0$ such that $(1 + n\,t^2)^{\frac{m}{2}} \le c\,(1 + t^m)$. We apply this observation in (6.26) to find

$$1 + \|\boldsymbol{\xi}\|^2 \le 1 + n\,\xi_{j_0}^2 \quad \Rightarrow \quad (1 + \|\boldsymbol{\xi}\|^2)^{\frac{m}{2}} \le (1 + n\,\xi_{j_0}^2)^{\frac{m}{2}} \le c\,(1 + |\xi_{j_0}|^m) \le c \sum_{|\boldsymbol{\alpha}| \le m} \|\boldsymbol{\xi}^{\boldsymbol{\alpha}}\|.$$

In fact, $1 = \xi^0$, and $|\xi_{j_0}|^m = \|\xi^\alpha\|$ if $\alpha = m\,e_{j_0}$. In other words, the term $1 + |\xi_{j_0}|^m$ contains two elements of the sum, which is a sum of non-negative terms, and therefore we have the inequality.

For every $\alpha \in \mathbb{N}^n$, if $|\alpha| \leq m$ we have $\partial^\alpha f \in L^2(\mathbb{R}^n;\mathbb{C})$. Therefore, we have also $\widehat{\partial^\alpha f} \in L^2(\mathbb{R}^n;\mathbb{C})$, and $\widehat{\partial^\alpha f} = (i\,\xi)^\alpha\,\hat{f}$. We obtain

$$(1 + \|\xi\|^2)^{\frac{m}{2}}\,|\hat{f}(\xi)| \leq c \sum_{|\alpha|\leq m} \|\xi^\alpha\,\hat{f}(\xi)\| = c \sum_{|\alpha|\leq m} |(i\,\xi)^\alpha\,\hat{f}(\xi)|.$$

The functions in the sum are all in $L^2(\mathbb{R}^n;\mathbb{C})$, and therefore the sum is also in $L^2(\mathbb{R}^n;\mathbb{C})$. We deduce that $(1 + \|\xi\|^2)^{\frac{m}{2}}\,\hat{f}(\xi) \in L^2(\mathbb{R}^n;\mathbb{C})$.

We now prove $\Leftarrow$. Let $f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$ satisfy $(1 + \|\xi\|^2)^{\frac{m}{2}}\,\hat{f} \in L^2(\mathbb{R}^n;\mathbb{C})$. We have to show that all the partial derivatives, of orders $\leq m$, are in $L^2(\mathbb{R}^n;\mathbb{C})$. To this end, it is suffices to show that the Fourier transforms of the partial derivatives of orders $\leq m$are in $L^2(\mathbb{R}^n;\mathbb{C})$, i.e.

$$\widehat{\partial^\alpha f} \in L^2(\mathbb{R}^n;\mathbb{C}) \qquad \text{if } |\alpha| \leq m.$$

Hence

$$\widehat{\partial^\alpha f} = (i\,\xi)^\alpha\,\hat{f} = \frac{(i\,\xi)^\alpha}{(1 + \|\xi\|^2)^{\frac{m}{2}}}\,\underbrace{(1 + \|\xi\|^2)^{\frac{m}{2}}\,\hat{f}}_{\in L^2(\mathbb{R}^n;\mathbb{C})}$$

We show that $\frac{(i\,\xi)^\alpha}{(1+\|\xi\|^2)^{\frac{m}{2}}}$ is bounded: this allows us to conclude the proof, because in the right hand-side of the latter we have a bounded function multiplying a function in $L^2(\mathbb{R}^n;\mathbb{C})$, and the product is clearly in $L^2(\mathbb{R}^n;\mathbb{C})$. To prove the boundedness, we write

$$|(i\,\xi)^\alpha| = |(i\,\xi_1)^{\alpha_1}\,(i\,\xi_2)^{\alpha_2}\,\ldots\,(i\,\xi_n)^{\alpha_n}| \leq \|\xi\|^{|\alpha|}.$$

It follows that

$$\frac{(i\,\xi)^\alpha}{(1 + \|\xi\|^2)^{\frac{m}{2}}} \leq \frac{\|\xi\|^{|\alpha|}}{(1 + \|\xi\|^\alpha)^{\frac{m}{2}}},$$

which is bounded for $|\alpha| \leq m$, since the function $t \mapsto \frac{t^{|\alpha|}}{(1+t^2)^{\frac{m}{2}}}$ is bounded in $[0, +\infty)$ if $|\alpha| \leq m$. This concludes the proof.

Consider again the Laplace equation Consider the Laplace equation

$$(\lambda - \nabla^2)\,u = f$$

where $\lambda \in \mathbb{R}_+^*$. We show that if $f \in \mathcal{H}^m(\mathbb{R}^n;\mathbb{C})$, then $u \in \mathcal{H}^{m+2}(\mathbb{R}^n;\mathbb{C})$. First of all, we consider the case $m = 0$, and we observe that in this case $\mathcal{H}^0(\mathbb{R}^n;\mathbb{C}) = L^2(\mathbb{R}^n;\mathbb{C})$. This means that $f \in L^2(\mathbb{R}^n;\mathbb{C})$, and we want to see that $u \in \mathcal{H}^2(\mathbb{R}^n;\mathbb{C})$. In other words, taking into account the result of Theorem 6.30, we need to show that $(1 + \|\xi\|^2)\,\hat{u} \in L^2(\mathbb{R}^n;\mathbb{C})$. We recall that we obtained

$$\hat{u}(\xi) = \frac{1}{\lambda + \|\xi\|^2}\,\hat{f}(\xi),$$

which leads to

$$(1 + \|\xi\|^2)\,\hat{u}(\xi) = \frac{(1 + \|\xi\|^2)}{\lambda + \|\xi\|^2}\,\hat{f}(\xi).$$

This function is in $L^2(\mathbb{R}^n;\mathbb{C})$, because $\hat{f} \in L^2(\mathbb{R}^n;\mathbb{C})$ (since $f \in L^2(\mathbb{R}^n;\mathbb{C})$) and $\frac{(1+\|\xi\|^2)}{\lambda+\|\xi\|^2}$ is bounded. Hence, we found that if $f \in L^2(\mathbb{R}^n;\mathbb{C})$, the solution is a function whose second-order derivatives are in

$L^2(\mathbb{R}^n;\mathbb{C})$. This is not true, in general, if we only assume f to be continuous. The same argument can be extended for any $m \in \mathbb{N}$:

$$(1+\|\boldsymbol{\xi}\|^2)^{\frac{m+2}{2}})\,\hat{u}(\boldsymbol{\xi}) = \frac{(1+\|\boldsymbol{\xi}\|^2)}{\lambda+\|\boldsymbol{\xi}\|^2}\,\underbrace{(1+\|\boldsymbol{\xi}\|^2)^{\frac{m}{2}})\hat{f}(\boldsymbol{\xi})}_{L^2(\mathbb{R}^n;\mathbb{C})},$$

6.9.1 Fourier transform of the convolution for tempered distributions

We have seen that, when f and g are functions in $L^1(\mathbb{R}^n;\mathbb{C})$, the Fourier transform of the convolution is well defined, and it is equal to the product of the Fourier transforms of the two functions:

$$\forall\,\boldsymbol{\xi}\in\mathbb{R}^n,\quad \widehat{f*g}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi}).$$

If we want to extend this result for tempered distributions, we need one of the Fourier transforms in the right hand-side to be in $O_M(\mathbb{R}^n;\mathbb{C})$, to ensure that the product is well defined. In the next theorem, we show that if f is a tempered distribution and $g \in O_M(\mathbb{R}^n;\mathbb{C})$, the convolution is well defined. In this case, $\hat{g}(\boldsymbol{\xi}) \in O_M(\mathbb{R}^n;\mathbb{C})$, and the product $\hat{f}(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi})$ is well defined.

Theorem 6.31 [Fourier Transform and Convolution - II]
Let $f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $g \in \mathcal{E}'(\mathbb{R}^n;\mathbb{C})$. Then:

(i) *f and g are convolvable;*

(ii) *$f * g \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$;*

(iii) *Let $\boldsymbol{U} \subseteq \mathbb{R}^n$ be an open set such that $\operatorname{supp} g \subseteq \boldsymbol{U}$, if $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, given $\zeta \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$ such that $\zeta(\boldsymbol{y}) = 1$ for $\boldsymbol{y} \in U$, then the function*

$$\begin{aligned}\mathbb{R}^n &\longrightarrow \mathbb{C}\\ \boldsymbol{x} &\mapsto \langle g(\boldsymbol{y}), \zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\end{aligned}$$

is in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Moreover,

$$\langle f*g,\varphi\rangle = \big\langle f(\boldsymbol{x}), \langle g(\boldsymbol{y}), \zeta(\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\big\rangle;$$

(iv) $\widehat{f*g}(\boldsymbol{\xi}) = \hat{f}(\boldsymbol{\xi})\,\hat{g}(\boldsymbol{\xi})$.

Proof We limit ourselves to proving the last point. We show that $\mathcal{F}^{-1}\{\hat{g}\,\hat{f}\} = f * g$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. We have

$$\begin{aligned}\langle \mathcal{F}^{-1}\{\hat{g}\,\hat{f}\},\varphi\rangle &= (2\pi)^{-n}\,\langle \widehat{\hat{g}\,\hat{f}},\varphi\rangle\\ &= (2\pi)^{-n}\,\langle \hat{g}\,\hat{f},\hat{\check{\varphi}}\rangle\\ &= (2\pi)^{-n}\,\langle \hat{g}\,\hat{f},\check{\hat{\varphi}}\rangle\\ &= (2\pi)^{-n}\,\langle \hat{f},\hat{g}\,\check{\hat{\varphi}}\rangle\\ &= (2\pi)^{-n}\,\langle f,\widehat{\hat{g}\,\check{\hat{\varphi}}}\rangle\end{aligned}$$

We have

$$\begin{aligned}
\widehat{\hat{g}\,\check{\hat{\varphi}}} &= \int_{\mathbb{R}^n} e^{-i\,x\cdot\xi}\,\hat{g}(\xi)\,\hat{\varphi}(-\xi)\,\mathrm{d}\xi \\
&= \int_{\mathbb{R}^n} \widehat{g(\cdot - x)}(\xi)\,\hat{\varphi}(-\xi)\,\mathrm{d}\xi \\
&= \langle \widehat{g(\cdot - x)}, \check{\hat{\varphi}}\rangle \\
&= \langle g(\cdot - x), \hat{\check{\varphi}}\rangle \\
&= (2\pi)^n \langle g(\cdot - x), \varphi\rangle \\
&= (2\pi)^n \langle g(y), \varphi(x+y)\rangle \\
&= (2\pi)^n \langle g(y), \zeta(y)\,\varphi(x+y)\rangle.
\end{aligned}$$

It follows that

$$\langle \mathcal{F}^{-1}\{\hat{g}\,\hat{f}\}, \varphi\rangle = (2\pi)^{-n} \left\langle f(x), (2\pi)^n \langle g(y), \zeta(y)\,\varphi(x+y)\rangle\right\rangle = \langle f * g, \varphi\rangle.$$

We also present the following result without proof.

Theorem 6.32 [Fourier Transform and Convolution - III]
Let $f \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. Let $g \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$. Then:

(i) *f and g are convolvable;*

(ii) $f * g \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$;

(iii) $f * g \in \mathcal{O}_M(\mathbb{R}^n;\mathbb{C})$;

(iv) $\widehat{f * g}(\xi) = \hat{f}(\xi)\,\hat{g}(\xi)$;

(v) $(f * g)(x) = \langle f(y), g(x-y)\rangle$.

Example 6.29

We want to find $u \in \mathcal{S}'(\mathbb{R};\mathbb{R})$ that satisfies

$$e^{-x^2} * u = x\,e^{-x^2}.$$

We know that e^{-x^2}, and therefore the convolution with a tempered distribution is well defined. We take the Fourier transform of both sides:

$$\widehat{e^{-x^2} * u} = \widehat{x\,e^{-x^2}},$$

and we obtain

$$\widehat{e^{-x^2}}\,\hat{u} = \widehat{x\,e^{-x^2}}.$$

We have already found that $\widehat{e^{-x^2}}(\xi) = \sqrt{\pi}\,e^{-\frac{\xi^2}{4}}$, so that

$$\begin{aligned}
\widehat{x\,e^{-x^2}}(\xi) &= i\,\widehat{(-i\,x)\,e^{-x^2}}(\xi) \\
&= i\,\partial\left(\mathcal{F}\{e^{-x^2}\}\right)(\xi) = i\,\sqrt{\pi}\,e^{-\frac{\xi^2}{4}}\left(-\frac{\xi}{2}\right).
\end{aligned}$$

We obtain

$$\sqrt{\pi}\,e^{-\frac{\xi^2}{4}}\,\hat{u}(\xi) = -i\,\sqrt{\pi}\,\frac{1}{2}\,e^{-\frac{\xi^2}{4}}\,\xi.$$

We can multiply both sides by $e^{\frac{\xi^2}{4}}$, which is of class C^∞, and we obtain

$$\hat{u}(\xi) = -\frac{i}{2}\,\xi = \widehat{-\tfrac{1}{2}\,\delta'}.$$

From the injectivity of the Fourier transform we conclude that

$$u = -\frac{1}{2}\,\delta'.$$

We check that indeed u satisfies the equation $e^{-x^2} * u = x\,e^{-x^2}$:

$$e^{-x^2} * \left(-\frac{1}{2}\,\delta'\right) = -\frac{1}{2}\,(e^{-x^2} * \delta)' = -\frac{1}{2}\,(e^{-x^2})' = -\frac{1}{2}\,e^{-x^2}\,(-2\,x) = x\,e^{-x^2}.$$

7 Linear partial differential equations

In this chapter, we consider a number of partial differential equations that arise in mathematical physics. In particular, we will study:

- the Laplace operator ∇^2
- the heat operator $\dfrac{\partial}{\partial t} - \nabla_x^2$
- the wave operator $\dfrac{\partial^2}{\partial t^2} - \dfrac{1}{c^2}\,\nabla_x^2$
- the Schroëndinger operator $i\,\dfrac{\partial}{\partial t} - \nabla_x^2$

7.1 Fundamental solution of differential operators

Consider the differential operator with constant coefficients of degree m

$$P(\partial) = \sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, \partial^{\alpha}. \tag{7.1}$$

For every $\boldsymbol{\alpha} \in \mathbb{N}^n$ such that $|\boldsymbol{\alpha}| \le m$, the coefficients a_{α} is a complex number, and we assume $a_{\boldsymbol{\beta}} \neq 0$ for at least one $\boldsymbol{\beta} \in \mathbb{N}^n$ with $|\boldsymbol{\beta}| = m$ (with no loss of generality: this is the condition that allows us to say that the degree of $P(\partial)$ is m: if for all $\boldsymbol{\beta} \in \mathbb{N}^n$ with $|\boldsymbol{\beta}| = m$ we have $a_{\boldsymbol{\beta}} = 0$, then the degree of $P(\partial)$ is smaller than m).

The partial differential equations that we will consider are equations in the form

$$\sum_{|\boldsymbol{\alpha}| \le m} a_{\boldsymbol{\alpha}}\, \partial^{\alpha} u = f,$$

where the solution, if it exists, is a distribution $u \in \mathcal{D}'(\Omega; \mathbb{C})$.

Definition 7.1 [Fundamental Solution]
Let $P(\partial)$ be a differential operator with constant coefficients of degree m. A fundamental solution of

$$P(\partial)\, u = f$$

is a distribution $\boldsymbol{E} \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$ such that

$$P(\partial)\, \boldsymbol{E} = \delta.$$

If $P(\partial) = \sum_{|\alpha| \leq m} a_\alpha \, \partial^\alpha$, this implies that for every $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$ we have

$$\left\langle \sum_{|\alpha| \leq m} a_\alpha \, \partial^\alpha \boldsymbol{E}, \varphi \right\rangle = \varphi(0),$$

which can also be written as

$$\left\langle \boldsymbol{E}, \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \, a_\alpha \, \partial^\alpha \varphi \right\rangle = \varphi(0).$$

Typically, the fundamental solution is not unique. If $\boldsymbol{E}$ is a fundamental solution of $P(\partial)\, u = f$ and if $\mathcal{F}$ is another distribution in the kernel of the operator, i.e., such that $P(\partial)\, \mathcal{F} = 0$, then $\boldsymbol{E} + \mathcal{F}$ is also a fundamental solution, i.e., we will have also $P(\partial)(\boldsymbol{E} + \mathcal{F}) = \delta$.

Example

We recall that we have calculated, in $\mathbb{R}^3$, the Laplacian of the function

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} \|\boldsymbol{x}\|^{-1} & \text{if } \boldsymbol{x} \neq 0 \\ 0 & \text{if } \boldsymbol{x} = 0 \end{cases} \end{aligned}$$

In particular, we found that $\nabla^2 f = -4\pi\delta$. We recall that in complex analysis (and in particular in the theory of functions of one complex variable), the so-called *harmonic functions* play an important role. These functions are the solutions of the equation $\nabla^2 f = 0$. Hence, the distribution

$$\boldsymbol{E}(\boldsymbol{x}) = -\frac{1}{4\pi} \|\boldsymbol{x}\|^{-1}$$

is a fundamental solution of ∇^2 in $\mathbb{R}^3$. Adding to $\boldsymbol{E}$ a harmonic function gives another fundamental solution.

If the polynomial P satisfies the condition

$$\forall\, \boldsymbol{\xi} \in \mathbb{R}^n, \quad |P(i\,\boldsymbol{\xi})| \geq c > 0, \tag{7.2}$$

then the equation $P(\delta)\, \boldsymbol{E} = \delta$ can straightforwardly be solved by using the Fourier transform. Indeed, we can take the Fourier transform of both sides of $P(\delta)\, \boldsymbol{E} = \delta$:

$$P(i\,\boldsymbol{\xi})\, \hat{\boldsymbol{E}} = 1.$$

The assumption (7.2) ensures that $P(i\,\boldsymbol{\xi})$ is never zero, so

$$\hat{\boldsymbol{E}} = \frac{1}{P(i\,\boldsymbol{\xi})}.$$

Moreover,

$$\left| \frac{1}{P(i\,\boldsymbol{\xi})} \right| \leq \frac{1}{c},$$

and therefore $\frac{1}{P(i\,\boldsymbol{\xi})} \in L^\infty(\mathbb{R}^3; \mathbb{R}) \subset \mathcal{S}'(\mathbb{R}^3; \mathbb{R})$. In other words, $\frac{1}{P(i\,\boldsymbol{\xi})}$ is a tempered distribution, and we can take the inverse Fourier transform, and

$$\boldsymbol{E} = \mathcal{F}^{-1} \left\{ \frac{1}{P(i\,\boldsymbol{\xi})} \right\}$$

is a fundamental solution.

Example Consider the differential operator $P(\delta) = 1 - \nabla^2$ in $\mathbb{R}^n$, which corresponds to the polynomial $P(\boldsymbol{\xi}) = 1 - \|\boldsymbol{\xi}\|^2 = 1 - \xi_1^2 - \xi_2^2 - \ldots - \xi_n^2$, with $\boldsymbol{\xi} \in \mathbb{R}^n$. In this case, $P(i\,\boldsymbol{\xi}) = 1 + \|\boldsymbol{\xi}\|^2 \geq 1$, which means that (7.2) is satisfied. It follows that a fundamental solution is

$$\boldsymbol{E} = \mathcal{F}^{-1}\left\{\frac{1}{1 + \|\boldsymbol{\xi}\|^2}\right\}.$$

Unfortunately, condition (7.2) is very restrictive in general.

We now consider the problems related to fundamental solutions: existence/uniqueness of solutions and the regularity of the solutions.

7.1.1 Existence and uniqueness of solutions

Theorem 7.1 [Existence of Solutions]
Let $\boldsymbol{E}$ be a fundamental solution of $P(\delta)$. Let $f \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$, and consider the equation

$$P(\delta)\, u = f. \tag{7.3}$$

If f and $\boldsymbol{E}$ are convolvable, then

$$u = \boldsymbol{E} * f$$

is a solution of (7.3).

Proof We need to show that $P(\partial)(\boldsymbol{E} * f) = f$. We recall that the convolution is distributive under the sum[a] and therefore, considering $P(\delta)$ in the form (7.1), we obtain

$$\begin{aligned} P(\partial)(\boldsymbol{E} * f) &= \sum_{|\alpha| \leq m} a_\alpha\, \partial^\alpha\, (\boldsymbol{E} * f) \\ &= \sum_{|\alpha| \leq m} a_\alpha\, (\partial^\alpha\, \boldsymbol{E} * f) \\ &= \left(\sum_{|\alpha| \leq m} a_\alpha\, \partial^\alpha\, \boldsymbol{E}\right) * f = P(\partial)\, \boldsymbol{E} * f = \delta * f = f. \end{aligned}$$

[a] If f_1 and f_2 are convolvable with g and $\lambda_1, \lambda_2 \in \mathbb{C}$, then $\lambda_1\, f_1 + \lambda_2\, f_2$ is convolvable with g and

$$(\lambda_1\, f_1 + \lambda_2\, f_2) * g = \lambda_1\, (f_1 * g) + \lambda_2\, (f_2 * g).$$

We now consider the problem of uniqueness of the solution.

Theorem 7.2 [Uniqueness of a Solution]
Let $\boldsymbol{E}$ be a fundamental solution of $P(\delta)$. Let $f \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$. Then, the equation

$$P(\delta)\, u = f \tag{7.4}$$

has at most one solution which is convolvable with $\boldsymbol{E}$.

Proof Let u_1, u_2 be two solutions of (7.4) which are convolvable with $\boldsymbol{E}$. We must show that $u_1 = u_2$. Since $P(\partial)\, u_1 = f$ and $P(\partial)\, u_2 = f$, by linearity

$$P(\partial)\,(u_1 - u_2) = 0.$$

Since both u_1 and u_2 are convolvable with $\boldsymbol{E}$, then $u_1 - u_2$ is also convolvable with $\boldsymbol{E}$, and the derivatives are also convolvable with$\boldsymbol{E}$. Therefore,

$$\boldsymbol{E} * P(\partial)\,(u_1 - u_2) = \boldsymbol{E} * 0 = 0.$$

On the other hand, this means that

$$\begin{aligned} 0 &= P(\partial)\,(u_1 - u_2) * \boldsymbol{E} \\ &= P(\partial)\,((u_1 - u_2) * \boldsymbol{E}) \\ &= P(\partial)\,(\boldsymbol{E} * (u_1 - u_2)) \\ &= P(\partial)\,\boldsymbol{E} * (u_1 - u_2) \\ &= \delta * (u_1 - u_2) = u_1 - u_2, \end{aligned}$$

and therefore $u_1 = u_2$.

The fundamental results regarding the existence of a fundamental solutions are the following. The first one was proved independently by Leon Ehrenpreis in 1954 and Bernard Malgrange in 1955.

Theorem 7.3 [Malgrange-Ehrenpreis Theorem]
If $P(\boldsymbol{\xi})$ is not the identically zero polynomial, then $P(\delta)$ has a fundamental solution.

The second result is the following, by Lars Valter Hörmander (1955).

Theorem 7.4 [Hörmander Theorem]
If $P(\boldsymbol{\xi})$ is not the identically zero polynomial, then a tempered fundamental solution always exists.

Let us consider again the equation

$$P(\partial)\, u = f,$$

and assume that $f \in \boldsymbol{E}'(\mathbb{R}^n; \mathbb{C})$. In other words, we are assuming that f has compact support and that P is not the identically zero polynomial. Then, this equation always has solutions. In fact, since P is not the identically zero polynomial, it will have a fundamental solution $\boldsymbol{E}$, and since f has compact support, it is convolvable with any distribution, and in particular with $\boldsymbol{E}$, so that $u = \boldsymbol{E} * f$ is a solution.

If f is a distribution, then $P(\partial)\, u = f$ is always locally solvable. Indeed, assume $f \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$, and let $\boldsymbol{x}_0 \in \mathbb{R}^n$. Let $\zeta \in \mathcal{D}(\mathbb{R}^n; \mathbb{C})$ be such that $\zeta(\boldsymbol{x}) = 1$ for all $\boldsymbol{x} \in \mathcal{B}_R(\boldsymbol{x}_0)$, where $R > 0$. Then, $\zeta\, f \in \boldsymbol{E}'(\mathbb{R}^n; \mathbb{C})$, and therefore $P(\partial)\, u = \zeta\, f$ is solvable. In other words, $P(\partial)\, u|_{\mathcal{B}_R(x_0)} = f|_{\mathcal{B}_R(x_0)}$.

7.1.2 Regularity of solutions

We begin this section with the definition of singular support of a distribution, which, in essence, is the set of points at which a distribution fails to be regular.

Definition 7.2 [Singular Support]
Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $T \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$. The singular support of T is defined as

$$\operatorname{singsupp} T = \Omega \setminus \{\boldsymbol{x} \in \Omega \mid \exists A \subseteq \Omega : \quad A \text{ is open, } \boldsymbol{x} \in A \text{ and } T|_A \in \boldsymbol{C}^\infty(A; \mathbb{C})\}.$$

It is possible to prove that, given $T \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$,

$$\operatorname{singsupp} T = \varnothing \quad \Leftrightarrow \quad T \in \boldsymbol{C}^\infty(A; \mathbb{C}).$$

Clearly, $T \in \boldsymbol{C}^\infty(A; \mathbb{C})$ implies that T is a distribution of function type.

Example We want to convince ourselves that $\operatorname{singsupp} \delta = \{0\}$. Indeed, let $\boldsymbol{x} \neq 0$. Let $A = \mathbb{R}^n \setminus \{0\}$. Hence, A is an open set which contains $\boldsymbol{x}$. We have $\delta|_A = 0$, and it is therefore of class C^∞. It follows that $\boldsymbol{x} \notin \operatorname{singsupp} \delta$. However, $0 \in \operatorname{singsupp} \delta$. In fact, it this was not the case, we would have $\operatorname{singsupp} \delta = \varnothing$, which would imply that $\delta \in \boldsymbol{C}^\infty(\mathbb{R}^n; \mathbb{R})$, and this contradicts the fact that δ is not of function type.

Consider a distribution $T \in \mathcal{D}'(\Omega; \mathbb{C})$ and let $P(\partial)$ be a differential operator. Then,

$$\operatorname{singsupp}(P(\partial)\, T) \subseteq \operatorname{singsupp} T.$$

To prove this, we can show that if $\boldsymbol{x}_0 \notin \operatorname{singsupp} T$, then $\boldsymbol{x}_0 \notin \operatorname{singsupp}(P(\partial)\, T)$. Let $\boldsymbol{x}_0 \notin \operatorname{singsupp} T$, so that there exists an open set $A \subseteq \Omega$ such that $\boldsymbol{x}_0 \in A$ and $T|_A \in \boldsymbol{C}^\infty(A; \mathbb{C})$. It follows that $P(\partial)\, T|_A \in \boldsymbol{C}^\infty(A; \mathbb{C})$, and therefore $\boldsymbol{x}_0 \notin \operatorname{singsupp}(P(\partial)\, T)$.

If the singular support does not reduce by applying the differential operator, the differential operator is said to be *hypoelliptic*.

Definition 7.3 [Hypoelliptic Operator]
Let $P(\partial)$ be a differential operator. We say that $P(\partial)$ is hypoelliptic *if, for all $T \in \mathcal{D}'(\Omega; \mathbb{C})$,*

$$\operatorname{singsupp}(P(\partial)\, T) = \operatorname{singsupp} T.$$

Example Consider $(x, y) \in \mathbb{R}^2$. Consider the polynomial $P(\xi_1, \xi_2) = \xi_1$ in $\mathbb{R}^2$, which corresponds to the differential operator

$$\frac{\partial}{\partial x}.$$

We show that $P(\partial)$ is not hypoelliptic. Consider the distribution

$$T(x, y) = 1(x) \otimes \delta(y).$$

Clearly, $P(\partial)T = \partial_x T = 0$, and therefore the singular support of $P(\partial)T$ is empty. We want to determine $\operatorname{singsupp} T$. Let $\varphi \in \mathcal{D}(\mathbb{R}^2; \mathbb{R})$. We get

$$\langle T(x, y), \varphi(x, y) \rangle = \int_{\mathbb{R}} \varphi(x, 0)\, \mathrm{d}x.$$

We proved (in a tutorial) that T is not of function type, and therefore its singular support cannot be empty.

It is also possible to see that the singular support is the entire x-axis.

Let us now consider the equation

$$P(\partial)\, T = f,$$

where $T, f \in \mathcal{D}'(\mathbb{R}^n; \mathbb{C})$, and let $P(\partial)$ by hypoelliptic. In this case,

$$\operatorname{singsupp} T = \operatorname{singsupp} f,$$

and therefore the singularities of the solution are the same as the singularities of f. In particular,

$$\boldsymbol{x} \notin \operatorname{singsupp} f \quad \Rightarrow \quad \boldsymbol{x} \notin \operatorname{singsupp} T.$$

This implies that if there is a neighbourhood where f is of class C^∞, then u in that neighbourhood is of class C^∞. In particular, if $f \in C^\infty(\mathbb{R}^n; \mathbb{C})$, we have also that $u \in C^\infty(\mathbb{R}^n; \mathbb{C})$.

Theorem 7.5 [Schwartz]
Let P be a non-identically zero polynomial. The following facts are equivalent:

(a) *$P(\partial)$ is hypoelliptic;*

(b) *every fundamental solution of $P(\partial)$ is of class C^∞ over $\mathbb{R}^n \setminus \{0\}$;*

(c) *there exists a fundamental solution of class C^∞ over $\mathbb{R}^n \setminus \{0\}$.*

We limit ourselves to the following remarks. The implication **(a)** $\Rightarrow$ **(b)** is clear: the singular support of the fundamental solution must coincide with the singular support of δ. The implication **(b)** $\Rightarrow$ **(c)** is also obvious.

The implication **(c)** $\Rightarrow$ **(a)** shows that the Laplacian ∇^2 in $\mathbb{R}^3$ is hypoelliptic. In fact,

$$\boldsymbol{E}(x) = -\frac{1}{4\pi} \|\boldsymbol{x}\|^{-1} \in C^\infty(\mathbb{R}^3 \setminus \{0\}).$$

It follows that if $\nabla^2 u = f$ with $u \in \mathcal{D}'(\mathbb{R}^3; \mathbb{R})$ and $f \in C^\infty(\mathbb{R}^3; \mathbb{R})$, then we have also $u \in C^\infty(\mathbb{R}^3; \mathbb{R})$. In particular, harmonic functions are all of class C^∞.

7.2 Ordinary differential operator

Let us consider the operator of degree $m \geq 1$ in $\mathbb{R}$

$$P(\partial) = \sum_{j=0}^{m} a_j \partial^j,$$

where $a_0, \ldots, a_m \in \mathbb{C}$ and $a_m \neq 0$. We want to construct a fundamental solution, i.e., a distribution $\boldsymbol{E} \in \mathcal{D}'(\mathbb{R}; \mathbb{C})$ such that $P(\partial)\, \boldsymbol{E} = \delta$. We now see that it is always possible to determine $\boldsymbol{E}$ with support in $[0, +\infty)$. In particular, we want to construct a fundamental solution with the structure

$$\boldsymbol{E}(x) = \begin{cases} S(x) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases}$$

where S is the "classic" solution over $\mathbb{R}$ of the ordinary differential equation $P(\partial)\, S = 0$: clearly, from Cauchy's Theorem, S is determined from the initial conditions $S(0), S'(0), \ldots, S^{m-1}(0)$. From $\boldsymbol{E}(x) =$

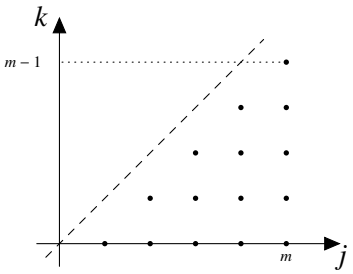
$S(x)\,H(x)$, we have

$$\begin{aligned}
\boldsymbol{E}'(x) &= S'(x)\,H(x) + S(0)\,\delta(x),\\
\boldsymbol{E}''(x) &= S''(x)\,H(x) + S'(0)\,\delta(x) + S(0)\,\delta'(x),\\
&\vdots\\
\boldsymbol{E}^{(j)}(x) &= S^{(j)}\,H(x) + S^{(j-1)}\,\delta(x) + \ldots + S(0)\,\delta^{(j-1)}(x)\\
&= S^{(j)}\,H(x) + \sum_{k=0}^{j-1} S^{(j-1-k)}(0)\,\delta^{(k)}(x).
\end{aligned}$$

Substituting these expressions into $P(\partial)\,\boldsymbol{E}$ yields

$$\begin{aligned}
P(\partial)\,\boldsymbol{E} &= \sum_{j=0}^{m} a_j\,\boldsymbol{E}^{(j)}\\
&= a_0\,H\,S + \sum_{j=1}^{m} a_j \left(S^{(j)}\,H + \sum_{k=0}^{j-1} S^{(j-1-k)}(0)\,\delta^{(k)} \right)\\
&= H\left(a_0\,S + \sum_{j=1}^{m} a_j\,S^{(j)} \right) + \sum_{j=1}^{m} a_j \left(\sum_{k=0}^{j-1} S^{(j-1-k)}(0)\,\delta^{(k)} \right)\\
&= \underbrace{H\,P(\partial)\,S}_{=0} + \sum_{j=1}^{m} a_j \left(\sum_{k=0}^{j-1} S^{(j-1-k)}(0)\,\delta^{(k)} \right),
\end{aligned}$$

where $H\,P(\partial)\,S$ is zero because S is the classic solution. We can re-write the sum as follows:



$$\sum_{j=1}^{m}\sum_{k=0}^{j-1} a_j\,S^{(j-1-k)}(0)\,\delta^{(k)} = \sum_{k=0}^{m-1}\sum_{j=k+1}^{m} a_j\,S^{(j-1-k)}(0)\,\delta^{(k)}$$

$$P(\partial)\,\boldsymbol{E} = \sum_{k=0}^{m-1}\sum_{j=k+1}^{m} a_j\,S^{(j-1-k)}(0)\,\delta^{(k)}.$$

Imposing the sum to be equal to δ requires imposing all the coefficients of $\delta^{(k)}$ to be zero except for $k = 0$: more precisely, we want

$$\sum_{j=1}^{m} a_j\,S^{(j-1)}(0) = 1 \qquad \text{and} \qquad \sum_{j=k+1}^{m} a_j\,S^{(j-1-k)}(0) = 1 \quad \text{for } k = 1, \ldots, m-1.$$

For $k = m - 1$ we have

$$\sum_{j=m}^{m} a_j\,S^{(j-1-m+1)}(0) = a_m\,S(0) = 0 \qquad \Rightarrow \qquad S(0) = 0$$

since by assumption $a_m \neq 0$. Likewise, for $k = m-2$

$$\sum_{j=m-1}^{m} a_j S^{(j-1-m+1)}(0) = a_{m-1} \underbrace{S(0)}_{=0} + a_m S'(0) = 0 \qquad \Rightarrow \qquad S'(0) = 0.$$

Continuing in this fashion, we obtain

$$S(0) = S'(0) = \ldots = S^{(m-2)} = 0.$$

For $k = 0$ we have

$$1 = \sum_{j=1}^{m} a_j S^{(j-1)}(0) = a_m S^{(m-1)}(0) \quad \Rightarrow \quad S^{(m-1)}(0) = \frac{1}{a_m}.$$

It follows that $\boldsymbol{E}(x) = H(x)\, S(x)$ is a fundamental solution if

$$\begin{cases} P(\partial)\, S(x) = 0 \\ S^{(j)}(0) = 0 \qquad \text{for } j = 0, \ldots, m-2 \\ S^{(m-1)}(0) = \frac{1}{a_m} \end{cases}$$

The support of the fundamental solution $\boldsymbol{E}$ is contained in the interval $[0, +\infty)$. Moreover, by construction, $\boldsymbol{E}|_{\mathcal{B}\mathbb{R}\setminus\{0\}} \in C^\infty(\mathbb{R}\setminus\{0\}; \mathbb{R})$. From Schwartz Theorem, the ordinary differential operator is hypoelliptic. Therefore, if

$$\sum_{j=0}^{m} a_j \partial^j u = f, \quad \text{where } f \in C^\infty(\mathbb{R}; \mathbb{R}),$$

the solution u, if it exists, is of class C^∞, and this means that the solutions of an ordinary differential equations are the "classical" ones.

Example

Consider the differential equation

$$\boldsymbol{E}'' + \boldsymbol{E} = \delta.$$

We are looking for a solution $\boldsymbol{E}(x) = H(x)\, S(x)$. We find

$$\begin{aligned} \boldsymbol{E}'(x) &= H S' + S(0)\, \delta \\ \boldsymbol{E}''(x) &= H S'' + S(0)\, \delta' + S'(0)\, \delta. \end{aligned}$$

Thus, $\boldsymbol{E}'' + \boldsymbol{E} = \delta$ implies

$$H\,(S'' + S'(0)\, \delta + S(0)\, \delta' + S) = \delta,$$

which gives

$$\begin{cases} S'' + S = 0 \\ S'(0) = 1 \\ S(0) = 0. \end{cases}$$

Since $\lambda^2 + 1 = 0$ gives $\lambda = \pm i$, the generic solution is of the form $S(x) = A \cos x + B \sin x$. Imposing the initial conditions yields $A = 0$ and $B = 1$, so that $S(x) = \sin x$. It follows that

$$\boldsymbol{E}(x) = H(x) \sin x.$$

Let us now consider the Cauchy problem

$$\begin{cases} P(\partial)\, u(x) = f(x), \qquad x \geq 0 \\ u(0) = u_0 \\ \vdots \\ u^{(m-1)}(0) = u_{m-1}, \end{cases}$$

where $P(\partial) = \sum_{j=0}^{m} a_j \partial^j$ with $a_m \neq 0$, $f \in C([0,+\infty);\mathbb{C})$, and $u_0, \ldots, u_{m-1} \in \mathbb{C}$. We define the operation of continuation of the solution over $\mathbb{R}$:

$$\begin{aligned} \tilde{u} : \mathbb{R} &\longrightarrow \mathbb{C} \\ x &\mapsto \begin{cases} u(x) & x \geq 0 \\ 0 & x < 0\,. \end{cases} \end{aligned}$$

We have

$$\tilde{u}'(x) = \widetilde{u'}(x) + u(0)\,\delta.$$

In other words, $\tilde{u}'$ coincides with the derivative of u plus the jump at 0. Hence

$$\begin{aligned} \tilde{u}''(x) &= \widetilde{u''}(x) + u'(0)\,\delta + u(0)\,\delta' \\ &\vdots \\ \tilde{u}^{(m)}(x) &= \widetilde{u^{(m)}}(x) + u^{(m-1)}(0)\,\delta + \ldots + u(0)\,\delta^{(m-1)} \\ &= \widetilde{u^{(m)}}(x) + \sum_{k=1}^{m-1} u^{(j-1-k)}(0)\,\delta^{(k)}. \end{aligned}$$

We now evaluate $P(\partial)\,\tilde{u}$:

$$\begin{aligned} P(\partial)\,\tilde{u} &= \sum_{j=0}^{m} a_j \partial^j \tilde{u} \\ &= a_0\,\tilde{u} + \sum_{j=1}^{m} a_j\,\tilde{u}^{(j)} \\ &= a_0\,\tilde{u} + \sum_{j=1}^{m} a_j \left(\widetilde{u^{(j)}} + \sum_{k=0}^{j-1} u^{(j-1-k)}(0)\,\delta^{(k)} \right) \\ &= \sum_{j=0}^{m} a_j\,\widetilde{u^{(j)}} + \sum_{j=1}^{m} a_j \left(\sum_{k=0}^{j-1} u^{(j-1-k)}(0)\,\delta^{(k)} \right) \\ &= \tilde{f} + \sum_{k=0}^{m-1} \left(\sum_{j=k+1}^{m} a_j\,u^{(j-1-k)}(0) \right) \delta^{(k)}. \end{aligned}$$

We can now use the fact that the fundamental solution that we have found has support in $[0,+\infty)$, and is therefore convolvable with any function:

$$\boldsymbol{E} * P(\partial)\,\tilde{u} = \boldsymbol{E} * \tilde{f} + \sum_{k=0}^{m-1} \left(\sum_{j=k+1}^{m} a_j\,u^{(j-1-k)}(0) \right) \boldsymbol{E} * \delta^{(k)}.$$

We can transfer the differential operator from one element of the convolution to the other in the left hand-side:

$$\boldsymbol{E} * P(\partial)\,\tilde{u} = P(\partial)\,\boldsymbol{E} * \tilde{u} = \delta * \tilde{u} = \tilde{u}.$$

Thus, we have obtained

$$\tilde{u} = \boldsymbol{E} * \tilde{f} + \sum_{k=0}^{m-1} \left(\sum_{j=k+1}^{m} a_j\,u^{(j-1-k)}(0) \right) \boldsymbol{E} * \delta^{(k)}.$$

Moreover, $\boldsymbol{E} * \delta^{(k)} = \boldsymbol{E}^{(k)} * \delta = \boldsymbol{E}^{(k)}$. Thus, for $x > 0$ we have

$$\boldsymbol{E} * \tilde{f}(x) = \int_{\mathbb{R}} \boldsymbol{E}(x-y)\,\tilde{f}(y)\,\mathrm{d}y = \int_0^{+\infty} \boldsymbol{E}(x-y)\,f(y)\,\mathrm{d}y = \int_0^x \boldsymbol{E}(x-y)\,f(y)\,\mathrm{d}y,$$

because both $\boldsymbol{E}$ and f have support in $[0, +\infty)$. We obtain, for $x > 0$,

$$u(x) = \int_0^x \boldsymbol{E}(x-y)\,f(y)\,\mathrm{d}y + \sum_{k=0}^{m-1} \left(\sum_{j=k+1}^{m} a_j\, u^{(j-1-k)}(0) \right) \boldsymbol{E}^{(k)}.$$

Example

Consider the Cauchy problem

$$\begin{cases} u''(x) + u(x) = f(x), & x \geq 0, \\ u(0) = u_0 \\ u'(0) = u_1 \end{cases}$$

We have already found the fundamental solution $\boldsymbol{E}(x) = H(x)\,S(x)$. Then,

$$\begin{aligned} u(x) &= \int_0^x \sin(x-y)\,f(y)\,\mathrm{d}y + \sum_{k=0}^{1} \left(\sum_{j=k+1}^{2} a_j\, u^{(j-1-k)}(0) \right) \sin^{(k)}(x) \\ &= \int_0^x \sin(x-y)\,f(y)\,\mathrm{d}y + \sum_{j=1}^{2} a_j\, u^{(j-1)}(0)\,\sin(x) + \sum_{j=2}^{2} a_j\, u^{(j-2)}(0)\,\cos(x) \\ &= \int_0^x \sin(x-y)\,f(y)\,\mathrm{d}y + u_1\,\sin(x) + u_2\,\cos(x), \end{aligned}$$

since $a_1 = 0$.

7.3 Heat equation

Consider the generic element $(t, \boldsymbol{x}) \in \mathbb{R}^{n+1}$, where $t \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$. The variable t in this equation is representative of time, while $\boldsymbol{x}$ represents spatial coordinates (typically $n = 3$). The heat equation describes the evolution of the temperature u at the time instant t at the point $\boldsymbol{x} \in \mathbb{R}^n$, where the function $f \in \mathcal{D}'(\mathbb{R}^{n+1}; \mathbb{R})$ represents the intensity of the heat source:

$$\frac{\partial u}{\partial t}(t, \boldsymbol{x}) - a^2\,\nabla^2_{\boldsymbol{x}} u(t, \boldsymbol{x}) = f(t, \boldsymbol{x}),$$

where $a > 0$ is a constant. The heat operator is the differential operator $P(\partial) = \partial_t - a^2\,\nabla^2_{\boldsymbol{x}}$, or, more explicitly,

$$P(\partial_t, \partial_{\boldsymbol{x}}) = \frac{\partial}{\partial t} - a^2\,\nabla^2_{\boldsymbol{x}}.$$

We want to find a fundamental solution $\boldsymbol{E} \in \mathcal{D}'(\mathbb{R}^{n+1}; \mathbb{C})$ that satisfies the equation

$$\frac{\partial \boldsymbol{E}}{\partial t} - a^2\,\nabla^2_{\boldsymbol{x}} \boldsymbol{E} = \delta(t, \boldsymbol{x}).$$

We now find a distribution $\boldsymbol{E}$, which we will then prove to be the sought fundamental solution.

We perform the Fourier transform only with respect to the spatial coordinates $\boldsymbol{x}$:

$$\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\boldsymbol{E}(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Hence,

$$\begin{aligned}\widehat{\frac{\partial \boldsymbol{E}}{\partial t}}(t,\boldsymbol{\xi}) &= \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\frac{\partial \boldsymbol{E}}{\partial t}(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \frac{\partial}{\partial t}\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\boldsymbol{E}(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \frac{\partial \hat{\boldsymbol{E}}}{\partial t}(t,\boldsymbol{\xi}).\end{aligned}$$

From $\delta(t,\boldsymbol{x}) = \delta(t)\otimes\delta(\boldsymbol{x})$, we find $\hat{\delta}(t,\boldsymbol{x}) = \delta(t)\otimes 1(\boldsymbol{\xi})$. Now, recalling that $\mathcal{F}\{\nabla^2 u\} = -\xi^2\,\hat{u}$, the Fourier transform of the equation $(\partial_t - a^2\,\nabla^2_{\xi})\boldsymbol{E}(t,\boldsymbol{x}) = \delta(t,\boldsymbol{x})$, which furnishes the fundamental solution, becomes

$$\frac{\partial \hat{\boldsymbol{E}}}{\partial t}(t,\boldsymbol{\xi}) + a^2\,\|\boldsymbol{\xi}\|^2\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = \delta(t)\otimes 1(\boldsymbol{\xi}).$$

We solve this equation as if it was an ordinary differential equation, by multiplying both sides by $\exp(a^2\,t\,\|\boldsymbol{\xi}\|^2)$:

$$e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\frac{\partial \hat{\boldsymbol{E}}}{\partial t}(t,\boldsymbol{\xi}) + a^2\,\|\boldsymbol{\xi}\|^2\,e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,(\delta(t)\otimes 1(\boldsymbol{\xi})). \tag{7.5}$$

We have $e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,(\delta(t)\otimes 1(\boldsymbol{\xi})) = \delta(t)\otimes 1(\boldsymbol{\xi})$. Indeed, given $\varphi\in\mathcal{D}(\mathbb{R}^{n+1};\mathbb{C})$ we have

$$\begin{aligned}\langle e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,(\delta(t)\otimes 1(\boldsymbol{\xi})),\varphi\rangle &= \langle(\delta(t)\otimes 1(\boldsymbol{\xi})), e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\varphi\rangle\\ &= \left\langle \delta(t), \int_{\mathbb{R}^n} e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\varphi(t,\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\right\rangle\\ &= \int_{\mathbb{R}^n}\varphi(0,\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} = \langle\delta(t)\otimes 1(\boldsymbol{\xi}),\varphi(t,\boldsymbol{\xi})\rangle.\end{aligned}$$

Moreover, the left hand-side of (7.5) is

$$e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\frac{\partial \hat{\boldsymbol{E}}}{\partial t}(t,\boldsymbol{\xi}) + a^2\,\|\boldsymbol{\xi}\|^2\,e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = \frac{\partial}{\partial t}\left(e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi})\right).$$

Thus, (7.5) can be written as

$$\frac{\partial}{\partial t}\left(e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi})\right) = \delta(t)\otimes 1(\boldsymbol{\xi}).$$

We obtain

$$e^{t\,a^2\,\|\boldsymbol{\xi}\|^2}\,\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = (H(t)+C)\otimes 1(\boldsymbol{\xi}),$$

where $C\in\mathbb{C}$ is a constant to be determined. This step is legitimate, because given two open sets $\Omega_1\subseteq\mathbb{R}^n$ and $\Omega_2\subseteq\mathbb{R}^m$ and two distributions $T_1\in\mathcal{D}'(\Omega_1;\mathbb{C})$ and $T_2\in\mathcal{D}'(\Omega_2;\mathbb{C})$, we have seen that for $\boldsymbol{x}\in\Omega_1$ and $\boldsymbol{y}\in\Omega_2$, and for all $\boldsymbol{\alpha}\in\mathbb{N}^n$ and $\boldsymbol{\beta}\in\mathbb{N}^m$

$$\partial^{\boldsymbol{\alpha}}_{\boldsymbol{x}}\,\partial^{\boldsymbol{\beta}}_{\boldsymbol{y}}\,(T_1\otimes T_2) = (\partial^{\boldsymbol{\alpha}}_{\boldsymbol{x}}\,T_1)\otimes(\partial^{\boldsymbol{\beta}}_{\boldsymbol{y}}\,T_2).$$

We obtain

$$\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2}\,(H(t)+C)\otimes 1(\boldsymbol{\xi}).$$

Thus,

$$\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = \begin{cases} (1+C)\, e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2} & \text{for } t > 0 \\ C\, e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2} & \text{for } t < 0. \end{cases}$$

If $t < 0$, we notice that the function $\boldsymbol{\xi} \mapsto e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2}$ is not tempered, and therefore we can only take the inverse Fourier transform with respect to $\boldsymbol{\xi}$ only if $C = 0$. We obtain

$$\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = \begin{cases} e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2} & \text{for } t > 0 \\ 0 & \text{for } t \leq 0, \end{cases}$$

or, more succinctly, $\hat{\boldsymbol{E}}(t,\boldsymbol{\xi}) = e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2}\, H(t)$. Taking the inverse Fourier transform with respect to the spatial coordinates gives

$$\begin{aligned} \boldsymbol{E}(t,\boldsymbol{x}) &= \frac{H(t)}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} e^{-t\,a^2\,\|\boldsymbol{\xi}\|^2}\, \mathrm{d}\boldsymbol{\xi} \\ &= \frac{H(t)}{(2\,\pi)^n} \int_{\mathbb{R}^n} e^{i\,\frac{\boldsymbol{x}}{a\,\sqrt{t}}\cdot\eta} e^{-\|\eta\|^2}\, (a\,\sqrt{t})^{-n} \mathrm{d}\eta \qquad (a\,\sqrt{t}\,\boldsymbol{\xi} = \eta) \\ &= \frac{H(t)}{(2\,\pi)^n}\, (a\,\sqrt{t})^n\, \pi^{\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}} \\ &= H(t)\, (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}}, \end{aligned}$$

where we have used the integral that we computed in the previous set of notes

$$\int_{\mathbb{R}^n} e^{-i\,\boldsymbol{y}\cdot\eta}\, e^{-\|\eta\|^2}\, \mathrm{d}\eta = \pi^{\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{y}\|^2}{4}}.$$

We now have to verify that

$$\boldsymbol{E}(t,\boldsymbol{x}) = \begin{cases} (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}} & \text{for } t > 0 \\ 0 & \text{for } t \leq 0 \end{cases}$$

is a fundamental solution. Since the exponent involves the norm of $\boldsymbol{x}$, the function $\boldsymbol{E}(t,\boldsymbol{x})$ is certainly of class C^∞ if we exclude the hyperplane $\boldsymbol{x} = 0$. However, we can show that it is of class C^∞ except for the origin. In other words, we prove that for all $\boldsymbol{x}_0 \neq 0$, we have

$$\lim_{(t,\boldsymbol{x}) \to (0,\boldsymbol{x}_0)} \boldsymbol{E}(t,\boldsymbol{x}) = 0.$$

In fact, if $\boldsymbol{x}$ is sufficiently close to $\boldsymbol{x}_0$, we will have $\|\boldsymbol{x}\| \geq \frac{\|\boldsymbol{x}_0\|}{2}$, so that

$$|\boldsymbol{E}(t,\boldsymbol{x})| \leq (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}_0\|^2}{16\,a^2\,t}},$$

and the limit

$$\lim_{t \to 0^+} (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}_0\|^2}{16\,a^2\,t}}$$

is of the kind $\lim_{t \to 0^+} t^{-m}\, e^{-\frac{\alpha}{t}} = \lim_{\tau \to +\infty} \tau^m\, e^{-\alpha\,\tau} = 0$ (with $\tau = 1/t$). This shows that the limit is 0, and this shows that $\boldsymbol{E}(t,\boldsymbol{x})$ is continuous in $\mathbb{R}^{n+1} \setminus \{0\}$. We show that the same is true for its derivatives. The partial derivative with respect to x_j is

$$\frac{\partial \boldsymbol{E}}{\partial x_j}(t,\boldsymbol{x}) = \begin{cases} (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\, e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}} \left(-\frac{x_j}{2\,a^2\,t}\right) & \text{for } t > 0 \\ 0 & \text{for } t \leq 0, \end{cases}$$

and that with respect to t is

$$\frac{\partial \boldsymbol{E}}{\partial t}(t,\boldsymbol{x}) = \begin{cases} (4\,\pi\,a^2)^{-\frac{n}{2}}\,t^{-\frac{n}{2}-1}\,e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}} + (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\,e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}}\left(\frac{\|\boldsymbol{x}\|}{4\,a^2\,t^2}\right) & \text{for } t > 0 \\ 0 & \text{for } t \le 0, \end{cases}$$

Using the same argument, it is easy to see that, if $\boldsymbol{x}_0 \neq 0$, we get

$$\lim_{t\to 0^+} \frac{\partial \boldsymbol{E}}{\partial t}(t,\boldsymbol{x}_0) = 0.$$

We find that all the first derivatives of $\boldsymbol{E}$ are continuous in $\mathbb{R}^{n+1} \setminus \{0\}$. For $t > 0$, the higher-order derivatives are linear combinations of terms in the form

$$t^{-m}\,e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}}\,P(\boldsymbol{x},t),$$

where P is a polynomial in $\boldsymbol{x}$ and t, and, again, the limit as $(t,\boldsymbol{x}) \to (0,\boldsymbol{x}_0)$, for $\boldsymbol{x}_0 \neq 0$, is 0. Since $\boldsymbol{E}$ is continuous at $(0,\boldsymbol{x}_0)$, and since the limit as $(t,\boldsymbol{x}) \to (0,\boldsymbol{x}_0)$ exists for all its derivatives, we conclude that $\boldsymbol{E}$ is differentiable an arbitrary number of times in $\mathbb{R}^{n+1} \setminus \{0\}$, i.e., $\boldsymbol{E} \in C^\infty(\mathbb{R}^{n+1} \setminus \{0\}; \mathbb{R})$.

Now, we show that $\boldsymbol{E}$ is locally integrable, i.e., $\boldsymbol{E} \in L^1_{\text{loc}}(\mathbb{R}^{n+1}; \mathbb{R})$. This says that $\boldsymbol{E}$ is a regular distribution. To this aim, we prove that $\boldsymbol{E}$ is integrable over every set in the form $\{(t,\boldsymbol{x}) \in \mathbb{R}^{n+1} \,|\, t \le T\}$ for all $T > 0$. Indeed, $\boldsymbol{E}(t,\boldsymbol{x}) > 0$ in $\mathbb{R}^{n+1}$, so that using Tonelli's Theorem

$$\begin{aligned} \int_{(-\infty,T]\times\mathbb{R}^n} \boldsymbol{E}(t,\boldsymbol{x})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{x} &= \int_0^T \left(\int_{\mathbb{R}^n} (4\,\pi\,a^2\,t)^{-\frac{n}{2}}\,e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,t}}\,\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t \\ &= \int_0^T (4\,\pi\,a^2\,t)^{-\frac{n}{2}} \left(\int_{\mathbb{R}^n} e^{-\|\boldsymbol{y}\|^2}\,(2\,a\,\sqrt{t})^n\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}t \qquad \text{where } \frac{\boldsymbol{x}}{\sqrt{4\,a^2\,t}} = \boldsymbol{y} \\ &= \int_0^T (4\,\pi\,a^2\,t)^{-\frac{n}{2}} \left(\int_{\mathbb{R}^n} e^{-\|\boldsymbol{y}\|^2}\,(4\,a^2\,t)^{\frac{n}{2}}\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}t \\ &= \pi^{-\frac{n}{2}} \int_0^T \left(\int_{\mathbb{R}^n} e^{-\|\boldsymbol{y}\|^2}\,\mathrm{d}\boldsymbol{y}\right)\mathrm{d}t = T, \end{aligned}$$

where we used the fact that $\int_{\mathbb{R}} e^{-y^2}\,\mathrm{d}y = \sqrt{\pi}$. It follows that $\boldsymbol{E} \in L^1_{\text{loc}}(\mathbb{R}^{n+1}; \mathbb{R})$. We are now ready to prove that $\boldsymbol{E}$ is the fundamental solution of the heat operator. In other words, we have to verify that

$$\frac{\partial \boldsymbol{E}}{\partial t} - a^2\,\nabla^2_{\boldsymbol{x}}\boldsymbol{E} = \delta(t,\boldsymbol{x}).$$

Let $\varphi \in \mathcal{D}(\mathbb{R}^{n+1}; \mathbb{R})$. We need to show that

$$\left\langle \frac{\partial \boldsymbol{E}}{\partial t} - a^2\,\nabla^2_{\boldsymbol{x}}\boldsymbol{E}, \varphi \right\rangle = \varphi(0,0).$$

Noting that $\frac{\partial}{\partial t}$ is a derivative of odd order while $\nabla^2_{\boldsymbol{x}}$ is a derivative of even order, we get

$$\begin{aligned} \left\langle \frac{\partial \boldsymbol{E}}{\partial t} - a^2\,\nabla^2_{\boldsymbol{x}}\boldsymbol{E}, \varphi \right\rangle &= -\left\langle \boldsymbol{E}, \frac{\partial \varphi}{\partial t} + a^2\,\nabla^2_{\boldsymbol{x}}\varphi \right\rangle \\ &= -\int_{\mathbb{R}^{n+1}} \boldsymbol{E}(t,\boldsymbol{x}) \left(\frac{\partial \varphi}{\partial t}(t,\boldsymbol{x}) + a^2\,\nabla^2_{\boldsymbol{x}}\varphi(t,\boldsymbol{x})\right)\mathrm{d}t\,\mathrm{d}\boldsymbol{x} \\ &= -\int_0^{+\infty} \left(\int_{\mathbb{R}^n} \boldsymbol{E}(t,\boldsymbol{x}) \left(\frac{\partial \varphi}{\partial t}(t,\boldsymbol{x}) + a^2\,\nabla^2_{\boldsymbol{x}}\varphi(t,\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t \\ &= -\lim_{\varepsilon\to 0^+} \int_\varepsilon^{+\infty} \left(\int_{\mathbb{R}^n} \boldsymbol{E}(t,\boldsymbol{x}) \left(\frac{\partial \varphi}{\partial t}(t,\boldsymbol{x}) + a^2\,\nabla^2_{\boldsymbol{x}}\varphi(t,\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t, \end{aligned}$$

where we consider the limit to avoid the singularity at $(0,0)$. Consider the first addend. Integration by parts yields:

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\int_\varepsilon^{+\infty}\left(\int_{\mathbb{R}^n}\boldsymbol{E}(t,\boldsymbol{x})\,\frac{\partial\varphi}{\partial t}(t,\boldsymbol{x})\,\mathrm{d}x\right)\mathrm{d}\boldsymbol{t} &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}\left(\int_\varepsilon^{+\infty}\boldsymbol{E}(t,\boldsymbol{x})\,\frac{\partial\varphi}{\partial t}(t,\boldsymbol{x})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\\ &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}\left(\left[\boldsymbol{E}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\right]_{t=\varepsilon}^{t=+\infty}-\int_\varepsilon^{+\infty}\frac{\partial\boldsymbol{E}}{\partial t}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\\ &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}\left(-\boldsymbol{E}(\varepsilon,\boldsymbol{x})\,\varphi(\varepsilon,\boldsymbol{x})-\int_\varepsilon^{+\infty}\frac{\partial\boldsymbol{E}}{\partial t}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\end{aligned}$$

Consider the second addend:

$$\lim_{\varepsilon\to 0^+}\int_\varepsilon^{+\infty}\left(\int_{\mathbb{R}^n}\boldsymbol{E}(t,\boldsymbol{x})\,\left(a^2\,\nabla_{\boldsymbol{x}}^2\varphi(t,\boldsymbol{x})\right)\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t = \lim_{\varepsilon\to 0^+}\int_\varepsilon^{+\infty}\left(\int_{\mathbb{R}^n}a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t$$

Putting these two terms back together gives

$$\begin{aligned}\left\langle\frac{\partial\boldsymbol{E}}{\partial t}-a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E},\varphi\right\rangle &= \lim_{\varepsilon\to 0^+}\left[\int_{\mathbb{R}^n}\left(\boldsymbol{E}(\varepsilon,\boldsymbol{x})\,\varphi(\varepsilon,\boldsymbol{x})+\int_\varepsilon^{+\infty}\frac{\partial\boldsymbol{E}}{\partial t}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\right.\\ &\qquad\left.-\int_\varepsilon^{+\infty}\left(\int_{\mathbb{R}^n}a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E}(t,\boldsymbol{x})\,\varphi(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)\mathrm{d}t\right]\\ &= \lim_{\varepsilon\to 0^+}\left[\int_{\mathbb{R}^n}\boldsymbol{E}(\varepsilon,\boldsymbol{x})\,\varphi(\varepsilon,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}+\int_{\mathbb{R}^n}\left(\int_\varepsilon^{+\infty}\left(\frac{\partial\boldsymbol{E}}{\partial t}(t,\boldsymbol{x})-a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E}(t,\boldsymbol{x})\right)\varphi(t,\boldsymbol{x})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{x}\right].\end{aligned}$$

Since $\frac{\partial\boldsymbol{E}}{\partial t}(t,\boldsymbol{x})-a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E}(t,\boldsymbol{x})=0$ if $t>0$, we find

$$\begin{aligned}\left\langle\frac{\partial\boldsymbol{E}}{\partial t}-a^2\,\nabla_{\boldsymbol{x}}^2\boldsymbol{E},\varphi\right\rangle &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}\boldsymbol{E}(\varepsilon,\boldsymbol{x})\,\varphi(\varepsilon,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}(4\,\pi\,a^2\,\varepsilon)^{-\frac{n}{2}}\,e^{-\frac{\|\boldsymbol{x}\|^2}{4\,a^2\,\varepsilon}}\,\varphi(\varepsilon,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n}(4\,\pi\,a^2\,\varepsilon)^{-\frac{n}{2}}\,e^{-\|\boldsymbol{y}\|^2}\,\varphi(\varepsilon,\sqrt{4\,a^2\,\varepsilon})\,(4\,\pi\,a^2\,\varepsilon)^{\frac{n}{2}}\,\mathrm{d}\boldsymbol{y} \qquad \frac{\boldsymbol{x}}{\sqrt{4\,a^2\,\varepsilon}}=\boldsymbol{y}\\ &= \lim_{\varepsilon\to 0^+}\pi^{-\frac{n}{2}}\int_{\mathbb{R}^n}e^{-\|\boldsymbol{y}\|^2}\,\varphi(\varepsilon,\sqrt{4\,a^2\,\varepsilon})\,\mathrm{d}\boldsymbol{y}\\ &= \pi^{-\frac{n}{2}}\int_{\mathbb{R}^n}\lim_{\varepsilon\to 0^+}e^{-\|\boldsymbol{y}\|^2}\,\varphi(\varepsilon,\sqrt{4\,a^2\,\varepsilon})\,\mathrm{d}\boldsymbol{y}\\ &= \pi^{-\frac{n}{2}}\int_{\mathbb{R}^n}e^{-\|\boldsymbol{y}\|^2}\,\varphi(0,0)\,\mathrm{d}\boldsymbol{y}=\varphi(0,0),\end{aligned}$$

where it was legitimate to use Lebesgue's Dominated Convergence Theorem since

$$\left|e^{-\|\boldsymbol{y}\|^2}\,\varphi(\varepsilon,\sqrt{4\,a^2\,\varepsilon})\right|\le\left|e^{-\|\boldsymbol{y}\|^2}\right|\sup_{\mathbb{R}^{n+1}}|\varphi|$$

from $\varphi\in\mathcal{D}(\mathbb{R}^{n+1};\mathbb{R})$, and this quantity is integrable.

The fundamental solution $\boldsymbol{E}(t,\boldsymbol{x})$ is of class C^∞ everywhere except at $(0,0)$, by giving it the value 0 over the set $\{(t,\boldsymbol{x})\in\mathbb{R}^{n+1}\,|\,t=0\}$, and therefore, from Schwartz Theorem, the operator $P(\partial)=\partial_t-a^2\,\nabla_{\boldsymbol{x}}^2$ is hypoelliptic.

7.4 Wave equation

Consider the equation

$$\frac{\partial^2 u}{\partial t^2} - a^2 \nabla_{\boldsymbol{x}}^2 u = f,$$

where the variable $\boldsymbol{x}$ lives in $\mathbb{R}^3$ and $t \in \mathbb{R}$. The wave operator is

$$P(\partial) = \partial_t^2 - a^2 \nabla_{\boldsymbol{x}}^2.$$

We are looking for a distribution $\boldsymbol{E} \in \mathcal{D}'(\mathbb{R}^4; \mathbb{R})$ such that

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2}(t, \boldsymbol{x}) - a^2 \nabla_{\boldsymbol{x}}^2 \boldsymbol{E}(t, \boldsymbol{x}) = \delta(t, \boldsymbol{x}).$$

We proceed as we did for the heat equation, by formally determining a distribution $\boldsymbol{E}(t, \boldsymbol{x})$, which, subsequently, we will prove to be a fundamental solution. Taking the Fourier transform with respect to the spatial variable $\boldsymbol{x}$

$$\hat{\boldsymbol{E}}(t, \boldsymbol{\xi}) = \int_{\mathbb{R}^3} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} \boldsymbol{E}(t, \boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

and writing $\delta(t, \boldsymbol{x}) = \delta(t) \otimes \delta(\boldsymbol{x})$, we find the equation

$$\frac{\partial^2 \hat{\boldsymbol{E}}}{\partial t^2}(t, \boldsymbol{\xi}) + a^2 \|\boldsymbol{\xi}\|^2 \hat{\boldsymbol{E}}(t, \boldsymbol{\xi}) = \delta(t) \otimes 1(\boldsymbol{\xi}).$$

Viewing $\boldsymbol{\xi}$ as a parameter, this equation can be seen as an ordinary differential equation in t of the form $w''_{\boldsymbol{\xi}}(t) + a^2 \|\boldsymbol{\xi}\|^2 w_{\boldsymbol{\xi}}(t) = \delta(t)$. The solution is the fundamental solution of the ordinary differential operator $P(\partial) = \partial^2 + a^2 \|\boldsymbol{\xi}\|^2$, and it is sought in the form $w_{\boldsymbol{\xi}}(t) = H(t)\, S(t)$, and the function S is the solution of the Cauchy problem

$$\begin{cases} S''(t) + a^2 \|\boldsymbol{\xi}\|^2 S(t) = 0 \\ S(0) = 0 \\ S'(0) = 1. \end{cases}$$

The characteristic equation is $\lambda^2 + a^2 \|\boldsymbol{\xi}\|^2 = 0$, which leads to $\lambda = \pm i\, a \|\boldsymbol{\xi}\|$. It follows that the general solution is

$$S(t) = A \cos(a\, t \|\boldsymbol{\xi}\|) + B \sin(a\, t \|\boldsymbol{\xi}\|).$$

The constants A and B are obtained by imposing the initial conditions. The condition $S(0) = 0$ gives $A = 0$. Hence

$$S'(t) = -A\, a \|\boldsymbol{\xi}\| \sin(a\, t \|\boldsymbol{\xi}\|) + a \|\boldsymbol{\xi}\| \cos(a\, t \|\boldsymbol{\xi}\|),$$

and imposing $S'(0) = a \|\boldsymbol{\xi}\| B = 1$ we find $B = \frac{1}{a \|\boldsymbol{\xi}\|}$. We obtain

$$w_{\boldsymbol{\xi}}(t) = H(t) \frac{\sin(a\, t \|\boldsymbol{\xi}\|)}{a \|\boldsymbol{\xi}\|}.$$

This means that, formally,

$$\hat{\boldsymbol{E}}(t, \boldsymbol{\xi}) = H(t) \frac{\sin(a\, t \|\boldsymbol{\xi}\|)}{a \|\boldsymbol{\xi}\|}.$$

We now take the inverse Fourier transform with respect to the coordinate $\boldsymbol{\xi}$. To this end, we recall the example of the Fourier transform of a distributed mass over a spherical surface

$$\langle T_R, \varphi \rangle = \int_{\|\boldsymbol{x}\| = R} \varphi(\boldsymbol{x}) \, \mathrm{d}S(\boldsymbol{x}),$$

which is

$$\hat{T}_R(\boldsymbol{\xi}) = 4\pi R \frac{\sin(R\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}.$$

Now, taking $T = a\,t$, we find

$$\hat{T}_{a\,t}(\boldsymbol{\xi}) = \frac{4\pi a\,t\,\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}.$$

We use this result by writing

$$\hat{\boldsymbol{E}}(t, \boldsymbol{\xi}) = \frac{H(t)}{4\pi a^2 t} \frac{4\pi a\,t\,\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|} \frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{a\,\|\boldsymbol{\xi}\|} = \frac{H(t)}{4\pi a^2 t} \hat{T}_{a\,t}(\boldsymbol{\xi}).$$

Therefore,

$$\boldsymbol{E}(t, \boldsymbol{x}) = \frac{H(t)}{4\pi a^2 t} T_{a\,t}(\boldsymbol{x}).$$

This distribution acts in the following way: given $\varphi \in \mathcal{D}(\mathbb{R}^4; \mathbb{R})$,

$$\langle \boldsymbol{E}(t, \boldsymbol{x}), \varphi(t, \boldsymbol{x}) \rangle = \int_0^{+\infty} \frac{1}{4\pi a^2 t} \left(\int_{\|\boldsymbol{x}\| = a\,t} \varphi(t, \boldsymbol{x}) \, \mathrm{d}S(\boldsymbol{x}) \right) \mathrm{d}t.$$

The domain of integration is a sphere whose radius increases with t. The problem is the division by t. However, considering the inner integral, we find that

$$\left| \int_{\|\boldsymbol{x}\| = a\,t} \varphi(t, \boldsymbol{x}) \, \mathrm{d}S(\boldsymbol{x}) \right| \leq \sup |\varphi| \, 4\pi a^2 t^2,$$

and therefore

$$\left| \frac{1}{4\pi a^2 t} \int_{\|\boldsymbol{x}\| = a\,t} \varphi(t, \boldsymbol{x}) \, \mathrm{d}S(\boldsymbol{x}) \right| \leq \frac{1}{4\pi a^2 t} \sup |\varphi| \, 4\pi a^2 t^2 = t \sup |\varphi|,$$

which shows that at 0 the integrand is asymptotically equivalent to t. This proves that there are no problems of integrability at 0. At ∞, we have no problems of integrability because φ has compact support. It is also possible to show that $\boldsymbol{E}$ is a distribution over $\mathbb{R}^4$. Certainly, $\boldsymbol{E}$ cannot be a distribution of function type. In fact, from

$$\langle \boldsymbol{E}(t, \boldsymbol{x}), \varphi(t, \boldsymbol{x}) \rangle = \int_0^{+\infty} \frac{1}{4\pi a^2 t} \left(\int_{\|\boldsymbol{x}\| = a\,t} \varphi(t, \boldsymbol{x}) \, \mathrm{d}S(\boldsymbol{x}) \right) \mathrm{d}t$$

we see immediately that

$$\operatorname{supp} \boldsymbol{E} = \{(t, \boldsymbol{x}) \in \mathbb{R}^4 \,|\, t \geq 0, \ \|\boldsymbol{x}\| = a\,t\},$$

and therefore, loosely speaking, for a given t, this set is a sphere. More precisely,

$$\mu_4(\operatorname{supp} \boldsymbol{E}) \leq \int_0^{+\infty} \mu_3(S_{a\,t}(0)) \, \mathrm{d}t,$$

and this upper bound is zero, because $\mu_3(S_{a\,t}(0)) = 0$: this is the measure of a sphere in $\mathbb{R}^3$ is zero. Stated differently, $\operatorname{supp} \boldsymbol{E}$ has zero measure in $\mathbb{R}^4$ because each section has measure zero in $\mathbb{R}^3$. It is also easy to

show that if $T \in \mathcal{D}'(\mathbb{R}^n;\mathbb{C})$ with $\mu_n(\operatorname{supp} T) = 0$, then if T is of function type, we have $T(\boldsymbol{x}) = 0$ almost everywhere: this comes from the fact that the support of T as a distribution coincides with the support of T as a function. Applying this result to our case, we obtain that if $\boldsymbol{E}$ was of function type, it would be the zero distribution. We will show that

$$\frac{\partial^2 \boldsymbol{E}}{\partial t^2} - a^2\,\nabla_{\boldsymbol{x}}^2 \boldsymbol{E} = \delta,$$

and therefore $\boldsymbol{E}$ cannot be the zero distribution. We can conclude from this fact that $\boldsymbol{E}$ is not hypoelliptic.

To prove that $\boldsymbol{E}$ is a fundamental solution, we use the following result.

Lemma 7.1 *Let $\varphi \in \mathcal{S}(\mathbb{R}^4;\mathbb{R})$. Let*

$$\hat{\varphi}(t,\boldsymbol{\xi}) = \int_{\mathbb{R}^3} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\varphi(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Then:

(i) $\hat{\varphi} \in \mathcal{S}(\mathbb{R}^4;\mathbb{R})$;

(ii) there holds

$$\frac{\partial\hat{\varphi}}{\partial t}(t,\boldsymbol{\xi}) = \int_{\mathbb{R}^3} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\frac{\partial\varphi}{\partial t}(t,\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We are now ready to show that $\boldsymbol{E}$ is a fundamental solution of the wave equation in $\mathbb{R}^3$. Let $\varphi \in \mathcal{D}(\mathbb{R}^4;\mathbb{R})$. Using the "hat" to denote the Fourier transform with respect to $a\,\boldsymbol{x}$, by considering t as a parameter, we can write in general $\check{w}(t,\boldsymbol{x}) = w(t,-\boldsymbol{x})$. We have

$$\begin{aligned}
\langle T_{a\,t}(\boldsymbol{x}), \varphi(t,\boldsymbol{x})\rangle &= \frac{1}{(2\,\pi)^3}\,\langle T_{a\,t}(\boldsymbol{x}), \check{\hat{\varphi}}(t,\boldsymbol{x})\rangle \\
&= \frac{1}{(2\,\pi)^3}\,\langle T_{a\,t}(\boldsymbol{x}), \hat{\check{\varphi}}(t,\boldsymbol{x})\rangle \\
&= \frac{1}{(2\,\pi)^3}\,\langle \hat{T}_{a\,t}(\boldsymbol{\xi}), \check{\varphi}(t,\boldsymbol{x})\rangle \\
&= \frac{1}{(2\,\pi)^3}\int_{\mathbb{R}^3} \frac{4\,\pi\,a\,t\,\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\hat{\varphi}(t,-\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\langle \boldsymbol{E}(t,\boldsymbol{x}), \varphi(t,\boldsymbol{x})\rangle &= \int_0^{+\infty} \frac{1}{4\,\pi\,a^2\,t}\,\langle T_{a\,t}(\boldsymbol{x}), \varphi(t,\boldsymbol{x})\rangle\,\mathrm{d}t \\
&= \int_0^{+\infty} \frac{1}{4\,\pi\,a^2\,t}\,\frac{1}{(2\,\pi)^3}\int_{\mathbb{R}^3} \frac{4\,\pi\,a\,t\,\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\hat{\varphi}(t,-\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\,\mathrm{d}t \\
&= \frac{1}{8\,\pi^3\,a}\int_0^{+\infty}\left(\int_{\mathbb{R}^3} \frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\hat{\varphi}(t,-\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\right)\mathrm{d}t.
\end{aligned}$$

We now find how the distribution $(\partial_t^2 - a^2\,\nabla_{\boldsymbol{x}}^2)\,\boldsymbol{E} \in \mathcal{D}'(\mathbb{R}^4;\mathbb{R})$ operates on $\varphi \in \mathcal{D}(\mathbb{R}^4;\mathbb{R})$, using the second point of the Lemma:

$$\begin{aligned}
\langle \partial_t^2 \boldsymbol{E} - a^2\,\nabla_{\boldsymbol{x}}^2\,\boldsymbol{E}, \varphi\rangle &= \langle \boldsymbol{E}, \partial_t^2\varphi - a^2\,\nabla_{\boldsymbol{x}}^2\,\varphi\rangle \\
&= \frac{1}{8\,\pi^3\,a}\int_0^{+\infty}\left(\int_{\mathbb{R}^3} \frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\left[\frac{\partial^2}{\partial t^2}\hat{\varphi}(t,-\boldsymbol{\xi}) + a^2\,\|\boldsymbol{\xi}\|^2\,\hat{\varphi}(t,-\boldsymbol{\xi})\right]\mathrm{d}\boldsymbol{\xi}\right)\mathrm{d}t.
\end{aligned}$$

Consider the first addend, and we change variable $\boldsymbol{\xi} := -\boldsymbol{\xi}$ and integrate by parts:

$$\begin{aligned}
\frac{1}{8\,\pi^3\,a}\int_0^{+\infty}\left(\int_{\mathbb{R}^3}\frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\frac{\partial^2}{\partial t^2}\hat{\varphi}(t,-\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\right)\mathrm{d}t &= \frac{1}{8\,\pi^3\,a}\int_0^{+\infty}\left(\int_{\mathbb{R}^3}\frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\frac{\partial^2}{\partial t^2}\hat{\varphi}(t,\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}\right)\mathrm{d}t \\
&= \frac{1}{8\,\pi^3\,a}\int_{\mathbb{R}^3}\left(\int_0^{+\infty}\frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\frac{\partial^2}{\partial t^2}\hat{\varphi}(t,\boldsymbol{\xi})\,\mathrm{d}t\right)\mathrm{d}\boldsymbol{\xi} \\
&= \frac{1}{8\,\pi^3\,a}\int_{\mathbb{R}^3}\left(\left[\frac{\sin(a\,t\,\|\boldsymbol{\xi}\|)}{\|\boldsymbol{\xi}\|}\,\frac{\partial}{\partial t}\hat{\varphi}(t,\boldsymbol{\xi})\right]_{t=0}^{t=+\infty}\right. \\
&\quad \left.-\int_0^{+\infty} a\,\cos(a\,t\,\|\boldsymbol{\xi}\|)\,\frac{\partial}{\partial t}\hat{\varphi}(t,\boldsymbol{\xi})\,\mathrm{d}t\right)\boldsymbol{\xi}
\end{aligned}$$

8 Laplace transform

In 1782 Pierre-Simon Laplace considered the integral equation

$$f(x) = \int_{-\infty}^{+\infty} e^{-x\,y}\, g(y)\, \mathrm{d}y$$

where the unknown is $g(y)$. The closed-form solution to this problem was obtained in 1823 by Siméon Denis Poisson:

$$g(y) = \frac{1}{2\,\pi\,j} \int_{a-j\,\infty}^{a+j\,\infty} e^{x\,y}\, f(x)\, \mathrm{d}x,$$

where a is sufficiently big.

In 1811 Joseph Fourier considered the integral equation

$$f(x) = \int_{0}^{+\infty} \cos(x\,y)\, g(y)\, \mathrm{d}y$$

and its solution

$$g(y) = \frac{2}{\pi} \int_{0}^{+\infty} \cos(x\,y)\, f(x)\, \mathrm{d}x.$$

The names of *Laplace transform* and *Fourier transform* were therefore given to integral transformations which are similar to those introduced by Laplace and by Fourier, and which are largely employed in several fields of physics and engineering.

8.1 Definition of Laplace transform

The Fourier transform allows to transform ordinary differential equations into algebraic equations, and from partial differential equations with suitable structure to ordinary differential equations. On the other hand, in the solution of a Cauchy problem of an ordinary differential equation on the real line, the Fourier transform is not the ideal instrument. Indeed, we have seen that the biggest limitation of the Fourier transform is the fact that it can be applied only to functions (or distributions) that do not increase too fast. On the other hand, the solutions of differential equations are often not temperate.

Example Consider the Cauchy problem

$$\begin{cases} u'(x) = u(x) & x \geq 0 \\ u(0) = 1 \end{cases} \tag{8.1}$$

We already know that the solution is the function $u(x) = e^x$. Let us pretend not to know this, and let us attempt to solve this problem using the Fourier transform. First, the Fourier transform is performed on functions

defined over $\mathbb{R}$, and therefore we must consider the trivial continuation of u:

$$\tilde{u}(x) = \begin{cases} u(x) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

We recall that $\tilde{u}'$ coincides with the derivative of u plus the jump at 0, and therefore

$$\begin{aligned} \tilde{u}'(x) &= \widetilde{u'}(x) + u(0)\,\delta(x) \\ &= \widetilde{u'}(x) + \delta(x) \\ &= \tilde{u}(x) + \delta(x), \end{aligned}$$

where the last equality follows from the differential equation of the given Cauchy problem for the trivial continuation of $u(x)$, which reads as $\widetilde{u'} = \tilde{u}$. We can now apply the Fourier transform to $\tilde{u}' = \tilde{u} + \delta$, and we obtain

$$i\,\xi\,\mathcal{F}\{\tilde{u}\}(\xi) = \mathcal{F}\{\tilde{u}\} + 1 \quad \Rightarrow \quad \mathcal{F}\{\tilde{u}\}(\xi) = \frac{1}{i\,\xi - 1}.$$

Clearly, if we use the inverse Fourier transform, we will not obtain $u(x)$, because this function is not tempered. We ask ourselves what we obtain then. Let us denote by z the result of the inverse Fourier transform:

$$z = \mathcal{F}^{-1}\left\{\frac{1}{i\,\xi - 1}\right\}.$$

This can be re-written as

$$\hat{z} = \frac{1}{i\,\xi - 1} \quad \Rightarrow \quad (i\,\xi - 1)\,\hat{z} = 1 \quad \Rightarrow \quad \mathcal{F}\{z' - z\} = \hat{\delta} \quad \Rightarrow \quad z' = z + \delta.$$

The analogy with the equation $\tilde{u}' = \tilde{u} + \delta$ seems to suggest we have really obtained the solution of the given Cauchy problem. Let us try to solve the linear differential equation $z' = z + \delta$ as if it was an ordinary differential equation: we multiply both sides by an exponential whose exponent is a primitive of -1:

$$e^{-x}\,z'(x) - e^{-x}\,z(x) = e^{-x}\,\delta(x) = \delta(x) \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}(e^{-x}\,z(x)) = \delta(x).$$

It follows that $e^{-x}\,z(x) = H(x) + C$, where C is the integration constant, and therefore

$$z(x) = (H(x) + C)\,e^{x}.$$

We have to choose the constant C in such a way that z is tempered. From

$$z(x) = \begin{cases} (C+1)\,e^{x} & \text{if } x > 0 \\ C\,e^{x} & \text{if } x < 0 \end{cases}$$

Thus, we must have $C = -1$, which leads to

$$z(x) = -e^{x}\,H(-x) \quad \Rightarrow \quad \hat{z}(\xi) = \frac{1}{i\,\xi - 1}.$$

On the other hand, this is not the trivial continuation of the solution to the Cauchy problem, and the reason is that the Fourier transform only delivers tempered objects.

We can now understand why it is useful to introduce another kind of transformation, which ideally should have similar properties to the Fourier transform (it still needs to transform an ordinary differential equation into an algebraic one), but which include exponential functions: the Laplace transform fulfils these requirements.

One of the fundamental limitations of the Laplace transform is that it appears to be less useful in the

case of functions of more than one variable.

There are several different definitions in the literature of the Laplace transform. Here we consider the so-called *unilateral Laplace transform*.

Definition 8.1 [$\mathcal{L}$-Transformable Functions]
Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$. We say that f is Laplace transformable *or $\mathcal{L}$-transformable if*

(i) $\operatorname{supp} f \subseteq [0,+\infty)$;

(ii) $\exists\, \sigma \in \mathbb{R} : \quad e^{-\sigma t} f(t) \in L^1(\mathbb{R};\mathbb{C})$.

We observe that the second condition is a short way of expressing the fact that the function

$$\begin{aligned} A &\longrightarrow \mathbb{C} \\ t &\mapsto e^{-\sigma t} f(t), \end{aligned}$$

where A is the domain of f, is in $L^1(\mathbb{R};\mathbb{C})$. This abuse of notation is very widespread, and usually does not lead to ambiguities.

Now let us consider the set of σ for which the second property in Definition 8.1 holds, i.e., $\{\sigma \in \mathbb{R} \,|\, e^{-\sigma t} f(t) \in L^1(\mathbb{R};\mathbb{C})\}$. If f is Laplace transformable, this set is not empty, and therefore it admits infimum in $\overline{\mathbb{R}}$. Such infimum of this set is called *abscissa of absolute convergence*:

Definition 8.2 [Abscissa of Absolute Convergence]
Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. We define the abscissa of absolute convergence *as*

$$\sigma_a(f) \stackrel{def}{=} \inf\{\sigma \in \mathbb{R} \,|\, e^{-\sigma t} f(t) \in L^1(\mathbb{R};\mathbb{C})\}.$$

Let us now consider an $\mathcal{L}$-transformable function f, and let σ denote a real number such that $e^{-\sigma t} f(t) \in L^1(\mathbb{R};\mathbb{C})$. If we consider $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma$, then

$$|e^{-s t} f(t)| = e^{-\Re\mathrm{e}\{s\} t} |f(t)| \leq e^{-\sigma t} |f(t)|.$$

Since $e^{-\sigma t} f(t) \in L^1(\mathbb{R};\mathbb{C})$, then also $e^{-s t} f(t) \in L^1(\mathbb{R};\mathbb{C})$, since

$$\int_{\mathbb{R}} |e^{-\sigma t} f(t)| \, \mathrm{d}t \leq \int_{\mathbb{R}} e^{-\sigma t} |f(t)| \, \mathrm{d}t < +\infty.$$

Therefore, recalling that since $f(t)$ is zero for $t < 0$ because it is Laplace transformable, if we take any value $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma$, then the integral

$$\int_0^{+\infty} f(t)\, e^{-s t} \, \mathrm{d}t \tag{8.2}$$

exists. Also, if we take any value $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma_a(f)$, then $\Re\mathrm{e}\{s\}$ is greater than at least one $\sigma \in \mathbb{R}$ for which $e^{-\sigma t} f(t) \in L^1(;\mathbb{R};\mathbb{C})$, and therefore for such s the integral (8.2) exists.

Definition 8.3 [Laplace Transform]
Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable, and let $\sigma_a(f)$ be its abscissa of absolute convergence. For every $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma_a(f)$, we define the Laplace Transform of f as

$$\mathcal{L}\{f\}(s) = \int_0^{+\infty} f(t)\, e^{-s\,t}\, \mathrm{d}t \tag{8.3}$$

Sometimes, with some abuse of notation, we will write also $\mathcal{L}\{f(t)\}$.

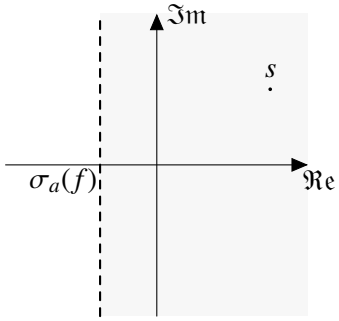
Remark 8.1 Since the support of f in Definition 8.3 is contained in $[0, +\infty)$, we can re-write (8.3) as

$$\mathcal{L}\{f\}(s) = \int_{-\infty}^{+\infty} f(t)\, e^{-s\,t}\, \mathrm{d}t.$$

Remark 8.2 The integral in Definition 8.3 is a Lebesgue integral (and this follows from the second condition in Definition 8.1). We recall that a measurable function is summable if and only if the integral of its modulus is finite. For this reason, the functions that we call "$\mathcal{L}$-transformable" are referred to, by some authors, as "absolutely $\mathcal{L}$-transformable"; these authors will then call "$\mathcal{L}$-transformable" those functions for which the integral converges as an improper integral.

We have, however, maintained the term "absolute" in the definition of $\sigma_a(f)$. The more general case where (8.3) is to be intended as an improper integral is covered by the extension of the definition of Laplace transform to distributions. In that context, we will talk simply of "abscissa of convergence".

The set of points $s \in \mathbb{C}$ for which $\Re\mathrm{e}\{s\} > \sigma_a(f)$ is called *region of absolute convergence*. This region is usually a half-plane, or the entire plane $\mathbb{C}$ when $\sigma_a(f) = -\infty$. In this latter case, the restriction $\Re\mathrm{e}\{s\} > \sigma_a(f)$ can be omitted.



An example of a function for which the abscissa of absolute convergence is $-\infty$ is $f(t) = e^{-t^2}\, H(t)$. Indeed, for every $\sigma \in \mathbb{R}$, we have $e^{-\sigma\,t}\, f(t) \in L^1(\mathbb{R};\mathbb{R})$, because the integral

$$\int_0^{+\infty} |e^{-\sigma\,t}\, e^{-t^2}|\, \mathrm{d}t$$

is convergent for every $\sigma \in \mathbb{R}$. In this case, $\{\sigma \in \mathbb{R} \,|\, e^{-\sigma\,t}\, f(t) \in L^1(\mathbb{R};\mathbb{R})\} = \mathbb{R}$, so that its infimum is $-\infty$.

Example Consider the Heaviside function $H : \mathbb{R} \longrightarrow \mathbb{R}$. We show that

(i) its abscissa of absolute convergence is $\sigma_0 = 0$;

(ii) for all s with real part > 0, the Laplace transform of H is

$$\boxed{\mathcal{L}\{H\}(s) = \frac{1}{s}}$$

To show that the abscissa of absolute convergence is 0, we consider the integral

$$\int_0^{+\infty} |H(t)|\, e^{-\sigma t}\, \mathrm{d}t = \int_0^{+\infty} e^{-\sigma t}\, \mathrm{d}t.$$

This integral is convergent if and only if $\sigma > 0$ (i.e., if and only if the exponential decays in $[0, +\infty)$).

For every $s \in \mathbb{C}$ such that $\mathfrak{Re}\{s\} > 0$, we find

$$\begin{aligned}\mathcal{L}\{H\}(s) &= \int_0^{+\infty} e^{-s t}\, H(t)\, \mathrm{d}t = \int_0^{+\infty} e^{-s t}\, \mathrm{d}t = \lim_{c \to +\infty} \int_0^c e^{-s t}\, \mathrm{d}t \\ &= \lim_{c \to +\infty} \left[-\frac{1}{s}\, e^{-s t} \right]_0^c = \lim_{c \to +\infty} \frac{1}{s}\, (1 - e^{-s c}) = \frac{1}{s}.\end{aligned}$$

Example Let $a \in \mathbb{C}$. Consider the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\mapsto e^{a t}\, H(t). \end{aligned}$$

The support of f is in $[0, \infty)$. Moreover, f is locally summable, since for every $t \in \mathbb{R}$, writing $a = x + i\, y$ with $x, y \in \mathbb{R}$, we have $|e^{a t}\, H(t)| = e^{x t}\, H(t)$, which is continuous over any closed bounded interval. We want to find out whether there exists $\sigma > 0$ such that $e^{-\sigma t}\, f(t) = e^{-(\sigma - a) t}\, H(t) \in L^1(\mathbb{R}; \mathbb{C})$, i.e.,

$$\int_{\mathbb{R}} |e^{-(\sigma - a) t}\, H(t)|\, \mathrm{d}t = \int_0^{+\infty} |e^{-(\sigma - x - i y) t}|\, \mathrm{d}t = \int_0^{+\infty} e^{-(\sigma - x) t}\, \mathrm{d}t < +\infty.$$

We observe that this happens if and only if $\sigma - x > 0$. Thus, $\sigma_a(f) = \mathfrak{Re}\{a\}$. We now compute the Laplace transform of f for $s \in \mathbb{C}$ such that $\mathfrak{Re}\{s\} > \mathfrak{Re}\{a\}$ as

$$F(s) = \int_0^{+\infty} e^{a t}\, e^{-s t}\, \mathrm{d}t = \int_0^{+\infty} e^{-(s - a) t}\, \mathrm{d}t = \lim_{c \to +\infty} \left[\frac{e^{(a - s) t}}{a - s} \right]_{t=0}^{t=c} = \lim_{c \to +\infty} \frac{e^{(x - \mathfrak{Re}\{s\})\, c}\, e^{i\, (y - \mathfrak{Im}\{s\})\, c}}{a - s} - \frac{1}{a - s} = \frac{1}{s - a},$$

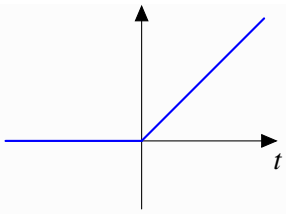
since for $\mathfrak{Re}\{s\} > \mathfrak{Re}\{a\} = x$ the exponential $e^{(x - \mathfrak{Re}\{s\})\, c}$ vanishes as $c \to +\infty$. In particular, when $a = 0$, we re-obtain the Laplace transform of the Heaviside function

$$\mathcal{L}\{H\} = \frac{1}{s}$$

for $s \in \mathbb{C}$ such that $\mathfrak{Re}\{s\} > 0$.

Example We consider the Laplace transform of the so-called *unit ramp* function, $r(t)$, defined by

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto t\,H(t) \end{aligned}$$



We need to determine the abscissa of absolute convergence. To this end, we determine for which values of $\sigma \in \mathbb{R}$ we have $e^{\sigma t}\,t\,H(t) \in L^1(\mathbb{R};\mathbb{R})$, and

$$\int_{\mathbb{R}} |e^{\sigma t}\,t\,H(t)|\,\mathrm{d}t = \int_0^{+\infty} e^{\sigma t}\,t\,\mathrm{d}t$$

converges if and only if $\sigma < 0$, so that the abscissa of absolute convergence is 0. We compute the Laplace transform for $\mathfrak{Re}\{s\} > 0$:

$$F(s) = \int_0^{+\infty} t\,e^{-st}\,H(t)\,\mathrm{d}t = \int_0^{+\infty} t\,e^{-st}\,\mathrm{d}t.$$

We integrate by parts:

$$\int t\,e^{-st}\,\mathrm{d}t = -\frac{1}{s}\,e^{-st}\,t - \int -\frac{1}{s}\,e^{-st}\,\mathrm{d}t = -\frac{1}{s}\,e^{-st}\,t + \frac{1}{s}\left(-\frac{1}{s}\,e^{-st}\right) = -\frac{1}{s^2}\,(s\,t+1)\,e^{-st}.$$

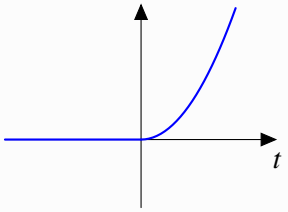
We obtain

$$F(s) = \lim_{c\to+\infty}\left[-\frac{1}{s^2}\,(s\,t+1)\,e^{-st}\right]_0^c = \lim_{c\to+\infty}\left(-\frac{1}{s^2}\,(s\,c+1)\,e^{-\mathfrak{Re}\{s\}\,c}\,e^{-i\,\mathfrak{Im}\{s\}\,c}\right) + \frac{1}{s^2} = \frac{1}{s^2}.$$

Example

We consider the Laplace transform of the so-called *unit parabola* function, $p(t)$, defined by

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \frac{1}{2}\, t^2\, H(t) \end{aligned}$$



We need to determine the abscissa of absolute convergence. To this end, we determine for which values of $\sigma \in \mathbb{R}$ we have $\frac{1}{2}\, e^{\sigma t}\, t^2\, H(t) \in L^1(\mathbb{R};\mathbb{R})$, and

$$\int_{\mathbb{R}} \left| e^{\sigma t}\, \frac{1}{2}\, t^2\, H(t) \right| \mathrm{d}t = \int_0^{+\infty} \frac{1}{2}\, e^{\sigma t}\, t^2 \, \mathrm{d}t$$

converges if and only if $\sigma < 0$, so that the abscissa of absolute convergence is 0. We compute the Laplace transform for $\mathfrak{Re}\{s\} > 0$:

$$F(s) = \int_0^{+\infty} \frac{1}{2}\, t^2\, e^{-s\,t}\, H(t)\, \mathrm{d}t = \int_0^{+\infty} \frac{1}{2}\, t^2\, e^{-s\,t}\, \mathrm{d}t.$$

We integrate by parts:

$$\begin{aligned} \int t^2\, e^{-s\,t}\, \mathrm{d}t &= -\frac{1}{s}\, t^2\, e^{-s\,t} - \int -\frac{1}{s}\, e^{-s\,t}\, 2\, t\, \mathrm{d}t \\ &= -\frac{1}{s}\, t^2\, e^{-s\,t} + \frac{2}{s} \int t\, e^{-s\,t}\, \mathrm{d}t \\ &= -\frac{1}{s}\, t^2\, e^{-s\,t} + \frac{2}{s} \left(-\frac{1}{s}\, e^{-s\,t}\, t - \int -\frac{1}{s}\, e^{-s\,t}\, \mathrm{d}t \right) \\ &= -\frac{1}{s}\, t^2\, e^{-s\,t} - \frac{2}{s^2}\, e^{-s\,t}\, t - \frac{2}{s^3}\, e^{-s\,t} \\ &= -\frac{1}{s^3}\, (t^2\, s^2 + 2\, s\, t + 2)\, e^{-s\,t}. \end{aligned}$$

We obtain

$$F(s) = \lim_{c \to +\infty} -\frac{1}{2\, s^3} \left[-(t^2\, s^2 + 2\, s\, t + 2)\, e^{-s\,t} \right]_0^c = \lim_{c \to +\infty} \left(-\frac{1}{s^3}\, (c^2\, s^2 + 2\, s\, c + 2)\, e^{-s\,c} \right) + \frac{1}{2\, s^3}\, 2 = \frac{1}{s^3}.$$

The Laplace transform is a linear operation: this means that:

- the set of $\mathcal{L}$-transformable functions is a vector space, with the usual operations of addition and scalar multiplications;
- the Laplace transform of the linear combination of two $\mathcal{L}$-transformable functions f_1 and f_2 can be obtained from the same linear combination of the Laplace transform of the two functions, i.e., the Laplace transform is a linear operator over the space of $\mathcal{L}$-transformable functions.

Theorem 8.1 [Linearity]
Let $k_1, k_2 \in \mathbb{C}$. Let $f_1, f_2 \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{C})$ be Laplace transformable with abscissas of absolute convergence $\sigma_a(f_1)$ and $\sigma_a(f_2)$, respectively. Then, $k_1\, f_1 + k_2\, f_2$ is Laplace transformable with abscissa of

absolute convergence $\max\{\sigma_a(f_1), \sigma_a(f_2)\}$, *and*

$$\mathcal{L}\{k_1\, f_1 + k_2\, f_2\}(s) = k_1\, F_1(s) + k_2\, F_2(s).$$

Proof First, notice that $k_1\, f_1 + k_2\, f_2 \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$, since for every compact set of $\mathbb{R}$ we have

$$\int_K |k_1\, f_1(t) + k_2\, f_2(t)|\, \mathrm{d}t \leq |k_1| \int_K f_1(t)\, \mathrm{d}t + |k_2| \int_K f_2(t)\, \mathrm{d}t < +\infty.$$

We show that $k_1\, f_1 + k_2\, f_2$ is Laplace transformable. Since f_1 is Laplace transformable, there exists $\sigma_1 \in \mathbb{R}$ such that $e^{-\sigma_1 t}\, f_1(t) \in L^1(\mathbb{R};\mathbb{C})$. Likewise, since f_2 is Laplace transformable, there exists $\sigma_2 \in \mathbb{R}$ such that $e^{-\sigma_2 t}\, f_2(t) \in L^1(\mathbb{R};\mathbb{C})$. Taking $\sigma > \max\{\sigma_1, \sigma_2\}$, we find

$$\int_{\mathbb{R}} \left| e^{-\sigma t}\, (k_1\, f_1(t) + k_2\, f_2(t)) \right| \mathrm{d}t \leq |k_1| \int_{\mathbb{R}} e^{-\sigma t}\, |f_1(t)|\, \mathrm{d}t + |k_2| \int_{\mathbb{R}} e^{-\sigma t}\, |f_2(t)|\, \mathrm{d}t < +\infty.$$

Moreover, the support of $k_1\, f_1 + k_2\, f_2$ is contained in $[0, +\infty)$. Hence, $k_1\, f_1 + k_2\, f_2$ is Laplace transformable. Moreover, taking $s \in \mathbb{C}$ such that $\mathfrak{Re}\{s\} > \max\{\sigma_a(f_1), \sigma_a(f_2)\}$, we find

$$\mathcal{L}\{k_1\, f_1 + k_2\, f_2\}(s) = \int_0^{+\infty} e^{-st}(k_1\, f_1(t) + k_2\, f_2(t))\, \mathrm{d}t = k_1 \int_0^{+\infty} e^{-st}\, f_1(t)\, \mathrm{d}t + k_2 \int_0^{+\infty} e^{-st}\, f_2(t)\, \mathrm{d}t = k_1\, F_1(s) + k_2\, F_2(s).$$

Example Let $\omega \in \mathbb{R}$. We want to find the Laplace transform of the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \sin(\omega\, t)\, H(t), \end{aligned}$$

after showing that it is Laplace transformable. First, we notice that f is locally summable and its support is contained in $[0, +\infty)$. We can exploit the linearity of the Laplace transform, by using Euler's formula for the sine

$$\sin(\omega\, t) = \frac{e^{i\omega t} - e^{-i\omega t}}{2\, i} \quad \Rightarrow \quad \sin(\omega\, t)\, H(t) = \frac{1}{2\, i}\, e^{i\omega t}\, H(t) - \frac{1}{2\, i}\, e^{-i\omega t}\, H(t).$$

Both the functions $e^{i\omega t}\, H(t)$ and $e^{-i\omega t}\, H(t)$ are Laplace transformable and their abscissas of absolute convergence is $\mathfrak{Re}\{i\,\omega\} = 0$ and $\mathfrak{Re}\{-i\,\omega\} = 0$, so that by Theorem 8.1 the function f is Laplace transformable and its abscissa of absolute convergence is 0. We find

$$\begin{aligned} \mathcal{L}\{f\}(s) &= \mathcal{L}\left\{ \frac{1}{2\, i}\, e^{i\omega t}\, H(t) - \frac{1}{2\, i}\, e^{-i\omega t}\, H(t) \right\} \\ &= \frac{1}{2\, i}\, \mathcal{L}\left\{ e^{i\omega t}\, H(t) \right\} - \frac{1}{2\, i}\, \mathcal{L}\left\{ e^{-i\omega t}\, H(t) \right\} \\ &= \frac{1}{2\, i}\, \frac{1}{s - i\,\omega} - \frac{1}{2\, i}\, \frac{1}{s + i\,\omega} = \frac{\omega}{s^2 + \omega^2}. \end{aligned}$$

It is easy to see that the Laplace transforms of

$$g(t) = \cos(\omega\, t)\, H(t)$$

is

$$\mathcal{L}\{g\}(s) = \frac{s}{s^2 + \omega^2}$$

Example The function $f : \mathbb{R} \longrightarrow \mathbb{R}$ defined by

$$f(t) = e^{t^2} H(t)$$

is not $\mathcal{L}$-transformable. Indeed, for every $\sigma \in \mathbb{R}$, we have

$$e^{-\sigma t} f(t) \notin L^1(\mathbb{R}; \mathbb{R}).$$

In all the examples that we have introduced so far, three common properties stand out:

- the Laplace transform is infinitesimal as $\mathfrak{Re}\{s\} \to +\infty$;
- the Laplace transform is bounded in every half-plane of the form $\mathfrak{Re}\{s\} \geq \sigma$, where $\sigma > \sigma_a(f)$;
- the Laplace transform is holomorphic in the half plane of absolute convergence.

The next result shows that these properties are, in fact, general.

Theorem 8.2 *Let $f \in L^1_{\text{loc}}(\mathbb{R}; \mathbb{R})$ be $\mathcal{L}$-transformable. Let $F = \mathcal{L}\{f\}$. Then,*

(i) F is bounded in the half-plane $\{s \in \mathbb{C} \,|\, \mathfrak{Re}\{s\} \geq \sigma\}$ for every $\sigma > \sigma_a(f)$;

(ii) $\lim_{\mathfrak{Re}\{s\} \to +\infty} F(s) = 0$;[a]

(iii) F is holomorphic in $\{s \in \mathbb{C} \,|\, \mathfrak{Re}\{s\} > \sigma_a(f)\}$.

[a] The meaning of this statement is the following: for every $\epsilon > 0$, there exists $\delta_\epsilon > \sigma_a(f)$ such that if $\mathfrak{Re}\{s\} \geq \delta_\epsilon$, then $|F(s)| \leq \epsilon$.

Proof We start from the first point. The Laplace transform is defined at the right of the abscissa of absolute convergence, and in that half-place the Laplace transform is bounded. Indeed, let $\sigma > \sigma_a(f)$. If $\mathfrak{Re}\{s\} \geq \sigma$, for $t \geq 0$ we have

$$|e^{-st}| = e^{-\mathfrak{Re}\{s\} t} \leq e^{-\sigma t},$$

so that

$$|F(s)| = \left| \int_0^{+\infty} e^{-st} f(t) \, \mathrm{d}t \right| \leq \int_0^{+\infty} e^{-\sigma t} |f(t)| \, \mathrm{d}t,$$

and the right hand-side does not depend on s.

We now prove *(ii)*. Let $(s_k)_{k \in \mathbb{N}}$ be a sequence of complex numbers satisfying the condition $\mathfrak{Re}\{s_k\} > \sigma_a(f)$ and such that

$$\lim_{k \to \infty} \mathfrak{Re}\{s_k\} = +\infty.$$

This is always possible because the region of the complex plane where the terms of the sequence live is unbounded. We show that $\lim_{k \to +\infty} \mathcal{L}\{f\}(s_k) = 0$. Indeed,

$$\mathcal{L}\{f\}(s_k) = \int_0^{+\infty} e^{-s_k t} f(t) \, \mathrm{d}t.$$

We use Lebesgue's dominated convergence: for every $t \in \mathbb{R}$

$$|e^{-s_k t} f(t)| = e^{-\mathfrak{Re}\{s_k\} t} |f(t)| \xrightarrow{k \to \infty} 0,$$

because we assumed that $\Re\mathrm{e}\{s_k\} \xrightarrow{k\to\infty} +\infty$. Hence, the integrand is pointwise convergent to the zero function. We now bound the absolute value of the integrand with a summable function that does not depend on k. To this end, we observe that the set $\{\Re\mathrm{e}\{s_k\} \,|\, k \in \mathbb{N}\}$ admits minimum. In fact, consider $\Re\mathrm{e}\{s_0\}$. The set $S = \{k \in \mathbb{N} \,|\, \Re\mathrm{e}\{s_k\} \leq \Re\mathrm{e}\{s_0\} + 1\}$ is non-empty (it contains at least $k = 0$) and finite (because $\Re\mathrm{e}\{s_k\} \xrightarrow{k\to\infty} +\infty$). The minimum of this set corresponds to the minimum of $\{\Re\mathrm{e}\{s_k\} \,|\, k \in \mathbb{N}\}$, i.e.,

$$\min_{k\in S} \Re\mathrm{e}\{s_k\} = \min_{k\in\mathbb{N}} \Re\mathrm{e}\{s_k\}.$$

If we take $\sigma = \min\{\Re\mathrm{e}\{s_k\} \,|\, k \in \mathbb{N}\}$, then $\Re\mathrm{e}\{s_k\} \geq \sigma > \sigma_a(f)$, so that for all $k \in \mathbb{N}$ we have

$$|e^{-s_k t} f(t)| = e^{-\Re\mathrm{e}\{s_k\} t} |f(t)| \leq e^{-\sigma t} |f(t)|,$$

which is summable. We can therefore apply the dominated convergence theorem:

$$\lim_{k\to+\infty} \mathcal{L}\{f\}(s_k) = \lim_{k\to+\infty} \int_0^{+\infty} e^{-s_k t} f(t)\,\mathrm{d}t = 0.$$

We now prove the last point. We show that F is holomorphic over $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \sigma_a(f)\}$. We show that the limit

$$\lim_{h\to 0} \frac{F(s+h) - F(s)}{h}$$

exists in $\mathbb{C}$, where $h \in \mathbb{C}$ and $\Re\mathrm{e}\{s\} > \sigma_a(f)$. We have

$$\frac{F(s+h) - F(s)}{h} = \int_0^{+\infty} \frac{e^{-ht} - 1}{h} e^{-st} f(t)\,\mathrm{d}t.$$

We first study the pointwise limit of the integrand function as $h \to 0$. The only factor in the integrand that depends on h is

$$\frac{e^{-ht} - 1}{h}. \tag{8.4}$$

Notice that $e^{-ht} - 1 = 1 - ht + \mathrm{O}(h^2) - 1 = -ht + \mathrm{O}(h^2)$, where the "big" O notation is used with respect to $h \to 0$. We obtain the Taylor power series expansion at $h \to 0$ of (8.4) as

$$\frac{e^{-ht} - 1}{h} = -t + \mathrm{O}(h) \xrightarrow{h\to 0} -t.$$

Hence, if we are able to interchange the limit with the integral and use the pointwise limit obtained above, we formally obtain the integral

$$\int_0^{+\infty} e^{-st}(-t) f(t)\,\mathrm{d}t.$$

Before we establish if it is possible to interchange the limit with the integral, using as usual the dominated convergence theorem, we have to make sure that this integral exists. We show in particular that the function $\mathbb{R} \longrightarrow \mathbb{C}$ that maps t into $e^{-st}(-t) f(t)$ is in $L^1([0,+\infty);\mathbb{C})$. Let $\sigma > \sigma_a(f)$. Take s satisfying $\sigma_a(f) < \sigma < \Re\mathrm{e}\{s\}$. Hence, $e^{-\sigma t} f(t) \in L^1([0,+\infty);\mathbb{C})$ because $\sigma > \sigma_a(f)$. Now,

$$|e^{-st} t f(t)| = |e^{-st} f(t)|\,|e^{(\sigma-s)t} t| = |e^{-st} f(t)|\, t\, e^{(\sigma - \Re\mathrm{e}\{s\})t},$$

where $e^{-st} f(t)$ is summable over $[0, +\infty)$ as seen above and where $t\, e^{(\sigma - \Re\mathrm{e}\{s\})t}$ is bounded on $[0, +\infty)$. Thus, the integral exists. We now want to justify the application of the dominated convergence. Let $\varepsilon > 0$ such that $\Re\mathrm{e}\{s\} - 2\varepsilon > \sigma_a(f)$. Let $h \neq 0$ be such that $|h| \leq \varepsilon$. From $\frac{\mathrm{d}}{\mathrm{d}t} e^{-ht} = -h\, e^{-ht}$, we find

$$e^{-ht} - 1 = -h \int_0^t e^{-h\tau}\, \mathrm{d}\tau \quad \Rightarrow \quad |e^{-ht} - 1| = |h| \left| \int_0^t e^{-h\tau}\, \mathrm{d}\tau \right| \leq |h| \int_0^t e^{-\tau \Re\mathrm{e}\{h\}}\, \mathrm{d}\tau.$$

Moreover,

$$-\Re\mathrm{e}\{h\} \leq |\Re\mathrm{e}\{h\}| \leq |h| \leq \varepsilon \quad \Rightarrow \quad e^{-\tau \Re\mathrm{e}\{h\}} \leq e^{\tau |h|} \leq e^{t\varepsilon},$$

because $\tau \leq t$. Hence,

$$|e^{-ht} - 1| \leq |h| \int_0^t e^{t\varepsilon}\, \mathrm{d}\tau = |h|\, e^{t\varepsilon} \int_0^t \mathrm{d}t = |h|\, e^{t\varepsilon}\, t.$$

We obtain

$$\left| \frac{e^{-ht} - 1}{h} \right| \leq e^{t\varepsilon}\, t.$$

Hence,

$$\left| \frac{e^{-ht} - 1}{h} e^{-st} f(t) \right| \leq e^{t\varepsilon}\, t\, e^{-\Re\mathrm{e}\{s\} t} |f(t)| = e^{-(\Re\mathrm{e}\{s\} - \varepsilon)t}\, t\, |f(t)|.$$

We have already observed that the Laplace transform of $t f$ is defined to the right of $\sigma_a(f)$, i.e., for $\Re\mathrm{e}\{s\} - \varepsilon > \sigma_a(f)$. It follows that $e^{-(\Re\mathrm{e}\{s\} - \varepsilon)t}\, t\, |f(t)|$ is summable.

Remark 8.3 We are not allowed to replace the first statement of Theorem 8.2 with one which requires the Laplace transform to be bounded in the half-plane $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} \geq \sigma_a(f)\}$. In fact, consider the Heaviside function H. We have found that the abscissa of absolute convergence is 0. Moreover, we know that $\mathcal{L}\{H\}(s) = 1/s$, which is not bounded in the half-plane $\Re\mathrm{e}\{s\} \geq 0$.

Remark 8.4 We have found in a previous example that, for $a \in \mathbb{C}$, for $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \Re\mathrm{e}\{a\}$ we have

$$\mathcal{L}\{e^{at} H(t)\}(s) = \frac{1}{s - a}.$$

The right hand-side of this equation defines a function which is holomorphic in $\mathbb{C} \setminus \{a\}$, and not just in the half-plane $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \Re\mathrm{e}\{a\}\}$. Nevertheless, the right hand-side coincides with the Laplace transform only over this half-plane, because for every other value of s the integral does not converge.

Example

The function $F(s) = 1/\Re\mathrm{e}\{s\}$ is not the $\mathcal{L}$-transform of a function, because $\Re\mathrm{e}\{s\}$ is not holomorphic, and therefore neither is $1/\Re\mathrm{e}\{s\}$.
The function $F(s) = \frac{1}{s \sin s}$ is not the $\mathcal{L}$-transform of a function because if is not bounded. Indeed

$$\left| \frac{1}{s \sin s} \right| = \frac{1}{|s|\, |\sin s|} \geq \frac{1}{|s|},$$

which is not bounded.
The function $F(s) = s \sin \frac{1}{s}$ is not the $\mathcal{L}$-transform of a function because $s \sin \frac{1}{s}$ does not go to zero as $\Re\mathrm{e}\{s\}$ goes to infinity. Finally, the function $F(s) = e^{-s^2}$ is not the $\mathcal{L}$-transform of a function because e^{-s^2} does not go to zero as $\Re\mathrm{e}\{s\}$ goes to infinity; this can be seen by writing $s = x + i\, y$ where $x, y \in \mathbb{R}$ and $|e^{-s^2}| = e^{y^2 - x^2}$, which does not go to zero as $x \to \infty$.

Theorem 8.3 *Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $f|_{[0,+\infty)} \in C^1([0,+\infty))$. XXX Let*

$$\widetilde{f'}(t) = \begin{cases} f'(t) & t > 0 \\ 0 & t \le 0 \end{cases}$$

Let $\widetilde{f'}$ be $\mathcal{L}$-transformable. Then, for $s \in \mathbb{C}$ such that

$$\Re\mathrm{e}\{s\} > \max\{\sigma_a(f), \sigma_a(\widetilde{f'})\},$$

there holds

$$\mathcal{L}\{\widetilde{f'}\}(s) = s\,F(s) - f(0^+),$$

where $f(0^+) = \lim_{t\to 0^+} f(t)$.

Proof Integrating by parts we find

$$\begin{aligned}\mathcal{L}\{\widetilde{f'}\}(s) &= \int_0^{+\infty} \widetilde{f'}(t)\,e^{-st}\,\mathrm{d}t = \lim_{\alpha\to\infty}\int_0^{\alpha} \widetilde{f'}(t)\,e^{-st}\,\mathrm{d}t \\ &= \lim_{\alpha\to\infty}\Big(\big[e^{-st}f(t)\big]_{t=0}^{t=\alpha} + s\int_0^{\alpha} e^{-st}f(t)\,\mathrm{d}t\Big) \\ &= \lim_{\alpha\to\infty}\Big(e^{-s\alpha}f(\alpha) - f(0^+) + s\int_0^{\alpha} e^{-st}f(t)\,\mathrm{d}t\Big).\end{aligned}$$

This limit exists because we assumed that $\widetilde{f'}$ is $\mathcal{L}$-transformable. Moreover, for $\Re\mathrm{e}\{s\} > \sigma_a(f)$, the limit

$$\lim_{\alpha\to\infty}\int_0^{\alpha} e^{-st}f(t)$$

also exists, and is equal to $F(s)$. By difference, also the limit

$$\lim_{\alpha\to\infty} e^{-s\alpha}f(\alpha)$$

exists and it is finite. Let us denote this limit by ℓ_s. We show that $\ell_s = 0$. Take $\sigma \in \mathbb{R}$ such that $\Re\mathrm{e}\{s\} > \sigma > \max\{\sigma_a(f), \sigma_a(\widetilde{f'})\}$. By replacing σ with s (which is legitimate, as it is a point with the same properties), we can say that the limit

$$\lim_{\alpha\to\infty} e^{-\sigma\alpha}f(\alpha)$$

exists and is finite, and we denote it by ℓ_σ. Thus,

$$e^{-s\alpha}f(\alpha) = e^{-(s-\sigma)\alpha}\,e^{-\sigma\alpha}f(\alpha).$$

Then, $\lim_{\alpha\to\infty} e^{-(s-\sigma)\alpha} = 0$ since $\Re\mathrm{e}\{s\} > \sigma$, while $\lim_{\alpha\to\infty} e^{-\sigma\alpha}f(\alpha) = \ell_\sigma$. Thus, $\lim_{\alpha\to\infty} e^{-s\alpha}f(\alpha) = 0$. It follows that

$$\int_0^{+\infty} \tilde{f}'(t)\,e^{-st}\,\mathrm{d}t = s\,F(s) - f(0^+),$$

Remark 8.5 The term $f(0^+)$ is due to the way that the function $\tilde{f}'$ has been defined, which does not take into account the fact that f may not be differentiable at the origin. When f is differentiable at the origin, it is also continuous at the origin, and therefore also equal to zero, and in this case

$f(0^+) = 0$.

Example

If a function $f \in C^1([0,+\infty))$ is $\mathcal{L}$-transformable, it is not always true that f' is $\mathcal{L}$-transformable. For example, consider the function $f(t) = H(t)\,\sin(e^{t^2})$ is $\mathcal{L}$-transformable, because it is bounded, and its restriction to $[0,+\infty)$ is of class $C^1([0,+\infty))$. However, for $t \geq 0$, its derivative $f'(t) = 2\,t\,e^{t^2}\,\cos e^{t^2}\,H(t)$ is not $\mathcal{L}$-transformable (why?).

Remark 8.6 Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $m \in \mathbb{N}$. Let $f|_{[0,+\infty)} \in C^m([0,+\infty))$. XXXLet, for all $m \in \mathbb{N}$,

$$\widetilde{f^{(m)}}(t) = \begin{cases} \frac{\mathrm{d}^m f}{\mathrm{d}t^m}(t) & t > 0 \\ 0 & t \leq 0. \end{cases}$$

Let $\widetilde{f'}, \ldots, \widetilde{f^{(m)}}$ be $\mathcal{L}$-transformable. Then, for $s \in \mathbb{C}$ such that

$$\Re\mathrm{e}\{s\} > \max\{\sigma_a(f), \sigma_a(\widetilde{f'}), \ldots, \sigma_a(\widetilde{f^{(m)}})\},$$

there holds

$$\mathcal{L}\{\widetilde{f^{(m)}}\}(s) = s^m\,\mathcal{L}\{f\}(s) - \sum_{k=0}^{m-1} s^{m-k-1}\,f^{(k)}(0^+),$$

where $f^{(k)}(0^+) = \lim_{t\to 0^+} f^{(k)}(t)$ for $k \in \{0,\ldots,m-1\}$.

Theorem 8.4 [Properties of the Laplace Transform]
Let $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $F(s) = \mathcal{L}\{f\}(s)$. Then:

(i) *Let $c > 0$. Let us denote by $f(c\,t)$ the function*

$$\begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{C} \\ t & \mapsto & f(c\,t). \end{array} \tag{8.5}$$

The function $f(c\,t)$ is $\mathcal{L}$-transformable, and for all $\Re\mathrm{e}\{s\} > c\,\sigma_a(f)$ there holds

$$\mathcal{L}\{f(c\,t)\}(s) = \frac{1}{c}\,F\left(\frac{s}{c}\right);$$

(ii) *Let $t_0 \geq 0$. Let us denote by $f(t-t_0)$ the function*

$$\begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{C} \\ t & \mapsto & f(t-t_0). \end{array} \tag{8.6}$$

The function $f(t-t_0)$ is $\mathcal{L}$-transformable, and for all $\Re\mathrm{e}\{s\} > \sigma_a(f)$ there holds

$$\mathcal{L}\{f(t-t_0)\}(s) = e^{-t_0\,s}F(s);$$

(iii) *Let $s_0 \in \mathbb{C}$. Let us denote by $e^{s_0 t} f(t)$ the function*

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\mapsto e^{s_0 t} f(t). \end{aligned} \tag{8.7}$$

The function $f(t - t_0)$ is $\mathcal{L}$-transformable, and for all $\mathfrak{Re}\{s\} > \mathfrak{Re}\{s_0\} + \sigma_a(f)$ there holds

$$\mathcal{L}\{e^{s_0 t} f(t)\}(s) = F(s - s_0).$$

Proof We prove **(i)**. We want to find $\sigma \in \mathbb{R}$ such that $e^{-\sigma t} f(c\,t) \in L^1(\mathbb{R}; \mathbb{C})$. By changing variable we find

$$\int_0^{+\infty} e^{-\sigma t} |f(c\,t)|\, \mathrm{d}t = \int_0^{+\infty} e^{-\frac{\sigma}{c}\tau} |f(\tau)|\, c^{-1}\, \mathrm{d}\tau.$$

Thus, $\sigma > c\,\sigma_a(f)$, i.e., $\sigma_a(f(c\,t)) \leq c\,\sigma_a(f)$. Let now $\mathfrak{Re}\{s\} > c\,\sigma_a(f)$. Then

$$\mathcal{L}\{f(c\,t)\}(s) = \int_0^{+\infty} e^{-s t} f(c\,t)\, \mathrm{d}t = \int_0^{+\infty} e^{-\frac{s}{c}\tau} f(\tau)\, c^{-1}\, \mathrm{d}\tau = \frac{1}{c} F\left(\frac{s}{c}\right).$$

Consider **(ii)**. Clearly, $\operatorname{supp} f(t - t_0) \subseteq [t_0, +\infty) \subseteq [0, +\infty)$. We find

$$\int_0^{+\infty} e^{-\sigma t} |f(t - t_0)|\, \mathrm{d}t = \int_{t_0}^{+\infty} e^{-\sigma(\tau + t_0)} |f(\tau)|\, \mathrm{d}\tau = e^{-\sigma t_0} \int_0^{+\infty} e^{-\sigma\tau} |f(\tau)|\, \mathrm{d}\tau,$$

which is finite if $\sigma > \sigma_a(f)$. Thus, for $\sigma > \sigma_a(f)$, the function $e^{-\sigma t} f(t - t_0)$ is in $L^1(\mathbb{R}; \mathbb{C})$. For every $\mathfrak{Re}\{s\} > \sigma_a(f)$, we get

$$\int_0^{+\infty} e^{-s t} f(t - t_0)\, \mathrm{d}t = \int_{t_0}^{+\infty} e^{-s(\tau + t_0)} f(\tau)\, \mathrm{d}\tau = \int_0^{+\infty} e^{-s t_0} e^{-s\tau} f(\tau)\, \mathrm{d}\tau = e^{-t_0 s} F(s).$$

Consider **(iii)**. Clearly, $\operatorname{supp}(e^{s_0 t} f(t)) \subseteq [0, +\infty)$. To ascertain that $e^{s_0 t} f(t)$ is $\mathcal{L}$-transformable, we must find $\sigma \in \mathbb{R}$ such that $\displaystyle\int_0^{+\infty} e^{-\sigma t} |e^{s_0 t} f(t)|\, \mathrm{d}t$. On the other hand,

$$\int_0^{+\infty} e^{-\sigma t} |e^{-s_0 t} f(t)|\, \mathrm{d}t = \int_0^{+\infty} e^{-(\sigma - s_0) t} |f(t)|\, \mathrm{d}t,$$

which is finite if $\mathfrak{Re}\{\sigma - s_0\} > \sigma_a(f)$. In other words, for $\sigma > \sigma_a(f) + \mathfrak{Re}\{s_0\}$, the function $e^{s_0 t} f(t)$ is in $L^1(\mathbb{R}; \mathbb{C})$. For all $\mathfrak{Re}\{s\} > \mathfrak{Re}\{s_0\} + \sigma_a(f)$ we get

$$\int_0^{+\infty} e^{-s t} e^{-s_0 t} f(t)\, \mathrm{d}t = \int_0^{+\infty} e^{-(s - s_0) t} f(t)\, \mathrm{d}t = F(s - s_0).$$

Theorem 8.5 [Laplace Transform of the Convolution]
Let $f, g \in L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{C})$ be $\mathcal{L}$-transformable. Let $F(s) = \mathcal{L}\{f\}(s)$ and $G(s) = \mathcal{L}\{g\}(s)$. Then, $f * g$ is $\mathcal{L}$-transformable, and for $\mathfrak{Re}\{s\} > \max\{\sigma_a(f), \sigma_a(g)\}$, there holds

$$\mathcal{L}\{f * g\}(s) = F(s)\, G(s).$$

Proof The two functions are certainly convolvable, being elements of $L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$. Moreover,

$$(f * g)(t) = \int_{\mathbb{R}} f(t-\tau)\, g(\tau)\, \mathrm{d}\tau \quad \text{almost everywhere.}$$

Moreover, $\operatorname{supp}(f * g) \subseteq [0, +\infty)$. Since $f(t-\tau)\, g(\tau) \neq 0$ if $0 \leq \tau \leq t$, we find

$$(f * g)(t) = \int_0^t f(t-\tau)\, g(\tau)\, \mathrm{d}\tau \quad \text{almost everywhere.}$$

Since we already know that $\operatorname{supp}(f * g) \subseteq [0, +\infty)$, to prove that $f * g$ is $\mathcal{L}$-transformable it is sufficient to show that there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma t}(f * g) \in L^1(\mathbb{R};\mathbb{C})$. Let us choose $\sigma > \max\{\sigma_a(f), \sigma_a(g)\}$.

