Lecture Notes on Measure Theory and Functional Analysis

P. Cannarsa & T. D'Aprile

Dipartimento di Matematica
Università di Roma "Tor Vergata"

cannarsa@mat.uniroma2.it daprile@mat.uniroma2.it

aa 2006/07

Contents

Chapter 1

Measure Spaces

1.1 Algebras and σ–algebras of sets

1.1.1 Notation and preliminaries

We shall denote by X a nonempty set, by $\mathcal{P}(X)$ the set of all parts (i.e., subsets) of X, and by $\varnothing$ the empty set.

For any subset A of X we shall denote by A^c its complement, i.e.,

$$A^c = \{x \in X \mid \; x \notin A\} \; .$$

For any $A, B \in \mathcal{P}(X)$ we set $A \setminus B = A \cap B^c$.

Let (A_n) be a sequence in $\mathcal{P}(X)$. The following *De Morgan* identity holds

$$\left(\bigcup_{n=1}^{\infty} A_n \right)^c = \bigcap_{n=1}^{\infty} A_n^c.$$

We define [(1)]

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \, , \qquad \liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m \, .$$

If $L := \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$, then we set $L = \lim_{n\to\infty} A_n$, and we say that (A_n) converges to L (in this case we shall write write $A_n \to L$).

[(1)] Observe the relationship with inf and sup limits for a sequence (a_n) of real numbers. We have $\limsup_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} \sup_{m\geq n} a_m$ and $\liminf_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} \inf_{m\geq n} a_m$.

Remark 1.1 (a) As easily checked, $\limsup_{n\to\infty} A_n$ (resp. $\liminf_{n\to\infty} A_n$) consists of those elements of X that belong to infinite elements of (A_n) (resp. that belong to all elements of (A_n) except perhaps a finite number). Therefore,

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$$

(b) It easy also to check that, if (A_n) is increasing $(A_n \subset A_{n+1},\ n \in \mathbb{N})$, then

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n,$$

whereas, if (A_n) is decreasing $(A_n \supset A_{n+1},\ n \in \mathbb{N})$, then

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

In the first case we shall write $A_n \uparrow L$, and in the second $A_n \downarrow L$.

1.1.2 Algebras and σ–algebras

Let $\mathcal{A}$ be a nonempty subset of $\mathcal{P}(X)$.

Definition 1.2 *$\mathcal{A}$ is said to be an* algebra *if*

(a) $\varnothing, X \in \mathcal{A}$

(b) $A, B \in \mathcal{A} \quad \Longrightarrow \quad A \cup B \in \mathcal{A}$

(c) $A \in \mathcal{A} \quad \Longrightarrow \quad A^c \in \mathcal{A}$

Remark 1.3 It is easy to see that, if $\mathcal{A}$ is an algebra and $A, B \in \mathcal{A}$, then $A \cap B$ and $A \setminus B$ belong to $\mathcal{A}$. Therefore, the symmetric difference

$$A \Delta B := (A \setminus B) \cup (B \setminus A)$$

also belongs to $\mathcal{A}$. Moreover, $\mathcal{A}$ is stable under finite union and intersection, that is

$$A_1, \ldots, A_n \in \mathcal{A} \quad \Longrightarrow \quad \begin{cases} A_1 \cup \cdots \cup A_n \in \mathcal{A} \\ A_1 \cap \cdots \cap A_n \in \mathcal{A}\,. \end{cases}$$

Definition 1.4 *An algebra $\mathcal{E}$ in $\mathcal{P}(X)$ is said to be a σ–*algebra *if, for any sequence (A_n) of elements of $\mathcal{E}$, we have $\bigcup_{n=1}^{\infty} A_n \in \mathcal{E}$.*

We note that, if $\mathcal{E}$ is a σ–algebra and $(A_n) \subset \mathcal{E}$, then $\bigcap_{n=1}^{\infty} A_n \in \mathcal{E}$ owing to the De Morgan identity. Moreover,

$$\liminf_{n\to\infty} A_n \in \mathcal{E}\,, \quad \limsup_{n\to\infty} A_n \in \mathcal{E}\,.$$

Example 1.5 The following examples explain the difference between algebras and a σ–algebras.

1. Obviously, $\mathcal{P}(X)$ and $\mathcal{E} = \{\varnothing, X\}$ are σ–algebras in X. Moreover, $\mathcal{P}(X)$ is the largest σ–algebra in X, and $\mathcal{E}$ the smallest.

2. In $X = [0, 1)$, the class $\mathcal{A}_0$ consisting of $\varnothing$, and of all finite unions

$$A = \cup_{i=1}^{n} [a_i, b_i) \text{ with } 0 \le a_i < b_i \le a_{i+1} \le 1\,, \tag{1.1}$$

is an algebra. Indeed, for A as in (1.1), we have

$$A^c = [0, a_1) \cup [b_1, a_2) \cup \cdots \cup [b_n, 1) \in \mathcal{A}_0$$

Moreover, in order to show that $\mathcal{A}_0$ is stable under finite union, it suffices to observe that the union of two (not necessarily disjoint) intervals $[a, b)$ and $[c, d)$ in $[0, 1)$ belongs to $\mathcal{A}_0$.

3. In an infinite set X consider the class

$$\mathcal{E} = \{A \in \mathcal{P}(X) \mid A \text{ is finite, or } A^c \text{ is finite }\}\,.$$

Then, $\mathcal{E}$ is an algebra. Indeed, the only point that needs to be checked is that $\mathcal{E}$ is stable under finite union. Let $A, B \in \mathcal{E}$. If A and B are both finite, then so is $A \cup B$. In all other cases, $(A \cup B)^c$ is finite.

4. In an uncountable set X consider the class

$$\mathcal{E} = \{A \in \mathcal{P}(X) \mid A \text{ is countable, or } A^c \text{ is countable }\}$$

(here, 'countable' stands for 'finite or countable'). Then, $\mathcal{E}$ is a σ–algebra. Indeed, $\mathcal{E}$ is stable under countable union: if (A_n) is a sequence in $\mathcal{E}$ and all A_n are countable, then so is $\cup_n A_n$; otherwise, $(\cup_n A_n)^c$ is countable.

Exercise 1.6 1. Show that algebra $\mathcal{A}_0$ in Example 1.5.2 fails to be a σ–algebra.

2. Show that algebra $\mathcal{E}$ in Example 1.5.3 fails to be a σ–algebra.

3. Show that σ–algebra $\mathcal{E}$ in Example 1.5.4 is strictly smaller than $\mathcal{P}(X)$.

4. Let $\mathcal{K}$ be a subset of $\mathcal{P}(X)$. Show that the intersection of all σ–algebras including $\mathcal{K}$, is a σ–algebra (the minimal σ–algebra including $\mathcal{K}$).

Let $\mathcal{K}$ be a subset of $\mathcal{P}(X)$.

Definition 1.7 *The intersection of all σ–algebras including $\mathcal{K}$ is called the* σ–algebra generated by $\mathcal{K}$, *and will be denoted by $\sigma(\mathcal{K})$.*

Exercise 1.8 In the following, let $\mathcal{K}\,,\ \mathcal{K}' \subset \mathcal{P}(X)$.

1. Show that, if $\mathcal{E}$ is a σ–algebra, then $\sigma(\mathcal{E}) = \mathcal{E}$.

2. Find $\sigma(\mathcal{K})$ for $\mathcal{K} = \{\varnothing\}$ and $\mathcal{K} = \{X\}$.

3. Show that, if $\mathcal{K} \subset \mathcal{K}' \subset \sigma(\mathcal{K})$, then $\sigma(\mathcal{K}') = \sigma(\mathcal{K})$.

Example 1.9 1. Let E be a metric space. The σ–algebra generated by all open subsets of E is called the *Borel* σ–algebra of E, and is denoted by $\mathcal{B}(E)$. Obviously, $\mathcal{B}(E)$ coincides with the σ–algebra generated by all closed subsets of E.

2. Let $X = \mathbb{R}$, and $\mathcal{I}$ be the class of all semi–closed intervals $[a, b)$ with $a \leq b$. Then, $\sigma(\mathcal{I})$ coincides with $\mathcal{B}(\mathbb{R})$. For let $[a, b) \subset \mathbb{R}$. Then, $[a, b) \in \mathcal{B}(\mathbb{R})$ since

$$[a, b) = \bigcap_{n=n_0}^{\infty} \left(a - \frac{1}{n}, b\right).$$

So, $\sigma(\mathcal{I}) \subset \mathcal{B}(\mathbb{R})$. Conversely, let A be an open set in $\mathbb{R}$. Then, as is well-known, A is the countable union of some family of open intervals [(2)]. Since any open interval (a, b) can be represented as

$$(a, b) = \bigcup_{n=n_0}^{\infty} \left[a + \frac{1}{n}, b\right),$$

(2) Indeed, each point $x \in A$ has an open interval $(p_x, q_x) \subset A$ with $p_x, q_x \in \mathbb{Q}$. Hence, A is contained in the union of the family $\{(p, q) \mid p, q \in \mathbb{Q}\,,\ (p, q) \subset A\}$, and this family is countable.

where $\frac{1}{n_0} < b - a$, we conclude that $A \in \sigma(\mathcal{I})$. Thus, $\mathcal{B}(\mathbb{R}) \subset \sigma(\mathcal{I})$.

Exercise 1.10 Let $\mathcal{E}$ be a σ–algebra in X, and $X_0 \subset X$.

1. Show that $\mathcal{E}_0 = \{A \cap X_0 \mid A \in \mathcal{E}\}$ is a σ–algebra in X_0.

2. Show that, if $\mathcal{E} = \sigma(\mathcal{K})$, then $\mathcal{E}_0 = \sigma(\mathcal{K}_0)$, where

$$\mathcal{K}_0 = \{A \cap X_0 \mid A \in \mathcal{K}\}.$$

HINT: $\mathcal{E}_0 \supset \sigma(\mathcal{K}_0)$ follows from point 1. To prove the converse, show that

$$\mathcal{F} := \left\{A \in \mathcal{P}(X) \mid A \cap X_0 \in \sigma(\mathcal{K}_0)\right\}$$

is a σ–algebra in X including $\mathcal{K}$.

1.2 Measures

1.2.1 Additive and σ–additive functions

Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra.

Definition 1.11 *Let $\mu : \mathcal{A} \to [0, +\infty]$ be such that $\mu(\varnothing) = 0$.*

- *We say that μ is* additive *if, for any family $A_1, ..., A_n \in \mathcal{A}$ of mutually disjoint sets, we have*

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$

- *We say that μ is* σ–additive *if, for any sequence $(A_n) \subset \mathcal{A}$ of mutually disjoint sets such that $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$, we have*

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

Remark 1.12 Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra.

1. Any σ–additive function on $\mathcal{A}$ is also additive.

2. If μ is additive, $A, B \in \mathcal{A}$, and $A \supset B$, then $\mu(A) = \mu(B) + \mu(A \setminus B)$. Therefore, $\mu(A) \geq \mu(B)$.

3. Let μ be additive on $\mathcal{A}$, and let $(A_n) \subset \mathcal{A}$ be mutually disjoint sets such that $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$. Then,

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \geq \sum_{k=1}^{n} \mu(A_k), \quad \text{for all } n \in \mathbb{N}.$$

   Therefore,

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \geq \sum_{k=1}^{\infty} \mu(A_k).$$

4. Any σ-additive function μ on $\mathcal{A}$ is also *countably subadditive*, that is, for any sequence $(A_n) \subset \mathcal{A}$ such that $\bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$,

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

5. In view of points 3 and 4 an additive function is σ-additive if and only if it is *countably subadditive.*

Definition 1.13 *A σ–additive function μ on an algebra $\mathcal{A} \subset \mathcal{P}(X)$ is said to be:*

- finite *if $\mu(X) < \infty$;*
- σ–finite *if there exists a sequence $(A_n) \subset \mathcal{A}$, such that $\bigcup_{n=1}^{\infty} A_n = X$ and $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.*

Example 1.14 In $X = \mathbb{N}$, consider the algebra

$$\mathcal{E} = \{A \in \mathcal{P}(X) \mid A \text{ is finite, or } A^c \text{ is finite }\}$$

of Example 1.5. The function $\mu : \mathcal{E} \to [0, \infty]$ defined as

$$\mu(A) = \begin{cases} \#(A) & \text{if } A \text{ is finite} \\ \infty & \text{if } A^c \text{ is finite} \end{cases}$$

(where $\#(A)$ stands for the number of elements of A) is σ–additive (**Exercise**). On the other hand, the function $\nu : \mathcal{E} \to [0, \infty]$ defined as

$$\nu(A) = \begin{cases} \sum_{n \in A} \frac{1}{2^n} & \text{if } A \text{ is finite} \\ \infty & \text{if } A^c \text{ is finite} \end{cases}$$

is additive but not σ–additive (**Exercise**).

For an additive function, the σ–additivity of μ is equivalent to continuity in the sense of the following proposition.

Proposition 1.15 *Let μ be additive on $\mathcal{A}$. Then (i) $\Leftrightarrow$ (ii), where:*

(i) *μ is σ–additive;*

(ii) *(A_n) and $A \subset \mathcal{A}$, $A_n \uparrow A \implies \mu(A_n) \uparrow \mu(A)$.*

Proof. (i)⇒(ii) Let $(A_n), A \subset \mathcal{A}$, $A_n \uparrow A$. Then,

$$A = A_1 \cup \bigcup_{n=1}^{\infty} (A_{n+1} \setminus A_n) ,$$

the above being disjoint unions. Since μ is σ–additive, we deduce that

$$\mu(A) = \mu(A_1) + \sum_{n=1}^{\infty} (\mu(A_{n+1}) - \mu(A_n)) = \lim_{n \to \infty} \mu(A_n),$$

and (ii) follows.

(ii)⇒(i) Let $(A_n) \subset \mathcal{A}$ be a sequence of mutually disjoint sets such that $A = \bigcup_{k=1}^{\infty} A_k \in \mathcal{A}$. Define

$$B_n = \bigcup_{k=1}^{n} A_k .$$

Then, $B_n \uparrow A$. So, in view of (ii), $\mu(B_n) = \sum_{k=1}^{n} \mu(A_k) \uparrow \mu(A)$. This implies (i). □

Proposition 1.16 *Let μ be σ–additive on $\mathcal{A}$. If $\mu(A_1) < \infty$ and $A_n \downarrow A$ with $A \in \mathcal{A}$, then $\mu(A_n) \downarrow \mu(A)$.*

Proof. We have

$$A_1 = \bigcup_{k=1}^{\infty} (A_k \setminus A_{k+1}) \cup A$$

the above being disjoint unions. Consequently,

$$\mu(A_1) = \sum_{k=1}^{\infty} (\mu(A_k) - \mu(A_{k+1})) + \mu(A) = \mu(A_1) - \lim_{n\to+\infty} \mu(A_n) + \mu(A)\,.$$

Since $\mu(A_1) < +\infty$, the conclusion follows. □

Example 1.17 The conclusion of Proposition 1.16 above may be false without assuming $\mu(A_1) < \infty$. This is easily checked taking $\mathcal{E}$ and μ as in Example 1.14, and $A_n = \{m \in \mathbb{N} \mid m \geq n\}$.

Corollary 1.18 *Let μ be a finite σ–additive function on a σ–algebra $\mathcal{E}$. Then, for any sequence (A_n) of subsets of $\mathcal{E}$, we have*

$$\mu\left(\liminf_{n\to\infty} A_n\right) \leq \liminf_{n\to\infty} \mu(A_n) \leq \limsup_{n\to\infty} \mu(A_n) \leq \mu\left(\limsup_{n\to\infty} A_n\right). \quad (1.2)$$

In particular, $A_n \to A \Longrightarrow \mu(A_n) \to \mu(A)$.

Proof. Set $L = \limsup_{n\to\infty} A_n$. Then we can write $L = \bigcap_{n=1}^{\infty} B_n$, where $B_n = \bigcup_{m=n}^{\infty} A_m \downarrow L$. Now, by Proposition 1.16 it follows that

$$\mu(L) = \lim_{n\to\infty} \mu(B_n) = \inf_{n\in\mathbb{N}} \mu(B_n) \geq \inf_{n\in\mathbb{N}} \sup_{m\geq n} \mu(A_m) = \limsup_{n\to\infty} \mu(A_n).$$

Thus, we have proved that

$$\limsup_{n\to\infty} \mu(A_n) \leq \mu\left(\limsup_{n\to\infty} A_n\right).$$

The remaining part of (1.2) can be proved similarly. □

1.2.2 Borel–Cantelli Lemma

The following result is very useful as we shall see later.

Lemma 1.19 *Let μ be a finite σ–additive function on a σ–algebra $\mathcal{E}$. Then, for any sequence (A_n) of subsets of $\mathcal{E}$ satisfying $\sum_{n=1}^{\infty} \mu(A_n) < +\infty$, we have*

$$\mu \left(\limsup_{n\to\infty} A_n \right) = 0\,.$$

Proof. Set $L = \limsup_{n\to\infty} A_n$. Then, $L = \bigcap_{n=1}^{\infty} B_n$, where $B_n = \bigcup_{m=n}^{\infty} A_m$ decreases to L. Consequently,

$$\mu(L) \le \mu(B_n) \le \sum_{m=n}^{\infty} \mu(A_m)$$

for all $n \in \mathbb{N}$. As $n \to \infty$, we obtain $\mu(L) = 0$. □

1.2.3 Measure spaces

Definition 1.20 *Let $\mathcal{E}$ be a σ–algebra of subsets of X.*

- *We say that the pair $(X, \mathcal{E})$ is a* measurable space*.*
- *A σ–additive function $\mu : \mathcal{E} \to [0, +\infty]$ is called a* measure *on $(X, \mathcal{E})$.*
- *The triple $(X, \mathcal{E}, \mu)$, where μ is a measure on a measurable space $(X, \mathcal{E})$, is called a* measure *space.*
- *A measure μ is called a* probability measure *if $\mu(X) = 1$.*
- *A measure μ is said to be* complete *if*

  $$A \in \mathcal{E}\,,\ B \subset A\,,\ \mu(A) = 0 \quad \Rightarrow \quad B \in \mathcal{E}$$

  (and so $\mu(B) = 0$).
- *A measure μ is said to be* concentrated on a set *$A \in \mathcal{E}$ if $\mu(A^c) = 0$. In this case we say that A is a support of μ.*

Example 1.21 Let X be a nonempty set and $x \in X$. Define, for every $A \in \mathcal{P}(X)$,

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

Then, δ_x is a measure in X, called the *Dirac measure in* x. Such a measure is concentrated on the singleton $\{x\}$.

Example 1.22 In a set X let us define, for every $A \in \mathcal{P}(X)$,

$$\mu(A) = \begin{cases} \#(A) & \text{if } A \text{ is finite} \\ \infty & \text{if } A \text{ is infinite} \end{cases}$$

(see Example 1.14). Then, μ is a measure in X, called the *counting measure* on X. It is easy to see that μ is finite if and only if X is finite, and that μ is σ–finite if and only if X is countable.

Let $(X, \mathcal{E}, \mu)$ be a measure space and let $A \in \mathcal{E}$.

Definition 1.23 *The* restriction *of* μ *to* A *(or* μ restricted to A*), written* $\mu \llcorner A$*, is the set function*

$$(\mu \llcorner A)(B) = \mu(A \cap B) \qquad \forall B \in \mathcal{E}\,. \tag{1.3}$$

Exercise 1.24 Show that $\mu \llcorner A$ is a measure on $(X, \mathcal{E})$.

1.3 The basic extension theorem

A natural question arising both in theory and applications is the following.

Problem 1.25 *Let* $\mathcal{A}$ *be an algebra in* X*, and* μ *be an additive function in* $\mathcal{A}$*. Does there exist a* σ*–algebra* $\mathcal{E}$ *including* $\mathcal{A}$*, and a measure* $\overline{\mu}$ *on* $(X, \mathcal{E})$ *that extends* μ*, i.e.,*

$$\overline{\mu}(A) = \mu(A) \qquad \forall A \in \mathcal{A}? \tag{1.4}$$

Should the above problem have a solution, one could assume $\mathcal{E} = \sigma(\mathcal{A})$ since $\sigma(\mathcal{A})$ would be included in $\mathcal{E}$ anyways. Moreover, for any sequence $(A_n) \subset \mathcal{A}$ of mutually disjoint sets such that $\bigcup_{k=1}^\infty A_k \in \mathcal{A}$, we would have

$$\mu\Big(\bigcup_{k=1}^\infty A_k\Big) = \overline{\mu}\Big(\bigcup_{k=1}^\infty A_k\Big) = \sum_{k=1}^\infty \overline{\mu}(A_k) = \sum_{k=1}^\infty \mu(A_k)\,.$$

Thus, for Problem 1.25 to have a positive answer, μ must be σ–additive. The following remarkable result shows that such a property is also sufficient for the existence of an extension, and more. We shall see an important application of this result to the construction of the Lebesgue measure later on in this chapter.

Theorem 1.26 *Let $\mathcal{A}$ be an algebra, and $\mu\colon \mathcal{A} \to [0, +\infty]$ be σ–additive. Then, μ can be extended to a measure on $\sigma(\mathcal{A})$. Moreover, such an extension is unique if μ is σ–finite.*

To prove the above theorem we need to develop suitable tools, namely Halmos' Monotone Class Theorem for uniqueness, and the concepts of outer measure and additive set for existence. This is what we shall do in the next sessions.

1.3.1 Monotone classes

Definition 1.27 *A nonempty class $\mathcal{M} \subset \mathcal{P}(X)$ is called a* monotone class *in X if, for any sequence $(A_n) \subset \mathcal{M}$,*

- $A_n \uparrow A \quad \Longrightarrow \quad A \in \mathcal{M}$
- $A_n \downarrow A \quad \Longrightarrow \quad A \in \mathcal{M}$

Remark 1.28 Clearly, any σ–algebra in X is a monotone class. Conversely, if a monotone class $\mathcal{M}$ in X is also an algebra, then $\mathcal{M}$ is a σ–algebra (**Exercise**).

Let us prove now the following result.

Theorem 1.29 (Halmos) *Let $\mathcal{A}$ be an algebra, and $\mathcal{M}$ be a monotone class in X including $\mathcal{A}$. Then, $\sigma(\mathcal{A}) \subset \mathcal{M}$.*

Proof. Let $\mathcal{M}_0$ be the minimal monotone class including $\mathcal{A}$[3]. We are going to show that $\mathcal{M}_0$ is an algebra, and this will prove the theorem in view of Remark 1.28. To begin with, we note that $\varnothing$ and X belong to $\mathcal{M}_0$.

Now, define, for any $A \in \mathcal{M}_0$,

$$\mathcal{M}_A = \Big\{ B \in \mathcal{M}_0 \;\Big|\; A \cup B\,,\ A \setminus B\,,\ B \setminus A \in \mathcal{M}_0 \Big\}\,.$$

---

[3]**Exercise:** show that the intersection of all monotone classes including $\mathcal{A}$ is also a monotone class in X.

We claim that $\mathcal{M}_A$ is a monotone class. For let $(B_n) \subset \mathcal{M}_A$ be an increasing sequence such that $B_n \uparrow B$. Then,

$$A \cup B_n \uparrow A \cup B\,, \quad A \backslash B_n \downarrow A \backslash B\,, \quad B_n \setminus A \uparrow B \setminus A\,.$$

Since $\mathcal{M}_0$ is a monotone class, we conclude that

$$B\,,\ A \cup B\,,\ A \backslash B\,,\ B \setminus A \in \mathcal{M}_0\,.$$

Therefore, $B \in \mathcal{M}_A$. By a similar reasoning, one can check that

$$(B_n) \subset \mathcal{M}_A\,, \quad B_n \downarrow B \qquad \Longrightarrow \quad B \in \mathcal{M}_A\,.$$

So, $\mathcal{M}_A$ is a monotone class as claimed.

Next, let $A \in \mathcal{A}$. Then, $\mathcal{A} \subset \mathcal{M}_A$ since any $B \in \mathcal{A}$ belongs to $\mathcal{M}_0$ and satisfies

$$A \cup B\,,\ A \backslash B\,,\ B \setminus A \in \mathcal{M}_0\,. \tag{1.5}$$

But $\mathcal{M}_0$ is the minimal monotone class including $\mathcal{A}$, so $\mathcal{M}_0 \subset \mathcal{M}_A$. Therefore, $\mathcal{M}_0 = \mathcal{M}_A$ or, equivalently, (1.5) holds true for any $A \in \mathcal{A}$ and $B \in \mathcal{M}_0$.

Finally, let $A \in \mathcal{M}_0$. Since (1.5) is satisfied by any $B \in \mathcal{A}$, we deduce that $\mathcal{A} \subset \mathcal{M}_A$. Then, $\mathcal{M}_A = \mathcal{M}_0$. This implies that $\mathcal{M}_0$ is an algebra. □

Proof of Theorem 1.26: uniqueness. Let $\mathcal{E} = \sigma(A)$, and let μ_1, μ_2 be two measures extending μ to $\mathcal{E}$. We shall assume, first, that μ is finite, and set

$$\mathcal{M} = \Big\{ A \in \mathcal{E} \;\Big|\; \mu_1(A) = \mu_2(A) \Big\}\,.$$

We claim that $\mathcal{M}$ is a monotone class including $\mathcal{A}$. Indeed, for any sequence $(A_n) \subset \mathcal{M}$, using Propositions 1.15 and 1.16 we have that

$$A_n \uparrow A \quad \Longrightarrow \quad \mu_1(A) = \lim_n \mu_i(A_n) = \mu_2(A) \quad (i = 1, 2)$$

$$A_n \downarrow A,\ \mu_1(X), \mu_2(X) < \infty \quad \Longrightarrow \quad \mu_1(A) = \lim_n \mu_i(A_n) = \mu_2(A) \quad (i = 1, 2)$$

Therefore, in view of Halmos' Theorem, $\mathcal{M} = \mathcal{E}$, and this implies that $\mu_1 = \mu_2$.

In the general case of a σ–finite function μ, we have that $X = \bigcup_{k=1}^{\infty} X_k$ for some $(X_k) \subset \mathcal{A}$ such that $\mu(X_k) < \infty$ for all $k \in \mathbb{N}$. It is not restrictive to assume that the sequence (X_k) is increasing. Now, define $\mu_k(A) = \mu(A \cap X_k)$ for all $A \in \mathcal{A}$, and

$$\left. \begin{array}{l} \mu_{1,k}(A) = \mu_1(A \cap X_k) \\ \mu_{2,k}(A) = \mu_2(A \cap X_k) \end{array} \right\} \qquad \forall A \in \mathcal{E}.$$

Then, as easily checked, μ_k is a finite σ–additive function on $\mathcal{A}$, and $\mu_{1,k}, \mu_{2,k}$ are measures extending μ_k to $\mathcal{E}$. So, by the conclusion of first part of this proof, $\mu_{1,k}(A) = \mu_{2,k}(A)$ for all $A \in \mathcal{E}$ and any $k \in \mathbb{N}$. Therefore, since $A \cap X_k \uparrow A$, using again Proposition 1.15 we obtain

$$\mu_1(A) = \lim_k \mu_{1,k}(A) = \lim_k \mu_{2,k}(A) = \mu_2(A) \qquad \forall A \in \mathcal{E}\,.$$

The proof is thus complete. □

1.3.2 Outer measures

Definition 1.30 *A function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ is called an* outer measure *in X if $\mu^*(\varnothing) = 0$, and if μ^* is monotone and countably subadditive, i.e.,*

$$A \subset B \quad \Longrightarrow \quad \mu^*(A) \leq \mu^*(B)$$
$$\mu^*\Big(\bigcup_{i=1}^{\infty} E_i\Big) \leq \sum_{i=1}^{\infty} \mu^*(E_i) \qquad \forall (E_i) \subset \mathcal{P}(X)$$

The following proposition studies an example of outer measure that will be essential for the proof of Theorem 1.26.

Proposition 1.31 *Let μ be a σ–additive function on an algebra $\mathcal{A}$. Define, for any $E \in \mathcal{P}(X)$,*

$$\mu^*(E) = \inf\Big\{ \sum_{i=1}^{\infty} \mu(A_i) \;\Big|\; A_i \in \mathcal{A},\ E \subset \bigcup_{i=1}^{\infty} A_i \Big\}\,. \tag{1.6}$$

Then,

1. *μ^* is finite whenever μ is finite;*
2. *μ^* is an extension of μ, that is,*

$$\mu^*(A) = \mu(A), \quad \forall\, A \in \mathcal{A}. \tag{1.7}$$

3. *μ^* is an outer measure.*

Proof. The first assertion being obvious, let us proceed to check (5.11). Observe that the iniequality $\mu^*(A) \le \mu(A)$ for any $A \in \mathcal{A}$ is trivial. To prove the converse inequality, let $A_i \in \mathcal{A}$ be a countable covering of a set $A \in \mathcal{A}$. Then, $A_i \cap A \in \mathcal{A}$ is also a countable covering of A satisfying $\cup_i (A_i \cap A) \in \mathcal{A}$. Since μ is countably subadditive (see point 4 in Remark 1.12),

$$\mu(A) \le \sum_{i=1}^{\infty} \mu(A_i \cap A) \le \sum_{i=1}^{\infty} \mu(A_i)\,.$$

Thus, taking the infimum as in (1.6), we conclude that $\mu^*(A) \ge \mu(A)$.

Finally, we show that μ^* is countably subadditive. Let $(E_i) \subset \mathcal{P}(X)$, and set $E = \bigcup_{i=1}^{\infty} E_i$. Assume, without loss of generality, that all $\mu(E_i)$'s are finite (otherwise the assertion is trivial). Then, for any $i \in \mathbb{N}$ and any $\varepsilon > 0$ there exists $(A_{i,j}) \subset \mathcal{A}$ such that

$$\sum_{j=1}^{\infty} \mu(A_{i,j}) < \mu^*(E_i) + \frac{\varepsilon}{2^i}\,, \quad E_i \subset \bigcup_{j=1}^{\infty} A_{i,j}\,.$$

Consequently,

$$\sum_{i,j=1}^{\infty} \mu(A_{i,j}) \le \sum_{i}^{\infty} \mu^*(E_i) + \varepsilon.$$

Since $E \subset \bigcup_{i,j=1}^{\infty} A_{i,j}$ we have that $\mu^*(E) \le \sum_{i=1}^{\infty} \mu^*(E_i) + \varepsilon$ for any $\varepsilon > 0$. The conclusion follows. □

Exercise 1.32 1. Let μ^* be an outer measure in X, and $A \in \mathcal{P}(X)$. Show that

$$\nu^*(B) = \mu^*(A \cap B) \qquad \forall B \in \mathcal{P}(X)$$

is an outer measure in X.

2. Let μ_n^* be outer measures in X for all $n \in \mathbb{N}$. Show that

$$\mu^*(A) = \sum_n \mu_n^*(A) \quad \text{and} \quad \mu_\infty^*(A) = \sup_n \mu^*(A) \qquad \forall A \in \mathcal{P}(X)$$

are outer measures in X.

Given an outer measure μ^* in X, we now proceed to define additive sets.

Definition 1.33 *A subset $A \in \mathcal{P}(X)$ is called* additive *(or μ^*–*measurable*) if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \forall E \in \mathcal{P}(X). \tag{1.8}$$

We denote by $\mathcal{G}$ the family of all additive sets.

Notice that, since μ^* is countable subadditive, (1.8) is equivalent to

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \forall E \in \mathcal{P}(X). \tag{1.9}$$

Also, observe that $A^c \in \mathcal{G}$ for all $A \in \mathcal{G}$. Other important properties of $\mathcal{G}$ are listed in the next proposition.

Theorem 1.34 (Caratheodory) *Let μ^* be an outer measure in X. Then, $\mathcal{G}$ is a σ–algebra, and μ^* is σ–additive on $\mathcal{G}$.*

Before proving Caratheodory's Theorem, let us use it to complete the proof of Theorem 1.26.

Proof of Theorem 1.26: existence. Given a σ–additive function μ on an algebra $\mathcal{A}$, define the outer measure μ^* as in Example 1.31. Then, as noted above, $\mu^*(A) = \mu(A)$ for any $A \in \mathcal{A}$. Moreover, in light of Theorem 1.34, μ^* is a measure on the σ–algebra $\mathcal{G}$ of additive sets. So, the proof will be complete if we show that $\mathcal{A} \subset \mathcal{G}$. Indeed, in this case, $\sigma(\mathcal{A})$ turns out to be included in $\mathcal{G}$, and it suffices to take the restriction of μ^* to $\sigma(\mathcal{A})$ to obtain the required extension.

Now, let $A \in \mathcal{A}$ and $E \in \mathcal{P}(X)$. Assume $\mu^*(E) < \infty$ (otherwise (1.9) trivially holds), and fix $\varepsilon > 0$. Then, there exists $(A_i) \subset \mathcal{A}$ such that $E \subset \cup_i A_i$, and

$$\begin{aligned}
\mu^*(E) + \varepsilon &> \sum_{i=1}^{\infty} \mu(A_i) \\
&= \sum_{i=1}^{\infty} \mu(A_i \cap A) + \sum_{i=1}^{\infty} \mu(A_i \cap A^c) \\
&\geq \mu^*(E \cap A) + \mu^*(E \cap A^c).
\end{aligned}$$

Since ε is arbitrary, $\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$. Thus $A \in \mathcal{G}$. □

We now proceed with the proof of Caratheodory's Theorem.

Proof of Theorem 1.34. We will split the reasoning into four steps.

1. $\boxed{\mathcal{G} \text{ is an algebra}}$ We note that $\varnothing$ and X belong to $\mathcal{G}$. We already know that $A \in \mathcal{G}$ implies $A^c \in \mathcal{G}$. Let us now prove that, if $A, B \in \mathcal{G}$, then $A \cup B \in \mathcal{G}$. For any $E \in \mathcal{P}(X)$, we have

$$\begin{aligned} \mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c) \\ &= \mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \qquad (1.10) \\ &= [\mu^*(E \cap A) + \mu^*(E \cap A^c \cap B)] + \mu^*(E \cap (A \cup B)^c). \end{aligned}$$

Since

$$(E \cap A) \cup (E \cap A^c \cap B) = E \cap (A \cup B),$$

the subadditivity of μ^* implies that

$$\mu^*(E \cap A) + \mu^*(E \cap A^c \cap B) \geq \mu^*(E \cap (A \cup B)).$$

So, by (1.10),

$$\mu^*(E) \geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c),$$

and $A \cup B \in \mathcal{G}$ as required.

2. $\boxed{\mu^* \text{ is additive on } \mathcal{G}}$ Let us prove that, if $A, B \in \mathcal{G}$ and $A \cap B = \varnothing$, then

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap A) + \mu^*(E \cap B). \qquad (1.11)$$

Indeed, replacing E with $E \cap (A \cup B)$ in (1.8), yields

$$\mu^*(E \cap (A \cup B)) = \mu^*(E \cap (A \cup B) \cap A) + \mu^*(E \cap (A \cup B) \cap A^c),$$

which is equivalent to (1.11) since $A \cap B = \varnothing$. In particular, taking $E = X$, it follows that μ^* is additive on $\mathcal{G}$.

3. $\boxed{\mathcal{G} \text{ is a } \sigma\text{–algebra}}$ Let $(A_n) \subset \mathcal{G}$. We will show that $S := \bigcup_{i=1}^{\infty} A_i \in \mathcal{G}$. Since $\mathcal{G}$ is an algebra, it is not restrictive to assume that all the sets in (A_n) are mutually disjoint. Set $S_n := \bigcup_{i=1}^{n} A_i,\ n \in \mathbb{N}$. For any $n \in \mathbb{N}$ we have, by the subadditivity of μ^*,

$$\begin{aligned} \mu^*(E \cap S) + \mu^*(E \cap S^c) &\leq \sum_{i=1}^{\infty} \mu^*(E \cap A_i) + \mu^*(E \cap S^c) \\ &= \lim_{n \to \infty} \Big[\sum_{i=1}^{n} \mu^*(E \cap A_i) + \mu^*(E \cap S^c) \Big] \\ &= \lim_{n \to \infty} \Big[\mu^*(E \cap S_n) + \mu^*(E \cap S^c) \Big] \end{aligned}$$

in view of (1.11). Since $S^c \subset S_n^c$, it follows that

$$\mu^*(E \cap S) + \mu^*(E \cap S^c) \leq \limsup_{n\to\infty} \left[\mu^*(E \cap S_n) + \mu^*(E \cap S_n^c)\right] = \mu^*(E).$$

So, $S \in \mathcal{G}$, and $\mathcal{G}$ is a σ–algebra.

4. $\boxed{\mu^* \text{ is } \sigma\text{–additive on } \mathcal{G}}$ Since μ^* is countably sub–additive, and additive by Step 2, then point 5 in Remark 1.12 gives the conclusion.

The proof is now complete. □

Remark 1.35 The σ–algebra $\mathcal{G}$ of additive sets is *complete*, that is, it contains all the sets with outer measure 0. Indeed, for any $M \subset X$ with $\mu^*(M) = 0$, and any $E \in \mathcal{P}(X)$, we have

$$\mu^*(E \cap M) + \mu^*(E \cap M^c) = \mu^*(E \cap M^c) \leq \mu^*(E).$$

Thus, $M \in \mathcal{G}$.

Remark 1.36 In our proof of Theorem 1.26 we have constructed the σ–algebra $\mathcal{G}$ of additive sets such that

$$\sigma(\mathcal{A}) \subset \mathcal{G} \subset \mathcal{P}(X)\,. \tag{1.12}$$

We shall see later on that the above inclusions are both strict, in general.

1.4 Borel measures in $\mathbb{R}^N$

Let (X, d) be a metric space. We recall that $\mathcal{B}(X)$ denotes the Borel σ–algebra in X.

Definition 1.37 *A measure μ on the measurable space $(X, \mathcal{B}(X))$ is called a* Borel measure. *A Borel measure μ is called a* Radon measure *if $\mu(K) < \infty$ for every compact set $K \subset X$.*

In this section we will study specific properties of Borel measures on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. We begin by introducing Lebesgue measure on the unit interval.

1.4.1 Lebesgue measure in $[0,1)$

Let $\mathcal{I}$ be the class of all semi–closed intervals $[a,b) \subset [0,1)$, and $\mathcal{A}_0$ be the algebra of all finite disjoint unions of elements of $\mathcal{I}$ (see Example 1.5.2). Then, $\sigma(\mathcal{I}) = \sigma(\mathcal{A}_0) = \mathcal{B}([0,1))$.

On $\mathcal{I}$, consider the additive set function

$$\lambda([a,b)) = b - a, \quad 0 \le a \le b \le 1. \tag{1.13}$$

If $a = b$ then $[a,b)$ reduces to the empty set, and we have $\lambda([a,b)) = 0$.

Exercise 1.38 Let $[a,b) \in \mathcal{I}$ be contained in $[a_1,b_1) \cup \cdots \cup [a_n,b_n)$, with $[a_i,b_i) \in \mathcal{I}$. Prove that

$$b - a \le \sum_{i=1}^{n} (b_i - a_i)\,. \qquad \square$$

Proposition 1.39 *The set function λ defined in* (1.13) *is σ–additive on $\mathcal{I}$.*

Proof. Let (I_i) be a disjoint sequence of sets in $\mathcal{I}$, with $I_i = [a_i,b_i)$, and suppose $I = [a_0,b_0) = \cup_i I_i \in \mathcal{I}$. Then, for any $n \in \mathbb{N}$, we have

$$\sum_{i=1}^{n} \lambda(I_i) = \sum_{i=1}^{n} (b_i - a_i) \le b_n - a_1 \le b_0 - a_0 = \lambda(I)\,.$$

Therefore,

$$\sum_{i=1}^{\infty} \lambda(I_i) \le \lambda(I).$$

To prove the converse inequality, suppose $a_0 < b_0$. For any $\varepsilon < b_0 - a_0$ and $\delta > 0$, we have [4]

$$[a_0, b_0 - \varepsilon] \subset \bigcup_{i=1}^{\infty} \big((a_i - \delta 2^{-i}) \vee 0, b_i\big)\ .$$

Then, the Heine–Borel Theorem implies that, for some $i_0 \in \mathbb{N}$,

$$[a_0, b_0 - \varepsilon) \subset [a_0, b_0 - \varepsilon] \subset \bigcup_{i=1}^{i_0} \big((a_i - \delta 2^{-i}) \vee 0, b_i\big)\ .$$

[4] If $a, b \in \mathbb{R}$ we set $\min\{a,b\} = a \wedge b$ and $\max\{a,b\} = a \vee b$.

Consequently, in view of Exercise 1.38,

$$\lambda(I) - \varepsilon = (b_0 - a_0) - \varepsilon \le \sum_{i=1}^{i_0} \left(b_i - a_i + \delta 2^{-i}\right) \le \sum_{i=1}^{\infty} \lambda(I_i) + \delta.$$

Since ε and δ are arbitrary, we obtain

$$\lambda(I) \le \sum_{i=1}^{\infty} \lambda(I_i). \qquad \square$$

We now proceed to extend λ to $\mathcal{A}_0$. For any set $A \in \mathcal{A}_0$ such that $A = \cup_i I_i$, where $I_1, \ldots, I_n$ are disjoint sets in $\mathcal{I}$, let us define

$$\lambda(A) := \sum_{i=1}^{n} \lambda(I_i)\,.$$

It is easy to see that the above definition is independent of the representation of A as a finite disjoint union of elements of $\mathcal{I}$.

Exercise 1.40 Show that, if $J_1, \ldots, J_j$ is another family of disjoint sets in $\mathcal{I}$ such that $A = \cup_j J_j$, then

$$\sum_{i=1}^{n} \lambda(I_i) = \sum_{j=1}^{m} \lambda(J_j) \qquad \square$$

Theorem 1.41 *λ is σ–additive on $\mathcal{A}_0$.*

Proof. Let $(A_n) \subset \mathcal{A}_0$ be a sequence of disjoint sets in $\mathcal{A}$ such that

$$A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}_0\,.$$

Then

$$A = \bigcup_{i=1}^{k} I_i \qquad A_n = \bigcup_{j=1}^{k_n} I_{n,j} \qquad (\forall n \in \mathbb{N})$$

for some disjoint families $I_1, \ldots, I_n$ and $I_{n,1}, \ldots, I_{n,k_n}$ in $\mathcal{I}$. Now, observe that, for any $i \in \mathbb{N}$,

$$I_i = I_i \cap A = \bigcup_{n=1}^{\infty} (I_i \cap A_n) = \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{k_n} (I_i \cap I_{n,j})$$

(with $I_i \cap I_{n,j} \in \mathcal{I}$), and apply Proposition 1.39 to obtain

$$\lambda(I_i) = \sum_{n=1}^{\infty} \sum_{j=1}^{k_n} \lambda(I_i \cap I_{n,j}) = \sum_{n=1}^{\infty} \lambda(I_i \cap A_n).$$

Hence,

$$\lambda(A) = \sum_{i=1}^{k} \lambda(I_i) = \sum_{i=1}^{k} \sum_{n=1}^{\infty} \lambda(A_n \cap I_i) = \sum_{n=1}^{\infty} \sum_{i=1}^{k} \lambda(A_n \cap I_i) = \sum_{n=1}^{\infty} \lambda(A_n) \quad \square$$

Summing up, thanks to Theorem 1.26, we conclude that λ can be uniquely extended to a measure on the σ–algebra $\mathcal{B}([0,1))$. Such an extension is called *Lebesgue measure.*

1.4.2 Lebesgue measure in $\mathbb{R}$

We now turn to the construction of Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Usually, this is done by an intrinsic procedure, applying an extension result for σ–additive set functions on *semirings*. In these notes, we will follow a shortcut, based on the following simple observations.

Proceeding as in the above section, one can define Lebesgue measure on $\mathcal{B}([a,b))$ for any interval $[a,b) \subset \mathbb{R}$. Such a measure will be denoted by $\lambda_{[a,b)}$. Let us begin by characterizing the associated Borel sets as follows.

Proposition 1.42 *A set A belongs to $\mathcal{B}([a,b))$ if and only if $A = B \cap [a,b)$ for some $B \in \mathcal{B}(\mathbb{R})$.*

Proof. Consider the class $\mathcal{E} := \{A \in \mathcal{P}([a,b)) \mid \exists B \in \mathcal{B}(\mathbb{R}) \ : \ A = B \cap [a,b)\}$. Let us check that $\mathcal{E}$ is a σ–algebra in $[a,b)$.

1. By the definition of $\mathcal{E}$ we have that $\varnothing, [a,b) \in \mathcal{E}$.

2. Let $A \in \mathcal{E}$ and $B \in \mathcal{B}(\mathbb{R})$ be such that $A = B \cap [a,b)$. Then, $[a,b) \setminus B \in \mathcal{B}(\mathbb{R})$. So, $[a,b) \setminus A = [a,b) \cap ([a,b) \setminus B) \in \mathcal{E}$.

3. Let $(A_n) \subset \mathcal{E}$ and $(B_n) \subset \mathcal{B}(\mathbb{R})$ be such that $A_n = B_n \cap [a,b)$. Then, $\cup_n B_n \in \mathcal{B}(\mathbb{R})$. So, $\cup_n A_n = (\cup_n B_n) \cap [a,b) \in \mathcal{E}$.

Since $\mathcal{E}$ contains all open subsets of $[a,b)$, we conclude that $\mathcal{B}([a,b)) \subset \mathcal{E}$. This proves the 'only if' part of the conclusion.

Next, to prove the 'if' part, let $\mathcal{F} := \{B \in \mathcal{P}(\mathbb{R}) \mid B \cap [a,b) \in \mathcal{B}([a,b))\}$. Then, arguing as in the first part of the proof, we have that $\mathcal{F}$ is a σ-algebra in $\mathbb{R}$.

1. $\varnothing, \mathbb{R} \in \mathcal{F}$ by definition.

2. Let $B \in \mathcal{F}$. Since $B \cap [a,b) \in \mathcal{B}([a,b))$, we have that $B^c \cap [a,b) = [a,b) \setminus (B \cap [a,b)) \in \mathcal{B}([a,b))$. So, $B^c \in \mathcal{F}$.

3. Let $(B_n) \subset \mathcal{F}$. Then, $(\cup_n B_n) \cap [a,b) = \cup_n (B_n \cap [a,b)) \in \mathcal{B}([a,b))$. So, $\cup_n B_n \in \mathcal{F}$.

Since $\mathcal{F}$ contains all open subsets of $\mathbb{R}$, we conclude that $\mathcal{B}(\mathbb{R}) \subset \mathcal{F}$. The proof is thus complete. □

Thus, for any pair of nested intervals $[a,b) \subset [c,d) \subset \mathbb{R}$, we have that $\mathcal{B}([a,b)) \subset \mathcal{B}([c,d))$. Moreover, a unique extension argument yields

$$\lambda_{[a,b)}(E) = \lambda_{[c,d)}(E) \qquad \forall E \in \mathcal{B}([a,b))\,. \tag{1.14}$$

Now, since $\mathbb{R} = \bigcup_{k=1}^{\infty} [-k,k)$, it is natural to define Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ as

$$\lambda(E) = \lim_{k\to\infty} \lambda_{[-k,k)}(E \cap [-k,k)) \qquad \forall E \in \mathcal{B}(\mathbb{R})\,. \tag{1.15}$$

Our next exercise is intended to show that the definition of λ would be the same taking any other sequence of intervals invading $\mathbb{R}$.

Exercise 1.43 Let (a_n) and (b_n) be real sequences satisfying

$$a_k < b_k\,, \qquad a_k \downarrow -\infty\,, \qquad b_k \uparrow \infty\,.$$

Show that

$$\lambda(E) = \lim_{k\to\infty} \lambda_{[a_k,b_k)}(E \cap [a_k,b_k)) \qquad \forall E \in \mathcal{B}(\mathbb{R})\,. \qquad \square$$

In order to show that λ is a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, we still have to check σ–additivity.

Proposition 1.44 *The set function defined in* (1.15) *is* σ*–additive in* $\mathcal{B}(\mathbb{R})$*.*

Proof. Let $(E_n) \subset \mathcal{B}(\mathbb{R})$ be a sequence of disjoint Borel sets satisfying $E := \cup_n E_n \in \mathcal{B}(\mathbb{R})$. Then, by the σ–additivity of $\lambda_{[-k,k)}$,

$$\lambda(E) = \lim_{k\to\infty} \lambda_{[-k,k)}(E \cap [-k,k)) = \lim_{k\to\infty} \sum_{n=1}^{\infty} \lambda_{[-k,k)}(E_n \cap [-k,k)).$$

Now, observe that, owing to (1.14),

$$\begin{aligned}\lambda_{[-k,k)}(E_n \cap [-k,k)) &= \lambda_{[-k-1,k+1)}(E_n \cap [-k,k)) \\ &\leq \lambda_{[-k-1,k+1)}(E_n \cap [-k-1,k+1))\end{aligned}$$

So, for any $n \in \mathbb{N}$, $k \mapsto \lambda_{[-k,k)}(E_n \cap [-k,k))$ is nondecreasing. The conclusion follows applying Lemma 1.45 below. □

Lemma 1.45 *Let* $(a_{nk})_{n,k\in\mathbb{N}}$ *be a sequence in* $[0,\infty]$ *such that, for any* $n \in \mathbb{N}$*,*

$$h \leq k \quad \Longrightarrow \quad a_{nh} \leq a_{nk}\,. \tag{1.16}$$

Set, for any $n \in \mathbb{N}$*,*

$$\lim_{k\to\infty} a_{nk} =: \alpha_n \in [0,\infty]\,. \tag{1.17}$$

Then,

$$\lim_{k\to\infty} \sum_{n=1}^{\infty} a_{nk} = \sum_{n=1}^{\infty} \alpha_n$$

Proof. Suppose, first, $\sum_n \alpha_n < \infty$, and fix $\varepsilon > 0$. Then, there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\sum_{n=n_\varepsilon+1}^{\infty} \alpha_n < \varepsilon\,.$$

Recalling (1.17), for k sufficiently large, say $k \geq k_\varepsilon$, we have $\alpha_n - \frac{\varepsilon}{n_\varepsilon} < a_{nk}$ for $n = 1, \ldots, n_\varepsilon$. Therefore,

$$\sum_{n=1}^{\infty} a_{nk} \geq \sum_{n=1}^{n_\varepsilon} \alpha_n - \varepsilon > \sum_{n=1}^{\infty} \alpha_n - 2\varepsilon$$

for any $k \geq k_\varepsilon$. Since $\sum_n a_{nk} \leq \sum_n \alpha_n$, the conclusion follows.

The analysis of the case $\sum_n \alpha_n = \infty$ is similar. Fix $M > 0$, and let $n_M \in \mathbb{N}$ be such that

$$\sum_{n=1}^{n_M} \alpha_n > 2M\,.$$

For k sufficiently large, say $k \geq k_M$, $\alpha_n - \frac{M}{n_M} < a_{nk}$ for $n = 1, \ldots, n_M$. Therefore, for all $k \geq k_M$,

$$\sum_{n=1}^{\infty} a_{nk} \geq \sum_{n=1}^{n_M} a_{nk} > \sum_{n=1}^{n_M} \alpha_n - M > M\,. \qquad \square$$

Example 1.46 The monotonicity assumption of the above lemma is essential. Indeed, (1.16) fails for the sequence

$$a_{nk} = \delta_{nk} = \begin{cases} 1 & \text{if } n = k \\ 0 & \text{if } n \neq k \end{cases} \qquad \text{[Kroneker delta]}$$

since

$$\lim_{k\to\infty} \sum_{n=1}^{\infty} a_{nk} = 1 \neq 0 = \sum_{n=1}^{\infty} \lim_{k\to\infty} a_{nk} \qquad \square$$

Since λ is bounded on bounded sets, Lebesgue measure is a Radon measure. Another interesting property of Lebesgue measure is *translation invariance*.

Proposition 1.47 *Let $A \in \mathcal{B}(\mathbb{R})$. Then, for every $x \in \mathbb{R}$,*

$$A + x : = \{a + x \mid a \in A\} \in \mathcal{B}(\mathbb{R}) \tag{1.18}$$

$$\lambda(A + x) = \lambda(A)\,. \tag{1.19}$$

Proof. Define, for any $x \in \mathbb{R}$,

$$\mathcal{E}_x = \{A \in \mathcal{P}(\mathbb{R}) \mid A + x \in \mathcal{B}(\mathbb{R})\}\,.$$

Let us check that $\mathcal{E}_x$ is a σ–algebra in $\mathbb{R}$.

1. $\varnothing, \mathbb{R} \in \mathcal{E}_x$ by direct inspection.

2. Let $A \in \mathcal{E}_x$. Since $A^c + x = (A + x)^c \in \mathcal{B}(\mathbb{R})$, we conclude that $A^c \in \mathcal{E}_x$.

3. Let $(A_n) \subset \mathcal{E}_x$. Then, $(\cup_n A_n) + x = \cup_n (A_n + x) \in \mathcal{B}(\mathbb{R})$. So, $\cup_n A_n \in \mathcal{E}_x$.

Since $\mathcal{E}_x$ contains all open subsets of $\mathbb{R}$, $\mathcal{B}(\mathbb{R}) \subset \mathcal{E}_x$ for any $x \in \mathbb{R}$. This proves (1.18).

Let us prove (1.19). Fix $x \in \mathbb{R}$, and define

$$\lambda_x(A) = \lambda(A + x) \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

It is straightforward to check that λ_x and λ agree on the class

$$\mathcal{I}_\mathbb{R} := \big\{(-\infty, a) \mid -\infty < a \leq \infty\big\} \bigcup \big\{[a, b) \mid -\infty < a \leq b \leq \infty\big\}.$$

So, λ_x and λ also agree on the algebra $\mathcal{A}_\mathbb{R}$ of all finite disjoint unions of elements of $\mathcal{I}_\mathbb{R}$. By the uniqueness result of Theorem 1.26, we conclude that $\lambda_x(A) = \lambda(A)$ for all $A \in \mathcal{B}(\mathbb{R})$. □

1.4.3 Examples

In this section we shall construct three examples of sets that are hard to visualize but possess very interesting properties.

Example 1.48 (Two unusual Borel sets) Let $\{r_n\}$ be an enumeration of $\mathbb{Q} \cap [0, 1]$, and fix $\varepsilon > 0$. Set

$$A = \bigcup_{n=1}^{\infty} \Big(r_n - \frac{\varepsilon}{2^n}, r_n + \frac{\varepsilon}{2^n}\Big).$$

Then, $A \cap [0, 1]$ is an *open* (with respect to the relative topology) dense Borel set. By subadditivity, $0 < \lambda(A \cap [0, 1]) < 2\varepsilon$. Moreover, the *compact* set $B := [0, 1] \setminus A$ has no interior and measure nearly 1. □

Example 1.49 (Cantor triadic set) To begin with, let us note that any $x \in [0, 1]$ has a triadic expansion of the form

$$x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \qquad a_i = 0, 1, 2. \tag{1.20}$$

Such a representation is not unique due to the presence of periodic expansions. We can, however, choose a unique representation of the form (1.20) picking the expansion with less digits equal to 1. Now, observe that the set

$$C_1 := \Big\{x \in [0, 1] \;\Big|\; x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_1 \neq 1\Big\}$$

is obtained from $[0,1]$ removing the *'middle third'* $(\frac{1}{3},\frac{2}{3})$. It is, therefore, the union of 2 closed intervals, each of which has measure $\frac{1}{3}$. More generally, for any $n \in \mathbb{N}$,

$$C_n := \Big\{ x \in [0,1] \;\Big|\; x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_1, \ldots, a_n \neq 1 \Big\}$$

is the union of 2^n closed intervals, each of which has measure $\left(\frac{1}{3}\right)^n$. So,

$$C_n \downarrow C := \Big\{ x \in [0,1] \;\Big|\; x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_i \neq 1 \; \forall i \in \mathbb{N} \Big\},$$

where C is the so-called *Cantor set.* It is a closed set by construction, with measure 0 since

$$\lambda(C) \leq \lambda(C_n) \leq \left(\frac{2}{3}\right)^n \qquad \forall n \in \mathbb{N}\,.$$

Nevertheless, C is *uncountable.* Indeed, the function

$$f\Big(\sum_{i=1}^{\infty} \frac{a_i}{3^i} \Big) = \sum_{i=1}^{\infty} a_i 2^{-(i+1)} \tag{1.21}$$

maps C *onto* $[0,1]$. $\square$

Exercise 1.50 Show that f in (1.21) is onto.

Remark 1.51 Observe that $\mathcal{B}(\mathbb{R})$ has the cardinality of $\mathcal{P}(\mathbb{Q})$. On the other hand, the σ–algebra $\mathcal{G}$ of Lebesgue measurable sets is complete. So, $\mathcal{P}(C) \subset \mathcal{G}$, where C is Cantor set. Since C is uncountable, $\mathcal{G}$ must have a higher cardinality than the σ–algebra of Borel sets. In other terms, $\mathcal{B}(\mathbb{R})$ is *strictly* included in $\mathcal{G}$.

Example 1.52 (A nonmeasurable set) We shall now show that $\mathcal{G}$ is also strictly included in $\mathcal{P}([0,1))$. For $x, y \in [0,1)$ define

$$x \oplus y = \begin{cases} x + y & \text{if} \quad x + y < 1 \\ x + y - 1 & \text{if} \quad x + y \geq 1 \end{cases}$$

Observe that, if $E \subset [0,1)$ is a measurable set, then $E \oplus x \subset [0,1)$ is also measurable, and $\lambda(E \oplus x) = \lambda(E)$ for any $x \in [0,1)$. Indeed,

$$E \oplus x = \Big((E + x) \cap [0,1) \Big) \bigcup \Big((E + x) \setminus [0,1) - 1 \Big)\,.$$

In $[0,1)$, define x and y to be equivalent if $x-y\in\mathbb{Q}$. By the Axiom of Choice, there exists a set $P\subset[0,1)$ such that P consists of exactly one representative point from each equivalence class. We claim that P provides the required example of a nonmeasurable set. Indeed, consider the countable family $(P_n)\subset\mathcal{P}([0,1))$, where $P_n=P\oplus r_n$ and (r_n) is an enumeration of $\mathbb{Q}\cap[0,1)$. Observe the following.

1. (P_n) is a disjoint family for if there exist $p,q\in P$ such that $p\oplus r_n=q\oplus r_m$ with $n\neq m$, then $p-q\in\mathbb{Q}$. So, $p=q$ and the fact that $p\oplus r_n=p\oplus r_m$ with $r_n,r_m\in[0,1)$ implies that $r_n=r_m$, a contradiction.
2. $\cup_n P_n=[0,1)$. Indeed, let $x\in[0,1)$. Since x is equivalent to some element of P, $x-p=r$ for some $p\in P$ and some $r\in\mathbb{Q}$ satisfying $|r|<1$. Now, if $r\geq 0$, then $r=r_n$ for some $n\in\mathbb{N}$ whence $x\in P_n$. On the other hand, for $r<0$, we have $1+r=r_n$ for some $n\in\mathbb{N}$. So, $x\in P_n$ once again.

