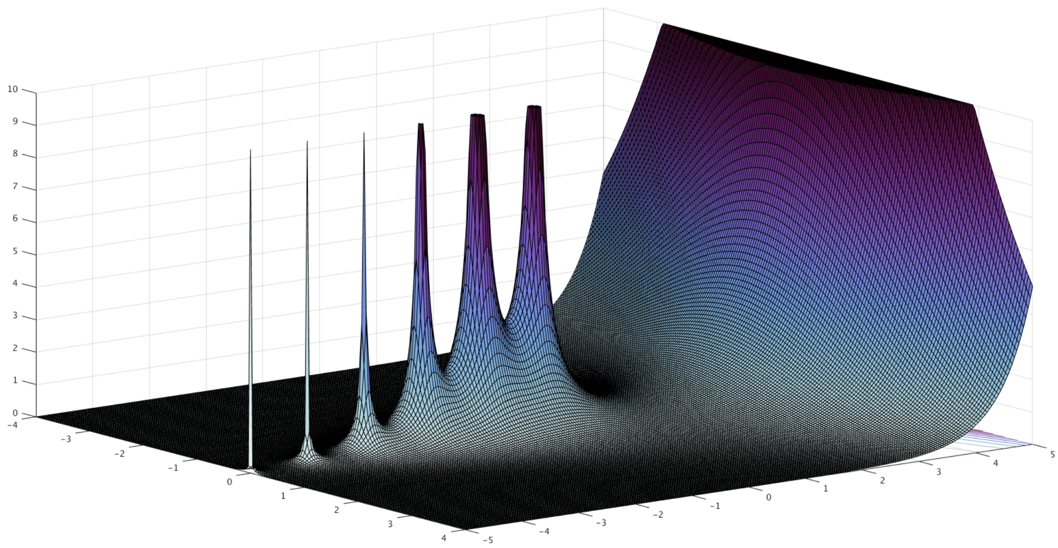
New York University
Courant Institute of Mathematical Sciences
Spring 2021

Introduction to Math Analysis II

MATH-GA.1420

Lecture notes



Samuel Boury

Spring 2021

Contents

Foreword

This document is a compilation of the lecture notes of *Introduction to Math Analysis* (Math-GA.1420) taught at NYU in Spring 2021. It is meant to help students to follow the course and provide exercises related to the notions covered. The lectures are articulated around the construction of Lebesgue's measure and theory of integration, with some applications such as the product of convolution and the Fourier transform.

The course is divided in ten chapters. Chapters 1, 2, 3, 4, 7, and 9, focus on the theoretical development of Lebesgue's integration (measure, integral, functional spaces), while chapters 5, 6, 8, and 10, are more related to the application of this theory. Although the class is restricted to the Lebesgue measure, some key elements on the general measure theory are also introduced. Similarly, an introduction to the theory of distributions is also provided in chapter 8 (with application in chapter 10). These topics require, of course, a more thorough discussion that is not proposed in these notes.

A few concepts and key results have not been discussed in the lectures, such as theorems on the existence and the unicity of the Lebesgue measure. If you are interested in delving more into these and willing to go beyond what we have seen in class, I can recommend you, among others, the following resources that have been used to prepare this class:

- **W. Appel**, *Mathematics for Physics and Physicists*, Princeton Univiversity Press, 2007;
- **F. Jones**, *Lebesgue Integration on Euclidean Space*, Jones and Barlett Publishers, 2001;
- **F. Burk**, *Lebesgue Measure and Integration: an Introduction*, John Wiley & Sons, 1998;
- **W. Rudin**, *Principles of Mathematical Analysis (*3^rd^ *ed.)*, MGraw-Hill, 1976;
- **P.R. Halmos**, *Measure Theory*, D. Van Nostrand Company, 1950.

Some other resources (in French) can also prove helpful: Thierry Gallay's notes from his lectures at Université Joseph Fourier (Grenoble, France), and Cédric Villani's notes from his lectures at École Normale Supérieure de Lyon (Lyon, France).

Finally, these notes being quite recent, they are likely to contain mistakes. If you find any typo, mistake, or comment that would be good to add, please let me knw by sending me an email at `sb7018@nyu.edu`. Many thanks in advance.

CHAPTER 1

Introduction to $\mathbb{R}^n$ and Elements of Topology

In this first chapter, we will review fundamental properties of the $\mathbb{R}^n$ set as well as basic topological concepts (open sets, closed sets, ...). By no means is this chapter a complete lecture on these notions, to which the reader is assumed to be somewhat familiar. Topological notions are defined on sets, without any further assumptions, and do not require precise knowledge of linear algebra although most of the concepts can be extended to vector spaces with many interesting properties.

1.1 Ensembles, Sets, and Subsets

Definition 1. Number systems (1/2).

1. $\mathbb{N}$: positive (or natural) integers, i.e. $\{0, 1, 2, 3, ...\}$.
2. $\mathbb{Z}$: (relative) integers, i.e. $\{..., -3, -2, -1, 0, 1, 2, 3...\}$.
3. $\mathbb{Q}$: rational numbers, i.e. $\{p/q | \, p \in \mathbb{Z}, q \in \mathbb{N}, q \neq 0\}$.
4. $\mathbb{R}$: real numbers.
5. $\mathbb{C}$: complex numbers, i.e. $\{x + iy | \, x \in \mathbb{R}, y \in \mathbb{R}\}$.
6. $\mathbb{K}$: either $\mathbb{R}$ or $\mathbb{C}$.
7. $\mathbb{U}$: unit circle, i.e. $\{e^{ix} | \, x \in \mathbb{R}\}$.

Definition 2. Number systems (2/2).

1. Let A be either $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. We note $A^\star = \{x \in A | \, x \neq 0\}$.
2. Let A be either $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. We note $A^+ = \{x \in A | \, x \geq 0\}$.
3. Let A be either $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. We note $A^- = \{x \in A | \, x \leq 0\}$.

Definition 3. Extended real system.
We note $\overline{\mathbb{R}}$ the extended real system, defined by $\overline{\mathbb{R}} = \{x \in \mathbb{R} \text{ or } x = \pm\infty\}$.

1.1.1 Sets and Subsets

In what follows, Ω is a set of elements.

Definition 4. Subset.
Let A and B be two sets of elements of Ω. If all elements of A are in B, we say that A is a *subset* of B, and we note $A \subset B$.

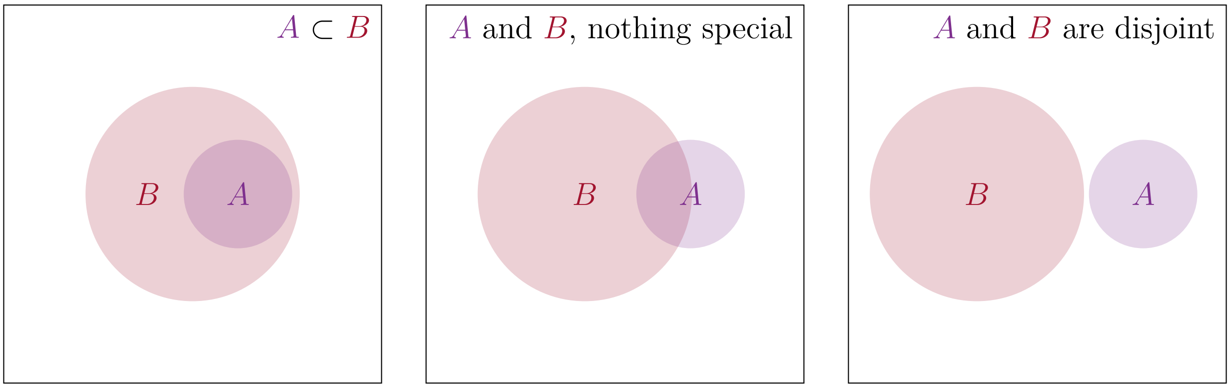


Figure 1.1: A can be a subset of B, or A and B can be disjoint, but A and B can have only some part in common and be neither subsets nor disjoint sets.

Remark 1. If two sets A and B satisfy $A \subset B$ and $B \subset A$, then $A = B$. Note that a usual way to prove that two sets A and B are equal is to prove that they are mutual subsets.

Definition 5. Set operators.

1. Union: $A \cup B = \{x \in \Omega | \, x \in A \text{ or } x \in B\}$
2. Intersection: $A \cap B = \{x \in \Omega | \, x \in A \text{ and } x \in B\}$
3. Difference: $A \setminus B = \{x \in \Omega | \, x \in A \text{ and } x \notin B\}$
4. Symmetric difference: $A \Delta B = \{x \in \Omega | \, x \in A \setminus B \text{ or } x \in B \setminus A\}$
5. Complement: $A^c = \{x \in \Omega | \, x \notin A\}$

Remark 2. The complement is always defined with respect to a set.

Definition 6. Disjoint sets.
We say that A and B are *disjoint* if $A \cap B = \emptyset$.

Definition 7. Cartesian product.
We call *Cartesian product* of A and B and we note $A \times B$ the set

$$A \times B = \{(x, y) | \, x \in A \text{ and } y \in B\}.$$

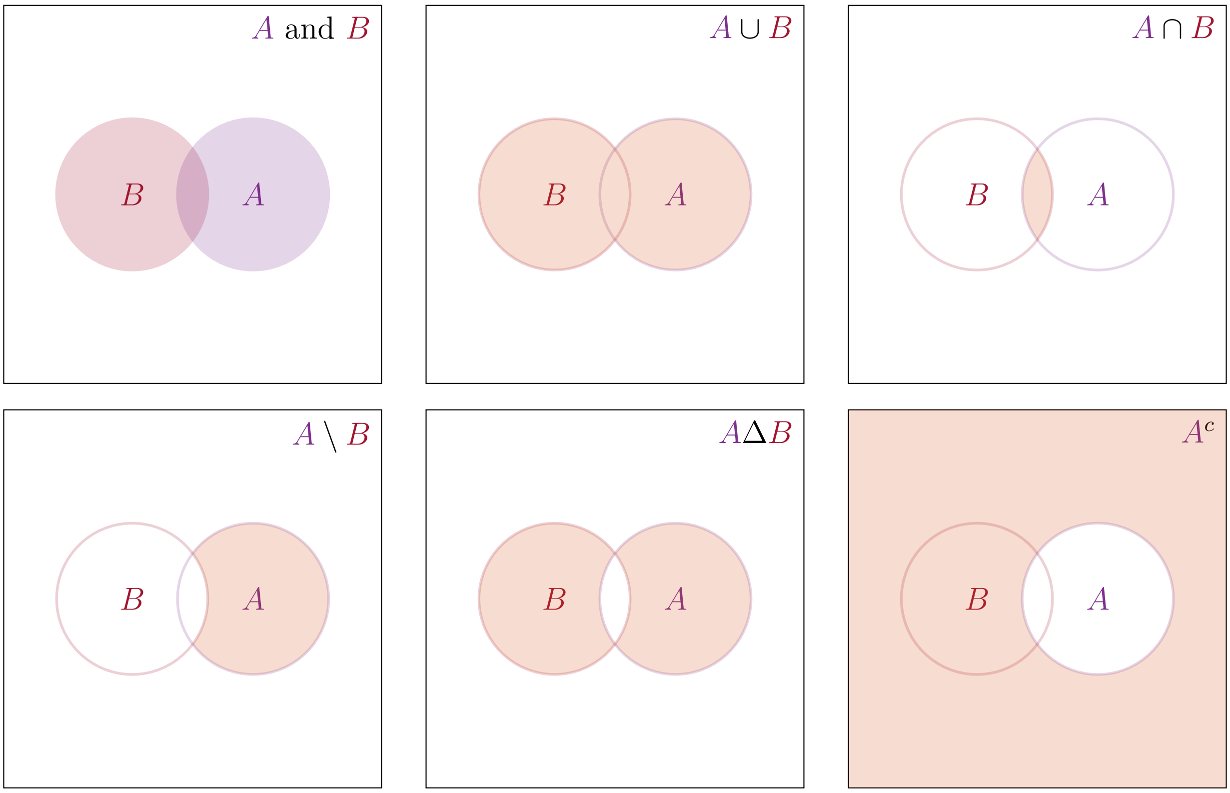


Figure 1.2: A and B are two subsets of Ω. This panel illustrates the set operators from definition 6 above. Note that A^c is understood as the complement of A with respect to Ω.

Remark 3. This definition naturally extends to a multi-product of sets and, notably, given $N \in \mathbb{N}$ subsets of Ω, A_1, A_2, ..., A_N, we can define their Cartesian product $A_\pi = A_1 \times A_2 \times \dots \times A_N$. The elements of A_π are ordered N-tuples. If the subsets A_1, A_2, ..., A_N, are all equal to a subset A, we will note $A_\pi = A^N$ the associated product space. In particular, given $n \in \mathbb{N}$, we note $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$.

Definition 8. Union and intersection of multiple sets.
Let I be an indexing set, and $\{A_k\}_{k \in I}$ a family of sets of Ω. The *union* and the *intersection* of these sets are defined, respectively, by

$$\bigcup_{k \in I} A_k = \{x \in \Omega | \, \exists k \in I, x \in A_k\}$$

$$\bigcap_{k \in I} A_k = \{x \in \Omega | \, \forall k \in I, x \in A_k\}$$

Remark 4. If the indexing set $I = \mathbb{N}$, we will write

$$\bigcup_{k \in \mathbb{N}} A_k = \bigcup_{k=0}^{+\infty} A_k \quad \text{and} \quad \bigcap_{k \in \mathbb{N}} A_k = \bigcap_{k=0}^{+\infty} A_k.$$

PROPOSITION 1. Morgan's Laws.
Let I be an indexing set, and $\{A_k\}_{k\in I}$ a family of sets of Ω. Then

$$\left(\bigcup_{k\in I} A_k\right)^c = \bigcap_{k\in I} A_k^c \quad \text{and} \quad \left(\bigcap_{k\in I} A_k\right)^c = \bigcup_{k\in I} A_k^c$$

EXERCISE 1. ★ Morgan's Laws.
Prove Morgan's Laws.

1.1.2 Applications

DEFINITION 9. Application.
Let A and B be two sets. An *application* f associates to each element of A an element of B. We note

$$\begin{aligned} f : A &\to B \\ x &\mapsto f(x) \end{aligned}$$

DEFINITION 10. Injectivity, sujectivity, and bijectivity.
Let A and B be two sets, and f and applicaton defined on A with values in B.

1. f is *injective* if $\forall (x, y) \in A^2, f(x) = f(y) \Leftrightarrow x = y$.
2. f is *surjective* if $\forall y \in B, \exists x \in A, f(x) = y$.
3. f is *bijective* if f is surjective and injective.

1.1.3 Countable Sets

DEFINITION 11. Countable set.
A set A is said to be *countable* if there is a one-to-one correspondance between it and a set of natural numbers, i.e. if there is a bijective application $f : I \to A$ where $I \subset \mathbb{N}$.

REMARK 5. A countable set can be finite, but it is not mandatory!

EXAMPLE 1.

1. The set of positive integers, $\mathbb{N}^\star$, is countable: we can think of a bijective application $f : n \mapsto n + 1$.
2. The sets of even and odd numbers are countable: we can use, respectively, the two applications $f : n \mapsto 2n$ and $f : n \mapsto 2n + 1$.
3. The set of prime numbers is countable: we can define by induction a sequence of prime numbers $(p_n)_{n\in\mathbb{N}}$, with $p_0 = 2$ and p_{n+1} the smallest prime number grater than p_n. Then, we have the bijective application $f : n \mapsto p_n$.

EXERCISE 2. ★ Countable sets.
Show that the following sets are countable:
$\mathbb{N} \times \mathbb{N}$, $\mathbb{Q}$, $\mathbb{N} \times \mathbb{Z} \times \mathbb{Q}$, and for $n \in \mathbb{N}$, $A_n = \{z \in \mathbb{U} |\, z^n = 1\}$.

1.2 Fundamentals of Topology

1.2.1 Norms

DEFINITION 12. Norm.
A norm on E is an application $N : E \to \mathbb{R}$ such that

1. $\forall x \in E, N(x) \geq 0$
2. $\forall x \in E, N(x) = 0 \Leftrightarrow x = 0$
3. $\forall \lambda \in \mathbb{K}, \forall x \in E, N(\lambda x) = |\lambda| N(x)$
4. $\forall (x, y) \in E^2, N(x + y) \leq N(x) + N(y)$ (triangle inequality)

and (E, N) is said to be a normed space.