Since f, g are $\mathcal{L}$-transformable,

$$(f * g)(t) = \int_{\mathbb{R}} f(t-\tau)\, g(\tau)\, d\tau = \begin{cases} 0 & t < 0 \\ \displaystyle\int_0^t f(t-\tau)\, g(\tau)\, \mathrm{d}\tau & t \geq 0, \end{cases}$$

which implies that

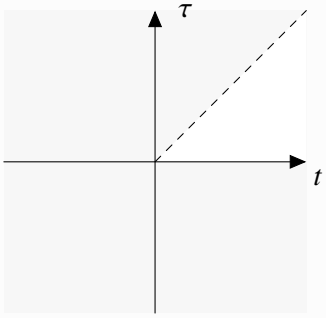
$$e^{-\sigma t}(f * g)(t) = \begin{cases} 0 & t < 0 \\ \displaystyle e^{-\sigma t} \int_0^t f(t-\tau)\, g(\tau)\, \mathrm{d}\tau & t \geq 0 \end{cases}$$

$$= \begin{cases} 0 & t < 0 \\ \displaystyle\int_0^t \left(e^{-\sigma(t-\tau)} f(t-\tau)\right)\left(e^{-\sigma\tau} g(\tau)\right) \mathrm{d}\tau & t \geq 0 \end{cases}$$

Hence, $e^{-\sigma t}(f * g) = (e^{-\sigma t} f) * (e^{-\sigma t} g)$. If $\Re\mathrm{e}\{s\} > \max\{\sigma_a(f), \sigma_a(g)\}$, then $e^{-\sigma t} f$ and $e^{-\sigma t} g$ are both in $L^1(\mathbb{R};\mathbb{C})$, and therefore so does their convolution (see Theorem 5.39). Let us now consider $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \max\{\sigma_a(f), \sigma_a(g)\}$. Then

$$\begin{aligned} \mathcal{L}\{f * g\}(s) &= \int_0^{+\infty} e^{-st} \left(\int_0^t f(t-\tau)\, g(\tau)\, \mathrm{d}\tau \right) \mathrm{d}t \\ &= \int_0^{+\infty} \left(\int_0^t e^{-s(t-\tau)} f(t-\tau)\, e^{-s\tau}\, g(\tau)\, \mathrm{d}\tau \right) \mathrm{d}t. \end{aligned}$$

The function $(t, \tau) \mapsto e^{-s(t-\tau)} f(t-\tau)\, e^{-s\tau}\, g(\tau)$ is in $L^1(\mathbb{R}^2;\mathbb{C})$. Applying Fubini's theorem and considering the region of integration



we obtain

$$\begin{aligned}\mathcal{L}\{f*g\}(s) &= \int_0^{+\infty}\left(\int_0^t e^{-s(t-\tau)}f(t-\tau)\,\mathrm{d}t\right)e^{-s\tau}g(\tau)\,\mathrm{d}\tau\\ &= \int_0^{+\infty}\left(\int_\tau^{+\infty} e^{-s(t-\tau)}f(t-\tau)\,\mathrm{d}t\right)e^{-s\tau}g(\tau)\,\mathrm{d}\tau.\end{aligned}$$

On the other hand,

$$\int_\tau^{+\infty} e^{-s(t-\tau)}f(t-\tau)\,\mathrm{d}t = \int_0^{+\infty} e^{-st}f(t)\,\mathrm{d}t = F(s),$$

so that

$$\begin{aligned}\mathcal{L}\{f*g\}(s) &= \int_0^{+\infty} F(s)\,e^{-s\tau}g(\tau)\,\mathrm{d}\tau\\ &= F(s)\int_0^{+\infty} e^{-s\tau}g(\tau)\,\mathrm{d}\tau = F(s)\,G(s).\end{aligned}$$

Remark 8.7 In general, the convolution of distributions is not an associative operation. However, if the distributions are $\mathcal{L}$-transformable, it is associative. In other words, if $T, S, R \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ and T, S, R are $\mathcal{L}$-transformable, then

$$(T*S)*R = T*(S*R).$$

It is sufficient, in fact, to apply the Laplace transform, to obtain

$$\mathcal{L}\{(T*S)*R\} = \mathcal{L}\{T*(S*R)\},$$

and the associativity follows from the injectivity of the Laplace transform.

Example Let $f \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let

$$g(t) = (H*f)(t) = \int_0^t f(\tau)\,\mathrm{d}\tau.$$

Since H is $\mathcal{L}$-transformable, g is also $\mathcal{L}$-transformable, and if $\Re\mathrm{e}\{s\} > \max\{\sigma_a(f),\sigma_a(H)\} = \max\{\sigma_a(f),0\}$, we find

$$G(s) = \mathcal{L}\{g\}(s) = \mathcal{L}\{H\}(s) = \frac{F(s)}{s}.$$

Example We want to generalize the formula, valid for $a \in \mathbb{C}$,

$$\mathcal{L}\left\{e^{at}H(t)\right\} = \frac{1}{s-a}.$$

In particular, we want to show that

$$\mathcal{L}\left\{\frac{1}{(n-1)!}\,t^{n-1}\,e^{at}H(t)\right\} = \frac{1}{(s-a)^n}. \tag{8.8}$$

Consider the case $n = 2$. From the Laplace transform of the convolution we find

$$\frac{1}{(s-a)^2} = \mathcal{L}\left\{(e^{at}\,H(t)) * (e^{at}\,H(t))\right\}.$$

If $t \geq 0$, we find

$$(e^{at}\,H(t)) * (e^{at}\,H(t)) = \int_0^t e^{a(t-\tau)}\,e^{a\tau}\,\mathrm{d}\tau = \int_0^t e^{at}\,\mathrm{d}\tau = e^{at}\int_0^t \mathrm{d}\tau = t\,e^{at}\,H(t).$$

It follows that

$$\mathcal{L}^{-1}\left\{\frac{1}{(s-a)^2}\right\} = t\,e^{at}\,H(t).$$

In general, we have

$$\frac{1}{(s-a)^n} = \mathcal{L}\Big\{\underbrace{(e^{at}\,H(t)) * (e^{at}\,H(t)) * \ldots * (e^{at}\,H(t))}_{n \text{ times}}\Big\},$$

and by induction we can show that

$$\underbrace{(e^{at}\,H(t)) * (e^{at}\,H(t)) * \ldots * (e^{at}\,H(t))}_{n \text{ times}} = \frac{1}{(n-1)!}\,t^{n-1}\,e^{at}\,H(t).$$

In fact, this is true for $n = 1$. Assume it is true for $n = m$, and we prove it for $n = m+1$. We find

$$\begin{aligned}
\underbrace{(e^{at}\,H(t)) * (e^{at}\,H(t)) * \ldots * (e^{at}\,H(t))}_{m+1 \text{ times}} &= \underbrace{(e^{at}\,H(t)) * (e^{at}\,H(t)) * \ldots * (e^{at}\,H(t))}_{m \text{ times}} * (e^{at}\,H(t)) \\
&= \frac{1}{(m-1)!}\,t^{m-1}\,e^{at}\,H(t) * (e^{at}\,H(t)) \\
&= \int_0^t \frac{1}{(m-1)!}\,(t-\tau)^{m-1}\,e^{a(t-\tau)}\,e^{a\tau}\,\mathrm{d}\tau\,H(t) \\
&= \frac{1}{(m-1)!}\,e^{at}\int_0^t (t-\tau)^{m-1}\,\mathrm{d}\tau\,H(t) \\
&= \frac{1}{(m-1)!}\,e^{at}\,(-1)\int_t^0 y^{m-1}\,\mathrm{d}y\,H(t) \\
&= \frac{1}{(m-1)!}\,e^{at}\int_0^t y^{m-1}\,\mathrm{d}y\,H(t) \\
&= \frac{1}{(m-1)!}\,e^{at}\left[\frac{1}{m}\,y^m\right]_0^t H(t) = \frac{1}{m!}\,t^m\,e^{at}\,H(t),
\end{aligned}$$

where we have used the substitution $y = t - \tau$. It follows that if $\Re\mathrm{e}\{s\} > \Re\mathrm{e}\{a\}$

$$\frac{1}{(s-a)^n} = \mathcal{L}\left\{\frac{1}{(n-1)!}\,t^{n-1}\,e^{at}\,H(t)\right\}.$$

In particular, when $a = 0$,

$$\frac{1}{s^n} = \mathcal{L}\left\{\frac{1}{(n-1)!}\,t^{n-1}\,H(t)\right\} \quad \Rightarrow \quad \mathcal{L}\{t^n\,H\}(s) = \frac{n!}{s^{n+1}}. \tag{8.9}$$

Theorem 8.6 *Let $f \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ be $\mathcal{L}$-transformable. Then, for every $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} >$*

$\overline{\sigma_a(f)}$, there holds

$$\mathcal{L}\{f\}(\bar{s}) = \overline{\mathcal{L}\{f\}(s)}.$$

Proof Indeed,

$$\overline{\mathcal{L}\{f\}(s)} = \overline{\int_0^{+\infty} f(t)\, e^{-st}\, \mathrm{d}t} = \int_0^{+\infty} \overline{f(t)}\, \overline{e^{-st}}\, \mathrm{d}t = \int_0^{+\infty} f(t)\, e^{-\bar{s}t}\, \mathrm{d}t,$$

where we have used the fact that $f(t)$ is real and the fact that for any complex number z we have $\overline{e^z} = e^{\bar{z}}$. This second point can be proved straightforwardly setting $z = x + iy$ with $x, y \in \mathbb{R}$:

$$\begin{aligned} \overline{e^z} &= \overline{e^{x+iy}} = \overline{e^x\,(\cos y + i\, \sin y)} \\ &= e^x\,(\cos y - i\, \sin y) \\ &= e^x\,(\cos(-y) + i\, \sin(-y)) = e^{x-iy} = e^{\bar{z}}. \end{aligned}$$

8.1.1 Gamma function and initial/final value theorems

We turn our attention to two results that are very useful in applications: the final value and the initial value theorems. These theorems relate the behavior of the Laplace transform $\mathcal{L}\{f\}$ as s approaches zero or infinity to the behavior of f as t approaches infinity or zero, respectively. Informally, the initial value theorem relates the asymptotic behavior of $f(t)$ as $t \to 0^+$ to the asymptotic behavior of $\mathcal{L}\{f\}(s)$ as $s \to +\infty$, considering s to be real. More specifically, the initial value theorem asserts that

$$\lim_{t \to 0^+} f(t) = \lim_{s \to +\infty} s\, \mathcal{L}\{f\}(s),$$

where s is real, whenever the first limit exists and it is finite. The final value theorem connects the behavior of $f(t)$ as $t \to +\infty$ to the behavior of $\mathcal{L}\{f\}(s)$ as $s \to 0^+$, i.e.,

$$\lim_{t \to +\infty} f(t) = \lim_{\sigma \to 0^+} \sigma\, \mathcal{L}\{f\}(\sigma),$$

whenever the first limit exists and it is finite. These these theorems allow us to deduce the behavior of the Laplace transform from the behavior of the function.

Remark 8.8 Since f is defined except at points constituting a null set, we have to specify what we mean by limit. A function f has limit ℓ as $t \to 0^+$ if

$$\forall\, \varepsilon > 0, \quad \exists\, \delta > 0 : \quad |f(t) - \ell| < \varepsilon \ \text{ for almost all } t \in (0, \delta).$$

Similarly, a function f has limit ℓ as $t \to +\infty$ if

$$\forall\, \varepsilon > 0, \quad \exists\, p > 0 : \quad |f(t) - \ell| < \varepsilon \ \text{ for almost all } t \in (p, +\infty).$$

Before we present our main results on the initial and final value theorems, we must introduce the Euler Gamma function, which plays an important role in many areas of mathematics. The Gamma function is defined by

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1}\, e^{-t}\, \mathrm{d}t \qquad \text{if } \alpha > 0 \tag{8.10}$$

and

$$\Gamma(\alpha) = \frac{\Gamma(\alpha+1)}{\alpha} \text{if } -n-1 < \alpha < -n \text{ for } n \in \mathbb{N} \tag{8.11}$$

The domain of Γ is $\mathbb{R} \setminus \{-n \,|\, n \in \mathbb{N}\}$. In fact, when $\alpha = 0$, the integral does not converge (as the integrand becomes asymptotically equivalent to $1/t$), and (8.11) does not allow us to define $\Gamma(-1)$, $\Gamma(-2)$, etc. Hence, the value $\Gamma(\alpha)$ is well defined for every $\alpha \in \mathbb{R}$ such that $-\alpha \notin \mathbb{N}$, and it is strictly positive if $\alpha > 0$. Integration by parts also shows that (8.11) holds also for $\alpha > 0$. Since $\Gamma(1) = 1$, evaluating Γ at $n \in \mathbb{N}^*$ yields

$$\Gamma(n+1) = \int_0^{+\infty} t^n\, e^{-t}\, \mathrm{d}t = \Big[-t^n\, e^{-t}\Big]_0^{+\infty} - \int_0^{+\infty} -n\, t^{n-1}\, e^{-t}\, \mathrm{d}t = n\,\Gamma(n).$$

Hence, for all $n \in \mathbb{N}$ we find $\Gamma(n+1) = n!$, which shows that the Gamma function can be considered a generalization of the factorial. The value of $\Gamma(\alpha)$ can therefore be computed in finite terms if α is a natural number. It can also be computed in finite terms if $\alpha - \frac{1}{2}$ is an integer. Indeed, using the change of variable $\tau = t^2$,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{+\infty} t^{-\frac{1}{2}}\, e^{-t}\, \mathrm{d}t = \int_0^{+\infty} \frac{1}{\tau}\, e^{-\tau^2}\, 2\,\tau\, \mathrm{d}\tau = \int_0^{+\infty} e^{-\tau^2}\, \mathrm{d}\tau = \sqrt{\pi}.$$

Hence, for example,

$$\Gamma\left(-\frac{3}{2}\right) = \frac{\Gamma\left(-\frac{1}{2}\right)}{-\frac{3}{2}} = \frac{\Gamma\left(\frac{1}{2}\right)}{\left(-\frac{3}{2}\right)\left(-\frac{1}{2}\right)} = \frac{4}{3}\,\Gamma\left(\frac{1}{2}\right) = \frac{4}{3}\,\sqrt{\pi},$$

and

$$\Gamma\left(\frac{5}{2}\right) = \frac{\Gamma\left(\frac{3}{2}\right)}{\frac{3}{2}} = \frac{3}{2}\,\frac{1}{2}\,\Gamma\left(\frac{1}{2}\right) = \frac{3}{4}\,\sqrt{\pi}.$$

It is possible to generalize these expressions for a generic $n \in \mathbb{N}$:

$$\Gamma\left(n+\frac{1}{2}\right) = \frac{(2\,n)!}{4^n\, n!}\,\sqrt{\pi} \quad \text{and} \quad \Gamma\left(-n+\frac{1}{2}\right) = \frac{(-1)^n\, 4^n\, n!}{(2\,n)!}\,\sqrt{\pi}.$$

Remark 8.9 The same definition can be adapted to the case where $\alpha \in \mathbb{C}$, and the domain remains the same. Indeed, (8.10) is defined for $\mathfrak{Re}\{\alpha\} > 0$, since $|t^\alpha| = t^{\mathfrak{Re}\{\alpha\}}$. Thus, if $\mathfrak{Re}\{\alpha\} \in (-1, 0]$, (8.11) defines $\Gamma(\alpha)$ if we add the requirement $\alpha \neq 0$.

It is possible to show that Γ is holomorphic $\mathbb{C} \setminus \{-n \,|\, n \in \mathbb{N}\}$. The singularities $\{-n \,|\, n \in \mathbb{N}\}$ are simple poles, and we can also compute the corresponding residues. In fact, from (8.11) we find that if α is not integer and $n \in \mathbb{N}$,

$$\begin{aligned}\Gamma(\alpha+n+1) &= (\alpha+n)\,\Gamma(\alpha+n) = (\alpha+n)\,(\alpha+n-1)\,\Gamma(\alpha+n-1) \\ &= \ldots = (\alpha+n)\,(\alpha+n-1)\cdots\alpha\,\Gamma(\alpha) = \Gamma(\alpha)\prod_{k=0}^{n}(\alpha+k).\end{aligned}$$

Hence, for all $n \in \mathbb{N}$

$$\operatorname{Res}(\Gamma, -n) = \lim_{\alpha \to -n} (\alpha + n)\,\Gamma(\alpha) = \lim_{\alpha \to -n} \Gamma(\alpha + n + 1) \cdot (\alpha + n) \prod_{k=0}^{n} \frac{1}{\alpha + k} = \frac{(-1)^n}{n!}.$$

Now, we define

$$t_+^{\alpha-1} \stackrel{\text{def}}{=} \begin{cases} t^{\alpha-1} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0, \end{cases}$$

with $\alpha > 0$. With some abuse of notation, we denote the function $t \mapsto t_+^{\alpha-1}$ as $t_+^{\alpha-1}$. We notice that

$$\int_{-\infty}^{+\infty} e^{-\sigma t}\, t^{\alpha-1}\, H(t)\, \mathrm{d}t = \int_0^{+\infty} e^{-\sigma t}\, \frac{1}{t^{1-\alpha}}\, \mathrm{d}t$$

converges if and only if $\alpha > 0$ and $\sigma > 0$. It follows that $t_+^{\alpha-1}$ is $\mathcal{L}$-transformable and $\sigma_a(t_+^{\alpha-1}) = 0$. Moreover, for s real and positive, we have

$$\mathcal{L}\{t_+^{\alpha-1}\}(s) = \int_0^{+\infty} t^{\alpha-1}\, e^{-st}\, \mathrm{d}t = \int_0^{+\infty} e^{-\tau} \left(\frac{\tau}{s}\right)^{\alpha-1} \frac{1}{s}\, \mathrm{d}\tau = \frac{\Gamma(\alpha)}{s^\alpha}. \tag{8.12}$$

Since $\sigma_a(t_+^{\alpha-1}) = 0$, the function thus obtained can be extended analytically over the half-plane $\Re\mathrm{e}\{s\} > 0$. In this case, s^α is to be interpreted as $e^{\alpha \operatorname{Log} s}$, where Log represents the principal logarithm function. We notice how this formula reduces to (8.9) when α is a natural number ≥ 1.

Theorem 8.7 [Generalized Initial Value Theorem]
Let $f \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ be $\mathcal{L}$-transformable. Let $\alpha > 0$. If the limit $\lim_{t \to 0^+} \frac{f(t)}{t^{\alpha-1}}$ exists and is finite, then the limit

$$\lim_{\substack{s \in \mathbb{R} \\ s \to +\infty}} \frac{s^\alpha\, \mathcal{L}\{f\}(s)}{\Gamma(\alpha)}$$

exists and is finite, and the two limits coincide.

Proof Let $\ell = \lim_{t \to 0^+} \frac{f(t)}{t^{\alpha-1}}$. Let $\varepsilon > 0$. From Remark 8.8, there exists $\delta > 0$ such that

$$\left| \frac{f(t)}{t^{\alpha-1}} - \ell \right| < \frac{\varepsilon}{2}$$

for almost all $t \in (0, \delta)$. Let $\sigma > \max\{\sigma_a(f), \delta\}$, using (8.12) we find

$$\begin{aligned}
\frac{s^\alpha\, \mathcal{L}\{f\}(s)}{\Gamma(\alpha)} - \ell &= \frac{s^\alpha}{\Gamma(\alpha)} \left(\mathcal{L}\{f\}(s) - \ell\, \frac{\Gamma(\alpha)}{s^\alpha} \right) \\
&= \frac{s^\alpha}{\Gamma(\alpha)} \left(\mathcal{L}\{f\}(s) - \ell\, \mathcal{L}\{t_+^{\alpha-1}\}(s) \right) \\
&= \frac{s^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} (f(t) - \ell\, t^{\alpha-1})\, e^{-st}\, \mathrm{d}t \\
&= \frac{s^\alpha}{\Gamma(\alpha)} \left(\int_0^{\delta} t^{\alpha-1} \left(\frac{f(t)}{t^{\alpha-1}} - \ell \right) e^{-st}\, \mathrm{d}t + \int_\delta^{+\infty} (f(t) - \ell\, t^{\alpha-1})\, e^{-(s-\sigma)t}\, e^{-\sigma t}\, \mathrm{d}t \right).
\end{aligned}$$

Hence,

$$\begin{aligned}
\left|\frac{s^\alpha\,\mathcal{L}\{f\}(s)}{\Gamma(\alpha)}-\ell\right| &\le \frac{s^\alpha}{\Gamma(\alpha)}\left(\int_0^\delta t^{\alpha-1}\left|\frac{f(t)}{t^{\alpha-1}}-\ell\right| e^{-s\,t}\,\mathrm{d}t+\int_\delta^{+\infty}\left|f(t)-\ell\,t^{\alpha-1}\right| e^{-(s-\sigma)\,t}\,e^{-\sigma\,t}\,\mathrm{d}t\right)\\
&< \frac{s^\alpha}{\Gamma(\alpha)}\left(\frac{\varepsilon}{2}\int_0^\delta t^{\alpha-1}\,e^{-s\,t}\,\mathrm{d}t+e^{-(s-\sigma)\,\delta}\int_\delta^{+\infty}\left|f(t)-\ell\,t^{\alpha-1}\right| e^{-\sigma\,t}\,\mathrm{d}t\right)\\
&\le \frac{s^\alpha}{\Gamma(\alpha)}\left(\frac{\varepsilon}{2}\int_0^{+\infty} t^{\alpha-1}\,e^{-s\,t}\,\mathrm{d}t+e^{-(s-\sigma)\,\delta}\int_0^{+\infty}\left|f(t)-\ell\,t^{\alpha-1}\right| e^{-\sigma\,t}\,\mathrm{d}t\right)\\
&= \frac{s^\alpha}{\Gamma(\alpha)}\left(\frac{\varepsilon}{2}\,\mathcal{L}\{t_+^{\alpha-1}\}(s)+e^{-(s-\sigma)\,\delta}\,\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma)\right)\\
&= \frac{\varepsilon}{2}+\frac{s^\alpha}{\Gamma(\alpha)}\,e^{-(s-\sigma)\,\delta}\,\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma).
\end{aligned}$$

Choosing s to be sufficiently large, we can guarantee $\frac{s^\alpha}{\Gamma(\alpha)}\,e^{-(s-\sigma)\,\delta}\,\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma)\le\frac{\varepsilon}{2}$. Indeed, if $\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma)$ this is trivial. If $\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma)\neq 0$, it suffices to take $s>\sigma-\frac{1}{\delta}\,\ln\left(\frac{\varepsilon}{2}\,\frac{\Gamma(\alpha)}{s^\alpha}\,\frac{1}{\mathcal{L}\{|f(t)-\ell\,t_+^{\alpha-1}|\}(\sigma)}\right)$.

Corollary 8.1 *Let $f\in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ be $\mathcal{L}$-transformable. If the limit $\lim_{t\to 0^+} f(t)$ exists and is finite, then the limit*

$$\lim_{\substack{s\in\mathbb{R}\\ s\to+\infty}} s\,\mathcal{L}\{f\}(s)$$

exists and is finite, and the two limits coincide.

Proof The statement follows immediately from Theorem 8.8 by setting $\alpha=1$.

Theorem 8.8 [Generalized Final Value Theorem]
Let $f\in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ be $\mathcal{L}$-transformable. Let $\alpha>0$. If the limit $\lim_{t\to+\infty}\frac{f(t)}{t^{\alpha-1}}$ exists and is finite, then:

(i) $\sigma_a(f)\le 0$;

(ii) the limit

$$\lim_{\substack{s\in\mathbb{R}\\ s\to 0^+}}\frac{s^\alpha\,\mathcal{L}\{f\}(s)}{\Gamma(\alpha)} \tag{8.13}$$

exists and is finite;

(iii) and the two limits coincide.

Proof Let $\ell=\lim_{t\to+\infty}\frac{f(t)}{t^{\alpha-1}}$. It follows that f is asymptotically equivalent to $t^{\alpha-1}$ as $t\to+\infty$, so that the integral

$$\int_0^{+\infty} e^{-\sigma\,t}\,|f(t)|\,\mathrm{d}t$$

converges for every $\sigma>0$. This implies $\sigma_a(f)\le 0$, thus proving the first point. Notice also that since $\sigma_a(f)\le 0$, the limit (8.13) is defined. We now prove that the limit in (8.13) coincides with ℓ. Let $\varepsilon>0$. From

Remark 8.8, there exists $p > 0$ such that

$$\left|\frac{f(t)}{t^{\alpha-1}} - \ell\right| < \frac{\varepsilon}{2}$$

for almost all $t \in (p, +\infty)$. Let $\sigma > \max\{\sigma_a(f), p\}$, using (8.12) and repeating some of the steps of the proof in Theorem 8.8 we find

$$\begin{aligned}\frac{s^\alpha\,\mathcal{L}\{f\}(s)}{\Gamma(\alpha)} - \ell &= \frac{s^\alpha}{\Gamma(\alpha)}\int_0^{+\infty}(f(t) - \ell\, t^{\alpha-1})\,e^{-s\,t}\,\mathrm{d}t \\ &= \frac{s^\alpha}{\Gamma(\alpha)}\left(\int_p^{+\infty} t^{\alpha-1}\left(\frac{f(t)}{t^{\alpha-1}} - \ell\right)e^{-s\,t}\,\mathrm{d}t + \int_0^p (f(t) - \ell\, t^{\alpha-1})\,e^{-(s-\sigma)\,t}\,e^{-\sigma\,t}\,\mathrm{d}t\right).\end{aligned}$$

Hence,

$$\begin{aligned}\left|\frac{s^\alpha\,\mathcal{L}\{f\}(s)}{\Gamma(\alpha)} - \ell\right| &\le \frac{s^\alpha}{\Gamma(\alpha)}\left(\frac{\varepsilon}{2}\int_p^{+\infty} t^{\alpha-1}\,e^{-s\,t}\,\mathrm{d}t + \int_0^p e^{-s\,t}\,|f(t) - \ell\, t^{\alpha-1}|\,\mathrm{d}t\right) \\ &\le \frac{s^\alpha}{\Gamma(\alpha)}\left(\frac{\varepsilon}{2}\int_0^{+\infty} t^{\alpha-1}\,e^{-s\,t}\,\mathrm{d}t + \int_0^p |f(t) - \ell\, t^{\alpha-1}|\,\mathrm{d}t\right) = \frac{\varepsilon}{2} + \frac{s^\alpha}{\Gamma(\alpha)}\int_0^p |f(t) - \ell\, t^{\alpha-1}|\,\mathrm{d}t.\end{aligned}$$

Choosing $s > 0$ to be sufficiently small, we can render $\dfrac{s^\alpha}{\Gamma(\alpha)}\displaystyle\int_0^p |f(t) - \ell\, t^{\alpha-1}|\,\mathrm{d}t$ smaller than $\dfrac{\varepsilon}{2}$. This concludes the proof.

Corollary 8.2 [Final Value Theorem]
Let $f \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$ be $\mathcal{L}$-transformable. If the limit $\lim_{t\to+\infty} f(t)$ exists and is finite, then

(i) $\sigma_a(f) \le 0$

(ii) the limit

$$\lim_{\substack{s\to 0^+ \\ s\in\mathbb{R}}} s\,\mathcal{L}\{f\}(s)$$

exists and is finite;

(iii) the two limits coincide, i.e., $\lim_{t\to+\infty} f(t) = \lim_{\substack{s\to 0^+ \\ s\in\mathbb{R}}} s\,\mathcal{L}\{f\}(s)$.

Proof The proof follows from Theorem 8.8 by taking $\alpha = 1$.

Remark 8.10 In Theorem 8.2, it is essential to require that the limit $\lim_{t\to+\infty} f(t)$ exists. Indeed, the function $f(t) = (\sin t)\,H(t)$, defined for $t \in \mathbb{R}$, oscillates as $t \to +\infty$; on the other hand,

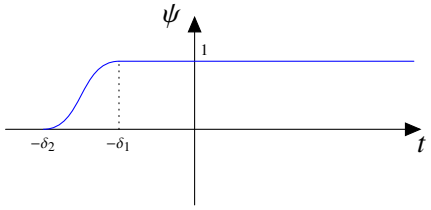
$$\lim_{s\to 0} s\,\mathcal{L}\{f\}(s) = \lim_{s\to 0}\frac{s}{s^2+1} = 0.$$

8.2 Laplace transform of distributions

Our aim is now to extend the notion of Laplace transform to a class of distributions. We have seen that, given an $\mathcal{L}$-transformable function $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$, if $\Re\mathrm{e}\{s\} > \sigma_a(f)$, we defined

$$\mathcal{L}\{f\}(s) = \int_0^{+\infty} e^{-s\,t}\, f(t)\,\mathrm{d}t.$$

Let us choose $\sigma \in \mathbb{R}$ so that $\sigma_a(f) < \sigma < \Re\mathrm{e}\{s\}$, and construct a function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$ of class C^∞ such that $\psi(t) = 1$ if $t \geq -\delta_1$ and $\psi(t) = 0$ for $t \leq \delta_0$, where $0 < \delta_1 < \delta_2$.



We have

$$\mathcal{L}\{f\}(s) = \int_{\mathbb{R}} \psi(t)\, f(t)\, e^{-s\,t}\,\mathrm{d}t = \int_{\mathbb{R}} \underbrace{e^{-\sigma\,t}\, f(t)}_{\in L^1(\mathbb{R};\mathbb{C})}\, e^{-(s-\sigma)\,t}\,\psi(t)\,\mathrm{d}t,$$

where $\sigma \in \mathbb{R}$ satisfies $\sigma_a(f) < \sigma < \Re\mathrm{e}\{s\}$. Since the function that maps t into $e^{-\sigma\,t}\, f(t)$ is in $L^1(\mathbb{R};\mathbb{C})$, it can be interpreted as a tempered distribution of function type. Accordingly, the function that maps t into $e^{-(s-\sigma)\,t}\,\psi(t)$ lives in $\mathcal{S}(\mathbb{R};\mathbb{C})$. First of all, it is of class C^∞ because such is ψ. Moreover, this function with all its derivatives multiplied by a polynomial is bounded over $\mathbb{R}$. To see this, define $v(t) = e^{-(s-\sigma)\,t}\,\psi(t)$. We show that $t^m\, v^{(j)}$ is bounded for all $m \in \mathbb{N}$ and for all $j \in \mathbb{N}$. Clearly, $v(t) = 0$ for $t < -\delta_2$, and therefore all its derivatives are also zero, so that $v^{(j)}(t) = 0$ for all $t < -\delta_2$, If $t \geq -\delta_1$, we have $v(t) = e^{-(s-\sigma)\,t}$, so that $v^{(j)}(t) = (\sigma - s)^j\, e^{-(s-\sigma)\,t}$, which, multiplied by t^m, remains bounded. Hence, for all $m \in \mathbb{N}$, the function $t^m\, v^{(j)}$ is bounded over $(-\delta_1, +\infty)$. Finally, $-\delta_2 \leq t \leq -\delta_2$ is bounded in view of Weierstrass extreme value theorem. These considerations suggest that we could consider

$$\mathcal{L}\{f\}(s) = \langle e^{-\sigma\,t}\, f(t), e^{-(s-\sigma)\,t}\,\psi(t)\rangle.$$

We introduce the following notation:

$$\mathcal{D}'_+ = \{T \in \mathcal{D}'(\mathbb{R};\mathbb{C}) \,|\, \operatorname{supp} T \subseteq [0, +\infty)\}.$$

Definition 8.4 [$\mathcal{L}$-Transformable Distribution]
Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$. We say that T is $\mathcal{L}$-transformable if

(i) $T \in \mathcal{D}'_+$;

(ii) there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$.

Remark 8.11 Let $T \in \mathcal{D}'_+$ and $e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Let $s \in \mathbb{C}$ with $\Re\mathrm{e}\{s\} > \sigma$.[a] Then, also $e^{-s\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Indeed, we can define $\psi \in C^\infty(\mathbb{R};\mathbb{R})$ in such a way that $\psi(t) = 1$ when $t > -\delta_1$ and $\psi(t) = 0$ for $t \leq -\delta_2$ with $0 < \delta_1 < \delta_2$. Therefore $\psi(t)$ is equal to one in an open set that contains

supp T, which implies $T = \psi T$. Hence

$$e^{-s\,t}\,T = (e^{-(s-\sigma)\,t}\,\psi)(e^{-\sigma\,t}\,T) \in \mathcal{S}'(\mathbb{R};\mathbb{C})$$

since $e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$ by assumption, and $e^{-(s-\sigma)\,t}\,\psi \in \mathcal{S}(\mathbb{R};\mathbb{C}) \subseteq \mathcal{O}_M$, because this function and its derivatives are bounded, since, for $t \geq 0$, we have

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^m (e^{-(s-\sigma)\,t}\,\psi(t)) = (\sigma - s)^m\,e^{-(s-\sigma)\,t}.$$

[a] It is easy to prove that one can also choose $\mathfrak{Re}\{s\} \geq \sigma$.

Definition 8.5 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. We define the* abscissa of convergence *as*

$$\sigma(T) \stackrel{def}{=} \inf\left\{\sigma \in \mathbb{R} \,|\, e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})\right\}.$$

In view of the considerations above, if $\sigma > \sigma(T)$, we have $e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Moreover, given $\mu \in \mathbb{R}$ satisfying $\sigma(T) < \mu < \sigma$, we have $e^{-\mu\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. We will shortly define the Laplace transform for those $s \in \mathbb{C}$ satisfying $\mathfrak{Re}\{s\} > \sigma(T)$ (which identifies, again, a half-plane in the complex plane), exactly as in the case of regular functions. Even if we talk about a half-plane, we recall that in some instances this could be the entire complex plane.

Remark 8.12 It is worth observing that the Laplace transform in the case of a distribution cannot certainly be a Lebesgue integral if T is a generic $\mathcal{L}$-transformable distribution. Therefore, the term "abscissa of *convergence*" may not seem appropriate. It would make more sense to refer to it as "abscissa of transformability" (and half-plane of transformability). However, we will continue to use the standard nomenclature, which has the advantage of uniformity.
However, in Definition 8.5, we omitted the adjective "absolute". In fact, since $L^1 \subset \mathcal{S}'$, every Laplace transformable function f is also a transformable distribution, and every value of σ that satisfies *(ii)* in Definition 8.1 (i.e., $t \mapsto e^{-\sigma\,t}\,f(t)$ is in $L^1(\mathbb{R};\mathbb{C})$) also satisfies *(ii)* in Definition 8.4 (i.e., $t \mapsto e^{-\sigma\,t}\,f(t)$ is in $\mathcal{S}'(\mathbb{R};\mathbb{C})$). It follows that $\sigma(f) \leq \sigma_a(f)$. On the other hand, the two abscissae can be different, as we will see via examples.

Definition 8.6 [Laplace Transform of a Distribution]
Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $s \in \mathbb{C}$ be such that $\mathfrak{Re}\{s\} > \sigma(T)$. Let $\psi \in C^\infty(\mathbb{R};\mathbb{R})$ be such that $\psi(t) = 1$ when $t > -\delta_1$ and $\psi(t) = 0$ for $t \leq -\delta_2$ with $0 < \delta_1 < \delta_2$. Finally, let $\sigma \in \mathbb{R}$ be such that $\sigma(T) < \sigma < \mathfrak{Re}\{s\}$. We define

$$\mathcal{L}\{T\}(s) \stackrel{def}{=} \langle e^{-\sigma\,t}\,T, e^{-(s-\sigma)\,t}\,\psi(t)\rangle.$$

The definition given for the Laplace transform of a distribution is well posed because $e^{-\sigma\,t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$ and $e^{-(s-\sigma)\,t}\,\psi(t) \in \mathcal{S}(\mathbb{R};\mathbb{C})$ (notice that the latter is obtained by applying a temperate distribution to an element of $\mathcal{S}(\mathbb{R};\mathbb{C})$).

On the other hand, in this definition there is arbitrariness in the choice of σ and ψ. We need to prove that the definition of Laplace transform of a distribution does not depend on the choice of σ and ψ.

(i) We show that the definition of Laplace transform of a distribution does not depend on the choice of σ. Let $\sigma(T) < \sigma_1 < \sigma_2 < \Re\mathrm{e}\{s\}$. We want to prove that

$$\langle e^{-\sigma_2 t}\, T, e^{-(s-\sigma_2)t}\, \psi(t)\rangle = \langle e^{-\sigma_1 t}\, T, e^{-(s-\sigma_1)t}\, \psi(t)\rangle.$$

We define $g \in C^\infty(\mathbb{R};\mathbb{R})$ to be such that $g(t) = 1$ if $t \geq -\delta_2$ and $g(t) = 0$ for $t \leq -\delta_3$, where $\delta_2 < \delta_3$, so that $\psi\, g = \psi$.

We have

$$e^{-\sigma_2 t}\, T = e^{-\sigma_2 t}\, g(t)\, T = e^{-(\sigma_2-\sigma_1)t}\, g(t)\, e^{-\sigma_1 t}\, T$$

where $e^{-(\sigma_2-\sigma_1)t}\, g(t) \in \mathcal{S}(\mathbb{R};\mathbb{C})$ since $\sigma_2 > \sigma_1$. Thus

$$\begin{aligned}\langle e^{-\sigma_2 t}\, T, e^{-(s-\sigma_2)t}\, \psi(t)\rangle &= \langle (e^{-(\sigma_2-\sigma_1)t}\, g(t))\, e^{-\sigma_1 t}\, T, e^{-(s-\sigma_2)t}\, \psi(t)\rangle \\ &= \langle e^{-\sigma_1 t}\, T, e^{-(\sigma_2-\sigma_1)t}\, g(t)\, e^{-(s-\sigma_2)t}\, \psi(t)\rangle \\ &= \langle e^{-\sigma_1 t}\, T, e^{-(s-\sigma_1)t}\, \psi(t)\rangle.\end{aligned}$$

(ii) We show that the definition of Laplace transform of a distribution does not depend on the choice of ψ. Let $\psi_1, \psi_2 \in C^\infty(\mathbb{R};\mathbb{R})$ be such that, for $j \in \{1,2\}$, we have $\psi_j(t) = 1$ for $t \geq -\alpha_j$ and $\psi_j(t) = 0$ for $t \leq -\beta_j$ with $0 < \alpha_j < \beta_j$. Let $\Re\mathrm{e}\{s\} > \sigma(T)$ and let σ be such that $\sigma(T) < \sigma < \Re\mathrm{e}\{s\}$. We want to show that

$$\langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}\, \psi_1(t)\rangle = \langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}\, \psi_2(t)\rangle$$

or, equivalently, such that

$$\langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}\, (\psi_1(t) - \psi_2(t))\rangle = 0.$$

Consider the term $e^{-(s-\lambda)t}\, (\psi_1(t) - \psi_2(t))$, or, more precisely, the function

$$\begin{aligned}\zeta : \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\mapsto e^{-(s-\lambda)t}\, (\psi_1(t) - \psi_2(t)).\end{aligned}$$

This function is in $\mathcal{D}(\mathbb{R};\mathbb{C})$. Indeed, $\zeta(t) = 0$ if $t \leq \min\{-\beta_1, -\beta_2\}$ and $\zeta(t) = 0$ if $t \geq \max\{-\alpha_1, -\alpha_2\}$. Thus, $\operatorname{supp}\zeta \subseteq [\min\{-\beta_1, -\beta_2\}, \max\{-\alpha_1, -\alpha_2\}]$.
The interval $[\min\{-\beta_1, -\beta_2\}, \max\{-\alpha_1, -\alpha_2\}]$ has empty intersection with $[0, +\infty)$, while $\operatorname{supp}(e^{-\sigma t}\, T) \subseteq [0, +\infty)$. Thus, since the supports are disjoint, i.e.,

$$\operatorname{supp}(e^{-\sigma t}\, T) \cap \operatorname{supp}\, (e^{-(s-\sigma)t}(\psi_1 - \psi_2)) = \varnothing,$$

we get $\langle e^{-\sigma t}\, T, \zeta\rangle = \langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}(\psi_1 - \psi_2)\rangle = 0$.

The Laplace transform of a distribution is a function defined over a certain half-plane in the complex plane. For example, a distribution $T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$ with $\operatorname{supp} T \subseteq [0, +\infty)$ is clearly $\mathcal{L}$-transformable, since $e^{-0\cdot t}\, T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$, and therefore $\sigma(T) \leq 0$.

Things become considerably easier for distributions with compact support.

Theorem 8.9 [Laplace Transform for Distributions with Compact Support]
Let $T \in \mathcal{D}'_+ \cap \boldsymbol{E}'(\mathbb{R};\mathbb{C})$. Then, T is $\mathcal{L}$-transformable and $\sigma(T) = -\infty$. Moreover,

$$\mathcal{L}\{T\}(s) = \langle T, e^{-st}\rangle.$$

Proof We start proving that T is $\mathcal{L}$-transformable and $\sigma(T) = -\infty$. For all $\sigma \in \mathbb{R}$, the distribution $e^{-\sigma t}\,T$ has compact support, because its support is contained in $\operatorname{supp} T$. Therefore, it is a tempered distribution, and the infimum is $-\infty$. In other words, the Laplace transform is defined for all $s \in \mathbb{C}$.

Let $s \in \mathbb{C}$ such that $\sigma < \Re\mathrm{e}\{s\}$. We have

$$\mathcal{L}\{T\}(s) = \langle e^{-\sigma t}\,T, e^{-(s-\sigma)t}\,\psi(t)\rangle.$$

Let $g \in \mathcal{D}(\mathbb{R};\mathbb{C})$ be such that $g(t) = 1$ on an open set U containing $\operatorname{supp} T$. Then

$$\mathcal{L}\{T\}(s) = \langle e^{-\sigma t}\,T, e^{-(s-\sigma)t}\,\psi(t)\,g(t)\rangle = \langle T, e^{-st}\,\psi(t)\,g(t)\rangle.$$

Function g is such that $\psi\,g = g$, and since there exists $R > 0$ such that $\operatorname{supp} T \subseteq [0, R]$, we choose g in such a way that g has compact support in $[-\delta_1, \infty)$, where ψ is equal to 1. Thus, $\mathcal{L}\{T\}(s) = \langle T, e^{-st}\,g(t)\rangle$.

Example Let $s \in \mathbb{C}$. We want to determine the Laplace transform of the Dirac δ distribution, whose support is compact (we recall that $\operatorname{supp}\delta = \{0\}$). We have

$$\mathcal{L}\{\delta\}(s) = \langle \delta, e^{-st}\rangle = \langle \delta, e^{-st}\,g(t)\rangle$$

where $g \in \mathcal{D}(\mathbb{R};\mathbb{C})$ is such that $g(t) = 1$ in a neighbourhood of 0. Thus,

$$\mathcal{L}\{\delta\}(s) = e^{-s\,0}\,g(0) = 1.$$

Example Let $s \in \mathbb{C}$. We want to determine the Laplace transform of δ', whose support is compact (we recall that $\operatorname{supp}\delta' = \{0\}$). We have

$$\begin{aligned}\mathcal{L}\{\delta'\}(s) &= \langle \delta', e^{-st}\rangle = \langle \delta', e^{-st}\,g(t)\rangle \\ &= -\langle \delta, -s\,e^{-st}\,g(t) + e^{-st}\,g'(t)\rangle \\ &= s\,g(0) - g'(0) = s,\end{aligned}$$

because $g(0) = 1$ and $g'(0) = 0$ (the function g is equal to 1 in an interval centered at 0).

Example Consider the distribution

$$T = \sum_{k=0}^{\infty} c_k\,\delta(t-k),$$

where, for all $k \in \mathbb{N}$, we have $c_k \in \mathbb{C}$. This definition means that T is given by the limit

$$\lim_{h\to\infty} \sum_{k=0}^{h} c_k\,\delta(t-k)$$

in $\mathcal{D}'(\mathbb{R};\mathbb{C})$. We show that a functional of this type is a distribution. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$. We have

$$\langle T, \varphi \rangle = \lim_{h\to\infty} \sum_{k=0}^{h} c_k \langle \delta(t-k), \varphi(t) \rangle = \lim_{h\to\infty} \sum_{k=0}^{h} c_k \varphi(k).$$

We know that φ has compact support. From a certain value of k onward, we will have that the value of φ is zero (or, in other words, the set $\{k \in \mathbb{N} \,|\, \varphi(k) \neq 0\}$ is finite) and therefore the series converges, so that

$$\langle T, \varphi \rangle = \sum_{k=0}^{\infty} c_k \varphi(k),$$

where the summation is made up of a finite number of terms. From the theorem of completeness, this is a distribution, independently from c_k. We now want to find what conditions we must impose on the coefficients to ensure that this is a tempered distribution. We show that a sufficient condition is the existence of a constant $A \geq 0$ and or a number $m \in \mathbb{N}^*$ such that $|c_k| \leq A\,(1+k)^m$ for all $k \in \mathbb{N}$. In other words, the coefficients must have at most a polynomial growth. Indeed, given $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$, the series

$$\sum_{k=0}^{\infty} c_k \varphi(k) \tag{8.14}$$

is absolutely convergent, because $(1+k)^p\,|\varphi(k)|$ is bounded for all $p \in \mathbb{N}$. It follows that for all $k \in \mathbb{N}$ we have $(1+k)^p\,|\varphi(k)| \leq c_p$, and, therefore,

$$|c_k \varphi(k)| \leq A\,(1+k)^m\,(1+k)^p\,(1+k)^{-p}\,|\varphi(k)| \leq A\,c_p\,(1+k)^{m-p}.$$

Letting $p > m+1$, the series

$$\sum_{k=0}^{\infty} (1+k)^{m-p}$$

is convergent, and therefore (8.14) is absolutely convergent. We show that T is tempered. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence in $\mathcal{S}(\mathbb{R};\mathbb{C})$ such that $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R};\mathbb{C})$, and verify that

$$\lim_{h\to\infty} \sum_{k=0}^{h} c_k \varphi_h(k) = \sum_{k=0}^{\infty} c_k \varphi(k).$$

To this end, we study the quantity $\left|\sum_{k=0}^{\infty} c_k\,(\varphi_h(k) - \varphi(k)\right|$. The convergence of $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{S}(\mathbb{R};\mathbb{C})$ implies that $(1+|x|)^p\,|\varphi_h(x) - \varphi(x)|$ goes to zero uniformly in $\mathbb{R}$ for all $p \in \mathbb{N}$ as $h \to \infty$. Hence, if we fix $p > m+1$, for all $\varepsilon > 0$ there exists $k_\varepsilon \in \mathbb{N}$ such that if $k \geq k_\varepsilon$, then $(1+|x|)^p\,|\varphi_h(x) - \varphi(x)| \leq \varepsilon$ for all $x \in \mathbb{R}$. This implies that, for all $k \in \mathbb{N}$, we have also $(1+k)^p\,|\varphi_h(k) - \varphi(k)| \leq \varepsilon$ if $k \geq k_\varepsilon$, so that

$$\left|\sum_{k=0}^{\infty} c_k\,(\varphi_h(k) - \varphi(k)\right| \leq \sum_{k=0}^{\infty} |c_k|\,|\varphi_h(k) - \varphi(k)| \leq A \sum_{k=0}^{\infty} (1+k)^{m-p}\,(1+k)^p\,|\varphi_h(k) - \varphi(k)| \leq A\,\varepsilon \sum_{k=0}^{\infty} |(1+k)^{m-p}$$

if $k \geq k_\varepsilon$. Since the series $\sum_{k=0}^{\infty} (1+k)^{m-p}$ is convergent in view of our choice of p, we obtain the result.

Hence, now we assume that $|c_k| \leq A\,(1+k)^m$. Then, T is $\mathcal{L}$-transformable and $\sigma(T) \leq 0$.

It is also possible to find a less restrictive conditions than imposing $T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. It is possible to verify that T is $\mathcal{L}$-transformable if there exists $\alpha \in \mathbb{R}$ such that $|c_k| \leq A\, e^{\alpha k}$ for all $k \in \mathbb{N}$, where $A \geq 0$. In fact, given $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$, we find

$$\langle e^{-\sigma t}\, T, \varphi\rangle = \langle T, e^{-\sigma t}\, \varphi\rangle = \sum_{k=0}^{\infty} c_k\, e^{-\sigma k}\, \varphi(k),$$

where $|c_k|\, e^{-\sigma k} \leq A\, e^{(\alpha-\sigma)k}$ from the previous argument, and we deduce that $e^{-\sigma t}\, T$ is tempered. Hence T is $\mathcal{L}$-transformable if $\sigma \geq \alpha$, from which we find $\sigma(T) \leq \alpha$. We are now ready to determine $\mathcal{L}\{T\}$ for $s \in \mathbb{C}$ with $\Re\mathrm{e}\{s\} > \alpha$:

$$\mathcal{L}\{T\}(s) = \langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}\, \psi(t)\rangle, \quad \text{with } \alpha < \sigma < \Re\mathrm{e}\{s\},$$

so that

$$\mathcal{L}\{T\}(s) = \sum_{k=0}^{\infty} c_k\, e^{-\sigma k}\, e^{-(s-\sigma)k}\, \psi(k) = \sum_{k=0}^{\infty} c_k\, e^{-sk}.$$

Theorem 8.10 *Let $T \in \mathcal{D}'_+(\mathbb{R};\mathbb{C})$ and $(T_k)_{k\in\mathbb{N}}$ be a sequence of elements of $\mathcal{D}'_+(\mathbb{R};\mathbb{C})$. Let $\sigma \in \mathbb{R}$ be such that*

- *$e^{-\sigma t}\, T_k \in \mathcal{S}'(\mathbb{R};\mathbb{C})$ for all $k \in \mathbb{N}$;*
- *$e^{-\sigma t}\, T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$;*
- *$e^{-\sigma t}\, T = \lim_{k\to\infty} e^{\sigma t}\, T_k$ in $\mathcal{S}'(\mathbb{R};\mathbb{C})$.*

Then, T is $\mathcal{L}$-transformable, and T_k is $\mathcal{L}$-transformable for all $k \in \mathbb{N}$. Moreover, for $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \sigma$ we have

$$\lim_{k\to\infty} \mathcal{L}\{T_k\}(s) = \mathcal{L}\{T\}(s).$$

Proof There holds

$$\mathcal{L}\{T_k\}(s) = \langle e^{-\sigma t}\, T_k(t), e^{-(s-\sigma)t}\psi(t)\rangle \xrightarrow{k\to\infty} \langle e^{-\sigma t}\, T, e^{-(s-\sigma)t}\psi(t)\rangle = \mathcal{L}\{T\}(s).$$

since $e^{-\sigma t}\, T_k$ converges in $\mathcal{S}'(\mathbb{R};\mathbb{C})$ and $e^{-(s-\sigma)t}\psi(t) \in \mathcal{S}(\mathbb{R};\mathbb{C})$.

Example

Consider the distribution

$$T(t) = \delta(t) + 2\sum_{k=0}^{\infty}(-1)^k\, \delta(t-2k).$$

This distribution is similar to the one of the previous example, if we define

$$c_k = \begin{cases} 0 & \text{if } k \text{ is odd} \\ 1 & \text{if } k = 0 \\ (-1)^h & \text{if } h = 2k \end{cases}$$

Notice that the sequence $(c_k)_{k\in\mathbb{N}}$ is bounded. With the same argument used earlier, considering the partial sums we see that

$$T_k(t) = \sum_{j=0}^{k} c_j\, \delta(t-\xi) \xrightarrow{k\to\infty} T$$

in $\mathcal{S}'(\mathbb{R};\mathbb{C})$. If $s \in \mathbb{C}$ is such that $\mathfrak{Re}\{s\} > 0$, then

$$\mathcal{L}\{T_k\}(s) \xrightarrow{k\to\infty} \mathcal{L}\{T\}(s)$$

pointwise.

We now examine the theorem of the Laplace transform of the distributional derivative.

Theorem 8.11 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Then, T' is $\mathcal{L}$-transformable and for all $s \in \mathbb{C}$ satisfying $\mathfrak{Re}\{s\} > \sigma(T)$ we have*

$$\mathcal{L}\{T'\}(s) = s\,\mathcal{L}\{T\}(s).$$

Proof We prove that T' is $\mathcal{L}$-transformable. We need to verify two conditions: firstly, that $\operatorname{supp} T' \subseteq [0,+\infty)$; this condition is satisfied because $\operatorname{supp} T' \subseteq \operatorname{supp} T \subseteq [0,+\infty)$. Secondly, we need to show that there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma t}T'$ is tempered. We know that there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma t}T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Thus, $(e^{-\sigma t}T)' \in \mathcal{S}'(\mathbb{R};\mathbb{C})$, which in turn means that $-\sigma\, e^{-\sigma t}T + e^{-\sigma t}T' \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Thus,

$$e^{-\sigma t}T' = (e^{-\sigma t}T)' + \sigma\, e^{-\sigma t}T.$$

Since both $(e^{-\sigma t}T)'$ and $\sigma\, e^{-\sigma t}T$ are in $\mathcal{S}'(\mathbb{R};\mathbb{C})$, we conclude that $e^{-\sigma t}T' \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. We have proved that T' is $\mathcal{L}$-transformable. We also have that $\sigma(T') \le \sigma(T)$. Let $\mathfrak{Re}\{s\} > \sigma(T)$. Let us choose σ to satisfy $\sigma(f) < \sigma < \mathfrak{Re}\{s\}$. We find

$$\begin{aligned}
\mathcal{L}\{T'\}(s) &= \langle e^{-\sigma t}T'(t), e^{-(s-\sigma)t}\psi(t)\rangle \\
&= \langle (e^{-\sigma t}T)', e^{-(s-\sigma)t}\psi(t)\rangle + \lambda\,\langle e^{-\sigma t}T, e^{-(s-\sigma)t}\psi(t)\rangle \\
&= -\langle e^{-\sigma t}T, (\sigma - s)\,e^{-(s-\sigma)t}\psi(t) + e^{-(s-\sigma)t}\psi'(t)\rangle + \lambda\,\mathcal{L}\{T\}(s) \\
&= (s-\sigma)\,\mathcal{L}\{T\}(s) - \langle e^{-\sigma t}T, e^{-(s-\sigma)t}\psi'(t)\rangle + \sigma\,\mathcal{L}\{T\}(s) \\
&= s\,\mathcal{L}\{T\}(s) - \langle e^{-\sigma t}T, e^{-(s-\sigma)t}\psi'(t)\rangle.
\end{aligned}$$

We prove that $\langle e^{-\sigma t}T, e^{-(s-\sigma)t}\psi'(t)\rangle = 0$. This is a consequence of the fact that $\psi'(t) = 0$ for $t \notin [-\delta_2, -\delta_1]$ and $\operatorname{supp}(e^{-\sigma t}T) \subseteq [0,+\infty)$. Therefore, the supports of the distribution and of the test function to which the distribution is applied are disjoint:

$$\operatorname{supp}(e^{-\sigma t}T) \cap \operatorname{supp}\left(e^{-(s-\sigma)t}\psi'(t)\right) = \varnothing,$$

so that

$$\langle e^{-\sigma t}T, e^{-(s-\sigma)t}\psi'(t)\rangle = 0,$$

and this completes the proof.

Remark 8.13 When $f \in L^1_{\text{loc}}(\mathbb{R};\mathbb{C})$, $\operatorname{supp} f \subseteq [0,+\infty)$ and f is $\mathcal{L}$-transformable, we recall that $f|_{[0,+\infty)} \in C^1([0,+\infty);\mathbb{C})$ and

$$\widetilde{f'}(t) = \begin{cases} f'(t) & t > 0 \\ 0 & t \le 0. \end{cases}$$

If $\widetilde{f'}$ is $\mathcal{L}$-transformable, then $\mathcal{L}\{\widetilde{f'}\}(s) = s\,\mathcal{L}\{f\}(s) - f(0^+)$. If we now think about f as a function-type distribution and we compute its derivative in the sense of distributions, we find

$$f' = \widetilde{f'} + f(0^+)\,\delta$$

from which we immediately obtain $\widetilde{f'} = f' - f(0^+)\,\delta$. It follows that

$$\mathcal{L}\{\widetilde{f'}\}(s) = \mathcal{L}\{f'\}(s) - f(0^+)\,\mathcal{L}\{\delta\}(s) = \mathcal{L}\{f'\}(s) - f(0^+),$$

and we re-obtain the formula already found previously.

Remark 8.14 By iterating the result of Theorem 8.11, it is easy to see that if $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ is $\mathcal{L}$-transformable, then for all $k \in \mathbb{N}$ the distribution $T^{(k)}$ is $\mathcal{L}$-transformable and for all $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \sigma(T)$ we have

$$\mathcal{L}\{T^{(k)}\}(s) = s^k\,\mathcal{L}\{T\}(s).$$

Example

We have for all $s \in \mathbb{C}$ and for all $k \in \mathbb{N}$

$$\mathcal{L}\{\delta^{(k)}\}(s) = s^k.$$

Likewise, for every $t_0 \geq 0$

$$\mathcal{L}\{\delta^{(k)}(t - t_0)\}(s) = s^k\,e^{-a\,t_0}.$$

We notice that from the Theorem 8.11 it follows that $\sigma(T') \leq \sigma(T)$. As anticipated in Remark 8.12, for a distribution of function type the abscissa of absolute convergence may be different from the abscissa of convergence, as the following example shows.

Example

Consider the distribution of function type

$$T = \sin(e^t - 1)\,H \in \mathcal{D}'(\mathbb{R};\mathbb{R}).$$

It is easy to see that it is $\mathcal{L}$-transformable and its abscissa of absolute convergence is $\sigma_a(T) = 0$. Indeed,

$$\int_0^{+\infty} e^{-\sigma\,t}\left|\sin(e^t - 1)\right|\mathrm{d}t$$

is finite if and only if $\sigma > 0$. Let us consider the distributional derivative of T:

$$T' = e^t\,\cos(e^t - 1)\,H(t),$$

which is still $\mathcal{L}$-transformable and its abscissa of absolute convergence is $\sigma_a(T') = 1$, since

$$\int_0^{+\infty} e^{-\sigma\,t}\,e^t\left|\cos(e^t - 1)\right|\mathrm{d}t = \int_0^{+\infty} e^{(1-\sigma)\,t}\left|\cos(e^t - 1)\right|\mathrm{d}t$$

is finite if and only if $\sigma > 1$. Recalling that for a distribution of function type we have $\sigma(T) \leq \sigma_a(T)$, we have the chain of inequalities

$$\sigma(T') \leq \sigma(T) \leq \sigma_a(T) < \sigma_a(T').$$

The first inequality follows from Theorem 8.11. The second is a consequence of the discussion in Remark 8.12. This example shows once again that there exist distributions of function type whose abscissa of convergence is strictly smaller than the corresponding abscissa of absolute convergence; indeed, we found a distribution of function type T' for which $\sigma(T') < \sigma_a(T')$. This shows, in particular, that the "ordinary" Laplace transform is not the restriction of the Laplace transform for distribution to the set of $\mathcal{L}$-transformable functions.

The following example shows that it is possible to have a function which is not Laplace transformable in the regular sense, but which is still Laplace transformable in the sense of distributions: in this case the abscissa of absolute convergence is not defined, but the abscissa of convergence is.

Example Consider the distribution of function type

$$T = \sin(e^{t^2} - 1)\, H \in \mathcal{D}'(\mathbb{R};\mathbb{R}).$$

From

$$\int_0^{+\infty} e^{-\sigma t} \left|\sin(e^{t^2} - 1)\right| \mathrm{d}t < +\infty \quad \Leftrightarrow \quad \sigma > 0$$

it follows immediately that T is $\mathcal{L}$-transformable as a regular function, and $\sigma_a(T) = 0$. The distributional derivative

$$T' = 2\, t\, e^{t^2} \cos(e^{t^2} - 1)\, H,$$

which is still of function type, is therefore $\mathcal{L}$-transformable as a distribution from Theorem 8.11. However, it is not $\mathcal{L}$-transformable as a regular function, because

$$\nexists\, \sigma \in \mathbb{R} : \quad \int_0^{+\infty} e^{-\sigma t - t^2}\, e^{-\sigma t} \left|\cos(e^{t^2} - 1)\right| \mathrm{d}t < +\infty.$$

Hence, in this case $\sigma_a(T')$ is not defined, while $\sigma(T')$ is defined.

Theorem 8.12 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $s_0 \in \mathbb{C}$. Then, $e^{s_0 t}\, T$ is $\mathcal{L}$-transformable and, for $s \in \mathbb{C}$ satisfying $\mathfrak{Re}\{s\} > \mathfrak{Re}\{s_0\} + \sigma(T)$, there holds*

$$\mathcal{L}\{e^{s_0 t}\, T\}(s) = \mathcal{L}\{T\}(s - s_0).$$

Proof First we show that $e^{s_0 t}\, T$ is $\mathcal{L}$-transformable. We have

$$\mathrm{supp}(e^{s_0 t}\, T) \subseteq \mathrm{supp}\, T \subseteq [0, +\infty).$$

We only need to establish if there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma t}(e^{s_0 t}\, T) \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. We have $e^{-\sigma t}(e^{s_0 t}\, T) = e^{-(\sigma - s_0) t}\, T$, which is tempered if $\sigma - \mathfrak{Re}\{s_0\} > \sigma(T)$, i.e., if $\sigma > \sigma(T) + \mathfrak{Re}\{s_0\}$.

Let now $s \in \mathbb{C}$ satisfy $\mathfrak{Re}\{s\} > \mathfrak{Re}\{s_0\} + \sigma(T)$, and let $\sigma \in \mathbb{R}$ satisfy $\mathfrak{Re}\{s_0\} + \sigma(T) < \sigma < \mathfrak{Re}\{s\}$. We find

$$\mathcal{L}\{e^{s_0 t}\}(s) = \langle e^{-\sigma t}\, e^{s_0 t}\, T, e^{-(s-\sigma) t}\, \psi(t)\rangle.$$

Since the function $\mathbb{R} \longrightarrow \mathbb{C}$ that maps t into $e^{i\, \mathfrak{Im}\{s_0\} t}$ is in $\boldsymbol{O}_M$, because it is bounded with its derivatives in $\mathbb{R}$, we can write

$$\mathcal{L}\{e^{s_0 t}\}(s) = \langle e^{-(\sigma - \mathfrak{Re}\{s_0\}) t}\, T, e^{-(s - \sigma - i\, \mathfrak{Im}\{s_0\}) t}\, \psi(t)\rangle.$$

Let $\mu = \sigma - \mathfrak{Re}\{s_0\}$, so that $\mu > \sigma(T)$. Then,

$$s - \sigma - i\, \mathfrak{Im}\{s_0\} = s - i\, \mathfrak{Im}\{s_0\} - \sigma - \mu + \mu = s - i\, \mathfrak{Im}\{s_0\} - \sigma + \sigma - \mathfrak{Re}\{s_0\} - \mu = s - s_0 - \mu,$$

and therefore

$$\mathcal{L}\{e^{s_0 t}\}(s) = \langle e^{-\mu t}\, T, e^{-(s - s_0 - \mu) t}\, \psi(t)\rangle = \mathcal{L}\{T\}(s - s_0).$$

We have seen that, for regular functions, the Laplace transform is holomorphic. This is also true in the case of distributions. However, while for regular functions the Laplace transform is bounded for $s \in \mathbb{C}$ satisfying the condition $\mathfrak{Re}\{s\} > \sigma > \sigma_a(f)$, this is not generally the case for distributions. These results can be obtained as a consequence of the following important result (that we will not prove), which shows that all tempered distributions can be obtained by taking the derivative of a suitable order of a certain continuous function with slow growth.

Theorem 8.13 [Schwartz Representation Theorem for Tempered Distributions]
Let $T \in \mathcal{S}'(\mathbb{R};\mathbb{C}$. Then, there exists a continuous function $f : \mathbb{R} \longrightarrow \mathbb{C}$ of slow (polynomial) growth (i.e., such that there exists $c \geq 0$ and $m \in \mathbb{N}$ for which we can write $|f(t)| \leq c\,(1 + |t|)^m$) for which there exists $n \in \mathbb{N}$ satisfying

$$T = f^{(n)}.$$

Moreover, if $\operatorname{supp} T \subseteq [0, +\infty)$, *the function f can be chosen to satisfy* $\operatorname{supp} f \subseteq [0, +\infty)$.

We use this result to prove the analyticity of the Laplace transform.

Theorem 8.14 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. The function*

$$\begin{aligned} \mathcal{L}\{T\} : \{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \sigma(T)\} &\longrightarrow \mathbb{C} \\ s &\mapsto \mathcal{L}\{T\}(s) \end{aligned}$$

is holomorphic. Moreover, if $\sigma > \sigma(T)$, there exists $c \geq 0$ and $m \in \mathbb{N}$ for which $|\mathcal{L}\{T\}(s)| \leq c\,(1 + |s|)^m$ for all $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \sigma$.

Proof Let $\sigma > \sigma(T)$ and take $\sigma' \in \mathbb{R}$ to satisfy $\sigma(T) < \sigma' < \sigma$. Then, $e^{-\sigma' t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$, and therefore it is $\mathcal{L}$-transformable. From Theorem 8.13, there exists a continuous function $f : \mathbb{R} \longrightarrow \mathbb{C}$ with slow growth and $\operatorname{supp} f \subseteq [0, +\infty)$ (because the support of $e^{-\sigma' t}\,T$ is in $[0, +\infty)$) such that $e^{-\sigma' t}\,T = f^{(n)}$ for some $n \in \mathbb{N}$. Hence,

$$T = e^{\sigma' t}\,(e^{-\sigma' t}\,T) = e^{\sigma' t}\,f^{(n)}.$$

Since f has slow growth, $\sigma_a(f) \leq 0$. Thus, choosing $\mu > 0$, we have $e^{-\mu t}\,f \in L^1(\mathbb{R};\mathbb{C})$, because the exponential prevails over the polynomial. It follows that the Laplace transform of f is defined for all $s \in \mathbb{C}$ with $\Re\mathrm{e}\{s\} > 0$. Therefore, if $\Re\mathrm{e}\{s\} > 0$, the Laplace transform $\mathcal{L}\{f\}(s)$ is defined, $\mathcal{L}\{f^{(n)}\}(s)$ is defined, and

$$\mathcal{L}\{f^{(n)}\}(s) = s^n\,\mathcal{L}\{f\}(s).$$

Now, let $s \in \mathbb{C}$ satisfy $\Re\mathrm{e}\{s\} > \sigma'$. We have

$$\mathcal{L}\{T\}(s) = \mathcal{L}\{e^{\sigma' t}\,f^{(n)}\}(s) = \mathcal{L}\{f^{(n)}\}(s - \sigma') = (s - \sigma')^n\,\mathcal{L}\{f\}(s - \sigma').$$

Since $\mathcal{L}\{f\}$ is holomorphic in $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > 0\}$, we find that $\mathcal{L}\{T\}$ is holomorphic in $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \sigma'\}$. Since we chose σ' to satisfy $\sigma(T) < \sigma' < \sigma$, we conclude that $\mathcal{L}\{T\}$ is holomorphic in $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \sigma(T)\}$.

We now prove the second part of the theorem. We know that $\mathcal{L}\{f\}$ is bounded over $\{s \in \mathbb{C} \,|\, \Re\mathrm{e}\{s\} > \mu\}$, for all $\mu > \sigma_a(f) = 0$. Let us take $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma$. Then,

$$\Re\mathrm{e}(s - \sigma') > \sigma - \sigma' > 0.$$

It follows that if $\Re\mathrm{e}\{s\} > \sigma$, we have $|\mathcal{L}\{f\}(s - \sigma')| \leq c$. Hence, if $\Re\mathrm{e}\{s\} > \sigma$,

$$|\mathcal{L}\{T\}(s)| \leq c\,|s - \sigma'|^n \leq c\,(|s| - |\sigma'|)^n = c\,\sum_{j=0}^{n} \binom{n}{j} |\sigma'|^{n-j}\,|s|^j,$$

and, since $|s|^j \leq 1 + |s|^n \leq (1 + |s|)^n$, we finally obtain

$$|\mathcal{L}\{T\}(s)| \leq c\,|s - \sigma'|^n \leq \underbrace{c\left[\sum_{j=0}^{n} \binom{n}{j} |\sigma'|^{n-j}\right]}_{A} (1 + |s|)^n,$$

for $\Re\mathrm{e}\{s\} > \sigma$, as required.

The difference between what happens for regular functions and the statement of Theorem 8.14 is the fact that for regular functions we can take $m = 0$. In the case of distributions, the Laplace transform is no longer bounded, but its growth is at most polynomial. Hence, the Laplace transform of a distribution is not controllable by a constant, but it has at most polynomial growth over any half-plane that excludes the abscissa of convergence.

We now extend another result that we have established for regular functions.

Theorem 8.15 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $c > 0$. Then, $T(c\,t)$ is $\mathcal{L}$-transformable and, for all $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > c\,\sigma(T)$,*

$$\mathcal{L}\{T(c\,\cdot)\}(s) = \frac{1}{c}\,\mathcal{L}\{T\}\left(\frac{s}{c}\right).$$

Proof First, we show that $T(c\cdot)$ is $\mathcal{L}$-transformable. We begin by showing that $\operatorname{supp}(T(c\cdot)) \subseteq [0, +\infty)$. Let $t_0 < 0$. We show that $t_0 \notin \operatorname{supp}(T(c\,\cdot))$. To this end, we show that there is a neighborhood $(t_0 - r, t_0 + r)$ of t_0 with $t_0 + r < 0$ such that, taking $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$ satisfying $\operatorname{supp}\varphi \subseteq (t_0 - r, t_0 + r)$, we have $\langle T(c\,t), \varphi(t)\rangle = 0$. Using the composition with diffeomorphisms we find

$$\langle T(c\,t), \varphi(t)\rangle = \left\langle T(\tau), \varphi(c^{-1}\,\tau)\,\frac{1}{c}\right\rangle = c^{-1}\,\langle T(\tau), \varphi(c^{-1}\,\tau)\rangle.$$

If $\varphi(c^{-1}\,\tau) \neq 0$, then $t_0 - r \leq c^{-1}\,\tau \leq t_0 + r$, so that $c\,(t_0 - r) \leq \tau \leq c\,(t_0 + r)$. It follows that $\operatorname{supp}(\varphi(c^{-1}\,\cdot)) \subseteq [c\,(t_0 - r), c\,(t_0 + r)]$. Since $c\,(t_0 + r) < 0$, the interval $[c\,(t_0 - r), c\,(t_0 + r)]$ has empty intersection with $\operatorname{supp} T$, which is contained in $[0, +\infty)$. It follows that $\langle T(c\,t), \varphi(t)\rangle = 0$, so that $\operatorname{supp}(T(c\dot{)}) \subseteq [0, +\infty)$.

We now check the second condition: we show that there exists $\sigma \in \mathbb{R}$ such that $e^{-\sigma\,t}\,T(c\,t) \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. We observe that

$$e^{-\sigma\,t}\,T(c\,t) = e^{-\frac{\sigma}{c}(c\,t)}\,T(c\,t) = \left(e^{-\frac{\sigma}{c}\cdot}\,T(\cdot)\right) \circ \Phi(t),$$

where $\Phi(t) = c\,t$ is a diffeomorphism (even more, it is a linear isomorphism), and therefore if $\frac{\sigma}{c} > \sigma(T)$, i.e., if $\sigma > c\,\sigma(T)$, we have $e^{-\sigma\,t}\,T(c\,t) \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. It follows that $T(c\,t)$ is $\mathcal{L}$-transformable if $\sigma > c\,\sigma(T)$. Take $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > c\,\sigma(T)$. Then, if $\sigma \in \mathbb{R}$ satisfies $\sigma(T) < \sigma < \frac{\Re\mathrm{e}\{s\}}{c}$, i.e., $c\,\sigma(T) < c\,\sigma < \Re\mathrm{e}\{s\}$, we have

$$\begin{aligned}
\mathcal{L}\{T(c\,\cdot)\}(s) &= \langle e^{-\sigma\,t}\,T(c\,t), e^{-(s-\sigma)\,t}\,\psi(t)\rangle \\
&= \langle e^{-\frac{\sigma}{c}\,c\,t}\,T(c\,t), e^{-(s-\sigma)\,t}\,\psi(t)\rangle \\
&= \left\langle e^{-\frac{\sigma}{c}\,\tau}\,T(\tau), e^{-\frac{s-\sigma}{c}\,\tau}\,\psi\left(\frac{\tau}{c}\right)\right\rangle \frac{1}{c},
\end{aligned}$$

where we applied the theorem of composition with a diffeomorphism ($c\,t = \tau$). Since $\psi\left(\frac{\tau}{c}\right)$ is equal to 1 if $\frac{\tau}{c} > -\delta_1$, i.e., if $\tau > -c\,\delta_1$, and it is equal to 0 if $\frac{\tau}{c} < -\delta_2$, i.e., if $\tau < -c\,\delta_2$, we obtain

$$\mathcal{L}\{T(c\,\cdot)\}(s) = \frac{1}{c}\,\mathcal{L}\{T\}(s).$$

Theorem 8.16 *Let $T, S \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Then, $T * S$ is $\mathcal{L}$-transformable, and for all $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \max\{\sigma(T), \sigma(S)\}$, we have*

$$\mathcal{L}\{T * S\}(s) = \mathcal{L}\{T\}(s)\,\mathcal{L}\{S\}(s).$$

Proof Since the support of the distributions T and S is contained in $[0, +\infty)$, they are certainly convolvable, and the support of $T * S$ is also contained in $[0, +\infty)$. We need to find $\sigma \in \mathbb{R}$ for which the distribution $e^{-\sigma t}(T * S)$ is tempered. Let $\sigma > \max\{\sigma(T), \sigma(S)\}$. We show that

$$e^{-\sigma t}(T * S) = (e^{-\sigma t} T) * (e^{-\sigma t} S).$$

In fact, consider a sequence $(\eta_k)_{k\in\mathbb{N}}$ of functions of $\mathcal{D}(\mathbb{R}^2;\mathbb{C})$ converging to unity, and let $w \in \mathcal{D}(\mathbb{R};\mathbb{C})$ be a function. We have

$$\begin{aligned}
\langle e^{-\sigma t}(T * S), w\rangle &= \langle T * S, e^{-\sigma t}\, w(t)\rangle \\
&= \lim_{k\to\infty} \langle T(t), \langle S(\tau), \eta_k(t,\tau)\, e^{-\sigma(t+\tau)}\, w(t+\tau)\rangle\rangle \\
&= \lim_{k\to\infty} \langle e^{-\sigma t} T(t), \langle e^{-\sigma\tau} S(\tau), \eta_k(t,\tau)\, w(t+\tau)\rangle\rangle \\
&= \langle (e^{-\sigma t} T) * (e^{-\sigma t} S), w\rangle,
\end{aligned}$$

since the distributions $e^{-\sigma t} T$ and $e^{-\sigma t} S$ are both tempered for this choice of σ. From Schwartz Representation Theorem (Theorem 8.13), there exist two continuous functions f and g of slow growth defined over $\mathbb{R}$ and with support in $[0, +\infty)$ and $n, m \in \mathbb{N}$ such that

$$e^{-\sigma t} T = f^{(m)} \quad \text{and} \quad e^{-\sigma t} S = g^{(n)}.$$

We can express the fact that f and g have slow growth by saying that there exist $A, B \geq 0$ and $\alpha, \beta \in \mathbb{N}$ such that

$$|f(t)| \leq A\,(1 + |t|)^{\alpha} \quad \text{and} \quad |g(t)| \leq B\,(1 + |t|)^{\beta}. \tag{8.15}$$

It follows that

$$e^{-\sigma t}(T * S) = (e^{-\sigma t} T) * (e^{-\sigma t} S) = f^{(m)} * g^{(n)} = (f * g)^{(m+n)},$$

where the function $t \mapsto (f * g)(t)$ is well defined, and in particular

$$(f * g)(t) = \begin{cases} \displaystyle\int_0^t f(t-\tau)\, g(\tau)\, \mathrm{d}\tau & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

If $t \geq 0$, we have in view of (8.15)

$$|(T * S)(t)| \leq \int_0^t |f(t-\tau)|\,|g(\tau)|\,\mathrm{d}\tau \leq A\,B \int_0^t (1 + |t-\tau|)^{\alpha}(1+\tau)^{\beta}\,\mathrm{d}\tau,$$

where $1 + |t-\tau| \leq 1 + |t| + |\tau| = 1 + t + \tau \leq (1+t)(1+\tau)$, so that

$$(1 + |t-\tau|)^{\alpha} \leq (1+t)^{\alpha}\,(1+\tau)^{\beta}.$$

We obtain

$$|(T * S)(t)| \leq A\,B \int_0^t (1+t)^{\alpha}\,(1+\tau)^{\alpha+\beta}\,\mathrm{d}\tau = A\,B\,(1+t)^{\alpha} \int_0^t (1+\tau)^{\alpha+\beta}\,\mathrm{d}\tau \leq A\,B\,C\,(1+t)^{2\,\alpha+\beta+1}.$$

The distribution of function type $T * S$ is tempered because the function

$$t \mapsto \frac{(T * S)(t)}{(1 + |t|)^p}$$

is in $L^1(\mathbb{R}; \mathbb{C})$ if $p > 2\alpha + \beta + 2$. We deduce that the distirbution $e^{-\sigma t}(T * S)$ is also tempered, with $\sigma > \max\{\sigma(T), \sigma(S)\}$, and therefore $T * S$ is $\mathcal{L}$-transformable. Let $s \in \mathbb{C}$ and $\sigma \in \mathbb{R}$ such that $\max\{\sigma(T), \sigma(S)\} < \sigma < \Re\mathrm{e}\{s\}$. We have

$$\begin{aligned}
\mathcal{L}\{T * S\}(s) &= \mathcal{L}\left\{e^{\sigma t}\left(e^{-\sigma t}(T * S)\right)\right\}(s) \\
&= \mathcal{L}\left\{e^{-\sigma t}(T * S)\right\}(s - \sigma) \\
&= \mathcal{L}\{(f * g)^{(m+n)}\}(s - \sigma) \\
&= (s - \sigma)^{m+n}\,\mathcal{L}\{f * g\}(s - \sigma) = (s - \sigma)^{m+n}\mathcal{L}\{f\}(s - \sigma)\,\mathcal{L}\{g\}(s - \sigma).
\end{aligned}$$

We also have

$$\begin{aligned}
\mathcal{L}\{T\}(s) &= \mathcal{L}\left\{e^{\sigma t}\left(e^{-\sigma t}T\right)\right\}(s) = \mathcal{L}\{e^{-\sigma t}\,T\}(s - \sigma) = \mathcal{L}\{f^{(m)}\}(s - \sigma) = (s - \sigma)^m\,\mathcal{L}\{f\}(s - \sigma) \\
\mathcal{L}\{S\}(s) &= \mathcal{L}\left\{e^{\sigma t}\left(e^{-\sigma t}S\right)\right\}(s) = \mathcal{L}\{e^{-\sigma t}\,S\}(s - \sigma) = \mathcal{L}\{g^{(n)}\}(s - \sigma) = (s - \sigma)^n\,\mathcal{L}\{g\}(s - \sigma).
\end{aligned}$$

We obtain $\mathcal{L}\{T * S\}(s) = \mathcal{L}\{T\}(s)\,\mathcal{L}\{S\}(s)$, as required.

Theorem 8.17 *Let $T \in \mathcal{D}'(\mathbb{R}; \mathbb{C})$ be $\mathcal{L}$-transformable. Let $t_0 \geq 0$. Then, $T(t - t_0)$ is $\mathcal{L}$-transformable, and for all $s \in \mathbb{C}$ satisfying $\Re\mathrm{e}\{s\} > \sigma(T)$, we have*

$$\mathcal{L}\{T(\cdot - t_0)\}(s) = e^{-t_0 s}\,\mathcal{L}\{T\}(s).$$

Proof Since $\operatorname{supp} T \subseteq [0, +\infty)$, we have $\operatorname{supp}(T(\cdot - t_0)) \subseteq [t_0, +\infty) \subseteq [0, +\infty)$. We now verify that there exists $\sigma \in \mathbb{R}$ such that the distribution $e^{-\sigma t}\,T(\cdot - t_0)$ is tempered. We have

$$e^{-\sigma t}\,T(t - t_0) = e^{-\sigma t_0}\,e^{-\sigma(t - t_0)}\,T(t - t_0) = e^{-\sigma t_0}\left(e^{-\sigma \cdot}\,T(\cdot)\right) \circ \Phi(t),$$

where Φ is the diffeomorphism $\Phi(t) = t - t_0$. Therefore, if $\sigma > \sigma(T)$, the distribution $e^{-\sigma t}\,T(t - t_0)$ is tempered, because it is the composition of a tempered distribution with a translation. Let $s \in \mathbb{C}$ and $\sigma \in \mathbb{R}$ be such that $\sigma(T) < \sigma < \Re\mathrm{e}\{s\}$. We have

$$\begin{aligned}
\mathcal{L}\{T(\cdot - t_0)\}(s) &= \langle e^{-\sigma t}\,T(t - t_0), e^{-(s-\sigma)t}\,\psi(t)\rangle \\
&= e^{-\sigma t_0}\,\langle (e^{-\sigma(t - t_0)}\,T(t - t_0), e^{-(s-\sigma)t}\,\psi(t)\rangle \\
&= e^{-\sigma t_0}\,\langle (e^{-\sigma \tau}\,T(\tau), e^{-(s-\sigma)(t_0+\tau)}\,\psi(t_0 + \tau)\rangle \\
&= e^{-\sigma t_0}\,e^{(\sigma - s)t_0}\,\langle (e^{-\sigma \tau}\,T(\tau), e^{-(s-\sigma)\tau}\,\psi(t_0 + \tau)\rangle \\
&= e^{-s t_0}\,\langle (e^{-\sigma \tau}\,T(\tau), e^{-(s-\sigma)\tau}\,\psi(t_0 + \tau)\rangle = e^{-s t_0}\,\mathcal{L}\{T\}(s),
\end{aligned}$$

where we used the theorem of the composition with a diffeomorphism (that corresponds to the change of variable $t - t_0 = \tau$) and where the last equality follows from the fact that $\psi(t_0 + \tau)$ has the same properties of ψ: indeed, $\psi(t_0 + \tau) = 1$ if $\tau \geq -\delta_1 - t_0$, i.e., if $t_0 + \tau \geq -\delta_1$, and it is 0 if $t_0 + \tau \leq -\delta_2$, i.e., $\tau \leq -\delta_2 - t_0$.

We now consider the problem of the inversion of the Laplace transform. An important result that will be needed states that if the Laplace transform of a distribution is zero on a straight vertical line to the right of the abscissa of convergence, the distribution is zero.

Theorem 8.18 *Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be $\mathcal{L}$-transformable. Let $\xi > \sigma(T)$ be such that $\mathcal{L}\{T\}(\xi + i\,\tau) = 0$ for all $\tau \in \mathbb{R}$. Then, $T = 0$.*

Proof We know that $\mathcal{L}\{T\}$ is holomorphic to the right of $\sigma(T)$. We also know that if we have a function which is holomorphic over an arc connected set that becomes zero over a set with an accumulation point, then the function is identically zero. It follows that $\mathcal{L}\{T\}(s) = 0$ for all $s \in \mathbb{C}$ with $\Re\mathrm{e}\{s\} > \sigma(T)$ Let $\sigma > \sigma(T)$. Then, $e^{-\sigma t}\,T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Using Schwartz Representation Theorem (Theorem 8.13), we write $e^{-\sigma t}\,T = f^{(m)}$, where f is a continuous function of slow growth and $f(t) = 0$ for $t < 0$. Hence, $T = e^{\sigma t}\,f^{(m)}$. If $\Re\mathrm{e}\{s\} > \sigma$, we have

$$0 = \mathcal{L}\{T\}(s) = \mathcal{L}\{e^{\sigma t}\,f^{(m)}\}(s) = \mathcal{L}\{f^{(m)}\}(s-\sigma) = (s-\sigma)^m\,\mathcal{L}\{f\}(s-\sigma).$$

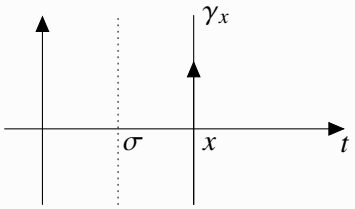
It follows that $\mathcal{L}\{T\}(s-\sigma) = 0$ if $\Re\mathrm{e}\{s\} > \sigma$. In other words, $\mathcal{L}\{f\}$ is zero over $\{z \in \mathbb{C}\,|\,\Re\mathrm{e}\{s\} > 0\}$. For a function, we know that the result is true, so that $f(t) = 0$ for all $t \in \mathbb{R}$, and therefore $f^{(m)} = 0$, which implies $T = e^{\sigma t}\,f^{(m)} = 0$.

Theorem 8.19 [Injectivity of the Laplace Transform]
Let $\sigma \in \mathbb{R}$. Let $f : \{s \in \mathbb{C}\,|\,\Re\mathrm{e}\{s\} > \sigma\} \longrightarrow \mathbb{C}$ be holomorphic and such that there exists $m \in \mathbb{R}$ and $c \geq 0$ satisfying

$$|f(s)| \leq c\,(1+|s|)^m$$

for all $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma$. Then, there exists one and only one $\mathcal{L}$-transformable distribution $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ such that $\sigma(T) \leq \sigma$ and $\mathcal{L}\{T\}(s) = f(s)$ for all $s \in \mathbb{C}$ such that $\Re\mathrm{e}\{s\} > \sigma$. Moreover, if $m < -1$, then T is a continuous function over $\mathbb{R}$.

Proof Let us start from the case $m < -1$. Let $x > \sigma$ and consider the straight line $\gamma_x = \{x + i\,\tau\,|\,\tau \in \mathbb{R}\}$.



For $t \in \mathbb{R}$, define

$$T(t) = \frac{1}{2\pi i}\int_{\gamma_x} e^{s\,t}\,f(s)\,\mathrm{d}s = \frac{1}{2\pi i}\int_{\mathbb{R}} e^{(x+i\tau)\,t}\,f(x+i\,\tau)\,i\,\mathrm{d}\tau = \frac{1}{2\pi}\int_{\mathbb{R}} e^{(x+i\tau)\,t}\,f(x+i\,\tau)\,\mathrm{d}\tau.$$

We start by showing that the integral is defined:

$$|e^{(x+i\tau)\,t}\,f(x+i\,\tau)| = e^{t\,x}\,|f(x+i\,\tau)| \leq C\,e^{t\,x}\,(1+|x+i\,\tau|)^m.$$

We have $|x+i\,\tau| \geq |\tau|$, and therefore, since m is negative,

$$(1+|x+i\,\tau|)^m \leq (1+|\tau|)^m.$$

It follows that $|e^{(x+i\tau)\,t}\,f(x+i\,\tau)| \leq C\,e^{t\,x}\,(1+|\tau|)^m$. Here, $C\,e^{t\,x}$ is a constant that does not depend on the integration variable, and since $m < -1$ the factor $(1+|\tau|)^m$ is summable. Hence, the integral exists. We now

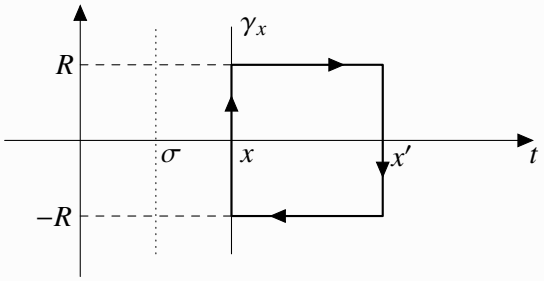
show that the integral does not depend on x. Let x and x', and let $\sigma < x < x'$. We show that

$$\frac{1}{2\pi i}\int_{\gamma_x} e^{st} f(s)\,\mathrm{d}s = \frac{1}{2\pi i}\int_{\gamma_{x'}} e^{st} f(s)\,\mathrm{d}s.$$

Let $R > 0$, and consider the contour

$$\Gamma_R = [x - iR, x + iR] + [x + iR, x' + iR, x' + iR] + [x' - iR, x' + iR, x' + iR] - [x - iR, x' - iR],$$

represented below:



Since there are no singularities, Cauchy's theorem yields

$$\oint_{\Gamma_R} e^{st} f(s)\,\mathrm{d}s = 0.$$

We now write the integral as

$$\begin{aligned} 0 = &\int_{-R}^{R} e^{t(x+i\tau)} f(x + i\tau)\, i\,\mathrm{d}\tau - \int_{-R}^{R} e^{t(x'+i\tau)} f(x' + i\tau)\, i\,\mathrm{d}\tau \\ &\int_{x}^{x'} e^{t(\xi+iR)} f(\xi + iR)\,\mathrm{d}\xi - \int_{x}^{x'} e^{t(\xi-iR)} f(\xi - iR)\,\mathrm{d}\xi. \end{aligned}$$

We now let R go to $+\infty$. The first integral tends to the integral over γ_x, and the second to the integral over $\gamma_{x'}$:

$$\begin{aligned} \lim_{R\to+\infty}\int_{-R}^{R} e^{t(x+i\tau)} f(x + i\tau)\, i\,\mathrm{d}\tau &= \int_{\gamma_x} e^{ts} f(s)\,\mathrm{d}s \\ \lim_{R\to+\infty}\int_{-R}^{R} e^{t(x'+i\tau)} f(x' + i\tau)\, i\,\mathrm{d}\tau &= \int_{\gamma_{x'}} e^{ts} f(s)\,\mathrm{d}s. \end{aligned}$$

We now show that the other two integrals tend to zero as $R \to +\infty$: we find

$$\begin{aligned} \left|\int_{x}^{x'} e^{t(\xi+iR)} f(\xi + iR)\,\mathrm{d}\xi\right| &\le \int_{x}^{x'} e^{t\xi}\,|f(\xi + iR)|\,\mathrm{d}\xi \\ &\le C\int_{x}^{x'} e^{t\xi}\,(1 + |\xi + iR|)^m\,\mathrm{d}\xi \le C\,(1 + R)^m \int_{x}^{x'} e^{t\xi}\,\mathrm{d}\xi \xrightarrow{R\to+\infty} 0. \end{aligned}$$

The remaining integral can be handled similarly. This proves that the definition of T does not depend on x. We now show that T is continuous over $\mathbb{R}$. Let $t_0 \in \mathbb{R}$. Consider a sequence $(t_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty} t_k = t_0$, and we show that $\lim_{k\to\infty} T(t_k) = T(t_0)$. We have

$$T(t_k) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{(x+i\tau)t_k} f(x + i\tau)\,\mathrm{d}\tau,$$

where $e^{(x+i\tau)t_k} f(x+i\tau)$ is pointwise convergent, as $k \to +\infty$, to $e^{(x+i\tau)t_0} f(x+i\tau)$. Moreover,

$$\left|e^{(x+i\tau)t_k} f(x+i\tau)\right| \leq e^{xt_k} C (1+|\tau|)^m.$$

Since the sequence $(t_k)_{k\in\mathbb{N}}$ admits limit, it is also bounded, so that $x t_k < d$ for some $d \in \mathbb{R}$, and for all $k \in \mathbb{N}$, so that

$$\left|e^{(x+i\tau)t_k} f(x+i\tau)\right| \leq C e^d (1+|\tau|)^m,$$

and the function in the right hand-side is summable over $\mathbb{R}$ because $m < -1$. From the dominated convergence theorem, we can bring the limit inside the integral, thus obtaining that T is continuous.

We now show that $T(t) = 0$ if $t < 0$. From what we have shown above,

$$T(t) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(x+i\tau)t} f(x+i\tau)\, \mathrm{d}\tau$$

for all $x > \sigma$. We need to show that this quantity goes to 0 as $x \to +\infty$ if $t < 0$:

$$|e^{(x+i\tau)t} f(x+i\tau)| \leq e^{tx} C (1+|\tau|)^m,$$

so that, if $t < 0$,

$$|T(t)| \leq \frac{C}{2\pi} e^{tx} \int_{\mathbb{R}} (1+|\tau|)^m \,\mathrm{d}\tau \xrightarrow{x\to+\infty} 0$$

Hence, T is continuous over $\mathbb{R}$ and it is zero for $t < 0$. We show that T is $\mathcal{L}$-transformable as a function. To this end, it suffices to show that T has exponential growth. In particular, we show that if $x > \sigma$, then

$$|T(t)| \leq M(x)\, e^{x(t)}$$

for some M. Indeed,

$$|T(t)| = \underbrace{\frac{C}{2\pi} \int_{\mathbb{R}} |1+|\tau|)^m \,\mathrm{d}\tau}_{=M(x)}\ e^{tx}.$$

Notice that $M(x)$ does not depend on x.

We now show that, if $\mu > \sigma$, we have $e^{-\sigma t} T \in L^1(\mathbb{R};\mathbb{C})$. Let us take x satisfying $\sigma < x < \mu$. Then

$$|e^{-\mu t}| \leq \begin{cases} 0 & \text{if } t < 0 \\ M(x)\, e^{(x-\mu)t} & \text{if } t \geq 0. \end{cases}$$

It follows that if we take $\mu > \sigma$, then $e^{-\sigma t} T$ is summable. Therefore, T is $\mathcal{L}$-transformable, and $\sigma(T) \leq \sigma$. Then, the Laplace transform of T is defined for $s \in \mathbb{C}$ satisfying $\mathfrak{Re}\{s\} > \sigma$.

We now need to prove that the Laplace transform of T is f for $s \in \mathbb{C}$ satisfying $\mathfrak{Re}\{s\} > \sigma$, i.e., that

$$\forall\, s \in \mathbb{C}, \qquad \mathfrak{Re}\{s\} > \sigma \quad \Rightarrow \quad \mathcal{L}\{T\}(s) = f(s).$$

By definition of T, we have

$$\begin{aligned} \mathcal{L}\{T\}(s) &= \int_0^{+\infty} e^{-ts}\, T(t)\, \mathrm{d}t \\ &= \frac{1}{2\pi i} \int_0^{+\infty} e^{-ts} \left(\int_{\gamma_x} e^{iv} f(v)\, \mathrm{d}v \right) \mathrm{d}v \\ &= \frac{1}{2\pi} \int_0^{+\infty} e^{-ts} \left(\int_{\mathbb{R}} e^{t(x+i\tau)} f(x+i\tau)\, \mathrm{d}\tau \right) \mathrm{d}t, \end{aligned}$$

where we have chosen x, which is arbitrary, to satisfy $\sigma < x < \Re\mathrm{e}\{s\}$. We need to verify that we are allowed to change the order of integration. To the end of using Fubini's theorem, we show that the function of two variables

$$(t, \tau) \mapsto e^{-t\,s} e^{t\,(x+i\,\tau)}\, f(x + i\,\tau)$$

is summable over $[0, +\infty) \times \mathbb{R}$. We have

$$|e^{-t\,s} e^{t\,(x+i\,\tau)}\, f(x + i\,\tau)| \leq e^{t\,(x-\Re\mathrm{e}\{s\})}\, C\,(1 + |\tau|)^m,$$

and, since $x - \Re\mathrm{e}\{s\} < 0$, the right hand-side is integrable over $[0, +\infty) \times \mathbb{R}$. We can therefore invoke Fubini's theorem, as we obtain

$$\mathcal{L}\{T\}(s) \;=\; \frac{1}{2\,\pi} \int_{\mathbb{R}} \left(\int_0^{+\infty} e^{-t\,s+t\,(x+i\,\tau)}\, \mathrm{d}t \right) f(x + i\,\tau)\, \mathrm{d}\tau.$$

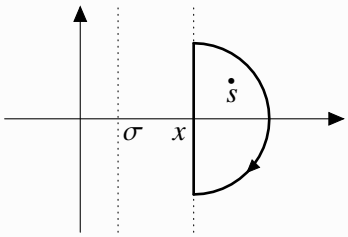
We now focus on the inner integral:

$$\int_0^{+\infty} e^{-t\,s+t\,(x+i\,\tau)}\, \mathrm{d}t \;=\; \left[\frac{e^{-t\,(s-x-i\,\tau)}}{-(s - x - i\,\tau)} \right]_{t=0}^{t=+\infty} \;=\; \frac{1}{s - x - i\,\tau}.$$

Hence,

$$\mathcal{L}\{T\}(s) \;=\; \frac{1}{2\,\pi} \int_{\mathbb{R}} \frac{f(x + i\,\tau)}{s - x - i\,\tau}\, \mathrm{d}\tau.$$

With $x < \Re\mathrm{e}\{s\}$, we consider the following contour Γ_R:



The contour Γ_R is defined as

$$\Gamma_R = [s - i\,R, s + i\,R] - \Gamma'_R,$$

where Γ'_R is parameterized by

$$\begin{aligned} \gamma : \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] &\longrightarrow \mathbb{C} \\ \theta &\mapsto x + R\,e^{i\,\theta}. \end{aligned}$$

We can apply the Residue theorem, taking into account the clockwise orientation of the contour, and that

$$\int_{\Gamma_R} \frac{f(w)}{s - w}\, \mathrm{d}w = -\frac{1}{2\,\pi\,i}\, \mathrm{Res}\left\{ \frac{f(w)}{s - w}; w = s \right\} = \lim_{w \to s} \frac{(w - s)\, f(w)}{s - w} = -f(s),$$

because the function $w \mapsto \frac{f(w)}{s-w}$ has a simple pole at s (remember that f is holomorphic). We find

$$\int_{\Gamma_R} \frac{f(w)}{s - w}\, \mathrm{d}w = 2\,\pi\, f(z) = \int_{-R}^{R} \frac{f(x + i\,\tau)}{s - x - i\,\tau}\, i\, \mathrm{d}\tau - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{f(x + R\,e^{i\,\theta})}{s - x - R\,e^{i\,\theta}}\, R\,i\,e^{i\,\theta}\, \mathrm{d}\theta \tag{8.16}$$

We now let R go to $+\infty$. We have

$$\lim_{R\to+\infty}\int_{-R}^{R}\frac{f(x+i\,\tau)}{s-x-i\,\tau}\,i\,\mathrm{d}\tau = 2\,\pi\,i\,\mathcal{L}\{T\}(s).$$

We show that the second integral in the right hand-side of (8.16) goes to 0:

$$\left|\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{f(x+R\,e^{i\,\theta})}{s-x-R\,e^{i\,\theta}}\,R\,i\,e^{i\,\theta}\,\mathrm{d}\theta\right| \leq R\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{|f(x+R\,e^{i\,\theta})|}{|s-x-R\,e^{i\,\theta}|}\,\mathrm{d}\theta.$$

Let us consider the denominator:

$$|s-x-R\,e^{i\,\theta}| = |(R\,e^{i\,\theta})-(x-s)| \geq |R\,e^{i\,\theta}|-|x-s| = R-|x-s|,$$

which is greater than 0 for large values of R. Let us consider the numerator:

$$|f(x+R\,e^{i\,\theta})| \leq C\,(1+|x+R\,e^{i\,\theta}|)^m,$$

and

$$|x+R\,e^{i\,\theta}| = |R\,e^{i\,\theta}-(-x)| \geq |R\,e^{i\,\theta}|-|-x| = R-|x|,$$

which is also positive for large values of R. It follows that

$$1+|x+R\,e^{i\,\theta}| \geq 1+R-|x| \quad\Rightarrow\quad |f(x+R\,e^{i\,\theta})| \leq C\,(1+R-|x|)^m.$$

Thus,

$$\left|\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{f(x+R\,e^{i\,\theta})}{s-x-R\,e^{i\,\theta}}\,R\,i\,e^{i\,\theta}\,\mathrm{d}\theta\right| \leq R\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{C\,(1+R-|s|)^m}{R-|x-s|}\,\mathrm{d}\theta = \pi\,R\,\frac{C\,(1+R-|s|)^m}{R-|x-s|}\xrightarrow{R\to+\infty}0.$$

We conclude that $\mathcal{L}\{T\}(s) = f(s)$. We recall that we have assumed that $m < -1$, and therefore we have exploited the fact that

$$|f(s)| \leq C\,(1+|s|)^m.$$

We now consider the case $m \geq -1$. Let $\sigma' < \sigma$. Let $p \in \mathbb{N}$ be such that $m-p < -1$, which is always possible by taking a sufficiently large value of p. Let

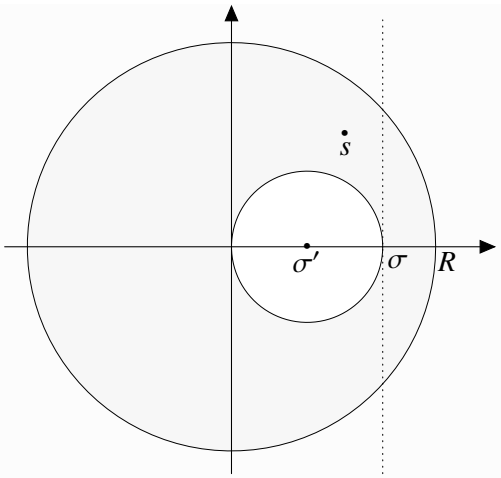
$$g(s) = (s-\sigma')^{-p}\,f(s).$$

We have

$$h(s) = \frac{|s-\sigma'|}{1+|s|} \geq \frac{|s|-|\sigma'|}{1+|s|}\xrightarrow{|s|\to+\infty}1.$$

It follows that if $|s|$ is sufficiently large (i.e., if $|s| \geq R$), we have $h(s) \geq \frac{1}{2}$. Consider the compact set

$$K = \{s \in \mathbb{C}\,|\,|s-\sigma'| \geq \sigma-\sigma' \text{ and } |s| \leq R\}.$$



Since h is continuous and is never zero over K, using Weierstrass extreme value theorem we find

$$h(s) \geq \mu > 0 \quad \forall\, s \in K.$$

Hence, if $|s-\sigma'| \geq |\sigma-\sigma'|$, then $h(s) \geq \min\left\{\mu, \frac{1}{2}\right\}$. Let us denote by c the value of this minimum. We have $c > 0$. Since

$$h(s) = \frac{|s-\sigma'|}{1+|s|} \geq c,$$

we obtain $|s-\sigma'| \geq c\,(1+|s|)$ if $|s-\sigma'| \geq \sigma-\sigma'$. Therefore, if $\Re\mathrm{e}\{s\} > \sigma$, we obtain

$$|s-\sigma'|^{-p} \leq c^{-p}\,(1+|s|)^{-p}$$

and

$$|g(s)| \leq c^{-p}\,(1+|s|)^{m-p}.$$

In this way, we are back to the previous case: g is the Laplace transform of a certain continuous function, i.e., $g(s) = \mathcal{L}\{T\}(s)$, where T is continuous over $\mathbb{R}$. We take the derivative:

$$\mathcal{L}\{\partial_t T - \sigma' T\}(s) = (s-\sigma')\,\mathcal{L}\{T\}(s).$$

We apply the operator $\partial_t - \sigma'$ for p times to obtain

$$\begin{aligned}\mathcal{L}\{(\partial_t-\sigma')^p T\}(s) &= (s-\sigma')^p\,\mathcal{L}\{T\}(s)\\ &= (s-\sigma')^p\,g(s) = f(s).\end{aligned}$$

To summarize the result of the previous theorem, if we have a holomorphic function $f : \{s \in \mathbb{C}\,|\,\Re\mathrm{e}\{s\} > \sigma\} \longrightarrow \mathbb{C}$ such that $|f(s)| \leq c\,(1+|s|)^m$ for some $m \in \mathbb{R}$ and $c \geq 0$, then:

- if $m < -1$,

$$\mathcal{L}^{-1}\{f\}(t) = \frac{1}{2\,\pi\,i}\int_{\gamma_x} e^{t\,s}\,f(s)\,\mathrm{d}s.$$

- if $m \geq -1$,

$$\mathcal{L}^{-1}\{f\}(t) = \left(\frac{\mathrm{d}}{\mathrm{d}t}-\sigma'\right)^p\left(\frac{1}{2\,\pi\,i}\int_{\gamma_x} e^{t\,s}\,(s-\sigma')^p\,f(s)\,\mathrm{d}s\right),$$

  with $m-p < -1$.

This is, at least in principle, a way of obtaining the inverse Laplace transform.

Example

Consider the function

$$g(s) = \frac{s^2\, e^{-2\,s}}{(s-1)^2},$$

which is holomorphic in $\mathbb{C} \setminus \{1\}$. We are looking for a half-plane where the function is holomorphic, and the natural choice is $\{s \in \mathbb{C} \mid \Re\mathfrak{e}\{s\} > 1\}$. From the result of a previous exercise, we know that

$$\frac{1}{s-1} = \mathcal{L}\{e^t\, H\}(s).$$

However, in $g(s)$ we have the square of the factor $(s-1)$. We can then recall that the Laplace transform transforms convolutions into products, so that

$$\frac{1}{(s-1)^2} = \mathcal{L}((e^t\, H) * (e^t\, H))(s).$$

For $t \geq 0$ we have

$$((e^{\cdot}\, H) * (e^{\cdot}\, H))(t) = \int_0^t e^{t-\tau}\, e^{\tau}\, \mathrm{d}\tau = t\, e^t.$$

It follows that

$$\frac{1}{(s-1)^2} = \mathcal{L}\{t\, e^t\, H\}(s).$$

We find

$$\frac{s^2}{(s-1)^2} = \mathcal{L}\left\{\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^2 (t\, e^t\, H)\right\}(s),$$

where the derivative is in the sense of distributions. To find the value of this derivative, we define the function

$$\begin{aligned} u : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\mapsto \begin{cases} t\, e^t & \text{if } t \geq 0 \\ 0 & \text{if } t < 0. \end{cases} \end{aligned}$$

Notice that there is no jump at 0. The derivative in the sense of distributions is

$$u'(t) = e^t\, (t+1)\, H(t).$$

Here, there is a jump at 0: indeed $\lim_{t \to 0^+} u'(t) = 1$ and $\lim\limits_{t \to 0^-} u'(t) = 0$. Hence, the second derivative is

$$u''(t) = e^t\, (t+2)\, H(t) + \delta(t).$$

We obtain

$$\frac{s^2}{(s-1)^2} = \mathcal{L}\left\{e^t\, (t+2)\, H(t) + \delta(t)\right\}(s).$$

Recalling that $\mathcal{L}\{f(t-t_0)\}(s) = e^{-t_0\, s}\, \mathcal{L}\{f\}(s)$, we obtain

$$g(s) = \mathcal{L}\left\{e^{t-2}\, t\, H(t-2) + \delta(t-2)\right\}(s).$$

Example

Consider the rational function

$$g(s) = \frac{s-4}{(s-4)^2+9}.$$

We recall that, for $\omega \in \mathbb{R}$, we have found

$$\mathcal{L}\{\sin(\omega\, t)\, H\}(s) = \frac{\omega}{s^2+\omega^2} \quad \text{and} \quad \mathcal{L}\{\cos(\omega\, t)\, H\}(s) = \frac{s}{s^2+\omega^2}.$$

Taking $\omega = 3$ we find

$$\frac{z}{z^2+9} = \mathcal{L}\{\cos(3\, t)\, H\}(s).$$

Using $\mathcal{L}\{e^{s_0\, t}\, f\}(s) = \mathcal{L}\{f\}(s-s_0)$, we find

$$\frac{z-4}{(z-4)^2+9} = \mathcal{L}\{e^{4\, t}\, \cos(3\, t)\, H\}(s).$$

8.2.1 Inverse Laplace transform of rational functions

We want to find general rules for the calculation of the inverse Laplace transform in a particularly interesting and useful case, that very often arises in applications: the case of rational functions. Consider

$$F(s) = \frac{N(s)}{D(s)}, \tag{8.17}$$

where $N(s)$ and $D(s)$ are polynomials with coefficients, of degree m and n, respectively:

$$F(s) = \frac{a_0 + a_1\, s + a_2\, s^2 + \ldots + a_m\, s^m}{b_0 + b_1\, s + b_2\, s^2 + \ldots + b_n\, s^n},$$

where $a_m, b_n \neq 0$. We do not lose any generality by assuming that the denominator is a monic polynomial, i.e., its leading coefficient b_n can be assumed to be equal to 1. Indeed, if this is not the case, we can always divide both the numerator and the denominator by b_n and obtain a monic polynomial in the denominator. We can also assume that the polynomials $N(s)$ and $D(s)$ are coprime (i.e., we assume to have cancelled the common factors between numerator and denominator, if present). We say that the rational function $F(s)$ is

- **proper** if the degree of the numerator is smaller or equal to the degree of the denominator: $m \leq n$;
- **strictly proper** if the degree of the numerator is strictly smaller than the degree of the denominator: $m < n$;
- **nonstrictly proper** or **biproper** if the degree of the denominator is equal to the degree of the numerator: $n = m$;
- **improper** if the degree of the numerator is strictly greater than the degree of the denominator: $m > n$.

The difference $n - m$ is called *relative degree* of $F(s)$.

If $m \geq n$ (i.e., if $F(s)$ is not strictly proper), we can perform the polynomial division: this means that there exist two polynomials $Q(s)$ (the quotient), of degree $m-n$, and $R(s)$ (the remainder) of degree $< n$, such that

$$N(s) = Q(s)\, D(s) + R(s). \tag{8.18}$$

The polynomials $Q(s)$ and $R(s)$ can be written, respectively, as

$$\begin{aligned} Q(s) &= q_0 + q_1\, s + \ldots + q_{m-n}\, s^{m-n} \\ R(s) &= r_0 + r_1\, s + \ldots + r_{n-1}\, s^{n-1}, \end{aligned}$$

where $q_{m-n} = a_m$, where the coefficients $q_0, \ldots, q_{m-n}, r_0, \ldots, r_{n-1}$ are complex. The coefficients of $Q(s)$ and $R(s)$ can be obtained by applying the principle of identity of polynomials to (8.18).

Example

Suppose that

$$F(s) = \frac{a_3\, s^3 + a_2\, s^2 + a_1\, s + a_0}{s + b_0},$$

so that the relative degree is $n - m = 1 - 3 = -2$. We want to determined $Q(s) = q_0 + q_1\, s + q_2\, s^2$ and $R(s) = r_0$ such that (8.18) holds true (the degree of $Q(s)$ is $m - n - 2$ and that of $R(s)$ is $n - 1 = 0$). We write (8.18) as

$$a_3\, s^3 + a_2\, s^2 + a_1\, s + a_0 = (q_0 + q_1\, s + q_2\, s^2)(s + b_0) + r_0,$$

i.e.,

$$a_3\, s^3 + a_2\, s^2 + a_1\, s + a_0 = q_2\, s^3 + (q_1 + q_2\, b_0)\, s^2 + (q_0 + q_1\, b_0)\, s + q_0\, b_0 + r_0.$$

Equating the coefficients that correspond to the same powers of the indeterminate s gives:

$$q_2 = a_3, \quad q_1 = a_2 - a_3\, b_0, \quad q_0 = a_1 - a_2\, b_0 + a_3\, b_0^2, \quad r_0 = a_0 - a_1\, b_0 + a_2\, b_0^2 - a_3\, b_0^3.$$

Using (8.18) into $F(s)$ yields

$$F(s) = \frac{Q(s)\, D(s) + R(s)}{D(s)} = Q(s) + \frac{R(s)}{D(s)},$$

for all $s \in \mathbb{C}$ such that $D(s) \neq 0$. The rational function $F(s)$ is analytic in the half-plane given by the condition $\Re\mathrm{e}\{s\} > \lambda$, where λ is any real number larger than the real part of all the roots of $D(s)$. We can find the inverse Laplace transform of $Q(s)$:

$$Q(s) = q_0 + q_1\, s + \ldots + q_{m-n}\, s^{m-n} \quad (\Re\mathrm{e}\{s\} > \lambda) \qquad \Rightarrow \qquad \mathcal{L}^{-1}\{Q\}(t) = q_0\, \delta + q_1\, \delta' + q_2\, \delta'' + \ldots + q_{m-n}\, \delta^{(m-n)}.$$

We are left with the calculation of the inverse Laplace transform of the strictly proper rational function $R(s)/D(s)$, which is in the same form as (8.17) (but, again, with the degree of the numerator strictly smaller than the degree of the denominator).[1]

Let $p_1, \ldots, p_\ell$ denote the distinct roots of the denominator $D(s)$, with multiplicities, respectively, $\nu_1, \ldots, \nu_\ell$, so that $\nu_1 + \ldots + \nu_\ell = n$. Since $D(s)$ is assumed to be monic , we can write $D(s) = (s - p_1)^{\nu_1}(s - p_2)^{\nu_2} \cdots (s - p_\ell)^{\nu_\ell}$. Hence, the strictly proper part $\tilde{F}(s)$ of $F(s)$ is

$$\tilde{F}(s) \stackrel{\text{def}}{=} \frac{R(s)}{D(s)} = \frac{R(s)}{(s - p_1)^{\nu_1}(s - p_2)^{\nu_2} \cdots (s - p_\ell)^{\nu_\ell}}.$$

[1] Note that $R(s)$ and $D(s)$ are coprime. If, in fact, we have $R(s) = \tilde{R}(s)\, \alpha(s)$ and $D(s) = \tilde{D}(s)\, \alpha(s)$, where $\alpha(s)$ is a polynomial with complex coefficients, then

$$Q(s) + \frac{\tilde{R}(s)\, \alpha(s)}{\tilde{D}(s)\, \alpha(s)} = \frac{Q(s)\, \tilde{D}(s)\, \alpha(s) + \tilde{R}(s)\, \alpha(s)}{\tilde{D}(s)\, \alpha(s)},$$

so that $N(s)$ and $D(s)$ also have a common factor: this is not possible because we assumed that $N(s)$ and $D(s)$ are coprime.

The rational function $\tilde{F}(s)$ is defined for all $s \in \mathbb{C}$ such that $\Re\{s\} > \max\{\Re\{p_1\}, \Re\{p_2\}, \ldots, \Re\{p_\ell\}\}$. We can decompose $\tilde{F}(s)$ using the so-called *partial fraction expansion*:

$$\tilde{F}(s) = \sum_{h=1}^{\ell} \sum_{k=1}^{\nu_h} \frac{A_{h,k}}{(s-p_h)^k} = \frac{A_{1,1}}{s-p_1} + \ldots + \frac{A_{1,\mu_1}}{(s-p_1)^{\mu_1}} + \ldots + \frac{A_{m,1}}{s-p_m} + \ldots + \frac{A_{m,\mu_m}}{(s-p_m)^{\mu_m}}, \tag{8.19}$$

where the complex numbers $A_{h,k}$ are determined using the principle of identity of polynomials. We can determine the inverse Laplace transform of (8.19):

$$\mathcal{L}^{-1}\{\tilde{F}\}(t) = \sum_{h=1}^{\ell} \sum_{k=1}^{\nu_h} A_{h,k} \frac{t^{k-1}}{(k-1)!} e^{p_h t} H(t).$$

We have proved the following important result.

Theorem 8.20 (Heaviside Theorem)
Let $N(s)$ and $D(s)$ be coprime polynomials of degrees m and n, respectively in the indeterminate $s \in \mathbb{C}$, and let $D(s)$ be monic. Let $p_1, \ldots, p_\ell$ be the distinct roots of $D(s)$ with multiplicities $\nu_1, \ldots, \nu_\ell$, respectively. Then:

(i) there exist $q_0, \ldots, q_{m-n} \in \mathbb{C}$, and $A_{h,k} \in \mathbb{C}$ for $h \in \{1, \ldots, \ell\}$ and $k \in \{1, \ldots, \nu_h\}$ such that

$$\frac{N(s)}{D(s)} = \sum_{k=0}^{m-n} q_k s^k + \sum_{h=1}^{\ell} \sum_{k=1}^{\nu_h} \frac{A_{h,k}}{(s-p_h)^k},$$

which is analytic for all $s \in \mathbb{C}$ such that $\Re\{s\} > \max_{h \in \{1,\ldots,\ell\}} \Re\{p_h\}$;

(ii) the inverse Laplace transform of $N(s)/D(s)$ is

$$\mathcal{L}^{-1}\left\{\frac{N(s)}{D(s)}\right\}(t) = \sum_{k=0}^{m-n} q_k \delta^{(k)} + \sum_{h=1}^{\ell} \sum_{k=1}^{\nu_h} A_{h,k} \frac{t^{k-1}}{(k-1)!} e^{p_h t} H(t);$$

(iii) The inverse Laplace transform $\mathcal{L}^{-1}\left\{\frac{N(s)}{D(s)}\right\}$ is of function type if and only if $\frac{N(s)}{D(s)}$ is strictly proper

Remark 8.15 Consider the contribution of each pole of $D(s)$ to the inverse Laplace transform. In other words, let us fix $h \in \{1, \ldots, \ell\}$ and define

$$g_h(t) = \sum_{k=1}^{\nu_h} A_{h,k} \frac{t^{k-1}}{(k-1)!} e^{p_h t} H(t).$$

The asymptotic behavior of $g_h(t)$ depends, loosely speaking, on the real part of the pole p_h. More precisely:

(i) if $\Re\{p_h\} < 0$, the modulus $|g_h(t)|$ vanishes exponentially as $t \to +\infty$;

(ii) if $\Re\{p_h\} > 0$, the modulus $|g_h(t)|$ grows exponentially as $t \to +\infty$;

(iii) if $\Re\{p_h\} = 0$ and $\nu_h = 1$ (so that p_h is a simple root of $D(s)$), the modulus $|g_h(t)|$ is bounded

(iv) if $\Re\{p_h\} = 0$ and $\nu_h > 1$, the modulus $|g_h(t)|$ grows polynomially as $t \to +\infty$.

From this discussion, considering that the overall inverse Laplace transform is the sum of the contributions of the inverse Laplace transforms of the terms corresponding to each pole separately, we deduce that the inverse Laplace transform $f(t)$ of $N(s)/D(s)$

(i) goes to 0 as $t \to +\infty$ if and only if all the poles have strictly negative real parts, i.e., for all $h \in \{1, \dots, \ell\}$, we have $\Re\mathrm{e}\{p_h\} < 0$;

(ii) is bounded if and only if all the poles have either strictly negative real part, or if they lie on the imaginary axis they must be simple poles;

(iii) diverges if and only if there exists at least a pole p_h with $\Re\mathrm{e}\{p_h\} > 0$, or at least a pole with $\Re\mathrm{e}\{p_h\} = 0$ with multiplicity ν_h greater than 1.

These considerations are at the basis of the stability theory of linear dynamical systems.

8.2.2 Rational functions with real coefficients

Up to this point we have considered the problem of finding the inverse Laplace transform of a rational function in $s \in \mathbb{C}$ given by the ration of polynomials with complex coefficients. We now focus our attention to the case of polynomials with real coefficients. In other words, we consider $N(s) = a_0 + a_1\, s + a_2\, s^2 + \ldots + a_m\, s^m$ and $D(s) = b_0 + b_1\, s + b_2\, s^2 + \ldots + b_{n-1}\, s^{n-1} + s^n$ (so that $D(s)$ is still monic) where $a_0, \dots, a_m, b_0, \dots, b_{n-1} \in \mathbb{R}$, and we continue to assume with no loss of generality that $N(s)$ and $D(s)$ are coprime. In this case, the roots of the denominator (and of the numerator) come in complex conjugate pairs. In other words, since the denominator is a polynomial with real coefficients, if $\lambda \in \mathbb{C}$ is a root, then its complex conjugate $\bar{\lambda}$ is also a root of $D(s)$.[2] As a result, complex roots of polynomials with real coefficients come in complex conjugate pairs.

Exploiting this property, we can assume thatthe poles of $D(s)$ are $p_1, \dots, p_\ell, \sigma_1 \pm i\,\omega_1, \dots, \sigma_r \pm i\,\omega_r$, and the corresponding multiplicities are $\nu_1, \dots, \nu_\ell, \mu_1, \dots, \mu_r$, for some $\ell, r \in \mathbb{N}$, where $p_1, \dots, p_\ell$ are the real roots, $\sigma_i, \omega_i \in \mathbb{R}$ for all $i \in \{1, \dots, r\}$ and $\omega_i \neq 0$. Obviously, $\ell + 2\,r = n$. In this case, it is always possible to determine a partial fraction expansion of $F(s)$ as

$$
\begin{aligned}
F(s) \;=\; \frac{N(s)}{D(s)} &= \sum_{k=1}^{m-n} q_k\, s^k + \sum_{h=1}^{\ell}\sum_{k=1}^{\nu_h} \frac{A_{h,k}}{(s-p_h)^k} + \sum_{h=1}^{\ell}\sum_{k=1}^{\nu_h} \frac{B_{h,k}\, s + C_{h,k}}{((s-\sigma_h)^2+\omega_h^2)^k} \\
&= \frac{A_{1,1}}{s-p_1} + \ldots + \frac{A_{1,\nu_1}}{(s-p_1)^{\nu_1}} + \ldots + \frac{A_{\ell,1}}{s-p_\ell} + \ldots + \frac{A_{\ell,\nu_\ell}}{(s-p_\ell)^{\nu_\ell}} \\
&\quad + \frac{B_{1,1}\, s + C_{1,1}}{(s-\sigma_1)^2+\omega_1^2} + \ldots + \frac{B_{1,\mu_1}\, s + C_{1,\mu_1}}{((s-\sigma_1)^2+\omega_1^2)^{\mu_1}} + \ldots + \frac{B_{r,1}\, s + C_{r,1}}{(s-\sigma_r)^2+\omega_r^2} + \ldots + \frac{B_{r,\mu_r}\, s + C_{r,\mu_r}}{((s-\sigma_r)^2+\omega_r^2)^{\mu_r}}.
\end{aligned}
$$

The problem now is to find the inverse Laplace transform of the generic term

$$
\frac{B_{h,k}\, s + C_{h,k}}{((s-\sigma_h)^2+\omega_h^2)^k}.
$$

[2] Indeed, if $\lambda \in \mathbb{C}$ is a root of a generic polynomial $P(s) = a_0 + a_1\, s + \ldots + a_n\, s^n$ with real coefficients, then $P(\lambda) = a_0 + a_1\,\lambda + a_2\,\lambda^2 + \ldots + a_n\,\lambda^n = 0$, so that

$$
P(\bar{\lambda}) = a_0 + a_1\,\bar{\lambda} + a_2\,\bar{\lambda}^2 + \ldots + a_n\,\bar{\lambda}^n = \overline{a_0 + a_1\,\lambda + a_2\,\lambda^2 + \ldots + a_n\,\lambda^n} = \overline{P(\lambda)} = \bar{0} = 0,
$$

which shows that $\bar{\lambda}$ is also a root of $P(s)$.