Should P be measurable, it would follow that $\lambda([0,1))=\sum_n\lambda(P_n)$. But this is impossible: the right-hand side is either 0 or $+\infty$.

1.4.4 Regularity of Radon measures

In this section, we shall prove regularity properties of a Radon measure in $\mathbb{R}^N$. We begin by studying finite measures.

Proposition 1.53 *Let μ be a finite measure on $(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N))$. Then, for any $B\in\mathcal{B}(\mathbb{R}^N)$,*

$$\mu(B)=\sup\{\mu(C):\ C\subset B,\ \text{closed}\}=\inf\{\mu(A):\ A\supset B,\ \text{open}\}. \tag{1.22}$$

Proof. Let us set

$$\mathcal{K}=\{B\in\mathcal{B}(\mathbb{R}^N)\mid (1.22)\text{ holds}\}.$$

It is enough to show that $\mathcal{K}$ is a σ-algebra of parts of $\mathbb{R}^n$ including all open sets. Obviously, $\mathcal{K}$ contains $\mathbb{R}^N$ and $\varnothing$. Moreover, if $B\in\mathcal{K}$ then its complement B^c belongs to $\mathcal{K}$. Let us now prove that $(B_n)\subset\mathcal{K}\Rightarrow\bigcup_{n=1}^\infty B_n\in\mathcal{K}$. We are going to show that, for any $\varepsilon>0$, there is a closed set C and an open set A such that

$$C\subset\bigcup_{n=1}^\infty B_n\subset A\,,\qquad \mu(A\backslash C)\leq\varepsilon\,. \tag{1.23}$$

Since $B_n \in \mathcal{K}$ for any $n \in \mathbb{N}$, there is an open set A_n and a closed set C_n such that

$$C_n \subset B_n \subset A_n\,, \qquad \mu(A_n \backslash C_n) \leq \frac{\varepsilon}{2^{n+1}}\,.$$

Now, take $A = \bigcup_{n=1}^\infty A_n$ and $S = \bigcup_{n=1}^\infty C_n$ to obtain $S \subset \bigcup_{n=1}^\infty B_n \subset A$ and

$$\mu(A\backslash S) \leq \sum_{n=1}^\infty \mu(A_n - S) \leq \sum_{n=1}^\infty \mu(A_n - C_n) \leq \frac{\varepsilon}{2}\,.$$

However, A is open but S is not necessarily closed. To overcome this difficulty, let us approximate S by the sequence $S_n = \bigcup_{k=1}^n C_k$. For any $n \in \mathbb{N}$, S_n is obviously closed, $S_n \uparrow S$, and so $\mu(S_n) \uparrow \mu(S)$. Therefore, there exists $n_\varepsilon \in \mathbb{N}$ such that $\mu(S\backslash S_{n_\varepsilon}) < \frac{\varepsilon}{2}$. Now, $C := S_{n_\varepsilon}$ satisfies $C \subset \bigcup_{n=1}^\infty B_n \subset A$ and $\mu(A\backslash C) = \mu(A\backslash S) + \mu(S\backslash C) < \varepsilon$. Therefore, $\bigcup_{n=1}^\infty B_n \in \mathcal{K}$. We have thus proved that $\mathcal{K}$ is a σ–algebra.

It remains to show that $\mathcal{K}$ contains the open subsets of $\mathbb{R}^N$. For this, let A be open, and set

$$C_n = \left\{ x \in \mathbb{R}^N \;\middle|\; d_{A^c}(x) \geq \frac{1}{n} \right\},$$

where $d_{A^c}(x)$ is the distance of x from A^c. Since d_{A^c} is continuous, C_n is a closed subsets of A. Moreover, $C_n \uparrow A$. So, recalling that μ is finite, we conclude that $\mu(A\backslash C_n) \downarrow 0$. □

The following result is a straightforward consequence of Proposition 1.53.

Corollary 1.54 *Let μ and ν be finite measures on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ such that $\mu(C) = \nu(C)$ for any closed subset C of $\mathbb{R}^N$. Then, $\mu = \nu$.*

Finally, we will extend Proposition 1.53 to Radon measures.

Theorem 1.55 *Let μ be a Radon measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, and let B be a Borel set. Then,*

$$\mu(B) = \inf\{\mu(A) \mid A \supset B\,,\ A \text{ open}\} \tag{1.24}$$
$$\mu(B) = \sup\{\mu(K) \mid K \subset B\,,\ K \text{ compact}\} \tag{1.25}$$

Proof. Since (1.24) is trivial if $\mu(B) = \infty$, we shall assume that $\mu(B) < \infty$. For any $n \in \mathbb{N}$, denote by Q_n the cube $(-n, n)^N$, and consider the finite measure $\mu \llcorner Q_n$ [5]. Fix $\varepsilon > 0$ and apply Proposition 1.53 to conclude that, for any $n \in \mathbb{N}$, there exists an open set $A_n \supset B$ such that

$$(\mu \llcorner Q_n)(A_n \setminus B) < \frac{\varepsilon}{2^n}\,.$$

Now, consider the open set $A := \cup_n (A_n \cap Q_n) \supset B$. We have

$$\begin{aligned} \mu(A \setminus B) &\leq \sum_{n=1}^{\infty} \mu\big((A_n \cap Q_n) \setminus B\big) \\ &= \sum_{n=1}^{\infty} (\mu \llcorner Q_n)(A_n \setminus B) < 2\varepsilon \end{aligned}$$

which in turn implies (1.24).

Next, let us prove (1.25) for $\mu(B) < \infty$. Fix $\varepsilon > 0$, and apply Proposition 1.53 to $\mu \llcorner \overline{Q}_n$ to construct, for any $n \in \mathbb{N}$, a closed set $C_n \subset B$ satisfying

$$(\mu \llcorner \overline{Q}_n)(B \setminus C_n) < \varepsilon\,.$$

Consider the sequence of compact sets $K_n = C_n \cap \overline{Q}_n$. Since

$$\mu(B \cap \overline{Q}_n) \uparrow \mu(B)\,,$$

for some $n_\varepsilon \in \mathbb{N}$ we have that $\mu(B \cap \overline{Q}_{n_\varepsilon}) > \mu(B) - \varepsilon$. Therefore,

$$\begin{aligned} \mu(B \setminus K_{n_\varepsilon}) &= \mu(B) - \mu(K_{n_\varepsilon}) \\ &< \mu(B \cap \overline{Q}_{n_\varepsilon}) - \mu(C_{n_\varepsilon} \cap \overline{Q}_{n_\varepsilon}) + \varepsilon \\ &= (\mu \llcorner \overline{Q}_{n_\varepsilon})(B \setminus C_{n_\varepsilon}) + \varepsilon < 2\varepsilon \end{aligned}$$

If $\mu(B) = +\infty$, then, setting $B_n = B \cap Q_n$, we have $B_n \uparrow B$, and so $\mu(B_n) \to +\infty$. Since $\mu(B_n) < +\infty$, for every n there exists a compact set K_n such that $K_n \subset B_n$ and $\mu(K_n) > \mu(B_n) - 1$, by which $K_n \subset B$ and $\mu(K_n) \to +\infty = \mu(B)$. □

Exercise 1.56 A Radon measure μ on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ is obviously σ–finite. Conversely, is a σ–finite Borel measure on $\mathbb{R}^N$ necessarily Radon?

[5] that is, μ restricted to Q_n (see Definition 1.23).

HINT: consider $\mu = \sum_n \delta_{1/n}$ on $\mathcal{B}(\mathbb{R})$, where $\delta_{1/n}$ is the Dirac measure at $1/n$. To prove σ–additivity observe that, if (B_k) are disjoint Borel sets, then

$$\begin{aligned}\mu\Big(\bigcup_{k=1}^{\infty} B_k\Big) &= \sum_{n=1}^{\infty} \delta_{1/n}\Big(\bigcup_{k=1}^{\infty} B_k\Big) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \delta_{1/n}(B_k) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n}\sum_{k=1}^{\infty} \delta_{1/i}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{\infty}\sum_{i=1}^{n} \delta_{1/i}(B_k) \\ &= \sum_{k=1}^{\infty}\sum_{i=1}^{\infty} \delta_{1/i}(B_k) = \sum_{i=1}^{\infty} \mu(B_k)\end{aligned}$$

where we have used Lemma 1.45. □

In the subsections 1.4.1-1.4.2 we constructed the Lebesgue measure on $\mathbb{R}$, starting from an additive function defined on the algebra of the finite disjoint union of semi-closed intervals $[a, b)$. This construction can be carried out in $\mathbb{R}^N$ provided that we substitute the semi-closed intervals by the semiclosed rectangles of the type

$$\prod_{i=1}^{N}[a_i, b_i), \quad a_i \le b_i \quad i = 1, \dots, N$$

whose measure is given by

$$\lambda\Big(\prod_{i=1}^{N}[a_i, b_i)\Big) = \prod_{i=1}^{N}(b_i - a_i).$$

In what follows λ will denote the Lebesgue measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$. Then λ is clearly a Radon measure and, by the analogue of Proposition 1.47, λ is translation invariant. Next proposition characterizes all the Radon measures having the property of translation invariance.

Proposition 1.57 *Let μ be a Radon measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ such that μ is translation invariant, that is*

$$\mu(E + x) = \mu(E) \quad \forall E \in \mathcal{B}(\mathbb{R}^N), \ \ \forall x \in E.$$

Then there exists $c \ge 0$ such that $\mu(E) = c\lambda(E)$ for every $E \in \mathcal{B}(\mathbb{R}^N)$.

Proof. For every $n \in \mathbb{N}$ define the set

$$\Omega_n = \left\{ \prod_{k=1}^{N} \left[\frac{a_k}{2^n}, \frac{a_k+1}{2^n} \right), a_k \in \mathbb{Z} \right\},$$

that is Ω_n is the set of the semi-closed cubes with edge of length $\frac{1}{2^n}$ and with vertexes having coordinates multiple of $\frac{1}{2^n}$. The sets Ω_n have the following properties:

a) for every n $\mathbb{R}^N = \cup_{Q\in\Omega_n} Q$ with disjoint union;

b) if $Q \in \Omega_n$ and $P \in \Omega_r$ with $r \le n$, then $Q \subset P$ or $P \cap Q = \emptyset$;

c) If $Q \in \Omega_n$, then $\lambda(Q) = 2^{-nN}$.

Observe that $[0,1)^N$ is the union of 2^{nN} disjoint cubes of Ω_n, and these cubes are identical up to a translation. Setting $c = \mu([0,1)^N)$ and using the translation invariance of μ and λ, for every $Q \in \Omega_n$ we have

$$2^{nN}\mu(Q) = \mu([0,1)^N) = c\lambda([0,1)^N) = 2^{nN} c\lambda([0,1)^N)$$

Then μ and $c\lambda$ coincide on the cubes of the sets Ω_n. If A is an open nonempty set of $\mathbb{R}^N$, then by property a) we have $A = \cup_n \cup_{Q\in\Omega_n, Q\subset A} Q = \cup_n Z_n$ where $Z_n = \cup_{Q\in\Omega_n, Q\subset A} Q$. By property b) we deduce that if $Q \in \Omega_n$ and $Q \subset A$, then $Q \subset Z_{n-1}$ or $Q \cap Z_{n-1} = \emptyset$. Then A can be rewritten as

$$A = \bigcup_n \bigcup_{Q\in\Omega_n, Q\subset A\setminus Z_{n-1}} Q$$

and the above union is disjoint. Then the σ-additivity of μ and λ gives $\mu(A) = c\lambda(A)$; finally, by (1.24), $\mu(B) = c\lambda(B)$ for every $B \in \mathcal{B}(\mathbb{R}^N)$. □

Next theorem shows how the Lebesgue measure changes under the linear non-singular transformations.

Theorem 1.58 *Let $T : \mathbb{R}^N \to \mathbb{R}^N$ be a linear non-singular transformation. Then*

i) $T(E) \in \mathcal{B}(\mathbb{R}^N)$ *for every* $E \in \mathcal{B}(\mathbb{R}^N)$*;*

ii) $\lambda(T(E)) = |\det T|\, \lambda(E)$ *for every* $E \in \mathcal{B}(\mathbb{R}^N)$.

Proof. Consider the family

$$\mathcal{E} = \{E \in \mathcal{B}(\mathbb{R}^N) \,|\, T(E) \in \mathcal{B}(\mathbb{R}^N)\}.$$

Since T is non-singular, then $T(\emptyset) = \emptyset$, $T(\mathbb{R}^N) = \mathbb{R}^N$, $T(E^c) = (T(E))^c$, $T(\cup_n E_n) = \cup_n T(E_n)$ for all $E, E_n \subset \mathbb{R}^N$. Hence $\mathcal{E}$ is an σ-algebra. Furthermore T maps open sets into open sets; so $\mathcal{E} = \mathcal{B}(\mathbb{R}^N)$ and i) follows.

Next define

$$\mu(B) = \lambda(T(B)) \quad \forall E \in \mathcal{B}(\mathbb{R}^N).$$

Since T maps compact sets into compact sets, we deduce that μ is a Radon measure. Furthermore if $B \in \mathcal{B}(\mathbb{R}^N)$ and $x \in \mathbb{R}^N$, since λ is translation invariant, we have

$$\mu(B+x) = \lambda(T(B+x)) = \lambda(T(B)+T(x)) = \lambda(T(B)) = \mu(B),$$

and so μ is translation invariant too. Proposition 1.57 implies that there exists $\Delta(T) \geq 0$ such that

$$\mu(B) = \Delta(T)\lambda(B) \quad \forall B \in \mathcal{B}(\mathbb{R}^N). \tag{1.26}$$

It remains to show that $\Delta(T) = |\det T|$. To prove this, let $\{\mathbf{e}_1, \ldots, \mathbf{e}_N\}$ denote the standard basis in $\mathbb{R}^N$, i.e. $\mathbf{e}_i$ has the j-th coordinate equal to 1 if $j = i$ and equal to 0 if $j \neq i$. We begin with the case of the following elementary transformations:

a) there exists $i \neq j$ such that $T(\mathbf{e}_i) = \mathbf{e}_j$, $T(\mathbf{e}_j) = \mathbf{e}_i$ and $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i, j$.

   In this case $T([0,1)^N) = [0,1)^N$ and $\det T = 1$. By taking $B = [0,1)^N$ in (1.26), we deduce $\Delta(T) = 1 = |\det T|$.

b) there exist $\alpha \neq 0$ and i such that $T(\mathbf{e}_i) = \alpha \mathbf{e}_i$ and $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i$.

   Assume $i = 1$. Then $T([0,1)^N) = [0,\alpha) \times [0,1)^{N-1}$ if $\alpha > 0$ and $T([0,1)^N) = (\alpha, 0] \times [0,1)^{N-1}$ if $\alpha < 0$. Therefore by taking $B = [0,1)^N$ in (1.26), we obtain $\Delta(T) = \lambda(T([0,1)^N)) = |\alpha| = |\det T|$.

c) there exist $i \neq j$ and $\alpha \neq 0$ such that $T(\mathbf{e}_i) = \mathbf{e}_i + \alpha \mathbf{e}_j$, $T(\mathbf{e}_k) = \mathbf{e}_k$ for $k \neq i$.

Assume $i = 1$ and $j = 2$ and set $Q_\alpha = \{\alpha x_2 \mathbf{e}_2 + \sum_{i \neq 2} x_i \mathbf{e}_i \,|\, 0 \leq x_i < 1\}$. Then we have

$$\begin{aligned} T(Q_\alpha) &= \{x_1 \mathbf{e}_1 + \alpha(x_1 + x_2)\mathbf{e}_2 + \ldots + x_N \mathbf{e}_N) \,|\, 0 \leq x_i < 1\} \\ &= \left\{ \alpha\xi_2 \mathbf{e}_2 + \sum_{i \neq 2} \xi_i \mathbf{e}_i \,\middle|\, \xi_1 \leq \xi_2 < \xi_1 + 1,\, 0 \leq \xi_i < 1 \text{ for } i \neq 2 \right\} \\ &= E_1 \cup E_2 \end{aligned}$$

with disjoint union, where

$$\begin{aligned} E_1 &= \left\{ \alpha\xi_2 \mathbf{e}_2 + \sum_{i \neq 2} \xi_i \mathbf{e}_i \,\middle|\, \xi_1 \leq \xi_2 < 1,\, 0 \leq \xi_i < 1 \text{ for } i \neq 2 \right\}, \\ E_2 &= \left\{ \alpha\xi_2 \mathbf{e}_2 + \sum_{i \neq 2} \xi_i \mathbf{e}_i \,\middle|\, 1 \leq \xi_2 < \xi_1 + 1,\, 0 \leq \xi_i < 1 \text{ for } i \neq 2 \right\}. \end{aligned}$$

Observe that $E_1 \subset Q_\alpha$ and $E_2 - \mathbf{e}_2 = Q_\alpha \setminus E_1$; then

$$\lambda(T(Q_\alpha)) = \lambda(E_1) + \lambda(E_2) = \lambda(E_1) + \lambda\big(E_2 - \mathbf{e}_2\big) = \lambda(Q_\alpha).$$

By taking $B = Q_\alpha$ in (1.26) we deduce $\Delta(T) = 1 = |\det T|$.

If $T = T_1 \cdot \ldots \cdot T_k$ with T_i elementary transformations of type a)-b)-c), since (1.26) implies $\Delta(T) = \Delta(T_1) \cdot \ldots \cdot \Delta(T_k)$, then we have

$$\Delta(T) = |\det T_1| \cdot \ldots \cdot |\det T_k| = |\det T|.$$

Therefore the thesis will follow if we prove the following claim: any linear non-singular transformation T is the product of elementary transformations of type a)-b)-c). We proceed by induction on the dimension N. The claim is trivially true for $N = 1$; assume that the claim holds for $N - 1$. Set $T = (a_{i,j})_{i,j=1,\ldots,N}$, that is

$$T(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij} \mathbf{e}_j \quad i = 1, \ldots, N.$$

For $k = 1, \ldots, N$, consider $T_k = (a_{i,j})_{j=1,\ldots,N-1,\, i=1,\ldots,N,\, i \neq k}$. Since $\det T = \sum_{k=1}^{N} (-1)^{k+N} a_{kN} \det T_k$, possibly changing two variables by a transformation of type a), we may assume $\det T_N \neq 0$. Then by induction the following transformation $S_1 : \mathbb{R}^N \to \mathbb{R}^N$

$$S_1(\mathbf{e}_i) = T_N(\mathbf{e}_i) = \sum_{j=1}^{N-1} a_{ij}\mathbf{e}_j \;\; i = 1, \dots, N-1, \quad S_1(\mathbf{e}_N) = \mathbf{e}_N$$

is the product of elementary transformations. By applying transformations of type c) we add $a_{iN}S_1(\mathbf{e}_N)$ to $S_1(\mathbf{e}_i)$ for $i = 1, \dots, N-1$ and we arrive at $S_2 : \mathbb{R}^N \to \mathbb{R}^N$ defined by

$$S_2(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \;\; i = 1, \dots, N-1, \quad S_2(\mathbf{e}_N) = \mathbf{e}_N.$$

Next we compose S_2 with a transformation of type b) to obtain

$$S_3(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \;\; i = 1, \dots, N-1, \quad S_3(\mathbf{e}_N) = b\mathbf{e}_N$$

where b will be chosen later. Now set $T_N^{-1} = (m_{ki})_{k,i=1,\dots N-1}$. By applying again transformations of type c), for every $i = 1, \dots, N-1$ we multiply $S_3(\mathbf{e}_i)$ by $\sum_{k=1}^{N-1} a_{Nk}m_{ki}$ and add the results to $S_3(\mathbf{e}_N)$; then we obtain:

$$S_4(\mathbf{e}_i) = \sum_{j=1}^{N} a_{ij}\mathbf{e}_j \;\; i = 1, \dots, N-1, \quad S_4(\mathbf{e}_N) = b\mathbf{e}_N + \sum_{i,k=1}^{N-1} a_{Nk}m_{ki} \sum_{j=1}^{N} a_{ij}\mathbf{e}_j.$$

Since $\sum_{i,k=1}^{N-1} a_{Nk}m_{ki} \sum_{j=1}^{N-1} a_{ij}\mathbf{e}_j = \sum_{k=1}^{N-1} a_{Nk}\mathbf{e}_k$, by choosing $b = a_{NN} - \sum_{i,k=1}^{N-1} a_{Nk}m_{ki}a_{iN}$ we have that $T = S_4$. □

Remark 1.59 As a corollary of Theorem 1.58 we obtain that the Lebesgue measure is *rotation invariant.*

Chapter 2

Integration

2.1 Measurable functions

2.1.1 Inverse image of a function

Let X, Y be non empty sets. For any map $\varphi\colon X \to Y$ and any $A \in \mathcal{P}(Y)$ we set

$$\varphi^{-1}(A) := \{x \in X \mid \varphi(x) \in A\} = \{\varphi \in A\}.$$

$\varphi^{-1}(A)$ is called the *inverse image* of A.

Let us recall some elementary properties of φ^{-1}. The easy proofs are left to the reader as an exercise.

(i) $\varphi^{-1}(A^c) = (\varphi^{-1}(A))^c$ for all $A \in \mathcal{P}(Y)$.

(ii) If $A, B \in \mathcal{P}(Y)$, then $\varphi^{-1}(A \cap B) = \varphi^{-1}(A) \cap \varphi^{-1}(B)$. In particular, if $A \cap B = \varnothing$, then $\varphi^{-1}(A) \cap \varphi^{-1}(B) = \varnothing$.

(iii) If $\{A_k\} \subset \mathcal{P}(Y)$ we have

$$\varphi^{-1}\left(\bigcup_{k=1}^{\infty} A_k\right) = \bigcup_{k=1}^{\infty} \varphi^{-1}(A_k).$$

Consequently, if $(Y, \mathcal{F})$ is a measurable space, then the family of parts of X

$$\varphi^{-1}(\mathcal{F}) := \{\varphi^{-1}(A) : \ A \in \mathcal{F}\}$$

is a σ–algebra in X.

Exercise 2.1 Let $\varphi\colon X \to Y$ and let $A \in \mathcal{P}(X)$. Set

$$\varphi(A) := \{\varphi(x) \mid x \in A\}.$$

Show that properties like (i), (ii) fail, in general, for $\varphi(A)$.

2.1.2 Measurable maps and Borel functions

Let $(X, \mathcal{E})$ and $(Y, \mathcal{F})$ be measurable spaces.

Definition 2.2 *We say that a map $\varphi\colon X \to Y$ is* measurable *if $\varphi^{-1}(\mathcal{F}) \subset \mathcal{E}$. When Y is a metric space and $\mathcal{F} = \mathcal{B}(Y)$, we also call φ a* Borel map. *If, in addition, $Y = \mathbb{R}$, then we say that φ is a* Borel function.

Proposition 2.3 *Let $\mathcal{A} \subset \mathcal{F}$ be such that $\sigma(\mathcal{A}) = \mathcal{F}$. Then $\varphi\colon X \to Y$ is measurable if and only if $\varphi^{-1}(\mathcal{A}) \subset \mathcal{E}$.*

Proof. Clearly, if φ is measurable, then $\varphi^{-1}(\mathcal{A}) \subset \mathcal{E}$. Conversely, suppose $\varphi^{-1}(\mathcal{A}) \subset \mathcal{E}$, and consider the family

$$\mathcal{G} := \{B \in \mathcal{F} \mid \varphi^{-1}(B) \in \mathcal{E}\}.$$

Using properties (i), (ii), and (iii) of φ^{-1} from the previous section, one can easily show that $\mathcal{G}$ is a σ–algebra in Y including $\mathcal{A}$. So, $\mathcal{G}$ coincides with $\mathcal{F}$ and the proof is complete. □

Exercise 2.4 Show that a function $\varphi\colon X \to \mathbb{R}$ is Borel if any of the following conditions holds:

(i) $\varphi^{-1}((-\infty, t]) \subset \mathcal{E}$ for all $t \in \mathbb{R}$.

(ii) $\varphi^{-1}((-\infty, t)) \subset \mathcal{E}$ for all $t \in \mathbb{R}$.

(iii) $\varphi^{-1}([a, b]) \subset \mathcal{E}$ for all $a, b \in \mathbb{R}$.

(iv) $\varphi^{-1}([a, b)) \subset \mathcal{E}$ for all $a, b \in \mathbb{R}$.

(v) $\varphi^{-1}((a, b)) \subset \mathcal{E}$ for all $a, b \in \mathbb{R}$.

Exercise 2.5 Let $\varphi(X)$ be countable. Show that φ is measurable if, for any $y \in Y$, $\varphi^{-1}(y) \in \mathcal{E}$.

Proposition 2.6 *Let X, Y be metric spaces, $\mathcal{E} = \mathcal{B}(X)$, and $\mathcal{F} = \mathcal{B}(Y)$. Then, any continuous map $\varphi\colon X \to Y$ is measurable.*

Proof. Let $\mathcal{A}$ be the family of all open subsets of Y. Then, $\sigma(\mathcal{A}) = \mathcal{B}(Y)$ and $\varphi^{-1}(\mathcal{A}) \subset \mathcal{B}(X)$. So, the conclusion follows from Proposition 2.3. □

Proposition 2.7 *Let $\varphi\colon X \to Y$ be measurable, let $(Z, \mathcal{G})$ be a measurable space, and let $\psi : Y \to Z$ be another measurable map. Then $\psi \circ \varphi$ is measurable.*

Example 2.8 Let $(X, \mathcal{E})$ be a measurable space, and let $\varphi : X \to \mathbb{R}^N$. We regard $\mathbb{R}^N$ as a measurable space with the Borel σ–algebra $\mathcal{B}(\mathbb{R}^N)$. Denoting by φ_i the components of φ, that is, $\varphi = (\varphi_1, \dots, \varphi_N)$, let us show that

$$\varphi \quad \text{is Borel} \quad \Longleftrightarrow \quad \varphi_i \quad \text{is Borel} \quad \forall i \in \{1, \dots, N\}. \tag{2.1}$$

Indeed, for any $y \in \mathbb{R}^N$ let

$$A_y = \prod_{i=1}^{N} (-\infty, y_i] = \{z \in \mathbb{R}^N \mid z_i \le y_i \ \forall i\},$$

and define $\mathcal{A} = \{A_y \mid y \in \mathbb{R}^N\}$. Observe that $\mathcal{B}(\mathbb{R}^N) = \sigma(\mathcal{A})$ to deduce, from Proposition 2.3, that φ is measurable if and only if $\varphi^{-1}(\mathcal{A}) \subset \mathcal{E}$. Now, for any $y \in \mathbb{R}^N$,

$$\varphi^{-1}(A_y) = \bigcap_{i=1}^{N} \{x \in X \mid \varphi_i(x) \le y_i\} = \bigcap_{i=1}^{N} \varphi_i^{-1}((-\infty, y_i]).$$

This shows the $\Leftarrow$ part of (2.1). To complete the reasoning, assume that φ is Borel and let $i \in \{1, \dots, N\}$ be fixed. Then, for any $t \in \mathbb{R}$

$$\varphi_i^{-1}((-\infty, t]) = \varphi^{-1}(\{z \in \mathbb{R}^N \mid z_i \le t\})$$

which implies $\varphi_i^{-1}((-\infty, t]) \in \mathcal{E}$, and so φ_i is Borel. □

Exercise 2.9 Let $\varphi, \psi\colon X \to \mathbb{R}$ be Borel. Then $\varphi + \psi$, $\varphi\psi$, $\varphi \wedge \psi$, and $\varphi \vee \psi$ are Borel.
HINT: define $f(x) = (\varphi(x), \psi(x))$ and $g(y_1, y_2) = y_1 + y_2$. Then, f is a Borel map owing to Example 2.8, and g is a Borel function since it is continuous. Thus, $\varphi + \psi = g \circ f$ is also Borel. The remaining assertions can be proved similarly. □

Exercise 2.10 Let $\varphi\colon X \to \mathbb{R}$ be Borel. Prove that the function

$$\psi(x) = \begin{cases} \frac{1}{\varphi(x)} & \text{if} \quad \varphi(x) \neq 0 \\ \\ 0 & \text{if} \quad \varphi(x) = 0 \end{cases}$$

is also Borel.
HINT: show, by a direct argument, that $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is a Borel function. □

Proposition 2.11 *Let (φ_n) be a sequence of Borel functions in $(X, \mathcal{E})$ such that $|\varphi_n(x)| \leq M$ for all $x \in X$, all $n \in \mathbb{N}$, and some $M > 0$. Then, the functions*

$$\sup_{n \in \mathbb{N}} \varphi_n(x), \quad \inf_{n \in \mathbb{N}} \varphi_n(x), \quad \limsup_{n \to \infty} \varphi_n(x), \quad \liminf_{n \to \infty} \varphi_n(x),$$

are Borel.

Proof. Let us prove that $\phi(x) \colon = \sup_{n \in \mathbb{N}} \varphi_n(x)$ is Borel. Let $\mathcal{F}$ be the set of all intervals of the form $(-\infty, a]$ with $a \in \mathbb{R}$. Since $\sigma(\mathcal{F}) = \mathcal{B}(\mathbb{R})$, we have that ϕ is Borel. In fact

$$\phi^{-1}((-\infty, a]) = \bigcap_{n=1}^{\infty} \varphi_n^{-1}((-\infty, a]) \in \mathcal{E}.$$

In a similar way one can prove the other assertions. □

It is convenient to consider functions with values on the extended space $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$. These are called *extended functions*. We say that a mapping $\varphi\colon X \to \overline{\mathbb{R}}$ is Borel if

$$\varphi^{-1}(-\infty), \varphi^{-1}(\infty) \in \mathcal{E}$$

and $\varphi^{-1}(I) \in \mathcal{E}$ for all $I \in \mathcal{B}(\mathbb{R})$.

All previous results can be generalized, with obvious modifications, to extended Borel functions. In particular, the following result holds.

Proposition 2.12 *Let (φ_n) be a sequence of Borel functions on $(X, \mathcal{E})$. Then the following functions:*

$$\sup_{n\in\mathbb{N}} \varphi_n(x), \quad \inf_{n\in\mathbb{N}} \varphi_n(x), \quad \limsup_{n\to\infty} \varphi_n(x), \quad \liminf_{n\to\infty} \varphi_n(x),$$

are Borel.

Exercise 2.13 Let $\varphi, \psi \colon X \to \overline{\mathbb{R}}$ be Borel functions on $(X, \mathcal{E})$. Prove that $\{\varphi = \psi\} \in \mathcal{E}$.

Exercise 2.14 Let (φ_n) be a sequence of Borel functions in $(X, \mathcal{E})$. Show that $\{x \in X \mid \exists \lim_n \varphi_n(x)\} \in \mathcal{E}$.

Exercise 2.15 Let $\varphi \colon X \to \overline{\mathbb{R}}$ be a Borel function on $(X, \mathcal{E})$, and let $A \in \mathcal{E}$. Prove that

$$\varphi_A(x) = \begin{cases} \varphi(x) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

is Borel.

Exercise 2.16

1. Let X be a metric space and $\mathcal{E} = \mathcal{B}(X)$. Then, any lower semicontinuous map $\varphi : X \to \mathbb{R}$ is Borel.
2. Any monotone function $\varphi : \mathbb{R} \to \mathbb{R}$ is Borel.

Exercise 2.17 Let $\mathcal{E}$ be a σ–algebra in $\mathbb{R}$. Show that $\mathcal{E} \supset \mathcal{B}(\mathbb{R})$ if and only if any continuous function $\varphi : \mathbb{R} \to \mathbb{R}$ is $\mathcal{E}$-measurable, that is, $\varphi^{-1}(B) \in \mathcal{E}$ for any $B \in \mathcal{B}(\mathbb{R})$.

Exercise 2.18 Show that Borel functions on $\mathbb{R}$ are the smallest class of functions which includes all continuous functions and is stable under pointwise limits.

Definition 2.19 *A Borel function $\varphi \colon X \to \mathbb{R}$ is said to be* simple*, if its range $\varphi(X)$ is a finite set. The class of all simple functions $\varphi \colon X \to \mathbb{R}$ is denoted by $\mathcal{S}(X)$.*

It is immediate that the class $\mathcal{S}(X)$ is closed under sum, product, $\wedge$, $\vee$, and so on.

We recall that $\chi_A : X \to \mathbb{R}$ stands for the *characteristic function* of a set $A \subset X$, i.e.,

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then $\chi_A \in \mathcal{S}(X)$ if and only if $A \in \mathcal{E}$.

Remark 2.20 1. We note that $\varphi\colon X \to \mathbb{R}$ is simple if an only if there exist disjoint sets $A_1, \ldots, A_n \in \mathcal{E}$ and real numbers $a_1, \ldots, a_n$ such that

$$X = \bigcup_{i=1}^{n} A_i \qquad \text{and} \qquad \varphi(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x) \quad \forall x \in X\,. \tag{2.2}$$

Indeed, any function given by (2.2) is simple. Conversely, if φ is simple, then

$$\varphi(X) = \{a_1, \ldots, a_n\} \quad \text{with} \quad a_i \neq a_j \quad \text{for} \quad i \neq j\,.$$

So, taking $A_i := \varphi^{-1}(a_i)\,,\ i \in \{1, \ldots, n\}\,$, we obtain a representation of φ of type (2.2).

Obviously, the choice of sets $A_1, \ldots, A_n \in \mathcal{E}$ and real numbers $a_1, \ldots, a_n$ is far from being unique.

2. Given two simple functions φ and ψ, they can always be represented as linear combinations of the characteristic functions of the same family of sets. To see this, let φ be given by (2.2), and let

$$X = \bigcup_{h=1}^{m} B_h \qquad \text{and} \qquad \psi(x) = \sum_{h=1}^{m} b_h \chi_{B_h}(x), \quad \forall x \in X.$$

Since $A_i = \bigcup_{h=1}^{m} (A_i \cap B_h)\,$, we have that

$$\chi_{A_i} = \sum_{h=1}^{m} \chi_{A_i \cap B_h}(x) \qquad i \in \{1, \ldots, n\}\,.$$

So,

$$\varphi(x) = \sum_{i=1}^{n} \sum_{h=1}^{m} a_i \chi_{A_i \cap B_h}(x), \quad x \in X\,.$$

Similarly,

$$\psi(x) = \sum_{h=1}^{m} \sum_{i=1}^{n} b_h \chi_{A_i \cap B_h}(x), \quad x \in X\,.$$

Now, we show that any positive Borel function can be approximated by simple functions.

Proposition 2.21 *Let φ be a positive extended Borel function on a measurable space $(X, \mathcal{E})$. Define for any $n \in \mathbb{N}$*

$$\varphi_n(x) = \begin{cases} \frac{i-1}{2^n} & \text{if } \frac{i-1}{2^n} \le \varphi(x) < \frac{i}{2^n}, \ i = 1, 2, \ldots, n2^n, \\ \\ n & \text{if } \varphi(x) \ge n. \end{cases} \tag{2.3}$$

Then, $(\varphi_n)_n \subset \mathcal{S}(X)$, $(\varphi_n)_n$ is increasing, and $\varphi_n(x) \to \varphi(x)$ for every $x \in X$. If, in addition, φ is bounded, then the convergence is uniform.

Proof. For every $n \in \mathbb{N}$ and $i = 1, \ldots, n2^n$ set

$$E_{n,i} = \varphi^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right), \quad F_n = \varphi^{-1}([n, +\infty)).$$

Since φ is Borel, we have $E_{n,i}$, $F_n \in \mathcal{B}(\mathbb{R}^N)$ and

$$\varphi_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n\chi_{F_n}.$$

Then, by Remark 2.20, $\varphi_n \in \mathcal{S}(X)$. Let $x \in X$ be such that $\frac{i-1}{2^n} \le \varphi(x) < \frac{i}{2^n}$. Then, $\frac{2i-2}{2^{n+1}} \le \varphi(x) < \frac{2i}{2^{n+1}}$ and we have

$$\varphi_{n+1}(x) = \frac{2i-2}{2^{n+1}} \quad \text{or} \quad \varphi_{n+1}(x) = \frac{2i-1}{2^{n+1}}.$$

In any case, $\varphi_n(x) \le \varphi_{n+1}(x)$. If $\varphi(x) \ge n$, then we have $\varphi(x) \ge n+1$ or $n \le \varphi(x) < n+1$. In the first case $\varphi_{n+1}(x) = n+1 > n = \varphi_n(x)$. In the second case let $i = 1, \ldots, (n+1)2^{n+1}$ be such that $\frac{i-1}{2^{n+1}} \le \varphi(x) < \frac{i}{2^{n+1}}$. Since $\varphi(x) \ge n$, we deduce $\frac{i}{2^{n+1}} > n$, by which $i = (n+1)2^{n+1}$. Then $\varphi_{n+1}(x) = n + 1 - \frac{1}{2^{n+1}} > n = \varphi_n(x)$. This proves that $(\varphi_n)_n$ is increasing.

Next, fix $x \in X$ and let $n > \varphi(x)$. Then,

$$0 \le \varphi(x) - \varphi_n(x) < \frac{1}{2^n}. \tag{2.4}$$

So, $\varphi_n(x) \to \varphi(x)$ as $n \to \infty$. Finally, if $0 \le \varphi(x) \le M$ for all $x \in X$ and some $M > 0$, then (2.4) holds for any $x \in X$ provided that $n > M$. Thus $\varphi_n \to \varphi$ uniformly. □

Definition 2.22 *Let $(X, \mathcal{E}, \mu)$ be a measure space and $\varphi_n, \varphi : X \to \mathbb{R}$. We say that $(\varphi_n)_n$ converges to a function φ*

- almost everywhere (a.e.) *if there exists a set $F \in \mathcal{E}$, of measure 0, such that*
$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x) \qquad \forall x \in X \setminus F;$$
- almost uniformly (a.u.) *if, for any $\varepsilon > 0$, there exists $F_\varepsilon \in \mathcal{E}$ such that $\mu(F_\varepsilon) < \varepsilon$ and $\varphi_n \to \varphi$ uniformly on $X \setminus F_\varepsilon$.*

Exercise 2.23 Let $(\varphi_n)_n$ be a sequence of Borel functions on a measure space $(X, \mathcal{E}, \mu)$.

1. Show that the pointwise limit of φ_n, when it exists, is also a Borel function.
2. Show that, if $\varphi_n \stackrel{a.u.}{\longrightarrow} \varphi$, then $\varphi_n \stackrel{a.e.}{\longrightarrow} \varphi$.
3. Show that, if $\varphi_n \stackrel{a.e.}{\longrightarrow} \varphi$ and $\varphi_n \stackrel{a.e.}{\longrightarrow} \psi$, then $\varphi = \psi$ except on a set of measure 0.
4. We say that $\varphi_n \to \varphi$ uniformly almost everywhere if there exists $F \in \mathcal{E}$ of measure 0 such that $\varphi_n \to \varphi$ uniformly in $X \setminus F$. Show that almost uniform convergence does not imply uniform convergence almost everywhere.

   HINT: consider $\varphi_n(x) = x^n$ for $x \in [0, 1]$.

Example 2.24 Observe that the a.e. limit of Borel functions may not be Borel. Indeed, in the trivial sequence $\varphi_n \equiv 0$ defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ (denoting λ the Lebesgue measure) converges a.e. to χ_C, where C is Cantor set (see Example 1.49), and also to χ_A where A is any subset of C which is not a Borel set. This is a consequence of the fact that Lebesgue measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is not complete. On the other hand, if the domain $(X, \mathcal{E}, \mu)$ of $(\varphi_n)_n$ is a *complete* measure space, then the a.e. limit of $(\varphi_n)_n$ is always a Borel function.

The following result establishes a suprising consequence of a.e. convergence on sets of finite measure.

Theorem 2.25 (Severini-Egorov) *Let $(\varphi_n)_n$ be a sequence of Borel functions on a measure space $(X, \mathcal{E}, \mu)$. If $\mu(X) < \infty$ and φ_n converges a.e. to a Borel function φ, then $\varphi_n \xrightarrow{a.u.} \varphi$.*

Proof. For any $k, n \in \mathbb{N}$ define

$$A_n^k = \bigcup_{i=n}^{\infty} \left\{ x \in X \;\middle|\; |\varphi(x) - \varphi_i(x)| > \frac{1}{k} \right\}.$$

Observe that $(A_n^k)_n \in \mathcal{E}$ because φ_n and φ are Borel functions. Also,

$$A_n^k \downarrow \limsup_{n\to\infty} \left\{ x \in X \;\middle|\; |\varphi(x) - \varphi_n(x)| > \frac{1}{k} \right\} =: A^k \quad (n \to \infty).$$

So, $A^k \in \mathcal{E}$. Moreover, for every $x \in A^k$, $|\varphi(x) - \varphi_n(x)| > \frac{1}{k}$ for infinitely many indeces n. Thus, $\mu(A^k) = 0$ by our hypothesis. Recalling that μ is finite, we conclude that, for every $k \in \mathbb{N}$, $\mu(A_n^k) \downarrow 0$ as $n \to \infty$. Therefore, for any fixed $\varepsilon > 0$, the exists an increasing sequence of integers $(n_k)_k$ such that $\mu(A_{n_k}^k) < \frac{\varepsilon}{2^k}$ for all $k \in \mathbb{N}$. Let us set

$$F_\varepsilon := \bigcup_{k=1}^{\infty} A_{n_k}^k.$$

Then, $\mu(F_\varepsilon) \leq \sum_k \mu(A_{n_k}^k) < \varepsilon$. Moreover, for any $k \in \mathbb{N}$, we have that

$$i \geq n_k \quad \Longrightarrow \quad |\varphi(x) - \varphi_i(x)| \leq \frac{1}{k} \qquad \forall x \in X \setminus F_\varepsilon.$$

This shows that $\varphi_n \to \varphi$ uniformly on $X \setminus F_\varepsilon$. $\square$

Example 2.26 The above result is false when $\mu(X) = \infty$. For instance, let $\varphi_n = \chi_{[n,\infty)}$ defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Then, $\varphi_n \to 0$ pointwise, but $\lambda(\{x \in \mathbb{R} \,|\, |\varphi_n| = 1\}) = +\infty$.

2.1.3 Approximation by continuous functions

The object of this section is to prove that a Borel function can be approximated in a measure theoretical sense by a continuous function, as shown by the following result known as Lusin's theorem.

Theorem 2.27 (Lusin) *Let μ be a Radon measure on $\mathbb{R}^N$ and $\varphi : \mathbb{R}^N \to \mathbb{R}$ be a Borel function. Assume $A \subset \mathbb{R}^N$ is a Borel set such that*

$$\mu(A) < \infty \qquad \& \qquad \varphi(x) = 0 \quad \forall x \notin A.$$

Then, for every $\varepsilon > 0$, there exists a continuous function $f_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ with compact support[(1)] *such that*

$$\mu\big(\{x \in \mathbb{R}^N \,|\, \varphi(x) \neq f_\varepsilon(x)\}\big) < \varepsilon \tag{2.5}$$

$$\sup_{x\in\mathbb{R}^N} |f_\varepsilon(x)| \leq \sup_{x\in\mathbb{R}^N} |\varphi(x)| \tag{2.6}$$

Proof. We split the reasoning into several steps.

1. Assume A is compact and $0 \leq \varphi < 1$, and let V be a bounded open set such that $A \subset V$. Consider the sequence (T_n) of measurable sets defined by

$$\begin{aligned} T_1 &= \left\{ x \in A \,\middle|\, \frac{1}{2} \leq \varphi(x) < 1 \right\} \\ T_n &= \left\{ x \in A \,\middle|\, \frac{1}{2^n} \leq \varphi(x) - \sum_{i=1}^{n-1} \frac{1}{2^i}\chi_{T_i}(x) < \frac{1}{2^{n-1}} \right\} \qquad \forall n \geq 2 \end{aligned}$$

   Arguing by induction, it is soon realized that, for any $x \in A$ and any $i \in \mathbb{N}$, $\chi_{T_i}(x) = a_i$, where a_i is the i-th digit in the binary expansion of $\varphi(x)$, i.e., $\varphi(x) = 0.a_1a_2\ldots a_i\ldots$. Therefore,

$$0 \leq \varphi(x) - \sum_{i=1}^{n} \frac{1}{2^i}\chi_{T_i}(x) < \frac{1}{2^n} \quad \forall x \in \mathbb{R}^N,\ \forall n \in \mathbb{N}\,.$$

   Hence,

$$\varphi(x) = \sum_{n=1}^{\infty} \frac{1}{2^n}\chi_{T_n}(x) \qquad \forall x \in \mathbb{R}^N\,, \tag{2.7}$$

   where the series converges uniformly in $\mathbb{R}^N$.

(1) For any continuous function $f : \mathbb{R}^N \to \mathbb{R}$, the *support* of f is the closure of the set $\{x \in \mathbb{R}^N \mid f(x) \neq 0\}$. Such a set will be denoted by supp(f).

2. Fix $\varepsilon > 0$. Owing to Theorem 1.55, for every n there exist a compact set K_n and an open set V_n such that

$$K_n \subset T_n \subset V_n \qquad \& \qquad \mu(V_n \setminus K_n) < \frac{\varepsilon}{2^n}$$

Possibly replacing V_n by $V_n \cap V$, we may assume $V_n \subset V$. Define

$$f_n(x) = \frac{d_{V_n^c}(x)}{d_{K_n}(x) + d_{V_n^c}(x)} \qquad \forall x \in \mathbb{R}^N$$

It is immediate to check that f_n is continuous for all $n \in \mathbb{N}$ and

$$0 \le f_n(x) \le 1 \quad \forall x \in \mathbb{R}^N \qquad \& \qquad f_n \equiv \begin{cases} 1 \text{ on } K_n \\ 0 \text{ on } V_n^c \end{cases}$$

So, in some sense, f_n approximates χ_{T_n}.

3. Now, let us set

$$f_\varepsilon(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} f_n(x) \qquad \forall x \in \mathbb{R}^N \tag{2.8}$$

Since $\sum_{n=1}^{\infty} \frac{1}{2^n} f_n$ is totally convergent, f_ε is continuous. Moreover,

$$\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq 0\} \subset \bigcup_{n=1}^{\infty} \{x \in \mathbb{R}^N \mid f_n(x) \neq 0\} \subset \bigcup_{n=1}^{\infty} V_n \subset V\,,$$

and so $\text{supp}(f_\varepsilon) \subset \overline{V}$. Consequently, $\text{supp}(f_\varepsilon)$ is compact. Furthermore, by (2.7) and (2.8),

$$\begin{aligned} \big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\} \quad &\subset \quad \bigcup_{n=1}^{\infty} \big\{x \in \mathbb{R}^N \mid f_n(x) \neq \chi_{T_n}(x)\big\} \\ &\subset \quad \bigcup_{n=1}^{\infty} (V_n \setminus K_n) \end{aligned}$$

which implies, in turn,

$$\mu\Big(\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\}\Big) \le \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon$$

Thus, conclusion (2.5) holds when A is compact and $0 \le \varphi < 1$.

4. Obviously, (2.5) also holds when A is compact and $0 \le \varphi < M$ for some $M > 0$ (it suffices to replace φ by φ/M). Moreover, if A is compact and φ is bounded, then $|\varphi| < M$ for some $M > 0$. So, in order to derive (2.5) in this case it suffices to decompose $\varphi = \varphi^+ - \varphi^-$ (2) and observe that $0 \le \varphi^+, \varphi^- < M$.

5. We will now remove the compactness assumption for A. By Theorem 1.55, there exists a compact set $K \subset A$ such that $\mu(A \setminus K) < \varepsilon$. Let us set
$$\bar{\varphi} = \chi_K \varphi$$
Since $\bar{\varphi}$ vanishes outside K, we can approximate $\bar{\varphi}$, in the above sense, by a continuous function with compact support, say f_ε. Then,
$$\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\} \subset \big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \bar{\varphi}(x)\big\} \cup (A \setminus K)\,,$$
since, for any $x \in K \cup A^c$, $f_\varepsilon(x) \neq \varphi(x)$ implies that $f_\varepsilon(x) \neq \bar{\varphi}(x)$. Hence,
$$\mu\Big(\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\}\Big) < 2\varepsilon$$

6. In order to remove the boundedness hypothesis for f, define measurable sets (B_n) by
$$B_n = \{x \in A \mid |\varphi(x)| \ge n\} \qquad n \in \mathbb{N}$$
Clearly,
$$B_{n+1} \subset B_n \qquad \& \qquad \bigcap_{n\in\mathbb{N}} B_n = \varnothing$$
Since $\mu(A) < \infty$, Proposition 1.16 yields $\mu(B_n) \to 0$. Therefore, for some $\bar{n} \in \mathbb{N}$, we have $\mu(B_{\bar{n}}) < \varepsilon$. Proceeding as above, we define
$$\bar{\varphi} = (1 - \chi_{B_{\bar{n}}})\varphi$$
Since $\bar{\varphi}$ is bounded (by $\bar{n}$), we can approximate $\bar{\varphi}$, in the above sense, by a continuous function with compact support, that we again label f_ε. Then,
$$\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\} \subset \big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \bar{\varphi}(x)\big\} \cup B_{\bar{n}}$$

(2) $\varphi^+ = \max\{\varphi, 0\}\,,\ \varphi^- = \max\{-\varphi, 0\}$

So,

$$\mu\Big(\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) \neq \varphi(x)\big\}\Big) < 2\varepsilon$$

The proof of (2.5) is thus complete.

7. Finally, in order to prove (2.6), suppose $R := \sup_{\mathbb{R}^N} |\varphi| < \infty$. Define

$$\theta_R : \mathbb{R} \to \mathbb{R} \qquad \theta_R(t) = \begin{cases} t & \text{if } |t| < R, \\ R\dfrac{t}{|t|} & \text{if } |t| \geq R \end{cases}$$

and $\bar{f}_\varepsilon = \theta_R \circ f_\varepsilon$ to obtain $|\bar{f}_\varepsilon| \leq R$. Since θ_R is continuous, so is $\bar{f}_\varepsilon$. Furthermore, $\operatorname{supp}(\bar{f}_\varepsilon) = \operatorname{supp}(f_\varepsilon)$ and

$$\big\{x \in \mathbb{R}^N \mid f_\varepsilon(x) = \varphi(x)\big\} \subset \big\{x \in \mathbb{R}^N \mid \bar{f}_\varepsilon(x) = \varphi(x)\big\}.$$

This completes the proof. □

2.2 Integral of Borel functions

Let $(X, \mathcal{E}, \mu)$ be a given measure space. In this section we will construct the integral of a Borel function $\varphi\colon X \to \overline{\mathbb{R}}$ with respect to μ. We will first consider the special case of positive functions, and then the case of functions with variable sign. We begin with what can rightfully be considered the central notion of Lebesgue integration.

2.2.1 Repartition function

Let $\varphi\colon X \to [0, \infty]$ be a Borel function. The *repartition function* F of φ is defined by

$$F(t)\colon = \mu(\{\varphi > t\}) = \mu(\varphi > t), \quad t \geq 0.$$

By definition, $F : [0, \infty) \to [0, \infty]$ is a decreasing [(3)] function; then F possesses limit at ∞. Moreover, since

$$\{\varphi = \infty\} = \bigcap_{n=1}^{\infty} \{\varphi > n\},$$

[(3)] A function $f\colon \mathbb{R} \to \mathbb{R}$ is decreasing if $t_1 < t_2 \implies f(t_1) \geq f(t_2)$, positive if $f(t) \geq 0$ for all $t \in \mathbb{R}$.

we have

$$\lim_{t\to\infty} F(t) = \lim_{n\to\infty} F(n) = \lim_{n\to\infty} \mu(\varphi > n) = \mu(\varphi = \infty)$$

whenever μ is finite. Other important properties of F are provided by the following result.

Proposition 2.28 *Let $\varphi\colon X \to [0,\infty]$ be a Borel function and let F be its repartition function. Then, the following properties hold:*

(i) *For any $t_0 \ge 0$,*

$$\lim_{t\downarrow t_0} F(t) = F(t_0),$$

(that is, F is right continuous).

(ii) *If $\mu(X) < \infty$, then, for any $t_0 > 0$,*

$$\lim_{t\uparrow t_0} F(t) = \mu(\varphi \ge t_0)$$

(that is, F possesses left limits[(4)]).

Proof. First observe that, since F is a monotonic function, then F possesses left limit at any $t > 0$ and right limit at any $t \ge 0$. Let us prove (i). We have

$$\lim_{t\downarrow t_0} F(t) = \lim_{n\to\infty} F\Big(t_0 + \frac{1}{n}\Big) = \lim_{n\to\infty} \mu\Big(\varphi > t_0 + \frac{1}{n}\Big) = \mu(\varphi > t_0) = F(t_0),$$

since

$$\Big\{\varphi > t_0 + \frac{1}{n}\Big\} \uparrow \{\varphi > t_0\}.$$

Now, to prove (ii), we note that

$$\Big\{\varphi > t_0 - \frac{1}{n}\Big\} \downarrow \{\varphi \ge t_0\}\,.$$

Thus, recalling that μ is finite, we have

$$\lim_{t\uparrow t_0} F(t) = \lim_{n\to\infty} F\Big(t_0 - \frac{1}{n}\Big) = \lim_{n\to\infty} \mu\Big(\varphi > t_0 - \frac{1}{n}\Big) = \mu(\varphi \ge t_0)\,,$$

and (ii) follows. □

From Proposition 2.28 it follows that, when μ is finite, F is continuous at t_0 iff $\mu(\varphi = t_0) = 0$.

[(4)] In the literature, a function that is right-continuous and has left limits is called a *cadlag* function.

2.2.2 Integral of positive simple functions

We now proceed to define the integral in the class $\mathcal{S}_+(X)$ of positive simple functions. Let $\varphi \in \mathcal{S}_+(X)$. Then, according to Remark 2.20.1,

$$\varphi(x) = \sum_{k=1}^{n} a_k \chi_{A_k}(x) \qquad x \in X,$$

where $a_1, \dots, a_n \geq 0$, and $A_1, \dots, A_n$ are mutually disjoint sets of $\mathcal{E}$ such that $A_1 \cup \cdots \cup A_n = X$. Using the convention $0 \cdot \infty = 0$, we define the *integral* of φ over X with respect to μ by

$$\int_X \varphi(x)\mu(dx) = \int_X \varphi d\mu = \sum_{k=1}^{n} a_k \mu(A_k). \tag{2.9}$$

It is easy to see that the above definition is independent of the representation of φ. Indeed, given disjoint sets $B_1, \dots, B_m \in \mathcal{E}$ with $B_1 \cup \cdots \cup B_m = X$ and real numbers $b_1, \dots, b_m \geq 0$ such that

$$\varphi(x) = \sum_{i=1}^{m} b_i \chi_{B_i}(x) \qquad x \in X,$$

we have that

$$A_k = \bigcup_{i=1}^{m} (A_k \cap B_i) \qquad B_i = \bigcup_{k=1}^{n} (A_k \cap B_i)$$

and

$$A_k \cap B_i \neq \varnothing \quad \Longrightarrow \quad a_k = b_i\,.$$

Therefore,

$$\begin{aligned} \sum_{k=1}^{n} a_k \mu(A_k) &= \sum_{k=1}^{n} \sum_{i=1}^{m} a_k \mu(A_k \cap B_i) \\ &= \sum_{i=1}^{m} \sum_{k=1}^{n} b_i \mu(A_k \cap B_i) = \sum_{i=1}^{m} b_i \mu(B_i)\,. \end{aligned}$$

Proposition 2.29 *Let $\varphi, \psi \in \mathcal{S}_+(X)$ and let $\alpha, \beta \geq 0$. Then,*

$$\int_X (\alpha\varphi + \beta\psi) d\mu = \alpha \int_X \varphi d\mu + \beta \int_X \psi d\mu$$

Proof. Owing to Remark 2.20.2, φ and ψ can be represented using the same family of mutually disjoint sets $A_1, \dots, A_n$ of $\mathcal{E}$ as

$$\varphi = \sum_{k=1}^{n} a_k \chi_{A_k} \qquad \psi = \sum_{k=1}^{n} b_k \chi_{A_k}.$$

Then,

$$\begin{aligned}
\int_X (\alpha\varphi + \beta\psi) d\mu &= \sum_{k=1}^{n} (\alpha a_k + \beta b_k)\mu(A_k) \\
&= \alpha \sum_{k=1}^{n} a_k \mu(A_k) + \beta \sum_{k=1}^{n} b_k \mu(A_k) \\
&= \alpha \int_X \varphi d\mu + \beta \int_X \psi d\mu
\end{aligned}$$

as required. □

Example 2.30 Let us choose a representation of a positive simple function φ of the form

$$\varphi(x) = \sum_{k=1}^{n} a_k \chi_{A_k} \qquad x \in X,$$

with $0 < a_1 < a_2 < \cdots < a_n$. Then, the repartition function F of φ is given by

$$F(t) = \begin{cases}
\mu(A_1) + \mu(A_2) + \cdots + \mu(A_n) = F(0) & \text{if } 0 \le t < a_1, \\
\cdots\cdots\cdots\cdots\cdots & \cdots\cdots \\
\mu(A_k) + \mu(A_{k+1}) + \cdots + \mu(A_n) = F(a_{k-1}) & \text{if } a_{k-1} \le t < a_k, \\
\cdots\cdots & \cdots \\
\mu(A_n) = F(a_{n-1}) & \text{if } a_{n-1} \le t < a_n, \\
0 = F(a_n) & \text{if } t \ge a_n.
\end{cases}$$

Thus, setting $a_0 = 0$, we have $\mu(A_k) = F(a_{k-1}) - F(a_k)$ and $F(t) = \sum_{k=1}^{n} F(a_{k-1})\chi_{[a_{k-1},a_k)}(t)$. Then F is a simple function itself on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and

$$\begin{aligned}
\int_X \varphi d\mu &= \sum_{k=1}^{n} a_k \mu(A_k) = \sum_{k=1}^{n} a_k (F(a_{k-1}) - F(a_k)) \\
&= \sum_{k=1}^{n} F(a_{k-1})(a_k - a_{k-1}) = \int_0^{\infty} F(t) dt,
\end{aligned}$$

where $\int_0^\infty F(t)dt$ denotes the integral of the simple function F with respect to the Lebesgue measure.

2.2.3 The archimedean integral

The identity we have obtained in Example 2.30 for simple functions, that is,

$$\int_X \varphi d\mu = \int_0^\infty \mu(\varphi > t)dt \tag{2.10}$$

makes perfectly sense because the repartition function of a simple function is a simple function itself (and even a step function). In order to be able to take such an identity as the definition of the integral of φ when φ is a positive Borel function, we first have to give its right-hand side a meaning. For this, we need to define, first, the integral of any positive decreasing function $f : [0, \infty) \to [0, \infty]$.