EXAMPLE 2. The following applications are norms:

1. If $E = \mathbb{R}$ or $\mathbb{C}$, the modulus (absolute value) is a norm.
2. If $E = \mathbb{K}^n$, we have three norms:
   (a) Norm-∞ or uniform:
   $$||x||_\infty = \sup_{1 \leq i \leq n} |x_i|$$
   (b) Norm-1:
   $$||x||_1 = \sum_{i=1}^{n} |x_i|$$
   (c) Norm-2:
   $$||x||_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$
3. If $E = \mathcal{C}([a; b], \mathbb{K})$, we have three norms:
   (a) Norm-∞ or uniform:
   $$||f||_\infty = \sup_{x \in [a;b]} |f(x)|$$
   (b) Norm-1:
   $$||f||_1 = \int_a^b |f(t)| \mathrm{d}t$$
   (c) Norm-2:
   $$||f||_2 = \sqrt{\int_a^b |f(t)|^2 \mathrm{d}t}$$

EXERCISE 3. ★ A few norms.
Prove that the above examples are norms.

PROPOSITION 2. 1-Lipschitzian.
Let N be a norm on E. Then

$$\forall (x,y) \in E^2, |N(x) - N(y)| \leq N(x-y).$$

N is said to be 1-Lipschitzian.

PROOF.
Let $(x,y) \in E^2$.
We have $N(x) = N(x-y+y) \leq N(x-y) + N(y)$. So $N(x) - N(y) \leq N(x-y)$.
Similarly, we show that $N(y) - N(x) \leq N(x-y)$.
Therfore, $|N(x) - N(y)| \leq N(x-y)$.

□

1.2.2 Distance

DEFINITION 13. Distance.
A distance on E is an application $d : E^2 \to \mathbb{R}$ such that

1. $\forall (x,y) \in E^2, d(x,y) \geq 0$
2. $\forall (x,y) \in E^2, d(x,y) = d(y,x)$
3. $\forall (x,y) \in E^2 \in E, d(x,y) = 0 \Leftrightarrow x = y$
4. $\forall (x,y,z) \in E^3, d(x,y) \leq d(x,z) + d(z,y)$ (triangle inequality)

and (E,d) is said to be a metric space.

PROPOSITION 3. Metric & normed spaces.
Let (E,N) be a normed space. Then (E,d) is a metric space, where d is the distance associated to the norm N

$$\begin{array}{rcl} d : E \times E & \to & \mathbb{R}^+ \\ (x,y) & \mapsto & N(x-y) \end{array}$$

PROOF. Trivial, since the norm is compatible with the definition of a distance.

□

EXERCISE 4. ★ Comparison of distances.
Show that $\forall (x,y,z) \in E^3, |d(x,y) - d(x,z)| \leq d(y,z)$.

DEFINITION 14. Cauchy sequence.
A sequence $(u_n)_{n\in\mathbb{N}}$ in a metric space (E,d) is a *Cauchy sequence* if $d(u_m,u_n) \to 0$ when $(m,n) \to +\infty$, i.e.

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \forall m \geq N, d(u_m,u_n) \leq \varepsilon.$$

DEFINITION 15. Complete metric space.
A metric space is said to be *complete* if and only if every Cauchy sequence is convergent.

Example 3.

1. $\mathbb{Q}$, with the usual distance $d(p,q) = |p-q|$ for $(p,q) \in \mathbb{Q}^2$, is not complete: if we consider the sequence $(u_n)_{n\in\mathbb{N}}$ defined by

$$u_0 = 2 \text{ and } \forall n \in \mathbb{N}, u_{n+1} = \frac{u_n}{2} + \frac{1}{u_n}.$$

This is a Cauchy sequence, with all its terms in $\mathbb{Q}$, but is does not converge in $\mathbb{Q}$ since its limit is $\sqrt{2}$.

2. $\mathbb{R}$ and $\mathbb{C}$, with the absolute value / modulus norm, are complete. This is convenient for us, since we will work in $\mathbb{R}$ and $\mathbb{C}$.

1.2.3 Open and Closed Sets

We consider that (E,d) is a metric space.

Definition 16. Open ball.
Let $x \in E$ and $r \in \mathbb{R}$. The *open ball* of center x and radius r, noted $\mathcal{B}(x,r)$, is defined by

$$\mathcal{B}(x,r) = \{y \in E \mid d(x,y) < r\}.$$

Exercise 5. ★ Distance of two open balls.
Let $(x,x') \in E^2$ and $(r,r') \in \mathbb{R}^2$. Show that

1. $\mathcal{B}(x,r)$ and $\mathcal{B}(x',r')$ are disjoints if and only if $r + r' \leq d(x,x')$
2. $\mathcal{B}(x,r) \subset \mathcal{B}(x',r')$ if and only if $d(x,x') \leq r' - r$

Definition 17. Bounded set.
A subset A of (E,d) is *bounded* if and only if it is contained in an open ball.

Exercise 6. ★ Diameter of a set.
Let A be a non-empty bounded set. We define its diameter δ by

$$\delta(A) = \sup_{(x,y)\in A^2} d(x,y).$$

Show that δ exists. What is the diameter of an open ball of center x_0 and radius r, $\mathcal{B}(x_0,r)$?

Definition 18. Neighbourhood.
Let $x \in E$. We call *neighbourood* of x a subset $v \subset E$ such that there exists $r \in \mathbb{R}^{+\star}$ so that $\mathcal{B}(x,r) \subset v$.

Definition 19. Open set.
Let Ω be a subset of E. We say that Ω is open if and only if it is a neighbourhood of all its points, i.e.

$$\Omega \text{ open } \Leftrightarrow \forall x \in \Omega, \exists r \in \mathbb{R}^{+\star}, \mathcal{B}(x,r) \subset \Omega.$$

Example 4.

1. $\emptyset$ is open.
2. E is open.

DEFINITION 20. Closed set.
A subset F of E is closed if and only if its complementary F^c is open.

DEFINITION 21. Closure.
Let A be a subset of E. The closure of A, noted $\overline{A}$, is the intersection of all closed sets containing A (it is also the smallest closed set containing A).

PROPOSITION 4. Closed sets and closure.
A set E is closed if and only if it is equal to its closure.

$$\overline{E} = E.$$

THEOREM 1. Sequential characterization of the closure.
Let A be a subset of E. $a \in \overline{A}$ if and only if there exists a sequence $(a_n)_{n\in\mathbb{N}}$ of elements in A such that its limit is a, i.e.

$$a \in \overline{A} \Leftrightarrow \exists\,(a_n) \in A^{\mathbb{N}},\ \lim_{n\to+\infty} a_n = a.$$

COROLLARY 1. Sequential characterization of a closed set.
A subset A of E is closed if and only if for all converging sequence $(a_n)_{n\in\mathbb{N}}$ of elements in A, its limit is in A.

EXAMPLE 5.

1. $\emptyset$ is closed since E is open.
2. E is closed since $\emptyset$ is open.
3. $[0\,;1[\subset \mathbb{R}$ is neither open nor closed.

DEFINITION 22. Interior point.
Let $x \in E$. The point x is said to be an *interior point* of E if there exists $r \in \mathbb{R}^{+\star}$ such that $\mathcal{B}(x,r) \subset E$.

DEFINITION 23. Interior.
The *interior* of E, noted $\mathring{E}$, is the union of all open sets of E (it is also the largest open set contained in E), and can be defined by

$$\mathring{E} = \left\{ x \in E \middle|\, \exists r \in \mathbb{R}^{+\star}, \mathcal{B}(x,r) \subset E \right\}$$

PROPOSITION 5. Open sets and interior.
A set Ω is open if and only if it is equal to its interior.

$$\mathring{\Omega} = \Omega.$$

1.3 Compacity

Definition 24. Compact set (sequential definition).
A subset K of (E, N) is *compact* if and only if for every sequence of elements in K, one can extract a subsequence converging in K.

Definition 25. Compact set (topological definition).
A subset K of (E, N) is *compact* if and only if whenever K is contained in a union of open sets, then K is contained in a finite union of some of these sets.

Remark 6. These definitions are equivalent. I will mostly use the first one when necessary.

Exercise 7. ★ Elementary compacts.
Prove that:

1. $\emptyset$ is compact.
2. Any finite set is compact.
3. If A and B are compact, then so is $A \cup B$.
4. Any finite union of compact sets is compact.
5. $\mathcal{B}(x, r)$ is not compact.
6. $\mathbb{R}^n$ is not compact.

Proposition 6. Properties of a compact set.

1. A compact set is closed.
2. A compact set is bounded.

Proof.

1. Let K be a compact set, and $(a_n)_{n\in\mathbb{N}}$ a converging sequence of elements in K. Since K is compact, we can extract a subsequence $\left(a_{\phi(n)}\right)_{n\in\mathbb{N}}$ that converges in K. Yet, since $(a_n)_{n\in\mathbb{N}}$ converges, its limit is the same as $\left(a_{\phi(n)}\right)_{n\in\mathbb{N}}$, i.e.

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} a_{\phi(n)},$$

and is therefore still in K. Hence, every converging sequence of elements of K converges in K, so K is a closed set.

2. Let A be a non-empty unbounded subset of (E, N); we will show that A cannot be compact. In order to do so, we will build a non-converging sequence $(a_n)_{n\in\mathbb{N}}$ of elements of A. We proceed by induction:

   (a) Since A in non-empty, we can pick at least one element $a_0 \in A$.

(b) Let $n \in \mathbb{N}$ such that we have built a sequence of elements $(a_0, ...a_n) \in A^{n+1}$, satisfying $\forall i \neq j, N(a_i - a_j) \geq 1$. We note

$$R = \max_{0 \leq i \leq n} N(a_0 - a_i).$$

Since A is unbounded, $\exists a_{n+1} \in A, N(a_{n+1} - a_0) \geq R + 1$. Then

$$\begin{array}{rcl} \forall i \in [\![0\,;n]\!], N(a_{n+1} - a_i) & = & N(a_{n+1} - a_0 + a_0 - a_i) \\ & \geq & N(a_{n+1} - a_0) - N(a_0 - a_i) \\ & \geq & R + 1 - R \\ & \geq & 1. \end{array}$$

Adding an element to the the proposed sequence.

(c) By induction, we build a sequence $(a_n)_{n\in\mathbb{N}}$ of elements of A that satisfies the inequality $\forall (i,j) \in \mathbb{N}^2, i \neq j, N(a_i - a_j) \geq 1$.

Due to the previous inequality, we cannot extract any Cauchy sequence from $(a_n)_{n\in\mathbb{N}}$, i.e. we cannot extract any converging subsequence. Hence, A is not compact. □

Theorem 2. Bounded sequences of $\mathbb{K}^n$.
If we consider the normed space $(\mathbb{K}^n, ||\bullet||_\infty)$, where $||x||_\infty = \sup_{1\leq i\leq n} |x_i|$, a converging subsequence can be extracted from every bounded sequence of elements in $\mathbb{K}^n$.

Proof.
Let $(x_p)_{p\in\mathbb{N}}$ be a bounded sequence of elements in $\mathbb{K}^n$. For every $p \in \mathbb{N}$, we can write x_p as $(x_{p,1}, x_{p,2}, ..., x_{p,n})$. Since $(x_p)_{p\in\mathbb{N}}$ is bounded, the sequences $(x_{p,i})_{p\in\mathbb{N}}$ are all bounded. We can can extract from $(x_{p,1})_{p\in\mathbb{N}}$ a converging subsequence $\left(x_{\phi_1(p),1}\right)_{p\in\mathbb{N}}$ [1]; then, from $(x_{p,2})_{p\in\mathbb{N}}$, we can extract a converging subsequence $\left(x_{\phi_1(\phi_2(p)),2}\right)_{p\in\mathbb{N}}$, and so on.

We note $\phi = \phi_1 \circ \phi_2 \circ ... \circ \phi_n$. Then, $\left(x_{\phi(p)}\right)_{p\in\mathbb{N}}$ is converging. □

Theorem 3. Heine-Borel.
Let A be a set of $(\mathbb{K}^n, ||\bullet||_\infty)$. A is compact if and only if A is closed and bounded.

Proof. [2]

1. We already know one implication: if $K \subset \mathbb{K}^n$ is compact, then K is closed and bounded.
2. We now assume that $K \subset \mathbb{K}^n$ is closed and bounded. Let $(a_p)_{p\in\mathbb{N}}$ be a sequence of elements of K. This sequence is bounded (since K is bounded) and theorem 2 tells us that we can extract a converging subsequence $\left(a_{\phi(p)}\right)_{p\in\mathbb{N}}$. The limit of this (converging) subsequence is in K, since K is closed. Therefore, from every sequence of elements of K we can extract a subsequence converging in K, and K is compact. □

Remark 7. The previous theorem is actually a restriction of Heine-Borel theorem to the normed space $(\mathbb{K}^n, ||\bullet||_\infty)$, since it is true for any vector space of finite dimension. It is false, however, in infinite dimension. This falls beyond the scope of this class.

[1] This comes from the Bolzano-Weiestrass theorem: from any bounded sequence in $\mathbb{R}$ or $\mathbb{C}$, one can extract a converging subsequence.

[2] The equivalent proof, using the topological definition of a compact instead of its sequential characterization, can be found in F. Jones, *Lebesgue Integration on Euclidean Space.*

1.4 Applications and Functions

1.4.1 Functions

DEFINITION 26. Function.
We call *function* an application f defined on $\mathbb{K}^n$ and with values in $\mathbb{K}^n$, and we will note

$$\begin{array}{rcl} f : \mathbb{K}^n & \to & \mathbb{K}^n \\ x & \mapsto & f(x) \end{array}$$

REMARK 8.

1. A 1-variable function is a function defined on $\mathbb{K}$. In these lectures, however, we will use the wording "1-variable" function to refer to functions defined on $\mathbb{R}$.

2. A real (valued) function with values in $\mathbb{R}$, whereas a complex (valued) function is a function with values in $\mathbb{C}$. In these lectures, we will mostly focus on real functions.

DEFINITION 27. Characteristic function.
Let A be a subset of $\mathbb{K}^n$. The *characteristic function* of A, χ_A, is an application defined on $\mathbb{K}^n$ and with values in $\{0; 1\}$, such that

$$\begin{array}{rcl} \chi_A : \mathbb{K}^n & \to & \{0; 1\} \\ x & \mapsto & \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases} \end{array}$$

1.4.2 Asymptotics

Let f and g be two functions defined on $\mathbb{R}^n$ and with values in $\mathbb{R}^n$.

DEFINITION 28. Asymptotically bounded ("domination").
We say that f is bounded by g in $a \in \mathbb{R}^n$, and we write $f = \underset{x \to a}{O}(g)$ if

$$\exists M \in \mathbb{R}^+, \exists v \in v(a), \forall x \in v \cap \mathbb{R}^n, ||f(x)|| \leq M||g(x)||.$$

DEFINITION 29. Asymptotically dominated ("preponderance").
We say that f is negligible compared to g in $a \in \mathbb{R}^n$, and we write $f = \underset{x \to a}{o}(g)$ if

$$\forall \varepsilon > 0, \exists v \in v(a), \forall x \in v \cap \mathbb{R}^n, ||f(x)|| \leq \varepsilon||g(x)||.$$

DEFINITION 30. Asymptotically equivalent ("equivalence").
We say that f is equivalent to g in $a \in \mathbb{R}^n$, and we write $f \underset{x \to a}{\sim} g$ if

$$(f - g) = \underset{x \to a}{o}(||f||).$$

REMARK 9. When f and g are defined on $\mathbb{R}$, these definitions are often used in $+\infty$ and can be written as follows

1. $f = O(g) \Leftrightarrow \exists M \in \mathbb{R}^{+}, \exists \alpha \in \mathbb{R}^{+\star}, \forall x \in \mathbb{R}, x > \alpha, ||f(x)|| \leq M||g(x)||$
2. $f = o(g) \Leftrightarrow \forall \varepsilon > 0, \exists \alpha \in \mathbb{R}^{+\star}, \forall x \in \mathbb{R}, x > \alpha, ||f(x)|| \leq \varepsilon||g(x)||$
3. $f \sim g \Leftrightarrow (f - g) = o(||f||)$.

1.4.3 Continuity

Let f be a function defined on a subset A of the metric space $(\mathbb{R}^n, d)$ and with values in $(\mathbb{R}^n, d)$.

Definition 31. Continuity at a point.
Let $a \in A$. We say that f is *continuous at a* if

$$\forall \epsilon > 0, \exists \eta > 0, \forall x \in A, d(a, x) < \eta \Rightarrow d(f(a), f(x)) < \epsilon.$$

Remark 10. Compared to 1-variable functions defined on a subset of $\mathbb{R}$, this definition is a generalisation to the multi-variable case. It means, in particular, that the function f is continuous at $a \in A$ if f has a unique limit in a (and this limit is reached).

Definition 32. Continuity on an interval.
We say that f is *continuous on A* if f is continuous at every $a \in A$.