We can solve the problem by linearity if we can find the inverse Laplace transform of the two rational functions

$$\frac{1}{((s-\sigma_h)^2+\omega_h^2)^k} \qquad \text{and} \qquad \frac{s}{((s-\sigma_h)^2+\omega_h^2)^k}.$$

Consider first the case $k = 1$. From the expressions of the Laplace transform of the sine and cosine functions and recalling that $\mathcal{L}\{e^{s_0 t} f\} = \mathcal{L}\{f\}(s - s_0)$ we have

$$\frac{1}{(s-\sigma_h)^2+\omega_h^2} = \mathcal{L}\left\{\frac{1}{\omega_h} e^{\sigma_h t} \sin(\omega_h t) H\right\}(s)$$

$$\frac{s}{(s-\sigma_h)^2+\omega_h^2} = \frac{s-\sigma_h}{(s-\sigma_h)^2+\omega_h^2} + \frac{\sigma_h}{(s-\sigma_h)^2+\omega_h^2} = \mathcal{L}\left\{e^{\sigma_h t} \cos(\omega_h t) H + \frac{\sigma_h}{\omega_h} e^{\sigma_h t} \sin(\omega_h t) H\right\}(s).$$

For $k > 1$, we can exploit the properties of the convolution:

$$\frac{1}{((s-\sigma_h)^2+\omega_h^2)^k} = \mathcal{L}\Big\{\underbrace{\Big(\frac{1}{\omega_h} e^{\sigma_h t} \sin(\omega_h t) H\Big) * \cdots * \Big(\frac{1}{\omega_h} e^{\sigma_h t} \sin(\omega_h t) H\Big)}_{k \text{ times}}\Big\}(s)$$

$$\frac{s}{((s-\sigma_h)^2+\omega_h^2)^j} = \frac{s}{(s-\sigma_h)^2+\omega_h^2} \underbrace{\frac{1}{(s-\sigma_h)^2+\omega_h^2} \cdot \ldots \cdot \frac{1}{(s-\sigma_h)^2+\omega_h^2}}_{k-1 \text{ times}}.$$

Example We want to determine the partial fraction decomposition of the rational function

$$F(s) = \frac{2\,s+3}{7\,s-5}. \tag{8.20}$$

We immediately observe that the relative degree is 0. Therefore, we can apply the polynomial division, and we obtain the constant polynomial $q(s) = 2/7$ as the quotient and the constant polynomial $r(s) = 31/7$ as remainder. Hence,

$$F(s) = \frac{2}{7} + \frac{\frac{31}{7}}{7\,s-5} = \frac{2}{7} + \frac{31}{49}\,\frac{1}{s-\frac{5}{7}}.$$

We find, exploiting the linearity of the Laplace transform and its inverse,

$$\mathcal{L}^{-1}\{F\}(t) = \frac{2}{7}\,\delta + \mathcal{L}^{-1}\left\{\frac{31}{49}\,\frac{1}{s-\frac{5}{7}}\right\}(t) = \frac{2}{7}\,\delta + \frac{31}{49}\,e^{\frac{5}{7}\,t}\,H(t).$$

Example Consider the improper rational function

$$F(s) = \frac{s^4+3\,s^3+11\,s^2+7\,s-7}{s^2+2\,s+10}. \tag{8.21}$$

We can apply the polynomial division, and we obtain $Q(s) = s^2 + s - 1$ as the quotient and $3 - s$ as the remainder, so that

$$F(s) = s^2 + s - 1 + \frac{3-s}{s^2+2\,s+10}.$$

We find, exploiting the linearity of the Laplace transform and its inverse and the fact that the roots of the denominator are $1 \pm 3\,i$

$$\begin{aligned}
\mathcal{L}^{-1}\{F\}(t) &= \delta^{(2)} + \delta^{(1)} - \delta + \mathcal{L}^{-1}\left\{\frac{3-s}{s^2+2\,s+10}\right\}(t) \\
&= \delta^{(2)} + \delta^{(1)} - \delta + 3\,\mathcal{L}^{-1}\left\{\frac{1}{s^2+2\,s+10}\right\}(t) - \mathcal{L}^{-1}\left\{\frac{s}{s^2+2\,s+10}\right\}(t) \\
&= \delta^{(2)} + \delta^{(1)} - \delta + 3\left(\frac{1}{3}\,e^{-t}\,\sin(3\,t)\,H(t)\right) - \left(e^{-t}\,\cos(3\,t)\,H(t) + \frac{-1}{3}\,e^{-t}\,\sin(3\,t)\,H(t)\right) \\
&= \delta^{(2)} + \delta^{(1)} - \delta + \frac{4}{3}\,e^{-t}\,\sin(3\,t)\,H(t) - e^{-t}\,\cos(3\,t)\,H(t).
\end{aligned}$$

8.2.3 Residues

An alternative way to obtain the inverse Laplace transform of a rational function (with real or complex coefficients) is to exploit Cauchy's Residue theorem. We assume that the rational function is strictly proper: we have already seen how the non-strictly proper case can be brought back to the strictly proper one.

Case 1: all the roots of the denominator are simple. In other words, we can write

$$F(s) = \frac{R(s)}{(s-p_1)(s-p_2)\cdots(s-p_n)},$$

where $p_1, \ldots, p_n \in \mathbb{C}$ are distinct: these are (simple) poles of $F(s)$. Since the degree of the numerator is smaller than the degree of the denominator, we can write

$$F(s) = \frac{R_1}{s - p_1} + \frac{R_2}{s - p_2} + \ldots + \frac{R_n}{s - p_n}.$$

The complex numbers $R_1, \ldots, R_n$ are the residues of $F(s)$ at p_1, p_2,..., p_n. We recall that, for a simple pole p_i, the corresponding residue R_i is given by

$$\boxed{R_i = \lim_{s \to p_i} \left[(s - p_i)\, F(s) \right]} \tag{8.22}$$

for $i \in \{1, \ldots, n\}$. By linearity, we can write

$$\mathcal{L}^{-1}\left\{ \frac{R_1}{s - p_1} + \frac{R_2}{s - p_2} + \ldots + \frac{R_n}{s - p_n} \right\}(t) = \left(R_1\, e^{p_1 t} + R_2\, e^{p_2 t} + \ldots + R_n\, e^{p_n t} \right) H(t).$$

Example

We want to determine the partial fraction decomposition of the rational function

$$F(s) = \frac{4\,s - 2}{s^3 - s^2 - 2\,s}. \tag{8.23}$$

We compute the roots of the denominator: $s^3 - s^2 - 2\,s = s\,(s^2 - s - 2) = s\,(s + 1)(s - 2)$. Thus, we write

$$F(s) = \frac{R_1}{s} + \frac{R_2}{s + 1} + \frac{R_3}{s - 2}, \tag{8.24}$$

where R_1, R_2, R_3 are the residues to be determined:

$$\begin{aligned}
R_1 &= \lim_{s \to 0} s\, f(s) = \lim_{s \to 0} \cancel{s}\, \frac{4\,s - 2}{\cancel{s}\,(s + 1)(s - 2)} \\
&= \lim_{s \to 0} \frac{4\,s - 2}{(s + 1)(s - 2)} = 1 \\
R_2 &= \lim_{s \to -1} (s + 1)\, f(s) = \lim_{s \to 0} \cancel{(s+1)}\, \frac{4\,s - 2}{s\,\cancel{(s+1)}(s - 2)} \\
&= \lim_{s \to -1} \frac{4\,s - 2}{s\,(s - 2)} = \frac{-4 - 2}{-1 \cdot (-3)} = -2 \\
R_3 &= \lim_{s \to 2} (s - 2)\, f(s) = \lim_{s \to 0} \cancel{(s-2)}\, \frac{4\,s - 2}{s\,(s + 1)\cancel{(s-2)}} \\
&= \lim_{s \to 2} \frac{4\,s - 2}{s\,(s + 1)} = \frac{8 - 2}{2 \cdot 3} = 1.
\end{aligned}$$

It follows that

$$F(s) = \frac{1}{s} - \frac{2}{s + 1} + \frac{1}{s - 2}.$$

The inverse Laplace transform of $F(s)$ is

$$\mathcal{L}^{-1}\{F(s)\}(t) = \left(1 - 2\,e^{-t} + e^{2t}\right) H(t).$$

This result holds both in the case where the poles are real and when there are complex poles. However, in the case where the poles are not all real, some of these exponentials are not real-valued. Combining these exponentials in a suitable way, and using Euler's laws, it is possible

to obtain a real-valued function $f(t)$. However, there is a more efficient way of performing the same task, by exploiting the fact that the roots of the denominator (and of the numerator) come in complex conjugate pairs. Indeed, since the polynomials have real coefficients, if $\lambda \in \mathbb{C}$ is a root of the denominator d, then its complex conjugate $\bar{\lambda}$ is also a root of d.[3] As a result, complex roots of polynomials with real coefficients come in complex conjugate pairs. Thus, assume now that we have pairs of complex conjugate poles. For example, let

$$p_1 = \sigma + i\,\omega, \qquad p_2 = \sigma - i\,\omega,$$

with the corresponding residues R_1 and R_2 given by

$$R_1 = \lim_{s \to p_1} [(s - p_1)\,F(s)] \qquad \text{and} \qquad R_2 = \lim_{s \to p_2} [(s - p_2)\,F(s)] \tag{8.25}$$

However, in the case where p_1 and p_2 are a complex conjugate pair, the residues R_1 and R_2 are complex conjugate. To see this, let us write $F(s)$ as

$$F(s) = \frac{K(s)}{(s - p_1)(s - p_2)} = \frac{K(s)}{(s - \sigma - i\,\omega)(s - \sigma + i\,\omega)},$$

where k is also a rational function. Equation (8.25) delivers

$$R_1 = \lim_{s \to \sigma + i\,\omega} \cancel{(s - \sigma - i\,\omega)} \frac{K(s)}{\cancel{(s - \sigma - i\,\omega)}(s - \sigma + i\,\omega)} = \frac{K(\sigma + i\,\omega)}{2\,i\,\omega}$$

$$R_2 = \lim_{s \to \sigma - i\,\omega} \cancel{(s - \sigma + i\,\omega)} \frac{K(s)}{(s - \sigma - i\,\omega)\cancel{(s - \sigma + i\,\omega)}} = \frac{K(\sigma - i\,\omega)}{-2\,i\,\omega} = \frac{K(\overline{\sigma + i\,\omega})}{-2\,i\,\omega} = \frac{\overline{K(\sigma + i\,\omega)}}{-2\,i\,\omega} = \bar{R}_1,$$

where we have used the fact that, from Theorem 8.6, $K(\bar{s}) = \overline{K(s)}$.

Now, since R_1 and R_2 are complex conjugate, we can write them as

$$R_1 = u + i\,v, \qquad R_2 = u - i\,v,$$

and we obtain

$$\frac{R_1}{s - p_1} + \frac{R_2}{s - p_2} = \frac{u + i\,v}{s - \sigma - i\,\omega} + \frac{u - i\,v}{s - \sigma + i\,\omega}. \tag{8.26}$$

We write the complex number $R_1 = u + i\,v$ in polar form as $R_1 = M\,e^{i\,\varphi}$, where obviously $M = \sqrt{u^2 + v^2}$ and $\varphi \in \arg(u + i\,v)$, so that we can also write $R_2 = M\,e^{-i\,\varphi}$, and (8.26) becomes

$$\frac{R_1}{s - p_1} + \frac{R_2}{s - p_2} = M \left(\frac{e^{i\,\varphi}}{s - \sigma - i\,\omega} + \frac{e^{-i\,\varphi}}{s - \sigma + i\,\omega} \right).$$

Taking the inverse Laplace transform and using Euler's formula yields

$$M \left(e^{i\,\varphi}\,e^{(\sigma + i\,\omega)\,t} + e^{-i\,\varphi}\,e^{(\sigma - i\,\omega)\,t} \right) H(t) = 2\,M\,e^{\sigma\,t}\,\frac{e^{i\,(\omega\,t + \varphi)} + e^{-i\,(\omega\,t + \varphi)}}{2}\,H(t) = 2\,M\,e^{\sigma\,t}\,\cos(\omega\,t + \varphi)\,H(t).$$

[3] Indeed, if $\lambda \in \mathbb{C}$ is a root of $d(s) = a_0 + a_1\,s + \ldots + a_n\,s^n$, then $d(\lambda) = a_0 + a_1\,\lambda + a_2\,\lambda^2 + \ldots + a_n\,\lambda^n = 0$, so that

$$d(\bar{\lambda}) = a_0 + a_1\,\bar{\lambda} + a_2\,\bar{\lambda}^2 + \ldots + a_n\,\bar{\lambda}^n = \overline{a_0 + a_1\,\lambda + a_2\,\lambda^2 + \ldots + a_n\,\lambda^n} = \overline{d(\lambda)} = \bar{0} = 0,$$

which shows that $\bar{\lambda}$ is also a root.

Example

We want to determine the inverse Laplace transform of the rational function

$$F(s) = \frac{4\,s - 8}{s^2 + 2\,s + 5}. \tag{8.27}$$

First, we determine the poles by solving $s^2 + 2\,s + 5 = 0$. We obtain $p_{1,2} = -1 \pm 2\,i$. Thus, we need to find the residues R_1 and R_2 such that

$$F(s) = \frac{R_1}{s - (-1 + 2\,i)} + \frac{R_2}{s - (-1 - 2\,i)}.$$

We find

$$R_1 = \lim_{s \to -1+2\,i} \left[(s - (-1 + 2\,i))\, F(s) \right] = \frac{-4 + 8\,i - 8}{-1 + 2\,i + 1 + 2\,i} = \frac{-12 + 8\,i}{4\,i} = 2 + 3\,i.$$

We know that $R_2 = \bar{R}_1 = 2 - 3\,i$.

Now, we need to determine the inverse Laplace transform of

$$F(s) = \frac{2 + 3\,i}{s - (-1 + 2\,i)} + \frac{2 - 3\,i}{s - (-1 - 2\,i)}.$$

Let $M = |2 + 3\,i| = \sqrt{4 + 9} = \sqrt{13}$ and $\varphi = \operatorname{Arg}(2 + 3\,i) = \arctan \frac{3}{2}$. We find

$$\mathcal{L}^{-1}\left\{ \frac{4\,s - 8}{s^2 + 2\,s + 5} \right\}(t) = (2\,\sqrt{13}\,e^{-t}\,\cos(2\,t + \varphi))\,H(t).$$

Case 2: one or more roots of d have multiplicity > 1. We now want to generalise the previous discussion to the case where the poles are not all necessarily simple. In other words, we are considering a rational function in the form

$$F(s) = \frac{R(s)}{(s - p_1)^{\nu_1}\,(s - p_2)^{\nu_2}\,\ldots\,(s - p_\ell)^{\nu_\ell}}.$$

In this expression, we are assuming that

- the poles $p_1, \ldots, p_\ell$ are distinct, i.e., $i \neq j$ implies $p_i \neq p_j$;
- the pole p_k has multiplicity ν_k, for all $k \in \{1, \ldots, \ell\}$.

We are still assuming that $F(s)$ is a strictly proper rational function. In this case, $F(s)$ has a partial fraction decomposition in the form

$$\begin{aligned} F(s) = {} & \frac{R_{1,\nu_1}}{s - p_1} + \frac{R_{1,\nu_1 - 1}}{(s - p_1)^2} + \ldots + \frac{R_{1,1}}{(s - p_1)^{\nu_1}} + \frac{R_{2,\nu_2}}{s - p_2} + \frac{R_{2,\nu_2 - 1}}{(s - p_2)^2} + \ldots + \frac{R_{2,1}}{(s - p_2)^{\nu_2}} \\ & + \ldots + \frac{R_{\ell,\nu_\ell}}{s - p_\ell} + \frac{R_{\ell,\nu_\ell - 1}}{(s - p_\ell)^2} + \ldots + \frac{R_{\ell,1}}{(s - p_\ell)^{\nu_\ell}} = \sum_{k=1}^{\ell} \sum_{j=1}^{\nu_k} \frac{R_{k,j}}{(s - p_k)^{\nu_k - j + 1}}. \end{aligned} \tag{8.28}$$

Moreover, if $a \in \mathbb{C}$, we recall that (8.8) holds:

$$\mathcal{L}^{-1}\left\{ \frac{1}{(s - a)^n} \right\} = \frac{1}{(n - 1)!}\,t^{n-1}\,e^{a\,t}\,H(t).$$

In view of the linearity of the Laplace transform we find

$$\mathcal{L}^{-1}\{F\}(t) = \sum_{k=1}^{\ell} \sum_{j=1}^{\nu_k} \frac{R_{k,j}}{(\nu_k - j)!}\,t^{\nu_k - j}\,e^{p_k\,t}.$$

The residues in this case can be obtained by

$$\boxed{R_{k,j} = \frac{1}{(j-1)!}\lim_{s\to p_k}\left[\frac{\mathrm{d}^{j-1}}{\mathrm{d}s^{j-1}}\left((s-p_k)^{\nu_k}\,f(s)\right)\right]} \tag{8.29}$$

for $k \in \{1,\ldots,\ell\}$ and $j \in \{1,\ldots,\nu_k\}$, which generalizes (8.22).

Example

Consider the rational function

$$F(s) = \frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{s^3\,(s+1)^2} \tag{8.30}$$

in the complex variable s. This rational function is strictly proper, since the degree of the numerator is 4, and it is strictly smaller than the degree of the denominator, which is 5. We want to obtain the inverse Laplace transform of $f(s)$.

Denoting the poles of $F(s)$ as $p_1 = 0$ and $p_1 = -1$, we can write $F(s)$ in the form (8.28), i.e.,

$$F(s) = \frac{R_{1,3}}{s} + \frac{R_{1,2}}{s^2} + \frac{R_{1,1}}{s^3} + \frac{R_{2,2}}{s+1} + \frac{R_{2,1}}{(s+1)^2}. \tag{8.31}$$

We determine the residues as follows:

$$R_{1,1} = \lim_{s\to 0}\left[s^3\,F(s)\right] = \lim_{s\to 0}\left[\cancel{s^3}\,\frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{\cancel{s^3}\,(s+1)^2}\right] = \lim_{s\to 0}\left[\frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{(s+1)^2}\right] = \frac{-1}{1} = -1.$$

To find $R_{1,2}$, we use the expression

$$R_{1,2} = \lim_{s\to 0}\left[\frac{\mathrm{d}}{\mathrm{d}s}\cancel{s^3}\,\frac{4\,s^4+6\,s^3-2\,s-1}{\cancel{s^3}\,(s+1)^2}\right] = \lim_{s\to 0}\left[\frac{\mathrm{d}}{\mathrm{d}s}\,\frac{4\,s^4+6\,s^3-2\,s-1}{(s+1)^2}\right] = \lim_{s\to 0}\frac{2\,s\,(4\,s^3+11\,s^2+9\,s+1)}{(s+1)^3} = 0.$$

Finally,

$$R_{1,1} = \lim_{s\to 0}\frac{1}{2}\left[\frac{\mathrm{d}^2}{\mathrm{d}s^2}\cancel{s^3}\,\frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{\cancel{s^3}\,(s+1)^2}\right] = \frac{1}{2}\lim_{s\to 0}\left[\frac{\mathrm{d}^2}{\mathrm{d}s^2}\,\frac{2\,(4\,s^4 + 16\,s^3 + 24\,s^2 + 16\,s + 1)}{(s+1)^4}\right] = \frac{1}{2}\,\frac{2}{1} = 1.$$

Now we compute the residues associated with $s = -1$:

$$R_{2,1} = \lim_{s\to -1}\left[(s+1)^2 F(s)\right] = \lim_{s\to -1}\left[\frac{\cancel{(s+1)^2}\,4\,s^4+6\,s^3-2\,s-1}{s^3\,\cancel{(s+1)^2}}\right] = \lim_{s\to -1}\frac{4\,s^4+6\,s^3-2\,s-1}{s^3} = \frac{4-6+2-1}{-1} = 1.$$

Finally

$$R_{2,2} = \lim_{s\to -1}\left[\frac{\mathrm{d}}{\mathrm{d}s}(s+1)^2\,F(s)\right] = \lim_{s\to -1}\left[\frac{\mathrm{d}}{\mathrm{d}s}\,\frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{s^3}\right] = \lim_{s\to -1}\frac{4\,s^4 + 4\,s + 3}{s^4} = \frac{4-4+3}{1} = 3.$$

Thus,

$$F(s) = \frac{1}{s} - \frac{1}{s^3} + \frac{3}{s+1} + \frac{1}{(s+1)^2}. \tag{8.32}$$

The inverse Laplace transform is

$$\mathcal{L}^{-1}\{F\}(t) = H(t) - \frac{1}{2}\,t^2\,H(t) + 3\,e^{-t}\,H(t) + t\,e^{-t}\,H(t) = \left(1 - \frac{1}{2}\,t^2 + 3\,e^{-t} + t\,e^{-t}\right) H(t).$$

We can also solve this problem by determining the partial fraction decomposition of $F(s)$ without computing the residues explicitly, i.e., we write $f(s)$ as

$$F(s) = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s^3} + \frac{D}{s+1} + \frac{E}{(s+1)^2}. \tag{8.33}$$

Equating (8.36) and (8.33) yields

$$\begin{aligned}\frac{4\,s^4 + 6\,s^3 - 2\,s - 1}{s^3\,(s+1)^2} &= \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s^3} + \frac{D}{s+1} + \frac{E}{(s+1)^2} \\ &= \frac{A\,s^2(s+1)^2 + B\,s\,(s+1)^2 + C\,(s+1)^2 + D\,s^3\,(s+1) + E\,s^3}{s^3\,(s+1)^2}.\end{aligned}$$

Hence, A, B, C, D and E must satisfy the equation

$$A\,s^2(s+1)^2 + B\,s\,(s+1)^2 + C\,(s+1)^2 + D\,s^3\,(s+1) + E\,s^3 = 4\,s^4 + 6\,s^3 - 2\,s - 1. \tag{8.34}$$

Using the principle of identity of polynomials, one can easily obtain $A = 1$, $B = 0$, $C = -1$, $D = 3$ and $E = 1$.

The case of repeated complex poles can be dealt as in the case of simple poles. Consider for the sake of argument the parts of a Laplace transform involving the pair of complex conjugate poles (p_1, p_2), with $p_2 = \bar{p}_1$, with the same multiplicity ν:

$$F(s) = \sum_{j=1}^{\nu}\left[\frac{R_{1,j}}{(s-p_1)^{\nu-j+1}} + \frac{\bar{R}_{2,j}}{(s-\bar{p}_1)^{\nu-j+1}}\right]. \tag{8.35}$$

Let $p_1 = \sigma + i\,\omega$, so that $p_2 = \sigma - i\,\omega$. Moreover, let $R_{1,j} = M_j\,e^{i\,\varphi_j}$, so that $R_{2,j} = M_j\,e^{-i\,\varphi_j}$ (it can be seen that the residues associated with complex conjugate poles are still complex conjugate). Thus,

$$F(s) = \sum_{j=1}^{\nu}\left[\frac{M_j\,e^{i\,\varphi_j}}{(s-\sigma-i\,\omega)^{\nu-j+1}} + \frac{M_j\,e^{-i\,\varphi_j}}{(s-\sigma+i\,\omega)^{\nu-j+1}}\right].$$

Consider the generic j-th term of this sum. Using again (8.8) we find

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{M_j\,e^{i\,\varphi_j}}{(s-\sigma-i\,\omega)^{\nu-j+1}} + \frac{M_j\,e^{-i\,\varphi_j}}{(s-\sigma+i\,\omega)^{\nu-j+1}}\right\}(t) &= M_j\,e^{i\,\varphi_j}\,\frac{1}{(\nu-j)!}\,t^{\nu-j}\,e^{(\sigma+i\,\omega)\,t}\,H(t) \\ &\quad + M_j\,e^{-i\,\varphi_j}\,\frac{1}{(\nu-j)!}\,t^{\nu-j}\,e^{(\sigma-i\,\omega)\,t}\,H(t) \\ &= \frac{1}{(\nu-j)!}\,M_j\,t^{\nu-j}\,e^{\sigma\,t}\left(e^{i\,(\omega\,t+\varphi_j)} + e^{-i\,(\omega\,t+\varphi_j)}\right) H(t) \\ &= \frac{2}{(\nu-j)!}\,M_j\,t^{\nu-j}\,e^{\sigma\,t}\,\frac{e^{i\,(\omega\,t+\varphi_j)} + e^{-i\,(\omega\,t+\varphi_j)}}{2}\,H(t) \\ &= \frac{2}{(\nu-j)!}\,M_j\,t^{\nu-j}\,e^{\sigma\,t}\,\cos(\omega\,t+\varphi_j)\,H(t).\end{aligned}$$

It follows that

$$\mathcal{L}^{-1}\{F\}(t) = 2\sum_{j=1}^{\nu} \frac{1}{(\nu-j)!} M_j\, t^{\nu-j}\, e^{\sigma t}\, \cos(\omega t + \varphi_j)\, H(t).$$

Example

Consider the rational function

$$F(s) = \frac{s-1}{(s^2+9)^2}. \tag{8.36}$$

Let $p_1 = -3\,j$, so that $(p_1, \bar{p}_1)$ is the pair of complex conjugate poles of $F(s)$ with double multiplicity. Writing $F(s)$ as in (8.35) yields

$$F(s) = \frac{R_{1,2}}{s+3\,j} + \frac{R_{1,1}}{(s+3\,j)^2} + \frac{\bar{R}_{1,2}}{s-3\,j} + \frac{\bar{R}_{1,1}}{(s-3\,j)^2}.$$

Using (8.29) we find

$$R_{1,1} = \lim_{s\to p_1} (s-p_1)^2\, F(s) = \lim_{s\to -3\,j} (s+3\,j)^2\, \frac{s-1}{(s+3\,j)^2\,(s-3\,j)^2} = \lim_{s\to -3\,j} \frac{s-1}{(s-3\,j)^2} = \frac{-3\,j-1}{(-6\,j)^2} = \frac{1}{36} + \frac{1}{12}\,j.$$

We find $R_{1,2}$ using the same formula:

$$R_{1,2} = \left[\frac{\mathrm{d}}{\mathrm{d}s}\left((s+3\,j)^2 \frac{s-1}{(s^2+9)^2}\right)\right]_{s=-3\,j} = \left[\frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{s-1}{(s-3\,j)^2}\right)\right]_{s=-3\,j} = \left[\frac{(s-3\,j)-2\,(s-1)}{(s-3\,j)^2}\right]_{s=-3\,j} = -\frac{1}{108}\,j.$$

Hence,

$$F(s) = -\frac{1}{108}\,j\,\frac{1}{s+3\,j} + \left(\frac{1}{36} + \frac{1}{12}\,j\right)\frac{1}{(s+3\,j)^2} + \frac{1}{108}\,j\,\frac{1}{s-3\,j} + \left(\frac{1}{36} - \frac{1}{12}\,j\right)\frac{1}{(s-3\,j)^2}.$$

We now write the residues in polar form. In other words, we need to determine their magnitude and angle:

$$\begin{aligned} M_{1,1} &= |R_{1,1}| = \sqrt{\left(\frac{1}{36}\right)^2 + \left(\frac{1}{12}\right)^2} = \sqrt{\frac{5}{648}} \\ \varphi_{1,1} &= \arg R_{1,1} = \arctan \frac{\frac{1}{12}}{\frac{1}{36}} = \arctan 3 \\ M_{1,2} &= |R_{1,2}| = \frac{1}{108} \\ \varphi_{1,2} &= \arg R_{1,2} = -\frac{\pi}{2}. \end{aligned}$$

The inverse Laplace transform is, for $t \geq 0$, given by

$$\begin{aligned} f(t) &= 2\sqrt{\frac{5}{648}}\, t\, \cos(-3\,t + \arctan 3) + 2\,\frac{1}{108}\,\cos\left(-3\,t - \frac{\pi}{2}\right) \\ &= \sqrt{\frac{5}{162}}\, t\, \cos(3\,t - \arctan 3) + 2\,\frac{1}{108}\,\cos\left(3\,t + \frac{\pi}{2}\right), \end{aligned}$$

and it is obviously zero for $t < 0$.

Example Consider the rational function

$$F(s) = \frac{2\,s\,(s^4 + s^3 + 4\,s^2 - 3\,s + 3)}{(s^2+1)^3}.$$

Its poles are $\pm i$, with multiplicity $\nu = 3$. We can therefore write

$$F(s) = \frac{R_{1,3}}{s-i} + \frac{R_{1,2}}{(s-i)^2} + \frac{R_{1,1}}{(s-i)^3} + \frac{\bar{R}_{1,3}}{s+i} + \frac{\bar{R}_{1,2}}{(s+i)^2} + \frac{\bar{R}_{1,1}}{(s+i)^3}.$$

We find for example

$$\begin{aligned} R_{1,1} &= 1 \cdot \lim_{s\to i}(s-i)^3\, f(s) = \lim_{s\to i} \frac{2\,s\,(s^4 + s^3 + 4\,s^2 - 3\,s + 3)}{(s+i)^3} \\ &= \frac{2\,i\,(i^4 + i^3 + 4\,i^2 - 3\,i + 3)}{(i+i)^3} = \frac{2\,i\,(1 - i + 4 - 3\,i + 3)}{-8\,i} = i. \end{aligned}$$

Thus, $M_1 = 1$ and $\varphi_1 = \pi/2$. Similarly, we find $R_{1,2} = 1 - i$ (so that $M_2 = \sqrt{2}$ and $\varphi_2 = -\pi/4$) and $R_{1,3} = 1$ (leading to $M_3 = 1$ and $\varphi_3 = 0$). The inverse Laplace transform is therefore

$$\begin{aligned} \mathcal{L}^{-1}\{F\}(t) &= 2\sum_{j=1}^{3} \frac{1}{(3-j)!}\, M_j\, t^{3-j}\, \cos(t + \varphi_j)\, H(t) \\ &= 2\left(\frac{1}{2}\,t^2\,\cos\left(t + \frac{\pi}{2}\right) + \sqrt{2}\,t\,\cos\left(t - \frac{\pi}{4}\right) + \cos(t)\right) H(t) \\ &= \left(t\,(2-t)\,\sin t + 2\,(t+1)\,\cos t\right) H(t). \end{aligned}$$

In the following examples, we show how the Laplace transform can be a powerful instrument for the solution of ordinary differential equations and Cauchy problems.

Example Consider the Cauchy problem

$$\begin{cases} w''(x) + w'(x) + 5\,w(x) = 1 \\ w(0) = 3 \\ w'(0) = 4 \end{cases}$$

First, we consider the extension of w to $\mathbb{R}$: we define, for all $x \in \mathbb{R}$, the function

$$\tilde{w}(x) = \begin{cases} w(x) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

The function $\tilde{w}$ is now considered as a distribution. We have

$$\tilde{w}'(x) = \widetilde{w'}(x) + 3\,\delta \quad \Rightarrow \quad \widetilde{w'}(x) = \tilde{w}'(x) - 3\,\delta.$$

Likewise,

$$\tilde{w}''(x) = \widetilde{w''}(x) + 4\,\delta + 3\,\delta' \quad \Rightarrow \quad \widetilde{w''}(x) = \tilde{w}''(x) - 4\,\delta - 3\,\delta'.$$

Re-writing the differential equation as

$$\widetilde{w''}(x) + \widetilde{w'}(x) + \tilde{w}(x) = H(x),$$

where H is the trivial extension of the constant function 1, we find

$$\tilde{w}''(x) - 4\,\delta - 3\,\delta' + \tilde{w}'(x) - 3\,\delta + \tilde{w}(x) = H,$$

i.e,

$$\tilde{w}''(x) + \tilde{w}'(x) + 5\,\tilde{w}(x) = H + 7\,\delta + 3\,\delta'.$$

We obtain

$$(s^2 + s + 5)\,\mathcal{L}\{\tilde{w}\}(s) = \frac{1}{s} + 7 + 3\,s \quad \Rightarrow \quad \mathcal{L}\{\tilde{w}\}(s) = \frac{3\,s^2 + 7\,s + 1}{s\,(s^2 + s + 5)}.$$

We have to determine the inverse Laplace transform. The poles are $p_1 = 0$ and $p_{2,3} = \frac{-1 \pm i\,\sqrt{19}}{2}$. We determine the corresponding residues:

$$\begin{aligned} R_1 &= \lim_{s\to 0} s\,\frac{3\,s^2 + 7\,s + 1}{s\,(s^2 + s + 5)} = \lim_{s\to 0} \frac{3\,s^2 + 7\,s + 1}{s^2 + s + 5} = \frac{1}{5} \\ R_2 &= \lim_{s\to \frac{-1+i\,\sqrt{19}}{2}} \left(s - \frac{-1 + i\,\sqrt{19}}{2}\right) \frac{3\,s^2 + 7\,s + 1}{s\,(s^2 + s + 5)} \\ &= \lim_{s\to \frac{-1+i\,\sqrt{19}}{2}} \frac{3\,s^2 + 7\,s + 1}{s\left(s - \frac{-1-i\,\sqrt{19}}{2}\right)} = \frac{3\left(\frac{-1+i\,\sqrt{19}}{2}\right)^2 + 7\left(\frac{-1+i\,\sqrt{19}}{2}\right) + 1}{\left(\frac{-1+i\,\sqrt{19}}{2}\right) i\,\sqrt{19}} = \frac{133 - 27\,i\,\sqrt{19}}{95}, \end{aligned}$$

and $R_3 = \bar{R}_2$. We easily find the corresponding magnitude and a corresponding argument:

$$M = |R_2| = \frac{2\,\sqrt{7885}}{95} \qquad \varphi = \operatorname{Arg} R_2 = -\arctan\frac{27\,\sqrt{19}}{133}.$$

We obtain

$$\tilde{w}(t) = \frac{1}{5}\,H(t) + \frac{34\,\sqrt{7885}}{95}\,e^{-\frac{t}{2}}\,\cos\left(\frac{\sqrt{19}}{2}\,t - \arctan\frac{27\,\sqrt{19}}{133}\right) H(t).$$

Example Consider the same Cauchy problem of the previous example:

$$\begin{cases} w''(x) + w'(x) + 5\,w(x) = 1 \\ w(0) = 3 \\ w'(0) = 4 \end{cases}$$

We found that the solution can be obtained by finding the inverse Laplace transform of the rational function

$$\frac{3\,s^2 + 7\,s + 1}{s\,(s^2 + s + 5)}.$$

To this end, we use the partial fraction expansion

$$\frac{3\,s^2 + 7\,s + 1}{s\,(s^2 + s + 5)} = \frac{A}{s} + \frac{B\,s + C}{\left(s + \frac{1}{2}\right)^2 + \frac{19}{4}}.$$

Using the principle of identity of polynomials we find

$$A = \frac{1}{5}, \quad B = \frac{14}{5}, \quad C = \frac{34}{5}.$$

Thus,

$$\frac{3\,s^2+7\,s+1}{s\,(s^2+s+5)}=\frac{1}{5\,s}+\frac{1}{5}\,\frac{14\,s+34}{\left(s+\frac{1}{2}\right)^2+\frac{19}{4}}.$$

Using the results of this section we find

$$\frac{1}{\left(s+\frac{1}{2}\right)^2+\frac{19}{4}}=\mathcal{L}\left\{\frac{2}{\sqrt{19}}\,\sin\left(\sqrt{\frac{19}{4}}\,t\right)e^{-\frac{t}{2}}\,H\right\}(s)$$

and

$$\frac{s}{\left(s+\frac{1}{2}\right)^2+\frac{19}{4}}=\frac{s+\frac{1}{2}}{\left(s+\frac{1}{2}\right)^2+\frac{19}{4}}-\frac{1}{2}\,\frac{1}{\left(s+\frac{1}{2}\right)^2+\frac{19}{4}}=\mathcal{L}\left\{e^{-\frac{t}{2}}\,\cos\left(\frac{\sqrt{19}}{2}\,t\right)H-\frac{1}{\sqrt{19}}\,\sin\left(\frac{\sqrt{19}}{2}\,t\right)e^{-\frac{t}{2}}\,H\right\}(s).$$

We obtain the solution, for $t\geq 0$,

$$\frac{1}{5}+\frac{1}{5}\,e^{-\frac{t}{2}}\left[14\left(\cos\left(\frac{\sqrt{19}}{2}\,t\right)-\frac{1}{\sqrt{19}}\,\sin\left(\frac{\sqrt{19}}{2}\,t\right)\right)+34\,\frac{2}{\sqrt{19}}\,\sin\left(\frac{\sqrt{19}}{2}\,t\right)\right].$$

The Laplace transform can also be used to solve integro-differential equations.

Example We want to determine the $\mathcal{L}$-transformable distributions $T\in\mathcal{D}'(\mathbb{R};\mathbb{R})$ that satisfy the integro-differential equation

$$T'=-T+(H\,\sin)*T+5\,H.$$

Applying the Laplace transform on both sides gives

$$s\,\mathcal{L}\{T\}(s)=-\mathcal{L}\{T\}(s)+\mathcal{L}\{H\,\sin\}(s)\,\mathcal{L}\{T\}(s)+\frac{5}{s}=-\mathcal{L}\{T\}(s)+\frac{1}{s^2+1}\,\mathcal{L}\{T\}(s)+\frac{5}{s},$$

i.e.,

$$\mathcal{L}\{T\}(s)=\frac{\frac{5}{s}}{s+1-\frac{1}{s^2+1}}=\frac{5\,(s^2+1)}{s^2\,(s^2+s+1)}.$$

The denominator has complex roots $-\frac{1}{2}\pm\frac{3}{2}\,i$. Using the partial fraction expansion

$$\mathcal{L}\{T\}(s)=\frac{A}{s}+\frac{B}{s^2}+\frac{C\,s+D}{s^2+s+1}$$

gives $A=-5$ and $B=C=D=5$. We can therefore write

$$\mathcal{L}\{T\}(s)=-\frac{5}{s}+\frac{5}{s^2}+5\,\frac{s+1}{s^2+s+1}=-\frac{5}{s}+\frac{5}{s^2}+5\,\frac{s+1}{\left(s+\frac{1}{2}\right)^2+\frac{3}{4}}.$$

We have

$$\frac{1}{\left(s+\frac{1}{2}\right)^2+\frac{3}{4}}=\mathcal{L}\left\{\frac{2}{\sqrt{3}}\,e^{-\frac{t}{2}}\,\sin\left(\frac{\sqrt{3}}{2}\,t\right)H(t)\right\}$$

and

$$\begin{aligned}\frac{s}{\left(s+\frac{1}{2}\right)^2+\frac{3}{4}} &= \frac{s+\frac{1}{2}}{\left(s+\frac{1}{2}\right)^2+\frac{3}{4}} - \frac{1}{2}\frac{1}{\left(s+\frac{1}{2}\right)^2+\frac{3}{4}} \\ &= \mathcal{L}\left\{e^{-\frac{t}{2}}\cos\left(\frac{\sqrt{3}}{2}t\right)H(t) - \frac{1}{\sqrt{3}}e^{-\frac{t}{2}}\sin\left(\frac{\sqrt{3}}{2}t\right)H(t)\right\}(s).\end{aligned}$$

We conclude that

$$T = -5\,H(t) + 5\,t\,H(t) + \frac{10}{\sqrt{3}}e^{-\frac{t}{2}}\sin\left(\frac{\sqrt{3}}{2}t\right)H(t) + 5\left(e^{-\frac{t}{2}}\cos\left(\frac{\sqrt{3}}{2}t\right)H(t) - \frac{1}{\sqrt{3}}e^{-\frac{t}{2}}\sin\left(\frac{\sqrt{3}}{2}t\right)H(t)\right).$$

Example We want to solve the equation

$$f'' + 4\,f' + 5\,f = H$$

in the set of $\mathcal{L}$-transformable distributions. We can express this equation as $f = k * H$ with

$$K(s) = \mathcal{L}\{k\} = \frac{1}{s^2 + 4\,s + 5} = \frac{1}{(s+2)^2 + 1}.$$

Thus, $k(t) = e^{-2\,t} \sin t\, H(t)$. We find

$$\begin{aligned} f &= k * H = \int_{\mathbb{R}} H(t-\tau)\, e^{-2\,\tau} \sin\tau\, H(\tau)\, d\tau \\ &= \int_0^t e^{-2\,\tau} \sin\tau\, d\tau \\ &= \int_0^t e^{-2\,\tau}\, \frac{e^{2\,\tau} - e^{-i\,\tau}}{2\,i}\, d\tau \\ &= \frac{1}{2\,i} \int_0^t e^{\tau\,(i-2)}\, d\tau - \frac{1}{2\,i} \int_0^t e^{-\tau\,(i+2)}\, d\tau \\ &= \frac{1}{2\,i} \left[\frac{e^{\tau\,(i-2)}}{i-2} \right]_0^t + \frac{1}{2\,i} \left[\frac{e^{-\tau\,(i+2)}}{i+2} \right]_0^t \\ &= \frac{1}{2\,i} (2\,e^{-2\,t}\,(e^{i\,t} - e^{-i\,t}) + i\,e^{-i\,t}(e^{i\,t} + e^{-i\,t}) - 2\,i) \frac{1}{5} \end{aligned}$$

so that $f(t) = -\frac{1}{5}\,(2\,e^{-2\,t} \sin t + e^{-2\,t} \cos t - 1)\, H(t)$.

Example We want to solve the equation

$$f'' + 4\,f = \delta'.$$

in the set of $\mathcal{L}$-transformable distributions. We can express this equation as $f = k * \delta' = k'$ with

$$K(s) = \mathcal{L}\{k\} = \frac{1}{s^2 + 4} = \frac{1}{(s+2)^2 + 1} = \frac{1}{2}\,\frac{1}{2}\,\frac{1}{\left(\frac{s}{2}\right)^2 + 1},$$

which gives $k(t) = \frac{1}{2}\, H(t)\, \sin(2\,t)$, so that $f(t) = \cos(2\,t)\, H(t)$.

Example We want to solve the equation

$$f' - f = t^{-\frac{1}{2}}\, H(t).$$

in the set of $\mathcal{L}$-transformable distributions. We can express this equation as $f = k * t^{-\frac{1}{2}}\, H(t)$ with

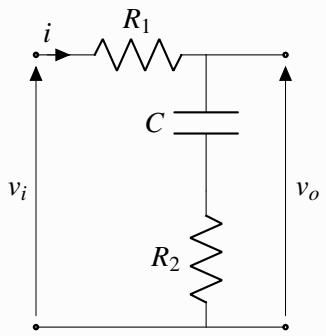
$$K(s) = \mathcal{L}\{k\} = \frac{1}{z-1},$$

which gives $k(t) = e^t\, H(t)$. Therefore

$$f = k * t^{-\frac{1}{2}}\, H(t) = \int_{\mathbb{R}} e^{t-\tau}\, H(t-\tau)\, \tau - \frac{1}{2}\, H(\tau)\, d\tau = \int_0^t e^t\, e^{-\tau}\, \tau^{-\frac{1}{2}}\, d\tau = e^t \int_0^t e^{-\tau}\, \tau^{-1/2}\, d\tau.$$

Example

Consider the electric circuit in the figure



The relation between input voltage v_i and output voltage v_o is given by

$$C\,(R_1 + R_2)\,\frac{\mathrm{d}v_o(t)}{\mathrm{d}t} + v_o(t) = C\,R_2\,\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} + v_i(t).$$

Suppose that the voltage v_o is constant and equal to 2V for all $t < 0$. This results in an initial condition $v_o(0^-) = 2\mathrm{V}$, and we obtain the differential equation with initial condition

$$\begin{cases} a\,\dfrac{\mathrm{d}y(t)}{\mathrm{d}t} + y(t) = b\,\dfrac{\mathrm{d}u(t)}{\mathrm{d}t} + u(t) \\ y_{0^-}^{(0)} = \lim\limits_{t\to 0^-} y(t) = 2, \end{cases} \tag{8.37}$$

where $y = v_o$ and $u = v_i$, and where we have defined for convenience $a = C\,(R_1 + R_2)$ and $b = C\,R_2$. Suppose now that u is equal to the Heaviside function, i.e., $u = H$.

We assume that the input is an $\mathcal{L}$-transformable distribution. We want to solve (8.37) in the space of distributions, i.e., we need to determine the distributions y that satisfy this equation, where the derivatives are in the sense of distributions. Eq. (8.37) can be assumed to be defined over the entire real axis (i.e., $t \in \mathbb{R}$), but we are only interested in $y(t)$ for $t \geq 0$. We need to understand the meaning behind the assignment of the initial condition $\lim\limits_{t\to 0^-} y(t) = 2$.

First, we observe that every solution of (8.37) can be expressed as the sum of a particular solution y_d and of a solution y_h of the associated homogeneous equation

$$a\,\frac{\mathrm{d}y(t)}{\mathrm{d}t} + y(t) = 0. \tag{8.38}$$

Indeed, if $y = y_h + y_d$ we find

$$a\,\frac{\mathrm{d}y(t)}{\mathrm{d}t} + y(t) = \underbrace{a\,\frac{\mathrm{d}y_h(t)}{\mathrm{d}t} + y_h(t)}_{=0} + a\,\frac{\mathrm{d}y_p(t)}{\mathrm{d}t} + y_d(t) = b\,\frac{\mathrm{d}u(t)}{\mathrm{d}t} + u(t).$$

Eq. (8.38) has to be considered in the sense of distributions, but it is not difficult to see that the solutions of (8.38) in the sense of distributions coincide with the standard (infinitely differentiable) solutions. In other words, considering the space of distributions does not increase the number of solutions of (8.38). We assume that an $\mathcal{L}$-transformable particular solution y_d exists (and we will verify afterwards that such solution exists). In this case, both sides of (8.37) can be transformed: using the rules of differentiation for the Laplace transform of a distribution gives

$$a\,s\,\mathcal{L}\{y_p\}(s) + \mathcal{L}\{y_p\}(s) = (b\,s + 1)\,\frac{1}{s}\mathcal{L}\{y_p\}(s) = \frac{b\,s + 1}{a\,s + 1}\,\frac{1}{s}.$$

The expression in the right hand-side of the latter is the Laplace transform of a distribution, because $\frac{b\,s+1}{s\,(a\,s+1)}$ is holomorphic for $\Re\mathrm{e}\{s\} > \max\{-1/a, 0\}$ (in this example, since $a > 0$, this condition becomes $\Re\mathrm{e}\{s\} > -1/a$. This shows that $\mathcal{L}\{y_p\}(s)$ is the Laplace transform of a distribution

$$y_p = \mathcal{L}^{-1}\left\{\frac{b\,s+1}{s\,(a\,s+1)}\right\}.$$

The limit

$$\lim_{t\to 0^-} y(t)$$

in the initial condition has to be interpreted as the corresponding limit of y_h, which is a distribution of function type (and, being the solution of a homogeneous differential equation, it is of class C^∞). Hence,

$$\lim_{t\to 0^-} y(t) = \lim_{t\to 0^+} y_h(t) = 2. \tag{8.39}$$

Moreover, defining

$$\tilde{y}(t) = \begin{cases} y_h(t) & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

we find that $y(t)$, for $t \geq 0$, is equal to $\tilde{y}(t) + y_p$. This means that, defining also

$$y(t) = \begin{cases} y(t) & \text{if } t \geq 0 \\ 0 & \text{if } t < 0, \end{cases}$$

we can write $\tilde{y} = \tilde{y}_h + y_p$. To find y_h that solves (8.38) with the initial condition (8.39), we consider $\tilde{y}_h$ that solves (8.38) for every $t > 0$ in the sense of regular functions. Since the equation is linear, the growth of $\tilde{y}_h$ is at most exponential, and therefore $\tilde{y}_h$ is $\mathcal{L}$-transformable. Taking the Laplace transform in both sides of (8.38) gives

$$a\,(s\,\mathcal{L}\{\tilde{y}_h\}(s) - \tilde{y}_h(0^+)) + \mathcal{L}\{\tilde{y}_h\}(s) = 0,$$

so that

$$\mathcal{L}\{\tilde{y}_h\}(s) = \frac{1}{a\,s+1}\,\tilde{y}_h(0^+).$$

We obtain

$$\begin{aligned}
\mathcal{L}\{\tilde{y}\}(s) &= \mathcal{L}\{y_p\}(s) + \mathcal{L}\{\tilde{y}_h\}(s) \\
&= \frac{b\,s+1}{s\,(a\,s+1)} + \frac{2\,a}{a\,s+1} \\
&= \frac{1}{a}\,\frac{b\,s+1}{s\left(s+\frac{1}{a}\right)} + \frac{2}{s+\frac{1}{a}} \\
&= \frac{1}{a}\left(\frac{a}{s} + \frac{b-a}{s+\frac{1}{a}}\right) + \frac{2}{s+\frac{1}{a}} \\
&= \frac{1}{s} + \frac{b-a}{a}\,\frac{1}{s+\frac{1}{a}} + \frac{2}{s+\frac{1}{a}}.
\end{aligned}$$

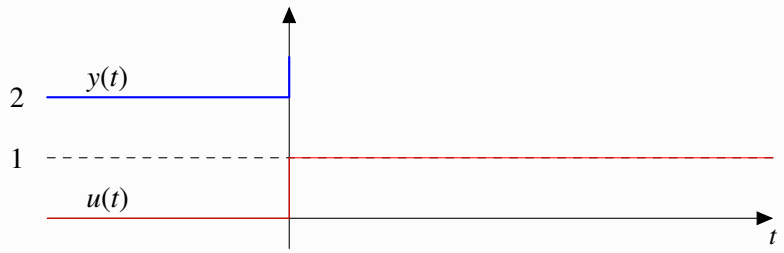
so that

$$\tilde{y}(t) = H(t) + \left(\frac{b-a}{a} + 2\right) e^{-\frac{1}{a}t}\,H(t).$$

Using for example the values $C = 0.1$ F, $R_1 = 1\ \Omega$ and $R_2 = 2\ \Omega$, we find $a = 0.3$ and $b = 0.2$. The output $v_o(t)$ is given by

$$v_o(t) = \begin{cases} 2 & \text{if } t < 0 \\ 1 + \frac{5}{3}\,e^{-\frac{10}{3}t} & \text{if } t \geq 0. \end{cases}$$

The figure below shows the response $v_o(t)$ (blue line) and the input $v_i(t)$ (red line).



9 Fourier series

9.1 Introduction

We recall that $L^2([-\pi,\pi];\mathbb{C})$ is the set of equivalence classes of Lebesgue measurable functions f defined almost everywhere in $[-\pi,\pi]$ such that

$$\int_{[-\pi,\pi]} |f(x)|^2 \,\mathrm{d}x < +\infty. \tag{9.1}$$

The equivalence class $[f]$ containing the function f which satisfies (9.1) consists of all measurable functions defined almost everywhere in $[-\pi,\pi]$ which coincide with f almost everywhere. Hence, if $g \in [f]$, there exist two null sets N_1 and N_2 such that $f : [-\pi,\pi] \setminus N_1 \longrightarrow \mathbb{C}$ and $g : [-\pi,\pi] \setminus N_2 \longrightarrow \mathbb{C}$, and the set $N_3 = \{x \in [-\pi,\pi] \setminus (N_1 \cup N_2) \,|\, f(x) \neq g(x)\}$ is measurable and has measure zero. The inner product between the equivalence classes $[f]$ and $[g]$ is given by

$$\langle [f], [g] \rangle = \int_{[-\pi,\pi]} \overline{g(x)}\, f(x) \,\mathrm{d}x. \tag{9.2}$$

It is easy to see that this definition is independent of the particular representatives of the equivalent classes $[f]$ and $[g]$. The inner product (9.2) induces the norm

$$\|[f]\| = \sqrt{\langle [f], [f] \rangle} = \left(\int_{[-\pi,\pi]} |f(x)|^2 \,\mathrm{d}x \right)^{1/2}.$$

In the sequel, with some (well established) abuse of notation we will write f and g instead of $[f]$ and $[g]$, thus identifying functions which coincide almost everywhere.

We now introduce an orthonormal system in $L^2([-\pi,\pi];\mathbb{C})$. To this end, we make use of the complex exponential function $\exp z = e^z$. In particular, we consider the case $z = i\,x$, where i is the imaginary unit and $x \in \mathbb{R}$. We recall some well-known properties of the function

$$\begin{aligned} h : \mathbb{R} &\longrightarrow \mathbb{C} \\ x &\mapsto e^{i\,x}. \end{aligned}$$

The first is the Euler's formula

$$e^{i\,x} = \cos x + i\, \sin x, \qquad \forall\, x \in \mathbb{R}. \tag{9.3}$$

From (9.3), one can easily deduce the following properties:

Theorem 9.1 *Let $x, y \in \mathbb{R}$. Then:*

(i) $|e^{i\,x}| = 1$;

(ii) $e^{i\,x}\, e^{i\,y} = e^{i\,(x+y)}$;

(iii) $\forall\, k \in \mathbb{Z}, \quad (e^{i\,x})^k = e^{i\,k\,x}$;

(iv) $\overline{e^{i\,x}} = e^{-i,\boldsymbol{x}} = e^{i\,(-x)}$;

(v) $\forall\, k \in \mathbb{Z}, \quad e^{i\,(x+2\,k\,\pi)} = e^{i\,x}$;

(vi) *in particular,* $\forall\, k \in \mathbb{Z}, \quad e^{i\,2\,k\,\pi} = 1$.

Theorem 9.2 *For all $x \in \mathbb{R}$, we have $h'(x) = i\,e^{i\,x} = i\,h(x)$.*

We now introduce the following notation. We define, for $k \in \mathbb{Z}$ and $x \in [-\pi, \pi]$,

$$x_k(t) \stackrel{\text{def}}{=} \frac{e^{i\,k\,t}}{\sqrt{2\,\pi}}. \tag{9.4}$$

From now on, we will focus on the Hilbert space $(L^2([-\pi,\pi];\mathbb{C}), \langle\cdot,\cdot\rangle)$, where the inner product is the one in (9.2).

Theorem 9.3 *The sequence $(x_k)_{k\in\mathbb{Z}}$ is an orthonormal system in the Hilbert space $(L^2([-\pi,\pi];\mathbb{C}), \langle\cdot,\cdot\rangle)$, where $\langle\cdot,\cdot\rangle$ is the inner product in (9.2).*

Proof Let $k, h \in \mathbb{Z}$. If $k = h$, using Theorem 9.1, we have

$$\langle x_k, x_k\rangle = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} e^{i\,k\,t}\,e^{-i\,k\,t}\,\mathrm{d}t = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} 1\,\mathrm{d}t = 1.$$

If $k \neq h$,

$$\langle x_k, x_h\rangle = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} e^{i\,k\,t}\,e^{-i\,h\,t}\,\mathrm{d}t = \frac{1}{2\,\pi}\int_{-\pi}^{\pi} e^{i\,(k-h)\,t}\,\mathrm{d}t = \frac{1}{2\,\pi}\left[\frac{e^{i\,(k-h)\,t}}{i\,(k-h)}\right]_{t=-\pi}^{t=\pi} = 0,$$

where we used *(v)* in Theorem 9.1.

From Theorem 2.32 we have the following result.

Corollary 9.1 *Let $f \in L^2([-\pi,\pi];\mathbb{C})$ and $(x_k)_{k\in\mathbb{Z}}$ as in (9.4). Then, the limit*

$$\lim_{n\to\infty}\sum_{k=-n}^{n}\langle f, x_k\rangle\,x_k$$

exists in $L^2([-\pi,\pi];\mathbb{C})$.

The question is whether

$$f = \lim_{n\to\infty}\sum_{k=-n}^{n}\langle f, x_k\rangle\,x_k.$$

We will see that this is the case for any $f \in L^2([-\pi,\pi];\mathbb{C})$, and it depends on whether $(x_i)_{i\in\mathbb{Z}}$ is a maximal orthonormal system, see Theorem 2.33 and Theorem2 2.34. For all $n \in \mathbb{N}$, we define

$$S_n(f) = \lim_{n\to\infty}\sum_{k=-n}^{n}\langle f, x_k\rangle\,x_k, \tag{9.5}$$

and we define in this way a sequence $(S_n(f))_{n\in\mathbb{N}}$. Obviously

$$S_n(f)(t) = \lim_{n\to\infty} \sum_{k=-n}^{n} \langle f, x_k\rangle\, x_k(t) = \int_{-\pi}^{\pi} \frac{1}{2\pi} \sum_{k=-n}^{n} e^{i\,k\,(t-\tau)} f(\tau)\,\mathrm{d}\tau.$$

We calculate the sum within the integral.

Lemma 9.1 *Let $n \in \mathbb{N}$ and let $t \in [-\pi, \pi]$. Then,*

$$\sum_{k=-n}^{n} e^{i\,k\,(t-\tau)} = \begin{cases} \dfrac{\cos(n\,t) - \cos\left((n+1)\,t\right)}{2\,\sin^2\frac{t}{2}} = \dfrac{\sin\left(\left(n+\frac{1}{2}\right)\,t\right)}{\sin\frac{t}{2}} & \text{if } t \neq 0 \\ 2\,n+1 & \text{if } t = 0 \end{cases}$$

Proof The case $t = 0$ is straightforward. Let $t \in [-\pi, \pi] \setminus \{0\}$. We have

$$\sum_{k=-n}^{n} e^{i\,k\,t} = e^{-i\,n\,t} \sum_{k=-n}^{n} e^{i\,(k+n)\,t} = e^{-i\,n\,t} \sum_{h=0}^{2\,n} e^{i\,h\,t}.$$

We have $e^{i\,t} \neq 1$. We recall that, if $q \in \mathbb{C} \setminus \{1\}$,

$$\forall\, m \in \mathbb{N}, \quad \sum_{k=0}^{m} q^k = \frac{1 - q^{m+1}}{1-q}.$$

We find

$$\begin{aligned} e^{-i\,n\,t} \sum_{h=0}^{2\,n} e^{i\,h\,t} &= e^{-i\,n\,t}\, \frac{1 - e^{i\,(2\,n+1)\,t}}{1 - e^{i\,t}} \\ &= -\frac{e^{i\,(n+1)\,t} - e^{-i\,n\,t}}{1 - e^{i\,t}} \\ &= \frac{(e^{-i\,n\,t} - e^{i\,(n+1)\,t})(1 - e^{-i\,t})}{(1 - e^{i\,t})(1 - e^{-i\,t})} \\ &= \frac{(e^{i\,n\,t} + e^{-i\,n\,t}) - (e^{i\,(n+1)\,t} + e^{-i\,(n+1)\,t})}{|1 - e^{-i\,t}|^2}. \end{aligned}$$

We have

$$|1 - e^{i\,t}|^2 = (1 - \cos t)^2 + \sin^2 t = 2\,(1 - \cos t) = 4\,\sin^2\frac{t}{2}$$

and

$$\begin{aligned} (e^{i\,n\,t} + e^{-i\,n\,t}) - (e^{i\,(n+1)\,t} + e^{-i\,(n+1)\,t}) &= 2\Big(\cos(n\,t) - \cos\left((n+1)\,t\right)\Big) \\ &= 4\,\sin\left(\left(n + \frac{1}{n}\right)t\right)\,\sin\frac{t}{2}, \end{aligned}$$

where we have used the prosthaphaeresis formula.[a] This leads immediately to the conclusion.

[a]The prosthaphaeresis formula for the cosine is

$$\cos p - \cos q = -2 \sin \frac{p+q}{2} \sin \frac{p-q}{2}.$$

We have the following simple but important consequence.

Corollary 9.2 *Let $f \in L^2([-\pi,\pi];\mathbb{C})$. Then, for all $n \in \mathbb{N}$ and for all $t \in [-\pi,\pi]$,*

$$S_n(f)(t) = \int_{[-\pi,\pi]} K_n(t-\tau)\, f(\tau)\, \mathrm{d}\tau,$$

with

$$K_n(t) = \begin{cases} \dfrac{\cos(n\,t) - \cos\left((n+1)\,t\right)}{4\,\pi\,\sin^2\frac{t}{2}} = \dfrac{\sin\left(\left(n+\frac{1}{2}\right)t\right)}{2\,\pi\,\sin\frac{t}{2}} & \text{if } t \neq 0 \\ \dfrac{2\,n+1}{2\,\pi} & \text{if } t = 0 \end{cases}$$

We want to establish whether

$$\forall\, f \in L^2([-\pi,\pi];\mathbb{C}), \qquad \lim_{n\to\infty} S_n(f) = f.$$

Before studying the convergence in $L^2([-\pi,\pi];\mathbb{C})$, we consider the pointwise and uniform convergence of the sequence $(S_n(f))_{n\in\mathbb{N}}$ to f. We observe that for all $n \in \mathbb{N}$ we have $S_n(f) \in C([-\pi,\pi];\mathbb{C})$ and $S_n(f)(-\pi) = S_n(f)(\pi)$. Thus, a necessary condition for the uniform convergence is that

$$f \in C([-\pi,\pi];\mathbb{C}), \quad f(-\pi) = f(\pi). \tag{9.6}$$

Give f satisfying (9.6), we extend it to a periodic function with period $2\,\pi$ in $\mathbb{R}$, by defining

$$\tilde{f} : \mathbb{R} \longrightarrow \mathbb{C}$$

such that

$$\tilde{f}(2\,k\,\pi + t) = f(t), \quad k \in \mathbb{Z},\ t \in [-\pi,\pi].$$

We observe that $\tilde{f} \in C(\mathbb{R};\mathbb{C})$ and $\tilde{f}((2\,k+1)\,\pi) = f(-\pi) = f(\pi)$ for all $k \in \mathbb{Z}$.

Example Show that if f satisfies (9.6), then $\tilde{f}$ is uniformly continuous.

Remark 9.1 The functions

$$\tilde{K}_n(t) = \frac{1}{2\,\pi} \sum_{k=-n}^{n} e^{i\,k\,t}, \quad t \in \mathbb{R},$$

are periodic with period $2\,\pi$. In the following, even if $t \notin [-\pi,\pi]$, we will write $K_n(t)$ instead of $\tilde{K}_n(t)$.

Lemma 9.2 *Let $\tilde{f} : \mathbb{R} \longrightarrow \mathbb{R}$ be summable in $[-\pi, \pi]$ and such that $f(t) = f(t + 2\pi)$ almost everywhere in $\mathbb{R}$. Then, f is summable in any bounded measurable subset of $\mathbb{R}$. Moreover, for all $a \in \mathbb{R}$*

$$\int_{[a-\pi,a+\pi]} f(t)\,\mathrm{d}t = \int_{[-\pi,\pi]} f(t)\,\mathrm{d}t.$$

Proof A simple application of the change of variable shows that f is summable in $[(2k-1)\pi, (2k+1)\pi]$ for any $k \in \mathbb{Z}$ and

$$\int_{[(2k-1)\pi,(2k+1)\pi]} f(t)\,\mathrm{d}t = \int_{[-\pi,\pi]} f(t)\,\mathrm{d}t.$$

This implies that f is summable in any bounded measurable subset of $\mathbb{R}$, because each of these sets is contained in a finite number of intervals in the form $[(2k-1)\pi, (2k+1)\pi]$ with $k \in \mathbb{Z}$. We introduce the following notation. If $a, b \in \mathbb{R}$, we define

$$\int_a^b f(t)\,\mathrm{d}t = \begin{cases} \displaystyle\int_{[a,b]} f(t)\,\mathrm{d}t & \text{if } a \leq b \\ \displaystyle -\int_{[b,a]} f(t)\,\mathrm{d}t & \text{if } b \leq a \end{cases}$$

It is straightforward to check that, if a, b, c are real numbers, we have

$$\int_a^c f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t + \int_b^c f(t)\,\mathrm{d}t.$$

We deduce that

$$\int_{[a-\pi,a+\pi]} f(t)\,\mathrm{d}t = \int_{a-\pi}^{-\pi} f(t)\,\mathrm{d}t + \int_{-\pi}^{\pi} f(t)\,\mathrm{d}t + \int_{\pi}^{a+\pi} f(t)\,\mathrm{d}t.$$

Using the change of variable theorem,

$$\begin{aligned} \int_{a-\pi}^{-\pi} f(t)\,\mathrm{d}t + \int_{\pi}^{a+\pi} f(t)\,\mathrm{d}t &= \int_{a-\pi}^{-\pi} f(t+2\pi)\,\mathrm{d}t + \int_{\pi}^{a+\pi} f(t)\,\mathrm{d}t \\ &= \int_{a+\pi}^{\pi} f(t)\,\mathrm{d}t + \int_{\pi}^{a+\pi} f(t)\,\mathrm{d}t \\ &= \int_{a+\pi}^{a+\pi} f(t)\,\mathrm{d}t = 0. \end{aligned}$$

Lemma 9.3 *Let $n \in \mathbb{N}$. Then,*

$$\int_{[-\pi,\pi]} K_n(t)\,\mathrm{d}t = 1.$$

Proof We have

$$\int_{[-\pi,\pi]} K_n(t)\,\mathrm{d}t = \sum_{k=-n}^{n} \int_{[-\pi,\pi]} e_k(t)\,\frac{1}{\sqrt{2\pi}}\,\mathrm{d}t = \sum_{k=-n}^{n} \langle e_k, e_0\rangle = 1,$$

where $\langle \cdot, \cdot, \rangle$ denote the inner product in $L^2([-\pi, \pi]; \mathbb{C})$ and $(e_k)_{k\in\mathbb{Z}}$ is orthonormal.

The key result for our proof of the pointwise convergence is the following.

Theorem 9.4 [Riemann-Lebesgue Lemma]
Let $-\infty \le a < b \le +\infty$ and let $f : (a,b) \longrightarrow \mathbb{C}$ be summable. Then,

$$\lim_{|\lambda|\to\infty} \int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t = 0. \tag{9.7}$$

Proof We have We notice that for every $\lambda \in \mathbb{R}$, the function $t \mapsto \sin(\lambda\, t)\, f(t)$ is summable in (a,b), since

$$|\sin(\lambda\, t)\, f(t)| \le |f(t)|\, \mathrm{d}t \qquad t \in (a,b).$$

We assume first that f is of class C^1 with compact support, i.e., there exist c and d such that $a < c < d < b$ and $f(t) = 0$ if $t \in (a,b) \setminus (c,d)$. Then, integrating by parts yields

$$\begin{aligned}\int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t &= \int_{(c,d)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t \\ &= \left[-\frac{\cos(\lambda\, t)}{\lambda} f(t)\right]_{t=c}^{t=d} + \frac{1}{\lambda}\int_c^d \cos(\lambda\, t)\, f'(t)\, \mathrm{d}t \\ &= \frac{1}{\lambda}\int_c^d \cos(\lambda\, t)\, f'(t)\, \mathrm{d}t,\end{aligned}$$

so that

$$\left|\int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t\right| \le \frac{1}{\lambda}\int_c^d |f'(t)|\, \mathrm{d}t,$$

which tends to 0 as $|\lambda|$ goes to ∞.

In order to treat the general case, we use the fact that the set of equivalence classes containing a function g of class C^1 with compact support is dense in the Banach space $L^1((a,b);\mathbb{C})$. Hence, if $f \in L^1((a,b);\mathbb{C})$ and $\varepsilon > 0$, there exists g of class C^1 with compact support such that

$$\int_{(a,b)} |f(t) - g(t)|\, \mathrm{d}t < \frac{\varepsilon}{2}, \tag{9.8}$$

see Theorem 5.5. Let $\varepsilon > 0$. Pick g as before, satisfying (9.8). From the first part of the proof, there exists $T_\varepsilon > 0$ such that, if $|\lambda| > T_\varepsilon$,

$$\left|\int_{(a,b)} \sin(\lambda\, t)\, g(t)\, \mathrm{d}t\right| < \frac{\varepsilon}{2}.$$

We deduce that if $|\lambda| > T_\varepsilon$

$$\begin{aligned}\left|\int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t\right| &= \left|\int_{(a,b)} \sin(\lambda\, t)\, (f(t) - g(t))\, \mathrm{d}t + \int_{(a,b)} \sin(\lambda\, t)\, g(t)\, \mathrm{d}t\right| \\ &\le \int_{(a,b)} |f(t) - g(t)|\, \mathrm{d}t + \left|\int_{(a,b)} \sin(\lambda\, t)\, g(t)\, \mathrm{d}t\right| \\ &\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

In the study of uniform convergence, we will employ the following refinement of the Riemann-Lebesgue Lemma.

Theorem 9.5 *Let $-\infty \le a < b \le +\infty$. Then, (9.7) holds uniformly for f in a compact subset of*

$L^1((a,b);\mathbb{C})$.

Proof Let K be a compact subset of $L^1((a,b);\mathbb{C})$. Let $\varepsilon > 0$. We have to show that there exists $T_\varepsilon > 0$ such that, if $|\lambda| > T_\varepsilon$,

$$\forall\, f \in K, \qquad \left| \int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t \right| < \varepsilon. \tag{9.9}$$

Let $f \in K$, let $g \in L^1((a,b);\mathbb{C})$ and

$$\|f - g\|_{L^1((a,b);\mathbb{C})} < \frac{\varepsilon}{2}.$$

From Theorem 9.4, there exists $T(f,\varepsilon) > 0$ such that, if $|\lambda| > T(f,\varepsilon)$,

$$\left| \int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t \right| < \frac{\varepsilon}{2}.$$

If $g \in \mathcal{B}_{\frac{\varepsilon}{2}}(f)$, the open ball centred at f with radius $\varepsilon/2$ in $L^1((a,b);\mathbb{C})$, and $|\lambda| > T(f,\varepsilon)$, then

$$\left| \int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t \right| \leq \int_{(a,b)} |g(t) - f(t)|\, \mathrm{d}t + \left| \int_{(a,b)} \sin(\lambda\, t)\, f(t)\, \mathrm{d}t \right| < \varepsilon.$$

Obviously $K \subset \bigcup_{f \in K} \mathcal{B}_{\frac{\varepsilon}{2}}(f)$. Since K is compact, there exist $f_1, \ldots, f_N$ in K such that $K \subseteq \bigcup_{j=1}^{N} \mathcal{B}_{\frac{\varepsilon}{2}}(f_j)$. We conclude that if $|\lambda| > \max_{j \in \{1,\ldots,N\}} T(f_j,\varepsilon)$, then (9.9) holds.

Now we are able to state and prove the first result on pointwise convergence.

Theorem 9.6 *Let f satisfy (9.6). Let $t \in [-\pi,\pi]$. Let the function*

$$\begin{aligned} h_t : [-\pi,\pi] &\longrightarrow \mathbb{C} \\ z &\mapsto \begin{cases} \frac{\tilde{f}(t-z)-\tilde{f}(t)}{\sin\frac{z}{2}} & \text{if } z \neq 0 \\ 0 & \text{if } z = 0 \end{cases} \end{aligned}$$

be summable in $[-\pi,\pi]$. Then,

$$\lim_{n\to\infty} S_n(f)(t) = f(t).$$

Proof Employing the change of variable theorem and Lemma 9.2, together with Remark 9.1, we have

$$\begin{aligned} S_n(f)(t) &= \int_{[-\pi,\pi]} K_n(t-\tau)\, \tilde{f}(\tau)\, \mathrm{d}\tau \\ &= \int_{[t-\pi,t+\pi]} K_n(z)\, \tilde{f}(t-z)\, \mathrm{d}z = \int_{[-\pi,\pi]} K_n(z)\, \tilde{f}(t-z)\, \mathrm{d}z. \end{aligned}$$

From Lemma 9.3, we deduce

$$\begin{aligned} S_n(t)(t) - f(t) &= \int_{[-\pi,\pi]} K_n(z)\,(\tilde{f}(t-z) - \tilde{f}(t))\,\mathrm{d}z \\ &= \frac{1}{2\pi}\int_{[-\pi,\pi]} \sin\left(\left(n+\frac{1}{2}\right)z\right) h_t(z)\,\mathrm{d}z, \end{aligned}$$

and the conclusion follows from Theorem 9.4.

We have the following consequence.

Corollary 9.3 *Let f satisfy (9.6). Let $t \in [-\pi,\pi]$. Suppose there exist $\alpha > 0$ and $M > 0$ such that*

$$\forall\, y \in \mathbb{R}, \qquad |\tilde{f}(\tau) - \tilde{f}(t)| \leq M\,|\tau - t|^{\alpha}.$$

Then,

$$\lim_{n\to\infty} S_n(f)(t) = f(t).$$

Proof Since

$$\lim_{\tau\to 0} \frac{\sin\tau}{\tau} = 1,$$

there exists $\delta > 0$ such that

$$\forall\, z \in [-\pi,\pi], \quad \left|\sin\frac{z}{2}\right|.$$

Therefore, if $z \in [-\pi,\pi] \setminus \{0\}$,

$$|h_t(z)| \leq \frac{M}{\delta}\,|z|^{\alpha-1}$$

and we can apply Theorem 9.6.

Definition 9.1 *Let I be an interval of $\mathbb{R}$. Let $\alpha \in (0,1)$ and $f : I \longrightarrow \mathbb{C}$. We shall write $f \in C^{\alpha}(I;\mathbb{C})$ if there exists $L \geq 0$ such that*

$$|f(t) - f(\tau)| \leq L\,|t-\tau|^{\alpha}. \tag{9.10}$$

In this case, we say that f is α-Hölder continuous. If (9.10) is satisfied with $\alpha = 1$, we say that f is Lipschitz *or* Lipschitz continuous*, and we write $f \in \mathrm{Lip}(I;\mathbb{C})$.*

Let I be an interval of $\mathbb{R}$. Show the following statements:

(i) if f is differentiable in I and f' is bounded, then f is Lipschitz;

(ii) if f satisfies (9.10) for some $\alpha > 1$, then f is a constant function (Hint: compute the derivative f');

(iii) if $f \in C^{\alpha}([-\pi,\pi])$ for some $\alpha \in (0,1)$ and $f(-\pi) = f(\pi)$, then $\tilde{f} \in C^{\alpha}(\mathbb{R};\mathbb{C})$;

(iv) if $f \in \mathrm{Lip}([-\pi,\pi];\mathbb{C})$ and $f(-\pi) = f(\pi)$, then $\tilde{f} \in \mathrm{Lip}(\mathbb{R};\mathbb{C})$;

(v) if I is a bounded interval and $0 < \alpha \leq \beta < 1$, then $\mathrm{Lip}(I;\mathbb{C}) \subseteq C^{\beta}(I;\mathbb{C}) \subseteq C^{\alpha}(I;\mathbb{C})$.

Theorem 9.7 *Let $f \in C^\alpha([-\pi,\pi];\mathbb{C})$ for some $\alpha \in (0,1)$, and let $f(-\pi) = f(\pi)$, Then,*

$$\lim_{n\to\infty} S_n(f)(t) = f(t)$$

uniformly in $[-\pi,\pi]$.

Proof From *(iii)* of the previous exercise, $\tilde{f} \in C^\alpha(\mathbb{R};\mathbb{C})$. From Corollary 9.3,

$$\lim_{n\to\infty} S_n(f)(t) = f(t) \quad \forall\, t \in [-\pi,\pi].$$

To reach the conclusion, we can show, in force of Theorem 9.5, that the mapping $t \to h_t$ is continuous from $[-\pi,\pi]$ to $L^1([-\pi,\pi])$, since the image of a compact set through a continuous mapping is compact. We take $t,\tau \in [-\pi,\pi]$, and try to estimate $\|h_t - h_\tau\|_{L^1([-\pi,\pi];\mathbb{C})}$. Let $z \in [-\pi,\pi]\setminus\{0\}$. Then,

$$|h_t(z) - h_\tau(z)| \le \frac{2\,L\,|t-\tau|^\alpha}{\delta\,|z|}$$

if

$$\forall\, u,v \in \mathbb{R}, \qquad |\tilde{f}(u) = \tilde{f}(v)| \le L\,|u-v|^\alpha.$$

On the other hand,

$$|h_t(z) - h_\tau(z)| \le |h_t(z)| + |h_t(z)| \le \frac{2\,L}{\delta}\,|z|^{\alpha-1},$$

so that

$$|h_t(z) - h_\tau(z)| \le \frac{2\,L}{\delta}\,\frac{|t-\tau|^{\frac{\alpha}{2}}}{|z|^{1-\frac{\alpha}{2}}},$$

which proves the continuity of $t \mapsto h_t$.

The following result is a simple corollary.

Theorem 9.8 *Let $(x_k)_{k\in\mathbb{Z}}$ be as in (9.4). Then, $(x_k)_{k\in\mathbb{Z}}$ is a maximal orthonormal system in the Hilbert space $L^2([-\pi,\pi];\mathbb{C})$.*

Proof Let $\boldsymbol{V} = \text{span}(\{x_k\,|\,k\in\mathbb{Z}\})$. By Theorem 9.7, since the uniform convergence implies the convergence in $L^2([-\pi,\pi];\mathbb{C})$,

$$\mathcal{D}((-\pi,\pi);\mathbb{C}) \subseteq \overline{\boldsymbol{V}}.$$

Since $\mathcal{D}((-\pi,\pi);\mathbb{C})$ is dense in $L^2([-\pi,\pi];\mathbb{C})$, we conclude that

$$L^2([-\pi,\pi];\mathbb{C}) = \overline{\mathcal{D}((-\pi,\pi);\mathbb{C})} \subseteq \overline{\boldsymbol{V}} \subseteq L^2([-\pi,\pi];\mathbb{C}),$$

which implies $\overline{\boldsymbol{V}} = L^2([-\pi,\pi];\mathbb{C})$, and the conclusion follows from Theorem 2.34.

Lemma 9.4 *Let $f \in C^1([-\pi,\pi];\mathbb{C})$, with $f(-\pi) = f(\pi)$. Let $(x_k)_{k\in\mathbb{Z}}$ be as in (9.4). Then,*

$$f' = \sum_{k\in\mathbb{Z}} \langle f', x_k\rangle\, x_k = \sum_{k\in\mathbb{Z}} i\,k\,\langle f, x_k\rangle\, x_k = \sum_{k\in\mathbb{Z}} \langle f, x_k\rangle\, x_k', \tag{9.11}$$

with convergence in $L^2([-\pi,\pi];\mathbb{C})$. *Moreover,*

$$\sum_{k\in\mathbb{Z}} k^2\,|\langle f, x_k\rangle|^2 = \|f'\|^2_{L^2([-\pi,\pi];\mathbb{C})}. \tag{9.12}$$

Proof The first identity follows from Theorem 9.8. Moreover, for $k \in \mathbb{Z}$,

$$\langle f', x_k\rangle = \frac{1}{\sqrt{2\,\pi}} \int_{[-\pi,\pi]} e^{-i\,k\,t}\, f'(t)\,\mathrm{d}t = \frac{1}{\sqrt{2\,\pi}} \left(\left[e^{-i\,k\,t}\, f(t)\right]_{t=-\pi}^{t=\pi} + i\,k \int_{[-\pi,\pi]} e^{-i\,k\,t}\, f(t)\,\mathrm{d}t \right) = i\,k\,\langle f, x_k\rangle.$$

Finally, we observe that $i\,k\,x_k = x_k'$.

Iterating Lemma 9.4, we obtain the following.

Theorem 9.9 *Let* $m \in \mathbb{N}$, *and let* $f \in C^m([-\pi,\pi];\mathbb{C})$, *with* $f^{(k)}(-\pi) = f^{(k)}(\pi)$ *for each* $k \in \{0, 1, \ldots, m-1\}$. *Let* $(x_k)_{k\in\mathbb{Z}}$ *be as in (9.4). Then,*

$$f^{(m)} = \sum_{k\in\mathbb{Z}} \langle f^{(m)}, x_k\rangle\, x_k = \sum_{k\in\mathbb{Z}} (i\,k)^m\, \langle f, x_k\rangle\, x_k = \sum_{k\in\mathbb{Z}} \langle f, x_k\rangle\, x_k^{(m)}, \tag{9.13}$$

with convergence in $L^2([-\pi,\pi];\mathbb{C})$. *Moreover,*

$$\sum_{k\in\mathbb{Z}} k^{2\,m}\,|\langle f, x_k\rangle|^2 = \|f^{(m)}\|^2_{L^2([-\pi,\pi];\mathbb{C})}.$$

If f satisfies (9.6), the corresponding Fourier series $(S_k(f)(t))_{k\in\mathbb{N}}$ (see (9.5)) does not converge in general, not even pointwise. For a discussion on this, see W. Rudin, Real and complex analysis, Mc Graw-Hill, Chapter 5. Nevertheless, by suitably modifying $S_k(f)$, it is possible to construct a series uniformly converting in $[-\pi,\pi]$ for every f that satisfies (9.6). This result is due to L. Fejer (L. Fejer, "Untersuchungen über Fouriersche Reihen", Math. Annales vol 58 (1904), 51-69.). Fejer's idea is to replace $(S_k(f))_{k\in\mathbb{N}}$ with

$$\left(\frac{1}{k+1} \sum_{j=0}^{n} S_j(f) \right)_{k\in\mathbb{N}}.$$

We begin with a simple computation.

Lemma 9.5 *Let* f *satisfy (9.6). If* $t \in [-\pi,\pi]$ *and* $k \in \mathbb{N}$, *we define*

$$F_k(f)(t) = \frac{1}{k+1} \sum_{j=0}^{k} S_j(f)(t).$$

Then,

$$F_k(f)(t) = \int_{[-\pi,\pi]} D_k(t-\tau)\, f(\tau)\,\mathrm{d}\tau,$$

where

$$D_k(t) = \begin{cases} \frac{1}{k+1} \frac{\sin^2\left(\frac{(k+1)t}{2}\right)}{2\pi \sin^2 \frac{t}{2}} & \text{if } t \neq 0 \\ \frac{k+1}{2\pi} & \text{if } t = 0. \end{cases}$$

Proof From Corollary 9.2,

$$F_k(f)(t) = \frac{1}{k+1} \sum_{j=0}^{k} S_j(f)(t) = \int_{[-\pi,\pi]} \frac{1}{k+1} \sum_{j=0}^{k} K_j(t-\tau) f(\tau) \, \mathrm{d}\tau,$$

and

$$\frac{1}{k+1} \sum_{j=0}^{k} K_j(0) = \frac{1}{k+1} \sum_{j=0}^{k} \frac{2j+1}{2\pi} = \frac{1}{2\pi(k+1)} (k(k+1)+k+1) = \frac{k+1}{2\pi}.$$

Moreover, if $z \in [-\pi, \pi] \setminus \{0\}$,

$$\begin{aligned} \frac{1}{k+1} \sum_{j=0}^{k} K_j(z) &= \frac{1}{k+1} \sum_{j=0}^{k} \frac{\cos(jt) - \cos((j+1)t)}{2\pi \sin^2 \frac{t}{2}} \\ &= \frac{1}{k+1} \frac{1 - \cos((j+1)t)}{4\pi \sin^2 \frac{t}{2}} \\ &= \frac{1}{k+1} \frac{\sin^2\left(\frac{(k+1)t}{2}\right)}{2\pi \sin^2 \frac{t}{2}}. \end{aligned}$$

Theorem 9.10 *Let f satisfy (9.6). Then,*

$$\lim_{k \to \infty} F_k(f)(t) = f(t)$$

uniformly for $t \in [-\pi, \pi]$.

Proof Applying Lemma 9.2, we have, for $t \in [-\pi, \pi]$

$$F_k(f)(t) = \int_{[-\pi,\pi]} D_k(t-\tau) f(\tau) \, \mathrm{d}\tau = \int_{[-\pi,\pi]} D_k(\tau) \tilde{f}(t-\tau) \, \mathrm{d}\tau.$$

From

$$\int_{[-\pi,\pi]} D_k(\tau) \, \mathrm{d}\tau = \frac{1}{k+1} \sum_{j=0}^{k} \int_{[-\pi,\pi]} K_j(\tau) \, \mathrm{d}\tau = 1,$$

we deduce

$$F_k(f)(t) - f(t) = \int_{[-\pi,\pi]} D_k(t) \left(\tilde{f}(t-\tau) - \tilde{f}(t)\right) \mathrm{d}\tau.$$

Let $\varepsilon > 0$. From a previous exercise, $\tilde{f}$ is uniformly continuous. It follows that there exists $\delta > 0$ such that

$$\left|\tilde{f}(t-\tau) - \tilde{f}(t)\right| \leq \varepsilon.$$

if $|\tau| \leq \delta$, for any $t \in \mathbb{R}$. Suppose $\delta \leq \pi$. We find

$$\begin{aligned} |F_k(f)(t) - f(t)| &\leq \varepsilon \int_{[-\delta,\delta]} D_k(\tau)\,\mathrm{d}\tau + 2 \int_{[-\pi,\pi]\setminus[-\delta,\delta]} D_k(\tau)\,\mathrm{d}\tau \max_{[-\pi,\pi]} |f| \\ &\leq \varepsilon + 2 \int_{[-\pi,\pi]\setminus[-\delta,\delta]} D_k(\tau)\,\mathrm{d}\tau \max_{[-\pi,\pi]} |f|. \end{aligned}$$

Next, we observe that

$$\int_{[-\pi,\pi]\setminus[-\delta,\delta]} D_k(\tau)\,\mathrm{d}\tau \leq \frac{2\,\pi}{(k+1)\,\sin^2(2\,\delta)},$$

and we choose k sufficiently large so that

$$\frac{2\,\pi}{(k+1)\,\sin^2(2\,\delta)} \leq \varepsilon.$$

We conclude that, if k is sufficiently large,

$$\forall\, t \in [-\pi,\pi], \quad |F_k(f)(t) - f(t)| \leq 2\,\varepsilon.$$

Remark 9.2 We have

$$\begin{aligned} F_k(f) &= \frac{1}{k+1} \sum_{j=0}^{k} S_j(f) = \frac{1}{k+1} \sum_{j=0}^{k} \sum_{\ell=-j}^{j} \langle f, e_j\rangle\, e_j \\ &= \frac{1}{k+1} \sum_{\substack{\ell\in\mathbb{Z} \\ |\ell|\leq k}} \Big(\sum_{j=|\ell|}^{k} 1 \Big) \langle f, e_j\rangle\, e_j \\ &= \sum_{\ell=-k}^{k} \left(1 - \frac{|\ell|}{k+1}\right) \langle f, e_j\rangle\, e_j. \end{aligned}$$

9.2 Other orthonormal systems

Let $a \in \mathbb{R}$, and let $T > 0$. For $j \in \mathbb{Z}$, we define

$$e_j(t) = \frac{1}{\sqrt{2\,T}}\, e^{\frac{i\,j\,\pi\,(t-a)}{T}}, \quad t \in [a-T, a+T].$$

Show that $(e_j)_{j\in\mathbb{Z}}$ is a maximal orthonormal system in $L^2([a-T, a+T])$.

Theorem 9.11 *For $j \in \mathbb{N}^*$ and $t \in [0,\pi]$, we define*

$$w_j(t) = \sqrt{\frac{2}{\pi}}\, \sin(j\,t).$$

Then, $(w_j)_{j\in\mathbb{N}^*}$ is a maximal orthonormal system in $L^2([0,\pi])$.

Proof We begin by showing that $(w_j)_{j\in\mathbb{N}^*}$ is an orthonormal system, First of all, denoting by $\langle\cdot,\cdot\rangle$ the scalar product in $L^2([0,\pi])$, we have

$$\langle w_j, w_j\rangle = \frac{2}{\pi}\int_0^\pi \sin^2(j\,t)\,\mathrm{d}t = \frac{2}{\pi}\int_0^\pi \frac{1-\cos(2\,j\,t)}{2}\,\mathrm{d}t = 1.$$

Now, if $j,k\in\mathbb{N}^*$, using prosthaphaeresis formulae, we have

$$\langle w_j, w_k\rangle = \frac{2}{\pi}\int_0^\pi \int_0^\pi \sin(j\,t)\,\sin(k\,t)\,\mathrm{d}t = \frac{1}{\pi}\int_0^\pi \Big(\cos\left((j-k)\,t\right) - \cos\left((j+k)\,t\right)\Big)\,\mathrm{d}t = 0$$

if $j\neq k$. Next, we show that it is maximal. Let $g\in L^2([-\pi,\pi])$. We suppose that g is odd almost everywhere, i.e., $g(-t) = -g(t)$ for almost all t. Let $(e)j)_{j\in\mathbb{Z}}$ be as in (9.4). Then,

$$\begin{aligned}\langle g, e_{-j}\rangle &= \frac{1}{\sqrt{2\,\pi}}\int_{[-\pi,\pi]} g(t)\,e^{i\,j\,t}\,\mathrm{d}t\\ &= \frac{1}{\sqrt{2\,\pi}}\int_{[-\pi,\pi]} g(-t)\,e^{-i\,j\,t}\,\mathrm{d}t\\ &= -\frac{1}{\sqrt{2\,\pi}}\int_{[-\pi,\pi]} g(t)\,e^{-i\,j\,t}\,\mathrm{d}t = -\langle g, e_j\rangle.\end{aligned}$$

It follows that

$$\langle g, e_j\rangle\,e_j(t) + \langle g, e_{-j}\rangle\,e_{-j}(t) = \langle g, e_j\rangle\,(e_j(t) - e_{-j}(t)) = \sqrt{\frac{2}{\pi}}\,i\,\langle g, e_j\rangle). \tag{9.14}$$

We deduce that

$$\sum_{j\in\mathbb{N}^*}\sqrt{\frac{2}{\pi}}\,i\,\langle g, e_j\rangle\,\sin(j\,\cdot) = g$$

in $L^2([-\pi,\pi])$. Let $f\in L^2([0,\pi])$. We extend f to an odd element g of $L^2([-\pi,\pi])$ by setting

$$g(t) = \begin{cases} f(t) & \text{if } t\in[0,\pi]\\ -f(-t) & \text{if } t\in[-\pi,0).\end{cases}$$

By (9.14), we have $g\in\overline{\boldsymbol{W}}$, with

$$\boldsymbol{W} = \operatorname{span}\{\sin(j\,\cdot)\,|\,j\in\mathbb{N}^*\}$$

in $L^2([-\pi,\pi])$. This implies $f\in\overline{\boldsymbol{V}}$, where

$$\boldsymbol{V} = \operatorname{span}\{\sin(j\,\cdot)\,|\,j\in\mathbb{N}^*\}$$

in $L^2([0,\pi])$. The conclusion follows from Theorem **??**.

Let, for $j\in\mathbb{N}$,

$$v_j(t) = \begin{cases} \frac{1}{\sqrt{\pi}} & \text{if } j=0\\ \sqrt{\frac{2}{\pi}}\,\cos(j\,t) & \text{if } j\in\mathbb{N}^*\end{cases}$$

Show that $(v_j)_{j\in\mathbb{N}}$ is a maximal orthonormal system in $L^2([0,\pi])$. Hint: extend $f\in L^2([0,\pi])$ to an even element of $L^2([-\pi,\pi])$.

9.3 Fourier series in several real variables

We introduce the following notation. If $n \in \mathbb{N}^*$, we define

$$Q_n \stackrel{\text{def}}{=} [-\pi,\pi]^n = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid x_1, x_2, \ldots, x_n \in [-\pi,\pi]\}. \tag{9.15}$$

We recall the notation (9.4). We shall indicate by $\boldsymbol{j} = (j_1, j_2, \ldots, j_n)$ the generic element of $\mathbb{Z}^n$ and define

$$e^{\boldsymbol{j}}(\boldsymbol{t}) = e^{\boldsymbol{j}}(t_1, t_2, \ldots, t_n) = v_{j_1}(t_1)\, v_{j_2}(t_2) \cdots v_{j_n}(t_n) = \frac{1}{(2\pi)^{\frac{n}{2}}}\, e^{i\, j_1\, t_1}\, e^{i\, j_2\, t_2} \cdots e^{i\, j_n\, t_n}. \tag{9.16}$$

A simple consequence of Fubini's theorem is that $(e^{\boldsymbol{j}})_{\boldsymbol{j}\in\mathbb{Z}^n}$ is an orthonormal system in the Hilbert space $L^2(Q_n;\mathbb{C})$. It is of interest to us to know if it is maximal. To this end, we generalize Theorem 9.10. If $f \in L^2(Q_n;\mathbb{C})$, we define, for $\boldsymbol{t} = (t_1, t_2, \ldots, t_n) \in Q_n$,

$$\begin{aligned} F_k(f)(\boldsymbol{t}) &= F_k(f)(t_1, t_2, \ldots, t_n) \\ &\stackrel{\text{def}}{=} \int_{Q_n} D_k(t_1-\tau_1) \cdots D_k(t_n-\tau_n)\, f(\tau_1, \tau_2, \ldots, \tau_n)\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2 \ \ldots\ \mathrm{d}\tau_n. \end{aligned} \tag{9.17}$$

Employing Remark 9.2 and Fubini's theorem, we observe that

$$\begin{aligned} F_k(f)(t_1, t_2, \ldots, t_n) &= \int_{[-\pi,\pi]} D_k(t_1-\tau_1) \left(\int_{Q_{n-1}} D_k(t_2-\tau_2) \cdots D_k(t_n-\tau_n)\, f(\tau_1, \tau_2, \ldots, \tau_n)\, \mathrm{d}\tau_2 \ \ldots\ \mathrm{d}\tau_n \right) \mathrm{d}\tau_1 \\ &= \sum_{j_1=-k}^{k} \left(1 - \frac{|j_1|}{k+1}\right) \int_{[-\pi,\pi]} \left(\int_{Q_{n-1}} D_k(t_2-\tau_2) \cdots D_k(t_n-\tau_n)\, f(\tau_1, \tau_2, \ldots, \tau_n)\, \mathrm{d}\tau_2 \ \ldots\ \mathrm{d}\tau_n \right) \overline{v_{j_1}(\tau_1)}\, \mathrm{d} \\ &= \sum_{j_1=-k}^{k} \sum_{j_2=-k}^{k} \left(1 - \frac{|j_1|}{k+1}\right)\left(1 - \frac{|j_2|}{k+1}\right) \int_{[-\pi,\pi]^2} \left(\int_{Q_{n-2}} D_k(t_3-\tau_3) \cdots D_k(t_n-\tau_n)\, f(\tau_1, \tau_2, \ldots, \tau_n)\, \mathrm{d}\tau \right. \\ &\vdots \\ &= \sum_{j_1=-k}^{k} \sum_{j_2=-k}^{k} \cdots \sum_{j_n=-k}^{k} \left(1 - \frac{|j_1|}{k+1}\right)\left(1 - \frac{|j_2|}{k+1}\right) \cdots \left(1 - \frac{|j_n|}{k+1}\right) \\ &\quad \times \int_{Q_n} f(\tau_1, \tau_2, \ldots, \tau_n)\, \overline{v_{j_1}(\tau_1)\, v_{j_2}(\tau_2) \cdots v_{j_n}(\tau_n)}\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2 \ \ldots\ \mathrm{d}\tau_n\, v_{j_1}(\tau_1)\, v_{j_2}(\tau_2) \cdots v_{j_n}(\tau_n). \end{aligned}$$

This implies that $F_k(f) \in \operatorname{span}\{e^{\boldsymbol{j}} \mid \boldsymbol{j} \in \mathbb{Z}^n\}$ in the Hilbert space $L^2(Q_n;\mathbb{C})$. Now we introduce the n-dimensional generalization of condition (9.6). We will write that f satisfies condition H_n if

$$\begin{aligned} &f \in C(Q_n);\ \forall\, j \in \{1, \ldots, n\},\ \forall\, (t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n) \in Q_{n-1} \\ &\quad f(t_1, \ldots, t_{j-1}, -\pi, t_{j+1}, \ldots, t_n) = f(t_1, \ldots, t_{j-1}, \pi, t_{j+1}, \ldots, t_n). \end{aligned}$$

If f satisfies this condition, we introduce its periodic extension $\tilde{f}$ as the function $\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{C}$ satisfying

$$\forall\, (k_1, \ldots, k_n) \in \mathbb{Z}^n,\ \forall\, (t_1, \ldots, t_n) \in Q_n, \quad \tilde{f}(2\,k_1\,\pi + t_1, \ldots, 2\,k_n\,\pi + t_n) = f(t_1, \ldots, t_n).$$

Let f satisfy the condition H_n. Check that $\tilde{f}$ is well defined and it is uniformly continuous in $\mathbb{R}^n$. Check that for all $j \in \{1, \ldots, n\}$ and for all $(t_1, \ldots, t_{j-1}, t_{j+1}, \ldots, t_n) \in Q_{n-1}$, the mapping

$$t_j \mapsto f(t_1, \ldots, t_{j-1}, t_j, t_{j+1}, \ldots, t_n)$$

is periodic with period 2π in $\mathbb{R}$.

We are now ready to extend Theorem 9.10.

Theorem 9.12 *Let f satisfy condition H_n. Then, $\lim_{k\to\infty} F_k(f)(t) = f(t)$ uniformly for $t \in Q_n$.*

Proof Applying Lemma 9.2 componentwise, and using the identity

$$\int_{Q_n} D_k(\tau_1)\cdots D_k(\tau_n)\,\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\,tau_n = 1,$$

we obtain

$$F_k(f)(t_1,\ldots,t_n) - f(t_1,\ldots,t_n) = \int_{Q_n} D_k(\tau_1)\cdots D_k(\tau_n)\,(\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{f}(t_1,\ldots,t_n))\,\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n.$$

We take ε and η. Using the fact that $\tilde{f}$ is uniformly continuous (see previous exercise), we see that there exists $\delta \in (0,\pi]$ such that

$$\forall\,(\tau_1,\ldots,\tau_n) \in [-\delta,\delta]^n, \quad |\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{f}(t_1,\ldots,t_n)| \le \eta.$$

We define $A = [-\delta,\delta]$, and $B = [-\pi,\pi] \setminus [-\delta,\delta]$. Then,

$$\begin{aligned}
&\left|\int_{Q_n} D_k(\tau_1)\cdots D_k(\tau_n)\,(\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{(}t_1,\ldots,t_n))\,\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n\right| \\
\le\ &\int_{[-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\left|\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{(}t_1,\ldots,t_n)\right|\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n \\
&+\int_{Q^n\setminus[-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\left|\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{(}t_1,\ldots,t_n)\right|\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n.
\end{aligned}$$

We have

$$\begin{aligned}
&\int_{[-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\left|\tilde{f}(t_1-\tau_1,\ldots,t_n-\tau_n) - \tilde{(}t_1,\ldots,t_n)\right|\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n \\
&\qquad \eta\int_{[-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\,\mathrm{d}\tau_1\ \ldots\ \mathrm{d}\tau_n \le \eta.
\end{aligned}$$

Moreover, $Q_n \setminus [-\delta,\delta]^n$ can be written as the union of $2^n - 1$ subsets which are Cartesian products of factors which are A or B, and at least one of them is B. We deduce from the inequality

$$D_k(t) \le \frac{1}{2\,\pi\,(k+1)\,\sin^2\frac{\delta}{2}}$$

if $t \in B$, choosing k sufficiently large so that

$$\frac{1}{(k+1)\,\sin^2\frac{\delta}{2}} \le 1$$

the estimate

$$\int_{Q^n \setminus [-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\left|\tilde{f}(t_1-\tau_1,\dots,t_n-\tau_n)-\tilde{(}t_1,\dots,t_n)\right| \mathrm{d}\tau_1 \dots \mathrm{d}\tau_n$$
$$\frac{2\pi}{(k+1)\,\sin^2\frac{\delta}{2}} \max_{Q_n}|f|.$$

We conclude that

$$\left|\int_{Q^n \setminus [-\delta,\delta]^n} D_k(\tau_1)\cdots D_k(\tau_n)\left|\tilde{f}(t_1-\tau_1,\dots,t_n-\tau_n)-\tilde{(}t_1,\dots,t_n)\right| \mathrm{d}\tau_1 \dots \mathrm{d}\tau_n\right|$$
$$\leq \eta + \frac{2\pi}{(k+1)\,\sin^2\frac{\delta}{2}} \max_{Q_n}|f| \leq \varepsilon$$

if we take $\eta \leq \varepsilon/2$, and k sufficiently large so that

$$\frac{2\pi}{(k+1)\,\sin^2\frac{\delta}{2}} \leq \frac{\varepsilon}{2}.$$

Corollary 9.4 *$(e^j)_{j\in\mathbb{Z}^n}$ is a maximal orthonormal system in $L^2(Q_n;\mathbb{C})$.*

Proof The proof follows the lines of the proof of Theorem 9.8, using the fact that $\mathcal{D}((-\pi,\pi)^n;\mathbb{C})$ is dense in $L^2(Q_n;\mathbb{C})$ (theorem ... together with theorem ...).

(Adjust symbols from now)

Theorem 9.13 *Let f satisfy condition H_n. Suppose that, for some $k \in \{1,\dots,n\}$, f admits in Q_n the partial derivative $\mathrm{D}_k f$, also satisfying condition H_n. Then,*

$$\lim_{k\to\infty} \mathrm{D}_k(F_k(f))(t) = \mathrm{D}_k f(t)$$

uniformly for $t \in Q_n$.

Proof Using (9.19), we find

$$\mathrm{D}_k(F_n(f))(t_1,\dots,t_k,\dots,t_N)$$
$$= \sum_{j_1=-n}^{n} \cdots \sum_{j_N=-n}^{n} \left(1-\frac{|j_1|}{n+1}\right)\cdots\left(1-\frac{|j_N|}{n+1}\right) \int_{Q_N} f(\tau_1,\dots,\tau_N)\,\overline{v_{j_1}(\tau_1)\cdots v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)}\,\mathrm{d}\tau_1\dots\mathrm{d}\tau_N\, v_{j_1}(\tau_1)\cdots \mathrm{D}\, v_{j_k}(\tau_k$$

and

$$\int_{Q_N} f(\tau_1,\dots,\tau_N)\,\overline{v_{j_1}(\tau_1)\cdots v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)}\,\mathrm{d}\tau_1\dots\mathrm{d}\tau_N\, v_{j_1}(\tau_1)\cdots \mathrm{D}\, v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)$$
$$= -\int_{Q_N} f(\tau_1,\dots,\tau_N)\,\overline{v_{j_1}(\tau_1)\cdots \mathrm{D}_k v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)}\,\mathrm{d}\tau_1\dots\mathrm{d}\tau_N\, v_{j_1}(\tau_1)\cdots\, v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)$$
$$= \int_{Q_N} \mathrm{D}_k f(\tau_1,\dots,\tau_N)\,\overline{v_{j_1}(\tau_1)\cdots v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)}\,\mathrm{d}\tau_1\dots\mathrm{d}\tau_N\, v_{j_1}(\tau_1)\cdots\, v_{j_k}(\tau_k)\cdots v_{j_N}(\tau_N)$$

using the periodicity in the variable τ_k of f to integrate by parts. The conclusion follows from Theorem 9.12.

Corollary 9.5 *Let $f \in \mathcal{D}((-\pi,\pi)^n;\mathbb{C})$. Let f be extended with zero on the boundary ∂Q_N. Then, for any multi-index $\boldsymbol{\alpha} \in \mathbb{N}^N$,*

$$\lim_{n\to\infty} \mathrm{D}_{\boldsymbol{t}}^{\boldsymbol{\alpha}}(F_n(f))(\boldsymbol{t}) = \mathrm{D}_{\boldsymbol{t}}^{\boldsymbol{\alpha}} f(\boldsymbol{t})$$

uniformly for $\boldsymbol{t} \in Q_N$.

Proof It suffices to observe that for any multi-index $\boldsymbol{\alpha} \in \mathbb{N}^N$, the function $\mathrm{D}_{\boldsymbol{t}}^{\boldsymbol{\alpha}} f$ is extensible to a function satisfying condition H_N, so that Theorem 9.13 can be applied.

Corollary 9.5 can be applied to the theory of distributions. Let A be an open subset of $\mathbb{R}^n$. We denote with $\mathcal{D}(A;\mathbb{C})$ the class of functions of class C^∞ defined over A with compact support in A. This vector space is usually equipped with a structure of topological vector space (which is not metrizable) in such a way that, if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in it, it converges to $\varphi \in \mathcal{D}(A;\mathbb{C})$ if and only if there exists a compact subset K of A containing the support of φ_k for every $k \in \mathbb{N}$, and

$$\lim_{k\to\infty} \mathrm{D}^{\boldsymbol{\alpha}}\varphi_k(\boldsymbol{x}) = \mathrm{D}^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x})$$

uniformly in A, for each $\boldsymbol{\alpha} \in \mathbb{N}^N$. Moreover, if we denote by $\mathcal{D}'(A;\mathbb{C})$ the space of distributions in A (that is, the class of continuous linear functionals in $\mathcal{D}(A;\mathbb{C})$, then T belongs to $\mathcal{D}'(A;\mathbb{C})$ if and only if, for any sequence $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence converging to φ in $\mathcal{D}(A;\mathbb{C})$, then

$$\lim_{k\to\infty} T(\varphi_k) = T(\varphi).$$

Let A, B be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1 \in \mathcal{D}'(A;\mathbb{C})$ and $T_2 \in \mathcal{D}'(B;\mathbb{C})$, and let $\varphi \in \mathcal{D}(A\times B;\mathbb{C})$. Then we showed that

- $\varphi(\boldsymbol{x},\cdot) \in \mathcal{D}(B;\mathbb{C})$ for all $\boldsymbol{x} \in A$;
- the function $\boldsymbol{x} \mapsto \langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle$ belongs to $\mathcal{D}(A;\mathbb{C})$.

These facts allowed us to define the linear functional

$$\langle T_1 \otimes T_2, \varphi\rangle \stackrel{\text{def}}{=} \langle T_1(\boldsymbol{x}), \langle T_2(\boldsymbol{y}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \tag{9.20}$$

for all $\varphi \in \mathcal{D}(A\times B;\mathbb{C})$. We also showed that $T_1 \otimes T_2 \in \mathcal{D}'(A\times B;\mathbb{C})$. We also observed in Remark 5.8 that another natural definition of $T_1 \otimes T_2$ could be

$$\langle T_1 \otimes T_2, \varphi\rangle \stackrel{\text{def}}{=} \langle T_2(\boldsymbol{y}), \langle T_1(\boldsymbol{x}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle. \tag{9.21}$$

In the same remark, we anticipated that the two definitions are equivalent. Now we prove that this is indeed the case. To this end, we introduce the class of functions $\mathcal{D}(A;\mathbb{C}) \otimes \mathcal{D}(B;\mathbb{C})$.

Definition 9.2 *Let A, B be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We denote by $\mathcal{D}(A;\mathbb{C}) \otimes \mathcal{D}(B;\mathbb{C})$ the span of the class of products*

$$(\varphi_1 \otimes \varphi_2)(\boldsymbol{x},\boldsymbol{y}) = \varphi_1(\boldsymbol{x})\,\varphi_2(\boldsymbol{y})$$

with $\varphi_1 \in \mathcal{D}(A;\mathbb{C})$ and $\varphi_2 \in \mathcal{D}(B;\mathbb{C})$.

It follows that the elements of $\mathcal{D}(A;\mathbb{C}) \otimes \mathcal{D}(B;\mathbb{C})$ are the functions $\varphi \in \mathcal{D}(A\times B;\mathbb{C})$ which can be written in the form

$$\varphi(\boldsymbol{x},\boldsymbol{y}) = \sum_{j=1}^{m} \varphi_{1,j}(\boldsymbol{x})\,\varphi_{2,j}(\boldsymbol{y})$$

with $\varphi_{j,1} \in \mathcal{D}(A;\mathbb{C})$ and $\varphi_{j,2} \in \mathcal{D}(B;\mathbb{C})$ for some $m \in \mathbb{N}^*$. We prove the following result.

Theorem 9.14 *Let A, B be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Then, $\mathcal{D}(A;\mathbb{C}) \otimes \mathcal{D}(B;\mathbb{C})$ is dense in $gD(A \times B;\mathbb{C})$, in the sense that, for all $\varphi \in \mathcal{D}(A \times B;\mathbb{C})$, there exists a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(A;\mathbb{C}) \otimes \mathcal{D}(B;\mathbb{C})$ such that:*

(a) *there exists compact subsets K_1 and K_2 of A and B, respectively, such that the support of φ_k is contained in $K_1 \times K_2$ for all $k \in \mathbb{N}$;*

(b)
$$\forall\, z = (\boldsymbol{x}, \boldsymbol{y}), \quad \lim_{k\to\infty} \mathrm{D}^{\alpha}\varphi_k(z) = \mathrm{D}^{\alpha}\varphi(z)$$
uniformly in $A \times B$ for any $\boldsymbol{\alpha} \in \mathbb{N}^{m+n}$.

Proof Partitioning as usual $\boldsymbol{x}$ as $(x_1, \ldots, x_m)$ and $\boldsymbol{y}$ as $(y_1, \ldots, y_n)$, for each $z = (\boldsymbol{x}, \boldsymbol{y}) \in A \times B$, there exists $\delta(\boldsymbol{x}, \boldsymbol{y}) > 0$ such that

$$\prod_{i=1}^{m} [x_i - 2\,\delta(\boldsymbol{x},\boldsymbol{y}), x_i + 2\,\delta(\boldsymbol{x},\boldsymbol{y})] \times \prod_{j=1}^{m} \left[y_j - 2\,\delta(\boldsymbol{x},\boldsymbol{y}), y_j + 2\,\delta(\boldsymbol{x},\boldsymbol{y})\right] \subseteq A \times B.$$

Let K be a compact subset of $A \times B$. Then,

$$K \subseteq \bigcup_{(\boldsymbol{x},\boldsymbol{y})\in K} \prod_{i=1}^{m} (x_i - \delta(\boldsymbol{x},\boldsymbol{y}), x_i + \delta(\boldsymbol{x},\boldsymbol{y})) \times \prod_{j=1}^{m} \left(y_j - \delta(\boldsymbol{x},\boldsymbol{y}), y_j + \delta(\boldsymbol{x},\boldsymbol{y})\right).$$

Since K is compact, there exist $(x^1, y^1), \ldots, (x^p, y^p)$ in K such that

$$K \subseteq \bigcup_{1\le k\le p} \prod_{i=1}^{m} \left(x_i^k - \delta(\boldsymbol{x}^k,\boldsymbol{y}^k), x_i^k + \delta(\boldsymbol{x}^k,\boldsymbol{y}^k)\right) \times \prod_{j=1}^{m} \left(y_j^k - \delta(\boldsymbol{x}^k,\boldsymbol{y}^k), y_j^k + \delta(\boldsymbol{x}^k,\boldsymbol{y}^k)\right).$$

From the theorem of partition of unity, there exist $\psi_1, \ldots, \psi_m \in \mathcal{D}(A \times B;\mathbb{C})$ such that, for each $k \in \{1, \ldots, m\}$,

(a) $\operatorname{supp}\psi_k \subseteq \prod_{i=1}^{m} \left(x_i^k - \delta(\boldsymbol{x}^k,\boldsymbol{y}^k), x_i^k + \delta(\boldsymbol{x}^k,\boldsymbol{y}^k)\right) \times \prod_{j=1}^{m} \left(y_j^k - \delta(\boldsymbol{x}^k,\boldsymbol{y}^k), y_j^k + \delta(\boldsymbol{x}^k,\boldsymbol{y}^k)\right)$;

(b)
$$\forall\, (\boldsymbol{x},\boldsymbol{y}) \in K, \quad \sum_{k=1}^{p} \psi_k(\boldsymbol{x},\boldsymbol{y}) = 1.$$

It follows that, if $\varphi \in \mathcal{D}(A \times B;\mathbb{C})$ and $\operatorname{supp}\varphi \subseteq K$, then

$$\varphi = \sum_{k=1}^{p} \psi_k\, \varphi.$$

This means that it suffices to show the statement for each $\psi_k\,\varphi$, which in turn means that we only need to consider the case

$$\operatorname{supp}\varphi \subseteq \prod_{i=1}^{m} (x_i - \delta, x_i + \delta) \times \prod_{j=1}^{m} (y_j - \delta, y_j + \delta)$$

for some $(\boldsymbol{x}, \boldsymbol{y}) \in A \times B$ and $\delta > 0$ such that

$$\prod_{i=1}^{m} [x_i - \delta, x_i + \delta] \times \prod_{j=1}^{m} [y_j - \delta, y_j + \delta] \subseteq A \times B.$$

For simplicity of notation, we assume that $(x_1, \dots, x_m, y_1, \dots, y_n) = (0, \dots, 0, 0, \dots, 0)$ and $\delta = \pi$. We can reduce ourselves to this case using an affine change of variable. We set $N = m + n$. We have $\varphi \in \mathcal{D}(A \times B; \mathbb{C})$ with support in $(-\pi, \pi)^N$. By Corollary 9.5

$$\lim_{n \to \infty} \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}}(F_n(\varphi))(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\varphi(\boldsymbol{x}, \boldsymbol{y})$$

uniformly for $\boldsymbol{x} \in Q_N$ for all $\boldsymbol{\alpha} \in \mathbb{N}^N$. Since the support of φ is a compact subset of $(-\pi, \pi)^N$, there exists $\varepsilon \in (0, \pi)$ such that $\operatorname{supp} \varphi \subseteq [\varepsilon - \pi, \varepsilon + \pi]^N$. We take $\zeta \in \mathcal{D}((-\pi, \pi)^n, \mathbb{C})$ with $\zeta(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in [\varepsilon - \pi, \varepsilon + \pi]^m$ and $\eta \in \mathcal{D}((-\pi, \pi)^m, \mathbb{C})$ with $\eta(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in (\varepsilon - \pi, \varepsilon + \pi)^n$, so that

$$\forall\, (\boldsymbol{x}, \boldsymbol{y}) \in (-\pi, \pi)^N, \quad \zeta(\boldsymbol{x})\, \eta(y)\, \varphi(\boldsymbol{x}, \boldsymbol{y}) = \varphi(\boldsymbol{x}, \boldsymbol{y}).$$

Then it is clear from (9.19) that for all $n \in \mathbb{N}$ the function $g_n(\boldsymbol{x}, \boldsymbol{y}) = \zeta(\boldsymbol{x})\, \eta(\boldsymbol{y})\, F_n(f)(\boldsymbol{x}, \boldsymbol{y})$ belongs to $\mathcal{D}(A; \mathbb{C}) \otimes \mathcal{D}(B; \mathbb{C})$. Moreover, for any $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \in \mathbb{N}^m \times \mathbb{N}^n$,

$$\mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}}\, g_n(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\substack{\boldsymbol{\gamma} \leq \boldsymbol{\alpha} \\ \boldsymbol{\delta} \leq \boldsymbol{\beta}}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}} \binom{\boldsymbol{\beta}}{\boldsymbol{\delta}} \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}-\boldsymbol{\gamma}} \zeta(\boldsymbol{x})\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}-\boldsymbol{\delta}} \eta(\boldsymbol{y})\, \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\gamma}}\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\delta}} F_n(f)(\boldsymbol{x}, \boldsymbol{y})$$

converges uniformly as $n \to \infty$ to

$$\mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}}\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}}\, g_n(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\substack{\boldsymbol{\gamma} \leq \boldsymbol{\alpha} \\ \boldsymbol{\delta} \leq \boldsymbol{\beta}}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}} \binom{\boldsymbol{\beta}}{\boldsymbol{\delta}} \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}-\boldsymbol{\gamma}} \zeta(\boldsymbol{x})\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}-\boldsymbol{\delta}} \eta(\boldsymbol{y})\, \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\gamma}}\, \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\delta}} \varphi(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}} (\zeta(\boldsymbol{x})\, \eta(\boldsymbol{y})\, \varphi(\boldsymbol{x}, \boldsymbol{y})) = \mathrm{D}_{\boldsymbol{x}}^{\boldsymbol{\alpha}} \mathrm{D}_{\boldsymbol{y}}^{\boldsymbol{\beta}} \varphi(\boldsymbol{x}, \boldsymbol{y}).$$

Corollary 9.6 *Let A, B be open sets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Let $T_1 \in \mathcal{D}'(A; \mathbb{C})$ and $T_2 \in \mathcal{D}'(B; \mathbb{C})$. Then, there exists a unique distribution $T_1 \otimes T_2$ in $\mathcal{D}'(A \times B; \mathbb{C})$ such that, for any $\varphi_1 \in \mathcal{D}(A; \mathbb{C})$ and $\varphi_2 \in \mathcal{D}(B; \mathbb{C})$*

$$(T_1 \otimes T_2)(\varphi_1 \otimes \varphi_2) = T_1(\varphi_1)\, T_2(\varphi_2).$$

Proof The existence follows from (9.20). In order to show the uniqueness, let S and T be two elements of $\mathcal{D}'(A \times B; \mathbb{C})$ such that

$$S(\varphi_1 \otimes \varphi_2) = T(\varphi_1 \otimes \varphi_2)$$

for any $\varphi_1 \in \mathcal{D}(A; \mathbb{C})$ and $\varphi_2 \in \mathcal{D}(B; \mathbb{C})$. By linearity

$$\forall\, \varphi \in \mathcal{D}(A; \mathbb{C}) \otimes \mathcal{D}(B; \mathbb{C}), \qquad S(\varphi) = T(\varphi).$$

9.4 An application

We apply the previous results to the study of a simple model of evolution of the temperature in a bar which is kept at a constant temperature at the end-points, in the absence of heating sources. We study in particular the following mixed Cauchy-Dirichlet problem for the heat equation

$$\begin{cases} \mathrm{D}_t u(t, x) = \mathrm{D}_x^2 u(t, x), & (t, x) \in [0, +\infty) \times [0, \pi] \\ u(t, 0) = u(t, \pi) = 0, & t \in [0, +\infty) \\ u(0, x) = u_0(x), & x \in [0, \pi]. \end{cases} \tag{9.22}$$

We begin by specifying what we mean by solution of (9.22).

Definition 9.3 *A solution of (9.22) is a real-valued continuous function with domain* $[0,+\infty)\times[0,\pi]$*, equipped at any point of* $(0,\pi)\times[0,\pi]$ *with continuous derivatives* $\mathrm{D}_t u(t,x)$*,* $\mathrm{D}_x u(t,x)$ *and* $\mathrm{D}_x^2 u(t,x)$ *and satisfying all conditions in (9.22).*

Lemma 9.6 *A necessary condition for the existence of solutions of (9.22) are:*

(i) $u_0 \in C([0,\pi];\mathbb{R})$*;*

(ii) $u_0(0) = u_0(\pi) = 0$*.*

The proof is straightforward and is left as an exercise. We now consider the orthonormal system $(w_j)_{k\in\mathbb{N}}$ introduced in Theorem 9.11. This is a maximal orthonormal system. We try to represent a solution u of (9.22) in the form

$$u(t,\cdot) = \sum_{j=1}^{\infty} c_j(t)\, w_j,$$

with

$$c_j(t) = \int_0^\pi u(t,x)\, w_j(x)\,\mathrm{d}x = \sqrt{\frac{2}{\pi}} \int_0^\pi u(t,x)\ \sin(j\,x)\,\mathrm{d}x.$$

If u is a solution, we can differentiate under the sign of integral, to get

$$c_j'(t) = \int_0^\pi \mathrm{D}_t u(t,x)\, w_j(x)\,\mathrm{d}x = \int_0^\pi \mathrm{D}_x^2 u(t,x)\, w_j(x)\,\mathrm{d}x = \sqrt{\frac{2}{\pi}} \int_0^\pi \mathrm{D}_x^2 u(t,x)\ \sin(j\,x)\,\mathrm{d}x.$$

Integrating by parts and using the fact that $\sin(j\,0) = \sin(j\,\pi) = 0 = u(t,0) = u(t,\pi)$, we find

$$\begin{aligned} c_j'(t) &= \int_0^\pi \mathrm{D}_x^2 u(t,x)\, w_j(x)\,\mathrm{d}x \\ &= \left[\mathrm{D}_x u(t,x)\, w_j(x)\right]_{x=0}^{x=\pi} - \int_0^\pi \mathrm{D}_x u(t,x)\, w_j'(x)\,\mathrm{d}x \\ &= -\left[u(t,x)\, w_j'(x)\right]_{x=0}^{x=\pi} + \int_0^\pi u(t,x)\, w_j''(x)\,\mathrm{d}x \\ &= -j^2 \int_0^\pi u(t,x)\, w_j(x)\,\mathrm{d}x = -j^2\, c_j(t). \end{aligned}$$

Hence, c_j is the solution of the Cauchy problem

$$\begin{cases} c_j'(t) = -j^2\, c_j(t) \\ c_j(0) = c_j, \end{cases} \tag{9.23}$$

with

$$c_j = \int_0^\pi u_0(x)\, w_j(x)\,\mathrm{d}x.$$

It is well-known that (9.23) admits the unique solution

$$c_j(t) = c_j\, e^{-j^2 t},$$

so that we obtain

$$u(t,\cdot) = \sum_{j=1}^{\infty} c_j\, e^{-j^2 t}\, w_j, \tag{9.24}$$

with the series converging in $L^2([0,\pi];\mathbb{R})$ for all $t \in [0,+\infty)$. These arguments imply, in particular, that if a solution exists, it can be only represented by (9.24). It is easy to see that the series in (9.24) converges uniformly in $[t_0,+\infty)$ for any $t_0 > 0$. In fact, if $t \geq t_0 > 0$,

$$\sum_{j=1}^{\infty} |c_j|\, e^{-j^2 t}\, |w_j(x)| \leq \sqrt{\frac{2}{\pi}} \sum_{j=1}^{\infty} |c_j|\, e^{-j^2 t_0} \leq \sqrt{\frac{2}{\pi}} \Big(\sum_{j=1}^{\infty} |c_j|^2 \Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{\infty} e^{-2\, j^2\, t_0} \Big)^{\frac{1}{2}}.$$

This implies that u is continuous in $(0,+\infty)\times[0,\pi]$. Moreover, if we differentiate the terms of the series in (9.24) k times with respect to t and ℓ times with respect to x, we obtain

$$\mathrm{D}_t^k \mathrm{D}_x^\ell u(t,x) = \sqrt{\frac{2}{\pi}} \sum_{j=1}^{\infty} c_j\, (-j^2)^k\, e^{-j^2 t}\, j^\ell\, \sin^{(\ell)}(j\,x),$$

which is again uniformly convergent in $(t_0,+\infty)\times[0,\pi]$ for any $t_0 > 0$. We deduce that $u \in C^\infty((0,+\infty)\times[0,\pi];\mathbb{R})$. Moreover, if $t > 0$,

$$\mathrm{D}_t u(t,x) = \sqrt{\frac{2}{\pi}} \sum_{j=1}^{\infty} c_j (-j^2)\, e^{-j^2 t}\, \sin(j\,x) = \mathrm{D}_x^2 u(t,x).$$

It is also clear that

$$u(t,0) = u(t,\pi) = 0$$

for all $t \in (0,+\infty)$.