Let Σ be the family of all finite sets of points $\sigma = \{t_0, \dots, t_N\}$ of $[0, \infty]$, where $N \in \mathbb{N}$ and $0 = t_0 < t_1 < \cdots < t_N < \infty$. For any decreasing function $f \colon [0, \infty) \to [0, \infty]$ and any $\sigma = \{t_0, t_1, ..., t_N\} \in \Sigma$, we set

$$I_f(\sigma) = \sum_{k=0}^{N-1} f(t_{k+1})(t_{k+1} - t_k).$$

The *archimedean integral* of f is defined by

$$\int_0^\infty f(t)dt := \sup\{I_f(\sigma) : \ \sigma \in \Sigma\}\,.$$

Exercise 2.31 Let $f \colon [0, \infty) \to [0, \infty]$ be a decreasing function.

1. Show that, if $\sigma, \zeta \in \Sigma$ and $\sigma \subset \zeta$, then $I_f(\sigma) \leq I_f(\zeta)$.

2. Show that, for any pair of decreasing functions $f, g \colon [0, \infty) \to [0, \infty]$ such that $f(x) \leq g(x)$, we have $\int_0^\infty f(t)dt \leq \int_0^\infty g(t)dt$.

3. Show that, if $f(t) = 0$ for all $t > 0$, then $\int_0^\infty f(t)dt = 0$.

Proposition 2.32 *Let* $f_n\colon [0,\infty) \to [0,\infty]$ *be a sequence of decreasing functions such that*

$$f_n(t) \uparrow f(t) \qquad (n \to \infty) \qquad \forall t \geq 0\,.$$

Then,

$$\int_0^\infty f_n(t)dt \quad \uparrow \quad \int_0^\infty f(t)dt\,.$$

Proof. According to Exercise 2.31.2, since $f_n \leq f_{n+1} \leq f$, we obtain $\int_0^\infty f_n(t)dt \leq \int_0^\infty f_{n+1}(t)dt \leq \int_0^\infty f(t)dt$ for every n. Then the inequality $\lim_n \int_0^\infty f_n(t)dt \leq \int_0^\infty f(t)dt$ is clear. To prove the opposite inequality, let $L < \int_0^\infty f(t)dt$. Then there exists $\sigma = \{t_0, \dots, t_N\} \in \Sigma$ such that

$$\sum_{k=0}^{N-1} f(t_{k+1})(t_{k+1} - t_k) > L.$$

Therefore, for n sufficiently large, say $n \geq n_L$,

$$\int_0^\infty f_n(t)dt \geq \sum_{k=0}^{N-1} f_n(t_{k+1})(t_{k+1} - t_k) > L\,.$$

Thus, $\lim_{n\to\infty} \int_0^\infty f_n(t)dt > L$. Since L is any number less than $\int_0^\infty f(t)dt$, we conclude that

$$\lim_{n\to\infty} \int_0^\infty f_n(t)dt \geq \int_0^\infty f(t)dt\,. \qquad \square$$

The definition of archimedean integral can be easily adapted to the case of a bounded interval $[0,a]$. Given a decreasing function $f\colon [0,a] \to [0,\infty]$ it suffices to set

$$\int_0^a f(t)dt = \int_0^\infty f^*(t)dt$$

where

$$f^*(t) = \begin{cases} f(t) & \text{if} \quad t \in [0,a], \\ 0 & \text{if} \quad t > a. \end{cases} \tag{2.11}$$

Exercise 2.33 1. Given a decreasing function $f\colon [0,a] \to [0,\infty]$, show that

$$\int_0^a f(t)dt \geq af(a).$$

2. Given a decreasing function $f\colon [0,\infty) \to [0,\infty]$, show that

$$\int_0^\infty f(t)dt \geq \int_0^a f(t)dt \quad \forall a > 0.$$

2.2.4 Integral of positive Borel functions

Given a measure space $(X, \mathcal{E}, \mu)$ and an extended positive Borel function φ, we can now define the *integral* of φ over X with respect to μ according to (2.10), that is,

$$\int_X \varphi d\mu = \int_X \varphi(x)\mu(dx)\colon = \int_0^\infty \mu(\varphi > t)dt\,, \tag{2.12}$$

where the integral in the right-hand side is the archimedean integral of the decreasing positive function $t \mapsto \mu(\varphi > t)$. If the integral of φ is finite we say that φ is *μ–summable.*

Proposition 2.34 (Markov) *Let $\varphi : X \to [0,\infty]$ be a Borel function. Then, for any $a \in (0,\infty)$,*

$$\mu(\varphi > a) \leq \frac{1}{a}\int_X \varphi d\mu\,. \tag{2.13}$$

Proof. Recalling Exercise 2.33, we have that, for any $a \in (0,\infty)$,

$$\int_X \varphi d\mu = \int_0^{+\infty} \mu(\varphi > t)dt \geq \int_0^a \mu(\varphi > t)dt \geq a\mu(\varphi > a)\,.$$

The conclusion follows. □

Markov's inequality has important consequences. Generalizing the notion of a.e. convergence (see Definition 2.22), we say that a property concerning the points of X holds *almost everywere* (a.e.), if it holds for all points of X except for a set $E \in \mathcal{E}$ with $\mu(E) = 0$.

Proposition 2.35 *Let $\varphi : X \to [0,\infty]$ be a Borel function.*

(i) *If φ is μ–summable, then the set $\{\varphi = \infty\}$ has measure 0, that is, φ is a.e. finite.*

(ii) *The integral of φ vanishes if and only if φ is equal to 0 a.e.*

Proof.

(i) From (2.13) it follows that $\mu(\varphi > a) < \infty$ for all $a > 0$ and

$$\lim_{a\to\infty} \mu(\varphi > a) = 0.$$

Since

$$\{\varphi > n\} \downarrow \{\varphi = \infty\},$$

we have that

$$\mu(\varphi = \infty) = \lim_{n\to\infty} \mu(\varphi > n) = 0.$$

(ii) If $\varphi \overset{a.e.}{=} 0$, we have $\mu(\varphi > t) = 0$ for all $t > 0$. Then $\int_X \varphi d\mu = \int_0^{+\infty} \mu(\varphi > t)dt = 0$ (see Exercise 2.31.3). Conversely, let $\int_X \varphi d\mu = 0$. Then, Markov's inequality yields $\mu(\varphi > a) = 0$ for all $a > 0$. Since $\{\varphi > \frac{1}{n}\} \uparrow \{\varphi > 0\}$, so

$$\mu(\varphi > 0) = \lim_{n\to\infty} \mu\Big(\varphi > \frac{1}{n}\Big) = 0\,.$$

The proof is complete. □

The following is a first result studying the passage to the limit under the integral sign. It is referred to as the Monotone Convergence Theorem.

Proposition 2.36 (Beppo Levi) *Let $\varphi_n : X \to [0,\infty]$ be an increasing sequence of Borel functions, and set*

$$\varphi(x) = \lim_{n\to\infty} \varphi_n(x) \qquad \forall x \in X\,.$$

Then,

$$\int_X \varphi_n d\mu \;\; \uparrow \;\; \int_X \varphi d\mu\,.$$

Proof. Observe that, in consequence of the assumptions,

$$\{\varphi_n > t\} \uparrow \{\varphi > t\} \qquad \forall t > 0\,.$$

Therefore, $\mu(\varphi_n > t) \uparrow \mu(\varphi > t)$ for any $t > 0$. The conclusion follows from Proposition 2.32. □

Combining Propositions 2.21 and 2.36 we deduce the following result.

Proposition 2.37 *Let $\varphi : X \to [0,\infty]$ be a Borel function. Then, there exist positive simple functions $\varphi_n : X \to [0,\infty)$ such that $\varphi_n \uparrow \varphi$ pointwise and*

$$\int_X \varphi_n d\mu \ \ \uparrow \ \ \int_X \varphi d\mu\,.$$

Let us state some basic properties of the integral.

Proposition 2.38 *Let $\varphi, \psi : X \to [0,\infty]$ be μ–summable. Then the following properties hold.*

(i) *If $a, b \geq 0$, then $\int_X (a\varphi + b\psi) d\mu = a \int_X \varphi d\mu + b \int_X \psi d\mu$.*

(ii) *If $\varphi \geq \psi$, then $\int_X \varphi d\mu \geq \int_X \psi d\mu$.*

Proof. The conclusion of point (i) holds for $\varphi, \psi \in \mathcal{S}_+(X)$, thanks to Proposition 2.29. To obtain it for Borel functions it suffices to apply Proposition 2.37.

To justify (ii), observe that the trivial inclusion $\{\psi > t\} \subset \{\varphi > t\}$ yields $\mu(\psi > t) \leq \mu(\varphi > t)$. The conclusion follows (see also Exercise 2.31.3). □

Proposition 2.39 *Let $\varphi_n : X \to [0,\infty]$ be a sequence of Borel functions and set*

$$\varphi(x) = \sum_{n=1}^{\infty} \varphi_n(x) \quad \forall x \in X.$$

Then

$$\sum_{n=1}^{\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu.$$

Proof. For every n set

$$f_n = \sum_{k=1}^{n} \varphi_k.$$

Then $f_n \uparrow \varphi$. By applying Proposition 2.36 we get

$$\int_X f_n d\mu \to \int_X \varphi d\mu.$$

On the other hand (i) of Proposition 2.38 implies

$$\int_X f_n d\mu = \sum_{k=1}^{n} \int_X \varphi_k d\mu \to \sum_{k=1}^{\infty} \int_X \varphi_k d\mu. \qquad \square$$

The following basic result, known as *Fatou's Lemma*, provides a semicontinuity property of the integral.

Lemma 2.40 (Fatou) *Let* $\varphi_n : X \to [0, \infty]$ *be a sequence of Borel functions and set* $\varphi = \liminf_{n\to\infty} \varphi_n$. *Then,*

$$\int_X \varphi d\mu \le \liminf_{n\to\infty} \int_X \varphi_n d\mu\,. \tag{2.14}$$

Proof. Setting $\psi_n(x) = \inf_{m\ge n} \varphi_m(x)$, we have that $\psi_n(x) \uparrow \varphi(x)$ for every $x \in X$. Consequently, by the Monotone Convergence Theorem,

$$\int_X \varphi d\mu = \lim_{n\to\infty} \int_X \psi_n d\mu = \sup_{n\in\mathbb{N}} \int_X \psi_n d\mu\,.$$

On the other hand, since $\psi_n \le \varphi_m$ for every $m \ge n$, we have

$$\int_X \psi_n d\mu \le \inf_{m\ge n} \int_X \varphi_m d\mu\,.$$

So,

$$\int_X \varphi d\mu \le \sup_{n\in\mathbb{N}} \inf_{m\ge n} \int_X \varphi_m d\mu = \liminf_{n\to\infty} \int_X \varphi_n d\mu\,. \quad \square$$

Corollary 2.41 *Let $\varphi_n : X \to [0, \infty]$ be a sequence of Borel functions converging to φ pointwise. If, for some $M > 0$,*

$$\int_X \varphi_n d\mu \le M \qquad \forall n \in \mathbb{N}\,,$$

then $\int_X \varphi d\mu \le M$.

Remark 2.42 Proposition 2.36 and Corollary 2.41 can be given a version that applies to a.e. convergence. In this case, the fact that the limit φ is a Borel function is no longer guaranteed (see Example 2.24). Therefore, such a property must be assumed a priori, or else measure μ must be complete.

Exercise 2.43 State and prove the analogues of Proposition 2.36 and of Corollary 2.41 for a.e. convergence.

Example 2.44 Consider the counting measure μ on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Then any function $x : i \mapsto x(i)$ is Borel and $x = \sum_{i=1}^{\infty} x(i)\chi_{\{i\}}$. Then, by Proposition 2.39, if x is positive we have

$$\int_{\mathbb{N}} x\, d\mu = \sum_{i=1}^{\infty} x(i)\mu(\{i\}) = \sum_{i=1}^{\infty} x(i).$$

Example 2.45 Consider the measure space $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ of the previous example. Let $(x_n)_n$ be a sequence of positive functions such that, for every $i \in \mathbb{N}$, $x_n(i) \uparrow x(i)$ as $n \to \infty$. Then, Beppo Levi's Theorem ensures that

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} x_n(i) = \sum_{i=1}^{\infty} x(i)\,.$$

Compare with Lemma 1.45.

Exercise 2.46 Let $a_{ni} \ge 0$ for n, $i \in \mathbb{N}$. Show that

$$\sum_{n=1}^{\infty}\sum_{i=1}^{\infty} a_{ni} = \sum_{i=1}^{\infty}\sum_{n=1}^{\infty} a_{ni}.$$

Hint: Set $x_n : i \mapsto a_{ni}$. Then x_n is a sequence of positive Borel functions on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Use Proposition 2.39 to conclude.

Exercise 2.47 Let $\varphi, \psi : X \to [0,\infty]$ be Borel functions.

1. Show that, if $\varphi \le \psi$ a.e., then $\int_X \varphi d\mu \le \int_X \psi d\mu$.

2. Show that, if $\varphi = \psi$ a.e., then $\int_X \varphi d\mu = \int_X \psi d\mu$.

3. Show that the monotonicity of φ_n is an essential hypothesis for Beppo Levi's Theorem.

   HINT: consider $\varphi_n(x) = \chi_{[n,n+1)}(x)$ for $x \in \mathbb{R}$.

4. Give an example to show that the inequality in Fatou's Lemma can be strict.

   HINT: set $\varphi_{2n}(x) = \chi_{[0,1)}(x)$ and $\varphi_{2n+1}(x) = \chi_{[1,2)}(x)$ for $x \in \mathbb{R}$.

Exercise 2.48 Let $(X, \mathcal{E}, \mu)$ be a measure space. The following statements are equivalent:

1. μ is σ-finite;

2. there exists a μ-summable function φ on X such that $\varphi(x) > 0$ for all $x \in X$.

2.2.5 Integral of functions with variable sign

Let $\varphi\colon X \to \overline{\mathbb{R}}$ be a Borel function. We say that φ is *μ–summable* if there exist two μ–summable Borel functions $f, g : X \to [0,\infty]$ such that

$$\varphi(x) = f(x) - g(x) \qquad \forall x \in X\,. \tag{2.15}$$

In this case, the number

$$\int_X \varphi d\mu := \int_X f d\mu - \int_X g d\mu\,. \tag{2.16}$$

is called the *integral of φ over X with respect to μ*. Let us check, as usual, that the integral of φ is independent of the choice of functions f, g used to represent φ as in (2.15). Indeed, let $f_1, g_1 : X \to [0,\infty]$ be μ–summable Borel functions such that

$$\varphi(x) = f_1(x) - g_1(x) \qquad \forall x \in X\,.$$

Then, f, g, f_1 and g_1 are finite a.e., and

$$f(x) + g_1(x) = f_1(x) + g(x) \qquad x \in X \text{ a.e.}$$

Therefore, owing to Exercise 2.47.2 and Proposition 2.38, we have

$$\int_X f d\mu + \int_X g_1 d\mu = \int_X f_1 d\mu + \int_X g d\mu\,.$$

Since the above integrals are all finite, we deduce that

$$\int_X f d\mu - \int_X g d\mu = \int_X f_1 d\mu - \int_X g_1 d\mu$$

as claimed.

Remark 2.49 Let $\varphi \colon X \to \overline{\mathbb{R}}$ be a μ–summable function.

1. The positive and negative parts

$$\varphi^+(x) = \max\{\varphi(x), 0\}, \quad \varphi^-(x) = \max\{-\varphi(x), 0\}.$$

are positive Borel functions such that $\varphi = \varphi^+ - \varphi^-$. We claim that φ^+ and φ^- are μ–summable. Indeed let $f, g : X \to [0, \infty]$ be Borel functions satisfying (2.15). If $x \in X$ is such that $\varphi(x) \geq 0$, then $\varphi^+(x) = \varphi(x) \leq f(x)$. So, $\varphi^+(x) \leq f(x)$ for all $x \in X$ and, recalling Exercise 2.47.1, we conclude that φ^+ is μ–summable. Similarly, one can show that φ^- is μ–summable. Therefore,

$$\int_X \varphi d\mu = \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu.$$

2. From the above remark we deduce that φ is μ–summable iff both φ^+ and φ^- are summable. Since $|\varphi| = \varphi^+ + \varphi^-$, it is also true that φ is μ–summable iff $|\varphi|$ is μ–summable. Moreover,

$$\left| \int_X \varphi d\mu \right| \leq \int_X |\varphi| d\mu\,. \tag{2.17}$$

Indeed,

$$\begin{aligned} \left| \int_X \varphi d\mu \right| &= \left| \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu \right| \leq \\ &\leq \int_X \varphi^+ d\mu + \int_X \varphi^- d\mu = \int_X |\varphi| d\mu\,. \end{aligned}$$

Remark 2.50 The notion of integral can be further extended allowing infinite values. A Borel function $\varphi\colon X \to \overline{\mathbb{R}}$ is said to be *μ–integrable* if at least one of the two functions φ^+ and φ^- is μ–summable. In this case, we define

$$\int_X \varphi d\mu = \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu.$$

Notice that $\int_X \varphi d\mu \in \overline{\mathbb{R}}$, in general.

In order to state the analogous of Proposition 2.38, we point out that the sum of two functions with values on the extended space $\overline{\mathbb{R}}$ may not be well defined; thus we need to assume that at least one of the function is real-valued.

Proposition 2.51 *Let $\varphi, \psi : X \to \overline{\mathbb{R}}$ be μ–summable functions. Then, the following properties hold.*

(i) *If $\varphi : X \to \mathbb{R}$, then, for any $\alpha, \beta \in \mathbb{R}$, $\alpha\varphi + \beta\psi$ is μ–summable and*

$$\int_X (\alpha\varphi + \beta\psi) d\mu = \alpha \int_X \varphi d\mu + \beta \int_X \psi d\mu\,.$$

(ii) *If $\varphi \le \psi$, then $\int_X \varphi d\mu \le \int_X \psi d\mu$.*

Proof.

(i) Assume first $\alpha, \beta > 0$ and let f, g, f_1, g_1 be positive μ-summable functions such that

$$\left.\begin{array}{l} \varphi(x) = f(x) - g(x) \\ \psi(x) = f_1(x) - g_1(x) \end{array}\right\} \qquad \forall x \in X\,.$$

Then, since f and g are finite, we have $\alpha\varphi+\beta\psi = (\alpha f+\beta f_1)-(\alpha g+\beta g_1)$ and so,

$$\int_X (\alpha\varphi + \beta\psi) d\mu = \int_X (\alpha f + \beta f_1) d\mu - \int_X (\alpha g + \beta g_1) d\mu\,.$$

The conclusion follows from Proposition 2.38(i). The case when α, β have different signs can be handled similarly.

(ii) Let $\varphi \le \psi$. It is immediate that $\varphi^+ \le \psi^+$ and $\psi^- \le \varphi^-$. Then by Proposition 2.38(ii) we obtain

$$\int_X \psi d\mu = \int_X \psi^+ d\mu - \int_X \psi^- d\mu \ge \int_X \varphi^+ d\mu - \int_X \varphi^- d\mu = \int_X \varphi d\mu.$$

□

Let $\varphi\colon X \to \overline{\mathbb{R}}$ be μ–summable and let $A \in \mathcal{E}$. Then, $\chi_A\varphi$ is μ–summable because $|\chi_A\varphi| \le |\varphi|$. Let us define

$$\int_A \varphi d\mu := \int_X \chi_A \varphi d\mu\,.$$

Since $\varphi = \chi_A\varphi + \chi_{A^c}\varphi$, from Proposition 2.51.(i) we obtain

$$\int_A \varphi d\mu + \int_{A^c} \varphi d\mu = \int_X \varphi d\mu\,. \tag{2.18}$$

Notation 2.52 *If $X = \mathbb{R}^N$ and $A \in \mathcal{B}(\mathbb{R}^N)$, we will write "$\int_A \varphi(x)dx$", "$\int_A \varphi(y)dy$" etc. rather then $\int_A \varphi d\mu$ when the integrals are taken with respect to the Lebesgue measure.*

Proposition 2.53 *Let $\varphi : X \to \overline{\mathbb{R}}$ be a μ-summable function.*

(i) *The set $\{|\varphi| = \infty\}$ has measure 0;*

ii) If $\varphi = 0$ a.e., then $\int_X \varphi d\mu = 0$;

(ii) *If $A \in \mathcal{E}$ has measure 0, then $\int_A \varphi d\mu = 0$;*

(iv) *If $\int_E \varphi d\mu = 0$ for every $E \in \mathcal{E}$, then $\varphi = 0$ a.e.*

Proof. Parts (i), (ii) and (iii) follow immediately from Proposition 2.35. Let us prove (iv). Set $E = \{\varphi^+ > 0\}$. Then we have

$$0 = \int_E \varphi d\mu = \int_X \varphi^+ d\mu.$$

Proposition 2.35(ii) implies $\varphi^+ = 0$ a.e. In a similar way we obtain $\varphi^- = 0$ a.e. □

The key result provided by the next proposition is referred to as the *absolute continuity* property of the integral.

Proposition 2.54 *Let $\varphi\colon X \to \overline{\mathbb{R}}$ be μ–summable. Then, for any $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that*

$$\mu(A) < \delta_\varepsilon \quad \Longrightarrow \quad \int_A |\varphi| d\mu \le \varepsilon \,. \tag{2.19}$$

Proof. Without loss of generality, φ may be assumed to be positive. Then,

$$\varphi_n(x) := \min\{\varphi(x), n\} \uparrow \varphi(x) \qquad \forall x \in X \,.$$

Therefore, by Beppo Levi's Theorem, $\int_X \varphi_n d\mu \uparrow \int_X \varphi d\mu$. So, for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that

$$0 \le \int_X (\varphi - \varphi_n) d\mu < \frac{\varepsilon}{2} \qquad \forall n \ge n_\varepsilon \,.$$

Then, for $\mu(A) < \frac{\varepsilon}{2n_\varepsilon}$, we have

$$\int_A \varphi d\mu \le \int_A \varphi_{n_\varepsilon} d\mu + \int_X (\varphi - \varphi_{n_\varepsilon}) d\mu < \varepsilon \,.$$

We have thus obtained (2.19) with $\delta_\varepsilon = \frac{\varepsilon}{2n_\varepsilon}$. □

Exercise 2.55 Let $\varphi\colon X \to \overline{\mathbb{R}}$ be μ–summable. Show that

$$\lim_{n\to\infty} \int_{\{|\varphi|>n\}} |\varphi| d\mu = 0 \,.$$

2.3 Convergence of integrals

We have already obtained two results that allow passage to the limit in integrals, namely Beppo Levi's Theorem and Fatou's Lemma. In this section, we will further analyze the problem.

2.3.1 Dominated Convergence

We begin with the following classical result, also known as Lebesgue's Dominated Convergence Theorem .

Proposition 2.56 (Lebesgue) *Let $\varphi_n : X \to \overline{\mathbb{R}}$ be a sequence of Borel functions converging to φ pointwise. Assume that there exists a positive μ-summable function $\psi : X \to [0, \infty]$ such that*

$$|\varphi_n(x)| \le \psi(x) \qquad \forall x \in X\,,\ \forall n \in \mathbb{N}. \tag{2.20}$$

Then, φ_n, φ are μ–summable and

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu\,. \tag{2.21}$$

Proof. First, we note that φ_n, φ are μ-summable because they are Borel and, in view of (2.20), $|\varphi(x)| \le \psi(x)$ for any $x \in X$. Let us prove (2.21) when $\psi : X \to [0, +\infty)$. Since $\psi + \varphi_n$ is positive, Fatou's Lemma yields

$$\int_X (\psi + \varphi) d\mu \le \liminf_{n\to\infty} \int_X (\psi + \varphi_n) d\mu = \int_X \psi d\mu + \liminf_{n\to\infty} \int_X \varphi_n d\mu\,.$$

Consequently, since $\int_X \psi d\mu$ is finite, we deduce

$$\int_X \varphi d\mu \le \liminf_{n\to\infty} \int_X \varphi_n d\mu\,. \tag{2.22}$$

Similarly,

$$\int_X (\psi - \varphi) d\mu \le \liminf_{n\to\infty} \int_X (\psi - \varphi_n) d\mu = \int_X \psi d\mu - \limsup_{n\to\infty} \int_X \varphi_n d\mu\,.$$

Whence,

$$\int_X \varphi d\mu \ge \limsup_{n\to\infty} \int_X \varphi_n d\mu\,. \tag{2.23}$$

The conclusion follows from (2.22) and (2.23).

In the general case $\psi : X \to [0, \infty]$, consider $E = \{x \in X \,|\, \psi(x) = \infty\}$. Then (2.21) holds over E^c and, by Proposition 2.35(i), we have $\mu(E) = 0$. Hence we deduce

$$\int_X \varphi_n d\mu = \int_{E^c} \varphi_n d\mu \to \int_{E^c} \varphi d\mu = \int_X \varphi d\mu.$$

□

Exercise 2.57 Derive (2.21) if (2.20) is satisfied a.e. and $\varphi_n \xrightarrow{a.e.} \varphi$, with φ Borel.

Exercise 2.58 Let φ, $\psi : X \to \overline{\mathbb{R}}$ be Borel functions such that φ is μ-summable and ψ in μ-integrable. Assume that φ or ψ is finite. Prove that $\varphi + \psi$ is μ-integrable and

$$\int_X (\varphi + \psi) d\mu = \int_X \varphi d\mu + \int_X \psi d\mu.$$

Exercise 2.59 Let $\varphi_n : X \to \overline{\mathbb{R}}$ be Borel functions satisfying, for some summable function $\psi : X \to \overline{\mathbb{R}}$ and some (Borel) function φ,

$$\left.\begin{array}{l} \varphi_n(x) \geq \psi(x) \\ \varphi_n(x) \uparrow \varphi(x) \end{array}\right\} \qquad \forall x \in X\,.$$

Show that φ_n, φ are μ-integrable and

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu\,.$$

Exercise 2.60 Let $\varphi_n : X \to \overline{\mathbb{R}}$ be Borel functions satisfying, for some μ–summable function ψ and some (Borel) function φ,

$$\left.\begin{array}{l} \varphi_n(x) \geq \psi(x) \\ \varphi_n(x) \to \varphi(x) \end{array}\right\} \qquad \forall x \in X\,.$$

Show that φ_n, φ are μ-integrable and

$$\int_X \varphi d\mu \leq \liminf_{n\to\infty} \int_X \varphi_n d\mu\,.$$

Exercise 2.61 Let $\varphi_n : X \to \mathbb{R}$ be Borel functions. Prove that, if μ is finite and, for some constant M and some (Borel) function φ,

$$\left.\begin{array}{l} |\varphi_n(x)| \leq M \\ \varphi_n(x) \to \varphi(x) \end{array}\right\} \qquad \forall x \in X\,,$$

then φ_n and φ are μ-summable and

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu\,.$$

Exercise 2.62 Let $\varphi_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$\varphi_n(x) = \begin{cases} 0 & x \leq 0; \\ (x|\log x|)^{-\frac{1}{n}} & 0 < x \leq 1; \\ (x \log x)^{-n} & x > 1; \end{cases}$$

Prove that

i) φ_n is summable (with respect to the Lebesgue measure) for every $n \geq 2$;

ii) $\lim_{n\to+\infty} \int_{\mathbb{R}} \varphi_n(x)dx = 1$.

Exercise 2.63 Let $(\varphi_n)_n$ be defined by

$$\varphi_n(x) = \frac{n}{x^{3/2}} \log\left(1 + \frac{x}{n}\right), \quad x \in [0, 1].$$

Prove that

i) φ_n is summable for every $n \geq 1$;

ii) $\lim_{n\to+\infty} \int_0^1 \varphi_n(x)dx = 2$.

Exercise 2.64 Let $(\varphi_n)_n$ be defined by

$$\varphi_n(x) = \frac{n\sqrt{x}}{1 + n^2x^2}, \quad x \in [0, 1].$$

Prove that:

i) $\varphi_n(x) \leq \frac{1}{\sqrt{x}}$ for every $n \geq 1$;

ii) $\lim_{n\to+\infty} \int_0^1 \varphi_n(x)dx = 0$.

Exercise 2.65 Let $(\varphi_n)_n$ be defined by

$$\varphi_n(x) = \frac{1}{x^{3/2}} \sin\frac{x}{n}, \quad x > 0.$$

Prove that

1. φ_n is summable for every $n \geq 1$;
2. $\lim_{n\to+\infty} \int_0^{+\infty} \varphi_n(x)dx = 0$.

2.3.2 Uniform integrability

Definition 2.66 *A sequence $\varphi_n : X \to \overline{\mathbb{R}}$ of μ-summable functions is said to be* uniformly μ–summable *if for any $\varepsilon > 0$ there exists $\delta_\varepsilon > 0$ such that*

$$\mu(A) < \delta_\varepsilon \quad \Longrightarrow \quad \int_A |\varphi_n| d\mu \leq \varepsilon \qquad \forall n \in \mathbb{N}\,. \tag{2.24}$$

In other terms, $(\varphi_n)_n$ is uniformly summable iff

$$\lim_{\mu(A)\to 0} \int_A |\varphi_n| d\mu = 0 \quad \text{uniformly in } n.$$

Notice that such a property holds for a single summable function, see Proposition 2.54.

The following theorem due to Vitali uses the notion of uniform summability to provide another sufficient condition for taking limits behind the integral sign.

Theorem 2.67 (Vitali) *Let $\varphi_n : X \to \mathbb{R}$ be a sequence of uniformly μ-summable functions satisfying*

$$\forall \varepsilon > 0 \;\; \exists B_\varepsilon \in \mathcal{E} \text{ such that } \mu(B_\varepsilon) < +\infty \text{ and } \int_{B_\varepsilon^c} |\varphi_n| d\mu < \varepsilon \;\; \forall n. \tag{2.25}$$

If $(\varphi_n)_n$ converges to $\varphi : X \to \mathbb{R}$ pointwise, then φ is μ-summable and

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu\,.$$

Proof. Let $\varepsilon > 0$ be fixed and let $\delta_\varepsilon > 0$, $B_\varepsilon \in \mathcal{E}$ be such that (2.24)-(2.25) hold true. Since, by Theorem 2.25, $\varphi_n \xrightarrow{a.u.} \varphi$ in B_ε, there exists a measurable set $A_\varepsilon \subset B_\varepsilon$ such that $\mu(A_\varepsilon) < \delta_\varepsilon$ and

$$\varphi_n \to \varphi \text{ uniformly in } B_\varepsilon \setminus A_\varepsilon. \tag{2.26}$$

So,

$$\begin{aligned}\int_X |\varphi_n - \varphi| d\mu &= \int_{B_\varepsilon^c} |\varphi_n - \varphi| d\mu + \int_{A_\varepsilon} |\varphi_n - \varphi| d\mu + \int_{B_\varepsilon\setminus A_\varepsilon} |\varphi_n - \varphi| d\mu \\ &\leq \int_{B_\varepsilon^c} |\varphi_n| d\mu + \int_{B_\varepsilon^c} |\varphi| d\mu + \int_{A_\varepsilon} |\varphi_n| d\mu + \int_{A_\varepsilon} |\varphi| d\mu + \mu(B_\varepsilon) \sup_{B_\varepsilon\setminus A_\varepsilon} |\varphi_n - \varphi|\,.\end{aligned}$$

Notice that $\int_{A_\varepsilon} |\varphi_n| d\mu \le \varepsilon$, $\int_{B_\varepsilon^c} |\varphi_n| d\mu \le \varepsilon$ by (2.24)-(2.25). Also, owing to Corollary 2.41, $\int_{A_\varepsilon} |\varphi| d\mu \le \varepsilon$, $\int_{B_\varepsilon^c} |\varphi| d\mu \le \varepsilon$. Thus,

$$\int_X |\varphi_n - \varphi| d\mu \le 4\varepsilon + \mu(B_\varepsilon) \sup_{B_\varepsilon \setminus A_\varepsilon} |\varphi_n - \varphi| .$$

Since $\mu(B_\varepsilon) < +\infty$, by (2.26) we deduce

$$\int_X |\varphi_n - \varphi| d\mu \to 0. \tag{2.27}$$

Then $\varphi_n - \varphi$ is μ-summable; consequently, since $\varphi = (\varphi - \varphi_n) + \varphi_n$, by Proposition 2.51(i) φ is μ-summable. The conclusion follows by (2.17) and (2.27). □

Exercise 2.68 Derive (2.21) if $\varphi_n, \varphi : X \to \overline{\mathbb{R}}$, φ_n, φ are a.e. finite, $\varphi_n \xrightarrow{a.e.} \varphi$ and φ is Borel.

For finite measures, (2.25) is always satisfied by taking $B_\varepsilon = X$; hence Vitali's Theorem states that uniform summability is a sufficient condition to pass to the limit under the integral sign.

Corollary 2.69 *Let $\mu(X) < \infty$ and let $\varphi_n : X \to \mathbb{R}$ be a sequence of uniformly μ-summable functions converging to $\varphi : X \to \mathbb{R}$ pointwise. Then*

$$\lim_{n\to\infty} \int_X \varphi_n d\mu = \int_X \varphi d\mu .$$

Exercise 2.70 Give an example to show that when $\mu(X) = \infty$ (2.25) is an essential condition for Vitali's Theorem.
HINT: consider $\varphi_n(x) = \chi_{[n,n+1)}(x)$ in $\mathbb{R}$.

Remark 2.71 We note that property (2.25) holds for a single summable function φ. Indeed, by Proposition 2.34, the sets $\{|\varphi| > \frac{1}{n}\}$ have finite measure and, by Lebesgue's Theorem,

$$\int_{\{|\varphi| \le \frac{1}{n}\}} |\varphi| d\mu = \int_X \chi_{\{|\varphi| \le \frac{1}{n}\}} |\varphi| d\mu \to 0 \text{ as } n \to \infty.$$

Remark 2.72 We point out that Vitali's Theorem can be regarded as a generalization of Lebesgue's Theorem. Indeed, by Proposition 2.54 and Remark 2.71 it follows that properties (2.24)-(2.25) hold for a single summable function. Therefore, if $(\varphi_n)_n$ is a sequence of Borel functions satisfying (2.20) for some summable function ψ, then φ_n is uniformly summable and satisfies (2.25). The converse is not true, in general. To see this, consider the sequence $\varphi_n = n\chi_{[\frac{1}{n},\frac{1}{n}+\frac{1}{n^2})}$; since $\int_{\mathbb{R}} \varphi_n dx = \frac{1}{n}$, then $(\varphi_n)_n$ satisfies (2.24)-(2.25); on the other hand $\sup_n \varphi_n = \psi$ where $\psi = \sum_{n=1}^{+\infty} n\chi_{[\frac{1}{n},\frac{1}{n}+\frac{1}{n^2})}$ and $\int_{\mathbb{R}} \psi dx = \sum_{n=1}^{+\infty} \frac{1}{n} = \infty$; consequently the sequence $(\varphi_n)_n$ cannot be dominated by any summable function.

2.3.3 Integrals depending on a parameter

Let $(X, \mathcal{E}, \mu)$ be a σ–finite measure space. In this section we shall see how to differentiate the integral on X of a function $\varphi(x, y)$ depending on the extra variable y, which is called a parameter. We begin with a continuity result.

Proposition 2.73 *Let (Y, d) be a metric space, let $y_0 \in Y$, let U be a neighbourhood of y_0, and let*

$$\varphi : X \times Y \to \mathbb{R}$$

be a function such that

(a) $x \mapsto \varphi(x, y)$ *is Borel for every* $y \in Y$;

(b) $y \mapsto \varphi(x, y)$ *is continuous at* y_0 *for every* $x \in X$;

(c) *for some μ–summable function ψ*

$$|\varphi(x, y)| \le \psi(x) \qquad \forall x \in X\,,\ \forall y \in U$$

Then, $\Phi(y) := \int_X \varphi(x, y)\mu(dx)$ is continuous at y_0.

Proof. Let (y_n) be any sequence in Y that converges to y_0. Suppose, further, $y_n \in U$ for every $n \in \mathbb{N}$. Then,

$$\forall x \in X \qquad \begin{cases} \varphi(x, y_n) \to \varphi(x, y_0) & \text{as} \quad n \to \infty \\ |\varphi(x, y_n)| \le \psi(x) & \forall n \in \mathbb{N}\,. \end{cases}$$

Therefore, by Lebesgue's Theorem,

$$\int_X \varphi(x, y_n)\mu(dx) \quad \longrightarrow \quad \int_X \varphi(x, y_0)\mu(dx) \quad \text{as} \quad n \to \infty$$

Since (y_n) is arbitrary, the conclusion follows. □

Exercise 2.74 Let $p > 0$ be given. For $t \in \mathbb{R}$ define

$$\varphi_t(x) = \begin{cases} \frac{1}{|t|}\, x^p e^{-x/|t|} & x \in [0,1] \qquad (t \neq 0) \\ 0 & \qquad\qquad\qquad (t = 0) \end{cases}$$

For what values of p does each of the following hold true?

(a) $\varphi_t(x) \xrightarrow{a.e.} 0$ as $t \to 0$;

(b) $\varphi_t \to 0$ uniformly in $[0,1]$ as $t \to 0$;

(c) $\int_0^1 \varphi_t(x)dx \longrightarrow 0$ as $t \to 0$.

For differentiability, we shall restrict the analysis to a real parameter.

Proposition 2.75 *Assume $\varphi : X \times (a,b) \to \mathbb{R}$ satisfies the following:*

(a) *$x \mapsto \varphi(x,y)$ is Borel for every $y \in (a,b)$;*

(b) *$y \mapsto \varphi(x,y)$ is differentiable in (a,b) for every $x \in X$;*

(c) *for some μ–summable function ψ,*

$$\sup_{a<y<b} \left| \frac{\partial \varphi}{\partial y}(x,y) \right| \leq \psi(x) \qquad \forall x \in X\,.$$

Then, $\Phi(y) := \int_X \varphi(x,y)\mu(dx)$ is differentiable on (a,b) and

$$\Phi'(y) = \int_X \frac{\partial \varphi}{\partial y}(x,y)\,\mu(dx)\,, \qquad \forall y \in (a,b)\,.$$

Proof. We note, first, that $x \mapsto \frac{\partial\varphi}{\partial y}(x,y)$ is Borel for every $y \in (a,b)$ because

$$\frac{\partial\varphi}{\partial y}(x,y) = \lim_{n\to\infty} n\Big[\varphi\Big(x, y+\frac{1}{n}\Big) - \varphi(x,y)\Big] \qquad \forall (x,y) \in X \times (a,b)\,.$$

Now, fix $y_0 \in (a,b)$ and let (y_n) be any sequence in (a,b) converging to y_0. Then,

$$\frac{\Phi(y_n)-\Phi(y_0)}{y_n - y_0} = \int_X \underbrace{\frac{\varphi(x,y_n)-\varphi(x,y_0)}{y_n-y_0}}_{\stackrel{n\to\infty}{\longrightarrow} \frac{\partial\varphi}{\partial y}(x,y_0)} \mu(dx)$$

and

$$\left|\frac{\varphi(x,y_n)-\varphi(x,y_0)}{y_n-y_0}\right| \le \psi(x) \qquad \forall x \in X\,,\ \forall n \in \mathbb{N}$$

thanks to the mean value theorem. Therefore, Lebesgue's Theorem yields

$$\frac{\Phi(y_n)-\Phi(y_0)}{y_n-y_0} \quad \longrightarrow \quad \int_X \frac{\partial\varphi}{\partial y}(x,y_0)\,\mu(dx) \quad \text{as} \quad n\to\infty$$

Since (y_n) is arbitrary, the conclusion follows. □

Remark 2.76 Note that assumption (b) above must be satisfied on the whole interval (a,b) (not just a.e.) in order to be able to differentiate under the integral sign. Indeed, for $X = (a,b) = (0,1)$, let

$$\varphi(x,y) = \begin{cases} 1 & \text{if} \quad y \ge x \\ 0 & \text{if} \quad y < x\,. \end{cases}$$

Then, $\frac{\partial\varphi}{\partial y}(x,y) = 0$ for all $y \ne x$, but

$$\Phi(y) = \int_0^1 \varphi(x,y)\,dx = y \qquad \Longrightarrow \qquad \Phi'(y) = 1\,.$$

Example 2.77 Let us compute the integral

$$\Phi(y) := \int_0^\infty e^{-x^2-\frac{y^2}{x^2}}\,dx \qquad y \in \mathbb{R}\,.$$

Since $\Phi(-y) = \Phi(y)$, without loss of generality we can suppose $y \geq 0$. Observe that

$$\begin{aligned}\left|\frac{\partial}{\partial y} e^{-x^2-\frac{y^2}{x^2}}\right| &= \frac{2y}{x^2} e^{-x^2-\frac{y^2}{x^2}} \\ &= \frac{2e^{-x^2}}{y} \underbrace{\frac{y^2}{x^2} e^{-\frac{y^2}{x^2}}}_{\leq 1/e} \leq \frac{2e^{-x^2}}{r} \quad \text{for } y \geq r > 0\end{aligned}$$

Therefore, for any $y > 0$,

$$\begin{aligned}\Phi'(y) &= -\int_0^\infty \frac{2y}{x^2} e^{-x^2-\frac{y^2}{x^2}} dx \\ &\overset{t=y/x}{=} -2\int_0^\infty y \frac{t^2}{y^2} e^{-t^2-\frac{y^2}{t^2}} \frac{y}{t^2} dt = -2\Phi(y)\,.\end{aligned}$$

Since

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}\,,$$

solving the Cauchy problem

$$\begin{cases}\Phi'(y) = -2\Phi(y) \\ \Phi(0) = \frac{\sqrt{\pi}}{2}\end{cases}$$

and recalling that Φ is an even function, we obtain

$$\Phi(y) = \frac{\sqrt{\pi}}{2} e^{-2|y|} \qquad (y \in \mathbb{R})$$

□

Example 2.78 Applying Lebesgue's Theorem to counting measure, we shall compute

$$\lim_{n\to\infty} n \sum_{i=1}^\infty \sin\left(\frac{2^{-i}}{n}\right).$$

Indeed, observe that

$$\varphi_n(i) := n \sin\left(\frac{2^{-i}}{n}\right)$$

satisfies $|\varphi_n(i)| \leq 2^{-1}$. Then, by Lebesgue's Theorem we have

$$\lim_{n\to\infty} \sum_{i=1}^\infty \varphi_n(i) = \sum_{i=1}^\infty \lim_{n\to\infty} \varphi_n(i) = \sum_{i=1}^\infty 2^{-i} = 1$$

□

Exercise 2.79 Compute the integral

$$\int_0^\infty \frac{\sin x}{x}\,dx$$

proceeding as follows.

(i) Show that

$$\Phi(t) := \int_0^\infty e^{-tx}\,\frac{\sin x}{x}\,dx$$

is differentiable for all $t > 0$.

HINT: recall

$$|e^{-tx}\sin x| \le e^{-tx} \le e^{-rx} \qquad \forall t \ge r > 0\,,\ \forall x \in \mathbb{R}_+$$

(ii) Compute $\Phi'(t)$ for $t \in]0, \infty[$.

HINT: proceed as in Example 2.77 noting that

$$\int e^{-tx}\sin x = -\frac{t\sin x + \cos x}{1+t^2}\,e^{-tx}$$

(iii) Compute $\Phi(t)$ (up to an additive constant) for all $t \in]0, \infty[$.

(iv) Show that Φ continuous at 0 and conclude that

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$$

HINT: observe that, for any $\varepsilon > 0$,

$$|\Phi(t)| \le \varepsilon + \Big|\int_\varepsilon^\infty e^{-tx}\,\frac{\sin x}{x}\,dx\Big|$$

Chapter 3

L^p spaces

3.1 Spaces $\mathcal{L}^p(X, \mathcal{E}, \mu)$ and $L^p(X, \mathcal{E}, \mu)$

For any $p \in [1, \infty)$, we denote by $\mathcal{L}^p(X, \mathcal{E}, \mu)$ the class of all Borel functions $\varphi : X \to \overline{\mathbb{R}}$ such that $|\varphi|^p$ is μ–summable, and we define

$$\|\varphi\|_p = \left(\int_X |\varphi|^p d\mu\right)^{1/p} \qquad \forall \varphi \in \mathcal{L}^p(X, \mathcal{E}, \mu).$$

Remark 3.1 It is easy to check that $\mathcal{L}^p(X, \mathcal{E}, \mu)$ is closed under the following operations: sum of two functions (provided that at least one is finite everywhere) and multiplication of a function by a real number. Indeed,

$$\alpha \in \mathbb{R}, \varphi \in \mathcal{L}^p(X, \mathcal{E}, \mu) \quad \Longrightarrow \quad \alpha\varphi \in \mathcal{L}^p(X, \mathcal{E}, \mu) \ \& \ \|\alpha\varphi\|_p = |\alpha| \, \|\varphi\|_p \,.$$

Moreover, if $\varphi, \psi \in \mathcal{L}^p(X, \mathcal{E}, \mu)$ and $\varphi : X \to \mathbb{R}$, then we have

$$|\varphi(x) + \psi(x)|^p \le 2^{p-1}(|\varphi(x)|^p + |\psi(x)|^p)^{(1)} \qquad \forall x \in X\,,$$

and so $\varphi + \psi \in \mathcal{L}^p(X, \mathcal{E}, \mu)$.

Example 3.2 Let μ be the counting measure on $\mathbb{N}$. Then, we will use the notation ℓ^p for space $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$. We have

$$\ell^p = \left\{(x_n)_n \,\middle|\, x_n \in \mathbb{R}, \sum_{n=1}^{\infty} |x_n|^p < \infty\right\}.$$

(1) Since $f(t) = t^p$ is convex on $[0, \infty)$, we have that $\left|\frac{a+b}{2}\right|^p \le \frac{|a|^p + |b|^p}{2}$ for all $a, b \ge 0$.

Observe that

$$1 \le p \le q \implies \ell^p \subset \ell^q.$$

Indeed, since $\sum_n |x_n|^p < \infty$, $(x_n)_n$ is bounded, say $|x_n| \le M$ for all $n \in \mathbb{N}$. Then, $|x_n|^q \le M^{q-p}|x_n|^p$. So, $\sum_n |x_n|^q < \infty$.

Example 3.3 Consider Lebesgue measure λ on $((0,1], \mathcal{B}((0,1]))$. We use the abbreviated notation $\mathcal{L}^p(0,1)$ for space $\mathcal{L}^p((0,1], \mathcal{B}((0,1]), \lambda)$. Let us set, for any $\alpha \in \mathbb{R}$

$$\varphi_\alpha(x) = x^\alpha \qquad \forall x \in (0,1]\,.$$

Then, $\varphi_\alpha \in \mathcal{L}^p(0,1)$ iff $\alpha p + 1 > 0$. Thus, $\mathcal{L}^p(0,1)$ fails to be an algebra. For instance, $\varphi_{-1/2} \in \mathcal{L}^1(0,1)$ but $\varphi_{-1} = \varphi_{-1/2}^2 \notin \mathcal{L}^1(0,1)$.

We have already observed that $\|\cdot\|_p$ is positively homogeneous of degree one. However, $\|\cdot\|_p$ in general is not a norm[(2)] since $\|\varphi\|_p = 0$ if and only if $\varphi(x) = 0$ for a.e. $x \in X$.

In order to construct a vector space on which $\|\cdot\|_p$ is a norm, let us consider the following equivalence relation on $\mathcal{L}^p(X, \mathcal{E}, \mu)$:

$$\varphi \sim \psi \iff \varphi \overset{a.e.}{=} \psi \tag{3.1}$$

Let us denote by $L^p(X, \mathcal{E}, \mu)$ the quotient space $\mathcal{L}^p(X, \mathcal{E}, \mu)/\sim$. For any $\varphi \in \mathcal{L}^p(X, \mathcal{E}, \mu)$ we shall denote by $\widetilde{\varphi}$ the equivalence class determined by φ. It is easy to check that $L^p(X, \mathcal{E}, \mu)$ is a vector space. Indeed, the precise definition of addition of two elements $\widetilde{\varphi}_1, \widetilde{\varphi}_2 \in L^p(X, \mathcal{E}, \mu)$ is the following: let f_1, f_2 be "representatives" of $\widetilde{\varphi}_1$ and $\widetilde{\varphi}_2$ respectively, i.e. $f_1 \in \widetilde{\varphi}_1$, $f_2 \in \widetilde{\varphi}_2$, such that f_1, f_2 are finite everywhere (such representatives exist by Proposition 2.53(i)). Then $\widetilde{\varphi}_1 + \widetilde{\varphi}_2$ is the class containing $f_1 + f_2$.

We set

$$\|\widetilde{\varphi}\|_p = \|\varphi\|_p \qquad \forall \widetilde{\varphi} \in L^p(X, \mathcal{E}, \mu)\,.$$

It is easy to see that this definition is independent of the particular element φ chosen in $\widetilde{\varphi}$. Then, since the zero element of $L^p(X, \mathcal{E}, \mu)$ is the class consisting of all functions vanishing almost everywhere, it is clear that $\|\widetilde{\varphi}\|_p = 0$ iff

[(2)] Let Y be a vector space. A *norm* on Y is a mapping $Y \to [0, +\infty)$, $y \mapsto \|y\|$ such that: (i) $\|y\| = 0$ iff $y = 0$, (ii) $\|\alpha y\| = |\alpha|\,\|y\|$ for all $\alpha \in \mathbb{R}$ and $y \in Y$, (iii) $\|y_1 + y_2\| \le \|y_1\| + \|y_2\|$ for all $y_1, y_2 \in Y$. The space Y, endowed with the norm $\|\cdot\|$, is called a *normed space*. It is a metric space with the distance $d(y_1, y_2) = \|y_1 - y_2\|$, $y_1, y_2 \in Y$. If it is a complete metric space, then Y is called a *Banach space*.

$\widetilde{\varphi} = 0$. To simplify notation, we will hereafter identify $\widetilde{\varphi}$ with φ and we will talk about "functions in $L^p(X, \mathcal{E}, \mu)$" when there is no danger of confusion, with the understanding that we regard equivalent functions (i.e. functions differing only on a set of measure zero) as identical elements of the space $L^p(X, \mathcal{E}, \mu)$.

In order to check that $\|\cdot\|_p$ is a norm we need only to verify that $\|\cdot\|_p$ is sublinear. First we derive two classical inequalities that play an essential role in real analysis. Let $1 < p, q < \infty$. We say that p and q are *conjugate exponents* if

$$\frac{1}{p} + \frac{1}{q} = 1\,.$$

Proposition 3.4 (Hölder) *Let $p, q \in (1, \infty)$ be conjugate exponents. Then, for any $\varphi \in L^p(X, \mathcal{E}, \mu)$ and $\psi \in L^q(X, \mathcal{E}, \mu)$, we have that $\varphi\psi \in L^1(X, \mathcal{E}, \mu)$ and*

$$\|\varphi\psi\|_1 \le \|\varphi\|_p \, \|\psi\|_q\,. \tag{3.2}$$

Proof. The conclusion is trivial if $\|\varphi\|_p = 0$ or $\|\psi\|_q = 0$. Assume next $\|\varphi\|_p > 0$ and $\|\psi\|_q > 0$, and set

$$f(x) = \frac{|\varphi(x)|}{\|\varphi\|_p} \qquad g(x) = \frac{|\psi(x)|}{\|\psi\|_q} \qquad \forall x \in X.$$

Then, by Young's inequality (A.4),

$$f(x)g(x) \le \frac{f(x)^p}{p} + \frac{g(x)^q}{q} \qquad \forall x \in X. \tag{3.3}$$

Integrating over X with respect to μ yields

$$\frac{\int_X |\varphi\psi| d\mu}{\|\varphi\|_p \, \|\psi\|_q} = \int_X fg \, d\mu \le \frac{1}{p}\int_X f^p d\mu + \frac{1}{q}\int_X g^q d\mu = 1\,.$$

□

Remark 3.5 Suppose equality holds in (3.2). Then, equality must hold in (3.3) for a.e. $x \in X$. Therefore, recalling Example A.4, $f(x)^p = g(x)^q$ for a.e. $x \in X$. We conclude that equality holds in (3.2) iff $|\varphi(x)|^p = \alpha|\psi(x)|^q$ for a.e. $x \in X$ and some $\alpha \ge 0$.

Corollary 3.6 *Let $\mu(X) < \infty$. If $1 \le p < q$, then*

$$L^q(X, \mathcal{E}, \mu) \subset L^p(X, \mathcal{E}, \mu)$$

and

$$\|\varphi\|_p \le (\mu(X))^{\frac{1}{p}-\frac{1}{q}} \|\varphi\|_q \qquad \forall \varphi \in L^q(X, \mathcal{E}, \mu)\,. \tag{3.4}$$

Proof. By hypothesis, $|\varphi|^p \in L^{\frac{q}{p}}(X, \mathcal{E}, \mu)$. Therefore, Hölder's inequality yields

$$\int_X |\varphi|^p d\mu \le (\mu(X))^{1-\frac{p}{q}} \Big(\int_X |\varphi|^q d\mu \Big)^{\frac{p}{q}}.$$

The conclusion follows. □

Exercise 3.7 Let $\varphi_1, \varphi_2, \ldots, \varphi_k$ be functions such that

$$\varphi_i \in L^{p_i}(X, \mathcal{E}, \mu), \quad \frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_k} \le 1.$$

Then $\varphi_1 \varphi_2 \ldots \varphi_k \in L^p(X, \mathcal{E}, \mu)$ and

$$\|\varphi_1 \varphi_2 \ldots \varphi_k\|_p \le \|\varphi_1\|_{p_1} \|\varphi_2\|_{p_2} \ldots \|\varphi_k\|_{p_k}.$$

Exercise 3.8 Let $1 \le p < r < q$ and let $\varphi \in L^p(X, \mathcal{E}, \mu) \cap L^q(X, \mathcal{E}, \mu)$. Then $\varphi \in L^r(X, \mathcal{E}, \mu)$ and

$$\|\varphi\|_r \le \|\varphi\|_p^{\theta} \, \|\varphi\|_q^{1-\theta}$$

where $\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{q}$.

Proposition 3.9 (Minkowski) *Let $p \in [1, \infty)$ and let $\varphi, \psi \in L^p(X, \mathcal{E}, \mu)$. Then, $\varphi + \psi \in L^p(X, \mathcal{E}, \mu)$ and*

$$\|\varphi + \psi\|_p \le \|\varphi\|_p + \|\psi\|_p. \tag{3.5}$$

Proof. The thesis is immediate if $p = 1$. Assume $p > 1$. We have

$$\int_X |\varphi + \psi|^p d\mu \le \int_X |\varphi + \psi|^{p-1} |\varphi| d\mu + \int_X |\varphi + \psi|^{p-1} |\psi| d\mu.$$

Since $|\varphi + \psi|^{p-1} \in L^q(X, \mathcal{E}, \mu)$, where $q = \frac{p}{p-1}$, using Hölder's inequality we find

$$\int_X |\varphi + \psi|^p d\mu \le \left(\int_X |\varphi + \psi|^p d\mu \right)^{1/q} (\|\varphi\|_p + \|\psi\|_p),$$

and the conclusion follows. □

Then space $L^p(X, \mathcal{E}, \mu)$, endowed with the norm $\|\cdot\|_p$, is a normed space. Our next result shows that $L^p(X, \mathcal{E}, \mu)$ is a Banach space.

Proposition 3.10 (Riesz-Fischer) *Let $(\varphi_n)_n$ be a Cauchy sequence*[(3)] *in the normed space $L^p(X, \mathcal{E}, \mu)$. Then, a subsequence $(\varphi_{n_k})_{k\in\mathbb{N}}$ and a function φ in $L^p(X, \mathcal{E}, \mu)$ exist such that*

(i) $\varphi_{n_k} \xrightarrow{a.e.} \varphi$;

(ii) $\varphi_n \xrightarrow{L^p} \varphi$.

Proof. Since $(\varphi_n)_n$ is a Cauchy sequence in $L^p(X, \mathcal{E}, \mu)$, for any $i \in \mathbb{N}$ there exists $n_i \in \mathbb{N}$ such that

$$\|\varphi_n - \varphi_m\|_p < 2^{-i} \qquad \forall n, m \geq n_i\,. \tag{3.6}$$

Consequently, we can construct an increasing sequence n_i such that

$$\|\varphi_{n_{i+1}} - \varphi_{n_i}\|_p < 2^{-i} \qquad \forall i \in \mathbb{N}\,.$$

Next, let us set

$$g(x) = \sum_{i=1}^{\infty} |\varphi_{n_{i+1}}(x) - \varphi_{n_i}(x)|, \qquad g_k(x) = \sum_{i=1}^{k} |\varphi_{n_{i+1}}(x) - \varphi_{n_i}(x)|, \quad k \geq 1.$$

Minkowski's inequality shows that $\|g_k\|_p < 1$ for every k; since $g_k \uparrow g$, the Monotone Convergence Theorem ensures that

$$\int_X |g|^p d\mu = \lim_{k\to\infty} \int |g_k|^p d\mu \leq 1.$$

Then, owing to Proposition 2.35, g is finite a.e.; therefore the series

$$\sum_{i=1}^{\infty} (\varphi_{n_{i+1}} - \varphi_{n_i}) + \varphi_{n_1}$$

converges almost everywhere on X to some function φ. Since

$$\sum_{i=1}^{k} (\varphi_{n_{i+1}} - \varphi_{n_i}) + \varphi_{n_1} = \varphi_{n_{k+1}},$$

(3) that is for any $\varepsilon > 0$ there exists $n_\varepsilon \in \mathbb{N}$ such that $n, m > n_\varepsilon \Rightarrow \|\varphi_n - \varphi_m\|_p < \varepsilon$.

then

$$\varphi(x) = \lim_{k\to\infty} \varphi_{n_k}(x) \text{ for a.e. } x \in X.$$

Observe that φ is a Borel function; moreover, $|\varphi(x)| \le g(x) + |\varphi_{n_1}(x)|$ for a.e. $x \in X$. So, $\varphi \in L^p(X, \mathcal{E}, \mu)$. This concludes the proof of point (i).

Next, to derive (ii), fix $\varepsilon > 0$; there exists $N \in \mathbb{N}$ such that

$$\|\varphi_n - \varphi_m\|_p \le \varepsilon \quad \forall n, m \ge N.$$

Taking $m = n_k$ and passing to the limit as $k \to \infty$, Fatou's Lemma yields

$$\int_X |\varphi_n - \varphi|^p d\mu \le \liminf_{k\to\infty} \int_X |\varphi_n - \varphi_{n_k}|^p d\mu \le \varepsilon^p \qquad \forall n \ge N\,.$$

The proof is thus complete. □

Notation 3.11 *If $A \in \mathcal{B}(\mathbb{R}^N)$, we will use the abbreviated notation $L^p(A)$ for space $L^p(A, \mathcal{B}(A), \lambda)$ where λ is the Lebesgue measure.*

Example 3.12 We note that the conclusion of point (i) in Proposition 3.10 only holds for a subsequence. Indeed, given any positive integer k, consider the function

$$\varphi_i^k(x) = \begin{cases} 1 & \dfrac{i-1}{k} \le x < \dfrac{i}{k}, \\ 0 & \text{otherwise}, \end{cases} \qquad 1 \le i \le k,$$

defined on the interval $[0, 1)$. The sequence

$$\varphi_1^1, \varphi_1^2, \varphi_2^2, \ldots, \varphi_1^k, \varphi_2^k, \ldots, \varphi_k^k, \ldots$$

converges to 0 in $L^p([0, 1))$, but does not converge at any point whatsoever. Observe that the subsequence $\varphi_1^k = \chi_{[0,\frac{1}{k})}$ converges to 0 a.e.