Exercise 8. ★ Continuity of 2-variable functions.
Discuss the continuity of these functions at $(0, 0)$:

$$f : (x, y) \mapsto x^2 + 3xy^2 \qquad \text{and} \qquad g : (x, y) \mapsto \frac{x + y^2}{x^2 + y^2}$$

Definition 33. Sets of continuous functions.
Let $I \subset \mathbb{R}$. We define

$$\begin{aligned}
\mathcal{C}^0(I, \mathbb{K}) &= \{f : I \to \mathbb{K} |\, f \text{ is continuous on } I\}, \\
\mathcal{C}^1(I, \mathbb{K}) &= \{f : I \to \mathbb{K} |\, f \text{ and } f' \text{ are continuous on } I\}, \\
\mathcal{C}^k(I, \mathbb{K}) &= \left\{f : I \to \mathbb{K} |\, f, f', ...f^{(k)} \text{ are continuous on } I\right\}, \\
\mathcal{C}^\infty(I, \mathbb{K}) &= \left\{f : I \to \mathbb{K} |\, \forall k \in \mathbb{N}, f^{(k)} \text{ is continuous on } I\right\}.
\end{aligned}$$

Definition 34. Image and inverse image.
The *image* of a set $X \subset A$ is formed by all the images of every point of this subset, i.e.

$$f(X) = \{f(x) |\, x \in X\}.$$

The *inverse image* of a set $Y \in f(A)$ is formed by all the points whose image is in Y, i.e.

$$f^{-1}(Y) = \{x \in A |\, f(x) \in Y\}.$$

Theorem 4. Image of a compact.
Let $f : A \to \mathbb{R}^n$ be a continuous fonction. The image by f of a compact set is compact.

Proof.

Let $K \subset A$ be a compact set and $f : A \to \mathbb{R}^n$ be a continuous fonction. We consider a sequence $(y_n)_{n\in\mathbb{N}}$ with $\forall n \in \mathbb{N}, y_n \in f(K)$. We can build a corresponding sequence $(x_n)_{n\in\mathbb{N}}$ defined by

$$\forall n \in \mathbb{N}, \exists x_n \in K, y_n = f(x_n).$$

Since all terms of this sequence are in K, which is compact, we can extract a subsequence $\left(x_{\phi(n)}\right)_{n\in\mathbb{N}}$ converging towards $x \in K$. Then, since f is continuous,

$$\lim_{n\to+\infty} y_{\phi(n)} = \lim_{n\to+\infty} f(x_{\phi(n)}) = f(\lim_{n\to+\infty} x_{\phi(n)}) = f(x) \in f(K).$$

Therefore, $\left(y_{\phi(n)}\right)_{n\in\mathbb{N}}$ converges in $f(K)$, i.e. for every sequence of elements in $f(K)$, one can extract a subsequence converging in $f(K)$, and $f(K)$ is compact. □

Theorem 5. Heine.
All real-valued continuous fonction f defined on a (non-empty) compact K is bounded, and its bounds are reached (i.e. inf and sup exist and are reached).

Proof.

Let $K \subset A$ be a non-empty compact. Since f is continuous, we know that $f(K)$ is a compact too and so $f(K)$ is closed and bounded. $f(K)$ is also non-empty because $K \neq \emptyset$, so $\inf f(K)$ and $\sup f(K)$ exist. We know that $\inf f(K) \in \overline{f(K)}$ (by definition), but $f(K)$ is closed so $\inf f(K) \in f(K)$. The same goes for $\sup f(K)$. □

1.4.4 Uniform Continuity

Let f be a function defined on a subset A of the metric space $(\mathbb{R}^n, d)$ and with values in $(\mathbb{R}^n, d)$.

Definition 35. Uniform continuity.
We say that f is *uniformly continuous* if and only if

$$\forall \varepsilon > 0, \exists \eta > 0, \forall (x, y) \in A^2, d(x, y) < \eta \Rightarrow d(f(x), f(y)) < \varepsilon.$$

Proposition 7. Uniform continuity & continuity.
If f is uniformly continuous on A, then f is continuous on A.

Remark 11. The reciprocal is not true!

Definition 36. Lipschitzian.
The function f is said to be k-Lipschitzian, $k \in \mathbb{R}^{+\star}$, if and only if

$$\forall (x, y) \in E^2, d(f(x), f(y)) \leq kd(x, y).$$

Remark 12. For $f : E \subset \mathbb{R} \to \mathbb{R}$, with the natural norm on $\mathbb{R}$, we have

$$\forall (x, y) \in E^2, |f(x) - f(y)| \leq k|x - y|.$$

Exercise 9. ★ Lipschitz.

1. Let $f \in \mathcal{C}^1(I \subset \mathbb{R}, \mathbb{K})$. Show that f is Lipschitzian if and only if f' is bounded.

2. Let $(E, ||\bullet||)$ be a normed space. Show that $||\bullet||$ is Lipschitzian.

3. Show that, if f and g are Lipschitzian, then $f \circ g$ is Lipschitzian.

Exercise 10. ★★ Norms and Lipschitz.
We define $\mathcal{L}([a;b], \mathbb{R})$ the set of Lipschitzian applications from $[a;b] \subset \mathbb{R}$ to $\mathbb{R}$.

1. Is $N_1(f) = \inf\{k | \forall (x,y) \in [a;b]^2, |f(x) - f(y)| \leq k|x-y|\}$ a norm?

2. Is $N_2(f) = |f(0)| + N_1(f)$ a norm?

Proposition 8. Uniform continuity & Lipschitz.
If f is Lipschitzian on A, then f is uniformly continuous on A.

Proof. This follows from taking $\eta = \varepsilon/(k+1)$ in the definition.

□

Remark 13. Again, the reciprocal is not true!

Theorem 6. Uniform continuity on a compact.
If f is continuous on a compact $K \subset A$, then f is uniformly continuous on K.

Proof.
Let assume that $f : K \subset A \to F$ is continuous, but not uniformly continuous. Then,

$$\exists \varepsilon > 0, \forall \eta > 0, \exists (x,y) \in K^2, d(x,y) < \eta \text{ and } d(f(x), f(y)) \geq \varepsilon.$$

We chose $\eta = 1/2^n$ with $n \in \mathbb{N}$. Then,

$$\exists \varepsilon > 0, \forall n \in \mathbb{N}, \exists (x_n, y_n) \in K^2, d(x_n, y_n) < 1/2^n \text{ and } d(f(x_n), f(y_n)) \geq \varepsilon.$$

However, K^2 is compact since it is a product of compacts, and we can extract from $((x_n, y_n))_{n\in\mathbb{N}}$ a subsequence $\left((x_{\phi(n)}, y_{\phi(n)})\right)_{n\in\mathbb{N}}$ that converges towards $(a,b) \in K^2$. f is continuous, and the distance d is continuous too (e.g. take the norm, which is 1-lipschitzian), so we can take the limit $n \to +\infty$. Therefore, we obtain

$$d(a,b) = 0 \qquad \text{and} \qquad d(f(a), f(b)) \geq \varepsilon,$$

i.e. $a = b$ and $f(a) \neq f(b)$, which is contradictory.

□

Remark 14.

1. A periodic function is uniformly continuous.

2. A continuous function on $\mathbb{R}$ with limits in $\pm\infty$ in uniformly continuous ($\overline{\mathbb{R}}$ is compact).

1.4.5 Convergence

Definition 37. Pointwise convergence.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $I \subset \mathbb{K}$. $(f_n)_{n\in\mathbb{N}}$ converges *pointwise* to the function f if and only if

$$\forall x \in \mathbb{K}, \lim_{n\to+\infty} f_n(x) = f(x).$$

Definition 38. Uniform convergence.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $I \subset \mathbb{K}$. $(f_n)_{n\in\mathbb{N}}$ converges *uniformly* to the function f if and only if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \Rightarrow \forall x \in \mathbb{K}, d(f_n(x), f(x)) \leq \varepsilon.$$

Remark 15.

1. Since we are considering $\mathbb{K}$, the uniform convergence of a sequence of functions is equivalent to say that the sequence of functions satisfy the Cauchy criterion.
2. The uniform convergence is a stronger convergence criterion than the pointwise convergence: a uniformly converging sequence of function also converges pointwise, but the reciprocal is not true.

1.5 Summary

This first chapter provides an introduction to some topological concepts and to the real ensemble, $\mathbb{R}$. The motivation for this is to begin with precise definitions of open and closed sets, with the basic operations (union, intersection, etc) that will be useful in the following chapters to build the Lebesgue measure and the notion of Lebesgue-measurable sets. In particular, we have discussed various categories of sets such as normed spaces, metric spaces, and compact spaces. They provide the "good" framework since they are stable with respect to the basic operations (on sets and on elements) and preserve convergence properties, especially when working with real-valued functions. With the norm and the distance, the notions of open balls and adherence, we might even get an idea of what could be the "size" of a set, or its "measure".

CHAPTER 2

Lebesgue Measure on $\mathbb{R}^n$

If we consider a function f and a domain $D \subset \mathbb{R}^n$ (we will make no further assumptions on f and D as of now), when we write the integral

$$\int_D f(x)\mathrm{d}x,$$

how can we interpret $\mathrm{d}x$? Newton and Leibniz already proposed an approach to this mathematical object by introducing infinitesimal calculus: $\mathrm{d}x$ is a small part of the integration domain D, and we assume that $\mathrm{d}x \to 0$. Similarly, we write the derivative of a function as

$$f'(x) = \lim_{\mathrm{d}x \to 0} \frac{f(x + \mathrm{d}x) - f(x)}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}x}(x).$$

This means that $\mathrm{d}x$ can be an infinitely small portion of a length, of an area, of a volume, etc (see figure 2.1). A priori, we can easily measure the "size" of a domain D in $\mathbb{R}^n$ by summing all the infinitely small portions that consitute this domain. This approach is notably used by Riemann to define the integral as a limit of a sum (discussed in a following chapter) and by physicists to derive continuous theories (soft matter, fluid dynamics, etc).

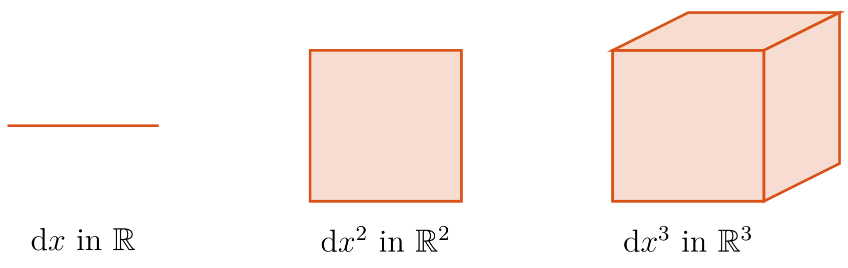


Figure 2.1: Elementary infinitesimal elements in $\mathbb{R}$, $\mathbb{R}^2$, and $\mathbb{R}^3$.

Such a definition seems to work, but there are two main problems here:

1. What happens when the domain D has a complicated shape? For example, if it is not connex, or if it is a collection of sets of various shapes, or if some sets are point-sets, or if point-sets are removed from some sets, or if some parts of the domain have a different weight? We want to use the characteristic function of the set, but what would be the meaning of "$\mathrm{d}\chi_D(x)$ "?

2. What happens when the function f does not have "good properties"? For example, if it is discontinuous on an interval, or if it is not discontinuous in a simple way but from point to point?

Hence, before delving into integration theory, we shall define a rigorous way to quantify the *area* of subsets of $\mathbb{R}^n$, which is called *measure*. The mathematical concept of *measure* is, formally, the generalisation of the notions of length (on $\mathbb{R}$), area (on $\mathbb{R}^2$), and volume (on $\mathbb{R}^3$), of which we all have an intuitive definition.

This is not an easy problem, and we can consider the following argument to see that we have to be careful while defining such a *measure*. Let us define ξ the "naive" length of subsets of $\mathbb{R}$. If we consider an interval of $\mathbb{R}$, say $I = [0\,;1]$, then we want its length to be $\xi(I) = 1$. If we now consider the open interval $]0\,;1]$, its length cannot be greater than 1 since it is contained in I (we reasonnably assume that the length application ξ is monotonic). For all $n \in \mathbb{N}, n \geq 2$, we define the intervals $I_n = [1/n\,;1-1/n]$, which are all contained in $]0\,;1]$. Using our intuitive definition of the length of an interval, we have $\forall n \in \mathbb{N}, n \geq 2, \xi(I_n) = 1 - 2/n$. As we expect ξ to be monotonic, we deduce that

$$\lim_{n \to +\infty} \xi(I_n) = 1,$$

hence we should have $\xi\left(\,]0\,;1]\,\right) = 1$. Since $I = \{0\} \cup]0;1]$ and $\xi(I) = \xi(\,]0\,;1]\,) = 1$, this naturally leads to an intuitive definition of the length of a singleton $\{x\}$ for $x \in \mathbb{R}$, as $\xi(\{x\}) = 0$, consistent with the statement that "points are dimensionless". This, however, leads to paradox[1]: if we define I as

$$I = \bigcup_{x \in I} \{x\}\,,$$

then the length of the left-hand side is 1, whereas the length of the right-hand side is a sum of lengths of singletons, i.e. 0. This simple example shows that uncountable sums are an issue with the "naive" definition of a length, ξ, and that the length of some subsets of $\mathbb{R}$ might not be well-defined (or not defined at all).

This second chapter is dedicated to the construction and characterisation of a proper way to measure the sizes of subsets of $\mathbb{R}^n$, called the *Lebesgue measure*. Note that there are different approaches to define the Lebesgue measure, for example: (1) by a step-by-step construction (see Jones), in which larger and larger subsets of $\mathbb{R}^n$ are considered with an appropriate application to measure their size; or (2) by defining rigorously the right family of subsets of an ensemble and showing that its measure, when restricted to $\mathbb{R}^n$, happens to be the Lebesgue measure. Here, the approach is slightly different (Burk, Rudin) as we first construct the Lebesgue measure, then construct the relevant subsets.

2.1 Sets, Subsets, and Set Functions

2.1.1 Sets and Subsets

Definition 39. Ring of sets.
A family $\mathcal{S}$ of sets is called a ring if

1. $\mathcal{S} \neq \emptyset$
2. $\forall (A, B) \in \mathcal{S}^2, A \cup B \in \mathcal{S}$
3. $\forall (A, B) \in \mathcal{S}^2, A \setminus B \in \mathcal{S}$

[1] This "paradox" is taken from F. Burk's book, *Lebesgue Measure and Integration*.

Proposition 9. Stability of a ring of sets.
Let $\mathcal{S}$ be a ring of sets. The following properties are satisfied:

1. The empty set is an element of $\mathcal{S}$: $\emptyset \in \mathcal{S}$
2. $\mathcal{S}$ is stable by intersection: $\forall (A, B) \in \mathcal{S}^2, A \cap B \in \mathcal{S}$
3. $\mathcal{S}$ is stable by symmetric difference: $\forall (A, B) \in \mathcal{S}^2, A\Delta B \in \mathcal{S}$

Proof.
We shall use the definition of a ring of sets to re-write the basic set operations in terms of stable operations. This yields

1. Let $A \in \mathcal{S}$, we write $\emptyset = A \setminus A$ and since a ring is stable by difference, then $\emptyset \in \mathcal{S}$.
2. Let $(A, B) \in \mathcal{S}^2$, we write $A \cap B = A \setminus (A \setminus B)$ and since a ring is stable by difference, then $A \cap B \in \mathcal{S}$.
3. Let $(A, B) \in \mathcal{S}^2$, we write $A\Delta B = (A \setminus B) \cup (B \setminus A)$ and since a ring is stable by finite union and by difference, then $A\Delta B \in \mathcal{S}$.

$\square$

Remark 16. A ring of sets can also be defined as a non-empty set of sets, stable by finite intersection and by symmetric difference. The stability by finite union and by difference are then consequences of the first two.

Definition 40. σ-ring of sets.
Let $\mathcal{S}$ be a ring of sets. $\mathcal{S}$ is said to be a *σ-ring of sets* if it is stable by countable union, i.e. if $\{A_k\}_{k\in\mathbb{N}}$ is a family of sets in $\mathcal{S}$, then

$$\bigcup_{k=0}^{+\infty} A_k \in \mathcal{S}.$$

Proposition 10. Stability by countable intersection.
A σ-ring is stable by countable intersection, i.e. if $\mathcal{S}$ is a σ-ring of sets and $\{A_k\}_{k\in\mathbb{N}}$ a family of sets in $\mathcal{S}$, then

$$\bigcap_{k=0}^{+\infty} A_k \in \mathcal{S}.$$

Proof.
We write the countable intersection in terms of a countable union

$$\bigcap_{k=0}^{+\infty} A_k = A_0 \setminus \bigcup_{k=0}^{+\infty} (A_0 \setminus A_k),$$

and since a σ-ring is stable by difference (ring) and countable union (σ-ring), then it is stable by countable intersection.