Now we study the behavior of u as t tends to 0. We assume that the following condition holds:

- $u_0 \in C^1([0,\pi];\mathbb{R})$ and $u_0(0) = u_0(\pi) = 0$.

Consider the orthonormal system $(v_j)_{j\in\mathbb{N}}$ in (**??**). We have

$$\int_0^\pi u_0'(x)\, v_0(x)\, \mathrm{d}x = 0$$

and, for $j \in \mathbb{N}^*$,

$$\int_0^\pi u_0'(x)\, v j 0(x)\, \mathrm{d}x = -\int_0^\pi u_0 x)\, v_j'((x)\, \mathrm{d}x = j\, c_j.$$

Thus,

$$u_0' = \sum_{j=1}^{\infty} j\, c_j\, v_j$$

in $L^2([0,\pi];\mathbb{R})$. It follows that

$$\sum_{j=0}^{\infty} j^2\, |c_j|^2 = \int_0^\pi |u_0'(x)|^2\, \mathrm{d}x < +\infty,$$

so that

$$\sum_{j=1}^{\infty} j^2\, |c_j|\, |w_j(x)| \leq \sqrt{\frac{2}{\pi}} \sum_{j=1}^{\infty} (j\,|c_j|)\, j^{-1} \leq \sqrt{\frac{2}{\pi}} \Big(\sum_{j=1}^{\infty} j^2\, |c_j|^2 \Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{\infty} j^{-2} \Big)^{\frac{1}{2}} < +\infty.$$

This implies that the series

$$\sum_{j=1}^{\infty} j\, c_j\, w_j(x)$$

converges uniformly in $[0,\pi]$. Moreover, if $(t,x) \in \mathbb{R}_+ \times [0,\pi]$,

$$|u(t,x) - u_0(x)| = \left| \sum_{j=1}^{\infty} (e^{-j^2 t} - 1 \big| c_j w_j(x) \right| \leq \sum_{j=1}^{\infty} (1 - e^{-j^2 t}) |c_j|,$$

which tends to 0 as $t \to 0$. This implies that, if we set $u(0,x) = u_0(x)$, then u is continuous in $[0,+\infty) \times [0,\pi]$, and is it is a solution of (9.22) in the sense of Definition 9.3. Hence, (9.22) has a unique solution in the sense of Definition 9.3.

It is also possible to prove that the conditions in Lemma 9.23 are also sufficient for a solution.

XXXX

9.5 Fourier series for distributions

We recall that a distribution $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ is t_0-periodic if $t_0 > 0$ and $T(x + t_0) = T(x)$. This means, as usual, that defining the diffeomorphism

$$\begin{aligned} \Phi : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x + t_0, \end{aligned}$$

we have $T \circ \Phi = T$.

The following result shows that every periodic distribution is tempered, and that the convergence in the sense of distributions coincides with the convergence in the sense of tempered distributions.

Theorem 9.15 *Let $t_0 > 0$. Let $T \in \mathcal{D}'(\mathbb{R};\mathbb{C})$ be t_0-periodic. Then, $T \in \mathcal{S}'(\mathbb{R};\mathbb{C})$. Moreover, given a sequence $(T_k)_{k\in\mathbb{N}}$ of t_0-periodic distributions such that $\lim_{k\to\infty} T_k = T$ in $\mathcal{D}'(\mathbb{R};\mathbb{C})$, then $\lim_{k\to\infty} T_k = T$ in $\mathcal{S}'(\mathbb{R};\mathbb{C})$.*

Proof Consider for the sake of simplicity $t_0 = 1$. We construct a function $\zeta \in \mathcal{D}(\mathbb{R};\mathbb{R})$ such that

$$\forall\, x \in \mathbb{R}, \quad \sum_{k=-\infty}^{+\infty} \zeta(x+k) = 1.$$

Let $\phi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ be a function with support contained in $[-1,1]$ and which satisfies the condition $\int_{\mathbb{R}} \phi(x)\,\mathrm{d}x = 1$. Let

$$\zeta(x) = (\phi * \chi_{[0,1]})(x) = \int_0^1 \phi(x-y)\,\mathrm{d}y = \int_{x-1}^{x} \phi(z)\,\mathrm{d}z, \tag{9.25}$$

where we used the change of variable $z = x - y$. The function ζ thus defined is of class C^∞ from the fundamental theorem of integral calculus. Moreover, from the same theorem we have

$$\zeta'(x) = \phi(x) - \phi(x-1).$$

We now show that ζ has compact support. Let $R > 0$ be such that $\operatorname{supp}\phi \subseteq [-R,R]$. For $x > R+1$, we have $\zeta(x) = 0$; likewise, for $x < -R-1$, we have $\zeta(x) = 0$. It follows that $\operatorname{supp}\zeta \subseteq [-R-1, R+1]$.

From (9.25) we have $\zeta(x+k) = \int_{x+k-1}^{x+k} \phi(z)\,\mathrm{d}z$, which gives, for every $n \in \mathbb{N}$,

$$\begin{aligned}\sum_{k=-n}^{n} \zeta(x+k) &= \sum_{k=-n}^{n} \int_{x+k-1}^{x+k} \phi(z)\,\mathrm{d}z \\ &= \int_{-n+k-1}^{-n+k} \phi(z)\,\mathrm{d}z + \int_{-n+k}^{-n+k+1} \phi(z)\,\mathrm{d}z + \ldots + \int_{n+k-2}^{n+k-1} \phi(z)\,\mathrm{d}z + \int_{n+k-1}^{n+k} \phi(z)\,\mathrm{d}z \\ &= \int_{-n+k-1}^{n+k} \phi(z)\,\mathrm{d}z \xrightarrow{n\to\infty} \int_{\mathbb{R}} \phi(z)\,\mathrm{d}z = 1.\end{aligned}$$

Let now $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$.

10 Solutions to the Exercises in Chapter 1

Solution 1.1 *(Exercise 1.1).* Let $x, y \in E$. Let $f(x) = f(y)$. We must prove that $x = y$. From $f(x) = f(y)$, we find immediately that $\mathrm{d}_F(f(x), f(y)) = 0$ from the identity of indiscernibles. From the definition of isometry we find $\mathrm{d}_E(x, y) = \mathrm{d}_F(f(x), f(y)) = 0$. Using the identity of indiscernibles again, we obtain $x = y$. ■

Solution 1.2 *(Exercise 1.2).* The only condition that needs working is the subadditivity: we need to use the fact that it holds for the absolute value. Let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$. Let $k \in \{1, 2, \ldots, n\}$ be an index at which the maximum $\max\limits_{1 \le i \le n} |x_i - y_i|$ is attained, so that $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = |x_k - y_k|$. It follows that

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = |x_k - y_k| \le |x_k - z_k| + |z_k - y_k| \le \max_{1 \le i \le n} |x_i - z_i| + \max_{1 \le i \le n} |z_i - y_i| = \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y}).$$

■

Solution 1.3 *(Exercise 1.3).* We prove that (E, d) is a metric space:

Identity of indiscernibles: trivially, for all $\boldsymbol{x}, \boldsymbol{y} \in E$, we have $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{y}$.

Symmetry: trivially, for all $\boldsymbol{x}, \boldsymbol{y} \in E$, we have

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{d}(\boldsymbol{y}, \boldsymbol{x}) = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{y} \\ 1 & \text{if } \boldsymbol{x} \neq \boldsymbol{y} \end{cases}$$

Triangle inequality: let $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in E$. We must consider two cases:

Case 1: $\boldsymbol{x} = \boldsymbol{y} = \boldsymbol{z}$. In this case obviously $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) = \mathrm{d}(\boldsymbol{z}, \boldsymbol{y}) = 0$, and therefore the triangle inequality $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) \le \mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y})$ reduces to $0 \le 0$, which is true.

Case 2: $\boldsymbol{x} \neq \boldsymbol{y}$ or $\boldsymbol{x} \neq \boldsymbol{z}$. In this case, at least one term in the sum $\mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y})$ will be equal to 1, so that

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{z}) + \mathrm{d}(\boldsymbol{z}, \boldsymbol{y}) \ge 1 \ge \mathrm{d}(\boldsymbol{x}, \boldsymbol{y}).$$

Therefore, (E, d) is a metric space. ■

Solution 1.4 *(Exercise 1.4).* Consider the distance d_1:

$$\begin{aligned} \mathrm{d}_1(\boldsymbol{x} + \boldsymbol{w}, \boldsymbol{y} + \boldsymbol{w}) &= |x_1 + w_1 - y_1 - w_1| + \ldots + |x_n + w_n - y_n - w_n| \\ &= |x_1 - y_1| + \ldots + |x_n - y_n| = \mathrm{d}_1(\boldsymbol{x}, \boldsymbol{y}). \end{aligned}$$

This proves the translation invariance;

$$\mathrm{d}_1(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = |k\,x_1 - k\,y_1| + \ldots + |k\,x_n - k\,y_n| = |k|\,|x_1 - y_1| + \ldots + |k|\,|x_n - y_n| = |k|\,\mathrm{d}_1(\boldsymbol{x}, \boldsymbol{y}).$$

This proves the homogeneity. ■

Solution 1.5 *(Exercise 1.5).* We denote the cardinality of a set A by $\#A$. We show the identity of indiscernibles. Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$:

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad \Leftrightarrow \quad \#\{i \in \{1, \dots, n\} \,|\, x_i \neq y_i\} = 0 \quad \Leftrightarrow \quad \forall\, i \in \{1, \dots, n\},\ x_i = y_i \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{y}.$$

We prove the symmetry. Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$ and $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$:

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = \#\{i \in \{1, \dots, n\} \,|\, x_i \neq y_i\} = \#\{i \in \{1, \dots, n\} \,|\, y_i \neq x_i\} = \mathrm{d}(\boldsymbol{y}, \boldsymbol{x}).$$

We prove the triangle inequality. Let $\boldsymbol{x} = (x_1, x_2, \dots, x_n)$, $\boldsymbol{y} = (y_1, y_2, \dots, y_n)$ and $\boldsymbol{z} = (z_1, z_2, \dots, z_n)$. Let $i \in \{1, \dots, n\}$. We consider the following cases:

- if $x_i = y_i$, then the i-th term does not contribute to the left hand-side. If
  - $z_i = x_i = y_i$, then there is no contribution to the right hand-side;
  - $z_i \neq x_i$ (so that also $z_i \neq y_i$), the i-th term adds 2 to the right hand-side.
- if $x_i \neq y_i$, the i-th term adds 1 to the left hand-side, and for the right hand-side we have three cases:
  - $z_i = x_i$, then $z_i \neq y_i$, and the i-th term contributes 1 to the right hand-side;
  - $z_i = y_i$, then $z_i \neq x_i$, and the i-th term contributes 1 to the right hand-side;
  - $z_i \neq x_i$ and $z_i \neq y_i$: the i-th term adds 2 to the right hand-side.

■

Solution 1.6 *(Exercise 1.6).* We have

$$\mathrm{d}_2(\boldsymbol{x}+\boldsymbol{w}, \boldsymbol{y}+\boldsymbol{w}) = \sqrt{\sum_{i=1}^{n} |x_i + w_i - y_i - w_i|^2} = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} = \mathrm{d}_2(\boldsymbol{x}, \boldsymbol{y})$$

and

$$\mathrm{d}_2(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} |k\,x_i - k\,y_i|^2} = \sqrt{\sum_{i=1}^{n} k^2\,|x_i - y_i|^2} = \sqrt{k^2 \sum_{i=1}^{n} |x_i - y_i|^2} = |k| \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} = |k|\,\mathrm{d}_2(\boldsymbol{x}, \boldsymbol{y}).$$

■

Solution 1.7 *(Exercise 1.7).* We have

$$\mathrm{d}(\boldsymbol{x}_1 + \boldsymbol{x}, \boldsymbol{x}_2 + \boldsymbol{x}) = \|(\boldsymbol{x}_1 + \boldsymbol{x}) - (\boldsymbol{x}_2 + \boldsymbol{x})\| = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = \mathrm{d}(\boldsymbol{x}_1, \boldsymbol{x}_2)$$

and

$$\mathrm{d}(k\,\boldsymbol{x}, k\,\boldsymbol{y}) = \|k\,\boldsymbol{x} - k\,\boldsymbol{y}\| = \|k\,(\boldsymbol{x} - \boldsymbol{y})\| = |k|\,\|\boldsymbol{x} - \boldsymbol{y}\| = |k|\,\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}).$$

■

Solution 1.8 *(Exercise 1.8).* We prove that this is a metric. We start with the identity of indiscernibles. If $\boldsymbol{x} = \boldsymbol{y}$, we have $\|\boldsymbol{x} - \boldsymbol{y}\| = 0$, and therefore $\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = 0$. Conversely, using the positivity of the norm,

$$\mathrm{d}(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad \Rightarrow \quad \frac{\|\boldsymbol{x} - \boldsymbol{y}\|}{1 + \|\boldsymbol{x} - \boldsymbol{y}\|} = 0 \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{y}| = 0 \quad \Rightarrow \quad \boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0} \quad \Rightarrow \quad \boldsymbol{x} = \boldsymbol{y}.$$

The symmetry follows from noting that $||\boldsymbol{x}-\boldsymbol{y}|| = ||\boldsymbol{y}-\boldsymbol{x}||$. We prove the triangle inequality. From the subadditivity of the norm

$$||\boldsymbol{x}-\boldsymbol{y}|| \leq ||\boldsymbol{x}-\boldsymbol{z}|| + ||\boldsymbol{z}-\boldsymbol{y}||.$$

Thus,

$$\begin{aligned} ||\boldsymbol{x}-\boldsymbol{y}|| &\leq ||\boldsymbol{x}-\boldsymbol{z}|| + ||\boldsymbol{z}-\boldsymbol{y}|| \\ &\leq ||\boldsymbol{x}-\boldsymbol{z}|| + ||\boldsymbol{z}-\boldsymbol{y}|| + ||\boldsymbol{x}-\boldsymbol{z}||\,||\boldsymbol{z}-\boldsymbol{y}|| + ||\boldsymbol{x}-\boldsymbol{z}||\,||\boldsymbol{x}-\boldsymbol{y}||\,||\boldsymbol{z}-\boldsymbol{y}|| + ||\boldsymbol{z}-\boldsymbol{y}||\,||\boldsymbol{x}-\boldsymbol{z}|| \end{aligned}$$

Adding on both sides $||\boldsymbol{x}-\boldsymbol{y}||\,||\boldsymbol{z}-\boldsymbol{y}||$, $||\boldsymbol{x}-\boldsymbol{y}||\,||\boldsymbol{x}-\boldsymbol{z}||$ and $||\boldsymbol{x}-\boldsymbol{y}||\,||\boldsymbol{x}-\boldsymbol{z}||\,||\boldsymbol{z}-\boldsymbol{y}||$ we can write

$$||\boldsymbol{x}-\boldsymbol{y}||\,(1+||\boldsymbol{x}-\boldsymbol{z}||)(1+||\boldsymbol{z}-\boldsymbol{y}||) \leq ||\boldsymbol{x}-\boldsymbol{z}||\,(1+||\boldsymbol{x}-\boldsymbol{y}||)(1+||\boldsymbol{z}-\boldsymbol{y}||) + ||\boldsymbol{z}-\boldsymbol{y}||\,(1+||\boldsymbol{x}-\boldsymbol{y}||)(1+||\boldsymbol{x}-\boldsymbol{z}||),$$

and dividing both sides by $(1+||\boldsymbol{x}-\boldsymbol{y}||)(1+||\boldsymbol{x}-\boldsymbol{z}||)(1+||\boldsymbol{z}-\boldsymbol{y}||)$ gives

$$\frac{||\boldsymbol{x}-\boldsymbol{y}||}{1+||\boldsymbol{x}-\boldsymbol{y}||} \leq \frac{||\boldsymbol{x}-\boldsymbol{z}||}{1+||\boldsymbol{x}-\boldsymbol{z}||} + \frac{||\boldsymbol{z}-\boldsymbol{y}||}{1+||\boldsymbol{z}-\boldsymbol{y}||}.$$

We show that this metric is translation invariant:

$$\mathrm{d}(\boldsymbol{x}+\boldsymbol{w},\boldsymbol{y}+\boldsymbol{w}) = \frac{||\boldsymbol{x}+\boldsymbol{w}-\boldsymbol{y}-\boldsymbol{w}||}{1+||\boldsymbol{x}+\boldsymbol{w}-\boldsymbol{y}-\boldsymbol{w}||} = \frac{||\boldsymbol{x}-\boldsymbol{y}||}{1+||\boldsymbol{x}-\boldsymbol{y}||} = \mathrm{d}(\boldsymbol{x},\boldsymbol{y}).$$

Consider the homogeneity: using the homogeneity of the norm

$$\mathrm{d}(k\,\boldsymbol{x},k\,\boldsymbol{y}) = \frac{||k\,\boldsymbol{x}-k\,\boldsymbol{y}||}{1+||k\,\boldsymbol{x}-k\,\boldsymbol{y}||} = \frac{|k|\,||\boldsymbol{x}-\boldsymbol{y}||}{1+|k|\,||\boldsymbol{x}-\boldsymbol{y}||}.$$

This is equal to $\mathrm{d}(\boldsymbol{x},\boldsymbol{y})$ if and only if

$$\frac{|k|\,||\boldsymbol{x}-\boldsymbol{y}||}{1+|k|\,||\boldsymbol{x}-\boldsymbol{y}||} = \frac{||\boldsymbol{x}-\boldsymbol{y}||}{1+||\boldsymbol{x}-\boldsymbol{y}||},$$

i.e., if and only if

$$|k|\,||\boldsymbol{x}-\boldsymbol{y}|| + |k|\,||\boldsymbol{x}-\boldsymbol{y}||^2 = ||\boldsymbol{x}-\boldsymbol{y}|| + |k|\,||\boldsymbol{x}-\boldsymbol{y}||^2,$$

i.e., if and only if $|k| = 1$. Thus, it does not hold for all $k \in \mathbb{F}$. ■

Solution 1.9 *(Exercise 1.9).* First, let us consider $\mathbb{C}$ as a vector space over the field $\mathbb{R}$. We recall that if $z = a + i\,b$ is a complex number and $a, b \in \mathbb{R}$, the modulus of z is defined as

$$|z| \stackrel{\text{def}}{=} \sqrt{a^2+b^2}.$$

We start checking the homogeneity. Let $k \in \mathbb{R}$. Then, $k\,z = k\,a + i\,k\,b$, so that

$$|k\,z| = \sqrt{k^2\,a^2 + k^2\,b^2} = \sqrt{k^2\,(a^2+b^2)} = \sqrt{k^2}\,\sqrt{a^2+b^2} = |k|\,|z|,$$

as required. Consider the subadditivity. Let $z = a + i\,b$ and $w = c + i\,d$, where $a, b, c, d \in \mathbb{R}$. We find

$$\begin{aligned} |z+w| \leq |z| + |w| \quad &\Leftrightarrow \quad |(a+c)+i\,(b+d)| \leq |a+i\,b| + |c+i\,d| \\ &\Leftrightarrow \quad \sqrt{(a+c)^2+(b+d)^2} \leq \sqrt{a^2+b^2} + \sqrt{c^2+d^2} \\ &\Leftrightarrow \quad \sqrt{a^2+2\,a\,c+c^2+b^2+2\,b\,d+d^2} \leq \sqrt{a^2+b^2} + \sqrt{c^2+d^2} \\ &\Leftrightarrow \quad a^2+2\,a\,c+c^2+b^2+2\,b\,d+d^2 \leq a^2+b^2+c^2+d^2+2\,\sqrt{a^2+b^2}\,\sqrt{c^2+d^2} \\ &\Leftrightarrow \quad a\,c+b\,d \leq \sqrt{a^2+b^2}\,\sqrt{c^2+d^2} \\ &\Leftrightarrow \quad (a\,c+b\,d)^2 \leq (a^2+b^2)(c^2+d^2) \\ &\Leftrightarrow \quad a^2\,c^2+b^2\,d^2+2\,a\,b\,c\,d \leq a^2\,c^2+a^2\,d^2+b^2\,c^2+b^2\,d^2 \\ &\Leftrightarrow \quad a^2\,d^2+b^2\,c^2-2\,a\,b\,c\,d \geq 0 \\ &\Leftrightarrow \quad (a\,d-b\,c)^2 \geq 0. \end{aligned}$$

Since the last inequality is true for any $a, b, c, d \in \mathbb{R}$, the triangle inequality holds true for all $z, w \in \mathbb{C}$. The positivity is straightforward; indeed,

$$|z| = 0 \quad \Leftrightarrow \quad \sqrt{a^2 + b^2} = 0 \quad \Leftrightarrow \quad a = b = 0 \quad \Leftrightarrow \quad z = 0.$$

This shows that the modulus of a complex number is a norm in $\mathbb{C}$ over the field $\mathbb{R}$. If we change the field to $\mathbb{C}$, the only axiom that needs to be checked is the homogeneity (it is the only one which involves scalars, so that the proofs of the other axioms is exactly the same). Let $k \in \mathbb{C}$. We write it in Cartesian form as $k = x + i\,y$, where $x, y \in \mathbb{R}$. Then,

$$k\,z = (x + i\,y)(a + i\,b) = x\,a - y\,b + i\,(a\,y + b\,x).$$

Hence,

$$\begin{aligned}
|k\,z| &= \sqrt{(x\,a - y\,b)^2 + (a\,y + b\,x)^2} \\
|k\,z| &= \sqrt{x^2\,a^2 - 2\,x\,a\,y\,b + y^2\,b^2 + a^2\,y^2 + 2\,a\,y\,b\,x + b^2\,x^2} \\
&= \sqrt{x^2\,a^2 + y^2\,b^2 + a^2\,y^2 + b^2\,x^2} \\
&= \sqrt{x^2\,(a^2 + b^2) + y^2\,(a^2 + b^2)} = \sqrt{(x^2 + y^2)\,(a^2 + b^2)} = \sqrt{x^2 + y^2}\,\sqrt{a^2 + b^2} = |k|\,|z|.
\end{aligned}$$

This concludes the exercise. ■

Solution 1.10 *(Exercise 1.10).* We check the first-order positive homogeneity: let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $k \in \mathbb{R}$. Then, from the properties of the absolute value

$$\|k\,\boldsymbol{x}\|_1 = |k\,x_1| + |k\,x_2| + \ldots + |k\,x_n| = |k|\,|x_1| + |k|\,|x_2| + \ldots + |k|\,|x_n| = |k|\,(|x_1| + |x_2| + \ldots + |x_n|) = |k|\,\|\boldsymbol{x}\|_1$$

We test the subadditivity. Let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $\boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Then,

$$\|\boldsymbol{x} + \boldsymbol{y}\|_1 = |x_1 + y_1| + |x_2 + y_2| + \ldots + |x_n + y_n| \leq |x_1| + |y_1| + |x_2| + |y_2| + \ldots + |x_n| + |y_n| = \|\boldsymbol{x}\|_1 + \|\boldsymbol{y}\|_1.$$

The positivity is obvious: the absolute value is always non-negative, and therefore such is also the sum of absolute values. Moreover, the sum of non-negative quantities can be zero only if those quantities are all zero: in other words, $\|\boldsymbol{x}\|_1 = 0$ implies $|x_1| + |x_2| + \ldots + |x_n| = 0$. Hence, $|x_j| = 0$ for all $j \in \{1, 2, \ldots, n\}$, which implies $x_j = 0$ for all $j \in \{1, 2, \ldots, n\}$.

Another solution is to show that this norm is associated with the taxi-cab metric introduced in an example in the lecture notes. ■

Solution 1.11 *(Exercise 1.11).* First, we prove the homogeneity: let $k \in \mathbb{R}$. We have

$$\|k\,\boldsymbol{x}\| = \sqrt{(k\,x_1)^2 + (k\,x_2)^2 + \ldots + (k\,x_n)^2} = \sqrt{k^2\,(x_1^2 + x_2^2 + \ldots + x_n^2)} = |k|\,\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = |k|\,\|\boldsymbol{x}\|.$$

We prove the positivity: if $\|\boldsymbol{x}\| = 0$, from

$$\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = 0$$

we obtain $x_1^2 + x_2^2 + \ldots + x_n^2 = 0$. This is a sum of non-negative quantities, and it is zero if and only if $x_i = 0$ for all $i = 1, \ldots, n$. Finally, we prove the triangle inequality. We have

$$\begin{aligned} \|\boldsymbol{x} + \boldsymbol{y}\|^2 &= \sum_{i=1}^{n} (x_i + y_i)^2 = \sum_{i=1}^{n} (x_i^2 + 2\, x_i\, y_i + y_i^2) \\ &= \sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} y_i^2 + 2 \sum_{i=1}^{n} x_i\, y_i \\ &= \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\, \boldsymbol{x} \cdot \boldsymbol{y} \\ &\leq \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 + 2\, \|\boldsymbol{x}\|\, \|\boldsymbol{y}\| = (\|\boldsymbol{x}\| + \|\boldsymbol{y}\|)^2, \end{aligned}$$

and taking the square root gives $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, as required. ■

Solution 1.12 *(Exercise 1.12).* Let

$$\begin{aligned} B_1(A, \mathbb{R}) &= \{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M \geq 0 : \quad \forall\, x \in A,\ |f(x)| \leq M\} \\ B_2(A, \mathbb{R}) &= \{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M > 0 : \quad \forall\, x \in A,\ |f(x)| \leq M\} \\ B_3(A, \mathbb{R}) &= \{f : A \longrightarrow \mathbb{R} \,|\, \exists\, M \in \mathbb{R} : \quad \forall\, x \in A,\ |f(x)| < M\}. \end{aligned}$$

We start proving that $B_1(A, \mathbb{R}) = B(A, \mathbb{R})$. Clearly, $B_1(A, \mathbb{R}) \subseteq B(A, \mathbb{R})$. We show the opposite inclusion. Let $f \in B(A, \mathbb{R})$, so that there exists $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in A$. Since $|f(x)| \geq 0$ for all $x \in A$, this implies that $M \geq 0$. Hence, $f \in B_1(A, \mathbb{R})$.

We show that $B_2(A, \mathbb{R}) = B(A, \mathbb{R})$. Again, $B_2(A, \mathbb{R}) \subseteq B(A, \mathbb{R})$, and we only need to show the opposite inclusion. Let $f \in B(A, \mathbb{R})$, so that there exists $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in A$. From the previous point, $M \geq 0$. If $M > 0$, we conclude immediately that $f \in B_2(A, \mathbb{R})$. If $M = 0$, then $|f(x)| \leq M < M + 1$, so that $|f(x)| \leq M + 1$, with $M + 1 > 0$, and therefore again $f \in B_2(A, \mathbb{R})$.

Finally, we have $B_3(A, \mathbb{R}) \subseteq B(A, \mathbb{R})$, and we show the opposite inclusion. Let $f \in B(A, \mathbb{R})$, so that there exists $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $x \in A$. It follows that $|f(x)| < M + 1$ for all $x \in A$, which implies that $f \in B_3(A, \mathbb{R})$.

We show that $B(A, \mathbb{R})$ is a subspace of $\mathbb{R}^A$. We only need to show that $B(A, \mathbb{R})$ is non-empty, and it is closed under addition and scalar multiplication. To show that $B(A, \mathbb{R})$ is not empty, we can simply notice that the identically zero function

$$\begin{aligned} A &\longrightarrow \mathbb{R} \\ x &\mapsto 0 \end{aligned}$$

is in $B(A, \mathbb{R})$. Let us consider $f, g \in B(A, \mathbb{R})$. Thus, there exist $M, N \in \mathbb{R}$ such that $|f(x)| \leq M$ and $|g(x)| \leq N$ for all $x \in A$. It follows that $|(f + g)(x)| = |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq M + N$ for all $x \in A$. This implies that $f + g \in B(A, \mathbb{R})$, i.e., $B(A, \mathbb{R})$ is closed under addition. Let $k \in \mathbb{R}$ and $f \in B(A, \mathbb{R})$, so that there exists $M \in \mathbb{R}$ such that $|f(x)| \leq M$. Then, $|(k\, f)(x)| = |k\, f(x)| = |k|\, |f(x)| \leq |k|\, M$, which shows that $k\, f \in B(A, \mathbb{R})$. This shows that $B(A, \mathbb{R})$ is closed under scalar multiplication, and therefore $B(A, \mathbb{R})$ is a subspace of the real vector space $\mathbb{R}^A$, and it is therefore a real vector space itself. ■

Solution 1.13 *(Exercise 1.13).* First, notice that if $\lambda = 0$ the equality holds true. Assume $\lambda \neq 0$. By definition of maximum, we have

$$\max_{1 \leq i \leq n} (|\lambda|\, |x_i|) = \max\Big\{|\lambda|\, |x_i| \;\Big|\; 1 \leq i \leq n\Big\} = |\lambda|\, |x_s|,$$

where $s \in \{1, \dots, n\}$ is such that $|\lambda|\,|x_i| \le |\lambda|\,|x_s|$ for all $i \in \{1, \dots, n\}$ (i.e., it is the index at which the maximum is attained). Since $|\lambda| \ge 0$, this is the same as $|x_i| \le |x_s|$ for all $i \in \{1, \dots, n\}$. This shows that

$$\max_{1 \le i \le n} (|\lambda|\,|x_i|) = \max\Big\{|\lambda|\,|x_i| \;\Big|\; 1 \le i \le n\Big\} = |\lambda|\,|x_s| = |\lambda| \max_{1 \le i \le n} |x_i|.$$

■

Solution 1.14 *(Exercise 1.14).* We prove the positive homogeneity. We have to show that for $k \in \mathbb{R}$ and $f \in C([a,b], \mathbb{R})$ we have $\|k\,f\|_\infty = |k|\,\|f\|_\infty$. This is indeed the case since from the property of the absolute value

$$\max_{x \in [a,b]} |k\,f(x)| = \max_{x \in [a,b]} |k|\,|f(x)| = |k| \max_{x \in [a,b]} |f(x)|.$$

The last equality follows from generalizing the previous result: we have $\max_{x \in [a,b]} |k|\,|f(x)| = |k| \max_{x \in [a,b]} |f(x)|$. Indeed, if $k = 0$, this is trivial; assuming $k \ne 0$, if we denote by x_0 the value of $x \in [a,b]$ such that

$$\max_{x \in [a,b]} |k|\,|f(x)| = |k|\,|f(x_0)|, \tag{10.1}$$

then by definition of maximum we have

$$\forall\, x \in [a,b], \quad |k|\,|f(x)| \le |k|\,|f(x_0)|,$$

which, since $|k| > 0$, is equivalent to

$$\forall\, x \in [a,b], \quad |f(x)| \le |f(x_0)| \quad \Rightarrow \quad |f(x_0)| = \max_{x \in [a,b]} |f(x)|.$$

Using this in (10.1) gives the result.

Consider the subadditivity. Let $f, g \in C([a,b], \mathbb{R})$. Then

$$\|f+g\|_\infty = \max_{x \in [a,b]} |(f+g)(x)| = \max_{x \in [a,b]} |f(x)+g(x)| \le \max_{x \in [a,b]} (|f(x)|+|g(x)|) \le \max_{x \in [a,b]} |f(x)| + \max_{x \in [a,b]} |g(x)| \le \|f\|_\infty + \|g\|_\infty.$$

We now prove the positivity. Clearly,

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)| = 0$$

implies that $f(x) = 0$ for all $x \in [a,b]$. ■

Solution 1.15 *(Exercise 1.15).* First, we re-write the statement as

$$\sup\Big\{|k|\,|f(x)| \;\Big|\; x \in A\Big\} = |k|\,\sup\Big\{|f(x)| \;\Big|\; x \in A\Big\}, \tag{10.2}$$

recalling that the supremum of a function is, by definition, the supremum of its image/range. The supremum of a set is the least upper bound of that set.

First, we notice that (10.2) trivially holds for $k = 0$, so we can focus only on the case $k \ne 0$. Then, given $M > 0$, we have

$$|k|\,|f(x)| \le M \quad \Leftrightarrow \quad |f(x)| \le \frac{M}{|k|},$$

which says that M is an upper bound for $\Big\{|k|\,|f(x)| \;\Big|\; x \in A\Big\}$ if and only if $M/|k|$ is an upper bound for $\Big\{|f(x)| \;\Big|\; x \in A\Big\}$. Thus, s is the least upper bound for $\Big\{|f(x)| \;\Big|\; x \in A\Big\}$ if and only if $|k|\,s$ is the least upper bound of $\Big\{|k|\,|f(x)| \;\Big|\; x \in A\Big\}$, which proves the result. ■

Solution 1.16 *(Exercise 1.16).* We have

$$\|\boldsymbol{v}\| = \sum_{j=1}^{n} |c_j| = 0 \quad \Rightarrow \quad |c_j| = 0, \ \ \forall\, j \in \{1, 2, \ldots, n\} \quad \Rightarrow \quad c_j = 0, \ \ \forall\, j \in \{1, 2, \ldots, n\} \quad \Rightarrow \quad \boldsymbol{v} = \boldsymbol{0}.$$

The converse is trivially true. Consider the homogeneity:

$$\|k\,\boldsymbol{v}\| = \sum_{j=1}^{n} |k\,c_j| = \sum_{j=1}^{n} |k|\,|c_j| = |k| \sum_{j=1}^{n} |c_j| = |k|\,\|\boldsymbol{v}\|.$$

Finally,

$$\|\boldsymbol{v} + \boldsymbol{w}\| = \sum_{j=1}^{n} |c_j + d_j| \leq \sum_{j=1}^{n} (|c_j| + |d_j|) = \sum_{j=1}^{n} |c_j| + \sum_{j=1}^{n} |d_j| = \|\boldsymbol{v}\| + \|\boldsymbol{w}\|.$$

■

Solution 1.17 *(Exercise 1.17).* We have

$$\|\boldsymbol{x}\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{\frac{1}{p}}.$$

Recall that, for $0 < b < a$, the function $x \mapsto a^x + b^x$ is asymptotically equivalent, as $x \to +\infty$, to a^x, since

$$\lim_{x \to +\infty} \frac{a^x + b^x}{a^x} = \lim_{x \to +\infty} \left[1 + \left(\frac{b}{a}\right)^x\right] = 1,$$

because $0 < b/a < 1$. Suppose that the maximum of $|x_i|$ is attained at the indices $j_1, j_2, \ldots, j_q$, so that $|x_{j_1}| = |x_{j_2}| = \ldots = |x_{j_q}|$. We find

$$\lim_{p \to +\infty} |x_1|^p + |x_2|^p + \ldots + |x_n|^p = |x_{j_1}|^p + |x_{j_2}|^p + \ldots + |x_{j_q}|^p = q\,|x_{j_1}|^p.$$

Thus,

$$\lim_{p \to +\infty} \|\boldsymbol{x}\|_p = \lim_{p \to +\infty} \left(q\,|x_{j_1}|^p\right)^{\frac{1}{p}} = \lim_{p \to +\infty} q^{\frac{1}{p}}\,|x_{j_1}| = |x_{j_1}| = \|\boldsymbol{x}\|_\infty.$$

■

Solution 1.18 *(Exercise 1.18).* We have

$$\max_{0 \leq x \leq 1} |e^x\,f(x)| = 0$$

and since $e^x \neq 0$ we obtain $f(x) = 0$ for all $x \in [0, 1]$. Let $k \in \mathbb{R}$. We have

$$\|k\,f\| = \max_{0 \leq x \leq 1} |e^x\,k\,f(x)| = \max_{0 \leq x \leq 1} |k|\,|e^x\,f(x)| = |k| \max_{0 \leq x \leq 1} |e^x\,f(x)| = |k|\,\|f\|.$$

Finally,

$$\begin{aligned}
\|f + g\| &= \max_{0 \leq x \leq 1} \left|e^x\,(f(x) + g(x))\right| \\
&= \max_{0 \leq x \leq 1} \left|e^x\,f(x) + e^x\,g(x)\right| \\
&\leq \max_{0 \leq x \leq 1} \left|e^x\,f(x)\right| + \max_{0 \leq x \leq 1} \left|e^x\,g(x)\right| \\
&= \|f\| + \|g\|.
\end{aligned}$$

■

Solution 1.19 *(Exercise 1.19).* First, from Weierstrass extreme value theorem we have

$$\forall\, f \in C([0,1],\mathbb{R}), \qquad \|f\| = \sup_{0\le x\le 1} w(x)\,|f(x)| = \max_{0\le x\le 1} w(x)\,|f(x)|,$$

because $w(x)\,|f(x)|$ is a continuous function (it is the product of continuous functions). Let

$$\max_{0\le x\le 1} w(x)\,|f(x)| = 0$$

and since $w(x) \neq 0$ we obtain $f(x) = 0$ for all $x \in [0,1]$. Let $k \in \mathbb{R}$. We have

$$\|k\,f\| = \max_{0\le x\le 1} w(x)\,|k\,f(x)| = \max_{0\le x\le 1} |k|\,w(x)\,|f(x)| = |k| \max_{0\le x\le 1} w(x)\,|f(x)| = |k|\,\|f\|.$$

Finally, given $f, g \in C([0,1],\mathbb{R})$ we have

$$\begin{aligned}\|f+g\| &= \max_{0\le x\le 1} w(x)\big|f(x)+g(x)\big| \\ &\le \max_{0\le x\le 1} w(x)\,(|f(x)|+|g(x)|) \\ &\le \max_{0\le x\le 1} w(x)\,|f(x)| + \max_{0\le x\le 1} w(x)\,|g(x)| = \|f\| + \|g\|.\end{aligned}$$

■

Solution 1.20 *(Exercise 1.20).*

(i) We show that the norm

$$\|(\boldsymbol{v},\boldsymbol{w})\| = \|\boldsymbol{v}\|_V + \|\boldsymbol{w}\|_W, \qquad \boldsymbol{v}\in\boldsymbol{V},\ \boldsymbol{w}\in\boldsymbol{W}$$

satisfies all the axioms of a norm.

Homogeneity: Let $\boldsymbol{v}\in\boldsymbol{V}$, $\boldsymbol{w}\in\boldsymbol{W}$ and $\lambda\in\mathbb{F}$. Then

$$\begin{aligned}\|\lambda\,(\boldsymbol{v},\boldsymbol{w})\| &= \|(\lambda\,\boldsymbol{v},\lambda\,\boldsymbol{w})\| = \|\lambda\,\boldsymbol{v}\|_V + \|\lambda\,\boldsymbol{w}\|_W = |\lambda|\,\|\boldsymbol{v}\|_V + |\lambda|\,\|\boldsymbol{w}\|_W \\ &= |\lambda|\,(\|\boldsymbol{v}\|_V + \|\boldsymbol{w}\|_W) = |\lambda|\,\|(\boldsymbol{v},\boldsymbol{w})\|.\end{aligned}$$

Subadditivity: Let $\boldsymbol{v}_1,\boldsymbol{v}_2\in\boldsymbol{V}$, $\boldsymbol{w}_1,\boldsymbol{w}_2\in\boldsymbol{W}$. Then

$$\begin{aligned}\|(\boldsymbol{v}_1,\boldsymbol{w}_1)+(\boldsymbol{v}_2,\boldsymbol{w}_2)\| &= \|(\boldsymbol{v}_1+\boldsymbol{v}_2,\boldsymbol{w}_1+\boldsymbol{w}_2)\| = \|\boldsymbol{v}_1+\boldsymbol{v}_2\|_V + \|\boldsymbol{w}_1+\boldsymbol{w}_2\|_W \\ &\le \|\boldsymbol{v}_1\|_V + \|\boldsymbol{v}_2\|_V + \|\boldsymbol{w}_1\|_W + \|\boldsymbol{w}_2\|_W \\ &= (\|\boldsymbol{v}_1\|_V + \|\boldsymbol{w}_1\|_W) + (\|\boldsymbol{v}_2\|_V + \|\boldsymbol{w}_2\|_W) \\ &= \|(\boldsymbol{v}_1,\boldsymbol{w}_1)\| + \|(\boldsymbol{v}_2,\boldsymbol{w}_2)\|.\end{aligned}$$

Positivity: $\|(\boldsymbol{v},\boldsymbol{w})\| = 0$ implies $\|\boldsymbol{v}\|_V + \|\boldsymbol{w}\|_W = 0$, and since the norms are non-negative we find $\|\boldsymbol{v}\|_V = 0$ and $\|\boldsymbol{w}\|_W = 0$. Hence, $\boldsymbol{v} = \boldsymbol{0}_V$ and $\boldsymbol{w} = \boldsymbol{0}_W$.

(ii) Consider the second norm.

Homogeneity: Let $\boldsymbol{v}\in\boldsymbol{V}$, $\boldsymbol{w}\in\boldsymbol{W}$ and $\lambda\in\mathbb{F}$. Then

$$\begin{aligned}\|\lambda\,(\boldsymbol{v},\boldsymbol{w})\| &= \|(\lambda\,\boldsymbol{v},\lambda\,\boldsymbol{w})\| = \max\{\|\lambda\,\boldsymbol{v}\|_V, \|\lambda\,\boldsymbol{w}\|_W\} \\ &= \max\{|\lambda|\,\|\boldsymbol{v}\|_V, |\lambda|\,\|\boldsymbol{w}\|_W\} \\ &= |\lambda|\,\max\{\|\boldsymbol{v}\|_V, \|\boldsymbol{w}\|_W\} = |\lambda|\,\|(\boldsymbol{v},\boldsymbol{w})\|.\end{aligned}$$

Subadditivity: Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \boldsymbol{V}$, $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \boldsymbol{W}$. Then

$$\begin{aligned} \|(\boldsymbol{v}_1, \boldsymbol{w}_1) + (\boldsymbol{v}_2, \boldsymbol{w}_2)\| &= \|(\boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{w}_1 + \boldsymbol{w}_2)\| = \max\{\|\boldsymbol{v}_1 + \boldsymbol{v}_2\|_V, \|\boldsymbol{w}_1 + \boldsymbol{w}_2\|_W\} \\ &\leq \max\{\|\boldsymbol{v}_1\|_V + \|\boldsymbol{v}_2\|_V, \|\boldsymbol{w}_1\|_W + \|\boldsymbol{w}_2\|_W\} \\ &\leq \max\{\|\boldsymbol{v}_1\|_V, \|\boldsymbol{w}_1\|_W\} + \max\{\|\boldsymbol{v}_2\|_V, \|\boldsymbol{w}_2\|_W\} \\ &= \|(\boldsymbol{v}_1, \boldsymbol{w}_1)\| + \|(\boldsymbol{v}_2, \boldsymbol{w}_2)\|. \end{aligned}$$

Here, we have used the obvious fact that $\max\{a + b, c + d\} \leq \max\{a, c\} + \max\{b, d\}$.

Positivity: $\|(\boldsymbol{v}, \boldsymbol{w})\| = 0$ implies $\max\{\|\boldsymbol{v}\|_V, \|\boldsymbol{w}\|_W\} = 0$, and therefore $\|\boldsymbol{v}\|_V = 0$ and $\|\boldsymbol{w}\|_W = 0$. Hence, $\boldsymbol{v} = \boldsymbol{0}_V$ and $\boldsymbol{w} = \boldsymbol{0}_W$.

■

Solution 1.21 *(Exercise 1.21).*

(i) Trivial

(ii) Using the fact that $\|\cdot\|$ is a norm,

$$p(k\,\boldsymbol{x}) = \|A\,(k\,\boldsymbol{x})\| = \|k\,A\,\boldsymbol{x}\| = |k|\,\|A\,\boldsymbol{x}\| = |k|\,p(\boldsymbol{x}),$$

$$p(\boldsymbol{x} + \boldsymbol{y}) = \|A(\boldsymbol{x} + \boldsymbol{y})\| = \|A\,\boldsymbol{x} + A\,\boldsymbol{y}\| \leq \|A\,\boldsymbol{x}\| + \|A\,\boldsymbol{y}\| = p(\boldsymbol{x}) + p(\boldsymbol{y})$$

and

$$p(\boldsymbol{x}) = 0 \quad \Leftrightarrow \quad \|A\,\boldsymbol{x}\| = 0 \quad \Leftrightarrow \quad A\,\boldsymbol{x} = \boldsymbol{0} \quad \Leftrightarrow \quad \boldsymbol{x} \in \ker A = \{0\} \quad \Leftrightarrow \quad \boldsymbol{x} = \boldsymbol{0},$$

because from the rank-nullity theorem, if A is full row rank, its null-space is zero.

(iii) Let $p(\boldsymbol{x}) = \alpha\,\|\boldsymbol{x}\|_{(1)} + \beta\,\|\boldsymbol{x}\|_{(2)}$, then

$$p(k\,\boldsymbol{x}) = \alpha\,\|k\,\boldsymbol{x}\|_{(1)} + \beta\,\|k\,\boldsymbol{x}\|_{(2)} = \alpha\,|k|\,\|\boldsymbol{x}\|_{(1)} + \beta\,|k|\,\|\boldsymbol{x}\|_{(2)} = |k|\,p(\boldsymbol{x}).$$

$$\begin{aligned} p(\boldsymbol{x} + \boldsymbol{y}) &= \alpha\,\|\boldsymbol{x} + \boldsymbol{y}\|_{(1)} + \beta\,\|\boldsymbol{x} + \boldsymbol{y}\|_{(2)} \leq \alpha\,(\|\boldsymbol{x}\|_{(1)} + \|\boldsymbol{y}\|_{(1)}) + \beta\,(\|\boldsymbol{x}\|_{(2)} + \|\boldsymbol{y}\|_{(2)}) \\ &= (\alpha\,\|\boldsymbol{x}\|_{(1)} + \beta\,\|\boldsymbol{x}\|_{(2)}) + (\alpha\,\|\boldsymbol{y}\|_{(1)} + \beta\,\|\boldsymbol{y}\|_{(2)}) = p(\boldsymbol{x}) + p(\boldsymbol{y}). \end{aligned}$$

Finally $p(\boldsymbol{x}) = 0$ iff $\alpha\,\|\boldsymbol{x}\|_{(1)} + \beta\,\|\boldsymbol{x}\|_{(2)} = 0$ iff $\|\boldsymbol{x}\|_{(1)} = \|\boldsymbol{x}\|_{(2)} = 0$ iff $\boldsymbol{x} = 0$.

- It is not a norm: for example $\min(0, 1, 2, 3) = 0$ but $\boldsymbol{x} \neq 0$.
- It is not a norm. Take $\boldsymbol{x} = (-2, -1, -2)$ and $\boldsymbol{y} = (-1, 1, 0)$ so that

  $$p(\boldsymbol{x}) = 2 + 1 = 3, \quad p(\boldsymbol{y}) = 1 + 0 = 1, \quad p(\boldsymbol{x} + \boldsymbol{y}) = p(-3, 0, -2) = 3 + 2 = 5,$$

  so that the triangle inequality is not satisfied.
- This is a norm:

  $$p(k\,(x_1, x_2)) = |k\,x_1 - k\,x_2| + |k\,x_1| = |k|\,(|x_1 - x_2| + |x_1|),$$

  $$p(x_1 + x_2, y_1 + y_2) = |x_1 + x_2 - y_1 - y_2| + |x_1 + x_2| \leq |x_1 - y_1| + |x_1| + |x_2 - y_2| + |x_2|$$

  and $p(x_1, x_2) = 0$ iff $|x_1 - x_2| + |x_1| = 0$ iff $x_1 = x_2$ and $x_1 = 0$ iff $\boldsymbol{x} = 0$. In $\mathbb{R}^3$ we can define $(x_1, x_2, x_3) \mapsto |x_3 - x_1 + x_2| + |x_1 - x_2| + |x_1|$.

■

Solution 1.22 *(Exercise 1.22).*

$$\begin{aligned} \mathrm{d}_1(\boldsymbol{x},\boldsymbol{y}) &= |x_1-y_1|+|x_2-y_2|+|x_3-y_3|+|x_4-y_4| = |1-2|+|2-2|+|-2+0|+|0-1| = 4 \\ \mathrm{d}_2(\boldsymbol{x},\boldsymbol{y}) &= \sqrt{|x_1-y_1|^2+|x_2-y_2|^2+|x_3-y_3|^2+|x_4-y_4|^2} = \sqrt{1+0+4+1} = \sqrt{6} \\ \mathrm{d}_d(\boldsymbol{x},\boldsymbol{y}) &= 0. \end{aligned}$$

■

Solution 1.23 *(Exercise 1.23).* We have

$$\begin{aligned} \|z\|_1 &= |1+2\,i|+|-1+i| = \sqrt{5}+\sqrt{2} \\ \|z\|_2 &= \sqrt{|1+2\,i|^2+|-1+i|^2} = \sqrt{5+2} = \sqrt{7} \\ \|z\|_\infty &= \max\{\sqrt{5},\sqrt{2}\} = \sqrt{5}. \end{aligned}$$

■

Solution 1.24 *(Exercise 1.24).* This function is not a norm. One way to see this is to show that the positivity does not hold. For example, if $\boldsymbol{x} = (1,1,0,0,\ldots,0)$, we obtain

$$\|\boldsymbol{x}\| = -1+1+0+0+\ldots+0 = 0.$$

■

Solution 1.25 *(Exercise 1.25).* This function satisfies the homogeneity. Indeed, given $\boldsymbol{x} \in \mathbb{R}^n$ and $k \in \mathbb{R}$,

$$\begin{aligned} p(k\,\boldsymbol{x}) &= (|k\,x_1|^{\frac{1}{2}}+|k\,x_2|^{\frac{1}{2}}+\ldots+|k\,x_n|^{\frac{1}{2}})^2 \\ &= (|k|^{\frac{1}{2}}(|x_1|^{\frac{1}{2}}+|x_2|^{\frac{1}{2}}+\ldots+|x_n|^{\frac{1}{2}}))^2 \\ &= |k|\,(|x_1|^{\frac{1}{2}}+|x_2|^{\frac{1}{2}}+\ldots+|x_n|^{\frac{1}{2}})^2 = |k|\,p(\boldsymbol{x}) \end{aligned}$$

The positivity is satisfied: if $p(\boldsymbol{x}) = 0$, then $|x_1|^{\frac{1}{2}}+|x_2|^{\frac{1}{2}}+\ldots+|x_n|^{\frac{1}{2}} = 0$ implies $|x_1| = \ldots = |x_2| = 0$, which in turn implies $x_1 = x_2 = \ldots = x_n = 0$. Therefore, the axiom that is violated in the triangle inequality. Consider e.g. $n = 2$, $\boldsymbol{x} = (1,2)$ and $\boldsymbol{y} = (1,0)$. Then,

$$p(\boldsymbol{x}+\boldsymbol{y}) = p(2,2) = (\sqrt{2}+\sqrt{2})^2 = 2+2+4 = 8$$

and

$$p(\boldsymbol{x})+p(\boldsymbol{y}) = (1+\sqrt{2})^2+1^2 = 1+2+2\sqrt{2}+1 = 4+2\sqrt{2}.$$

We have $p(\boldsymbol{x}+\boldsymbol{y}) > p(\boldsymbol{x})+p(\boldsymbol{y})$ because $2 > \sqrt{2}$, and therefore p is not a norm. ■

Solution 1.26 *(Exercise 1.26).* For all $\boldsymbol{x} \in \mathbb{R}^n$, we have

$$p(\boldsymbol{x}) = |x_1|+2\,|x_2|+3\,|x_3|+\ldots+n\,|x_n|.$$

We prove the homogeneity. Given $k \in \mathbb{R}$

$$\begin{aligned} p(k\,\boldsymbol{x}) &= |k\,x_1|+2\,|k\,x_2|+3\,|k\,x_3|+\ldots+n\,|k\,x_n| \\ &= |k|\,|x_1|+2\,|k|\,|x_2|+3\,|k|\,|x_3|+\ldots+n\,|k|\,|x_n| = |k|\,p(\boldsymbol{x}). \end{aligned}$$

We prove the subadditivity. Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$p(\boldsymbol{x}+\boldsymbol{y}) = |x_1+y_1| + 2\,|x_2+y_2| + \ldots + n\,|x_n+y_n| \leq |x_1| + |y_1| + 2\,|x_2| + 2\,|y_2| + \ldots + n\,|x_n| + n\,|y_n| = \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$

We prove the positivity:

$$p(\boldsymbol{x}) = 0 \quad \Leftrightarrow \quad |x_1| + 2\,|x_2| + 3\,|x_3| + \ldots + n\,|x_n| = 0.$$

Since the sum involves non-negative terms, we get zero if and only if $x_1, x_2, \ldots, x_n = 0$, i.e., $\boldsymbol{x} = \boldsymbol{0}$. ■

Solution 1.27 *(Exercise 1.27).* The implication *(i)* ⇒ *(ii)* is trivial: if $\varepsilon > 0$, then $\ell \leq a < a + \varepsilon$, and so $\ell < a + \varepsilon$. We show that *(ii)* ⇒ *(i)*. Suppose by contradiction that $\ell > a$. Take $\varepsilon = \ell - a$. This value of ε contradicts *(ii)*.

Now we show that *(i)* ⇔ *(iii)*. The implication *(i)* ⇒ *(iii)* is trivial: for all $\varepsilon > 0$ we have $\ell \leq a < a + \varepsilon$ which implies $\ell < a + \varepsilon$. We need to show that *(i)* ⇐ *(iii)*. By contradiction suppose that *(iii)* does not holds, and that $\ell > a$. Taking

$$\varepsilon = \frac{\ell - a}{2} > 0$$

we find $\ell - a = 2\,\varepsilon$ and so $\ell = a + 2\,\varepsilon > a + \varepsilon$, contradicting *(i)*.

Finally, we have to show the equivalence of *(ii)* and *(iii)*. The fact that *(iii)* implies *(ii)* is obvious. We show that *(ii)* implies *(iii)*. We reason again by contradiction. Suppose that *(iii)* is false, so there exists $\bar{\varepsilon} > 0$ such that $\ell \geq a + \varepsilon$. Then, $\ell \geq a + \bar{\varepsilon} > a + \frac{\bar{\varepsilon}}{2}$, which contradicts *(ii)*: we found a $\varepsilon > 0$ (equal to $\bar{\varepsilon}/2$) such that $\ell > a + \varepsilon$. ■

Solution 1.28 *(Exercise 1.28).* Consider the case where $X = \mathbb{R}$ and the norm is the absolute value. Then,

$$\mathcal{B}_r(x_0) = \{x \in \mathbb{R} \,|\, |x - x_0| < r\} = \{x \in \mathbb{R} \,|\, -r < x - x_0 < r\} = \{x \in \mathbb{R} \,|\, x_0 - r < x < x_0 + r\} = (x_0 - r, x_0 + r).$$

This shows that in this case the ball is simply an open interval centered at x_0. Consider $X = \mathbb{R}^2$ with the Euclidean norm and a point $(x_0, y_0) \in X$:

$$\begin{aligned}\mathcal{B}_r((x_0, y_0)) &= \{(x, y) \in X \,|\, \|(x - x_0, y - y_0)\| < r\} \\ &= \left\{(x, y) \in X \,\middle|\, \sqrt{(x - x_0)^2 + (y - y_0)^2} < r\right\}.\end{aligned}$$

The inequality $\sqrt{(x - x_0)^2 + (y - y_0)^2} < r$ is satisfied by the points $(x, y) \in \mathbb{R}^2$ whose distance from (x_0, y_0) is strictly smaller than r. When $X = \mathbb{R}^3$, we have

$$\mathcal{B}_r((x_0, y_0, z_0)) = \left\{(x, y) \in X \,\middle|\, \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} < r\right\},$$

which geometrically represents the points of a sphere centered at (x_0, y_0, z_0) of radius r. ■

Solution 1.29 *(Exercise 1.29).* Let $A \subseteq E$ be arbitrary. Let $x_0 \in A$. Let $r = \frac{1}{2}$. We find

$$\mathcal{B}_{\frac{1}{2}}(x_0) = \left\{x \in E \,\middle|\, \mathrm{d}(x, x_0) < \frac{1}{2}\right\} = \{x_0\} \subseteq A.$$

It follows that A is open. We now show that $\partial A = \varnothing$. Indeed, choosing again $r = \frac{1}{2}$, given any point $x_0 \in E$, we have two possibilities: either $x_0 \in A$, or $x_0 \notin A$. If $x_0 \in A$, we find

$$\mathcal{B}_{\frac{1}{2}}(x_0) \cap A = \{x_0\} \quad \text{but} \quad \mathcal{B}_{\frac{1}{2}}(x_0) \cap (E \setminus A) = \varnothing.$$

If $x_0 \notin A$, we find

$$\mathcal{B}_{\frac{1}{2}}(x_0) \cap A = \varnothing \quad \text{but} \quad \mathcal{B}_{\frac{1}{2}}(x_0) \cap (E \setminus A) = \{x_0\}.$$

It follows that $x_0 \notin \partial A$. From the arbitrariness of $x_0 \in E$ we conclude that $\partial A = \varnothing$. Since $\partial A = \varnothing \subseteq A$, we also conclude that A is closed. ■

Solution 1.30 *(Exercise 1.30).* Since $\boldsymbol{U}$ is a neighbourhood of $\boldsymbol{x}_0$, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \boldsymbol{U}$. Then, $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \boldsymbol{U} \subseteq \boldsymbol{W}$ implies $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \boldsymbol{W}$, which says that $\boldsymbol{W}$ is a neighbourhood of $\boldsymbol{x}_0$. ■

Solution 1.31 *(Exercise 1.31).* The first point is trivial. Let $j \in \mathcal{I}$. Since $\boldsymbol{U}_j$ is a neighborhood of $\boldsymbol{x}_0$, there exists $r_j > 0$ such that $\mathcal{B}_{r_j}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_j$. Thus,

$$\mathcal{B}_{r_j}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_j \subseteq \bigcup_{i \in \mathcal{I}} \boldsymbol{U}_i$$

is enough to show that $\bigcup_{i \in \mathcal{I}} \boldsymbol{U}_i$ is a neighborhood of $\boldsymbol{x}_0$.

We now want to prove the second point. To simplify, we start with the intersection of two neighborhoods. Let $\boldsymbol{U}_1, \boldsymbol{U}_2$ be neighbourhoods of $\boldsymbol{x}_0$. Then, there exists $r_1 > 0$ such that $\mathcal{B}_{r_1}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_1$, and, likewise, there exists $r_2 > 0$ such that $\mathcal{B}_{r_2}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_2$. Take $r = \min\{r_1, r_2\}$. We have $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \mathcal{B}_{r_1}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_1$ and $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \mathcal{B}_{r_2}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_2$, so that $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_1 \cap \boldsymbol{U}_2$. It follows that $\boldsymbol{U}_1 \cap \boldsymbol{U}_2$ is a neighbourhood of $\boldsymbol{x}_0$.

The same argument can be used for a finite set of neighborhoods. Indeed, in the case of a finite set of neighborhoods, there exists $p \in \mathbb{N}^*$ such that $\mathcal{I} = \{i_1, i_2, \ldots, i_p\}$. For every $k \in \{1, \ldots, p\}$, the set $\boldsymbol{U}_{i_k}$ is a neighborhood of $\boldsymbol{x}_0$, so there exists $r_{i_k} > 0$ such that $\mathcal{B}_{r_{i_k}}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_{i_k}$. Let $r = \min_{k \in \{1,\ldots,p\}} r_{i_k}$: the minimum is well defined because the set $\{r_{i_k} \,|\, k \in \{1, \ldots, p\}\}$ is finite. We have $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \mathcal{B}_{r_{i_k}}(\boldsymbol{x}_0) \subseteq \boldsymbol{U}_{i_k}$ for all $k \in \{1, \ldots, p\}$, so that

$$\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \bigcap_{k=1}^{p} \boldsymbol{U}_{i_k},$$

which implies that $\bigcap_{k=1}^{p} \boldsymbol{U}_{i_k}$ is a neighborhood of $\boldsymbol{x}_0$.

Now we show that the statement is not true if we consider an infinite set of neighbourhoods. For example, let $\boldsymbol{X} = \mathbb{R}$, and consider the point $x_0 = 0$. Let $\mathcal{I} = \mathbb{N}^*$, and consider $U_i = \left(-\frac{1}{i}, \frac{1}{i}\right)$ for all $i \in \mathcal{I}$. Each U_i is a neighborhood of 0, but the intersection

$$\bigcap_{i \in \mathbb{N}^*} \left(-\frac{1}{i}, \frac{1}{i}\right)$$

is $\{0\}$. In fact, $0 \in \left(-\frac{1}{i}, \frac{1}{i}\right)$ for every $i \in \mathbb{N}^*$, and if by contradiction $\left(-\frac{1}{i}, \frac{1}{i}\right)$ contained a point $s \neq 0$, we would easily find a value $i \in \mathbb{N}^*$ such that $s \notin U_i$ (for example, if $s > 0$, we could take i to be the smallest natural number $\geq 1/s$). However, $\{0\}$ is not a neighbourhood of $x_0 = 0$ (why?). ■

Solution 1.32 *(Exercise 1.32).* We recall that A is closed if, by definition, $\partial A \subseteq A$. Moreover, $\overline{A} = A \cup \partial A$. Thus,

$$A \text{ is closed} \quad \Leftrightarrow \quad \partial A \subseteq A \quad \Rightarrow \quad \partial A \cup A \subseteq A \cup A = A \quad \Rightarrow \quad \partial A \cup A \subseteq A \quad \Rightarrow \quad \overline{A} \subseteq A.$$

Since obviously $A \subseteq A \cup \partial A = \overline{A}$, we conclude that $A = \overline{A}$. Conversely,

$$\overline{A} = A \quad \Rightarrow \quad A \cup \partial A = A \quad \Rightarrow \quad \partial A \subseteq A \quad \Rightarrow \quad A \text{ is closed.}$$

■

Solution 1.33 *(Exercise 1.33).* We have to show that

$$\partial A \subseteq A \quad \Leftrightarrow \quad X \setminus A \text{ is open.}$$

We start proving $\Rightarrow$. We assume that $\partial A \subseteq A$ and we show that for every $x \in X \setminus A$ there exists $r > 0$ such that $\mathcal{B}_r(x) \subseteq X \setminus A$. Let $x \in X \setminus A$. Since $\partial A \subseteq A$, certainly $x \notin \partial A$ (otherwise x would be an element of A). This means that it is **not** true that

$$\forall\, r > 0, \quad \mathcal{B}_r(x) \cap A \neq \varnothing \quad \text{and} \quad \mathcal{B}_r(x) \cap (X \setminus A) \neq \varnothing.$$

Hence,

$$\exists\, r > 0, \quad \mathcal{B}_r(x) \cap A = \varnothing \quad \text{or} \quad \mathcal{B}_r(x) \cap (X \setminus A) = \varnothing.$$

On the other hand, the second ($\mathcal{B}_r(x) \cap (X \setminus A) = \varnothing$) is certainly false, because x belongs to both $\mathcal{B}_r(x)$ and $X \setminus A$. This implies that

$$\exists\, r > 0, \quad \mathcal{B}_r(x) \cap A = \varnothing.$$

It follows that with this r we have $\mathcal{B}_r(x) \subseteq X \setminus A$, which is what we needed to prove.

We now prove $\Leftarrow$. We assume that $X \setminus A$ is open, and we show that $\partial A \subseteq A$. Let $x \in \partial A$. We show that $x \notin X \setminus A$, so that we must have $x \in A$. By contradiction, suppose that $x \in X \setminus A$, and since we are assuming that $X \setminus A$ is open, there exists $r > 0$ such that $\mathcal{B}_r(x) \subseteq X \setminus A$, i.e., $\mathcal{B}_r(x) \cap A = \varnothing$. On the other hand, since $x \in \partial A$, for all $r > 0$ we have

$$\mathcal{B}_r(x) \cap A \neq \varnothing \quad \text{and} \quad \mathcal{B}_r(x) \cap (X \setminus A) \neq \varnothing,$$

and we reach a contradiction since $\mathcal{B}_r(x) \cap A = \varnothing$. ■

Solution 1.34 *(Exercise 1.34).*

(i) We start with the *if* part. Let $x_0 \in A$ and $x_0 \notin \mathcal{D}(A)$. Suppose that it is not true that there exists $r > 0$ such that the ball $\mathcal{B}_r(x_0)$ contains only the element x_0. Hence, for all $r > 0$, the ball $\mathcal{B}_r(x_0)$ contains other elements of A other than x_0. It follows that $x_0 \in \mathcal{D}(A)$, against our assumptions. We prove the *only if* part. We assume that $r > 0$ is such that $\mathcal{B}_r(x_0) \cap A = \{x_0\}$, so that $x_0 \in A$ and x_0 is not a limit point of A, so that $x_0 \in A \setminus \mathcal{D}(A)$.

(ii) Suppose $x_0 \in A$ is not isolated. Then, there is no $r > 0$ such that $\mathcal{B}_r(x_0) \cap A = \{x_0\}$. Since $x_0 \in A$, it follows that for every $r > 0$ we have $\{x_0\} \subset \mathcal{B}_r(x_0) \cap A$. Hence, for every $r > 0$, the set $\mathcal{B}_r(x_0) \cap A$ contains other points of A other than x_0, i.e., $x_0 \in \mathcal{D}(A)$. The converse is obviously true as well: if $x_0 \in \mathcal{D}(A)$, for every $r > 0$ the set $\mathcal{B}_r(x_0) \cap A$ contains other points of A other than x_0, and therefore $\mathcal{B}_r(x_0) \cap A \neq \{x_0\}$.

(iii) This follows from the first two:

$$(\boldsymbol{A} \setminus \mathcal{D}(\boldsymbol{A})) \cap (\boldsymbol{A} \cap \mathcal{D}(\boldsymbol{A})) = \varnothing.$$

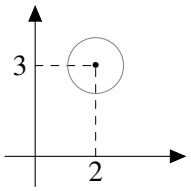
Solution 1.35 *(Exercise 1.35).* We have $\boldsymbol{x}_0 \in \boldsymbol{X} \setminus \mathring{\boldsymbol{A}}$ if and only if $\boldsymbol{x}_0 \notin \mathring{\boldsymbol{A}}$, i.e., if and only if

$$\nexists\, r > 0 : \quad \mathcal{B}_r(\boldsymbol{x}_0) \subseteq \boldsymbol{A} \quad \Leftrightarrow \quad \forall\, r > 0, \quad \mathcal{B}_r(\boldsymbol{x}_0) \cap (\boldsymbol{X} \setminus \boldsymbol{A}) \neq \varnothing \quad \Leftrightarrow \quad \boldsymbol{x}_0 \in \boldsymbol{X} \setminus \boldsymbol{A} \text{ or } \boldsymbol{x}_0 \in \partial(\boldsymbol{X} \setminus \boldsymbol{A}) \quad \Leftrightarrow \quad \boldsymbol{x}_0 \in \overline{\boldsymbol{X} \setminus \boldsymbol{A}}.$$

■

Solution 1.36 *(Exercise 1.36).* We can easily check that it is impossible to find a ball around the point $(2, 3)$ (which is the only point of $\boldsymbol{A}$) entirely contained in $\boldsymbol{A}$. This is very intuitive geometrically, but it can also be verified analytically. Let $r > 0$. Let $\mathcal{B}_r((2, 3))$. We prove that there exists $\boldsymbol{x} \in \mathcal{B}_r((2, 3))$ such that $\boldsymbol{x} \notin \boldsymbol{A}$ (which is the same as saying that $\boldsymbol{x} \neq (2, 3)$). Take for example $\boldsymbol{x} = (x, y) = (2 + \frac{r}{2}, 3)$, which is in $\mathcal{B}_r((2, 3))$ because

$$\|(x, y) - (2, 3)\| = \|(x - 2, y - 3)\| = \left\|\left(2 + \frac{r}{2} - 2, 3 - 3\right)\right\| = \left\|\left(\frac{r}{2}, 0\right)\right\| = \sqrt{\left(\frac{r}{2}\right)^2 + 0} = \frac{r}{2} < r$$



Notice that $\boldsymbol{x} = (x, y) \notin \boldsymbol{A}$, since it is not equal to $(2, 3)$. We conclude that $\boldsymbol{A}$ is not open. Its interior is $\varnothing$.

The boundary of $\boldsymbol{A}$ is $\boldsymbol{A}$ itself: in fact, for any $r > 0$, the ball $\mathcal{B}_r((2, 3))$ contains the point $(2, 3)$, which is in $\boldsymbol{A}$, and we have already seen that it must also contain points, such as $\boldsymbol{x} = (x, y) = (2 + \frac{r}{2}, 3)$, that do not belong to A. No other point can belong to $\partial\boldsymbol{A}$. Consider for example $(x_0, y_0) \neq (2, 3)$. Hence, its distance from $(2, 3)$ is

$$d = \|(x_0, y_0) - (2, 3)\| = \sqrt{(x_0 - 2)^2 + (y_0 - 3)^2} > 0.$$

Consider a radius $r < d$, and the points in the ball $\mathcal{B}_r((x_0, y_0))$: none of this points can be a point of $\boldsymbol{A}$, and therefore if a point is not in $\boldsymbol{A}$, it cannot be a boundary point. Hence, we can conclude that $\partial\boldsymbol{A} = \{(2, 3)\}$. Since $\partial\boldsymbol{A} \subseteq \boldsymbol{A}$, we can conclude that $\boldsymbol{A}$ is closed.

The point $(2, 3)$ is not a limit point: every ball $\mathcal{B}_r((2, 3))$ contains only one point of $\boldsymbol{A}$, which is $(2, 3)$ itself, and no other point (there are no other points). We say that $(2, 3)$ is an *isolated point*. ■

Solution 1.37 *(Exercise 1.37).* For every $r > 0$, the ball $\mathcal{B}_r(\mathbf{0})$ contains the element $(r/2, 0, 0, \ldots, 0)$ of $\boldsymbol{A}$, which is distinct from $\mathbf{0}$.

■

Solution 1.38 *(Exercise 1.38).* Let $x_0 \in [a, +\infty)$. Let $r > 0$. In this case $\mathcal{B}_r(x_0) = (x_0 - r, x_0 + r)$. The point $x_0 + \frac{r}{2}$ is in $[a, +\infty)$ because

$$x_0 + \frac{r}{2} > x_0 \geq a.$$

Since $x_0 + \frac{r}{2} \neq x_0$, we found that any ball centred at x_0 contains points of $[a, +\infty)$ (and also $(a, +\infty)$) different from x_0.

Consider the second part, involving a subset of $\mathbb{R}^2$. We show that

$$\mathcal{D}(\boldsymbol{U}) = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 \geq a^2\}.$$

Let $(x_0, y_0) \in \boldsymbol{U}$. Let $r > 0$. If $x_0 \geq 0$, then we can consider the point $\left(x_0 + \frac{r}{2}, y_0\right)$. This point is in $\mathcal{B}_r((x_0, y_0))$ because

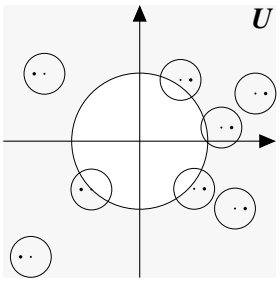
$$\left\|\left(x_0 + \frac{r}{2}, y_0\right) - (x_0, y_0)\right\| = \left\|\left(\frac{r}{2}, 0\right)\right\| = \frac{r}{2} < r.$$

Moreover, this point is in $\boldsymbol{U}$ because

$$\left(x_0 + \frac{r}{2}\right)^2 + y_0^2 = x_0^2 + x_0\, r + \frac{r^2}{4} + y_0^2 \geq x_0^2 + y_0^2 > a.$$

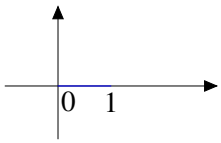
Thus, we found a point $\left(x_0 + \frac{r}{2}, y_0\right)$ in that ball and in $\boldsymbol{U}$, but distinct from (x_0, y_0). If $x_0 < 0$, we can take the point $\left(x_0 - \frac{r}{2}, y_0\right)$. In this case the point is still in $\mathcal{B}_r((x_0, y_0))$, and it is in $\boldsymbol{U}$ since $x_0\, r < 0$ and therefore

$$\left(x_0 - \frac{r}{2}\right)^2 + y_0^2 = x_0^2 - x_0\, r + \frac{r^2}{4} + y_0^2 \geq x_0^2 + y_0^2 > a.$$



■

Solution 1.39 *(Exercise 1.39).* First, no points of $\boldsymbol{A}$ are interior points. In fact, take a point $(x_0, y_0) \in \boldsymbol{A}$. We must show that we cannot find $r > 0$ such that the ball $\mathcal{B}_r((x_0, y_0))$ is entirely contained in $\boldsymbol{A}$. Since $(x_0, y_0) \in A$, we must have $y_0 = 0$, so we can write this point simply as $(x_0, 0)$, where $x_0 \in (0, 1)$.



Let $r > 0$ be arbitrary. Consider the ball $\mathcal{B}_r((x_0, 0))$. This ball contains the point $\left(x_0, \frac{r}{2}\right)$, since

$$\left\|\left(x_0, \frac{r}{2}\right) - (x_0, 0)\right\| = \left\|\left(0, \frac{r}{2}\right)\right\| = \sqrt{0 + \left(\frac{r}{2}\right)^2} = \frac{r}{2} < r.$$

However, $\left(x_0, \frac{r}{2}\right) \notin \boldsymbol{A}$ because its second coordinate is $\neq 0$. We conclude that $\mathring{\boldsymbol{A}} = \varnothing$.

Now, we want to find the set of boundary points. First, notice that every point of $\boldsymbol{A}$ is a boundary point: indeed, for each point $(x_0, 0)$, where $x_0 \in (0, 1)$, we showed that there is a ball containing points that are not in $\boldsymbol{A}$, but they contain the center of the ball which is in $\boldsymbol{A}$, and therefore these are all boundary points. It follows that $\boldsymbol{A} \subseteq \partial\boldsymbol{A}$.

We now wonder if $\boldsymbol{A} = \partial\boldsymbol{A}$, or if there are points outside of $\boldsymbol{A}$ that belongs to the boundary. Hence, we now consider points outside $\boldsymbol{A}$:

- First, we consider $(x_0, y_0) = (1, 0)$. This is a boundary point. Indeed, for every $r > 0$, the ball $\mathcal{B}_r((1, 0))$ contains points of $\boldsymbol{A}$, e.g., $\left(1 - \frac{r}{2}, 0\right)$ if $\frac{r}{2} \leq 1$ or $\left(\frac{1}{2}, 0\right)$ if $\frac{r}{2} > 1$. On the other hand, $\mathcal{B}_r((1, 0))$ contains also the point $\left(1 + \frac{r}{2}, 0\right)$, which is not a point of $\boldsymbol{A}$.
- We can adapt the same argument to see that $(0, 0)$ is also a point of $\partial\boldsymbol{A}$. For any $r > 0$, the ball $\mathcal{B}_r((0, 0))$ contains points of $\boldsymbol{A}$: for example, $\left(\frac{r}{2}, 0\right)$ if $\frac{r}{2} \leq 1$ or $\left(\frac{1}{2}, 0\right)$ if $\frac{r}{2} > 1$. On the other hand, $\mathcal{B}_r((0, 0))$ contains also the point $\left(-\frac{r}{2}, 0\right)$, which is not a point of $\boldsymbol{A}$.
- Let us now consider points (x_0, y_0) with $y_0 \neq 0$. Let $d = |y_0|$, i.e., its distance from the x-axis. Taking $\mathcal{B}_r(x_0, y_0)$ with $r < d$ ensures that the ball has empty intersection with the x-axis, and therefore with $\boldsymbol{A}$: these points are not in $\partial\boldsymbol{A}$.
- Let us consider points $(x_0, 0)$ with $x_0 > 1$. Let $r < x_0 - 1$. The ball $\mathcal{B}_r((x_0, 0))$ does not contain points of $\boldsymbol{A}$. In fact, the only points $(x, 0)$ on the x-axis that it contains (those which can be potentially in common with $\boldsymbol{A}$) must satisfy $|x - x_0| < r$, i.e., $x_0 - r < x < x_0 + r$. Since $r < x_0 - 1$, we find $x > x_0 - r > x_0 - x_0 + 1$, i.e., $x > 1$. A similar argument can be applied for the points $(x_0, 0)$ with $x_0 < 0$. Indeed, choosing $r < -x_0$, the points of the ball on the x-axis must satisfy $x_0 - r < x < x_0 + r$, and therefore $x < x_0 + r < x_0 - x_0 = 0$, i.e., $x < 0$, and therefore they do not belong to $\boldsymbol{A}$.

We can conclude that $\partial\boldsymbol{A} = [0, 1] \times \{0\}$. Thus, $\boldsymbol{A}$ is not closed because $\partial\boldsymbol{A}$ is not contained in $\boldsymbol{A}$. The closure of $\boldsymbol{A}$ coincides with $\partial\boldsymbol{A} = [0, 1] \times \{0\}$. ■

Solution 1.40 *(Exercise 1.40).* We have to show that for every point $(x_0, y_0) \in \boldsymbol{A}$, there exists $r > 0$ such that $\mathcal{B}_r((x_0, y_0)) \subseteq \boldsymbol{A}$. Since $(x_0, y_0) \in \boldsymbol{A}$, obviously $f(x_0) > y_0$. Let $\varepsilon = \frac{f(x_0) - y_0}{2}$. From the continuity of f, there exists $\delta > 0$ such that if, $|x - x_0| < \delta$, we have $|f(x) - f(x_0)| < \varepsilon$. Let $r = \min\{\varepsilon, \delta\}$. To show that $\mathcal{B}_r((x_0, y_0)) \subseteq \boldsymbol{A}$, we take $(x, y) \in \mathcal{B}_r((x_0, y_0))$. We obtain

$$|x - x_0| \leq \|(x, y) - (x_0, y_0)\| < r \leq \delta \quad \Rightarrow \quad |x - x_0| < \delta$$

and

$$|y - y_0| \leq \|(x, y) - (x_0, y_0)\| < r \leq \varepsilon \quad \Rightarrow \quad -\varepsilon < y - y_0 < \varepsilon.$$

From the first, we find $|f(x) - f(x_0)| < \varepsilon$, i.e., $-\varepsilon < f(x) - f(x_0) < \varepsilon$, so that $f(x) > f(x_0) - \varepsilon$, and from the second $y < y_0 + \varepsilon$. Subtracting gives $f(x) - y > f(x_0) - \varepsilon - y_0 - \varepsilon$, i.e.,

$$f(x) - y > f(x_0) - y_0 - 2\,\varepsilon.$$

Since $\varepsilon = \frac{f(x_0) - y_0}{2}$, we obtain

$$f(x) - y > f(x_0) - y_0 - 2\,\varepsilon \geq 0 \quad \Rightarrow \quad f(x) > y \quad \Rightarrow \quad (x, y) \in \boldsymbol{A}.$$

■

Solution 1.41 *(Exercise 1.41).*

(i) We have

$$\mathring{A} \subseteq A \quad \Rightarrow \quad \mathring{A} \cup \partial A \subseteq A \cup \partial A.$$

We need to prove the opposite inclusion: $A \cup \partial A \subseteq \mathring{A} \cup \partial A$. It suffices to show that

$$A \subseteq \mathring{A} \cup \partial A, \tag{10.3}$$

because in this case, taking the union of both sides of (10.3) with ∂A gives the desired inclusion. Let $x \in A$. It follows that for all $r > 0$ we have $\mathcal{B}_r(x) \cap A \neq \emptyset$. We can be in one and only one of the following two cases:

- $\forall\, r > 0, \quad \mathcal{B}_r(x) \cap (X \setminus A) \neq \emptyset$;
- $\exists\, r > 0 : \quad \mathcal{B}_r(x) \cap (X \setminus A) = \emptyset$.

In the first case, we have $x \in \partial A$, and in the second case we obtain that there exists $r > 0$ such that $\mathcal{B}_r(x) \cap A \neq \emptyset$ and $\mathcal{B}_r(x) \cap (X \setminus A) = \emptyset$, which imply $\mathcal{B}_r(x) \subseteq A$. In fact,

$$\mathcal{B}_r(x) = (\mathcal{B}_r(x) \cap A) \cup (\mathcal{B}_r(x) \cap (X \setminus A)) = (\mathcal{B}_r(x) \cap A),$$

which indeed implies $\mathcal{B}_r(x) \subseteq A$. Hence, $x \in \mathring{A}$. This concludes the proof of the first point.

(ii) Let $x \in \mathring{A} \cap \partial A$. Since $x \in \mathring{A}$, there exists $r > 0$ such that $\mathcal{B}_r(x) \subseteq A$, so that $\mathcal{B}_r(x) \cap (X \setminus A) = \emptyset$. This says that $x \notin \partial A$.

■

Solution 1.42 *(Exercise 1.42).* We have to prove that

$$\overline{A} = \{x \in X \mid \forall\, r > 0, \quad \mathcal{B}_r(x_0) \cap A \neq \emptyset\}.$$

Recall that $\overline{A} = A \cup \partial A$, and also $\overline{A} = \mathring{A} \cup \partial A$ (in this second case the union involves disjoint subsets of X). Thus, we have to show that

$$\mathring{A} \cup \partial A = \{x \in X \mid \forall\, r > 0, \quad \mathcal{B}_r(x) \cap A \neq \emptyset\}.$$

The inclusion $\subseteq$ is obvious: if $x \in \mathring{A}$, then $x \in A$, and therefore $\mathcal{B}_r(x) \cap A$ contains at least x, and it is therefore not empty. If $x \in \partial A$, for all $r > 0$ we have $\mathcal{B}_r(x) \cap A \neq \emptyset$ and $\mathcal{B}_r(x) \cap (X \setminus A) \neq \emptyset$, and the first guarantees that x is a point of closure.

We prove $\supseteq$. Let $x \in \{x \in X \mid \forall\, r > 0, \quad \mathcal{B}_r(x) \cap A \neq \emptyset\}$. There can be two and only two cases:

- $\mathcal{B}_r(x) \cap (X \setminus A) \neq \emptyset \quad \forall\, r > 0$;
- $\exists\, r > 0 : \quad \mathcal{B}_r(x) \cap (X \setminus A) = \emptyset$.

In the first case, since $\mathcal{B}_r(x) \cap A \neq \emptyset$, we can conclude that $x \in \partial A$. In the second case, we conclude that for $r > 0$ such that $\mathcal{B}_r(x) \cap (X \setminus A) = \emptyset$, we have $\mathcal{B}_r(x) \subseteq A$, and therefore $x \in \mathring{A}$. ■

Solution 1.43 *(Exercise 1.43).*

(i) In Question 1.42 we proved that

$$\overline{A} = \{x \in X \mid \forall\, r > 0, \quad \mathcal{B}_r(x_0) \cap A \neq \emptyset\}.$$

The negation of the statement $\forall\, r > 0, \quad \mathcal{B}_r(x_0) \cap A \neq \emptyset$ is the statement $\exists\, r > 0, \quad \mathcal{B}_r(x_0) \cap A = \emptyset$. Thus, $x_0 \notin \overline{A}$ implies $x_0 \in \mathrm{ext}(A)$, i.e., $x_0 \in X \setminus A$ implies $x_0 \in \mathrm{ext}(A)$, and *vice-versa*.

(ii) In Theorem 1.10 we proved that $\operatorname{int}(X \setminus A) = X \setminus \overline{A}$. It follows that $\operatorname{ext}(A) = \operatorname{int}(X \setminus A)$, and since the interior of a set is open, the exterior is open;

(iii) Suppose $x_0 \in \operatorname{ext}(A) \cap A$. Then, for all $r > 0$ the set $\mathcal{B}_r(x_0) \cap A$ would contain at least x_0, which is in A, and therefore it would not be empty.

(iv) Let $x_0 \in \operatorname{ext}(\mathbf{B})$. Then, there exists $r > 0$ such that $\mathcal{B}_r(x_0) \cap \mathbf{B} = \varnothing$. The intersection of both sides of this identity with A gives $\mathcal{B}_r(x_0) \cap \mathbf{B} \cap A = \varnothing \cap A$, i.e., $\mathcal{B}_r(x_0) \cap A = \varnothing$ since $A \subseteq \mathbf{B}$, and therefore $x_0 \in \operatorname{ext}(A)$;

(v) Let $G = \{G \subseteq X \mid G \text{ is open and } G \cap A = \varnothing\}$. We prove that

$$\operatorname{ext}(A) = \bigcup_{G \in G} G.$$

Since $\operatorname{ext}(A)$ is an element of G, the inclusion

$$\operatorname{ext}(A) \subseteq \bigcup_{G \in G} G$$

is straightforward. We prove the opposite inclusion. Let $x_0 \in \bigcup_{G \in G} G$. Then, there exists an open set G disjoint from A (i.e., $G \cap A = \varnothing$) such that $x_0 \in G$. Since $G \cap A = \varnothing$, it follows that $G \subseteq X \setminus A$. Since G is open, there exists $r > 0$ such that $\mathcal{B}_r(x_0) \subseteq G \subseteq X \setminus A$, so that $\mathcal{B}_r(x_0) \cap A = \varnothing$. Hence, $x_0 \in \operatorname{ext}(A)$.

(vi) Let $\mathbf{S}$ denote the largest open subset of X that is disjoint from A. From the previous point,

$$\operatorname{ext}(A) = \bigcup_{G \in G} G,$$

where $G = \{G \subseteq X \mid G \text{ is open and } G \cap A = \varnothing\}$. Since $\mathbf{S}$ is an element of G, then $\mathbf{S} \subseteq \bigcup_{G \in G} G = \operatorname{ext}(A)$. However, $\operatorname{ext}(A)$ is an open set disjoint from A, so that $\operatorname{ext}(A) \subseteq \mathcal{S}$. Therefore, they must be equal.

■

Solution 1.44 *(Exercise 1.44).* We prove that the following two statements are equivalent:

(A) A is closed, i.e., $\partial A \subseteq A$;

(B) for all $x_0 \in X$, if x_0 is a limit point of A, then $x_0 \in A$.

We show that **(A)** implies **(B)**. Let x_0 be a limit point of A. Then, for all $r > 0$,

$$(A \cap \mathcal{B}_r(x_0)) \setminus \{x_0\} \neq \varnothing. \tag{10.4}$$

We must show that $x_0 \in A$. Suppose by contradiction that $x_0 \notin A$, so that $x_0 \in X \setminus A$. Since A is closed, $X \setminus A$ is open, and therefore there exists $r > 0$ such that $\mathcal{B}_r(x_0) \subseteq X \setminus A$. For this value of r, (10.4) is violated.

We now prove that **(B)** implies **(A)**. Suppose by contradiction that A is not closed, so that ∂A is not contained in A. Thus, there exists $x_0 \in \partial A$ which is not in A. On the other hand, x_0 is a limit point[1], and therefore we found a limit point that is not in A, contradicting **(B)**. ■

[1] Indeed, since $x_0 \in \partial A$, for every $r > 0$ we have $\mathcal{B}_r(x_0) \cap A \neq \varnothing$. Since $x_0 \notin A$, we have $x_0 \notin \mathcal{B}_r(x_0) \cap A$, so that $\mathcal{B}_r(x_0) \cap A$ must contain some other element different from x_0, which means that x_0 is a limit point.

Solution 1.45 *(Exercise 1.45).* Since $\mathring{\boldsymbol{A}} \subseteq \boldsymbol{A}$ and $\mathring{\boldsymbol{A}}$ is open, we have $\mathring{\boldsymbol{A}} \in \{G \subseteq \boldsymbol{A} \,|\, G \text{ is open}\}$, so that

$$\mathring{\boldsymbol{A}} \subseteq \max\{G \subseteq \boldsymbol{A} \,|\, G \text{ is open}\}.$$

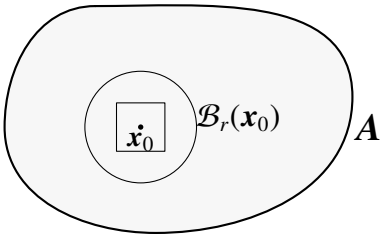
To show the converse, let $\mathcal{S} = \max\{G \subseteq \boldsymbol{A} \,|\, G \text{ is open}\}$. Clearly, $\mathcal{S} \subseteq \boldsymbol{A}$. Take $\boldsymbol{x}_0 \in \mathcal{S}$. Then, $\boldsymbol{x}_0 \in \boldsymbol{A}$ and there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}_0) \subseteq \mathcal{S} \subseteq \boldsymbol{A}$. Hence, $\boldsymbol{x}_0 \in \mathring{\boldsymbol{A}}$. ■

Solution 1.46 *(Exercise 1.46).* Since $\boldsymbol{A}$ is open, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq \boldsymbol{A}$. Then, k only needs to satisfy

$$\prod_{j=1}^{n} \left[x_{0,j} - \frac{1}{k}, x_{0,j} + \frac{1}{k} \right] \subseteq \mathcal{B}_r(\boldsymbol{x}_0).$$

Consider for example the case $n = 2$. Then, the latter becomes

$$\left[x_{0,1} - \frac{1}{k}, x_{0,1} + \frac{1}{k} \right] \times \left[x_{0,2} - \frac{1}{k}, x_{0,2} + \frac{1}{k} \right] \subseteq \left\{ (x_1, x_2) \in \mathbb{R}^2 \,\middle|\, \sqrt{(x_1 - x_{0,1})^2 + (x_2 - x_{0,2})^2} < r \right\}.$$



For this to hold, it is necessary and sufficient that all the vertexes of the square be contained in $\mathcal{B}_r(x_{0,1}, x_{0,2})$, i.e.,

$$\left(x_{0,1} - \frac{1}{k}, x_{0,2} - \frac{1}{k} \right), \left(x_{0,1} - \frac{1}{k}, x_{0,2} + \frac{1}{k} \right), \left(x_{0,1} + \frac{1}{k}, x_{0,2} - \frac{1}{k} \right), \left(x_{0,1} + \frac{1}{k}, x_{0,2} + \frac{1}{k} \right) \in \mathcal{B}_r(x_{0,1}, x_{0,2}).$$

This is equivalent to imposing that just one point is in $\mathcal{B}_r(x_{0,1}, x_{0,2})$. Consider the first point. We must have

$$\sqrt{\left(x_{0,1} - \frac{1}{k} - x_{0,1} \right)^2 + \left(x_{0,2} - \frac{1}{k} - x_{0,2} \right)^2} < r$$

leads to

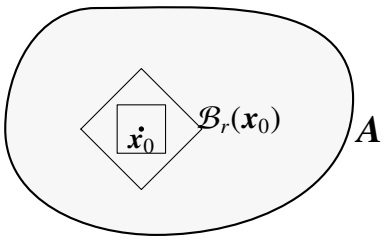
$$\frac{1}{k^2} + \frac{1}{k^2} < r^2 \quad \Rightarrow \quad k > \frac{\sqrt{2}}{r}.$$

In general we will have the condition

$$k > \frac{\sqrt{n}}{r}.$$

If, instead of the Euclidean norm, we use the 1-norm, then

$$\mathcal{B}_r(x_{0,1}, x_{0,2}) = \left\{ (x_1, x_2) \in \mathbb{R}^2 \,\middle|\, |x_1 - x_{0,1}| + |x_2 - x_{0,2}| < r \right\}.$$



In this case, again, the condition is only one vertex, and choosing the same vertex chosen before, we have

$$\left|x_{0,1} - \frac{1}{k} - x_{0,1}\right| + \left|x_{0,2} - \frac{1}{k} - x_{0,2}\right| < r,$$

which gives $k > \frac{2}{r}$. In the case of $\mathbb{R}^n$, this condition becomes $k > \frac{n}{r}$.

It is easy to see that in the case of the p-norm, we have the condition $k > \frac{n^{\frac{1}{p}}}{r}$, since

$$\left(\left|x_1 - x_{0,1}\right|^p + \left|x_2 - x_{0,2}\right|^p + \ldots + \left|x_n - x_{0,n}\right|^p\right)^{\frac{1}{p}} < r$$

which gives

$$\left(\frac{1}{k^p} + \ldots + \frac{1}{k^p}\right)^{\frac{1}{p}} < r$$

i.e.,

$$\frac{n}{k^p} < r^p \quad \Rightarrow \quad k > \frac{n^{\frac{1}{p}}}{r}.$$

In the case $p = +\infty$, this becomes

$$\max\left\{\frac{1}{k}, \ldots, \frac{1}{k}\right\} < r,$$

which gives $k > 1/r$. ■

Solution 1.47 *(Exercise 1.47).* The segment $[\boldsymbol{x}_0, \boldsymbol{y}_0]$ belongs to the ball $\mathcal{B}_1(\boldsymbol{0}, \boldsymbol{X}, \|\cdot\|)$ if and only if the condition

$$\forall\, \lambda \in [0,1], \quad \|\lambda\, \boldsymbol{x}_0 + (1-\lambda)\, \boldsymbol{y}_0\| = 1 \tag{10.5}$$

is satisfied. In particular for $\lambda = 1/2$ we find

$$\frac{1}{2}\,\|\boldsymbol{x}_0 + \boldsymbol{y}_0\| = 1,$$

and, moreover, $\|\boldsymbol{x}_0\| = \|\boldsymbol{y}_0\| = 1$. It follows that $\|\boldsymbol{x}_0 + \boldsymbol{y}_0\| = 2 = 1 + 1 = \|\boldsymbol{x}_0\| + \|\boldsymbol{y}_0\|$, as required. Now we show that $\boldsymbol{x}_0$ and $\boldsymbol{y}_0$ are linearly independent. Suppose by contradiction that they are linearly dependent. The linear dependence of two vectors implies that one of the two is a multiple of the other, say for the sake of argument $\boldsymbol{x}_0 = \kappa\, \boldsymbol{y}_0$. Condition (10.5) can be written as

$$1 = \|\lambda\, \kappa\, \boldsymbol{y}_0 + (1-\lambda)\, \boldsymbol{y}_0\| = |\lambda\, \kappa + 1 - \lambda|\, \|\boldsymbol{y}_0\| = |\lambda\, \kappa + 1 - \lambda|.$$

If $\lambda = 1$ we find $|\kappa| = 1$. If $\lambda = 1/2$ we find

$$\left|\frac{1}{2}\,\kappa + \frac{1}{2}\right| = \frac{1}{2}\,|k+1| = 1 \quad \Rightarrow \quad |\kappa + 1| = 2.$$

It follows that $\kappa = 1$, and so $\boldsymbol{x}_0 = \boldsymbol{y}_0$. ■

Solution 1.48 *(Exercise 1.48).* Since every vector space contains the origin, if $\mathbf{0} \notin \mathcal{B}_r(\boldsymbol{x}_0)$ there are no subspaces of $\boldsymbol{X}$ contained in $\mathcal{B}_r(\boldsymbol{x}_0)$. Consider the case $\mathbf{0} \in \mathcal{B}_r(\boldsymbol{x}_0)$. Let $\boldsymbol{W}$ be a subspace of $\boldsymbol{X}$ contained in $\mathcal{B}_r(\boldsymbol{x}_0)$, so that

$$\forall\, \boldsymbol{x} \in \boldsymbol{W}, \quad \|\boldsymbol{x} - \boldsymbol{x}_0\| < r.$$

Now, if $\boldsymbol{x} \in \boldsymbol{W}$, we know that $k\,\boldsymbol{x} \in \boldsymbol{W}$ for all $k \in \mathbb{F}$ (where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$). It follows that

$$\forall\, k \in \mathbb{F}, \quad \|k\,\boldsymbol{x} - \boldsymbol{x}_0\| < r.$$

Since $\|k\,\boldsymbol{x}\| \leq \|k\,\boldsymbol{x} - \boldsymbol{x}_0\| + \|\boldsymbol{x}_0\| < r + \|\boldsymbol{x}_0\|$, we find

$$\forall\, k \in \mathbb{F}, \quad \|\boldsymbol{x}\| < \frac{r + \|\boldsymbol{x}_0\|}{k}.$$

The right hand-side can be made as small as we like by taking k sufficiently large, and so we obtain $\|\boldsymbol{x}\| = 0$, which implies $\boldsymbol{x} = \mathbf{0}$.

■

Solution 1.49 *(Exercise 1.49).* Let $\boldsymbol{W}$ be a subspace of $\boldsymbol{X}$ containing the ball $\mathcal{B}_r(\boldsymbol{x}_0)$. We prove that $\boldsymbol{W}$ contains also the ball $\mathcal{B}_r(\mathbf{0})$. Indeed,

$$\boldsymbol{x} \in \mathcal{B}_r(\mathbf{0}) \quad \Leftrightarrow \quad \boldsymbol{x} + \boldsymbol{x}_0 \in \mathcal{B}_r(\boldsymbol{x}_0) \quad \Rightarrow \quad \boldsymbol{x} + \boldsymbol{x}_0 \in \boldsymbol{W} \quad \Rightarrow \quad \boldsymbol{x} \in \boldsymbol{W},$$

where the last implication is due to the fact that $\boldsymbol{x}_0 \in \boldsymbol{W}$ and $\boldsymbol{W}$ is closed under addition. Now, for every $\boldsymbol{x} \in \boldsymbol{X}$, we have

$$\frac{r}{2}\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in \mathcal{B}_r(\mathbf{0}) \subseteq \boldsymbol{W},$$

and since $\boldsymbol{W}$ is a subspace we obtain $\boldsymbol{x} \in \boldsymbol{W}$. We conclude that $\boldsymbol{X} \subseteq \boldsymbol{W}$, and so $\boldsymbol{W} = \boldsymbol{X}$.

■

Solution 1.50 *(Exercise 1.50).* Let $\varepsilon > 0$. Let $n = \lceil 1/\varepsilon \rceil + 1$. Then, $\mathcal{B}_\varepsilon(0) \cap A$ contains $1/n \in A$. This shows that 0 is a limit point. Moreover, A has no other limit points, since for an arbitrary point $1/n \in A$ and for

$$\varepsilon = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+1}\right)$$

we have $\mathcal{B}_\varepsilon(0) \cap A = \varnothing$. ■

Solution 1.51 *(Exercise 1.51).* The condition "$\boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}$ such that $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$" is the same as the condition "$\boldsymbol{x} \in \boldsymbol{A}$ such that $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$", because in both cases $\boldsymbol{x}$ is not allowed to be equal to $\boldsymbol{x}_0$: in the first case we explicitly take at the outset $\boldsymbol{x}$ in $\boldsymbol{A}$ "minus" $\{\boldsymbol{x}_0\}$, while in the second case this is ensured by imposing the norm $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X$ to be different from zero (recall that $\|\boldsymbol{y}\| = 0$ if and only if $\boldsymbol{y} = 0$). ■

Solution 1.52 *(Exercise 1.52).* To say that (1.28) and (1.29) are equivalent is the same as saying that (1.28) holds if and only if (1.29) holds. Thus, we have to prove two things:

(i) that (1.28) implies (1.29)

(ii) that (1.29) implies (1.28).

We show that (1.28) implies (1.29). Thus, we assume that (1.28) holds. We have to prove (1.29). We re-write (1.29) to remind ourselves of what we have to prove:

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon.$$

This means choosing an arbitrary $\varepsilon > 0$, and then proving that $\delta > 0$ exists which satisfies the conditions. We cannot give ε a numerical value: we simply assume that it has been chosen, and it is now an assigned positive quantity, that we continue to call ε. We will express this fact by saying "Let $\varepsilon > 0$". Now, we must prove the existence of a value $\delta > 0$ such that for all $\boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}$, if we have $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$, then $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$. The only piece of information to show that such value of δ exists is (1.28). We therefore re-write (1.28), but we use different symbol for its ε, to avoid confusion between the (already assigned) value of ε, and we do the same for δ:

$$\forall\, \tilde{\varepsilon} > 0,\ \exists\, \tilde{\delta} > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \tilde{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \tilde{\varepsilon}. \tag{10.6}$$

Since in this equation $\tilde{\varepsilon}$ is chosen arbitrarily, we can choose it in particular to be equal to the pre-assigned ε. Thus, we know that there exists $\tilde{\delta} > 0$ such that choosing $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \tilde{\delta}$ will give $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \tilde{\varepsilon} = \varepsilon$, and therefore we will have also $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$. This means that the desired value of δ can be exactly equal to $\tilde{\delta}$.

We show that (1.29) implies (1.28). Thus, now we need to show that

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$$

knowing that

$$\forall\, \hat{\varepsilon} > 0,\ \exists\, \hat{\delta} > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \hat{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \hat{\varepsilon},$$

where, again, we used the symbols $\hat{\varepsilon}$ and $\hat{\delta}$ to avoid confusion between the two conditions. Let $\varepsilon > 0$. We must find $\delta > 0$ such that, if $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$, we have $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. This time we cannot select $\hat{\varepsilon} = \varepsilon$, because $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \varepsilon$ does not imply $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. Since $\hat{\varepsilon}$ is arbitrary, we could, however, select $\hat{\varepsilon} = \varepsilon/2$, because we can say that $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y \leq \frac{\varepsilon}{2}$ implies $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$ (since $\frac{\varepsilon}{2} < \varepsilon$, and ε is a strictly positive quantity). Therefore, choosing $\hat{\varepsilon} = \varepsilon/2$, there exists $\hat{\delta} > 0$ such that $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \hat{\delta}$ implies $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. Taking $\delta = \hat{\delta}$ therefore allows us to conclude the argument. Notice that we could have taken $\hat{\varepsilon} = \varepsilon/3$, or $\hat{\varepsilon} = \varepsilon/4$, etc. ■

Solution 1.53 *(Exercise 1.53).* We show that (1.30) implies (1.31). Hence, we must show that

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

Let $\varepsilon > 0$. We know that, using once again different symbols for the variables ε and δ to avoid confusion,

$$\forall\, \check{\varepsilon} > 0,\ \exists\, \check{\delta} > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \check{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \check{\varepsilon}.$$

We can take $\check{\varepsilon} = \varepsilon$. Let $\tilde{\delta} > 0$ be such that $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \tilde{\delta}$ implies $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. It follows that

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \frac{\tilde{\delta}}{2} \ \Rightarrow\ \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \tilde{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

Therefore, δ exists, and a possible value is $\tilde{\delta}/2$.

Now, we show that (1.31) implies (1.30). Hence, we must show that

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

Let $\varepsilon > 0$. We know that

$$\forall\, \tilde{\varepsilon} > 0,\ \exists\, \tilde{\delta} > 0 :\ \forall\, \boldsymbol{x} \in \boldsymbol{A} \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \tilde{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \tilde{\varepsilon}.$$

We can choose $\tilde{\varepsilon} = \varepsilon$. Let $\tilde{\delta} > 0$ be such that $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \tilde{\delta} \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$. Therefore, δ exists, and a possible value is $\tilde{\delta}/2$, since

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \tilde{\delta} \quad \Rightarrow \quad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X \leq \tilde{\delta} \quad \Rightarrow \quad \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

■

Solution 1.54 *(Exercise 1.54).* Consider the first: since $\varepsilon \geq 0$ is arbitrary, we can choose for example $\varepsilon = 0$, and in this case $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$ becomes $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < 0$, which is never satisfied. In this case, a function never has a limit at any point.

Consider the second. Suppose that, given $\varepsilon > 0$, the value $\delta = 0$ satisfies the implication

$$0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta = 0 \quad \Rightarrow \quad \|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon.$$

Since $0 < \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < 0$ is always false, the conclusion $\|f(\boldsymbol{x}) - \boldsymbol{\ell}\|_Y < \varepsilon$ is always true, and this means that every $\boldsymbol{\ell}$ is the limit (thus contradicting the uniqueness of the limit). ■

Solution 1.55 *(Exercise 1.55).* Since $\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) = \infty$,

$$\forall\, M > 0,\ \ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0) :\ \ \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad |f(\boldsymbol{x})| > M. \tag{10.7}$$

Any $M > 0$ can be written as $1/\varepsilon$, with $\varepsilon > 0$, and *vice-versa*, so that (10.7) is equivalent to

$$\forall\, \varepsilon > 0,\ \ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0) :\ \ \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad |f(\boldsymbol{x})| > \frac{1}{\varepsilon}, \tag{10.8}$$

which is obviously equivalent to

$$\forall\, \varepsilon > 0,\ \ \exists\, \boldsymbol{V} \in \mathcal{N}(\boldsymbol{x}_0) :\ \ \forall\, \boldsymbol{x} \in (\boldsymbol{V} \cap \boldsymbol{A}) \setminus \{\boldsymbol{x}_0\}, \qquad \frac{1}{|f(\boldsymbol{x})|} < \varepsilon. \tag{10.9}$$

This shows that $\frac{1}{f(\boldsymbol{x})}$ goes to 0 as $\boldsymbol{x} \to \boldsymbol{x}_0$. ■

Solution 1.56 *(Exercise 1.56).* We re-write (1.32) as

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(2)} \geq \frac{1}{c_1} \|\boldsymbol{x}\|_{(1)} \ \text{ and } \ \|\boldsymbol{x}\|_{(2)} \leq c_2 \|\boldsymbol{x}\|_{(1)},$$

which is the same as

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \frac{1}{c_1} \|\boldsymbol{x}\|_{(1)} \leq \|\boldsymbol{x}\|_{(2)} \leq c_2 \|\boldsymbol{x}\|_{(1)}.$$

This implies $c_2 \geq \frac{1}{c_1}$. Using the substitutions $m = 1/c_1$ and $M = c_2$, (1.34) is seen to be equivalent to (1.32). ■

Solution 1.57 *(Exercise 1.57).*

(i) The condition

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \leq c_1 \|\boldsymbol{x}\|_{(1)} \text{ and } \|\boldsymbol{x}\|_{(1)} \leq c_2 \|\boldsymbol{x}\|_{(1)}$$

is satisfied using $c_1 = c_2 = 1$;

(ii) Suppose $\|\cdot\|_1, \|\cdot\|_2$, so that there exist $c_1, c_2 \in \mathbb{R}$ such that

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \leq c_1 \|\boldsymbol{x}\|_{(2)} \text{ and } \|\boldsymbol{x}\|_{(2)} \leq c_2 \|\boldsymbol{x}\|_{(1)}.$$

Swapping c_1 and c_2 yields the result;

(iii) Suppose $\|\cdot\|_1, \|\cdot\|_2$ and $\|\cdot\|_2, \|\cdot\|_3$, so that there exist $c_1', c_2' \in \mathbb{R}$ such that

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(1)} \leq c_1' \|\boldsymbol{x}\|_{(2)} \text{ and } \|\boldsymbol{x}\|_{(2)} \leq c_2' \|\boldsymbol{x}\|_{(1)},$$

and there exist $c_1'', c_2'' \in \mathbb{R}$ such that

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_{(2)} \leq c_1'' \|\boldsymbol{x}\|_{(3)} \text{ and } \|\boldsymbol{x}\|_{(3)} \leq c_2'' \|\boldsymbol{x}\|_{(2)},$$

It follows that for all $\boldsymbol{x} \in \boldsymbol{X}$

$$\|\boldsymbol{x}\|_{(1)} \leq c_1' \|\boldsymbol{x}\|_{(2)} \leq \underbrace{c_1' c_1''}_{c_1} \|\boldsymbol{x}\|_{(3)}$$

and

$$\|\boldsymbol{x}\|_{(3)} \leq c_2'' \|\boldsymbol{x}\|_{(2)} \leq \underbrace{c_2'' c_2'}_{c_2} \|\boldsymbol{x}\|_{(1)}.$$

■

Solution 1.58 *(Exercise 1.58).* If $p = \infty$ or $q = \infty$, the statement has been proved in the lecture notes. We assume that both p and q are finite (and both ≥ 1). We have seen that $\|\cdot\|_p$ and $\|\cdot\|_\infty$ in $\mathbb{R}^n$ are equivalent, so that

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty.$$

Likewise,

$$\forall\, \boldsymbol{x} \in \mathbb{R}^n, \quad \|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_q \leq n^{\frac{1}{q}} \|\boldsymbol{x}\|_\infty.$$

It follows that

$$n^{-\frac{1}{q}} \|\boldsymbol{x}\|_q \leq \|\boldsymbol{x}\|_\infty \leq \|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}} \|\boldsymbol{x}\|_\infty \leq n^{\frac{1}{p}} \|\boldsymbol{x}\|_q,$$

so that $n^{-\frac{1}{q}} \|\boldsymbol{x}\|_q \leq \|\boldsymbol{x}\|_p \leq n^{\frac{1}{p}} \|\boldsymbol{x}\|_q$, which shows that the two norms are equivalent. Taking $c_1 = n^{\frac{1}{p}}$ and $c_2 = n^{\frac{1}{q}}$, we obtain

$$\forall\, \boldsymbol{x} \in \boldsymbol{X}, \qquad \|\boldsymbol{x}\|_p \leq c_1 \|\boldsymbol{x}\|_q \text{ and } \|\boldsymbol{x}\|_q \leq c_2 \|\boldsymbol{x}\|_p.$$

■

Solution 1.59 *(Exercise 1.59).* We start considering the first norm. We prove the positive homogeneity. Let $f \in C^1([a,b], \mathbb{R})$ and $k \in \mathbb{R}$:

$$\begin{aligned} \|k\, f\|_1 &= \max_{a \leq x \leq b} |k\, f(x)| + \max_{a \leq x \leq b} |k\, f'(x)| \\ &= \max_{a \leq x \leq b} |k|\, |f(x)| + \max_{a \leq x \leq b} |k|\, |f'(x)| \\ &= |k| \left(\max_{a \leq x \leq b} |f(x)| + \max_{a \leq x \leq b} |f'(x)| \right) = |k|\, \|f\|_1. \end{aligned}$$

We prove the subadditivity. Let $f, g \in C^1([a,b],\mathbb{R})$:

$$\begin{aligned}\|f+g\|_1 &= \max_{a\le x\le b}|(f+g)(x)| + \max_{a\le x\le b}|(f+g)'(x)| \\ &= \max_{a\le x\le b}|f(x)+g(x)| + \max_{a\le x\le b}|f'(x)+g'(x)| \\ &\le \max_{a\le x\le b}(|f(x)|+|g(x)|) + \max_{a\le x\le b}(|f'(x)|+|g'(x)|) \\ &\le \max_{a\le x\le b}|f(x)| + \max_{a\le x\le b}|g(x)| + \max_{a\le x\le b}|f'(x)| + \max_{a\le x\le b}|g'(x)| = \|f\|_1 + \|g\|_1.\end{aligned}$$

We prove the positivity: $\|f\|_1 = 0$ implies

$$\max_{a\le x\le b}|f(x)| + \max_{a\le x\le b}|f'(x)| = 0,$$

which implies $f(x) = 0$ and $f'(x) = 0$ for all $x \in [a,b]$, so that $f(x) = 0$ for all $x \in [a,b]$. We consider the second norm. We prove the positive homogeneity. Let $f \in C^1([a,b],\mathbb{R})$ and $k \in \mathbb{R}$:

$$\|k\,f\|_2 = |k\,f(a)| + \max_{a\le x\le b}|k\,f'(x)| = |k|\,|f(a)| + \max_{a\le x\le b}|k|\,|f'(x)| = |k|\,|f(a)| + |k|\max_{a\le x\le b}|f'(x)| = |k|\,\|f\|_2.$$

We prove the subadditivity. Let $f, g \in C^1([a,b],\mathbb{R})$:

$$\begin{aligned}\|f+g\|_2 &= |(f+g)(a)| + \max_{a\le x\le b}|(f+g)'(x)| \\ &= |f(a)+g(a)| + \max_{a\le x\le b}|f'(x)+g'(x)| \\ &\le |f(a)| + |g(a)| + \max_{a\le x\le b}(|f'(x)|+|g'(x)|) \\ &\le |f(a)| + |g(a)| + \max_{a\le x\le b}|f'(x)| + \max_{a\le x\le b}|g'(x)| = \|f\|_2 + \|g\|_2.\end{aligned}$$

We prove the positivity: $\|f\|_2 = 0$ implies

$$|f(a)| + \max_{a\le x\le b}|f'(x)| = 0,$$

which implies $f(a) = 0$ and $f'(x) = 0$ for all $x \in [a,b]$. From the second condition f is constant in $[a,b]$, and from the first condition that constant is zero, i.e., $f(x) = 0$ for all $x \in [a,b]$.

We now show that the two norms are equivalent. First, notice that $\|f\|_1 \ge \|f\|_2$, so that the first constant is $c_1 = 1$. We now show that there exists $c_2 > 0$ such that

$$\max_{a\le x\le b}|f(x)| + \max_{a\le x\le b}|f'(x)| \le c_2\,|f(a)| + c_2\max_{a\le x\le b}|f'(x)|.$$

We use the fundamental theorem of calculus and write

$$f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t.$$

We find

$$\begin{aligned}|f(x)| &\le |f(a)| + \left|\int_a^x f'(t)\,\mathrm{d}t\right| \\ &\le |f(a)| + \int_a^x |f'(t)|\,\mathrm{d}t \\ &\le |f(a)| + (x-a)\max_{t\in[a,x]}|f'(t)|.\end{aligned}$$

Taking the maximum over $[a, b]$ on both sides gives

$$\begin{aligned}
\max_{a \le x \le b} |f(x)| &\le |f(a)| + \max_{x \in [a,b]} \left((x-a) \max_{t \in [a,x]} |f'(t)|\right) \\
&\le |f(a)| + \left(\max_{x \in [a,b]} (x-a)\right)\left(\max_{t \in [a,x]} |f'(t)|\right) \\
&\le |f(a)| + \left(\max_{x \in [a,b]} (x-a)\right)\left(\max_{x \in [a,b]} \max_{t \in [a,x]} |f'(t)|\right) \\
&= |f(a)| + (b-a) \max_{x \in [a,b]} |f'(x)| \\
&\le (b-a+1)\,|f(a)| + (b-a) \max_{x \in [a,b]} |f'(x)|.
\end{aligned}$$

Here, we used the fact that $\max |f\, g| \le \max |f| \max |g|$ (why?). Adding $\max_{x \in [a,b]} |f'(x)|$ on both sides gives

$$\max_{x \in [a,b]} |f(x)| + \max_{x \in [a,b]} |f'(x)| \le (b-a+1)\,|f(a)| + (b-a+1) \max_{x \in [a,b]} |f'(x)|.$$

Therefore, $c_2 = b - a + 1$. ■

Solution 1.60 *(Exercise 1.60).* Let $M_1 > 0$ be given. We have to show that, for all $\boldsymbol{x} \in \boldsymbol{X}$, if $\|\boldsymbol{x}\|_{(1)} \le M_1$, then $\|\boldsymbol{x}\|_{(2)} \le M_2$ for a certain M_2. We have

$$\|\boldsymbol{x}\|_{(2)} \le c_1 \|\boldsymbol{x}\|_{(1)} < c_1 M_1.$$

Taking $M_2 = M_1 c_1$ solves the problem. ■

Solution 1.61 *(Exercise 1.61).* We denote by p the function

$$p(\boldsymbol{x}) = \alpha_1 \|\boldsymbol{x}\|_{(1)} + \alpha_2 \|\boldsymbol{x}\|_{(2)} + \ldots + \alpha_q \|\boldsymbol{x}\|_{(q)}.$$

We need to show that p is a norm. We prove the homogeneity. Let $k \in \mathbb{F}$ and $\boldsymbol{x} \in \boldsymbol{X}$:

$$\begin{aligned}
p(k\,\boldsymbol{x}) &= \alpha_1 \|k\,\boldsymbol{x}\|_{(1)} + \alpha_2 \|k\,\boldsymbol{x}\|_{(2)} + \ldots + \alpha_q \|k\,\boldsymbol{x}\|_{(q)} \\
&= \alpha_1 |k| \|\boldsymbol{x}\|_{(1)} + \alpha_2 |k| \|\boldsymbol{x}\|_{(2)} + \ldots + \alpha_q |k| \|\boldsymbol{x}\|_{(q)} \\
&= |k| \left(\alpha_1 \|\boldsymbol{x}\|_{(1)} + \alpha_2 \|\boldsymbol{x}\|_{(2)} + \ldots + \alpha_q \|\boldsymbol{x}\|_{(q)}\right) = |k|\, p(\boldsymbol{x}).
\end{aligned}$$

We prove the triangle inequality. Let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. Then,

$$\begin{aligned}
p(\boldsymbol{x}+\boldsymbol{y}) &= \alpha_1 \|\boldsymbol{x}+\boldsymbol{y}\|_{(1)} + \alpha_2 \|\boldsymbol{x}+\boldsymbol{y}\|_{(2)} + \ldots + \alpha_q \|\boldsymbol{x}+\boldsymbol{y}\|_{(q)} \\
&\le \alpha_1 \|\boldsymbol{x}\|_{(1)} + \alpha_1 \|\boldsymbol{y}\|_{(1)} + \alpha_2 \|\boldsymbol{x}\|_{(2)} + \alpha_2 \|\boldsymbol{y}\|_{(2)} + \ldots + \alpha_n \|\boldsymbol{x}\|_{(n)} + \alpha_n \|\boldsymbol{y}\|_{(n)} \\
&= \alpha_1 \|\boldsymbol{x}\|_{(1)} + \alpha_2 \|\boldsymbol{x}\|_{(2)} + \ldots + \alpha_n \|\boldsymbol{x}\|_{(n)} + \alpha_1 \|\boldsymbol{y}\|_{(1)} + \alpha_2 \|\boldsymbol{y}\|_{(2)} + \ldots + \alpha_n \|\boldsymbol{y}\|_{(n)} = p(\boldsymbol{x}) + p(\boldsymbol{y}).
\end{aligned}$$

We prove the positivity. Suppose $p(\boldsymbol{x}) = 0$, so that

$$\alpha_1 \|\boldsymbol{x}\|_{(1)} + \alpha_2 \|\boldsymbol{x}\|_{(2)} + \ldots + \alpha_q \|\boldsymbol{x}\|_{(q)} = 0.$$

Since $\alpha_i \ge 0$ and $\|\boldsymbol{x}\|_{(i)} \ge 0$, we must have $\alpha_i \|\boldsymbol{x}\|_{(i)} = 0$ for all $i = 1, \ldots, q$. Since there is at least one $\alpha_i \ne 0$, the corresponding norm $\|\boldsymbol{x}\|_{(i)}$ must be zero, and this implies that $\boldsymbol{x} = \boldsymbol{0}$. ■

Solution 1.62 *(Exercise 1.62).* The first is a norm. The second is not because $p(0) \neq 0$. The third is not because $p(1,0,1) = 0$. The fourth is not because $p(1,1,0) = 0$. The fifth is not because $p(1,1,1) = 0$. The sixth is a norm. In fact, the homogeneity holds:

$$p(k\,x_1, k\,x_2, k\,x_3) = \max\{|k\,x_1|, |k\,x_2|, |k\,x_3|\} + |k\,x_1 + k\,x_2| = |k|\,(\max\{|x_1|, |x_2|, |x_3|\} + |x_1 + x_2|).$$

The triangle inequality also holds:

$$\begin{aligned} p(\boldsymbol{x} + \boldsymbol{y}) &= \max\{|x_1 + y_1|, |x_2 + y_2|, |x_3 + y_3|\} + |x_1 + y_1 + x_2 + y_2| \\ &\leq \max\{|x_1|, |x_2|, |x_3|\} + |x_1 + x_2| + \max\{|y_1|, |y_2|, |y_3|\} + |y_1 + y_2|. \end{aligned}$$

Finally, since $p(x_1, x_2, x_3)$ is the sum of two non-negative quantities, we have $p(x_1, x_2, x_3) = 0$ if and only if $\max\{|x_1|, |x_2|, |x_3|\} = 0$, which implies $x_1 = x_2 = x_3 = 0$. The seventh is not a norm because $p(0,1,2) = 0$. The eighth is not a norm because $p(-1,-2,-3) < 0$. The last is not a norm because the homogeneity is not satisfied. ■

Solution 1.63 *(Exercise 1.63).* The proof that $p(\mathbf{0}) = 0$ and $p(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \boldsymbol{X}$ can be carried out exactly as in the proof of Theorem 2: given any $\boldsymbol{x} \in \boldsymbol{X}$

$$p(\mathbf{0}) = p(0\,\boldsymbol{x}) = |0|\,p(\boldsymbol{x}) = 0$$

and

$$p(-\boldsymbol{x}) = p(-1\,\boldsymbol{x}) = |-1|\,p(\boldsymbol{x}) = p(\boldsymbol{x}).$$

Thus,

$$0 = p(\mathbf{0}) = p(\boldsymbol{x} + (-x)) \leq p(\boldsymbol{x}) + p(-\boldsymbol{x}) = p(\boldsymbol{x}) + p(\boldsymbol{x}) = 2\,p(\boldsymbol{x}) \quad \Rightarrow \quad p(\boldsymbol{x}) \geq 0.$$

Now, given a norm $\|\cdot\|$ in $\mathbb{R}^m$, and let us consider the function

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \|A\,\boldsymbol{x}\|. \end{aligned}$$

Using the fact that $\|\cdot\|$ is a norm, we find

$$p(k\,\boldsymbol{x}) = \|A\,(k\,\boldsymbol{x})\| = \|k\,A\,\boldsymbol{x}\| = |k|\,\|A\,\boldsymbol{x}\| = |k|\,p(\boldsymbol{x}),$$

and

$$p(\boldsymbol{x} + \boldsymbol{y}) = \|A(\boldsymbol{x} + \boldsymbol{y})\| = \|A\,\boldsymbol{x} + A\,\boldsymbol{y}\| \leq \|A\,\boldsymbol{x}\| + \|A\,\boldsymbol{y}\| = p(\boldsymbol{x}) + p(\boldsymbol{y}).$$

Finally, we have $p(\boldsymbol{x}) = \|A\,\boldsymbol{x}\| \geq 0$. It follows that p is a seminorm in $\mathbb{R}^n$. We prove that it is a norm if and only if $\operatorname{rank} A = n$. If $\operatorname{rank} A = n$, then $\ker A = \{0\}$ (rank-nullity theorem), and therefore

$$p(\boldsymbol{x}) = 0 \quad \Leftrightarrow \quad \|A\,\boldsymbol{x}\| = 0 \quad \Leftrightarrow \quad A\,\boldsymbol{x} = \mathbf{0} \quad \Leftrightarrow \quad \boldsymbol{x} \in \ker A = \{0\} \quad \Leftrightarrow \quad \boldsymbol{x} = \mathbf{0}.$$

We prove the converse, i.e., that if p is a norm, then $\operatorname{rank} A = n$. We use the contrapositive: we prove that if $\operatorname{rank} A \neq n$, then p is not a norm. If $\operatorname{rank} A \neq n$, then $\ker A$ is not the origin, so that there exists $\boldsymbol{x} \in \ker A$ different from the zero vector. For this vector $p(\boldsymbol{x}) = \|A\,\boldsymbol{x}\| = \|\mathbf{0}\| = 0$, and therefore p is not a norm.

We prove the last point. Let $\|\cdot\|$ be a norm and let p be a seminorm over a vector space $\boldsymbol{X}$. Then, for all $k \in \mathbb{F}$ and $\boldsymbol{x} \in \boldsymbol{X}$,

$$p(k\,\boldsymbol{x}) + \|k\,\boldsymbol{x}\| = |k|\,p(\boldsymbol{x}) + |k|\,\|\boldsymbol{x}\| = |k|\,(p(\boldsymbol{x}) + \|\boldsymbol{x}\|).$$

For every $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$,

$$p(\boldsymbol{x}+\boldsymbol{y}) + \|\boldsymbol{x}+\boldsymbol{y}\| \leq p(\boldsymbol{x}) + p(\boldsymbol{y}) + \|\boldsymbol{x}\| + \|\boldsymbol{y}\| = p(\boldsymbol{x}) + \|\boldsymbol{x}\| + p(\boldsymbol{y}) + \|\boldsymbol{y}\|.$$

Finally,

$$p(\boldsymbol{x}) + \|\boldsymbol{x}\| = 0$$

implies $p(\boldsymbol{x}) = \|\boldsymbol{x}\| = 0$, because both these quantities are non-negative, and from $\|\boldsymbol{x}\| = 0$ we obtain $\boldsymbol{x} = \boldsymbol{0}$. ■

Solution 1.64 *(Exercise 1.64).* We have

$$\|\boldsymbol{x}_k - \boldsymbol{x}\|_{(2)} \leq c_1 \|\boldsymbol{x}_k - \boldsymbol{x}\|_{(1)} \leq c_1 \varepsilon.$$

Choosing $\varepsilon = \varepsilon_1 / c_1$ we obtain $\|\boldsymbol{x}_k - \boldsymbol{x}\|_{(2)} < \varepsilon_1$, so that it goes to zero as $k \to \infty$.