Exercise 3.13 Generalize Exercise 2.55 showing that, if $\varphi_n \xrightarrow{L^1} \varphi$, then

$$\lim_{k\to\infty} \sup_{n\in\mathbb{N}} \int_{\{|\varphi_n|\ge k\}} |\varphi_n| d\mu = 0\,.$$

Hint: observe that

$$\begin{aligned} \int_{\{|\varphi_n|\ge 2k\}} |\varphi_n| d\mu &\le 2\int_{\{|\varphi_n-\varphi|\vee|\varphi|\ge k\}} |\varphi_n - \varphi| \vee |\varphi| d\mu \\ &\le 2\int_{\{|\varphi_n-\varphi|\ge k\}} |\varphi_n - \varphi| d\mu + 2\int_{\{|\varphi|\ge k\}} |\varphi| d\mu. \end{aligned}$$

Example 3.14 There are measure spaces $(X, \mathcal{E}, \mu)$ such that

$$L^p(X, \mathcal{E}, \mu) \not\subset L^q(X, \mathcal{E}, \mu)$$

for $p \neq q$. For instance, consider Lebesgue measure λ in $[0, 1)$ and set

$$\mu = \lambda + \sum_{n=1}^{\infty} \delta_{1/n}$$

where δ_y denotes the Dirac measure concentrated at y. Then, $\varphi(x) := x$ is in $L^2(X, \mathcal{E}, \mu) \setminus L^1(X, \mathcal{E}, \mu)$ because

$$\begin{aligned} \int_{[0,1)} x^2 d\mu &= \frac{1}{3} + \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty, \\ \int_{[0,1)} x d\mu &= \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1}{n} = \infty \,. \end{aligned}$$

On the other hand,

$$\psi(x) := \begin{cases} \frac{1}{\sqrt{x}} & \text{if} \quad x \in [0,1) \setminus \mathbb{Q} \\ 0 & \text{if} \quad x \in [0,1) \cap \mathbb{Q} \end{cases}$$

belongs to $L^1(X, \mathcal{E}, \mu) \setminus L^2(X, \mathcal{E}, \mu)$ since

$$\begin{aligned} \int_{[0,1)} \psi(x) d\mu &= \int_0^1 \frac{dx}{\sqrt{x}} < \infty, \\ \int_{[0,1)} \psi^2(x) d\mu &= \int_0^1 \frac{dx}{x} = \infty \,. \end{aligned}$$

Exercise 3.15 Show that $L^p(\mathbb{R}) \not\subset L^q(\mathbb{R})$ for $p \neq q$.
HINT: consider $f(x) = |x(\log^2 |x| + 1)|^{-1/p}$ and show that $f \in L^p(\mathbb{R})$ but $f \notin L^q(\mathbb{R})$ for $q \neq p$.

Exercise 3.16 Let $(\varphi_n)_n$ be a sequence in $L^1(X, \mathcal{E}, \mu)$. If

$$\sum_{n=1}^{\infty} \int_X |\varphi_n| d\mu < \infty \,,$$

then

$$\text{(i)} \qquad \sum_{n=1}^{\infty} |\varphi_n(x)| < \infty \quad \text{a.e.,}$$

$$\text{(ii)} \qquad \sum_{n=1}^{\infty} \varphi_n \in L^1(X, \mathcal{E}, \mu),$$

$$\text{(iii)} \qquad \sum_{n=1}^{\infty} \int_X \varphi_n d\mu = \int_X \sum_{n=1}^{\infty} \varphi_n d\mu \, .$$

Exercise 3.17 Let $1 \leq p < \infty$. Show that if $\varphi \in L^p(\mathbb{R}^N)$ and φ is uniformly continuous, then

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

HINT: if, by contradiction, $(x_n)_n \subset \mathbb{R}^N$ is such that $|x_n| \to \infty$ and $|\varphi(x_n)| \geq \delta > 0$ for every n, then the uniform continuity of φ implies the existence of $\eta > 0$ such that $|\varphi(x)| \geq \frac{\delta}{2}$ in $B(x_n, \eta)$. Show that this yields $\int_{\mathbb{R}^N} |\varphi|^p dx = \infty$.

Exercise 3.18 Show that the result in Exercise 3.17 is false in general if one only assumes that φ is continuous.
HINT: Consider

$$f_n(x) = \begin{cases} nx + 1 & \text{if} \quad -\dfrac{1}{n} \leq x \leq 0, \\ 1 - nx & \text{if} \quad 0 \leq x \leq \dfrac{1}{n}, \\ 0 & \text{if} \quad x \notin \Big(-\dfrac{1}{n}, \dfrac{1}{n}\Big), \end{cases}$$

defined on $\mathbb{R}$ and set $\varphi(x) = \sum_{n=1}^{\infty} n^{1/p} f_n(x - n)$.

3.2 Space $L^\infty(X, \mathcal{E}, \mu)$

Let $\varphi \colon X \to \overline{\mathbb{R}}$ be a Borel function. We say that φ is *essentially bounded* if there exists $M > 0$ such that $\mu(|\varphi| > M) = 0$. In this case, we set

$$\|\varphi\|_\infty = \inf\{M \geq 0 \mid \mu(|\varphi| > M) = 0\} \, . \tag{3.7}$$

We denote by $\mathcal{L}^\infty(X, \mathcal{E}, \mu)$ the class of all essentially bounded functions.

Example 3.19 The function $\varphi : (0,1] \to \mathbb{R}$ defined by

$$\varphi(x) = \begin{cases} 1 & \text{if} \quad x \neq \dfrac{1}{n} \\ \\ n & \text{if} \quad x = \dfrac{1}{n} \end{cases}$$

is essentially bounded and $\|\varphi\|_\infty = 1$.

Example 3.20 Let μ be the counting measure on $\mathbb{N}$. In the following we will use the notation ℓ^∞ for space $\mathcal{L}^\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$. We have

$$\ell^\infty = \left\{(x_n)_n \,\middle|\, x_n \in \mathbb{R}, \sup_n |x_n| < \infty\right\}.$$

Observe that

$$\ell^p \subset \ell^\infty \quad \forall p \in [1, \infty).$$

Remark 3.21 Recalling that $t \to \mu(|\varphi| > t)$ is right continuous (see Proposition 2.28), we conclude that

$$M_n \downarrow M_0 \quad \& \quad \mu(|\varphi| > M_n) = 0 \quad \Longrightarrow \quad \mu(|\varphi| > M_0) = 0\,.$$

So, the infimum in (3.7) is actually a minimum. In particular, for any $\varphi \in \mathcal{L}^\infty(X, \mathcal{E}, \mu)$,

$$|\varphi(x)| \leq \|\varphi\|_\infty \qquad \text{for a.e.} \quad x \in X. \tag{3.8}$$

In order to construct a vector space on which $\|\cdot\|_\infty$ is a norm we proceed as in the previous section defining $L^\infty(X, \mathcal{E}, \mu)$ as the quotient space of $\mathcal{L}^\infty(X, \mathcal{E}, \mu)$ modulo the equivalence relation introduced in (3.1). So, $L^\infty(X, \mathcal{E}, \mu)$ is obtained by identifying functions in $\mathcal{L}^\infty(X, \mathcal{E}, \mu)$ that coincide almost everywhere.

Exercise 3.22 Show that $L^\infty(X, \mathcal{E}, \mu)$ is a vector space and $\|\cdot\|_\infty$ is a norm in $L^\infty(X, \mathcal{E}, \mu)$.
HINT: use (3.8). For instance, for any $\alpha \neq 0$, we have $|\alpha\varphi(x)| \leq |\alpha|\,\|\varphi\|_\infty$ for a.e. $x \in X$. So, $\|\alpha\varphi\|_\infty \leq |\alpha|\,\|\varphi\|_\infty$. Hence, we also have

$$\|\varphi\|_\infty = \left\|\frac{1}{\alpha}\alpha\varphi\right\|_\infty \leq \frac{1}{|\alpha|}\,\|\alpha\varphi\|_\infty\,.$$

Thus, $\|\alpha\varphi\|_\infty = |\alpha|\,\|\varphi\|_\infty$.

Proposition 3.23 *$L^\infty(X, \mathcal{E}, \mu)$ is a Banach space.*

Proof. For a given Cauchy sequence $(\varphi_n)_n$ in $L^\infty(X, \mathcal{E}, \mu)$, let us set, for any $n, m \in \mathbb{N}$,

$$\begin{aligned} A_n &= \{|\varphi_n| > \|\varphi_n\|_\infty\}, \\ B_{m,n} &= \{|\varphi_n - \varphi_m| > \|\varphi_n - \varphi_m\|_\infty\}. \end{aligned}$$

Observe that, in view of Remark 3.21,

$$\mu(A_n) = 0 \quad \& \quad \mu(B_{m,n}) = 0 \qquad \forall m, n \in \mathbb{N}.$$

Therefore,

$$X_0 := (\cup_n A_n) \cup (\cup_{m,n} B_{m,n})$$

has measure zero and $(\varphi_n)_n$ is a Cauchy sequence for uniform convergence on X_0^c. Thus, a Borel function $\varphi : X \to \mathbb{R}$ exists such that $\varphi_n \to \varphi$ uniformly on X_0^c. This suffices to get the conclusion. □

Corollary 3.24 *Let $(\varphi_n)_n \subset L^\infty(X, \mathcal{E}, \mu)$ be such that $\varphi_n \xrightarrow{L^\infty} \varphi$. Then $\varphi_n \xrightarrow{a.e.} \varphi$.*

Exercise 3.25 Show that

$$\varphi \in L^p(X, \mathcal{E}, \mu),\ \psi \in L^\infty(X, \mathcal{E}, \mu) \quad \Longrightarrow \quad \varphi\psi \in L^p(X, \mathcal{E}, \mu)$$

and

$$\|\varphi\psi\|_p \le \|\varphi\|_p \, \|\psi\|_\infty.$$

Notation 3.26 *If $A \in \mathcal{B}(\mathbb{R}^N)$, we will use the abbreviated notation $L^\infty(A)$ for space $L^\infty(A, \mathcal{B}(A), \lambda)$ where λ is the Lebesgue measure.*

Example 3.27 It is easy to realize that spaces $L^\infty([0,1])$ and ℓ^∞ fail to be separable(4).

1. Set

$$\varphi_t(x) = \chi_{[0,t]}(x) \qquad \forall t, x \in [0,1].$$

We have

$$t \neq s \quad \Longrightarrow \quad \|\varphi_t - \varphi_s\|_\infty = 1.$$

(4) A metric space is said to be separable if it contains a countable dense subset.

Let us argue by contradiction: assume that $(\varphi_n)_n$ is a dense countable set in $L^\infty([0,1])$. Then,

$$L^\infty([0,1]) \subset \cup_n B_{1/2}(\varphi_n)^{(5)},$$

in contrast with the fact no pair of functions of the family $(\varphi_t)_{t\in[0,1]}$ belongs to the same ball $B_{1/2}(\varphi_n)$.

2. Let $(x_n)_n$ be a countable set in ℓ^∞ and define the function

$$x : \mathbb{N} \to \mathbb{R}, \quad x(k) = \begin{cases} 0 & \text{if } |x_k(k)| \geq 1, \\ 1 + x_k(k) & \text{if } |x_k(k)| < 1. \end{cases}$$

We have $x \in \ell^\infty$ and $\|x\|_\infty \leq 2$. Furthermore, for every $n \in \mathbb{N}$

$$\|x - x_n\|_\infty = \sup_k |x(k) - x_n(k)| \geq |x(n) - x_n(n)| \geq 1;$$

consequently $(x_n)_n$ is not dense in ℓ^∞.

Proposition 3.28 *Let* $p \in [1, +\infty)$ *and* $\varphi \in L^p(X, \mathcal{E}, \mu) \cap L^\infty(X, \mathcal{E}, \mu)$. *Then,*

$$\varphi \in \bigcap_{q \geq p} L^p(X, \mathcal{E}, \mu) \qquad \& \qquad \lim_{q\to\infty} \|\varphi\|_q = \|\varphi\|_\infty.$$

Proof. For $q \geq p$ we have

$$|\varphi(x)|^q \leq \|\varphi\|_\infty^{q-p} |\varphi(x)|^p \quad \text{ for a.e. } x \in X,$$

by which, after integration,

$$\|\varphi\|_q \leq \|\varphi\|_p^{\frac{p}{q}} \, \|\varphi\|_\infty^{1-\frac{p}{q}}.$$

Consequently $\varphi \in \cap_{q\geq p} L^q(X, \mathcal{E}, \mu)$ and

$$\limsup_{q\to\infty} \|\varphi\|_q \leq \|\varphi\|_\infty. \tag{3.9}$$

Conversely, let $0 < a < \|\varphi\|_\infty$ (for $\|\varphi\|_\infty = 0$ the conclusion is trivial). By Markov's inequality

$$\mu(|\varphi| > a) = \mu(|\varphi|^p > a^p) \leq a^{-p} \|\varphi\|_p^p.$$

(5) Given a metric space (Y, d), for any $y_0 \in Y$ and $r > 0$ we denote by $B_r(y_0)$ the open ball of radius r centered at y_0, i.e. $B_r(y_0) = \{y \in Y \mid d(y, y_0) < r\}$.

Consequently,

$$\|\varphi\|_p \geq a\mu(|\varphi| > a)^{1/p},$$

whence, since $\mu(|\varphi| > a) > 0$,

$$\liminf_{p\to\infty} \|\varphi\|_p \geq a.$$

Since a is any number less than $\|\varphi\|_\infty$, we conclude that

$$\liminf_{p\to\infty} \|\varphi\|_p \geq \|\varphi\|_\infty. \tag{3.10}$$

From (3.9) and (3.10) the conclusion follows. □

Corollary 3.29 *Let μ be finite and let $\varphi \in L^\infty(X, \mathcal{E}, \mu)$. Then,*

$$\varphi \in \bigcap_{p\geq 1} L^p(X, \mathcal{E}, \mu) \qquad \& \qquad \lim_{p\to\infty} \|\varphi\|_p = \|\varphi\|_\infty. \tag{3.11}$$

Proof. For $p \geq 1$ we have

$$\int_X |\varphi(x)|^p \mu(dx) \leq \mu(X)\|\varphi\|_\infty^p.$$

So, $\varphi \in \cap_p L^p(X, \mathcal{E}, \mu)$. The conclusion follows from Proposition 3.28. □

It is noteworthy that

$$\bigcap_{p\geq 1} L^p(X, \mathcal{E}, \mu) \neq L^\infty(X, \mathcal{E}, \mu)\,.$$

Exercise 3.30 Show that

$$\varphi(x) := \log x \qquad \forall x \in (0, 1]$$

belongs to $L^p((0, 1])$ for all $p \in [1, \infty)$, but $\varphi \notin L^\infty((0, 1])$.

3.3 Convergence in measure

We now present a kind of convergence for sequences of Borel functions which is of considerable importance in probability theory.

Definition 3.31 *A sequence $\varphi_n : X \to \mathbb{R}$ of Borel functions is said to* converge in measure *to a Borel function φ if for every $\varepsilon > 0$:*

$$\mu(|\varphi_n - \varphi| \geq \varepsilon) \to 0 \text{ as } n \to +\infty.$$

Let us compare the convergence in measure with other kind of convergences.

Proposition 3.32 *Let $\varphi_n, \varphi : X \to \mathbb{R}$ be Borel functions. The following holds:*

1. *If $\varphi_n \xrightarrow{a.e.} \varphi$ and $\mu(X) < +\infty$, then $\varphi_n \to \varphi$ in measure;*
2. *If $\varphi_n \to \varphi$ in measure, then there exists a subsequence $(\varphi_{n_k})_k$ such that $\varphi_{n_k} \xrightarrow{a.e.} \varphi$;*
3. *If $1 \leq p \leq +\infty$ and $\varphi_n \xrightarrow{L^p} \varphi$, then $\varphi_n \to \varphi$ in measure.*

Proof. 1. Fix $\varepsilon, \eta > 0$. According to Theorem 2.25 there exists $E \in \mathcal{E}$ such that $\mu(E) < \eta$ and $\varphi_n \to \varphi$ uniformly in $X \setminus E$. Then, for n sufficiently large

$$\{|\varphi_n - \varphi| \geq \varepsilon\} \subset E,$$

by which

$$\mu(|\varphi_n - \varphi| \geq \varepsilon) \leq \mu(E) < \eta.$$

2. For every $k \in \mathbb{N}$ we have

$$\mu\Big(|\varphi_n - \varphi| \geq \frac{1}{k}\Big) \to 0 \text{ as } n \to \infty;$$

consequently, we can construct an increasing sequence $(n_k)_k$ of positive integers such that

$$\mu\Big(|\varphi_{n_k} - \varphi| \geq \frac{1}{k}\Big) < \frac{1}{2^k} \quad \forall k \in \mathbb{N}.$$

Now set

$$A_k = \bigcup_{i=k}^{\infty} \Big\{|\varphi_{n_i} - \varphi| \geq \frac{1}{i}\Big\}, \quad A = \bigcap_{k=1}^{\infty} A_k.$$

Observe that $\mu(A_k) \le \sum_{i=k}^{\infty} \frac{1}{2^i}$ for every $k \in \mathbb{N}$. Since $A_k \downarrow A$, Proposition 1.16 implies

$$\mu(A) = \lim_{k\to\infty} \mu(A_k) = 0.$$

For any $x \in A^c$ there exists $k \in \mathbb{N}$ such that $x \in A_k^c$, that is

$$|\varphi_{n_i}(x) - \varphi(x)| < \frac{1}{i} \quad \forall i \ge k.$$

This shows that $(\varphi_{n_k})_k$ converges to φ in A^c.

3. Let $\varepsilon > 0$ be fixed. First assume $1 \le p < \infty$. Then Markov's inequality implies

$$\mu(|\varphi_n - \varphi| > \varepsilon) \le \frac{1}{\varepsilon^p} \int_X |\varphi_n - \varphi|^p d\mu \to 0 \text{ as } n \to +\infty.$$

If $p = \infty$, for large n we have $|\varphi_n - \varphi| \le \varepsilon$ a.e. in X, by which $\mu(|\varphi_n - \varphi| > \varepsilon) = 0$. □

Exercise 3.33 Show that the conclusion of Part 1 in Proposition 3.32 is false in general if $\mu(X) = \infty$.

HINT: Consider $f_n = \chi_{[n,+\infty)}$ in $\mathbb{R}$.

Example 3.34 Consider the sequence constructed in Example 3.12: it converges to 0 in $L^1([0,1))$ and, consequently, in measure. This example shows that Part 2 of Proposition 3.32 and Part (i) of Proposition 3.10 only hold for a subsequence.

Exercise 3.35 Give an example to show that the converse of Part 3 in Proposition 3.32 is not true in general.

HINT: Consider the sequence $f_n = n\chi_{[0,\frac{1}{n})}$ in $[0,1]$.

3.4 Convergence and approximation in L^p

In this section, we will exhibit techniques to derive convergence in mean of order p from a.e. convergence. Then, we will show that all elements of $L^p(X, \mathcal{E}, \mu)$ can be approximated in mean by continuous functions.

3.4.1 Convergence results

In this section we shall use the abbreviated notation $L^p(X)$ for $L^p(X, \mathcal{E}, \mu)$ when there is no danger of confusion.

The following is a direct consequence of Fatou's Lemma and Lebesgue's Theorem.

Corollary 3.36 *Let $1 \le p < \infty$ and let $(\varphi_n)_n$ be a sequence in $L^p(X)$ such that $\varphi_n \xrightarrow{a.e.} \varphi$.*

(i) *If $(\varphi_n)_n$ is bounded in $L^p(X)$, then $\varphi \in L^p(X)$ and*

$$\|\varphi\|_p \le \liminf_{n\to\infty} \|\varphi_n\|_p.$$

(ii) *If, for some $\psi \in L^p(X)$, $|\varphi_n(x)| \le \psi(x)$ for all $n \in \mathbb{N}$ and a.e. $x \in X$, then $\varphi \in L^p(X)$ and $\varphi_n \xrightarrow{L^p} \varphi$.*

Exercise 3.37 Show that, for $p = \infty$, point (i) above is still true, while (ii) is false.
HINT: consider the sequence $\varphi_n(x) = \chi_{\left(\frac{1}{n},1\right)}(x)$ in $(0,1)$.

Now, observe that, since $|\,\|\varphi_n\|_p - \|\varphi\|_p\,| \le \|\varphi_n - \varphi\|_p$, the following holds:

$$\varphi_n \xrightarrow{L^p} \varphi \quad \Longrightarrow \quad \|\varphi_n\|_p \to \|\varphi\|_p\,.$$

Then a necessary condition for convergence in $L^p(X)$ is convergence of L^p–norms. Our next result shows that, if $\varphi_n \xrightarrow{a.e.} \varphi$, such a condition is also sufficient.

Proposition 3.38 *Let $1 \le p < \infty$ and let $(\varphi_n)_n$ be a sequence in $L^p(X)$ such that $\varphi_n \xrightarrow{a.e.} \varphi$. If $\varphi \in L^p(X)$ and $\|\varphi_n\|_p \to \|\varphi\|_p$, then $\varphi_n \xrightarrow{L^p} \varphi$.*

Proof. [(6)] Define

$$\psi_n(x) = \frac{|\varphi_n(x)|^p + |\varphi(x)|^p}{2} - \left|\frac{\varphi_n(x) - \varphi(x)}{2}\right|^p \qquad \forall x \in X\,.$$

[(6)] By Novinger, 1972.

Since $p \geq 1$, a simple convexity argument shows that $\psi_n \geq 0$. Moreover, $\psi_n \xrightarrow{a.e.} |\varphi|^p$. Therefore, Fatou's Lemma yields

$$\begin{aligned}\int_X |\varphi|^p d\mu &\leq \liminf_{n\to\infty} \int_X \psi_n d\mu \\ &= \int_X |\varphi|^p d\mu - \limsup_{n\to\infty} \int_X \left|\frac{\varphi_n(x)-\varphi(x)}{2}\right|^p d\mu\,.\end{aligned}$$

So, $\limsup_n \|\varphi_n - \varphi\|_p \leq 0$, by which $\varphi_n \xrightarrow{L^p} \varphi$. □

The results below generalize Vitali's uniform summability condition, and give applications to $L^p(X)$ for $p \geq 1$. We begin by giving the following definition.

Definition 3.39 *Let $1 \leq p < \infty$. A sequence $(\varphi_n)_n$ in $L^p(X)$ is said to be* tight *if for any $\varepsilon > 0$ there exists $A_\varepsilon \in \mathcal{E}$ such that*

$$\mu(A_\varepsilon) < \infty \qquad \& \qquad \int_{A_\varepsilon^c} |\varphi_n|^p d\mu < \varepsilon \qquad \forall n \in \mathbb{N}\,.$$

Corollary 3.40 *Let $1 \leq p < \infty$ and let $(\varphi_n)_n$ be a sequence in $L^p(X)$ satisfying the following:*

(i) $\varphi_n \xrightarrow{a.e.} \varphi$

(ii) *for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\mu(A) < \delta \quad \Longrightarrow \quad \int_A |\varphi_n|^p d\mu < \varepsilon\,.$$

(iii) *$(\varphi_n)_n$ is tight.*

Then, $\varphi \in L^p(X)$ and $\varphi_n \xrightarrow{L^p} \varphi$.

Proof. Let us set $\psi_n = |\varphi_n|^p$. Then, $(\psi_n)_n$ is uniformly μ-summable, satisfies (2.25) and converges to $|\varphi|^p$ a.e. in X. Therefore, Theorem 2.67 implies $\varphi \in L^p(X)$ and

$$\|\varphi_n\|_p^p = \int_X \psi_n d\mu \quad \longrightarrow \quad \|\varphi\|_p^p.$$

The conclusion now follows from Proposition 3.38. □

Remark 3.41 If μ is finite, then, by taking $A_\varepsilon = X$ we deduce that every sequence is tight; hence, (i) and (ii) of Corollary 3.40 provide sufficient conditions for convergence in $L^p(X)$.

Corollary 3.42 *Assume $\mu(X) < \infty$. Let $1 < q < \infty$ and let $(\varphi_n)_n$ be a bounded sequence in $L^q(X)$ such that $\varphi_n \xrightarrow{a.e.} \varphi$. Then, $\varphi \in \cap_{1\le p\le q} L^p(X)$ and $\varphi_n \xrightarrow{L^p} \varphi$ for any $p \in [1, q)$.*

Proof. Let $M \ge 0$ be such that $\|\varphi_n\|_q \le M$ for any $n \in \mathbb{N}$. Point (i) of Corollary 3.36 implies $\varphi \in L^q(X)$; consequently, by Corollary 3.6, $\varphi \in \cap_{1\le p\le q} L^p(X)$. Let $1 \le p < q$: by Hölder's inequality for any $A \in \mathcal{E}$ we have

$$\int_A |\varphi_n|^p d\mu \le \left(\int_A |\varphi_n|^q d\mu\right)^{\frac{p}{q}} (\mu(A))^{1-\frac{p}{q}} \le M^p (\mu(A))^{1-\frac{p}{q}} .$$

The conclusion follows from Corollary 3.40. □

Corollary 3.43 *Assume $\mu(X) < \infty$. Let $(\varphi_n)_n$ be a sequence in $L^1(X)$ such that $\varphi_n \xrightarrow{a.e.} \varphi$ and suppose that, for some $M \ge 0$,*

$$\int_X |\varphi_n| \log^+ (|\varphi_n|)\, d\mu \le M^{(7)} \qquad \forall n \in \mathbb{N} .$$

Then, $\varphi \in L^1(X)$ and $\varphi_n \xrightarrow{L^1} \varphi$.

Proof. Fix $\varepsilon \in (0,1)$, $t \in X$, and apply estimate (A.5) with $x = \frac{1}{\varepsilon}$ and $y = \varepsilon|\varphi_n(t)|$ to obtain

$$|\varphi_n(t)| \le \ \varepsilon|\varphi_n(t)| \log(\varepsilon|\varphi_n(t)|) + e^{\frac{1}{\varepsilon}} \le \varepsilon|\varphi_n(t)| \log^+(|\varphi_n(t)|) + e^{\frac{1}{\varepsilon}}.$$

Consequently, for any $A \in \mathcal{E}$,

$$\int_A |\varphi_n| d\mu \le M\varepsilon + \mu(A)e^{\frac{1}{\varepsilon}} \qquad \forall n \in \mathbb{N}.$$

This implies that $(\varphi_n)_n$ is uniformly μ-summable. The conclusion follows from Theorem 2.67. □

Exercise 3.44 Show how Corollary 3.43 can be adapted to generic measures for tight sequences.

---

(7) Here, $\log^+(x) = (\log x) \vee 0$ for any $x \ge 0$.

3.4.2 Dense subsets of L^p

Let $\Omega \subset \mathbb{R}^N$ be an open set and denote by $\mathcal{C}_c(\Omega)$ the space of all real-valued continuous functions on Ω which are zero outside a compact set $K \subset \Omega$. Clearly, if μ be a Radon measure on $(\Omega, \mathcal{B}(\Omega))$, then

$$\mathcal{C}_c(\Omega) \subset L^p(\Omega, \mu)^{(8)} \quad \forall p \in [1, \infty].$$

Theorem 3.45 *Let $\Omega \subset \mathbb{R}^N$ be an open set and let μ be a Radon measure on $(\Omega, \mathcal{B}(\Omega))$. Then, for any $p \in [1, +\infty)$, $\mathcal{C}_c(\Omega)$ is dense in $L^p(\Omega, \mu)$.*

Proof. We begin by proving the theorem when $\Omega = \mathbb{R}^N$. We shall start imposing additional assumptions and split the reasoning into several steps, each of which will achieve a higher degree of generality.

1. Let us show how to approximate, by continuous functions with compact support, any function $\varphi \in L^p(\mathbb{R}^N, \mu)$ that satisfies, for some $M, r > 0^{(9)}$,

$$0 \leq \varphi(x) \leq M \quad x \in \mathbb{R}^N \text{ a.e.} \tag{3.12}$$

$$\varphi(x) = 0 \quad x \in \mathbb{R}^N \setminus \overline{B}_r \text{ a.e.} \tag{3.13}$$

   Let $\varepsilon > 0$. Since μ is Radon, we have $\mu(\overline{B}_r) < \infty$. Then, by Lusin's Theorem (Theorem 2.27), there exists a function $\varphi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that

$$\mu(\varphi_\varepsilon \neq \varphi) < \frac{\varepsilon}{(2M)^p} \qquad \& \qquad \|\varphi_\varepsilon\|_\infty \leq M\,.$$

   Then,

$$\int_{\mathbb{R}^N} |\varphi - \varphi_\varepsilon|^p d\mu \leq (2M)^p \mu(\varphi_\varepsilon \neq \varphi) < \varepsilon\,.$$

2. We now proceed to remove assumption (3.13). Let $\varphi \in L^p(\mathbb{R}^N, \mu)$ be a function satisfying (3.12) and fix $\varepsilon > 0$. Since $\overline{B}_n \uparrow \mathbb{R}^N$, owing to Lebesgue's Theorem, $\int_{\overline{B}_n^c} |\varphi|^p d\mu = \int_{\mathbb{R}^N} |\varphi|^p \chi_{\overline{B}_n^c} d\mu \to 0$ as $n \to \infty$. Then, there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\int_{\overline{B}_{n_\varepsilon}^c} |\varphi|^p d\mu < \varepsilon^p. \tag{3.14}$$

(8) Hereafter we shall use the abbreviated notation (Ω, μ) for measure space $(\Omega, \mathcal{B}(\Omega), \mu)$.
(9) Hereafter, $B_r = B_r(0)$.

Set $\varphi_\varepsilon := \varphi\chi_{\overline{B}_{n_\varepsilon}}$. In view of Step 1, there exists $\psi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\|\varphi_\varepsilon - \psi_\varepsilon\|_p < \varepsilon$. Then, by (3.14) we conclude that

$$\|\varphi - \psi_\varepsilon\|_p \leq \|\varphi - \varphi_\varepsilon\|_p + \|\varphi_\varepsilon - \psi_\varepsilon\|_p = \|\varphi\chi_{\overline{B}^c_{n_\varepsilon}}\|_p + \|\varphi_\varepsilon - \psi_\varepsilon\|_p < 2\varepsilon.$$

3. Next, let us dispense with the upper bound in (3.12). Since

$$0 \leq \varphi_n(x) := \min\{\varphi(x), n\} \uparrow \varphi(x) \quad x \in \mathbb{R}^N \text{ a.e.},$$

we have that $\varphi_n \xrightarrow{L^p} \varphi$. Therefore, there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\|\varphi - \varphi_{n_\varepsilon}\|_p < \varepsilon\,.$$

In view of Step 2, there exists $\psi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\|\varphi_{n_\varepsilon} - \psi_\varepsilon\|_p < \varepsilon$. Then, $\|\varphi - \psi_\varepsilon\|_p \leq \|\varphi - \varphi_{n_\varepsilon}\|_p + \|\varphi_{n_\varepsilon} - \psi_\varepsilon\|_p < 2\varepsilon$.

Finally, the extra assumption that $\varphi \geq 0$ can be disposed of applying Step 3 to φ^+ and φ^-. The proof is thus complete in the case $\Omega = \mathbb{R}^N$.

Next consider $\Omega \subset \mathbb{R}^N$ an open set and $\varphi \in L^p(\Omega, \mu)$. The function

$$\tilde{\varphi}(x) = \begin{cases} \varphi(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in \mathbb{R}^N \setminus \Omega \end{cases}$$

belongs to $L^p(\mathbb{R}^N, \tilde{\mu})$ where $\tilde{\mu}(A) = \mu(A \cap \Omega)$ for every $A \in \mathcal{B}(\mathbb{R}^N)$. Since $\tilde{\mu}$ is a Radon measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$, then there exists $\varphi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that

$$\int_{\mathbb{R}^N} |\tilde{\varphi} - \varphi_\varepsilon|^p d\tilde{\mu} \leq \varepsilon.$$

Let $(V_n)_n$ be a sequence of open sets of $\mathbb{R}^N$ such that

$$\overline{V}_n \text{ is compact}, \quad \overline{V}_n \subset V_{n+1}, \quad \cup_n V_n = \Omega \tag{3.15}$$

(for example, we can choose $V_n = B_n \cap \{x \in \Omega \,|\, d_{\Omega^c}(x) > \frac{1}{n}\}$ [10]) and set

$$\psi_n(x) = \varphi_\varepsilon(x) \frac{d_{V^c_{n+1}}(x)}{d_{V^c_{n+1}}(x) + d_{V_n}(x)}, \quad x \in \Omega.$$

[10] We recall that, given a nonempty set $S \subset \mathbb{R}^N$, $d_S(x)$ denotes the distance function of x from S, see Appendix A.1

We have $\psi_n = 0$ outside $\overline{V}_{n+1}$, by which $\psi_n \in \mathcal{C}_c(\Omega)$. Furthermore $\psi_n = \varphi_\varepsilon$ in V_n and $|\psi_n| \le \varphi_\varepsilon$; then, since $V_n \uparrow \Omega$, we deduce $\psi_n \to \varphi_\varepsilon$ in $L^p(\Omega,\mu)$. Therefore, there exists $n_\varepsilon \in \mathbb{N}$ such that

$$\int_\Omega |\varphi_\varepsilon - \psi_{n_\varepsilon}|^p d\mu < \varepsilon.$$

Then,

$$\begin{aligned}\int_\Omega |\varphi - \psi_{n_\varepsilon}|^p d\mu &\le 2^{p-1}\int_\Omega |\varphi - \varphi_\varepsilon|^p d\mu + 2^{p-1}\int_\Omega |\varphi_\varepsilon - \psi_{n_\varepsilon}|^p d\mu \\ &= 2^{p-1}\int_{\mathbb{R}^N} |\tilde{\varphi} - \varphi_\varepsilon|^p d\tilde{\mu} + 2^{p-1}\int_\Omega |\varphi_\varepsilon - \psi_{n_\varepsilon}|^p d\mu \le 2^p\varepsilon.\end{aligned}$$

□

Exercise 3.46 Given $\Omega \subset \mathbb{R}^N$ an open set, explain why $\mathcal{C}_c(\Omega)$ is not dense in $L^\infty(\Omega)$ (with respect to the Lebesgue measure), and characterize the closure of $\mathcal{C}_c(\Omega)$ in $L^\infty(\Omega)$.
HINT: show that the closure is given by the set $\mathcal{C}_0(\Omega)$ of the continuous functions $\varphi : \Omega \to \mathbb{R}$ satisfying

$$\forall \varepsilon > 0 \;\; \exists K \subset \Omega \text{ compact s.t. } \sup_{x \in \Omega \setminus K} |\varphi(x)| \le \varepsilon.$$

In particular, if $\Omega = \mathbb{R}^N$, we have

$$\mathcal{C}_0(\mathbb{R}^N) = \{\varphi : \mathbb{R}^N \to \mathbb{R} \,|\, \varphi \text{ continuous } \& \lim_{|x|\to\infty} \varphi(x) = 0\},$$

while, if Ω is bounded,

$$\mathcal{C}_0(\Omega) = \{\varphi : \Omega \to \mathbb{R} \,|\, \varphi \text{ continuous } \& \lim_{d_{\Omega^c}(x)\to 0} \varphi(x) = 0\},$$

Proposition 3.47 *Let $A \in \mathcal{B}(\mathbb{R}^N)$ and μ a Radon measure on $(A, \mathcal{B}(A))$. Then $L^p(A,\mu)$ is separable for $1 \le p < \infty$.*

Proof. First assume $\Omega = \mathbb{R}^N$. Denote by $\mathcal{R}$ the set of the rectangles in $\mathbb{R}^N$ of the form

$$R = \prod_{k=1}^N [a_k, b_k), \quad a_k, b_k \in \mathbb{Q},\;\; a_k < b_k.$$

Let $\mathcal{F}$ the vector space on $\mathbb{Q}$ generated by $(\chi_R)_{R\in\mathcal{R}}$, that is

$$\mathcal{F} = \Big\{ \sum_{i=1}^{n} c_i \chi_{R_i} \Big| n \in N, \, c_i \in \mathbb{Q}, \, R_i \in \mathcal{R} \Big\}.$$

Then $\mathcal{F}$ is countable. We are going to verify that $\mathcal{F}$ is dense in $L^p(\mathbb{R}^N, \mu)$ for $1 \le p < \infty$. Indeed, let $\varphi \in L^p(\mathbb{R}^N, \mu)$ and $\varepsilon > 0$. According to Theorem 3.45 there exists $\varphi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\|\varphi - \varphi_\varepsilon\|_p \le \varepsilon$. Let $m \in \mathbb{N}$ be sufficiently large such that, setting $Q = [-m, m)^N$, it results $\text{supp}(\varphi_\varepsilon) \subset Q$. Since μ is Radon, we have $\mu(Q) < \infty$. By the uniform continuity of φ_ε we get the existence of $\delta > 0$ such that

$$|\varphi_\varepsilon(x) - \varphi_\varepsilon(y)| \le \frac{\varepsilon}{(\mu(Q))^{1/p}} \quad \forall x, \, y \in \mathbb{R}^N \text{ s.t. } |x - y| \le \delta$$

Next split the cube Q in a finite number of disjoint cubes $Q_1, \dots, Q_n \in \mathcal{R}$ such that $diam(Q_i) \le \delta$, and define

$$\psi_\varepsilon = \sum_{i=1}^{n} c_i \chi_{Q_i}$$

where $c_i \in \mathbb{Q}$ is chosen in the interval $(\inf_{Q_i} \varphi_\varepsilon, \frac{\varepsilon}{(\mu(Q))^{1/p}} + \inf_{Q_i} \varphi_\varepsilon)$. Then $\psi_\varepsilon \in \mathcal{F}$ and $\|\varphi_\varepsilon - \psi_\varepsilon\|_\infty \le \frac{\varepsilon}{(\mu(Q))^{1/p}}$, by which we have

$$\|\varphi - \psi_\varepsilon\|_p \le \|\varphi - \varphi_\varepsilon\|_p + \|\varphi_\varepsilon - \psi_\varepsilon\|_p \le \varepsilon + (\mu(Q))^{1/p} \|\varphi_\varepsilon - \psi_\varepsilon\|_\infty \le 2\varepsilon.$$

If $A \in \mathcal{B}(\mathbb{R}^N)$, then the set

$$\mathcal{F}\big|_A = \Big\{ \sum_{i=1}^{n} c_i \chi_{R_i \cap A} \Big| n \in \mathbb{N}, \, c_i \in \mathbb{Q}, \, R_i \in \mathcal{R} \Big\}$$

is dense in $L^p(A, \mu)$. □

Remark 3.48 If $A \in \mathcal{B}(\mathbb{R})$ and μ a Radon measure on $(A, \mathcal{B}(A))$, then the set

$$\Big\{ \sum_{k=0}^{n-1} c_i \chi_{[t_k, t_{k+1}) \cap A} \Big| n \in \mathbb{N}, \, c_i, \, t_i \in \mathbb{Q}, \, t_0 < t_1 < \ldots < t_n \Big\}$$

is countable and dense in $L^p(A, \mu)$ for $1 \le p < \infty$.

Exercise 3.49 ℓ^p is separable for $1 \leq p < \infty$.

HINT: show that the set

$$\mathcal{F} = \left\{ (x_n)_n \,\middle|\, x_n \in \mathbb{Q}, \sup_{x_n \neq 0} n < \infty \right\}$$

is countable and dense in ℓ^p.

Our next result shows that the integral with respect to Lebesgue measure is translation continuous.

Proposition 3.50 *Let $p \in [1, +\infty)$ and let $\varphi \in L^p(\mathbb{R}^N)$ (with respect to the Lebesgue measure). Then,*

$$\lim_{|h| \to 0} \int_{\mathbb{R}^N} |\varphi(x+h) - \varphi(x)|^p dx = 0\,.$$

Proof. Let $\varepsilon > 0$. Theorem 3.45 ensures the existence of $\varphi_\varepsilon \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\|\varphi_\varepsilon - \varphi\|_p^p < \varepsilon$. Let $A_\varepsilon = \mathrm{supp}(\varphi_\varepsilon)$. Then, $B_\varepsilon := \{x \in \mathbb{R}^N \mid d_{A_\varepsilon}(x) \leq 1\}$ is a compact set and, since the Lebesgue's measure λ is translation invariant, for $|h| \leq 1$ we have

$$\begin{aligned} \int_{\mathbb{R}^N} |\varphi(x+h) - \varphi(x)|^p dx &\leq 3^{p-1} \int_{\mathbb{R}^N} |\varphi(x+h) - \varphi_\varepsilon(x+h)|^p dx \\ +3^{p-1} \int_{\mathbb{R}^N} |\varphi_\varepsilon(x+h) - \varphi_\varepsilon(x)|^p dx &+ 3^{p-1} \int_{\mathbb{R}^N} |\varphi_\varepsilon(x) - \varphi(x)|^p dx \\ &\leq 3^p \varepsilon + 3^{p-1} \lambda(B_\varepsilon) \sup_{|x-y| \leq |h|} |\varphi_\varepsilon(x) - \varphi_\varepsilon(y)|^p\,. \end{aligned}$$

Therefore,

$$\limsup_{|h| \to 0} \int_{\mathbb{R}^N} |\varphi(x+h) - \varphi(x)|^p dx \leq 3^p \varepsilon\,.$$

Since ε is arbitrary, the conclusion follows. □

Chapter 4

Hilbert spaces

4.1 Definitions and examples

Let H be a real vector space.

Definition 4.1 *A* scalar product $\langle\cdot,\cdot\rangle$ *in* H *is a mapping* $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{R}$ *with the following properties:*

1. $\langle x, x\rangle \geq 0$ *for all* $x \in H$ *and* $\langle x, x\rangle = 0$ *iff* $x = 0$*;*

2. $\langle x, y\rangle = \langle y, x\rangle$ *for all* $x, y \in H$*;*

3. $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$ *for all* $x, y, z \in H$ *and* $\alpha, \beta \in \mathbb{R}$.

A real pre-Hilbert space *is a pair* $(H, \langle\cdot,\cdot\rangle)$.

Remark 4.2 Since, for any $y \in H$, $0y = 0$, we have

$$\langle x, 0\rangle = 0\langle x, y\rangle = 0 \qquad \forall x \in H\,.$$

Let us set

$$\|x\| = \sqrt{\langle x, x\rangle} \qquad \forall x \in H\,. \tag{4.1}$$

The following inequality is fundamental.

Proposition 4.3 (Cauchy-Schwarz) *Let* $(H, \langle\cdot,\cdot\rangle)$ *be a pre–Hilbert space. Then*

$$|\langle x, y\rangle| \leq \|x\|\,\|y\| \qquad \forall x, y \in H \tag{4.2}$$

Moreover, equality holds iff x *and* y *are linearly dependent.*

Proof. The conclusion is trivial if $y = 0$. So, we will suppose $y \neq 0$. In fact, to begin with, let $\|y\| = 1$. Then,

$$0 \leq \|x - \langle x, y\rangle y\|^2 = \|x\|^2 - \langle x, y\rangle^2, \tag{4.3}$$

whence the conclusion follows. In the general case, it suffices to apply the above inequality to $y/\|y\|$.

If x and y are linearly dependent, then it is clear that $|\langle x, y\rangle| = \|x\|\,\|y\|$. Conversely, if $\langle x, y\rangle = \pm\|x\|\,\|y\|$ and $y \neq 0$, then (4.3) implies that x and y are linear dependent. □

Exercise 4.4 Define

$$F(\lambda) = \|x + \lambda y\|^2 = \lambda^2\|y\|^2 + 2\lambda\langle x, y\rangle + \|x\|^2 \qquad \forall \lambda \in \mathbb{R}\,.$$

Observing that $F(\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$, give an alternative proof of (4.2).

Corollary 4.5 *Let* $(H, \langle\cdot,\cdot\rangle)$ *be a pre-Hilbert space. Then the function* $\|\cdot\|$ *defined in* (4.1) *has the following properties:*

1. $\|x\| \geq 0$ *for all* $x \in H$ *and* $\|x\| = 0$ *iff* $x = 0$*;*
2. $\|\alpha x\| = |\alpha|\|x\|$ *for any* $x \in H$ *and* $\alpha \in \mathbb{R}$*;*
3. $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in H$*.*

Function $\|\cdot\|$ *is called the* norm associated with $\langle\cdot,\cdot\rangle$.

Proof. The only assertion that needs a justification is property 3. For this, observe that for all $x, y \in H$ we have, by(4.2),

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\|\,\|y\| = (\|x\| + \|y\|)^2 \qquad \square \end{aligned}$$

Remark 4.6 It is easy to see that, in a pre-Hilbert space $(H, \langle\cdot,\cdot\rangle)$, the function

$$d(x, y) = \|x - y\| \qquad \forall x, y \in H \tag{4.4}$$

is a metric.

Definition 4.7 *A pre-Hilbert space* $(H, \langle\cdot,\cdot\rangle)$ *is called an* Hilbert space *if it is complete with respect to the metric defined in* (4.4) *.*

Example 4.8 1. $\mathbb{R}^N$ is a Hilbert space with the scalar product

$$\langle x, y\rangle = \sum_{k=1}^{N} x_k y_k,$$

where $x = (x_1, \dots, x_N),\ y = (y_1, \dots, y_N) \in \mathbb{R}^N$.

2. Let $(X, \mathcal{E}, \mu)$ be a measure space. Then $L^2(X, \mathcal{E}, \mu)$, endowed with the scalar product

$$\langle \varphi, \psi\rangle = \int_X \varphi(x)\psi(x)\mu(dx), \quad \varphi, \psi \in L^2(X, \mathcal{E}, \mu),$$

is a Hilbert space (completeness follows from Proposition 3.10).

3. Let ℓ^2 be the space of all sequences of real numbers $x = (x_k)$ such that

$$\sum_{k=1}^{\infty} x_k^2 < \infty.$$

ℓ^2 is a vector space with the usual operations,

$$a(x_k) = (ax_k), \quad (x_k) + (y_k) = (x_k + y_k), \quad a \in \mathbb{R},\ (x_k), (y_k) \in \ell^2.$$

The space ℓ^2, endowed with the scalar product

$$\langle x, y\rangle = \sum_{k=1}^{\infty} x_k y_k, \quad x = (x_k), y = (y_k) \in \ell^2.$$

is a Hilbert space. This is a special case of the above example, with $X = \mathbb{N}$, $\mathcal{E} = \mathcal{P}(\mathbb{N})$, and μ given by counting measure.

Exercise 4.9 1. Show that ℓ^2 is complete arguing as follows. Take a Cauchy sequence $(x^{(n)})$ in ℓ^2, that is, $x^{(n)} = (x_k^{(n)})$.

(a) Show that, for any $k \in \mathbb{N}$, $(x_k^{(n)})_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, and deduce that the limit $x_k := \lim_{n\to\infty} x_k^{(n)}$ does exist.

(b) Show that $(x_k) \in \ell^2$.

(c) Show that $x^{(n)} \to (x_k)$ as $n \to \infty$.

2. Let $H = \mathcal{C}([-1,1])$ the linear space of all real continuous functions on $[0,1]$. Show that
   (a) H is a pre–Hilbert space with the scalar product
   $$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$$
   (b) H is not a Hilbert space.
   HINT: let
   $$f_n(t) = \begin{cases} 1 & \text{if} \quad t \in [1/n, 1] \\ nt & \text{if} \quad t \in (-1/n, 1/n) \\ -1 & \text{if} \quad t \in [-1, -1/n] \end{cases}$$
   and show that (f_n) is a Cauchy sequence in H. Observe that, if $f_n \stackrel{H}{\to} f$, then
   $$f(t) = \begin{cases} 1 & \text{if} \quad t \in (0, 1] \\ -1 & \text{if} \quad t \in [-1, 0) \end{cases}$$
3. In a pre-Hilbert space H, show that the following *parallelogram identity* holds:
   $$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad \forall x, y \in H\,. \tag{4.5}$$
   (One can prove that parallelogram identity characterizes the norms that are associated with a scalar product.)

4.2 Orthogonal projections

Let H be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$.

Definition 4.10 *Two elements x and y of H are said to be* orthogonal *if $\langle x, y\rangle = 0$. In this case, we write $x \perp y$. Two subsets A, B of H are said to be orthogonal ($A \perp B$) if $x \perp y$ for all $x \in A$ and $y \in B$.*

The following proposition is the Hilbert space version of the *Pythagorean Theorem* .

Proposition 4.11 *If $x_1, \ldots, x_n$ are pairwise orthogonal vectors in H, then*
$$\|x_1 + x_2 \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2\,.$$

Exercise 4.12 Prove Proposition 4.11

4.2.1 Projection onto a closed convex set

Definition 4.13 *A set $K \subset H$ is said to be* convex *if, for any $x, y \in K$,*

$$[x, y] := \{\lambda x + (1-\lambda)y \mid \lambda \in [0,1]\} \subset K\,.$$

For instance, any subspace of H is convex. Similarly, for any $x_0 \in H$ and $r > 0$ the ball

$$B_r(x_0) = \big\{x \in H \mid \|x - x_0\| < r\big\}$$

is a convex set. We shall also use the notation $B(x_0, r)$ to denote such a set.

Exercise 4.14 Show that, if $(K_i)_{i\in I}$ are convex subsets of H, then $\cap_i K_i$ is convex.

We know that, in a finite dimensional space, a point x has a nonempty projection onto a closed set, see Proposition A.2. The following result extends such a property to convex subsets of a Hilbert space.

Theorem 4.15 *Let $K \subset H$ be a nonempty closed convex set. Then, for any $x \in H$ there exists a unique element $y_x = p_K(x) \in K$, called the* orthogonal projection *of x onto K, such that*

$$\|x - y_x\| = \inf_{y \in K} \|x - y\|. \tag{4.6}$$

Moreover, $p_K(x)$ is the unique solution of the problem

$$\begin{cases} y \in K \\ \langle x - y, z - y\rangle \le 0 \qquad \forall z \in K. \end{cases} \tag{4.7}$$

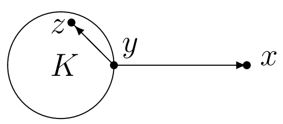

Figure 4.1: inequality (4.7) has a simple geometric meaning

Proof. Let $d = \inf_{y\in Y} \|x - y\|$. We shall split the reasoning into 4 steps.

1. Let $y_n \in K$ be a minimizing sequence, that is,

$$\|x - y_n\| \to d \quad \text{as} \quad n \to \infty \tag{4.8}$$

We claim that (y_n) is a Cauchy sequence. Indeed, for any $m, n \in Y$, parallelogram identity (4.5) yields

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

Hence, since K is convex and $\frac{y_n+y_m}{2} \in K$,

$$\begin{aligned} \|y_n - y_m\|^2 &= 2\|x-y_n\|^2 + 2\|x-y_m\|^2 - 4\left\|x - \frac{y_n+y_m}{2}\right\|^2 \\ &\leq 2\|x-y_n\|^2 + 2\|x-y_m\|^2 - 4d^2 \end{aligned}$$

So, $\|y_n - y_m\| \to 0$ as $m, n \to \infty$, as claimed.

2. Since H is complete and K is closed, (y_n) converges to some $y_x \in K$ satisfying $\|x - y_x\| = d$. The existence of y_x is thus proved.

3. We now proceed to show that (4.7) holds for any point $y \in K$ at which the infimum in (4.6) is attained. Let $z \in K$ and let $\lambda \in (0, 1]$. Since $\lambda z + (1-\lambda)y \in K$, we have that $\|x - y\| \leq \|x - y - \lambda(z-y)\|$. So,

$$\begin{aligned} 0 &\geq \frac{1}{\lambda}\left[\|x-y\|^2 - \|x-y-\lambda(z-y)\|^2\right] \\ &= 2\langle x-y, z-y\rangle - \lambda\|z-y\|^2 . \end{aligned} \tag{4.9}$$

Taking the limit as $\lambda \downarrow 0$ yields (4.6).

4. We will complete the proof showing that (4.6) has at most one solution. Let y be another solution of (4.6). Then,

$$\langle x - y_x, y - y_x\rangle \leq 0 \quad \text{and} \quad \langle x - y, y_x - y\rangle \leq 0$$

The above inequalities imply that $\|y - y_x\|^2 \leq 0$, or $y = y_x$. □

Exercise 4.16 Let $K \subset H$ be a nonempty closed convex set. Show that

$$\langle x - y, p_K(x) - p_K(y)\rangle \geq \|p_K(x) - p_K(y)\|^2 \qquad \forall x, y \in H$$

HINT: apply (4.7) to $z = p_K(x)$ and $z = p_K(y)$.

Example 4.17 In an infinite dimensional Hilbert space the projection of a point onto a closed set may be empty (in absence of convexity). Indeed, let Q consist of all sequences $x^{(n)} = (x_k^{(n)})_{k\in\mathbb{N}} \in \ell^2$ such that

$$x_k^{(n)} = \begin{cases} 0 & \text{if} \quad k \neq n \\ 1+\frac{1}{n} & \text{if} \quad k = n \end{cases} \qquad (n \geq 1)$$

Then, Q is closed. Indeed, since

$$n \neq m \quad \Longrightarrow \quad \|x^{(n)} - x^{(m)}\|_2 > \sqrt{2}\,,$$

Q has no cluster points in H. On the other hand, Q has no element of minimal norm (i.e., 0 has no projection onto Q) as well, for

$$d_Q(0) = \inf_{n\geq 1} \|x^{(n)}\|_2 = \inf_{n\geq 1} \Big(1 + \frac{1}{n}\Big) = 1\,,$$

but $\|x^{(n)}\|_2 > 1$ for every $n \geq 1$.

4.2.2 Projection onto a closed subspace

Theorem 4.15 applies, in particular, to subspaces of H. In this case, however, the variational inequality in (4.7) takes a special form.

Corollary 4.18 *Let M be a closed subspace of a Hilbert space H. Then $p_M(x)$ is the unique solution of*

$$\begin{cases} y \in M \\ \langle x - y, v\rangle = 0 \qquad \forall v \in M. \end{cases} \tag{4.10}$$

Proof. It suffices to show that (4.6) and (4.10) are equivalent when M is a subspace. If y is a solution of (4.10), then (4.6) follows taking $v = z - y$. Conversely, suppose y satisfies (4.6). Then, taking $z = y + \lambda v$ with $\lambda \in \mathbb{R}$ and $v \in M$ we obtain

$$\lambda\langle x - y, v\rangle \leq 0 \qquad \forall \lambda \in \mathbb{R}.$$

Since λ is any real number, necessarily $\langle x - y, v\rangle = 0$. □

Exercise 4.19 1. It is well known that any subspace of a finite dimensional space H is closed. Show that this is not the case if H is infinite dimensional.

HINT: consider the set of all sequences $x = (x_k) \in \ell^2$ such that $x_k = 0$ but for a finite number of subscripts k, and show that this is a dense subspace of ℓ^2.

2. Show that, if M is a closed subspace of H and $M \neq H$, then there exists $x_0 \in H \setminus \{0\}$ such that $\langle x_0, y \rangle = 0$ for all $y \in M$.

3. Let Y be a subspace of H. Show that $\overline{Y}$ is a (closed) subspace of H.

4. For any $A \subset H$ let us set

$$A^\perp = \{x \in H \mid x \perp A\}. \tag{4.11}$$

Show that, if $A, B \subset H$, then

(a) $A^\perp$ is a closed subspace of H

(b) $A \subset B \quad \Longrightarrow \quad B^\perp \subset A^\perp$

(c) $(A \cup B)^\perp = A^\perp \cap B^\perp$

$A^\perp$ is called the *orthogonal complement* of A in H.

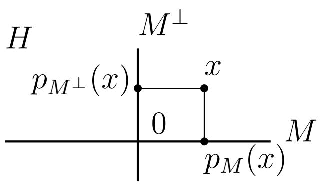

Figure 4.2: Riesz orthogonal decomposition

Proposition 4.20 *Let M be a closed subspace of a Hilbert space H. Then, the following properties hold.*

(i) For any $x \in H$ there exists a unique pair $(y_x, z_x) \in M \times M^\perp$ giving the Riesz orthogonal decompisition $x = y_x + z_x$. *Moreover,*

$$y_x = p_M(x) \qquad \textit{and} \qquad z_x = p_{M^\perp}(x) \tag{4.12}$$

(ii) $p_M : H \to H$ *is linear and* $\|p_M(x)\| \leq \|x\|$ *for all* $x \in H$.

(iii) *(a)* $p_M \circ p_M = p_M$

(b) $\ker p_M = M^\perp$

(c) $p_M(H) = M$

Proof. Let $x \in H$.

(i): define $y_x = p_M(x)$ and $z_x = x - y_x$ to obtain, by (4.10), that $z_x \perp M$ and

$$\langle x - z_x, v\rangle = \langle y_x, v\rangle = 0 \qquad \forall v \in M^\perp .$$

Therefore, $z_x = p_{M^\perp}(x)$ in view of (4.10). Suppose $x = y + z$ for some $y \in M$ and $z \in M^\perp$. Then,

$$y_x - y = z - z_x \in M \cap M^\perp = \{0\} .$$

(ii): for any $x_1, x_2 \in H$, $\alpha_1, \alpha_2 \in \mathbb{R}$ and $y \in M$, we have

$$\begin{aligned}&\langle(\alpha_1 x_1 + \alpha_2 x_2) - (\alpha_1 p_M(x_1) + \alpha_2 p_M(x_2)), y\rangle \\ &\quad = \alpha_1\langle x_1 - p_M(x_1), y\rangle + \alpha_2\langle x_2 - p_M(x_2), y\rangle = 0\end{aligned}$$

Then, by Corollary 4.18 $p_M(\alpha_1 x_1 + \alpha_2 x_2) = (\alpha_1 p_M(x_1) + \alpha_2 p_M(x_2))$. Moreover, since $\langle x - p_M(x), p_M(x)\rangle = 0$ for any $x \in H$, we obtain

$$\|p_M(x)\|^2 = \langle x, p_M(x)\rangle \leq \|x\|\,\|p_M(x)\| .$$

(iii): the first assertion follows from the fact that $p_M(x) = x$ for any $x \in Y$. The rest is a consequence of (i). □

Exercise 4.21 1. In the Hilbert space $H = L^2(0,1)$ consider sets

$$N = \left\{u \in H \;\Big|\; \int_0^1 u(x)dx = 0\right\}$$

and

$$M = \{u \in H \mid u \text{ is constant a.e. on } (0,1)\}$$

(a) Show that N and M are closed subspaces of H.