$\square$

Exercise 11. ★ Examples of σ-rings.
Show that the following collections of sets C form a σ-ring on a set Ω

1. $C = \{\emptyset, \Omega\}$.

2. $C = 2^\Omega$, the collection of all subsets of Ω.

3. $C = \{\emptyset, \{2k | k \in \mathbb{N}\}, \{2k+1 | k \in \mathbb{N}\}, \mathbb{N}\}$, with $\Omega = \mathbb{N}$.

Remark 17. Finite additivity vs. countable additivity.

The difference between a ring of sets and a σ-ring of sets is that a ring of sets is stable by union, and finite union (can be shown by induction from the stability by union), whereas a σ-ring of sets is also stable by countable union. The issue with the countable union is that one may easily get infinite sets.

We consider the two following examples:

1. We define $\mathcal{A} = \{A_n = [\![1; n]\!]\}$. The set of sets $\mathcal{A}$ is not empty, since the singleton $\{1\} \in \mathcal{A}$. Moreover, $\mathcal{A}$ is stable by union: $\forall (m, n) \in \mathbb{N}^2, A_n \cup A_m = A_{\max(m,n)} \in \mathcal{A}$. Hence, any finite union of elements of $\mathcal{A}$ is still in $\mathcal{A}$, since it will be of the form $[\![1; N]\!]$ with $N \in \mathbb{N}$. If we now consider the countable union
$$\bigcup_{n=1}^{+\infty} A_k = \mathbb{N},$$
then, this is not an element of $\mathcal{A}$ since this set cannot be written in the form $[\![1; N]\!]$ with $N \in \mathbb{N}$. $\mathcal{A}$ is stable by finite union, but not by countable union.

2. We consider $\mathcal{N}$ the set of finite sets of $\mathbb{N}$. We can show that $\mathcal{N}$ is a ring of sets, since it is not empty (it contains singletons), it is stable by finite union, and it is stable by difference. Yet, $\mathcal{N}$ is not a σ-ring of sets, because the countable union of elements of $\mathcal{N}$, such as the A_k previously defined, can be an infinite subset of $\mathbb{N}$.

Definition 41. σ-algebra of sets.
A σ-ring of sets $\mathcal{S}$ is said to be a σ-algebra of sets if it is stable by complement, i.e. if
$$\forall A \in \mathcal{S}, A^c \in \mathcal{S}.$$

Remark 18. Let define a σ-ring of sets $\mathcal{R}$ and a σ-algebra of sets $\mathcal{A}$, on a set Ω. The difference between $\mathcal{R}$ and $\mathcal{A}$ is that the σ-algebra $\mathcal{A}$ contains Ω, whereas the σ-ring $\mathcal{R}$ does not necessarily contains Ω. For example, $\mathcal{S} = \{\emptyset\}$ is a (trivial) σ-ring of Ω, but not a σ-algebra (since it is not stable by complement).

Exercise 12. ★ Examples of σ-algebras.
Show that the following collections of sets C form a σ-algebra:

1. $C = \{\emptyset, \Omega\}$.

2. $C = 2^\Omega$, the collection of all subsets of Ω.

3. $C = \{\emptyset, \{2k | k \in \mathbb{N}\}, \{2k+1 | k \in \mathbb{N}\}, \mathbb{N}\}$, with $\Omega = \mathbb{N}$.

2.1.2 Set Functions

Definition 42. Set function.
Let $\mathcal{S}$ be a family of sets. A *set function* is an application ϕ defined as
$$\begin{array}{rcl} \phi : \mathcal{S} & \to & \overline{\mathbb{R}} \\ A & \mapsto & \phi(A) \end{array}$$

Definition 43. Positivity.
A set function ϕ defined on a family of sets $\mathcal{S}$ is said to be *positive* if

$$\forall A \in \mathcal{S}, \phi(A) \geq 0.$$

Definition 44. Additivity.
A set function ϕ defined on a family of sets $\mathcal{S}$ is said to be *additive* if

$$\forall (A, B) \in \mathcal{S}^2, A \cap B = \emptyset \Rightarrow \phi(A \cup B) = \phi(A) + \phi(B).$$

Definition 45. Countable additivity.
A set function ϕ defined on a family of sets $\mathcal{S}$ is said to be *countably additive* if, given a family $\{A_k\}_{k \in \mathbb{N}}$ of non-overlapping elements of $\mathcal{S}$ (i.e. $\forall (k, l) \in \mathbb{N}^2, k \neq l \Rightarrow A_k \cap A_l = \emptyset$), we have

$$\phi\left(\bigcup_{k=0}^{+\infty} A_k\right) = \sum_{k=0}^{+\infty} \phi\left(A_k\right).$$

Remark 19.
If we have an inequality instead of the equality, like

$$\phi\left(\bigcup_{k=0}^{+\infty} A_k\right) \leq \sum_{k=0}^{+\infty} \phi\left(A_k\right),$$

then the set function ϕ is said to be *countably subadditive*.

2.2 Construction of the Lebesgue Measure

2.2.1 Length of an Interval

Definition 46. Length of an interval.
Let $(a, b) \in \mathbb{R}^2, a \leq b$. The *length* l of an interval I of endpoints a and b is defined by

$$l(I) = b - a$$

Remark 20. The length application has its values in $\overline{\mathbb{R}}^+$. It is a positive set function.

Example 6. A direct application of the length definition yields:

1. $l\,(\,]0\,;1[\,) = l\,(\,]0\,;1]\,) = l\,(\,[0\,;1]\,) = 1$
2. Even though it is not rigorously defined, if $a = b$ then the length of a singleton is 0.
3. $l\,(\,[1\,;+\infty[\,) = +\infty$

Exercise 13. ★ Length of intervals.
Calculate the lengths of the following intervals:

1. $I_1 =]-1;\, 1[)$
2. $I_2 = \{y | \, x \in I_1, y = 2x + 1\}$

3. $I_3 = \{y | \, x \in I_1, y = x^2\}$

EXERCISE 14. ★ Monotonicity.

Show that the length application l is monotonic, i.e. if I_1 and I_2 are two real intervals such that $I_1 \subset I_2$, then $l(I_1) \leq l(I_2)$. Show that if $\{I_k\}_{k\in\mathbb{N}}$ is a sequence of intervals of $\mathbb{R}$ such that $\forall k \in \mathbb{N}, I_k \subset I_{k+1}$, then $\forall (p,q) \in \mathbb{N}^2, p < q \Rightarrow l(I_p) \leq l(I_q)$.

EXERCISE 15. ★ Invariance by translation.

Show that the length is invariant by translation, i.e. if $I \subset \mathbb{R}$ is an interval and $c \in \mathbb{R}$, then $l(I) = l(I+c)$ where $I + c = \{x \, | x - c \in I\}$.

EXERCISE 16. ★★ Finite covers.

Let I be an interval of $\mathbb{R}$. Let $N \in \mathbb{N}$ and $\{I_0, I_1, ...I_N\}$ a finite family of bounded open intervals such that

$$I \subset \bigcup_{k=0}^{N} I_k.$$

Show that the length of I is smaller than the length of its finite cover, i.e.

$$l(I) \leq \sum_{k=0}^{N} l(I_k).$$

EXERCISE 17. ★★ Lengths of sub-intervals.

Let I be a bounded open interval of $\mathbb{R}$.

1. If $\{I_0, I_1, ...I_N\}$ is a finite family of mutually disjoint open intervals such that

$$\bigcup_{k=0}^{N} I_k \subset I,$$

show that

$$\sum_{k=0}^{N} l(I_k) \leq l(I).$$

2. If $\{I_k\}_{k\in\mathbb{N}}$ is a countable collection of mutually disjoint open intervals such that

$$\bigcup_{k=0}^{+\infty} I_k \subset I,$$

show that

$$\sum_{k=0}^{+\infty} l(I_k) \leq l(I).$$

3. If I_1 and I_2 are two disjoint intervals such that $I = I_1 \cup I_2$, show that $l(I) = l(I_1) + l(I_2)$.

LEMMA 1. Lindelöf.

Any open cover of a set of real numbers contains a countable subcover.

PROOF.

Let $A \subset \mathbb{R}$ and $\{A_x\}_{x\in X\subset\mathbb{R}}$ a cover of A.

If we pick $a \in A$, then there exists at least one $x_a \in X$ such that $a \in A_{x_a}$. Since A_{x_0} is open, then we can find $\varepsilon > 0$ such that $]a-\varepsilon;\, a+\varepsilon[\subset A_{x_a}$.

Let $(p, q) \in \mathbb{Q}$ such that $a - \varepsilon < p < a < q < a + \varepsilon$. We can associate to each point a of A an open interval with rational endpoints $I(a) =]p; q[$ that contains a and that is a subset of (at least) one open set of the cover $\{A_x\}_{x \in X \subset \mathbb{R}}$ (in this construction, it is A_{x_a}). We define

$$\mathcal{I} = \{I(a) \,|a \in A\},$$

and since for each $a \in A$ we can find at least one open set of the cover $\{A_x\}_{x \in X \subset \mathbb{R}}$ that contains $I(a)$ we construct $\mathcal{A}$ the collection of such sets, so that there is a correspondance one-to-one between $\mathcal{I}$ and $\mathcal{A}$. We note that the collection of intervals $\mathcal{I}$ is a cover of A, as well as $\mathcal{A}$.

With this construction, it is likely that there exists two elements $(a_1, a_2) \in A^2$ such that $I(a_1) = I(a_2)$: there is no bijection, a priori, between A and $\mathcal{I}$.

Then, since $\mathbb{Q}$ is countable, so is the set of all possible open intervals with rational endpoints, and so is $\mathcal{I}$. As we have a bijection between $\mathcal{I}$ and $\mathcal{A}$, the collection of sets $\mathcal{A}$ is also countable. Hence, $\mathcal{A}$ is a countable subcover of A. □

Proposition 11. Comparison of lengths.
Let I be a bounded open interval of $\mathbb{R}$ and $\{I_k\}_{k \in \mathbb{N}}$ a countable family of bounded open intervals of $\mathbb{R}$ such that $I \subset \bigcup I_k$. Then

$$l(I) \leq \sum_{k=1}^{+\infty} l(I_k).$$

Corollary 2. Comparison of lengths.
If I be a bounded open interval of $\mathbb{R}$, then

$$l(I) = \inf \left\{ \sum_{k=1}^{+\infty} l(I_k) \middle| I \subset \bigcup_{k=1}^{+\infty} I_k \text{ and } \forall k \in \mathbb{N}, I_k \text{ is an open interval} \right\}.$$

Proof.
Let $I =]a; b[$ be a bounded open interval of $\mathbb{R}$ and $\{I_k\}_{k \in \mathbb{N}}$ a countable family of bounded open intervals of $\mathbb{R}$ covering I. If the cover family is not countable, we use Lindelöf's lemma to reduce it to a countable subcover family.

We introduce $K = [a; b]$. $K \subset \mathbb{R}$ is closed and bounded, so K is compact (Heine-Borel theorem). Using the topological definition of a compact (definition 25) we know that from any countable cover of open sets of K we can extract a finite subcover of open sets.

Let $\varepsilon > 0$. Then, $\{]a - \varepsilon/4; a + \varepsilon/4[,]b - \varepsilon/4; b + \varepsilon/4[\} \cup \{I_k\}_{k \in \mathbb{N}}$ is a countable family of bounded open intervals covering K. Using the monotonicity of the length application and the result from the exercise on finite covers, we have

$$l(I) \leq l(K) \leq \varepsilon + \sum_{k=1}^{N} l(I_k) \leq \varepsilon + \sum_{k=1}^{+\infty} l(I_k).$$

Since this is true for all $\varepsilon > 0$, we conclude that

$$l(I) \leq \sum_{k=1}^{+\infty} l(I_k).$$

The corollary follows, since it is true for all family of bounded open intervals covering I, it is in particular true for the family that contains I and empty intervals

$$l(I) = \inf \left\{ \sum_{k=1}^{+\infty} l(I_k) \middle| I \subset \bigcup_{k=1}^{+\infty} I_k \text{ and } \forall k \in \mathbb{N}, I_k \text{ is an open interval} \right\}.$$ □

2.2.2 Lebesgue Outer Measure on $\mathbb{R}$

DEFINITION 47. Lebesgue Outer Measure.
Let A be a subset of $\mathbb{R}$. The *Lebesgue outer measure* $\mu^\star$ of A is given by

$$\mu^\star(A) = \inf\left\{ \sum_{k=1}^{+\infty} l(I_k) \middle| A \subset \bigcup_{k=1}^{+\infty} I_k \text{ and } \forall k \in \mathbb{N}, I_k \text{ is an open interval} \right\}.$$

THEOREM 7. Properties of the Lebesgue outer measure $\mu^\star$.

1. $\mu^\star$ is defined for every set of real numbers;
2. Definite: $\mu^\star(\emptyset) = 0$;
3. Positivity: for any subset $A \subset \mathbb{R}$, $0 \leq \mu^\star(A) \leq +\infty$;
4. Monotonicity: for any subsets $A \subset \mathbb{R}$ and $B \subset \mathbb{R}$, if $A \subset B$ then $\mu^\star(A) \leq \mu^\star(B)$;
5. Points are dimensionless: $\forall a \in \mathbb{R}, \mu^\star(\{a\}) = 0$;
6. Length of an interval: if $I \subset \mathbb{R}$ is an interval, then $\mu^\star(I) = l(I)$;
7. Invariance by translation: $\forall A \subset \mathbb{R}, \forall a \in \mathbb{R}, \mu^\star(A + a) = \mu^\star(A)$;
8. Countable subadditivity: for any sequence $(A_k)_{k\in\mathbb{N}}$ of subsets of $\mathbb{R}$,

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

PROOF.

1. Let A be a set of real numbers. The Lebesgue outer measure of A is defined if and only if there exists an open cover of A in $\mathbb{R}$. By definition of A, there is at least one open subset of $\mathbb{R}$ that contains A ($\mathbb{R}$ itself), and the inf is well defined.

2. The empty set is a subset of every subset of $\mathbb{R}$. In particular, if $a \in \mathbb{R}$, then $\emptyset \subset \{a\}$. Using the positivity and the monotonicity of $\mu^\star$, we have $0 \leq \mu^\star(\emptyset) \leq \mu^\star(\{a\})$, and since points are dimensionless we deduce $\mu^\star(\emptyset) = 0$.

3. Let A be a set of real numbers. As mentioned before, the collection of countable open sets covering A is not empty (it contains at least $\mathbb{R}$) so the set

$$\left\{ \sum_{k=1}^{+\infty} l(I_k) \middle| A \subset \bigcup_{k=1}^{+\infty} I_k \text{ and } \forall k \in \mathbb{N}, I_k \text{ is an open interval} \right\},$$

is not empty and contains a lower bound. Since the length of an interval is positive, this lower bound is positive, and $0 \leq \mu^\star(A)$.

4. Let A and B be two subsets of $\mathbb{R}$ with $A \subset B$. Their Lebesgue outer measure exist (subsets of $\mathbb{R}$). We note $\{I_{B,k}\}_{k\in\mathbb{N}}$ a family of open intervals of $\mathbb{R}$ covering B, and S_B the sum of their length. By definition, $\mu^\star(B) \leq S_B$. Since $A \subset B$, any open cover of B is also an open cover of A, so $\mu^\star(A) \leq S_B$. This is true for any family of open intervals covering B, hence $\mu^\star(A) \leq \mu^\star(B)$.

5. Let $a \in \mathbb{R}$ and $\varepsilon > 0$. The point-set $\{a\}$ is contained in the open set $I_\varepsilon =]a-\varepsilon;\, a+\varepsilon[$. Since $\mu^\star$ is monotonic, $0 \leq \mu^\star(\{a\}) \leq \mu^\star(I_\varepsilon)$. This inequality is true for all $\varepsilon > 0$ so it follows $\mu^\star(\{a\}) = 0$.

6. We proceed by steps, because we should prove that for any interval of $\mathbb{R}$ (open, closed), $\mu^\star(I) = l(I)$.

   (a) Let $I =]a;\, b[\subset \mathbb{R}$ be a bounded open interval. I is a cover of itself, so by definition of $\mu^\star$ we have $\mu^\star(I) \leq l(I) = b-a$. Yet, corollary 2 tells us that $l(I)$ is smaller than the inf, so $l(I) \leq \mu^\star(I)$. Therefore, $\mu^\star(I) = l(I) = b-a$.