■

Solution 1.65 *(Exercise 1.65).* From $\lim_{k\to\infty} \boldsymbol{x}_k = \hat{\boldsymbol{x}}$, we know that

$$\forall\, \varepsilon > 0,\ \exists\, p > 0 : \qquad k \geq p \ \Rightarrow\ \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \leq \varepsilon.$$

We need to show that

$$\forall\, \varepsilon > 0,\ \exists\, p > 0 : \qquad k \geq p \ \Rightarrow\ \big|\|\boldsymbol{x}_k\| - \|\hat{\boldsymbol{x}}\|\big| \leq \varepsilon.$$

This is an immediate consequence of

$$\big|\|\boldsymbol{x}_k\| - \|\hat{\boldsymbol{x}}\|\big| \leq \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \leq \varepsilon.$$

■

Solution 1.66 *(Exercise 1.66).* By contradiction, suppose that there exists $h \in \mathbb{N}$ such that $x_h \neq 0$, so that $x_h > 0$ (since $x_k \geq 0$ for all $k \in \mathbb{N}$). We prove that the sequence cannot converge to 0. Choose $\varepsilon \in (0, x_h)$. Since $(x_k)_{k\in\mathbb{N}}$ is non-decreasing, for every $k > h$ we have $x_k \geq x_h$. Then, there does not exist $p \in \mathbb{N}$ such that for all $k > p$ we have $|x_k| = x_k < \varepsilon < x_h$, and this means that the limit is not 0. ■

Solution 1.67 *(Exercise 1.67).* We have

$$\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N} :\ \ \forall\, k > N,\ \ |x_k - \ell| < \varepsilon.$$

Then,

$$\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N} :\ \ \forall\, k > N,\ \ |\ell| \leq |\ell - x_k| + |x_k| \leq a + \varepsilon.$$

We conclude that

$$\forall\, \varepsilon > 0,\ |\ell| \leq a + \varepsilon.$$

From Exercise 1.27, we deduce that $|\ell| \leq a$. Since $\ell \geq 0$, we find $\ell \leq a$. ■

Solution 1.68 *(Exercise 1.68).*

(i) For all $x \in A$, we have $\inf A \leq x \leq \sup A$, which leads immediately to the result.

(ii) Suppose $A \subseteq B$. Hence, $\sup B$ is an upper bound of A, and therefore $\sup A \leq \sup B$. Similarly, $\inf B$ is a lower bound of A, from which we find $\inf A \geq \inf B$.

■

Solution 1.69 *(Exercise 1.69).* Consider the reflexive property. If $a \in \overline{\mathbb{R}}$, this property holds, because it holds in $\mathbb{R}$ for the standard order $\leq$. Since the definition of ordering in $\overline{\mathbb{R}}$ gives, in particular, $+\infty \leq +\infty$ and $-\infty \leq \infty$, then we have $a \leq a$ if $a = \pm\infty$.

Let us consider the transitive property, and assume $a, b, c \in \overline{\mathbb{R}}$ with $a \leq b$ and $b \leq c$. Then:

- if $a, b, c \in \mathbb{R}$, we know that the transitive property holds (as it holds in $\mathbb{R}$);
- if $a = +\infty$, then we must have $b = +\infty$ (otherwise $a \leq b$ is not true) and $c = +\infty$ (otherwise $b \leq c$ is not true), so that $+\infty = a \leq c = +\infty$;
- if $a = -\infty$, then we can have $b, c \in \overline{\mathbb{R}}$ with $b \leq c$, but in all cases $-\infty = a \leq c$;
- if $a \in \mathbb{R}$, then we have the following cases:
  - if $b \in \mathbb{R}$, then c can be real (and in this case we already said that $a \leq c$ is true), $c = +\infty$, and in this case we certainly have $a \leq c$, and the case $c = -\infty$ violates $b \leq c$;
  - if $b = +\infty$, to have $b \leq c$ we must have $c = +\infty$, and in this case obviously $a \leq c$;
  - if $b = -\infty$, then to have $a \leq b$ we must have $a = -\infty$, and we have already seen that in this case $a \leq c$.

We now prove the anti-symmetric property. Assume $a, b \in \overline{\mathbb{R}}$ with $a \leq b$ and $b \leq a$. Then:

- if $a, b \in \mathbb{R}$, we know that $a = b$;
- if $a = +\infty$, since $a \leq b$ we also have $b = +\infty$, so that $a = b$;
- if $a = -\infty$, since $b \leq a$ we also have $b = -\infty$, so that $a = b$.

We now prove the trichotomy law. Assume $a, b \in \overline{\mathbb{R}}$.

- if $a, b \in \mathbb{R}$, the property is true;
- if $a = +\infty$, then we have $a \leq b$ for any $b \in \overline{\mathbb{R}}$;
- if $a = -\infty$, then we have $b \leq a$ for any $b \in \overline{\mathbb{R}}$.

■

Solution 1.70 *(Exercise 1.70).* The commutativity is obvious from the definition of addition: all the cases are symmetric. Let us consider the associativity. We want to study under what conditions $(a + b) + c = a + (b + c)$. We distinguish the cases as follows[2]

- if $a = +\infty$,
  - if $b = +\infty$,
    - if $c = +\infty$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$
    - if $c = -\infty$ then $(a + b) + c = 0$ and $a + (b + c) = +\infty$
    - if $c \in \mathbb{R}$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$
  - if $b = -\infty$,
    - if $c = +\infty$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$

[2]Some of these are redundant, but we still list all of them to ensure we are not missing anything.

        * if $c = -\infty$ then $(a + b) + c = +\infty$ and $a + (b + c) = 0$
        * if $c \in \mathbb{R}$, then $c = (a + b) + c \neq a + (b + c) = 0$ (false in general, unless $c = 0$)
    - if $b \in \mathbb{R}$,
        * if $c = +\infty$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$
        * if $c = -\infty$ then $(a + b) + c = a + (b + c) = 0$
        * if $c \in \mathbb{R}$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$
- if $a = -\infty$,
    - if $b = +\infty$,
        * if $c = +\infty$, then $+\infty = (a + b) + c \neq a + (b + c) = 0$
        * if $c = -\infty$ then $-\infty = (a + b) + c = a + (b + c)$
        * if $c \in \mathbb{R}$, then $c = (a + b) + c \neq a + (b + c) = 0$ (false in general, unless $c = 0$)
    - if $b = -\infty$,
        * if $c = +\infty$, then $0 = (a + b) + c \neq a + (b + c) = -\infty$
        * if $c = -\infty$ then $-\infty = (a + b) + c = a + (b + c) = -\infty$
        * if $c \in \mathbb{R}$, then $-\infty = (a + b) + c = a + (b + c) = -\infty$
    - if $b \in \mathbb{R}$,
        * if $c = +\infty$, then $0 = (a + b) + c = a + (b + c) = 0$
        * if $c = -\infty$ then $-\infty = (a + b) + c = a + (b + c) = -\infty$
        * if $c \in \mathbb{R}$, then $-\infty = (a + b) + c = a + (b + c) = -\infty$
- if $a \in \mathbb{R}$,
    - if $b = +\infty$,
        * if $c = +\infty$, then $+\infty = (a + b) + c = a + (b + c)$
        * if $c = -\infty$ then $0 = (a + b) + c \neq a + (b + c) = a$ (false in general, unless $a = 0$)
        * if $c \in \mathbb{R}$, then $+\infty = (a + b) + c = a + (b + c) = +\infty$
    - if $b = -\infty$,
        * if $c = +\infty$, then $0 = (a + b) + c \neq a + (b + c) = a$ (false in general, unless $a = 0$)
        * if $c = -\infty$ then $-\infty = (a + b) + c = a + (b + c) = -\infty$
        * if $c \in \mathbb{R}$, then $-\infty = (a + b) + c = a + (b + c) = -\infty$
    - if $b \in \mathbb{R}$,
        * if $c = +\infty$, then $+\infty = (a + b) + c = a + (b + c)$
        * if $c = -\infty$ then $-\infty = (a + b) + c = a + (b + c) = -\infty$
        * if $c \in \mathbb{R}$, then we are in the standard case.

■

Solution 1.71 *(Exercise 1.71).* Once again, the commutative property follows by the symmetry in the definition of the multiplication in $\overline{\mathbb{R}}$. We consider the associative property $(a\,b)\,c = a\,(b\,c)$, by considering

all the cases (some of which are redundant):

- if $a = +\infty$,
  - if $b > 0$,
    - if $c > 0$, then $+\infty = (a\,b)\,c = a\,(b\,c) = +\infty$
    - if $c < 0$, then $-\infty = (a\,b)\,c = a\,(b\,c) = -\infty$
    - if $c = 0$, then $0 = (a\,b)\,c = a\,(b\,c) = 0$
  - if $b < 0$, same as for the case $b > 0$, but with opposite signs;
  - if $b = 0$, for any value of c we have $0 = (a\,b)\,c = a\,(b\,c) = 0$.
- if $a = -\infty$,
  - if $b > 0$,
    - if $c > 0$, then $-\infty = (a\,b)\,c = a\,(b\,c) = -\infty$
    - if $c < 0$, then $+\infty = (a\,b)\,c = a\,(b\,c) = +\infty$
    - if $c = 0$, then $0 = (a\,b)\,c = a\,(b\,c) = 0$
  - if $b < 0$, same as for the case $b > 0$, but with opposite signs;
  - if $b = 0$, for any value of c we have $0 = (a\,b)\,c = a\,(b\,c) = 0$.
- if $a \in (0, +\infty)$,
  - if $b = +\infty$,
    - if $c > 0$, then $+\infty = (a\,b)\,c = a\,(b\,c) = +\infty$
    - if $c < 0$, then $-\infty = (a\,b)\,c = a\,(b\,c) = -\infty$
    - if $c = 0$, then $0 = (a\,b)\,c = a\,(b\,c) = 0$
  - if $b = -\infty$, same as for the case $b > 0$, but with opposite signs;
  - if $b = 0$, for any value of c we have $0 = (a\,b)\,c = a\,(b\,c) = 0$;
  - $b \in (0, +\infty)$,
    - if $c + \infty$, then $+\infty = (a\,b)\,c = a\,(b\,c) = +\infty$
    - if $c = -\infty$, then $-\infty = (a\,b)\,c = a\,(b\,c) = -\infty$
    - if $c = 0$, then $0 = (a\,b)\,c = a\,(b\,c) = 0$
  - $b \in (-\infty, 0)$, same as for the case $b \in (0, +\infty)$, but with opposite signs;
- if $a \in (-\infty, 0)$, same as for the case $a \in (0, +\infty)$, but with opposite signs.

■

Solution 1.72 *(Exercise 1.72).* Again, we enumerate all the possibilities:

- if $a = +\infty$, then
  - if $b = +\infty$, then
    - if $c = +\infty$, then (1.44) becomes $+\infty = +\infty$;
    - if $c = -\infty$, it becomes $0 = 0$;

        * if $c \in (0, +\infty)$, it becomes $+\infty = +\infty$;
        * if $c \in (-\infty, 0)$, it becomes $+\infty = +\infty - \infty = 0$ (FAIL);
        * if $c = 0$, it becomes $+\infty = +\infty$;
    - if $b = -\infty$, then
        * if $c = +\infty$, then (1.44) becomes $0 = 0$;
        * if $c = -\infty$, it becomes $-\infty = -\infty$;
        * if $c \in (0, +\infty)$, it becomes $-\infty = 0$ (FAIL);
        * if $c \in (-\infty, 0)$, it becomes $-\infty = -\infty$;
        * if $c = 0$, it becomes $-\infty = -\infty$;
    - if $b = 0$, then
        * if $c = +\infty$, then (1.44) becomes $+\infty = +\infty$;
        * if $c = -\infty$, it becomes $-\infty = -\infty$;
        * if $c \in (0, +\infty)$, it becomes $+\infty = +\infty$;
        * if $c \in (-\infty, 0)$, it becomes $-\infty = -\infty$;
        * if $c = 0$, it becomes $0 = 0$;
- if $a = -\infty$, then
    - if $b = +\infty$, then
        * if $c = +\infty$, then (1.44) becomes $-\infty = -\infty$;
        * if $c = -\infty$, it becomes $0 = 0$;
        * if $c \in (0, +\infty)$, it becomes $-\infty = -\infty$;
        * if $c \in (-\infty, 0)$, it becomes $-\infty = 0$ (FAIL);
        * if $c = 0$, it becomes $-\infty = -\infty$;
    - if $b = -\infty$, then
        * if $c = +\infty$, then (1.44) becomes $0 = 0$;
        * if $c = -\infty$, it becomes $+\infty = +\infty$;
        * if $c \in (0, +\infty)$, it becomes $+\infty = 0$ (FAIL);
        * if $c \in (-\infty, 0)$, it becomes $+\infty = +\infty$;
        * if $c = 0$, it becomes $+\infty = - + \infty$;
    - if $b = 0$, then
        * if $c = +\infty$, then (1.44) becomes $-\infty = -\infty$;
        * if $c = -\infty$, it becomes $+\infty = +\infty$;
        * if $c \in (0, +\infty)$, it becomes $-\infty = -\infty$;
        * if $c \in (-\infty, 0)$, it becomes $+\infty = +\infty$;
        * if $c = 0$, it becomes $0 = 0$;

- if $a = 0$, we always have $0 = 0$;
- if $a \in (0, +\infty)$, then
  - if $b = +\infty$, then
    - if $c = +\infty$, then (1.44) becomes $+\infty = +\infty$;
    - if $c = -\infty$, it becomes $0 = 0$;
    - if $c \in (0, +\infty)$, it becomes $+\infty = +\infty$;
    - if $c \in (-\infty, 0)$, it becomes $+\infty = +\infty$;
    - if $c = 0$, it becomes $+\infty = +\infty$;
  - if $b = -\infty$, then
    - if $c = +\infty$, then (1.44) becomes $0 = 0$;
    - if $c = -\infty$, it becomes $-\infty = -\infty$;
    - if $c \in (0, +\infty)$, it becomes $-\infty = -\infty$;
    - if $c \in (-\infty, 0)$, it becomes $-\infty = -\infty$;
    - if $c = 0$, it becomes $-\infty = -\infty$;
  - if $b = 0$, then
    - if $c = +\infty$, then (1.44) becomes $+\infty = +\infty$;
    - if $c = -\infty$, it becomes $-\infty = -\infty$;
    - if $c \in \mathbb{R}$, the distributivity works because all variables are in $\mathbb{R}$.

We do not need to consider the case $a \in (-\infty, 0)$ because all the results work as for the case $a \in (0, +\infty)$ with the sign changed. We do not need to consider the case $b \in (0, +\infty)$ and $b \in (-\infty, 0)$ because we can swap b and c (and for c we have considered these cases).

We observe that the distributivity fails if at least one of the variables is negative (possibly $-\infty$). Hence, in $[0, +\infty]$ the distributivity holds.

The distributive property

$$(a + b)\, c = a\, c + b\, c \tag{10.10}$$

does not present differences because of the commutativity of the multiplication. ■

Solution 1.73 *(Exercise 1.73).*

(i) For all $x \in A$, we have $\inf A \le x \le \sup A$, which leads immediately to the result.

(ii) Suppose $A \subseteq B$. Hence, $\sup B$ is an upper bound of A, and therefore $\sup A \le \sup B$. Similarly, $\inf B$ is a lower bound of A, from which we find $\inf A \ge \inf B$.

■

Solution 1.74 *(Exercise 1.74).* The terms of $(a_k)_{k \in \mathbb{N}}$ are

$$1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6, \quad 7, \quad 8, \quad 9, \quad 10, \quad \dots$$

In this case $(y_k)_{k\in\mathbb{N}}$ is the sequence

$$+\infty, \quad +\infty, \quad +\infty, \quad +\infty, \quad \dots$$

and therefore the limit superior is $+\infty$. The sequence $(z_k)_{k\in\mathbb{N}}$ is given by

$$1, \quad 2, \quad 3, \quad 4, \quad 5, \quad 6, \quad 7, \quad 8, \quad 9, \quad 10, \quad \dots$$

whose supremum is $+\infty$, so that the limit inferior is also $+\infty$. ■

Solution 1.75 *(Exercise 1.75).*

(i) Using the definition of liminf, we find

$$-\liminf_{k\to\infty}(-a_k) = -\lim_{k\to\infty}\inf_{h\geq k}(-a_h) = \lim_{k\to\infty}\Big(-\inf_{h\geq k}(-a_h)\Big) = \lim_{k\to\infty}\sup_{h\geq k}(a_h) = \limsup_{k\to\infty}(a_k).$$

We used the fact that $\inf_{k\in\mathbb{N}}(-a_k) = -\sup_{k\in\mathbb{N}}(a_k)$. To prove this, we recall that the infimum is the maximum of the lower bounds of a set (and if applied to a function, it is the set is the image of that function), so that

$$\inf_{k\in\mathbb{N}}(-a_k) = \max\{m\in\mathbb{R}\,|\,m\leq -a_k, \quad \forall\, k\in\mathbb{N}\}.$$

Let us denote the maximum by μ. By definition, the maximum μ of the set $\{m\in\mathbb{R}\,|\,m\leq -a_k, \quad \forall\, k\in\mathbb{N}\}$ is an element of the set (so that $\mu\leq -a_k$) and it is greater or equal to any other element of the set. This means that, given $x\in\{m\in\mathbb{R}\,|\,m\leq -a_k, \quad \forall\, k\in\mathbb{N}\}$, we have $x\leq\mu$. We have obtained the three conditions

$$\mu\leq -a_k, \qquad x\leq -a_k, \qquad x\leq\mu.$$

Defining $y=-x$ and multiplying all the relations by -1 gives

$$-\mu\geq a_k, \qquad y\geq a_k, \qquad y\geq -\mu.$$

These say that $-\mu$ is the minimum of the set $\{M\in\mathbb{R}\,|\,M\geq a_k, \quad \forall\, k\in\mathbb{N}\}$. This minimum is the supremum of a_k, so that $-\mu=\sup_{k\in\mathbb{N}}(a_k)$.

(ii) We have

$$\limsup_{k\to\infty}(b+a_k) = \lim_{k\to\infty}\sup_{h\geq k}(b+a_k) = \lim_{k\to\infty} b+\sup_{h\geq k} a_k = b+\lim_{k\to\infty}\sup_{h\geq k} a_k = b+\limsup_{k\to\infty}(a_k),$$

where we used the fact that $\sup_{k\in\mathbb{N}}(b+a_k) = b+\sup_{k\in\mathbb{N}} a_k$. To prove this, let $s=\sup_{k\in\mathbb{N}}(b+a_k)$, which means that $s=\min\{M\in\mathbb{R}\,|\,M\geq b+a_k\}$. In other words, $s\geq b+a_k$ for all $k\in\mathbb{N}$, and for all x such that $x\geq b+a_k$ for all $k\in\mathbb{N}$, we have $s\leq x$. This is equivalent to writing $s-b\geq a_k$ for all $k\in\mathbb{N}$, and for all x such that $x-b\geq a_k$ for all $k\in\mathbb{N}$, we have $s-b\leq x-b$. Defining $y=x-b$, this is in turn equivalent to saying that $s-b\geq a_k$ for all $k\in\mathbb{N}$, and for all y such that $y\geq a_k$ for all $k\in\mathbb{N}$, we have $s-b\leq y$. Hence, $s-b$ is the minimum of the upper bounds of a_k, i.e., $s-b=\sup_{k\in\mathbb{N}} a_k$, which implies $s=b+\sup_{k\in\mathbb{N}} a_k$.

(iii) This can be proved as the previous point *mutatis mutandis*.

(iv) If $(a_k)_{k\in\mathbb{N}}$ is not upper bounded, then $\limsup_{k\to\infty} a_k = \inf\varnothing = +\infty$, and the result follows from the rules of addition in $\overline{\mathbb{R}}$. If $(a_k)_{k\in\mathbb{N}}$ is not lower bounded, then

$$\liminf_{k\to\infty} a_k = \sup(-\infty) = \sup\varnothing = -\infty \le \limsup_{k\to\infty} a_k.$$

Now we suppose that $(a_k)_{k\in\mathbb{N}}$ is bounded. Using the definitions, we have to show that

$$\lim_{k\to\infty}\inf_{h\ge k} a_h \le \lim_{k\to\infty}\sup_{h\ge k} a_h.$$

To this end, it is sufficient to prove that $\inf_{h\ge k} a_h \le \sup_{h\ge k} a_h$. In fact, if this is the case the sequence $\sup_{h\ge k} a_h - \inf_{h\ge k} a_h$ is ≥ 0, and taking the limit preserves the sign and we can conclude. On the other hand, $\inf_{h\ge k} a_h \le \sup_{h\ge k} a_h$ follows immediately from the fact that, given a bounded set A, we have $\inf A \le \sup A$. Indeed, $s = \inf A$ is the maximum of the lower bounds, and therefore s is itself a lower bound. This means that s satisfies $s \le x$ for all $x \in A$. Likewise, $\sigma = \sup A$ is an upper bound, and therefore $\sigma \ge x$ for all $x \in A$. Thus, for all $x \in A$ we have $s \le x \le \sigma$, which implies $\inf A \le \sup A$.

(v) Assume $\limsup_{k\to\infty} a_k = \liminf_{k\to\infty} a_k = \ell$. Recalling that, in the definition of limit superior, $b_k = \sup_{h\ge k} a_h$, for all $k \in \mathbb{N}$ we have $a_k \le b_k$. Likewise, in the definition of limit inferior, $c_k = \inf_{h\ge k} a_h$, and for all $k \in \mathbb{N}$ we have $c_k \le a_k$. Thus, for all $k \in \mathbb{N}$ we have $c_k \le a_k \le b_k$. We use the squeeze theorem, taking into account the fact that $\ell = \limsup_{k\to\infty} a_k = \lim_{k\to\infty} b_k$ and $\ell = \liminf_{k\to\infty} a_k = \lim_{k\to\infty} c_k$. Since the limits of c_k and b_k as $k \to \infty$ coincide with ℓ, then a_k is convergent as $k \to \infty$ and $\lim_{k\to\infty} a_k = \ell$.
Now, we assume that $\lim_{k\to\infty} a_k = \ell$.

We prove that $(b_k)_{k\in\mathbb{N}}$ and $(c_k)_{k\in\mathbb{N}}$ converge to ℓ as $k \to \infty$. In particular, we prove that for all $\varepsilon > 0$ there exists $p \in \mathbb{N}$ such that for all $k \ge p$ we have $|b_k - \ell| \le \varepsilon$ and $|c_k - \ell| \le \varepsilon$, or, equivalently, $\ell - \varepsilon \le b_k \le \ell + \varepsilon$ and $\ell - \varepsilon \le c_k \le \ell + \varepsilon$.

Let $\varepsilon > 0$. Since $\lim_{k\to\infty} a_k = \ell$, there exists $p \ge 0$ such that for all $k \ge p$ we have $\ell - \varepsilon \le a_k \le \ell + \varepsilon$. It follows that $\ell - \varepsilon \le b_k \le \ell + \varepsilon$ and $\ell - \varepsilon \le c_k \le \ell + \varepsilon$ for $k \ge p$, and therefore $(b_k)_{k\in\mathbb{N}}$ and $(c_k)_{k\in\mathbb{N}}$ converge to ℓ as $k \to \infty$.

■

Solution 1.76 *(Exercise 1.76).* We have

$$\liminf_{k\to\infty} a_k = -1 \qquad \text{and} \qquad \limsup_{k\to\infty} a_k = 1.$$

The upper and lower limits have different values, and in fact we know that this sequence does not have a limit in $\overline{\mathbb{R}}$. ■

Solution 1.77 *(Exercise 1.77).* We have, for all $k \in \mathbb{N}^*$,

$$x_k = \inf_{h\ge k} a_h = \inf_{h\ge k} \frac{1}{h} = 0.$$

Thus,

$$\liminf_{k\to\infty} a_k = \lim_{k\to\infty} x_k = \lim_{k\to\infty} 0 = 0.$$

Since the sequence is not bounded from above, we have

$$\limsup_{k\to\infty} a_k = +\infty.$$

■

Solution 1.78 *(Exercise 1.78).* Since $\frac{1}{k}$ is decreasing, we have

$$x_k = \inf_{h\geq k} a_h = \begin{cases} 1-\frac{1}{k} & \text{if } k \text{ is odd} \\ 1-\frac{1}{k+1} & \text{if } k \text{ is even} \end{cases}$$

and

$$y_k = \sup_{h\geq k} a_h = \begin{cases} 1+\frac{1}{k+1} & \text{if } k \text{ is odd} \\ 1+\frac{1}{k} & \text{if } k \text{ is even.} \end{cases}$$

Both x_k and y_k go to 1 as $k\to\infty$, and they coincide with the limit of $(a_k)_{k\in\mathbb{N}}$. In this case, rather than computing the limit inferior and limit superior, we could have computed the limit and conclude that the latter two are the same. ■

Solution 1.79 *(Exercise 1.79).* For all $\ell \geq k$, we have

$$a_\ell + b_\ell \leq \sup_{h\geq k} a_h + \sup_{h\geq k} b_h.$$

It follows that $\sup_{h\geq k} a_h + \sup_{h\geq k} b_h$ is an upper bound for $a_\ell + b_\ell$ for all $\ell \geq k$. Since an upper bound is larger or equal than the supremum by definition, for all $k\in\mathbb{N}$ we obtain

$$\sup_{\ell\geq k}(a_\ell + b_\ell) \leq \sup_{h\geq k} a_h + \sup_{h\geq k} b_h.$$

By definition of limit superior we obtain

$$\limsup_{k\to\infty}(a_k + b_k) \leq \limsup_{k\to\infty} a_k + \limsup_{k\to\infty} b_k.$$

■

Solution 1.80 *(Exercise 1.80).* Given any subset Ω of $\mathbb{R}$, we have $\sup\Omega = -\inf(-\Omega)$ (why? we have established it in a previous exercise). In particular, for all $k\in\mathbb{N}$ we have

$$\sup_{h\geq k} a_h = -\inf_{h\geq k}(-a_h).$$

Taking the limits we find

$$\limsup_{k\to\infty} a_k = \lim_{k\to\infty}\sup_{h\geq k} a_h = -\lim_{k\to\infty}\inf_{h\geq k}(-a_h) = -\liminf_{k\to\infty}(-a_k).$$

■

Solution 1.81 *(Exercise 1.81).* Suppose for the sake of argument that $(a_k)_{k\in\mathbb{N}}$ is convergent. From a previous example we know that

$$\limsup_{k\to\infty}(a_k + b_k) \leq \limsup_{k\to\infty} a_k + \limsup_{k\to\infty} b_k = \lim_{k\to\infty} a_k + \limsup_{k\to\infty} b_k.$$

We need to prove that the opposite inequality holds. We use the last example to find

$$\limsup_{k\to\infty} b_k = \limsup_{k\to\infty}(a_k + b_k - a_k) \leq \limsup_{k\to\infty}(a_k + b_k) - \liminf_{k\to\infty} a_k.$$

We obtain

$$\liminf_{k\to\infty} a_k + \limsup_{k\to\infty} b_k \leq \limsup_{k\to\infty}(a_k + b_k),$$

and since $(a_k)_{k\in\mathbb{N}}$ is assumed to converge, this is the same as

$$\lim_{k\to\infty} a_k + \limsup_{k\to\infty} b_k \leq \limsup_{k\to\infty}(a_k + b_k),$$

which is indeed what we needed to prove.

The strict equality can be therefore obtained only when neither sequence is convergent. For example, consider the sequences

$$\begin{array}{rcl} a : \mathbb{N} & \longrightarrow & \mathbb{R} \\ k & \mapsto & (-1)^k \end{array} \qquad \text{and} \qquad \begin{array}{rcl} b : \mathbb{N} & \longrightarrow & \mathbb{R} \\ k & \mapsto & (-1)^{k+1} \end{array}$$

We have

$$\limsup_{k\to\infty} a_k = \limsup_{k\to\infty} b_k = 1.$$

However, for all $k \in \mathbb{N}$ we have $a_k + b_k = (-1)^k - (-1)^{k+1} = (-1)^k(1-1) = 0$. Hence,

$$\limsup_{k\to\infty}(a_k + b_k) = 0.$$

■

Solution 1.82 *(Exercise 1.82).* Let, for a generic $k \in \mathbb{N}$,

$$y_k = \sup_{h\geq k} x_h = \sup\left\{3 + (-1)^k \frac{k}{k+8}, 3 + (-1)^{k+1}\frac{k+1}{k+9}, 3 + (-1)^{k+2}\frac{k+2}{k+10}, \ldots\right\}.$$

If k is even,

$$y_k = \sup_{h\geq k} x_h = \sup\left\{3 + \frac{k}{k+8}, 3 - \frac{k+1}{k+9}, 3 + \frac{k+2}{k+10}, \ldots\right\},$$

and if k is odd,

$$y_k = \sup_{h\geq k} x_h = \sup\left\{3 - \frac{k}{k+8}, 3 + \frac{k+1}{k+9}, 3 - \frac{k+2}{k+10}, \ldots\right\}.$$

The supremum is, in both cases, $3 + 1 = 4$.
Likewise, we have for a generic $k \in \mathbb{N}$,

$$z_k = \inf_{h\geq k} x_h = \inf\left\{3 + (-1)^k \frac{k}{k+8}, 3 + (-1)^{k+1}\frac{k+1}{k+9}, 3 + (-1)^{k+2}\frac{k+2}{k+10}, \ldots\right\},$$

and regardless of whether k is odd or even we have that the infimum is $3 - 1 = 2$. It follows that

$$\limsup_{k\to\infty} x_k = 4 \quad \text{and} \quad \liminf_{k\to\infty} x_k = 2.$$

We also deduce that the sequence is not convergent, because the limit superior and inferior do not coincide. ■

Solution 1.83 *(Exercise 1.83).*

☒ If $(x_k)_{k\in\mathbb{N}}$ is convergent, then $(x_{k_j})_{j\in\mathbb{N}}$ is convergent and the limits coincide;

☐ If $(x_{k_j})_{j\in\mathbb{N}}$ is convergent, then $(x_k)_{k\in\mathbb{N}}$ is convergent and the limits coincide;

☒ If $(x_k)_{k\in\mathbb{N}}$ is bounded, then $(x_{k_j})_{j\in\mathbb{N}}$ is bounded;

☐ If $(x_{k_j})_{j\in\mathbb{N}}$ is bounded, then $(x_k)_{k\in\mathbb{N}}$ is bounded;

☐ If $(x_k)_{k\in\mathbb{N}}$ oscillates, then $(x_{k_j})_{j\in\mathbb{N}}$ does not converge;

☐ If $(x_k)_{k\in\mathbb{N}}$ is bounded, then $(x_{k_j})_{j\in\mathbb{N}}$ converges;

☐ $(x_k)_{k\in\mathbb{N}}$ always has a limit inferior and limit superior in $\mathbb{R}$;

☒ $(x_k)_{k\in\mathbb{N}}$ always has a limit inferior and limit superior in $\overline{\mathbb{R}}$;

☒ If $\limsup_{k\to\infty} x_k = \liminf_{k\to\infty} x_k$, then $(x_k)_{k\in\mathbb{N}}$ is convergent in $\overline{\mathbb{R}}$;

☒ If $(x_k)_{k\in\mathbb{N}}$ is convergent in $\mathbb{R}$, then $\limsup_{k\to\infty} x_k = \liminf_{k\to\infty} x_k$.

■

Solution 1.84 *(Exercise 1.84).*

(1) Since $x_k, y_k \geq 0$ for all $k \in \mathbb{N}$, for all $\ell \geq k$ we have

$$x_\ell\, y_\ell \leq (\sup_{h\geq k} x_h)(\sup_{h\geq k} y_h),$$

which shows that $(\sup_{h\geq k} x_h)(\sup_{h\geq k} y_h)$ is an upper bound for $x_\ell\, y_\ell$ for all $\ell \geq k$. Since an upper bound is larger or equal than the supremum, we obtain for all $k \in \mathbb{N}$

$$\sup_{\ell\geq k}(x_\ell\, y_\ell) \leq (\sup_{h\geq k} x_h)(\sup_{h\geq k} y_h).$$

Taking the limit and using the limit rules we obtain

$$\limsup_{k\to\infty}(x_k\, y_k) \leq (\limsup_{k\to\infty} x_k)(\limsup_{k\to\infty} y_k).$$

(2) Consider the sequences

$$\begin{array}{rcl} x : \mathbb{N} & \longrightarrow & \mathbb{R} \\ k & \mapsto & 2 + (-1)^k \end{array} \qquad \text{and} \qquad \begin{array}{rcl} y : \mathbb{N} & \longrightarrow & \mathbb{R} \\ k & \mapsto & 2 + (-1)^{k+1}. \end{array}$$

Clearly $x_k, y_k > 0$ for all $k \in \mathbb{N}$, and

$$\limsup_{k\to\infty} x_k = \limsup_{k\to\infty} y_k = 3.$$

Nevertheless, for all $k \in \mathbb{N}$ we have

$$x_k\, y_k = (2 + (-1)^k)(2 + (-1)^{k+1}) = 4 + \underbrace{2\,(-1)^{k+1} + 2\,(-1)^k}_{=0} + \underbrace{(-1)^{2\,k+1}}_{=-1} = 3,$$

so that

$$\limsup_{k\to\infty}(x_k\, y_k) = 3.$$

It follows that

$$3 = \limsup_{k\to\infty}(x_k\, y_k) < (\limsup_{k\to\infty} x_k)(\limsup_{k\to\infty} y_k) = 3 \cdot 3 = 9.$$

(3) Let $a_k = -1$ and $y_k = (-1)^k$ for all $k \in \mathbb{N}$. Then, $x_k\, y_k = (-1)^{k+1}$, so that

$$\limsup_{k\to\infty}(x_k\, y_k) = 1, \quad \lim_{k\to\infty} x_k = -1, \quad \limsup_{k\to\infty} y_k = -1.$$

(4) Assume that $(y_k)_{k\in\mathbb{N}}$ converges to the limit $\hat{y} \in \mathbb{R}$. If $(y_k)_{k\in\mathbb{N}}$ is infinitesimal and $(x_k)_{k\in\mathbb{N}}$ is bounded, clearly $x_k\, y_k$ converges to 0, and therefore the equality is satisfied. If $\hat{y} > 0$, we find

$$\begin{aligned}
(\limsup_{k\to\infty} x_k)(\limsup_{k\to\infty} y_k) &= (\limsup_{k\to\infty} x_k)\,\hat{y} \\
&= \limsup_{k\to\infty}(x_k\,\hat{y}) \\
&= \limsup_{k\to\infty}\left((x_k\, y_k)\,\frac{\hat{y}}{y_k}\right) \\
&\le \limsup_{k\to\infty}(x_k\, y_k)\ \limsup_{k\to\infty}\frac{\hat{y}}{y_k} \\
&= \limsup_{k\to\infty}(x_k\, y_k)\ \lim_{k\to\infty}\frac{\hat{y}}{y_k} = \limsup_{k\to\infty}(x_k\, y_k).
\end{aligned}$$

This, together with (1.45) gives the equality. This argument holds also when $(x_k)_{k\in\mathbb{N}}$ is not bounded from above.

■

Solution 1.85 *(Exercise 1.85).* For all $k \in \mathbb{N}$ and all $x \in A$, we have $f_k(x) \ge 0$. By Theorem 1.25, for all $x \in A$

$$\forall\, k \in \mathbb{N}, \quad f_k(x) \ge 0 \quad \Rightarrow \quad \lim_{k\to\infty} f_k(x) \ge 0,$$

which is the same as saying that $f(x) \ge 0$ for all $x \in A$. ■

Solution 1.86 *(Exercise 1.86).* Let $x, y \in I$, and take $\lambda \in [0, 1]$. Since f_k is convex, we have

$$f_k(\lambda\, x + (1-\lambda)\, y) \le \lambda\, f_k(x) + (1-\lambda)\, f_k(y).$$

Since $(f_k)_{k\in\mathbb{N}}$ converges pointwise to f, we have

$$\lim_{k\to\infty} f_k(\lambda\, x + (1-\lambda)\, y) = f(\lambda\, x + (1-\lambda)\, y)$$

and

$$\lim_{k\to\infty} (\lambda\, f_k(x) + (1-\lambda)\, f_k(y)) = \lambda\, f(x) + (1-\lambda)\, f(y).$$

From Theorem 1.25, the limit preserves inequalities, and since $f_k(\lambda\, x + (1-\lambda)\, y) \le \lambda\, f_k(x) + (1-\lambda)\, f_k(y)$ holds for every $k \in \mathbb{N}$, taking the limit yields

$$f(\lambda\, x + (1-\lambda)\, y) \le \lambda\, f(x) + (1-\lambda)\, f(y).$$

■

Solution 1.87 *(Exercise 1.87).* Choosing $\epsilon = 1$ in the definition of uniform convergence, there exists $p \in \mathbb{R}$ such that for all $k \ge p$

$$\forall\, x \in A, \quad |f_k(x) - f(x)| \le 1.$$

Let $k \geq p$. Since f_k is bounded, there exists a constant $M_k > 0$ such that for all $x \in A$

$$|f_k(x)| < M_k.$$

It follows that for all $x \in A$

$$|f(x)| = |f(x) - f_k(x) + f_k(x)| \leq |f(x) - f_k(x)| + |f_k(x)| \leq 1 + M_k.$$

We can conclude that f is bounded in A. ■

Solution 1.88 *(Exercise 1.88).* We first establish the pointwise convergence, by finding the limit for $x \in [c, +\infty)$:

$$\lim_{k \to \infty} f_k(x) = \lim_{k \to \infty} \frac{\sin(k\,x)}{1 + k\,x} = 0,$$

since $1 + k\,x \to \infty$ as $k \to \infty$. It follows that $(f_k)_{k \to \mathbf{N}}$ is pointwise convergent to the function

$$\begin{aligned} f : [c, +\infty) &\longrightarrow \mathbb{R} \\ x &\mapsto 0. \end{aligned}$$

The sequence $(f_k)_{k \to \mathbf{N}}$ converges to f uniformly. In fact, let $\varepsilon > 0$. For every $x \in [c, +\infty)$ we have

$$\left|\frac{\sin(k\,x)}{1 + k\,x}\right| \leq \frac{1}{1 + k\,x} \leq \frac{1}{1 + k\,c}.$$

We find $\left|\frac{\sin(k\,x)}{1+k\,x}\right| \leq \varepsilon$ by choosing k such that

$$\frac{1}{1 + k\,c} \leq \varepsilon,$$

i.e.,

$$k \geq \frac{1}{c}\left(\frac{1}{\varepsilon} - 1\right).$$

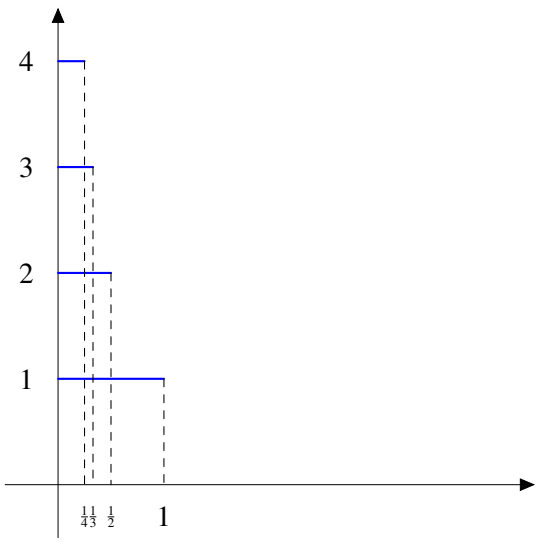
Alternatively,

$$\sup_{x \in [c, +\infty)} |f_k(x) - 0| = \sup_{x \in [c, +\infty)} \left|\frac{\sin(k\,x)}{1 + k\,x}\right| = \frac{1}{1 + k\,c},$$

which goes to 0 as $k \to \infty$. ■

Solution 1.89 *(Exercise 1.89).* The pointwise limit of $(f_k)_{k \in \mathbb{N}}$ is the zero function.

$$\begin{aligned} f_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} k & \text{if } 0 < x < \frac{1}{k} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$



The sequence $(f_k)_{k\in\mathbb{N}}$ does not converge uniformly to its pointwise limit. For example, take $\varepsilon = 1/2$. Then, for all $p \in \mathbb{N}$, there exists $k = p + 1 > \varepsilon$ such that for $x = \frac{1}{2k}$ we have

$$|f_k(x) - 0| = |k - 0| > \varepsilon.$$

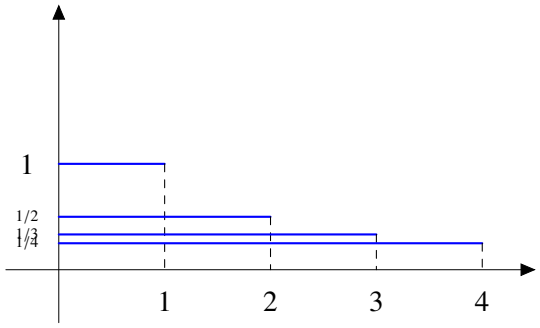
Another way to see this is to find

$$\sup_{x\in\mathbb{R}} |f_k(x) - 0| = \sup_{0<x<\frac{1}{k}} |k| = k,$$

which does not converge to zero. ■

Solution 1.90 *(Exercise 1.90).* The sequence $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the zero function.

$$f_k : \mathbb{R} \longrightarrow \mathbb{R}$$
$$x \mapsto \begin{cases} \frac{1}{k} & \text{if } 0 < x < k \\ 0 & \text{otherwise} \end{cases}$$



To see that $(f_k)_{k\in\mathbb{N}}$ converges uniformly to the zero function, take $\varepsilon > 0$. There exists $p = \lceil \frac{1}{\varepsilon} \rceil \in \mathbb{N}$ such that for all $k \geq p$ and all $x \in \mathbb{R}$

$$|f_k(x) - f(x)| \leq \left|\frac{1}{k} - 0\right| \leq \frac{1}{k} \leq \varepsilon,$$

since $|f_k(x)| \leq \left|\frac{1}{k}\right|$ for all $x \in \mathbb{R}$ and since $k \geq p \geq \frac{1}{\varepsilon}$ implies $\frac{1}{k} \leq \varepsilon$.

Another way to see this is to find

$$\sup_{x\in\mathbb{R}} |f_k(x) - 0| = \sup_{0<x<k} \left|\frac{1}{k}\right| = \frac{1}{k},$$

which converges to zero. ■

Solution 1.91 *(Exercise 1.91).* For every $x \in \mathbb{R}$, the function f_k goes to zero, so $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the zero function. We now need to establish the uniform convergence. We have

$$|f_k(x) - f(x)| = \left|\frac{1}{k\,(1+x^2)} - 0\right| = \left|\frac{1}{k}\right| \cdot \left|\frac{1}{1+x^2}\right| \leq \frac{1}{k}$$

since for all $x \in \mathbb{R}$ we have

$$\left|\frac{1}{1+x^2}\right| \leq 1.$$

Thus, for any $\varepsilon > 0$, we can choose $p \geq 1/\varepsilon$ such that, for every $x \in \mathbb{R}$ we have $|f_k(x) - f(x)| \leq \varepsilon$ if $k \geq p$. It follows that $(f_k)_{k\in\mathbb{N}}$ converges uniformly to the zero function.

Consider the second sequence. For all $x \in \mathbb{R}$ we have $f_k(x) \longrightarrow x$ as $k \to \infty$. Thus, the pointwise limit is the identity function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x. \end{aligned}$$

We now establish the uniform convergence. We see that

$$|f_k(x) - f(x)| = \left| \frac{x\,(x+k)}{k} - x \right| = \frac{x^2}{k}.$$

Asking $|f_k(x) - f(x)| \leq \varepsilon$ for every $x \in \mathbb{R}$ requires choosing $p > \frac{x^2}{\varepsilon}$, which says that there is no value of $p \in \mathbb{N}$ that works for every $x \in \mathbb{R}$. Therefore, the sequence does not converge uniformly.

If we consider the restriction of the sequence to the interval $[-a, a]$, we notice that choosing $p > a^2/\varepsilon$ we find that for all $x \in [-a, a]$

$$|f_k(x) - f(x)| = \left| \frac{x\,(x+k)}{k} - x \right| = \frac{x^2}{k} \leq \frac{a^2}{k} \leq \frac{a^2}{p} \leq \varepsilon$$

for all $k \geq p$. Thus, we have uniform convergence.

■

Solution 1.92 *(Exercise 1.92).* For every $x \in \mathbb{R}$

$$k \left(x + \frac{1}{k} \right)^2 - k\,x^2 = 2\,x + \frac{1}{k} \longrightarrow 2\,x$$

as $k \to \infty$. It follows that the pointwise limit is the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 2\,x. \end{aligned}$$

We check the uniform convergence:

$$|f_k(x) - f(x)| = \left| 2\,x + \frac{1}{k} - 2\,x \right| = \frac{1}{k}.$$

Now, given $\varepsilon > 0$, taking $p \geq 1/\varepsilon$, we have $|f_k(x) - f(x)| \leq \varepsilon$ for all $x \in \mathbb{R}$ for $k \geq p$.

■

Solution 1.93 *(Exercise 1.93).* For every $x \in \mathbb{R}$

$$\frac{k\,x^2}{k + x} \longrightarrow x^2$$

as $k \to \infty$. It follows that the pointwise limit is the function

$$\begin{aligned} f : \mathbb{R}_+ &\longrightarrow \mathbb{R} \\ x &\mapsto x^2. \end{aligned}$$

We check the uniform convergence:

$$|f_k(x) - f(x)| = \left| \frac{k\,x^2}{k+x} - x^2 \right| = \left| \frac{-x^3}{k+x} \right| = \frac{x^3}{k+x}$$

for all $x \in \mathbb{R}_+$. We do not have uniform convergence since for every $k \in \mathbb{N}$

$$\sup_{x \in \mathbb{R}_+} |f_k(x) - f(x)| = \sup_{x \in \mathbb{R}_+} \frac{x^3}{k+x} = \infty.$$

If $x \in [0, a]$ we have

$$|f_k(x) - f(x)| = \left| \frac{k\,x^2}{k+x} - x^2 \right| = \left| \frac{-x^3}{k+x} \right| = \frac{x^3}{k+x} \leq \frac{x^3}{k} \leq \frac{a^3}{k}$$

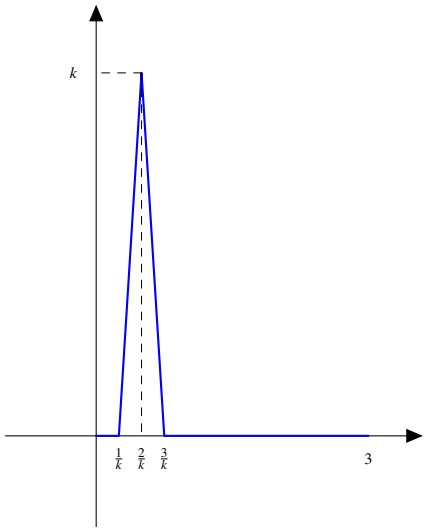
and a^3/k goes to 0 as $k \to \infty$ independently from $x \in [0, a]$, so we have uniform convergence. ■

Solution 1.94 *(Exercise 1.94).* The pointwise limit is

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 1 & \text{if } x = k\,\pi \text{ for some } k \in \mathbb{Z} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The limit is not continuous, while for every $k \in \mathbb{N}$ the function f_k is continuous. Thus, the convergence is not uniform. ■

Solution 1.95 *(Exercise 1.95).* The graph of $f_k(x)$ is:



For every $k \in \mathbb{N}^*$, the function f_k is continuous since

$$\lim_{x \to \frac{1}{k}^-} f_k(x) = 0 \quad \text{and} \quad \lim_{x \to \frac{1}{k}^+} f_k(x) = k^2 \frac{1}{k} - k = 0$$

and

$$\lim_{x \to \frac{2}{k}^-} f_k(x) = k^2 \frac{2}{k} - k = k \quad \text{and} \quad \lim_{x \to \frac{2}{k}^+} 3k - k^2 \frac{2}{k} = k,$$

and, finally,

$$\lim_{x \to \frac{3}{k}^-} f_k(x) = 3k - k^2 \frac{3}{k} = 0 \quad \text{and} \quad \lim_{x \to \frac{3}{k}^+} f_k(x) = 0.$$

Thus, the function is continuous. The function converges pointwise to the zero function. Indeed, given $x \in [0, 3]$:

- if $x \neq 0$, there exists $k = k_1$ sufficiently large such that $3/k_1 < x$, so that for each k greater than k_1 we have that x belongs to the last interval, and therefore $f_k(x)$, as $k \to +\infty$, is 0.
- if $x = 0$, then by definition $f_k(x) = 0$.

For every $k \in \mathbb{N}^*$, we have

$$\begin{aligned}\int_0^3 f_k(x)\,\mathrm{d}x &= \int_0^{1/k} 0\,\mathrm{d}x + \int_{1/k}^{2/k} (k^2 x - k)\,\mathrm{d}x + \int_{2/k}^{3/k} (3k - k^2 x)\,\mathrm{d}x + \int_{3/k}^3 0\,\mathrm{d}x \\ &= \left[k^2 \frac{x^2}{2} - kx\right]_{1/k}^{2/k} + \left[3kx - k^2 \frac{x^2}{2}\right]_{2/k}^{3/k} = \frac{1}{2} + \frac{1}{2} = 1.\end{aligned}$$

It follows also that

$$\lim_{k \to +\infty} \int_0^3 f_k(x)\,\mathrm{d}x = 1.$$

On the other hand, if $(f_k)_{k \in \mathbb{N}^*}$ was uniformly convergent, its uniform limit would be the zero function, but for the zero function the integral is zero. Thus,

$$\lim_{k \to +\infty} \int_0^3 f_k(x)\,\mathrm{d}x = 1 \neq 0 = \int_0^3 \lim_{k \to +\infty} f_k(x)\,\mathrm{d}x.$$

Thus, we cannot interchange the limit with the integral, and this shows that $(f_k)_{k \in \mathbb{N}^*}$ is not uniformly convergent. ■

Solution 1.96 *(Exercise 1.96).* We start with the pointwise convergence: for every $x \in [0, 1]$ we have

$$\lim_{k \to \infty} f_k(x) = \lim_{k \to \infty} e^{\frac{x}{k}} = 1.$$

The sequence is therefore pointwise convergent to the function

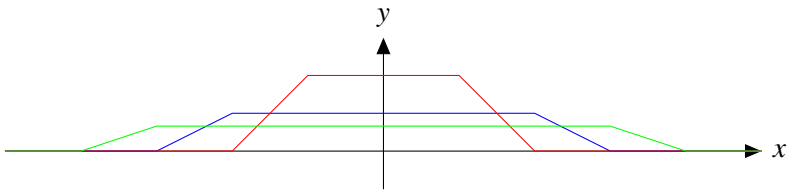
$$\begin{aligned} f : [0, 1] &\longrightarrow \mathbb{R} \\ x &\mapsto 1. \end{aligned}$$

For the uniform convergence:

$$\sup_{x \in [0,1]} |f_k(x) - f(x)| = \sup_{x \in [0,1]} |e^{\frac{x}{k}} - 1| = |e^{\frac{1}{k}} - 1|,$$

and the sequence $(|e^{\frac{1}{k}} - 1|)_{k \in \mathbb{N}^*}$ is infinitesimal. It follows that we have uniform convergence to f. ■

Solution 1.97 *(Exercise 1.97).* We draw the graphs for $k = 1$ (red), $k = 2$ (blue) and $k = 3$ (green).



We notice that for every $k \in \mathbb{N}^*$ the function f_k is continuous. Moreover, the sequence is pointwise convergent to the zero function, since for every $t \in \mathbb{R}$

$$\lim_{k\to\infty} f_k(t) = \lim_{k\to\infty} \frac{1}{k} = 0.$$

We study the uniform convergence:

$$\sup_{x\in\mathbb{R}} |f_k(t)| = \frac{1}{k},$$

which tends to 0, so that the sequence is uniformly convergent to the zero function. It follows that we have convergence in $C(\mathbb{R};\mathbb{R})$. ■

Solution 1.98 *(Exercise 1.98).* We can use induction. We have $k_0 \geq 0$. Assume $k_j \geq j$. We must show that $k_{j+1} \geq j+1$. Since $(k_j)_{j\in\mathbb{N}}$ is strictly increasing,

$$k_{j+1} > k_j \geq j \quad \Rightarrow \quad k_{j+1} \geq j+1,$$

as required. ■

Solution 1.99 *(Exercise 1.99).* We start proving the **if** part. Let $(\mathfrak{Re}\{z_k\})_{k\in\mathbb{N}}$ and $(\mathfrak{Im}\{z_k\})_{k\in\mathbb{N}}$ converge to $\mathfrak{Re}\{\hat{z}\}$ and $\mathfrak{Im}\{\hat{z}\}$, respectively. We must show that $(z_k)_{k\in\mathbb{N}}$ converges to $\hat{z} = \mathfrak{Re}\{\hat{z}\} + i\,\mathfrak{Im}\{\hat{z}\}$, i.e., that for all $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that, for all $k > p$, we have

$$|z_k - \hat{z}| < \varepsilon.$$

Let $\varepsilon > 0$. Since $(\mathfrak{Re}\{z_k\})_{k\in\mathbb{N}}$ converges to $\mathfrak{Re}\{\hat{z}\}$, there exists $p_1 \in \mathbb{N}$ such that for all $k > p_1$ we have $|\mathfrak{Re}\{z_k\} - \mathfrak{Re}\{\hat{z}\}| < \varepsilon/2$. Likewise, since $(\mathfrak{Im}\{z_k\})_{k\in\mathbb{N}}$ converges to $\mathfrak{Im}\{\hat{z}\}$, there exists $p_2 \in \mathbb{N}$ such that for all $k > p_2$ we have $|\mathfrak{Im}\{z_k\} - \mathfrak{Im}\{\hat{z}\}| < \varepsilon/2$. It follows that for all $k > \max\{p_1, p_2\}$ we have

$$\begin{aligned}
|z_k - \hat{z}| &= |\mathfrak{Re}\{z_k\} + i\,\mathfrak{Im}\{z_k\} - \mathfrak{Re}\{\hat{z}\} - i\,\mathfrak{Im}\{\hat{z}\}| \\
&= |(\mathfrak{Re}\{z_k\} - \mathfrak{Re}\{\hat{z}\}) + i\,(\mathfrak{Im}\{z_k\} - \mathfrak{Im}\{\hat{z}\})| \leq |\mathfrak{Re}\{z_k\} - \mathfrak{Re}\{\hat{z}\}| + |i\,(\mathfrak{Im}\{z_k\} - \mathfrak{Im}\{\hat{z}\})| \\
&= |\mathfrak{Re}\{z_k\} - \mathfrak{Re}\{\hat{z}\}| + |\mathfrak{Im}\{z_k\} - \mathfrak{Im}\{\hat{z}\}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

We prove the **only if** part. The sequence $(z_k)_{k\in\mathbb{N}}$ converges to the complex number $\hat{z}$, so that for all $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that for all $k > p$ we have

$$|\mathfrak{Re}\{z_k\} + i\,\mathfrak{Im}\{z_k\} - \mathfrak{Re}\{\hat{z}\} - i\,\mathfrak{Im}\{\hat{z}\}| < \varepsilon,$$

which can be re-written as

$$|(\mathfrak{Re}\{z_k\} - \mathfrak{Re}\{\hat{z}\}) + i\,(\mathfrak{Im}\{z_k\} - \mathfrak{Im}\{\hat{z}\})| < \varepsilon.$$

Since $|\Re\mathrm{e}\{z\}| \leq |z|$ for every $z \in \mathbb{C}$, we find

$$|(\Re\mathrm{e}\{z_k\} - \Re\mathrm{e}\{\hat{z}\})| \leq |(\Re\mathrm{e}\{z_k\} - \Re\mathrm{e}\{\hat{z}\}) + i\,(\Im\mathrm{m}\{z_k\} - \Im\mathrm{m}\{\hat{z}\})| < \varepsilon.$$

It follows that $(\Re\mathrm{e}\{z_k\})_{k\in\mathbb{N}}$ converges to $\Re\mathrm{e}\{\hat{z}\}$. Likewise, since $|\Im\mathrm{m}\{z\}| \leq |z|$ for all $z \in \mathbb{C}$, we find

$$|(\Im\mathrm{m}\{z_k\} - \Im\mathrm{m}\{\hat{z}\})| \leq |(\Re\mathrm{e}\{z_k\} - \Re\mathrm{e}\{\hat{z}\}) + i\,(\Im\mathrm{m}\{z_k\} - \Im\mathrm{m}\{\hat{z}\})| < \varepsilon.$$

It follows that $(\Im\mathrm{m}\{z_k\})_{k\in\mathbb{N}}$ converges to $\Im\mathrm{m}\{\hat{z}\}$. ■

Solution 1.100 *(Exercise 1.100).* Since $(a_k)_{k\in\mathbb{N}}$ is strictly increasing, we have

$$h > k \quad \Rightarrow \quad a_h > a_k.$$

Since $(\beta)_{k\in\mathbb{N}}$ is strictly increasing and its terms are natural numbers, we find

$$h > k \quad \Rightarrow \quad \beta_h > \beta_k \quad \Rightarrow \quad a_{\beta_h} > a_{\beta_k}$$

i.e., $(a \circ \beta)(h) > (a \circ \beta)(k)$. Hence, $a \circ \beta$ is strictly increasing. ■

Solution 1.101 *(Exercise 1.101).* If (**??**) holds true, then (**??**) is true, because $\mathbb{N}^*$ is a proper subset of $\mathbb{R}^*_+$. We now assume that (**??**) is true. Let $c \in \mathbb{R}^*_+$. Let k be the smallest natural number greater than c, i.e., $k = \lceil c \rceil = \min\{i \in \mathbb{N} \,|\, i > c\}$. We have $k > 0$, and in view of (**??**) there exists $\boldsymbol{x} \in \boldsymbol{X}$ such that $\|\boldsymbol{x}\| > k\,\|\boldsymbol{x}\|_\star$. Hence,

$$\|\boldsymbol{x}\| > k\,\|\boldsymbol{x}\|_\star > c\,\|\boldsymbol{x}\|_\star,$$

which proves (**??**). ■

Solution 1.102 *(Exercise 1.102).* The first example is $x_k = (-1)^k$. The subsequence obtained choosing $k_j = 2\,j$ for every $j \in \mathbb{N}$ leads to $x_{k_j} = (-1)^{2\,j} = 1$, so that the subsequence is a sequence whose elements are all equal to 1. Thus, it converges to 1.

The second example is the sequence defined by

$$x_k = \begin{cases} k & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}$$

The terms of the sequence are $0, 2, 0, 4, 0, 6, 0, 8, \ldots$. Therefore, the sequence is not bounded. However, choosing $k_j = 2\,j + 1$ for every $j \in \mathbb{N}$ leads to $x_{k_j} = 0$, and the subsequence converges to 0.

The third example is $x_k = k$, for $k \in \mathbb{N}$. In this case, any subsequence would be strictly increasing, and therefore it would not converge.

The last example cannot be formulated: every bounded sequence has a convergent subsequence in view of Bolzano-Weierstrass Theorem. ■

Solution 1.103 *(Exercise 1.103).* Let, for all $k \in \mathbb{N}$,

$$\boldsymbol{x}_k = x_{k,1}\,\boldsymbol{b}_1 + \ldots + x_{k,n}\,\boldsymbol{b}_n$$

and let

$$\hat{\boldsymbol{x}} = \hat{x}_1\,\boldsymbol{b}_1 + \ldots + \hat{x}_n\,\boldsymbol{b}_n.$$

We have to prove that the following two statements are equivalent:

(A) $\sum_{k=0}^{\infty} \boldsymbol{x}_k$ is convergent and the sum of the series is $\hat{\boldsymbol{x}}$;

(B) for all $i = 1, \ldots, n$, the series $\sum_{k=0}^{\infty} x_{k,i}$ is convergent and the sum of the series is $\hat{x}_i$.

We define the partial sums as

$$\begin{aligned} \boldsymbol{s}_k &= \boldsymbol{x}_0 + \boldsymbol{x}_1 + \ldots + \boldsymbol{x}_k \\ &= (x_{0,1}\,\boldsymbol{b}_1 + \ldots + x_{0,n}\,\boldsymbol{b}_n) + (x_{1,1}\,\boldsymbol{b}_1 + \ldots + x_{1,n}\,\boldsymbol{b}_n) + \ldots + (x_{k,1}\,\boldsymbol{b}_1 + \ldots + x_{k,n}\,\boldsymbol{b}_n) \\ &= (x_{0,1} + x_{1,1} + \ldots + x_{k,1})\,\boldsymbol{b}_1 + \ldots + (x_{0,n} + x_{1,n} + \ldots + x_{k,n})\,\boldsymbol{b}_n. \end{aligned}$$

Then, $(\boldsymbol{s}_k)_{k\in\mathbb{N}}$ is convergent and its limit is $\hat{\boldsymbol{x}}$ if and only if, for all $i = 1, \ldots, n$, the sequence of its i-th component $(s_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$, i.e., if and only if for all $i = 1, \ldots, n$

$$\lim_{k\to\infty} (x_{0,i} + x_{1,i} + \ldots + x_{k,i}) = \hat{x}_i,$$

which is equivalent to saying that for all $i = 1, \ldots, n$ the series $\sum_{k=0}^{\infty} x_{k,i}$ is convergent and the sum of the series is $\hat{x}_i$. ■

Solution 1.104 *(Exercise 1.104).* Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a convergent sequence in $\boldsymbol{X}$, and let $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ be the limit. We have to show that $|f(\boldsymbol{x}_k) - f(\hat{\boldsymbol{x}})| \overset{k\to\infty}{\longrightarrow} 0$. The fact that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent to $\hat{\boldsymbol{x}} \in \boldsymbol{X}$ is equivalent to writing

$$\|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} 0.$$

Thus, using the reverse triangle inequality we find

$$\left|f(\boldsymbol{x}_k) - f(\hat{\boldsymbol{x}})\right| = \left|\|\boldsymbol{x}_k\| - \|\hat{\boldsymbol{x}}\|\right| \leq \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}\| \overset{k\to\infty}{\longrightarrow} 0.$$

From the Squeeze Theorem, we find that $\left|f(\boldsymbol{x}_k) - f(\hat{\boldsymbol{x}})\right| \overset{k\to\infty}{\longrightarrow} 0$, which implies that the sequence $(f(\boldsymbol{x}_k))_{k\in\mathbb{N}}$ converges to $f(\hat{\boldsymbol{x}})$. ■

Solution 1.105 *(Exercise 1.105).* Recall that f is continuous if and only if its m components $f_1, \ldots, f_m$ are continuous (Theorem 52 in the lecture notes on normed spaces). The i-th component of f (for $i \in \{1, \ldots, m\}$) is

$$\begin{aligned} f_i : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto a_{i,1}\,x_1 + a_{i,2}\,x_2 + \ldots + a_{i,n}\,x_n. \end{aligned}$$

We show that it is continuous at any $\boldsymbol{x}_0 \in \mathbb{R}^n$. To this end, we consider the $\|\cdot\|_1$ norm in $\mathbb{R}^n$ to show that

$$\forall\,\varepsilon > 0,\ \exists\,\delta > 0 :\ \forall\,\boldsymbol{x} \in \mathbb{R}^n,$$
$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_1 < \delta \ \Rightarrow\ |a_{i,1}\,x_1 + a_{i,2}\,x_2 + \ldots + a_{i,n}\,x_n - (a_{i,1}\,x_{0,1} + a_{i,2}\,x_{0,2} + \ldots + a_{i,n}\,x_{0,n})| < \varepsilon,$$

where we denoted by $x_{0,1}, \ldots, x_{0,n}$ the components of $\boldsymbol{x}_0$, i.e., $\forall\,\varepsilon > 0,\ \exists\,\delta > 0 :\ \forall\,\boldsymbol{x} \in \mathbb{R}^n$,

$$|x_1 - x_{0,1}| + |x_2 - x_{0,2}| + \ldots + |x_n - x_{0,n}| < \delta \ \Rightarrow\ |a_{i,1}\,(x_1 - x_{0,1}) + a_{i,2}\,(x_2 - x_{2,0}) + \ldots + a_{i,n}\,(x_n - x_{0,n})| < \varepsilon.$$

If the i-th row of A is zero, there is nothing to prove. Assume it is not zero. Let $\alpha = \max_{j=1,\dots,n} |a_{i,j}|$, which is strictly positive because the i-th row is not zero. Then,

$$\begin{aligned} |a_{i,1}(x_1 - x_{0,1}) + a_{i,2}(x_2 - x_{2,0}) + \dots + a_{i,n}(x_n - x_{0,n})| &\leq |a_{i,1}||x_1 - x_{0,1}| + |a_{i,2}||x_2 - x_{0,2}| + \dots + |a_{i,n}||x_n - x_{0,n}| \\ &\leq \alpha |x_1 - x_{0,1}| + \alpha |x_2 - x_{0,2}| + \dots + \alpha |x_n - x_{0,n}| \\ &\leq \alpha \delta. \end{aligned}$$

The statement then follows by taking $\delta < \varepsilon/\alpha$. ■

Solution 1.106 *(Exercise 1.106).* We notice that f is continuous because it is convergent at $x = 0$

$$\lim_{x \to 0^-} f(x) = \lim_{x \to 0^-} (x^2 + 2x + 1) = 1 \quad \text{and} \quad \lim_{x \to 0^+} f(x) = \lim_{x \to 0^+} (-x + 1) = 1$$

and the value of the limit (which is 1) coincides with $f(0) = 1$. Now,

$$\begin{aligned} \|f\|_1 &= \int_{-1}^{0} |x^2 + 2x + 1| \, \mathrm{d}x + \int_{0}^{2} |-x + 1| \, \mathrm{d}x \\ &= \int_{-1}^{0} (x^2 + 2x + 1) \, \mathrm{d}x + \int_{0}^{1} (-x + 1) \, \mathrm{d}x + \int_{1}^{2} (x - 1) \, \mathrm{d}x \\ &= \left[\frac{x^3}{3} + 2\frac{x^2}{2} + x \right]_{-1}^{0} + \left[-\frac{x^2}{2} + x \right]_{0}^{1} + \left[\frac{x^2}{2} - x \right]_{1}^{2} = \frac{4}{3}. \end{aligned}$$

To find the norm $\|f\|_\infty$, we notice that f is increasing in $[-1, 0]$ (since $f'(x) = 2x + 2$ is always non-negative in $[-1, 0]$) and it is decreasing in $[0, 2]$. The maximum is attained at $x = 0$, and it is equal to $f(0) = 1$, so that

$$\|f\|_\infty = 1.$$

■

Solution 1.107 *(Exercise 1.107).* For $k = 1, 2, 3$, we have, for $x \in [0, 1]$,

$$f_1(x) = \begin{cases} 2 & \text{for } x \in \left[\frac{1}{2}, 1\right] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_2(x) = \begin{cases} 4 & \text{for } x \in \left[\frac{1}{4}, \frac{1}{2}\right] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f_3(x) = \begin{cases} 6 & \text{for } x \in \left[\frac{1}{6}, \frac{1}{3}\right] \\ 0 & \text{otherwise} \end{cases}$$

The pointwise limit is the zero function in $[0, 1]$. We have

$$\lim_{k \to \infty} \int_0^1 f_k(x) \, \mathrm{d}x = \lim_{k \to \infty} \int_{\frac{1}{2k}}^{\frac{1}{k}} 2k \, \mathrm{d}x = \lim_{k \to \infty} 2k \left(\frac{1}{k} - \frac{1}{2k} \right) = 1,$$

but

$$\int_0^1 f(x) \, \mathrm{d}x = 0.$$

To prove that $(f_k)_{k \in \mathbb{N}^*}$ is not uniformly convergent to the zero function, it suffices to show that the sequence

$$\left(\sup_{x \in [0,1]} |f_k(x) - 0| \right)_{k \in \mathbb{N}^*}$$

is not infinitesimal. Let $k \in \mathbb{N}^*$. Thus,

$$\sup_{x \in [0,1]} |f_k(x)| = 2k,$$

and the sequence $(2\,k)_{k\in\mathbb{N}^*}$ is divergent (and therefore not infinitesimal).

The pointwise limit of the moving bump is still the zero function. Thus,

$$\int_{\mathbb{R}} f_k(x)\,\mathrm{d}x = \lim_{N\to+\infty}\int_{-N}^{N} f_k(x)\,\mathrm{d}x = \lim_{N\to+\infty}\int_{k}^{k+1} 1\,\mathrm{d}x = \lim_{N\to+\infty}(k+1-k) = 1.$$

On the other hand,

$$\int_{\mathbb{R}} f(x)\,\mathrm{d}x = \lim_{N\to+\infty}\int_{-N}^{N} f(x)\,\mathrm{d}x = \lim_{N\to+\infty}\int_{-N}^{N} 0\,\mathrm{d}x = 0.$$

Again, the moving bump is not uniformly convergent, because, given $k\in\mathbb{N}$, we have $\sup_{x\in\mathbb{R}} f_k(x) = 1$, and the sequence $(1)_{k\in\mathbb{N}}$ is not infinitesimal. ■

11 Solutions to the Exercises in Chapter 2

Solution 2.1 *(Exercise 2.1).* Indeed, assume that $\overline{\boldsymbol{A}} \cap \mathbf{B} = \boldsymbol{A}$, and suppose that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a sequence of points of $\boldsymbol{A}$ and convergent to a point of $\mathbf{B}$. The limit of the sequence is also in $\overline{\boldsymbol{A}}$, or else there would be a ball around the limit entirely contained in $\boldsymbol{X} \setminus \boldsymbol{A}$, and as $k \to \infty$ the terms of the sequence would not get into that ball. This implies that the limit is in $\overline{\boldsymbol{A}} \cap \mathbf{B}$, and therefore the limit is in $\boldsymbol{A}$. Conversely, suppose that given any sequence $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ of points of $\boldsymbol{A}$ and convergent to a point of $\mathbf{B}$, the limit is in $\boldsymbol{A}$. Since $\overline{\boldsymbol{A}} \supseteq \boldsymbol{A}$ and $\mathbf{B} \supseteq \boldsymbol{A}$, we have $\overline{\boldsymbol{A}} \cap \mathbf{B} \supseteq \boldsymbol{A}$. We need to prove the opposite inclusion: take $\boldsymbol{x} \in \overline{\boldsymbol{A}} \cap \mathbf{B}$ and suppose by contradiction that $\boldsymbol{x} \notin \boldsymbol{A}$. Consider a ball $\mathcal{B}_r(\boldsymbol{x})$. Since $\boldsymbol{x} \in \overline{\boldsymbol{A}}$, this ball contains points of $\boldsymbol{A}$. Taking $r = \frac{1}{k}$ with $k \in \mathbb{N}^*$, we construct a sequence that converges to $\boldsymbol{x} \in \mathbf{B}$, but its limit must be in $\boldsymbol{A}$ by assumption, and we reach a contradiction. ■

Solution 2.2 *(Exercise 2.2).* The given sequence is not convergent, and therefore it is not a Cauchy sequence. However, the condition above is satisfied. In fact,

$$|x_{n+1} - x_n| = \ln\left(1 + \frac{1}{n}\right),$$

which tends to 0 as $n \to \infty$. ■

Solution 2.3 *(Exercise 2.3).* Since $(z_k)_{k\in\mathbb{N}}$ is a Cauchy sequence,

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\quad \forall\, i, j \in \mathbb{N}, \qquad i, j > p \ \Rightarrow\ |z_i - z_j| < \varepsilon.$$

If, for all $k \in \mathbb{N}$, we define $x_k = \Re\mathrm{e}\{z_k\}$ and $y_k = \Im\mathrm{m}\{z_k\}$, we can write $z_k = x_k + i\, y_k$ for all $k \in \mathbb{N}$. Hence,[1]

$$|z_i - z_j| = |(x_i + i\, y_i) - (x_j + i\, y_j)| = |(x_i - x_j) + i\,(y_i - y_j)| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}.$$

On the other hand,

$$\begin{aligned} |x_i - x_j| &\le \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \\ |y_i - y_j| &\le \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}, \end{aligned}$$

which imply that for all $i, j > p$ we have $|x_i - x_j| < \varepsilon$ and $|y_i - y_j| < \varepsilon$. Hence, the two real sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ are Cauchy sequences. Therefore, they are convergent from Theorem 2 in the lecture notes. Let $\hat{x}$ and $\hat{y}$ denote their limits. We show that for every $\varepsilon > 0$ there exists $p \in \mathbb{N}$ such that for all $k > p$ we have $|z_k - (\hat{x} + i\,\hat{y})| < \varepsilon$, which shows that $\hat{x} + i\,\hat{y}$ is the limit of $(z_k)_{k\in\mathbb{N}}$. Since $(x_k)_{k\in\mathbb{N}}$ converges to $\hat{x}$, there exists $p_1 \in \mathbb{N}$ such that for all $k > p_1$ we have $|x_k - \hat{x}| < \varepsilon/2$. Similarly, since $(y_k)_{k\in\mathbb{N}}$ converges to $\hat{y}$, there exists $p_2 \in \mathbb{N}$ such that for all $k > p_2$ we have $|y_k - \hat{y}| < \varepsilon/2$. It follows that for all $k > \max\{p_1, p_2\}$ we have

$$\begin{aligned} |z_k - (\hat{x} + i\,\hat{y})| &= |x_k + i\, y_k - \hat{x} - i\,\hat{y}| \\ &= |(x_k - \hat{x}) + i\,(y_k - \hat{y})| \le |x_k - \hat{x}| + |i\,(y_k - \hat{y})| \\ &= |x_k - \hat{x}\}| + |y_k - \hat{y}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

[1] The meaning of this equation should be clear despite the abuse of notation of using the letter i both for the imaginary unit and for one of the two indices in the condition defining a Cauchy sequence.

Then, $\hat{z} \stackrel{\text{def}}{=} \hat{x} + i\,\hat{y}$ is the limit of $(z_k)_{k\in\mathbb{N}}$. ■

Solution 2.4 *(Exercise 2.4).* Consider the first sequence $(k^2)_{k\in\mathbb{N}}$. The sequence is a Cauchy sequence if

$$\forall\,\varepsilon > 0,\ \ \exists\, p \in \mathbb{N} :\ \ \forall\, n, m > p, \qquad |n^2 - m^2| < \varepsilon.$$

Let $\varepsilon > 0$, say $\varepsilon = 1/2$, and suppose by contradiction that such p exists. Take for example $n = p + 2$ and $m = p + 1$, so that both n and m are greater than p. We find

$$|(p+2)^2 - (p+1)^2| < \frac{1}{2} \quad \Rightarrow \quad |p^2 + 4\,p + 4 - p^2 - 2\,p - 1| < \frac{1}{2} \quad \Rightarrow \quad 2\,p + 3 < \frac{1}{2} \quad \Rightarrow \quad p < -\frac{5}{2},$$

which leads to a contradiction because there are no natural numbers smaller than $-5/2$.

Consider the second sequence. The Cauchy's condition is

$$\forall\,\varepsilon > 0,\ \ \exists\, p \in \mathbb{N} :\ \ \forall\, n, m > p, \qquad |(-1)^n - (-1)^m| < \varepsilon.$$

Suppose that ε is assigned, and that p exists. Choosing again $n = p + 2$ and $m = p + 1$, we find

$$|(-1)^{p+2} - (-1)^{p+1}| < \varepsilon. \tag{11.1}$$

If p is odd, then $p + 1$ is even and $p + 2$ is odd, so that (11.1) becomes

$$|-1-1| < \varepsilon,$$

which says that $\varepsilon > 2$: hence, ε cannot be chosen to be arbitrarily small. If p is even, then $p + 1$ is odd and $p + 2$ is even, so that (11.1) becomes

$$|1 - (-1)| = 2 < \varepsilon,$$

which leads to the same contradiction. ■

Solution 2.5 *(Exercise 2.5).*

(i) The set

$$U = \{f \in C([0,1],\mathbb{R}) \mid f(0) = 0\}$$

is a subspace of $C([0,1],\mathbb{R})$: linearly combining functions that are zero at zero gives functions that are zero at zero. Indeed, given $f, g \in U$ and $\alpha, \beta \in \boldsymbol{U}$, we know that the function $\alpha\,f + \beta\,g$ is continuous, and from the definition of addition and scalar multiplication of functions

$$(\alpha\,f + \beta\,g)(0) = \alpha\,f(0) + \beta\,g(0) = \alpha \cdot 0 + \beta \cdot 0 = 0 \quad \Rightarrow \quad \alpha\,f + \beta\,g \in U.$$

Consider a sequence $(f_k)_{k\in\mathbb{N}}$ in U, which converges to a function f in $C([0,1],\mathbb{R})$. Therefore, $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. We must prove that f is in U, i.e., that it is zero at 0. The uniform convergence implies the pointwise convergence, and therefore for all $x \in [0,1]$, we have $\lim_{k\to\infty} f_k(x) = f(x)$, which implies that for $x = 0$ we have $\lim_{k\to\infty} f_k(0) = f(0)$. Since for all $k \in \mathbb{N}$ we have $f_k(0) = 0$, we obtain $f(0) = 0$, and therefore $f \in U$. Therefore, U is closed.

(ii) The set

$$U = \{f \in C([0,1],\mathbb{R}) \mid f \geq 0\}$$

is not a subspace of $C([0,1],\mathbb{R})$: if you multiply a function in U by a negative scalar, the result is no longer non-negative, and so it is not an element of U. Consider a sequence $(f_k)_{k\in\mathbb{N}}$ in U, which converges to a function f in $C([0,1],\mathbb{R})$. Therefore, $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. We must prove that f is in U, i.e., that it is non-negative. The uniform convergence implies the pointwise convergence, and the pointwise limit is the same as the uniform limit f. We have seen that if, for all $k \in \mathbb{N}$, the function f_k is non-negative, the pointwise limit is also non-negative. Thus, $f(x) \geq 0$ for all $x \in [0,1]$, and therefore $f \in U$. Therefore, U is closed.

(iii) Let

$$U = \left\{ f \in C([0,1],\mathbb{R}) \;\middle|\; \int_0^1 f(x)\,\mathrm{d}x = 0 \right\}$$

is a subspace. In fact, we know that a linear combination $\alpha f + \beta g$ of two of its elements with scalars $\alpha, \beta \in \mathbb{R}$ is still continuous, and using the linearity of the Riemann integral

$$\int_0^1 (\alpha f+\beta g)(x)\,\mathrm{d}x = \int_0^1 (\alpha f(x)+\beta g(x))\,\mathrm{d}x = \alpha \int_0^1 f(x)\,\mathrm{d}x+\beta \int_0^1 g(x)\,\mathrm{d}x = \alpha\cdot 0+\beta\cdot 0 = 0 \quad \Rightarrow \quad \alpha f+\beta g \in U$$

Consider a sequence $(f_k)_{k\in\mathbb{N}}$ in U (so that $\int_0^1 f_k(x)\,\mathrm{d}x = 0$ for all $k \in \mathbb{N}$), which converges to a function f in $C([0,1],\mathbb{R})$. Therefore, $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. We must prove that f is in U, i.e., that $\int_0^1 f(x)\,\mathrm{d}x = 0$. Since the domain is bounded and the functions of the sequence are continuous, by Theorem 1.58 the uniform convergence implies that limits can be interchanged with integrals, so that

$$0 = \lim_{k\to\infty} \underbrace{\int_0^1 f_k(x)\,\mathrm{d}x}_{=0} = \int_0^1 \lim_{k\to\infty} f_k(x)\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x.$$

This shows that $f \in U$.

(iv) The subset

$$U = \left\{ f \in C([0,1],\mathbb{R}) \;\middle|\; \int_0^1 f(x)\,\mathrm{d}x \geq 0 \right\}$$

of the vector space $C([0,1],\mathbb{R})$ is not a subspace. In fact, it is not closed under scalar multiplication (consider e.g. the constant function $f(x) = 1$ for $x \in [0,1]$. The value of the integral is 1, which is positive, but the integral of αf, for $\alpha = -2$, is equal to -2).
Consider a sequence $(f_k)_{k\in\mathbb{N}}$ in U (so that $\int_0^1 f_k(x)\,\mathrm{d}x \geq 0$ for all $k \in \mathbb{N}$), which converges to a function f in $C([0,1],\mathbb{R})$. Therefore, $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. We must prove that f is in U, i.e., that $\int_0^1 f(x)\,\mathrm{d}x \geq 0$. The uniform convergence implies that limits can be interchanged with integrals, so that, remembering that limits preserve the sign, we find

$$0 \leq \lim_{k\to\infty} \int_0^1 f_k(x)\,\mathrm{d}x = \int_0^1 \lim_{k\to\infty} f_k(x)\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x.$$

This shows that $f \in U$.

(v) Let us denote by U the subset of functions differentiable at $1/2$ and with derivative at $1/2$ equal to 0. This is a subspace of $C([0,1],\mathbb{R})$, because the differentiability is linear, and if $f, g \in U$ and $\alpha, \beta \in \mathbb{R}$ we find

$$(\alpha f + \beta g)'\left(\frac{1}{2}\right) = \alpha f'\left(\frac{1}{2}\right) + \beta g'\left(\frac{1}{2}\right) = \alpha \cdot 0 + \beta \cdot 0 = 0.$$

To find that U is not a subspace, we can modify an example in the Lecture Notes on normed spaces to see that the sequence

$$f_k(x) = \sqrt{\left(x - \frac{1}{2}\right)^2 + \frac{1}{k^2}}$$

is uniformly convergent to $f(x) = \left|x - \frac{1}{2}\right|$. These functions $f_k(x)$ live in U, because

$$f_k'(x) = \frac{1}{2\sqrt{\left(x - \frac{1}{2}\right)^2 + \frac{1}{k^2}}} 2\left(x - \frac{1}{2}\right),$$

and for $x = 1/2$ we get 0. However, f is not differentiable at $x = 1/2$, and therefore $f \notin U$. Hence, U is not closed.

(vi) The set of functions differentiable in $(0,1)$ and such that f' is identically zero is the set of constant functions in $[0,1]$. Linearly combining constant functions yield constant functions, and therefore this set U is a subspace. Take a sequence $(f_k)_{k\in\mathbb{N}}$ in U (so that each f_k is constant in $[0,1]$) uniformly convergent to a function f. We show that f is also constant. Since f is also the pointwise limit, for every $x \in [0,1]$ we have $\lim_{k\to\infty} f_k(x) = f(x)$. The left hand-side does not depend on x, and therefore the right hand-side also does not depend on x, and this implies that f is constant. It follows that U is closed.

(vii) The set of polynomial functions of degree ≤ 2

$$U = \{a_0 + a_1 x + a_2 x^2 \,|\, a_0, a_1, a_2 \in \mathbb{R}\}$$

is a subset of $C([0,1],\mathbb{R})$, because polynomial functions are continuous. This subspace has dimension 3, because a basis is given by the real functions $1, x, x^2$. Since it is a finite-dimensional subspace, we can conclude that it is closed.

(viii) The set of polynomial functions of degree exactly 2 is not a subspace: combining the functions $f(x) = x^2 + 1$ with $g(x) = 2x^2 + x + 3$ with coefficients -2 and 1 gives a polynomial of degree 1. We can use this consideration to see that this set is not closed. Consider the sequence of functions given by $f_k(x) = \frac{1}{k} x^2$. This sequence of functions is uniformly convergent to the zero function, because

$$\sup_{x\in[0,1]} \left|\frac{1}{k} x^2 - 0\right| = \frac{1}{k},$$

and this sequence goes to 0 as $k \to \infty$. However, f is not a polynomial function of degree 2.

(ix) The set of polynomial functions in $[0,1]$ is a subspace (let us call it U) of the set of continuous functions in $[0,1]$. Take for example the sequence in U

$$f_k(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots + \frac{x^k}{k!}.$$

For each $k \in \mathbb{N}$, these functions are elements of U because they are polynomial functions. However, f_k is uniformly convergent to the function e^x, since

$$\sup_{x\in[0,1]} \left|f_k(x) - e^x\right| = \sup_{x\in[0,1]} \left|1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \ldots + \frac{x^k}{k!} - \sum_{i=0}^{\infty} \frac{x^i}{i!}\right|$$
$$= \sum_{i=k+1}^{\infty} \frac{x^i}{i!} = \frac{1}{(k+1)!} + \frac{1}{(k+2)!} + \ldots,$$

which converges to 0 as $k \to \infty$. Therefore, $f \notin U$ and U is not closed.

(x) The set U of continuous functions in $[0, 1]$ such that $f(0) = f(1)$ is a subspace (why?). Let $(f_k)_{k\in\mathbb{N}}$ be a sequence in U which converges uniformly to f. Then, it converges to f pointwise, so that for all $x \in [0, 1]$, we have $\lim_{k\to\infty} f_k(x) = f(x)$. It follows that

$$f(0) - f(1) = \lim_{k\to\infty} f_k(0) - \lim_{k\to\infty} f_k(1) = \lim_{k\to\infty} (f_k(0) - f_k(1)) = \lim_{k\to\infty} 0 = 0.$$

Thus, $f(0) = f(1)$, and therefore $f \in U$. Hence, U is closed.

■

Solution 2.6 *(Exercise 2.6).* This set U of all continuous decreasing functions in $[0, 1]$ is not a subspace of $C([0, 1], \mathbb{R})$: if you multiply a function in U by a negative scalar, the result is no longer a decreasing function. Consider a sequence $(f_k)_{k\in\mathbb{N}}$ in U, which converges to a function f in $C([0, 1], \mathbb{R})$. Therefore, $(f_k)_{k\in\mathbb{N}}$ is uniformly convergent to f. We must prove that f is in U, i.e., that it is decreasing. The uniform convergence implies the pointwise convergence, and the pointwise limit is the same as the uniform limit f. We have seen that if, for all $k \in \mathbb{N}$, the function f_k is decreasing, the pointwise limit is also decreasing (Theorem 22 of the Lecture Notes on sequences). Thus, f is decreasing, and therefore $f \in U$. Therefore, $\boldsymbol{U}$ is closed. ■

Solution 2.7 *(Exercise 2.7).* We need to show that, knowing that $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_1$, for all $\varepsilon > 0$ there exists p such that for all $n, m > p$ we have $\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_2 < \varepsilon$. Let $\varepsilon > 0$. Moreover, since the two norms are equivalent, there exists $c_2 > 0$ such that for all $\boldsymbol{\xi} \in \boldsymbol{X}$ we have $\|\boldsymbol{\xi}\|_2 \leq c_2 \|\boldsymbol{\xi}\|_1$, so that in particular

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_2 \leq c_2 \|\boldsymbol{x}_n - \boldsymbol{x}_m\|_1.$$

Since $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to $\|\cdot\|_1$, there exists $p \in \mathbb{N}$ such that for all $n, m > p$ we have

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_1 < \frac{\varepsilon}{c_2}.$$

Thus, for all $n, m > p$ we obtain

$$\|\boldsymbol{x}_n - \boldsymbol{x}_m\|_2 \leq c_2 \|\boldsymbol{x}_n - \boldsymbol{x}_m\|_1 < c_2 \frac{\varepsilon}{c_2} = \varepsilon.$$

■

Solution 2.8 *(Exercise 2.8).* The only sequence of points in $\{\boldsymbol{x}\}$ is the constant sequence, and so its limit is $\boldsymbol{x}$, which is in $\{\boldsymbol{x}\}$. ■

Solution 2.9 *(Exercise 2.9).* Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}\cap\boldsymbol{H}$. Then, $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}$, and since $\boldsymbol{K}$ is sequentially compact, there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ that converges in $\boldsymbol{K}$. This subsequence is also a sequence of points of $\boldsymbol{H}$, and since $\boldsymbol{H}$ is sequentially compact, there exists a subsequence $(\boldsymbol{x}_{k_{j_h}})_{h\in\mathbb{N}}$ that converges in $\boldsymbol{H}$. Since

$$\lim_{h\to\infty}\boldsymbol{x}_{k_{j_h}}=\lim_{j\to\infty}\boldsymbol{x}_{k_j},$$

the subsequence $(\boldsymbol{x}_{k_{j_h}})_{h\in\mathbb{N}}$ is also convergent in $\boldsymbol{K}$, and therefore it is a sequence that converges in $\boldsymbol{K}\cap\boldsymbol{H}$. This shows that $\boldsymbol{K}\cap\boldsymbol{H}$ is compact.

Now, we show that $\boldsymbol{K}\cup\boldsymbol{H}$ is compact. Let $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ be a sequence of points in $\boldsymbol{K}\cup\boldsymbol{H}$. Then, $\{k\in\mathbb{N}\,|\,\boldsymbol{x}_k\in\boldsymbol{K}\}$ is infinite, or $\{k\in\mathbb{N}\,|\,\boldsymbol{x}_k\in\boldsymbol{H}\}$ is infinite. Suppose for the sake of argument that $\{k\in\mathbb{N}\,|\,\boldsymbol{x}_k\in\boldsymbol{K}\}$ is infinite. Then, there exists a subsequence $(\boldsymbol{x}_{k_j})_{j\in\mathbb{N}}$ such that $\boldsymbol{x}_{k_j}\in\boldsymbol{K}$ for all $j\in\mathbb{N}$. Since A is compact, there exists a subsequence $(\boldsymbol{x}_{k_{j_h}})_{h\in\mathbb{N}}$ that has limit in $\boldsymbol{K}$. This is a subsequence of $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ that converges in $\boldsymbol{K}$, and therefore it converges also in $\boldsymbol{K}\cup\boldsymbol{H}$. This shows that $\boldsymbol{K}\cup\boldsymbol{H}$ is sequentially compact. ■

Solution 2.10 *(Exercise 2.10).* For this question, we can use the same argument in the proof of Theorem 2.1. Let $\varepsilon>0$. Since $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly,

$$\exists\, p>0:\ \forall\, k>p,\ \forall\, x\in A,\quad \|f_k(x)-f(x)\|<\frac{\varepsilon}{2}.$$

Therefore, for every $n,m>p$ and for every $x\in A$ we have

$$\|f_n(x)-f_m(x)\|=\|f_n(x)-f(x)+f(x)-f_m(x)\|\le\|f_n(x)-f(x)\|+\|f(x)-f_m(x)\|\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

where we used the triangle inequality. ■

Solution 2.11 *(Exercise 2.11).* Let, for all $k\in\mathbb{N}$,

$$\boldsymbol{x}_k=x_{k,1}\,\boldsymbol{b}_1+\ldots+x_{k,n}\,\boldsymbol{b}_n$$

and let

$$\hat{\boldsymbol{x}}=\hat{x}_1\,\boldsymbol{b}_1+\ldots+\hat{x}_n\,\boldsymbol{b}_n.$$

We have to prove that the following two statements are equivalent:

(A) $\displaystyle\sum_{k=0}^{\infty}\boldsymbol{x}_k$ is convergent and the sum of the series is $\hat{\boldsymbol{x}}$;

(B) for all $i\in\{1,2,\ldots,n\}$, the series $\displaystyle\sum_{k=0}^{\infty}x_{k,i}$ is convergent and the sum of the series is $\hat{x}_i$.

We define the partial sums as

$$\begin{aligned}\boldsymbol{s}_k&=\boldsymbol{x}_0+\boldsymbol{x}_1+\ldots+\boldsymbol{x}_k\\&=(x_{0,1}\,\boldsymbol{b}_1+\ldots+x_{0,n}\,\boldsymbol{b}_n)+(x_{1,1}\,\boldsymbol{b}_1+\ldots+x_{1,n}\,\boldsymbol{b}_n)+\ldots+(x_{k,1}\,\boldsymbol{b}_1+\ldots+x_{k,n}\,\boldsymbol{b}_n)\\&=(x_{0,1}+x_{1,1}+\ldots+x_{k,1})\,\boldsymbol{b}_1+\ldots+(x_{0,n}+x_{1,n}+\ldots+x_{k,n})\,\boldsymbol{b}_n.\end{aligned}$$

Then, $(\boldsymbol{s}_k)_{k\in\mathbb{N}}$ is convergent and its limit is $\hat{\boldsymbol{x}}$ if and only if, for all $i\in\{1,2,\ldots,n\}$, the sequence of its i-th component $(s_{k,i})_{k\in\mathbb{N}}$ converges to $\hat{x}_i$ (see Theorem 1.45), i.e., if and only if for all $i\in\{1,2,\ldots,n\}$

$$\lim_{k\to\infty}(x_{0,i}+x_{1,i}+\ldots+x_{k,i})=\hat{x}_i,$$

which is equivalent to saying that for all $i \in \{1, 2, \ldots, n\}$ the series $\sum_{k=0}^{\infty} x_{k,i}$ is convergent and the sum of the series is $\hat{x}_i$. ■

Solution 2.12 *(Exercise 2.12).*

(i) The addition of sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ and the scalar multiplication of the sequence $(x_k)_{k\in\mathbb{N}}$ by the scalar $\lambda \in \mathbb{R}$ are defined as usual:

$$x + y : \mathbb{N} \longrightarrow \mathbb{R} \quad k \mapsto x_k + y_k \qquad \text{and} \qquad \lambda x : \mathbb{N} \longrightarrow \mathbb{R} \quad k \mapsto \lambda x_k.$$

We have also established the linearity of the limit for convergent sequences (and for some divergent sequences) in Theorem 1.33. Thus, if $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ are in c_0 and if $\lambda, \mu \in \mathbb{R}$, we have

$$\lim_{k\to\infty} (\lambda x_k + \mu y_k) = \lim_{k\to\infty} \lambda x_k + \mu \lim_{k\to\infty} y_k = 0 + 0 = 0,$$

which says that $(\lambda x_k + \mu y_k)_{k\in\mathbb{N}}$ is also in c_0. Therefore, c_0 is a subspace of $\mathbb{R}^{\mathbb{N}}$ (which is a real vector space because, as you might recall from your Linear Algebra courses, in general the space of functions $f : S \longrightarrow \boldsymbol{V}$, where S is an arbitrary set and $\boldsymbol{V}$ is a vector space over a field $\mathbb{F}$, is a vector space over $\mathbb{F}$). The zero of c_0 is the zero sequence $(0, 0, 0, \ldots)$.

(ii) Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in c_0, so that it is convergent. Convergent sequences are bounded (see Theorem 1.32), and therefore $\sup_{k\in\mathbb{N}} |x_k|$ exists in $\mathbb{R}$. We now check the norm axioms. Let $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ be sequences in c_0 and let $\lambda \in \mathbb{R}$. Using the result of Exercise 1.15 we have

$$\|\lambda x\|_\infty = \sup_{k\in\mathbb{N}} |\lambda x_k| = \sup_{k\in\mathbb{N}} (|\lambda| \cdot |x_k|) \overset{\text{since } |\lambda| \geq 0}{=} |\lambda| \sup_{k\in\mathbb{N}} |x_k| = |\lambda| \cdot \|x\|_\infty$$

We check the subadditivity:

$$\|x + y\|_\infty = \sup_{k\in\mathbb{N}} |x_k + y_k| \leq \sup_{k\in\mathbb{N}} (|x_k| + |y_k|) \leq \sup_{k\in\mathbb{N}} |x_k| + \sup_{k\in\mathbb{N}} |y_k| = \|x\|_\infty + \|y\|_\infty.$$

The positivity is trivial: if $\|x\|_\infty = \sup_{k\in\mathbb{N}} |x_k| = 0$, the sequence x must be the zero sequence.

(iii) We prove that $(c_0, \|\cdot\|_\infty)$ is a Banach space. We use the definition: we take a Cauchy sequence in c_0, and we prove that it is convergent with respect to the norm $\|\cdot\|_\infty$. Let $(x^{(j)})_{j\in\mathbb{N}}$ denote a sequence of elements of c_0. In other words, $(x^{(j)})_{j\in\mathbb{N}}$ is a sequence of convergent sequences. Let $(x^{(j)})_{j\in\mathbb{N}}$ be a Cauchy sequence in c_0. We now prove that, for every $k \in \mathbb{N}$, the sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence. This involves showing that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, n, m > p, \quad |x_k^{(n)} - x_k^{(m)}| < \varepsilon. \tag{11.2}$$

Let $\varepsilon > 0$. Since $(x^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence,

$$\exists\, p \in \mathbb{N} :\ \forall\, n, m > p, \quad \|x^{(n)} - x^{(m)}\|_\infty < \epsilon,$$

which is the same, by definition of this norm, as

$$\exists\, p \in \mathbb{N} :\ \forall\, n, m > p, \quad \sup_{k\in\mathbb{N}} |x_k^{(n)} - x_k^{(m)}| < \varepsilon. \tag{11.3}$$

This value of p works in (11.2), since $\sup_{k\in\mathbb{N}} |x_k^{(n)} - x_k^{(m)}| \leq \varepsilon$ implies that for every $k \in \mathbb{N}$ we have $|x_k^{(n)} - x_k^{(m)}| < \varepsilon$.

Now, for a fixed $k \in \mathbb{N}$, the Cauchy sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ is a sequence of $\mathbb{R}$, and we know that $\mathbb{R}$ is complete. It follows that $(x_k^{(j)})_{j\in\mathbb{N}}$ is also convergent. Let

$$\ell_k = \lim_{j\to\infty} x_k^{(j)}. \tag{11.4}$$

In this way, we have defined the sequence $(\ell_k)_{k\in\mathbb{N}}$. The following diagram shows the argument that we are constructing:

$$\begin{array}{rcl}
x^{(0)} & = & (\ x_0^{(0)},\ x_1^{(0)},\ x_2^{(0)},\ \dots\ ) \\
x^{(1)} & = & (\ x_0^{(1)},\ x_1^{(1)},\ x_2^{(1)},\ \dots\ ) \\
x^{(2)} & = & (\ x_0^{(2)},\ x_1^{(2)},\ x_2^{(2)},\ \dots\ ) \\
 & \vdots & \\
x^{(j)} & = & (\ x_0^{(j)},\ x_1^{(j)},\ x_2^{(j)},\ \dots\ ) \\
 & \vdots & \\
 & \downarrow\ \ \downarrow\ \ \downarrow\ \ \dots & \\
 & \ell_0\ \ \ell_2\ \ \ell_2\ \ \dots &
\end{array}$$

Now, we show that this sequence $\ell : \mathbb{N} \longrightarrow \mathbb{R}$ thus defined satisfies

$$\ell = \lim_{j\to\infty} x^{(j)},$$

i.e., that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, j > p, \quad \|x^{(j)} - \ell\|_\infty \leq \varepsilon.$$

Let $\varepsilon > 0$. From (11.4)

$$\exists\, p \in \mathbb{N} :\ \forall\, j > p, \quad |x_k^{(j)} - \ell_k| < \varepsilon. \tag{11.5}$$

This value of p is the one we need. Indeed, if $j > p$, from (11.5) we see that ε is an upper bound of

$$A = \{|x_k^{(j)} - \ell_k| \,|\, k \in \mathbb{N}\},$$

and by definition $\|x^{(j)} - \ell\|_\infty$ is the least upper bound of A, and so $\|x^{(j)} - \ell\|_\infty \leq \varepsilon$.

To conclude, we must show that the sequence $(\ell_k)_{k\in\mathbb{N}}$ is in c_0, i.e., that it converges to 0. This means that we must verify that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k > N, \quad |\ell_k| < \varepsilon.$$

Let $\varepsilon > 0$. Since $(x^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence,

$$\exists\, N \in \mathbb{N} :\ \forall\, n, m > N, \quad \|x^{(n)} - x^{(m)}\|_\infty < \frac{\varepsilon}{2},$$

and so for $n, m > N$ we have

$$\sup_{k\in\mathbb{N}} |x_k^{(n)} - x_k^{(m)}| < \frac{\varepsilon}{2}, \tag{11.6}$$

Let $m > N$. Since $(x_k^{(m)})_{k\in\mathbb{N}}$ is in c_0, it is infinitesimal, and so

$$\exists\, p_1(m)\, in\mathbb{N} :\ \forall\, k > p_1, \quad |x_k^{(n)}| < \frac{\varepsilon}{2},$$

Taking $p = p_1(m)$, for all $k > p$ we find

$$\begin{aligned} |\ell_k| &= \lim_{j\to\infty} |x_k^{(j)}| \le \lim_{j\to\infty} (|x_k^{(j)} - x_k^{(m)}| + |x_k^{(m)}|) \\ &= |x_k^{(m)}| + \lim_{j\to\infty} |x_k^{(j)} - x_k^{(m)}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Note that $\lim_{j\to\infty} |x_k^{(j)} - x_k^{(m)}| < \frac{\varepsilon}{2}$ is a consequence of (11.6) and Exercise 1.67.

(iv) The set c_{00} contains those sequences $(x_k)_{k\in\mathbb{N}}$ (and only those) for which there exists $N \in \mathbb{N}$ such that, for all $k > N$, there holds $x_k = 0$. From this fact, and noting that a sequence in c_{00} converges to 0, it follows immediately that c_{00} is closed under addition and scalar multiplication, and therefore it is a subspace of c_0. We now show that c_{00} is not closed. Consider the sequence $(x^{(j)})_{j\in\mathbb{N}}$ in c_{00} defined as follows:

$$x_k^{(j)} = \begin{cases} \frac{1}{k+1} & \text{if } k \le j \\ 0 & \text{if } k > j, \end{cases}$$

or, more explicitly,

$$\begin{aligned} x^{(0)} &= (1, 0, 0, \ldots) \\ x^{(1)} &= \left(1, \frac{1}{2}, 0, \ldots\right) \\ x^{(3)} &= \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right) \\ &\vdots \\ x^{(j)} &= \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{j+1}, 0, \ldots\right) \\ &\vdots \end{aligned}$$

We have $x^{(j)} \in c_{00}$ for all $j \in \mathbb{N}$. Now, we define $(\hat{x}_k)_{k\in\mathbb{N}}$ as $\hat{x}_k = \frac{1}{k+1}$ for all $k \in \mathbb{N}$, so that $\hat{x} \in c_0$. We show that $\lim_{j\to\infty} x^{(j)} = \hat{x}$. To this end, we show that

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, j > p, \quad \|x^{(j)} - \hat{x}\|_\infty < \varepsilon.$$

Let $\varepsilon > 0$. Let $p > \frac{1}{\varepsilon} - 1$. For all $j > p$ we have

$$\|x^{(j)} - \hat{x}\|_\infty = \sup_{k\in\mathbb{N}} |x_k^{(j)} - \hat{x}_k| \le \frac{1}{p+1} < \varepsilon.$$

This shows that $\lim_{j\to\infty} x^{(j)} = \hat{x}$. On the other hand, $\hat{x} \notin c_{00}$, and so c_{00} is not a closed subspace of c_0.

(v) We show that the sequence $(x^{(j)})_{j\in\mathbb{N}}$ defined in the previous point is a Cauchy sequence. In fact, if this is the case, knowing that it converges to $\hat{x}$, which is not in c_{00}, we show that $(c_{00}, \|\cdot\|_\infty)$ is not complete. We show that

$$\forall\, \varepsilon > 0,\ \exists\, N \in \mathbb{N} :\ \forall\, n, m > N, \quad \|x^{(n)} - x^{(m)}\|_\infty < \varepsilon.$$

Let $\varepsilon > 0$. For the sake of argument, assume $n > m$, so

$$x^{(n)} = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{m+1}, \frac{1}{m+2}, \ldots, \frac{1}{n+1}, 0, 0, \ldots, 0\right)$$
$$x^{(m)} = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{m+1}, 0, \ldots, 0, 0, 0, \ldots, 0\right).$$

Thus,

$$x^{(n)} - x^{(m)} = \left(0, 0, 0, \ldots, 0, \frac{1}{m+2}, \ldots, \frac{1}{n+1}, 0, 0, \ldots, 0\right),$$

which leads to

$$\|x^{(n)} - x^{(m)}\|_\infty = \sup_{k \in \mathbb{N}} |x_k^{(n)} - x_k^{(m)}| = \frac{1}{m+2}.$$

Let $N > 1/\varepsilon$. Then,

$$n > m > N \quad \Rightarrow \quad \|x^{(n)} - x^{(m)}\|_\infty = \frac{1}{m+2} < \frac{1}{N+2} < \frac{1}{N} < \varepsilon.$$

Therefore, $(x^{(j)})_{j \in \mathbb{N}}$ is a Cauchy sequence. However, it is not convergent in c_{00} as proved earlier, and this allows us to conclude that $(c_{00}, \|\cdot\|_\infty)$ is not a Banach space.

■

Solution 2.13 *(Exercise 2.13).*

(i) To see that $\ell^p(\mathbb{N}, \mathbb{F})$ is a vector space it suffices to show that it is closed under addition and scalar multiplication:

- Let $x, y \in \ell^p(\mathbb{N}, \mathbb{F})$. We have to prove that $x + y \in \ell^p(\mathbb{N}, \mathbb{F})$. We observe that for any $k \in \mathbb{N}$ we have

  $$|x_k + y_k|^p \leq (2 \max\{|x_k|, |y_k|\})^p = 2^p \max\{|x_k|^p, |y_k|^p\} \leq 2^p (|x_k|^p + |y_k|^p),$$

  so that, by definition of the norm (2.18),

  $$\|x + y\|_p^p = \sum_{k=0}^{\infty} |x_k + y_k|^p \leq \sum_{k=0}^{\infty} 2^p (|x_k|^p + |y_k|^p) = 2^p (\|x\|^p + \|y\|^p) < +\infty.$$

- for $\lambda \in \mathbb{F}$,

  $$|\lambda x_k| = |\lambda| |x_k| \quad \Rightarrow \quad \|\lambda x\|_p = \left(\sum_{k=0}^{\infty} |\lambda x_k|^p\right)^{\frac{1}{p}} = \left(\sum_{k=0}^{\infty} |\lambda|^p |x_k|^p\right)^{\frac{1}{p}} = |\lambda| \left(\sum_{k=0}^{\infty} |x_k|^p\right)^{\frac{1}{p}} = |\lambda| \|x\|_p < +\infty.$$

Now, we only need to observe that every sequence in $\ell^p(\mathbb{N}, \mathbb{F})$ is convergent to 0. This is a consequence of the fact that if a series $\sum_{k=0}^{\infty} |x_k|^p$ is convergent, the sequence of the terms $(|x_k|^p)_{k \in \mathbb{N}}$ must go to zero as $k \to \infty$. Indeed, let s be the sum of the series. For $k \in \mathbb{N}$, let s_k denote the k-th partial sum. We have $\lim_{k \to \infty} s_k = s$ and also $\lim_{k \to \infty} s_{k+1} = s$, so that

$$|x_{k+1}|^p = s_{k+1} - s_n \xrightarrow{k \to \infty} s - s = 0.$$

If $\lim_{k \to \infty} |x_k|^p = 0$, then also $\lim_{k \to \infty} x_k = 0$. In fact, for all $\varepsilon > 0$, there exists $p \in \mathbb{N}$ such that for all $k > p$ we have $|x_k|^p < \varepsilon^p$, and therefore also $|x_k| < \varepsilon$.

(ii) We now to show that $\ell^p(\mathbb{N},\mathbb{F})$ is a normed space, we need to verify that $\|\cdot\|_p$ is indeed a norm. We start from the homogeneity. Let $k \in \mathbb{F}$ and $x \in \ell^p(\mathbb{N},\mathbb{F})$. Then,

$$\|k\,x\|_p = \Big(\sum_{i=1}^{\infty} |k\,x_i|^p\Big)^{\frac{1}{p}} = \Big(\sum_{i=1}^{\infty} |k|^p\,|x_i|^p\Big)^{\frac{1}{p}} = \Big(|k|^p \sum_{i=1}^{\infty} |x_i|^p\Big)^{\frac{1}{p}} = |k|\Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{\frac{1}{p}} = |k|\,\|x\|_p.$$

By definition, the norm $\|x\|_p$ is non-negative, and $\|x\|_p = 0$ implies $x_k = 0$ for all $k \in \mathbb{N}$, which in turn implies $x = 0$. What remains to be proved is the triangle inequality. When $p = 1$, this is a consequence of the usual triangle inequality:

$$\|x + y\|_1 = \sum_{k=0}^{\infty} |x_k + y_k| \le \sum_{k=0}^{\infty} (|x_k| + |y_k|) = \|x\|_1 + \|y\|_1.$$

For $p \in (1,+\infty)$, we use the Young's inequality and then Hölder's and Minkowski's inequalities. In particular, if $p, q \in (1,+\infty)$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, we show that for two sequences $(x_k)_{k\in\mathbb{N}}$ and $(x_k)_{k\in\mathbb{N}}$ we have

$$\sum_{k=0}^{\infty} |x_i\,y_i| \le \|x\|_p\,\|y\|_q. \tag{11.7}$$

If x is the zero sequence, or if y is the zero sequence, then both sides of this inequality are zero, and therefore in these cases Hölder's inequality trivially holds. Suppose that both x and y are non-zero sequences, so that $\|x\|_p$ and $\|y\|_q$ are both non-zero. Let $k \in \mathbb{N}$. Applying Young's inequality with $a = \frac{|x_k|}{\|x\|_p}$ and $b = \frac{|y_k|}{\|y\|_q}$

$$\frac{|x_k\,y_k|}{\|x\|_p\,\|y\|_q} \le \frac{|x_k|^p}{p\,\|x\|_p^p} + \frac{|y_k|^q}{q\,\|y\|_q^q},$$

and adding over k gives

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{|x_k\,y_k|}{\|x\|_p\,\|y\|_q} &\le \sum_{k=0}^{\infty}\left(\frac{|x_k|^p}{p\,\|x\|_p^p} + \frac{|y_k|^q}{q\,\|y\|_q^q}\right) = \frac{1}{p\,\|x\|_p^p}\sum_{k=0}^{\infty}|x_k|^p + \frac{1}{q\,\|y\|_q^q}\sum_{k=0}^{\infty}|y_k|^q \\ &= \frac{1}{p\,\|x\|_p^p}\,\|x\|_p^p + \frac{1}{q\,\|y\|_q^q}\,\|y\|_q^q = \frac{1}{p} + \frac{1}{q} = 1.\end{aligned}$$

Multiplying both sides by $\|\boldsymbol{x}\|_p\,\|\boldsymbol{y}\|_q$ gives (11.7).

Now we derive the Minkowski's inequality. Choosing q so that p and q are conjugate exponents (i.e., $\frac{1}{p} + \frac{1}{q} = 1$), we find

$$\begin{aligned}\|x + y\|_p^p &= \sum_{k=0}^{\infty} |x_k + y_k|^p = \sum_{k=0}^{\infty} |x_k + y_k|\,|x_k + y_k|^{p-1} \\ &\le \sum_{k=0}^{\infty} (|x_k| + |y_k|)\,|x_k + y_k|^{p-1} = \sum_{k=0}^{\infty} |x_k|\,|x_k + y_k|^{p-1} + \sum_{k=0}^{\infty} |y_k|\,|x_k + y_k|^{p-1} \\ &\le \|x\|^p \Big(\sum_{k=0}^{\infty} |x_k + y_k|^{q\,(p-1)}\Big)^{\frac{1}{q}} + \|y\|^p \Big(\sum_{k=0}^{\infty} |x_k + y_k|^{q\,(p-1)}\Big)^{\frac{1}{q}} \\ &= \|x\|^p \Big(\sum_{k=0}^{\infty} |x_k + y_k|^p\Big)^{1-\frac{1}{p}} + \|y\|^p \Big(\sum_{k=0}^{\infty} |x_k + y_k|^p\Big)^{1-\frac{1}{p}} \\ &= \|x\|^p \frac{\|x + y\|_p^p}{\|x + y\|_p} + \|y\|^p \frac{\|x + y\|_p^p}{\|x + y\|_p}\end{aligned}$$

Diving both sides by $\|x+y\|_p^p$ and multiplying by $\|x+y\|_p$ gives the triangle inequality. Dividing by the first factor on the right we obtain the Minkowski's inequality

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

This shows that $\ell^p(\mathbb{N},\mathbb{F})$ is indeed a normed vector space.

Now, consider for example $p = 1/2$. Take

$$x = \left(\frac{1}{4}, 0, 0, \ldots\right) \quad \text{and} \quad y = \left(0, \frac{1}{4}, 0, \ldots\right).$$

We have

$$\|x+y\|_{1/2} = \left(\left|\frac{1}{4}\right|^{1/2} + \left|\frac{1}{4}\right|^{1/2}\right)^2 = 1$$

and

$$\|x\|_{1/2} + \|y\|_{1/2} = \left(\left|\frac{1}{4}\right|^{1/2}\right)^2 + \left(\left|\frac{1}{4}\right|^{1/2}\right)^2 = \frac{1}{2},$$

so that $\|x+y\|_{1/2} > \|x\|_{1/2} + \|y\|_{1/2}$.

(iii) To prove that $(\ell^p(\mathbb{N},\mathbb{F}), \|\cdot\|_p)$ is a Banach space, we take a Cauchy sequence $(x^{(j)})_{j\in\mathbb{N}}$ in $(\ell^p(\mathbb{N},\mathbb{F}), \|\cdot\|_p)$, and we show that it is convergent. First, we show that for every $k \in \mathbb{N}$ the sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ of elements of $\mathbb{F}$ is a Cauchy sequence, i.e., we show that

$$\forall\, \varepsilon > 0,\ \exists N \in \mathbb{N} :\ \forall\, n, m > N, \quad |x_k^{(n)} - x_k^{(m)}| < \varepsilon. \tag{11.8}$$

Let $\varepsilon > 0$. Since $(x^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence,

$$\exists\, p \in \mathbb{N} :\ \forall\, n, m > p, \quad \|x^{(n)} - x^{(m)}\|_p < \varepsilon. \tag{11.9}$$

We can take $N = p$. Indeed, for any $k \in \mathbb{N}$,

$$|x_k^{(n)} - x_k^{(m)}|^p \leq \sum_{i=0}^{\infty} |x_i^{(n)} - x_i^{(m)}|^p = \|x^{(n)} - x^{(m)}\|_p^p < \varepsilon^p,$$

which implies $|x_k^{(n)} - x_k^{(m)}| < \varepsilon$, as required. This shows that for every $k \in \mathbb{N}$ the sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ of elements of $\mathbb{F}$ is a Cauchy sequence. Since $\mathbb{F}$ is complete, for each $k \in \mathbb{N}$ the sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ has limit $\hat{x}_k$:

$$\lim_{j\to\infty} x_k^{(j)} = \hat{x}_k.$$

The following diagram clarifies what we have obtained:

$$\begin{array}{lcccccc}
x^{(0)} & = & (x_0^{(0)} & x_1^{(0)} & x_2^{(0)} & \ldots) \\
x^{(1)} & = & (x_0^{(1)} & x_1^{(1)} & x_2^{(1)} & \ldots) \\
x^{(2)} & = & (x_0^{(2)} & x_1^{(2)} & x_2^{(2)} & \ldots) \\
 & & \vdots & \vdots & \vdots & \ddots \\
 & & \downarrow & \downarrow & \downarrow & \\
 & & \hat{x}_0 & \hat{x}_1 & \hat{x}_2 & \ldots
\end{array}$$

Each "horizontal" sequence is a Cauchy sequence, and every vertical sequence converges to a certain limit. Hence, if we fix $k \in \mathbb{N}$, the sequence $(x_k^{(j)})_{j\in\mathbb{N}}$ converges to $\hat{x}_k$. Let $h \in \mathbb{N}$ be fixed. Let $n, m > N$. From (11.8) we obtain

$$\sum_{k=0}^{h} |x_k^{(n)} - x_k^{(m)}|^p \le \sum_{k=0}^{\infty} |x_k^{(n)} - x_k^{(m)}|^p = \|x^{(n)} - x^{(m)}\|_p^p < \varepsilon^p.$$

Since $\lim_{m\to\infty} \|x^{(m)} - \hat{x}\|_p = 0$,

$$\forall\, \delta > 0, \quad \exists\, p_1 \in \mathbb{N} : \quad \forall\, m > p_1, \quad \|x^{(m)} - \hat{x}\|_p < \delta.$$

It follows that for $n > N$

$$\|x^{(n)} - \hat{x}\|_p \le \|x^{(n)} - x^{(m)}\|_p + \|x^{(m)} - \hat{x}\|_p < \varepsilon + \delta$$

Since this is true for every $\delta > 0$, we can apply Exercise 1.27 to conclude that for all $n > N$ we have $\|x^{(n)} - \hat{x}\|_p \le \varepsilon$. Therefore, for $n > N$ we obtain

$$\sum_{k=0}^{h} |x_k^{(n)} - \hat{x}_k|^p < \|x^{(n)} - \hat{x}\|_p^p < \varepsilon^p. \tag{11.10}$$

This shows that $\lim_{n\to\infty} \|x^{(n)} - \hat{x}\|_p = 0$. Now, we only need to show that $\hat{x} \in \ell^p(\mathbb{N};\mathbb{F})$, i.e., that $\|\hat{x}\|_p$ is bounded. We can use Minkowski's inequality for the p-norm in $\mathbb{R}^h$ on $\hat{x}_k = \hat{x}_k + x_k^{(n)} - x_k^{(n)}$ and we use (11.10):

$$\begin{aligned}
\Big(\sum_{k=0}^{h} |\hat{x}_k|^p\Big)^{\frac{1}{p}} &\le \Big(\sum_{k=0}^{h} |x_k^{(n)} - \hat{x}_k|^p\Big)^{\frac{1}{p}} + \Big(\sum_{k=0}^{h} |x_k^{(n)}|^p\Big)^{\frac{1}{p}} \\
&< (\varepsilon^p)^{\frac{1}{p}} + \Big(\sum_{k=0}^{h} |x_k^{(n)}|^p\Big)^{\frac{1}{p}} \\
&= \varepsilon + \Big(\sum_{k=0}^{h} |x_k^{(n)}|^p\Big)^{\frac{1}{p}}.
\end{aligned}$$

We let $h \to \infty$ and we obtain

$$\|\hat{x}\|_p < \varepsilon + \|x^{(n)}\|_p.$$

Since $x^{(n)}$ is a Cauchy sequence, it is bounded (see first part of the proof of Theorem 2.2), and so $\|\hat{x}\|_p$ is bounded. This proves that $\hat{x} \in \ell^p(\mathbb{N};\mathbb{F})$. This concludes the proof.

■

Solution 2.14 *(Exercise 2.14).*

(i) We prove that $\|\cdot\|_\infty$ in $\ell^\infty(\mathbb{N},\mathbb{F})$ is a norm. It is easy to establish the positivity: if $x \in \ell^\infty(\mathbb{N},\mathbb{F})$ is such that $\|x\|_\infty = 0$, we have $\sup_{i\in\mathbb{N}} |x_i| = 0$, from which we immediately find that $|x_i| = 0$ for all $i \in \mathbb{N}$, and therefore $x_i = 0$ for all $i \in \mathbb{N}$. The homogeneity is equally simple. Given $x \in \ell^\infty(\mathbb{N},\mathbb{F})$ and $k \in \mathbb{F}$

$$\|k\,x\|_\infty = \sup_{i\in\mathbb{N}} |k\,x_i| = |k| \sup_{i\in\mathbb{N}} |x_i| = |k|\,\|x\|_\infty,$$

where we used the result in Exercise 2.7. We only need to prove the triangle inequality (subadditivity). Let $x, y \in \ell^\infty(\mathbb{N}, \mathbb{F})$. Consider the inequality $|x_i + y_i| \leq |x_i| + |y_i|$. Taking the supremum over all i, we find

$$\sup_{i\in\mathbb{N}} |x_i + y_i| \leq \sup_{i\in\mathbb{N}} |x_i| + \sup_{i\in\mathbb{N}} |y_i|,$$

which is exactly $\|x + y\|_\infty \leq \|x\|_\infty + \|y\|_\infty$. Therefore, $(\ell^\infty(\mathbb{N}, \mathbb{F}), \|\cdot\|_\infty)$ is a normed space.

(ii) Let $(x^{(j)})_{j\in\mathbb{N}}$ be a Cauchy sequence in $\ell^\infty(\mathbb{N}, \mathbb{F})$, we want to show that it has a limit $\hat{x}$ in $\ell^\infty(\mathbb{N}, \mathbb{F})$. Thus, we must first construct $\hat{x}$ and then show that

$$\lim_{j\to\infty} \|x^{(j)} - \hat{x}\|_\infty = 0. \tag{11.11}$$

Since $(x^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence in $\ell^\infty(\mathbb{N}, \mathbb{F})$,

$$\forall\, \varepsilon > 0,\ \exists N \in \mathbb{N} :\ \forall\, n, m > N, \quad \|x^{(n)} - x^{(m)}\|_\infty < \varepsilon,$$

which, by definition of the norm, is the same as

$$\forall\, \varepsilon > 0,\ \exists N \in \mathbb{N} :\ \forall\, n, m > N, \quad \sup_{k\in\mathbb{N}} |x_k^{(n)} - x_k^{(m)}| < \varepsilon,$$

from which we find immediately that $|x_k^{(n)} - x_k^{(m)}| < \varepsilon$ for all $k \in \mathbb{N}$ and all $n, m > N$. This says that $(x_k^{(j)})_{j\in\mathbb{N}}$ is a Cauchy sequence of elements of $\mathbb{F}$, and since $\mathbb{F}$ is complete, $(x_k^{(j)})_{j\in\mathbb{N}}$ is also convergent. Let us define

$$\hat{x}_k \stackrel{\text{def}}{=} \lim_{j\to\infty} x_k^{(j)}.$$

The sequence $\hat{x} = (\hat{x}_k)_{k\in\mathbb{N}}$ thus obtained is in $\ell^\infty(\mathbb{N}, \mathbb{F})$. To see this, we show that it is bounded. Since $x^{(j)} \in \ell^\infty(\mathbb{N}, \mathbb{F})$ for all $j \in \mathbb{N}$, the sequence $x^{(j)}$ is bounded for all $j \in \mathbb{N}$. Thus,

$$\forall\, j \in \mathbb{N},\ \exists\, M \geq 0 : \|x^{(j)}\|_\infty \leq M.$$

This implies that for all $j \in \mathbb{N}$

$$\sup_{k\in\mathbb{N}} |x_k^{(j)}| \leq M \quad \Rightarrow \forall\, k \in \mathbb{N},\ |x_k^{(j)}| \leq M.$$

Taking the limit as $j \to \infty$ gives

$$\forall\, k \in \mathbb{N},\ |\hat{x}_k| \leq M.$$

Thus, M is an upper bound for the set $\{|\hat{x}_k| \mid k \in \mathbb{N}\}$. But $\|\hat{x}\|_\infty$ is the least upper bound for the same set, and so $\|\hat{x}\|_\infty \leq M$.