(b) Prove that $N = M^\perp$.

(c) Does $u(x) := 1/\sqrt[3]{x}, 0 < x < 1$, belong to H? If so, Find the Riesz orthogonal decomposition of u with respect to N and M.

2. For any $A \subset H$, show that the intersection of all closed linear subspaces including A is a closed linear subspace of H. Such a subspace, the so-called *closed linear subspace generated by* A, will be denoted by $\overline{\text{sp}}(A)$.

Given $A \subset H$, we will denote by $\text{sp}(A)$ the *linear subspace generated by* A, that is,

$$\text{sp}(A) = \Big\{ \sum_{k=1}^{n} c_k x_k \mid n \geq 1\,,\ c_k \in \mathbb{R}\,,\ x_k \in A \Big\}\,.$$

Exercise 4.22 Show that $\overline{\text{sp}}(A)$ is the closure of $\text{sp}(A)$.
HINT: since $\overline{\text{sp}(A)}$ is a closed subspace containing A, we have that $\overline{\text{sp}}(A) \subset \overline{\text{sp}(A)}$. Conversely, $\text{sp}(A) \subset \overline{\text{sp}}(A)$ yields $\overline{\text{sp}(A)} \subset \overline{\text{sp}}(A)$.

Corollary 4.23 *In a Hilbert space H the following properties hold.*

(i) *If M is a closed linear subspace of H, then $(M^\perp)^\perp = M$.*

(ii) *For any $A \subset H$, $(A^\perp)^\perp = \overline{\text{sp}}(A)$.*

(iii) *If N is a subspace of H, then N is dense iff $N^\perp = \{0\}$.*

Proof. We will show each point of the conclusion in sequence.

(i): from point (i) of Proposition 4.20 we deduce that

$$p_{M^\perp} = I - p_M\,.$$

Similarly, $p_{(M^\perp)^\perp} = I - p_{M^\perp} = p_M$. Thus, owing to point (iii) of the same proposition,

$$(M^\perp)^\perp = p_{(M^\perp)^\perp}(H) = p_M(H) = M\,.$$

(ii): let $M = \overline{\text{sp}}(A)$. Since $A \subset M$, we have $A^\perp \supset M^\perp$ (recall Exercise 4.19.4). So, $(A^\perp)^\perp \subset (M^\perp)^\perp = M$. Conversely, observe that A is included in the closed subspace $(A^\perp)^\perp$. So, $M \subset (A^\perp)^\perp$.

(iii): first, observe that, since $\overline{N}$ is a closed subspace, $\overline{N} = \overline{\text{sp}}(N)$. So, in view of point (ii) above,

$$\overline{N} = H \iff (N^\perp)^\perp = H \iff N^\perp = \{0\} \qquad \square$$

Exercise 4.24 1. Using Corollary 4.23 show that

$$\ell^1 := \left\{ (x_n)_{n\in\mathbb{N}} \,\middle|\, x_n \in \mathbb{R}\,,\ \sum_{n=1}^{\infty} |x_n| < \infty \right\}$$

is a dense subspace of ℓ^2.

2. Let $x, y \in H$ be linearly independent unit vectors. Show that

$$\|\lambda x + (1-\lambda)y\| < 1 \qquad \forall \lambda \in (0,1)\,.$$

HINT: observe that

$$\| \underbrace{\lambda x + (1-\lambda)y}_{x_\lambda} \|^2 = 1 + 2\lambda(1-\lambda)\big(\langle x, y\rangle - 1\big) \tag{4.13}$$

and recall the Cauchy-Schwarz inequality. (Property (4.13), recast as $\|\lambda x + (1-\lambda)y\|^2 = 1 - \lambda(1-\lambda)\|x-y\|^2$, implies that a Hilbert space is *uniformly convex*, see [3].)

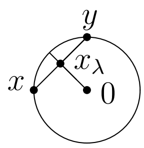

Figure 4.3: uniform convexity

4.3 The Riesz Representation Theorem

Let H be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$.

4.3.1 Bounded linear functionals

A *linear functional* F on H is a linear mapping $F : H \to \mathbb{R}$.

Definition 4.25 *A linear functional F on H is said to be* bounded *if*

$$|F(x)| \le C\|x\| \qquad \forall x \in H$$

for some constant $C \ge 0$.

Proposition 4.26 *For any linear functional F on H the following properties are equivalent.*

(a) F is continuous.

(b) F is continuous at 0.

(c) F is continuous at some point.

(d) F is bounded.

Proof. The implications (a)⇒(b)⇒(c) and (d)⇒(b) are trivial. So, it suffices to show that (c)⇒(a) and (b)⇒(d).

(c)⇒(a): let F be continuous at x_0 and let $y_0 \in H$. For any sequence (y_n) in H, converging to y_0, we have that

$$x_n = y_n - y_0 + x_0 \to x_0 \,.$$

Then, $F(x_n) = F(y_n) - F(y_0) + F(x_0) \to F(x_0)$. Therefore, $F(y_n) \to F(y_0)$. So, F is continuous at y_0.

(b)⇒(d): by hypothesis, for some $\delta > 0$ we have that $|F(x)| < 1$ for every $x \in H$ satisfying $\|x\| < \delta$. Now, let $\varepsilon > 0$ and $x \in H$. Then,

$$\Big|F\Big(\frac{\delta x}{\|x\| + \varepsilon}\Big)\Big| < 1 \,.$$

So, $|F(x)| < \frac{1}{\delta}(\|x\| + \varepsilon)$. Since ε is arbitrary, the conclusion follows. □

Definition 4.27 *The family of all bounded linear functionals on H is called the (topolgical)* dual *of H and is denoted by H^*. For any $F \in H^*$ we set*

$$\|F\|_* = \sup_{\|x\| \le 1} |F(x)| \,.$$

Exercise 4.28 1. Show that H^* is a vector space on $\mathbb{R}$, and that $\|\cdot\|_*$ is a norm in H^*.

2. For any $F \in H^*$ show that

$$\|F\|_* = \sup_{\|x\|=1} |F(x)| = \sup_{x \neq 0} \frac{|F(x)|}{\|x\|} = \inf\big\{C \ge 0 \;\big|\; |F(x)| \le C\|x\|\big\} \,.$$

4.3.2 Riesz Theorem

Example 4.29 For any fixed vector $y \in H$ define the linear functional F_y by

$$F_y(x) = \langle x, y \rangle \qquad \forall x \in H\,.$$

Then, $|F_y(x)| \le \|y\|\,\|x\|$ for any $x \in H$. So, $F_y \in H^*$ and $\|F_y\|_* \le \|y\|$. We have thus defined a map

$$\begin{cases} j : H \to H^* \\ j(y) = F_y \qquad \forall y \in H \end{cases} \tag{4.14}$$

It is easy to check that j is linear. Also, since $|F_y(y)| = \|y\|^2$ for any $y \in H$, we conclude that $\|F_y\|_* = \|y\|$ Therefore, j is a *linear isometry.*

Our next result will show that map j above is onto. So, j is an isometric isomorphism, called the *Riesz isomorphism.*

Theorem 4.30 (Riesz-Fréchet) *Let F be a bounded linear functional on H. Then there is a unique vector $y_F \in H$ such that*

$$F(x) = \langle x, y_F \rangle, \quad \forall x \in H. \tag{4.15}$$

Moreover, $\|F\|_* = \|y_F\|$.

Proof. To show the existence of a vector y satisfying (4.15), suppose $F \ne 0$ (otherwise the conclusion is trivial taking $y_F = 0$) and let $M = \ker F$. Since M is a closed proper[(1)] subspace of H, there exists $y_0 \in M^\perp \setminus \{0\}$. We can also assume, without loss of generality, that $F(y_0) = 1$. Thus, for any $x \in H$ we have that $F(x - F(x)y_0) = 0$. So, $x - F(x)y_0 \in M$. Hence, $\langle x - F(x)y_0, y_0 \rangle = 0$ or

$$F(x)\|y_0\|^2 = \langle x, y_0 \rangle \qquad \forall x \in H$$

This implies that $y_F := y_0/\|y_0\|^2$ satisfies (4.15). The rest of the conclusion follows from the fact that the map j of Example 4.29 is an isometry. □

Example 4.31 From the above theorem we deduce that, if $(X, \mathcal{E}, \mu)$ is a measure space and $F : L^2(X) \to \mathbb{R}$ is a bounded linear functional, then there exists a unique $\psi \in L^2(X)$ such that

$$F(\varphi) = \int_X \varphi\psi \, d\mu \qquad \forall \varphi \in L^2(X)\,.$$

(1) that is, $M \ne H$

A *hyperplane* Π in H is an affine subspace of codimension [(2)] 1. Given a bounded linear functional $F \in H^*$, for any $c \in \mathbb{R}$ let

$$\Pi_c = \{x \in H \mid F(x) = c\}.$$

From the proof Theorem 4.30 it follows that $\ker F = \Pi_0^\perp = \{\lambda y_F \mid \lambda \in \mathbb{R}\}$. So, Π_0 can be viewed as a closed hyperplane through the origin. Moreover, fixed any $x_c \in \Phi_c$, we have that $\Pi_c = x_c + \Pi_0$ Therefore, Π_c is a closed hyperplane in H.

Our next result provides sufficient conditions for two convex sets to be strictly separated by closed hyperplanes.

Proposition 4.32 *Let A and B be nonempty closed convex subsets of a Hilbert space H such that $A \cap B = \varnothing$. Suppose further that A is compact. Then there exist a bounded linear functional $F \in H^*$ and two constants c_1, c_2 such that*

$$F(x) \leq c_1 < c_2 \leq F(y) \qquad \forall x \in A\,,\ \forall y \in B$$

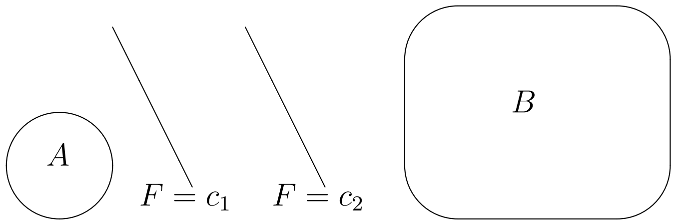

Figure 4.4: separation of convex subsets

Proof. Let $C = B - A := \big\{z \in H \mid z = y - x\,,\ x \in A\,,\ y \in B\big\}$. Then, it is easy to see that C is a nonempty convex set such that $0 \notin C$. We claim that C is closed. For let $C \ni y_n - x_n \to z$. Since A is compact, there exists a subsequence (x_{k_n}) such that $x_{k_n} \to x \in A$. Therefore,

$$y_{k_n} \underbrace{-x_{k_n} + x}_{\to 0} \to z + x =: y$$

and so $y_{k_n} \to y \in B$ since B is closed. Then, $z_0 := p_C(0)$ satisfies $z_0 \neq 0$ and

$$\langle 0 - z_0, y - x - z_0\rangle \leq 0 \qquad \forall x \in A\,,\ \forall y \in B$$

[(2)]Here, $\operatorname{codim}\Pi = \dim \Pi^\perp$.

Hence,

$$\langle x, z_0\rangle + \|z_0\|^2 \le \langle y, z_0\rangle \qquad \forall x \in A\,,\ \forall y \in B$$

and the conclusion follows taking

$$F = F_{z_0}\,, \qquad c_1 = \sup_{x\in A}\langle x, z_0\rangle\,, \qquad c_2 = \inf_{y\in B}\langle y, z_0\rangle \qquad \square$$

Exercise 4.33 Let $H = \ell^2$.

1. For $N \ge 1$ let us set $F((x_n)_n) = x_N$. Find $y \in H$ satisfying (4.15).
2. Show that, for any $x = (x_n)_n \in H$, the power series $\sum_n x_n z^n$ has radius of convergence at least 1.
3. For a given $z \in (-1, 1)$, set $F((x_n)_n) = \sum_n x_n z^n$. Find $y \in H$ representing F, and determine $\|F\|_*$.
4. Consider the sets

$$A := \big\{(x_n) \in H \mid n|x_n - n^{-2/3}| \le x_1 \quad \forall n \ge 2\big\}$$

and

$$B := \big\{(x_n) \in H \mid x_n = 0 \quad \forall n \ge 2\big\}\,.$$

   (a) Prove that A and B are disjoint closed convex subsets of H.

   (b) Show that

$$A - B = \big\{(x_n) \in H \mid \exists C \ge 0 \,:\, n|x_n - n^{-2/3}| \le C \quad \forall n \ge 2\big\}\,.$$

   (c) Deduce that $A - B$ is *dense* in H.
   Hint: fix $x = (x_n) \in H$ and define the sequence $(x^{(k)})$ in $A - B$ by

$$x_n^{(k)} = \begin{cases} x_n & \text{if } k \le n \\ 1/n^{2/3} & \text{if } k \ge n+1\,. \end{cases}$$

   (d) Prove that A and B cannot be separated by a closed hyperplane.
   Hint: otherwise $A - B$ would be included in a closed half-space.

(This example shows that the compactness assumption of Proposition 4.32 cannot be dropped.)

4.4 Orthonormal sets and bases

Let H be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$.

Definition 4.34 *A sequence* $(e_k)_{k\in\mathbb{N}}$ *is called* orthonormal *if*

$$\forall h, k \in \mathbb{N} \qquad \langle e_h, e_k \rangle = \begin{cases} 1 & \text{if} \quad h = k \\ 0 & \text{if} \quad h \neq k \end{cases}$$

Example 4.35

1. The sequence of vectors

$$e_k = (\overbrace{0, \dots, 0}^{k-1}, 1, 0, \dots) \qquad k = 1, 2 \dots$$

is orthonornal in ℓ^2.

2. Let $(e_k)_{k\in\mathbb{N}}$ be the sequence of functions in $L^2(-\pi, \pi)$ given by

$$\forall t \in [-\pi, \pi] \begin{cases} e_0(t) = \dfrac{1}{\sqrt{2\pi}} \\ e_{2j-1}(t) = \dfrac{\sin(jt)}{\sqrt{\pi}} \quad e_{2j}(t) = \dfrac{\cos(jt)}{\sqrt{\pi}} \quad (j \geq 1) \end{cases} \tag{4.16}$$

Since, for any $j, h \geq 1$,

$$\begin{aligned} \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(jt) \sin(ht)\, dt &= 0 \\ \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(jt) \sin(ht)\, dt &= \begin{cases} 0 & \text{if } j \neq h \\ 1 & \text{if } j = h \end{cases} \\ \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(jt) \cos(ht)\, dt &= \begin{cases} 0 & \text{if } j \neq h \\ 1 & \text{if } j = h\,, \end{cases} \end{aligned}$$

it is easy to check that $(e_k)_{k\in\mathbb{N}}$ is an orthonormal sequence in $L^2(-\pi, \pi)$. Such a sequence is called the *trigonometric system*.

4.4.1 Bessel's inequality

Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal sequence in H.

Proposition 4.36 *1. For any $N \in \mathbb{N}$* Bessel's identity *holds*

$$\Big\| x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k \Big\|^2 = \|x\|^2 - \sum_{k=1}^{N} \big|\langle x, e_k \rangle\big|^2 \qquad \forall x \in H \tag{4.17}$$

2. Bessel's inequality *holds*

$$\sum_{k=1}^{\infty} \big|\langle x, e_k \rangle\big|^2 \le \|x\|^2 \qquad \forall x \in H \tag{4.18}$$

In particular, the series in the left-hand side converges.

3. For any sequence $(c_k) \in \mathbb{R}$

$$\sum_{k=1}^{\infty} c_k e_k \in H \quad \Longleftrightarrow \quad \sum_{k=1}^{\infty} |c_k|^2 < \infty$$

Proof. Let $x \in H$. Bessel's identity can be easily checked by induction on N. For $N = 1$, (4.17) is true[(3)]. Suppose it holds for some $N \ge 1$. Then,

$$\begin{aligned} &\Big\| x - \sum_{k=1}^{N+1} \langle x, e_k \rangle e_k \Big\|^2 \\ &= \Big\| x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k \Big\|^2 + \big|\langle x, e_{N+1} \rangle\big|^2 - 2\Big\langle x - \sum_{k=1}^{N} \langle x, e_k \rangle e_k, \langle x, e_{N+1} \rangle e_{N+1} \Big\rangle \\ &= \|x\|^2 - \sum_{k=1}^{N} \big|\langle x, e_k \rangle\big|^2 - \big|\langle x, e_{N+1} \rangle\big|^2 \end{aligned}$$

So, (4.17) holds for any $N \ge 1$. Moreover, Bessel's identity implies that all the partial sums of the series in (4.18) are bounded above by $\|x\|^2$. So, Bessel's inequality holds as well. Finally, for all $n \in \mathbb{N}$ we have

$$\Big\| \sum_{k=n+1}^{n+p} c_k e_k \Big\|^2 = \sum_{k=n+1}^{n+p} |c_k|^2 \qquad \forall p = 1, 2, \dots$$

Therefore, Cauchy's convergence test amounts to the same condition for the two series of point 3. □

For any $x \in H$, $\langle x, e_k \rangle$ are called the *Fourier coefficients* of x, and $\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ is called the *Fourier series* of x.

(3) indeed, we used it to prove Cauchy's inequality (4.2)

Remark 4.37 Fix $n \in \mathbb{N}$ and let $M_n := \mathrm{sp}\big(\{e_1, \ldots, e_n\}\big)$. Then

$$p_{M_n}(x) = \sum_{k=1}^{n} \langle x, e_k \rangle e_k \qquad \forall x \in H$$

Indeed, for any $x \in H$ and any point $\sum_{k=1}^{n} c_k e_k \in M_n$, we have

$$\begin{aligned}
\Big\| x - \sum_{k=1}^{n} c_k e_k \Big\|^2 &= \|x\|^2 - 2 \sum_{k=1}^{n} c_k \langle x, e_k \rangle + \sum_{k=1}^{n} |c_k|^2 \\
&= \Big(\|x\|^2 - \sum_{k=1}^{n} \big|\langle x, e_k \rangle\big|^2 \Big) + \sum_{k=1}^{n} \big| c_k - \langle x, e_k \rangle \big|^2 \\
&= \Big\| x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k \Big\|^2 + \sum_{k=1}^{n} \big| c_k - \langle x, e_k \rangle \big|^2
\end{aligned}$$

thanks to Bessel's identity (4.17).

4.4.2 Orthonormal bases

To begin this section, let us characterize situations where a vector $x \in H$ is given by the sum of its Fourier series. This fact has important consequences.

Theorem 4.38 *Let $(e_k)_{k \in \mathbb{N}}$ be an orthonormal sequence in H. Then the following properties are equivalent.*

(a) $\mathrm{sp}(e_k \mid k \in \mathbb{N})$ *is dense in H.*

(b) Every $x \in H$ is given by the sum of its Fourier series, that is,

$$x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k \,.$$

(c) Every $x \in H$ satisfies Parseval's identity

$$\|x\|^2 = \sum_{k=1}^{\infty} \big|\langle x, e_k \rangle\big|^2 \,. \tag{4.19}$$

(d) If $x \in H$ and $\langle x, e_k \rangle = 0$ for every $k \in \mathbb{N}$, then $x = 0$.

Proof. We will show that $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$.

$\boxed{(a) \Rightarrow (b)}$: for any $n \in \mathbb{N}$ let $M_n := \mathrm{sp}\big(\{e_1, \ldots, e_n\}\big)$. Then, by hypothesis, $d(x, M_n) \to 0$ as $n \to \infty$ for any $x \in H$. Thus, owing to Remark 4.37,

$$\Big\|x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k \Big\|^2 = \|x - p_{M_n}(x)\|^2 = d^2(x, M_n) \to 0 \quad (n \to \infty)\,.$$

This yields (b).

$\boxed{(b) \Rightarrow (c)}$: this part of the conclusion follows from Bessel's identity.

$\boxed{(c) \Rightarrow (d)}$: obviuos.

$\boxed{(d) \Rightarrow (a)}$: let $N := \mathrm{sp}(e_k \mid k \in \mathbb{N})$. Then, $N^{\perp} = \{0\}$ owing to (d). So, N is dense on account of point (iii) of Corollary 4.23. □

Definition 4.39 *The orthonormal sequence* $(e_k)_{k \in \mathbb{N}}$ *is called* complete *if* $\mathrm{sp}(e_k \mid k \in \mathbb{N})$ *is dense in* H *(or any of the four equivalent conditions of Theorem 4.38 holds). In this case,* $(e_k)_{k \in \mathbb{N}}$ *is also said to be an* ortonormal basis *of* H*.*

Exercise 4.40 1. Prove that, if H possesses an orthonormal basis $(e_k)_{k \in \mathbb{N}}$, then H is *separable*, that is, H contains a dense countable set.

HINT: Consider all linear combinations of the e_k's with rational coefficients.

2. Let $(y_n)_{n \in \mathbb{N}}$ be a sequence in H. Show that there exists an at most countable set of *linearly independent* vectors $(x_j)_{j \in J}$ in H such that

$$\mathrm{sp}(y_n \mid n \in \mathbb{N}) = \mathrm{sp}(x_j \mid j \in J)\,.$$

HINT: for any $j = 0, 1, \ldots$, let n_j be the first integer $n \in \mathbb{N}$ such that

$$\dim \mathrm{sp}(\{y_1, \ldots, y_n\}) = j\,.$$

Set $x_j := y_{n_j}$. Then, $\mathrm{sp}(\{x_1, \ldots, x_j\}) = \mathrm{sp}(\{y_1, \ldots, y_{n_j}\}) \ldots$

3. Let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of H. Show that

$$\langle x,y\rangle = \sum_{k=1}^{\infty}\langle x,e_k\rangle\langle y,e_k\rangle \qquad \forall x,y\in H\,.$$

Hint: observe that

$$\langle x,y\rangle = \frac{\|x+y\|^2-\|x\|^2-\|y\|^2}{2}\,.$$

Our next result shows the converse of the property described in Exercise 4.40.1.

Proposition 4.41 *Let H be a separable Hilbert space. Then H possesses an orthonormal basis.*

Proof. Let $(y_n)_{n\in\mathbb{N}}$ be a dense subset of H and let $(x_j)_{j\in J}$ be linearly independent vectors such that $\overline{\mathrm{sp}}(x_j \mid j\in J) = H$ (constructed, e.g., as in Exercise 4.40.2). Define

$$e_1 = \frac{x_1}{\|x_1\|} \quad \text{and} \quad e_k = \frac{x_k - \sum_{j<k}\langle x_k,e_j\rangle e_j}{\left\|x_k - \sum_{j<k}\langle x_k,e_j\rangle e_j\right\|} \quad (k\ge 2)\,^{(4)}\,.$$

Then, (e_k) is an orthonormal sequence by construction. Moreover,

$$\mathrm{sp}(\{e_1,\dots,e_k\}) = \mathrm{sp}(\{x_1,\dots,x_k\}) \qquad \forall k\ge 1\,.$$

So, $\mathrm{sp}(e_k \mid k\ge 1)$ is dense in H. □

Example 4.42 In $H=\ell^2$, it is immediate to check that the orthonormal sequence $(e_k)_{k\in\mathbb{N}}$ of Example 4.35.1 is complete.

4.4.3 Completeness of the trigonometric system

In this section we will show that the orthonormal sequence $(e_k)_{k\in\mathbb{N}}$ defined in (4.16), that is,

$$\forall t\in[-\pi,\pi]\begin{cases} e_0(t) = \dfrac{1}{\sqrt{2\pi}} \\ e_{2j-1}(t) = \dfrac{\sin(jt)}{\sqrt{\pi}} \quad e_{2j}(t) = \dfrac{\cos(jt)}{\sqrt{\pi}} \quad (j\ge 1)\end{cases}$$

(4) This is the so-called Gram-Schmidt orthonormalization process.

is an orthonormal basis of $L^2(-\pi,\pi)$.

We begin by constructing a sequence of trigonometric polynomials with special properties. We recall that a *trigonometric polynomial* $q(t)$ is a linear combination of the above functions, i.e., an element of $\mathrm{sp}(e_k \mid k \in \mathbb{N})$. Any trigonometric polynomial q is a continuous 2π-periodic function.

Lemma 4.43 *There exists a sequence of trigonometric polynomials $(q_k)_{k\in\mathbb{N}}$ such that, for any $k \in \mathbb{N}$,*

$$\begin{cases} (a) & q_k(t) \geq 0 \qquad \forall t \in \mathbb{R} \\ (b) & \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} q_k(t)dt = 1 \\ (c) & \forall \delta > 0 \quad \displaystyle\lim_{k\to\infty} \sup_{\delta\leq|t|\leq\pi} q_k(t) = 0\,. \end{cases} \tag{4.20}$$

Proof. For any $k \in \mathbb{N}$ define

$$q_k(t) = c_k\Big(\frac{1+\cos t}{2}\Big)^k \qquad \forall t \in \mathbb{R}$$

where c_k is chosen so as to satisfy property (b). Recalling that

$$\cos(kt)\cos t \;=\; \frac{1}{2}\Big[\cos\big((k+1)t\big) + \cos\big((k-1)t\big)\Big]$$

it is easy to check that each q_k is a linear combination of $(\cos(kt))_{k\in\mathbb{N}}$. So q_k is a trigonometric polynomial.

Since (a) is immediate, it only remains to check (c). Observe that, since q_k is even,

$$\begin{aligned} 1 = \frac{c_k}{\pi}\int_0^{\pi}\Big(\frac{1+\cos t}{2}\Big)^k dt &\geq \frac{c_k}{\pi}\int_0^{\pi}\Big(\frac{1+\cos t}{2}\Big)^k \sin t\,dt \\ &= \frac{c_k}{\pi(k+1)}\Big[-2\Big(\frac{1+\cos t}{2}\Big)^{k+1}\Big]_0^{\pi} = \frac{2c_k}{\pi(k+1)} \end{aligned}$$

to conclude that

$$c_k \leq \frac{\pi(k+1)}{2} \qquad \forall k \in \mathbb{N}\,.$$

Now, fix $0 < \delta < \pi$. Since q_k is even on $[-\pi,\pi]$ and decreasing on $[0,\pi]$, using the above estimate for c_k we obtain

$$\sup_{\delta\leq|t|\leq\pi} q_k(t) = q_k(\delta) \leq \frac{\pi(k+1)}{2}\Big(\frac{1+\cos\delta}{2}\Big)^k \overset{k\to\infty}{\longrightarrow} 0\,. \qquad \square$$

Our next step is to derive a classical uniform approximation theorem by trigonometric polynomials.

Theorem 4.44 (Weierstrass) *Let f be a continuous 2π-periodic function. Then there exists a sequence of trigonometric polynomials $(p_n)_{n\in\mathbb{N}}$ such that $\|f-p_n\|_\infty \to 0$ as $n\to\infty$.*

Proof. [(5)] Let (q_n) be a sequence of trigonometric polynomials enjoying properties (4.20), e.g. the sequence given by Lemma 4.43. For any $n\in\mathbb{N}$ and $t\in\mathbb{R}$, a simple periodicity argument shows that

$$\begin{aligned} p_n(t) &:= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t-s)q_n(s)ds \\ &= \frac{1}{2\pi}\int_{t-\pi}^{t+\pi} f(\tau)q_n(t-\tau)d\tau = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)q_n(t-\tau)d\tau\,. \end{aligned}$$

This implies that p_n is a trigonometric polynomial. Indeed, if

$$q_n(t) = a_0 + \sum_{k=1}^{k_n}\Big[a_k\cos(kt)+b_k\sin(kt)\Big]\,,$$

then

$$\begin{aligned} &p_n(t) - \frac{a_0}{2\pi}\int_{-\pi}^{\pi} f(\tau)d\tau \\ &= \frac{1}{2\pi}\sum_{k=1}^{k_n}\int_{-\pi}^{\pi} f(\tau)\Big[a_k\cos\big(k(t-\tau)\big)+b_k\sin\big(k(t-\tau)\big)\Big]d\tau \\ &= \frac{1}{2\pi}\sum_{k=1}^{k_n} a_k\Big[\cos(kt)\int_{-\pi}^{\pi} f(\tau)\cos(k\tau)d\tau + \sin(kt)\int_{-\pi}^{\pi} f(\tau)\sin(k\tau)d\tau\Big] \\ &\quad + \frac{1}{2\pi}\sum_{k=1}^{k_n} b_k\Big[\sin(kt)\int_{-\pi}^{\pi} f(\tau)\cos(k\tau)d\tau - \cos(kt)\int_{-\pi}^{\pi} f(\tau)\sin(k\tau)d\tau\Big]\,. \end{aligned}$$

Next, for any $\delta>0$ let

$$\omega_f(\delta) = \sup_{|x-y|<\delta} |f(x)-f(y)|\,.$$

[(5)] This proof, based on a *convolution* method, is due to de la Vallée Poussin.

Since f is uniformly continuous, $\omega_f(\delta) \to 0$ as $\delta \to 0$. Now, for $\delta \in (0, \pi]$ properties (4.20) (a) and (b) ensure that

$$
\begin{aligned}
|f(t) - p_n(t)| &= \Big|\frac{1}{2\pi}\int_{-\pi}^{\pi}\big[f(t) - f(t-s)\big]q_n(s)ds\Big| \\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\big|f(t) - f(t-s)\big|q_n(s)ds \\
&\leq \frac{1}{2\pi}\int_{-\delta}^{\delta}\omega_f(\delta)q_n(s)ds + \frac{1}{2\pi}\int_{\delta\leq|s|\leq\pi}2\|f\|_\infty q_n(s)ds \\
&\leq \omega_f(\delta) + 2\|f\|_\infty \sup_{\delta\leq|s|\leq\pi} q_n(s)
\end{aligned}
$$

for any $t \in \mathbb{R}$. Now, fix $\varepsilon > 0$ and let $\delta_\varepsilon \in (0, \pi]$ be such that that $\omega_f(\delta_\varepsilon) < \varepsilon$. Owing to (4.20) (c), $n_\varepsilon \in \mathbb{N}$ exists such that $\sup_{\delta_\varepsilon\leq|s|\leq\pi} q_n(s) < \varepsilon$ for all $n \geq n_\varepsilon$. Thus,

$$\|f - p_n\|_\infty < (1 + 2\|f\|_\infty)\varepsilon \qquad \forall n \geq n_\varepsilon\,. \qquad \square$$

We are now ready to deduce the announced completeness of the trigonometric system. We recall that $\mathcal{C}_c(a, b)$ denotes the space of all continuous functions in (a, b) with compact support.

Theorem 4.45 *$(e_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $L^2(-\pi, \pi)$.*

Proof. We will show that trigonometric polynomials are dense in $L^2(-\pi, \pi)$. Let $f \in L^2(-\pi, \pi)$ and fix $\varepsilon > 0$. Since $\mathcal{C}_c(-\pi, \pi)$ is dense in $L^2(-\pi, \pi)$ on account of Theorem 3.45, there exists $f_\varepsilon \in \mathcal{C}_c(-\pi, \pi)$ such that $\|f - f_\varepsilon\|_2 < \varepsilon$. Clearly, we can extend f_ε, by periodicity, to a continuous function on whole real line. Also, by Weierstrass' Theorem 4.44 we can find a trigonometric polynomial p_ε such that $\|f_\varepsilon - p_\varepsilon\|_\infty < \varepsilon$. Then,

$$\|f - p_\varepsilon\|_2 \leq \|f - f_\varepsilon\|_2 + \|f_\varepsilon - p_\varepsilon\|_2 \leq \varepsilon + \varepsilon\sqrt{2\pi}\,. \qquad \square$$

Exercise 4.46 Applying (4.19) to the function

$$x(t) = t \qquad t \in [-\pi, \pi]\,,$$

derive Euler's identity

$$\sum_{k=1}^{\infty}\frac{1}{k^2} = \frac{\pi^2}{6}\,.$$

Chapter 5

Banach spaces

5.1 Definitions and examples

Let X be a real vector space.

Definition 5.1 *A* norm $\langle \cdot, \cdot \rangle$ *in* X *is a map* $\| \cdot \| : X \times H \to \mathbb{R}$ *with the following properties:*

1. $\|x\| \geq 0$ *for all* $x \in H$ *and* $\|x\| = 0$ *iff* $x = 0$*;*
2. $\|\alpha x\| = |\alpha| \|x\|$ *for any* $x \in H$ *and* $\alpha \in \mathbb{R}$*;*
3. $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in H$*.*

A normed space *is a pair* $(X, \| \cdot \|)$*.*

As we already observed in Chapter 4, in a normed space $(X, \|\cdot\|)$, the function

$$d(x, y) = \|x - y\| \qquad \forall x, y \in X \tag{5.1}$$

is a metric.

Definition 5.2 *Two norms in* X*,* $\| \cdot \|_1$ *and* $\| \cdot \|_2$*, are said to be equivalent if there exist constants* $C \geq c > 0$ *such that*

$$c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \qquad \forall x \in X \, .$$

Exercise 5.3 1. Show that two norms are equivalent if and only if they induce the same topology on X.

2. In $\mathbb{R}^n$, show that the following norms are equivalent

$$\|x\|_p = \Big(\sum_{k=1}^{n} |x_k|^p\Big)^{1/p} \quad \text{and} \quad \|x\|_\infty = \max_{1\le k\le n} |x_k|\,.$$

Definition 5.4 *A normed space* $(X, \|\cdot\|)$ *is called a* Banach space *if it is complete with respect to the metric defined in* (5.1) .

Example 5.5 1. Every Hilbert space is a Banach space.

2. Given any set $\mathcal{S} \neq \varnothing$, the family $\mathcal{B}(S)$ of all bounded functions $f : S \to \mathbb{R}$ is a vector space on $\mathbb{R}$ with the usual sum and product defined, for any $f, g \in \mathcal{B}(S)$ and $\alpha \in \mathbb{R}$, by

$$\forall x \in S \quad \begin{cases} (f+g)(x) = f(x) + g(x) \\ (\alpha f)(x) = \alpha f(x)\,. \end{cases}$$

Moreover, $\mathcal{B}(S)$ equipped with the norm

$$\|f\|_\infty = \sup_{x\in S} |f(x)| \qquad \forall f \in \mathcal{B}(S)\,,$$

is a Banach space.

3. Let (M, d) be a metric space. The family, $\mathcal{C}_b(M)$, of all bounded continuous functions on M is a *closed* subspace of $\mathcal{B}(M)$. So, $\big(\mathcal{C}_b(M), \|\cdot\|_\infty\big)$ is a Banach space.

4. Let $(X, \mathcal{E}, \mu)$ be a measure space. For any $p \in [1, \infty]$, spaces $L^p(X, \mathcal{E}, \mu)$, introduced in Chapter 3, are some of the main examples of Banach spaces with norm defined by

$$\|\varphi\|_p = \Big(\int_X |\varphi|^p d\mu\Big)^{1/p} \qquad \forall \varphi \in \mathcal{L}^p(X, \mathcal{E}, \mu)$$

for $p \in [1, \infty)$, and, for $p = \infty$, by

$$\|\varphi\|_\infty = \inf\{m \ge 0 \mid \mu(|\varphi| > m) = 0\} \qquad \forall \varphi \in \mathcal{L}^\infty(X, \mathcal{E}, \mu)\,.$$

We recall that, when μ is the counting measure on $\mathbb{N}$, we use the abbreviated notation ℓ^p for $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$. In this case we have

$$\|x\|_p = \Big(\sum_{n=1}^{\infty} |x_n|^p\Big)^{1/p} \quad \text{and} \quad \|x\|_\infty = \sup_n |x_n|\,.$$

The case of $p = 2$ was studied in Chapter 4.

Exercise 5.6 1. Let (M, d) be a locally compact metric space. Show that the set, $\mathcal{C}_0(M)$, of all functions $f \in \mathcal{C}_b(M)$ such that, for all $\varepsilon > 0$,

$$\{x \in M \mid |f(x)| \geq \varepsilon\}$$

is compact, is a closed subspace of $\mathcal{C}_b(M)$ (so, it is a Banch space).

2. Show that

$$c_0 := \{(x_n) \in \ell^\infty | \lim_{n\to\infty} x_n = 0\} \tag{5.2}$$

is a closed subspace of ℓ^∞.

3. Show that $\|\cdot\|$ (in $\mathcal{B}(S)$, $\mathcal{C}_b(M)$ or ℓ^∞) is not induced by a scalar product.

4. In a Banach space X, let (x_n) be a sequence such that $\sum_n \|x_n\| < \infty$. Show that

$$\sum_{n=1}^{\infty} x_n := \lim_{k\to\infty} \sum_{n=1}^{k} x_n \in X .$$

5.2 Bounded linear operators

Let X, Y be normed spaces. We denote by $\mathcal{L}(X, Y)$ the space of all continuous linear mappings $\Lambda : X \to Y$. The elements of $\mathcal{L}(X, Y)$ are also called *bounded operators* between X and Y. In the special case of $X = Y$, we abbreviate $\mathcal{L}(X, X) = \mathcal{L}(X)$ and any $\Lambda \in \mathcal{L}(X)$ is called a bounded operator on X. Another special case of interest is when $Y = \mathbb{R}$. As in the Hilbert space case, $\mathcal{L}(X, \mathbb{R})$ is called the topological dual of X and will be denoted by X^*. The elements of X^* are called bounded linear functionals.

Arguing exactly as in the proof of Proposition 4.26 one can show the following.

Proposition 5.7 *For any linear mapping $\Lambda : X \to Y$ the following properties are equivalent.*

(a) Λ is continuous.

(b) Λ is continuous at 0.

(c) Λ is continuous at some point.

(d) There exists $C \geq 0$ such that $\|\Lambda x\| \leq C\|x\|$ for all $x \in X$.

As in Definition 4.27, let us set

$$\|\Lambda\| = \sup_{\|x\| \leq 1} \|\Lambda x\| \qquad \forall \Lambda \in \mathcal{L}(X, Y)\,. \tag{5.3}$$

Then, for any $\Lambda \in \mathcal{L}(X, Y)$, we have

$$\|\Lambda\| = \sup_{\|x\|=1} \|\Lambda x\| = \sup_{x \neq 0} \frac{\|\Lambda x\|}{\|x\|} = \inf \left\{ C \geq 0 \;\middle|\; \|\Lambda x\| \leq C\|x\|\,,\ \forall x \in X \right\}.$$

(see also Exercise 4.28).

Exercise 5.8 Show that $\|\cdot\|$ is a norm in $\mathcal{L}(X, Y)$.

Proposition 5.9 *If Y is complete, then $\mathcal{L}(X, Y)$ is a Banach space. In particular, the topological dual of X, X^*, is a Banach space.*

Proof. Let (Λ_n) be a Cauchy sequence in $\mathcal{L}(X, Y)$. Then, for any $x \in X$, $(\Lambda_n x)$ is a Cauchy sequence in Y. Since Y is complete, $(\Lambda_n x)$ converges to a point in Y that we label Λx. We have thus defines a mapping $\Lambda : X \to Y$ which is easily checked to be linear. Moreover, since (Λ_n) is bounded in $\mathcal{L}(X, Y)$, say $\|\Lambda_n\| \leq M$ for all $n \in \mathbb{N}$, we also have that $\|\Lambda\| \leq M$. Thus, $\Lambda \in \mathcal{L}(X, Y)$. Finally, to show that $\Lambda_n \to \Lambda$ in $\mathcal{L}(X, Y)$, fix $\varepsilon > 0$ and let $n_\varepsilon \in \mathbb{N}$ be such that $\|\Lambda_n - Ł_m\| < \varepsilon$ for all $n, m \geq n_\varepsilon$. Then,

$$\|\Lambda_n x - Ł_m x\| < \varepsilon\|x - y\| \qquad \forall x \in X\,.$$

Taking the limit as $m \to \infty$, we obtain

$$\|\Lambda_n x - \Lambda x\| < \varepsilon\|x - y\| \qquad \forall x \in X\,.$$

Hence, $\|\Lambda_n - Ł\| \leq \varepsilon$ for all $n \geq n_\varepsilon$ and the proof is complete. □

Exercise 5.10 Given $f \in C([a, b])$, define $\Lambda : L^1(a, b) \to L^1(a, b)$ by

$$\Lambda g(t) = f(t)g(t) \qquad t \in [a, b]\,.$$

Show that Λ is a bounded operator and $\|\Lambda\| = \|f\|_\infty$.
HINT: $\|\Lambda\| \leq \|f\|_\infty$ follows from Hölder's inequality; to prove the equality, suppose $|f(x)| > \|f\|_\infty - \varepsilon$ for all $x \in [x_0, x_1]$ and let $g(x) = \chi_{[x_0, x_1]}$ be the characteristic function of such interval ...

5.2.1 The principle of uniform boundedness

Theorem 5.11 (Banach-Steinhaus) *Let X be a Banach space, Y be a normed space, and let $\{\Lambda_i\}_{i\in I} \subset \mathcal{L}(X,Y)$. Then,*

either *a number $M \geq 0$ exists such that*

$$\|\Lambda_i\| \leq M \qquad \forall i \in I\,, \tag{5.4}$$

or *a dense set $D \subset X$ exists such that*

$$\sup_{i\in I} \|\Lambda_i x\|_Y = \infty \qquad \forall x \in D\,. \tag{5.5}$$

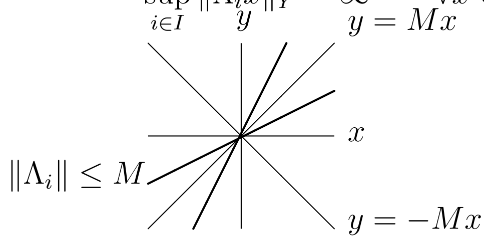

Figure 5.1: the Banach-Steinhaus Theorem

Proof. Define

$$\alpha(x) := \sup_{i\in I} \|\Lambda_i x\| \qquad \forall x \in X\,.$$

Since $\alpha : X \to [0,\infty]$ is a lower semicontinuous function, for any $n \in \mathbb{N}$

$$A_n := \{x \in X \mid \alpha(x) > n\}$$

is an open set [(1)]. If all sets A_n are dense, then (5.5) holds on $D := \cap_n A_n$ which is, in turn, a dense set owing to Baire's Lemma, see Proposition A.6. Now, suppose A_N fails to be dense for some $N \in \mathbb{N}$. Then, there exists a closed ball $\overline{B}_r(x_0) \subset X \setminus \overline{A}_N$. Therefore,

$$\|x\| \leq r \quad \Longrightarrow \quad x_0 + x \notin A_N \quad \Longrightarrow \quad \alpha(x_0 + x) \leq N\,.$$

Consequently, $\|\Lambda_i x\| \leq 2N$ for all $i \in I$ and $\|x\| \leq r$. Hence, for all $i \in I$,

$$\|\Lambda_i x\| = \frac{\|x\|}{r}\Big\|\Lambda_i \frac{rx}{\|x\|}\Big\| \leq \frac{2N}{r}\|x\| \qquad \forall x \in X \setminus \{0\}\,.$$

We have thus shown that (5.4) holds with $M = 2N/r$. □

[(1)] Alternatively, let $x \in A_n$. Then, for some $i_x \in I$, we have that $\|\Lambda_{i_x} x\| > n$. Since Λ_{i_x} is contionuous, there exists a neighbourhood V of x such that $\|\Lambda_{i_x} y\| > n$ for all $y \in V$. Thus, $\alpha(y) > n$ for all $y \in V$. So, $V \subset A_n$.

Exercise 5.12 1. Let $y = (y_n)$ be a real sequence and let $1 \le p, q \le \infty$ be conjugate exponents. Show that, if $\sum_n x_n y_n$ converges for all $x = (x_n) \in \ell^p$, then $y \in \ell^q$.

2. Let $1 \le p, q \le \infty$ be conjugate exponents, and let $f \in L^p_{loc}(\mathbb{R})$ (2). Show that, if

$$\int_{-\infty}^{\infty} f(x)g(x)dx \qquad \forall g \in L^q(\mathbb{R}),$$

then $f \in L^p(\mathbb{R})$.

5.2.2 The open mapping theorem

Bounded operators between two Banach spaces, X and Y, enjoy topological properties—closely related one another—that are very useful for applications, for instance, to differential equations. The first and main of these results is the so-called Open Mapping Theorem that we give below.

Theorem 5.13 (Schauder) *If $\Lambda \in \mathcal{L}(X, Y)$ is onto, then Λ is* open(3).

Proof. We split the reasoning into four steps.

1. Let us show that a radius $r > 0$ exists such that

$$B_{2r} \subset \overline{\Lambda(B_1)}. \tag{5.6}$$

Observe that, since Λ is onto,

$$Y = \bigcup_k \overline{\Lambda(B_k)}.$$

Therefore, by Proposition A.6 (Baire's Lemma), at least one of the closed sets $\overline{\Lambda(B_k)}$ must contain a ball, say $B_s(y) \subset \overline{\Lambda(B_k)}$. Since $\Lambda(B_k)$ is symmetric with respect to 0,

$$B_s(-y) \subset -\overline{\Lambda(B_k)} = \overline{\Lambda(B_k)}.$$

(2) We denote by $L^p_{loc}(\mathbb{R})$ the vector space of all measurable functions $f : \mathbb{R} \to \mathbb{R}$ such that $f \in L^p(a, b)$ for every interval $[a, b] \subset \mathbb{R}$.

(3) that is, U open in $X \implies \Lambda(U)$ open in Y.

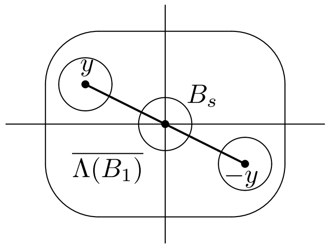

Figure 5.2: the Open Mapping Theorem

Hence, for any $x \in B_s$, we have that $x \pm y \in B_s(\pm y) \subset \overline{\Lambda(B_k)}$. Since $\overline{\Lambda(B_k)}$ is convex, we conclude that

$$x = \frac{(x+y)+(x-y)}{2} \in \overline{\Lambda(B_k)}\,.$$

Thus, $B_s \subset \overline{\Lambda(B_k)}$. Let us show how (5.6) follows with $r = s/2k$ by a rescaling argument. Indeed, let $z \in B_{2r} = B_{s/k}$. Then, $kz \in B_s$ and there exists a sequence (x_n) in B_k such that $\Lambda x_n \to kz$. So, $x_n/k \in B_1$ and $\Lambda(x_n/k) \to z$ as claimed.

2. Note that, by linearity, (5.6) yields the family of inclusions

$$B_{2^{1-n}r} \subset \overline{\Lambda(B_{2^{-n}})} \qquad \forall n \in \mathbb{N}\,. \tag{5.7}$$

3. We now proceed to show that

$$B_r \subset \Lambda(B_1)\,. \tag{5.8}$$

Let $y \in B_r$. We have to prove that $y = \Lambda x$ for some $x \in B_1$. Applying (5.7) with $n = 1$, we can find a point

$$x_1 \in B_{2^{-1}} \quad \text{such that} \quad \big\|y - \Lambda x_1\big\| < \frac{r}{2}\,.$$

Thus, $y - \Lambda x_1 \in B_{2^{-1}r}$. So, applying (5.7) with $n = 2$ we find a point

$$x_2 \in B_{2^{-2}} \quad \text{such that} \quad \big\|y - \Lambda(x_1 + x_2)\big\| < \frac{r}{2^2}\,.$$

Repeated application of this construction yield a sequence (x_n) in X such that

$$x_n \in B_{2^{-n}} \quad \text{and} \quad \big\|y - \Lambda(x_1 + \cdots + x_n)\big\| < \frac{r}{2^n}\,.$$

Since

$$\sum_{n=1}^{\infty} \|x_n\| < \sum_{n=1}^{\infty} \frac{1}{2^n} = 1\,,$$

recalling Exercise 5.6.5 we conclude that $x := \sum_n x_n \in B_1$, and, by the continuity of Λ, $\Lambda x = \sum_n \Lambda x_n = y$.

4. To conclude the proof, let $U \subset X$ be open and let $x \in U$. Then, for some $\rho > 0$, $B_\rho(x) \subset U$, whence $\Lambda x + \Lambda(B_\rho) \subset \Lambda(U)$. Therefore,

$$B_{r\rho}(\Lambda x) = \underbrace{\Lambda x + B_{r\rho} \subset \Lambda x + \Lambda(B_\rho)}_{\text{by (5.8)}} \subset \Lambda(U)\,. \qquad \square$$

The first consequence we deduce from the above result is the following Inverse Mapping Theorem.

Corollary 5.14 (Banach) *If $\Lambda \in \mathcal{L}(X,Y)$ is bijective, then Λ^{-1} is continuous.*

Proof. We have to show that, for any open set $U \subset X$, $(\Lambda^{-1})^{-1}(U)$ is open. But this follows from Theorem refth:omt since $(\Lambda^{-1})^{-1} = \Lambda$. $\square$

Exercise 5.15 1. Let $\Lambda \in \mathcal{L}(X,Y)$ be bijective. Show that a constant $\lambda > 0$ exists such that

$$\|\Lambda x\|_Y \geq \lambda \|x\|_X \qquad \forall x \in X\,. \tag{5.9}$$

Hint: use Corollary 5.14 and apply Proposition 5.7 to Λ^{-1}.

2. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on a vector space Z such that Z is complete with respect to both $\|\cdot\|_1$ and $\|\cdot\|_2$. If a constant $c \geq 0$ exists such that $\|x\|_2 \leq c\|x\|_1$ for any $x \in X$, then there also exists $C \geq 0$ such that $\|x\|_1 \leq C\|x\|_2$ for any $x \in X$ (so, $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms).

   Hint: apply (5.9) to the identity map $(Z, \|\cdot\|_1) \to (Z, \|\cdot\|_2)$.

To introduce our next result, let us observe that the Cartesian product $X \times Y$ is naturally equipped with the *product norm*

$$\|(x,y)\|_{X\times Y} := \|x\|_X + \|y\|_Y \qquad \forall (x,y) \in X \times Y\,.$$

Exercise 5.16 $\big(X \times Y, \|(\cdot,\cdot)\|_{X\times Y}\big)$ is a Banach space.

We conclude with the so-called Closed Graph Theorem.

Corollary 5.17 (Banach) *Let* $\Lambda : X \to Y$ *be a linear mapping. Then* $\Lambda \in \mathcal{L}(X,Y)$ *if and only if the* graph *of* Λ*, that is*

$$\mathrm{Graph}(\Lambda) := \big\{(x,y) \in X \times Y \;\big|\; y = \Lambda x\big\}\,,$$

is closed in $X \times Y$*.*

Proof. Suppose $\Lambda \in \mathcal{L}(X,Y)$. Then, it is easy to see that

$$\Delta : X \times Y \to Y \qquad \Delta(x,y) = y - \Lambda x$$

is continuous. Therefore, $\mathrm{Graph}(\Lambda) = \Delta^{-1}(0)$ is closed.

Conversely, let $\mathrm{Graph}(\Lambda)$ be a closed subspace of the Banach space $X \times Y$. Then, $\mathrm{Graph}(\Lambda)$ is, in turn, a Banach space with the product norm. Moreover, the linear map

$$\Pi_\Lambda : \mathrm{Graph}(\Lambda) \to X \qquad \Pi_\Lambda(x,\Lambda x) := x$$

is bounded and bijective. Therefore, owing to Corollary 5.14,

$$\Pi_\Lambda^{-1} : X \to \mathrm{Graph}(\Lambda) \qquad \Pi_\Lambda^{-1} x = (x,\Lambda x)$$

is continuous, and so is $\Lambda = \Pi_Y \circ \Pi_\Lambda^{-1}$, where

$$\Pi_Y : X \times Y \to Y \qquad \Pi_Y(x,y) := y\,. \qquad \square$$

Example 5.18 Let $X = \mathcal{C}^1([0,1])$ and $Y = \mathcal{C}([0,1])$ be both equipped with the sup norm $\|\cdot\|_\infty$. Define

$$\Lambda x(t) = x'(t) \qquad \forall x \in X\,,\ \forall t \in [0,1]\,.$$

Then $\mathrm{Graph}(\Lambda)$ is closed since

$$\begin{cases} x_n \xrightarrow{L^\infty} x_\infty \\ x'_k \xrightarrow{L^\infty} y_\infty \end{cases} \qquad \Longrightarrow \qquad x_\infty \in \mathcal{C}^1([0,1]) \quad \& \quad x'_\infty = y_\infty\,.$$

On the other hand, Λ fails to be a bounded operator. Indeed, taking

$$x_n(t) = t^n \qquad \forall t \in [0,1]\,,$$

we have that

$$x_n \in X\,, \qquad \|x_n\|_\infty = 1\,, \qquad \|\Lambda x_n\|_\infty = n \qquad \forall n \geq 1\,.$$

This shows the necessity of X being a Banach space in Thorem 5.13.

Exercise 5.19 1. For a given operator $\Lambda \in \mathcal{L}(X,Y)$ show that the following properties are equivalent:

(a) there exists $c > 0$ such that $\|\Lambda x\| \geq c\|x\|$ for all $x \in X$;

(b) $\ker \Lambda = \{0\}$ and $\Lambda(X)$ is closed in Y.

Hint: use Exercise 5.15.1.

2. Let H be a Hilbert space and let $A, B : H \to H$ be two linear maps such that

$$\langle Ax, y\rangle = \langle x, By\rangle \qquad \forall x, y \in H\,. \tag{5.10}$$

Show that $A, B \in \mathcal{L}(H)$.

Hint: use (5.10) to deduce that Graph(A) and Graph(B) are closed in $X \times X$; then, apply Corollary 5.17.

3. Let $(X, \|\cdot\|)$ be an infinite dimensional separable Banach space and let $(e_i)_{i\in I}$ be a *Hamel basis* of X [4] such that $\|e_i\| = 1$ for all $i \in I$.

(a) Show that I is uncountable.

Hint: use Baire's Lemma.

(b) Prove that

$$\|x\|_1 = \sum_{i\in I} |\lambda_i| \quad \text{if} \quad x = \sum_{i\in I} \lambda_i e_i$$

is a norm in X and that $\|x\| \leq \|x\|_1$ for every $x \in X$ (observe that both series above are finite sums).

(c) Show that X is not complete with respect to $\|\cdot\|_1$.

Hint: should $(X, \|\cdot\|_1)$ be a Banach space, then $\|\cdot\|$ and $\|\cdot\|_1$ would be equivalent norms by Exercise 5.15.2, but, for any $i \neq j$, we have $\|e_i - e_j\|_1 = \ldots$

[4] that is, a maximal linearly independent subset of X. Let us recall that, applying Zorn's Lemma, one can show that any linearly independent subset of X can be completed to a Hamel basis. Moreover, given a Hamel basis $(e_i)_{i\in I}$, we have that $X = \mathrm{sp}\{e_i \mid i \in I\}$.

5.3 Bounded linear functionals

In this section we shall study a special case of bounded linear operators, namely $\mathbb{R}$-valued operators or—as we usually say—*bounded linear functionals.* We will see that functionals enjoy an important extension property described by the Hahn-Banach Theorem. Then we will derive useful analytical and geometric consequences of such a property. These results will be essential for the analysis of dual spaces that we shall develop in the next section. Finally, we will characterize the duals of the Banach spaces ℓ^p.

5.3.1 The Hahn-Banach Theorem

Let us consider the following extension problem: given a subspace $M \subset X$ (not necessarily closed) and a continuous linear functional $f : M \to \mathbb{R}$,

$$\text{find} \quad F \in X^* \quad \text{such that} \quad \begin{cases} F\big|_M = f \\ \|F\| = \|f\| . \end{cases} \tag{5.11}$$

Remark 5.20 1. Observe that a bounded linear functional f defined on a subspace M can be extended to the closure $\overline{M}$ by a standard completeness argument. For let $\overline{x} \in \overline{M}$ and let $(x_n) \subset M$ be such that $x_n \to \overline{x}$. Since

$$|f(x_n) - f(x_m)| \leq \|f\| \, \|x_n - x_m\| ,$$

$(f(x_n))$ is a Cauchy sequence in $\mathbb{R}$. Therefore, $(f(x_n))$ converges. Then, it is easy to see that $F(\overline{x}) := \lim_n f(x_n)$ is the required extension of f. So, the problem of finding an extension of f satifying (5.11) has a *unique solution* when M is dense in X.

2. Another case where the extension satifying (5.11) is unique is when X is a Hilbert space. Indeed, let us still denote by f the extension of the given functional to $\overline{M}$, obtained by the procedure described at point 1. Note that $\overline{M}$ is also a Hilbert space. So, by the Riesz-Fréchet Theorem, there exists a unique vector $y_f \in \overline{M}$ such that $\|x_f\| = \|f\|$ and

$$f(x) = \langle x, y_f \rangle \qquad \forall x \in \overline{M} .$$

Define

$$F(x) = \langle x, y_f \rangle \qquad \forall x \in X .$$

Then, $f \in X^*$ satisfies (5.11) and $\|F\| = \|y_f\| = \|f\|$. We claim that F is the *unique extension* of f with these properties. For let G be another bounded linear functional satifying (5.11) and let y_G be the vector in X associated with G in the Riesz representation of G. Consider the Riesz orthogonal decomposition of y_G, that is,

$$y_G = y_G' + y_G'' \quad \text{where} \quad y_G' \in \overline{M} \quad \text{and} \quad y_G'' \perp \overline{M}\,.$$

Then,

$$\langle x, y_G' \rangle = G(x) = f(x) = \langle x, y_f \rangle \qquad \forall x \in \overline{M}\,.$$

So, $y_G' = y_f$. Moreover,

$$\|y_G''\|^2 = \|y_G\|^2 - \|y_G'\|^2 = \|f\|^2 - \|y_f\|^2 = 0\,.$$

In general, the following classical result ensures the existence of an extension of f satisfying (5.11) even though its uniqueness is no longer guaranteed.

Theorem 5.21 (Hahn-Banach: first analytic form) *Let $(X, \|\cdot\|)$ be a normed space and let M be a subspace of X. If $f : M \to \mathbb{R}$ is a continuous linear functional on M, then there is a functional $F \in X^*$ such that*

$$F\big|_M = f \quad \text{and} \quad \|F\| = \|f\|\,.$$

Proof. To begin with, let us suppose that $\|f\| \neq 0$ for otherwise one can take $F \equiv 0$ and the conclusion becomes trivial. Then we can assume, without loss of generality, that $\|f\| = 1$. We will show, first, how to extend f to a subspace of X which *strictly* includes M. The general case will be treated later—in steps 2 and 3—by a maximality argument.