   (b) Let $I = [a;\, b[\subset \mathbb{R}$. For all $\varepsilon > 0$, we can write that $]a;\, b[\subset I \subset]a-\varepsilon;\, b[$. With the monotonicity of $\mu^\star$, we now have $\forall \varepsilon > 0, b-a \leq \mu^\star(I) \leq b-a+\varepsilon$, from which we deduce that $\mu^\star(I) = b-a = l(I)$. The same reasoning applies for $I =]a;\, b] \subset \mathbb{R}$.

   (c) Let $I = [a;\, b] \subset \mathbb{R}$ be a bounded closed interval. Similarly to $[a;\, b[$ and $]a;\, b]$, we show that $\mu^\star(I) = b-a = l(I)$.

   (d) Let $I =]a;\, +\infty[\subset \mathbb{R}$. Then, by definition of $\mu^\star$, we have $\mu^\star(I) \leq +\infty = l(I)$. For all $b \in \mathbb{R}, b > a$, we have $]a;\, b[\subset I$ so, by monotonicity of $\mu^\star$, we can write $\forall b \in \mathbb{R}, b > a, b-a \leq \mu^\star(I) \leq +\infty$. This is true for any b as large as possible, so we deduce $\mu^\star(I) = +\infty = l(I)$. The same goes if we consider an interval of the form $]-\infty;\, b[\subset \mathbb{R}$ and, as shown before, if we change the open bounds by closed bounds.

7. We already know (see previous exercises) that the length is invariant by translation. Let A be a subset of $\mathbb{R}$ and $c \in \mathbb{R}$. If we consider a family of open intervals $\{I_k\}_{k\in\mathbb{N}}$ of $\mathbb{R}$ covering A, then we also have

$$(A+c) \subset \bigcup_{k=1}^{+\infty} (I_k + c).$$

   Then, by definition of $\mu^\star$, we have

$$\mu^\star(A+c) \leq \sum_{k=1}^{+\infty} l(I_k + c) = \sum_{k=1}^{+\infty} l(I_k).$$

   This means that $\mu^\star(A+c)$ is a lower bound for the lengths of covers of A, from what we deduce $\mu^\star(A+c) \leq \mu^\star(A)$ using the definition of $\mu^\star$.

   If we now consider a a family of open intervals $\{J_k\}_{k\in\mathbb{N}}$ of $\mathbb{R}$ covering $B = A+c$, then the same reasoning holds for $B - c = A$ and we show that $\mu^\star(A) \leq \mu^\star(A+c)$. Therefore, $\mu^\star(A) = \mu^\star(A+c)$ and $\mu^\star$ is invariant by translation.

8. Let consider a sequence $(A_k)_{k\in\mathbb{N}}$ of subsets of $\mathbb{R}$.

   (a) If the sum of their Lebesgue outer measure diverges, then the proof is trivial by definition of $\mu^\star$ and we have

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

   (b) We now assume that this sum does not diverge and has a finite limit. We set $\varepsilon > 0$. For each $k \in \mathbb{N}$, we pick a family of open intervals $\{I_{k,i}\}_{i\in\mathbb{N}}$ covering A_k, i.e.

$$A_k \subset \bigcup_{i=1}^{+\infty} I_{k,i},$$

and such that
$$\mu^\star(A_k) \leq \sum_{i=1}^{+\infty} l(I_{k,i}) \leq \mu^\star(A_k) + \frac{\varepsilon}{2^k}.$$
Note that such a cover of A_k exists by definition of $\mu^\star$ as an infimum. Then, the collection of intervals formed by the reunion of all these families is a countable (bijection from $\mathbb{N} \times \mathbb{N}$) set of open intervals covering the union of the sets A_k, i.e.
$$\bigcup_{k=1}^{+\infty} A_k \subset \bigcup_{k=1}^{+\infty} \left(\bigcup_{i=1}^{+\infty} I_{k,i} \right),$$
and, by monotonicity of $\mu^\star$, we can write
$$\mu^\star \left(\bigcup_{k=1}^{+\infty} A_k \right) \leq \mu^\star \left(\bigcup_{k=1}^{+\infty} \left(\bigcup_{i=1}^{+\infty} I_{k,i} \right) \right).$$
The next step is to calculate the right-hand side, the measure of a double union of intervals. The union of a union of intervals is, in the end, simply a union of intervals and since the family formed by the $\{I_{k,i}\}_{(i,k)\in\mathbb{N}^2}$ is countable, we can label them differently as $\{J_j\}_{j\in\mathbb{N}}$ where $\forall j \in \mathbb{N}, \exists (k,i) \in \mathbb{N}^2, J_j = I_{k,i}$, and
$$\mu^\star \left(\bigcup_{k=1}^{+\infty} \left(\bigcup_{i=1}^{+\infty} I_{k,i} \right) \right) = \mu^\star \left(\bigcup_{j=1}^{+\infty} J_j \right) \leq \sum_{j=1}^{+\infty} l(J_j) = \sum_{k=1}^{+\infty} \sum_{i=1}^{+\infty} l(I_{k,i}).$$
We now use the definition of the families of open intervals $\{I_{k,i}\}_{i\in\mathbb{N}}$ and
$$\sum_{k=1}^{+\infty} \sum_{i=1}^{+\infty} l(I_{k,i}) \leq \sum_{k=1}^{+\infty} \left(\mu^\star(A_k) + \frac{\varepsilon}{2^k} \right) = \sum_{k=1}^{+\infty} \mu^\star(A_k) + \varepsilon.$$
This being true for all $\varepsilon > 0$ we conclude that $\mu^\star$ is countably subadditive
$$\mu^\star \left(\bigcup_{k=1}^{+\infty} A_k \right) \leq \sum_{k=1}^{+\infty} \mu^\star(A_k).$$
□

REMARK 21. Use of infimum.

1. The infimum is a *lower bound.*
2. The infimum is the *greatest lower bound.*

REMARK 22. Why is this set application called an *outer* measure? We will see, in the rest of this chapter, that we have to define a particular set of ensembles that are measurables, and that these sets form a σ-algebra. A σ-ring is a particular case of what is called *hereditary rings* (we will not go into the details in this class). A *measure* is, formally, an extended real valued, non-negative, and countably additive set function defined on a ring. An *outer measure* is, this time, an extended real valued, non-negative, and countably subadditive set function defined on a hereditary σ-ring.

EXERCISE 18. ★ Similar sets.

1. Let $A \subset \mathbb{R}$, and $a \in \mathbb{R}$. We define $B = A \cup \{a\}$. Show that $\mu^\star(A) = \mu^\star(B)$.
2. Let A and B be two subsets of $\mathbb{R}$. Show that, if $\mu^\star(A) = 0$ then $\mu^\star(A \cup B) = \mu^\star(B)$.

Exercise 19. ★ Measures of a few sets.
Calculate the measure of the following sets:

1. $A_1 =]a; b[\cup]c; d[$ where $(a, b, c, d) \in \mathbb{R}^4$ and $a < b$ and $c < d$.
2. $A_2 = \{1/k \,|\, k \in \mathbb{N}^\star\}$.
3. $A_3 = \mathbb{Q}$.
4. $A_4 = \mathbb{R} \setminus \mathbb{Q}$.

Exercise 20. ★★ A *sur*mesure?

1. We define $\mu_i^\star$ such that $\forall A \subset \mathbb{R}, \mu_i^\star(A) = \inf\{\mu^\star(B) \,|\, A \subset B \subset \mathbb{R}\}$. Is $\mu_i^\star$ a measure? And, if so, is it a new measure or do we have $\forall A \subset \mathbb{R}, \mu_i^\star(A) = \mu^\star(A)$?
2. We define $\mu_s^\star$ such that $\forall A \subset \mathbb{R}, \mu_s^\star(A) = \sup\{\mu^\star(B) \,|\, A \subset B \subset \mathbb{R}\}$. Is $\mu_s^\star$ a measure? And, if so, is it a new measure or do we have $\forall A \subset \mathbb{R}, \mu_s^\star(A) = \mu^\star(A)$?

2.2.3 Concluding Note

In this section, we started by defining the length l of an interval and we showed that it naturally extends to the Lebesgue outer measure for any subset of $\mathbb{R}$. It would be nice, however, to have a property stronger than the countable subadditivity, namely a countable additivity i.e. we would like property 8 to be an equality instead of a sub-equality.

2.3 Lebesgue-measurable σ-Algebra

2.3.1 Carathéodory's Measurability Criteria

Definition 48. Carathéodory's measurability criteria.
Let E be a set of real numbers. E is said to be Lebesgue-measurable if, for every set X of real numbers,

$$\mu^\star(X) = \mu^\star(X \cap E) + \mu^\star(X \setminus E).$$

Exercise 21. ★ Elementary Lebesgue-measurable sets.

1. Show that $\emptyset$ and $\mathbb{R}$ are Lebesgue-measurable.
2. Show that, if E is a Lebesgue-measurable set of real numbers, so is $\mathbb{R} \setminus E$.

2.3.2 Carathéodory's σ-Algebra

Lemma 2. Stability of Lebesgue-measurable sets.
Let $\mathcal{S}$ be a collection of sets that satisfy Carathéodory's measurability criteria. $\mathcal{S}$ is stable with respect to the basic set operations:

1. $\forall A \in \mathcal{S}, A^c \in \mathcal{S}$
2. $\forall (A, B) \in \mathcal{S}^2, A \cap B \in \mathcal{S}$
3. $\forall (A, B) \in \mathcal{S}^2, A \cup B \in \mathcal{S}$
4. $\forall (A, B) \in \mathcal{S}^2, A \setminus B \in \mathcal{S}$

Proof.

Let $\mathcal{S}$ be a collection of sets that satisfy Carathéodory's measurability criteria, and $(A, B) \in \mathcal{S}^2$ two Lebesgue-measurable sets.

1. We consider $A^c = \mathbb{R} \setminus A$. We note that

$$\begin{aligned} \forall X \subset \mathbb{R}, X \cap A &= X \setminus (\mathbb{R} \setminus A)\,, \\ \forall X \subset \mathbb{R}, X \setminus A &= X \cap (\mathbb{R} \setminus A)\,, \end{aligned}$$

hence we prove that $\mathbb{R} \setminus A \in \mathcal{S}$ since it satisfies Carathéodory's criteria

$$\forall X \subset \mathbb{R}, \mu^\star (X) = \mu^\star (X \cap A) + \mu^\star (X \setminus A) = \mu^\star (X \setminus (\mathbb{R} \setminus A)) + \mu^\star (X \cap (\mathbb{R} \setminus A))\,.$$

2. Let $X \subset \mathbb{R}$. We write

$$X = (X \cap (A \cap B)) \cup (X \setminus (A \cap B))\,.$$

Since $\mu^\star$ is subadditive, we have

$$\mu^\star (X) \leq \mu^\star (X \cap (A \cap B)) + \mu^\star (X \setminus (A \cap B))\,.$$

Yet, we can also write

$$X \setminus (A \cap B) = (X \setminus A) \cup ((X \cap A) \setminus B)\,.$$

Again, using the subadditivity of $\mu^\star$ we have

$$\mu^\star (X \setminus (A \cap B)) \leq \mu^\star (X \setminus A) + \mu^\star ((X \cap A) \setminus B)\,,$$

leading to, with Carathéodory's criteria, since B and A are Lebesgue-measurable

$$\begin{aligned} \mu^\star (X \setminus (A \cap B)) &\leq \mu^\star (X \setminus A) + \mu^\star (X \cap A) - \mu^\star ((X \cap A) \setminus B)\,, \\ &\leq \mu^\star(X) - \mu^\star ((X \cap A) \setminus B)\,. \end{aligned}$$

Therefore, we deduce that

$$\mu^\star (X) \leq \mu^\star (X \cap (A \cap B)) + \mu^\star (X \setminus (A \cap B)) \leq \mu^\star(X),$$

so

$$\mu^\star (X) = \mu^\star (X \cap (A \cap B)) + \mu^\star (X \setminus (A \cap B))\,,$$

and $A \cap B$ satisfies Carathéodory's criteria, so $A \cap B \in \mathcal{S}$ and $\mathcal{S}$ is stable by intersection.

3. Showing that $A \cup B \in \mathcal{S}$ is equivalent to show that $(A \cup B)^c \in \mathcal{S}$. We can write $(A \cup B)^c = A^c \cap B^c$. Since $\mathcal{S}$ is stable by complement, $A^c \in \mathcal{S}$ and $B^c \in \mathcal{S}$, and since $\mathcal{S}$ is stable by intersection, $A^c \cap B^c \in \mathcal{S}$. Hence, $(A \cup B)^c \in \mathcal{S}$ and again, since $\mathcal{S}$ is stable by complement, $A \cup B \in \mathcal{S}$. Thus $\mathcal{S}$ is stable by union.

4. We note that $A \setminus B = A \cap B^c$. Since $\mathcal{S}$ is stable by complement, $B^c \in \mathcal{S}$, and since $\mathcal{S}$ is stable by intersection, $A \cap B^c \in \mathcal{S}$. Hence, $A \setminus B \in \mathcal{S}$ and $\mathcal{S}$ is stable by difference. $\square$

Theorem 8. Carathéodory's σ-algebra.
The collection of sets that satisfy Carathéodory's measurability criteria forms a σ-algebra.

Proof.

Let $\mathcal{S}$ be a collection of sets that satisfy Carathéodory's measurability criteria, and $\{A_k\}_{k\in\mathbb{N}}$ a family of mutually disjoint sets of $\mathcal{S}$. In order to show that $\mathcal{S}$ is a σ-algebra, we have to show that $\mathcal{S}$ is a σ-ring stable by complement. We will procede step by step, using the results from the previous lemma.

1. We first show that $\mathcal{S}$ is a ring of sets.
   (a) We verify that $\mathcal{S}$ is not empty. This is true since $\emptyset \in \mathcal{S}$.
   (b) We have shown that $\mathcal{S}$ is stable by union: $\forall (A,B) \in \mathcal{S}^2, A \cup B \in \mathcal{S}$.
   (c) We have shown that $\mathcal{S}$ is stable by difference: $\forall (A,B) \in \mathcal{S}^2, A \setminus B \in \mathcal{S}$.

   Hence, $\mathcal{S}$ is a ring of sets.

2. We now show that $\mathcal{S}$ is a σ-ring of sets, i.e. that $\mathcal{S}$ is stable by countable union.

   In order to do so, we will first show a preliminary result:

$$\forall n \in \mathbb{N}, \forall \{B_k\}_{k\leq n} \in \mathcal{S}^{n+1}, \forall X \subset \mathbb{R}, \mu^\star\left(X \cap \left(\bigcup_{k=0}^{n} B_k\right)\right) = \sum_{k=0}^{n} \mu^\star\left(X \cap B_k\right).$$

   This can be shown be induction.

   Now, we can write

$$\begin{aligned}
\mu^\star(X) &= \mu^\star\left(X \cap \left(\bigcup_{k=0}^{n} A_k\right)\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{n} A_k\right)\right) && \text{(by definition)}\\
&\geq \mu^\star\left(X \cap \left(\bigcup_{k=0}^{n} A_k\right)\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) && \text{(by monotonicity of } \mu^\star)\\
&\geq \sum_{k=0}^{n} \mu^\star\left(X \cap A_k\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) && \text{(inductive result)}
\end{aligned}$$

   The previous result is valid for all $n \in \mathbb{N}$, so we can take the limit $n \to +\infty$ and write

$$\begin{aligned}
\mu^\star(X) &\geq \sum_{k=0}^{+\infty} \mu^\star\left(X \cap A_k\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right)\\
&\geq \mu^\star\left(X \cap \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) && \text{(by subadditivity)}
\end{aligned}$$

   If we write, however, that

$$X = \left(X \cap \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) \cup \left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right),$$

   then using the subadditivity of $\mu^\star$ we obtain the reverse inequality

$$\mu^\star(X) \leq \mu^\star\left(X \cap \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right).$$

   We therefore proved that any countable union of sets satifying Carathéordory's criteria also satifies this criteria, so $\mathcal{S}$ is stable by countable union

$$\mu^\star(X) = \mu^\star\left(X \cap \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) + \mu^\star\left(X \setminus \left(\bigcup_{k=0}^{+\infty} A_k\right)\right).$$

3. Finally, we have shown that the σ-ring of sets $\mathcal{S}$ is stable by complement, therefore it is a σ-alebra of sets.

□

Theorem 9. Countable additivity of $\mu^\star$.
If $\mathcal{S}$ is a collection of sets that satisfy Carathéodory's measurability criteria, then the Lebesgue outer measure $\mu^\star$ is countably additive on $\mathcal{S}$.