We show (11.11). Let $\varepsilon > 0$. We know that

$$\exists N \in \mathbb{N} :\ \forall\, n, m > N, \quad \|x^{(n)} - x^{(m)}\|_\infty < \varepsilon.$$

Let $n = j$ and let us choose $m \geq N$. We find

$$\|x^{(j)} - \hat{x}\|_\infty \leq \|x^{(j)} - x^{(m)}\|_\infty + \|x^{(m)} - \hat{x}\|_\infty.$$

We can bound both terms in the right hand-side. Indeed, from definition of $\hat{x}_k$ we know that there exists $N_1 \in \mathbb{N}$ such that for all $j > N_1$ we have

$$\forall\, k \in \mathbb{N},\ |x_k^{(j)} - \hat{x}_k| < \frac{\varepsilon}{2} \quad \Rightarrow \quad \sup_{k\in\mathbb{N}} |x_k^{(j)} - \hat{x}_k| < \frac{\varepsilon}{2} \quad \Rightarrow \quad \|x^{(j)} - \hat{x}\|_\infty < \frac{\varepsilon}{2}.$$

Moreover, from the Cauchy property if $j, m > N$

$$\|x^{(j)} - x^{(m)}\|_\infty < \frac{\varepsilon}{2},$$

and so for $m > \max\{N, N_1\}$

$$\|x^{(j)} - \hat{x}\|_\infty \leq \|x^{(j)} - x^{(m)}\|_\infty + \|x^{(m)} - \hat{x}\|_\infty = \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

(iii) First, note that c_0 is a subspace of $\ell^\infty(\mathbb{N}, \mathbb{F})$. Indeed, a sequence $(x_{(k)})_{k\in\mathbb{N}}$ in c_0 converges to 0 as $k \to \infty$, and therefore it is bounded (see Theorem 1.32). Let $(x^{(j)})_{j\in\mathbb{N}}$ be a sequence of elements of c_0 which converges, as $j \to \infty$, to the sequence $\hat{x}$ in $\ell^\infty(\mathbb{N}, \mathbb{F})$. We must show that $\hat{x} \in c_0$, i.e., that $\hat{x}$ is infinitesimal:

$$\forall\, \varepsilon > 0,\ \exists\, p \in \mathbb{N} :\ \forall\, k \geq N,\ \|\hat{x}\|_\infty < \varepsilon.$$

Let $\varepsilon > 0$. Since $(x^{(j)})_{j\in\mathbb{N}}$ converges to $\hat{x}$ in $\ell^\infty(\mathbb{N}, \mathbb{F})$,

$$\exists\, N > 0 :\ \forall\, j \geq N,\ \ \|\boldsymbol{x}^{(j)} - \hat{x}\|_\infty < \frac{\varepsilon}{2}.$$

In particular,

$$\|x^{(N)} - \hat{x}\|_\infty < \frac{\varepsilon}{2}.$$

On the other hand, $x^{(N)} \in c_0$, and so $x_k^{(N)} \xrightarrow{k\to\infty} 0$. This implies that

$$\exists\, p > 0 :\ \forall\, k \geq p,\ \ |x_k^{(N)}| < \frac{\varepsilon}{2}.$$

It follows that

$$\forall\, k \geq p,\ \ |\hat{x}_k| \leq |x_k^{(N)} - \hat{x}_k| + |x_k^{(N)}| \leq \|x^{(N)} - \hat{x}\|_\infty + \|x^{(N)}\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(iv) Since $c_{00} \subseteq c_0$ and c_0 is closed, we have $\overline{c_{00}} \subseteq c_0$. We have to show that $\overline{c_{00}} \supseteq c_0$. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in c_0. We have to show that this sequence is also in $\overline{c_{00}}$, i.e., $(x_k)_{k\in\mathbb{N}}$ is an adherent point of c_{00} using the norm $\|\cdot\|_\infty$. From Definition 1.13, $\boldsymbol{x}$ is an adherent point of c_{00} if

$$\forall\, r > 0,\ \ \mathcal{B}_r(x, \ell^\infty(\mathbb{N}, \mathbb{F}), \|\cdot\|_\infty) \cap c_{00} \neq \varnothing.$$

Consider the sequence $(y_k^{(n)})_{k\in\mathbb{N}}$, where

$$y_k = \begin{cases} x_k & \text{if } k \leq n \\ 0 & \text{if } k > n. \end{cases}$$

Thus, in essence, $y^{(n)} = (x_0, x_1, \ldots, x_n, 0, 0, 0, \ldots)$. For every $n \in \mathbb{N}$, we have $y^{(n)} \in c_{00}$. It remains to show that, by taking a sufficiently large value of n, the sequence $(y_k^{(n)})_{k\in\mathbb{N}}$ is also in $\mathcal{B}_r(x, \ell^\infty(\mathbb{N}, \mathbb{F}), \|\cdot\|_\infty)$, i.e.,

$$\forall\, r > 0,\ \exists\, N \in \mathbb{N} :\ \forall n > N,\ \|y^{(n)} - x\|_\infty = \sup_{i>n} |x_i| < r.$$

Let $r > 0$. Since x is infinitesimal (it is an element of c_0), there exists $N \in \mathbb{N}$ such that for all $k > N$ we have $|x_k| < r$. Taking $n > N$ implies

$$\|y^{(n)} - x\|_\infty = \sup_{i>n} |x_i| < r.$$

■

Solution 2.15 *(Exercise 2.15).* Let $p \geq 1$ and let $x \in \ell^p(\mathbb{N};\mathbb{F})$, so that $\sum_{k\in\mathbb{N}} |x_k|^p < +\infty$. This implies that the sequence is infinitesimal, i.e., $|x_k|^p$ tends to 0 as $k \to \infty$ and therefore it is bounded. We conclude that $x \in \ell^\infty(\mathbb{N};\mathbb{F})$.

To see that the inclusion is strict, we only need an example of a sequence in $\ell^\infty(\mathbb{N};\mathbb{F})$ which is not in $\ell^p(\mathbb{N};\mathbb{F})$. We can take for example the sequence $(x_k)_{k\in\mathbb{N}}$ defined by $x_k = 1$ for all $k \in \mathbb{N}$. This sequence is bounded, but $x \notin \ell^p(\mathbb{N};\mathbb{F})$ for any $p \geq 1$. ■

Solution 2.16 *(Exercise 2.16).* Let $p_2 \neq +\infty$. If $p = p_1$ or $p = p_2$, the result is obvious. Let $p \in (p_1, p_2)$. There exists $t \in (0,1)$ such that $p = t\,p_1 + (1-t)\,p_2$. We apply Hölder's inequality with exponent $1/t$ and conjugate exponent q, where

$$\frac{1}{\frac{1}{t}} + \frac{1}{q} = 1 \quad \Rightarrow \quad q = \frac{1}{1-t}.$$

We obtain

$$\begin{aligned}
\sum_{k=0}^{\infty} |x_k|^p &= \sum_{k=0}^{\infty} |x_k|^{t\,p_1+(1-t)\,p_2} \\
&= \sum_{k=0}^{\infty} |x_k|^{t\,p_1} + |x_k|^{(1-t)\,p_2} \\
&\leq \left(\sum_{k=0}^{\infty} \left(|x_k|^{t\,p_1}\right)^{\frac{1}{t}}\right)^{\frac{1}{t}} \left(\sum_{k=0}^{\infty} \left(|x_k|^{(1-t)\,p_2}\right)^{\frac{1}{1-t}}\right)^{\frac{1}{1-t}} \\
&= \left(\sum_{k=0}^{\infty} |x_k|^{p_1}\right)^{t} \left(\sum_{k=0}^{\infty} |x_k|^{p_2}\right)^{1-t} = \|x\|_{p_1}^{t\,p_1}\, \|x\|_{p_2}^{(1-t)\,p_2} < +\infty.
\end{aligned}$$

Let us consider the case $p_2 = +\infty$. We find

$$\sum_{k=0}^{\infty} |x_k|^p = \sum_{k=0}^{\infty} |x_k|^p\, |x_k|^{p-p_1} \leq \left(\sum_{k=0}^{\infty} |x_k|^p\right) \|x\|_\infty^{p-p_1} < +\infty.$$

■

Solution 2.17 *(Exercise 2.17).* Let $x : \mathbb{N} \longrightarrow \mathbb{F}$ be a sequence in $\ell^p(\mathbb{N};\mathbb{F})$. We show that $x \in \ell^q(\mathbb{N};\mathbb{F})$ for every $q \in [p, +\infty]$. From Example 2.15, if $x \in \ell^p(\mathbb{N};\mathbb{F})$, then $x \in \ell^\infty(\mathbb{N};\mathbb{F})$. From Exercise 2.16, the sequence x belongs to $\ell^q(\mathbb{N};\mathbb{F})$ for every $q \in [p, +\infty]$.
We only need to show that it is always possible to find a sequence in $\ell^q(\mathbb{N};\mathbb{F})$ which is not in $\ell^p(\mathbb{N};\mathbb{F})$. Since $p < q$, we have $\frac{1}{q} < \frac{1}{p}$. Let $\alpha \in \left(\frac{1}{q}, \frac{1}{p}\right)$, and consider the sequence $(x_k)_{k\in\mathbb{N}^*}$ whose generic term is given by $x_k = 1/k^\alpha$. Then,

$$\sum_{k=1}^{\infty} \frac{1}{k^{\alpha\,p}} = +\infty \quad \text{and} \quad \sum_{k=1}^{\infty} \frac{1}{k^{\alpha\,q}} < +\infty,$$

because $\alpha\,p < 1$ and $\alpha\,q > 1$. Thus, $x \in \ell^q(\mathbb{N};\mathbb{F}) \setminus \ell^p(\mathbb{N};\mathbb{F})$.

■

Solution 2.18 *(Exercise 2.18).* Since T is bounded with respect to $(\boldsymbol{X}, \|\cdot\|_{X,1})$ and $(\boldsymbol{Y}, \|\cdot\|_{Y,1})$,

$$\exists\, M \geq 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|T(\boldsymbol{x})\|_{Y,1} \leq M \|\boldsymbol{x}\|_{X,1}.$$

We have to prove that

$$\exists\, N \geq 0 : \quad \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|T(\boldsymbol{x})\|_{Y,2} \leq M \|\boldsymbol{x}\|_{X,2}.$$

Since $\|\cdot\|_{X,1}$ and $\|\cdot\|_{X,2}$ are equivalent,

$$\exists\, c_1, c_2 > 0 : \ \forall\, \boldsymbol{x} \in \boldsymbol{X}, \quad \|\boldsymbol{x}\|_{X,1} \leq c_1 \|\boldsymbol{x}\|_{X,2} \quad \text{and} \quad \|\boldsymbol{x}\|_{X,2} \leq c_2 \|\boldsymbol{x}\|_{X,1}.$$

Likewise, since $\|\cdot\|_{Y,1}$ and $\|\cdot\|_{Y,2}$,

$$\exists\, d_1, d_2 > 0 : \ \forall\, \boldsymbol{y} \in \boldsymbol{Y}, \quad \|\boldsymbol{y}\|_{Y,1} \leq d_1 \|\boldsymbol{y}\|_{Y,2} \quad \text{and} \quad \|\boldsymbol{y}\|_{Y,2} \leq d_2 \|\boldsymbol{y}\|_{Y,1}.$$

It follows that

$$\|T(\boldsymbol{x})\|_{Y,2} \leq d_2 \|T(\boldsymbol{x})\|_{Y,1} \leq d_2\, M \|\boldsymbol{x}\|_{X,1} \leq d_2\, M\, c_1 \|\boldsymbol{x}\|_{X,2}.$$

Thus, the statement follows by taking $N = d_2\, M\, c_1$, which is non-negative because it is the product of three non-negative quantities. ■

Solution 2.19 *(Exercise 2.19).* First, we notice that T is a linear form: given $f, g \in C^1([0,1], \mathbb{R})$ and $\alpha, \beta \in \mathbb{R}$, using the linearity of the derivative

$$T(\alpha f + \beta g) = \frac{\mathrm{d}}{\mathrm{d}x}(\alpha f(x) + \beta g(x))\Big|_{x=0} = \alpha f'(x) + \beta g'(x)\Big|_{x=0} = \alpha f'(0) + \beta g'(0) = \alpha T(f) + \beta T(g).$$

From the Sequential Continuity Theorem, the linear form T is continuous if, for any sequence $(f_k)_{k\in\mathbb{N}}$ of $C^1([0,1], \mathbb{R})$, we have

$$f_k \xrightarrow{k\to\infty} f \qquad \Rightarrow \qquad T(f_k) \xrightarrow{k\to\infty} T(f),$$

or, using the given norms,

$$\|f_k - f\|_\infty \xrightarrow{k\to\infty} 0 \qquad \Rightarrow \qquad |T(f_k) - T(f)| \xrightarrow{k\to\infty} 0.$$

The first says that f is the uniform limit of f_k. The hint suggests to use $f_k(x) = \frac{\sin(k\pi x)}{\sqrt{k}}$, and therefore we need to find its uniform limit. It is easy to see that

$$\frac{\sin(k\pi x)}{\sqrt{k}} \xrightarrow{k\to\infty} 0$$

uniformly, since $\sup_{x\in[0,1]} \left|\frac{\sin(k\pi x)}{\sqrt{k}} - 0\right| = \frac{1}{\sqrt{k}}$ goes to zero as $k \to \infty$ (the value $\frac{1}{\sqrt{k}}$ is obtained by noting that for any $k \in \mathbb{N}^*$, the sine attains the maximum for $x = \frac{1}{2k}$, which is always in $[0,1]$). On the other hand,

$$\lim_{k\to\infty} |T(f_k) - T(f)| = \lim_{k\to\infty} |f_k'(0) - 0| = \lim_{k\to\infty} \left|\frac{k\pi}{\sqrt{k}} \cos(k\pi 0)\right| = \lim_{k\to\infty} |\sqrt{k}\pi| = +\infty.$$

■

Solution 2.20 *(Exercise 2.20).* First, we show that T is a linear transformation, i.e., that it preserves the addition and the scalar multiplication: given $f, g \in C([0,1], \mathbb{R})$, we find

$$T(f+g) = \int_0^1 (f+g)(x)\,\mathrm{d}x = \int_0^1 (f(x)+g(x))\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x + \int_0^1 g(x)\,\mathrm{d}x = T(f) + T(g).$$

Given $k \in \mathbb{R}$, we find

$$T(kf) = \int_0^1 (kf)(x)\,\mathrm{d}x = \int_0^1 kf(x)\,\mathrm{d}x = k\int_0^1 f(x)\,\mathrm{d}x = kT(f).$$

Thus, T is a linear transformation. We now show that T is continuous. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of $C([0,1],\mathbb{R})$ such that

$$\lim_{k\to\infty} \|f_k - f\|_\infty = 0,$$

which means that $(f_k)_{k\in\mathbb{N}}$ converges to f uniformly. It follows that

$$\begin{aligned}\lim_{k\to\infty} |T(f_k) - T(f)| &= \lim_{k\to\infty} \left| \int_0^1 f_k(x)\,\mathrm{d}x - \int_0^1 f(x)\,\mathrm{d}x \right| \\ &= \left| \lim_{k\to\infty} \int_0^1 f_k(x)\,\mathrm{d}x - \int_0^1 f(x)\,\mathrm{d}x \right| = \left| \int_0^1 \lim_{k\to\infty} (f_k(x) - f(x))\,\mathrm{d}x \right| = 0,\end{aligned}$$

where we have used the fact that uniform convergence allows the interchange of the limit with the integral, and that uniform convergence implies pointwise convergence, which in turn implies that for every $x \in [0,1]$ we have $\lim_{k\to\infty} f_k(x) = f(x)$. This proves that T is continuous. ■

Solution 2.21 *(Exercise 2.21).* Consider the linearity: let $f, g \in C([-2,2],\mathbb{R})$. Then

$$T(f+g) = (f+g)(0) + (f+g)(1) = f(0) + g(0) + f(1) + g(1) = T(f) + T(g).$$

Let $k \in \mathbb{R}$. Then

$$T(kf) = (kf)(0) + (kf)(1) = kf(0) + kf(1) = kT(f).$$

Therefore, it is linear. We now prove the continuity. Let $(f_k)_{k\in\mathbb{N}}$ be a sequence of $C([-2,2],\mathbb{R})$ such that

$$\lim_{k\to\infty} \|f_k - f\|_\infty = 0 \quad \Leftrightarrow \quad \lim_{k\to\infty} \max_{x\in[-2,2]} |f_k(x) - f(x)|$$

by the Weierstrass theorem. Recall that the convergence with respect to the norm $\|\cdot\|_\infty$ is equivalent to the uniform convergence, which in turns implies the pointwise convergence. Now, using the triangle inequality and the fact that the limit of the sum is the sum of the limits we find

$$\begin{aligned}0 \le \lim_{k\to\infty} |T(f_k) - T(f)| &= \lim_{k\to\infty} |f_k(0) + f_k(1) - f(0) - f(1)| \\ &\le \lim_{k\to\infty} (|f_k(0) - f(0)| + |f_k(1) - f(1)|) \\ &= \lim_{k\to\infty} |f_k(0) - f(0)| + \lim_{k\to\infty} |f_k(1) - f(1)| \xrightarrow{k\to\infty} 0.\end{aligned}$$

The result follows from the squeeze theorem and the pointwise convergence. ■

Solution 2.22 *(Exercise 2.22).* T is continuous from $(\boldsymbol{V}, \|\cdot\|_{V,1})$ to $(\boldsymbol{W}, \|\cdot\|_{W,1})$ implies

$$\|T(\boldsymbol{v}_k) - T(\boldsymbol{v})\|_{W,1} \leq M_1 \|\boldsymbol{v}_k - \boldsymbol{v}\|_{V,1}$$

We must show that T is continuous from $(\boldsymbol{V}, \|\cdot\|_{V,2})$ to $(\boldsymbol{W}, \|\cdot\|_{W,2})$, i.e., that there exists $M_2 > 0$ such that

$$\|T(\boldsymbol{v}_k) - T(\boldsymbol{v})\|_{W,2} \leq M_2 \|\boldsymbol{v}_k - \boldsymbol{v}\|_{V,2}.$$

On the other hand,

$$\|T(\boldsymbol{v}_k) - T(\boldsymbol{v})\|_{W,2} \leq c_1 \|T(\boldsymbol{v}_k) - T(\boldsymbol{v})\|_{W,1} \leq c_1 M_1 \|\boldsymbol{v}_k - \boldsymbol{v}\|_{V,1} \leq M_1 c_1 c_2 \|\boldsymbol{v}_k - \boldsymbol{v}\|_{V,2},$$

so that we can take $M_2 = M_1 c_1 c_2$ and we obtain the result. ■

Solution 2.23 *(Exercise 2.23).* First, we notice that

$$\lim_{k\to\infty} \frac{\sin(k\pi x)}{\sqrt{k}} = 0.$$

It follows that

$$\lim_{k\to\infty} \|\varphi(f_k) - \varphi(f)\| = \lim_{k\to\infty} \left\| \frac{k\pi}{\sqrt{k}} \cos(k\pi x) - 0 \right\| = \lim_{k\to\infty} \sqrt{k}\pi = +\infty.$$

This proves that the functional is not continuous.

■

Solution 2.24 *(Exercise 2.24).*

(i) Let $x, y \in \ell^\infty(\mathbb{N}, \mathbb{F})$ and let $\alpha, \beta \in \mathbb{F}$. Then,

$$\begin{aligned} T(\alpha x + \beta y) &= T(\alpha x_0 + \beta y_0, \alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots) \\ &= (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \alpha x_3 + \beta y_3, \ldots) \\ &= \alpha (x_1, x_2, x_3, x_4, \ldots) + \beta (y_1, y_2, y_3, y_4, \ldots). \end{aligned}$$

It follows that T is a linear operator. Now, we show that it is continuous. To this end, it is sufficient to show that it is bounded (see Theorem 2.17), i.e., that there exists $M \geq 0$ such that

$$\|T(x)\|_\infty \leq M \|x\|_\infty,$$

which can be re-written as

$$\sup_{k\geq 1} |x_k| \leq M \sup_{k\geq 0} |x_k|.$$

This inequality is satisfied for any $M \geq 1$, and so T is bounded. We conclude that T is a continuous linear operator.

(ii) Take $x, y \in S$. Therefore, there exist $a, b \in \ell^\infty(\mathbb{N}, \mathbb{F})$ such that $x = a - T(a)$ and $y = b - T(b)$. Let $\alpha, \beta \in \mathbb{F}$. Using the linearity of T we find immediately

$$\alpha x + \beta y = \alpha (a - T(a)) + \beta (b - T(b)) = \alpha a + \beta b - T(\alpha a + \beta b) \in S.$$

Therefore, S is a subspace of $\ell^\infty(\mathbb{N}, \mathbb{F})$.

(iii) First, we notice that $z \notin S$. In fact, there is no $x \in \ell^\infty(\mathbb{N}, \mathbb{F})$ such that $z = x - T(x)$. Suppose by contradiction that such x exists. Then, $x - T(x) = z$ can be written explicitly as

$$(x_0 - x_1, x_1 - x_2, x_2 - x_3, x_3 - x_4, \ldots) = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right),$$

which leads to

$$\begin{cases} x_1 = x_0 - 1 \\ x_2 = x_1 - \frac{1}{2} = x_0 - 1 - \frac{1}{2} \\ x_3 = x_2 - \frac{1}{3} = x_0 - 1 - \frac{1}{2} - \frac{1}{3} \\ \quad \vdots \\ x_k = x_0 - 1 - \frac{1}{2} - \frac{1}{3} - \ldots - \frac{1}{k} = x_0 - \displaystyle\sum_{i=0}^{k} \frac{1}{i+1} = x_0 - \sum_{i=1}^{k+1} \frac{1}{i} \\ \quad \vdots \end{cases}$$

Therefore, for any $x_0 \in \mathbb{F}$, the sequence $(x_k)_{k \in \mathbb{N}}$ is unbounded because the harmonic series is divergent. We conclude that $z \notin S$. Now, we show that, for every $j \in \mathbb{N}$, the sequence $(x_k^{(j)})_{k \in \mathbb{N}}$ defined by

$$x_k^{(j)} = \begin{cases} \frac{1}{k+1} & \text{if } k \leq j \\ 0 & \text{otherwise} \end{cases}$$

is in S. We have

$$\begin{aligned} x^{(0)} &= (1, 0, 0, \ldots) \\ x^{(1)} &= \left(1, \frac{1}{2}, 0, \ldots\right) \\ x^{(2)} &= \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right) \\ &\vdots \\ x^{(j)} &= \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{j+1}, 0, \ldots\right) \\ &\vdots \end{aligned}$$

These sequences are all in S.[2] We only need to show that

$$\lim_{j \to \infty} \|x^{(j)} - z\|_\infty = 0.$$

[2] You can repeat the previous argument:

$$\begin{cases} x_1^{(j)} = x_0^{(j)} - 1 \\ x_2^{(j)} = x_1^{(j)} - \frac{1}{2} = x_0^{(j)} - 1 - \frac{1}{2} \\ x_3^{(j)} = x_2^{(j)} - \frac{1}{3} = x_0^{(j)} - 1 - \frac{1}{2} - \frac{1}{3} \\ \quad \vdots \\ x_{j+1}^{(j)} - x_j = 0 - \frac{1}{j+1} \\ x_{j+2}^{(j)} - x_{j+1}^{(j)} = 0 \\ \quad \vdots \end{cases}$$

We obtain $x_{j+1}^{(j)} = x_{j+2}^{(j)} = x_{j+3}^{(j)} = \ldots$ which means that $(x_k^{(j)})_{k \in \mathbb{N}}$ is bounded, and therefore in $\ell^\infty(\mathbb{N}, \mathbb{F})$.

Noting that

$$x^{(j)} = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{j+1}, 0, 0, 0 \ldots\right)$$
$$z = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{j+1}, \frac{1}{j+2}, \frac{1}{j+3}, \frac{1}{j+4} \ldots\right)$$

the difference is

$$x^{(j)} - z = \left(0, 0, 0, \ldots, 0, \frac{1}{j+2}, \frac{1}{j+3}, \frac{1}{j+4} \ldots\right),$$

and so

$$\lim_{j\to\infty} \|x^{(j)} - z\|_\infty = \lim_{j\to\infty} \sup_{k\in\mathbb{N}} \frac{1}{j+2} = 0.$$

■

Solution 2.25 *(Exercise 2.25).* Using the definition of sesquilinear form we immediately find

$$\begin{aligned} \mathfrak{b}(\boldsymbol{x}+\boldsymbol{y}, \boldsymbol{z}+\boldsymbol{w}) &= \mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}+\boldsymbol{w}) + \mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}+\boldsymbol{w}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \mathfrak{b}(\boldsymbol{x}, \boldsymbol{w}) + \mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}) + \mathfrak{b}(\boldsymbol{y}, \boldsymbol{w}) \\ \mathfrak{b}(a\,\boldsymbol{x}, b\,\boldsymbol{y}) &= a\,\mathfrak{b}(\boldsymbol{x}, b\,\boldsymbol{y}) = a\,\bar{b}\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{y}). \end{aligned}$$

■

Solution 2.26 *(Exercise 2.26).* Let $f, g, h \in \boldsymbol{X}$ and $\alpha, \beta \in \mathbb{R}$. Then, using the linearity of the Riemann integral

$$\begin{aligned} \mathfrak{b}(\alpha\, f + \beta\, g, h) &= \int_0^1 (\alpha\, f(t) + \beta\, g(t))\, h(t)\, \mathrm{d}t = \int_0^1 (\alpha\, f(t)\, h(t) + \beta\, g(t)\, h(t))\, \mathrm{d}t \\ &= \alpha \int_0^1 f(t)\, h(t)\, \mathrm{d}t + \beta \int_0^1 g(t)\, h(t)\, \mathrm{d}t = \alpha\, \mathfrak{b}(f, h) + \beta\, \mathfrak{b}(g, h). \end{aligned}$$

Notice that all integrals are well defined because the product of continuous functions is a continuous function, and so Riemann integrability over compacts is guaranteed. We also have linearity in the second argument:

$$\begin{aligned} \mathfrak{b}(f, \alpha\, g + \beta\, h) &= \int_0^1 f(t)\, (\alpha\, g(t) + \beta\, h(t))\, \mathrm{d}t = \int_0^1 (\alpha\, f(t)\, g(t) + \beta\, f(t)\, h(t))\, \mathrm{d}t \\ &= \int_0^1 \alpha\, f(t)\, g(t)\, \mathrm{d}t + \int_0^1 \beta\, f(t)\, h(t)\, \mathrm{d}t = \alpha\, \mathfrak{b}(f, g) + \beta\, \mathfrak{b}(f, h). \end{aligned}$$

The corresponding quadratic form is $\mathfrak{q}(f) = \int_0^1 f^2(t)\, \mathrm{d}t$. ■

Solution 2.27 *(Exercise 2.27).* Let $f, g, h \in \boldsymbol{X}$ and $\alpha, \beta \in \mathbb{C}$. Then, using the linearity of the Riemann integral

$$\begin{aligned} \mathfrak{b}(\alpha\, f + \beta\, g, h) &= \int_0^1 \overline{h(t)}\, (\alpha\, f(t) + \beta\, g(t))\, \mathrm{d}t = \int_0^1 (\alpha\, \overline{h(t)}\, f(t) + \beta\, \overline{h(t)}\, g(t))\, \mathrm{d}t \\ &= \alpha \int_0^1 \overline{h(t)}\, f(t)\, \mathrm{d}t + \beta \int_0^1 \overline{h(t)}\, g(t)\, \mathrm{d}t = \alpha\, \mathfrak{b}(f, h) + \beta\, \mathfrak{b}(g, h). \end{aligned}$$

Notice that all integrals are well defined because the product of continuous functions is a continuous function, and so Riemann integrability over compacts is guaranteed. We also have linearity in the second argument:

$$\begin{aligned}\mathfrak{b}(f, \alpha\, g + \beta\, h) &= \int_0^1 \overline{(\alpha\, g(t) + \beta\, h(t))}\, f(t)\, \mathrm{d}t = \int_0^1 (\bar{\alpha}\, \overline{g(t)}\, f(t) + \bar{\beta}\, \overline{h(t)}\, f(t))\, \mathrm{d}t \\ &= \int_0^1 \bar{\alpha}\, \overline{g(t)}\, f(t)\, \mathrm{d}t + \bar{\beta} \int_0^1 \overline{h(t)}\, f(t)\, \mathrm{d}t = \bar{\alpha}\, \mathfrak{b}(f, g) + \bar{\beta}\, \mathfrak{b}(f, h).\end{aligned}$$

The corresponding quadratic form is $\mathfrak{q}(f) = \int_0^1 |f(t)|^2\, \mathrm{d}t$. ■

Solution 2.28 *(Exercise 2.28).* For all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^3$ and $\alpha, \beta \in \mathbb{F}$ we have

$$\begin{aligned}\mathfrak{b}(\alpha\, \boldsymbol{x} + \beta\, \boldsymbol{y}, \boldsymbol{z}) &= \mathfrak{b}((\alpha\, x_1 + \beta\, y_1, \alpha\, x_2 + \beta\, y_2, \alpha\, x_3 + \beta\, y_3), (z_1, z_2, z_3)) \\ &= (\alpha\, x_1 + \beta\, y_1)\, z_1 + (\alpha\, x_2 + \beta\, y_2)\, z_2 \\ &= \alpha\, (x_1\, z_1 + x_2\, z_2) + \beta\, (y_1\, z_1 + y_2\, z_2) = \alpha\, \mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \beta\, \mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}).\end{aligned}$$

We show that the corresponding quadratic form $\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x})$ is positive semidefinite. Indeed,

$$\mathfrak{q}(\boldsymbol{x}) = \mathfrak{q}((x_1, x_2, x_3)) = \mathfrak{b}((x_1, x_2, x_3), (x_1, x_2, x_3)) = x_1^2 + x_2^2.$$

Clearly, $\mathfrak{q}(\boldsymbol{x}) \geq 0$ for every $\boldsymbol{x} \in \mathbb{R}^3$. However, $\mathfrak{q}(\boldsymbol{x})$ can be zero even if $\boldsymbol{x}$ is not the zero vector: for example, consider $\boldsymbol{x} = (0, 0, -2)$. It follows that $\mathfrak{q}$ is not positive definite.

■

Solution 2.29 *(Exercise 2.29).* For all $\boldsymbol{z}, \boldsymbol{w}, \boldsymbol{u} \in \mathbb{C}^3$ and $\alpha, \beta \in \mathbb{F}$ we have

$$\begin{aligned}\mathfrak{b}(\alpha\, \boldsymbol{z} + \beta\, \boldsymbol{w}, \boldsymbol{u}) &= -\bar{u}_1\, (\alpha\, z_1 + \beta\, w_1) - \bar{u}_3\, (\alpha\, z_3 + \beta\, w_3) \\ &= \alpha\, (-\bar{u}_1\, z_1 - \bar{u}_3\, z_3) + \beta\, (-\bar{u}_1\, w_1 - \bar{u}_3\, w_3) \\ &= \alpha\, \mathfrak{b}(\boldsymbol{z}, \boldsymbol{u}) + \beta\, \mathfrak{b}(\boldsymbol{w}, \boldsymbol{u}) \\ \mathfrak{b}(\boldsymbol{z}, \alpha\, \boldsymbol{w} + \beta\, \boldsymbol{u}) &= -\overline{(\alpha\, w_1 + \beta\, u_1)}\, z_1 - \overline{(\alpha\, w_3 + \beta\, u_3)}\, z_3 \\ &= -\bar{\alpha}\, \bar{w}_1\, z_1 - \bar{\beta}\, \bar{u}_1\, z_1 - \bar{\alpha}\, \bar{w}_3\, z_3 - \bar{\beta}\, \bar{u}_3\, z_3 \\ &= \bar{\alpha}\, (-\bar{w}_1\, z_1 - \bar{w}_3\, z_3) + \bar{\beta}\, (-\bar{u}_1\, z_1 - \bar{u}_3\, z_3) \\ &= \bar{\alpha}\, \mathfrak{b}(\boldsymbol{z}, \boldsymbol{w}) + \bar{\beta}\, \mathfrak{b}(\boldsymbol{z}, \boldsymbol{u}).\end{aligned}$$

We show that the corresponding quadratic form $\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x})$ is negative semidefinite. Indeed,

$$\mathfrak{q}(\boldsymbol{z}) = \mathfrak{q}((z_1, z_2, z_3)) = \mathfrak{b}((z_1, z_2, z_3), (z_1, z_2, z_3)) = -\bar{z}_1\, z_1 - \bar{z}_3\, z_3 = -|z_1|^2 - |z_3|^2.$$

We have $\mathfrak{q}(\boldsymbol{z}) \leq 0$ for every $\boldsymbol{z} \in \mathbb{C}^3$. However, $\mathfrak{q}(\boldsymbol{z})$ can be zero even if $\boldsymbol{z}$ is not the zero vector: for example, consider $\boldsymbol{z} = (0, 1, 0)$. It follows that $\mathfrak{q}$ is not negative semi-definite. ■

Solution 2.30 *(Exercise 2.30).* We prove the linearity in the first argument, by taking $\boldsymbol{x} = (x_1, x_2), \boldsymbol{y} = (y_1, y_2)$ and $\boldsymbol{z} = (z_1, z_2)$, and $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned}\mathfrak{b}(\alpha\, \boldsymbol{x} + \beta\, \boldsymbol{y}, \boldsymbol{z}) &= \mathfrak{b}(\alpha\, (x_1, x_2) + \beta\, (y_1, y_2), (z_1, z_2)) \\ &= \mathfrak{b}((\alpha\, x_1 + \beta\, y_1, \alpha\, x_2 + \beta\, y_2), (z_1, z_2)) \\ &= (\alpha\, x_1 + \beta\, y_1)\, z_1 - (\alpha\, x_2 + \beta\, y_2)\, z_2 \\ &= \alpha\, (x_1\, z_1 - x_2\, z_2) + \beta\, (y_1\, z_1 - y_2\, z_2) = \alpha\, \mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \beta\, \mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}).\end{aligned}$$

The linearity in the second argument can be proved in the same way.

To see that the corresponding quadratic form is sign indefinite, we consider $\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x})$ where $\boldsymbol{x} = (x_1, x_2)$, so that

$$\mathfrak{q}(\boldsymbol{x}) = \mathfrak{q}((x_1, x_2)) = \mathfrak{b}((x_1, x_2), (x_1, x_2)) = x_1^2 - x_2^2.$$

Hence, if $\boldsymbol{x} = (1, 0)$ we have $\mathfrak{q}(\boldsymbol{x}) = 1$, and if $\boldsymbol{x} = (0, 1)$ we have $\mathfrak{q}(\boldsymbol{x}) = -1$. Thus, $\mathfrak{q}$ is an indefinite quadratic form. ■

Solution 2.31 *(Exercise 2.31).* We find immediately

$$\mathfrak{b}(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{n} \bar{v}_i \, u_i = \sum_{i=1}^{n} \overline{\bar{u}_i \, v_i} = \overline{\sum_{i=1}^{n} \bar{u}_i \, v_i} = \overline{\mathfrak{b}(\boldsymbol{v}, \boldsymbol{u})}.$$

■

Solution 2.32 *(Exercise 2.32).* We have

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

■

Solution 2.33 *(Exercise 2.33).* First, we show that if (2.30) is an inner product, it induces the norm $\|\cdot\|$, since

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{x} \rangle &= \frac{1}{4} \left(\|\boldsymbol{x} + \boldsymbol{x}\|^2 - \|\boldsymbol{x} - \boldsymbol{x}\|^2 + i \, \|\boldsymbol{x} + i \, \boldsymbol{x}\|^2 - i \, \|\boldsymbol{x} - i \, \boldsymbol{x}\|^2 \right) \\ &= \|\boldsymbol{x}\|^2 - \frac{i}{4} \, |1 - i|^2 \, \|\boldsymbol{x}\|^2 - \frac{i}{4} \, |1 - i|^2 \, \|\boldsymbol{x}\|^2) = \|\boldsymbol{x}\|^2, \end{aligned}$$

because $|1 + i| = |1 - i| = \sqrt{2}$. Now, we show that (2.30) is indeed an inner product. The positivity is simple: we have already established that $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \|\boldsymbol{x}\|^2$, and since $\|\cdot\|$ is a norm, we have $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geq 0$ for every $\boldsymbol{x} \in \boldsymbol{X}$, and $\langle \boldsymbol{x}, \boldsymbol{x} \rangle > 0$ if and only if $\boldsymbol{x} = 0$. We show the anti-symmetry: let $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{X}$. Using the homogeneity of the norm, we find

$$\|\boldsymbol{y} + i \, \boldsymbol{x}\| = \|i \, (y + i \, \boldsymbol{x})\| = \|i \, \boldsymbol{y} - \boldsymbol{x}\| \tag{11.12}$$

$$\|\boldsymbol{y} - i \, \boldsymbol{x}\| = \|i \, (y - i \, \boldsymbol{x})\| = \|i \, \boldsymbol{y} + \boldsymbol{x}\|. \tag{11.13}$$

Thus, recalling that $\|\boldsymbol{x}\| = \|-\boldsymbol{x}\|$ for every $\boldsymbol{x} \in \boldsymbol{X}$,

$$\begin{aligned} \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle} &= \frac{1}{4} \left(\|\boldsymbol{y} + \boldsymbol{x}\|^2 - \|\boldsymbol{y} - \boldsymbol{x}\|^2 - i \, \|\boldsymbol{y} + i \, \boldsymbol{x}\|^2 + i \, \|\boldsymbol{y} - i \, \boldsymbol{x}\|^2 \right) \\ &= \frac{1}{4} \left(\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2 - i \, \|i \, \boldsymbol{y} - \boldsymbol{x}\|^2 + i \, \|\boldsymbol{x} + i \, \boldsymbol{y}\|^2 \right) \\ &= \frac{1}{4} \left(\|\boldsymbol{x} + \boldsymbol{y}\|^2 - \|\boldsymbol{x} - \boldsymbol{y}\|^2 - i \, \|\boldsymbol{x} - i \, \boldsymbol{y}\|^2 + i \, \|\boldsymbol{x} + i \, \boldsymbol{y}\|^2 \right) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle. \end{aligned}$$

We follow the same idea of the proof of Theorem 2.21:

(a) $2 \left\langle \dfrac{\boldsymbol{x} + \boldsymbol{y}}{2}, z \right\rangle = \langle \boldsymbol{x}, z \rangle + \langle \boldsymbol{y}, z \rangle$;

(b) $\langle \boldsymbol{x}+\boldsymbol{y}, z\rangle = \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle$;

(c) $\forall\, q \in \mathbb{Q}, \quad \langle q\,\boldsymbol{x}, z\rangle = q\,\langle \boldsymbol{x}, z\rangle$;

(d) $\forall\, a \in \mathbb{R}, \quad \langle a\,\boldsymbol{x}, z\rangle = a\,\langle \boldsymbol{x}, z\rangle$.

We find

$$\begin{aligned}
2\left\langle \frac{\boldsymbol{x}+\boldsymbol{y}}{2}, z\right\rangle &= \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z\right\|^2 - \left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z\right\|^2 + i\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+i\,z\right\|^2 - i\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-i\,z\right\|^2\right)\\
&= \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right) - \frac{1}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right)\\
&\quad \frac{i}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}+i\,z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right) - \frac{i}{2}\left(\left\|\frac{\boldsymbol{x}+\boldsymbol{y}}{2}-i\,z\right\|^2 + \left\|\frac{\boldsymbol{x}-\boldsymbol{y}}{2}\right\|^2\right)\\
&= \frac{1}{4}\left(\|\boldsymbol{x}+z\|^2 + \|\boldsymbol{y}+z\|^2\right) - \frac{1}{4}\left(\|\boldsymbol{x}-z\|^2 + \|\boldsymbol{y}-z\|^2\right) + \frac{i}{4}\left(\|\boldsymbol{x}+i\,z\|^2 + \|\boldsymbol{y}+i\,z\|^2\right) - \frac{i}{4}\left(\|\boldsymbol{x}-i\,z\|^2 + \|\boldsymbol{y}-i\,z\|\right.\\
&= \frac{1}{4}\left(\|\boldsymbol{x}+z\|^2 - \|\boldsymbol{x}-z\|^2 + i\,\|\boldsymbol{x}+i\,z\|^2 - i\,\|\boldsymbol{x}-i\,z\|^2\right) + \frac{1}{4}\left(\|\boldsymbol{y}+z\|^2 - \|\boldsymbol{y}-z\|^2 + i\,\|\boldsymbol{y}+i\,z\|^2 - i\,\|\boldsymbol{y}-i\,z\|^2\right)\\
&= \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle,
\end{aligned}$$

Choosing $\boldsymbol{y} = \boldsymbol{0}$ in **(a)** gives, for all $\boldsymbol{x}, z \in \boldsymbol{X}$,

$$\begin{aligned}
2\left\langle \frac{1}{2}\,\boldsymbol{x}, z\right\rangle &= \langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{0}, z\rangle\\
&= \langle \boldsymbol{x}, z\rangle + \frac{1}{4}\left(\|z\|^2 - \|-z\|^2 + i\,\|i\,z\|^2 - i\,\|-i\,z\|^2\right) = \langle \boldsymbol{x}, z\rangle.
\end{aligned}$$

Re-writing this with $\boldsymbol{x}+\boldsymbol{y}$ in place of $\boldsymbol{x}$ gives

$$2\left\langle \frac{\boldsymbol{x}+\boldsymbol{y}}{2}, z\right\rangle \langle \boldsymbol{x}+\boldsymbol{y}, z\rangle.$$

Since we proved that the left hand-side is equal to $\langle \boldsymbol{x}, z\rangle + \langle \boldsymbol{y}, z\rangle$, we obtain **(b)**. Note that **(b)** with $\boldsymbol{x} = \boldsymbol{y}$ gives $2\,\langle \boldsymbol{x}, z\rangle = \langle 2\,\boldsymbol{x}, z\rangle$ for all $\boldsymbol{x}, z \in \boldsymbol{X}$. Using induction as in the proof of Theorem 2.21 yields

$$\forall\, \boldsymbol{x}, z \in \boldsymbol{X},\ \forall\, q \in \mathbb{Q}, \quad q\,\langle \boldsymbol{x}, z\rangle = \langle q\,\boldsymbol{x}, z\rangle.$$

Using (11.12-11.13) we also find

$$\begin{aligned}
\langle i\,\boldsymbol{x}, \boldsymbol{y}\rangle &= \frac{1}{4}\left(\|i\,\boldsymbol{x}+\boldsymbol{y}\|^2 - \|i\,\boldsymbol{x}-\boldsymbol{y}\|^2 + i\,\|i\,\boldsymbol{x}+i\,\boldsymbol{y}\|^2 - i\,\|i\,\boldsymbol{x}-\boldsymbol{i}\,\boldsymbol{y}\|^2\right)\\
&= \frac{1}{4}\left(\|\boldsymbol{x}-i\,\boldsymbol{y}\|^2 - \|\boldsymbol{x}+i\,\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-\boldsymbol{y}\|^2\right)\\
&= \frac{i}{4}\left(\|\boldsymbol{x}+\boldsymbol{y}\|^2 - \|\boldsymbol{x}-\boldsymbol{y}\|^2 - i\,\|\boldsymbol{x}-i\,\boldsymbol{y}\|^2 + i\,\|\boldsymbol{x}+i\,\boldsymbol{y}\|^2\right) = i\,\langle \boldsymbol{x}, \boldsymbol{y}\rangle.
\end{aligned}$$

Let $(a_k)_{k\in\mathbb{N}}$ be a sequence of complex numbers which converges to $\hat{a} \in \mathbb{C}$. Let $\boldsymbol{x}, z \in \boldsymbol{X}$. Using Theorem 2.20, the terms $a_k\,\boldsymbol{x}+z$, $a_k\,\boldsymbol{x}-z$, $a_k\,\boldsymbol{x}+i\,z$ and $a_k\,\boldsymbol{x}-i\,z$ converge, respectively, to $\hat{a}\,\boldsymbol{x}+z$, $\hat{a}\,\boldsymbol{x}-z$, $\hat{a}\,\boldsymbol{x}+i\,z$ and $\hat{a}\,\boldsymbol{x}-i\,z$. Thus,

$$\begin{aligned}
\lim_{k\to\infty}\langle a_k\,\boldsymbol{x}, z\rangle &= \frac{1}{4}\lim_{k\to\infty}\left(\|a_k\,\boldsymbol{x}+z\|^2 - \|a_k\,\boldsymbol{x}-z\|^2 + i\,\|a_k\,\boldsymbol{x}+i\,z\|^2 - i\,\|a_k\,\boldsymbol{x}-i\,z\|^2\right)\\
&\quad \frac{1}{4}\left(\|\hat{a}\,\boldsymbol{x}+z\|^2 - \|\hat{a}\,\boldsymbol{x}-z\|^2 + i\,\|\hat{a}\,\boldsymbol{x}+i\,z\|^2 - i\,\|\hat{a}\,\boldsymbol{x}-z\|^2\right) = \langle \hat{a}\,\boldsymbol{x}, z\rangle.
\end{aligned}$$

Finally, consider $w \in \mathbb{C}$, that we write as $w = \alpha + i\beta$, where $\alpha, \beta \in \mathbb{R}$. From the density of $\mathbb{Q}$ in $\mathbb{R}$, there exist two sequences of rational numbers $(q_k)_{k\in\mathbb{N}}$ and $(r_k)_{k\in\mathbb{N}}$ such that

$$\lim_{k\to\infty} q_k = \alpha \quad \text{and} \quad \lim_{k\to\infty} r_k = \beta.$$

It follows that

$$\begin{aligned}
\langle w\,\boldsymbol{x}, \boldsymbol{z}\rangle &= \langle(\alpha + i\beta)\,\boldsymbol{x}, \boldsymbol{y}\rangle = \langle\alpha\,\boldsymbol{x}, \boldsymbol{z}\rangle + \langle i\beta\,\boldsymbol{x}, \boldsymbol{z}\rangle \\
&= \langle\alpha\,\boldsymbol{x}, \boldsymbol{z}\rangle + i\,\langle\beta\,\boldsymbol{x}, \boldsymbol{z}\rangle \\
&= \left\langle \lim_{k\to\infty} q_k\,\boldsymbol{x}, \boldsymbol{z}\right\rangle + i\left\langle \lim_{k\to\infty} r_k\,\boldsymbol{x}, \boldsymbol{z}\right\rangle \\
&= \lim_{k\to\infty} \langle q_k\,\boldsymbol{x}, \boldsymbol{z}\rangle + i \lim_{k\to\infty} \langle r_k\,\boldsymbol{x}, \boldsymbol{z}\rangle \\
&= \lim_{k\to\infty} q_k\,\langle\boldsymbol{x}, \boldsymbol{z}\rangle + i \lim_{k\to\infty} r_k\,\langle\boldsymbol{x}, \boldsymbol{z}\rangle \\
&= \alpha\,\langle\boldsymbol{x}, \boldsymbol{z}\rangle + i\beta\,\langle\boldsymbol{x}, \boldsymbol{z}\rangle = w\,\langle\boldsymbol{x}, \boldsymbol{z}\rangle.
\end{aligned}$$

■

Solution 2.34 *(Exercise 2.34).* Denoting by an asterisk the entries that do not interest us, and recalling that given a complex number z we have $\bar{z}\,z = |z|^2$, we have

$$\begin{aligned}
\operatorname{trace}(A^*A) &= \operatorname{trace}\left(\begin{bmatrix} \bar{a}_{1,1} & \bar{a}_{2,1} & \dots & \bar{a}_{m,1} \\ \bar{a}_{1,2} & \bar{a}_{2,2} & \dots & \bar{a}_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{a}_{1,n} & \bar{a}_{2,n} & \dots & \bar{a}_{m,n} \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}\right) \\
&= \operatorname{trace}\begin{bmatrix} \sum_{i=1}^m \bar{a}_{i,1}\,a_{i,1} & * & \dots & * \\ * & \sum_{i=1}^m \bar{a}_{i,2}\,a_{i,2} & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \sum_{i=1}^m \bar{a}_{i,m}\,a_{i,m} \end{bmatrix} = \operatorname{trace}\begin{bmatrix} \sum_{i=1}^m |a_{i,1}|^2 & * & \dots & * \\ * & \sum_{i=1}^m |a_{i,2}|^2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \sum_{i=1}^m |a_{i,m}|^2 \end{bmatrix} \\
&= \sum_{j=1}^n \sum_{i=1}^m |a_{i,j}|^2.
\end{aligned}$$

The Hermitian product that induces the Frobenius norm is

$$\langle A, B\rangle = \operatorname{trace}(B^*A).$$

Using the same calculation above we find

$$\begin{aligned}
\operatorname{trace}(B^*A) &= \operatorname{trace}\left(\begin{bmatrix} \bar{b}_{1,1} & \bar{b}_{2,1} & \dots & \bar{b}_{m,1} \\ \bar{b}_{1,2} & \bar{b}_{2,2} & \dots & \bar{b}_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{b}_{1,n} & \bar{b}_{2,n} & \dots & \bar{b}_{m,n} \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{bmatrix}\right) \\
&= \operatorname{trace}\begin{bmatrix} \sum_{i=1}^m \bar{b}_{i,1}\, a_{i,1} & * & \dots & * \\ * & \sum_{i=1}^m \bar{b}_{i,2}\, a_{i,2} & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \sum_{i=1}^m \bar{b}_{i,m}\, a_{i,m} \end{bmatrix} \\
&= \sum_{j=1}^n \sum_{i=1}^m \bar{b}_{i,j}\, a_{i,j}
\end{aligned}$$

and it is easy to see that this is a Hermitian product. Recalling that the conjugate of the sum is the sum of the conjugates we prove the anti-symmetry:

$$\langle A, B\rangle = \operatorname{trace}(B^*A) = \sum_{j=1}^n \sum_{i=1}^m \bar{b}_{i,j}\, a_{i,j} = \sum_{j=1}^n \sum_{i=1}^m \overline{\bar{a}_{i,j}\, \bar{b}_{i,j}} = \overline{\sum_{j=1}^n \sum_{i=1}^m \bar{a}_{i,j}\, \bar{b}_{i,j}} = \overline{\operatorname{trace}(A^*B)} = \overline{\langle B, A\rangle}.$$

The linearity in the first argument comes from the properties of the matrix products: for $A, B, C \in \mathbb{C}^{m\times n}$

$$\langle A + B, C\rangle = \operatorname{trace}\left(C^*\,(A + B)\right) = \operatorname{trace}(C^*A) + \operatorname{trace}(C^*B) = \langle A, C\rangle + \langle B, C\rangle$$

and, given $k \in \mathbb{C}$

$$\langle k\,A, B\rangle = \operatorname{trace}\left(B^*(k\,A)\right) = \operatorname{trace}\left(k\,(B^*A)\right) = k\,\operatorname{trace}(B^*A) = k\,\langle A, B\rangle.$$

Finally, as we say above

$$\langle A, A\rangle = \operatorname{trace}(A^*A) = \sum_{j=1}^n \sum_{i=1}^m |a_{i,j}|^2,$$

which is ≥ 0, and if it is zero we have $a_{i,j} = 0$ for all $i = 1, \dots, m$ and $j = 1, \dots, n$, i.e., $A = 0$. We conclude that this is a Hermitian product. The Frobenius norm is induced by this inner product. The set $\mathbb{C}^{m\times n}$ is finite-dimensional (its dimension is $m \cdot n$), and therefore it is complete (every finite dimensional vector space is complete, i.e., Cauchy sequences are convergent). It follows that $(\mathbb{C}^{m\times n}, \langle\cdot,\cdot,\rangle)$, where the Hermitian product $\langle\cdot,\cdot,\rangle$ is the one defined above, is a Hilbert space. ■

Solution 2.35 *(Exercise 2.35).* This is a norm, because it is obtained by adding seminorms (linearly combining seminorms with positive coefficients gives a seminorm), and it is also positive, because $\|\boldsymbol{x}\| = 0$ implies $x_1 = x_2 = 0$, i.e., $\boldsymbol{x} = \boldsymbol{0}$. To find if it is induced by an inner product, we check the parallelogram law. Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$. We have

$$\begin{aligned}
\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 &= (x_1 + y_1)^2 + 2\,(x_2 + y_2)^2 + (x_1 - y_1)^2 + 2\,(x_2 - y_2)^2 \\
&= x_1^2 + 2\,x_1\,y_1 + y_1^2 + 2\,x_2^2 + 4\,x_2\,y_2 + 2\,y_2^2 + x_1^2 - 2\,x_1\,y_1 + y_1^2 + 2\,x_2^2 - 4\,x_2\,y_2 + 2\,y_2^2 \\
&= 2\,x_1^2 + 4\,x_2^2 + 2\,y_1^2 + 4\,y_2^2 = 2\,(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2).
\end{aligned}$$

It follows that this norm is induced by an inner product $\langle \cdot, \cdot \rangle$. Since $\boldsymbol{X}$ is finite dimensional, it is complete (i.e., every Cauchy sequence is convergent), and therefore $(\boldsymbol{X}, \langle \cdot, \cdot \rangle)$ is a Hilbert space. It is easy to see that given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^2$, the inner product is defined by

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1\, y_1 + 2\, x_2\, y_2.$$

Recall that the vector $\hat{\boldsymbol{y}}$ that minimizes $\|\boldsymbol{x} - \hat{\boldsymbol{y}}\|$ is $\hat{\boldsymbol{y}} = P_A(\boldsymbol{x})$, which can be written as

$$P_A(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{x}_1 \rangle\, \boldsymbol{x}_1 + \langle \boldsymbol{x}, \boldsymbol{x}_2 \rangle\, \boldsymbol{x}_2 + \ldots + \langle \boldsymbol{x}, \boldsymbol{x}_r \rangle\, \boldsymbol{x}_r,$$

where $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\}$ is an orthonormal basis of A. In this case, A is 1-dimensional, and we find an orthonormal basis if we divide the vector $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ by its norm:

$$\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| = \sqrt{1+2} = \sqrt{3}.$$

Hence,

$$\begin{aligned} P_A(\boldsymbol{x}) &= \left\langle \boldsymbol{x}, \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{3} \left\langle \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{3} \left(2 + 2\,(-2 \cdot 1)\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -\frac{2}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \end{aligned}$$

To check that this vector $\hat{\boldsymbol{y}}$ is correct, we can show that $\boldsymbol{x} - \hat{\boldsymbol{y}}$ is orthogonal to any vector of A. Representing the generic vector of A as $\left[\begin{smallmatrix} a \\ a \end{smallmatrix}\right]$, we have indeed,

$$\left\langle \begin{bmatrix} \frac{8}{3} \\ -\frac{4}{3} \end{bmatrix}, \begin{bmatrix} a \\ a \end{bmatrix} \right\rangle = \frac{8}{3}\, a + 2 \left(-\frac{4}{3}\, a \right) = 0,$$

as expected. ■

Solution 2.36 *(Exercise 2.36).* We find a counter-example of the parallelogram law. In particular, take $f(x) = 1$ and $g(x) = 2\,x$ for all $x \in [0, 1]$. The parallelogram law in this case reads as

$$\left(\int_0^1 |f + g| \right)^2 + \left(\int_0^1 |f - g| \right)^2 = 2 \left[\left(\int_0^1 |f| \right)^2 + \left(\int_0^1 |g| \right)^2 \right].$$

We find

$$\begin{aligned} \int_0^1 |f + g| &= \int_0^1 |1 + 2\,x|\, \mathrm{d}x = \int_0^1 (1 + 2\,x)\, \mathrm{d}x = [x + x^2]_0^1 = 2 \\ \int_0^1 |f - g| &= \int_0^1 |1 - 2\,x|\, \mathrm{d}x = \int_0^{\frac{1}{2}} (1 - 2\,x)\, \mathrm{d}x + \int_{\frac{1}{2}}^1 (2\,x - 1)\, \mathrm{d}x \\ &= [x - x^2]_0^{1/2} + [x^2 - x]_{1/2}^1 = \frac{1}{2} - \frac{1}{4} + 1 - 1 - \frac{1}{4} + \frac{1}{2} = \frac{1}{2} \\ \int_0^1 |f| &= \int_0^1 \mathrm{d}x = 1 \\ \int_0^1 |g| &= \int_0^1 |2\,x|\, \mathrm{d}x = \int_0^1 2\,x\, \mathrm{d}x = [x^2]_0^1 = 1. \end{aligned}$$

The left hand-side of the parallelogram law is

$$\left(\int_0^1 |f+g|\right)^2 + \left(\int_0^1 |f-g|\right)^2 = 4 + \frac{1}{4} = \frac{17}{4},$$

and the right hand-side is

$$2\left[\left(\int_0^1 |f|\right)^2 + \left(\int_0^1 |g|\right)^2\right] = 2\,(1+1) = 4.$$

Therefore, the parallelogram law is not satisfied, and therefore the norm $\|\cdot\|_1$ is not induced by an inner product.

■

Solution 2.37 *(Exercise 2.37).* We find a counter-example of the parallelogram law. In particular, take $f(x) = x$ and $g(x) = 1 - x$ for all $x \in [0,1]$. The parallelogram law in this case reads as

$$\left(\max_{x\in[0,1]} |f(x)+g(x)|\right)^2 + \left(\max_{x\in[0,1]} |f(x)-g(x)|\right)^2 = 2\left[\left(\max_{x\in[0,1]} |f(x)|\right)^2 + \left(\max_{x\in[0,1]} |g(x)|\right)^2\right].$$

We have

$$\begin{aligned}
\max_{x\in[0,1]} |f(x)+g(x)| &= \max_{x\in[0,1]} |x+1-x| = \max_{x\in[0,1]} 1 = 1\\
\max_{x\in[0,1]} |f(x)-g(x)| &= \max_{x\in[0,1]} |x-1+x| = \max_{x\in[0,1]} 2\,x-1 = 1\\
\max_{x\in[0,1]} |f(x)| &= \max_{x\in[0,1]} |x| = 1\\
\max_{x\in[0,1]} |g(x)| &= \max_{x\in[0,1]} |1-x| = 1.
\end{aligned}$$

The left hand-side of the parallelogram law is $1 + 1 = 2$, while the right hand-side is $2\,(1+1) = 4$. This is enough to conclude that the norm $\|\cdot\|_\infty$ is not induced by an inner product. ■

Solution 2.38 *(Exercise 2.38).* We have to show that for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for all $\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}$, if $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$, then $|\langle f(\boldsymbol{x}), g(\boldsymbol{x})\rangle - \langle \boldsymbol{\ell}_1, \boldsymbol{\ell}_2\rangle| < \varepsilon$.

Let $\varepsilon > 0$. First, since f converges to $\boldsymbol{\ell}_1 \in \boldsymbol{Y}$, there is a neighborhood of $\boldsymbol{x}_0$ where f is bounded (Lemma 1 in the lecture notes on normed spaces), i.e.,

$$\exists\,\delta_1 > 0 :\ \exists\, M > 0 :\quad \forall\,\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta_1 \ \Rightarrow\ \|f(\boldsymbol{x})\|_Y < M.$$

Moreover, since f converges to $\boldsymbol{\ell}_1$,

$$\exists\,\delta_2 > 0 :\quad \forall\,\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta_2 \ \Rightarrow\ \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y < \frac{\varepsilon}{2\,\|\boldsymbol{\ell}_2\|_Y}.$$

Since g converges to $\boldsymbol{\ell}_2$,

$$\exists\,\delta_3 > 0 :\quad \forall\,\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta_3 \ \Rightarrow\ \|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y < \frac{\varepsilon}{2\,M}.$$

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Then, for all $\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}$ satisfying $\|\boldsymbol{x} - \boldsymbol{x}_0\|_X < \delta$, we have

$$\begin{aligned}
|\langle f(\boldsymbol{x}), g(\boldsymbol{x})\rangle - \langle \boldsymbol{\ell}_1, \boldsymbol{\ell}_2\rangle| &= |\langle f(\boldsymbol{x}), g(\boldsymbol{x})\rangle - \langle f(\boldsymbol{x}), \boldsymbol{\ell}_2\rangle + \langle f(\boldsymbol{x}), \boldsymbol{\ell}_2\rangle - \langle \boldsymbol{\ell}_1, \boldsymbol{\ell}_2\rangle| \\
&\leq |\langle f(\boldsymbol{x}), g(\boldsymbol{x})\rangle - \langle f(\boldsymbol{x}), \boldsymbol{\ell}_2\rangle| + |\langle f(\boldsymbol{x}), \boldsymbol{\ell}_2\rangle - \langle \boldsymbol{\ell}_1, \boldsymbol{\ell}_2\rangle| \\
&= |\langle f(\boldsymbol{x}), g(\boldsymbol{x}) - \boldsymbol{\ell}_2\rangle| + |\langle f(\boldsymbol{x}) - \boldsymbol{\ell}_1, \boldsymbol{\ell}_2\rangle| \\
&\leq \|f(\boldsymbol{x})\|_Y \, \|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y + \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y \, \|\boldsymbol{\ell}_2\|_Y \\
&\leq M \, \|g(\boldsymbol{x}) - \boldsymbol{\ell}_2\|_Y + \|f(\boldsymbol{x}) - \boldsymbol{\ell}_1\|_Y \, \|\boldsymbol{\ell}_2\|_Y \\
&\leq M \, \frac{\varepsilon}{2\,M} + \frac{\varepsilon}{2\,\|\boldsymbol{\ell}_2\|_Y} \, \|\boldsymbol{\ell}_2\|_Y = \varepsilon,
\end{aligned}$$

where we used the triangle inequality for the absolute value, the linearity property of the inner product and the Cauchy-Schwarz inequality. ■

Solution 2.39 *(Exercise 2.39).* We prove the linearity in the first argument, by taking $\boldsymbol{x} = (x_1, x_2), \boldsymbol{y} = (y_1, y_2)$ and $\boldsymbol{z} = (z_1, z_2)$, and $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned}
\mathfrak{b}(\alpha\,\boldsymbol{x} + \beta\,\boldsymbol{y}, \boldsymbol{z}) &= \mathfrak{b}(\alpha\,(x_1, x_2) + \beta\,(y_1, y_2), (z_1, z_2)) \\
&= \mathfrak{b}((\alpha\,x_1 + \beta\,y_1, \alpha\,x_2 + \beta\,y_2), (z_1, z_2)) \\
&= (\alpha\,x_1 + \beta\,y_1)\,z_1 - (\alpha\,x_2 + \beta\,y_2)\,z_2 \\
&= \alpha\,(x_1\,z_1 - x_2\,z_2) + \beta\,(y_1\,z_1 - y_2\,z_2) = \alpha\,\mathfrak{b}(\boldsymbol{x}, \boldsymbol{z}) + \beta\,\mathfrak{b}(\boldsymbol{y}, \boldsymbol{z}).
\end{aligned}$$

The linearity in the second argument can be proved in the same way.

To see that the corresponding quadratic form is sign indefinite, we consider $\mathfrak{q}(\boldsymbol{x}) = \mathfrak{b}(\boldsymbol{x}, \boldsymbol{x})$ where $\boldsymbol{x} = (x_1, x_2)$, so that

$$\mathfrak{q}(\boldsymbol{x}) = \mathfrak{q}((x_1, x_2)) = \mathfrak{b}((x_1, x_2), (x_1, x_2)) = x_1^2 - x_2^2.$$

Hence, if $\boldsymbol{x} = (1, 0)$ we have $\mathfrak{q}(\boldsymbol{x}) = 1$, and if $\boldsymbol{x} = (0, 1)$ we have $\mathfrak{q}(\boldsymbol{x}) = -1$. Thus, $\mathfrak{q}$ is an indefinite quadratic form. ■

Solution 2.40 *(Exercise 2.40).* We check the parallelogram law:

$$\begin{aligned}
&\left\| \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix} + \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\|^2 + \left\| \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix} - \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\|^2 \\
&\quad = |z_1 + w_1|^2 + 5\,|z_2 + w_2|^2 + |z_3 + w_3|^2 + |z_1 - w_1|^2 + 5\,|z_2 - w_2|^2 + |z_3 - w_3|^2.
\end{aligned}$$

On the other hand, $(\mathbb{C}, |\cdot|)$ is a Hilbert space, because the modulus is the norm induced by the inner product $\langle z, w\rangle = \bar{w}\,z$ (since $\langle z, z\rangle = \bar{z}\,z = |z|^2$). It follows that for $(\mathbb{C}, |\cdot|)$ the parallelogram law holds true, so that

$$|z_1 + w_1|^2 + |z_1 - w_1|^2 = 2\,(|z_1|^2 + |w_1|^2)$$

and

$$|z_2 + w_2|^2 + |z_2 - w_2|^2 = 2\,(|z_2|^2 + |w_2|^2).$$

Using these identities gives

$$\begin{aligned}
&\left\| \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix} + \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\|^2 + \left\| \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix} - \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\|^2 \\
&\quad = 2\,|z_1|^2 + 2\,|w_1|^2 + 10\,|z_2|^2 + 10\,|w_2|^2 + 2\,|z_3|^2 + 2\,|w_3|^2 \\
&\quad = 2\left\| \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix} \right\|^2 + 2\left\| \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\|^2.
\end{aligned}$$

Since the vector space $\boldsymbol{X}$ is finite dimensional, it is complete. Since the norm is induced by an inner product, $\boldsymbol{X}$ with this inner product is a Hilbert space. Notice that the inner product is

$$\left\langle \begin{bmatrix} z_1 & z_2 \\ 0 & z_3 \end{bmatrix}, \begin{bmatrix} w_1 & w_2 \\ 0 & w_3 \end{bmatrix} \right\rangle = \bar{w}_1 z_1 + 5\,\bar{w}_2 z_2 + \bar{w}_3 z_3.$$

■

Solution 2.41 *(Exercise 2.41).*

- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☒ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is convergent, then it is a Cauchy sequence;
- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and it is also convergent, then $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space;
- ☐ If $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and it is also convergent, then $(\boldsymbol{X}, \|\cdot\|)$ is a Hilbert space;
- ☒ If $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space and if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☒ If $(\boldsymbol{X}, \|\cdot\|)$ is a Hilbert space and if $(\boldsymbol{x}_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, it is convergent;
- ☒ If $\boldsymbol{X}$ is finite dimensional, $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space;
- ☐ If $\boldsymbol{X}$ is finite dimensional, $(\boldsymbol{X}, \|\cdot\|)$ is a Hilbert space;
- ☐ If the norm is induced by an inner product, $(\boldsymbol{X}, \|\cdot\|)$ is a Hilbert space;
- ☐ If the norm is induced by an inner product, $(\boldsymbol{X}, \|\cdot\|)$ is a Banach space.

■

Solution 2.42 *(Exercise 2.42).* Consider first the case $p \in [1, +\infty)$. The sequences $(x_k)_{k\in\mathbb{N}}$ and $(y_k)_{k\in\mathbb{N}}$ defined by

$$x_k = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k \neq 0 \end{cases} \qquad \text{and} \qquad y_k = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{if } k \neq 1 \end{cases}$$

are in $\ell^p(\mathbb{N};\mathbb{C})$ for every $p \in \mathbb{N}$. Thus

$$\|x\|_p = \|y\|_p = \Big(\sum_{k=0}^{\infty} |x_k|^p\Big)^{\frac{1}{p}} = (1^p)^{\frac{1}{p}} = 1$$

$$\|x+y\|_p = \|x-y\|_p^2 = \Big(\sum_{k=0}^{\infty} |x_k + y_k|^p\Big)^{\frac{1}{p}} = (1^p + 1^p)^{\frac{1}{p}} = 2^{\frac{1}{p}}.$$

The parallelogram law $\|x+y\|_p^2 + \|x-y\|_p^2 = 2\,(\|x\|_p^2 + \|y\|_p^2)$ reads as

$$2^{\frac{2}{p}} + 2^{\frac{2}{p}} = 2\,(1+1) \quad \Leftrightarrow \quad 2^{\frac{2}{p}} = 2 \quad \Leftrightarrow \quad \frac{2}{p} = 1$$

and it is satisfied only for $p = 2$. This means that, unless $p = 2$, the normed space $\ell^p(\mathbb{N};\mathbb{C})$ is not a Hilbert space because the norm $\|\cdot\|_p$ is not induced by an inner product. This result holds for $p = +\infty$, as well. In fact, in this case the parallelogram law $\|x+y\|_\infty^2 + \|x-y\|_\infty^2 = 2\,(\|x\|_\infty^2 + \|y\|_\infty^2)$ reads as

$$1 + 1 = 2\,(1+1),$$

and it is not satisfied. To show that $\ell^2(\mathbb{N};\mathbb{C})$ is a Hilbert space, we introduce the inner product

$$\langle x, y\rangle = \sum_{k=0}^{\infty} \bar{y}_k\, x_k.$$

First, we have to make sure that this product is well defined for $x, y \in \ell^2(\mathbb{N};\mathbb{C})$. For a given $k \in \mathbb{N}$,

$$\frac{1}{2}\left(|x_k| - |y_k|\right)^2 \geq 0 \quad \Rightarrow \quad \frac{|x_k|^2}{2} + \frac{|y_k|^2}{2} \geq |x_k| \cdot |y_k|.$$

Adding over k gives

$$\sum_{k=0}^{\infty} |x_k| \cdot |y_k| \leq \frac{1}{2} \sum_{k=0}^{\infty} |x_k|^2 + \frac{1}{2} \sum_{k=0}^{\infty} |y_k|^2.$$

We find that if $x, y \in \ell^2(\mathbb{N};\mathbb{C})$ the right hand-side is finite, and thus $\displaystyle\sum_{k=0}^{\infty} |x_k| \cdot |y_k|$ is convergent, which implies that $\sum_{k=0}^{\infty} \bar{y}_k\, x_k$ is absolutely convergent. To see the inner product defined above is indeed an inner product, we take $x, y, z \in \ell^2(\mathbb{N};\mathbb{C})$ and $\alpha, \beta \in \mathbb{F}$:

- Anti-symmetry:

$$\overline{\langle y, x\rangle} = \overline{\sum_{k=0}^{\infty} \bar{x}_k\, y_k} = \sum_{k=0}^{\infty} \bar{y}_k\, x_k = \langle x, y\rangle;$$

- Linearity in the first argument:

$$\langle \alpha\, x + \beta\, y, z\rangle = \sum_{k=0}^{\infty} \bar{z}_k\, (\alpha\, x_k + \beta\, y_k) = \alpha \sum_{k=0}^{\infty} \bar{z}_k\, x_k + \beta \sum_{k=0}^{\infty} \bar{z}_k\, y_k \alpha\, \langle x, z\rangle + \beta\, \langle y, z\rangle.$$

- Positive-definiteness: $\displaystyle\sum_{k=0}^{\infty} |x_i|^2 \geq 0$, and $\displaystyle\sum_{k=0}^{\infty} |x_i|^2 = 0$ implies that x is the zero sequence.

■

Solution 2.43 *(Exercise 2.43).* Consider the mapping

$$\begin{aligned} P : \boldsymbol{A} &\longrightarrow \boldsymbol{H} \\ \boldsymbol{x} &\mapsto P_{\boldsymbol{A}}(\boldsymbol{x}). \end{aligned}$$

Let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{H}$ and $\alpha, \beta \in \mathbb{F}$. We show that $P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2) = \alpha\, P_{\boldsymbol{A}}(\boldsymbol{x}_1) + \beta\, P_{\boldsymbol{A}}(\boldsymbol{x}_2)$. From Theorem 2.24, we have $\langle \boldsymbol{x}_1 - P_{\boldsymbol{A}}(\boldsymbol{x}_1), \boldsymbol{u}\rangle = \langle \boldsymbol{x}_2 - P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle = 0$ and $\langle \alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2 - P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle = 0$ for any $\boldsymbol{u} \in \boldsymbol{A}$, so that

$$\begin{aligned} 0 &= \langle \alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2 - P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle \\ &= \langle \alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2, \boldsymbol{u}\rangle - \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle \\ &= \alpha\, \langle \boldsymbol{x}_1, \boldsymbol{u}\rangle + \beta\, \langle \boldsymbol{x}_2, \boldsymbol{u}\rangle - \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle \\ &= \alpha\, \cancel{\langle \boldsymbol{x}_1 - P_{\boldsymbol{A}}(\boldsymbol{x}_1), \boldsymbol{u}\rangle} + \alpha\, \langle P_{\boldsymbol{A}}(\boldsymbol{x}_1), \boldsymbol{u}\rangle + \beta\, \cancel{\langle \boldsymbol{x}_2 - P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle} + \beta\, \langle P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle - \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle \\ &= \alpha\, \langle P_{\boldsymbol{A}}(\boldsymbol{x}_1), \boldsymbol{u}\rangle + \beta\, \langle P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle - \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle \\ &= \langle \alpha\, P_{\boldsymbol{A}}(\boldsymbol{x}_1) + \beta\, P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle - \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle, \end{aligned}$$

so that

$$\langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2), \boldsymbol{u}\rangle = \langle \alpha\, P_{\boldsymbol{A}}(\boldsymbol{x}_1) + \beta\, P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle.$$

Hence,

$$\forall\, \boldsymbol{u} \in \boldsymbol{A}, \qquad \langle P_{\boldsymbol{A}}(\alpha\, \boldsymbol{x}_1 + \beta\, \boldsymbol{x}_2) - \alpha\, P_{\boldsymbol{A}}(\boldsymbol{x}_1) - \beta\, P_{\boldsymbol{A}}(\boldsymbol{x}_2), \boldsymbol{u}\rangle = 0,$$

so that the vector $P_{\boldsymbol{A}}(\alpha\,\boldsymbol{x}_1+\beta\,\boldsymbol{x}_2)-\alpha\,P_{\boldsymbol{A}}(\boldsymbol{x}_1)-\beta\,P_{\boldsymbol{A}}(\boldsymbol{x}_2)$ (which is a vector of $\boldsymbol{A}$) is orthogonal to every vector of $\boldsymbol{A}$, and therefore it is also orthogonal to itself. Since $\langle \boldsymbol{v},\boldsymbol{v}\rangle = 0$ if and only if $\boldsymbol{v}=0$, we conclude that $P_{\boldsymbol{A}}(\alpha\,\boldsymbol{x}_1+\beta\,\boldsymbol{x}_2)=\alpha\,P_{\boldsymbol{A}}(\boldsymbol{x}_1)+\beta\,P_{\boldsymbol{A}}(\boldsymbol{x}_2)$. ■

Solution 2.44 *(Exercise 2.44).* If $n=2$, we have immediately

$$\begin{aligned}\|\boldsymbol{x}_1+\boldsymbol{x}_2\|^2 &= \langle \boldsymbol{x}_1,\boldsymbol{x}_1\rangle+\langle \boldsymbol{x}_1,\boldsymbol{x}_2\rangle+\langle \boldsymbol{x}_2,\boldsymbol{x}_1\rangle+\langle \boldsymbol{x}_2,\boldsymbol{x}_2\rangle\\ &= \|\boldsymbol{x}_1\|^2+0+0+\|\boldsymbol{x}_2\|^2.\end{aligned}$$

We assume that the claim is true for n pairwise orthogonal vectors $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$, and we prove it for $n+1$ pairwise orthogonal vectors $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{n+1}$:

$$\begin{aligned}\|(\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n)+\boldsymbol{x}_{n+1}\|^2 &= \langle(\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n)+\boldsymbol{x}_{n+1},(\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n)+\boldsymbol{x}_{n+1}\rangle\\ &= \langle \boldsymbol{x}_1+\ldots+\boldsymbol{x}_n,\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n\rangle+\langle \boldsymbol{x}_1+\ldots+\boldsymbol{x}_n,\boldsymbol{x}_{n+1}\rangle+\langle \boldsymbol{x}_{n+1},\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n\rangle+\langle \boldsymbol{x}_{n+1},\boldsymbol{x}_{n+1}\rangle\\ &= \|\boldsymbol{x}_1\|^2+\ldots+\|\boldsymbol{x}_n\|^2+0+0+\|\boldsymbol{x}_{n+1}\|^2.\end{aligned}$$

In fact,

$$\langle \boldsymbol{x}_1+\ldots+\boldsymbol{x}_n,\boldsymbol{x}_{n+1}\rangle=\langle \boldsymbol{x}_1,\boldsymbol{x}_{n+1}\rangle+\ldots+\langle \boldsymbol{x}_n,\boldsymbol{x}_{n+1}\rangle=0$$

and the same is true for $\langle \boldsymbol{x}_{n+1},\boldsymbol{x}_1+\ldots+\boldsymbol{x}_n\rangle$ (i.e., it is also zero). ■

Solution 2.45 *(Exercise 2.45).* We prove the first statement. Let $\alpha=\sup A$. From the definition of supremum as the least upper bound (i.e., the minimum of the set $\{U\in\mathbb{R}\,|\,U \text{ is an upper bound for } A\}=\{U\in\mathbb{R}\,|\,\forall\, x\in A,\ x\le U\}$), the supremum is itself and upper bound, and therefore for every $x\in A$ we have $x\le\alpha$. This proves the first condition. Using again the definition of supremum as the least upper bound, we have that

$$\forall\, M\in\mathbb{R},\qquad M \text{ is an upper bound for } A \quad\Rightarrow\quad \alpha\le M. \tag{11.14}$$

We recall that, given two logical propositions p and q, the contrapositive of the implication $p\Rightarrow q$ is the statement $\neg q\Rightarrow\neg p$ (where the symbol "$\neg$" is the logical negation. The two statements $p\Rightarrow q$ and $\neg q\Rightarrow\neg p$ are logically equivalent (this property of formal logic is called *modus tollens*). Hence, the contrapositive of the statement (11.14) is the statement

$$\forall\, M\in\mathbb{R},\qquad \alpha>M \quad\Rightarrow\quad M \text{ is not an upper bound for } A. \tag{11.15}$$

The condition $\alpha>M$ is always satisfied if we take $M=\alpha-\varepsilon$, for any $\varepsilon>0$. Therefore

$$\forall\,\varepsilon>0,\quad \alpha-\varepsilon \text{ is not an upper bound for } A. \tag{11.16}$$

Saying that $\alpha-\varepsilon$ is not an upper bound for A is equivalent to saying that there exists an element of A greater than $\alpha-\varepsilon$, i.e.,

$$\forall\,\varepsilon>0,\quad \exists\, x\in A:\quad \alpha-\varepsilon<x. \tag{11.17}$$

This is exactly what we needed to prove. The second statement can be proved similarly.

■

Solution 2.46 *(Exercise 2.46).* Suppose that $s_1, s_2 \in X$ both satisfy (2.39). From Definition 2.22, for every $\varepsilon > 0$, there exist two finite subsets $F_\varepsilon^{(1)}$ and $F_\varepsilon^{(2)}$ such that, for any finite subset G of I satisfying $F_\varepsilon^{(1)} \cup F_\varepsilon^{(2)} \subseteq G$,

$$\left\| \sum_{\alpha \in G} x_\alpha - s_1 \right\| < \frac{\varepsilon}{2} \quad \text{and} \quad \left\| \sum_{\alpha \in G} x_\alpha - s_2 \right\| < \frac{\varepsilon}{2}.$$

From the triangle inequality

$$\|s_1 - s_2\| \leq \left\| \sum_{\alpha \in G} x_\alpha - s_1 \right\| + \left\| \sum_{\alpha \in G} x_\alpha - s_2 \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■

Solution 2.47 *(Exercise 2.47).* We prove **(a)**. We start by showing that, for all $\varepsilon > 0$, there exists $F_\varepsilon \subseteq I$, finite, such that for all $F \subseteq I$ and finite satisfying $F \cap F_\varepsilon = \emptyset$, we have

$$\sum_{i \in F} \|x_i\| \leq \varepsilon.$$

Suppose by contradiction that this is not true, i.e., for all $F_\varepsilon \subseteq I$, finite, there exist $F \subseteq I$ and finite satisfying $F \cap F_\varepsilon = \emptyset$, such that

$$\sum_{i \in F} \|x_i\| > \varepsilon.$$

Take any $F_1 \subseteq I$, finite, such that

$$\sum_{i \in F_1} \|x_i\| > \varepsilon.$$

Then, there exists F_2, finite, such that $F_1 \cap F_2 = \emptyset$ and

$$\sum_{i \in F_2} \|x_i\| > \varepsilon.$$

We deduce that

$$\sum_{i \in F_1 \cup F_2} \|x_i\| > 2\,\varepsilon.$$

There exists $F_3 \subseteq I$, finite, such that $(F_1 \cup F_2) \cap F_3 = \emptyset$ and

$$\sum_{i \in F_3} \|x_i\| > \varepsilon.$$

We deduce that

$$\sum_{i \in F_1 \cup F_2 \cup F_3} \|x_i\| > 3\,\varepsilon.$$

We can continue in this way, and we construct a finite subset F of I such that

$$\sum_{i \in F} \|x_i\| > M,$$

which is a contradiction.
Now, let us take a finite subset F_1 of I such that, if F is finite and $F_1 \cap F = \emptyset$,

$$\sum_{i \in F} \|x_i\| < 1.$$

Next, we take F_2 of I such that, if F is finite and $F_2 \cap F = \varnothing$,

$$\sum_{i \in F} \|\boldsymbol{x}_i\| < \frac{1}{2}.$$

No generality is lost by assuming $F_1 \subseteq F_2$. Continuing in this way, we construct a sequence of finite subsets $F_1 \subseteq F_2 \subseteq \ldots F_k \subseteq F_{k+1} \subseteq \ldots$ such that, if $F \cap F_k = \varnothing$, we have

$$\sum_{i \in F} \|\boldsymbol{x}_i\| < 2^{-k}.$$

Since for any $k, p \in \mathbb{N}^*$ it is always true that $F_k \cap (F_{k+p} \setminus F_k) = \varnothing$, we can take $F_{k+p} \setminus F_k$ to be F, and we obtain

$$\Big\| \sum_{i \in F_{k+p} \setminus F_k} \boldsymbol{x}_i \Big\| \leq \sum_{i \in F_{k+p} \setminus F_k} \|\boldsymbol{x}_i\| < 2^{-k}.$$

We obtain, for every $k, p \in \mathbb{N}^*$,

$$\Big\| \sum_{i \in F_{k+p}} \boldsymbol{x}_i - \sum_{i \in F_k} \boldsymbol{x}_i \Big\| < 2^{-k}. \tag{11.18}$$

Now, consider the sequence $(\boldsymbol{s}_k)_{k \in \mathbb{N}^*}$ of elements of $\boldsymbol{X}$, where

$$\boldsymbol{s}_k = \sum_{i \in F_k} \boldsymbol{x}_i.$$

Then, (11.18) ensures that $(\boldsymbol{s}_k)_{k \in \mathbb{N}^*}$ is a Cauchy sequence, and since $\boldsymbol{X}$ is complete, it has a limit. Let us denote the limit by $\boldsymbol{s}$:

$$\boldsymbol{s} = \lim_{k \to \infty} \sum_{i \in F_k} \boldsymbol{x}_i.$$

Now, we show that

$$\sum_{i \in I} \boldsymbol{x}_i = \boldsymbol{s}.$$

Let $\varepsilon > 0$. There exists k such that

$$\Big\| \sum_{i \in F_k} \boldsymbol{x}_i - \boldsymbol{s} \Big\| < \varepsilon.$$

If $F_k \subseteq F$ and F is finite,

$$\begin{aligned}
\Big\| \sum_{i \in F} \boldsymbol{x}_i - \boldsymbol{s} \Big\| &= \Big\| \sum_{i \in F} \boldsymbol{x}_i - \sum_{i \in F_k} \boldsymbol{x}_i \Big\| + \Big\| \sum_{i \in F_k} \boldsymbol{x}_i - \boldsymbol{s} \Big\| \\
&= \Big\| \sum_{i \in F \setminus F_k} \boldsymbol{x}_i \Big\| + \Big\| \sum_{i \in F_k} \boldsymbol{x}_i - \boldsymbol{s} \Big\| \\
&\leq \sum_{i \in F \setminus F_k} \|\boldsymbol{x}_i\| + \Big\| \sum_{i \in F_k} \boldsymbol{x}_i - \boldsymbol{s} \Big\| \leq 2^{-k} + \varepsilon.
\end{aligned}$$

Now, we only need to show that $\|\boldsymbol{s}\| \leq M$. Since

$$\Big\| \sum_{i \in F_k} \boldsymbol{x}_i \Big\| \leq \sum_{i \in F_k} \|\boldsymbol{x}_i\| \leq M \quad \text{and} \quad \|\boldsymbol{s}\| = \Big\| \lim_{k \to \infty} \sum_{i \in F_k} \boldsymbol{x}_i \Big\|,$$

we deduce that $\|\boldsymbol{s}\| \leq M$.

We prove **(b)**. We show that for every $\delta > 0$, the set $A_\delta = \{\alpha \in I \mid \|\boldsymbol{x}_\alpha\| \geq \delta\}$ is finite. Suppose by

contradiction that A_δ is not finite. Then, for every finite subset $F \subset A_\delta$ we have $\|\boldsymbol{x}_\alpha\| \geq \delta$ for every $\alpha \in F$, and so

$$\sum_{\alpha \in F} \|\boldsymbol{x}_\alpha\| \geq \sum_{\alpha \in F} \delta = \delta\,|F|,$$

where $|F|$ denotes the cardinality of F. Since A_δ is infinite, for any arbitrarily large $N \in \mathbb{N}$ we can choose a finite subset $F \subset A_\delta$ such that $|F| = N$, and so we can make $\displaystyle\sum_{\alpha \in F} \|\boldsymbol{x}_\alpha\|$ arbitrarily large, contradicting (2.40).

Now, we show that $B = \{\alpha \in I \,|\, \|\boldsymbol{x}_\alpha\| \neq 0\}$ is at most countable. Let $\delta_k = 1/k$ for $k \in \mathbb{N}^*$. From the previous result, for each $k \in \mathbb{N}^*$ the set $A_{\delta_k} = \{\alpha \in I \,|\, \|\boldsymbol{x}_\alpha\| \geq \delta_k\}$ is finite. Now, we show that

$$B = \bigcup_{k=1}^{\infty} A_{\delta_k}. \tag{11.19}$$

First, note that $B \subseteq \displaystyle\bigcup_{k=1}^{\infty} A_{\delta_k}$. In fact,

$$\alpha \in B \quad \Rightarrow \quad \|\boldsymbol{x}_\alpha\| \neq 0 \quad \Rightarrow \quad \|\boldsymbol{x}_\alpha\| > 0.$$

Take $k \geq 1/\|\boldsymbol{x}_\alpha\|$. We find $\alpha \in A_{\delta_k}$, and so α belongs to the union. Consider the opposite inclusion. Let $\alpha \in A_{\delta_k}$ for some $k \in \mathbb{N}^*$. Then, $\|\boldsymbol{x}_\alpha\| \geq \delta_k = 1/k$, and so $\|\boldsymbol{x}_\alpha\| \neq 0$, from which we deduce that $\alpha \in B$. From (11.19), the set B is the countable union of finite sets, and so B is finite or countable. ■

Solution 2.48 *(Exercise 2.48).* We have to show that for all $\varepsilon > 0$, there exists a finite set $F_\varepsilon \subseteq I$ such that for all finite subset $G \subseteq I$, if $F_\varepsilon \subseteq G$ we have

$$\Big\| \sum_{\alpha \in G} T(\boldsymbol{x}_\alpha) - T(\boldsymbol{s}) \Big\|_Y < \varepsilon.$$

Let $\varepsilon > 0$. Since T is continuous, there exists $M > 0$ such that for all $\boldsymbol{x} \in \boldsymbol{X}$, we have $\|T(\boldsymbol{x})\|_Y \leq M\,\|\boldsymbol{x}\|_X$. From Definition 2.22, there exists a finite set $F_\varepsilon \subseteq I$ such that for all finite subset $G \subseteq I$, if $F_\varepsilon \subseteq G$ we have

$$\Big\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \Big\|_X < \frac{\varepsilon}{M}.$$

Using the linearity of T

$$\begin{aligned} \Big\| \sum_{\alpha \in G} T(\boldsymbol{x}_\alpha) - T(\boldsymbol{s}) \Big\|_Y &= \Big\| T\Big(\sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \Big) \Big\|_Y \\ &\leq M \Big\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \Big\|_X < M\,\frac{\varepsilon}{M} = \varepsilon. \end{aligned}$$

■

Solution 2.49 *(Exercise 2.49).* Suppose that the unordered sum

$$\sum_{\alpha \in I} \boldsymbol{x}_\alpha$$

converge to $\boldsymbol{s}$. Let $\varepsilon > 0$. From Definition 2.22, there exists a finite subset F_ε of I such that, for any finite subset G of I,

$$F_\varepsilon \subseteq G \quad \Rightarrow \quad \Big\| \sum_{\alpha \in G} \boldsymbol{x}_\alpha - \boldsymbol{s} \Big\| < \frac{\varepsilon}{2}.$$

Now, given any finite subsets G' and H of I such that contain F_ε,

$$\left\| \sum_{\alpha \in G'} \boldsymbol{x}_\alpha - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| \leq \left\| \sum_{\alpha \in G'} \boldsymbol{x}_\alpha - \boldsymbol{s} \right\| + \left\| \boldsymbol{s} - \sum_{\alpha \in H} \boldsymbol{x}_\alpha \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■

Solution 2.50 *(Exercise 2.50).* We need to show that $\overline{\boldsymbol{V}}$ is non-empty and it is closed under addition and scalar multiplication. Recall that $\overline{\boldsymbol{V}}$ is the set of points of $\boldsymbol{X}$ that are limits of sequences of points from $\boldsymbol{V}$.

Let $\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}} \in \overline{\boldsymbol{V}}$. Then, there exist sequences $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ and $(\boldsymbol{y}_k)_{k \in \mathbb{N}}$ such that

$$\boldsymbol{x}_k \xrightarrow{k \to \infty} \hat{\boldsymbol{x}} \quad \text{and} \boldsymbol{y}_k \xrightarrow{k \to \infty} \hat{\boldsymbol{y}}.$$

Since $\boldsymbol{V}$ is a subspace of $\boldsymbol{X}$, it is closed under addition, and so $(\boldsymbol{x}_k + \boldsymbol{y}_k)_{k \in \mathbb{N}}$ is a sequence of points from $\boldsymbol{V}$. Since both sequences $(\boldsymbol{x}_k)_{k \in \mathbb{N}}$ and $(\boldsymbol{y}_k)_{k \in \mathbb{N}}$ converge, the sum also converges, and limit is $\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}}$. It follows that $\hat{\boldsymbol{x}} + \hat{\boldsymbol{y}} \in \overline{\boldsymbol{V}}$. Now, take $k \in \mathbb{F}$. Since $\boldsymbol{V}$ is closed under scalar multiplication, $(k\,\boldsymbol{x}_k)_{k \in \mathbb{N}}$ is a sequence of points from $\boldsymbol{V}$. From the continuity of scalar multiplication (Theorem 2.20, where the sequence in $\mathbb{F}$ is assumed to be constant), we have

$$k\,\boldsymbol{x}_k \xrightarrow{k \to \infty} k\,\hat{\boldsymbol{x}}.$$

Therefore, $k\,\hat{\boldsymbol{x}} \in \overline{\boldsymbol{V}}$.

Finally, since $\boldsymbol{V}$ is a subspace of $\boldsymbol{X}$, it contains the zero vector of $\boldsymbol{X}$. Moreover, $\boldsymbol{V}$ contains the constant sequence $\boldsymbol{x}_k = 0$ for all $k \in \mathbb{N}$. It follows that $0 \in \overline{\boldsymbol{V}}$. ■

12 Solutions to the Exercises in Chapter 3

Solution 3.1 *(Exercise 3.1).* Consider the reflexive property. If $a \in [0, +\infty)$, this property holds, because it holds in $\mathbb{R}$ for the standard order $\leq$. If $a = +\infty$, then $a \leq a$ holds since $a = a$.
Consider the transitive property. If $a, b, c \in [0, +\infty)$, the property holds. The other possible case compatible with the relations $a \leq b$ and $b \leq c$ are the following:

- $a = b = c = +\infty$. Then, $a \leq c$ is true;
- $b = c = +\infty$ and $a \in [0, +\infty)$. Then, again, $a \leq c$ is true;
- $c = +\infty$ and $a, b \in [0, +\infty)$. Again, $a \leq c$ is true.

Let us consider the anti-symmetric property. If $a, b \in [0, +\infty)$, the property holds true. If $a = +\infty$, then $a \leq b$ implies $b = +\infty$, so that $a = b$. If $b = +\infty$, then $b \leq a$ implies $a = +\infty$, so that, again, $a = b$.

Finally, consider the last point. Once again, if $a, b \in [0, +\infty)$, the property holds true. We then have the following cases:

- $a = +\infty$ and $b \in [0, +\infty)$. We obtain $b < a$;
- $a \in [0, +\infty)$ and $b = +\infty$. Then, $a < b$;
- $a = b = +\infty$. In this case, $a = b$.

In all these cases, we have $a \leq b$ or $b \leq a$. This concludes the argument. ■

Solution 3.2 *(Exercise 3.2).* We recall that, given $a, b \in [0, +\infty]$,

$$a + b = \begin{cases} a + b & \text{if } a, b \in [0, +\infty) \\ +\infty & \text{if } a = +\infty \text{ or } b = +\infty \end{cases}$$

where the $+$ sign in the right hand-side is the usual addition sign in $\mathbb{R}$. It follows that

$$\begin{aligned} a + b &= \begin{cases} b + a & \text{if } a, b \in [0, +\infty) \\ +\infty & \text{if } a = +\infty \text{ or } b = +\infty \end{cases} \\ &= b + a. \end{aligned}$$

This proves the commutative property. Consider the associative property. In other words, given $a, b, c \in [0, +\infty]$, we show that $(a + b) + c = a + (b + c)$. By definition of addition in $[0, +\infty]$ we have

$$\begin{aligned} (a + b) + c &= \begin{cases} (a + b) + c & \text{if } a + b, c \in [0, +\infty) \\ +\infty & \text{if } a + b = +\infty \text{ or } c = +\infty \end{cases} \\ &= \begin{cases} (a + b) + c & \text{if } a, b, c \in [0, +\infty) \\ +\infty & \text{if } (a = +\infty \text{ or } b = +\infty) \text{ or } c = +\infty \end{cases} \\ &= \begin{cases} (a + b) + c & \text{if } a, b, c \in [0, +\infty) \\ +\infty & \text{if } a = +\infty \text{ or } b = +\infty \text{ or } c = +\infty \end{cases} \\ &= \begin{cases} a + (b + c) & \text{if } a, b, c \in [0, +\infty) \\ +\infty & \text{if } a = +\infty \text{ or } b = +\infty \text{ or } c = +\infty \end{cases} \\ &= a + (b + c). \end{aligned}$$

■

Solution 3.3 *(Exercise 2.1).* We recall that, given $a, b \in [0, +\infty]$,

$$a \cdot b = \begin{cases} a \cdot b & \text{if } a, b \in (0, +\infty) \\ 0 & \text{if } a = 0 \text{ or } b = 0 \\ +\infty & \text{if } (a \neq 0 \text{ and } b = +\infty) \text{ or } (a = +\infty \text{ and } b \neq 0) \end{cases}$$

where the $\cdot$ sign in the right hand-side is the usual multiplication sign in $\mathbb{R}$. The symmetry of this definition and the fact that the commutative property holds for the standard multiplication in $\mathbb{R}$ allow us to conclude that the multiplication in $[0, +\infty]$ is commutative.

Let us now prove the associative property. First, notice that $a \cdot b = 0$ if and only if $a = 0$ or $b = 0$. As a result, $a \cdot b \in (0, +\infty)$ if and only if $a, b \in (0, +\infty)$. Let $a, b, c \in [0, +\infty]$.

$$\begin{aligned}
(a \cdot b) \cdot c &= \begin{cases} (a \cdot b) \cdot c & \text{if } a \cdot b, c \in (0, +\infty) \\ 0 & \text{if } a \cdot b = 0 \text{ or } c = 0 \\ +\infty & \text{if } (a \cdot b \neq 0 \text{ and } c = +\infty) \text{ or } (a \cdot b = +\infty \text{ and } c \neq 0) \end{cases} \\
&= \begin{cases} (a \cdot b) \cdot c & \text{if } a \cdot b, c \in (0, +\infty) \\ 0 & \text{if } a \cdot b = 0 \text{ or } c = 0 \\ +\infty & \text{if } a \cdot b \neq 0 \text{ and } c = +\infty \\ +\infty & \text{if } a \cdot b = +\infty \text{ and } c \neq 0 \end{cases} \\
&= \begin{cases} (a \cdot b) \cdot c & \text{if } a, b, c \in (0, +\infty) \\ 0 & \text{if } a = 0 \text{ or } b = 0 \text{ or } c = 0 \\ +\infty & \text{if } (a, b \neq 0 \text{ and } c = +\infty) \\ +\infty & \text{if } [(a \neq 0 \text{ and } b = +\infty) \text{ or } (a = +\infty \text{ and } b \neq 0)] \text{ and } c \neq 0 \end{cases} \\
&= \begin{cases} (a \cdot b) \cdot c & \text{if } a, b, c \in (0, +\infty) \\ 0 & \text{if } a = 0 \text{ or } b = 0 \text{ or } c = 0 \\ +\infty & \text{if } (a, b \neq 0 \text{ and } c = +\infty) \\ +\infty & \text{if } (a, c \neq 0 \text{ and } b = +\infty) \text{ or } (a = +\infty \text{ and } b, c \neq 0) \end{cases} \\
&= \begin{cases} (a \cdot b) \cdot c & \text{if } a, b, c \in (0, +\infty) \\ 0 & \text{if } a = 0 \text{ or } b = 0 \text{ or } c = 0 \\ +\infty & \text{if } (a, b \neq 0 \text{ and } c = +\infty) \text{ or } (a, c \neq 0 \text{ and } b = +\infty) \text{ or } (a = +\infty \text{ and } b, c \neq 0) \end{cases}
\end{aligned}$$

Likewise

$$
\begin{aligned}
a\cdot(b\cdot c) &= \begin{cases} a\cdot(b\cdot c) & \text{if } a, b\cdot c \in (0,+\infty) \\ 0 & \text{if } a=0 \text{ or } b\cdot c = 0 \\ +\infty & \text{if } (a\neq 0 \text{ and } b\cdot c = +\infty) \text{ or } (a=+\infty \text{ and } b\cdot c \neq 0) \end{cases} \\
&= \begin{cases} a\cdot(b\cdot c) & \text{if } a, b, c \in (0,+\infty) \\ 0 & \text{if } a=0 \text{ or } b=0 \text{ or } c=0 \\ +\infty & \text{if } a\neq 0 \text{ and } b\cdot c = +\infty \\ +\infty & \text{if } a=+\infty \text{ and } b\cdot c \neq 0 \end{cases} \\
&= \begin{cases} a\cdot(b\cdot c) & \text{if } a, b, c \in (0,+\infty) \\ 0 & \text{if } a=0 \text{ or } b=0 \text{ or } c=0 \\ +\infty & \text{if } a\neq 0 \text{ and } [(b\neq 0 \text{ and } c=+\infty) \text{ or } (b=+\infty \text{ and } c\neq 0)] \\ +\infty & \text{if } a=+\infty \text{ and } (b\neq 0 \text{ and } c\neq 0) \end{cases} \\
&= \begin{cases} a\cdot(b\cdot c) & \text{if } a, b, c \in (0,+\infty) \\ 0 & \text{if } a=0 \text{ or } b=0 \text{ or } c=0 \\ +\infty & \text{if } (a, b\neq 0 \text{ and } c=+\infty) \text{ or } (b=+\infty \text{ and } a, c\neq 0) \text{ or } (a=+\infty \text{ and } (b, c\neq 0) \end{cases}
\end{aligned}
$$

The two expressions coincide. ■

Solution 3.4 *(Exercise 3.4).* If $p = 1$, the statement is trivially true, since it reduces to the identity $\alpha_1 + \beta_1 = \alpha_1 + \beta_1$. Suppose that (11.14) holds for $p \in \mathbb{N}^*$. We show that it holds also for $p+1$:

$$
\begin{aligned}
\sum_{k=1}^{p+1}(\alpha_k+\beta_k) &= \sum_{k=1}^{p}(\alpha_k+\beta_k) + \alpha_{p+1} + \beta_{p+1} \\
&= \sum_{k=1}^{p}\alpha_k + \sum_{k=1}^{p}\beta_k + \alpha_{p+1} + \beta_{p+1} \\
&= \sum_{k=1}^{p+1}\alpha_k + \sum_{k=1}^{p+1}\beta_k.
\end{aligned}
$$

■

Solution 3.5 *(Exercise 3.5).* If $q = 1$, the statement is trivially true, since it reduces to the identity

$$
\sum_{k=1}^{p} a_{k,1} = \sum_{k=1}^{p} a_{k,1}.
$$

Suppose now that (11.15) holds for $q \in \mathbb{N}^*$. We show that it holds also for $q+1$:

$$
\begin{aligned}
\sum_{k=1}^{p}\left(\sum_{\ell=1}^{q+1} a_{k,\ell}\right) &= \sum_{k=1}^{p}\left(\sum_{\ell=1}^{q} a_{k,\ell} + a_{k,q+1}\right) \\
&= \sum_{k=1}^{p}\left(\sum_{\ell=1}^{q} a_{k,\ell}\right) + \sum_{k=1}^{p} a_{k,q+1} \\
&= \sum_{\ell=1}^{q}\left(\sum_{k=1}^{p} a_{k,\ell}\right) + \sum_{k=1}^{p} a_{k,q+1} = \sum_{\ell=1}^{q+1}\left(\sum_{k=1}^{p} a_{k,\ell}\right).
\end{aligned}
$$

■

Solution 3.6 *(Exercise 3.6).*

$$\begin{aligned} x \in A \setminus B \quad &\Leftrightarrow \quad x \in A \ \text{ and } \ x \notin B \\ &\Leftrightarrow \quad x \in A \ \text{ and } \ x \in B^{\complement} \\ &\Leftrightarrow \quad x \in A \cap B^{\complement}. \end{aligned}$$

■

Solution 3.7 *(Exercise 3.7).* We define $A_k = X$ for all $k > m$, and then use **(c)**:

$$\bigcap_{k=0}^{m} A_k = \bigcap_{k=0}^{\infty} A_k \in \mathcal{A}.$$

Indeed, for any subset A of X, there holds $A \cap X = A$. ■

Solution 3.8 *(Exercise 3.8).* Let $A = \varnothing$. In thus case, $x_0 \notin A$, and so $\mu(A) = 0$. Let now $(A_k)_{k\in\mathbb{N}}$ be a sequence of pairwise disjoint elements of $\mathcal{A}$. If

$$x_0 \in \bigcup_{k=0}^{\infty} A_k,$$

there exists one and only one $k_0 \in \mathbb{N}$ such that $x_0 \in A_{k_0}$, otherwise there would be more elements of the sequence containing x_0 and x_0 would be in the intersection of at least two elements of $\mathcal{A}$, contradicting the assumption that the elements of $(A_k)_{k\in\mathbb{N}}$ are pairwise disjoint. It follows that

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = 1 = \sum_{k=0}^{\infty} \mu(A_k).$$

■

Solution 3.9 *(Exercise 3.9).* The power set $\mathcal{P}(X)$ is a σ-algebra in X, because it contains all the subsets of X, and therefore in particular it contains the empty set, the complement of every subset of X and the countable unions of the subsets of X (which is itself a subset of X). We have

$$\mu(\varnothing) = 0,$$

because $x_0 \notin \varnothing$. Now, let $(A_k)_{k\in\mathbb{N}}$ be a sequence of pairwise disjoint subsets of X. Then,

$$\begin{aligned} \mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) &= \begin{cases} 1 & \text{if } x_0 \in \bigcup_{k=0}^{\infty} A_k \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} 1 & \text{if } \exists\, k \in \mathbb{N} : \ x_0 \in A_k \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Now, we show that $\sum_{k=0}^{\infty} \mu(A_k) \in \{0, 1\}$. In fact, since the sets of the sequence are pairwise disjoint, $x_0 \in A_k$ implies $x_0 \notin A_h$ if $h \neq k$, so that only one addend of the sum can be equal to 1. This happens when $x_0 \in A_k$ for some $k \in \mathbb{N}$. Otherwise, the sum is a sum of zeros:

$$\sum_{k=0}^{\infty} \mu(A_k) = \begin{cases} 1 & \text{if } \exists\, k \in \mathbb{N} : \ x_0 \in A_k \\ 0 & \text{otherwise} \end{cases}$$

Hence,

$$\mu\Big(\bigcup_{k=0}^{\infty} A_k\Big) = \sum_{k=0}^{\infty} \mu(A_k).$$

■

Solution 3.10 *(Exercise 3.11).* We have

$$I_{x_n} = \{(x_1, x_2, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} \,|\, (x_1, x_2, \ldots, x_n) \in I\}.$$

Let $(x_1, x_2, \ldots, x_n) \in I$. Then, if $x_n \notin I_n$, we have $I_{x_n} = \varnothing$. If $x_n \in I_n$, we have

$$I_{x_n} = \{(x_1, x_2, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} \,\big|\, \forall\, i \in \{1, 2, \ldots, n-1\},\ x_i \in I_i\} = I_1 \times I_2 \times \ldots \times I_{n-1}.$$

■

Solution 3.11 *(Exercise 3.12).* Let $(x, y) \in (A \times B) \cap (C \times D)$. Then, $(x, y) \in (A \times B)$ and $(x, y) \in (C \times D)$. It follows that $x \in A$ and $x \in C$, and so $x \in A \cap C$. Likewise, $y \in B \cap D$, and so $(x, y) \in (A \cap C) \times (B \cap D)$. Conversely, let $(x, y) \in (A \cap C) \times (B \cap D)$, which implies that $x \in A$ and $y \in B$, so $(x, y) \in A \times B$. Similarly, $(x, y) \in (A \cap C) \times (B \cap D)$ implies that $x \in C$ and $y \in D$, and so $(x, y) \in C \times D$. Therefore, $(x, y) \in (A \times B) \cap (C \times D)$. ■

Solution 3.12 *(Exercise 3.13).* We recall that I is degenerate if it is empty or if it contains a single point. Assume that I is non-degenerate, so it contains at least two distinct points, say $a, b \in I$. Then, $\sup I \geq \max\{a, b\}$ and $\inf I \leq \min\{a, b\}$, so that $\ell(I) = \sup I - \inf I = \max\{a, b\} - \min\{a, b\} = |a - b| > 0$. Hence, $\ell(I) > 0$. From the definition of interval of $\mathbb{R}$, every $x \in (a, b)$ is a point of I, and therefore I contains an infinite number of points. ■

Solution 3.13 *(Exercise 3.15).* We have to show that

$$\boldsymbol{x} + \boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} (\boldsymbol{x} + \boldsymbol{I}^k) \qquad \Leftrightarrow \qquad \boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k.$$

We start proving $\subseteq$. Let $\boldsymbol{e} \in \boldsymbol{E}$. Then,

$$\boldsymbol{x} + \boldsymbol{e} \in \boldsymbol{x} + \boldsymbol{E} \quad \Rightarrow \quad \boldsymbol{x} + \boldsymbol{e} \in \bigcup_{k=0}^{\infty} (\boldsymbol{x} + \boldsymbol{I}^k),$$

and therefore there exists $h \in \mathbb{N}$ such that $\boldsymbol{x} + \boldsymbol{e} \in \boldsymbol{x} + \boldsymbol{I}^h$. On the other hand, by definition of shifted set $\boldsymbol{x} + \boldsymbol{I}^h = \{\boldsymbol{x} + \boldsymbol{p} \,|\, \boldsymbol{p} \in \boldsymbol{I}^h\}$, so that $\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{x} + \boldsymbol{p}$ for some $\boldsymbol{p} \in \boldsymbol{I}^h$. This implies $\boldsymbol{e} = \boldsymbol{p}$, i.e., $\boldsymbol{e} \in \boldsymbol{I}^h$. From the arbitrariness of $\boldsymbol{e}$ we obtain $\boldsymbol{E} \subseteq \boldsymbol{I}^h$, and therefore

$$\boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k.$$

The opposite inclusion is trivial. Assume $\boldsymbol{E} \subseteq \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$. For any $\boldsymbol{e} \in \boldsymbol{E}$, we have $\boldsymbol{e} \in \bigcup_{k=0}^{\infty} \boldsymbol{I}^k$, and therefore there exists $h \in \mathbb{N}$ such that $\boldsymbol{e} \in \boldsymbol{I}^h$. Hence, $\boldsymbol{x} + \boldsymbol{e} \in \{\boldsymbol{x} + \boldsymbol{p} \,|\, \boldsymbol{p} \in \boldsymbol{I}^h\}$, and from the arbitrariness of $\boldsymbol{e} \in \boldsymbol{E}$ we find

$$\boldsymbol{x} + \boldsymbol{E} \subseteq \{\boldsymbol{x} + \boldsymbol{p} \,|\, \boldsymbol{p} \in \boldsymbol{I}^h\} = \boldsymbol{x} + \boldsymbol{I}^h \subseteq \bigcup_{k=0}^{\infty} (\boldsymbol{x} + \boldsymbol{I}^k).$$

■

Solution 3.14 *(Exercise 3.16).* Let $N, M \in \mathbb{R}_+$. Clearly, $(-N, N)\times(-M, M) \subseteq \mathbb{R}^2$. From the monotonicity of the outer measure,

$$\mu_2^*((-N, N) \times (-M, M)) \leq \mu_2^*(\mathbb{R}^2).$$

On the other hand, $(-N, N) \times (-M, M)$ is a 2-dimensional interval, given by the Cartesian product of two intervals $(-N, N)$ and $(-M, M)$ of $\mathbb{R}$, so that

$$\mu_2^*((-N, N) \times (-M, M)) = \mathrm{vol}_2((-N, N) \times (-M, M)) = \ell((-N, N)) \cdot \ell((-M, M)) = 4\, M\, N.$$

Combining the two, we find $\mu_2^*(\mathbb{R}^2) \geq 4\, M\, N$ for any $M, N > 0$. This implies $\mu_2^*(\mathbb{R}^2) = +\infty$. ■

Solution 3.15 *(Exercise 3.17).* We prove that $\mu_3^*(\mathbb{R} \times \mathbb{R} \times \{0\}) = 0$. We have

$$\mathbb{R} \times \mathbb{R} \times 0 = \bigcup_{k=0}^{\infty} [-k, k] \times [-k, k] \times \{0\}.$$

From the countable sub-additivity we have

$$\begin{aligned}
\mu_3^*(\mathbb{R} \times \mathbb{R} \times \{0\}) &\leq \sum_{k=0}^{\infty} \mu_3^*([-k, k] \times [-k, k] \times \{0\}) \\
&= \sum_{k=0}^{\infty} \mathrm{vol}_3([-k, k] \times [-k, k] \times \{0\}) \\
&= \sum_{k=0}^{\infty} \ell([-k, k])\, \ell([-k, k])\, \ell(\{0\}) = 0.
\end{aligned}$$

We could have written instead

$$\mathbb{R} \times \mathbb{R} \times \{0\} = \bigcup_{k=0}^{\infty} \bigcup_{h=0}^{\infty} [-k, k] \times [-h, h] \times \{0\}.$$

To apply the sub-additivity, in this case, we need to consider a bijection $\varphi : \mathbb{N} \longrightarrow \mathbb{N}\times\mathbb{N}$ that maps ℓ into $\varphi(\ell) = (k, h)$, so that by denoting $I^{(k,h)} = [-k, k] \times [-h, h]$, we have

$$\mathbb{R} \times \mathbb{R} \times \{0\} = \bigcup_{k=0}^{\infty} \bigcup_{h=0}^{\infty} I^{(k,h)} \times \{0\} = \bigcup_{\ell=0}^{\infty} I^{\varphi(\ell)} \times \{0\}.$$

Thus,

$$\begin{aligned}
\mu_3^*(\mathbb{R} \times \mathbb{R} \times \{0\}) &= \sum_{\ell=0}^{\infty} \mathrm{vol}_3(I^{\varphi(\ell)} \times \{0\}) \\
&= \sum_{(k,h)\in\mathbb{N}\times\mathbb{N}} \mathrm{vol}_3(I^{(k,h)} \times \{0\}) \\
&= \sum_{k\in\mathbb{N}} \Big(\sum_{h\in\mathbb{N}} \mathrm{vol}_3(I^{(k,h)} \times \{0\}) \Big) \\
&= \sum_{k=0}^{\infty} \Big(\sum_{h=0}^{\infty} \mathrm{vol}_3(I^{(k,h)} \times \{0\}) \Big) \\
&= \sum_{k=0}^{\infty} \Big(\sum_{h=0}^{\infty} 4\, k\, h \cdot 0 \Big) = 0.
\end{aligned}$$

■

Solution 3.16 *(Exercise 3.18).* For every $\boldsymbol{A} \subseteq \mathbb{R}^n$,

$$\mu_n^*(\boldsymbol{A} \cap \varnothing) + \mu_n^*(\boldsymbol{A} \cap \varnothing^{\complement}) = \mu_n^*(\varnothing) + \mu_n^*(\boldsymbol{A} \cap \mathbb{R}^n) = \mu_n^*(\boldsymbol{A})$$

from **(a)** in Theorem 3.14; ■

Solution 3.17 *(Exercise 3.19).* For every $\boldsymbol{A} \subseteq \mathbb{R}^n$,

$$\mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement\complement}) = \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}^{\complement}) + \mu_n^*(\boldsymbol{A} \cap \boldsymbol{E}) = \mu_n^*(\boldsymbol{A}).$$

■

Solution 3.18 *(Exercise 3.20).* Using (3.23), we only need to show that for all $A \subseteq \mathbb{R}^n$,

$$\forall\, \boldsymbol{A} \subseteq \mathbb{R}^n, \qquad \mu_n^*(\boldsymbol{A}) \geq \mu_n^*(\boldsymbol{A} \cap \{\boldsymbol{x}_0\}) + \mu_n^*(\boldsymbol{A} \cap \{\boldsymbol{x}_0\}^{\complement}).$$

Let $\boldsymbol{A} \subseteq \mathbb{R}^n$. We have $\mu_n^*(\boldsymbol{A} \cap \{\boldsymbol{x}_0\}) = 0$, because $\boldsymbol{A} \cap \{\boldsymbol{x}_0\}$ is either the empty set (if $\boldsymbol{x}_0 \notin \boldsymbol{A}$) or a singleton (if $\boldsymbol{x}_0 \in \boldsymbol{A}$), and in both cases the outer measure is 0. We have

$$\begin{aligned}\mu_n^*(\boldsymbol{A} \cap \{\boldsymbol{x}_0\}^{\complement}) &= \mu_n^*(\boldsymbol{A} \cap (\mathbb{R}^n \setminus \{\boldsymbol{x}_0\})) \\ &= \mu_n^*(\boldsymbol{A} \cap \mathbb{R}^n \cap \{\boldsymbol{x}_0\}^{\complement}) = \mu_n^*(\boldsymbol{A} \cap \{\boldsymbol{x}_0\}^{\complement}) = \mu_n^*(\boldsymbol{A} \setminus \{\boldsymbol{x}_0\}) \leq \mu_n^*(\boldsymbol{A}).\end{aligned}$$

It follows that $\{\boldsymbol{x}_0\}$ is Lebesgue measurable ■

Solution 3.19 *(Exercise 3.21).* We have already shown that $\{\boldsymbol{x}_0\}$ is Lebesgue measurable in Exercise 3.20. However, even without taking the result of Exercise 3.20 into account, if we prove that the outer measure $\mu_n^*(\{\boldsymbol{x}_0\})$ is zero, then we can apply Theorem 3.19 to conclude that $\{\boldsymbol{x}_0\}$ is Lebesgue measurable, and thus its Lebesgue measure is equal to its outer measure (which is zero). We have We write $\boldsymbol{x}_0 = (x_{0,1}, x_{0,2}, \ldots, x_{0,1})$. The box $B = [x_{0,1}, x_{0,1}] \times [x_{0,2}, x_{0,2}] \times \ldots \times [x_{0,n}, x_{0,n}]$ is a covering of $\boldsymbol{x}_0$, and $\operatorname{vol}_n B = 0$ because $\ell([x_{0,j}, x_{0,j}]) = 0$ for all $j \in \{1, 2, \ldots, n\}$. Therefore, $\mu_n^*(\{\boldsymbol{x}_0\}) = 0$. It follows that $\mu_n(\{\boldsymbol{x}_0\}) = 0$.
To prove the second point we can use Theorem 3.18. A countable set J in $\mathbb{R}^n$ can be written as the union of singletons. In fact, by enumeration of the elements of J as $j_0, j_1, \ldots$, we have

$$J = \bigcup_{k=0}^{\infty} \{j_k\},$$

and the countable union of Lebesgue measurable sets is Lebesgue measurable by Theorem 3.18. Now,

$$\mu_n(J) = \mu_n^*(J) = \sum_{k=0}^{\infty} \mu_n^*(\{j_k\}) = 0.$$

■

Solution 3.20 *(Exercise 3.22).* Since $\mu_n(\boldsymbol{A}) = 0$, there exists $(\boldsymbol{I}^k)_{k \in \mathbb{N}}$ such that

$$\boldsymbol{A} \subseteq \bigcup_{k \in \mathbb{N}} \boldsymbol{I}^k \qquad \text{and} \qquad \sum_{k \in \mathbb{N}} \operatorname{vol}_n(\boldsymbol{I}^k) \leq \varepsilon$$

for any $\varepsilon > 0$. Since $\boldsymbol{B} \subseteq \boldsymbol{A}$, it follows that

$$\boldsymbol{B} \subseteq \bigcup_{k \in \mathbb{N}} \boldsymbol{I}^k \qquad \text{and} \qquad \sum_{k \in \mathbb{N}} \operatorname{vol}_n(\boldsymbol{I}^k) \leq \varepsilon.$$

Thus, $\mu_n(\boldsymbol{B}) = 0$. ■

Solution 3.21 *(Exercise 3.23).* We can write

$$\boldsymbol{A} = \mathring{\boldsymbol{A}} \cup \boldsymbol{A}_1, \qquad \text{where } \boldsymbol{A}_1 \subseteq \partial \boldsymbol{A}.$$

Since $\mathring{\boldsymbol{A}}$ is open, it is measurable. Since $\mu_n^*(\partial \boldsymbol{A}) = 0$, by the monotonicity of the outer measure we have also $\mu_n^*(\boldsymbol{A}_1) = 0$. A set with outer measure zero is measurable, so that $\boldsymbol{A}_1 \in \mathcal{M}_n$. From the properties of σ-algebras, the union of measurable sets is measurable. Hence, we conclude that $\boldsymbol{A}$ is measurable. ■

Solution 3.22 *(Exercise 3.24).* Let $\boldsymbol{x} \in \boldsymbol{A}_i$. Then,

$$\sum_{k=1}^{p} \alpha_k \chi_{A_k}(\boldsymbol{x}) = \alpha_1 \underbrace{\chi_{A_1}(\boldsymbol{x})}_{=0} + \alpha_2 \underbrace{\chi_{A_2}(\boldsymbol{x})}_{=0} + \ldots + \alpha_i \underbrace{\chi_{A_i}(\boldsymbol{x})}_{=1} + \ldots + \alpha_p \underbrace{\chi_{A_p}(\boldsymbol{x})}_{=0} = \alpha_i.$$

■

Solution 3.23 *(Exercise 3.25).* The product λf can be viewed as the product of the constant function

$$\begin{aligned} \boldsymbol{A} : &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \lambda, \end{aligned}$$

which is a simple function, by f, which is simple by assumption. Since in Theorem 3.23 it is proved that the product of two simple functions is itself a simple function, this conclusion holds also in the case λf. ■

Solution 3.24 *(Exercise 3.26).* The function f is a simple function because there exist

- $m = 3$
- disjoint subsets $A_1 = [0, 3]$, $A_2 = (3, 5]$, $A_3 = (-\infty, 0) \cup (5, +\infty)$ of $\mathbb{R}$ such that $A_1 \cup A_2 \cup A_3 = \mathbb{R}$
- $\alpha_1 = 1$, $\alpha_2 = 3$ and $\alpha_3 = 0$

such that $f(x) = \alpha_i$ for all $x \in A_i$. We have

$$\begin{aligned} \int_{\boldsymbol{A}} f(x)\,\mathrm{d}x &= \sum_{i=1}^{3} \alpha_i \mu_1(A_i) = 1 \cdot \mu_1([0,3]) + 3 \cdot \mu_1((3,5]) + 0 \cdot \mu_1((-\infty, 0) \cup (5, +\infty)) \\ &= 1 \cdot 3 + 3 \cdot 2 + 0 \cdot (+\infty) = 9. \end{aligned}$$

■

Solution 3.25 *(Exercise 3.27).* The function f is a simple function. To verify this, we can take

$$\boldsymbol{A}_1 = [0,1] \times [0,1], \quad \boldsymbol{A}_2 = (1,2] \times [0,1], \quad \boldsymbol{A}_3 = \mathbb{R}^2 \setminus (A_1 \cup A_2).$$

It follows that $\boldsymbol{A}_1, \boldsymbol{A}_2, \boldsymbol{A}_3$ are measurable, pairwise disjoint and their union is $\mathbb{R}^2$. Moreover, $f(x_1, x_2) = \alpha_i$ for $\boldsymbol{x} \in \boldsymbol{A}_i$, where $\alpha_1 = 1$, $\alpha_2 = 2$ and $\alpha_3 = 0$. We have

$$\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1 \cdot \mathrm{vol}_2(\boldsymbol{A}_1) + 2 \cdot \mathrm{vol}_2(\boldsymbol{A}_2) + 0 \cdot \mu_2(\boldsymbol{A}_3) = 1 + 2 = 3.$$

■

Solution 3.26 *(Exercise 3.28).* The function f is a simple function because we can partition its domain as the union of two measurable (and obviously disjoint) subsets $A_1 = (-\infty, 0]$ and $A_2 = (0, +\infty)$ (the measurability follows from the fact that A_1 is closed and A_2 is open), and, moreover, $f(x) = \alpha_i$ for all $x \in A_i$ for $i \in \{0, 1\}$, where $\alpha_1 = 0$ and $\alpha_2 = 1$. We have

$$\int_{\boldsymbol{A}} f(x)\,\mathrm{d}x = 0 \cdot \mu_1(A_1) + 1 \cdot \mu_1(A_2) = \mu_1(A_2).$$

To find $\mu_1(A_2)$, we observe that for all $k \in \mathbb{N}$ we have $(0, k] \subseteq A_2$, so that

$$\mu_1(A_2) \geq \mu_1((0, k]) = k.$$

Since this is true for every $k \in \mathbb{N}$, we conclude that $\mu_1(A_2) = +\infty$. ■

Solution 3.27 *(Exercise 3.29).* By definition of simple function, there exist $p \in \mathbb{N}^*$, the sets $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_p \subseteq \boldsymbol{A}$ measurable and pairwise disjoint, whose union is $\boldsymbol{A}$, and coefficients $\alpha_1, \alpha_2, \ldots, \alpha_p \in \mathbb{R}$ such that for all $\boldsymbol{x} \in A$ we have

$$f(\boldsymbol{x}) = \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}).$$

We have

$$(\lambda f)(\boldsymbol{x}) = \lambda f(\boldsymbol{x}) = \lambda \sum_{k=1}^{p} \alpha_k \chi_{\boldsymbol{A}_k}(\boldsymbol{x}) = \sum_{k=1}^{p} (\lambda \alpha_k) \chi_{\boldsymbol{A}_k}(\boldsymbol{x}),$$

which shows that λf is simple. ■

Solution 3.28 *(Exercise 3.30).* We prove $\Rightarrow$. For all $k \in \mathbb{N}^*$ we have,

$$t - \frac{1}{k} < t \leq a,$$

so that $a > t - \frac{1}{k}$ for all $k \in \mathbb{N}^*$. We prove $\Leftarrow$. Suppose that for all $k \in \mathbb{N}^*$ we have $a > t - \frac{1}{k}$. We cannot have $a < t$, because in that case we would have $a = t - c$, for a certain real quantity $c > 0$, and there exists $k \in \mathbb{N}^*$ such that $k > 1/c$ (Archimede's axiom for real numbers: given any real number, there exists a natural number greater than it), so that $c > 1/k$, and therefore $a = t - c < t - \frac{1}{k}$ against our assumption. Any choice $a \geq t$ is consistent with $a > t - \frac{1}{k}$, since for all $k \in \mathbb{N}^*$ we have $a \geq t > t - \frac{1}{k}$. ■

Solution 3.29 *(Exercise 3.31).* Let $y = k\,x$. Then

$$\frac{1}{k} \int_{1/k}^{+\infty} \frac{\sin x}{x^2}\,\mathrm{d}x = \frac{1}{k} \int_{1}^{+\infty} \frac{\sin\left(\frac{y}{k}\right)}{\frac{y^2}{k^2}} \frac{1}{k}\,\mathrm{d}y = \int_{1}^{+\infty} \frac{\sin\left(\frac{y}{k}\right)}{y^2}\,\mathrm{d}y.$$

We can bring the limit inside because

$$\left| \frac{\sin\left(\frac{y}{k}\right)}{y^2} \right| \leq \left| \frac{1}{y^2} \right|,$$

which is summable over $[1, +\infty)$. Since the pointwise limit is 0, the value of the integral is 0. ■

Solution 3.30 *(Exercise 3.32).*

$$\begin{aligned}\lim_{k\to\infty}\int_0^k \frac{\sin x}{x}\,\mathrm{d}x &= \lim_{k\to\infty}\int_0^k \sin x\left(\int_0^{+\infty} e^{-x\,t}\,\mathrm{d}t\right)\mathrm{d}x\\ &= \lim_{k\to\infty}\int_0^k\int_0^{+\infty} \sin x\, e^{-x\,t}\,\mathrm{d}t\,\mathrm{d}x\\ &= \lim_{k\to\infty}\int_0^{+\infty}\int_0^{+\infty} (\sin x)\, e^{-x\,t}\,\chi_{[0,k]}(x)\,\mathrm{d}t\,\mathrm{d}x.\end{aligned}$$

The integrand is summable, since

$$\left|(\sin x)\, e^{-x\,t}\,\chi_{[0,k]}(x)\right| \le e^{-x\,t}$$

is summable. Then

$$\lim_{k\to\infty}\int_0^k \frac{\sin x}{x}\,\mathrm{d}x = \lim_{k\to\infty}\int_0^{+\infty}\left(\int_0^k \sin x\, e^{-x\,t}\,\mathrm{d}x\right)\mathrm{d}t.$$

We can integrate by parts:

$$\begin{aligned}\int (\sin x)\, e^{-t\,x}\,\mathrm{d}t &= -\frac{1}{t}\,\sin x\, e^{-t\,x} - \int (\cos x)\left(-\frac{1}{t}\,e^{-t\,x}\right)\mathrm{d}t\\ &= -\frac{1}{t}\,\sin x\, e^{-t\,x} - \frac{1}{t^2}\,(\cos x)\, e^{-t\,x} - \frac{1}{t^2}\int (\sin x)\, e^{-t\,x}\,\mathrm{d}x.\end{aligned}$$

Hence

$$\left(1+\frac{1}{t^2}\right)\int (\sin x)\, e^{-t\,x}\,\mathrm{d}t = -\frac{1}{t}\,\sin x\, e^{-t\,x} - \frac{1}{t^2}\,(\cos x)\, e^{-t\,x}$$

and so

$$\int (\sin x)\, e^{-t\,x}\,\mathrm{d}t = -\frac{t\,\sin x + \cos x}{t^2+1}\, e^{-t\,x}.$$

We obtain

$$\begin{aligned}\lim_{k\to\infty}\int_0^k \frac{\sin x}{x}\,\mathrm{d}x &= \lim_{k\to\infty}\int_0^{+\infty}\left(\int_0^k \sin x\, e^{-x\,t}\,\mathrm{d}x\right)\mathrm{d}t\\ &= \lim_{k\to\infty}\int_0^{+\infty}\left[-\frac{t\,\sin x + \cos x}{t^2+1}\, e^{-t\,x}\right]_{x=0}^{x=k}\mathrm{d}t\\ &= \lim_{k\to\infty}\int_0^{+\infty}\left(-\frac{t\,\sin k + \cos k}{t^2+1}\, e^{-k\,t} + \frac{1}{t^2+1}\right)\mathrm{d}t\\ &= \lim_{k\to\infty}\int_0^{+\infty} -\frac{(t\,\sin k + \cos k)\, e^{-k\,t} - 1}{t^2+1}\,\mathrm{d}t\end{aligned}$$

For all $t > 0$

$$\lim_{k\to\infty} -\frac{(t\,\sin k + \cos k)\, e^{-k\,t} - 1}{t^2+1} = \frac{1}{1+t^2}.$$

We want to apply Lebesgue dominated convergence theorem (Theorem 3.47), but we cannot use the triangle inequality twice, because we obtain a term that goes to zero as $\frac{t}{t^2} = \frac{1}{t}$, which is not integrable

at ∞. However, since $k > 0$ and $t \geq 0$, we have $e^{k\,t} \geq k\,t$ (for example from the Taylor expansion of the exponential). Thus, $e^{-k\,t} \leq \frac{1}{k\,t} \leq \frac{1}{t}$. It follows that

$$|t\,\sin k + \cos k|\,e^{-k\,t} \leq |t\,\sin k|\,e^{-k\,t} + |\cos k|\,e^{-k\,t} \leq t\,\frac{1}{t} + 1 \cdot 1 \leq 2,$$

and now we can use the triangle inequality:

$$\left|-\frac{(t\,\sin k + \cos k)\,e^{-k\,t} - 1}{t^2+1}\right| \leq \left|\frac{(t\,\sin k + \cos k)\,e^{-k\,t}}{t^2+1}\right| + \left|\frac{1}{t^2+1}\right| \leq \frac{2}{t^2+1} + \frac{1}{t^2+1},$$

which does not depend on k and is summable. We can apply Lebesgue dominated convergence theorem (Theorem 3.47):

$$\lim_{k\to\infty} \int_0^k \frac{\sin x}{x}\,\mathrm{d}x = \int_0^{+\infty} \frac{1}{1+t^2}\,\mathrm{d}t = \frac{\pi}{2}.$$

■

Solution 3.31 *(Exercise 3.33).* We can write

$$\frac{1}{2\,k}\int_{-k}^{k} f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \underbrace{\frac{1}{2\,k}\,f(x)\,\chi_{[-k,k]}(x)}_{f_k(x)}\,\mathrm{d}x.$$

The pointwise limit of the integrand is 0. Moreover,

$$|f_k(x)| \leq \frac{1}{2}\,|f(x)|,$$

which is summable, so that we can use Lebesgue dominated convergence theorem (Theorem 3.47):

$$\frac{1}{2\,k}\int_{-k}^{k} f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \lim_{k\to\infty} f_k(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} 0\,\mathrm{d}x = 0.$$

■

Solution 3.32 *(Exercise 3.34).* First, for every $x \in [0,a]$ we have

$$\lim_{n\to\infty} \sin\left(1 - e^{-\frac{x}{n}}\right) e^{-\frac{x}{n}} = \sin 0 = 0.$$

Therefore, the pointwise limit of the integrand function is 0. We apply Lebesgue dominated convergence theorem (Theorem 3.47): for $x \in [0,a]$ we have

$$\left|\sin\left(1 - e^{-\frac{x}{n}}\right) e^{-\frac{x}{n}}\right| \leq e^{-\frac{x}{n}} \leq 1,$$

which is summable in $[0,a]$. We conclude that

$$\lim_{n\to\infty} \int_0^a \sin\left(1 - e^{-\frac{x}{n}}\right) e^{-\frac{x}{n}}\,\mathrm{d}x = \int_0^a \lim_{n\to\infty} \sin\left(1 - e^{-\frac{x}{n}}\right) e^{-\frac{x}{n}}\,\mathrm{d}x = \int_0^a 0\,\mathrm{d}x = 0.$$

■

13 Solutions to the Exercises in Chapter 4

Solution 4.1 *(Exercise 4.1).* The first two are trivial. Since f, g, h are defined almost everywhere in A, there exist three zero measure sets $N_1, N_2, N_3 \subseteq A$ such that $f : A \setminus N_1 \longrightarrow \mathbb{C}$, $g : A \setminus N_2 \longrightarrow \mathbb{C}$ and $h : A \setminus N_3 \longrightarrow \mathbb{C}$. Now, we define

$$N_A = \{\boldsymbol{x} \in A \setminus (N_1 \cup N_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\} \quad \text{and} \quad N_B = \{\boldsymbol{x} \in A \setminus (N_2 \cup N_3) \,|\, g(\boldsymbol{x}) \neq h(\boldsymbol{x})\}.$$

Both these sets have Lebesgue measure zero because $f(\boldsymbol{x}) = g(\boldsymbol{x})$ almost everywhere in A and $g(\boldsymbol{x}) = h(\boldsymbol{x})$ almost everywhere in A. Their union is also a Lebesgue measure zero set. If $\boldsymbol{x} \notin (N_A \cup N_B) \cup (N_1 \cup N_2 \cup N_3)$, all three functions are defined at $\boldsymbol{x}$, and $f(\boldsymbol{x}) = g(\boldsymbol{x})$ and $g(\boldsymbol{x}) = h(\boldsymbol{x})$ hold simultaneously. Thus, if $\boldsymbol{x} \notin (N_A \cup N_B) \cup (N_1 \cup N_2 \cup N_3)$, we have $f(\boldsymbol{x}) = h(\boldsymbol{x})$. Since $(N_A \cup N_B) \cup (N_1 \cup N_2 \cup N_3)$ has measure zero, we deduce that $f = h$ almost everywhere. ■

Solution 4.2 *(Exercise 4.2).* Since f_1, f_2, g_1, g_2 be defined almost everywhere in A, there exist $N_1, N_2, N_3, N_4 \subset \mathbb{R}^n$, all of measure zero, such that

$$\begin{aligned} f_1 &: A \setminus N_1 \longrightarrow \mathbb{C} \\ f_2 &: A \setminus N_2 \longrightarrow \mathbb{C} \\ g_1 &: A \setminus N_3 \longrightarrow \mathbb{C} \\ g_2 &: A \setminus N_4 \longrightarrow \mathbb{C}. \end{aligned}$$

Moreover, we define the sets

$$\begin{aligned} N_a &= \{\boldsymbol{x} \in A \setminus (N_1 \cup N_3) \,|\, f_1(\boldsymbol{x}) \neq g_1(\boldsymbol{x})\} \\ N_b &= \{\boldsymbol{x} \in A \setminus (N_2 \cup N_4) \,|\, f_2(\boldsymbol{x}) \neq g_2(\boldsymbol{x})\}, \end{aligned}$$

which both have measure zero by assumption. We show that

$$\underbrace{\{\boldsymbol{x} \in A \setminus (N_1 \cup N_2 \cup N_3 \cup N_4) \,|\, f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) \neq g_1(\boldsymbol{x}) + g_2(\boldsymbol{x})\}}_{H} \subseteq N_a \cup N_b.$$

In fact, if $\boldsymbol{x} \in A \setminus (N_1 \cup N_2 \cup N_3 \cup N_4)$, then $\boldsymbol{x} \in A \setminus (N_1 \cup N_3)$ and $\boldsymbol{x} \in A \setminus (N_2 \cup N_4)$. Moreover, if $f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) \neq g_1(\boldsymbol{x}) + g_2(\boldsymbol{x})$, then $f_1(\boldsymbol{x}) \neq g_1(\boldsymbol{x})$ or $f_2(\boldsymbol{x}) \neq g_2(\boldsymbol{x})$.