1. Suppose $M \neq X$ and let $x_0 \in X \setminus M$. Let us construct an extension of f to the subspace

$$M_0 := \text{sp}(M \cup \{x_0\}) = \{x + \lambda x_0 \mid x \in M\,,\ \lambda \in \mathbb{R}\}\,.$$

Define

$$f_0(x + \lambda x_0) := f(x) + \lambda\alpha \qquad \forall x \in M\,,\ \forall \lambda \in \mathbb{R}\,, \tag{5.12}$$

where α is a real number to be fixed. Clearly, f_0 is a linear functional on M_0 that extends f. We must find $\alpha \in \mathbb{R}$ such that

$$|f_0(x + \lambda x_0)| \leq \|x + \lambda x_0\| \qquad \forall x \in M\,,\ \forall \lambda \in \mathbb{R}\,.$$

A simple re-scaling argument shows that the last inequality is equivalent to

$$|f_0(x_0 - y)| \le \|x_0 - y\| \qquad \forall y \in M\,.$$

Therefore, replacing f_0 by its definition (5.12), we conclude that $\alpha \in \mathbb{R}$ must satisfy $|\alpha - f(y)| \le \|x_0 - y\|$ for all $y \in M$, or

$$f(y) - \|x_0 - y\| \le \alpha \le f(y) + \|x_0 - y\| \qquad \forall y \in M\,.$$

Now, such a choice of α is possible because

$$f(y) - f(z) = f(y - z) \le \|y - z\| \le \|x_0 - y\| + \|x_0 - z\|$$

for all $y, z \in M$, and so

$$\sup_{y\in M} \big\{ f(y) - \|x_0 - y\| \big\} \le \inf_{z\in M} \big\{ f(z) - \|x_0 - z\| \big\}\,.$$

2. Denote by $\mathcal{P}$ the family of all pairs $(\widetilde{M}, \widetilde{f})$ where $\widetilde{M}$ is a subspace of X including M, and $\widetilde{f}$ is a bounded linear functional extending f to $\widetilde{M}$ such that $\|\widetilde{f}\| = 1$. $\mathcal{P} \ne \varnothing$ since it contains (M, f). We can turn $\mathcal{P}$ into a partially ordered set defining, for all pairs $(M_1, f_1), (M_2, f_2) \in \mathcal{P}$,

$$(M_1, f_1) \le (M_2, f_2) \quad \Longleftrightarrow \quad \begin{cases} M_1 \quad \text{subspace of} \quad M_2 \\ f_2 = f_1 \quad \text{on} \quad M_1\,. \end{cases} \tag{5.13}$$

We claim that $\mathcal{P}$ is inductive. For let $\mathcal{Q} = \{(M_i, f_i)_{i\in I}\}$ be a totally ordered subset of $\mathcal{P}$. Then, it is easy to check that

$$\begin{cases} \widetilde{M} := \bigcup_{i\in I} M_i \\ \widetilde{f}(x) := f_i(x) \quad \text{if} \quad x \in M_i \end{cases}$$

defines a pair $(\widetilde{M}, \widetilde{f}) \in \mathcal{P}$ which is an upper bound for $\mathcal{Q}$.

3. By Zorn's Lemma, $\mathcal{P}$ has a maximal element, say $(\mathcal{M}, F)$. We claim that $\mathcal{M} = X$ and F is the required extension. Indeed, $F = f$ on M and $\|F\| = 1$ by construction. Moreover $\mathcal{M} = X$, for if $\mathcal{M} \ne X$ then the first step of this proof would imply the existence of a proper extension of $(\mathcal{M}, F)$, contradicting its maximality. □

Example 5.22 In general, the extension provided by the Hahn-Banach Theorem is not unique. For instance, consider the space

$$c_1 := \left\{ x = (x_n) \in \ell^\infty \;\middle|\; \exists \lim_{n\to\infty} x_n \right\}.$$

As it is easy to see, c_1 is a closed subspace of ℓ^∞ and c_0 is a closed subspace of c_1. The map

$$f(x) := \lim_{n\to\infty} x_n \qquad \forall x = (x_n) \in c_1$$

is a bounded linear functional on c_1 such that $f \equiv 0$ on c_0. Then, f is a nontrivial extension of the null map on c_0.

We shall now discuss some useful consequences of the Hahn-Banach Theorem.

Corollary 5.23 *Let M be a closed subspace of X and let $x_0 \notin M$. Then there exists $F \in X^*$ such that*

$$\begin{cases} (a) & F(x_0) = 1 \\ (b) & F\big|_M = 0 \\ (c) & \|F\| = 1/d_M(x_0)\,. \end{cases} \tag{5.14}$$

Proof. Let $M_0 = \mathrm{sp}(M \cup \{x_0\}) = M + \mathbb{R}x_0$. Define $f : M_0 \to \mathbb{R}$ by

$$f(x + \lambda x_0) = \lambda \qquad \forall x \in M\,,\ \forall \lambda \in \mathbb{R}\,.$$

Then, $f(x_0) = 1$ and $f\big|_M = 0$. Also, since

$$\|x + \lambda x_0\| = |\lambda|\, \Big\|\frac{x}{\lambda} + x_0\Big\| \geq |\lambda| d_M(x_0)\,,$$

we have that $\|f\| \leq 1/d_M(x_0)$. Moreover, let (x_n) be a sequence in M such that

$$\|x_n - x_0\| < \Big(1 + \frac{1}{n}\Big) d_M(x_0) \qquad \forall n \geq 1\,.$$

Then,

$$\|f\|\, \|x_n - x_0\| \geq f(x_0 - x_n) = 1 > \frac{n}{n+1}\, \frac{\|x_n - x_0\|}{d_M(x_0)} \qquad \forall n \geq 1\,.$$

Therefore, $\|f\| = 1/d_M(x_0)$. Now, the existence of an extension $F \in X^*$ satisfying (5.14) follows from the Hahn-Banach Theorem. □

Corollary 5.24 *Let $x_0 \in X \setminus \{0\}$. Then there exists $F \in X^*$ such that*

$$F(x_0) = \|x_0\| \quad \textit{and} \quad \|F\| = 1\,. \tag{5.15}$$

Proof. Let $M = \mathbb{R}x_0$. Define $f : M \to \mathbb{R}$ by

$$f(\lambda x_0) = \lambda \|x_0\| \qquad \forall \lambda \in \mathbb{R}\,.$$

Then, one can easily check that $f(x_0) = \|x_0\|$ and $\|f\| = 1$. Now, the existence of an extension $F \in X^*$ satisfying (5.15) follows from the Hahn-Banach Theorem. □

Exercise 5.25 Hereafter, for any $f \in X^*$, we will use the standard notation

$$\langle f, x\rangle := f(x) \qquad \forall x \in X\,.$$

1. Let $x_1, \ldots, x_n$ be linearly independent vectors in X and let $\lambda_1, \ldots, \lambda_n$ be real numbers. Show that there exists $f \in X^*$ such that

   $$f(x_i) = \lambda_i \qquad \forall i = 1, \ldots, n\,.$$

2. Let M be a subspace of X.

   (a) Show that a point $x \in X$ belongs to $\overline{M}$ iff $f(x) = 0$ for every $f \in X^*$ such that $f_{|M} = 0$.

   (b) Show that M is dense iff the only functional $f \in X^*$ that vanishes on M is $f \equiv 0$.

3. Show that X^* separates the points of X, that is, for any $x_1, x_2 \in X$ with $x_1 \neq x_2$ there exists $f \in X^*$ such that $f(x_1) \neq f(x_2)$.

4. Show that $\|x\| = \max\big\{\langle f, x\rangle \;\big|\; f \in X^*\,,\ \|f\| \le 1\big\}\,.$

5.3.2 Separation of convex sets

It turns out that the Hahn-Banach Thoerem has significant geometric applications. To achieve this, we shall extend our analysis to vector spaces.

Definition 5.26 *A* sublinear functional *on a vector space X is a function $p : X \to \mathbb{R}$ such that* *(a)* $\quad p(\lambda x) = \lambda p(x) \qquad \forall x \in X \qquad \forall \lambda \geq 0$

(b) $\quad p(x+y) \leq p(x) + p(y) \qquad \forall x, y \in X.$

The Hahn-Banach Theorem can be extended in the following way.

Theorem 5.27 (Hahn-Banach: second analytic form) *Let p be a sublinear functional on a vector space X and let M be subspace of X. If $f : M \to \mathbb{R}$ is a linear functional such that*

$$f(x) \leq p(x) \qquad \forall x \in M\,, \tag{5.16}$$

then there is a linear functional $F : X \to \mathbb{R}$ such that

$$\begin{cases} F\big|_M = f \\ F(x) \leq p(x) \qquad \forall x \in M\,. \end{cases} \tag{5.17}$$

The proof of Theorem 5.27 will be omitted. The reader is invited to check that the proof of Theorem 5.21 can be easily adapted to the present context.

Theorem 5.28 (Hahn-Banach: first geometric form) *Let A and B be nonempty disjoint convex subsets of a normed space X. If A is open, then there is a functional $f \in X^*$ and a real number α such that*

$$f(x) < \alpha \leq f(y) \qquad \forall x \in A \qquad \forall y \in B\,. \tag{5.18}$$

Remark 5.29 Observe that (5.18) yields, in particular, $f \neq 0$. It can be proved that, given a functional $f \in X^* \setminus 0$, for any $\alpha \in \mathbb{R}$ the set

$$\Pi_\alpha := f^{-1}(\alpha) = \{x \in X \mid f(x) = \alpha\} \tag{5.19}$$

is a closed subspace of X. We will call any such set a closed *hyperplane* in X. Therefore, an equivalent way to state the conclusion of Theorem 5.28 is that A and B can be separated by a closed hyperplane.

Lemma 5.30 *Let C be a nonempty convex open subset of a normed space X such that $0 \in C$. Then*

$$p_C(x) := \inf\{\tau \geq 0 \mid x \in \tau C\} \qquad \forall x \in X \tag{5.20}$$

is a sublinear functional on X called the Minkowski function *of C or the* gauge *of C. Moreover.*

- $\exists c \geq 0$ *such that* $0 \leq p_C(x) \leq c\|x\| \qquad \forall x \in X \qquad (5.21)$
- $C = \{x \in X \mid p_C(x) < 1\}\,. \qquad (5.22)$

Proof. To begin with, observe that, being open, C contains a ball B_R.

1. Let us prove (5.21). For any $\varepsilon > 0$ we have that
$$\frac{Rx}{\|x\| + \varepsilon} \in B_R \subset C\,.$$
Since ε is arbitrary, this yields $0 \leq p_C(x) \leq \|x\|/R$.

2. We now proceed to show that p_C is a sublinear functional. Let $\lambda > 0$ and $x \in X$. Fix $\varepsilon > 0$ and let $0 \leq \tau_\varepsilon < p_C(x) + \varepsilon$ be such that $x \in \tau_\varepsilon C$. Then, $\lambda x \in \lambda \tau_\varepsilon C$. Thus, $p_C(\lambda x) \leq \lambda \tau_\varepsilon < \lambda(p_C(x) + \varepsilon)$. Since ε is arbitrary, we conclude that
$$p_C(\lambda x) \leq \lambda p_C(x) \qquad \forall \lambda \geq 0\,,\ \forall x \in X\,. \tag{5.23}$$
To obtain the converse inequality observe that, in view of (5.23),
$$p_C(x) = p_C\Big(\frac{1}{\lambda}\,\lambda x\Big) \leq \frac{1}{\lambda}\,p_C(\lambda x)\,.$$
Finally, let us check that p_C satisfies property (b) of Definition 5.26. Fix $x, y \in X$ and $\varepsilon > 0$. Let $0 < \tau_\varepsilon < p_C(x) + \varepsilon$ and $0 < \sigma_\varepsilon < p_C(y) + \varepsilon$ be such that $x \in \tau_\varepsilon C$ and $y \in \sigma_\varepsilon C$. Then, $x = \tau_\varepsilon x_\varepsilon$ and $y = \sigma_\varepsilon y_\varepsilon$ for some points $x_\varepsilon, y_\varepsilon \in C$. Since C is convex,
$$x + y = \tau_\varepsilon x_\varepsilon + \sigma_\varepsilon y_\varepsilon = (\tau_\varepsilon + \sigma_\varepsilon)\Big(\underbrace{\frac{\tau_\varepsilon}{\tau_\varepsilon + \sigma_\varepsilon}\,x_\varepsilon + \frac{\sigma_\varepsilon}{\tau_\varepsilon + \sigma_\varepsilon}\,y_\varepsilon}_{\in C}\Big)\,.$$
Thus,
$$p_C(x + y) \leq \tau_\varepsilon + \sigma_\varepsilon < p_C(x) + p_C(y) + 2\varepsilon \qquad \forall \varepsilon > 0\,,$$
whence $p_C(x + y) \leq p_C(x) + p_C(y)$.

3. Denote by $\widetilde{C}$ the set in the right-hand side of (5.22). Since $\tau C \subset C$ for every $\tau \in [0, 1]$, we have that $\widetilde{C} \subset C$. Conversely, since C is open, any point $x \in C$ belongs to some ball $B_r(x) \subset C$. Therefore, $(1 + r)x \in C$ and so $p_C(x) \leq 1/(1 + r) < 1$. □

Lemma 5.31 *Let C be a nonempty convex open subset of a normed space X and let $x_0 \in X \setminus C$. Then there is a functional $f \in X^*$ such that $f(x) < f(x_0)$ for all $x \in C$.*

Proof. First, we note that we can assume that $0 \in C$ without loos of generality. Indeed, this is always the case up to translation. Define $M := \mathbb{R}x_0$ and

$$g : M \to \mathbb{R} \quad \text{by} \quad g(\lambda x_0) = \lambda p_C(x_0) \qquad \forall \lambda \in \mathbb{R}\,,$$

where p_C is the Minkowski function of C. Observe that g satisfies condition (5.16) with respect to the sublinear functional p_C since, for any $x = \lambda x_0 \in M$, it is easy to see that

$$g(x) = \lambda p_C(x_0) \leq p_C(x) \qquad \forall \lambda \in \mathbb{R}\,.$$

Therefore, Theorem 5.27 ensures the existence of a linear extension of g, say f, which satisfies (5.17). Then, $f(x_0) = g(x_0) = 1$ and, owing to (5.22),

$$f(x) = g(x) \leq p_C(x) < 1 \qquad \forall x \in C\,. \qquad \square$$

Proof of Theorem 5.28. It is easy to see that

$$C := A - B = \{x - y \mid x \in A\,,\ y \in B\}$$

is a convex open set and that $0 \notin C$. Then, Lemma 5.31 ensures the existence of a linear functional $f \in X^*$ such that $f(z) < 0 = f(0)$ for all $z \in C$. Hence, $f(x) < f(y)$ for all $x \in A$ and $y \in B$. So,

$$\alpha := \sup_{x \in A} f(x) \leq f(y) \qquad \forall y \in B\,.$$

We claim that $f(x) < \alpha$ for all $x \in A$. For suppose $f(x_0) = \alpha$ for some $x_0 \in A$. Then, since A is open, $\overline{B}_r(x_0) \subset A$ for some $r > 0$. So,

$$f(x_0 + rx) \leq \alpha \qquad \forall x \in \overline{B}_1\,.$$

Now, taking $x \in \overline{B}_1$ such that $f(x) > \|f\|/2 (> 0)$, we obtain

$$f(x_0 + rx) = f(x_0) + rf(x) > \alpha + \frac{r\|f\|}{2}\,.$$

a contradiction that concludes the proof. $\square$

Theorem 5.32 (Hahn-Banach: second geometric form) *Let C and D be nonempty disjoint convex subsets of a normed space X. If C is closed and D is compact, then there is a functional $f \in X^*$ such that*

$$\sup_{x\in C} f(x) < \inf_{y\in D} f(y)\,. \tag{5.24}$$

Proof. Let us denote by d_C the distance function from A. Since C is closed and D is compact, the continuity of d_C yields

$$\delta := \min_{x\in D} d_C(x) > 0\,.$$

Define

$$C_\delta := C + B_{\delta/2} \quad \text{and} \quad D_\delta := D + B_{\delta/2}\,.$$

It is easy to see that C_δ and D_δ are nonempty open convex sets. They are aslo disjoint for if $c+x = d+y$ for some points $c \in C, d \in D$ and $x, y \in B_{\delta/2}$, then

$$d_C(d) \le \|c-d\| = \|y-x\| < \delta\,.$$

Then, by Theorem 5.28, there is a linear functional $f \in X^*$ and a number α such that

$$f\Big(c + \frac{\delta}{2}\,x\Big) \le \alpha \le f\Big(d + \frac{\delta}{2}\,y\Big) \qquad \forall c \in C,\ \forall d \in D,\ \forall x, y \in B_1\,.$$

Now, let $x \in B_1$ be such that $f(x) > \|f\|/2$ [5]. Then

$$f(c) + \frac{\delta\|f\|}{4} < f\Big(c + \frac{\delta}{2}\,x\Big) \le \alpha \le f\Big(d - \frac{\delta}{2}\,x\Big) < f(d) - \frac{\delta\|f\|}{4}$$

for all $c \in C$ and $d \in D$. The conclusion follows. □

5.3.3 The dual of ℓ^p

In this section we will study the dual of the Banach spaces c_0 and ℓ^p defined in Example 5.5.4. To begin, let $p \in [1, \infty]$ and let q be the conjugate exponent, that is, $1/p + 1/q = 1$. With any $y = (y_n) \in \ell^q$ we can associate the linear map $f_y : \ell^p \to \mathbb{R}$ defined by

$$f_y(x) = \sum_{n=1}^{\infty} x_n y_n \qquad \forall x = (x_n) \in \ell^p\,. \tag{5.25}$$

[5] Recall that $\|f\| > 0$, see Remark 5.29.

Hölder's inequality ensures that

$$|f_y(x)| \le \|y\|_q \|x\|_p \qquad \forall x \in \ell^p \,.$$

Hence, $f_y \in (\ell^p)^*$ and $\|f_y\| \le \|y\|_q$. Therefore,

$$\begin{cases} j_p : \ell^q \to (\ell^p)^* \\ j_p(y) := f_y \end{cases} \tag{5.26}$$

is a bounded linear operator such that $\|j_p\| \le 1$. Moreover, since c_0 is a subspace of ℓ^∞, f_y is also a bounded linear functional on c_0 for any $y \in \ell^1$. In this section, for $p = \infty$, we shall restrict our attention to the bounded linear operator $j_\infty : \ell^1 \to (c_0)^*$.

Proposition 5.33 *The bounded linear operator*

$$j_p : \begin{cases} \ell^q \to (\ell^p)^* & \text{if} \quad 1 \le p < \infty \\ \ell^1 \to (c_0)^* & \text{if} \quad p = \infty \end{cases}$$

is an isometric isomorphism.

Let us first prove the following

Lemma 5.34 *Let*

$$X = \begin{cases} \ell^p & \text{if} \quad 1 \le p < \infty \\ c_0 & \text{if} \quad p = \infty \,. \end{cases}$$

Then X is the closed linear subspace (with respect to $\|\cdot\|_p$) generated by the set of vectors

$$e_k = (\overbrace{0,\dots,0}^{k-1}, 1, 0, \dots) \qquad k = 1, 2 \dots \tag{5.27}$$

Consequently, X is separable.

Proof. For any $x = (x_n) \in \ell^p$, $1 \le p < \infty$ we have

$$\Big\| x - \sum_{k=1}^{n} x_k e_k \Big\|_p^p = \sum_{k=n+1}^{\infty} |x_k|^p \to 0 \quad (n \to \infty) \,.$$

Similarly, for any $x = (x_n) \in c_0$,

$$\Big\| x - \sum_{k=1}^{n} x_k e_k \Big\|_\infty = \max\{|x_k| \mid k > n\} \to 0 \quad (n \to \infty)$$

because $x_n \to 0$ by definition. The conclusion follows. □

Remark 5.35 We note that the conclusion of above lemma is false for ℓ^∞ since

$$\overline{\mathrm{sp}}(e_k \mid k \geq 1) = c_0 \subsetneq \ell^\infty\,. \tag{5.28}$$

In fact, we know that ℓ^∞ is not separable, see Exercise 5.6.4.

Proof of Proposition 5.33. Let us consider, first, the case of $1 < p < \infty$. Fix $f \in (\ell^p)^*$ and set

$$y_k := f(e_k) \qquad \forall k \geq 1 \tag{5.29}$$

where e_k is defined in (5.28). It suffices to show that $y := (y_k)$ satisfies

$$y \in \ell^q\,, \qquad \|y\|_q \leq \|f\|\,, \qquad f = f_y\,. \tag{5.30}$$

For any $n \geq 1$ let[6]

$$z_k^{(n)} = \begin{cases} |y_k|^{q-2} y_k & \text{if} \quad k \leq n \\ 0 & \text{if} \quad k > n\,. \end{cases}$$

Then $z^{(n)} \in \ell^p$, since all its components vanish but a finite number, and

$$\sum_{k=1}^{n} |y_k|^q = f(z^{(n)}) \leq \|f\|\,\|z^{(n)}\|_p = \|f\|\Big(\sum_{k=1}^{n} |y_k|^q\Big)^{1/p},$$

whence

$$\Big(\sum_{k=1}^{n} |y_k|^q\Big)^{1/q} \leq \|f\| \qquad \forall n \geq 1\,.$$

This yields the first two assertions in (5.30). To obtain the third one, fix $x \in \ell^p$ and let

$$x_k^{(n)} := \begin{cases} x_k & \text{if} \quad k \leq n \\ 0 & \text{if} \quad k > n\,. \end{cases}$$

Observe that

$$f(x^{(n)}) = \sum_{k=1}^{n} x_k f(e_k) = \sum_{k=1}^{n} x_k y_k\,.$$

Since $x^{(n)} \to x$ in ℓ^p and the series $\sum_k x_k y_k$ converges, we conclude that $f = f_y$. This completes the analysis of the case $1 < p < \infty$. The similar reasoning for the remaining cases is left as an exercise. □

---

[6] observe that $|y_k|^{q-2} y_k = 0$ if $y_k = 0$ since $q > 1$.

Exercise 5.36 1. Prove Proposition 5.33 for $p = 1$.

HINT: defining y as in (5.29) the bound $\|y\|_\infty \le \|f\|$ is immediate ...

2. Prove Proposition 5.33 for $p = \infty$.

HINT: define y as in (5.29) and

$$z_k^{(n)} = \begin{cases} \dfrac{y_k}{|y_k|} & \text{if} \quad k \le n \quad \text{and} \quad y_k \ne 0 \\ 0 & \text{if} \quad y_k = 0 \quad \text{or} \quad k > n\,. \end{cases}$$

Then $\|z^{(n)}\|_\infty \le 1$ ...

5.4 Weak convergence and reflexivity

Let $(X, \|\cdot\|)$ be a normed space. Then the dual space X^* is itself a Banach space with the dual norm.

Definition 5.37 *The space* $X^{**} = (X^*)^*$ *is called the* bidual *of* X.

Let $J_X : X \to X^{**}$ be the linear map defined by

$$\langle J_X(x), f\rangle := \langle f, x\rangle \qquad \forall x \in X\,,\ \forall f \in X^*\,. \tag{5.31}$$

Then, $|\langle J_X(x), f\rangle| \le \|f\|\,\|x\|$ by definition. So, $\|J_X(x)\| \le \|x\|$. Moreover, by Corollary 5.24, for any $x \in X$ a functional $f_x \in X^*$ exists such that $f_x(x) = \|x\|$ and $\|f_x\| = 1$. Thus, $\|x\| = |\langle J_X(x), f_x\rangle| \le \|J_X(x)\|$. Therefore, $\|J_X(x)\| = \|x\|$ for every $x \in X$, that is, J_X is a *linear isometry.*

5.4.1 Reflexive spaces

The above considerations imply that $J_X(X)$ is a subspace of X^{**}. It is useful to single out the case where such a subspace coincides with the bidual.

Definition 5.38 *A space* X *is called* reflexive *if the map* $J_X : X \to X^{**}$ *defined in* (5.31) *is onto.*

Recalling that J_X is a linear isometry, we conclude that any reflexive space X is isometrically isomorphic to its bidual X^{**}. Since X^{**} is complete, like every dual space, every reflexive normed space must also be complete.

Example 5.39 1. If X is a Hilbert space, then X^* is isometrically isomorphic to X by the Riesz-Fréchet Theorem. Therefore, so is X^{**}. In other words, any Hilbert space is reflexive.

2. Let $1 < p < \infty$. Then Proposition 5.33 ensures that $(\ell^p)^* = \ell^q$, where p and q are conjugate exponents. So, ℓ^p is reflexive for all $p \in (1, \infty)$.

Theorem 5.40 *Let X be a Banach space.*

(a) If X^* is separable, then X is separable.

(b) If X^* is reflexive, then X is reflexive.

Proof.

(a) Let (f_k) be a dense sequence in X^*. There exists a sequence (x_k) in X such that

$$\|x_k\| \quad \text{and} \quad |f_k(x_k)| \geq \frac{\|f_k\|}{2} \qquad \forall k \geq 1\,.$$

We claim that X coincides with the closed subspace generated by (x_k). For let $M = \overline{\text{sp}}(x_k \mid k \geq 1)$ and suppose there exists $x_0 \in X \setminus M$. Then, applying Corollary 5.23 we can find a functional $f \in X^*$ such that

$$f(x_0) = 1\,, \quad f_{\big|M} = 0\,, \quad \|f\| = \frac{1}{d_M(x)}\,.$$

So,

$$\frac{\|f_k\|}{2} \leq |f_k(x_k)| = |f_k(x_k) - f(x_k)| \leq \|f_k - f\|\,,$$

whence

$$\frac{1}{d_M(x)} = \|f\| \leq \|f - f_k\| + \|f_k\| \leq 3\|f - f_k\|\,.$$

Thus, (f_k) cannot be dense in X^*—a contradiction.

(b) Observe that, since X is a Banach space, $J_X(X)$ is a closed subspace of X^{**}. Suppose there exists $\phi_0 \in X^{**} \setminus J_X(X)$. Then, by Corollary 5.23 applied to the bidual, we can find a bounded linear functional on X^{**}, valued 1 at ϕ_0 and 0 on $J_X(X)$. Since X^* is reflexive, such a functional will belong to $J_{X^*}(X^*)$. So, for some $f \in X^*$,

$$\langle \phi_0, f \rangle = 1 \quad \text{and} \quad 0 = \langle J_X(x), f \rangle = \langle f, x \rangle \qquad \forall x \in X\,,$$

a contradiction that concludes the proof. □

Remark 5.41 1. From point (a) of Theorem 5.40 we conclude that, since ℓ^∞ is not separable, $(\ell^\infty)^*$ also fails to be separable. So, $(\ell^\infty)^*$ is not isomorphic to ℓ^1, and ℓ^1 is not reflexive. Moreover, ℓ^∞ also fails to be reflexive since otherwise ℓ^1 would be reflexive by point (b) above.

2. The result of point (b) of Theorem 5.40 is an equivalence since the implication
$$X \quad \text{reflexive} \quad \Longrightarrow \quad X^* \quad \text{reflexive}$$
in trivial. On the contrary, the implication of point (a) cannot be reversed. Indeed, ℓ^1 is separable, whereas $\ell^\infty = (\ell^1)^*$ is not.

Corollary 5.42 *A Banach space X is reflexive and separable iff X^* is reflexive and separable.*

Proof. The only part of the conclusion that needs to be justified is the fact that, if X is reflexive and separable, then X^* is separable. But this follows from the fact that X^{**} is separable, since it is isomorphic to X, and from Theorem 5.40 (a). □

We conclude this section with a result on the reflexivity of subspaces.

Proposition 5.43 *Let M be a closed linear subspace of a reflexive Banach space X. Then M is reflexive.*

Proof. Let ϕ be a bounded linear functional on M^*. Define a functional $\overline{\phi}$ on X^* by
$$\langle \overline{\phi}, f \rangle = \big\langle \phi, f_{|_M} \big\rangle \qquad \forall f \in X^* .$$
Since $\overline{\phi} \in X^{**}$, by hypothesis we have that $\overline{\phi} = J_X(\overline{x})$ for some $\overline{x} \in X$. The proof is completed by the following two steps.

1. We claim that $\overline{x} \in M$. For if $\overline{x} \in X \setminus M$, then by Corollary 5.23 there exists $\overline{f} \in X^*$ such that
$$\langle \overline{f}, \overline{x} \rangle = 1 \quad \text{and} \quad \overline{f}_{|_M} = 0 .$$
This yields a contradiction since
$$1 = \langle \overline{\phi}, \overline{f} \rangle = \big\langle \phi, \overline{f}_{|_M} \big\rangle = 0 .$$

2. We claim that $\overline{\phi} = J_M(\overline{x})$. Indeed, for any $f \in M^*$ let $\widetilde{f} \in X^*$ be the extension of f to X provided by the Hahn-Banach Theorem. Then,
$$\langle \phi, f \rangle = \langle \overline{\phi}, \widetilde{f} \rangle = \langle \widetilde{f}, \overline{x} \rangle = \langle f, \overline{x} \rangle \qquad \forall f \in X^* . \qquad \square$$

5.4.2 Weak convergence and BW property

It is well known that the unit ball B_1 of a finite dimesional Banach spaces is relatively compact. We refer to such a property as the *Bolzano-Weierstrass property*. One of the most striking phenomena that occur in infinite dimensions is that the Bolzano-Weierstrass property is no longer true. In fact, the following result holds.

Theorem 5.44 *Any Banach space with the Bolzano-Weierstrass property must be finite dimensional.*

Lemma 5.45 *Let M be a closed linear subspace of a Banach space X such that $M \neq X$. Then a sequence $(x_n) \subset X$ exists such that*

$$\|x_n\| = 1 \qquad \forall n \geq 1 \quad \textit{and} \quad d_M(x_n) \to 1 \quad \textit{as} \quad n \to \infty \,. \tag{5.32}$$

Proof. Invoking Corollary 5.23, we can find a functional $f \in X^*$ such that

$$\|f\| = 1 \quad \text{and} \quad f_{\big|M} = 0 \,.$$

Then, for every $n \geq 1$ there exists $x_n \in X$ such that

$$\|x_n\| = 1 \quad \text{and} \quad |f(x_n)| > 1 - \frac{1}{n} \,.$$

Therefore, for every $n \geq 1$,

$$1 - \frac{1}{n} < |f(x_n) - f(y)| \leq \|x_n - y\| \qquad \forall y \in M \,.$$

Taking the infinum over all $y \in M$ we obtain that $1 - 1/n \leq d_M(x_n) \leq 1$. The conclusion follows. □

Proof of Theorem 5.44. Suppose $\dim X = \infty$. Let x_1 be a fixed unit vector and let $V_1 := \mathbb{R}x_1 = \text{sp}(\{x_1\})$. Since $V_1 \neq X$, the above lemma implies the existence of a vector $x_2 \in X$ such that

$$\|x_2\| = 1 \quad \text{and} \quad d_{V_1}(x_2) > \frac{1}{2} \,.$$

Let $V_2 := \text{sp}(\{x_1, x_2\})$ and observe that $V_1 \subset V_2 \subsetneq X$. Again by Lemma 5.45 we can find a vector $x_3 \in X$ such that

$$\|x_3\| = 1 \quad \text{and} \quad d_{V_2}(x_3) > 1 - \frac{1}{3} \,.$$

Iterating this process we can construct a sequence (x_n) in X such that

$$\|x_n\| = 1 \quad \text{and} \quad d_{V_n}(x_{n+1}) > 1 - \frac{1}{n+1},$$

where $V_n = \mathrm{sp}(\{x_1, \ldots, x_n\}) \subsetneq X$. Then, (x_n) has no cluster point in X since, for any $1 \le m < n$, we have $1 - 1/n < d_{V_m}(x_n) \le \|x_n - x_m\|$. $\square$

A surrogate for the Bolzano-Weierstrass property in infinite dimensional spaces is ther notion of convergence we introduce below.

Definition 5.46 *A sequence $(x_n) \subset X$ is said to* converge weakly *to a point $x \in X$ if*

$$\lim_{n\to\infty} \langle f, x_n \rangle = \langle f, x \rangle \qquad \forall f \in X^*.$$

In this case we write $w - \lim_{n\to\infty} x_n = x$ or $x_n \rightharpoonup x$.

A sequence (x_n) that converges in norm is also said to *converge strongly.* Since $|\langle f, x_n \rangle - \langle f, x \rangle| \le \|f\| \, \|x_n - x\|$, n it is easy to see that any strongly convergent sequence is also weakly convergent. The conserve is not true as is shown by the following example.

Example 5.47 Let (e_n) be an orthonormal sequence in an infinite dimensional Hilbert space X. Then, owing to Bessel's inequality $\langle x, e_n \rangle \to 0$ as $n \to \infty$ for every $x \in X$. Therefore, $e_n \rightharpoonup 0$ as $n \to \infty$. On the other hand, $\|e_n\| = 1$ for every n. So, (e_n) does not converge strongly to 0.

Proposition 5.48 *Let $(x_n), (y_n)$ be sequences in a Banach space X, and let $x, y \in X$.*

(a) If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$, then $x = y$.

(b) If $x_n \rightharpoonup x$ and $y_n \rightharpoonup y$, then $x_n + y_n \rightharpoonup x + y$.

(c) If $x_n \rightharpoonup x$ and $(\lambda_n) \subset \mathbb{R}$ converges to λ, then $\lambda_n x_n \rightharpoonup \lambda x$.

(d) If $x_n \overset{X}{\rightharpoonup} x$ and $\Lambda \in \mathcal{L}(X, Y)$, then $\Lambda x_n \overset{Y}{\rightharpoonup} \Lambda x$.

(e) If $x_n \rightharpoonup x$, then (x_n) is bounded.

(f) If $x_n \rightharpoonup x$, then $\|x\| \le \liminf_{n\to\infty} \|x_n\|$.

Proof.

(a) By hypothesis we have that $\langle f, x-y\rangle = 0$ for every $f \in X^*$. Then, the conlusion follows recalling Exercise 5.25.3.

(b) The proof is left to the reader.

(c) Since (λ_n) is bounded, say $|\lambda_n| \le C$, for any $f \in X^*$ we have that

$$|\lambda_n\langle f, x_n\rangle - \lambda\langle f, x\rangle| \le \underbrace{|\lambda_n|}_{\le C}\,|\underbrace{\langle f, x_n - x\rangle}_{\downarrow 0}| + \underbrace{|\lambda_n - \lambda|}_{\downarrow 0}\,|\langle f, x\rangle|\,.$$

(d) Let $g \in Y^*$. Then $\langle g, \Lambda x_n\rangle = \langle g \circ \Lambda, x_n\rangle \to 0$ since $g \circ \Lambda \in X^*$.

(e) Consider the sequence $(J_X(x_n))$ in X^{**}. Since

$$\langle J_X(x_n), f\rangle = \langle f, x_n\rangle \to \langle f, x\rangle \qquad \forall f \in X^*\,.$$

we have that $\sup_n |\langle J_X(x_n), f\rangle| < \infty$ for all $f \in X^*$. So, the Banach-Steinhaus Theorem implies that

$$\sup_n \|x_n\| = \sup_n \|J_X(x_n)\| < \infty\,.$$

(f) Let $f \in X^*$ be such that $\|f\| \le 1$. Then,

$$\underbrace{|\langle f, x_n\rangle|}_{\to|\langle f,x\rangle|} \le \|x_n\| \quad \Longrightarrow \quad |\langle f, x\rangle| \le \liminf_{n\to\infty} \|x_n\|\,.$$

The conclusion follows recalling Exercise 5.25.4. □

Exercise 5.49 1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by

$$f_n(x) = \begin{cases} \dfrac{1}{2^n} & \text{if } x \in [2^n, 2^{n+1}], \\ 0 & \text{otherwise.} \end{cases}$$

Show that

- $f_n \to 0$ in $L^p(\mathbb{R})$ for all $1 < p \le +\infty$;
- $\{f_n\}$ does not converge weakly in $L^1(\mathbb{R})$.

2. Show that, in a Hilbert space X,

$$x_n \to x \quad \Longleftrightarrow \quad x_n \rightharpoonup x \quad \text{and} \quad \|x_n\| \to \|x\| \,.$$

   HINT: observe that $\|x_n - x\|^2 = \|x_n\|^2 + \|x\|^2 - 2\langle x_n, x\rangle \ldots$

3. Let C be a closed convex subset of X and let $(x_n) \subset C$. Show that, if $x_n \rightharpoonup x$, then $x \in C$.

   HINT: use Lemma 5.31.

Besides strong and weak convergence, in dual spaces one can define another notion of convergence.

Definition 5.50 *A sequence $(f_n) \subset X^*$ is said to converge* weakly$-*$ *to a functional $f \in X^*$ iff*

$$\langle f_n, x\rangle \to \langle f, x\rangle \quad \textit{as} \quad n \to \infty \qquad \forall x \in X \,. \tag{5.33}$$

In this case we write

$$w^* - \lim_{n\to\infty} f_n = f \quad \textit{or} \quad f_n \overset{*}{\rightharpoonup} f \quad (\textit{as} \quad n \to \infty) \,.$$

Remark 5.51 It is interesting to compare weak and weak$-*$ convergence on X^*. By definition, a sequence $(f_n) \subset X^*$ converges weakly to f iff

$$\langle \phi, f_n\rangle \to \langle \phi, f\rangle \quad \text{as} \quad n \to \infty \tag{5.34}$$

for all $\phi \in X^{**}$, whereas, $f_n \overset{*}{\rightharpoonup} f$ iff (5.34) holds for all $\phi \in J_X(X)$. Therefore, weak convergence is equivalent to weak$-*$ convergence if X is reflexive but, in general, weak convergence is stronger than weak$-*$ convergence.

Example 5.52 In $\ell^\infty = (\ell^1)^*$ consider the sequence $(x^{(n)})$ defined by

$$x_k^{(n)} := \begin{cases} 0 & \text{if} \quad k \le n \\ 1 & \text{if} \quad k > n \,. \end{cases}$$

Then $x^{(n)} \overset{*}{\rightharpoonup} 0$. Indeed, for every $y = (y_k) \in \ell^1$,

$$\big\langle j_1\big(x^{(n)}\big), y\big\rangle = \sum_{k=1}^{\infty} x_k^{(n)} y_k = \sum_{k=n+1}^{\infty} y_k \to 0 \quad (n \to \infty) \,.$$

where $j_1 : \ell^\infty \to (\ell^1)^*$ is the linear isometry defined in (5.26). On the other hand, we have that $x^{(n)} \not\rightharpoonup 0$. Indeed, define

$$f(x) := \lim_{k\to\infty} x_k \qquad \forall x = (x_k) \in c_1$$

where $c_1 := \{x = (x_k) \in \ell^\infty \mid \exists \lim_k x_k\}$ (see Example 5.22). Then, denoting by F any bounded linear functional extending f to ℓ^∞—for instance, the one provided by the Hahn-Banach Theorem—we have that

$$\langle F, x^{(n)} \rangle = \lim_{k\to\infty} x_k^{(n)} = 1 \qquad \forall n \geq 1\,.$$

Exercise 5.53 1. Show that any $(f_n) \subset X^*$ that converges weakly$-*$ is bounded in X^*.

2. Show that, if $x_n \rightharpoonup x$ and $f_n \to f$, then $\langle f_n, x_n\rangle \to \langle f, x\rangle$ as $n \to \infty$.

3. Show that, if $x_n \to x$ and $f_n \overset{*}{\rightharpoonup} f$, then $\langle f_n, x_n\rangle \to \langle f, x\rangle$ as $n \to \infty$.

One of the nice features of weak$-*$ convergence is the following result yielding a sort of weak$-*$ Bolzano-Weierstrass property of X^*.

Theorem 5.54 (Banach-Alaoglu) *Let X be a separable normed space. Then every bounded sequence $(f_n) \subset X^*$ has a weakly$-*$ convergent subsequence.*

Proof. Let (x_n) be a dense sequence in X and let $C \geq 0$ be an upper bound for $\|f_n\|$. Then $|f_n(x_1)| \leq C\|x_1\|$. So, there exists a subsequence of (f_n), say $(f_{1,n})$, such that $f_{1,n}(x_1)$ converges. Next, since $|f_{1,n}(x_2)| \leq C\|x_2\|$, there exists a subsequence $(f_{2,n}) \subset (f_{1,n})$, such that $f_{2,n}(x_2)$ converges. Iterating this process, for any $k \geq 1$ we can construct nested subsequences

$$(f_{k,n}) \subset (f_{k-1,n}) \subset \cdots \subset (f_{1,n}) \subset (f_n)$$

such that $|f_n(x_k)| \leq C\|x_k\|$ and $f_{k,n}(x_k)$ converges as $n \to \infty$ for every $k \geq 1$. Define, for $n \geq 1$, $g_n(x) := f_{n,n}(x)$ for all $x \in X$. Then, $(g_n) \subset (f_n)$, $\|g_n(x)\| \leq C\|x\|$, and $g_n(x_k)$ converges as $n \to \infty$ for every $k \geq 1$ since it is, for $n \geq k$, a subsequence of $f_{k,n}(x_k)$.

Let us complete the proof showing that $g_n(x)$ converges for every $x \in X$. Fix $x \in X$ and $\varepsilon > 0$. Then, there exist $k_\varepsilon, n_\varepsilon \geq 1$ such that

$$\begin{cases} \|x - x_{k_\varepsilon}\| < \varepsilon \\ |g_n(x_{k_\varepsilon}) - g_m(x_{k_\varepsilon})| < \varepsilon \quad \forall m, n \geq n_\varepsilon \end{cases}$$

Therefore, for all $m, n \geq n_\varepsilon$,

$$|g_n(x) - g_m(x)| \leq \underbrace{|g_n(x) - g_n(x_{k_\varepsilon})| + |g_m(x_{k_\varepsilon}) - g_m(x)|}_{\leq 2C\|x - x_{k_\varepsilon}\|} + |g_n(x_{k_\varepsilon}) - g_m(x_{k_\varepsilon})| \leq (2C+1)\varepsilon\,.$$

Thus, $(g_n(x))$ is a Cauchy sequence satisfying $|g_n(x)| \leq C\|x\|$ for all $x \in X$. This implies that $f(x) := \lim_n g_n(x)$ is an element of X^*. □

The main result of this section is that reflexive Banach space have the weak Bolzano-Weierstrass property as we show next.

Theorem 5.55 *In a reflexive Banach space, every bounded sequence has a weakly convergent subsequence.*

Proof. Define $M := \overline{\text{sp}}(x_n \mid n \geq 1)$. Observe that, in view of Proposition 5.43, M is a separable reflexive Banach space. Therefore, Corollary 5.42 ensures that M^* is separable and reflexive too. Consider the bounded sequence $(J_M(x_n)) \subset M^{**}$. Applying Alaoglu's Theorem, we can find a subsequence (x_{n_k}) such that $J_M(x_{n_k}) \overset{*}{\rightharpoonup} \overline{\phi} \in M^{**}$ as $n \to \infty$. The reflexivity of M guarantees that $\overline{\phi} = J_M(\overline{x})$ for some $\overline{x} \in M$. Therefore, for every $f \in M^*$,

$$f(x_{n_k}) = \langle J_M(x_{n_k}), f\rangle \to \langle J_M(\overline{x}), f\rangle = f(\overline{x}) \quad \text{as} \quad n \to \infty\,.$$

Finally, for any $F \in X^*$ we have that $F_{|_M} \in M^*$. So,

$$F(x_{n_k}) = F_{|_M}(x_{n_k}) \to F_{|_M}(\overline{x}) = f(\overline{x}) \quad \text{as} \quad n \to \infty\,. \qquad \square$$

Exercise 5.56 1. Let $1 < p < \infty$ and let $x^{(n)} = (x_k^{(n)})_{k\geq 1}$ be a *bounded* sequence in ℓ^p. Show that $x^{(n)} \rightharpoonup x = (x_k)_{k\geq 1}$ in ℓ^p if and only if, for every $k \geq 1$, $x_k^{(n)} \to x_k$ as $n \to \infty$.

HINT: suppose $x = 0$, fix $y \in \ell^q$, and let $\|x^{(n)}\|_p \leq C$ for all $n \geq 1$. For any $\varepsilon > 0$ let $k_\varepsilon \geq 1$ be such that

$$\Big(\sum_{k=k_\varepsilon+1}^{\infty} |y_k|^q\Big)^{\frac{1}{q}} < \varepsilon\,,$$

and let $n_\varepsilon \geq 1$ be such that

$$\Big(\sum_{k=1}^{k_\varepsilon} |x_k^{(n)}|^p\Big)^{\frac{1}{p}} < \varepsilon \qquad \forall n \geq n_\varepsilon\,.$$

Then, for all $n \geq n_\varepsilon$,

$$\big|\langle j_p(y), x^{(n)}\rangle\big| = \Big|\sum_{k=1}^{k_\varepsilon} y_k x_k^{(n)}\Big| + \Big|\sum_{k=k_\varepsilon+1}^{\infty} y_k x_k^{(n)}\Big|$$

$$\leq \underbrace{\Big(\sum_{k=1}^{k_\varepsilon} |y_k|^q\Big)^{\frac{1}{q}}}_{\leq \|y\|} \underbrace{\Big(\sum_{k=1}^{k_\varepsilon} |x_k^{(n)}|^p\Big)^{\frac{1}{p}}}_{\leq \varepsilon} + \underbrace{\Big(\sum_{k=k_\varepsilon+1}^{\infty} |y_k|^q\Big)^{\frac{1}{q}}}_{\leq \varepsilon} \underbrace{\Big(\sum_{k=k_\varepsilon+1}^{\infty} |x_k^{(n)}|^p\Big)^{\frac{1}{p}}}_{\leq C} .$$

2. Find a counterexample to show that the above conclusion is false if $x^{(n)}$ fails to be bounded.

   HINT: in ℓ^2 let $x^{(n)} = n^2 e_n$ where e_n is the sequence of vectors defined in (5.27). Then, for every $k \geq 1$, $x_k^{(n)} \to 0$ as $n \to \infty$. On the other hand, taking $y = (1/k)_{k\geq 1}$ we have that

   $$y \in \ell^2 \quad \text{and} \quad \langle j_2(y), x^{(n)}\rangle = n \to \infty .$$

3. Let $x^{(n)} = ()_{k\geq 1}$ be bounded in c_0. Show that $x^{(n)} \rightharpoonup x = (x_k)_{k\geq 1}$ in c_0 if and only if, for every $k \geq 1$, $x_k^{(n)} \to x_k$ as $n \to \infty$.

   HINT: argue as in point 1 above.

4. Let $1 < p < \infty$ and let x, $x^{(n)} \subset \ell^p$. Show that

   $$x^{(n)} \to x \quad \Longleftrightarrow \quad \begin{cases} x^{(n)} \rightharpoonup x \\ \|x^{(n)}\| \to \|x\| \end{cases} \tag{5.35}$$

   HINT: use point 1 of this exercise and adapt the proof of Proposition 3.38 observing that, for any $k \geq 1$,

   $$0 \leq \frac{|x_k^{(n)}|^p + |x_k|^p}{2} - \Big|\frac{x_k^{(n)} - x_k}{2}\Big|^p \to |x_k|^p \quad (n \to \infty) .$$

5. Show that property (5.35) fails c_0.

   HINT: consider the sequence $x^{(n)} = e_1 + e_n$ where (e_k) is the sequence of vectors defined in (5.27).

Remark 5.57 1. We say Banach space X has the *Radon-Riesz property* if (5.35) holds true for every sequence $x^{(n)}$ in X. By the above exercise, such a property holds in ℓ^p for all $1 < p < \infty$, but not in c_0. Owing to Exercise 5.49.2, all Hilbert spaces have the Radon-Riesz property.

2. A surprising result known as Schur's Theorem[7] ensures that, in ℓ^1, weak and strong convergence coincide, that is, for all $x^{(n)}$, $x \in \ell^1$, we have that

$$x^{(n)} \to x \quad \Longleftrightarrow \quad x^{(n)} \rightharpoonup x\,.$$

Then, in view of Schur's Theorem, ℓ^1 has the Radon-Riesz property. On the other hand, this very theorem makes it easy to check that the property described in Exercise 5.56.1 fails in ℓ^1. Indeed, the sequence (e_k) in (5.27) does not converge strongly—thus, weakly—to 0.

[7]see. for instance, Proposition 2.19 in [3].

Chapter 6

Product measures

6.1 Product spaces

6.1.1 Product measure

Let $(X, \mathcal{F})$ and $(Y, \mathcal{G})$ be measurable space. We will turn the product $X \times Y$ into a measurable space in a canonical way.

A set of the form $A \times B$, where $A \in \mathcal{F}$ and $B \in \mathcal{G}$, is called a *measurable rectangle*. Let us denote by $\mathcal{R}$ the family of all finite disjoint unions of measurable rectangles.

Proposition 6.1 *$\mathcal{R}$ is an algebra.*

Proof. Clearly, $\emptyset$ and $X \times Y$ are measurable rectangles. It is also easy to see that the intersection of any two measurable rectangles is again a measurable rectangle. Moreover, the intersection of any two elements of $\mathcal{R}$ stays in $\mathcal{R}$. Indeed, let $\dot\cup_i (A_i \times B_i)$ and $\dot\cup_j (C_j \times D_j)$ [(1)] be finite disjoint unions of measurable rectangles. Then,

$$\left(\dot\cup_i (A_i \times B_i)\right) \bigcap \left(\dot\cup_j (C_j \times D_j)\right) = \dot\cup_{i,j} \left((A_i \times B_i) \cap (C_j \times D_j)\right) \in \mathcal{R}.$$

Let us show that the complement of any set $E \in \mathcal{R}$ is again in $\mathcal{R}$. This is true if $E = A \times B$ is a measurable rectangle since

$$E^c = (A^c \times B) \dot\cup (A \times B^c) \dot\cup (A^c \times B^c).$$

[(1)] Hereafter the symbol $\dot\cup$ denotes a disjoint union.

Now, proceeding by induction, let

$$E = \underbrace{\Big(\bigcup_{i=1}^{n}(A_i \times B_i)\Big)}_{F} \bigcup (A_{n+1} \times B_{n+1}) \in \mathcal{R}$$

and suppose $F^c \in \mathcal{R}$. Then, $E^c = F^c \cap (A_{n+1} \times B_{n+1})^c \in \mathcal{R}$ because $(A_{n+1} \times B_{n+1})^c \in \mathcal{R}$ and we have already proven that $\mathcal{R}$ is closed under intersection. This completes the proof. □

Definition 6.2 *The σ–algebra generated by $\mathcal{R}$ is called the* product σ–algebra *of $\mathcal{F}$ and $\mathcal{G}$. It is denoted by $\mathcal{F} \times \mathcal{G}$.*

For any $E \in \mathcal{F} \times \mathcal{G}$ we define the *sections* of E putting, for $x \in X$ and $y \in Y$,

$$E_x = \{y \in Y : (x, y) \in E\}, \quad E^y = \{x \in X : (x, y) \in E\}.$$

Proposition 6.3 *Let $(X, \mathcal{F}, \mu)$, $(Y, \mathcal{G}, \nu)$ be σ-finite measure spaces and let $E \in \mathcal{F} \times \mathcal{G}$. Then the following statements hold.*

(a) *$E_x \in \mathcal{G}$ and $E^y \in \mathcal{F}$ for any $(x, y) \in X \times Y$.*

(b) *the functions*

$$\begin{cases} X \to \mathbb{R} \\ x \mapsto \nu(E_x) \end{cases} \quad \text{and} \quad \begin{cases} Y \to \mathbb{R} \\ y \mapsto \mu(E^y) \end{cases}$$

are μ–measurable and ν–measurable, respectively. Moreover,

$$\int_X \nu(E_x) d\mu = \int_Y \mu(E^y) d\nu. \tag{6.1}$$

Proof. Suppose, first, that $E = \dot\cup_{i=1}^{n}(A_i \times B_i)$ stays in $\mathcal{R}$. Then, for $(x, y) \in X \times Y$ we have $E_x = \dot\cup_{i=1}^{n}(A_i \times B_i)_x$ and $E^y = \dot\cup_{i=1}^{n}(A_i \times B_i)^y$, where

$$(A_i \times B_i)_x = \begin{cases} B_i & \text{if } x \in A_i, \\ \varnothing & \text{if } x \notin A_i, \end{cases} \qquad (A_i \times B_i)^y = \begin{cases} A_i & \text{if } y \in B_i, \\ \varnothing & \text{if } y \notin B_i. \end{cases}$$

Consequently,

$$\nu(E_x) = \sum_{i=1}^{n} \nu((A_i \times B_i)_x) = \sum_{i=1}^{n} \nu(B_i) \chi_{A_i}(x),$$

$$\mu(E^y) = \sum_{i=1}^{n} \mu((A_i \times B_i)^y) = \sum_{i=1}^{n} \mu(A_i)\chi_{B_i}(y)$$

so that (6.1) follows. Then the thesis holds true when E stays in $\mathcal{R}$.

Now, let $\mathcal{E}$ be the family of all sets $E \in \mathcal{F} \times \mathcal{G}$ satisfying (a). Clearly, $X \times Y \in \mathcal{E}$. Furthermore for any $E, (E_n)_n \subset \mathcal{F} \times \mathcal{G}$ and $(x, y) \in X \times Y$ we have

$$(E^c)_x = (E_x)^c, \quad (E^c)^y = (E^y)^c,$$

$$\cup_n (E_n)_x = (\cup_n E_n)_x, \quad \cup_n (E_n)^y = (\cup_n E_n)^y.$$

Hence $\mathcal{E}$ is a σ-algebra including $\mathcal{R}$ and, consequently, $\mathcal{E} = \mathcal{F} \times \mathcal{G}$.

We are going to prove (b). First assume that μ and ν are finite and define

$$\mathcal{M} = \big\{ E \in \mathcal{F} \times \mathcal{G} \,\big|\, E \text{ satisfies } (b) \big\}.$$

We claim that $\mathcal{M}$ is a monotone class. For let $(E_n)_n \subset \mathcal{M}$ be such that $E_n \uparrow E$. Then, for any $(x, y) \in X \times Y$,

$$(E_n)_x \uparrow E_x \quad \text{and} \quad (E_n)^y \uparrow E^y\,.$$

Thus

$$\nu\big((E_n)_x\big) \uparrow \nu(E_x) \quad \text{and} \quad \mu\big((E_n)^y\big) \uparrow \mu(E^y).$$

Since $x \mapsto \nu\big((E_n)_x\big)$ is μ–measurable for all $n \in \mathbb{N}$, we have that $x \mapsto \nu(E_x)$ is μ–measurable too. Similarly, $y \mapsto \mu(E^y)$ is ν–measurable. Furthermore, by the Monotone Convergence Theorem,

$$\int_X \nu(E_x) d\mu = \lim_{n\to\infty} \int_X \nu\big((E_n)_x\big) d\mu = \lim_{n\to\infty} \int_Y \mu\big((E_n)^y\big) d\nu = \int_Y \mu(E^y) d\nu.$$

Therefore, $E \in \mathcal{M}$. Next consider $(E_n)_n \subset \mathcal{M}$ such that $E_n \downarrow E$. Then, a similar argument as above shows that for every $(x, y) \in X \times Y$

$$\nu\big((E_n)_x\big) \downarrow \nu(E_x) \quad \text{and} \quad \mu\big((E_n)^y\big) \downarrow \mu(E^y).$$

Consequently the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are μ–measurable and ν–measurable, respectively. Furthermore,

$$\nu\big((E_n)_x\big) \le \nu(Y) \quad \forall x \in X, \qquad \mu\big((E_n)^y\big) \le \mu(X) \quad \forall y \in Y,$$

and, since μ and ν are finite, the constants are summable. Then, Lebesgues's Theorem yields

$$\int_X \nu(E_x)d\mu = \lim_{n\to\infty}\int_X \nu\big((E_n)_x\big)d\mu = \lim_{n\to\infty}\int_Y \mu\big((E_n)^y\big)d\nu = \int_Y \mu(E^y)d\nu,$$

which implies $E \in \mathcal{M}$. So, $\mathcal{M}$ is a monotone class as claimed. For the first part of the proof $\mathcal{R} \subset \mathcal{M}$. Theorem 1.29 implies that $\mathcal{M} = \mathcal{F} \times \mathcal{G}$. Then the thesis is proved when μ and ν are finite. Now assume that μ and ν are σ-finite; we have $X = \cup_k X_k$, $Y = \cup_k Y_k$ for some increasing sequences $(X_k)_k \subset \mathcal{F}$ and $(Y_k)_k \subset \mathcal{G}$ such that $X = \cup X_k$, $Y = \cup Y_k$ and

$$\mu(X_k) < \infty, \quad \nu(Y_k) < \infty \quad \forall k \in \mathbb{N}. \tag{6.2}$$

Define $\mu_k = \mu \llcorner X_k$, $\nu_k = \nu \llcorner Y_k$ and fix $E \in \mathcal{F} \times \mathcal{G}$. For any $(x, y) \in X \times Y$,

$$E_x \cap Y_k \uparrow E_x \quad \text{and} \quad E^y \cap X_k \uparrow E^y\,.$$

Thus

$$\nu_k(E_x) = \nu\big(E_x \cap Y_k\big) \uparrow \nu(E_x) \quad \text{and} \quad \mu_k(E^y) = \mu\big(E^y \cap X_k\big) \uparrow \mu(E^y).$$

Since μ_k and ν_k are finite measures, for all $k \in \mathbb{N}$ the function $x \mapsto \nu_k(E_x)$ is μ–measurable; consequently $x \mapsto \nu(E_x)$ is μ–measurable too. Similarly, $y \mapsto \mu(E^y)$ is ν–measurable. Furthermore, by the Monotone Convergence Theorem,

$$\int_X \nu(E_x)d\mu = \lim_{k\to\infty}\int_X \nu_k(E_x)d\mu = \lim_{k\to\infty}\int_Y \mu_k(E^y)d\nu = \int_Y \mu(E^y)d\nu.$$

□

Theorem 6.4 *Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be σ-finite measure spaces. The set function $\mu \times \nu$ defined by*

$$(\mu \times \nu)(E) = \int_X \nu(E_x)d\mu = \int_Y \mu(E^y)d\nu \qquad \forall E \in \mathcal{F} \times \mathcal{G} \tag{6.3}$$

is a σ-finite measure on $(X \times Y, \mathcal{F} \times \mathcal{G})$, called product measure *of μ and ν. Moreover, if λ is any measure on $(X \times Y, \mathcal{F} \times \mathcal{G})$ satisfying*

$$\lambda(A \times B) = \mu(A)\nu(B) \qquad \forall A \in \mathcal{F}\,,\ \forall B \in \mathcal{G}\,, \tag{6.4}$$

then $\lambda = \mu \times \nu$.