Proof.
Let $\mathcal{S}$ be a collection of sets that satisfy Carathéodory's measurability criteria, and $\{A_k\}_{k\in\mathbb{N}}$ a family of mutually disjoint sets of $\mathcal{S}$. In the previous proof, we showed that

$$\forall X \subset \mathbb{R}, \forall n \in \mathbb{N}, \mu^\star\left(X \cap \left(\bigcup_{k=0}^{n} A_k\right)\right) = \sum_{k=0}^{n} \mu^\star\left(X \cap A_k\right).$$

The choice $X = \mathbb{R}$ yields

$$\forall n \in \mathbb{N}, \mu^\star\left(\bigcup_{k=0}^{n} A_k\right) = \sum_{k=0}^{n} \mu^\star\left(A_k\right),$$

therefore finite additivity is satisfied.

We now write, for $n \in \mathbb{N}$,

$$\begin{aligned}
\sum_{k=0}^{n} \mu^\star\left(A_k\right) &= \mu^\star\left(\bigcup_{k=0}^{n} A_k\right) \\
&\leq \mu^\star\left(\bigcup_{k=0}^{+\infty} A_k\right) \quad \text{(using monotonicity)} \\
&\leq \sum_{k=0}^{+\infty} \mu^\star\left(A_k\right) \quad \text{(since } \mu^\star \text{ is countably subadditive).}
\end{aligned}$$

This result is valid for all $n \in \mathbb{N}$, therefore we can take the limit $n \to +\infty$ and we conclude that $\mu^\star$ is countably additive

$$\mu^\star\left(\bigcup_{k=0}^{+\infty} A_k\right) = \sum_{k=0}^{+\infty} \mu^\star\left(A_k\right).$$

□

Proposition 12. Lebesgue-measurability of intervals of $\mathbb{R}$.
Intervals of $\mathbb{R}$ satisfy Carathéodory's measurability criteria.

Proof.
Let $a \in \mathbb{R}$. We will show that $]a; +\infty[$ satisfies Carathéodory's criteria, i.e.

$$\forall X \subset \mathbb{R}, \mu^\star(X) = \mu^\star\left(X \cap]a; +\infty[\right) + \mu^\star\left(X \setminus]a; +\infty[\right).$$

First of all, using the subadditivity of $\mu^\star$, we already know that

$$\forall X \subset \mathbb{R}, \mu^\star(X) \leq \mu^\star\left(X \cap]a; +\infty[\right) + \mu^\star\left(X \setminus]a; +\infty[\right).$$

Let $X \subset \mathbb{R}$ and $\varepsilon > 0$. From the definition of the Lebesgue measure, we can find a cover of open intervals $\{I_k\}_{k\in\mathbb{N}}$ such that

$$\mu^\star(X) \leq \sum_{k=1}^{+\infty} l(I_k) < \mu^\star(X) + \varepsilon.$$

For $k \in \mathbb{N}$, if we now take a look at $I_k \cap]a; +\infty[$, this is either empty or an open interval, and $X \cap]a; +\infty[$ is a subset of the union over $k \in \mathbb{N}$ of the $I_k \cap]a; +\infty[$. Similarly, for $k \in \mathbb{N}$, $I_k \backslash]a; +\infty[$ is either empty or an interval (open or closed), and $X \backslash]a; +\infty[$ is a subset of the union over $k \in \mathbb{N}$ of the $I_k \backslash]a; +\infty[$. Hence, we write

$$\begin{aligned}
\mu^\star \left(X \cap]a; +\infty[\right) + \mu^\star \left(X \backslash]a; +\infty[\right) &\leq \mu^\star \left(\bigcup_{k=1}^{+\infty} \left(I_k \cap]a; +\infty[\right) \right) + \mu^\star \left(\bigcup_{k=1}^{+\infty} \left(I_k \backslash]a; +\infty[\right) \right) \\
&\leq \sum_{k=1}^{+\infty} l \left(I_k \cap]a; +\infty[\right) + \sum_{k=1}^{+\infty} l \left(I_k \backslash]a; +\infty[\right) \\
&\leq \sum_{k=1}^{+\infty} \left\{ l \left(I_k \cap]a; +\infty[\right) + l \left(I_k \backslash]a; +\infty[\right) \right\} \\
&\leq \sum_{k=1}^{+\infty} l(I_k) \\
&< \mu^\star(X) + \varepsilon,
\end{aligned}$$

in which we use the additivity of the length application over mutually disjoint intervals[2]. Since this is true for any $\varepsilon > 0$, we deduce

$$\forall X \subset \mathbb{R}, \mu^\star(X) \geq \mu^\star \left(X \cap]a; +\infty[\right) + \mu^\star \left(X \backslash]a; +\infty[\right),$$

and finally

$$\forall X \subset \mathbb{R}, \mu^\star(X) = \mu^\star \left(X \cap]a; +\infty[\right) + \mu^\star \left(X \backslash]a; +\infty[\right).$$

A similar reasoning shows that, for $b \in \mathbb{R}$, intervals of the form $]-\infty; b[$ satisfy Carathéodory's criteria. Using the stability of Carathéodory's σ-algebra regarding the elementary set operators, we show that intervals of the form $]a; b[=]-\infty; b[\cap]a; +\infty[$ also satisfy Carathéodory's criteria. □

2.3.3 Subtractivity and Continuity of the Lebesgue Measure

Theorem 10. Subtractivity. Let A and B be two subsets in $\mathcal{S}$ with $B \subset A$. Then, if $\mu^\star(A) < +\infty$, we have

$$\mu^\star(A \setminus B) = \mu^\star(A) - \mu^\star(B).$$

Exercise 22. ★ Subtractivity.
Prove the subtractivity theorem.

Exercise 23. ★ Measure of a subsets and elementary operations.
Let A and B be two subsets of $\Omega \subset \mathbb{R}$. What can you say about $\mu^\star(A \cup B)$, $\mu^\star(A \cap B)$, $\mu^\star(A \setminus B)$, $\mu^\star(A \Delta B)$, $\mu^\star(A^c)$, and $\mu^\star(A^c \cup B)$?

Definition 49. Increasing sequence of sets.
A sequence of sets $\{A_k\}_{k \in \mathbb{N}}$ is an *increasing sequence of sets* if $\forall k \in \mathbb{N}, A_k \subset A_{k+1}$.

Theorem 11. Continuity from below.
If $\{A_k\}_{k \in \mathbb{N}}$ is an increasing sequence of sets in $\mathcal{S}$ of which the limit A is in $\mathcal{S}$, then

$$\mu^\star(A) = \mu^\star \left(\lim_{k \to +\infty} A_k \right) = \lim_{k \to +\infty} \mu^\star(A_k).$$

[2] See result from exercise 17, question 3.

Proof.

Let$\{A_k\}_{k\in\mathbb{N}}$ be an increasing sequence of sets in $\mathcal{S}$ of which the limit A is in $\mathcal{S}$. Without changing our problem, we can assume that $A_0 = \emptyset$. Then we will use the countable additivity of Lebesgue outer measure on disjoint sets

$$\begin{aligned}
\mu^\star(A) = \mu^\star\left(\lim_{k\to+\infty} A_k\right) &= \mu\left(\bigcup_{i=1}^{+\infty} A_i\right) \\
&= \mu\left(\bigcup_{i=1}^{+\infty}(A_i \setminus A_{i-1})\right) \\
&= \sum_{i=1}^{+\infty} \mu^\star\left(A_i \setminus A_{i-1}\right) \quad \text{(additivity)} \\
&= \lim_{k\to+\infty} \sum_{i=1}^{k} \mu^\star\left(A_i \setminus A_{i-1}\right) \\
&= \lim_{k\to+\infty} \mu^\star\left(\bigcup_{i=1}^{k}(A_i \setminus A_{i-1})\right) \quad \text{(additivity)} \\
\mu^\star(A) = \mu^\star\left(\lim_{k\to+\infty} A_k\right) &= \lim_{k\to+\infty} \mu^\star(A_k)
\end{aligned}$$

□

Definition 50. Decreasing sequence of sets.
A sequence of sets $\{A_k\}_{k\in\mathbb{N}}$ is a *decreasing sequence of sets* if $\forall k \in \mathbb{N}, A_{k+1} \subset A_k$.

Theorem 12. Continuity from above.
If $\{A_k\}_{k\in\mathbb{N}}$ is a decreasing sequence of sets in $\mathcal{S}$ of which the limit A is in $\mathcal{S}$ and of which at least one of the sets has a finite measure, then

$$\mu^\star(A) = \mu^\star\left(\lim_{k\to+\infty} A_k\right) = \lim_{k\to+\infty} \mu^\star(A_k).$$

Proof.

Let $\{A_k\}_{k\in\mathbb{N}}$ be a decreasing sequence of sets in $\mathcal{S}$ of which the limit A is in $\mathcal{S}$ and of which at least one of the sets has a finite measure. Let note A_q be one of these terms of finite measure. Then, $\{A_q \setminus A_k\}_{k\geq q}$ is an increasing sequence. Using the previous theorem and the subtractivity of $\mu^\star$, we write

$$\begin{aligned}
\mu^\star\left(A_q\right) - \mu^\star\left(\lim_{k\to+\infty} A_k\right) &= \mu^\star\left(A_q \setminus \lim_{k\to+\infty} A_k\right) \\
&= \mu^\star\left(\lim_{k\to+\infty}\left(A_q \setminus A_k\right)\right) \\
&= \lim_{k\to+\infty} \mu^\star\left(A_q \setminus A_k\right) \\
&= \lim_{k\to+\infty}\left(\mu^\star\left(A_q\right) - \mu^\star\left(A_k\right)\right) \\
&= \mu^\star\left(A_q\right) - \lim_{k\to+\infty} \mu^\star\left(A_k\right),
\end{aligned}$$

from which we conclude, since $\mu^\star\left(A_q\right)$ is finite, that

$$\mu^\star(A) = \mu^\star\left(\lim_{k\to+\infty} A_k\right) = \lim_{k\to+\infty} \mu^\star(A_k).$$

□

REMARK 23. Continuity.

The outer Lebesgue measure is continuous from above and continuous from below for sets in Carathéodory's σ-algebra. We will simply say that the outer Lebesgue measure is continuous.

THEOREM 13. Continuity implies countable additivity.
Let λ be a finite, non-negative, and additive set function on $\mathcal{S}$. If λ is continuous (either from above or from below), then λ is *countably additive*.

PROOF.

We already know that λ is additive, i.e. for two disjoint sets A and B we have $\lambda(A \cup B) = \lambda(A) + \lambda(B)$. By induction, we can easily prove that λ is finitely additive. We now proceed with the countable additivity, and we consider a countable disjoint sequence of sets $\{A_k\}_{k\in\mathbb{N}}$ of $\mathbb{R}$. We note

$$A = \bigcup_{k=1}^{+\infty} A_k.$$

1. Case 1: If λ is continuous from above, then we define

$$\forall k \in \mathbb{N}, B_k = \bigcup_{i=1}^{k} A_i,$$

which is an increasing sequence of sets in $\mathbb{R}$, with limit A. Then, we can write

$$\begin{aligned}
\lambda(A) = \lambda\left(\bigcup_{k=1}^{+\infty} A_k\right) &= \lambda\left(\lim_{k\to+\infty} B_k\right) \\
&= \lim_{k\to+\infty} \lambda(B_k) \quad \text{(continuity)} \\
&= \lim_{k\to+\infty} \sum_{i=1}^{k} \lambda(A_i) \quad \text{(finite additivity)} \\
&= \sum_{k=1}^{+\infty} \lambda(A_k).
\end{aligned}$$

2. Case 2: If λ is continuous from below, then we define

$$\forall k \in \mathbb{N}, C_k = A \setminus \bigcup_{i=1}^{k} A_i,$$

which is a decreasing sequence of sets in $\mathbb{R}$, with limit $\emptyset$. Then, the subtractivity of λ (see previous theorems, λ being additive and finite) allows to write

$$\forall k \in \mathbb{N}, \lambda(A) = \lambda(C_k) + \lambda\left(\bigcup_{i=1}^{k} A_i\right).$$

Hence, we have

$$\begin{aligned}
\lambda(A) = \lambda\left(\bigcup_{k=1}^{+\infty} A_k\right) &= \lambda(C_k) + \lambda\left(\bigcup_{i=1}^{k} A_i\right) \\
&= \lim_{k\to+\infty}\left(\lambda(C_k) + \lambda\left(\bigcup_{i=1}^{k} A_i\right)\right) \\
&= \lim_{k\to+\infty}\left(\lambda(C_k) + \sum_{i=1}^{k}\lambda(A_i)\right) \quad \text{(finite additivity)} \\
&= \lim_{k\to+\infty}\lambda(C_k) + \lim_{k\to+\infty}\sum_{i=1}^{k}\lambda(A_i) \\
&= \lambda\left(\lim_{k\to+\infty} C_k\right) + \sum_{i=1}^{+\infty}\lambda(A_i) \quad \text{(continuity)} \\
&= \sum_{k=1}^{+\infty}\lambda(A_k).
\end{aligned}$$

□

2.4 Extension of the Lebesgue Measure to $\mathbb{R}^n$

Definition 51. Cuboid of $\mathbb{R}^n$.
A *cuboid* $\mathcal{P}$ of $\mathbb{R}^n$ is defined by a sequence of intervals $\{I_k\}_{1\leq k\leq n}$ such that

$$\mathcal{P} = I_1 \times I_2 \times \ldots \times I_n = \prod_{k=1}^{n} I_k.$$

Remark 24.
In $\mathbb{R}^2$, this corresponds to the intuitive definition of a rectangular domain, and in $\mathbb{R}^3$, of a rectangular parallelepiped. In $\mathbb{R}$, using this definition, we recover the definition of an interval.

Definition 52. Volume of a cuboid in $\mathbb{R}^n$.
Let $\mathcal{P}$ be a cuboid of $\mathbb{R}^n$ generated by the sequence of intervals $\{I_k\}_{1\leq k\leq n}$. Then, the volume of $\mathcal{P}$, noted $v(\mathcal{P})$, is defined by the product of all the lengths of the intervals of the sequence that generates $\mathcal{P}$, i.e.

$$v(\mathcal{P}) = \prod_{k=1}^{n} l(I_k).$$

Example 7.

1. $\mathcal{P}_1 =]0;\, 1[$ in $\mathbb{R}$: $v(\mathcal{P}_1) = l\left(]0;\, 1[\right) = 1$.
2. $\mathcal{P}_2 =]0;\, 1[\times[0;\, 2]$ in $\mathbb{R}^2$: $v(\mathcal{P}_2) = 1 \times 2 = 2$.
3. $\mathcal{P}_3 =]0;\, 1[\times[0;\, 2]\times[1;\, 2[$ in $\mathbb{R}^3$: $v(\mathcal{P}_3) = 1 \times 2 \times 1 = 2$.

Exercise 24. ★ Volumes of cuboids.
Calculate the volumes of the following cuboids:

1. $\mathcal{P}_1 =]0;\, 1[\times[1;\, 3[\times[-\pi;\, \pi]$ in $\mathbb{R}^3$.

2. $\mathcal{P}_2 =]1;\, 2[^n$ in $\mathbb{R}^n$.

3. $\mathcal{P}_3 = I_1 \times I_2 \times \dots \times I_n$ in $\mathbb{R}^n$ with $\forall k \in [\![0\,;n]\!], I_k =]-k;\, k[$.

Proposition 13. Volume of a compact of $\mathbb{R}^n$.
Let K be a compact of $\mathbb{R}^n$. Then, the *volume* of K is given by

$$v(K) = \inf \left\{ \sum_{k=1}^{+\infty} v(\mathcal{P}_k) \,\middle|\, K \subset \bigcup_{k=1}^{+\infty} \mathcal{P}_k \text{ and } \forall k \in \mathbb{N}, \mathcal{P}_k \text{ is an open cuboid} \right\}.$$

Definition 53. Lebesgue Outer Measure on $\mathbb{R}^n$.
Let A be a subset of $\mathbb{R}^n$. The *Lebesgue outer measure* $\mu^\star$ of A is given by

$$\mu^\star(A) = \inf \left\{ \sum_{k=1}^{+\infty} v(\mathcal{P}_k) \,\middle|\, A \subset \bigcup_{k=1}^{+\infty} \mathcal{P}_k \text{ and } \forall k \in \mathbb{N}, \mathcal{P}_k \text{ is an open cuboid} \right\}.$$

Definition 54. Lebesgue Inner Measure on $\mathbb{R}^n$.
Let A be a subset of $\mathbb{R}^n$. The *Lebesgue inner measure* $\mu_\star$ of A is given by

$$\mu_\star(A) = \sup \left\{ \sum_{k=1}^{+\infty} v(K_k) \,\middle|\, \bigcup_{k=1}^{+\infty} K_k \subset A \text{ and } \forall k \in \mathbb{N}, K_k \text{ is a compact} \right\}.$$

Theorem 14. Lebesgue outer and inner measures on $\mathbb{R}^n$.
Let A be a subset of $\mathbb{R}^n$. If A satisfies Carathéodory's measurability criteria (extended to $\mathbb{R}^n$), then $\mu^\star(A) = \mu_\star(A)$ (i.e. same inner and outer Lebesgue measures).