Since $N_a \cup N_b$ has measure zero (it is the union of Lebesgue null sets), we find that the set H has also measure zero, so that $(f_1 + f_2) \sim (g_1 + g_2)$.

■

Solution 4.3 *(Exercise 4.3).* Let $N_1, N_2 \subseteq \mathbb{R}^n$ be null sets such that

$$\begin{aligned} f &: A \setminus N_1 \longrightarrow \mathbb{C} \\ g &: A \setminus N_2 \longrightarrow \mathbb{C}. \end{aligned}$$

If $c = 0$, the statement is obvious: for all $\boldsymbol{x} \in A \setminus (N_1 \cup N_2)$ (and $N_1 \cup N_2$ is a null set), we have $c\, f(\boldsymbol{x}) = 0 = c\, g(\boldsymbol{x})$. Thus, $c\, f \sim c\, g$. If $c \neq 0$, we have

$$\{\boldsymbol{x} \in A \setminus (N_1 \cup N_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\} = \{\boldsymbol{x} \in A \setminus (N_1 \cup N_2) \,|\, c\, f(\boldsymbol{x}) \neq c\, g(\boldsymbol{x})\},$$

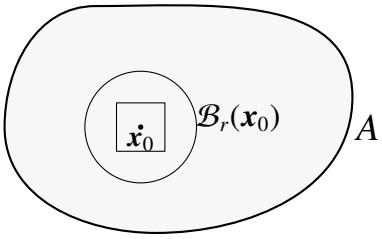
and the set in the left hand-side has measure zero, so the second also has measure zero. ■

Solution 4.4 *(Exercise 4.4).* Since A is open, there exists $r > 0$ such that $\mathcal{B}_r(\boldsymbol{x}) \subseteq A$. Then, k only needs to satisfy

$$\prod_{j=1}^{n} \left[x_{0,j} - \frac{1}{k}, x_{0,j} + \frac{1}{k} \right] \subseteq \mathcal{B}_r(\boldsymbol{x}_0).$$

Consider for example the case $n = 2$. Then, the latter becomes

$$\left[x_{0,1} - \frac{1}{k}, x_{0,1} + \frac{1}{k} \right] \times \left[x_{0,2} - \frac{1}{k}, x_{0,2} + \frac{1}{k} \right] \subseteq \left\{ (x_1, x_2) \in \mathbb{R}^2 \,\middle|\, \sqrt{(x_1 - x_{0,1})^2 + (x_2 - x_{0,2})^2} < r \right\}.$$



For this to hold, it is necessary and sufficient that all the vertexes of the square be contained in $\mathcal{B}_r(x_{0,1}, x_{0,2})$, i.e.,

$$\left(x_{0,1} - \frac{1}{k}, x_{0,2} - \frac{1}{k} \right), \left(x_{0,1} - \frac{1}{k}, x_{0,2} + \frac{1}{k} \right), \left(x_{0,1} + \frac{1}{k}, x_{0,2} - \frac{1}{k} \right), \left(x_{0,1} + \frac{1}{k}, x_{0,2} + \frac{1}{k} \right) \in \mathcal{B}_r(x_{0,1}, x_{0,2}).$$

This is equivalent to imposing that just one point is in $\mathcal{B}_r(x_{0,1}, x_{0,2})$. Consider the first point. We must have

$$\sqrt{\left(x_{0,1} - \frac{1}{k} - x_{0,1} \right)^2 + \left(x_{0,2} - \frac{1}{k} - x_{0,2} \right)^2} < r$$

leads to

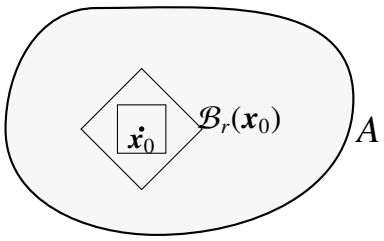
$$\frac{1}{k^2} + \frac{1}{k^2} < r^2 \quad \Rightarrow \quad k > \frac{\sqrt{2}}{r}.$$

In general we will have the condition

$$k > \frac{\sqrt{n}}{r}.$$

If, instead of the Euclidean norm, we use the 1-norm, then

$$\mathcal{B}_r(x_{0,1}, x_{0,2}) = \left\{ (x_1, x_2) \in \mathbb{R}^2 \,\middle|\, |x_1 - x_{0,1}| + |x_2 - x_{0,2}| < r \right\}.$$



In this case, again, the condition is only one vertex, and choosing the same vertex chosen before, we have

$$\left|x_{0,1} - \frac{1}{k} - x_{0,1}\right| + \left|x_{0,2} - \frac{1}{k} - x_{0,2}\right| < r,$$

which gives $k > \frac{2}{r}$. In the case of $\mathbb{R}^n$, this condition becomes $k > \frac{n}{r}$.

It is easy to see that in the case of the p-norm, we have the condition $k > \frac{n^{\frac{1}{p}}}{r}$, since

$$\left(\left|x_1 - x_{0,1}\right|^p + \left|x_2 - x_{0,2}\right|^p + \ldots + \left|x_n - x_{0,n}\right|^p\right)^{\frac{1}{p}} < r$$

which gives

$$\left(\frac{1}{k^p} + \ldots + \frac{1}{k^p}\right)^{\frac{1}{p}} < r$$

i.e.,

$$\frac{n}{k^p} < r^p \quad \Rightarrow \quad k > \frac{n^{\frac{1}{p}}}{r}.$$

In the case $p = +\infty$, this becomes

$$\max\left\{\frac{1}{k}, \ldots, \frac{1}{k}\right\} < r,$$

which gives $k > 1/r$.

■

Solution 4.5 *(Exercise 4.5).* Let the domains of f and g be $A \setminus N_1$ and $A \setminus N_2$, respectively, where N_1 and N_2 are null sets. Since f is equal almost everywhere to g, the set

$$N = \{\boldsymbol{x} \in A \setminus (N_1 \cup N_2) \,|\, f(\boldsymbol{x}) \neq g(\boldsymbol{x})\}$$

is a null set. The union $N \cup N_1 \cup N_2$ is also a null set, and for all $\boldsymbol{x} \in A \setminus (N \cup N_1 \cup N_2)$ both f and g are defined and they are equal, so that their real and imaginary parts are also equal. ■

Solution 4.6 *(Exercise 4.6).* We need to show that the operation of multiplication defined here does not depend on the particular representatives of the equivalence classes involved. In other words, given $f_1, f_2, g_1, g_2 \in L^p(A; \mathbb{C})$, we have to show that

$$f_1 \sim f_2 \;\text{ and }\; g_1 \sim g_2 \qquad \Rightarrow \qquad f_1\, g_1 \sim f_2\, g_2.$$

Let the domains of f_1, f_2, g_1, g_2 be $A \setminus N_1$, $A \setminus N_2$, $A \setminus N_3$, $A \setminus N_4$, respectively, where N_1, N_2, N_3, N_4 are null sets. Let

$$\begin{aligned} N_5 &= \{\boldsymbol{x} \in A \setminus (N_1 \cup N_2) \,|\, f_1(\boldsymbol{x}) \neq f_2(\boldsymbol{x}_2)\} \\ N_6 &= \{\boldsymbol{x} \in A \setminus (N_3 \cup N_4) \,|\, g_1(\boldsymbol{x}) \neq g_2(\boldsymbol{x}_2)\}. \end{aligned}$$

For all $\boldsymbol{x} \in A \setminus (N_1 \cup N_2 \cup N_3 \cup N_4 \cup N_5 \cup N_6)$, we have $f_1(\boldsymbol{x}) = f_2(\boldsymbol{x})$ and $g_1(\boldsymbol{x}) = g_2(\boldsymbol{x})$, so that $f_1(\boldsymbol{x})\, g_1(\boldsymbol{x}) = f_2(\boldsymbol{x})\, g_2(\boldsymbol{x})$. Since $N_1 \cup N_2 \cup N_3 \cup N_4 \cup N_5 \cup N_6$ is a null set, the claim is proved. ■

Solution 4.7 *(Exercise 4.7).* We have

$$u \notin L^1((0,+\infty),\mathbb{R}), \qquad u \in L^2((0,+\infty),\mathbb{R}), \qquad u \notin L^\infty((0,+\infty),\mathbb{R}).$$

■

Solution 4.8 *(Exercise 4.8).*

(i) We have

- $\max\limits_{x\in\mathbb{R}} f = \sup\limits_{x\in\mathbb{R}} f = 4$;
- $\min\limits_{x\in\mathbb{R}} f = \inf\limits_{x\in\mathbb{R}} f = -5$;
- $\operatorname{ess\,sup} f = 3$;
- $\operatorname{ess\,inf} f = 3$.

(ii) The function is unbounded from above and from below, and therefore the maximum and minimum are not defined, while

- $\sup\limits_{x\in\mathbb{R}} f = +\infty$;
- $\inf\limits_{x\in\mathbb{R}} f = -\infty$.

Since $\mathbb{Q}$ is of zero Lebesgue measure, what happens in $\mathbb{Q}$ does not matter. Thus,

- $\operatorname{ess\,sup} f = +\pi/2$;
- $\operatorname{ess\,inf} f = -\pi/2$.

(iii) The function is again unbounded, so maximum and minimum are not defined. We have

- $\sup\limits_{x\in\mathbb{R}} f = \operatorname{ess\,sup} f = +\infty$;
- $\inf\limits_{x\in\mathbb{R}} f = \operatorname{ess\,inf} f = -\infty$.

■

Solution 4.9 *(Exercise 4.9).* Since $(u_k)_{k\in\mathbb{N}}$ and $(v_k)_{k\in\mathbb{N}}$ are sequences of $L^2(A;\mathbb{C})$, the integrals

$$\int_A |u_k|^2 \,\mathrm{d}\boldsymbol{x} \quad \text{and} \quad \int_A |v_k|^2 \,\mathrm{d}\boldsymbol{x}$$

exist. Moreover, by assumption $u_k \overset{k\to\infty}{\longrightarrow} u$ and $v_k \overset{k\to\infty}{\longrightarrow} v$ in $L^2(A;\mathbb{C})$. Thus,

$$\begin{aligned}\int_A \left|u_k\,v_k - u\,v\right| \mathrm{d}\boldsymbol{x} &= \int_A \left|u_k\,v_k - u_k\,v + u_k\,v - u\,v\right| \mathrm{d}\boldsymbol{x} \le \int_A |u_k\,(v_k - v)|\,\mathrm{d}\boldsymbol{x} + \int_A |v\,(u_k - u)|\,\mathrm{d}\boldsymbol{x} \\ &\le \left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \cdot \left(\int_A |v_k - v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} + \left(\int_A |v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \cdot \left(\int_A |u_k - u|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \overset{k\to\infty}{\longrightarrow} 0\end{aligned}$$

since $\left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12}$ and $\left(\int_A |v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12}$ are bounded and the integrals $\left(\int_A |v_k - v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12}$ and $\left(\int_A |u_k - u|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12}$ tend to 0 as $k \to \infty$. In the latter, use used Hölder's inequality with conjugate exponents both equal to 2:

$$\begin{aligned}\int_A |u_k|\,|v_k - v|\,\mathrm{d}\boldsymbol{x} &= \|u_k\,(v_k - v)\|_1 \le \|u_k\|_2\,\|v_k - v\|_2 = \left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \cdot \left(\int_A |v_k - v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \\ \int_A |v|\,|u_k - u|\,\mathrm{d}\boldsymbol{x} &= \|v\,(u_k - u)\|_1 \le \|v\|_2\,\|u_k - u\|_2 = \left(\int_A |v|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12} \cdot \left(\int_A |u_k - u|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac12}.\end{aligned}$$

We prove the second point. Since $(u_k)_{k\in\mathbb{N}}$ converges to 0 in $L^2(A,\mathbb{C})$, using again Hölder's inequality, we find

$$\int_A |u_k\, v_k|\,\mathrm{d}\boldsymbol{x} = \|u_k\, v_k\|_1 \le \|u_k\|_2\, \|v_k\|_2 = \left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \cdot \left(\int_A |v_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \le \left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \cdot \left(\sup_A |v_k|^2 \cdot \mu(A)\right)^{\frac{1}{2}} \overset{k\to\infty}{\longrightarrow} 0$$

since $\left(\sup_A |v_k|^2 \cdot \mu(A)\right)^{\frac{1}{2}}$ is bounded, and $\left(\int_A |u_k|^2\,\mathrm{d}\boldsymbol{x}\right)^{\frac{1}{2}} \overset{k\to\infty}{\longrightarrow} 0$. ■

Solution 4.10 *(Exercise 4.10).* Consider the following cases:

- For $p = 1$ we find

$$\|u_k\|_1 = \int_{\mathbb{R}} |u_k|\,\mathrm{d}x = \int_{\mathbb{R}} \left|k^\lambda \chi_{[0,\frac{1}{k}]}\right|\,\mathrm{d}x = \int_0^{1/k} |k^\lambda|\,\mathrm{d}x = k^{\lambda-1},$$

  which converges to zero as $k \to \infty$ for $\lambda < 1$.

- Let $p = 2$. Then,

$$\|u_k\|_2 = \int_{\mathbb{R}} |u_k|^2\,\mathrm{d}x = \int_{\mathbb{R}} \left|k^\lambda \chi_{[0,\frac{1}{k}]}\right|^2\,\mathrm{d}x = \int_0^{1/k} |k^{2\,\lambda}|\,\mathrm{d}x = k^{\frac{2\lambda-1}{2}},$$

  which converges to zero as $k \to \infty$ for $\frac{2\lambda-1}{2} < 0$, i.e., for $\lambda < \frac{1}{2}$.

- Finally, consider $p = +\infty$. By definition $\|u_k\|_\infty = \inf\{M \ge 0\,|\,|u_k| \le M \text{ a.e.}\}$. If $\lambda > 0$, the infimum does not exist. If $\lambda = 0$, the infimum is equal to 1. If $\lambda < 0$, the infimum is 1. In none of these cases the infimum is 0, so that there are no values of λ for which $\|u_k\|_p \overset{k\to\infty}{\longrightarrow} 0$.

■

Solution 4.11 *(Exercise 4.11).*

(i) We have $e^{-k|x|} \overset{k\to\infty}{\longrightarrow} 0$ almost everywhere in $\mathbb{R}$ since the set of $x \in \mathbb{R}$ such that $e^{-k|x|}$ does not converge to 0 as $k \to \infty$ is $\{0\}$, which has measure zero. We now establish the convergence in $L^p(\mathbb{R},\mathbb{R})$ for

- $p = 1$: the sequence $(e^{-k|x|})_{k\in\mathbb{N}}$ converges in $L^1(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} \left|e^{-k|x|}\right|\,\mathrm{d}x = 2\int_0^{+\infty} e^{-k\,x}\,\mathrm{d}x = 2\left[\frac{e^{-k\,x}}{-k}\right]_0^{+\infty} = \frac{2}{k} \overset{k\to\infty}{\longrightarrow} 0.$$

- $p = 2$: the sequence $(e^{-k|x|})_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} \left|e^{-k|x|}\right|^2\,\mathrm{d}x = 2\int_0^{+\infty} e^{-2\,k\,x}\,\mathrm{d}x = 2\left[\frac{e^{-2\,k\,x}}{-2\,k}\right]_0^{+\infty} = \frac{1}{k} \overset{k\to\infty}{\longrightarrow} 0.$$

- $p = +\infty$: the sequence $(e^{-k|x|})_{k\in\mathbb{N}}$ does not converge in $L^\infty(\mathbb{R},\mathbb{R})$ because it does not converge uniformly to the zero function in $\mathbb{R}$ (we have $\sup_{x\in\mathbb{R}} e^{-k|x|} = 1$ for all $k \in \mathbb{N}$).

(ii) We now establish the convergence of the function $e^{-|x-k|}$ in $L^p(\mathbb{R},\mathbb{R})$ in the case

- $p = 1$: We have

$$\begin{aligned}\int_{\mathbb{R}} e^{-|x-k|}\,\mathrm{d}x &= \int_{-\infty}^{k} e^{x-k}\,\mathrm{d}x + \int_{k}^{\infty} e^{-(x-k)}\,\mathrm{d}x \\ &= e^{-k}\int_{-\infty}^{k} e^{x}\,\mathrm{d}x + e^{k}\int_{k}^{\infty} e^{-x}\,\mathrm{d}x \\ &= \lim_{r\to-\infty} e^{-k}\,[e^{x}]_{r}^{k} + \lim_{s\to\infty} e^{k}\,[-e^{-x}]_{k}^{s} \\ &= \lim_{r\to-\infty} e^{-k}\,(e^{k}-e^{r}) + \lim_{s\to\infty} e^{k}\,(-e^{-s}+e^{-k}) = e^{-k}\,e^{k} + e^{k}\,e^{-k} = 1+1 = 2.\end{aligned}$$

  Thus, we do not have convergence in $L^1(\mathbb{R},\mathbb{R})$.

- $p = 2$: we proceed as in the previous case:

$$\begin{aligned}\int_{\mathbb{R}} e^{-2\,|x-k|}\,\mathrm{d}x &= \int_{-\infty}^{k} e^{2\,(x-k)}\,\mathrm{d}x + \int_{k}^{\infty} e^{-2\,(x-k)}\,\mathrm{d}x \\ &= e^{-2\,k}\int_{-\infty}^{k} e^{2\,x}\,\mathrm{d}x + e^{2\,k}\int_{k}^{\infty} e^{-2\,x}\,\mathrm{d}x \\ &= \lim_{r\to-\infty} e^{-2\,k}\left[\frac{1}{2}\,e^{2\,x}\right]_{r}^{k} + \lim_{s\to\infty} e^{2\,k}\left[-\frac{1}{2}\,e^{-2\,x}\right]_{k}^{s} \\ &= e^{-2\,k}\left(\frac{1}{2}\,e^{2\,k}-0\right) + e^{2\,k}\left(0+\frac{1}{2}\,e^{-2\,k}-0\right) = \frac{1}{2}+\frac{1}{2} = 1.\end{aligned}$$

- $p = +\infty$: the sequence $(e^{-|x-k|})_{k\in\mathbb{N}}$ does not converge in $L^\infty(\mathbb{R},\mathbb{R})$. Indeed, for every $k \in \mathbb{N}$ we have

$$\sup_{x\in\mathbb{R}} \left|e^{-|x-k|}\right| = 1.$$

  Thus, we do not uniform convergence to the zero function (whereas the pointwise limit is the zero function).

(iii) We have $k^{1/2}\,e^{-k^2\,|x|} \overset{k\to\infty}{\longrightarrow} 0$ almost everywhere in $\mathbb{R}$ since the set of $x \in \mathbb{R}$ such that $k^{1/2}\,e^{-k^2\,|x|}$ does not converge to 0 as $k \to \infty$ is $\{0\}$, which has measure zero. We now establish the convergence in $L^p(\mathbb{R},\mathbb{R})$ for

- $p = 1$: the sequence $(k^{1/2}\,e^{-k^2\,|x|})_{k\in\mathbb{N}}$ converges in $L^1(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} k^{1/2}\,e^{-k^2\,|x|}\,\mathrm{d}x = 2\int_{0}^{+\infty} k^{1/2}\,e^{-k^2\,x}\,\mathrm{d}x = 2\,k^{1/2}\left[\frac{e^{-k^2\,x}}{-k^2}\right]_{0}^{+\infty} = 2\,\frac{k^{1/2}}{k^2} = \frac{2}{k^{3/2}} \overset{k\to\infty}{\longrightarrow} 0.$$

- $p = 2$: the sequence $(k^{1/2}\,e^{-k^2\,|x|})_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R},\mathbb{R})$:

$$\int_{\mathbb{R}} k\,e^{-2\,k^2\,|x|}\,\mathrm{d}x = 2\int_{0}^{+\infty} k\,e^{-2\,k^2\,x}\,\mathrm{d}x = 2\,k\left[-\frac{e^{-2\,k^2\,x}}{2\,k^2}\right]_{0}^{+\infty} = \frac{2\,k}{2\,k^2} = \frac{1}{k} \overset{k\to\infty}{\longrightarrow} 0.$$

- $p = +\infty$: the sequence $(k^{1/2}\,e^{-k^2\,|x|})_{k\in\mathbb{N}}$ does not converge in $L^\infty(\mathbb{R},\mathbb{R})$.

(iv) We have $e^{-k\,|x-k|} \le e^{-k\,|x|}\,e^{-k^2} \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We now establish the convergence in $L^p(\mathbb{R},\mathbb{R})$ for

- $p = 1$: the sequence $(e^{-k|x-k|})_{k\in\mathbb{N}}$ converges in $L^1(\mathbb{R},\mathbb{R})$ since

$$0 < \int_{\mathbb{R}} e^{-k|x-k|}\,\mathrm{d}x \le \int_{\mathbb{R}} e^{-k^2}\,e^{-k|x|}\,\mathrm{d}x = \int_0^{+\infty} e^{-k^2}\,e^{-kx}\,\mathrm{d}x = e^{-k^2}\left[\frac{e^{-kx}}{-k}\right]_0^{+\infty} = \frac{e^{-k^2}}{k} \overset{k\to\infty}{\longrightarrow} 0,$$

or, directly, for $k \neq 0$

$$\begin{aligned}\int_{\mathbb{R}} e^{-k|x-k|}\,\mathrm{d}x &= \int_{-\infty}^{k} e^{k(x-k)}\,\mathrm{d}x + \int_k^{+\infty} e^{-k(x-k)}\,\mathrm{d}x\\ &= e^{-k^2}\int_{-\infty}^{k} e^{kx}\,\mathrm{d}x + e^{k^2}\int_k^{+\infty} e^{-kx}\,\mathrm{d}x\\ &= e^{-k^2}\left(\frac{1}{k}\,e^{k^2} - 0\right) + e^{k^2}\left(0 + \frac{1}{k}\,e^{-k^2}\right) = \frac{1}{k} + \frac{1}{k} = \frac{2}{k} \overset{k\to\infty}{\longrightarrow} 0.\end{aligned}$$

- $p = 2$: using the same argument

$$\begin{aligned}\int_{\mathbb{R}} e^{-2k|x-k|}\,\mathrm{d}x &= \int_{-\infty}^{k} e^{2k(x-k)}\,\mathrm{d}x + \int_k^{+\infty} e^{-2k(x-k)}\,\mathrm{d}x\\ &= e^{-2k^2}\int_{-\infty}^{k} e^{2kx}\,\mathrm{d}x + e^{2k^2}\int_k^{+\infty} e^{-2kx}\,\mathrm{d}x\\ &= e^{-2k^2}\left(\frac{1}{2k}\,2\,e^{k^2} - 0\right) + e^{2k^2}\left(0 + \frac{1}{2k}\,e^{-2k^2}\right) = \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k} \overset{k\to\infty}{\longrightarrow} 0.\end{aligned}$$

Therefore, the sequence $(e^{-k|x-k|})_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R},\mathbb{R})$.

- $p = +\infty$: the sequence $(e^{-k|x-k|})_{k\in\mathbb{N}}$ does not converge in $L^\infty(\mathbb{R},\mathbb{R})$ since $\sup_{x\in\mathbb{R}} e^{-k|x-k|} = 1$ for all $k \in \mathbb{N}$.

(v) We have $\frac{1}{k}\,e^{-\frac{|x|}{k}} \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We now establish the convergence in $L^p(\mathbb{R},\mathbb{R})$ for

- $p = 1$: the sequence $(\frac{1}{k}\,e^{-\frac{|x|}{k}})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} \frac{1}{k}\,e^{-\frac{|x|}{k}}\,\mathrm{d}x = 2\int_0^{+\infty} \frac{1}{k}\,e^{-\frac{x}{k}}\,\mathrm{d}x = \frac{2}{k}\left[\frac{e^{-\frac{x}{k}}}{-\frac{1}{k}}\right]_0^{+\infty} = \frac{2}{k}\,k = 2;$$

- $p = 2$: the sequence $(\frac{1}{k}\,e^{-\frac{|x|}{k}})_{k\in\mathbb{N}}$ converges in $L^2(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} \frac{1}{k^2}\,e^{-2\frac{|x|}{k}}\,\mathrm{d}x = 2\int_0^{+\infty} \frac{1}{k^2}\,e^{-\frac{2x}{k}}\,\mathrm{d}x = \frac{2}{k^2}\left[\frac{e^{-\frac{2x}{k}}}{-\frac{2}{k}}\right]_0^{+\infty} = \frac{2}{k^2}\,\frac{1}{\frac{2}{k}} = \frac{1}{k} \overset{k\to\infty}{\longrightarrow} 0.$$

- $p = +\infty$: the sequence $(\frac{1}{k}\,e^{-\frac{|x|}{k}})_{k\in\mathbb{N}}$ converges in $L^\infty(\mathbb{R},\mathbb{R})$ because we have uniform convergence to the identically zero function: indeed, $\sup_{x\in\mathbb{R}} \frac{1}{k}\,e^{-\frac{|x|}{k}} = \frac{1}{k}$, which is infinitesimal.

(vi) We have $k\,e^{-k|x|} \overset{k\to\infty}{\longrightarrow} 0$ almost everywhere in $\mathbb{R}$ since the set of $x \in \mathbb{R}$ such that $k\,e^{-k|x|}$ does not converge to 0 as $k \to \infty$ is $\{0\}$, which has measure zero. We now establish the convergence in $L^p(\mathbb{R},\mathbb{R})$ for

- $p = 1$: the sequence $(k\,e^{-k\,|x|})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} k\,e^{-k\,|x|}\,\mathrm{d}x = 2\int_0^{+\infty} k\,e^{-k\,x}\,\mathrm{d}x = 2\,k\left[\frac{e^{-k\,x}}{-k}\right]_0^{+\infty} = 2.$$

- $p = 2$: the sequence $(k\,e^{-k\,|x|})_{k\in\mathbb{N}}$ does not converge in $L^2(\mathbb{R},\mathbb{R})$ since

$$\int_{\mathbb{R}} k^2\,e^{-2\,k\,|x|}\,\mathrm{d}x = 2\int_0^{+\infty} k^2\,e^{-2\,k\,x}\,\mathrm{d}x = k^2\left[\frac{e^{-2\,k\,x}}{-2\,k}\right]_0^{+\infty} = 2\,k^2\,\frac{1}{2\,k} = k \overset{k\to\infty}{\longrightarrow} +\infty.$$

- $p = +\infty$: the sequence $(k\,e^{-k\,|x|})_{k\in\mathbb{N}}$ does not converge in $L^\infty(\mathbb{R},\mathbb{R})$ because it does not converge uniformly in $\mathbb{R}$. In fact,

$$\sup_{x\in\mathbb{R}} |k\,e^{-k\,|x|}| = k$$

which is not infinitesimal.

■

Solution 4.12 *(Exercise 4.12).*

(i) First, notice that $k^{-\frac{1}{2}}\,\chi_{[k,k+1]}(x) \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We find

$$\int_{\mathbb{R}} k^{-\frac{1}{2}}\,\chi_{[k,k+1]}(x)\,\mathrm{d}x = \int_k^{k+1} k^{-\frac{1}{2}}\,\mathrm{d}x = k^{-\frac{1}{2}} \overset{k\to\infty}{\longrightarrow} 0.$$

It follows that $(k^{-\frac{1}{2}}\,\chi_{[k,k+1]})_{k\in\mathbb{N}}$ converges in $L^1(\mathbb{R},\mathbb{R})$, $L^2(\mathbb{R},\mathbb{R})$ and $L^\infty(\mathbb{R},\mathbb{R})$;

(ii) We have $k^{-\frac{1}{2}}\,\chi_{[k,2\,k]}(x) \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We find

$$\int_{\mathbb{R}} k^{-\frac{1}{2}}\,\chi_{[k,2\,k]}(x)\,\mathrm{d}x = \int_k^{2\,k} k^{-\frac{1}{2}}\,\mathrm{d}x = k^{-\frac{1}{2}}\,k = k^{\frac{1}{2}} \overset{k\to\infty}{\longrightarrow} +\infty.$$

It follows that $(k^{-\frac{1}{2}}\,\chi_{[k,2\,k]})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$ nor in $L^2(\mathbb{R},\mathbb{R})$, but it converges in $L^\infty(\mathbb{R},\mathbb{R})$;

(iii) We have $k\,\chi_{[0,\frac{1}{k}]}(x) \overset{k\to\infty}{\longrightarrow} 0$ almost everywhere in $\mathbb{R}$. We find

$$\int_{\mathbb{R}} k\,\chi_{[0,\frac{1}{k}]}(x)\,\mathrm{d}x = \int_0^{1/k} k\,\mathrm{d}x = k\,\frac{1}{k} = 1.$$

It follows that $(k\,\chi_{[0,\frac{1}{k}]})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$, it does not converge in $L^2(\mathbb{R},\mathbb{R})$ and it does not converge in $L^\infty(\mathbb{R},\mathbb{R})$;

(iv) We have $k\,\chi_{[k,k+\frac{1}{k}]}(x) \overset{k\to\infty}{\longrightarrow} 0$ almost everywhere in $\mathbb{R}$. We find

$$\int_{\mathbb{R}} k\,\chi_{[k,k+\frac{1}{k}]}(x)\,\mathrm{d}x = \int_k^{1+\frac{1}{k}} k\,\mathrm{d}x = k\,\frac{1}{k} = 1.$$

It follows that $(k\,\chi_{[k,k+\frac{1}{k}]})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$, it does not converge in $L^2(\mathbb{R},\mathbb{R})$ and it does not converge in $L^\infty(\mathbb{R},\mathbb{R})$;

(v) We have $k^{-\frac{2}{3}}\chi_{[0,k]}(x) \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We find

$$\int_{\mathbb{R}} k^{-\frac{2}{3}}\chi_{[0,k]}(x)\,\mathrm{d}x = \int_0^k k^{-\frac{2}{3}}\,\mathrm{d}x = k\,k^{-\frac{2}{3}} = k^{\frac{1}{3}} \overset{k\to\infty}{\longrightarrow} +\infty.$$

It follows that $(k^{-\frac{2}{3}}\chi_{[0,k]})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$, but it converges in $L^2(\mathbb{R},\mathbb{R})$ and in $L^\infty(\mathbb{R},\mathbb{R})$;

(vi) We have $k^{-\frac{1}{2}}\chi_{[-k,k^2]}(x) \overset{k\to\infty}{\longrightarrow} 0$ for every $x \in \mathbb{R}$. We find

$$\int_{\mathbb{R}} k^{-\frac{1}{2}}\chi_{[-k,k^2]}(x)\,\mathrm{d}x = \int_{-k}^{k^2} k^{-\frac{1}{2}}\,\mathrm{d}x = k^{-\frac{1}{2}}\,(k^2+k) \overset{k\to\infty}{\longrightarrow} +\infty.$$

It follows that $(k^{-\frac{1}{2}}\chi_{[-k,k^2]})_{k\in\mathbb{N}}$ does not converge in $L^1(\mathbb{R},\mathbb{R})$ nor in $L^2(\mathbb{R},\mathbb{R})$, but it converges in $L^\infty(\mathbb{R},\mathbb{R})$.

■

Solution 4.13 *(Exercise 4.13).*

(i) $f(x) = c$ is not summable, unless $c = 0$. Indeed, if $c = 0$ the integral is zero, and if $c \neq 0$ the improper integral

$$\int_{\mathbb{R}} |c|\,\mathrm{d}x = |c| \int_{\mathbb{R}} \mathrm{d}x = |c|\,\mu_n(\mathbb{R})$$

is divergent;

(ii) The characteristic function $f(x) = \chi_{[a,b]}(x)$ is certainly summable, since

$$\int_{\mathbb{R}} |\chi_{[a,b]}(x)|\,\mathrm{d}x = \int_{\mathbb{R}} \chi_{[a,b]}(x)\,\mathrm{d}x = \int_a^b \mathrm{d}x = (b-a).$$

(iii) The function $f(x) = e^{-x}$ is not summable. Indeed, since $|e^{-x}| = e^{-x}$, we have

$$\begin{aligned}\int_{\mathbb{R}} e^{-x}\,\mathrm{d}x &\geq \int_{-\infty}^0 e^{-x}\,\mathrm{d}x \\ &= \lim_{c\to-\infty}\int_c^0 e^{-x}\,\mathrm{d}x = \lim_{c\to-\infty}[-e^{-x}]_{x=0}^{x=c} = \lim_{c\to-\infty}(1-e^{-c}),\end{aligned}$$

and the limit does not exist in $\mathbb{R}$, so that the function is not summable;

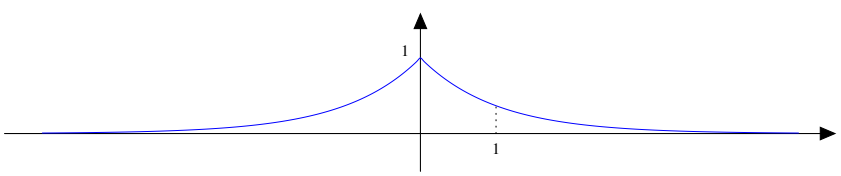
(iv) The function $f(x) = e^x$ is not summable. Indeed,

$$\int_{\mathbb{R}} e^x\,\mathrm{d}x \geq \int_0^{+\infty} e^x\,\mathrm{d}x = \lim_{c\to+\infty}[e^x]_{x=0}^{x=c} = \lim_{c\to+\infty}(e^c-1)$$

is divergent and the limit does not exist in $\mathbb{R}$.

(v) Consider the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto e^{-|x|}\end{aligned}$$



This function is summable, since

$$\int_{\mathbb{R}} |e^{-|x|}| \, \mathrm{d}x = \int_{\mathbb{R}} e^{-|x|} \, \mathrm{d}x = \int_{-\infty}^{0} e^{x} \, \mathrm{d}x + \int_{0}^{+\infty} e^{-x} \, \mathrm{d}x$$
$$= \lim_{a \to -\infty} [e^{x}]_{x=a}^{x=0} + \lim_{b \to +\infty} [-e^{-x}]_{x=0}^{x=b} = \lim_{a \to -\infty} (1 - e^{a}) + \lim_{b \to +\infty} (1 - e^{-b}) = 2.$$

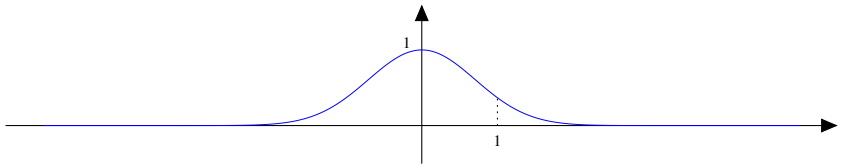
(vi) Consider the function $f(x) = e^{-x^2}$. We have

$$\int_{\mathbb{R}} |e^{-x^2}| \, \mathrm{d}x = \int_{-\infty}^{-1} e^{-x^2} \, \mathrm{d}x + \int_{-1}^{+1} e^{-x^2} \, \mathrm{d}x + \int_{1}^{+\infty} e^{-x^2} \, \mathrm{d}x.$$

The second integral in the right hand-side exists because the integrand is continuous over the closed interval $[-1, +1]$. Consider for example the last integral. For $x > 1$, we have $e^{-x^2} < e^{-x}$, so that

$$\int_{1}^{+\infty} e^{-x^2} \, \mathrm{d}x \leq \int_{1}^{+\infty} e^{-x} \, \mathrm{d}x,$$

and the integral in the right hand-side exists (see previous point). From the comparison test, the integral in the left hand-side also exists.



Alternatively, one can notice that this is the Gauss integral, whose value can be computed using polar coordinates and the polar coordinates $x = \rho \cos\theta$ and $y = \rho \sin\theta$, with $\rho \in \mathbb{R}_+$ and $\theta \in (0, 2\pi)$. We obtain, applying Fubini's Theorem

$$\int_{\mathbb{R}^2} e^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbb{R}_+ \times (0,2\pi)} e^{-\rho^2} \rho \, \mathrm{d}\rho \, \mathrm{d}\theta = 2\pi \int_{0}^{+\infty} e^{-\rho^2} \rho \, \mathrm{d}\rho = \frac{2\pi}{-2} \int_{0}^{+\infty} e^{-\rho^2} (-2\rho) \, \mathrm{d}\rho = -\pi \left[e^{-\rho^2} \right]_{0}^{+\infty} = \pi.$$

We now apply again Fubini's Theorem:

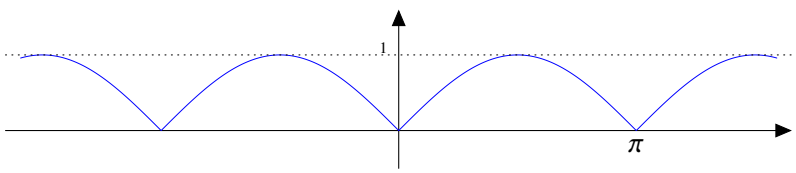
$$\pi = \int_{\mathbb{R}^2} e^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-y^2} \, \mathrm{d}y \right) e^{-x^2} \, \mathrm{d}x = \left(\int_{\mathbb{R}} e^{-x^2} \, \mathrm{d}x \right)^2.$$

The integral cannot take the value $-\sqrt{\pi}$ because the integrand is non-negative, and therefore the result is proved.

(vii) The integral

$$\int_{\mathbb{R}} |\sin x| \, \mathrm{d}x$$

does not converge.



We can see it for example in the following way:

$$\begin{aligned}\int_{\mathbb{R}} |\sin x|\,\mathrm{d}x &= \sum_{k=-\infty}^{+\infty} \left(\int_{2k\pi}^{(2k+1)\pi} \sin x\,\mathrm{d}x - \int_{(2k+1)\pi}^{(2k+2)\pi} \sin x\,\mathrm{d}x \right) \\ &= \sum_{k=-\infty}^{+\infty} \left([-\cos x]_{2k\pi}^{(2k+1)\pi} + [\cos x]_{(2k+1)\pi}^{(2k+2)\pi} \right) = \sum_{k=-\infty}^{+\infty} 4,\end{aligned}$$

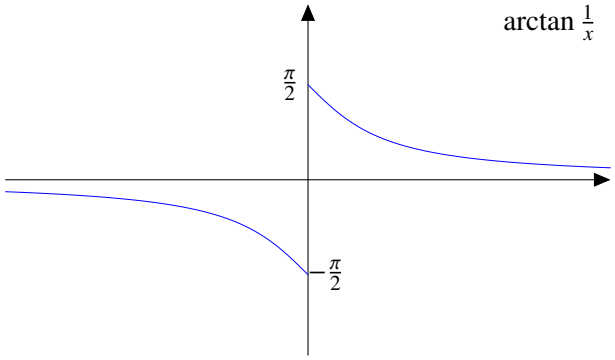
and the Laurent series does not converge.

■

Solution 4.14 *(Exercise 4.14).*

(i) First, observe that

$$\lim_{x\to 0^+} \arctan\frac{1}{x} = \frac{\pi}{2} \quad \text{and} \quad \lim_{x\to 0^-} \arctan\frac{1}{x} = -\frac{\pi}{2}.$$



We can find the primitive of $\arctan\frac{1}{x}$ by integrating by part:

$$\begin{aligned}\int 1\cdot \arctan\frac{1}{x}\,\mathrm{d}x &= x\arctan\frac{1}{x} - \int x\,\frac{1}{1+\frac{1}{x^2}}\left(-\frac{1}{x^2}\right)\mathrm{d}x \\ &= x\arctan\frac{1}{x} + \int \frac{x}{1+x^2}\,\mathrm{d}x \\ &= x\arctan\frac{1}{x} + \frac{1}{2}\int \frac{2x}{1+x^2}\,\mathrm{d}x \\ &= \arctan\frac{1}{x} + \frac{1}{2}\ln(x^2+1) + C,\end{aligned}$$

where C is a constant. Now, since $\arctan\frac{1}{x}$ is an odd function, we find

$$\int_{\mathbb{R}} \left|\arctan\frac{1}{x}\right| \mathrm{d}x = 2\int_0^{+\infty} \arctan\frac{1}{x}\,\mathrm{d}x = 2\lim_{b\to+\infty}\left[\arctan\frac{1}{x} + \frac{1}{2}\ln(x^2+1)\right]_{x=0}^{x=b},$$

which does not converge as $b\to+\infty$. Thus, the function is not in $L^1(\mathbb{R}\setminus\{0\};\mathbb{R})$. Another way of seeing this is to recall that, as $y\to 0$, the function $\arctan y$ is asymptotically equivalent to y, and therefore as $|x|\to\infty$ the function $\arctan\frac{1}{x}$ is asymptotically equivalent to $\frac{1}{x}$, which is not integrable at ∞.

(ii) Since we have established that e^{-x^2} is integrable when x is a real variable, we can apply Fubini's Theorem to decompose the integral:

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-\|\boldsymbol{x}\|^2}\,\mathrm{d}\boldsymbol{x} &= \int_{\mathbb{R}^n} e^{-x_1^2-x_2^2-\ldots-x_n^2}\,\mathrm{d}x_1\,\mathrm{d}x_2\ldots\,\mathrm{d}x_n\\ &= \int_{\mathbb{R}^n} e^{-x_1^2}\,e^{-x_2^2}\,\ldots\,e^{-x_n^2}\,\mathrm{d}x_1\,\mathrm{d}x_2\ldots\,\mathrm{d}x_n\\ &= \int_{\mathbb{R}} e^{-x_1^2}\,\mathrm{d}x_1\cdot\int_{\mathbb{R}} e^{-x_2^2}\,\mathrm{d}x_2\cdot\ldots\cdot\int_{\mathbb{R}} e^{-x_n^2}\,\mathrm{d}x_n\\ &= \left(\int_{\mathbb{R}} e^{-t^2}\,\mathrm{d}t\right)^n = (\sqrt{\pi})^n = \pi^{\frac{n}{2}}.\end{aligned}$$

Clearly, this implies that $f(\boldsymbol{x}) = e^{-\|\boldsymbol{x}\|^2}$ is in $L^1(\mathbb{R}^n;\mathbb{R})$.

■

Solution 4.15 *(Exercise 4.15).* We study the behavior at infinity of $\arctan\frac{1}{x}$. Since as $y\to 0$ the function $\arctan y$ is asymptotically equivalent to y, we find that as $|x|\to\infty$ the function $\arctan\frac{1}{x}$ is asymptotically equivalent to $\frac{1}{x}$. Therefore $|\arctan\frac{1}{x}^2|$ is asymptotically equivalent to $\frac{1}{x^2}$, which is integrable. We conclude that $\arctan\frac{1}{x}\in L^2(\mathbb{R}\setminus\{0\};\mathbb{R})$.

■

Solution 4.16 *(Exercise 4.16).* The first function is not in $L^1(\mathbb{R};\mathbb{R})$:

$$\int_1^{+\infty}\left|\frac{1}{x}\right|\,\mathrm{d}x = \int_1^{+\infty}\frac{1}{x}\,\mathrm{d}x$$

is a divergent integral. Indeed,

$$\int_1^{+\infty}\frac{1}{x}\,\mathrm{d}x = \lim_{c\to+\infty}\int_1^c\frac{1}{x}\,\mathrm{d}x = \lim_{c\to+\infty}[\ln x]_1^c = \lim_{c\to+\infty}\ln c = +\infty.$$

However, f is in $L^2(\mathbb{R};\mathbb{R})$, since

$$\int_1^{+\infty}\left|\frac{1}{x^2}\right|\,\mathrm{d}x = \int_1^{+\infty}\frac{1}{x^2}\,\mathrm{d}x = \lim_{c\to+\infty}\left[-\frac{1}{x}\right]_1^c = \lim_{c\to+\infty}\left(1-\frac{1}{c}\right) = 1.$$

We now consider g. This function is in $L^1(\mathbb{R};\mathbb{R})$:

$$\int_0^1\left|\frac{1}{\sqrt{x}}\right|\,\mathrm{d}x = \lim_{c\to 0}\int_c^1\frac{1}{\sqrt{x}}\,\mathrm{d}x = \lim_{c\to 0}[2\sqrt{x}]_c^1 = \lim_{c\to 0}(2-2\sqrt{c}) = 2.$$

However, it is not in $L^2(\mathbb{R};\mathbb{R})$, because we obtain the integral

$$\int_0^1\left|\frac{1}{x}\right|\,\mathrm{d}x$$

which is not convergent. ■

14 Solutions to the Exercises in Chapter 5

Solution 5.1 *(Exercise 5.1).* We recall that $\boldsymbol{x} \notin \operatorname{supp} f$ implies that there exists an open subset B of A such that $\boldsymbol{x} \in B$ and $f|_B = 0$ almost everywhere.

(i) Every $\boldsymbol{x} \in A$ is such that $\boldsymbol{x} \notin \operatorname{supp} f$; thus, $\operatorname{supp} f = \varnothing$;

(ii) We recall that if $S \subseteq T$, where S and T are subsets of A, then $A \setminus S \supseteq A \setminus T$. Moreover, $A \setminus (S \cap T) = (A \setminus S) \cup (A \setminus T)$ (De Morgan's law of the complement). Thus, we need to prove that $A \setminus \operatorname{supp}(f\,g) \supseteq (A \setminus \operatorname{supp} f) \cup (A \setminus \operatorname{supp} g)$. Let $\boldsymbol{x} \in (A \setminus \operatorname{supp} f) \cup (A \setminus \operatorname{supp} g)$. This means that $\boldsymbol{x} \in (A \setminus \operatorname{supp} f)$ or $\boldsymbol{x} \in (A \setminus \operatorname{supp} g)$. Consider for example the case where $\boldsymbol{x} \in A \setminus \operatorname{supp} f$. Then, there exists an open subset B of A such that $\boldsymbol{x} \in B$ and $f|_B = 0$ almost everywhere, which also implies that $(f\,g)|_B = 0$ almost everywhere. It follows that $\boldsymbol{x} \in A \setminus \operatorname{supp}(f\,g)$. Changing f with g, we find that if $\boldsymbol{x} \in A \setminus \operatorname{supp} g$, we also have $\boldsymbol{x} \in A \setminus \operatorname{supp}(f\,g)$. Hence, $(A \setminus \operatorname{supp} f) \cup (A \setminus \operatorname{supp} g) \subseteq A \setminus \operatorname{supp}(f\,g)$ as required.

■

Solution 5.2 *(Exercise 5.2).* Let $\varphi, \psi \in \mathcal{D}(A;\mathbb{C})$ and $\alpha, \beta \in \mathbb{F}$. We need to show that $\alpha\,\varphi + \beta\,\psi \in \mathcal{D}(A;\mathbb{C})$. It is easy to see that $\alpha\,\varphi + \beta\,\psi$ is C^∞, because such are φ and ψ (and the linear combinations of continuous/differentiable functions are continuous/differentiable). We only need to show that $\alpha\,\varphi + \beta\,\psi$ has compact support. Since φ has compact support, there exists a compact subset K of A such that $\operatorname{supp}\varphi \subseteq K$, so that for almost every $\boldsymbol{x} \in A \setminus K$ we have $\varphi(\boldsymbol{x}) = 0$. From the continuity of φ, the "almost every" becomes an "every" (see Example on page 9 of the Lecture Notes on Banach-Hilbert spaces). In other words,

$$\forall\, \boldsymbol{x} \in A \setminus K, \quad \varphi(\boldsymbol{x}) = 0.$$

Likewise, there exists a compact subset H of A such that $\operatorname{supp}\psi \subseteq H$, so that for every $\boldsymbol{x} \in A \setminus H$ we have $\psi(\boldsymbol{x}) = 0$. Thus,

$$\forall\, \boldsymbol{x} \in (A \setminus K) \cap (A \setminus H), \quad \varphi(\boldsymbol{x}) = \psi(\boldsymbol{x}) = 0.$$

In view of the de Morgan's laws for the complements, this is equivalent to

$$\forall\, \boldsymbol{x} \in A \setminus (K \cup H), \quad \varphi(\boldsymbol{x}) = \psi(\boldsymbol{x}) = 0,$$

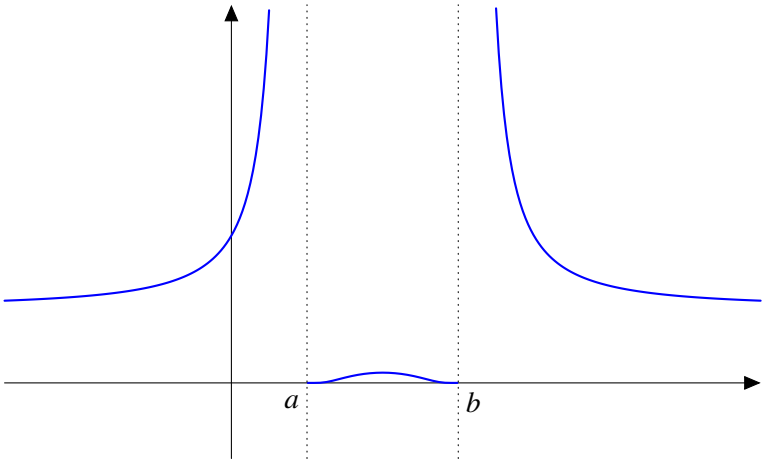
which is in turn equivalent to

$$\forall\, \boldsymbol{x} \in A \setminus (K \cup H), \quad \forall\, \alpha, \beta \in \mathbb{F}, \quad \alpha\,\varphi(\boldsymbol{x}) + \beta\,\psi(\boldsymbol{x}) = 0.$$

Since $K \cup H$ is compact (the union of a finite number of closed sets is closed, and the union of bounded sets is bounded), we find that $\alpha\,\varphi + \beta\,\psi$ has compact support.

The proof of the corresponding closure under scalar multiplication is trivial. ■

Solution 5.3 *(Exercise 5.3).* The graph of the function $\exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right)$ for $x \in \mathbb{R}$ is



The restriction of this function to the interval $[a, b]$ is continuous. The continuation of this restriction using the zero function is continuous. Indeed,

$$\lim_{x\to a^+} \exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right) = \lim_{x\to a^+} \exp\left(\frac{1}{a-b} - \frac{1}{x-a}\right) = 0,$$

and

$$\lim_{x\to b^-} \exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right) = \lim_{x\to b^-} \exp\left(\frac{1}{x-b} - \frac{1}{b-a}\right) = 0.$$

The derivative of φ is

$$\varphi'(x) = \begin{cases} \left(\frac{1}{(x-a)^2} - \frac{1}{(x-b)^2}\right) \exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right) & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

and the limits are still 0:

$$\lim_{x\to a^+} \varphi'(x) = 0 \quad \text{and} \quad \lim_{x\to b^-} \varphi'(x) = 0,$$

because the exponential term is the same, and it goes to 0 exponentially, while the factor in front of it will go to infinity as a polynomial function. We can continue in this way:

$$\varphi''(x) = \begin{cases} \left[\frac{2}{(x-b)^3} - \frac{1}{(x-a)^3} + \left(\frac{1}{(x-a)^2} - \frac{1}{(x-b)^2}\right)\right] \exp\left(\frac{1}{x-b} - \frac{1}{x-a}\right) & \text{if } a < x < b \\ 0 & \text{otherwise} \end{cases}$$

Notice that the graph is symmetric about the line $x = \frac{b+a}{2}$. Indeed, for all $x \in \mathbb{R}$ we have

$$\varphi\left(\frac{b+a}{2} - x\right) = \varphi\left(\frac{b+a}{2} + x\right).$$

Let us now consider the function ψ. This function is equal to 1 for $\|\boldsymbol{x}\|^2 \le a$ and zero for $\|\boldsymbol{x}\|^2 \ge b$. Since it is zero outside the compact $\overline{\mathcal{B}}_{\sqrt{|b|}}(0)$, the function ψ is in $\mathcal{D}(\mathbb{R}^n; \mathbb{R})$.

■

Solution 5.4 *(Exercise 5.4).* The inclusion $\mathcal{D}(A; \mathbb{C}) \subseteq L^\infty(A; \mathbb{C})$ means

$$f \in \mathcal{D}(A; \mathbb{C}) \quad \Rightarrow \quad [f] \in L^\infty(A; \mathbb{C}).$$

Let K be a compact set that contains the support of f. Since f is bounded by $\max_K |f|$ (which exists by the extreme value theorem), we find $[f] \in L^\infty(A; \mathbb{C})$. ■

Solution 5.5 *(Exercise 5.5).*

(i) Consider a compact set $K = [a, b]$, where $a < b$. Then

$$\int_K |c| \, \mathrm{d}x = \int_a^b |c| \, \mathrm{d}x = (b - a)\,|c|.$$

Therefore, the constant function $f = c$ is in $L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{R})$.

(ii) Consider a compact set $K = [a, b]$, where $a < b$. Then

$$\int_K |e^x| \, \mathrm{d}x = \int_a^b e^x \, \mathrm{d}x = [e^x]_a^b = e^b - e^a,$$

which is finite. Thus, e^x is in $L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{R})$.

(iii) Consider a compact set $K = [a, b]$, where $a < b$. Then

$$\int_K |\sin x| \, \mathrm{d}x = \int_a^b |\sin x| \, \mathrm{d}x.$$

The explicit value of this integral depends on the sign of a and b. However, since $|\sin x| \leq 1$ for all $x \in \mathbb{R}$, we have

$$\int_a^b |\sin x| \, \mathrm{d}x \leq \int_a^b 1 \, \mathrm{d}x = b - a.$$

Thus, $\sin x$ is in $L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{R})$.

(iv) We have seen that a primitive of $\arctan \frac{1}{x}$ is $x \arctan \frac{1}{x} + \frac{1}{2} \ln(x^2+1) + C$, where $C \in \mathbb{R}$ is a constant. Thus, considering a compact set $K = [a, b]$, where $a < b$, and taking into account that $\arctan \frac{1}{x}$ is an odd function, we can restrict our attention to the case where $a, b > 0$, and we have

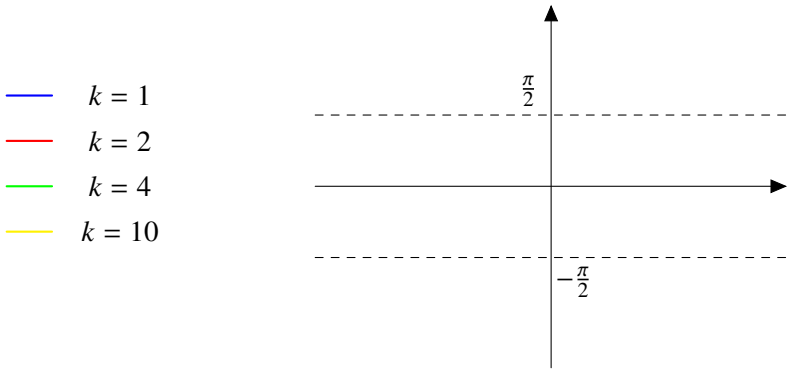
$$\int_K \left|\arctan \frac{1}{x}\right| \mathrm{d}x = \int_a^b \arctan \frac{1}{x} \, \mathrm{d}x = \left[x \arctan \frac{1}{x} + \frac{1}{2} \ln(x^2 + 1)\right]_a^b = \arctan \frac{a+b}{a\,b-1} + \frac{1}{2} \ln \frac{b^2+1}{a^2+1},$$

which is a finite quantity. Thus, $\arctan \frac{1}{x}$ is in $L^1_{\mathrm{loc}}(\mathbb{R} \setminus \{0\}; \mathbb{R})$. However, it is also in $L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{R})$. In fact, consider $K = [0, 1]$:

$$\begin{aligned}\int_K \left|\arctan \frac{1}{x}\right| \mathrm{d}x &= \int_0^1 \arctan \frac{1}{x} \, \mathrm{d}x = \lim_{c \to 0^+} \left[x \arctan \frac{1}{x} + \frac{1}{2} \ln(x^2 + 1)\right]_c^1 \\ &= \arctan 1 + \frac{1}{2} \ln 2 - \lim_{c \to 0^+} \left(c \arctan \frac{1}{c}\right) = \frac{\pi}{4} + \frac{1}{2} \ln 2.\end{aligned}$$

■

Solution 5.6 *(Exercise 5.6).* For all $k \in \mathbb{N}$ the function f_k is in $L^1_{\mathrm{loc}}(\mathbb{R}; \mathbb{R})$ because it is continuous. The graphs of $\arctan(k\,t)$ for $k = 1, 2, 4, 10$ are given below:



We show that the sequence $(f_k)_{k\in\mathbb{N}}$ converges to the sign function in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. The sequence $(f_k)_{k\in\mathbb{N}}$ converges pointwise to the sign function since for every $t \in \mathbb{R}$

$$\lim_{k\to\infty} \arctan(k\,t) = \frac{\pi}{2}\,\operatorname{sgn}(t).$$

In order to apply Lebesgue dominated convergence theorem (Theorem 3.47), we notice that

$$0 \le \left|\frac{2}{\pi}\arctan(k\,t) - \operatorname{sgn} t\right| \le \left|\frac{2}{\pi}\arctan(k\,t)\right| + 1 \le \frac{2}{\pi}\frac{\pi}{2} + 1 = 2.$$

Thus, given a compact set K in $\mathbb{R}$, we find

$$\begin{aligned}\lim_{k\to\infty}\int_K \left|\frac{2}{\pi}\arctan(k\,t) - \operatorname{sgn} t\right| \mathrm{d}t &= \int_K \lim_{k\to\infty}\left|\frac{2}{\pi}\arctan(k\,t) - \operatorname{sgn} t\right| \mathrm{d}t \\ &= \int_K 0\,\mathrm{d}t = 0.\end{aligned}$$

■

Solution 5.7 *(Exercise 5.7).* The function has two singularities: $x = \pm 1$. We show that $f \notin L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. To this end, we can choose for example the compact interval $K = [1, b]$, where $b > 1$:

$$\begin{aligned}\int_1^b \frac{1}{x^2-1}\,\mathrm{d}x &= \lim_{c\to 1^+}\int_c^b \frac{1}{(x+1)(x-1)}\,\mathrm{d}x \\ &= \lim_{c\to 1^+}\int_{c-1}^{b-1} \frac{1}{y\,(y+2)}\,\mathrm{d}y \\ &= \lim_{d\to 0^+}\left(\int_d^{b-1}\frac{1}{2\,y}\,\mathrm{d}y - \int_d^{b-1}\frac{1}{2\,(y+2)}\,\mathrm{d}y\right),\end{aligned}$$

and the first integral does not converge. We conclude that $f \notin L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. ■

Solution 5.8 *(Exercise 5.8).* The only problem of integrability is when x tends to $\pm\infty$. Let $[f] \in \boldsymbol{W}$. Then, $[f] \in L^2(\mathbb{R};\mathbb{R})$, so that

$$\int_{\mathbb{R}} |f|^2 < +\infty,$$

which implies that $|f|^2$ goes to zero faster than $1/x$ as $x \to \pm\infty$, i.e., $|f|$ goes to zero faster than $1/\sqrt{x}$ as $x \to \pm\infty$. Since $[x\,f] \in L^2(\mathbb{R};\mathbb{R})$, we have

$$\int_{\mathbb{R}} x^2\,|f(x)|^2 < +\infty,$$

and therefore $x^2\,|f|^2$ goes to zero faster than $1/x$, so that $|f|^2$ goes to zero faster than $1/x^3$, and so $|f|$ goes to zero faster than $1/(x^{3/2})$. This latter fact implies that $|f|$ goes to zero faster than $1/x$, and therefore $[f] \in L^1(\mathbb{R};\mathbb{R})$.

■

Solution 5.9 *(Exercise 5.9).* Let K be a compact set contained in A. Then, by Weierstrass theorem

$$\int_K |\psi(t)\,f(t)|\,\mathrm{d}t \leq \int_K (\max_{t\in K} \psi(t))\,f(t)\,\mathrm{d}t = (\max_{t\in K} \int_K f(t)\,\mathrm{d}t < +\infty,$$

and so $[\psi\, f] \in L^1_{\mathrm{loc}}(A;\mathbb{R})$. ■

Solution 5.10 *(Exercise 5.10).* Since φ and ψ are both C^∞-class functions, so is their product. Moreover, since φ has compact support, there exists $K \subseteq A$, compact, such that $\operatorname{supp}\varphi \subseteq K$. Therefore, $\operatorname{supp}(\varphi\,\psi) \subseteq K$. We conclude that $\psi\,\varphi \in \mathcal{D}(A;\mathbb{R})$. ■

Solution 5.11 *(Exercise 5.11).*

(i) Since $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, there exists a compact subset $\boldsymbol{K}$ of $\boldsymbol{A}$ such that $\operatorname{supp}\varphi \subset \boldsymbol{K}$, so that $\varphi(\boldsymbol{x}) = 0$ for every $\boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{K}$. Then,

$$\left|\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi(x)\,\mathrm{d}\boldsymbol{x}\right| = \left|\int_{\boldsymbol{K}} f(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right| \leq \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,|\varphi(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} \leq \sup_{\boldsymbol{K}} |\varphi| \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < \infty,$$

since $[f] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$.

(ii) Notice that the domain and codomain are both vector spaces. We need to show that it is linear. To this end, consider $\alpha,\beta \in \mathbb{F}$ and $\varphi_1,\varphi_2 \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then

$$\begin{aligned} L_f(\alpha\,\varphi_1 + \beta\,\varphi_2) &= \int_{\boldsymbol{A}} f(\boldsymbol{x})\,(\alpha\,\varphi_1(\boldsymbol{x}) + \beta\,\varphi_2(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \\ &= \alpha \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi_1(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + \beta \int_{\boldsymbol{A}} f(\boldsymbol{x})\,\varphi_2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \alpha\,L_f(\varphi_1) + \beta\,L_f(\varphi_2). \end{aligned}$$

■

Solution 5.12 *(Exercise 5.12).* Since $[f_1],[f_2] \in L^1_{\mathrm{loc}}(\boldsymbol{A};\mathbb{C})$, f_1 is defined over $\boldsymbol{A} \setminus \boldsymbol{N}_1$ and f_2 is defined over $\boldsymbol{A} \setminus \boldsymbol{N}_2$, where $\boldsymbol{N}_1, \boldsymbol{N}_2$ have measure zero. The set $\boldsymbol{N} = \{\boldsymbol{x} \in \boldsymbol{A} \setminus (\boldsymbol{N}_1 \cup \boldsymbol{N}_2) \,|\, f_1(\boldsymbol{x}) \neq f_2(\boldsymbol{x})\}$ has also measure zero. Thus, given $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$,

$$\begin{aligned} \int_{\boldsymbol{A}} f_1(x)\,\varphi(x)\,\mathrm{d}x &= \int_{\boldsymbol{A}\cap(\boldsymbol{N}_1\cup\boldsymbol{N}_2\cup\boldsymbol{N})} f_1(x)\,\varphi(x)\,\mathrm{d}x + \int_{\boldsymbol{A}\setminus(\boldsymbol{N}_1\cup\boldsymbol{N}_2\cup\boldsymbol{N})} f_1(x)\,\varphi(x)\,\mathrm{d}x \\ &= 0 + \int_{\boldsymbol{A}\setminus(\boldsymbol{N}_1\cup\boldsymbol{N}_2\cup\boldsymbol{N})} f_2(x)\,\varphi(x)\,\mathrm{d}x \\ &= \int_{\boldsymbol{A}\cap(\boldsymbol{N}_1\cup\boldsymbol{N}_2\cup\boldsymbol{N})} f_2(x)\,\varphi(x)\,\mathrm{d}x + \int_{\boldsymbol{A}\setminus(\boldsymbol{N}_1\cup\boldsymbol{N}_2\cup\boldsymbol{N})} f_2(x)\,\varphi(x)\,\mathrm{d}x = \int_{\boldsymbol{A}} f_2(x)\,\varphi(x)\,\mathrm{d}x. \end{aligned}$$

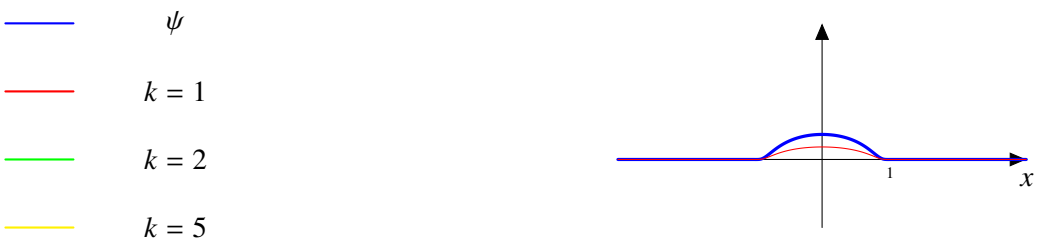
■

Solution 5.13 *(Exercise 5.13).* First, we notice that, for every $k \in \mathbb{N}^*$, the function φ_k is in $C^\infty(\mathbb{R};\mathbb{R})$ because it is the product of a constant 2^{-k} by the composition of two functions in $C^\infty(\mathbb{R};\mathbb{R})$. Moreover, since ψ is zero outside a compact set of $\mathbb{R}$, so is φ, because $\psi(k^5\,x) = 0$ implies $\varphi_k(x) = 0$. We now need to show that for every $k \in \mathbb{N}^*$ the function φ_k has compact support. We know that this is true for ψ because $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. In particular, there exists $R > 0$ such that $\psi(t) = 0$ for all $|t| > R$, which implies that $\psi(k^5\,x) = 0$ for all $|k^5\,x| > R$, i.e., for all $|x| > k^{-5}\,R$. It follows that the support of φ_k is contained in the compact $[-k^{-5}\,R, k^{-5}\,R]$, and therefore φ_k has compact support.

To prove the convergence in $\mathcal{D}(\mathbb{R};\mathbb{R})$, we find the compact set $K \subset \mathbb{R}$ that contains the supports of each φ_k. Using the same argument employed before, since $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, there exists $R > 0$ such that $\operatorname{supp}\psi \subseteq [-R, R]$. It follows that if $|k^5\,x| > R$, then $\varphi_k(x) = 0$, which means that if $|x| > k^{-5}\,R$, we have $\varphi_k(x) = 0$. Therefore, the supports of all φ_k are contained in the interval $[-k^{-5}\,R, k^{-5}\,R]$, i.e., for all $k \in \mathbb{N}^*$, we have $\operatorname{supp}\varphi_k \subseteq [-k^{-5}\,R, k^{-5}\,R] \subseteq [-R, R]$. Now, let us consider the second condition on the uniform convergence of the derivatives. For all $\alpha \in \mathbb{N}$ we have

$$\sup_{x\in\mathbb{R}} |\varphi_k^{(\alpha)}(x)| = \sup_{x\in\mathbb{R}} |2^{-k}\,k^{5\,\alpha}\,\psi^{(\alpha)}(k^5\,x)| = 2^{-k}\,k^{5\,\alpha} \max_{x\in[-R,R]} |\psi^{(\alpha)}(k^5\,x)| \xrightarrow{k\to\infty} 0.$$

Indeed, a sequence $\frac{k^\beta}{\gamma^k}$, for $\gamma > 1$, goes to 0 as $k \to +\infty$. This shows that the sequence $\partial^\alpha\,\varphi_k(t)$ converges to zero uniformly on $\mathbb{R}$. Thus, the sequence $(\varphi_k)_{k\in\mathbb{N}}$ converges to $\varphi = 0$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$. In the figure below, we considered as ψ the function $\psi(x) = e^{-\frac{1}{1-x^2}}\,\chi_{|x|<1}$.



■

Solution 5.14 *(Exercise 5.14).* First, we need to show that $(\varphi_k)_{k\in\mathbb{N}^*}$ is a sequence of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$. Let $k \in \mathbb{N}^*$. The function φ_k is $C^\infty(\mathbb{R};\mathbb{R})$ because it is the composition of functions in $C^\infty(\mathbb{R};\mathbb{R})$. Since $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, as in the previous exercise there exists $R > 0$ such that $\psi(t) = 0$ for all $|t| > R$, so that $\psi\left(\frac{x}{k}\right) = 0$ for all $\left|\frac{x}{k}\right| > R$, i.e., for all $|x| > k\,R$. It follows that $\operatorname{supp}\varphi_k \subseteq [-k\,R, k\,R]$. This proves that $(\varphi_k)_{k\in\mathbb{N}^*}$ is a sequence of functions in $\mathcal{D}(\mathbb{R};\mathbb{R})$. However, this argument also shows that we do not have convergence, because we cannot find a compact subset K of $\mathbb{R}$ such that $\operatorname{supp}\varphi_k \subseteq K$. Therefore, φ_k does not converge in $\mathcal{D}(\mathbb{R};\mathbb{R})$.

■

Solution 5.15 *(Exercise 5.15).*

■

Solution 5.16 *(Exercise 5.16).* The implication (**A**) $\Rightarrow$ (**B**) is obvious: we can just take φ equal to the zero function, which is an element of $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ (since $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ is a vector space), and in this case $T(\varphi) = T(0) = 0$ (recall that a linear transformation maps 0 of the domain into the 0 of the codomain).

Now we prove (**B**) $\Rightarrow$ (**A**). Let (**B**) hold, and let $\lim_{k\to+\infty} \varphi_k = \varphi$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. We have to show that $T(\varphi_k) \xrightarrow{k\to\infty} T(\varphi)$. First, since $\lim_{k\to\infty} \varphi_k = \varphi$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$, we have $\lim_{k\to\infty} (\varphi_k - \varphi) = 0$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. In fact, the conditions $\operatorname{supp}\varphi_k \subseteq K$ for all $k \in \mathbb{N}$ (where $\boldsymbol{K}$ is a compact subset of $\boldsymbol{A}$) and $\partial^\alpha \varphi_k \xrightarrow{k\to\infty} \partial^\alpha \varphi$ uniformly also imply that $\operatorname{supp}\varphi \subseteq \boldsymbol{K}$. It follows that $\operatorname{supp}(\varphi_k - \varphi) \subseteq \boldsymbol{K}$ and from the linearity of the limit $\partial^\alpha(\varphi_k - \varphi) \xrightarrow{k\to\infty} 0$. From **(B)**, we now have that $T(\varphi_k - \varphi) \xrightarrow{k\to\infty} 0$, and from the linearity of T this implies that $T(\varphi_k) - T(\varphi) \xrightarrow{k\to\infty} 0$. Again, from the linearity of the limit we obtain $T(\varphi_k) \xrightarrow{k\to\infty} T(\varphi)$.

■

Solution 5.17 *(Exercise 5.17).* Let $\alpha, \beta \in \mathbb{R}$ and $\varphi, \psi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We proved that $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$ is a real vector space, so that $\alpha\varphi + \beta\psi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$, where obviously

$$\begin{aligned} \alpha\varphi + \beta\psi : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \alpha\varphi(\boldsymbol{x}) + \beta\psi(\boldsymbol{x}). \end{aligned}$$

Then,

$$\delta(\alpha\varphi + \beta\psi) = \alpha\varphi(0) + \beta\psi(0) = \alpha\,\delta(\varphi) + \beta\,\delta(\psi).$$

■

Solution 5.18 *(Exercise 5.18).* First, we check linearity and continuity. Let $\varphi, \psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Let $\alpha, \beta \in \mathbb{R}$. From the linearity of the Lebesgue integral (and the fact that $\varphi, \psi, \alpha\varphi + \beta\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$), we obtain

$$T(\alpha\varphi+\beta\psi) = \int_0^{+\infty} x\,(\alpha\varphi'(x)+\beta\psi'(x))\,\mathrm{d}x = \alpha\int_0^{+\infty} x\varphi'(x)\,\mathrm{d}x + \beta\int_0^{+\infty} x\psi'(x)\,\mathrm{d}x = \alpha\,T(\varphi)+\beta\,T(\psi).$$

This proves the linearity. Now we check the continuity, by considering a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$ that converges to $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$. We need to show that

$$T(\varphi_k) \xrightarrow{k\to\infty} T(\varphi).$$

From the convergence of $(\varphi_k)_{k\in\mathbb{N}}$ to φ in $\mathcal{D}(\mathbb{R};\mathbb{R})$, we know that there exists a compact subset of $\mathbb{R}$ that contains the supports of every φ_k and also of φ (and, moreover, the support of φ' is contained in the support of φ. Why?), so that

$$T(\varphi_k) = \int_0^{+\infty} x\,\varphi_k'(x)\,\mathrm{d}x = \int_{K\cap\mathbb{R}_+} x\,\varphi_k'(x)\,\mathrm{d}x.$$

From the uniform convergence of $(\varphi_k)_{k\in\mathbb{N}}$ to φ (with all the derivatives of every order), we have the pointwise convergence, which guarantees that the integrand $x\,\varphi_k'(x)$ is pointwise convergent to $x\,\varphi'(x)$. Now, we only need to apply Lebesgue dominated convergence theorem (Theorem 3.47):

$$|x\,\varphi_k'(x)| \leq (\max_{x\in K} x)\cdot \sup_{k\in\mathbb{N}}(\max_K |\varphi_k'|) < +\infty,$$

since K is bounded, and where we have used Weierstrass Extreme Value Theorem. Thus, Lebesgue dominated convergence theorem (Theorem 3.47) can be applied and we find

$$\lim_{k\to\infty} T(\varphi_k) = \int_0^{+\infty} \lim_{k\to\infty} x\,\varphi_k'(x)\,\mathrm{d}x = \int_0^{+\infty} x\,\varphi'(x)\,\mathrm{d}x = T(\varphi).$$

We are now ready to find f. We integrate by parts:

$$T(\varphi) = \underbrace{\lim_{a\to+\infty}[x\,\varphi(x)]_0^a}_{=0} - \int_0^{+\infty}\varphi(x)\,\mathrm{d}x = -\int_0^{+\infty}\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}} f(x)\,\varphi(x)\,\mathrm{d}x,$$

by defining for all $x\in\mathbb{R}$

$$f(x) = \begin{cases} 0 & \text{if } x\leq 0\\ -1 & \text{if } x>0.\end{cases}$$

■

Solution 5.19 *(Exercise 5.19).* First, we observe that the balls $\mathcal{B}_{\frac{1}{k}}(0)$ decrease in size as k increases, but the integral is always equal to 1:

$$\forall\, k\in\mathbb{N},\qquad \int_{\mathbb{R}^n} f_k(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{\mathcal{B}_{\frac{1}{k}}(0)} \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\,\mathrm{d}\boldsymbol{x} = \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\int_{\mathcal{B}_{\frac{1}{k}}(0)}\mathrm{d}\boldsymbol{x} = 1.$$

Let us now consider $\varphi\in\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. We find

$$\langle L_{f_k},\varphi\rangle = \int_{\mathbb{R}^n} f_k(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\int_{\mathcal{B}_{\frac{1}{k}}(0)}\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

We want to find the limit, as $k\to+\infty$, of $\langle L_{f_k},\varphi\rangle$: we add and subtract $\varphi(0)$, and obtain

$$\langle L_{f_k},\varphi\rangle = \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\int_{\mathcal{B}_{\frac{1}{k}}(0)}(\varphi(\boldsymbol{x})-\varphi(0))\,\mathrm{d}\boldsymbol{x} + \varphi(0).$$

The first term in the right hand-side of the latter goes to 0 as $k\to+\infty$. Indeed, since φ is continuous, for every $\varepsilon>0$ there exists $\delta_\varepsilon>0$ such that $|\boldsymbol{x}|\leq\delta_\varepsilon$ implies $|\varphi(\boldsymbol{x})-\varphi(0)|\leq\varepsilon$. Thus, taking $k>\frac{1}{\delta_\varepsilon}$ we find

$$\left|\frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\int_{\mathcal{B}_{\frac{1}{k}}(0)}(\varphi(\boldsymbol{x})-\varphi(0))\,\mathrm{d}\boldsymbol{x}\right| \leq \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\int_{\mathcal{B}_{\frac{1}{k}}(0)}\left|\varphi(\boldsymbol{x})-\varphi(0)\right|\mathrm{d}\boldsymbol{x} \leq \frac{1}{\mu_n(\mathcal{B}_{\frac{1}{k}}(0))}\,\varepsilon\,\mu_n(\mathcal{B}_{\frac{1}{k}}(0)) = \varepsilon.$$

We have proved that for any $\varphi\in\mathcal{D}(A)$ we have

$$\langle L_{f_k},\varphi\rangle \xrightarrow{\quad k\to\infty\quad} \varphi(0) = \langle\delta,\varphi\rangle.$$

Thus, $\lim_{k\to\infty} L_{f_k} = \delta$ in $\mathcal{D}'(\mathbb{R}^n;\mathbb{R})$. ■

Solution 5.20 *(Exercise 5.20).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned}
\langle L_{f_k}, \varphi\rangle &= \int_{\mathbb{R}} f_k(x)\,\varphi(x)\,\mathrm{d}x = \sqrt{k}\int_0^{1/k} \varphi(x)\,\mathrm{d}x \\
&= \sqrt{k}\int_0^{1/k} (\varphi(x)-\varphi(0))\,\mathrm{d}x + \sqrt{k}\int_0^{1/k}\varphi(0)\,\mathrm{d}x \\
&= \sqrt{k}\int_0^{1/k} (\varphi(x)-\varphi(0))\,\mathrm{d}x + \frac{\varphi(0)}{\sqrt{k}} \\
&\leq \sqrt{k}\,\frac{1}{k}\sup_{x\in\left[0,\frac{1}{k}\right]}(\varphi(x)-\varphi(0)) + \frac{\varphi(0)}{\sqrt{k}} \xrightarrow{k\to\infty} 0.
\end{aligned}$$

Thus, L_{f_k} converges to zero in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. ■

Solution 5.21 *(Exercise 5.21).* Consider $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, and let $M > 0$ be such that $\operatorname{supp}\varphi \subseteq [-M, M]$. We wish to determine whether the following limit exists:

$$\begin{aligned}
\lim_{k\to\infty}\int_{\mathbb{R}} \sin^2(k\,t)\,\varphi(t)\,\mathrm{d}t &= \lim_{k\to\infty}\int_{\mathbb{R}} \frac{1-\cos(2\,k\,t)}{2}\varphi(t)\,\mathrm{d}t \\
&= \lim_{k\to\infty}\left(\frac{1}{2}\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t - \frac{1}{2}\int_{\mathbb{R}}\cos(2\,k\,t)\,\varphi(t)\,\mathrm{d}t\right) \\
&= \lim_{k\to\infty}\Big(\frac{1}{2}\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t - \underbrace{\frac{1}{2}\left[\frac{\sin(2\,k\,t)}{2\,k}\,\varphi(t)\right]_{-M}^{M}}_{=0} \\
&\quad + \frac{1}{2}\int_{-M}^{M}\frac{\sin(2\,k\,t)}{2\,k}\,\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}\,\mathrm{d}t\Big) \\
&= \lim_{k\to\infty}\Big(\frac{1}{2}\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t + \frac{1}{2}\int_{-M}^{M}\frac{\sin(2\,k\,t)}{2\,k}\,\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}\,\mathrm{d}t\Big) \\
&= \frac{1}{2}\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t + \frac{1}{2}\lim_{k\to\infty}\frac{\sin(2\,k\,t)}{2\,k}\,\frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}\,\mathrm{d}t = \frac{1}{2}\int_{\mathbb{R}}\varphi(t)\,\mathrm{d}t,
\end{aligned}$$

so that $\sin^2(k\,t) \xrightarrow{k\to\infty} \frac{1}{2}$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. ■

Solution 5.22 *(Exercise 5.22).*

(i) First, we want to convince ourselves that the definition of T makes sense. For $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$, we need to show that the function

$$\begin{aligned}
\psi : \mathbb{R} &\longrightarrow \mathbb{R} \\
x &\mapsto \varphi(x,0)
\end{aligned}$$

is in $\mathcal{D}(\mathbb{R};\mathbb{R})$. It is obviously of class C^∞. We must show that its support is contained in a compact set. Since $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$, there exists $R > 0$ such that $\operatorname{supp}\varphi \subseteq \overline{\mathcal{B}_R((0,0))}$, which implies that $\varphi(x,y) = 0$ for $(x,y) \in \mathbb{R}^2$ such that $x^2 + y^2 > R^2$. Thus, $\varphi(x,0) = 0$ for all $x \in \mathbb{R}$ such that $|x| > R$, which implies that $\operatorname{supp}\psi \subseteq [-R, R]$. This proves that $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$.

Let $\alpha, \beta \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$. The set $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$ is a real vector space, so that $\alpha\varphi_1 + \beta\varphi_2 \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$, where obviously

$$\begin{aligned} \alpha\varphi_1 + \beta\varphi_2 : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto \alpha\varphi_1(x,y) + \beta\varphi_2(x,y). \end{aligned}$$

We prove the linearity, by exploiting the property of linearity of the Lebesgue integral:

$$\begin{aligned} T(\alpha\varphi_1 + \beta\varphi_2) &= \int_{\mathbb{R}} h(x)\left(\alpha\varphi_1(x,0) + \beta\varphi_2(x,0)\right)\mathrm{d}x \\ &= \alpha\int_{\mathbb{R}} h(x)\,\varphi_1(x,0)\,\mathrm{d}x + \beta\int_{\mathbb{R}} h(x)\,\varphi_2(x,0)\,\mathrm{d}x = \alpha T(\varphi_1) + \beta T(\varphi_2). \end{aligned}$$

We now prove that T is continuous. Let $(\varphi_k)_{k\in\mathbb{N}}$ be a sequence of functions in $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$ that converges to $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$ in $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$. Denoting

$$\begin{aligned} \psi : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \varphi(x,0) \end{aligned} \qquad \text{and} \qquad \begin{aligned} \psi_k : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \varphi_k(x,0) \end{aligned}$$

we must show that $(\psi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{D}(\mathbb{R};\mathbb{R})$ that converges to $\psi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ in $\mathcal{D}(\mathbb{R};\mathbb{R})$. In particular, we need to show that there exists a compact set K of $\mathbb{R}$ which contains all the supports of ψ_k and that we have uniform convergence of the derivatives of every order. First, since $(\varphi_k)_{k\in\mathbb{N}}$ converges to $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$ in $\mathcal{D}(\mathbb{R}^2;\mathbb{R})$, there exists a compact H in $\mathbb{R}^2$ such that

$$\forall\, k \in \mathbb{N}, \quad \operatorname{supp}\varphi_k \subseteq H.$$

Since H is bounded, there exists a ball centred at $(0,0)$ which contains H:

$$\exists\, R > 0 : \quad \operatorname{supp}\varphi_k \subseteq H \subseteq \overline{\mathcal{B}_R((0,0))}.$$

Using the same argument above, we find that $\varphi_k(x,0) = 0$ for all $x \in \mathbb{R}$ satisfying $|x| > R$. The same is true for the limit: $\varphi(x,0) = 0$ for all $x \in \mathbb{R}$ satisfying $|x| > R$. The condition on the uniform convergence of the derivatives of every order is trivial: given $\alpha \in \mathbb{N}$ we have

$$|\psi_k^{(\alpha)}(x) - \psi^{(\alpha)}(x)| = |\partial_x^\alpha\varphi_k(x,0) - \partial_x^\alpha\varphi(x,0)| \xrightarrow{k\to\infty} 0$$

uniformly, where ∂_x^α denotes the partial derivative with respect to the variable x.

Now we must show that $\lim_{k\to+\infty} T(\varphi_k) = T(\varphi)$. We take the difference and show that it goes to 0:

$$\begin{aligned} |\langle T,\varphi_k\rangle - \langle T,\varphi\rangle| &= \left|\int_{\mathbb{R}} h(x)\left(\varphi_k(x,0) - \varphi(x,0)\right)\mathrm{d}x\right| \\ &= \left|\int_{-R}^{R} h(x)\left(\varphi_k(x,0) - \varphi(x,0)\right)\mathrm{d}x\right| \\ &\le \int_{-R}^{R} |h(x)|\left|\varphi_k(x,0) - \varphi(x,0)\right|\mathrm{d}x \le \sup_{x\in[-R,R]} |\psi_k(x) - \psi(x)| \int_{-R}^{R} |h(x)|\,\mathrm{d}x. \end{aligned}$$

The term $\int_{-R}^{R} |h(x)|\,\mathrm{d}x$ is finite because h is locally integrable. The term $\sup_{x\in[-R,R]} |\psi_k(x) - \psi(x)|$ goes to 0 as $k \to +\infty$ from the definition of convergence in $\mathcal{D}$. It follows that

$$|\langle T,\varphi_k\rangle - \langle T,\varphi\rangle| \xrightarrow{k\to\infty} 0.$$

(ii) We now prove that T is not of function type. To be of function type, there should exist $f \in L^1_{\text{loc}}(\mathbb{R}^2;\mathbb{R})$ such that

$$\forall\, \varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R}), \quad T(\varphi) = \int_{\mathbb{R}^2} f(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

In other words, we must have

$$\forall\, \varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R}), \quad \int_{\mathbb{R}^2} f(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x. \tag{14.1}$$

Suppose by contradiction that such function $f \in L^1_{\text{loc}}(\mathbb{R}^2;\mathbb{R})$ exists. Consider $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$ such that

$$\operatorname{supp}\varphi \subseteq A, \text{ where } A = \mathbb{R}^2 \setminus (\mathbb{R}\times\{0\}) = \{(x,y)\in\mathbb{R}^2 \,|\, y \neq 0\}.$$

In other words, this support is not intersecting the x-axis. Thus, if $\operatorname{supp}\varphi \subseteq A$, we have

$$\int_{\mathbb{R}^2} f(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x = 0.$$

Thus, eliminating the x-axis (which has measure zero in $\mathbb{R}^2$), the value of the integral is still the same:

$$\int_A f|_A(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0 \qquad \forall\, \varphi \in \mathcal{D}(A;\mathbb{R}),$$

Applying the Fundamental Lemma, this implies that $f|_A$ is zero almost everywhere in A, so that f is zero almost everywhere in $\mathbb{R}^2$. This means that

$$\int_{\mathbb{R}^2} f(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x = 0 \qquad \forall\, \varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R}).$$

In other words, T is the identically zero distribution. This is going to be true, in particular, for a function $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$ constructed in this way: consider an arbitrary $w \in \mathcal{D}(\mathbb{R};\mathbb{R})$ and $\alpha \in \mathcal{D}(\mathbb{R};\mathbb{R})$ such that $\alpha(0) = 1$, and consider $\varphi(x,y) = w(x)\,\alpha(y)$. Since both w and α have compact support in $\mathbb{R}$, the product $w(x)\,\alpha(y)$ is zero outside a compact rectangle. We obtain

$$\int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x = \int_{\mathbb{R}} h(x)\,w(x)\,\mathrm{d}x = 0, \qquad \forall\, w \in \mathcal{D}(\mathbb{R};\mathbb{R}).$$

Thus, the fundamental lemma guarantees that $h(x) = 0$ almost everywhere in $\mathbb{R}$, leading to a contradiction.

(iii) Now we show that T can be obtained as the limit of distributions of function type. Let $\varepsilon > 0$, and let us define

$$\begin{aligned} f_\varepsilon : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\mapsto \begin{cases} \frac{1}{\varepsilon}\,h(x) & \text{for } 0 \le y < \varepsilon \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

For every $\varepsilon > 0$, the function f_ε is measurable in $\mathbb{R}^2$, and also locally integrable (i.e., $f_\varepsilon \in L^1_{\text{loc}}(\mathbb{R}^2;\mathbb{R})$). Let us denote by T_ε the distribution associated with f_ε, and we prove that

$$\lim_{\varepsilon\to 0^+} T_\varepsilon = T \qquad \text{in } \mathcal{D}'(\mathbb{R}^2;\mathbb{R}).$$

Equivalently, using the definition of convergence in $\mathcal{D}'(\mathbb{R}^2;\mathbb{R})$ this means that we need to show that

$$\forall\,\varphi\in\mathcal{D}(\mathbb{R}^2;\mathbb{R}),\qquad \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^2} f_\varepsilon(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y=\langle T,\varphi\rangle.$$

Using Fubini's theorem (Theorem 3.8)

$$\begin{aligned}\int_{\mathbb{R}^2} f_\varepsilon(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{\mathbb{R}\times[0,\varepsilon)}\frac{1}{\varepsilon}\,h(x)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_{\mathbb{R}} h(x)\,\frac{1}{\varepsilon}\left(\int_0^\varepsilon \varphi(x,y)\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int_{[-M,M]} h(x)\,\frac{1}{\varepsilon}\left(\int_0^\varepsilon \varphi(x,y)\,\mathrm{d}y\right)\mathrm{d}x,\end{aligned}$$

where we assumed $\varphi(x,y)=0$ for $(x,y)\in[-M,M]\times[-M,M]$. Now we want to find the limit of $\frac{1}{\varepsilon}\int_0^\varepsilon\varphi(x,y)\,\mathrm{d}y$ as $\varepsilon\to 0^+$. We use the fundamental theorem of integral calculus: for a continuous function g, we have[1]

$$\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon}\int_0^\varepsilon g(y)\,\mathrm{d}y=g(0).$$

Thus,

$$h(x)\,\frac{1}{\varepsilon}\int_0^\varepsilon\varphi(x,y)\,\mathrm{d}y\xrightarrow{\varepsilon\to 0^+}h(x)\,\varphi(x,0).$$

We apply Lebesgue dominated convergence theorem (Theorem 3.47):

$$\begin{aligned}\left|\frac{1}{\varepsilon}\int_0^\varepsilon\varphi(x,y)\,\mathrm{d}y\right| &\le \frac{1}{\varepsilon}\int_0^\varepsilon|\varphi(x,y)|\,\mathrm{d}y\\ &\le \frac{1}{\varepsilon}\sup|\varphi|\,\varepsilon,\end{aligned}$$

so that

$$\left|h(x)\,\frac{1}{\varepsilon}\int_0^\varepsilon\varphi(x,y)\,\mathrm{d}y\right|\le|h(x)|\,\sup|\varphi|,$$

which is integrable over $[-M,M]$ since $h\in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. It follows that

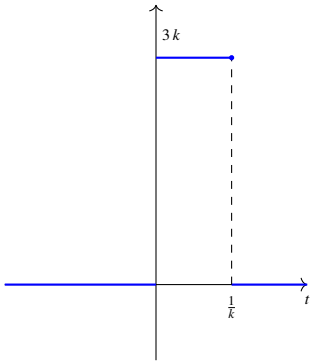
$$\begin{aligned}\lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^2} f_\varepsilon(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{[-M,M]} h(x)\,\varphi(x,0)\,\mathrm{d}x\\ &= \int_{\mathbb{R}} h(x)\,\varphi(x,0)\,\mathrm{d}x=\langle T,\varphi\rangle.\end{aligned}$$

■

Solution 5.23 *(Exercise 5.23).*

[1] Recall that, if f is a continuous function over $[a,b]$, then $F(x)=\int_a^x f(t)\,\mathrm{d}t$ implies that F is differentiable in (a,b) and $F'(x)=f(x)$. In our case, we would have $G(\varepsilon)=\int_0^\varepsilon g(y)\,\mathrm{d}y$ where G is a primitive of g, so that

$$\lim_{\varepsilon\to 0^+}\frac{1}{\varepsilon}[G(y)]_{y=0}^{y=\varepsilon}=\lim_{\varepsilon\to 0^+}\frac{G(\varepsilon)-G(0)}{\varepsilon}=G'(0)=g(0).$$



For every $k \in \mathbb{N}$, the function f_k is in $L^1_{\text{loc}}(\mathbb{R};\mathbb{R})$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned}\langle L_{f_k}, \varphi\rangle &= \int_{\mathbb{R}} f_k(t)\,\varphi(t)\,\mathrm{d}t = 3\,k\int_0^{1/k} \varphi(t)\,\mathrm{d}t \\ &= 3\,k\int_0^{1/k} (\varphi(t)-\varphi(0))\,\mathrm{d}t + 3\,\varphi(0).\end{aligned}$$

We have

$$\left|3\,k\int_0^{1/k} (\varphi(t)-\varphi(0))\,\mathrm{d}t\right| \le 3\,k\,\frac{1}{k}\sup_{t\in[0,\frac{1}{k}]} |\varphi(t)-\varphi(0)| \xrightarrow{k\to\infty} 0,$$

so that $\langle L_{f_k}, \varphi\rangle \xrightarrow{k\to\infty} 3\,\varphi(0)$. Stated differently, $\langle L_{f_k}, \varphi\rangle \xrightarrow{k\to\infty} \langle 3\,\delta, \varphi\rangle$ for all $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Therefore, $L_{f_k} \xrightarrow{k\to\infty} 3\,\delta$ in $\mathcal{D}'(\mathbb{R};\mathbb{R})$. ■

Solution 5.24 *(Exercise 5.24).*

(a) Since $[f] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$, for every compact set $\boldsymbol{K}$ of $\boldsymbol{A}$ we have

$$\int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty.$$

Since for all $z \in \mathbb{C}$ we have $|z| = |\bar{z}|$, then

$$\int_{\boldsymbol{K}} |\bar{f}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{K}} |\overline{f(\boldsymbol{x})}|\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{K}} |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} < +\infty.$$

Thus, $[\bar{f}] \in L^1_{\text{loc}}(\boldsymbol{A};\mathbb{C})$.

(b) From

$$\begin{aligned}\int_{\boldsymbol{A}} \overline{f(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x} &= \int_{\boldsymbol{A}} \mathfrak{Re}\{f(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} + i\int_{\boldsymbol{A}} -\mathfrak{Im}\{f(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} = \int_{\boldsymbol{A}} \mathfrak{Re}\{f(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} - i\int_{\boldsymbol{A}} \mathfrak{Im}\{f(\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} \\ &= \overline{\int_{\boldsymbol{A}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}},\end{aligned}$$

we have

$$\begin{aligned}\langle L_{\bar{f}}, \varphi\rangle &= \int_A \bar{f}(\boldsymbol{x})\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_A \overline{f(\boldsymbol{x})}\,\varphi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\\ &= \int_A \overline{f(\boldsymbol{x})\,\overline{\varphi(\boldsymbol{x})}}\,\mathrm{d}\boldsymbol{x}\\ &= \overline{\int_A f(\boldsymbol{x})\,\overline{\varphi(\boldsymbol{x})}\,\mathrm{d}\boldsymbol{x}} = \overline{\langle L_f, \bar{\varphi}\rangle}.\end{aligned}$$

(c) Given $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, we have $\bar{\varphi} \in C^\infty(\boldsymbol{A};\mathbb{C})$. Moreover, the support of φ is the same as the support of $\bar{\varphi}$ (we have $z = 0$ if and only if $\bar{z} = 0$ for all $z \in \mathbb{C}$). Thus, $\bar{\varphi} \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$.

We have to show that $\bar{T}$ defined in this way is linear and continuous. We start proving that it is linear. Let $\varphi, \psi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then

$$\begin{aligned}\bar{T}(\varphi+\psi) &= \langle \bar{T}, \varphi+\psi\rangle = \overline{\langle T, \overline{\varphi+\psi}\rangle}\\ &= \overline{\langle T, \overline{\varphi}+\overline{\psi}\rangle}\\ &= \overline{\langle T, \bar{\varphi}\rangle + \langle T, \bar{\psi}\rangle}\\ &= \overline{\langle T, \bar{\varphi}\rangle} + \overline{\langle T, \bar{\psi}\rangle} = \langle \bar{T}, \varphi\rangle + \langle \bar{T}, \psi\rangle = \bar{T}(\varphi) + \bar{T}(\psi).\end{aligned}$$

Let $\alpha \in \mathbb{C}$. Then,

$$\begin{aligned}\overline{T}(\alpha\,\varphi) &= \langle \bar{T}, \alpha\,\varphi\rangle = \overline{\langle T, \overline{\alpha\,\varphi}\rangle}\\ &= \overline{\langle T, \overline{\alpha}\,\overline{\varphi}\rangle}\\ &= \overline{\overline{\alpha}\,\langle T, \overline{\varphi}\rangle}\\ &= \alpha\,\overline{\langle T, \overline{\varphi}\rangle}\\ &= \alpha\,\langle \bar{T}, \varphi\rangle = \alpha\,\bar{T}(\varphi).\end{aligned}$$

We now prove the continuity. We must prove that if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$ that converges to $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$, then $\bar{T}(\varphi_k)$ converges to $\bar{T}(\varphi)$, i.e., that

$$\overline{\langle T, \bar{\varphi}_k\rangle} \xrightarrow{k\to\infty} \overline{\langle T, \bar{\varphi}\rangle}$$

The convergence of a sequence of complex numbers is equivalent to the convergence of the real and imaginary parts of the terms of the sequence. Thus, it suffices to show that $(\bar{\varphi}_k)_{k\in\mathbb{N}}$ converges to $\bar{\varphi}$ in $\mathcal{D}(\boldsymbol{A};\mathbb{C})$. Since the supports do not change, we only need to show that we have uniform convergence of the partial derivatives of any order.

(d) We have

$$\langle \bar{T}_k, \varphi\rangle = \overline{\langle T_k, \bar{\varphi}\rangle} \xrightarrow{k\to\infty} \overline{\langle T, \bar{\varphi}\rangle} = \langle \bar{T}, \varphi\rangle.$$

■

Solution 5.25 *(Exercise 5.25).* Let $\varphi, \psi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. We have

$$(T\,h)(\varphi+\psi) = T\big(h\,(\varphi+\psi)\big) = T(h\,\varphi + h\,\psi) = T(h\,\varphi) + T(h\,\psi) = (T\,h)(\varphi) + (T\,h)\,(\psi).$$

Given $k \in \mathbb{C}$, we have

$$(T\,h)(k\,\varphi) = T\,(h\,k\,\varphi) = k\,T(h\,\varphi) = k\,(T\,h)(\varphi).$$

■

Solution 5.26 *(Exercise 5.26).* Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. There holds

$$\langle h\,T_k, \varphi\rangle = \langle T_k, h\,\varphi\rangle \xrightarrow{k\to\infty} \langle T, h\,\varphi\rangle = \langle T\,h, \varphi\rangle.$$

■

Solution 5.27 *(Exercise 5.27).*

(i) Consider f_1, which is invertible and whose inverse is

$$\begin{aligned} f_1^{-1} : (1,\,\sqrt{2}) &\longrightarrow (1,2) \\ y &\mapsto y^2. \end{aligned}$$

Both f_1 and f_1^{-1} are C^∞. Thus, f_1 is a C^∞ diffeomorphism. The corresponding Jacobian matrix is the derivative:

$$\begin{aligned} f_1' : (1,2) &\longrightarrow (1,\,\sqrt{2}) \\ x &\mapsto \tfrac{1}{2\,\sqrt{x}} \end{aligned}$$

which is invertible for every $x \in (1,2)$.

(ii) Consider f_2, which is invertible and whose inverse is

$$\begin{aligned} f_2^{-1} : \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right) &\longrightarrow \mathbb{R} \\ y &\mapsto \tan y. \end{aligned}$$

Both f_2 and f_2^{-1} are C^∞. Thus, f_2 is a C^∞ diffeomorphism. The corresponding Jacobian matrix is the derivative:

$$\begin{aligned} f_2' : \mathbb{R} &\longrightarrow \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right) \\ x &\mapsto \tfrac{1}{x^2+1} \end{aligned}$$

which is invertible for every $x \in \mathbb{R}$.

(iii) Consider f_3, which is not invertible. In fact, it is not injective: $f_3(-\sqrt{3}) = f_3(\sqrt{3}) = 0$. Thus, f_3 is not a diffeomorphism. Correspondingly, the derivative

$$\begin{aligned} f_3' : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto 3\,x^2 - 3 \end{aligned}$$

is not invertible over $\mathbb{R}$, because $3\,x^2 - 3 = 0$ if and only if $x = \pm 1$. However, if we restrict f_3 to the open interval $U = (-1,1)$, then $f_3|_U$ is a C^∞ diffeomorphism from U to $V = (-2,2)$. In this case we say that $f_3|_U$ is a "local diffeomorphism" for the points of U.

■

Solution 5.28 *(Exercise 5.28).* For every $\boldsymbol{x}$ we have

$$((f+g)\circ h)=(f+g)(h(\boldsymbol{x}))=f(h(\boldsymbol{x}))+g(h(\boldsymbol{x}))=(f\circ h)(\boldsymbol{x})+(g\circ h)(\boldsymbol{x}).$$

Moreover,

$$((k\,f)\circ g)(\boldsymbol{x})=(k\,f)(g(\boldsymbol{x}))=k\,f(g(\boldsymbol{x}))=k\,(f\circ g)(\boldsymbol{x}).$$

Now, given $\varphi,\psi\in\mathcal{D}(\boldsymbol{A};\mathbb{C})$ we find immediately

$$\begin{aligned}\langle T\circ F,\varphi+\psi\rangle &= \langle T,((\varphi+\psi)\circ F^{-1})\,|\det J_{F^{-1}}|\rangle\\ &= \langle T,((\varphi\circ F^{-1})+(\psi\circ F^{-1}))\,|\det J_{F^{-1}}|\rangle\\ &= \langle T,(\varphi\circ F^{-1})\,|\det J_{F^{-1}}|\rangle+\langle T,(\psi\circ F^{-1})\,|\det J_{F^{-1}}|\rangle=\langle T\circ F,\varphi\rangle+\langle T\circ F,\psi\rangle.\end{aligned}$$

Finally, for any $k\in\mathbb{C}$ we have

$$\begin{aligned}\langle T\circ F,k\,\varphi\rangle &= \langle T,((k\,\varphi)\circ F^{-1})\,|\det J_{F^{-1}}|\rangle\\ &= \langle T,k\,(\varphi\circ F^{-1})\,|\det J_{F^{-1}}|\rangle\\ &= k\,\langle T,(\varphi\circ F^{-1})\,|\det J_{F^{-1}}|\rangle=k\,\langle T\circ F,\varphi\rangle.\end{aligned}$$

■

Solution 5.29 *(Exercise 5.29).* Let $\varphi\in\mathcal{D}(\boldsymbol{A}';\mathbb{C})$. Then,

$$\begin{aligned}\lim_{k\to\infty}\int_{\boldsymbol{A}'}T_k(F(\boldsymbol{x}'))\,\varphi(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}' &= \lim_{k\to\infty}\int_{\boldsymbol{A}}T_k(\boldsymbol{x})\,\varphi(F^{-1}(\boldsymbol{x}))\,|\det J_{F^{-1}}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{A}}T(\boldsymbol{x})\,\varphi(F^{-1}(\boldsymbol{x}))\,|\det J_{F^{-1}}(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}\\ &= \int_{\boldsymbol{A}'}T(F(\boldsymbol{x}'))\,\varphi(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}',\end{aligned}$$

since we have shown that $\varphi(F^{-1}(\boldsymbol{x}))\,|\det J_{F^{-1}}(\boldsymbol{x})|\in\mathcal{D}(\boldsymbol{A};\mathbb{C})$. ■

Solution 5.30 *(Exercise 5.30).* We can write the pair $(x_1+1,2\,x_1+2\,x_2+2)$ as a column vector

$$\begin{bmatrix}x_1+1\\2\,x_1+2\,x_2+2\end{bmatrix}=\underbrace{\begin{bmatrix}1&0\\2&2\end{bmatrix}}_{A}\begin{bmatrix}x_1\\x_2\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}.$$

Matrix A is invertible: its determinant is 2, and $A^{-1}=\left[\begin{smallmatrix}1&0\\-1&\frac{1}{2}\end{smallmatrix}\right]$. We find

$$\begin{aligned}\langle\delta(x_1+1,2\,x_1+2\,x_2+2),\varphi(x_1,x_2)\rangle &= \frac{1}{2}\left\langle\delta(y_1,y_2),\varphi\left(\begin{bmatrix}1&0\\-1&\frac{1}{2}\end{bmatrix}\begin{bmatrix}x_1-1\\x_2-2\end{bmatrix}\right)\right\rangle\\ &= \frac{1}{2}\left\langle\delta(y_1,y_2),\varphi\left(\begin{bmatrix}x_1-1\\-x_1+1+\frac{1}{2}\,x_2-1\end{bmatrix}\right)\right\rangle\\ &= \frac{1}{2}\left\langle\delta(y_1,y_2),\varphi(x_1-1,-x_1+\frac{1}{2}\,x_2)\right\rangle=\frac{1}{2}\,\varphi(-1,0).\end{aligned}$$

■

Solution 5.31 *(Exercise 5.31).* The function F is a C^∞ diffeomorphism, and

$$\begin{aligned} F^{-1} : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\mapsto \frac{y+3}{2}. \end{aligned}$$

We have

$$\begin{aligned} \langle T \circ F, \varphi \rangle &= \langle T, (\varphi \circ F^{-1}) \, |(F^{-1})'| \rangle \\ &= \lim_{\varepsilon \to 0^+} \int_{|y| \geq \varepsilon} \frac{1}{y} \, \varphi\left(\frac{y+3}{2}\right) \frac{1}{2} \, \mathrm{d}y \\ &= \lim_{\varepsilon \to 0^+} \int_{|2x-3| \geq \varepsilon} \frac{1}{2x-3} \, \varphi(x) \, \mathrm{d}x. \end{aligned}$$

■

Solution 5.32 *(Exercise 5.32).* Let $\varphi \in \mathcal{D}(\boldsymbol{\Theta}; \mathbb{C})$ and let

$$\tilde{\varphi}(\boldsymbol{x}) = \begin{cases} \varphi(\boldsymbol{x}) & \boldsymbol{x} \in \boldsymbol{\Theta} \\ 0 & \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{\Theta} \end{cases}$$

Then

$$\langle T_k|_{\boldsymbol{\Theta}}, \varphi \rangle = \langle T_k, \tilde{\varphi} \rangle \xrightarrow{k \to \infty} \langle T, \tilde{\varphi} \rangle = \langle T|_{\boldsymbol{\Theta}}, \varphi \rangle,$$

which shows that $T_k|_{\boldsymbol{\Theta}} \xrightarrow{k \to \infty} T|_{\boldsymbol{\Theta}}$.