Proof. First, to check that $\mu \times \nu$ is σ–additive let $(E_n)_n$ be a disjoint sequence in $\mathcal{F} \times \mathcal{G}$. Then, for any $(x, y) \in X \times Y$, $\big((E_n)_x\big)_n$ and $\big((E_n)^y\big)_n$ are disjoint families in $\mathcal{G}$ and $\mathcal{F}$, respectively. Therefore,

$$\begin{aligned}
(\mu \times \nu)(\cup_n E_n) &= \int_X \nu\big((\cup_n E_n)_x\big) d\mu \\
&= \int_X \nu\big(\cup_n (E_n)_x\big) d\mu = \int_X \sum_n \nu\big((E_n)_x\big) d\mu \\
[\text{Proposition 2.39}] \quad &= \sum_n \int_X \nu\big((E_n)_x\big) d\mu = \sum_n (\mu \times \nu)(E_n)\,.
\end{aligned}$$

To prove that $\mu \times \nu$ is σ-finite, observe that if $(X_k)_k \subset \mathcal{F}$ and $(Y_k)_k \subset \mathcal{G}$ are two increasing sequences such that

$$\mu(X_k) < \infty, \quad \nu(Y_k) < \infty \quad \forall k \in \mathbb{N},$$

then, setting $Z_k = X_k \times Y_k$, we have $Z_k \in \mathcal{F} \times \mathcal{G}$, $(\mu \times \nu)(Z_k) = \mu(X_k)\nu(Y_k) < \infty$ and $X \times Y = \cup_k Z_k$. Next, if λ is a measure on $(X \times Y, \mathcal{F} \times \mathcal{G})$ satisfying (6.4), then λ and $\mu \times \nu$ coincide on $\mathcal{R}$. Theorem 1.26 ensures that λ and $\mu \times \nu$ coincide on $\sigma(\mathcal{R})$. □

The following result is a straightforward consequence of (6.3).

Corollary 6.5 *Under the same assumptions of Theorem 6.4, let $E \in \mathcal{F} \times \mathcal{G}$ be such that $(\mu \times \nu)(E) = 0$. Then, $\mu(E^y) = 0$ for ν–a.e. $y \in Y$, and $\nu(E_x) = 0$ for μ–a.e. $x \in X$.*

Example 6.6 We note that $\mu \times \nu$ may not be a complete measure even when both μ and ν are complete. Indeed, let λ denote Lebesgue measure on $X = [0, 1]$ and take $\mathcal{G}$ to be the σ–algebra of all Lebesgue measurable sets in $[0, 1]$ (that is, $\mathcal{G}$ consists of all additive sets, see Definition 1.33). Let $A \subset [0, 1]$ be a nonempty negligible set and let $B \subset [0, 1]$ be a set that is not measurable (see Example 1.52). Then, $A \times B \subset A \times [0, 1]$ and $(\lambda \times \lambda)(A \times [0, 1]) = 0$. On the other hand, $A \times B \notin \mathcal{G} \times \mathcal{G}$ for otherwise one would get a contradiction with Proposition 6.3(a).

6.1.2 Fubini-Tonelli Theorem

In this section we will reduce the computation of a double integral with respect to $\mu \times \nu$ to the computation of two simple integrals. The following two theorems are basic in the theory of multiple integration.

Theorem 6.7 (Tonelli) *Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be σ–finite measure spaces. Let $F : X \times Y \to [0, \infty]$ be a $(\mu \times \nu)$–measurable function. Then the following statements hold true.*

(a) (i) *For every $x \in X$ the function $\underbrace{y \mapsto F(x, y)}_{F(x,\cdot)}$ is ν–measurable.*

(ii) *For every $y \in Y$ the function $\underbrace{x \mapsto F(x, y)}_{F(\cdot,y)}$ is μ–measurable.*

(b) (i) *The function $x \mapsto \int_Y F(x, y)d\nu(y)$ is μ–measurable.*

(ii) *The function $y \mapsto \int_X F(x, y)d\mu(x)$ is ν–measurable.*

(c) *We have the identities*

$$\int_{X\times Y} F(x, y)d(\mu \times \nu)(x, y) = \int_X \left[\int_Y F(x, y)d\nu(y)\right] d\mu(x) \qquad (6.5)$$

$$= \int_Y \left[\int_X F(x, y)d\mu(x)\right] d\nu(y) \qquad (6.6)$$

Proof. Assume first that $F = \chi_E$ with $E \in \mathcal{F} \times \mathcal{G}$. Then,

$$F(x, \cdot) = \chi_{E_x} \qquad \forall x \in X,$$
$$F(\cdot, y) = \chi_{E^y} \qquad \forall y \in Y .$$

So, properties (a) and (b) follow from Proposition 6.3, while (c) reduces to formula (6.3), used to define product measure. Consequently, (a), (b), and (c) hold true when F is a simple function. In the general case, owing to Proposition 2.37 we can approximate F pointwise by an increasing sequence of simple functions

$$F_n : X \times Y \to [0, \infty] .$$

The $F_n(x, \cdot)$'s are themselves simple functions on Y such that

$$F_n(x, \cdot) \uparrow F(x, \cdot) \text{ pointwise as } n \to \infty \;\; \forall x \in X .$$

So, the function $F(x, \cdot)$ is ν–measurable and (a)(i) is proven. Moreover, $x \mapsto \int_Y F_n(x, y)d\nu(y)$ is an increasing sequence of positive simple functions satisfying

$$\int_Y F_n(x, y)d\nu(y) \quad \uparrow \quad \int_Y F(x, y)d\nu(y) \qquad \forall x \in X ,$$

thanks to the Monotone Convergence Theorem. Therefore, (b)(i) holds true and, again by monotone convergence,

$$\int_X \left[\int_Y F_n(x,y)d\nu(y)\right] d\mu(x) \quad \uparrow \quad \int_X \left[\int_Y F(x,y)d\nu(y)\right] d\mu(x) \,.$$

Since we also have that

$$\underbrace{\int_{X\times Y} F_n(x,y)d(\mu\times\nu)(x,y)}_{=\int_X\left[\int_Y F_n(x,y)d\nu(y)\right]d\mu(x)} \quad \uparrow \quad \int_{X\times Y} F(x,y)d(\mu\times\nu)(x,y)\,,$$

we have obtained (6.5). By a similar reasoning one can show (a)-(b)(ii) and (6.6). The proof is thus complete. □

Theorem 6.8 (Fubini) *Let* $(X,\mathcal{F},\mu)$, $(Y,\mathcal{G},\nu)$ *be* σ*–finite measure spaces and let* F *be a* $(\mu\times\nu)$*–summable function on* $X\times Y$*. Then the following statements hold true.*

(a) (i) *For* μ*–a.e.* $x\in X$ *the function* $y\mapsto F(x,y)$ *is* ν*–summable on* Y*.*

(ii) *For* ν*–a.e.* $y\in Y$ *the function* $x\mapsto F(x,y)$ *is* μ*–summable on* X*.*

(b) (i) *The function* $x\mapsto\int_Y F(x,y)d\nu(y)$ *is* μ*–summable on* X*.*

(ii) *The function* $y\mapsto\int_X F(x,y)d\mu(x)$ *is* ν*–summable on* Y*.*

(c) *Identities* (6.5) *and* (6.6) *are valid.*

Proof. Let F^+ and F^- be the positive and negative parts of F. Theorem 6.7 (c), applied to F^+ and F^-, yields identities (6.5) and (6.6). Also, we have that

$$\int_X \left[\int_Y F^\pm(x,y)d\nu(y)\right] d\mu(x) < \infty \qquad \int_Y \left[\int_X F^\pm(x,y)d\mu(x)\right] d\nu(y) < \infty$$

Therefore, (b) holds true for F^+ and F^-, hence for F. So, on account of Proposition 2.35,

- $x\mapsto\int_Y F^\pm(x,y)d\nu(y)$ is μ–a.e. finite;
- $y\mapsto\int_X F^\pm(x,y)d\mu(x)$ is ν–a.e. finite.

Consequently, (a) holds true and the proof is complete. □

Example 6.9 Let $X = Y = [-1, 1]$ with the Lebesgue measure and set

$$f(x, y) = \frac{xy}{(x^2 + y^2)^2}.$$

Observe that the iterated integrals exist and are equal; indeed

$$\int_{-1}^{1} dy \int_{-1}^{1} f(x, y)dx = \int_{-1}^{1} dx \int_{-1}^{1} f(x, y)dy = 0.$$

On the other hand the double integral fails to exist, since

$$\int_{[-1,1]^2} |f(x, y)|dxdy \geq \int_0^1 dr \int_0^{2\pi} \frac{|\sin\theta\cos\theta|}{r} d\theta = 2\int_0^1 \frac{dr}{r} = \infty.$$

This example shows that the existence of the iterated integrals does not imply the existence of the double integral.

Example 6.10 Consider the spaces

$$([0, 1], \mathcal{P}([0, 1]), \mu), \quad ([0, 1], \mathcal{B}([0, 1]), \lambda)$$

where μ denotes the counting measure and λ the Lebesgue measure. Consider the diagonal of $[0, 1]^2$, that is

$$\Delta = \{(x, x) \,|\, x \in [0, 1]\}.$$

For every $n \in \mathbb{N}$, set

$$Q_n = \left[0, \frac{1}{n}\right]^2 \cup \left[\frac{1}{n}, \frac{2}{n}\right]^2 \cup \ldots \cup \left[\frac{n-1}{n}, 1\right]^2.$$

Q_n is a finite union of measurable rectangles and $\Delta = \cap_n Q_n$, by which $\Delta \in \mathcal{P}([0, 1]) \times \mathcal{B}([0, 1])$. So the function χ_D is $(\mu \times \lambda)$-measurable. We have

$$\int_0^1 dy \int_0^1 f(x, y)d\mu(x) = \int_0^1 1\, dy = 1,$$

$$\int_0^1 d\mu(x) \int_0^1 f(x, y)dy = \int_0^1 0\, d\mu = 0.$$

Then, since μ is not σ-finite, the thesis of Tonelli's theorem fails.

Exercise 6.11 Show that

$$\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}).$$

6.2 Compactness in L^p

In this section we shall characterize all *relatively compact* subsets of $L^p(\mathbb{R}^N)$ [(2)] for any $1 \le p < \infty$, that is, all families of functions $\mathcal{M} \subset L^p(\mathbb{R}^N)$ whose closure $\overline{\mathcal{M}}$ in $L^p(\mathbb{R}^N)$ is compact. We shall see that two properties that were examined in chapter 3, namely tightness and continuity under translations, characterize relatively compact sets in $L^p(\mathbb{R}^N)$.

Definition 6.12 *Let $1 \le p < \infty$. For any $r > 0$ and $\varphi \in L^p(\mathbb{R}^N)$ define $S_r\varphi : \mathbb{R}^N \to \mathbb{R}$ by the Steklov formula*

$$S_r\varphi(x) = \frac{1}{\omega_N r^N} \int_{B(0,r)} \varphi(x+y)dy \qquad \forall x \in \mathbb{R}^N ,$$

where ω_N is the surface measure of the unit sphere in $\mathbb{R}^N$.

Proposition 6.13 *Let $1 \le p < \infty$ and $\varphi \in L^p(\mathbb{R}^N)$. Then for every $r > 0$ $S_r\varphi$ is a continuous function. Furthermore $S_r\varphi \in L^p(\mathbb{R}^N)$ and, using the notation $\tau_h\varphi(x) = \varphi(x+h)$, the following hold:*

$$|S_r\varphi(x)| \le \frac{1}{(\omega_N r^N)^{1/p}} \|\varphi\|_p; \tag{6.7}$$

$$\|S_r\varphi\|_p \le \|\varphi\|_p;$$

$$|S_r\varphi(x) - S_r\varphi(x+h)| < \frac{1}{(\omega_N r^N)^{1/p}} \|\varphi - \tau_h\varphi\|_p; \tag{6.8}$$

$$\|\varphi - S_r\varphi\|_p \le \sup_{0 \le |h| \le r} \|\varphi - \tau_h\varphi\|_p. \tag{6.9}$$

Proof. (6.7) can be derived using Hölder's inequality:

$$|S_r\varphi(x)| \le \frac{1}{(\omega_N r^N)^{1/p}} \left(\int_{B(0,r)} |\varphi(x+y)|^p dy \right)^{1/p}. \tag{6.10}$$

(6.8) follows from (6.7) applied to $\varphi - \tau_h\varphi$. Thus, (6.8) and Proposition 3.50 imply that $S_r\varphi$ is a continuous function. By (6.10), using Fubini's theorem we get

$$\int_{\mathbb{R}^N} |S_r\varphi|^p \, dx \le \frac{1}{\omega_N r^N} \int_{B(0,r)} \left(\int_{\mathbb{R}^N} |\varphi(x+y)|^p dx \right) dy$$
$$= \frac{\|\varphi\|_p^p}{\omega_N r^N} \int_{B(0,r)} dx = \|\varphi\|_p^p.$$

(2) $L^p(\mathbb{R}^N) = L^p(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \lambda)$ where λ is Lebesgue measure.

To obtain (6.9) observe that $(\varphi - S_r\varphi)(x) = \frac{1}{\omega_N r^N}\int_{B(0,r)}(\varphi(x)-\varphi(x+y))dy$, by which

$$|(\varphi - S_r\varphi)(x)| \le \frac{1}{(\omega_N r^N)^{1/p}}\left(\int_{B(0,r)}|\varphi(x)-\varphi(x+y)|^p dy\right)^{1/p}.$$

Therefore, Fubini's Theorem yields

$$\begin{aligned}\int_{\mathbb{R}^N}|\varphi - S_r\varphi|^p\,dx &\le \frac{1}{\omega_N r^N}\int_{\mathbb{R}^N}\left(\int_{B(0,r)}|\varphi(x)-\varphi(x+y)|^p dy\right)dx\\ &= \frac{1}{\omega_N r^N}\int_{B(0,r)}\left(\int_{\mathbb{R}^N}|\varphi(x)-\varphi(x+y)|^p dx\right)dy\end{aligned}$$

and (6.9) follows. □

Theorem 6.14 (M. Riesz) *Let $1 \le p < \infty$ and let $\mathcal{M}$ be a bounded family in $L^p(\mathbb{R}^N)$. Then, $\mathcal{M}$ is relatively compact iff*

$$\sup_{\varphi\in\mathcal{M}}\int_{|x|>R}|\varphi|^p dx \to 0 \quad as \quad R\to\infty \tag{6.11}$$

$$\sup_{\varphi\in\mathcal{M}}\int_{\mathbb{R}^N}|\varphi(x+h)-\varphi(x)|^p dx \to 0 \quad as \quad h\to 0 \tag{6.12}$$

Proof. Let us set $\tau_h\varphi(x) = \varphi(x+h)$ for any $x, h \in \mathbb{R}^N$. We already know that (6.11) and (6.12) hold for a single element of $L^p(\mathbb{R}^N)$ ((6.11) follows from Lebesgue Theorem; see Proposition 3.50 for (6.12)). If $\mathcal{M}$ is relatively compact, then for any $\varepsilon > 0$ there exist functions $\varphi_1, \dots, \varphi_m \in \mathcal{M}$ such that $\mathcal{M} \subset B_\varepsilon(\varphi_1)\cup\cdots\cup B_\varepsilon(\varphi_m)$. As we have just recalled, each φ_i satisfies (6.11) and (6.12). So, there exist $R_\varepsilon, \delta_\varepsilon > 0$ such that, for every $i = 1, \dots, m$,

$$\int_{|x|>R_\varepsilon}|\varphi_i|^p dx < \varepsilon^p \quad \& \quad \|\varphi_i - \tau_h\varphi_i\|_p < \varepsilon \quad \forall |h| < \delta_\varepsilon\,. \tag{6.13}$$

Let $\varphi \in \mathcal{M}$ and let φ_i be such that $\varphi \in B_\varepsilon(\varphi_i)$. Therefore, recalling (6.13), we have

$$\begin{aligned}\left(\int_{|x|>R_\varepsilon}|\varphi|^p dx\right)^{1/p} &\le \left(\int_{|x|>R_\varepsilon}|\varphi-\varphi_i|^p dx\right)^{1/p} + \left(\int_{|x|>R_\varepsilon}|\varphi_i|^p dx\right)^{1/p}\\ &\le \|\varphi-\varphi_i\|_p + \left(\int_{|x|>R_\varepsilon}|\varphi_i|^p dx\right)^{1/p} < 2\varepsilon\end{aligned}$$

and

$$\|\varphi - \tau_h\varphi\|_p \le \|\varphi - \varphi_i\|_p + \|\varphi_i - \tau_h\varphi_i\|_p + \|\tau_h\varphi_i - \tau_h\varphi\|_p < 3\varepsilon.$$

The necessity of (6.11) and (6.12) is thus proved.

To prove sufficiency it will suffice to show that $\mathcal{M}$ is totally bounded. Let $\varepsilon > 0$ be fixed. On account of assumption (6.11),

$$\exists R_\varepsilon > 0 \quad \text{such that} \quad \int_{|x|>R_\varepsilon} |\varphi|^p dx < \varepsilon^p \qquad \forall \varphi \in \mathcal{M}\,. \tag{6.14}$$

Also, recalling (6.9), assumption (6.12) yields

$$\exists \delta_\varepsilon > 0 \quad \text{such that} \quad \|\varphi - S_{\delta_\varepsilon}\varphi\|_p < \varepsilon \qquad \forall \varphi \in \mathcal{M}\,, \tag{6.15}$$

where S_{δ_ε} is the Steklov operator introduced in Definition (6.12). Moreover, properties (6.7) and (6.8) ensure that $\{S_{\delta_\varepsilon}\varphi\}_{\varphi\in\mathcal{M}}$ is a bounded equicontinuous family on $\overline{B}(0, R_\varepsilon)$. Thus, $\{S_{\delta_\varepsilon}\varphi\}_{\varphi\in\mathcal{M}}$ is relatively compact thanks to Ascoli-Arzelà's Theorem. Consequently, there exists a finite set of continuous functions $\{\psi_1, \ldots, \psi_m\}$ on $\overline{B}(0, R_\varepsilon)$ such that for each $\varphi \in \mathcal{M}$ the function $S_{\delta_e}\varphi : \overline{B}(0, R_\varepsilon) \to \mathbb{R}$ belongs to a ball of sufficiently small radius centered at ψ_i, say

$$|S_{\delta_\varepsilon}\varphi(x) - \psi_i(x)| < \frac{\varepsilon}{(\omega_N R_\varepsilon^N)^{1/p}} \qquad \forall x \in \overline{B}(0, R_\varepsilon)\,. \tag{6.16}$$

Set

$$\varphi_i(x) := \begin{cases} \psi_i(x) & |x| \le R_\varepsilon \\ 0 & |x| > R_\varepsilon\,. \end{cases}$$

Then, $\varphi_i \in L^p(\mathbb{R}^N)$ and, by (6.14), (6.15), and (6.16)

$$\begin{aligned} \|\varphi - \varphi_i\|_p &= \left(\int_{|x|>R_\varepsilon} |\varphi|^p dx\right)^{1/p} + \left(\int_{\overline{B}(0,R_\varepsilon)} |\varphi - \psi_i|^p dx\right)^{1/p} \\ &< \varepsilon + \left(\int_{\overline{B}(0,R_\varepsilon)} |\varphi - S_{\delta_\varepsilon}\varphi|^p dx\right)^{1/p} + \left(\int_{\overline{B}(0,R_\varepsilon)} |S_{\delta_\varepsilon}\varphi - \psi_i|^p dx\right)^{1/p} \\ &< 3\varepsilon. \end{aligned}$$

This shows that $\mathcal{M}$ is totally bounded and completes the proof. □

6.3 Convolution and approximation

In this section we will develop a systematic procedure for approximating a L^p function by smooth functions. The operation of convolution provides the tool to build such smooth approximations. The measure space of interest is $\mathbb{R}^N$ with Lebesgue measure λ.

6.3.1 Convolution Product

Definition 6.15 *Let f, $g : \mathbb{R}^N \to \overline{\mathbb{R}}$ be two Borel functions such that for a.e. $x \in \mathbb{R}^N$ the function*

$$y \in \mathbb{R}^N \mapsto f(x-y)g(y) \tag{6.17}$$

is summable. We define the convolution product *of f and g by*

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x-y)g(y)\,dy \qquad x \in \mathbb{R}^N \text{ a.e.}$$

Remark 6.16 1. If f, $g : \mathbb{R}^N \to [0,\infty]$ are Borel functions, then, since the function (6.17) is positive and Borel, $f * g : \mathbb{R}^N \to [0,\infty]$ is well defined for every $x \in \mathbb{R}^N$.

2. By making the change of variable $z = x - y$ and using the translation invariance of the Lebesgue measure we obtain that the function (6.17) is summable iff the function $z \in \mathbb{R}^N \mapsto f(z)g(x-z)$ is summable and $(f * g)(x) = (g * f)(x)$. This proves that the convolution is commutative.

Next proposition gives a sufficient condition to guarantee that $f * g$ is well-defined a.e. in $\mathbb{R}^N$.

Proposition 6.17 (Young) *Let $p, q, r \in [1,\infty]$ be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1 \tag{6.18}$$

and let $f \in L^p(\mathbb{R}^N)$ and $g \in L^q(\mathbb{R}^N)$. Then for a.e. $x \in \mathbb{R}^N$ the function (6.17) *is summable. Furthermore $f * g \in L^r(\mathbb{R}^N)$ and*

$$\|f * g\|_r \le \|f\|_p \, \|g\|_q. \tag{6.19}$$

Moreover, if $r = \infty$, then $f * g$ is a continuous function on $\mathbb{R}^N$.

Proof. First assume $\boxed{r = \infty}$; then $\frac{1}{p} + \frac{1}{q} = 1$. By the translation invariance of the Lebesgue measure we have that for every $x \in \mathbb{R}^N$ the function $y \in \mathbb{R}^N \to f(x-y)$ stays in $L^p(\mathbb{R}^N)$ and has the same L^p-norm as f. Then, by Hölder's inequality and Exercise 3.25 we deduce that for every $x \in \mathbb{R}^N$ the function (6.17) is summable and

$$|(f * g)(x)| \leq \|f\|_p \|g\|_q \quad \forall x \in \mathbb{R}^N. \tag{6.20}$$

Since p and q are conjugate, at least one of them is finite and, since the convolution is commutative, without loss of generality we may assume $p < \infty$. Then, for any $x, h \in \mathbb{R}^N$, the above estimate yields

$$|(f * g)(x+h) - (f * g)(x)| = |((\tau_h f - f) * g)(x)| \leq \|\tau_h f - f\|_p \|g\|_q$$

where $\tau_h f(x) = f(x+h)$. Since $\|\tau_h f - f\|_p \to 0$ as $h \to 0$ by Proposition 3.50, the continuity of $f * g$ follows; (6.19) can be derived immediately from (6.20).

Thus, assume $\boxed{r < \infty}$ (whence $p, q < \infty$). We will get the conclusion in four steps.

1. Suppose $f, g \geq 0$. Then $f * g : \mathbb{R}^N \to [0, +\infty]$ (see Remark 6.16.1) is a Borel function.

   Indeed the function

   $$F : \mathbb{R}^N \times \mathbb{R}^N \to [0, \infty] \qquad (x, y) \mapsto f(x-y)g(y)$$

   is Borel in the product space $\mathbb{R}^N \times \mathbb{R}^N$. Then Tonelli's Theorem ensures that the function $x \in \mathbb{R}^N \mapsto \int_{\mathbb{R}^N} F(x,y)dy = (f * g)(x)$ is Borel.

2. Suppose $p = 1 = q$ (whence $r = 1$). Then, $|f| * |g| \in L^1(\mathbb{R}^N)$ and $\||f| * |g|\|_1 = \|f\|_1 \, \|g\|_1$.

   Indeed, according to Step 1, $|f| * |g|$ is a Borel function and Tonelli's Theorem ensures that

   $$\begin{aligned} \int_{\mathbb{R}^N} (|f| * |g|)(x)\, dx &= \int_{\mathbb{R}^N} \left[\int_{\mathbb{R}^N} |f(x-y)g(y)|\, dy \right] dx \\ &= \int_{\mathbb{R}^N} |g(y)| \left[\int_{\mathbb{R}^N} |f(x-y)|\, dx \right] dy = \|f\|_1 \, \|g\|_1. \end{aligned}$$

   Therefore the thesis of Step 2 follows.

3. We claim that, for all $f \in L^p(\mathbb{R}^N)$ and $g \in L^q(\mathbb{R}^N)$,

$$(|f| * |g|)^r(x) \le \|f\|_p^{r-p}\ \|g\|_q^{r-q}(|f|^p * |g|^q)(x) \qquad \forall x \in \mathbb{R}^N. \qquad (6.21)$$

First assume $1 < p, q < \infty$ and let p' and q' be the conjugate exponents of p and q, respectively. Then,

$$\frac{1}{p'} + \frac{1}{q'} = 2 - \frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{r}\,.$$

Thus,

$$\begin{aligned} 1 - \frac{p}{r} &= p\left(1 - \frac{1}{q}\right) = \frac{p}{q'}, \\ 1 - \frac{q}{r} &= q\left(1 - \frac{1}{p}\right) = \frac{q}{p'}. \end{aligned}$$

Using the above relations for every $x, y \in \mathbb{R}^N$ we obtain

$$|f(x-y)g(y)| = (|f(x-y)|^p)^{1/q'}\,(|g(y)|^q)^{1/p'}\,(|f(x-y)|^p|g(y)|^q)^{1/r}\ ,$$

whence, by Exercise 3.7,

$$(|f| * |g|)(x) \le \|f\|_p^{p/q'}\ \|g\|_q^{q/p'}(|f|^p * |g|^q)^{1/r}(x) \qquad \forall x \in \mathbb{R}^N.$$

Since $rp/q' = r - p$ and $rq/p' = r - q$, (6.21) follows.

(6.21) is immediate for $p = 1 = q$.

Consider the case $p = 1$ and $1 < q < \infty$ (whence $r = q$). We have

$$|f(x-y)g(y)| = |f(x-y)|^{1/q'}\,(|f(x-y)||g(y)|^q)^{1/q}\ ,$$

Thus, by Hölder's inequality we get

$$(|f| * |g|)(x) \le \|f\|_p^{1/q'}\ (|f| * |g|^q)^{1/q}(x) \qquad x \in \mathbb{R}^N \text{ a.e.}.$$

The last case $q = 1$, $1 < p < \infty$ follows from the previous one since the convolution is commutative.

4. Owing to Step 1, $|f| * |g|$ is a Borel function and

$$\int_{\mathbb{R}^N} (|f| * |g|)^r dx \underbrace{\leq \|f\|_p^{r-p} \, \|g\|_q^{r-q} \| \, |f|^p * |g|^q \, \|_1}_{\text{by (6.21)}} \underbrace{= \|f\|_p^r \, \|g\|_q^r}_{\text{by step2}} . \tag{6.22}$$

Then $|f| * |g| \in L^r(\mathbb{R}^N)$, that is,

$$\int_{\mathbb{R}^N} \Big(\int_{\mathbb{R}^N} |f(x-y)g(y)| \, dy \Big)^r dx < \infty .$$

Therefore, $y \mapsto f(x-y)g(y)$ is summable for a.e. $x \in \mathbb{R}^N$. Hence, $f * g$ is well defined and a.e. finite. Since $f^+, f^- \in L^p(\mathbb{R}^N)$ and $g^+, g^- \in L^q(\mathbb{R}^N)$, then the functions $f^+ * g^+$, $f^- * g^-$, $f^+ * g^-$, $f^- * g^+$ are finite a.e. and, according to part 1, are Borel. Moreover we have

$$f * g = f^+ * g^+ + f^- * g^- - (f^+ * g^- + f^- * g^+) \qquad \text{a.e. } x \in \mathbb{R}^N .$$

We deduce that $f * g$ is Borel and

$$\int_{\mathbb{R}^N} |f * g|^r dx \leq \| |f| * |g| \|_r^r \underbrace{\leq \|f\|_p^r \, \|g\|_q^r}_{\text{by (6.22)}} .$$

□

Remark 6.18 For $r = \infty$ and $1 < p, q < \infty$ in (6.18),

$$\lim_{|x| \to \infty} (f * g)(x) = 0 .$$

Indeed, for $\varepsilon > 0$ let $R_\varepsilon > 0$ be such that

$$\int_{|y| \geq R_\varepsilon} |f(y)|^p dy < \varepsilon^p \quad \& \quad \int_{|y| \geq R_\varepsilon} |g(y)|^q dy < \varepsilon^q .$$

Then,

$$\begin{aligned} |(f * g)(x)| &\leq \Big| \int_{|y| \geq R_\varepsilon} f(x-y)g(y) \, dy \Big| + \Big| \int_{|y| < R_\varepsilon} f(x-y)g(y) \, dy \Big| \\ &\leq \|f\|_p \Big(\int_{|y| \geq R_\varepsilon} |g(y)|^q \, dy \Big)^{1/q} + \|g\|_q \Big(\int_{B(x, R_\varepsilon)} |f(z)|^p \, dz \Big)^{1/p} . \end{aligned}$$

Therefore, for all $|x| \geq 2R_\varepsilon$,

$$|(f * g)(x)| \leq \varepsilon (\|f\|_p + \|g\|_q) .$$

Remark 6.19 As a particular case of Young's Theorem, if $f \in L^1(\mathbb{R}^N)$ and $g \in L^p(\mathbb{R}^N)$ with $1 \leq p \leq \infty$, then $f * g$ is well defined and, further $f * g \in L^p(\mathbb{R}^N)$ with

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p. \tag{6.23}$$

Remark 6.20 By taking $p = 1$ in Remark 6.19, we obtain that the operation of convolution

$$* : L^1(\mathbb{R}^N) \times L^1(\mathbb{R}^N) \to L^1(\mathbb{R}^N)$$

provides a multiplication structure for $L^1(\mathbb{R}^N)$. This operation is commutative (see Remark 6.16.2) and associative. Indeed, if $f, g, h \in L^1(\mathbb{R}^N)$, then, by using the change of variables $z = t - y$ and by Fubini's Theorem

$$\begin{aligned}
((f * g) * h)(x) &= \int_{\mathbb{R}^N} (f * g)(x - y) h(y) dy \\
&= \int_{\mathbb{R}^N} h(y) dy \int_{\mathbb{R}^N} f(x - y - z) g(z) dz \\
&= \int_{\mathbb{R}^N} f(x - t) dt \int_{\mathbb{R}^N} g(t - y) h(y) dy \\
&= \int_{\mathbb{R}^N} f(x - t)(g * h)(t) dt = (f * (g * h))(x),
\end{aligned}$$

which proves the associativity. Finally, it is apparent that convolution obeys the distributive laws. However, there is not unit in $L^1(\mathbb{R}^N)$ under this multiplication. Indeed, assume by absurd the existence of $g \in L^1(\mathbb{R}^N)$ such that $g * f = f$ for every $f \in L^1(\mathbb{R}^N)$. Then the absolute continuity of the integral implies the existence of $\delta > 0$ such that

$$A \in \mathcal{B}(\mathbb{R}^N) \ \& \ \lambda(A) \leq \delta \Rightarrow \int_A |g| dx < 1.$$

Let $\rho > 0$ be sufficiently small such that $\lambda(B(0, \rho)) < \delta$ and, taking $f = \chi_{B(0,\rho)} \in L^1(\mathbb{R}^N)$, for every $x \in \mathbb{R}^N$ we compute

$$\begin{aligned}
|f(x)| = |(g * f)(x)| &\leq \int_{\mathbb{R}^N} |g(x - y)| \, |f(y)| dy = \int_{B(0,\rho)} |g(x - y)| dy \\
&= \int_{B(x,\rho)} |g(z)| dz < 1
\end{aligned}$$

and the contradiction follows.

Exercise 6.21 Compute $f * g$ for $f(x) = \chi_{[-1,1]}(x)$ and $g(x) = e^{-|x|}$.

6.3.2 Approximation by smooth functions

Definition 6.22 *A family* $(f_\varepsilon)_\varepsilon$ *in* $L^1(\mathbb{R}^N)$ *is called an* approximate identity *if satisfies the following*

$$f_\varepsilon \geq 0, \qquad \int_{\mathbb{R}^N} f_\varepsilon(x)dx = 1 \quad \forall \varepsilon > 0, \tag{6.24}$$

$$\forall \delta > 0 : \quad \int_{|x|\geq\delta} f_\varepsilon(x)dx \to 0 \ \textit{as}\ \varepsilon \to 0^+. \tag{6.25}$$

Remark 6.23 A common way to produce approximate identities in $L^1(\mathbb{R}^N)$ is to take a function $f \in L^1(\mathbb{R}^N)$ such that $f \geq 0$ and $\int_{\mathbb{R}^N} f(x)dx = 1$ and to define for $\varepsilon > 0$

$$f_\varepsilon(x) = \varepsilon^{-N} f(\varepsilon^{-1}x).$$

Condition (6.24)-(6.25) are satisfied since, introducing the change of variables $y = \varepsilon^{-1}x$, we obtain

$$\int_{\mathbb{R}^N} f_\varepsilon(x)dx = \int_{\mathbb{R}^N} f(y)dy = 1$$

and

$$\int_{|x|\geq\delta} f_\varepsilon(x)dx = \int_{|y|\geq\varepsilon^{-1}\delta} f(y)dy \to 0 \text{ as } \varepsilon \to 0^+,$$

the latter convergence is by the Lebesgue dominated convergence theorem.

Proposition 6.24 *Let* $(f_\varepsilon)_\varepsilon \subset L^1(\mathbb{R}^N)$ *be an approximate identity. Then the following hold*

1. *If* $f \in L^\infty(\mathbb{R}^N)$ *and* f *is continuous in* $\mathbb{R}^N$, *then* $f * f_\varepsilon \to f$ *uniformly on compact sets of* $\mathbb{R}^N$ *as* $\varepsilon \to 0^+$;

2. *If* $f \in L^\infty(\mathbb{R}^N)$ *and* f *is uniformly continuous in* $\mathbb{R}^N$, *then* $f * f_\varepsilon \xrightarrow{L^\infty} f$ *as* $\varepsilon \to 0^+$;

3. *If* $1 \leq p < \infty$ *and* $f \in L^p(\mathbb{R}^N)$, *then* $f * f_\varepsilon \xrightarrow{L^p} f$ *as* $\varepsilon \to 0^+$.

Proof. 1. By Young's theorem we get that $f * f_\varepsilon$ is continuous and $f * f_\varepsilon \in L^\infty(\mathbb{R}^N)$. Let $K \subset \mathbb{R}^N$ be a compact set. Hence the set $\{x \in \mathbb{R}^N \mid d_K(x) \leq 1\}$

is compact and f is uniformly continuous over it; then, given $\eta > 0$, there exists $\delta \in (0,1)$ such that

$$|f(x-y) - f(x)| \le \eta \quad \forall x \in K \quad \forall y \in B(x,\delta).$$

Since $\int_{\mathbb{R}^N} f_\varepsilon(y)dy = 1$, for every $x \in K$ we have

$$\begin{aligned} |(f * f_\varepsilon)(x) - f(x)| &= \Big| \int_{\mathbb{R}^N} \big(f(x-y) - f(x)\big) f_\varepsilon(y)dy \Big| \\ &\le \int_{|y|<\delta} \big|f(x-y) - f(x)\big| f_\varepsilon(y)dy \\ &\quad + \int_{|y|\ge\delta} \big|f(x-y) - f(x)\big| f_\varepsilon(y)dy \qquad (6.26) \\ &\le \eta \int_{\mathbb{R}^N} f_\varepsilon(y)dy + 2\|f\|_\infty \int_{|y|\ge\delta} f_\varepsilon(y)dy \\ &= \eta + 2\|f\|_\infty \int_{|y|\ge\delta} f_\varepsilon(y)dy. \end{aligned}$$

The conclusion follows from (6.25).

2. The proof is the same as in Part 1 except that in this case estimate (6.26) holds for every $x \in \mathbb{R}^N$.

3. According to Remark 6.19 $f * f_\varepsilon \in L^p(\mathbb{R}^N)$ for all $\varepsilon > 0$. Since $\int_{\mathbb{R}^N} f_\varepsilon(y)dy = 1$, for every $x \in \mathbb{R}^N$ we have

$$\begin{aligned} |(f * f_\varepsilon)(x) - f(x)| &= \Big| \int_{\mathbb{R}^N} \big(f(x-y) - f(x)\big) f_\varepsilon(y)dy \Big| \\ &\le \int_{\mathbb{R}^N} |f(x-y) - f(x)| f_\varepsilon(y)dy. \end{aligned} \qquad (6.27)$$

If $p > 1$, let $p' \in (1,\infty)$ be the conjugate exponent of p. Then

$$|(f * f_\varepsilon)(x) - f(x)| \le \int_{\mathbb{R}^N} |f(x-y) - f(x)| (f_\varepsilon(y))^{1/p} (f_\varepsilon(y))^{1/p'} dy.$$

By applying Hölder's inequality we obtain

$$\begin{aligned} |(f * f_\varepsilon)(x) - f(x)|^p &\le \left(\int_{\mathbb{R}^N} |f(x-y) - f(x)|^p f_\varepsilon(y)dy \right) \left(\int_{\mathbb{R}^N} f_\varepsilon(y)dy \right)^{p/p'} \\ &= \int_{\mathbb{R}^N} |f(x-y) - f(x)|^p f_\varepsilon(y)dy. \end{aligned}$$

Combing this with (6.27) we deduce that the following inequality holds for $1 \le p < \infty$:

$$|(f * f_\varepsilon)(x) - f(x)|^p \le \int_{\mathbb{R}^N} |f(x-y) - f(x)|^p f_\varepsilon(y)dy.$$

After integration over $\mathbb{R}^N$, by applying Tonelli's Theorem, we have

$$\|f * f_\varepsilon - f\|_p^p \le \int_{\mathbb{R}^N} \|\tau_{-y} f - f\|_p^p f_\varepsilon(y)dy$$

where $\tau_{-y}f(x) = f(x-y)$. Setting $\Delta(y) = \|\tau_{-y}f - f\|_p$, the above inequality becomes

$$\|f * f_\varepsilon - f\|_p^p \le (\Delta^p * f_\varepsilon)(0)$$

For every y, $y_0 \in \mathbb{R}^N$ by using the translation invariance of the Lebesgue measure

$$\begin{aligned}|\Delta(y) - \Delta(y_0)| &= \big| \|\tau_{-y}f - f\|_p - \|\tau_{-y_0}f - f\|_p \big| \le \|\tau_{-y}f - \tau_{-y_0}f\|_p \\ &= \|\tau_{-y+y_0}f - f\|_p \to 0 \text{ as } y \to y_0;\end{aligned}$$

the latter fact follows by Proposition 3.50. Hence Δ is a continuous function. Since $\Delta^p(y) \le 2^p\|f\|_p^p$, then $\Delta^p \in L^\infty(\mathbb{R}^N)$. By part 1 we conclude $(\Delta^p * f_\varepsilon)(0) \to \Delta^p(0) = 0$. □

Notation 6.25 *Let $\Omega \subset \mathbb{R}^N$ be an open set and $k \in \mathbb{N}$. $\mathcal{C}^k(\Omega)$ is the space of the functions $f : \Omega \to \mathbb{R}$ which are k times continuously differentiable, $\mathcal{C}_c(\Omega)$ is the space of the continuous functions $f : \Omega \to \mathbb{R}$ which are zero outside a compact set $K \subset \Omega$, and*

$$\mathcal{C}^\infty(\Omega) = \cap_k \mathcal{C}^k(\Omega), \quad \mathcal{C}_c^k(\Omega) = \mathcal{C}^k(\Omega) \cap \mathcal{C}_c(\Omega), \quad \mathcal{C}_c^\infty(\Omega) = \mathcal{C}^\infty(\Omega) \cap \mathcal{C}_c(\Omega).$$

In particular, if $k = 0$, $\mathcal{C}^0(\Omega) = \mathcal{C}(\Omega)$ is the space of the continuous functions $f : \Omega \to \mathbb{R}$. If $f \in \mathcal{C}^k(\Omega)$ and $\alpha = (\alpha_1, \ldots, \alpha_N)$ is a multiindex such that $|\alpha| := \alpha_1 + \ldots + \alpha_N \le k$, then we set

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \ldots \partial x_N^{\alpha_N}}.$$

If $\alpha = (0, \ldots, 0)$, we set $D^0 f = f$.

Proposition 6.26 *Let $f \in L^1(\mathbb{R}^N)$ and $g \in \mathcal{C}^k(\mathbb{R}^N)$ such that $D^\alpha g \in L^\infty(\mathbb{R}^N)$ for every $\alpha \in \mathbb{N}^N$ such that $0 \le |\alpha| \le k$. Then $f * g \in \mathcal{C}^k(\mathbb{R}^N)$ and*

$$D^\alpha(f * g) = f * D^\alpha g \quad \forall \alpha \in \mathbb{N}^N \ s.t. \ 0 \le |\alpha| \le k.$$

Proof. The continuity of $f * g$ follows from Young's Theorem. By induction it will be sufficient to prove the thesis when $k = 1$. Setting

$$\varphi(x, y) = f(y)g(x - y),$$

we have

$$\Big|\frac{\partial \varphi}{\partial x_i}(x, y)\Big| = |f(y)\frac{\partial g}{\partial x_i}(x - y)| \le \Big\|\frac{\partial g}{\partial x_i}\Big\|_\infty |f(y)|$$

Since $(f * g)(x) = \int_{\mathbb{R}^N} \varphi(x, y)dy$, Proposition 2.75 implies that $f * g$ is differentiable and

$$\frac{\partial(f * g)}{\partial x_i}(x) = \int_{\mathbb{R}^N} f(y)\frac{\partial g}{\partial x_i}(x - y)dy = \Big(f * \frac{\partial g}{\partial x_i}\Big)(x).$$

By hypothesis $\frac{\partial g}{\partial x_i} \in \mathcal{C}(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$. Again Young's Theorem implies $f * \frac{\partial g}{\partial x_i} \in \mathcal{C}(\mathbb{R}^N)$; hence $f * g \in \mathcal{C}^1(\mathbb{R}^N)$. □

Thus convolution with a smooth function produces a smooth function. This fact provides us with a powerful technique to prove a variety of density theorems.

Definition 6.27 *For every $\varepsilon > 0$ define the function $\rho_\varepsilon : \mathbb{R}^N \to \mathbb{R}$ by*

$$\rho_\varepsilon(x) = \begin{cases} C\varepsilon^{-N} \exp\left(\dfrac{\varepsilon^2}{|x|^2 - \varepsilon^2}\right) & \text{if } |x| < \varepsilon, \\ 0 & \text{if } |x| \ge \varepsilon \end{cases}$$

where $C = \big(\int_{|x|<1} \exp\big(\frac{1}{|x|^2-1}\big)dx\big)^{-1}$. The family $(\rho_\varepsilon)_\varepsilon$ is called the standard mollifier.

Lemma 6.28 *The standard mollifier $(\rho_\varepsilon)_\varepsilon$ satisfies*

$$\rho_\varepsilon \in \mathcal{C}_c^\infty(\mathbb{R}^N), \quad \text{supp}(\rho_\varepsilon) = \overline{B}(0, \varepsilon) \quad \forall \varepsilon > 0;$$

$$(\rho_\varepsilon)_\varepsilon \text{ is an approximate identity.}$$

Proof. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(t) = \begin{cases} \exp\left(\dfrac{1}{t-1}\right) & \text{if } t < 1 \\ 0 & \text{if } t \geq 1 \end{cases}$$

Then f is a $\mathcal{C}^\infty$ function. Indeed we only need to check the smoothness at $t = 1$. As $t \uparrow 1$ all the derivatives are zero. As $t \uparrow 1$ the derivatives are finite linear combination of terms of the form $\frac{1}{(t-1)^l} \exp\left(\frac{1}{t-1}\right)$, l being an integer greater than or equal to zero and these terms tend to zero as $t \uparrow 1$.

Observe that for every $\varepsilon > 0$

$$\rho_\varepsilon(x) = \frac{1}{\varepsilon^N} \rho_1\left(\frac{x}{\varepsilon}\right) = C \frac{1}{\varepsilon^N} f\left(\frac{|x|^2}{\varepsilon}\right).$$

Then $\rho_\varepsilon \in \mathcal{C}_c^\infty(\mathbb{R}^N)$ and $\text{supp}(\rho_\varepsilon) = \overline{B}(0, \varepsilon)$. Further the definition of C implies $\int_{\mathbb{R}^N} \rho_1(x) dx = 1$. Remark 6.23 allows us to conclude. □

Lemma 6.29 *Let f, $g \in \mathcal{C}_c(\mathbb{R}^N)$. Then $f * g \in \mathcal{C}_c(\mathbb{R}^N)$ and*

$$\text{supp}(f * g) \subset \text{supp}(f) + \text{supp}(g),$$

where for sets A and B of $\mathbb{R}^N$: $A + B = \{x + y \,|\, x \in A,\, y \in B\}$.

Proof. By Proposition 6.26 we get $f * g \in \mathcal{C}(\mathbb{R}^N)$. Set $A = \text{supp}(f)$, $B = \text{supp}(g)$. For every $x \in \mathbb{R}^N$ we have

$$(f * g)(x) = \int_{(x - \text{supp}(f)) \cap \text{supp}(g)} f(x - y) g(y) dy.$$

In order to obtain that $(f * g)(x) \neq 0$, necessarily $(x - \text{supp}(f)) \cap \text{supp}(g) \neq \emptyset$, that is $x \in \text{supp}(f) + \text{supp}(g)$. □

Proposition 6.30 *Let $\Omega \subset \mathbb{R}^N$ be an open set. Then*

- *space $\mathcal{C}_c^\infty(\Omega)$ is dense in $\mathcal{C}_0(\Omega)$* [(3)]*;*
- *space $\mathcal{C}_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for every $1 \leq p < \infty$.*

[(3)] see Exercise 3.46 for the definition of $\mathcal{C}_0(\Omega)$.

Proof. According to Theorem 3.45 and Exercise 3.46 it is sufficient to prove that, given $f \in \mathcal{C}_c(\Omega)$, there exists a sequence $(f_n)_n \subset \mathcal{C}_c^\infty(\Omega)$ such that $f_n \xrightarrow{L^\infty} f$ and $f_n \xrightarrow{L^p} f$. Indeed, fixed $f \in \mathcal{C}_c(\Omega)$, set

$$\tilde{f} = \begin{cases} f(x) & \text{if } x \in \Omega, \\ 0 & \text{if } x \in \mathbb{R}^N \setminus \Omega. \end{cases}$$

Then $\tilde{f} \in \mathcal{C}_c(\mathbb{R}^N)$. Let $(\rho_\varepsilon)_\varepsilon$ be the mollifier constructed in Definition 6.27 and for every n define $f_n := f * \rho_{1/n}$. According to Proposition 6.26 $f_n \in \mathcal{C}^\infty(\mathbb{R}^N)$. Next let $K = \mathrm{supp}(f)$ and $\eta = \inf_{x \in K} d_{\partial\Omega}(x) > 0$. Then $\tilde{K} := \{x \in \mathbb{R}^N \mid d_K(x) \le \frac{\eta}{2}\}$ is a compact set and $\tilde{K} \subset \Omega$. By Proposition 6.29, if n is such that $\frac{1}{n} < \frac{\eta}{2}$ we obtain

$$\mathrm{supp}(f_n) \subset K + \overline{B}\Big(0, \frac{1}{n}\Big) = \Big\{x \in \mathbb{R}^N \,\Big|\, d_K(x) \le \frac{1}{n}\Big\} \subset \tilde{K}.$$

Then $f_n \in \mathcal{C}_c^\infty(\Omega)$ for n sufficiently large. Since f is uniformly continuous, Proposition 6.24.2 gives $f_n \xrightarrow{L^\infty} \tilde{f}$ in $L^\infty(\mathbb{R}^N)$, which implies

$$f_n \to f \text{ in } L^\infty(\Omega).$$

Finally, for large n,

$$\int_\Omega |f_n - f|^p dx = \int_{\tilde{K}} |f_n - f|^p dx \le \lambda(\tilde{K}) \|f_n - f\|_\infty^p \to 0.$$

□

An interesting consequence of smoothing properties of convolution is the following *Weierstrass approximation Theorem.*

Theorem 6.31 (Weierstrass) *Let $f \in \mathcal{C}_c(\mathbb{R}^N)$. Then there exists a sequence of polynomials $(p_n)_n$ such that $p_n \to f$ uniformly on compact sets of $\mathbb{R}^N$.*

Proof. For every $\varepsilon > 0$ define

$$u_\varepsilon(x) = \varepsilon^{-N} u(\varepsilon^{-1} x), \quad x \in \mathbb{R}^N,$$

where

$$u(x) = \pi^{-N/2} \exp(-|x|^2), \quad x \in \mathbb{R}^N.$$

The well-known *Poisson formula*

$$\int_{\mathbb{R}^N} \exp(-|x|^2)dx = \pi^{N/2}$$

and Remark 6.23 imply that $(u_\varepsilon)_\varepsilon$ is an approximate identity. Theorem 6.24.2 yields

$$u_\varepsilon * f \xrightarrow{L^\infty} f \text{ as } \varepsilon \to 0. \tag{6.28}$$

Fix $\varepsilon > 0$ and let $K \subset \mathbb{R}^N$ be a compact set. We claim that there exists a sequence of polynomials $(P_n)_n$ such that

$$P_n \to u_\varepsilon * f \text{ uniformly in } K. \tag{6.29}$$

Indeed the function u_ε is analytic and so on any compact set can be uniformly approximated by the partial sums of its Taylor series which are polynomials. The set $\tilde{K} := K - \text{supp}(f)$ is compact, then there exists a sequence $(p_n)_n$ of polynomials on $\mathbb{R}^N$ such that $p_n \to u_\varepsilon$ uniformly in $\tilde{K}$. Next set

$$P_n(x) = \int_{\mathbb{R}^N} p_n(x-y)f(y)dy. \tag{6.30}$$

Since f is compactly supported, then the integrand in (6.30) is bounded by $|f| \sup_{y\in \text{supp}(f)} |p_n(x-y)|$ which is summable for every $x \in \mathbb{R}^N$. Then P_n is well defined on $\mathbb{R}^N$. Observe that $p_n(x-y)$ is a polynomial in the variables (x, y), that is $p_n(x-y) = \sum_{k=1}^K q_k(x)s_k(y)$ with q_k, s_k polynomials in $\mathbb{R}^N$; substituting in (6.30) we obtain that each P_n is also a polynomial. Furthermore for every $x \in K$

$$\begin{aligned} |P_n(x) - (u_\varepsilon * f)(x)| &\le \int_{\text{supp}(f)} |p_n(x-y) - u_\varepsilon(x-y)||f(y)|dy \\ &\le \sup_{t\in\tilde{K}} |p_n(t) - u_\varepsilon(t)| \int_{\mathbb{R}^N} |f(y)|dy \end{aligned}$$

and (6.29) follows.

To conclude, consider a sequence $\varepsilon_n \to 0^+$. For every $n \in \mathbb{N}$, since the set $\{|x| \le n|\}$ is compact, we can find a polynomial Q_n such that

$$\sup_{|x|\le n} |Q_n(x) - (u_{\varepsilon_n} * f)(x)| \le \varepsilon_n.$$

If $K \subset \mathbb{R}^N$ is a compact set, then for n sufficiently large $K \subset B(0,n)$, which implies

$$\sup_{x\in K} |Q_n(x) - f(x)| \le \sup_{x\in K} |Q_n(x) - (u_{\varepsilon_n} * f)(x)| + \sup_{x\in K} |(u_{\varepsilon_n} * f)(x) - f(x)|$$
$$\le \varepsilon_n + \|(u_{\varepsilon_n} * f) - f\|_\infty \to 0$$

by (6.28). □

Corollary 6.32 *Let $A \in \mathcal{B}(\mathbb{R}^N)$ be a bounded set and $1 \le p < \infty$. Then the set $\mathcal{P}_A$ of all polynomials defined on A is dense in $L^p(A)$.*

Proof. Consider $f \in L^p(A)$ and let $\tilde{f}$ be the extension of f by zero outside A. Then $\tilde{f} \in L^p(\mathbb{R}^N)$; fixed $\varepsilon > 0$, Proposition 6.30 implies the existence of $g \in \mathcal{C}_c(\mathbb{R}^N)$ such that $\int_{\mathbb{R}^N} |\tilde{f} - g|^p dx \le \varepsilon$. Since $\bar{A}$ is a compact set, by Theorem 6.31 we get the existence of a polynomial p such that $\sup_{x\in\bar{A}} |p(x) - g(x)| \le \left(\frac{\varepsilon}{\lambda(A)}\right)^{1/p}$. Then

$$\int_A |g(x) - p(x)|^p dx \le \Big(\sup_{x\in\bar{A}} |g(x) - p(x)|\Big)^p \lambda(A) \le \varepsilon,$$

by which

$$\int_A |f(x) - p(x)|^p dx \le 2^{p-1} \int_A |f(x) - g(x)|^p dx + 2^{p-1} \int_A |g(x) - p(x)|^p dx$$
$$\le 2^{p-1} \int_{\mathbb{R}^N} |\tilde{f}(x) - g(x)|^p dx + 2^{p-1}\varepsilon \le 2^p \varepsilon.$$

□

Remark 6.33 By Corollary 6.32 we deduce that if $A \in \mathcal{B}(\mathbb{R}^N)$ is a bounded set, then the set of all polynomials defined on A with rational coefficients is countable and everywhere dense in $L^p(A)$ for $1 \le p < \infty$ (see Proposition 3.47).

Chapter 7

Functions of bounded variation and absolutely continuous functions

Let f and F be two functions on $[a, b]$ such that f is continuous and F has a continuous derivative. Then it will be recalled from elementary calculus that the connection between the operations of differentiation and integration is expressed by the familiar formulas

$$\frac{d}{dx}\int_a^x f(t)dt = f(x), \tag{7.1}$$

$$\int_a^x F'(t)dt = F(x) - F(a). \tag{7.2}$$

This immediately suggests:

1. Does (7.1) continue to hold almost everywhere for an arbitrary summable function f?

2. What is the largest class of functions for which (7.2) holds?

These questions will be answered in this chapter. We observe that if f is nonnegative, then the indefinite Lebesgue integral

$$\int_a^x f(t)dt, \quad x \in [a, b], \tag{7.3}$$

as a function of its upper limit, is nondecreasing. Moreover, since every summable function f is the difference of two nonnegative summable functions f^+ and f^-, the integral (7.3) is the difference between two nondecreasing functions. Hence, the study of the indefinite Lebesgue integral is closely related to the study of monotonic functions. Monotonic functions have a number of simple and important properties which we now discuss.

7.1 Monotonic functions

Definition 7.1 *A function $f : [a, b] \to \mathbb{R}$ is said to be* nondecreasing *if $a \leq x_1 \leq x_2 \leq b$ implies $f(x_1) \leq f(x_2)$ and* nonincreasing *if $a \leq x_1 \leq x_2 \leq b$ implies $f(x_1) \geq f(x_2)$. By a* monotonic *function is meant a function which is either nondecreasing or nonincreasing.*

Definition 7.2 *Given a monotonic function $f : [a, b] \to \mathbb{R}$ and $x_0 \in [a, b)$, the limit*

$$f(x_0^+) := \lim_{h \to 0,\, h>0} f(x_0 + h)$$

(which always exists) is said to be the right hand limit *of f at the point x_0. Similarly, if $x_0 \in (a, b]$, the limit*

$$f(x_0^-) = \lim_{h \to 0,\, h>0} f(x_0 - h)$$

is called the left-hand limit *of f at x_0.*

Remark 7.3 *Let f be nondecreasing on $[a, b]$. If $a \leq x < y \leq b$, then*

$$f(x^+) \leq f(y^-).$$

Analogously, if f is nonincreasing on $[a, b]$ and $a \leq x < y \leq b$, then

$$f(x^+) \geq f(y^-).$$

We now establish the basic properties of monotonic functions.

Theorem 7.4 *Every monotonic function f on $[a, b]$ is Borel and bounded, and hence summable.*

Proof. Assume that f is nondecreasing. Since $f(a) \leq f(x) \leq f(b)$ for all $x \in [a, b]$, f is obviously bounded. For every $c \in \mathbb{R}$ consider the set

$$E_c = \{x \in [a, b] \mid f(x) < c\}.$$

If E_c is empty, then E_c is (trivially) a Borel set. If E_c is nonempty, let y be the least upper bound of all $x \in E_c$. Then E_c is either the closed interval $[a, y]$, if $y \in E_c$, or the half-open interval $[a, y)$, if $y \notin E_c$. In either case, E_c is a Borel set; this proves that f is Borel. Finally we have

$$\int_a^b |f(x)|dx \leq \max\{|f(a)|, |f(b)|\}(b - a),$$

by which f is summable. □

Theorem 7.5 *Let $f : [a, b] \to \mathbb{R}$ be a monotonic function. Then the set of points of $[a, b]$ at which f is discontinuous is at most countable.*

Proof. Suppose, for the sake of definiteness, that f is nondecreasing, and let E be the set of points at which f is discontinuous. If $x \in E$ we have $f(x^-) < f(x^+)$; then with every point x of E we associate we associate a rational number $r(x)$ such that

$$f(x^-) < r(x) < f(x^+).$$

Since by Remark 7.3 $x_1 < x_2$ implies $f(x_1^+) \leq f(x_2^-)$, we see that $r(x_1) \neq r(x_2)$. We have thus established a 1-1 correspondence between the set E and a subset of the rational numbers. □

7.1.1 Differentiation of a monotonic function

The key result of this section will be to show that a monotonic function f defined on an interval $[a, b]$ has a finite derivative almost everywhere in $[a, b]$. Before proving this proposition, due to Lebesgue, we must first introduce some further notation. For every $x \in (a, b)$ the following four quantities (which may take infinite values) always exist:

$$D'_L f(x) = \liminf_{h \to 0,\, h<0} \frac{f(x+h) - f(x)}{h}, \quad D''_L f(x) = \limsup_{h \to 0,\, h<0} \frac{f(x+h) - f(x)}{h},$$

$$D'_R f(x) = \liminf_{h \to 0,\, h>0} \frac{f(x+h) - f(x)}{h}, \quad D''_R f(x) = \limsup_{h \to 0,\, h>0} \frac{f(x+h) - f(x)}{h}.$$

These four quantities are called the *derived numbers* of f at x. It is clear that the inequalities

$$D'_L f(x) \leq D''_L f(x), \quad D'_R f(x) \leq D''_R f(x) \tag{7.4}$$

always hold. If $D'_L f(x)$ and $D''_L f(x)$ are finite and equal, their common value is just the left-hand derivative of f at x. Similarly, if $D'_R f(x)$ and $D''_R f(x)$ are finite and equal, their common value is just the right-hand derivative of f at x. Moreover, f has a derivative at x if and only if all four derived numbers $D'_L f(x)$, $D''_L f(x)$, $D'_R f(x)$ and $D''_R f(x)$ are finite and equal.