Exercise 25. ★★★ Properties of the Lebesgue outer measure on $\mathbb{R}^n$.
Show that the Lebesgue outer measure we have defined on $\mathbb{R}^n$ and the Lebesgue outer measure we have defined on $\mathbb{R}$ satisfy the same properties (see theorem 7), i.e. positivity, monotonicity, etc.

Exercise 26. ★

1. Prove that if G is a bounded open set, then $\mu^\star(G) < +\infty$.

2. Prove that $\mu^\star(\mathbb{R}^n) = +\infty$

Exercise 27. ★ Measures of sets in $\mathbb{R}^2$.
Compute the measure of the following sets:

$$\begin{aligned}
A_1 &= \left\{(x, y) \,\middle|\, 1 \leq x \text{ and } 0 < y < \frac{1}{x}\right\} \\
A_2 &= \left\{(x, y) \,|\, 0 \leq x \text{ and } 0 < y < e^{-x}\right\} \\
A_3 &= \left\{(x, y) \,\middle|\, 1 < x \text{ and } 0 < y < e^{-a}\right\} \text{ with } a \in [1;\, +\infty[
\end{aligned}$$

Exercise 28. ★★ Frontier of a set.
Let A be a subset of $\mathbb{R}^n$. We note ∂A the frontier of A, defined by $\partial A = \overline{A} \setminus \mathring{A}$.

1. Is ∂A a closed set? Why?

2. Is ∂A an open set? Why?

3. If A is bounded, is ∂A a compact?

4. Assuming that $\overline{A}$ and $\mathring{A}$ are Lebesgue-measurable, show that ∂A is also Lebesgue-measurable.

5. What is the Lebesgue-measure of the frontier, $\mu(\partial A)$?

PROPOSITION 14. Lebesgue-measurability of cuboids of $\mathbb{R}^n$.
Cuboids of $\mathbb{R}^n$ satisfy Carathéodory's measurability criteria.

EXERCISE 29. ★ Carathéodory's measurability criteria for cuboids.
Prove that cuboids of $\mathbb{R}^n$ satisfy Carathéodory's measurability criteria.

EXERCISE 30. ★★ Intersection and countable additivity.
Let A be a measurable set, and $\{B_k\}_{k\in\mathbb{N}}$ a sequence of mutually disjoint measurable sets. We note

$$\forall N \in \mathbb{N}, \tilde{B}_N = \bigcup_{k=1}^{N} B_k \qquad \text{and} \qquad B = \bigcup_{k=1}^{+\infty} B_k.$$

1. Show that for all $N \in \mathbb{N}$, $\tilde{B}_N$ is Lebesgue-measurable.

2. Is B Lebesgue-measurable, and why?

3. Show that
$$\forall N \in \mathbb{N}, \mu^\star\left(A \cap \tilde{B}_N\right) = \sum_{k=1}^{N} \mu^\star\left(A \cap B_k\right).$$

4. Show that
$$\mu^\star\left(A \cap B\right) = \sum_{k=1}^{+\infty} \mu^\star\left(A \cap B_k\right).$$

2.5 Summary: Measuring in $\mathbb{R}^n$?

In this chapter, we have built the *Lebesgue measure* on $\mathbb{R}^n$: starting from the length of an interval and the volume of a cuboid, we defined the Lebesgue outer measure of a set A as an infimum on the lengths and volumes of open sets covering A. We then restricted the collection of measurable thanks to Carathéodory's measurability criteria, and we obtained a σ-algebra of measurable sets. We therefore did two important steps: first, we defined the Lebesgue measure; second, we defined (theoretically) a set of measurable sets stable with respect to the elementary set operations.

There is still a remaining question: what can we measure? Or, in other words: is everything measurable in $\mathbb{R}^n$? We have shown that open intervals of $\mathbb{R}$ and open cuboids of $\mathbb{R}^n$ are measurable, but these cannot describe all the possible sets of real numbers. Using the stability of Carathéodory's σ-algebra, we can show that unions of intervals, differences, etc, are measurable too, but this implies that we should characterise the measurability of many kinds of sets! Can we provide a more accurate description of the measurable sets of $\mathbb{R}^n$? This will be the aim of the next chapter.

CHAPTER 3

Measurable Sets and Functions

In chapter 2 we defined, through an inductive construction, the Lebesgue measure on $\mathbb{R}^n$, and we have shown that the set of Lebesgue measurable subsets of $\mathbb{R}^n$ is a σ-algebra of sets, characterised by Carathéodory's criteria, and therefore stable with respect to the elementary set operations. We may now ask the following question: what is contained in this σ-algebra of Lebesgue measurable sets? We will proceed as follows: first, we will show that there exist non-measurable sets of real numbers, which justifies the need of such a σ-algebra; then, we will define the relevant sets for the Lebesgue integration theory on $\mathbb{R}^n$, which are called *Borelians*. Later on, we will discuss the measurability of functions defined on such sets.

Throughout this chapter, we will note $\mathcal{S}$ Carathéodory's σ-algebra of measurable sets.

3.1 Is Everything Measurable?

3.1.1 Existence of Non-Measurable Sets

THEOREM 15. Existence of non-measurable sets.
There exists a set $E \subset \mathbb{R}^n$ such that E is not measurable.

PROOF.
We will construct a non-measurable subset of $\mathbb{R}^n$. In order to do so, we define, for all $\in \mathbb{R}^n$, the set $Q_x = \{a \in \mathbb{R}^n \,|\exists q \in \mathbb{Q}^n, a = x + q\}$. We can show (see next exercise) that for all $(x, y) \in (\mathbb{R}^n)^2$, we have either $Q_x = Q_y$ or $Q_x \cap Q_y = \emptyset$, i.e. every vector of real numbers belongs to one and only one set Q_x. Note that the such a set Q_x is not unique, since it is based on the choice of a representant $x \in \mathbb{R}^n$. Since every vector of real numbers is contained in one, and only one, of these sets Q_x, $\mathbb{R}^n$ can be covered by the translates of $\mathbb{Q}^n$[1].

Hence, we choose a set $E \subset \mathbb{R}^n$ such that we can construct such a cover of $\mathbb{R}^n$ with disjoint sets, i.e.[2]

$$\mathbb{R}^n = \bigcup_{x \in E} Q_x = \bigcup_{x \in E} (x + \mathbb{Q}^n).$$

[1] It is also possible to define an equivalence relation $\sim$ on $\mathbb{R}^n$ with $x \sim y \Leftrightarrow -y \in \mathbb{Q}^n$, and to consider the quotient ensemble formed by $\mathbb{R}^n$ and this equivalence relation. Then, the subset E defined in the proof can be choose to be in $[0; 1]$ and the last part of the proof is more straightforward.

[2] The reasoning here is based on the *axiom of choice*: we can choose such elements $x \in \mathbb{R}^n$ to build this cover, although the choice is not unique.

We can also write this covering property as follows

$$\mathbb{R}^n = \bigcup_{q \in \mathbb{Q}^n} (q + E).$$

We assume that E is measurable. From this equality, we can write

$$\mu^\star(\mathbb{R}^n) = +\infty = \sum_{q \in \mathbb{Q}^n} \mu^\star(q + E) = \sum_{q \in \mathbb{Q}^n} \mu^\star(E),$$

since $\mathbb{Q}^n$ is countable (we then use the countable additivity and the invariance by translation of the Lebesgue measure). The only way to satisfy this equality is to have $\mu^\star(E) > 0$[3].

Now, we consider an arbitrary compact $K \subset E$ and a subset $D = \mathcal{B}(0,1) \cap \mathbb{Q}^n$. D is a bounded infinite countable set. The set

$$A = \bigcup_{q \in D} (q + K),$$

is then an infinite countable disjoint union, contained in $D + K$ by definition; since D and K are bounded, the union has a finite measure. Yet, using the countable additivity and the invariance by translation of μ, we have

$$\mu_\star(A) = \sum_{q \in D} \mu_\star(q + K) = \sum_{q \in D} \mu_\star(K),$$

and since the sum is infinite, but the Lebesgue measure of A is not, the only possibility is that $\mu_\star(K) = 0$. K being an arbitrary compact, we deduce that $\mu_\star(E) = 0$ [4].

These two results are contradictory, therefore E is not Lebesgue-measurable[5]. □

Exercise 31. ★ Translated of $\mathbb{Q}^n$.

Let $(x, y) \in (\mathbb{R}^n)^2$. We define Q_x (respectively Q_y) by $Q_x = \{a \in \mathbb{R}^n \,|\, \exists q \in \mathbb{Q}^n, a = x + q\}$ (respectively Q_y). Show that either

$$Q_x = Q_y \qquad \text{or} \qquad Q_x \cap Q_y = \emptyset.$$

Corollary 3. Existence of non-measurable subsets.
If $A \subset \mathbb{R}^n$ is measurable and of measure $\mu(A) > 0$, then there exists a subset $B \subset A$ such that B is not measurable.

Proof.

Using the notations of the previous proof, if $A \subset \mathbb{R}^n$ is measurable and of measure $\mu(A) > 0$, then we can write

$$A = \bigcup_{q \in \mathbb{Q}^n} (q + E) \cap A.$$

We know that $\mu(A) > 0$ so, with the countable additivity of μ, we need a least one of the sets of the union to have a non-zero outer measure. Let note $B = (q + E) \cap A$ this set, with $q \in \mathbb{Q}^n$, and we have $\mu^\star(B) > 0$. Yet, we also have

$$\mu_\star(B) = \mu_\star((q + E) \cap A) \leq \mu_\star(q + E) = \mu_\star(E) = 0,$$

which is, again, contradictory. □

[3] To be more precise, the *outer* Lebesgue measure is strictly positive

[4] To be more precise, the *inner* Lebesgue measure is zero

[5] To be exact, we found that the inner Lebesgue measure of E is zero, and that the outer Lebesgue measure of E is strictly positive. We have shown, however, that if a set is measurable, its inner and outer Lebesgue measures are the same. This cannot be the case here, hence E is not Lebesgue-measurable.

3.1.2 Cantor Set

In this subsection, we will construct a particular set called a *Cantor*. This is motivated by the following observation: what has zero Lebesgue measure? A singleton for example, or a countable union of singletons, have zero Lebesgue measure; this is "intuitive" since the interior of such sets is equal to the empty set, as if "there is nothing to measure". On the contrary, we have seen many measurable sets of non-zero Lebesgue measure, and they all have a non-empty interior: intervals, $\mathbb{R}^n$, etc. As in the introduction of chapter 2, we will see that we should be careful with this intuitive reasoning.

Let consider $(\alpha_k)_{k\in\mathbb{N}}$ a sequence of strictly positive real numbers, such that

$$\sum_{k=1}^{+\infty} \alpha_k 2^k \leq 1.$$

We build a decreasing sequence of compacts A_k all included in $[0;\,1]$ as follows: for all $k \in \mathbb{N}$, A_{k+1} is obtained by removing open intervals of length α_k in each of the closed intervals composing A_k, in order to split them in two halfs of equal length. This means:

1. $A_0 = [0;\,1]$
2. $A_1 = [0;\,(1-\alpha_0)/2] \cup [(1+\alpha_0)/2;\,1]$
3. and so on.

This process is illustrated in figure 3.1.

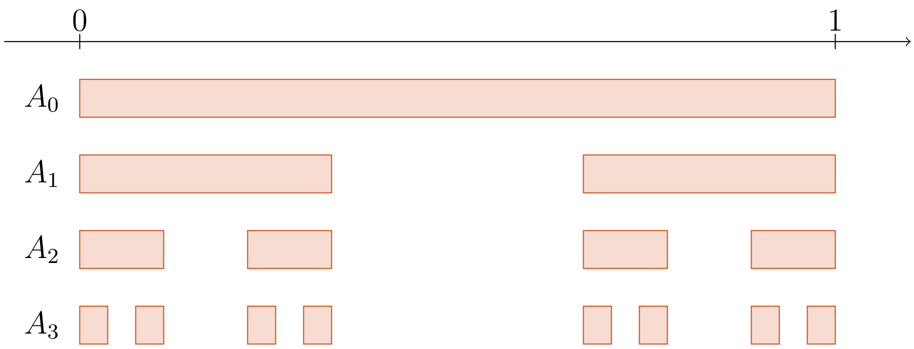


Figure 3.1: Construction of a Cantor set (here the ternary Cantor set).

Then, $\forall k \in \mathbb{N}, A_{k+1} \subset A_k$, A_k is constituted of 2^k disjoint closed intervals of the same length, and

$$\mu(A_k) = 1 - \sum_{p=0}^{k-1} 2^p \alpha_p > 0.$$

We can define K the intersection of all of these sets

$$K = \bigcap_{k=0}^{+\infty} A_k.$$

K is a non-empty compact, since it is the intersection of a decreasing sequence of non-empty compacts, and therefore $K = \overline{K}$ (Heine-Borel theorem, K is a compact of $\mathbb{R}$). Yet, $\mathring{K} = \emptyset$: $\forall k \in \mathbb{N}, K \subset A_k$ and A_k does not contain any interval whose length is larger than 2^{-k}. Hence, we deduce that $K = \partial K$ (its frontier). If we now compute the measure of K, we have

$$\mu(K) = \lim_{k\to+\infty} \mu(A_k) = 1 - \sum_{k=0}^{+\infty} \alpha_k 2^k \geq 0.$$

We now have two options:

1. if the sum of $\alpha_k 2^k$ is equal to 1, i.e.

$$\sum_{k=1}^{+\infty} \alpha_k 2^k = 1,$$

then $\mu(K) = 0$ and we call K a *"thin" Cantor set*;

2. if the sum of $\alpha_k 2^k$ is smaller than 1, i.e.

$$\sum_{k=1}^{+\infty} \alpha_k 2^k < 1,$$

then $\mu(K) > 0$ and we call K a *"fat" Cantor set*.

Remark 25.

1. The Cantor set is uncountable.
2. The Cantor set is equal to its frontier: $\mathring{K} = \emptyset$, $K = \overline{K} = \partial K$.
3. The Cantor set does not contain open intervals.

3.2 Borelians

3.2.1 Borelian Tribe

Definition 55. Tribe.
A *tribe* $\mathcal{T}$ of Ω is a family of subsets from Ω that satisfies the following properties:

1. $\mathcal{T}$ is stable by complement
$$\forall E \in \mathcal{T}, E^c \in \mathcal{T}.$$
2. $\mathcal{T}$ is stable by countable union
$$\forall \{E_i\}_{i\in I} \in \mathcal{T}, \bigcup_{i\in I} E_i \in \mathcal{T}.$$

Definition 56. Borelian tribe.
The *Borelian tribe* $\mathcal{B}$ of Ω is the smallest family of subsets from Ω that is a tribe and that contains all the open sets of Ω. We note $\mathcal{B}(\Omega)$ the Borelian tribe of Ω. Elements in $\mathcal{B}(\Omega)$ are called the *Borelians* of Ω.

Proposition 15. Borel σ-algebra.
Let Ω be a non empty set. Then, $\mathcal{B}(\Omega)$ is a σ-algebra of sets.

Proof.
First, we note that since $\mathcal{B}$ is stable by complement and by countable union, it is also stable by countable intersection (application of Morgan's Laws). Then, it is also stable by finite intersection.

We now proceed with the proof of the proposition, step by step:

1. $\mathcal{B}$ is a ring of sets:

(a) $\mathcal{B}$ is not empty, since it contains $\emptyset$ (it is a trivial open set of Ω).

(b) $\mathcal{B}$ is stable by finite union, since it is stable by countable union.

(c) $\mathcal{B}$ is stable by difference: $\forall (A, B) \in \mathcal{B}^2, A \setminus B = A \cap B^c$ and $\mathcal{B}$ is stable by intersection and complement.

2. $\mathcal{B}$ is a σ-ring of sets, since it is stable by countable union.

3. $\mathcal{B}$ is a σ-algebra of sets, since it is stable by complement. □

Example 8.
Let Ω be a non empty set.

1. $\mathcal{T}_1 = \{\emptyset, \Omega\}$ is the smallest tribe of Ω.

2. $\mathcal{T}_2 = \mathcal{P}(\Omega)$ (partition of Ω) is the largest tribe of Ω.