■

Solution 5.33 *(Exercise 5.33).* Let $\alpha, \beta \in \mathbb{C}$ and let $\varphi, \psi \in \mathcal{D}(\boldsymbol{\Theta}; \mathbb{C})$. First, for every $\boldsymbol{x} \in \boldsymbol{\Theta}$ we have

$$(\alpha \varphi + \beta \psi)(\boldsymbol{x}) = \alpha \varphi(\boldsymbol{x}) + \beta \psi(\boldsymbol{x}).$$

Defining $\xi = \alpha \varphi + \beta \psi$, and

$$\begin{aligned} \tilde{\xi} : \boldsymbol{A} &\longrightarrow \mathbb{C}, \\ \boldsymbol{x} &\mapsto \begin{cases} \xi(\boldsymbol{x}) & \boldsymbol{x} \in \boldsymbol{\Theta} \\ 0 & \boldsymbol{x} \in \boldsymbol{A} \setminus \boldsymbol{\Theta}. \end{cases} \end{aligned}$$

we find $\tilde{\xi} = \alpha \tilde{\varphi} + \beta \tilde{\psi}$. Thus,

$$\langle T|_{\boldsymbol{\Theta}}, \alpha \varphi + \beta \psi \rangle = \langle T, \tilde{\xi} \rangle = \langle T, \alpha \tilde{\varphi} + \beta \tilde{\psi} \rangle = \alpha \langle T, \tilde{\varphi} \rangle + \beta \langle T, \tilde{\psi} \rangle = \alpha \langle T|_{\boldsymbol{\Theta}}, \varphi \rangle + \beta \langle T|_{\boldsymbol{\Theta}}, \psi \rangle.$$

■

Solution 5.34 *(Exercise 5.34).* If $\varphi_k \xrightarrow{k \to \infty} \varphi$ in $\mathcal{D}(\boldsymbol{A}; \mathbb{C})$, there is a compact set of $\boldsymbol{A}$ that contains the supports of each φ_k. Moreover, $\partial^\alpha \varphi_k \xrightarrow{k \to \infty} \partial^\alpha \varphi$ uniformly on $\boldsymbol{A}$ for every $\alpha \in \mathbb{N}^n$. On the other hand,

$$\partial^\alpha \varphi_k = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_1^{\alpha_n}} \varphi_k \quad \text{and} \quad \partial^\alpha \varphi = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_1^{\alpha_n}} \varphi.$$

We need to prove that for every $\beta \in \mathbb{N}^n$ we have

$$\partial^\beta \left(\frac{\partial \varphi_k}{\partial x_j} \right) \xrightarrow{k \to \infty} \partial^\beta \left(\frac{\partial \varphi}{\partial x_j} \right)$$

uniformly on $\boldsymbol{A}$. In particular, choosing $\alpha = \boldsymbol{e}_j$ the latter is equivalent to

$$\partial^\beta \left(\partial^\alpha \varphi_k \right) \xrightarrow{k \to \infty} \partial^\beta \left(\partial^\alpha \varphi \right)$$

uniformly on $\boldsymbol{A}$ i.e., $\partial^{\alpha+\beta} \varphi_k \xrightarrow{k \to \infty} \partial^{\alpha+\beta} \varphi$ uniformly on $\boldsymbol{A}$, which is certainly true. Finally, we observe that

$$\operatorname{supp} \left(\frac{\partial \varphi_k}{\partial x_j} \right) \subseteq \operatorname{supp} \varphi_k \subseteq \boldsymbol{K}.$$

■

Solution 5.35 *(Exercise 5.35).* Let $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. Let $\alpha \in \mathbb{N}^n$. We have

$$\langle \partial^\alpha T_k, \varphi \rangle = (-1)^{|\alpha|} \langle T_k, \partial^\alpha \varphi \rangle \xrightarrow{k \to \infty} (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle = \langle \partial^\alpha T, \varphi \rangle,$$

in view of the convergence of $(T_k)_{k \in \mathbb{N}}$ to T, so that $\partial^\alpha T_k \xrightarrow{k \to \infty} \partial^\alpha T$ in $\mathcal{D}'(\boldsymbol{A}; \mathbb{C})$. ■

Solution 5.36 *(Exercise 5.36).* Let $\varphi \in \mathcal{D}(\boldsymbol{A}; \mathbb{C})$. We find

$$\begin{aligned} \langle \partial^\alpha T_1 + \partial^\alpha T_2, \varphi \rangle &= \langle \partial^\alpha T_1, \varphi \rangle + \langle \partial^\alpha T_2, \varphi \rangle \\ &= (-1)^{|\alpha|} \langle T_1, \varphi \rangle + (-1)^{|\alpha|} \langle T_2, \varphi \rangle \\ &= (-1)^{|\alpha|} \langle T_1 + T_2, \varphi \rangle = \langle \partial^\alpha (T_1 + T_2), \varphi \rangle. \end{aligned}$$

We conclude that $\partial^\alpha T_1 + \partial^\alpha T_2 = \partial^\alpha (T_1 + T_2)$. ■

Solution 5.37 *(Exercise 5.37).* Since $\sum_{k=0}^{\infty} T_k$ is convergent, defining the corresponding partial sums as

$$\begin{aligned} s_0 &= T_0 \\ s_1 &= T_0 + T_1 \\ &\vdots \\ s_k &= T_0 + T_1 + \ldots + T_k, \end{aligned}$$

the sequence $(s_k)_{k \in \mathbb{N}}$ is convergent and its limit is the value of the series, i.e.,

$$\sum_{k=0}^{\infty} T_k = \lim_{k \to \infty} s_k.$$

Let $\alpha \in \mathbb{N}^n$. We define the partial sums of the series $\sum_{k=0}^{\infty} \partial^\alpha T_k$ as

$$\begin{aligned} \sigma_0 &= \partial^\alpha T_0 \\ \sigma_1 &= \partial^\alpha T_0 + \partial^\alpha T_1 \\ &\vdots \\ \sigma_k &= \partial^\alpha T_0 + \partial^\alpha T_1 + \ldots + \partial^\alpha T_k, \end{aligned}$$

and we have to show that the sequence $(\sigma_k)_{k\in\mathbb{N}}$ is convergent in $\mathcal{D}'(\boldsymbol{A};\mathbb{C})$ and its limit is $\partial^\alpha \sum_{k=0}^{\infty} T_k$. Let $\varphi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then, noting that for any $k \in \mathbb{N}$ we have $\sigma_k = \partial^\alpha s_k$ from the result of the previous exercise, then

$$\begin{aligned}\langle \sigma_k, \varphi\rangle &= \langle \partial^\alpha s_k, \varphi\rangle \\ &= (-1)^{|\alpha|}\langle s_k, \partial^\alpha\varphi\rangle \xrightarrow{k\to\infty} (-1)^{|\alpha|}\left\langle \sum_{k=0}^{\infty} T_k, \partial^\alpha\varphi\right\rangle = \left\langle \partial^\alpha\left(\sum_{k=0}^{\infty} T_k\right), \varphi\right\rangle. \end{aligned} \tag{14.2}$$

It follows that $(\sigma_k)_{k\in\mathbb{N}}$ is convergent in $\mathcal{D}'(A;\mathbb{C})$. Thus,

$$\lim_{k\to\infty}\sigma_k = \sum_{k=0}^{\infty}\partial^\alpha T_k,$$

or, which is the same,

$$\langle \sigma_k, \varphi\rangle \xrightarrow{k\to\infty} \left\langle \sum_{k=0}^{\infty}\partial^\alpha T_k, \varphi\right\rangle. \tag{14.3}$$

From (14.2-14.3) and the uniqueness of the limit, we obtain

$$\sum_{k=0}^{\infty}\partial^\alpha T_k = \partial^\alpha\sum_{k=0}^{\infty} T_k.$$

■

Solution 5.38 *(Exercise 5.38).* Let $\alpha,\beta \in \mathbb{C}$ and let $\varphi,\psi \in \mathcal{D}(\boldsymbol{A};\mathbb{C})$. Then, using the linearity of the partial derivative, we obtain

$$\begin{aligned}\langle \partial_j T, \alpha\,\varphi + \beta\,\psi\rangle &= -\langle T, \partial_j(\alpha\,\varphi + \beta\,\psi)\rangle \\ &= -\langle T, \alpha\,\partial_j\varphi + \beta\,\partial_j\psi\rangle \\ &= -\alpha\,\langle T, \partial_j\varphi\rangle - \beta\,\langle T, \partial_j\psi\rangle \\ &= \alpha\,\langle \partial_j T, \varphi\rangle + \beta\,\langle \partial_j T, \psi\rangle.\end{aligned}$$

■

Solution 5.39 *(Exercise 5.39).* Since the sign function is locally integrable, we expect it to be differentiable in the distributional sense, and we will find that indeed $f' = 2\,\delta$ in the sense of $\mathcal{D}'(\mathbb{R};\mathbb{R})$. To see it, if $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, we know there exists $M > 0$ such that $\operatorname{supp}\varphi \subseteq [-M, M]$, and we find

$$\begin{aligned}\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi'\rangle \\ &= -\int_{\mathbb{R}}\varphi'(t)\,\operatorname{sgn} t\,\mathrm{d}t = -\left[\int_0^{\infty}\varphi'(t)\,\mathrm{d}t - \int_{-\infty}^{0}\varphi'(t)\,\mathrm{d}t\right] \\ &= -(-\varphi(0) - \varphi(0)) = 2\,\varphi(0) = \langle \delta, 2\,\varphi\rangle = \langle 2\,\delta, \varphi\rangle,\end{aligned}$$

so that $f' = 2\,\delta$. ■

Solution 5.40 *(Exercise 5.40).* We assume that T is even, so that $T(-\boldsymbol{x}) = T(\boldsymbol{x})$. We prove that $T'(-\boldsymbol{x}) = -T'(\boldsymbol{x})$. Stated differently, using the C^∞ diffeomorphism

$$\begin{aligned} F : A &\longrightarrow A \\ \boldsymbol{x} &\mapsto -\boldsymbol{x} \end{aligned}$$

we know that $T \circ F = T$ and we prove that $T' \circ F = -T'$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We find

$$\begin{aligned} \langle T' \circ F, \varphi \rangle &= \langle T', \varphi \circ F^{-1} \rangle \\ &= -\langle T, (\varphi' \circ F^{-1})(F^{-1})' \rangle \\ &= -\langle T, -(\varphi' \circ F^{-1}) \rangle \\ &= \langle T, \varphi' \circ F^{-1} \rangle \\ &= \langle T \circ F, \varphi' \rangle = \langle T, \varphi' \rangle = -\langle T', \varphi \rangle. \end{aligned}$$

where we have used the fact that $|(F^{-1})'| = 1$ and that $(F^{-1})' = -1$. Alternatively, we can use the more intuitive (but slightly sloppier) notation

$$\begin{aligned} \left\langle \frac{\partial T}{\partial x}(x), \varphi(x) \right\rangle &= -\left\langle T(x), \frac{\partial \varphi(x)}{\partial x} \right\rangle \\ &= -\left\langle T(-x), \frac{\partial \varphi(x)}{\partial x} \right\rangle \\ &= -\left\langle T(x), -\frac{\partial \varphi(-x)}{\partial x} \right\rangle \\ &= \left\langle T(x), \frac{\partial \varphi(-x)}{\partial x} \right\rangle \\ &= -\left\langle \frac{\partial T}{\partial x}(x), \varphi(-x) \right\rangle \\ &= -\left\langle \frac{\partial T}{\partial x}(-x), \varphi(x) \right\rangle = \left\langle -\frac{\partial T}{\partial x}(-x), \varphi(x) \right\rangle. \end{aligned}$$

■

Solution 5.41 *(Exercise 5.41).* We assume that T is odd, so that $T(-\boldsymbol{x}) = -T(\boldsymbol{x})$. We prove that $T'(-\boldsymbol{x}) = T'(\boldsymbol{x})$. Stated differently, using the C^∞ diffeomorphism

$$\begin{aligned} F : \Omega &\longrightarrow \Omega \\ \boldsymbol{x} &\mapsto -\boldsymbol{x} \end{aligned}$$

we know that $T \circ F = -T$ and we prove that $T' \circ F = T'$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We find

$$\begin{aligned} \langle T' \circ F, \varphi \rangle &= \langle T', \varphi \circ F^{-1} \rangle \\ &= -\langle T, (\varphi \circ F^{-1})' \rangle \\ &= -\langle T, (\varphi' \circ F^{-1})(F^{-1})' \rangle \\ &= -\langle T, -\varphi' \circ F^{-1} \rangle \\ &= \langle T, \varphi' \circ F^{-1} \rangle \\ &= \langle T \circ F, \varphi' \rangle \qquad \text{(and since } T \circ F = -T\text{)} \\ &= \langle -T, \varphi' \rangle = \langle T', \varphi \rangle. \end{aligned}$$

where we have used the fact that $|(F^{-1})'| = 1$ and that $(F^{-1})' = -1$. ■

Solution 5.42 *(Exercise 5.42).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then,

$$\begin{aligned}
\left\langle x^2 \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{pv}\,\frac{1}{x}\right), \varphi \right\rangle &= \left\langle \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{pv}\,\frac{1}{x}\right), x^2\,\varphi(x) \right\rangle \\
&= -\left\langle \mathrm{pv}\,\frac{1}{x}, 2\,x\,\varphi(x) + x^2\,\varphi'(x) \right\rangle \\
&= -\left\langle \mathrm{pv}\,\frac{1}{x}, 2\,x\,\varphi(x) \right\rangle - \left\langle \mathrm{pv}\,\frac{1}{x}, x^2\,\varphi'(x) \right\rangle \\
&= -\lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} 2\,\varphi(x)\,\mathrm{d}x - \lim_{\varepsilon\to 0^+} \int_{|x|\geq\varepsilon} x\,\varphi'(x)\,\mathrm{d}x \\
&= -\int_{-\infty}^{+\infty} 2\,\varphi(x)\,\mathrm{d}x - \int_{-\infty}^{+\infty} x\,\varphi'(x)\,\mathrm{d}x \\
&= -2\int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x - \Big(\underbrace{[x\,\varphi(x)]_{-\infty}^{+\infty}}_{=0} - \int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x\Big) = -\int_{-\infty}^{+\infty} \varphi(x)\,\mathrm{d}x = -\int_{\operatorname{supp}\varphi} \varphi(x)\,\mathrm{d}x = \langle -1, \varphi\rangle.
\end{aligned}$$

It follows that

$$x^2 \frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{pv}\,\frac{1}{x}\right) = -1.$$

■

Solution 5.43 *(Exercise 5.43).* We use the spherical coordinates

$$\begin{cases} x_1 = \rho\,\sin\theta\,\cos\varphi \\ x_2 = \rho\,\sin\theta\,\sin\varphi \\ x_3 = \rho\,\cos\theta, \end{cases} \qquad \text{with } \rho\in(0,1],\ \ \theta\in(0,\pi),\ \ \varphi\in(0,2\,\pi).$$

The determinant of the Jacobian is

$$\begin{vmatrix} \sin\theta\,\cos\varphi & \rho\,\cos\theta\,\cos\varphi & -\rho\,\sin\theta\,\sin\varphi \\ \sin\theta\,\sin\varphi & \rho\,\cos\theta\,\sin\varphi & \rho\,\sin\theta\,\cos\varphi \\ \cos\theta & -\rho\,\sin\theta & 0 \end{vmatrix} = \rho^2\,\sin\theta.$$

Using Fubini's theorem (Theorem 3.8), we find

$$\begin{aligned}
\int_{\mathbb{R}^3} \|\boldsymbol{x}\|^\alpha\,\mathrm{d}\boldsymbol{x} &= \int_{(0,1)\times(0,\pi)\times(0,2\,\pi)} \rho^\alpha\,\rho^2\,\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\
&= 2\,\pi \int_{(0,1)\times(0,\pi)} \rho^{\alpha+2}\,\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\,\pi \int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho \cdot \int_0^\pi \sin\theta\,\mathrm{d}\theta = 4\,\pi \int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho.
\end{aligned}$$

The last integral is improper if $\alpha \leq -3$, and it does not converge if $\alpha \leq -3$. If $\alpha > -3$, we have

$$\int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho = \left[\frac{\rho^{\alpha+3}}{\alpha+3}\right]_0^1 = \frac{1}{\alpha+3}.$$

■

Solution 5.44 *(Exercise 5.44).* Consider

$$\begin{aligned} F : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x + t_0. \end{aligned}$$

Then, T is t_0-periodic if $T \circ F = T$. We prove that if T is t_0-periodic, then T' is t_0-periodic, i.e., we have to prove that $T' \circ F = T'$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then

$$\begin{aligned}\langle T' \circ F, \varphi\rangle &= \langle T', (\varphi \circ F^{-1})\,|1|\rangle \\ &= \langle T', \varphi \circ F^{-1}\rangle \\ &= -\langle T, (\varphi' \circ F^{-1})\,(F^{-1})'\rangle \\ &= -\langle T, \varphi' \circ F^{-1}\rangle \\ &= -\langle T \circ F, \varphi'\rangle = -\langle T, \varphi'\rangle = \langle T', \varphi\rangle.\end{aligned}$$

Alternatively, we can write

$$\begin{aligned}\langle T', \varphi\rangle &= -\langle T, \varphi'\rangle \\ &= -\langle T(x + t_0), \varphi'(x)\rangle \\ &= -\langle T(x), \varphi'(x - t_0)\rangle \\ &= \langle T'(x), \varphi'(x - t_0)\rangle = \langle T'(x + t_0), \varphi(x)\rangle,\end{aligned}$$

which shows that $T'(x) = T'(x + t_0)$.

■

Solution 5.45 *(Exercise 5.45).* Consider the change of variable $\boldsymbol{y} = 2^{j-1}\,\boldsymbol{x}$, so that $\boldsymbol{x} = \frac{\boldsymbol{y}}{2^{1-j}}$. We obtain

$$\int_{2^{-j}<\|\boldsymbol{x}\|\le 2^{1-j}} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x} = \int_{\frac{1}{2}<\|\boldsymbol{y}\|\le 1} \frac{\|\boldsymbol{y}\|^{\alpha}}{|2^{j-1}|^{\alpha}}\,\frac{1}{(2^{j-1})^{n}}\,\mathrm{d}\boldsymbol{y} = 2^{-(n+\alpha)(j-1)} \int_{\frac{1}{2}<\|\boldsymbol{x}\|\le 1} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x}.$$

Since for every $j \in \mathbb{N}^*$ the integration sets are pairwise disjoint, because

$$\begin{aligned} j = 1 \quad &\Rightarrow \quad \frac{1}{2} < \|\boldsymbol{x}\| \le 1 \\ j = 2 \quad &\Rightarrow \quad \frac{1}{4} < \|\boldsymbol{x}\| \le \frac{1}{2} \\ j = 3 \quad &\Rightarrow \quad \frac{1}{8} < \|\boldsymbol{x}\| \le \frac{1}{4} \\ &\vdots \end{aligned}$$

we obtain

$$\begin{aligned}\int_{0<\|\boldsymbol{x}\|\le 1} \|\boldsymbol{x}\|\,\mathrm{d}\boldsymbol{x} &= \sum_{j=1}^{\infty} \int_{2^{-j}<\|\boldsymbol{x}\|\le 2^{1-j}} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\frac{1}{2}<\|\boldsymbol{x}\|\le 1} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x} \sum_{j=1}^{\infty} 2^{-(n+\alpha)(j-1)} \\ &= \int_{\frac{1}{2}<\|\boldsymbol{x}\|\le 1} \|\boldsymbol{x}\|^{\alpha}\,\mathrm{d}\boldsymbol{x}\, 2^{n+\alpha} \sum_{j=1}^{\infty} 2^{-(\alpha+n)\,j}.\end{aligned}$$

The geometric series converges if and only if $2^{(\alpha+n)} < 1$, i.e., if and only if $\alpha + n < 0$, i.e., if and only if $\alpha < -n$. ■

Solution 5.46 *(Exercise 5.46).* Let $T = \chi_{\{(x,y)\in\mathbb{R}^2 \,|\, x>0 \text{ and } y>0\}}$. Let $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$. Then, the partial derivative of T with respect to x is

$$\begin{aligned}
\left\langle \frac{\partial T}{\partial x}(x,y), \varphi(x,y) \right\rangle &= -\left\langle T(x,y), \frac{\partial \varphi}{\partial x}(x,y) \right\rangle \\
&= -\int_{\mathbb{R}^2} \chi_{\{(x,y)\in\mathbb{R}^2 \,|\, x>0 \text{ and } y>0\}}(x,y)\, \frac{\partial \varphi}{\partial x}(x,y)\, \mathrm{d}x\, \mathrm{d}y \\
&= -\int_0^{+\infty} \left(\int_0^{+\infty} \frac{\partial \varphi}{\partial x}(x,y)\, \mathrm{d}x \right) \mathrm{d}y \\
&= -\int_0^{+\infty} \left[\varphi(x,y) \right]_{x=0}^{x=+\infty} \mathrm{d}y \\
&= -\int_0^{+\infty} -\varphi(0,y)\, \mathrm{d}y \\
&= \int_0^{+\infty} \varphi(0,y)\, \mathrm{d}y \\
&= \int_{\mathbb{R}} H(y)\, \varphi(0,y)\, \mathrm{d}y \\
&= \langle H(y), \langle \delta(x), \varphi(x,y) \rangle \rangle \\
&= \langle (\delta \otimes H)(x,y), \varphi(x,y) \rangle \rangle.
\end{aligned}$$

It follows that $\frac{\partial T}{\partial x} = \delta \otimes H$. Now we compute the partial derivative of T with respect to y:

$$\begin{aligned}
\left\langle \frac{\partial T}{\partial y}(x,y), \varphi(x,y) \right\rangle &= -\left\langle T(x,y), \frac{\partial \varphi}{\partial y}(x,y) \right\rangle \\
&= -\int_{\mathbb{R}^2} \chi_{\{(x,y)\in\mathbb{R}^2 \,|\, x>0 \text{ and } y>0\}}(x,y)\, \frac{\partial \varphi}{\partial y}(x,y)\, \mathrm{d}x\, \mathrm{d}y \\
&= -\int_0^{+\infty} \left(\int_0^{+\infty} \frac{\partial \varphi}{\partial y}(x,y)\, \mathrm{d}y \right) \mathrm{d}x \\
&= -\int_0^{+\infty} \left[\varphi(x,y) \right]_{y=0}^{y=+\infty} \mathrm{d}x \\
&= -\int_0^{+\infty} -\varphi(x,0)\, \mathrm{d}x \\
&= \int_0^{+\infty} \varphi(x,0)\, \mathrm{d}x \\
&= \int_{\mathbb{R}} H(x)\, \varphi(x,0)\, \mathrm{d}x \\
&= \langle H(x), \langle \delta(y), \varphi(x,y) \rangle \rangle = \langle (H \otimes \delta)(x,y), \varphi(x,y) \rangle \rangle.
\end{aligned}$$

Therefore, $\frac{\partial T}{\partial y} = H \otimes \delta$. ■

Solution 5.47 *(Exercise 5.47).* Let $\varphi \in \mathcal{D}((0,3);\mathbb{R})$. Then

$$\begin{aligned}\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi'\rangle = -\int_0^3 \lceil t\rceil\, \varphi'(t)\,\mathrm{d}t \\ &= -\left(\int_0^1 1\cdot\varphi'(t)\,\mathrm{d}t + \int_1^2 2\cdot\varphi'(t)\,\mathrm{d}t + \int_2^3 3\cdot\varphi'(t)\,\mathrm{d}t\right) \\ &= -\varphi(1) + \cancel{\varphi(0)} - 2\,\varphi(2) + 2\,\varphi(1) - 3\,\cancel{\varphi(3)} + 3\,\varphi(2) \\ &= \varphi(1) + \varphi(2) = \langle \delta(t-1), \varphi(t)\rangle + \langle \delta(t-2), \varphi(t)\rangle.\end{aligned}$$

Thus, $f'(t) = \delta(t-1) + \delta(t-2)$. ■

Solution 5.48 *(Exercise 5.48).* The function f is in $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Given $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$, we have

$$\begin{aligned}\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi\rangle = -\int_{\mathbb{R}} x\,|x|\,\varphi'(x)\,\mathrm{d}x \\ &= -\int_{-\infty}^0 -x^2\,\varphi'(x)\,\mathrm{d}x - \int_0^{+\infty} x^2\,\varphi'(x)\,\mathrm{d}x \\ &= \int_{-\infty}^0 x^2\,\varphi'(x)\,\mathrm{d}x - \int_0^{+\infty} x^2\,\varphi'(x)\,\mathrm{d}x \\ &= [x^2\,\varphi(x)]_{-\infty}^0 - \int_{-\infty}^0 2\,x\,\varphi(x)\,\mathrm{d}x - \left([x^2\,\varphi(x)]_0^{+\infty} - \int_0^{+\infty} 2\,x\,\varphi(x)\,\mathrm{d}x\right) \\ &= -2\int_{-\infty}^0 x\,\varphi(x)\,\mathrm{d}x + 2\int_0^{+\infty} x\,\varphi(x)\,\mathrm{d}x = 2\int_{\mathbb{R}} x\,\operatorname{sgn} x\,\varphi(x)\,\mathrm{d}x.\end{aligned}$$

The first derivative is therefore $2\,x\,\operatorname{sgn} x = 2\,|x|$. The second derivative is

$$\begin{aligned}-\int_{\mathbb{R}} 2\,|x|\,\varphi'(x)\,\mathrm{d}x &= -\left(\int_{-\infty}^0 -2\,x\,\varphi'(x)\,\mathrm{d}x + \int_0^{+\infty} 2\,x\,\varphi'(x)\,\mathrm{d}x\right) \\ &= \int_{-\infty}^0 2\,x\,\varphi'(x)\,\mathrm{d}x - \int_0^{+\infty} 2\,x\,\varphi'(x)\,\mathrm{d}x \\ &= [2\,x\,\varphi(x)]_{-\infty}^0 - \int_{-\infty}^0 2\,\varphi(x)\,\mathrm{d}x - \left([2\,x\,\varphi(x)]_0^{+\infty} - \int_0^{+\infty} 2\,\varphi(x)\,\mathrm{d}x\right) \\ &= -2\int_{-\infty}^0 \varphi(x)\,\mathrm{d}x + 2\int_0^{+\infty} \varphi(x)\,\mathrm{d}x = 2\int_{\mathbb{R}} \operatorname{sgn} x\,\varphi(x)\,\mathrm{d}x.\end{aligned}$$

Hence, $f'' = 2\,\operatorname{sgn} x$. The third derivative is

$$\begin{aligned}-\int_{\mathbb{R}} 2\,\operatorname{sgn} x\,\varphi'(x)\,\mathrm{d}x &= 2\int_{-\infty}^0 \varphi'(x)\,\mathrm{d}x - 2\int_0^{+\infty} \varphi'(x)\,\mathrm{d}x \\ &= 2\,[\varphi(x)]_{-\infty}^0 - 2\,[\varphi(x)]_0^{+\infty} = 2\,\varphi(0) + 2\,\varphi(0) = 4\,\delta(\varphi).\end{aligned}$$

It follows that $f''' = 4\,\delta$. ■

Solution 5.49 *(Exercise 5.49).* Let $T(x) = e^{-x}\,H(x)$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then

$$\begin{aligned}\langle T', \varphi\rangle &= -\langle T, \varphi'\rangle \\ &= -\int_{\mathbb{R}} e^{-x}\,H(x)\,\varphi'(x)\,\mathrm{d}x \\ &= -\int_0^{+\infty} e^{-x}\,\varphi'(x)\,\mathrm{d}x \\ &= -\left[e^{-x}\,\varphi(x)\right]_0^{+\infty} - \int_0^{+\infty} e^{-x}\,\varphi(x)\,\mathrm{d}x \\ &= \varphi(0) + \int_{\mathbb{R}} -e^{-x}\,H(x)\,\varphi(x)\,\mathrm{d}x.\end{aligned}$$

It follows that

$$T'(x) = \delta(x) - e^{-x}\,H(x).$$

The same result can be obtained with the δ-jumps.

We now find the second derivative:

$$\begin{aligned}\langle T'', \varphi\rangle &= -\langle T', \varphi'\rangle \\ &= \langle \delta', \varphi\rangle + \int_{\mathbb{R}} e^{-x}\,H(x)\,\varphi'(x)\,\mathrm{d}x \\ &= \langle \delta', \varphi\rangle + \int_0^{+\infty} e^{-x}\,\varphi'(x)\,\mathrm{d}x \\ &= \langle \delta', \varphi\rangle + \left[e^{-x}\,\varphi(x)\right]_0^{+\infty} + \int_0^{+\infty} e^{-x}\,\varphi(x)\,\mathrm{d}x = \langle \delta', \varphi\rangle - \langle \delta, \varphi\rangle + \langle e^{-x}\,H(x), \varphi(x)\rangle,\end{aligned}$$

so that

$$T'' = \delta' - \delta + e^{-x}\,H.$$

The same result can be obtained using the δ-jumps. ■

Solution 5.50 *(Exercise 5.50).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{C})$. Then

$$\begin{aligned}\langle e^x\,\delta'', \varphi\rangle &= \langle \delta'', e^x\,\varphi\rangle = \left\langle \delta, \frac{\mathrm{d}^2}{\mathrm{d}x^2}(e^x\,\varphi)\right\rangle \\ &= \langle \delta, e^x\,\varphi + e^x\,\varphi' + e^x\,\varphi' + e^x\,\varphi''\rangle = \langle \delta, \varphi\rangle + 2\,\langle \delta, \varphi'\rangle + \langle \delta, \varphi''\rangle \\ &= \langle \delta - 2\,\delta' + \delta'', \varphi\rangle.\end{aligned}$$

■

Solution 5.51 *(Exercise 5.51).* First, note that $[f] \in L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. Notice that

$$H(x) - H(x-2) = \begin{cases} 1 & \text{if } 0 \le x < 2 \\ 0 & \text{otherwise}\end{cases}$$

Hence, given $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$,

$$\begin{aligned}\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi'\rangle = -\int_0^2 (x^2 - x - 2)\,\varphi'(x)\,\mathrm{d}x = \left[-(x^2 - x - 2)\,\varphi(x)\right]_0^2 + \int_0^2 (2\,x - 1)\,\varphi(x)\,\mathrm{d}x \\ &= (2 + 2 - 4)\,\varphi(2) - 2\,\varphi(0) + \int_0^2 (2\,x - 1)\,\varphi\,\mathrm{d}x.\end{aligned}$$

It follows that

$$L_{f'} = (2\,x - 1)\,(H(x) - H(x-2)) - 2\,\delta.$$

We now compute the second derivative:

$$\begin{aligned}
\langle L_{f''}, \varphi\rangle &= -\int_0^2 (2\,x-1)\,\varphi'(x)\,\mathrm{d}x - 2\,\langle \delta', \varphi\rangle \\
&= -\left[(2\,x-1)\,\varphi(x)\right]_0^2 + \int_0^2 2\,\varphi(x)\,\mathrm{d}x - 2\,\langle \delta', \varphi\rangle \\
&= -3\,\varphi(2) - \varphi(0) + 2\,\langle H(x) - H(x-2), \varphi\rangle - 2\,\langle \delta', \varphi\rangle.
\end{aligned}$$

It follows that

$$L_{f''} = 2\,(H(x) - H(x-2) - \delta') - 3\,\delta(x-2) - \delta.$$

■

Solution 5.52 *(Exercise 5.52).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then

$$\begin{aligned}
\langle L_{f'}, \varphi\rangle &= -\langle L_f, \varphi'\rangle = -\int_{\mathbb{R}} f(x)\,\varphi'(x)\,\mathrm{d}x \\
&= -\int_0^{+\infty} \ln(3\,x)\,\varphi'(x)\,\mathrm{d}x - \int_{-\infty}^0 \ln(-2\,x)\,\varphi'(x)\,\mathrm{d}x \\
&= -\int_0^{+\infty} (\ln 3 + \ln\,x)\,\varphi'(x)\,\mathrm{d}x - \int_{-\infty}^0 (\ln 2 + \ln(-x))\,\varphi'(x)\,\mathrm{d}x \\
&= -\ln 3 \int_0^{+\infty} \varphi'(x)\,\mathrm{d}x - \int_0^{+\infty} \ln\,x\,\varphi'(x)\,\mathrm{d}x - \ln 2 \int_{-\infty}^0 \varphi'(x)\,\mathrm{d}x - \int_{-\infty}^0 \ln(-x)\,\varphi'(x)\,\mathrm{d}x \\
&= \ln 3\,\varphi(0) - \int_{\mathbb{R}} \ln\,|x|\,\varphi'(x)\,\mathrm{d}x - \ln 2\,\varphi(0) \\
&= \ln\frac{3}{2}\,\varphi(0) - \int_{\mathbb{R}} \ln\,|x|\,\varphi'(x)\,\mathrm{d}x.
\end{aligned}$$

Hence,

$$f' = \mathrm{pv}\,\frac{1}{x} + \left(\ln\frac{3}{2}\right)\delta.$$

■

Solution 5.53 *(Exercise 5.53).* Consider $\varphi \in \mathcal{D}(\boldsymbol{A}_1 \times \boldsymbol{A}_2;\mathbb{C})$. We have

$$\begin{aligned}
\langle \partial_{\boldsymbol{x}}^{\alpha}\partial_{\boldsymbol{y}}^{\beta}(T_1 \otimes T_2), \varphi\rangle &= (-1)^{|\alpha|}\,(-1)^{|\beta|}\,\langle T_1 \otimes T_2, \partial_{\boldsymbol{y}}^{\beta}\,\partial_{\boldsymbol{x}}^{\alpha}\,\varphi\rangle \\
&= (-1)^{|\alpha|+|\beta|}\,\langle T_1(\boldsymbol{x}), \langle T_2(\boldsymbol{y}), \partial_{\boldsymbol{y}}^{\beta}\,\partial_{\boldsymbol{x}}^{\alpha}\,\varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \\
&= (-1)^{|\alpha|}\,\langle T_1(\boldsymbol{x}), \langle \partial_{\boldsymbol{y}}^{\beta}\,T_2(\boldsymbol{y}), \partial_{\boldsymbol{x}}^{\alpha}\,\varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \\
&= (-1)^{|\alpha|}\,\langle T_1(\boldsymbol{x}), \partial_{\boldsymbol{x}}^{\alpha}\,\langle \partial_{\boldsymbol{y}}^{\beta}T_2(\boldsymbol{y}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \\
&= \langle \partial_{\boldsymbol{x}}^{\alpha}\,T_1(\boldsymbol{x}), \langle \partial_{\boldsymbol{y}}^{\beta}T_2(\boldsymbol{y}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle\rangle \\
&= \langle \partial_{\boldsymbol{x}}^{\alpha}\,T_1(\boldsymbol{x}) \otimes \partial_{\boldsymbol{y}}^{\beta}T_2(\boldsymbol{y}), \varphi(\boldsymbol{x},\boldsymbol{y})\rangle = \langle \partial_{\boldsymbol{x}}^{\alpha}\,T_1 \otimes \partial_{\boldsymbol{y}}^{\beta}T_2, \varphi\rangle.
\end{aligned}$$

■

Solution 5.54 *(Exercise 5.54).* With the usual abuse of notation, we write

$$\frac{\partial^2(T\otimes S)}{\partial x\,\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}T(x)\,S(y)\right) = \frac{\partial}{\partial y}S(y)\,T'(x) = T'(x)\,S'(y) = T'\otimes S'.$$

We apply this formula to $T = S = H$:

$$\frac{\partial^2(T\otimes S)}{\partial x\,\partial y} = H'\otimes H' = \delta\otimes\delta = \delta(x)\,\delta(y) = \delta(x,y).$$

■

Solution 5.55 *(Exercise 5.55).* First, notice that for each $k\in\mathbb{N}$, the equivalence class that corresponds to the function $\frac{k}{k^2\,x^2+1}$ is in $L^1_{\mathrm{loc}}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned}\lim_{k\to\infty}\int_{\mathbb{R}}\frac{k}{k^2\,x^2+1}\,\varphi(x)\,\mathrm{d}x &= \lim_{k\to\infty}\int_{\mathbb{R}}\frac{k}{y^2+1}\,\varphi\left(\frac{y}{k}\right)\frac{1}{k}\,\mathrm{d}y\\ &= \lim_{k\to\infty}\int_{\mathbb{R}}\frac{1}{y^2+1}\,\varphi\left(\frac{y}{k}\right)\mathrm{d}y\\ &= \int_{\mathbb{R}}\frac{1}{y^2+1}\,\varphi(0)\,\mathrm{d}y = \varphi(0)\int_{\mathbb{R}}\frac{1}{y^2+1}\,\mathrm{d}y = \varphi(0)\,\pi,\end{aligned}$$

where Lebesgue dominated convergence theorem (Theorem 3.47) has been applied. In fact, denoting by K a compact set containing the support of φ we have

$$\left|\frac{1}{y^2+1}\,\varphi\left(\frac{y}{k}\right)\right| \le \max_K|\varphi|\,\frac{1}{y^2+1},$$

which is summable over $\mathbb{R}$. ■

Solution 5.56 *(Exercise 5.56).* Let $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$. The function

$$\begin{aligned}F:\mathbb{R} &\longrightarrow \mathbb{R}\\ x &\mapsto 2\,k\,x-1,\end{aligned}$$

is a C^∞-diffeomorphism, and

$$\begin{aligned}F^{-1}:\mathbb{R} &\longrightarrow \mathbb{R}\\ y &\mapsto \frac{y+1}{2\,k}.\end{aligned}$$

Then

$$\begin{aligned}\lim_{k\to\infty}\langle T_k,\varphi\rangle &= \lim_{k\to\infty}\langle k^2\,x\,\delta(2\,k\,x-1),\varphi(x)\rangle\\ &= \lim_{k\to\infty}\langle\delta\circ F, k^2\,x\,\varphi(x)\rangle\\ &= \lim_{k\to\infty}\langle\delta, k^2\,F^{-1}(y)\,\varphi(F^{-1}(y))\,|(F^{-1})'|\rangle\\ &= \lim_{k\to\infty}\left\langle\delta, \cancel{k^2}\,\frac{y+1}{2\cancel{k}}\,\varphi\left(\frac{y+1}{2\,k}\right)\frac{1}{2\cancel{k}}\right\rangle\\ &= \lim_{k\to\infty}\frac{1}{4}\left\langle\delta,(y+1)\,\varphi\left(\frac{y+1}{2\,k}\right)\right\rangle = \lim_{k\to\infty}\frac{1}{4}\left\langle\delta,\varphi\left(\frac{y}{2\,k}\right)\right\rangle = \frac{1}{4}\,\varphi(0).\end{aligned}$$

Then,

$$\lim_{k\to\infty} T_k = \frac{1}{4}\,\delta.$$

■

Solution 5.57 *(Exercise 5.57).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. The function

$$\begin{aligned} F_1 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x - \frac{1}{k}, \end{aligned}$$

is a C^∞-diffeomorphism, and

$$\begin{aligned} F_1^{-1} : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\mapsto y + \frac{1}{k}. \end{aligned}$$

Likewise,

$$\begin{aligned} F_2 : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x + \frac{5}{k}, \end{aligned}$$

is a C^∞-diffeomorphism, and

$$\begin{aligned} F_2^{-1} : \mathbb{R} &\longrightarrow \mathbb{R} \\ y &\mapsto y - \frac{5}{k}. \end{aligned}$$

Hence

$$\begin{aligned} \lim_{k\to\infty} \left\langle k\left(\delta\Big(x-\frac{1}{k}\Big) - \delta\Big(x+\frac{5}{k}\Big)\right), \varphi\right\rangle &= \lim_{k\to\infty} \left[\left\langle k\,\delta\Big(x-\frac{1}{k}\Big), \varphi\right\rangle + \left\langle k\,\delta\Big(x+\frac{5}{k}\Big), \varphi\right\rangle\right] \\ &= \lim_{k\to\infty} \left[\left\langle \delta, k\,\varphi\left(y+\frac{1}{y}\right)\right\rangle - \left\langle \delta, k\,\varphi\left(y-\frac{5}{y}\right)\right\rangle\right] \\ &= \lim_{k\to\infty} k\left[\varphi\Big(\frac{1}{k}\Big) - \varphi\Big(-\frac{5}{k}\Big)\right] \\ &= \lim_{k\to\infty} \frac{\varphi\Big(\frac{1}{k}\Big) - \varphi\Big(-\frac{5}{k}\Big)}{\frac{1}{k}} \\ &= \lim_{k\to\infty} \frac{\varphi\Big(\frac{1}{k}\Big) - \varphi(0)}{\frac{1}{k}} + 5\,\frac{\varphi\Big(-\frac{5}{k}\Big) - \varphi(0)}{-\frac{5}{k}} = 6\,\varphi'(0) = -\langle \delta', \varphi\rangle. \end{aligned}$$

It follows that

$$\lim_{k\to\infty} k\left(\delta\Big(x-\frac{1}{k}\Big) - \delta\Big(x+\frac{5}{k}\Big)\right) = -6\,\delta'.$$

■

Solution 5.58 *(Exercise 5.58).* Let $\varphi \in \mathcal{D}(\mathbb{R}^3;\mathbb{R})$. Then,

$$\lim_{\varepsilon\to 0}\langle f_\varepsilon,\varphi\rangle = \lim_{\varepsilon\to 0}\int_{1\le y^2+z^2\le(1+\varepsilon)^2}\frac{1}{\varepsilon}\,\varphi(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

We can use cylindrical coordinates:

$$\begin{cases} x=\xi \\ y=\rho\,\sin\theta \\ z=\rho\,\cos\theta \end{cases} \qquad \xi\in(-\infty,\infty),\quad \rho\in[1,1+\varepsilon],\quad \theta\in[0,2\pi].$$

The determinant of the Jacobian matrix is

$$\det\begin{bmatrix} 1 & 0 & 0 \\ 0 & \sin\theta & \rho\,\cos\theta \\ 0 & \cos\theta & -\rho\,\sin\theta \end{bmatrix} = -\rho.$$

We find

$$\lim_{\varepsilon\to 0}\langle f_\varepsilon,\varphi\rangle = \lim_{\varepsilon\to 0}\int_1^{1+\varepsilon}\Big(\int_0^{2\pi}\int_{-\infty}^{+\infty}\frac{\rho}{\varepsilon}\,\varphi(\xi,\rho\,\sin\theta,\rho\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta\Big)\,\mathrm{d}\rho.$$

We define

$$\begin{aligned}\Psi:\mathbb{R} &\longrightarrow \mathbb{R}\\ \rho &\mapsto \int_0^{2\pi}\int_{-\infty}^{+\infty}\rho\,\varphi(\xi,\rho\,\sin\theta,\rho\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta.\end{aligned}$$

We assume that Ψ is continuous, so that we can apply the fundamental theorem of calculus: there exists a function G satisfying

$$G(t)=\int_0^t\Psi(\rho)\,\mathrm{d}\rho.$$

Thus,

$$\begin{aligned}\lim_{\varepsilon\to 0}\langle f_\varepsilon,\varphi\rangle &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_1^{1+\varepsilon}\Psi(\rho)\,\mathrm{d}\rho\\ &= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\int_0^{1+\varepsilon}\Psi(\rho)\,\mathrm{d}\rho-\int_0^1\Psi(\rho)\,\mathrm{d}\rho\right)\\ &= \lim_{\varepsilon\to 0}\frac{G(1+\varepsilon)-G(1)}{\varepsilon}=G'(1)=\Psi(1)\\ &= \int_0^{2\pi}\int_{-\infty}^{+\infty}\frac{1}{\varepsilon}\,\varphi(\xi,\rho\,\sin\theta,\rho\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta.\end{aligned}$$

We now need to verify that Ψ is continuous. For all $\rho\in\mathbb{R}$ we can write

$$\Psi(\rho)=\rho\int_0^{2\pi}\int_{-\infty}^{+\infty}\varphi(\xi,\rho\,\sin\theta,\rho\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta,$$

we only need to prove the continuity of the integral. To this end, we can use the sequential continuity theorem, and show that if the sequence $(\rho_k)_{k\in\mathbb{N}}$ tends to $\hat\rho$, then

$$\int_0^{2\pi}\int_{-\infty}^{+\infty}\varphi(\xi,\rho_k\,\sin\theta,\rho_k\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta\xrightarrow{k\to\infty}\int_0^{2\pi}\int_{-\infty}^{+\infty}\varphi(\xi,\hat\rho\,\sin\theta,\hat\rho\,\cos\theta)\,\mathrm{d}\xi\,\mathrm{d}\theta.$$

Since φ has compact support, there exists $M > 0$ such that

$$\int_0^{2\pi} \int_{-\infty}^{+\infty} \varphi(\xi, \rho_k \sin\theta, \rho_k \cos\theta)\, \mathrm{d}\xi\, \mathrm{d}\theta = \int_0^{2\pi} \int_{-M}^{M} \varphi(\xi, \rho_k \sin\theta, \rho_k \cos\theta)\, \mathrm{d}\xi\, \mathrm{d}\theta.$$

We can now apply Lebesgue dominated convergence theorem (Theorem 3.47) to interchange the limit with the integral: this is possible because, by Weierstrass theorem, φ is bounded over bounded and closed sets. This allows us to conclude. ■

Solution 5.59 *(Exercise 5.59).* First, notice that the question is well posed because for every $n \in \mathbb{N}^*$ the function $x \mapsto \frac{n}{1+4\,n^2\,x^2}$ is summable over compact sets, since it is continuous. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then, changing variable we find

$$\begin{aligned}
\lim_{n\to\infty} \left\langle \frac{n}{1+4\,n^2\,x^2}, \varphi \right\rangle &= \lim_{n\to\infty} \int_{\mathbb{R}} \frac{n}{1+4\,n^2\,x^2}\, \varphi(x)\, \mathrm{d}x \\
&\overset{y=nx}{=} \lim_{n\to\infty} \int_{\mathbb{R}} \frac{n}{1+4\,y^2}\, \varphi\left(\frac{y}{n}\right) \frac{1}{n}\, \mathrm{d}y \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \frac{1}{1+4\,y^2}\, \varphi\left(\frac{y}{n}\right) \mathrm{d}y.
\end{aligned}$$

For every $y \in \mathbb{R}$, we have

$$\lim_{n\to\infty} \frac{1}{1+4\,y^2}\, \varphi\left(\frac{y}{n}\right) = \frac{1}{1+4\,y^2}\, \varphi(0).$$

We can apply Lebesgue dominated convergence theorem (Theorem 3.47), since

$$\left| \frac{1}{1+4\,y^2}\, \varphi\left(\frac{y}{n}\right) \right| \le \frac{1}{1+4\,y^2} \max_{\frac{y}{n} \in K} \left| \varphi\left(\frac{y}{n}\right) \right|,$$

where K is a compact containing the support of φ. The right hand-side of the latter is summable function that does not depend on n, and therefore we interchange the limit with the integral, and we obtain, changing variable again

$$\begin{aligned}
\lim_{n\to\infty} \left\langle \frac{n}{1+4\,n^2\,x^2}, \varphi \right\rangle &= \int_{\mathbb{R}} \frac{1}{1+4\,y^2}\, \varphi(0)\, \mathrm{d}y \\
&\overset{x=2y}{=} \frac{\varphi(0)}{2} \int_{\mathbb{R}} \frac{1}{1+x^2}\, \mathrm{d}x = \frac{\varphi(0)}{2}\, [\arctan x]_{-\infty}^{+\infty} = \frac{\varphi(0)}{2} \left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \frac{\pi}{2}\, \varphi(0).
\end{aligned}$$

We obtain

$$\lim_{n\to\infty} \frac{n}{1+4\,n^2\,x^2} = \frac{\pi}{2}\, \delta.$$

■

Solution 5.60 *(Exercise 5.60).* Suppose by contradiction that there exists $[f] \in L^1_{\mathrm{loc}}(\mathbb{R}^2;\mathbb{C})$ such that, for all $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$,

$$T(\varphi) = \int_{\mathbb{R}^2} f(x,y)\, \varphi(x,y)\, \mathrm{d}x\, \mathrm{d}y.$$

We denote by r the straight line described by the Cartesian equation $y = -3\,x$. Consider $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$ such that $\operatorname{supp}\varphi \subseteq \mathbb{R}^2 \setminus r$. Then, $T(\varphi) = 0$, because φ is zero over r. On the other hand, r has measure zero in $\mathbb{R}^2$, so that

$$\int_{\mathbb{R}^2 \setminus r} f(x,y)\, \varphi(x,y)\, \mathrm{d}x\, \mathrm{d}y = \int_{\mathbb{R}^2} f(x,y)\, \varphi(x,y)\, \mathrm{d}x\, \mathrm{d}y = 0.$$

The fundamental lemma then ensures that $f(x,y) = 0$ almost everywhere over $\mathbb{R}^2 \setminus r$, which implies that $f(x,y) = 0$ almost everywhere over $\mathbb{R}^2$. Let us now consider an arbitrary $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$. Then, $f(x,y)\,\varphi(x,y) = 0$ almost everywhere over $\mathbb{R}^2$, and this implies that $T(\varphi) = 0$ for all $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$, and therefore T is the zero distribution. This, however, cannot be true. In fact, we can always take $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$ such that $\varphi(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^2$, and $\varphi(0,0) > 0$. In this case,

$$T(\varphi) = \int_{\mathbb{R}^2} f(x,y)\,\varphi(x,y)\,\mathrm{d}x\,\mathrm{d}y > 0.$$

We reached a contradiction.

Now we address the second point. Let $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{C})$. We have

$$\begin{aligned}
\langle \partial_x T - 3\,\partial_y T, \varphi\rangle &= -\langle T, \partial_x\varphi\rangle + 3\,\langle T, \partial_y\varphi\rangle \\
&= \langle T, 3\,\partial_y\varphi - \partial_x\varphi\rangle \\
&= \int_{\mathbb{R}} (\partial_y(x,-3\,x)\varphi - \partial_x\varphi(x,-3\,x))\,\mathrm{d}x \\
&= \int_{\mathbb{R}} -\frac{\mathrm{d}}{\mathrm{d}x}\varphi(x,-3\,x)\,\mathrm{d}x = [\varphi(x,-3\,x)]_{-\infty}^{+\infty} = 0.
\end{aligned}$$

■

Solution 5.61 *(Exercise 5.61).* We can proceed in two ways. The first consists in observing that the derivative in the sense of distributions can be viewed as the limit of a difference quotient:

$$\begin{aligned}
\lim_{n\to\infty} n^\alpha \left(\delta\left(x+\frac{3}{n}\right) - \delta(x)\right) &= \lim_{n\to\infty} n^\alpha\,\frac{3}{n}\,\frac{\delta\left(x+\frac{3}{n}\right) - \delta(x)}{\frac{3}{n}} \\
&= \lim_{n\to\infty} 3\,n^{\alpha-1}\,\frac{\delta\left(x+\frac{3}{n}\right) - \delta(x)}{\frac{3}{n}}.
\end{aligned}$$

If $\alpha = 1$, we have

$$\lim_{n\to\infty} n^\alpha \left(\delta\left(x+\frac{3}{n}\right) - \delta(x)\right) = \lim_{n\to\infty} 3\,\frac{\delta\left(x+\frac{3}{n}\right) - \delta(x)}{\frac{3}{n}} = 3\,\delta'(x).$$

If $\alpha < 1$, we have 0, i.e., the limit is the zero distribution. Finally, if $\alpha > 1$, the limit does not exist in $\mathcal{D}'(\mathbb{R};\mathbb{R})$.

We now solve the same problem taking $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned}
\lim_{n\to\infty} \langle n^\alpha \left(\delta\left(x+\frac{3}{n}\right) - \delta(x)\right), \varphi(x)\rangle &= \lim_{n\to\infty} n^\alpha \left(\langle\delta\left(x+\frac{3}{n}\right), \varphi(x)\rangle - \langle\delta(x), \varphi(x)\rangle\right) \\
&= \lim_{n\to\infty} n^\alpha \left(\varphi\left(-\frac{3}{n}\right) - \varphi(0)\right).
\end{aligned}$$

We can use the Taylor expansion with initial point 0:

$$\varphi\left(-\frac{3}{n}\right) = \varphi(0) + \varphi'(0)\left(-\frac{3}{n}\right) + \mathrm{O}\left(\frac{1}{n^2}\right).$$

We obtain

$$\begin{aligned}\lim_{n\to\infty} n^\alpha \left(\varphi\left(-\frac{3}{n}\right)-\varphi(0)\right) &= \lim_{n\to\infty} n^\alpha \left(\varphi(0)+\varphi'(0)\left(-\frac{3}{n}\right)+\mathrm{O}\left(\frac{1}{n^2}\right)-\varphi(0)\right)\\ &= \lim_{n\to\infty}\left(n^\alpha\left(-\frac{3}{n}\right)\varphi'(0)+\mathrm{O}(n^{\alpha-2})\right)\\ &= \lim_{n\to\infty}\left(-3\,n^{\alpha-1}\,\varphi'(0)+\mathrm{O}(n^{\alpha-2})\right)\end{aligned}$$

If $\alpha = 1$, we obtain $-3\,\varphi'(0)$, so that the limit is $3\,\delta'$. If $\alpha < 1$, we obtain 0, and if $\alpha > 1$ the limit does not exist: taking for example $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$ such that $\varphi'(0) = 1$, we do not have convergence (to have convergence in $\mathcal{D}'$ we must have convergence for every φ in $\mathcal{D}$). ■

Solution 5.62 *(Exercise 5.62).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We cannot write

$$\lim_{\varepsilon\to 0^+}\langle \varepsilon\,|x|^{3\,\varepsilon-1},\varphi\rangle = \lim_{\varepsilon\to 0^+}\left(\int_0^{+\infty}\varepsilon\,x^{3\,\varepsilon-1}\,\varphi(x)\,\mathrm{d}x+\int_{-\infty}^{0}\varepsilon\,(-x)^{3\,\varepsilon-1}\,\varphi(x)\,\mathrm{d}x\right),$$

because as $\varepsilon\to 0$ the term $3\,\varepsilon-1$ tends to a negative number, and the base is non-negative, but could be zero. We have to write, instead,

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\langle \varepsilon\,|x|^{3\,\varepsilon-1},\varphi\rangle &= \lim_{\varepsilon\to 0^+}\lim_{\delta\to 0^+}\left(\int_\delta^{+\infty}\varepsilon\,x^{3\,\varepsilon-1}\,\varphi(x)\,\mathrm{d}x+\int_{-\infty}^{-\delta}\varepsilon\,(-x)^{3\,\varepsilon-1}\,\varphi(x)\,\mathrm{d}x\right)\\ &= \lim_{\varepsilon\to 0^+}\lim_{\delta\to 0^+}\left(\left[\frac{x^{3\,\varepsilon}}{3}\,\varphi(x)\right]_\delta^{+\infty}-\int_\delta^{+\infty}\frac{x^{3\,\varepsilon}}{3}\,\varphi'(x)\,\mathrm{d}x\right.\\ &\qquad\left.+\left[\frac{(-x)^{3\,\varepsilon}}{3}\,\varphi(x)\right]_{-\infty}^{-\delta}-\int_{-\infty}^{-\delta}\frac{(-x)^{3\,\varepsilon}}{3}\,\varphi'(x)\,\mathrm{d}x\right)\\ &= \lim_{\varepsilon\to 0^+}\lim_{\delta\to 0^+}\left(-\frac{\delta^{3\,\varepsilon}}{3}\,\varphi(\delta)+\frac{\delta^{3\,\varepsilon}}{3}\,\varphi(-\delta)-\frac{1}{3}\int_\delta^{+\infty}x^{3\,\varepsilon}\,\varphi'(x)\,\mathrm{d}x\right.\\ &\qquad\left.-\frac{1}{3}\int_{-\infty}^{-\delta}(-x)^{3\,\varepsilon}\,\varphi'(x)\,\mathrm{d}x\right)\\ &= \lim_{\varepsilon\to 0^+}-\frac{1}{3}\left(\int_0^{+\infty}x^{3\,\varepsilon}\,\varphi'(x)\,\mathrm{d}x+\int_{-\infty}^{0}(-x)^{3\,\varepsilon}\,\varphi'(x)\,\mathrm{d}x\right),\end{aligned}$$

where we used the continuity of φ. We want to bring the limit inside the integral, and to this end we use Lebesgue dominated convergence theorem (Theorem 3.47): we bound $|x^{3\,\varepsilon}\,\varphi'(x)|$ with a summable function that does not depend on ε. We can consider $|x^{3\,\varepsilon}\,\varphi'(x)| \le |\varphi'(x)|$ for $x \le 1$ and $|x^{3\,\varepsilon}\,\varphi'(x)| \le x^3\,|\varphi'(x)|$ for $x > 1$. Since

$$\lim_{\varepsilon\to 0^+} x^{3\,\varepsilon}\,\varphi'(x) = \varphi'(x) \qquad \text{and} \qquad \lim_{\varepsilon\to 0^+}(-x)^{3\,\varepsilon}\,\varphi'(x) = \varphi'(x),$$

we obtain

$$\begin{aligned}\lim_{\varepsilon\to 0^+}\langle \varepsilon\,|x|^{3\,\varepsilon-1},\varphi\rangle &= -\frac{1}{3}\left(\int_0^{+\infty}\varphi'(x)\,\mathrm{d}x+\int_{-\infty}^{0}\varphi'(x)\,\mathrm{d}x\right)\\ &= -\frac{1}{3}\,(\,-\varphi(0)-\varphi(0)) = \frac{2}{3}\,\varphi(0).\end{aligned}$$

It follows that

$$\lim_{\varepsilon\to 0^+}\varepsilon\,|x|^{3\,\varepsilon-1} = \frac{2}{3}\,\delta.$$

■

Solution 5.63 *(Exercise 5.63).* We write the equation as $x^2\,u'(x) + 3\,x\,u'(x) + 3\,u(x) = -1$. We start evaluating $x^2\,u'(x)$. Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. We have

$$\begin{aligned}
\left\langle x^2 \frac{\mathrm{d}}{\mathrm{d}x}\ \mathrm{p.\,v.}\ \frac{1}{x}, \varphi \right\rangle &= -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\,\varphi\right)\right\rangle \\
&= -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, 2\,x\,\varphi + x^2\,\varphi'\right\rangle \\
&= -\lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon} \left(2\,\varphi(x) + x\,\varphi'(x)\right)\mathrm{d}x \\
&= -2\lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon} \varphi(x)\,\mathrm{d}x \\
&\quad -\lim_{\varepsilon\to 0^+}\left(\left[x\,\varphi(x)\right]_{-\infty}^{-\varepsilon} - \int_{-\infty}^{-\varepsilon}\varphi(x)\,\mathrm{d}x + \left[x\,\varphi(x)\right]_{\varepsilon}^{+\infty} - \int_{\varepsilon}^{+\infty}\varphi(x)\,\mathrm{d}x\right) \\
&= -2\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, x\,\varphi\right\rangle - \lim_{\varepsilon\to 0^+}\left(\varphi(\varepsilon)(-\varepsilon) + \varphi(\varepsilon)\,\varepsilon - \int_{|x|\geq\varepsilon}\varphi(x)\,\mathrm{d}x\right) \\
&= -2\left\langle x\ \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle + \lim_{\varepsilon\to 0^+}\int_{|x|\geq\varepsilon}\varphi(x)\,\mathrm{d}x \\
&= -1\left\langle x\ \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle = -1.
\end{aligned}$$

where we used the continuity of φ, and also the fact that we have shown that $\left\langle x\ \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle = 1$ for all $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Now, we only need to verify that $x\,\frac{\mathrm{d}}{\mathrm{d}x}\left(\mathrm{p.\,v.}\ \frac{1}{x}\right) = -\,\mathrm{p.\,v.}\ \frac{1}{x}$. Again, let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then,

$$\begin{aligned}
\left\langle x \frac{\mathrm{d}}{\mathrm{d}x}\ \mathrm{p.\,v.}\ \frac{1}{x}, \varphi \right\rangle &= -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi + x\,\varphi'\right\rangle \\
&= -\lim_{\varepsilon\to 0^+}\left(\int_{|x|\geq\varepsilon}\frac{\varphi(x)}{x}\,\mathrm{d}x + \int_{|x|\geq\varepsilon}\varphi'(x)\,\mathrm{d}x\right) \\
&= -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle - \lim_{\varepsilon\to 0^+}\left(\left[\varphi(x)\right]_{-\infty}^{-\varepsilon} + \left[\varphi(x)\right]_{\varepsilon}^{+\infty}\right) \\
&= -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle - \lim_{\varepsilon\to 0^+}\left(\varphi(-\varepsilon) - \varphi(\varepsilon)\right) = -\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle,
\end{aligned}$$

where we used once again the continuity of φ. We conclude that

$$\left\langle x^2\,u'(x) + 3\,x\,u'(x) + 3\,u(x), \varphi\right\rangle = -1 - 3\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle + 3\left\langle \mathrm{p.\,v.}\ \frac{1}{x}, \varphi\right\rangle = -1.$$

■

Solution 5.64 *(Exercise 5.64).* Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then,

$$\lim_{k\to\infty}\left\langle k^\alpha\,\chi_{[0,\frac{1}{3k}]}, \varphi\right\rangle = \lim_{k\to\infty}\int_0^{\frac{1}{3k}} k^\alpha\,\varphi(x)\,\mathrm{d}x.$$

Since φ is continuous, we can define the primitive of φ using the fundamental theorem of calculus:

$$G(t) = \int_0^t \varphi(x)\,\mathrm{d}x.$$

Then,

$$\begin{aligned}\lim_{k\to\infty}\left\langle k^\alpha \chi_{[0,\frac{1}{3k}]},\varphi\right\rangle &= \lim_{k\to\infty} k^\alpha\left(G\Big(\frac{1}{3\,k}\Big)-G(0)\right)\\ &= \lim_{k\to\infty} k^\alpha\,\frac{1}{3\,k}\,\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}}\\ &= \lim_{k\to\infty}\frac{1}{3}\,k^{\alpha-1}\,\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}}.\end{aligned}$$

For $\alpha = 1$ we have

$$\lim_{k\to\infty}\left\langle k^\alpha \chi_{[0,\frac{1}{3k}]},\varphi\right\rangle = \lim_{k\to\infty}\frac{1}{3}\,\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}} = \frac{1}{3}\,G'(0) = \frac{1}{3}\,\varphi(0).$$

If $\alpha < 1$,

$$\lim_{k\to\infty}\left\langle k^\alpha \chi_{[0,\frac{1}{3k}]},\varphi\right\rangle = \lim_{k\to\infty}\frac{1}{3}\,k^{\alpha-1}\,\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}} = \frac{1}{3}\,G'(0) = \lim_{k\to\infty}\frac{1}{3}\cdot 0\cdot\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}} = 0.$$

If $\alpha > 1$, the limit does not exist, because $k^{\alpha-1}$ does not converge in $\mathcal{D}'(\mathbb{R};\mathbb{R})$.[2] We conclude that

$$\lim_{k\to\infty} k^\alpha \chi_{[0,\frac{1}{3k}]} = \begin{cases} 0 & \text{if } \alpha < 1\\ \frac{1}{3}\,\delta & \text{if } \alpha = 1.\end{cases}$$

An alternative solution can be obtained changing variable in the integral:

$$\begin{aligned}\lim_{k\to\infty}\int_0^{\frac{1}{3k}} k^\alpha\,\varphi(x)\,\mathrm{d}x &\overset{3kx=y}{=} \lim_{k\to\infty} k^\alpha\int_0^1 \varphi\Big(\frac{y}{3\,k}\Big)\,\frac{1}{3\,k}\,\mathrm{d}y\\ &= \lim_{k\to\infty}\frac{1}{3}\,k^{\alpha-1}\int_0^1\varphi\Big(\frac{y}{3\,k}\Big)\,\mathrm{d}y.\end{aligned}$$

If $\alpha = 1$, then we need to consider

$$\lim_{k\to\infty}\frac{1}{3}\int_0^1\varphi\Big(\frac{y}{3\,k}\Big)\,\mathrm{d}y.$$

We want to interchange the limit with the integral. To this end, we use Weierstrass extreme value theorem:

$$\left|\varphi\Big(\frac{y}{3\,k}\Big)\right| \le C,$$

[2] We could have convergence if we have

$$\lim_{k\to\infty}\frac{G\Big(\frac{1}{3\,k}\Big)-G(0)}{\frac{1}{3\,k}} = 0.$$

However, if we take $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$ satisfying $\varphi(0) = 1$, the limit does not exist. The convergence in $\mathcal{D}'(\mathbb{R};\mathbb{R})$ requires the existence of the limit for any $\varphi\in\mathcal{D}(\mathbb{R};\mathbb{R})$.

where C is a non-negative constant, which is summable over $[0, 1]$. We can conclude that

$$\lim_{k\to\infty} \frac{1}{3} \int_0^1 \varphi\left(\frac{y}{3\,k}\right) \mathrm{d}y = \frac{1}{3}\,\varphi(0).$$

The same argument shows that, for $\alpha < 1$,

$$\int_0^1 \varphi\left(\frac{y}{3\,k}\right) \mathrm{d}y \xrightarrow{k\to\infty} \varphi(0), \tag{14.4}$$

and so

$$\lim_{k\to\infty} \frac{1}{3}\,k^{\alpha-1} \int_0^1 \varphi\left(\frac{y}{3\,k}\right) \mathrm{d}y = \frac{1}{3}\cdot 0\cdot\varphi(0) = 0.$$

If $\alpha > 1$, it is still true that (14.4) holds, but when $\varphi(0) \neq 0$ the limit does not exist.

■

Solution 5.65 *(Exercise 5.65).* First, we observe that the question is well posed, because if $\alpha > -1$, for every ε the function $u_\varepsilon(x) = \varepsilon^{-5}\,H(x)\,H(\varepsilon - x)\,x^\alpha$ is locally summable. Moreover, we have

$$u_\varepsilon(x) = \begin{cases} \varepsilon^{-5}\,x^\alpha & \text{if } x \in (0,\varepsilon) \\ 0 & \text{elsewhere.} \end{cases}$$

Let $\varphi \in \mathcal{D}(\mathbb{R};\mathbb{R})$. Then,

$$\begin{aligned} \lim_{\varepsilon\to 0^+} \langle \varepsilon^{-5}\,H(x)\,H(\varepsilon - x)\,x^\alpha, \varphi\rangle &= \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon^5} \int_0^\varepsilon x^\alpha\,\varphi(x)\,\mathrm{d}x \\ &\overset{x=\varepsilon y}{=} \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon^5} \int_0^1 (y\,\varepsilon)^\alpha\,\varphi(y\,\varepsilon)\,\varepsilon\,\mathrm{d}y \\ &= \lim_{\varepsilon\to 0^+} \varepsilon^{\alpha-4} \int_0^1 y^\alpha\,\varphi(y\,\varepsilon)\,\mathrm{d}y. \end{aligned}$$

Pointwise, we have $y^\alpha\,\varphi(y\,\varepsilon) \xrightarrow{k\to\infty} y^\alpha\,\varphi(0)$. We can use Lebesgue dominated convergence theorem (Theorem 3.47) to interchange the limit with the integral:

$$|y^\alpha\,\varphi(y\,\varepsilon)| \le \sup|\varphi|\,y^\alpha$$

is summable over $(0, 1]$, and therefore

$$\lim_{\varepsilon\to 0^+} \varepsilon^{\alpha-4} \int_0^1 y^\alpha\,\varphi(y\,\varepsilon)\,\mathrm{d}y = \lim_{\varepsilon\to 0^+} \varepsilon^{\alpha-4}\,\varphi(0) \int_0^1 y^\alpha\,\mathrm{d}y = \lim_{\varepsilon\to 0^+} \varepsilon^{\alpha-4}\,\frac{\varphi(0)}{\alpha+1}.$$

Since

$$\lim_{\varepsilon\to 0^+} \varepsilon^{\alpha-4} = \begin{cases} 0 & \text{if } \alpha > 4 \\ 1 & \text{if } \alpha = 4 \\ +\infty & \text{if } \alpha < 4 \end{cases}$$

we obtain

$$\lim_{\varepsilon\to 0^+} \varepsilon^{-5}\,H(x)\,H(\varepsilon - x)\,x^\alpha = \begin{cases} 0 & \text{if } \alpha > 4 \\ \frac{1}{5}\,\delta & \text{if } \alpha = 4. \end{cases}$$

The limit does not exist for $\alpha < 4$. In fact, if φ is not zero at 0, the limit does not exist, and this is sufficient to conclude.

■

Solution 5.66 *(Exercise 5.66).* Let $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$. Then,

$$\begin{aligned}
\lim_{\varepsilon\to 0}\left\langle\left(\partial_x u_\varepsilon - \partial_y u_\varepsilon\right),\varphi\right\rangle &= \lim_{\varepsilon\to 0}\left(\langle\partial_x u_\varepsilon,\varphi\rangle - \left\langle\partial_y u_\varepsilon,\varphi\right\rangle\right)\\
&= \lim_{\varepsilon\to 0}\left(-\left\langle u_\varepsilon,\frac{\partial\varphi}{\partial x}\right\rangle + \left\langle u_\varepsilon,\frac{\partial\varphi}{\partial y}\right\rangle\right)\\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{R}^2} u_\varepsilon(x,y)\left(\frac{\partial\varphi}{\partial x}(x,y) - \frac{\partial\varphi}{\partial y}(x,y)\right)\mathrm{d}x\,\mathrm{d}y\\
&= \lim_{\varepsilon\to 0}\int_{\mathbb{R}\times[0,4\varepsilon]}\frac{3}{\varepsilon}\left(\frac{\partial\varphi}{\partial x}(x,y) - \frac{\partial\varphi}{\partial y}(x,y)\right)\mathrm{d}x\,\mathrm{d}y\\
&= \lim_{\varepsilon\to 0}\frac{3}{\varepsilon}\int_{\mathbb{R}}\int_0^{4\varepsilon}\left(\frac{\partial\varphi}{\partial x}(x,y) - \frac{\partial\varphi}{\partial y}(x,y)\right)\mathrm{d}y\,\mathrm{d}x\\
&= \lim_{\varepsilon\to 0}\frac{3}{\varepsilon}\int_0^{4\varepsilon}\left(\int_{\mathbb{R}}\left(\frac{\partial\varphi}{\partial x}(x,y) - \frac{\partial\varphi}{\partial y}(x,y)\right)\mathrm{d}x\right)\mathrm{d}y,
\end{aligned}$$

where we have applied Fubini's theorem (Theorem 3.8). We define

$$\Psi(y) = \int_{\mathbb{R}}\left(\frac{\partial\varphi}{\partial x}(x,y) - \frac{\partial\varphi}{\partial y}(x,y)\right)\mathrm{d}x.$$

From the fundamental theorem of calculus there exists a function G such that

$$G(t) = \int_0^t \Psi(y)\,\mathrm{d}y,$$

and therefore

$$\begin{aligned}
\lim_{\varepsilon\to 0}\left\langle\left(\partial_x u_\varepsilon - \partial_y u_\varepsilon\right),\varphi\right\rangle &= \lim_{\varepsilon\to 0}\frac{3}{\varepsilon}\int_0^{4\varepsilon}\Psi(y)\,\mathrm{d}y\\
&= \lim_{\varepsilon\to 0}\frac{3}{\varepsilon}\left(G(4\varepsilon) - G(0)\right)\\
&= \lim_{\varepsilon\to 0} 12\,\frac{G(4\varepsilon)-G(0)}{4\varepsilon} = 12\,G'(0) = 12\,\Psi(0)\\
&= 12\int_{\mathbb{R}}\left(\frac{\partial\varphi}{\partial x}(x,0) - \frac{\partial\varphi}{\partial y}(x,0)\right)\mathrm{d}x.
\end{aligned}$$

We have

$$\int_{\mathbb{R}}\frac{\partial\varphi}{\partial x}(x,0)\,\mathrm{d}x = \left[\varphi(x,0)\right]_{-\infty}^{+\infty} = 0$$

and therefore

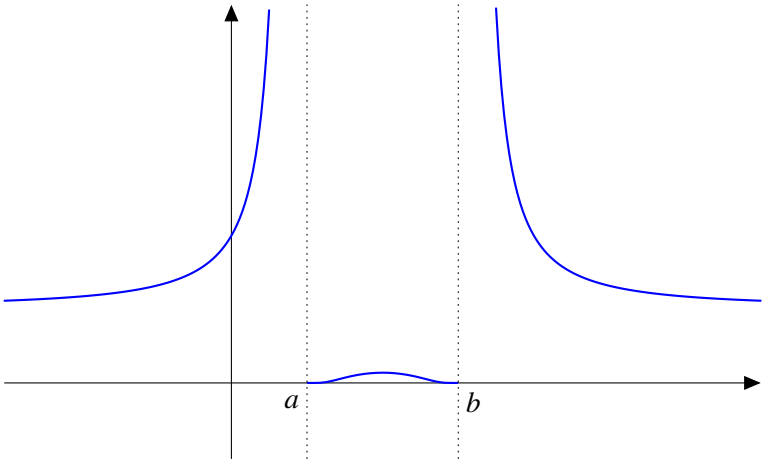
$$\begin{aligned}
\lim_{\varepsilon\to 0}\left\langle\left(\partial_x u_\varepsilon - \partial_y u_\varepsilon\right),\varphi\right\rangle &= 12\int_{\mathbb{R}}\frac{\partial\varphi}{\partial y}(x,0)\,\mathrm{d}x\\
&= 12\left\langle 1\otimes\delta,\frac{\partial\varphi}{\partial y}\right\rangle = -12\left\langle\partial_y(1\otimes\delta),\varphi\right\rangle.
\end{aligned}$$

■

Solution 5.67 *(Exercise 5.67).* Let $\varphi \in \mathcal{D}(\mathbb{R}^2;\mathbb{R})$. Then,

$$\begin{aligned}
\left\langle x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}, \varphi \right\rangle &= \left\langle \frac{\partial u}{\partial x}, x\,\varphi \right\rangle + \left\langle \frac{\partial u}{\partial y}, y\,\varphi \right\rangle \\
&= -\left\langle u, \frac{\partial}{\partial x}(x\,\varphi) \right\rangle - \left\langle u, \frac{\partial}{\partial y}(y\,\varphi) \right\rangle \\
&= -\left\langle u, \frac{\partial}{\partial x}(x\,\varphi) + \frac{\partial}{\partial y}(y\,\varphi) \right\rangle \\
&= -\int_{\mathbb{R}^2} u(x,y)\left(\frac{\partial}{\partial x}(x\,\varphi) + \frac{\partial}{\partial y}(y\,\varphi)\right) \mathrm{d}x\,\mathrm{d}y \\
&= -\int_{x^2+y^2\le 9} \left(\frac{\partial}{\partial x}(x\,\varphi) + \frac{\partial}{\partial y}(y\,\varphi)\right) \mathrm{d}x\,\mathrm{d}y \\
&= -\int_{C_3(0)} \left(x\,\varphi(x,y)\,\frac{x}{3} + y\,\varphi(x,y)\,\frac{y}{3}\right) \mathrm{d}s \\
&= -\int_0^{2\pi} 3\,\varphi(3\,\cos\theta, 3\,\sin\theta)\,3\,\mathrm{d}\theta \\
&= -3\int_0^{2\pi} 3\,\varphi(3\,\cos\theta, 3\,\sin\theta)\,\mathrm{d}\theta = -3\int_0^{2\pi} \varphi(s)\,\mathrm{d}s,
\end{aligned}$$

where we used Green's theorem and the change of variables $x = 3\,\cos\theta$ and $y = 3\,\sin\theta$. Thus, $c = -3$.



■

Solution 5.68 *(Exercise 5.68).* Let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{R})$. Let $M > 0$ such that $\operatorname{supp}\varphi \subseteq [-M, M]$. Then, we have also $\operatorname{supp}(\eta_k\,\varphi) \subseteq [-M, M]$.

Moreover, for every $\boldsymbol{\alpha} \in \mathbb{N}^n$ we have

$$\begin{aligned}
\sup|\partial^{\alpha}(\eta_k\,\varphi) - \partial^{\alpha}\varphi| &= \sup|\varphi\,\partial^{\alpha}\eta_k + \eta_k\,\partial^{\alpha}\varphi - \partial^{\alpha}\varphi| \\
&\le \sup|\varphi\,\partial^{\alpha}\eta_k| + \sup|(\eta_k - 1)\,\partial^{\alpha}\varphi| \\
&= \sup(|\varphi|\cdot|\,\partial^{\alpha}\eta_k|) + \sup(|\eta_k - 1)|\cdot|\partial^{\alpha}\varphi|) \xrightarrow{k\to\infty} 0
\end{aligned}$$

because φ and $\partial^{\alpha}\varphi$ are bounded, and $|\partial^{\alpha}\eta_k| \xrightarrow{k\to\infty} 0$ and $\eta_k - 1 \xrightarrow{k\to\infty} 0$. Therefore $\partial^{\alpha}(\eta_k\,\varphi) \xrightarrow{k\to\infty} \partial^{\alpha}\varphi$ uniformly on $\mathbb{R}$, which implies $\eta_k\,\varphi \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{R})$. ■

Solution 5.69 *(Exercise 5.69).* If $T_1 * = S$ and $T_2 * S$ exist, then $(\lambda T_1) * S$ and $(\mu T_2) * S$ exist, and this immediately implies that $(\lambda T_1 + \mu T_2) * S$ exists. Let $\varphi \in \mathcal{D}(\mathbb{R}^n; \mathbb{R})$, and let $(\eta_k)_{k\in\mathbb{N}}$ be converging to unity in $\mathbb{R}^{2n}$. Then

$$\begin{aligned}
\langle (\lambda T_1 + \mu T_2) * S, \varphi\rangle &= \lim_{k\to\infty} \Big\langle ((\lambda T_1 + \mu T_2)\otimes S)(\boldsymbol{x},\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\Big\rangle \\
&= \lim_{k\to\infty} \Big\langle \lambda T_1(\boldsymbol{x}) + \mu T_2(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle \\
&= \lim_{k\to\infty} \Big(\Big\langle \lambda T_1(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle + \Big\langle \mu T_2(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle\Big) \\
&= \lim_{k\to\infty} \Big(\lambda\Big\langle T_1(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle + \mu\Big\langle T_2(\boldsymbol{x}), \langle S(\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big\rangle\Big) \\
&= \lim_{k\to\infty} \Big(\lambda\langle (T_1\otimes S)(\boldsymbol{x},\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle + \mu\langle (T_2\otimes S)(\boldsymbol{x},\boldsymbol{y}), \eta_k(\boldsymbol{x},\boldsymbol{y})\,\varphi(\boldsymbol{x}+\boldsymbol{y})\rangle\Big) \\
&= \lambda\langle T_1 * S, \varphi\rangle + \mu\langle T_2 * S, \varphi\rangle = \langle \lambda(T_1 * S) + \mu(T_2 * S), \varphi\rangle.
\end{aligned}$$

■

Solution 5.70 *(Exercise 5.70).* For every compact set $\boldsymbol{K} \subseteq \mathbb{R}^n$, there exists $p_{\boldsymbol{K}} \in \mathbb{N}$ such that for all $k \geq p_{\boldsymbol{K}}$ we have $\eta_k(\boldsymbol{y},\boldsymbol{x}) = \tilde{\eta}_k(\boldsymbol{x},\boldsymbol{y}) = 1$ for all $(\boldsymbol{y},\boldsymbol{x}) \in \boldsymbol{K}$. Moreover, if $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{2n}$, then also $(\boldsymbol{y},\boldsymbol{x}) \in \mathbb{R}^{2n}$, so that

$$|\partial^\alpha \eta_k(\boldsymbol{x},\boldsymbol{y})| \leq c_\alpha \quad \Rightarrow \quad |\partial^\alpha \eta_k(\boldsymbol{y},\boldsymbol{x})| \leq c_\alpha \quad \Rightarrow \quad |\partial^\alpha \tilde{\eta}_k(\boldsymbol{x},\boldsymbol{y})| \leq c_\alpha,$$

so that $(\eta_k)_{k\in\mathbb{N}}$ converges to unity.

■

15 Solutions to the Exercises in Chapter 6

Solution 6.1 *(Exercise 6.1).* Recall that a function $f : A \longrightarrow \mathbb{R}$, with $A \subseteq \mathbb{R}$, is even if

$$\forall\, x \in A, \quad f(-x) = f(x),$$

and it is odd if

$$\forall\, x \in A, \quad f(-x) = -f(x)$$

An even function is symmetric about the vertical axis, and an odd function is symmetric about the origin. An example of even function is the cosine of x, and an example of odd function is the sine of x.

If $f, g : A \longrightarrow \mathbb{R}$ are both even, then for all $x \in A$ we have

$$(f\,g)(-x) = f(-x)\,g(-x) = f(x)\,g(x) = (f\,g)(x)$$

and if they are both odd we have

$$(f\,g)(-x) = \big(-f(-x)\big)\big(-g(-x)\big) = f(x)\,g(x) = (f\,g)(x).$$

Now suppose that f is even and g is odd. Then,

$$(f\,g)(-x) = \big(f(-x)\big)\big(-g(-x)\big) = -f(x)\,g(x) = -(f\,g)(x).$$

Thus, the product is odd.

Now we prove the result about the integrals of even/odd functions. If f is even,

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x &= \int_{-\infty}^{0} f(x)\,\mathrm{d}x + \int_{0}^{\infty} f(x)\,\mathrm{d}x \\
&\overset{y=-x}{=} -\int_{\infty}^{0} f(-y)\,\mathrm{d}y + \int_{0}^{\infty} f(x)\,\mathrm{d}x = \int_{0}^{\infty} f(y)\,\mathrm{d}y + \int_{0}^{\infty} f(x)\,\mathrm{d}x \\
&= 2\int_{0}^{\infty} f(x)\,\mathrm{d}x.
\end{aligned}$$

If f is odd, we have instead

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x &= \int_{-\infty}^{0} f(x)\,\mathrm{d}x + \int_{0}^{\infty} f(x)\,\mathrm{d}x \\
&\overset{y=-x}{=} -\int_{\infty}^{0} f(-y)\,\mathrm{d}y + \int_{0}^{\infty} f(x)\,\mathrm{d}x = \int_{0}^{\infty} -f(y)\,\mathrm{d}y + \int_{0}^{\infty} f(x)\,\mathrm{d}x \\
&= -\int_{0}^{\infty} f(y)\,\mathrm{d}y + \int_{0}^{\infty} f(x)\,\mathrm{d}x = 0.
\end{aligned}$$

Let f be even, so that $\cos(x\,\xi)\,f(x)$ is even and $\sin(x\,\xi)\,f(x)$ is odd. Then,

$$\begin{aligned}
\hat{f}(\xi) &= \int_{-\infty}^{+\infty} \cos(x\,\xi)\,f(x)\,\mathrm{d}x - i\int_{-\infty}^{+\infty} \sin(x\,\xi)\,f(x)\,\mathrm{d}x \\
&= 2\int_{0}^{+\infty} \cos(x\,\xi)\,f(x)\,\mathrm{d}x - i\cdot 0 = 2\int_{0}^{+\infty} \cos(x\,\xi)\,f(x)\,\mathrm{d}x.
\end{aligned}$$

Likewise, if f is odd, then $\cos(x\xi)\,f(x)$ is odd and $\sin(x\xi)\,f(x)$ is even. Then,

$$\begin{aligned}\hat{f}(\xi) &= \int_{-\infty}^{+\infty} \cos(x\xi)\,f(x)\,\mathrm{d}x - i\int_{-\infty}^{+\infty} \sin(x\xi)\,f(x)\,\mathrm{d}x \\ &= 0 - i\cdot 2\int_{0}^{+\infty} \sin(x\xi)\,f(x)\,\mathrm{d}x.\end{aligned}$$

■

Solution 6.2 *(Exercise 6.2).* We have to determine

$$\int_{\mathbb{R}_+} e^{-x(1+i\xi)}\,\mathrm{d}x, \qquad \xi\in\mathbb{R}.$$

We obtain[1]

$$\int_{\mathbb{R}_+} e^{-x(1+i\xi)}\,\mathrm{d}x = \lim_{c\to+\infty}\int_0^c e^{-x(1+i\xi)}\,\mathrm{d}x = \lim_{c\to+\infty}\frac{1-e^{-c(1+i\xi)}}{1+i\xi}.$$

Since $|e^{-c(1+i\xi)}| = e^{-c}$ goes to zero as $c\to+\infty$, we obtain

$$\hat{f}(\xi) = \frac{1}{1+i\xi} \qquad \forall\,\xi\in\mathbb{R}.$$

If $\xi\geq 1$, we have

$$|\hat{f}(\xi)| = \frac{1}{\sqrt{1+\xi^2}} = \frac{1}{\xi}\,\frac{1}{\sqrt{1+\frac{1}{\xi^2}}} \geq \frac{1}{\xi\,\sqrt{2}}.$$

Since $\displaystyle\int_{[1,+\infty)}\frac{1}{\xi\,\sqrt{2}}\,\mathrm{d}\xi = +\infty$, $\hat{f}$ is not summable. ■

Solution 6.3 *(Exercise 6.3).* Let $A=\lambda I$, where I denotes the $n\times n$ identity matrix. Then, A is invertible and $A^{-1}=\frac{1}{\lambda}I$. We have $|\det A| = |\lambda^n| = |\lambda|^n$, and $A^\top = A$, so that for every $\boldsymbol{\xi}\in\mathbb{R}^n$ we find

$$\begin{aligned}\mathcal{F}\{f(\lambda\boldsymbol{x})\}(\boldsymbol{\xi}) &= \frac{1}{|\det A|}\,\mathcal{F}\{f\}(A^{-\top}\boldsymbol{\xi}) \\ &= \frac{1}{|\lambda|^n}\,\mathcal{F}\{f\}(A^{-\top}\boldsymbol{\xi}) = \frac{1}{|\lambda|^n}\,\mathcal{F}\{f\}\left(\frac{\boldsymbol{\xi}}{\lambda}\right).\end{aligned}$$

■

[1] If $-\infty<a<b\leq+\infty$, let $f:[a,b)\longrightarrow\mathbb{R}$ be Riemann integrable over $[a,c)$ for every $c\in(a,b)$, and let

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

Then:

(i) the limit $\lim_{c\to b}\int_a^c |f(x)|\,\mathrm{d}x$ exists in $\mathbb{R}$;

(ii) f is integrable in $[a,b)$;

(iii) $\displaystyle\int_{[a,b)} f(x)\,\mathrm{d}x = \lim_{c\to b}\int_a^c |f(x)|\,\mathrm{d}x.$

Solution 6.4 *(Exercise 6.4).* If f is real, and, in addition

- if f is an even function, we have

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x = 2\int_{0}^{+\infty} \cos(x\,\xi)\, f(x)\,\mathrm{d}x.$$

  It follows that $\hat{f}(\xi)$ is also even, since $\hat{f}(\xi) = \hat{f}(-\xi)$.

- if f is an odd function, we have

$$\hat{f}(\xi) = -i\int_{-\infty}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x = -2\,i\int_{0}^{+\infty} \sin(x\,\xi)\, f(x)\,\mathrm{d}x.$$

  Thus, $\hat{f}(\xi)$ is also odd.

We can also consider the following argument. If f is odd, then $f(-\boldsymbol{x}) = -f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^n$. Defining $A \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^n)$ as $A\,\boldsymbol{x} = -\boldsymbol{x}$, and the identity $f(-\boldsymbol{x}) = -f(\boldsymbol{x})$ can be rewritten as $f \circ A = -f$. Taking the Fourier transform of both sides gives $\widehat{f \circ A} = -\hat{f}$. We can use property **(P3)** in Theorem 6.2, we find

$$\frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) = -\hat{f}(\boldsymbol{\xi}), \quad \forall\,\boldsymbol{\xi} \in \mathbb{R}^n.$$

On the other hand, $A\,\boldsymbol{x} = -\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$, which implies that A is a diagonal matrix with -1 on the diagonal. This means that A is orthogonal and symmetric, and therefore $A = A^{-1} = A^{\top}$ and $|\det A| = 1$. It follows that

$$\frac{1}{|\det A|}\,\hat{f}(A^{-\top}\boldsymbol{\xi}) = (\hat{f} \circ A)(\boldsymbol{\xi}) \quad \Rightarrow \quad \hat{f} \circ A = -\hat{f},$$

which implies that $\hat{f}$ is odd. ■

Solution 6.5 *(Exercise 6.5).* We use Euler's law:

$$f(x) = \frac{\cos x}{1+x^2} = \frac{1}{2}\,\frac{e^{i\,x} + e^{-i\,x}}{1+x^2} = \frac{1}{2}\,\frac{1}{1+x^2}\,e^{i\,x} + \frac{1}{2}\,\frac{1}{1+x^2}\,e^{-i\,x}.$$

We now use **(P2)** and the linearity of the Fourier transform to find

$$\hat{f}(\xi) = \frac{\pi}{2}\,e^{-|\xi-1|} + \frac{\pi}{2}\,e^{-|\xi+1|}.$$

Consider the second. The application of **(P1)** gives immediately

$$e^{-i\,\xi\,(-3)}\,\pi\,e^{-|\xi|} = \pi\,e^{3\,i\,\xi - |\xi|}.$$

Consider the last. We can write

$$\frac{1}{4+x^2} = \frac{1}{4}\,\frac{1}{1+\left(\frac{x}{2}\right)^2}$$

so that we can exploit the Fourier transform of a homothety:

$$\frac{1}{4}\cdot 2\,\pi\,e^{-|2\,\xi|} = \frac{1}{2}\,\pi\,e^{-2\,|\xi|}.$$

■

Solution 6.6 *(Exercise 6.6).* Let $f : \mathbb{R}^n \longrightarrow \mathbb{C}$. We can write $f = f_R + i\, f_I$, where f_R and f_I are real, i.e., $f_R, f_I : \mathbb{R}^n \longrightarrow \mathbb{R}$. We write

$$\begin{aligned}\lim_{\|\boldsymbol{\xi}\|\to\infty} \hat{f}(\boldsymbol{\xi}) &= \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} (\cos(\boldsymbol{x}\cdot\boldsymbol{\xi}) - i\,\sin(\boldsymbol{x}\cdot\boldsymbol{\xi}))\,(f_R(\boldsymbol{x}) + i\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \\ &= \lim_{\|\boldsymbol{\xi}\|\to\infty} \left(\int_{\mathbb{R}^n} (\cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x}) + \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} + i \int_{\mathbb{R}^n} (-\sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x}) + \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} \right) = 0,\end{aligned}$$

which implies

$$\begin{aligned}\lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} (\cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x}) + \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = 0 \\ \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} (-\sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x}) + \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = 0.\end{aligned}$$

We notice that we can exploit the linearity since

$$\begin{aligned}|\cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})| &\leq 1\cdot|f(\boldsymbol{x})| \\ |\sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})| &\leq 1\cdot|f(\boldsymbol{x})| \\ |\sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x})| &\leq 1\cdot|f(\boldsymbol{x})| \\ |\cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})| &\leq 1\cdot|f(\boldsymbol{x})|.\end{aligned}$$

We obtain

$$\begin{aligned}\lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \\ \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \\ \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0 \\ \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 0.\end{aligned}$$

In particular, the first and the last say that

$$\begin{aligned}0 &= \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_R(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} + i \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f_I(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\ &= \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\,(f_R(\boldsymbol{x}) + i\, f_I(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x} = \lim_{\|\boldsymbol{\xi}\|\to\infty} \int_{\mathbb{R}^n} \cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\, f(\boldsymbol{x}),\end{aligned}$$

and the second follows in the same way. ■

Solution 6.7 *(Exercise 6.7).* Consider the first. Clearly, it is a function in $L^1(\mathbb{R};\mathbb{R})$ since

$$\int_{\mathbb{R}} |e^{-x}\, H(x)|\,\mathrm{d}x = \int_0^{+\infty} e^{-x}\,\mathrm{d}x = [-e^{-x}]_0^{+\infty} = 1.$$

The Fourier transform is

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\,x\,\xi}\, e^{-x}\, H(x)\,\mathrm{d}x = \int_0^{+\infty} e^{-x(i\,\xi+1)}\,\mathrm{d}x = \left[\frac{e^{-x(i\,\xi+1)}}{-(i\,\xi+1)}\right]_0^{+\infty} = \frac{1}{1+i\,\xi}.$$

The second function is still in $L^1(\mathbb{R};\mathbb{R})$ since

$$\int_{\mathbb{R}} |e^x\, H(-x)|\,\mathrm{d}x = \int_{-\infty}^{0} e^x\,\mathrm{d}x = [e^x]_{-\infty}^{0} = 1 - 0 = 1.$$

The Fourier transform is

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\,x\,\xi}\, e^x\, H(-x)\,\mathrm{d}x = \int_{-\infty}^{0} e^{x\,(1-i\,\xi)}\,\mathrm{d}x = \left[\frac{e^{x\,(1-i\,\xi)}}{1-i\,\xi}\right]_{-\infty}^{0} = \frac{1}{1-i\,\xi}.$$

Finally, consider the third, which is the function

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto \begin{cases} e^{-x} & \text{if } x \geq 0 \\ e^{x} & \text{if } x < 0 \end{cases} \end{aligned}$$

Thus,

$$\int_{\mathbb{R}} |e^{-|x|}|\,\mathrm{d}x = \int_{-\infty}^{0} e^x\,\mathrm{d}x + \int_{0}^{+\infty} e^{-x}\,\mathrm{d}x = [e^x]_{-\infty}^{0} + [-e^{-x}]_{0}^{+\infty} = 1 - 0 + (0 - 1) = 2.$$

The Fourier transform is

$$\begin{aligned} \hat{f}(\xi) &= \int_{\mathbb{R}} e^{-i\,x\,\xi}\, e^{-|x|}\,\mathrm{d}x = \int_{0}^{+\infty} e^{-i\,x\,\xi} e^{-x}\,\mathrm{d}x + \int_{-\infty}^{0} e^{-i\,x\,\xi}\, e^{x}\,\mathrm{d}x = \left[\frac{e^{-x\,(i\,\xi+1)}}{-(i\,\xi+1)}\right]_{0}^{+\infty} + \left[\frac{e^{x\,(1-i\,\xi)}}{1-i\,\xi}\right]_{-\infty}^{0} \\ &= \frac{1}{1+i\,\xi} + \frac{1}{1-i\,\xi} = \frac{1-i\,\xi+1+i\,\xi}{1+\xi^2} = \frac{2}{1+\xi^2}. \end{aligned}$$

■

Solution 6.8 *(Exercise 6.8).* First, we show that $f_k \xrightarrow{k\to\infty} f$ in $L^1(\mathbb{R};\mathbb{C})$. Indeed,

$$\int_{\mathbb{R}} |f_k - f| = \int_{\mathbb{R}} |f\,\chi_{[-k,k]} - f| = \int_{\mathbb{R}} |f\,(\chi_{[-k,k]} - 1) = \int_{\mathbb{R}} |f|\,|\chi_{[-k,k]} - 1| = \int_{-\infty}^{-k} |f| + \int_{k}^{+\infty} |f| \xrightarrow{k\to\infty} 0.$$

We recall that, from Theorem 1 in the Lecture Notes on the Fourier transform, for all $\boldsymbol{\xi} \in \mathbb{R}^n$ we have $|\hat{f}(\boldsymbol{\xi})| \leq \|f\|_1$. In our case, this means that for all $\boldsymbol{\xi} \in \mathbb{R}^n$ we have $|\hat{f}_k(\boldsymbol{\xi}) - f(\boldsymbol{\xi})| \leq \|f_k - f\|_1$. This implies also that $\sup_{\mathbb{R}^n} |\hat{f}_k(\boldsymbol{\xi}) - f(\boldsymbol{\xi})| \leq \|f_k - f\|_1$, which in turn implies $\|\hat{f}_k - \hat{f}\|_\infty \leq \|f_k - f\|_1$ (we just found what we already knew: the convergence with respect to $L^\infty(\mathbb{R};\mathbb{C})$ is the uniform convergence). Thus, since we showed that $f_k \xrightarrow{k\to\infty} f$ in $L^1(\mathbb{R};\mathbb{C})$, then $\hat{f}_k \xrightarrow{k\to\infty} f$ uniformly. ■

Solution 6.9 *(Exercise 6.9).* We use the change of variable $\boldsymbol{y} = A\,\boldsymbol{x}$:

$$\int_{\mathbb{R}^n} |f(A\,\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}^n} |f(\boldsymbol{y})|\,\frac{1}{|\det A|}\,\mathrm{d}\boldsymbol{y} = \frac{1}{|\det A|}\int_{\mathbb{R}^n} |f(\boldsymbol{y})|\,\mathrm{d}\boldsymbol{y} < +\infty.$$

■

Solution 6.10 *(Exercise 6.10).* Let $A\,\boldsymbol{x} = -\boldsymbol{x}$, so that $A = -I$. Since f is odd

$$f \circ A = -f \quad \Rightarrow \quad \widehat{f \circ A} = -\hat{f} \quad \Rightarrow \quad \frac{1}{|\det A|}\,\hat{f} \circ A^{-\top} = -\hat{f} \quad \Rightarrow \quad \hat{f} \circ A = -\hat{f}.$$

This shows that $\hat{f}$ is odd. Then,

$$\hat{f}(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \underbrace{\int_{\mathbb{R}^n} \underbrace{\cos(\boldsymbol{x}\cdot\boldsymbol{\xi})\,f(\boldsymbol{x})}_{\text{odd}}\,\mathrm{d}\boldsymbol{x}}_{=0} - i \int_{\mathbb{R}^n} \underbrace{\sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\,f(\boldsymbol{x})}_{\text{even}}\,\mathrm{d}\boldsymbol{x} = -i \int_{\mathbb{R}^n} \sin(\boldsymbol{x}\cdot\boldsymbol{\xi})\,f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}.$$

Since f is real, the integrand in the right hand-side is real, and therefore $\hat{f}(\boldsymbol{\xi})$ is purely imaginary. ■

Solution 6.11 *(Exercise 6.11).* We use the spherical coordinates

$$\begin{cases} x_1 = \rho\,\sin\theta\,\cos\varphi \\ x_2 = \rho\,\sin\theta\,\sin\varphi \\ x_3 = \rho\,\cos\theta, \end{cases} \qquad \text{with } \rho \in (0,1],\quad \theta \in (0,\pi),\quad \varphi \in (0,2\,\pi).$$

The determinant of the Jacobian is

$$\begin{vmatrix} \sin\theta\,\cos\varphi & \rho\,\cos\theta\,\cos\varphi & -\rho\,\sin\theta\,\sin\varphi \\ \sin\theta\,\sin\varphi & \rho\,\cos\theta\,\sin\varphi & \rho\,\sin\theta\,\cos\varphi \\ \cos\theta & -\rho\,\sin\theta & 0 \end{vmatrix} = \rho^2\,\sin\theta.$$

Using Fubini's theorem (Theorem 3.8), we find

$$\begin{aligned} \int_{\mathbb{R}^3} \|\boldsymbol{x}\|^\alpha\,\mathrm{d}\boldsymbol{x} &= \int_{(0,1)\times(0,\pi)\times(0,2\,\pi)} \rho^\alpha\,\rho^2\,\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= 2\,\pi \int_{(0,1)\times(0,\pi)} \rho^{\alpha+2}\,\sin\theta\,\mathrm{d}\rho\,\mathrm{d}\theta = 2\,\pi \int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho \cdot \int_0^\pi \sin\theta\,\mathrm{d}\theta = 4\,\pi \int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho. \end{aligned}$$

The last integral is improper if $\alpha \le -3$, and it does not converge if $\alpha \le -3$. If $\alpha > -3$, we have

$$\int_0^1 \rho^{\alpha+2}\,\mathrm{d}\rho = \left[\frac{\rho^{\alpha+3}}{\alpha+3}\right]_0^1 = \frac{1}{\alpha+3}.$$

■

Solution 6.12 *(Exercise 6.12).* The function f is odd and real, so the Fourier transform will be odd and purely imaginary. In fact

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\,x\,\xi}\,e^{-|x|}\,\operatorname{sgn} x\,\mathrm{d}x = -\int_{-\infty}^0 e^{-i\,x\,\xi}\,e^{x}\,\mathrm{d}x + \int_0^{+\infty} e^{-i\,x\,\xi}\,e^{-x}\,\mathrm{d}x = -\frac{1}{1-i\,\xi} + \frac{1}{1+i\,\xi} = \frac{-2\,i\,\xi}{1+\xi^2}.$$

■

Solution 6.13 *(Exercise 6.13).* Let $\xi \in \mathbb{R}$. From (1) in the Lecture Notes we have

$$\int_{-\infty}^{+\infty} e^{-(i\,\eta\,x+x^2)}\,\mathrm{d}x = \sqrt{\pi}\,e^{-\frac{\eta^2}{4}}.$$

We find immediately

$$\begin{aligned} \mathcal{F}\left\{e^{-x^2}\right\}(\xi) &= \int_{-\infty}^{+\infty} e^{-i\,\xi\,x}\,e^{-x^2}\,\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} e^{-(i\,\xi\,x+x^2)}\,\mathrm{d}x = \sqrt{\pi}\,e^{-\frac{\xi^2}{4}}. \end{aligned}$$

■

Solution 6.14 *(Exercise 6.14).* The function f is odd and real, so the Fourier transform will be odd and purely imaginary. In fact

$$\begin{aligned}
\hat{f}(\xi) &= \int_{\mathbb{R}} x\, e^{-i\,x\,\xi}\, e^{-|x|}\, \operatorname{sgn} x\, \mathrm{d}x \\
&= -\int_{-\infty}^{0} x\, e^{x}\, e^{-i\,x\,\xi}\, \mathrm{d}x + \int_{0}^{+\infty} x\, e^{-x}\, e^{-i\,x\,\xi}\, \mathrm{d}x \\
&= -\int_{-\infty}^{0} x\, e^{x\,(1-i\,\xi)}\, \mathrm{d}x + \int_{0}^{+\infty} x\, e^{-x\,(1+i\,\xi)}\, \mathrm{d}x \\
&= -\underbrace{\left[x\,\frac{e^{x\,(1-i\,\xi)}}{1-i\,\xi}\right]_{-\infty}^{0}}_{=0} + \int_{-\infty}^{0} \frac{e^{x\,(1-i\,\xi)}}{1-i\,\xi}\, \mathrm{d}x + \underbrace{\left[x\,\frac{e^{-x\,(1+i\,\xi)}}{-(1+i\,\xi)}\right]_{0}^{+\infty}}_{=0} - \int_{0}^{+\infty} \frac{e^{-x\,(1+i\,\xi)}}{-(1+i\,\xi)}\, \mathrm{d}x \\
&= \frac{1}{1-i\,\xi} \int_{-\infty}^{0} e^{x\,(1-i\,\xi)}\, \mathrm{d}x + \frac{1}{1+i\,\xi} \int_{0}^{+\infty} e^{-x\,(1+i\,\xi)}\, \mathrm{d}x \\
&= \frac{1}{1-i\,\xi} \left[\frac{e^{x\,(1-i\,\xi)}}{(1-i\,\xi)}\right]_{-\infty}^{0} + \frac{1}{1+i\,\xi} \left[\frac{e^{-x\,(1+i\,\xi)}}{-(1+i\,\xi)}\right]_{0}^{+\infty} \\
&= \frac{1}{1-i\,\xi} \left(\frac{1}{1-i\,\xi} - 0\right) + \frac{1}{1+i\,\xi} \left(0 + \frac{1}{1+i\,\xi}\right) \\
&= \frac{1}{(1-i\,\xi)^2} + \frac{1}{(1+i\,\xi)^2} \\
&= \frac{1-\xi^2+2\,i\,\xi+1-\xi^2-2\,i\,\xi}{(1+\xi^2)^2} = 2\,\frac{1-\xi^2}{(1+\xi^2)^2}.
\end{aligned}$$

Another method is to use the Fourier transform of

$$h(x) = x\, e^{-|x|}\, \operatorname{sgn} x,$$

(see last Tutorial), which is equal to $-\dfrac{2\,i\,\xi}{1+\xi^2}$, and use the formula

$$\mathcal{F}\{x\,h(x)\} = i\,\hat{h}'(\xi),$$

which gives

$$\mathcal{F}\{x\,h(x)\} = i\,\frac{\mathrm{d}}{\mathrm{d}\,\xi}\left(\frac{-2\,i\,\xi}{1+\xi^2}\right) = i\,\frac{-2\,i\,(1-\xi^2)}{(1+\xi^2)^2} = 2\,\frac{1-\xi^2}{(1+\xi^2)^2}.$$

■

Solution 6.15 *(Exercise 6.15).* The function is integrable:

$$\int_{\mathbb{R}} \chi_{[a,b]}(x)\, \mathrm{d}x = \int_{a}^{b} \mathrm{d}x = b - a.$$

The Fourier transform is given by $b-a$ for $\xi = 0$, and for $\xi \neq 0$:

$$\hat{\chi}_{[a,b]}(\xi) = \int_{a}^{b} e^{-i\,x\,\xi}\, \mathrm{d}x = \left[\frac{e^{-i\,x\,\xi}}{-i\,\xi}\right]_{a}^{b} = \frac{e^{-i\,\xi\,a} - e^{-i\,\xi\,b}}{i\,\xi}.$$

We verify that it is continuous at $\xi = 0$. Indeed, using de L'Hôpital we find

$$\lim_{\xi \to 0} \frac{e^{-i\xi a} - e^{-i\xi b}}{i\xi} = \lim_{\xi \to 0} \frac{(-i\,a)\, e^{-i\xi a} - (-i\,b)\, e^{-i\xi b}}{i} = b - a.$$

We prove that the Fourier transform goes to 0 as $|\xi| \to +\infty$:

$$\left| \frac{e^{-i\xi a} - e^{-i\xi b}}{i\xi} \right| \leq \frac{|e^{-i\xi a}| + |e^{-i\xi b}|}{|\xi|} \leq \frac{2}{|\xi|} \to 0$$

as $|\xi| \to +\infty$. ■

Solution 6.16 *(Exercise 6.16).* Let $\xi \in \mathbb{R}^n$. Using Fubini's theorem (Theorem 3.8) we find

$$\begin{aligned}
\mathcal{F}\left\{e^{-\|x\|^2}\right\}(\xi) &= \int_{\mathbb{R}^n} e^{-i\,\langle x,\xi\rangle}\, e^{-\|x\|^2}\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} e^{-i\,(x_1\,\xi_1 + x_2\,\xi_2 + \ldots + x_n\,\xi_n)}\, e^{-(x_1^2 + x_2^2 + \ldots + x_n^2)}\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \prod_{i=1}^{n} e^{-i\,x_i\,\xi_i} \prod_{i=1}^{n} e^{-x_i^2}\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \prod_{i=1}^{n} (e^{-i\,x_i\,\xi_i}\, e^{-x_i^2})\,\mathrm{d}x \\
&= \prod_{i=1}^{n} \int_{-\infty}^{+\infty} e^{-i\,x_i\,\xi_i}\, e^{-x_i^2}\,\mathrm{d}x \\
&= \prod_{i=1}^{n} \sqrt{\pi}\, e^{-\frac{\xi_i^2}{4}} = \pi^{\frac{n}{2}}\, e^{-\frac{\|\xi\|^2}{4}}.
\end{aligned}$$

■

Solution 6.17 *(Exercise 6.17).* First, we notice that

$$\frac{\mathrm{d}}{\mathrm{d}x} \frac{1}{1 + x^2} = \frac{-2\,x}{(1 + x^2)^2}.$$

Thus, we can use $\mathcal{F}\{f'(x)\} = i\,\xi\,\hat{f}(\xi)$ and find

$$\mathcal{F}\left\{\frac{x}{(1 + x^2)^2}\right\} = -\frac{1}{2}\,\mathcal{F}\left\{\frac{-2\,x}{(1 + x^2)^2}\right\} = -\frac{1}{2}\,i\,\xi\,\mathcal{F}\left\{\frac{1}{1 + x^2}\right\} = -\frac{i\,\xi\,\pi}{2}\, e^{-|\xi|}.$$

Consider the second. In this case, we can exploit the previous result and apply $\mathcal{F}\{x\,f(x)\} = i\,\hat{f}\,'(\xi)$:

$$\begin{aligned}
\mathcal{F}\left\{\frac{x^2}{(1 + x^2)^2}\right\} &= i\,\frac{\mathrm{d}}{\mathrm{d}\xi}\left(-\frac{i\,\xi\,\pi}{2}\, e^{-|\xi|}\right) \\
&= i\left(-\frac{1}{2}\,i\,\pi\, e^{-|\xi|} - \frac{1}{2}\,i\,\xi\,\pi\,(-\operatorname{sgn}\xi)\, e^{-|\xi|}\right) \\
&= \frac{\pi}{2} e^{-|\xi|} - \frac{\pi}{2}\,(\xi\,\operatorname{sgn}\xi)\, e^{-|\xi|} = \frac{\pi}{2}\, e^{-|\xi|} - \frac{\pi}{2}\,|\xi|\, e^{-|\xi|} = \frac{\pi}{2}(1 - |\xi|)\, e^{-|\xi|}.
\end{aligned}$$

Consider the last one. Noting that

$$\frac{1}{(1+x^2)^2} = \frac{x^2+1}{(1+x^2)^2} - \frac{x^2}{(1+x^2)^2} = \frac{1}{1+x^2} - \frac{x^2}{(1+x^2)^2},$$

using the linearity of the Fourier transform we obtain

$$\begin{aligned}\mathcal{F}\left\{\frac{1}{(1+x^2)^2}\right\} &= \mathcal{F}\left\{\frac{1}{1+x^2}\right\} - \mathcal{F}\left\{\frac{x^2}{(1+x^2)^2}\right\} \\ &= \pi\, e^{-|\xi|} - \frac{\pi}{2}\,(1-|\xi|)\, e^{-|\xi|} = \frac{\pi}{2}\,(1+|\xi|)\, e^{-|\xi|}.\end{aligned}$$

■

Solution 6.18 *(Exercise 6.18).* By definition of tensor product, for all $\boldsymbol{x} \in \mathbb{R}^n$ we have

$$(f_1 \otimes \ldots \otimes f_n)(\boldsymbol{x}) = f_1(x_1) \cdots f_n(x_n).$$

The converse of Fubini's theorem (Theorem 3.8) ensures that $f_1 \otimes \ldots \otimes f_n$ is integrable in $\mathbb{R}^n$. In other words

$$\begin{aligned}\int_{\mathbb{R}} |f(x_1)|\,\mathrm{d}x_1 \cdot \cdots \cdot \int_{\mathbb{R}} |f(x_n)|\,\mathrm{d}x_n &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \left(\ldots \int_{\mathbb{R}} |f(x_1)|\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \right) \ldots \mathrm{d}x_n \\ &= \int_{\mathbb{R}^n} |f(x_1)| \cdots |f(x_n)|\,\mathrm{d}x_1 \ \ldots\ \mathrm{d}x_n = \int_{\mathbb{R}^n} |(f_1 \otimes \ldots \otimes f_n)(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}.\end{aligned}$$

We compute the Fourier transform using Fubini's theorem (Theorem 3.8):

$$\begin{aligned}\widehat{f_1 \otimes \ldots \otimes f_n}(\boldsymbol{\xi}) &= \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}} f_1(x_1) \cdots f_n(x_n)\,\mathrm{d}\boldsymbol{x} \\ &= \int_{\mathbb{R}} e^{-i\,x_1\,\xi_1}\, f_1(x_1)\,\mathrm{d}x_1 \cdots \int_{\mathbb{R}} e^{-i\,x_n\,\xi_n}\, f_n(x_n)\,\mathrm{d}x_n \\ &= \hat{f}_1(x_1) \cdots \hat{f}_n(x_n) = (\hat{f}_1 \otimes \ldots \otimes \hat{f}_n)(\boldsymbol{\xi}).\end{aligned}$$

The characteristic function of B is

$$\begin{aligned}\chi_B : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ \boldsymbol{x} &\mapsto \begin{cases} 1 & \text{if } \boldsymbol{x} \in B \text{ (i.e. if } x_j \in [-1,1] \text{ for all } j \in \{1,\ldots,n\} \\ 0 & \text{otherwise} \end{cases}\end{aligned}$$

We find

$$\hat{\chi}_B(\boldsymbol{\xi}) = \int_{\mathbb{R}^n} e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\, \chi_B(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_B e^{-i\,\boldsymbol{x}\cdot\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{x}.$$

Since

$$\chi_B = \chi_{[-1,1]} \otimes \chi_{[-1,1]} \otimes \ldots \otimes \chi_{[-1,1]},$$

from what we have found previously

$$\hat{\chi}_B = \hat{\chi}_{[-1,1]} \otimes \hat{\chi}_{[-1,1]} \otimes \ldots \otimes \hat{\chi}_{[-1,1]},$$

and we have

$$\hat{\chi}_B(\boldsymbol{\xi}) = \hat{\chi}_{[-1,1]}(\xi_1) \cdot \hat{\chi}_{[-1,1]}(\xi_2) \cdots \hat{\chi}_{[-1,1]}(\xi_n).$$

For example, if $\xi_j \neq 0$ for all $j \in \{1, \ldots, n\}$, we have

$$\hat{\chi}_B(\boldsymbol{\xi}) = \prod_{j=1}^{n} 2 \frac{\sin \xi_j}{\xi_j}.$$

If $\boldsymbol{\xi} = \mathbf{0}$, we have $\hat{\chi}_B(\boldsymbol{\xi}) = 2^n$. If some components of $\boldsymbol{\xi}$ are zero, we have intermediate expressions between these two.

■

Solution 6.19 *(Exercise 6.19).* First we prove that **(A)** implies **(B)**. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $m \in \mathbb{N}$. We use **(A)** with $\boldsymbol{\alpha}$ and m: there exists $c_1 > 0$ such that

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \leq c_1.$$

Next, we use **(A)** with the same $\boldsymbol{\alpha}$ and $m = 0$: there exists $c_2 > 0$ such that

$$|\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \leq c_2.$$

Adding these two inequalities together shows that there exists $c = c_1 + c_2$ such that

$$(1 + \|\boldsymbol{x}\|^m) \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \leq c_1 + c_2 = c.$$

This is exactly **(B)**.

Now, we prove that **(B)** implies **(A)**. In this case,

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \leq (1 + \|\boldsymbol{x}\|^m) \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| \leq c,$$

which shows that **(A)** holds. ■

Solution 6.20 *(Exercise 6.20).* Let $\varphi, \psi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. Let $a, b \in \mathbb{C}$. We have to prove that $a\varphi + b\psi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$. Clearly, $a\varphi + b\psi$ is of class C^∞. Let $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $m \in \mathbb{N}$. Since $\varphi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, there exists $c_1 > 0$ such that for all $\boldsymbol{x} \in \mathbb{R}^n$

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \varphi(\boldsymbol{x})| < c_1.$$

Likewise, since $\psi \in \mathcal{S}(\mathbb{R}^n; \mathbb{C})$, there exists $c_2 > 0$ such that for all $\boldsymbol{x} \in \mathbb{R}^n$

$$\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}} \psi(\boldsymbol{x})| < c_2.$$

Then, using the triangle inequality, for every $\boldsymbol{x} \in \mathbb{R}^n$ we can write

$$\begin{aligned}
\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}(a\varphi + b\psi)| &= \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}(a\varphi) + \partial^{\boldsymbol{\alpha}}(b\psi)| \\
&= \|\boldsymbol{x}\|^m \, |a\, \partial^{\boldsymbol{\alpha}}\varphi + b\, \partial^{\boldsymbol{\alpha}}\psi| \\
&\leq \|\boldsymbol{x}\|^m \, (|a\, \partial^{\boldsymbol{\alpha}}\varphi| + |b\, \partial^{\boldsymbol{\alpha}}\psi|) \\
&= |a| \, \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}\varphi| + |b| \, \|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}\psi| < |a|\, c_1 + |b|\, c_2,
\end{aligned}$$

which says that $\|\boldsymbol{x}\|^m \, |\partial^{\boldsymbol{\alpha}}(a\varphi + b\psi)|$ is bounded, as required.

■

Solution 6.21 *(Exercise 6.21).* We proceed by induction over $|\boldsymbol{\alpha}|$. Let $\boldsymbol{\alpha} = \boldsymbol{e}_j$, where $j \in \{1, \ldots, n\}$. Then

$$\frac{\partial}{\partial x_j} e^{-\|\boldsymbol{x}\|^2} = -2\, x_j\, e^{-\|\boldsymbol{x}\|^2},$$

and $-2\, x_j$ is indeed a polynomial of degree $|\boldsymbol{\alpha}| = 1$. Now, suppose that the result is true for $\boldsymbol{\alpha} \in \mathbb{N}^n$, i.e., there exists a polynomial $p(\boldsymbol{x})$ of degree $|\boldsymbol{\alpha}|$ such that $\partial^{\boldsymbol{\alpha}} e^{-\|\boldsymbol{x}\|^2} = p(\boldsymbol{x})\, e^{-\|\boldsymbol{x}\|^2}$. We prove it for $\boldsymbol{\beta} = \boldsymbol{\alpha} + \boldsymbol{e}_i$, where $i \in \{1, \ldots, n\}$. We have

$$\partial^{\boldsymbol{\beta}} e^{-\|\boldsymbol{x}\|^2} = \partial^{\boldsymbol{e}_i} \partial^{\boldsymbol{\alpha}} e^{-\|\boldsymbol{x}\|^2} = \partial^{\boldsymbol{e}_i} \left(p(\boldsymbol{x})\, e^{-\|\boldsymbol{x}\|^2} \right) = \frac{\partial}{\partial x_i} p(\boldsymbol{x})\, e^{-\|\boldsymbol{x}\|^2} + p(\boldsymbol{x})(-2\, x_i)\, e^{-\|\boldsymbol{x}\|^2} = \left(\frac{\partial}{\partial x_i} p(\boldsymbol{x}) - 2\, x_i\, p(\boldsymbol{x}) \right) e^{-\|\boldsymbol{x}\|^2},$$

and $\frac{\partial}{\partial x_i} p(\boldsymbol{x}) - 2\, x_i\, p(\boldsymbol{x})$ is a polynomial of degree $|\boldsymbol{\alpha}| + 1$. ■

Solution 6.22 *(Exercise 6.22).* Let $\boldsymbol{\alpha} \in \mathbb{N}^n$ and $m \in \mathbb{N}$. Let p be a polynomial in the indeterminate $\boldsymbol{x}$ of degree r. Then,

$$\begin{aligned}
\|\boldsymbol{x}\|^m \left| \partial^{\boldsymbol{\alpha}}\, p(\partial)\, \varphi_k - \partial^{\boldsymbol{\alpha}}\, p(\partial)\, \varphi \right| &= \|\boldsymbol{x}\|^m \left| \partial^{\boldsymbol{\alpha}}\, p(\partial)\, (\varphi_k - \varphi) \right| \\
&= \|\boldsymbol{x}\|^m \left| \partial^{\boldsymbol{\alpha}} \sum_{\boldsymbol{\beta} \le r} a_{\boldsymbol{\beta}}\, \partial^{\boldsymbol{\beta}}\, (\varphi_k - \varphi) \right| \\
&= \|\boldsymbol{x}\|^m \left| \sum_{\boldsymbol{\beta} \le r} a_{\boldsymbol{\beta}}\, \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, (\varphi_k - \varphi) \right| \\
&\le \|\boldsymbol{x}\|^m \sum_{\boldsymbol{\beta} \le r} a_{\boldsymbol{\beta}} \left| \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, \varphi_k - \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, \varphi \right| \\
&= \sum_{\boldsymbol{\beta} \le r} a_{\boldsymbol{\beta}}\, \|\boldsymbol{x}\|^m \left| \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, \varphi_k - \partial^{\boldsymbol{\alpha}+\boldsymbol{\beta}}\, \varphi \right|,
\end{aligned}$$

which converges to 0 in $\mathcal{S}(\mathbb{R}^n; \mathbb{C})$. ■

Solution 6.23 *(Exercise 6.23).* The function ϕ is not defined at the origin. For $x > 0$, the first derivative is

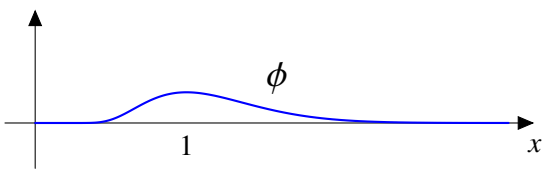
$$\phi'(x) = \left(\frac{2}{x^3} - 2\, x \right) e^{-\frac{1}{x^2} - x^2},$$

and the second derivative is

$$\phi''(x) = \left[\left(\frac{2}{x^3} - 2\, x \right)^2 - \frac{6}{x^4} - 2 \right] e^{-\frac{1}{x^2} - x^2}.$$

Continuing in this fashion, we see that each derivative is given by a rational function multiplied by $e^{-\frac{1}{x^2} - x^2}$. More rigorously, we can proceed by induction, to see that for $x > 0$

$$\phi^{(j)}(x) = \left[\frac{\mathrm{d}^{j-1}}{\mathrm{d}x^{j-1}} \left(\frac{2}{x^3} - 2\, x \right) + \left(\frac{2}{x^3} - 2\, x \right)^j \right] e^{-\frac{1}{x^2} - x^2}.$$

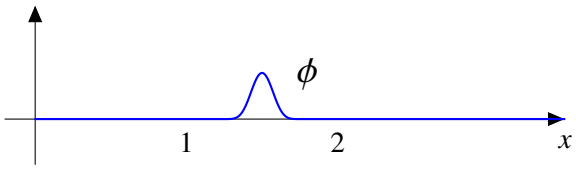


To prove that ϕ is of class C^∞, it is sufficient to show that the limit of all the derivatives, as $x \to 0^+$, are equal to 0. This is certainly the case, because

$$\lim_{x\to 0^+} \frac{1}{x^\alpha} e^{-\frac{1}{x^2}-x^2} = 0 \qquad \forall\, \alpha > 0.$$

Moreover, any derivative of ϕ will go to zero faster than any rational function in x, as $x \to +\infty$.

Consider the second statement. By definition, ϕ is zero outside the closed interval $[a, b]$, and therefore ϕ has compact support. We only need to convince ourselves that ϕ is of class C^∞. This follows from the fact that it is the product of compositions of a function of class C^∞ with a constant function, which is also of class C^∞.



■

Solution 6.24 *(Exercise 6.24).* Consider a sequence $(\varphi_k)_{k\in\mathbb{N}}$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, and let $\varphi \in \mathcal{D}(\mathbb{R}^n;\mathbb{C})$. If $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, then $\bar\varphi_k \xrightarrow{k\to\infty} \bar\varphi$ in $\mathcal{D}(\mathbb{R}^n;\mathbb{C})$, because the convergence of a sequence is equivalent to the convergence of the real and of the imaginary parts. Then,

$$\langle \bar T, \varphi_k\rangle = \overline{\langle T, \bar\varphi_k\rangle} \xrightarrow{k\to\infty} \overline{\langle T, \bar\varphi\rangle} = \langle \bar T, \varphi\rangle.$$

Thus, $\bar T \in \mathcal{D}'(\mathbb{R}^n;\mathbb{C})$.

Likewise, if $(\varphi_k)_{k\in\mathbb{N}}$ is a sequence in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$, then $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$ implies $\bar\varphi_k \xrightarrow{k\to\infty} \bar\varphi$ in $\mathcal{S}(\mathbb{R}^n;\mathbb{C})$, and therefore again

$$\langle \bar T, \varphi_k\rangle = \overline{\langle T, \bar\varphi_k\rangle} \xrightarrow{k\to\infty} \overline{\langle T, \bar\varphi\rangle} = \langle \bar T, \varphi\rangle.$$

Thus, $\bar T \in \mathcal{S}'(\mathbb{R}^n;\mathbb{C})$. ■

Solution 6.25 *(Exercise 6.25).* We must prove that $\hat T(-\boldsymbol{\xi}) = -\hat T(\boldsymbol{\xi})$, or, equivalently, that $\check{\hat T} = \widehat{-T}$. Since T is odd, $\check T = -T$. Given $\varphi \in \mathcal{S}(\mathbb{R}^n;\mathbb{C})$ we find

$$\langle \check{\hat T}, \varphi\rangle = \langle \check T, \hat\varphi\rangle = \langle -T, \hat\varphi\rangle = \langle \widehat{-T}, \varphi\rangle.$$

■

Solution 6.26 *(Exercise 6.26).* Since T is λ-homogeneous, for every $t > 0$ we have $T(t\,\boldsymbol{x}) = t^\lambda\, T(\boldsymbol{x})$. Taking the Fourier transform of both sides and using

$$\widehat{T \circ A} = \frac{1}{|\det A|} \hat T \circ A^{-\top}$$

with $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that $A\,\boldsymbol{x} = t\,\boldsymbol{x}$, i.e., $A = t\,I$ (which implies $|\det A| = t^n$), we obtain

$$t^{-n}\,\hat{T} \circ (t^{-1}\,I) = t^{\lambda}\,\hat{T},$$

which can be written as

$$t^{-n}\,\hat{T}\left(\frac{1}{t}\,\boldsymbol{\xi}\right) = t^{\lambda}\,\hat{T}(\boldsymbol{\xi}).$$

Since this is true for all $t > 0$, we can define a new parameter $\tau = 1/t$ and the latter becomes

$$\tau^n\,\hat{T}(\tau\,\boldsymbol{\xi}) = \tau^{-\lambda}\,\hat{T}(\boldsymbol{\xi}) \qquad \Rightarrow \qquad \hat{T}(\tau\,\boldsymbol{\xi}) = \tau^{-\lambda-n}\,\hat{T}(\boldsymbol{\xi}),$$

which shows that $\hat{T}$ is $(-\lambda - n)$-homogeneous. ■

Solution 6.27 *(Exercise 6.27).* We find immediately

$$\widehat{\delta^{(k)}} = (i\,\boldsymbol{\xi})^k\,\hat{\delta} = (i\,\boldsymbol{\xi})^k$$

and

$$\widehat{(-i\,x)^k\cdot 1} = \partial^k\hat{1} = 2\,\pi\,\delta^{(k)} \quad \Rightarrow \quad -i^k\,\widehat{x^k} = 2\,\pi\,\delta^{(k)} \quad \Rightarrow \quad \widehat{x^k} = i^k\,2\,\pi\,\delta^{(k)}.$$

■

Solution 6.28 *(Exercise 6.28).* We find

$$\cos^2 x = \left(\frac{e^{i\,x} + e^{-i\,x}}{2}\right)^2 = \frac{e^{2\,i\,x} + e^{-2\,i\,x} + 2}{4}.$$

The Fourier transform is, by linearity,

$$\widehat{\cos^2 x} = \frac{\widehat{e^{2\,i\,x}} + \widehat{e^{-2\,i\,x}} + \hat{2}}{4}.$$

We have

$$e^{2\,i\,x} = e^{2\,i\,x}\,\hat{\delta} = \widehat{\delta(\cdot + 2)} \quad \Rightarrow \quad \widehat{e^{2\,i\,x}} = \widehat{\widehat{\delta(\cdot + 2)}} = 2\,\pi\,\widetilde{\delta(\cdot + 2)} = 2\,\pi\,\delta(\xi - 2).$$

Likewise, $e^{-2\,i\,x} = 2\,\pi\,\delta(\xi + 2)$. Finally, $\hat{2} = 4\,\pi\,\delta(\xi)$. We obtain

$$\widehat{\cos^2 x} = \frac{2\,\pi\big(\delta(\xi - 2) + \delta(\xi + 2) + 2\,\delta(\xi)\big)}{4} = \pi\,\delta(\xi) + \frac{\pi}{2}\big(\delta(\xi - 2) + \delta(\xi + 2)\big).$$

In the same way, we can find

$$\widehat{\sin^2 x} = \pi\,\delta(\xi) - \frac{\pi}{2}\big(\delta(\xi - 2) + \delta(\xi + 2)\big).$$

■

Solution 6.29 *(Exercise 6.29).* From $\operatorname{sgn} x = 2\,H(x) - 1$ we find $\widehat{\operatorname{sgn}}(\xi) = 2\,\hat{H}(\xi) - \hat{1}$. It follows that

$$\frac{2}{i}\ \operatorname{pv}\frac{1}{\xi} = 2\,\hat{H}(\xi) - 2\,\pi\,\delta(\xi),$$

which gives

$$\hat{H}(\xi) = \frac{1}{i}\ \operatorname{pv}\frac{1}{\xi} + \pi\,\delta(\xi).$$

To prove the second, we notice that

$$\operatorname{sgn}(x) = -2\,H(-x) + 1 \quad \Rightarrow \quad 2\,H(-x) = 1 - \operatorname{sgn} x,$$

and taking the Fourier transform yields

$$2\,\widehat{H(-\cdot)} = \hat{1} - \widehat{\operatorname{sgn}}(\xi) \quad \Rightarrow \quad 2\,\widehat{H(-\cdot)} = 2\,\pi\,\delta(\xi) - \frac{2}{i}\ \operatorname{pv}\frac{1}{\xi} \quad \Rightarrow \quad \widehat{H(-\cdot)} = \pi\,\delta(\xi) - \frac{1}{i}\ \operatorname{pv}\frac{1}{\xi}.$$

■

Solution 6.30 *(Exercise 6.30).* Since the function

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\mapsto x \end{aligned}$$

is in C^∞ and $x\,T = 0$, we have $\operatorname{supp} T \subseteq \{x \in \mathbb{R}\,|\,x = 0\} = \{0\}$. Since $\operatorname{supp} T$ is compact, then $T \in \mathcal{S}'(\mathbb{R};\mathbb{R})$.

Let $T = \hat{S}$, where $S \in \mathcal{S}'(\mathbb{R};\mathbb{R})$. Let $\varphi \in \mathcal{S}(\mathbb{R};\mathbb{R})$. Then

$$\begin{aligned} \langle x\,T, \varphi\rangle &= \langle x\,\hat{S}, \varphi\rangle \\ &= \langle \hat{S}, x\,\varphi\rangle \\ &= \langle S, \widehat{x\,\varphi}\rangle \\ &= \langle S, i\,\hat{\varphi}'\rangle \\ &= \langle i\,S, \hat{\varphi}'\rangle \\ &= \langle -i\,S', \hat{\varphi}\rangle \\ &= \langle -i\,\widehat{S'}, \varphi\rangle, \end{aligned}$$

where we used the formulae $\widehat{-i\,x\,\varphi} = \hat{\varphi}'$ and $\widehat{x\,\varphi} = i\,\hat{\varphi}'$. Now, since it is assumed that $x\,T = 0$, the previous chain of identities gives

$$-i\,\widehat{S'} = 0 \quad \Rightarrow \quad \widehat{S'} = 0 \quad \Rightarrow \quad S' = 0 \quad \Rightarrow \quad S = c,$$

where c is a constant. We obtain $\hat{S} = c\,2\,\pi\,\delta$, and therefore

$$T(x) = c\,\delta(x).$$

■