Theorem 7.6 (Lebesgue) *Let $f : [a, b] \to \mathbb{R}$ be a monotonic function. Then f has a derivative almost everywhere on $[a, b]$. Furthermore $f' \in L^1([a, b])$ and*

$$\int_a^b |f'(t)| dt \leq |f(b) - f(a)|. \tag{7.5}$$

Proof. There is no loss of generality in assuming that f is nondecreasing, since if f is nonincreasing, we can apply the result to $-f$ which is obviously nondecreasing. We begin by proving that the derived numbers of f are equal (with possibly infinite value) almost everywhere on $[a, b]$. It will be enough to show that the inequality

$$D'_L f(x) \geq D''_R f(x) \tag{7.6}$$

holds almost everywhere on $[a, b]$. In fact, setting, $f^*(x) = -f(-x)$, we see that f^* in nondecreasing on $[-b, -a]$; moreover, it is easily verified that

$$D'_L f^*(x) = D'_R f(-x), \quad D''_L f^*(x) = D''_R f(-x).$$

Therefore, applying (7.6) to f^*, we get

$$D'_L f^*(x) \geq D''_R f^*(x)$$

or

$$D'_R f(x) \geq D''_L f(x).$$

Combining this inequality with (7.6), we obtain

$$D''_R f \leq D'_L f \leq D''_L f \leq D'_R f \leq D''_R f,$$

after using (7.4), and the equality of the four derived numbers follows. To prove that (7.6) holds almost everywhere, observe that the set of points where $D_L^- f < D_R^+ f$ can clearly be represented as the union over $u, v \in \mathbb{Q}$ with $v > u > 0$ of the sets

$$E_{u,v} = \{x \in (a, b) \mid D_R'' f(x) > v > u > D_L' f(x)\}.$$

It will then follow that (7.6) holds almost everywhere, if we succeed in showing that $\lambda(E_{u,v}) = 0$. Let $s = \lambda(E_{u,v})$. Then, given $\varepsilon > 0$, according to Proposition 1.53 there is an open set A such that $E_{u,v} \subset A$ and $\lambda(A) < s+\varepsilon$. For every $x \in E_{u,v}$ and $\delta > 0$, since $D_L' f(x) < u$, there exists $h_{x,\delta} \in (0, \delta)$ such that $[x - h_{x,\delta}, x] \subset A$ and

$$f(x) - f(x - h_{x,\delta}) < u h_{x,\delta}.$$

Since the collection of closed intervals $([x - h_{x,\delta}, x])_{x\in(a,b),\,\delta>0}$ is a fine cover of $E_{u,v}$, by Vitali's covering lemma there exists a finite number of disjoint intervals of such collection, say

$$I_1 := [x_1 - h_1, x_1], \ldots, I_N := [x_N - h_N, x_N],$$

such that, setting $B = E_{u,v} \cap \bigcup_{i=1}^N (x_i - h_i, x_i)$,

$$\lambda(B) = \lambda\Big(E_{u,v} \cap \bigcup_{i=1}^N I_k\Big) > s - \varepsilon.$$

Summing up over these intervals we get

$$\sum_{i=1}^N \big(f(x_i) - f(x_i - h_i)\big) < u \sum_{i=1}^N h_i < u\lambda(A) < u(s + \varepsilon). \tag{7.7}$$

Now we reason as above and use the inequality $D_R'' f(x) > v$; for every $y \in B$ and $\eta > 0$, since $D_R'' f(x) > v$, there exists $k_{y,\eta} \in (0, \eta)$ such that $[y, y+k_{y,\eta}] \subset I_i$ for some $i \in \{1, \ldots, N\}$ and

$$f(y + k_{y,\eta}) - f(y) > v k_{y,\eta}.$$

Since the collection of closed intervals $([y, y+k_{y,\eta}])_{y\in B,\,\eta>0}$ is a fine cover of B, by Vitali's covering lemma there exists a finite number of disjoint intervals of such collection, say

$$J_1 := [y_1, y_1 + k_1], \ldots, J_M := [y_M, y_M + k_M],$$

such that,

$$\lambda\Big(B \cap \bigcup_{j=1}^{M} J_j\Big) \geq \lambda(B) - \varepsilon > s - 2\varepsilon.$$

Summing up over these intervals we get

$$\sum_{j=1}^{M} \big(f(y_j + k_j) - f(y_j)\big) > v \sum_{j=1}^{M} k_j = v\lambda\Big(\bigcup_{j=1}^{M} J_j\Big) > v(s - 2\varepsilon). \qquad (7.8)$$

For every $i \in \{1, \dots, N\}$, we sum up over all the intervals J_j such that $J_j \subset I_i$, and, using that f is nondecreasing, we obtain

$$\sum_{j,\, J_j \subset I_i} \big(f(y_j + k_j) - f(y_j)\big) \leq f(x_i) - f(x_i - h_i)$$

by which, summing over i and taking into account that every interval J_j is contained in some interval I_i,

$$\sum_{i=1}^{N} \big(f(x_i) - f(x_i - h_i)\big) \geq \sum_{i=1}^{N} \sum_{j,\, J_j \subset I_i} \big(f(y_j + k_j) - f(y_j)\big) = \sum_{j=1}^{M} \big(f(y_j + k_j) - f(y_j)\big).$$

Combining this with (7.7)-(7.8),

$$u(s + \varepsilon) > v(s - 2\varepsilon).$$

The arbitrariness of ε implies $us \geq vs$; since $u < v$, then $s = 0$. This shows that $\lambda(E_{u,v}) = 0$, as asserted.

We have thus proved that the function

$$\Phi(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

is defined almost everywhere on $[a, b]$ and f has a derivative at x if and only if $\Phi(x)$ is finite. Let

$$\Phi_n(x) = n\bigg(f\Big(x + \frac{1}{n}\Big) - f(x)\bigg)$$

where, to make Φ_n meaningful for all $x \in [a, b]$, we get $f(x) = f(b)$ for $x \geq b$, by definition. Since f is summable on $[a, b]$, so is every Φ_n. Integrating Φ_n, we get

$$\begin{aligned}\int_a^b \Phi_n(x)dx &= n\int_a^b \Big(f\Big(x+\frac{1}{n}\Big)-f(x)\Big)dx = n\Big(\int_{a+\frac{1}{n}}^{b+\frac{1}{n}} f(x)dx - \int_a^b f(x)dx\Big)\\ &= n\Big(\int_b^{b+\frac{1}{n}} f(x)dx - \int_a^{a+\frac{1}{n}} f(x)dx\Big) = f(b) - n\int_a^{a+\frac{1}{n}} f(x)dx\\ &\le f(b)-f(a)\end{aligned}$$

where in the last step we use the fact that f is nondecreasing. From Fatou's lemma it follows that

$$\int_a^b \Phi(x)dx \le f(b) - f(a).$$

In particular Φ is summable, and, consequently, a.e. finite. Then f has a derivative almost everywhere and $f'(x) = \Phi(x)$ a.e. in $[a,b]$. □

Example 7.7 It is easy to find monotonic functions f for which (7.5) becomes a strict inequality. For example, given points $a = x_0 < x_1 < \ldots < x_n = b$ in the interval $[a,b]$ and $h_1, h_2, \ldots, h_n$ corresponding numbers, consider the function

$$f(x) = \begin{cases} h_1 & \text{if } a \le x < x_1,\\ h_2 & \text{if } x_1 \le x < x_2,\\ \cdots & \\ h_n & \text{if } x_{n-1} \le x \le b.\end{cases}$$

A function of this particularly simple type is called a *step function*. If $h_1 \le h_2 \le \ldots \le h_n$, then f is obviously nondecreasing and

$$0 = \int_a^b f'(x)dx < f(b) - f(a) = h_n - h_1.$$

Example 7.8 [Vitali's function] In the preceding example, f is discontinuous. However, it is also possible to find continuous nondecreasing functions satisfying the strict inequality (7.5). To this end let

$$(a_1^1, b_1^1) = \Big(\frac{1}{3}, \frac{2}{3}\Big)$$

be the middle third of the interval $[0, 1]$, let

$$(a_1^2, b_1^2) = \left(\frac{1}{9}, \frac{2}{9}\right), \quad (a_2^2, b_2^2) = \left(\frac{7}{9}, \frac{8}{9}\right)$$

be the middle thirds of the intervals remaining after deleting (a_1^1, b_1^1) from [0,1], let

$$(a_1^3, b_1^3) = \left(\frac{1}{27}, \frac{2}{27}\right), \quad (a_2^3, b_2^3) = \left(\frac{7}{27}, \frac{8}{27}\right),$$
$$(a_3^3, b_3^3) = \left(\frac{19}{27}, \frac{20}{27}\right), \quad (a_4^3, b_4^3) = \left(\frac{25}{27}, \frac{26}{27}\right)$$

be the middle thirds of the intervals remaining after deleting (a_1^1, b_1^1), (a_1^2, b_1^2), (a_2^2, b_2^2) from $[0, 1]$ and so on. Note that the complement of the union of all the intervals (a_k^n, b_k^n) is the Cantor set constructed in Example 1.49. Now define a function

$$f(0) = 0, \quad f(1) = 1, \quad f(t) = \frac{2k-1}{2^n} \text{ if } t \in (a_k^n, b_k^n),$$

so that

$$f(t) = \frac{1}{2} \quad \text{if } \frac{1}{3} < t < \frac{2}{3},$$

$$f(t) = \begin{cases} \frac{1}{4} & \text{if } \frac{1}{9} < t < \frac{2}{9}, \\ \frac{3}{4} & \text{if } \frac{7}{9} < t < \frac{8}{9}, \end{cases}$$

$$f(t) = \begin{cases} \frac{1}{8} & \text{if } \frac{1}{27} < t < \frac{2}{27}, \\ \frac{3}{8} & \text{if } \frac{7}{27} < t < \frac{8}{27}, \\ \frac{5}{8} & \text{if } \frac{19}{27} < t < \frac{20}{27}, \\ \frac{7}{8} & \text{if } \frac{25}{27} < t < \frac{26}{27}, \end{cases}$$

and so on. Then f is defined everywhere except at points of the Cantor set C; furthermore f is nondecreasing on $[0, 1] \setminus C$ and $f([0, 1] \setminus C) = \{\frac{2k-1}{2^n} \mid n \in \mathbb{N},\ 1 \leq k \leq 2^{n-1}\}$ which is dense in $[0, 1]$, that is

$$\overline{f([0, 1] \setminus C)} = [0, 1]. \tag{7.9}$$

Given any point $t^* \in C$, let $(t_n)_n$ be an increasing sequence of points in $[0,1] \setminus C$ converging to t^* and let $(t'_n)_n$ be a decreasing sequence of points in $[0,1]\setminus C$ converging to t^*. Such sequences exist since $[0,1]\setminus C$ is dense in $[0,1]$. Then the limits $\lim_n f(t_n)$ and $\lim_n f(t'_n)$ exist (since f is nondecreasing in $[0,1] \setminus C$); we claim that they are equal. Otherwise, setting $a = \lim_n f(t_n)$ and $b = \lim_n f(t'_n)$, then $(a,b) \subset [0,1] \setminus f([0,1] \setminus C)$, in contradiction with (7.9). Then let

$$f(t^*) = \lim_n f(t_n) = \lim_n f(t'_n).$$

Completing the definition of f in this way, we obtain a continuous nondecreasing function on the whole interval $[0,1]$, known as Vitali's function. The derivatives f' obviously vanishes at every interval (a_k^n, b_k^n), and hence vanishes almost everywhere, since the Cantor set has measure zero. It follows that

$$0 = \int_0^1 f'(x)dx < f(1) - f(0) = 1.$$

7.2 Functions of bounded variation

Definition 7.9 *A function f defined on an interval $[a,b]$ if said to be* of bounded variation *if there is a constant $C > 0$ such that*

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \le C \tag{7.10}$$

for every partition

$$a = x_0 < x_1 < \ldots < x_n = b \tag{7.11}$$

of $[a,b]$. By the total variation of f on $[a,b]$, denoted by $V_a^b(f)$, is meant the quantity:

$$V_a^b(f) = \sup \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \tag{7.12}$$

where the least upper bound is taken over all partitions (7.11) *of the interval $[a,b]$.*

Remark 7.10 It is an immediate consequence of the above definition that if $\alpha \in \mathbb{R}$ and f is a function of bounded variation on $[a,b]$, then so is αf and

$$V_a^b(\alpha f) = |\alpha| V_a^b(f).$$

Example 7.11 1. If f is a monotonic function on $[a, b]$, then the left-hand side of (7.10) equals $|f(b) - f(a)|$ regardless of the choice of partition. Then f is of bounded variation and $V_a^b(f) = |f(b) - f(a)|$.

2. If f is a step function of the type considered in Example 7.7 with $h_1, \ldots, h_n \in \mathbb{R}$, then f is of bounded variation, with total variation given by the sum of the jumps, i.e.

$$V_a^b(f) = \sum_{i=1}^{n-1} |h_{i+1} - h_i|.$$

Example 7.12 Suppose f is a Lipschitz function on $[a, b]$ with Lipschitz constant K; then for any partition (7.11) of $[a, b]$ we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \leq K \sum_{k=0}^{n-1} |x_{k+1} - x_k| = K(b - a).$$

Then f is of bounded variation and $V_a^b(f) \leq K(b - a)$.

Example 7.13 It is easy to find a continuous function which is not of bounded variation. Indeed consider the function

$$f(x) = \begin{cases} x \sin \dfrac{1}{x} & \text{if } 0 < x \leq 1, \\ 0 & \text{if } x = 0 \end{cases}$$

and, fixed $n \in \mathbb{N}$, take the following partition

$$0, \frac{2}{(4n-1)\pi}, \frac{2}{(4n-3)\pi}, \ldots, \frac{2}{3\pi}, \frac{2}{\pi}, 1.$$

The sum on the left-hand side of (7.10) associated to such partition is given by

$$\frac{4}{\pi} \sum_{k=1}^{2n-1} \frac{1}{2k+1} + \frac{2}{\pi} + \left| \sin 1 - \frac{2}{\pi} \right|.$$

Taking into account that $\sum_{k=1}^{\infty} \frac{1}{2k+1} = \infty$, we deduce that the least upper bound on the right-hand side of (7.12) over all partitions of $[a, b]$ is infinity.

Proposition 7.14 *If f and g are functions of bounded variation on $[a, b]$, then so is $f + g$ and*

$$V_a^b(f+g) \le V_a^b(f) + V_a^b(g).$$

Proof. For any partition of the interval $[a, b]$, we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) + g(x_{k+1}) - f(x_k) - g(x_k)| \\ \le \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| + \sum_{k=0}^{n-1} |g(x_{k+1}) - g(x_k)| \le V_a^b(f) + V_a^b(g).$$

Taking the least upper bound on the left-hand side over all partitions of $[a, b]$ we immediately get the thesis. □

It follows from Remark 7.10 and Proposition 7.14 that any linear combination of functions of bounded variation is itself a function of bounded variation. In other words, the set $BV([a, b])$ of all functions of bounded variation on the interval $[a, b]$ is a linear space (unlike the set of all monotonic functions).

Proposition 7.15 *If f is a function of bounded variation on $[a, b]$ and $a < c < b$, then*

$$V_a^b(f) = V_a^c(f) + V_c^b(f).$$

Proof. First we consider a partition of the interval $[a, b]$ such that c is one of the points of subdivision, say $x_r = c$. Then

$$\begin{aligned} &\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| \\ &\qquad = \sum_{k=0}^{r-1} |f(x_{k+1}) - f(x_k)| + \sum_{k=r}^{n-1} |f(x_{k+1}) - f(x_k)| \\ &\qquad \le V_a^c(f) + V_c^b(f). \end{aligned} \tag{7.13}$$

Now consider an arbitrary partition of $[a, b]$. It is clear that adding an extra point of subdivision to this partition can never decrease the sum $\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$. Therefore (7.13) holds for any subdivision of $[a, b]$, and hence

$$V_a^b(f) \le V_a^c(f) + V_c^b(f).$$

On the other hand, given any $\varepsilon > 0$, there are partitions of the intervals $[a, c]$ and $[c, b]$, respectively, such that

$$\sum_i |f(x'_{i+1}) - f(x'_i)| > V_a^c(f) - \frac{\varepsilon}{2},$$

$$\sum_j |f(x''_{j+1}) - f(x''_j)| > V_c^b(f) - \frac{\varepsilon}{2}.$$

Combining all points of subdivision x'_i, x''_j, we get a partition of the interval $[a, b]$, with points of subdivision x_k, such that

$$\begin{aligned} V_a^b(f) \geq \sum_k |f(x_{k+1}) - f(x_k)| &= \sum_i |f(x'_{i+1}) - f(x'_i)| + \sum_j |f(x''_{j+1}) - f(x''_j)| \\ &> V_a^c(f) + V_c^b(f) - \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, it follows that $V_a^b(f) \geq V_a^c(f) + V_c^b(f)$. □

Corollary 7.16 *If f is a function of bounded variation on $[a, b]$, then the function*

$$x \longmapsto V_a^x(f)$$

is nondecreasing.

Proof. If $a \leq x < y \leq b$, Proposition 7.15 implies

$$V_a^y(f) = V_a^x(f) + V_x^y(f) \geq V_a^x(f).$$

□

Proposition 7.17 *A function $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if f can be represented as the difference between two nondecreasing functions on $[a, b]$.*

Proof. Since, by Example 7.11, any monotonic function is of bounded variation and since the set $BV([a, b])$ is a linear space, we get that the difference of two nondecreasing functions is of bounded variation. To prove the converse, set

$$g_1(x) = V_a^x(f), \quad g_2(x) = V_a^x(f) - f(x).$$

By Corollary 7.16 g_1 is a nondecreasing function. We claim that g_2 is nondecreasing too. Indeed, if $x < y$, then, using Proposition 7.15, we get

$$g_2(y) - g_2(x) = V_x^y(f) - (f(y) - f(x)). \tag{7.14}$$

But from Definition 7.9

$$|f(y) - f(x)| \leq V_x^y(f)$$

and hence the right hand side of (7.14) is nonnegative. Writing $f = g_1 - g_2$, we get the desired representation of f as the difference between two nondecreasing functions. □

Theorem 7.18 *Let $f : [a, b] \to \mathbb{R}$ be a function of bounded variation. Then the set of points of $[a, b]$ at which f is discontinuous is at most countable. Furthermore f has a derivative almost everywhere on $[a, b]$, $f' \in L^1([a, b])$ and*

$$\int_a^b |f'(x)|dx \leq V_a^b(f). \tag{7.15}$$

Proof. Combining Theorem 7.5, Theorem 7.6 and Proposition 7.17 we immediately obtain that f has no more than countably many points of discontinuity, has a derivative almost everywhere on $[a, b]$ and $f' \in L^1([a, b])$. Since for $a \leq x < y \leq b$

$$|f(y) - f(x)| \leq V_x^y(f) = V_a^y(f) - V_a^x(f),$$

we get

$$|f'(x)| \leq (V_a^x(f))' \quad \text{a.e. in } [a, b].$$

Finally, using (7.5)

$$\int_a^b |f'(x)|dx \leq \int_a^b (V_a^x(f))'dx \leq V_a^b(f).$$

□

Remark 7.19 Any step function and the Vitali's function (see Example 7.8) provide examples of functions of bounded variation satisfying the strict inequality (7.15).

Proposition 7.20 *A function $f : [a, b] \to \mathbb{R}$ is of bounded variation if and only if the curve*

$$y = f(x) \quad a \le x \le b$$

is rectifiable, i.e. has finite lenght[1]*.*

Proof. For any partition of $[a, b]$ we get

$$\begin{aligned}\sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)| &\le \sum_{i=0}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2} \\ &\le (b-a) + \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|.\end{aligned}$$

Taking the least upper bound over all partitions we get the thesis. □

Exercise 7.21 Let $(a_n)_n$ be a sequence of positive numbers and let

$$f(x) = \begin{cases} a_n & x = \dfrac{1}{n},\ n \ge 1; \\ 0 & \text{otherwise.} \end{cases}$$

Prove that f is of bounded variation on $[0, 1]$ iff $\sum_{n=1}^{\infty} a_n < \infty$.

Exercise 7.22 Let f be a function of bounded variation on $[a, b]$ such that

$$f(x) \ge c > 0 \quad \forall x \in [a, b].$$

Prove that $\frac{1}{f}$ is of bounded variation and

$$V_a^b\left(\frac{1}{f}\right) \le \frac{1}{c^2} V_a^b(f).$$

Exercise 7.23 Prove that the function

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x^3} & 0 < x \le 1, \\ 0 & x = 0 \end{cases}$$

is not of bounded variation on $[0, 1]$.

---

[1] By the length of the curve $y = f(x)$ $(a \le x \le b)$ is meant the quantity

$$\sup \sum_{i=0}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + (f(x_{i+1}) - f(x_i))^2}$$

where the least upper bound is taken over all possible partitions of $[a, b]$.

7.3 Absolutely continuous functions

We now address ourselves to the problems posed at the beginning of the chapter. The object of this section is to describe the class of functions satisfying (7.2).

Definition 7.24 *A function f defined on an interval $[a,b]$ is said to be* absolutely continuous *if, given $\varepsilon > 0$, there is a $\delta > 0$ such that*

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon \tag{7.16}$$

for every finite system of pairwise disjoint subintervals

$$(a_k, b_k) \subset [a,b] \quad k = 1, \dots, n$$

of total length $\sum_{k=1}^{n}(b_k - a_k)$ less than δ.

Example 7.25 Suppose f is a Lipschitz function on $[a,b]$ with Lipschitz constant K; then, choosing $\delta = \frac{\varepsilon}{K}$, we immediately get that f is absolutely continuous.

Remark 7.26 Clearly every absolutely continuous function is uniformly continuous, as we see by choosing a single subinterval $(a_1, b_1) \subset [a,b]$. However, a uniformly continuous function need not be absolutely continuous. For example, the Vitali's function f constructed in Example 7.8 is continuous (and hence uniformly continuous) on $[0,1]$, but not absolutely continuous on $[0,1]$. In fact, for every n consider the set

$$C_n = \left\{ x \in [0,1] \,\middle|\, x = \sum_{i=1}^{\infty} \frac{a_i}{3^i} \text{ with } a_1, \dots, a_n \neq 1 \right\}$$

which is the union of 2^n pairwise disjoint closed intervals I_i, each of which has measure $\frac{1}{3^n}$ (then the total length is $(\frac{2}{3})^n$). Denoting by C the Cantor set (see Example 1.49), we have $C \subset C_n$; since, by construction, the Vitali's function is constant on the subintervals of $[0,1] \setminus C$, then the sum (7.16) associated to the system (I_i) is equal to 1. Hence the Cantor set C can be covered by a finite system of subintervals of arbitrarily small length, but the sum (7.16) associated to every such system is equal to 1. The same example shows that a function of bounded variation needs not be absolutely continuous. On the other hand, an absolutely continuous function is necessarily of bounded variation (see Proposition 7.27).

Proposition 7.27 *If f is absolutely continuous on $[a,b]$, then f is of bounded variation on $[a,b]$.*

Proof. Given any $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

for every finite system of pairwise disjoint subintervals $(a_k, b_k) \subset [a,b]$ such that

$$\sum_{k=1}^{n} (b_k - a_k) < \delta.$$

Hence if $[\alpha, \beta]$ is any subinterval of length less than δ, we have

$$V_\alpha^\beta(f) \leq \varepsilon.$$

Let $a = x_0 < x_1 < \ldots < x_N = b$ be a partition of $[a,b]$ into N subintervals $[x_k, x_{k+1}]$ all of length less than δ. Then, by Proposition 7.15,

$$V_a^b(f) \leq N\varepsilon.$$

□

An immediate consequence of Definition 7.24 and obvious properties of absolute value is the following.

Proposition 7.28 *If f is absolutely continuous on $[a,b]$, then so is αf, where α is any constant. Moreover, if f and g are absolutely continuous on $[a,b]$, then so is $f+g$.*

It follows from Proposition 7.28 together with Remark 7.26 that the set $AC([a,b])$ of all absolutely continuous functions on $[a,b]$ is a proper subspace of the linear space $BV([a,b])$ of all functions of bounded variation on $[a,b]$.

We now study the close connection between absolute continuity and the indefinite Lebesgue integral. To this aim we need the following result.

Lemma 7.29 *Let $g \in L^1([a,b])$ be such that $\int_I g(t)dt = 0$ for every subinterval $I \subset [a,b]$. Then $g(x) = 0$ a.e. in $[a,b]$.*

Proof. If we denote by $\mathcal{I}$ the family of all finite disjoint union of subintervals of $[a, b]$, it is immediate to see that $\mathcal{I}$ is an algebra and $\int_A g(t)dt = 0$ for every $A \in \mathcal{I}$. Let V be an open set in $[a, b]$; then $V = \cup_{n=1}^{\infty} I_n$ where $I_n \subset [a, b]$ is a subinterval. For every n, since $\cup_{i=1}^{n} I_i \in \mathcal{I}$, we have $\int_{\cup_{i=1}^{n} I_i} g(t)dt = 0$; Lebesgue Theorem implies

$$\int_V g(t)dt = \lim_{n\to\infty} \int_{\cup_{i=1}^{n} I_i} g(t)dt = 0$$

Assume by contradiction the existence of $E \in \mathcal{B}([a, b])$ such that $\lambda(E) > 0$ and $g(x) > 0$ in E. By Theorem 1.55 there exists a compact set $K \subset E$ such that $\lambda(K) > 0$. Setting $V = [a, b] \setminus K$, V is an open set in $[a, b]$; then

$$0 = \int_a^b g(t)dt = \int_V g(t)dt + \int_K g(t)dt = \int_K g(t)dt > 0,$$

and the contradiction follows. □

Returning to the problem of differentiating the indefinite Lebesgue integral, in the following Theorem we evaluate the derivative (7.1), thereby giving an affirmative answer to the first of the two questions posed at the beginning of the chapter.

Theorem 7.30 *Let $f \in L^1([a, b])$ and set*

$$F(x) = \int_a^x f(t)dt, \quad x \in [a, b].$$

Then F is absolutely continuous on $[a, b]$ and

$$F'(x) = f(x) \textit{ for a.e. } x \in [a, b]. \tag{7.17}$$

Proof. Given any finite collection of pairwise disjoint intervals (a_k, b_k), we have

$$\sum_{k=1}^{n} |F(b_k) - F(a_k)| = \sum_{k=1}^{n} \Big| \int_{a_k}^{b_k} f(t)dt \Big| \le \sum_{k=1}^{n} \int_{a_k}^{b_k} |f(t)|dt = \int_{\bigcup_k (a_k, b_k)} |f(t)|dt.$$

By the absolute continuity of the integral, the last expression on the right approaches zero as the total length of the intervals (a_k, b_k) approaches zero. This proves that F is absolutely continuous on $[a, b]$. By Proposition 7.27 F

is of bounded variation; consequently, by Theorem 7.18, F has a derivative almost everywhere on $[a, b]$ and $F' \in L^1([a, b])$. It remains to prove (7.17). First assume that there exists $K > 0$ such that $|f(x)| \le K$ for every $x \in [a, b]$ and let

$$g_n(x) = n\Big[F\Big(x + \frac{1}{n}\Big) - F(x)\Big]$$

where, to make g_n meaningful for all $x \in [a, b]$, we get $F(x) = F(b)$ for $b < x \le b + 1$, by definition. Clearly

$$\lim_{n\to\infty} g_n(x) = F'(x)$$

almost everywhere on $[a, b]$. Furthermore

$$|g_n(x)| = \Big|n \int_x^{x+\frac{1}{n}} f(t)dt\Big| \le K \quad \forall x \in [a, b].$$

Consider $a \le c < d \le b$ and, by using Lebesgue Theorem, we get

$$\int_c^d F'(x)dx = \lim_{n\to\infty} \int_c^d g_n(x)dx = \lim_{n\to\infty} n\Big[\int_{c+\frac{1}{n}}^{d+\frac{1}{n}} F(x)dx - \int_c^d F(x)dx\Big]$$
$$= \lim_{n\to\infty} \Big[\int_d^{d+\frac{1}{n}} F(x)dx - \int_c^{c+\frac{1}{n}} F(x)dx\Big] = F(d) - F(c)$$

where last equality follows from the mean value theorem. Hence we deduce

$$\int_c^d F'(x)dx = F(d) - F(c) = \int_c^d f(t)dt$$

by which, using Lemma 7.29, we conclude $F'(x) = f(x)$ a.e. in $[a, b]$.

Next we want to remove the hypothesis on the boundedness of f. Without loss of generality we may assume $f \ge 0$ (otherwise, we can consider separately f^+ and f^-). Then F is a nondecreasing function on $[a, b]$. Define f_n as follows:

$$f_n(x) = \begin{cases} f(x) & \text{if } 0 \le f(x) \le n, \\ n & \text{if } f(x) \ge n. \end{cases}$$

Since $f - f_n \ge 0$, the function $H_n(x) := \int_a^x (f(t) - f_n(t))dt$ in nondecreasing; hence, by Theorem 7.6, H_n has nonnegative derivative almost everywhere.

Since $0 \leq f_n \leq n$, by the first part of the proof we have $\frac{d}{dx}\int_a^x f_n(t)dt = f_n(x)$ a.e. in $[a,b]$; therefore for every $n \in \mathbb{N}$

$$F'(x) = H_n'(x) + \frac{d}{dx}\int_a^x f_n(t)dt \geq f_n(x) \quad \text{a.e. in } [a,b]$$

by which $F'(x) \geq f(x)$ for a.e. $x \in [a,b]$ and so, after integration,

$$\int_a^b F'(x)dx \geq \int_a^b f(x)dx = F(b) - F(a).$$

On the other hand, since F is nondecreasing on $[a,b]$, (7.5) gives $\int_a^b F'(x)dx \leq F(b) - F(a)$, and then

$$\int_a^b F'(x)dx = F(b) - F(a) = \int_a^b f(x)dx.$$

We obtain $\int_a^b (F'(x) - f(x))dx = 0$; since $F'(x) \geq f(x)$ a.e., we conclude $F'(x) = f(x)$ a.e. in $[a,b]$. □

We are going to give a definite answer to the second of the question posed at the beginning of the chapter.

Lemma 7.31 *Let f be an absolutely continuous function on $[a,b]$ such that $f'(x) = 0$ a.e. in $[a,b]$. Then f is constant on $[a,b]$.*

Proof. Fixed $c \in (a,b)$, we want to show that $f(c) = f(a)$. Let $E \subset (a,c)$ be such that $f'(x) = 0$ for every $x \in E$. Then $E \in \mathcal{B}([a,b])$ and $\lambda(E) = c - a$. Given $\varepsilon > 0$, there is a $\delta > 0$ such that

$$\sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon$$

for every finite system of pairwise disjoint subintervals $(a_k, b_k) \subset [a,b]$ such that

$$\sum_{k=1}^{n} (b_k - a_k) < \delta.$$

Fix $\eta > 0$. For every $x \in E$ and $\gamma > 0$, since $\lim_{y \to x} \frac{f(y)-f(x)}{y-x} = 0$, there exists $y_{x,\gamma} > x$ such that $[x, y_{x,\gamma}] \subset (a,c)$, $|y_{x,\gamma} - x| \leq \gamma$ and

$$|f(y_{x,\eta}) - f(x)| \leq (y_{x,\gamma} - x)\eta. \tag{7.18}$$

The intervals $([x, y_{x,\gamma}])_{x\in(a,c),\gamma>0}$ provide a fine cover of E; hence, by Vitali's covering Theorem, there exists a finite number of such disjoint subintervals of (a, c)

$$I_1 = [x_1, y_1], \ldots, I_n = [x_n, y_n]$$

with $x_k < x_{k+1}$, such that $\lambda(E \setminus \cup_{i=1}^n I_k) < \delta$. Then we have

$$y_0 := a < x_1 < y_1 < x_2 < \ldots < y_n < c := x_{n+1}, \quad \sum_{k=0}^n (x_{k+1} - y_k) < \delta.$$

From the absolute continuity of f we obtain

$$\sum_{k=0}^n |f(x_{k+1}) - f(y_k)| < \varepsilon, \tag{7.19}$$

while, by (7.18),

$$\sum_{k=1}^n |f(y_k) - f(x_k)| \le \eta \sum_{k=1}^n (y_k - x_k) \le \eta(b - a). \tag{7.20}$$

Combining (7.19)-(7.20) we deduce

$$|f(c) - f(a)| = \Big| \sum_{k=0}^n (f(x_{k+1}) - f(y_k)) + \sum_{k=1}^n (f(y_k) - f(x_k)) \Big| \le \varepsilon + \eta(b - a).$$

The arbitrariness of ε and η allows us to conclude. □

Theorem 7.32 *If f is absolutely continuous on $[a, b]$, then f has a derivative almost everywhere on $[a, b]$, $f' \in L^1([a, b])$ and*

$$f(x) = f(a) + \int_a^x f'(t)dt \quad \forall x \in [a, b]. \tag{7.21}$$

Proof. By Proposition 7.27 f is of bounded variation; hence, by Theorem 7.18, f has a derivative almost everywhere and $f' \in L^1([a, b])$. To prove (7.21) consider the function

$$g(x) = \int_a^x f'(t)dt.$$

Then, by Theorem 7.30, g is absolutely continuous on $[a, b]$ and $g'(x) = f'(x)$ a.e. in $[a, b]$. Setting $\Phi = g - f$, Φ is absolutely continuous, being the difference of two absolutely continuous functions, and $\Phi'(x) = 0$ a.e. in $[a, b]$. It follows from the previous lemma that Φ is constant, that is $\Phi(x) = \Phi(a) = f(a) - g(a) = f(a)$, by which

$$f(x) = \Phi(x) + g(x) = f(a) + \int_a^x f'(t)dt \quad \forall x \in [a, b].$$

□

Remark 7.33 Combining Theorem 7.30 and 7.32 we can now give a definitive answer to the second question posed at the beginning of the chapter: the formula

$$\int_a^x F'(t)dt = F(x) - F(a)$$

holds for all $x \in [a, b]$ if and only if F is absolutely continuous on $[a, b]$.

Proposition 7.34 *Let $f : [a, b] \to \mathbb{R}$. The following properties are equivalent:*

a) f is absolutely continuous on $[a, b]$;

b) f is of bounded variation on $[a, b]$ and

$$\int_a^b |f'(t)|dt = V_a^b(f).$$

Proof. $\boxed{a) \Rightarrow b)}$ For any partition $a = x_0 < x_1 < \ldots < x_n = b$ of $[a, b]$, by Theorem 7.32 we have

$$\sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)| = \sum_{k=0}^{n-1} \left| \int_{x_k}^{x_{k+1}} f'(t)dt \right| \le \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} |f'(t)|dt = \int_a^b |f'(t)|dt,$$

which implies

$$V_a^b(f) \le \int_a^b |f'(t)|dt.$$

On the other hand, by Theorem 7.18, $\int_a^b |f'(t)|dt \le V_a^b(f)$, and so $V_a^b(f) = \int_a^b |f'(t)|dt$.

b) $\Rightarrow$ a) For every $x \in [a,b]$, using (7.15), we have

$$V_a^x(f) \geq \int_a^x |f'(t)|dt = \int_a^b |f'(t)|dt - \int_x^b |f'(t)|dt = V_a^b(f) - \int_x^b |f'(t)|dt$$
$$\geq V_a^b(f) - V_x^b(f) = V_a^x(f)$$

where last equality follows from Proposition 7.15. Then we get

$$V_a^x(f) = \int_a^x |f'(t)|dt.$$

Since $f' \in L^1([a,b])$, Theorem 7.30 implies that the function $x \mapsto V_a^x(f)$ is absolutely continuous. Given any collection of pairwise disjoint intervals (a_k, b_k), we have

$$\sum_{k=1}^n |f(b_k) - f(a_k)| \leq \sum_{k=1}^n V_{a_k}^{b_k}(f) = \sum_{k=1}^n \big(V_a^{b_k}(f) - V_a^{a_k}(f)\big).$$

By the absolute continuity of $x \mapsto V_a^x(f)$, the last expression on the right approaches zero as the total length of the intervals (a_k, b_k) approaches zero. This proves that f is absolutely continuous on $[a,b]$. □

By applying the above proposition to the particular case of monotonic functions, we obtain the following result.

Corollary 7.35 *Let $f : [a,b] \to \mathbb{R}$ be a monotonic function. The following properties are equivalent:*

a) f is absolutely continuous on $[a,b]$;

b) $\int_a^b |f'(t)|dt = |f(b) - f(a)|$.

Remark 7.36 Let f, g absolutely continuous functions on $[a,b]$. Then the following formula of integration by parts holds:

$$\int_a^b f(x)g'(x)dx = f(b)g(b) - f(a)b(a) - \int_a^b f'(x)g(x)dx.$$

Indeed, by Tonelli's Theorem $\iint_{[a,b]^2} |f'(x)g'(y)|dxdy = \int_a^b |f'(x)|dx \int_a^b |g'(y)|dy < \infty$, that is $f'(x)g'(y) \in L^1([a,b]^2)$. Then consider the set

$$A = \{(x,y) \in [a,b]^2 \,|\, a \leq x \leq y \leq b\}$$

and let us evaluate the integral

$$I = \iint_A f'(x)g'(y)dxdy$$

in two ways using Fubini's theorem and formula (7.21). On the one hand

$$\begin{aligned} I &= \int_a^b g'(y)\Big(\int_a^y f'(x)dx\Big)dy = \int_a^b g'(y)f(y)dy - f(a)\int_a^b g'(y)dy \\ &= \int_a^b g'(y)f(y)dy - f(a)\big(g(b) - g(a)\big) \end{aligned}$$

and, on the other hand

$$\begin{aligned} I &= \int_a^b f'(x)\Big(\int_x^b g'(y)dy\Big)dy = g(b)\int_a^b f'(y)dy - \int_a^b f'(x)g(x)dx \\ &= g(b)\big(f(b) - f(a)\big) - \int_a^b f'(x)g(x)dy. \end{aligned}$$

Exercise 7.37 Prove that if f and g are absolutely continuous functions on $[a, b]$, then so is fg.

Exercise 7.38 Let $(f_n)_n$ be a sequence of absolutely continuous functions on $[0, 1]$, which converges pointwise to a function f on $[0, 1]$, such that

$$\int_0^1 |f_n'(x)|dx \leq M, \quad \forall n \in \mathbb{N},$$

where $M > 0$ is a constant.

- Show that $\lim_{n\to\infty} \int_0^1 f_n(x)dx = \int_0^1 f(x)dx$;
- Prove that f is of bounded variation on $[0, 1]$;
- Give an example to show that, in general, f is not absolutely continuous on $[0, 1]$.

Appendix A

A.1 Distance function

In this section we recall the basic properties of the distance function from a nonempty set $S \subset \mathbb{R}^N$.

Definition A.1 *The* distance function *from S is the function $d_S : \mathbb{R}^N \to \mathbb{R}$ defined by*

$$d_S(x) = \inf_{y \in S} \|x - y\| \qquad \forall x \in \mathbb{R}^N$$

The projection *of x onto S consists of those points (if any) at which the infimum defining $d_S(x)$ is attained. Such a set will be denoted by $\mathrm{proj}_S(x)$.*

Proposition A.2 *Let S be a nonempty subset of $\mathbb{R}^N$. Then the following properties hold true.*

1. *d_S is Lipschitz continuous of rank 1*[1]*.*
2. *For any $x \in \mathbb{R}^N$ we have that $d_S(x) = 0$ iff $x \in \overline{S}$.*
3. *$\mathrm{proj}_S(x) \neq \varnothing$ for every $x \in \mathbb{R}^N$ iff S is closed.*

Proof. We shall prove the three properties in sequence.

---

[1] A function $f : \Omega \subset \mathbb{R}^N \to \mathbb{R}$ is said to be Lipschitz of rank $L \geq 0$ in Ω iff

$$|f(x) - f(y)| \leq L\|x - y\| \qquad \forall x, y \in \Omega$$

1. Let $x, x' \in \mathbb{R}^N$ and $\varepsilon > 0$ be fixed. Then there exists $y_\varepsilon \in S$ such that $\|x - y_\varepsilon\| < d_S(x) + \varepsilon$. Thus, by the triangle inequality for Euclidean norm,
$$d_S(x') - d_S(x) \leq \|x' - y_\varepsilon\| - \|x - y_\varepsilon\| + \varepsilon \leq \|x' - x\| + \varepsilon$$
Since ε is arbitrary, $d_S(x') - d_S(x) \leq \|x' - x\|$. Exchanging the role of x and x' we conclude that $|d_S(x') - d_S(x)| \leq \|x' - x\|$ as desired.

2. For any $x \in \mathbb{R}^N$ we have that $d_S(x) = 0$ iff a sequence $(y_n) \subset S$ exists such that $\|x - y_n\| \to 0$ as $n \to \infty$, hence iff $x \in \overline{S}$.

3. Let S be closed and $x \in \mathbb{R}^N$ be fixed. Then
$$K := \Big\{ y \in S \mid \|x - y\| \leq d_S(x) + 1 \Big\}$$
is a nonempty compact set. Therefore, any point $\widehat{x} \in K$ such that
$$\|x - \widehat{x}\| = \min_{y \in K} \|x - y\|$$
lies in $\mathrm{proj}_S(x)$. Conversely, let $x \in \overline{S}$. Observe that, by point 2, $d_S(x) = 0$. Take $\widehat{x} \in \mathrm{proj}_S(x)$. Then $\|x - \widehat{x}\| = 0$. So $x = \widehat{x} \in S$. □

A.2 Legendre transform

Let $f : \mathbb{R}^N \to \mathbb{R}$ be a convex function. The function $f^* : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ defined by
$$f^*(y) = \sup_{x \in \mathbb{R}} \{x \cdot y - f(x)\} \qquad \forall y \in \mathbb{R}^N \tag{A.1}$$
is called the *Legendre transform* (and, sometimes, the *Fenchel transform* or *convex conjugate*) of f. From of the definition of f^* it follows that
$$x \cdot y \leq f(x) + f^*(y) \qquad \forall x, y \in \mathbb{R}^N \,. \tag{A.2}$$
Some properties of the Legendre transform of a superlinear function are described below.

Proposition A.3 *Let $f \in C^1(\mathbb{R}^N)$ be a convex function satisfying*
$$\lim_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \infty \,. \tag{A.3}$$
Then, the following properties hold:

(a) $\forall y \in \mathbb{R}^N \quad \exists x_y \in \mathbb{R}^N$ *such that* $f^*(y) = x_y \cdot y - f(x_y)$*;*

(b) $y = Df(x)$ *if and only if* $f^*(y) + f(x) = x \cdot y$*;*

(c) f^* *is convex;*

(d) f^* *is superlinear;*

(e) $f^{**} = f$.

Proof.

(a): the conclusion is a straightforward consequence of the continuity and superlinearity of f.

(b): let $x, y \in \mathbb{R}^N$ satisfy $f^*(y) + f(x) = x \cdot y$. Then, $F(x) := x \cdot y - f(x)$ attains its maximum at x, whence $y = Df(x)$. Conversely, being $F(x)$ concave, the supremum in (A.1) is attained at every point at which $0 = DF(x) = y - Df(x)$.

(c): take any $y_1, y_2 \in \mathbb{R}^N$ and $t \in [0, 1]$, and let x_t be a point such that

$$f^*(ty_1 + (1-t)y_2) = [ty_1 + (1-t)y_2]x_t - f(x_t)\,.$$

Since $f^*(y_i) \geq y_i \cdot x_t - f(x_t)$ for $i = 1, 2$, we conclude that

$$f^*(ty_1 + (1-t)y_2) \leq tf^*(y_1) + (1-t)f^*(y_2)\,,$$

i.e., f^* is convex.

(d): for all $M > 0$ and $y \in \mathbb{R}^N$, we have

$$f^*(y) \geq M\frac{y}{\|y\|} \cdot y - f\left(M\frac{y}{\|y\|}\right) \geq M\|y\| - \max_{\|x\|=M} f(x)\,.$$

So, for all $M > 0$,

$$\liminf_{\|y\|\to\infty} \frac{f^*(y)}{\|y\|} \geq M\,.$$

Since M is arbitrary, f^* must be superlinear.

(e): by definition, $f(x) \geq x \cdot y - f^*(y)$ for all $x, y \in \mathbb{R}^N$. So, $f \geq f^{**}$. To prove the converse inequality, fix $x \in \mathbb{R}^N$ and let $y_x = Df(x)$. Then, owing to point (b) above,

$$f(x) = x \cdot y_x - f^*(y_x) \leq f^{**}(x)\,. \qquad \square$$

Example A.4 (Young's inequality) Define, for $p > 1$,

$$f(x) = \frac{|x|^p}{p} \qquad \forall x \in \mathbb{R}\,.$$

Then, f is a superlinear function of class $C^1(\mathbb{R})$. Moreover,

$$f'(x) = |x|^{p-1}\text{sign}(x)$$

where

$$\text{sign}(x) = \begin{cases} \frac{x}{|x|} & \text{if} \quad x \neq 0 \\ 0 & \text{if} \quad x = 0 \end{cases}$$

So, f' is an increasing function, and f is convex.

In view of point (b) of Proposition A.3, we can compute $f^*(y)$ by solving $y = |x|^{p-1}\text{sign}(x)$. We find $x_y = |y|^{\frac{1}{p-1}}\text{sign}(y)$, whence

$$f^*(y) = x_y y - f(x_y) = \frac{|y|^q}{q} \qquad \forall y \in \mathbb{R}\,,$$

where $q = \frac{p}{p-1}$. Thus, on account of (A.2), we obtain the following estimate:

$$|xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q} \qquad \forall x, y \in \mathbb{R}\,, \tag{A.4}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Moreover, owing to point (b) above, we conclude that equality holds in (A.4) iff $|y|^q = |x|^p$.

Exercise A.5 Let $f(x) = e^x,\ x \in \mathbb{R}$. Show that

$$f^*(y) = \sup_{x \in \mathbb{R}}\{xy - e^x\} = \begin{cases} \infty & \text{if} \quad y < 0 \\ 0 & \text{if} \quad y = 0 \\ y \log y - y & \text{if} \quad y > 0\,. \end{cases}$$

Deduce the following estimate

$$xy \leq e^x + y \log y - y \qquad \forall x, y > 0\,. \tag{A.5}$$

A.3 Baire's Lemma

Let (X, d) be a nonempty metric space. The following result is often referred to as Baire's Lemma. It is a classical result in topology.

Proposition A.6 (Baire) *Let (X, d) be a complete metric space. Then the following properties hold.*

(a) Any countable intersection of dense open sets $G_n \subset X$ is dense.

(b) If X is the countable union of nonempty closed sets F_k, then at least one F_k has nonempty interior.

Proof. We shall use the closed balls

$$\overline{B}_r(x) := \{y \in X \mid d(x, y) \le r\} \qquad r > 0\,,\ x \in X\,.$$

(a) Let us fix any ball $\overline{B}_{r_0}(x_0)$. We shall prove that $(\cap_n G_n) \cap \overline{B}_{r_0}(x_0) \neq \varnothing$. Since G_1 is dense, there exists a point $x_1 \in G_1 \cap B_{r_0}(x_0)$. Since G_1 is open, there also exists $0 < r_1 < 1$ such that

$$\overline{B}_{r_1}(x_1) \subset G_1 \cap B_{r_0}(x_0)\,.$$

Since G_2 is dense, we can find a point $x_2 \in G_2 \cap B_{r_1}(x_1)$ and—since G_2 is open—a radius $0 < r_2 < 1/2$ such that

$$\overline{B}_{r_2}(x_2) \subset G_2 \cap B_{r_1}(x_1)\,.$$

Iterating the above procedure, we can construct a decreasing sequence of closed balls $\overline{B}_{r_k}(x_k)$ such that

$$\overline{B}_{r_k}(x_k) \subset G_k \cap B_{r_{k-1}}(x_{k-1}) \quad \text{and} \quad 0 < r_k < 1/k\,.$$

We note that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. Indeed, for any $h, k \ge n$ we have that $x_k, x_k \in B_{r_n}(x_n)$. So, $d(x_k, x_h) < 2/n$. Therefore, X being complete, $(x_n)_{n\in\mathbb{N}}$ converges to a point $x \in X$ which must belong to $\cap_n G_n$.

(b) Suppose, by contradiction, that all F_k's have empty interior. Applying point (a) to $G_k := X \setminus F_k$, we can find a point $x \in \cap_n G_n$. Then, $x \in X \setminus \cup_k F_k$ in contrast with the fact that the F_k's do cover X. □

A.4 Precompact families of continuous functions

Let K be a compact topological space. We denote by $\mathcal{C}(K)$ the Banach space of all continuous functions $f : K \to \mathbb{R}$ endowed with the uniform norm

$$\|f\|_\infty = \max_{x\in K} |f(x)| \qquad \forall f \in \mathcal{C}(K)\,.$$

We recall that convergence in $\mathcal{C}(K)$ is equivalent to uniform convergence.

Definition A.7 *A family $\mathcal{M} \subset \mathcal{C}(K)$ is said to be:*

(i) equicontinuous *if, for any $\varepsilon > 0$ and any $x \in K$ there exists a neighbourhood V of x in K such that*

$$|f(x) - f(y)| < \varepsilon \qquad \forall y \in V\,,\ \forall f \in \mathcal{M};$$

(ii) pointwise bounded *if, for any $x \in X$, $\{f(x) \mid f \in \mathcal{M}\}$ is a bounded subset of $\mathbb{R}$.*

Theorem A.8 (Ascoli-Arzelà) *A family $\mathcal{M} \subset \mathcal{C}(K)$ is relatively compact iff $\mathcal{M}$ is equicontinuous and pointwise bounded.*

Proof. Let $\mathcal{M}$ be relatively compact. Then, $\mathcal{M}$ is bounded, hence pointwise bounded, in $\mathcal{C}(K)$. So, it suffices to show that $\mathcal{M}$ is equicontinuous. For any $\varepsilon > 0$ there exist $f_1, \ldots, f_m \in \mathcal{M}$ such that $\mathcal{M} \subset B_\varepsilon(f_1) \cup \cdots \cup B_\varepsilon(f_m)$. Let $x \in K$. Since each function f_i is continuous in x, x possesses neighbourhoods $V_1, \ldots, V_n \subset K$ such that

$$|f_i(x) - f_i(y)| < \varepsilon \qquad \forall y \in V_i\,, \quad i = 1, \ldots, m\,.$$

Set $V := V_1 \cap \cdots \cap V_m$ and fix $f \in \mathcal{M}$. Let $i \in \{1, \ldots, m\}$ be such that $f \in B_\varepsilon(f_i)$. Thus, for any $y \in V$,

$$|f(y) - f(x)| \le |f(y) - f_i(y)| + |f_i(y) - f_i(x)| + |f_i(x) - f(x)| < 3\varepsilon\,.$$

This shows that $\mathcal{M}$ is equicontinuous.

Conversely, given a pointwise bounded equicontinuous family $\mathcal{M}$, since K is compact for any $\varepsilon > 0$ there exist points $x_1, \ldots, x_m \in K$ and corresponding neighbourhoods $V_1, \ldots, V_m$ such that $K = V_1 \cup \cdots \cup V_m$ and

$$|f(x) - f(x_i)| < \varepsilon \qquad \forall f \in \mathcal{M}\,, \quad \forall x \in V_i\,, \quad i = 1, \ldots, m\,. \tag{A.6}$$

Since $\{(f(x_1), \ldots, f(x_m)) \mid f \in \mathcal{M}\}$ is relatively compact in $\mathbb{R}^m$, there exist functions $f_1, \ldots, f_n \in \mathcal{M}$ such that

$$\{(f(x_1), \ldots, f(x_m)) \mid f \in \mathcal{M}\} \subset \bigcup_{j=1}^{n} B_\varepsilon(f_j(x_1), \ldots, f_j(x_m)) . \tag{A.7}$$

We claim that

$$\mathcal{M} \subset B_{3\varepsilon}(f_1) \cup \cdots \cup B_{3\varepsilon}(f_n) , \tag{A.8}$$

which implies that $\mathcal{M}$ is totally bounded[(2)] , hence relatively compact. To obtain (A.8), let $f \in \mathcal{M}$ and let $j \in \{1, \ldots, n\}$ be such that

$$(f(x_1), \ldots, f(x_m)) \in B_\varepsilon(f_j(x_1), \ldots, f_j(x_m)) .$$

Now, fix $x \in K$ and let $i \in \{1, \ldots, m\}$ be such that $x \in V_i$. Then, in view of (A.6) and (A.7),

$$|f(x) - f_j(x)| \leq |f(x) - f(x_i)| + |f(x_i) - f_j(x_i)| + |f_j(x_i) - f_j(x)| < 3\varepsilon .$$

This proves (A.8) and completes the proof. □

Remark A.9 The compactness of K is essential for the above result. Indeed, the sequence

$$f_n(x) := e^{-(x-n)^2} \qquad \forall x \in \mathbb{R}$$

is a bounded equicontinuous family in $\mathcal{C}(\mathbb{R})$. On the other hand,

$$n \neq m \quad \Longrightarrow \quad \|f_n - f_m\|_\infty \geq 1 - \frac{1}{e} .$$

So, $(f_n)_n$ fails to be relatively compact.

A.5 Vitali's covering theorem

We present in this section the fundamental covering theorem of Vitali.

[(2)] given a metric space X and a subset $M \subset X$, we say that M is *totally bounded* if for every $\varepsilon > 0$ there exist a finite set $\{x_1, \ldots, x_m\} \subset X$ such that $M \subset \cup_{i=1}^m B_\varepsilon(x_i)$. A subset M of a complete metric space X is relatively compact iff it is totally bounded.

Definition A.10 *A collection $\mathcal{F}$ of closed balls in $\mathbb{R}^N$ is a* fine cover *of a set $E \subset \mathbb{R}^N$ iff*

$$E \subset \bigcup_{B \in \mathcal{F}} B,$$

and, for every $x \in E$

$$\inf\{\mathrm{diam}(B) \,|x \in B,\, B \in \mathcal{F}\} = 0,$$

where $\mathrm{diam}(B)$ *denotes the diameter of the ball B.*

Theorem A.11 (Vitali) *Let $E \in \mathcal{B}(\mathbb{R}^N)$ such that $\lambda(E) < \infty$*(3)*. Assume that $\mathcal{F}$ is a fine cover of E. Then, for every $\varepsilon > 0$ there exists a finite collection of disjoint balls $B_1, \dots, B_n \in \mathcal{F}$ such that*

$$\lambda\Big(E \setminus \bigcup_{i=1}^{n} B_i\Big) < \varepsilon.$$

Proof. According to Proposition 1.53, there exists an open set V such that $E \subset V$ and $\lambda(V) < \infty$. Possibly substituting $\mathcal{F}$ by the subcollection $\tilde{\mathcal{F}} = \{B \in \mathcal{F} \,|\, B \subset V\}$, which is still a fine cover of E, we may assume without loss of generality all the balls of $\mathcal{F}$ are contained in V. This implies

$$\sup\{\mathrm{diam}(B) \,|\, B \in \mathcal{F}\} < \infty.$$

We describe by induction the choice of $B_1, B_2, \dots, B_k, \dots$. We choose B_1 so that $\mathrm{diam}(B_1) > \frac{1}{2}\sup\{\mathrm{diam}(B) \,|\, B \in \mathcal{F}\}$. Let us suppose that $B_1, \dots, B_k$ have already been chosen. There are two possibilities: either

a) $E \subset \cup_{i=1}^{k} B_k$;

or

b) there exists $\bar{x} \in E \setminus \cup_{i=1}^{k} B_k$.

In the case a), we terminate at B_k and the thesis immediately follows. Assume that b) holds true. Since $\cup_{i=1}^{k} B_k$ is a compact set, we denote by $\delta > 0$ the distance of $\bar{x}$ from $\cup_{i=1}^{k} B_k$. Since $\mathcal{F}$ is a fine cover of E, there exists a ball $B \in \mathcal{F}$ such that $\bar{x} \in B$ and $\mathrm{diam}(B) < \frac{\delta}{2}$. In particular B is disjoint from

(3) $\mathcal{B}(\mathbb{R}^N)$ is the σ-algebra of the Borel sets of $\mathbb{R}^N$ and λ denotes the Lebesgue measure.

$B_1, \dots, B_k$. Then the set $\{B \in \mathcal{F} \mid B \text{ disjoint from } B_1, \dots, B_k\}$ is nonempty, hence we can define

$$d_k = \sup\{\operatorname{diam}(B) \mid B \in \mathcal{F},\ B \text{ disjoint from } B_1, \dots, B_k\} > 0.$$

We choose $B_{k+1} \in \mathcal{F}$ such that B_{k+1} is disjoint from $B_1, \dots, B_k$ and $\operatorname{diam}(B_{k+1}) > \frac{d_k}{2}$. If the process does not terminate, we get a sequence $B_1, B_2, \dots, B_k, \dots,$ of disjoint balls in $\mathcal{F}$ such that

$$\frac{d_k}{2} < \operatorname{diam}(B_{k+1}) \le d_k.$$

Since $\cup_{k=1}^{\infty} B_k \subset V$, we have $\sum_{k=1}^{\infty} \lambda(B_k) \le \lambda(V) < \infty$. Then there exists $n \in \mathbb{N}$ such that

$$\sum_{k=n+1}^{\infty} \lambda(B_k) < \frac{\varepsilon}{5^N}.$$

We claim that

$$E \setminus \bigcup_{k=1}^{n} B_k \subset \bigcup_{k=n+1}^{\infty} B_k^*, \tag{A.9}$$

where B_k^* denotes the ball having the same center as B_k but whose diameter in five times as large. Indeed let $x \in E \setminus \cup_{k=1}^{n} B_k$. By reasoning as in case b), there exists a ball $B \in \mathcal{F}$ such that $x \in B$ and B is disjoint from $B_1, \dots, B_n$. We state that B must intersect at least one of the balls B_k (with $k > n$), otherwise from the definition of d_k for every k it would result

$$\operatorname{diam}(B) \le d_k \le 2 \operatorname{diam}(B_{k+1});$$

since $\sum_k \lambda(B_k) < \infty$, then $\lambda(B_k) \to 0$, by which $\operatorname{diam}(B_k) \to 0$; consequently the above inequality cannot be true for large k.

Then we take the first j such that $B \cap B_j \neq \emptyset$. We have $j > n$ and

$$\operatorname{diam}(B) \le d_{j-1} < 2 \operatorname{diam}(B_j).$$

From an obvious geometric consideration it is then evident that B is contained in the ball that has the same center as B_j and five times the diameter of B_j, i.e. $B \subset B_j^*$. Thus we have proved (A.9), and so

$$\lambda\Big(E \setminus \bigcup_{k=1}^{n} B_k\Big) \le \sum_{k=n+1}^{\infty} \lambda(B_k^*) = 5^N \sum_{k=n+1}^{\infty} \lambda(B_k) \le \varepsilon$$

which proves the theorem. □

Bibliography

[1] CONWAY J.B., *A course in functional analysis* (second edition), Graduate Texts in Mathematics 96, Springer, New York, 1990.

[2] EVANS L.C., GARIEPY R.F., *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, Ann Arbor, 1992.

[3] KOMORNIK V., *Précis d'analyse réelle. Analyse fonctionnelle, intégrale de Lebesgue, espaces fonctionnels* (volume 2), Mathématiques pour le 2^e cycle, Ellipses, Paris, 2002.

[4] RUDIN W., *Analisi reale e complessa*, Programma di matematica, fisica, elettronica, Boringhieri, Torino, 1974.

Index