Exercise 32. ★ Tribes-inter-tribes.
Let $I \subset \mathbb{N}$ and $\{\mathcal{T}_i\}_{i \in I}$ be a family of tribes on a non empty set Ω. Show that

$$\mathcal{T}_I = \bigcap_{i \in I} \mathcal{T}_i$$

is still a tribe of Ω.

3.2.2 Borelians of $\mathbb{R}$, $\mathbb{R}^n$, and $\overline{\mathbb{R}}$

Proposition 16. Borelian tribe $\mathcal{B}(\mathbb{R})$.
The Borelian tribe of $\mathbb{R}$, $\mathcal{B}(\mathbb{R})$, is generated by the open intervals $]a; +\infty[$, with $a \in \mathbb{R}$.

Proof.
Let $\mathcal{I}$ be the tribe generated by the open intervals $]a; +\infty[$, with $a \in \mathbb{R}$.

1. By construction, $\mathcal{I} \subset \mathcal{B}(\mathbb{R})$.

2. Let $a \in \mathbb{R}$. If we consider the semi-open interval $[a; +\infty[$, we can write

$$[a; +\infty[= \bigcap_{k \in \mathbb{N}} \left] a - \frac{1}{k}; +\infty \right[$$

and since $\mathcal{I}$ is a tribe, $[a; +\infty[\in \mathcal{I}$. Using the stability by complement, we also have $]-\infty; a[= [a; +\infty[^c \in \mathcal{I}$. Using the stability by intersection, we have for $(a, b) \in \mathbb{R}^2, a < b,]a; b[=]-\infty; b[\cap]a; +\infty[$.

Yet, any open set of $\mathbb{R}$ is a (countable) union of intervals of the form $]-\infty; a[$, $]a; b[$, and $]a; +\infty[$, so $\mathcal{I}$ contains all the open intervals of $\mathbb{R}$. We therefore have $\mathcal{B}(\mathbb{R}) \subset \mathcal{I}$.

Hence, $\mathcal{I} = \mathcal{B}(\mathbb{R})$. □

Remark 26.
$\mathcal{B}(\mathbb{R})$ is a σ-algebra of sets, called the *Borel σ-algebra of* $\mathbb{R}$.

Remark 27.
What is in $\mathcal{B}(\mathbb{R})$?

1. All the open sets of $\mathbb{R}$.
2. All the closed sets of $\mathbb{R}$.
3. All the countable unions of closed sets of $\mathbb{R}$.
4. All the countable intersections of open sets of $\mathbb{R}$.
5. ... and more!

Exercise 33. ★ Borelian tribe $\mathcal{B}\left(\overline{\mathbb{R}}\right)$.
Characterise the Borelian tribe of $\overline{\mathbb{R}}$.

Exercise 34. ★★ Borelian tribe $\mathcal{B}\left(\mathbb{R}^n\right)$.
Show that the Borelian tribe of $\mathbb{R}^n$, $\mathcal{B}\left(\mathbb{R}^n\right)$, is generated by the open cuboids of $\mathbb{R}^n$.

Theorem 16. Measurability of Borel sets.
Every Borel set of real numbers is Lebesgue measurable.

Proof.
The Borel tribe of $\mathbb{R}$ is generated by the open intervals of $\mathbb{R}$. Similarly, the Borel tribe of $\mathbb{R}^n$ is generated by the open cuboids of $\mathbb{R}^n$. These are the smallest σ-algebra generated as such. Since $\mathcal{S}$, the set of Lebesgue-measurable sets, is also a σ-algebra, we only have to show that open intervals and open cuboids are in $\mathcal{S}$, which we did in the previous chapter. The proof is complete.

$\square$

Remark 28.
We have proven that $\mathcal{B} \subset \mathcal{S}$, i.e. all Borel sets of $\mathbb{R}^n$ are Lebesgue-measurable in the sense of Carathéodory. This does not mean, however, that any Lebesgue measurable set of $\mathbb{R}^n$ is a Borelian. In fact, there exist such sets, but they are beyond the scope of this class[6].

3.2.3 Structure of Lebesgue Measurable Sets

Theorem 17. Approximations of Lebesgue measurable sets.
Let E be a subset of $\mathbb{R}^n$. The following statements are equivalent:

1. E is Lebesgue measurable in the sense of Carathéodory.
2. For any $\varepsilon > 0$, there exists an open set G such that $E \subset G \subset \mathbb{R}^n$ and $\mu^\star(G \setminus E) < \varepsilon$ (called *exterior* approximation by open sets).
3. For any $\varepsilon > 0$, there exists a closed set F such that $F \subset E \subset \mathbb{R}^n$ and $\mu^\star(E \setminus F) < \varepsilon$ (called *interior* approximation by closed sets).
4. There exists a countable intersection B of open sets of $\mathbb{R}^n$ such that $E \subset B \subset \mathbb{R}^n$ and $\mu^\star(B \setminus E) = 0$.
5. There exists a countable union C of closed sets of $\mathbb{R}^n$ such that $C \subset E \subset \mathbb{R}^n$ and $\mu^\star(E \setminus C) = 0$.

[6]See Appendix C of F. Burk's *Lebesgue Measure and Integration*.

Proof.
We proceed step by step:

1. $1 \Rightarrow 2$. We assume that E is Lebesgue measurable in the sense of Carathéodory. Let $\varepsilon > 0$.

   (a) Case 1: $\mu^\star(E) < +\infty$. From the definition of the Lebesgue measure, we can find a family of open intervals $\{I_k\}_{k\in\mathbb{N}}$ such that, introducing G as

   $$G = \bigcup_{k\in\mathbb{N}} I_k$$

   we have

   $$\mu^\star(G) < \mu^\star(E) + \varepsilon.$$

   G is Lebesgue measurable and $G = (G \setminus E) \cup E$ so $\mu^\star(G) = \mu^\star(G \setminus E) + \mu^\star(E)$. As $\mu^\star(E) < +\infty$, we can substract and finally obtain $\mu^\star(G \setminus E) < \varepsilon$.

   (b) Case 2: $\mu^\star(E) = +\infty$. For all $k \in \mathbb{N}$, we define $E_k = E \cap [-k; k]$. E_k is Lebesgue measurable, $\mu^\star(E_k) < +\infty$ (by construction), and we just showed that for any $k \in \mathbb{N}$ we can find open sets G_k such that $E_k \subset G_k$ and $\mu^\star(G_k \setminus E_k) < \varepsilon/2^k$ (as small as we want). E is the countable union of all E_k and E is a subset of G being the countable union of all G_k. Therefore, we have

   $$\begin{aligned} \mu^\star(G \setminus E) &\leq \mu^\star\left(\bigcup_{k=1}^{+\infty}(G_k \setminus E_k)\right) \\ &\leq \sum_{k=1}^{+\infty} \mu^\star(G_k \setminus E_k) \\ &< \varepsilon. \end{aligned}$$

2. $2 \Rightarrow 3$: this can be deduced using the complement. Using the previous notations with E^c, we have $E^c \subset G$ with $\mu^\star(G \setminus E^c) < \varepsilon$. Then, if we introduce $F = G^c$, we have $F \subset E$ and $\mu^\star(E \setminus F) = \mu^\star(G \setminus E^c) < \varepsilon$.

3. $3 \Rightarrow 4$: let $E \subset \mathbb{R}^n$. We apply point 3 to E^c to build a sequence of closed sets $\{B_k\}_{k\in\mathbb{N}}$ such that $\forall k \in \mathbb{N}, B_k \subset E$ and $\mu^\star(E^c \setminus B_k) < 1/k$. We note B the countable intersection of all the B_k^c. Then, B is a countable intersection of open sets and $E \subset B$ by construction. Then

   $$\mu^\star(B \setminus E) = \mu^\star\left(E^c \setminus \bigcup_{k=1}^{+\infty} B_k\right),$$

   so

   $$\forall k \in \mathbb{N}, \mu^\star(B \setminus E) \leq \mu^\star(E^c \setminus B_k) < 1/k.$$

   Hence, we deduce $\mu^\star(B \setminus E) = 0$.

4. $4 \Rightarrow 5$: deduced using the complement, similarly as for $2 \Rightarrow 3$.

5. $5 \Rightarrow 1$: let E and C be the sets from point 5, and let $X \subset \mathbb{R}^n$. Note that E does not necessarily satisfies Carathéodory's criteria at this point, but C does.

   (a) By subadditivity, we already have

   $$\mu^\star(X) = \mu^\star\left((X \cap E) \cup (X \setminus E)\right) \leq \mu^\star(X \cap E) + \mu^\star(X \setminus E).$$

(b) First of all, we can write

$$\begin{aligned}\mu^\star(X\cap E) &= \mu^\star\left(((X\cap E)\cap C)\cup((X\cap E)\setminus C)\right)\\ &= \mu^\star\left((X\cap E)\cap C\right)+\mu^\star\left((X\cap E)\setminus C\right)\\ &= \mu^\star\left(X\cap C\right)+\mu^\star\left((X\cap E)\setminus C\right)\end{aligned}$$

Since $(X\cap E)\setminus C$ is a subset of $E\setminus C$, and since $\mu^\star(E\setminus C)=0$, we deduce that $\mu^\star\left((X\cap E)\setminus C\right)=0$ and therefore $\mu^\star(X\cap E)=\mu^\star\left(X\cap C\right)$. Moreover, we have

$$\begin{aligned}\mu^\star(X\setminus E) &= \mu^\star\left(((X\setminus E)\cap C)\cup((X\setminus E)\setminus C)\right)\\ &= \mu^\star\left((X\setminus E)\cap C\right)+\mu^\star\left((X\setminus E)\setminus C\right)\\ &= \mu^\star(\emptyset)+\mu^\star\left((X\setminus E)\setminus C\right)\\ &\leq \mu^\star(X\setminus C).\end{aligned}$$

Hence, we have

$$\mu^\star(X\cap E)+\mu^\star(X\setminus E)\leq\mu^\star(X\cap C)+\mu^\star(X\setminus C)=\mu^\star(X)$$

Therefore, E satisfies Carathéodory's measurability criteria and

$$\forall X\subset\mathbb{R}^n,\mu^\star(X)=\mu^\star(X\cap E)+\mu^\star(X\setminus E).$$

□

Theorem 18. "Completion" of Borel measure.
Every Lebesgue measurable set of real numbers is the union of a Borel set and of a Lebesgue measurable set of measure 0.

Proof.
Let E be a Lebesgue measurable set of real numbers. Using the previous theorem, we know that there exists a Borel set B such that $B\subset E$ and $\mu^\star(E\setminus B)=0$. Thus, we can write $E=B\cup(E\setminus B)$ with B a Borelian and $E\setminus B$ a Lebesgue measurable set of measure 0. □

Exercise 35. ★★★
Suppose E is a subset of $\mathbb{R}^n$ with $\mu^\star(E)<+\infty$. Show that E is Lebesgue measurable if and only if there exists U, a finite union of open intervals, such that for any $\varepsilon>0$ we have

$$\mu^\star\left(E\setminus U\right)+\mu^\star\left(U\setminus E\right)<\varepsilon.$$

3.2.4 Negligeable Ensembles and the Notion of *Almost Everywhere*

In all this subsection, Ω is a metric space with a measure μ.

Definition 57. Negligeable ensemble.
An ensemble $X\subseteq\Omega$ is said to be μ-negligeable if there is a Borelian $\mathcal{B}\subseteq\Omega$ such that $X\subseteq\mathcal{B}$ and $\mu(\mathcal{B})=0$.

Example 9. Countable set of $\mathbb{R}$.
If $D\subset\mathbb{R}$ is a countable set of $\mathbb{R}$, then D is Lebesgue-negligeable. This is trivial, since the Lebesgue measure of a singleton is 0 and any countable set of $D\subset\mathbb{R}$ can be written as

$$D=\bigcup_{x\in D}\{x\}.$$

Definition 58. Completeness of a measure.
Let λ be a measure. λ is said to be *complete* if, for any subset A of a measurable negligeable set X, A is also negligeable.

Proposition 17. Non-completeness of the Lebesgue measure.
The Lebesgue measure is not complete on $\mathcal{B}(\mathbb{R})$.

Proof.
This comes from the Cantor set: a Cantor set of measure 0 contains non-measurable subsets, therefore non-negligeable subsets.

□

Definition 59. Notion of *almost everywhere*.
Let $x \mapsto P(x)$ be a property that depends on $x \in \Omega$, and X the ensemble containing all the $x \in \Omega$ such that $P(x)$ is not satisfied. P is said to be true *μ-almost everywhere* if the ensemble X is μ-negligeable.

Example 10.
We consider the function

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto |x| \end{aligned}$$

This function is continuous on $\mathbb{R}$, but not derivable on $\mathbb{R}$: indeed, the left and right derivatives at 0 are different, so f is not derivable in 0. The set of points for which f is not derivable is the singleton $\{0\}$, and $\mu(\{0\}) = 0$: f is derivable almost everywhere.

Definition 60. Convergence almost everywhere.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions defined on $I \subset \mathbb{K}$. $(f_n)_{n\in\mathbb{N}}$ converges *almost everywhere* to the function f if and only if it converges pointwise on a subset $A \subset I$ whose complement (in I) is μ-negligeable.

Remark 29.
This provides a third type of convergence, weaker than the pointwise convergence (and also weaker than the uniform convergence). With regards to the integration (or other “large scale” transforms), the convergence almost everywhere is often a sufficient condition for existence.

Exercise 36. ★ Polynomial function.
Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a polynomial function. Show that either $f = 0$, or $f \neq 0$ almost everywhere. What can you say about $|f|$?

Exercise 37. ★ Fractional function.
Let f and g be two polynomial functions defined on $\mathbb{R}$ and with values in $\mathbb{R}$. We define $r = f/g$. Show that:

1. If $g \neq 0$, then r is defined almost everywhere.
2. If $f \neq 0$, then $r \neq 0$ almost everywhere.
3. For any $a \in \mathbb{R}$, either $r = a$, or $r \neq a$ almost everywhere.

3.3 Measurable Functions

DEFINITION 61. Measurable function.
Let f be a function defined on the measurable space X, with values in $\mathbb{R}$. The function f is said to be *measurable* if, for every $a \in \mathbb{R}$, the set

$$\{x \in X | f(x) > a\}$$

is measurable.

EXERCISE 38. ★ Characteristic function.

Let $A \subset \mathbb{R}$. Prove that the characteristic function of the set A, χ_A, is measurable if and only if A is measurable.

THEOREM 19. Stability of measurable functions (1/3).
If f and g are measurable functions defined on $X \subset \mathbb{R}$, then

1. If $f \neq 0$, $1/f$ is measurable;
2. $f + g$ is measurable;
3. When defined, $f \circ g$ is measurable;
4. fg is measurable;
5. $\min(f, g)$ and $\max(f, g)$ are measurable.

THEOREM 20. Stability of measurable functions (2/3).
If f is a measurable function defined on $X \subset \mathbb{R}$, then

1. $f^+ = \max(f, 0)$ and $f^- = -\min(f, 0)$ are measurable;
2. $|f| = f^+ + f^-$ is measurable;
3. $\forall p \in \mathbb{R}^{+\star}$, $|f|^p$ is measurable.

THEOREM 21. Stability of measurable functions (3/3).
If $(f_k)_{k\in\mathbb{N}}$ is a sequence of measurable functions defined on $X \subset \mathbb{R}$, then

1. $\sup_k f_k$ and $\inf_k f_k$ are measurable;
2. $\limsup_{k\to+\infty} f_k$ and $\liminf_{k\to+\infty} f_k$ are measurable;
3. If it exists, $\lim_{k\to+\infty} f_k$ is measurable.

Corollary 4. Measurability in $\mathbb{C}$.
Let f and g be two functions defined on $X \subset \mathbb{R}$ with values in $\mathbb{C}$.

1. f is measurable if and only if $\Re(f)$ and $\Im(f)$ are measurable;
2. If f and g are measurable, then so are $|f|$, $f+g$, $f \circ g$, fg;
3. There exists a measurable function $\alpha : X \to \mathbb{C}$, with $\forall x \in X, |\alpha(x)| = 1$ and $f = \alpha|f|$.

Exercise 39. ★ Measurability in $\mathbb{C}$.
Prove the last point of the previous corollary: if $f : X \to \mathbb{C}$ is a measurable function, then there exists a measurable function $\alpha : X \to \mathbb{C}$, with $\forall x \in X, |\alpha(x)| = 1$ and $f = \alpha|f|$.

Exercise 40. ★ Measurable functions.
Are these functions measurable? (do it without the following proposition)

1. $f_1 : x \mapsto \cos(x)$
2. $f_2 : x \mapsto \cos(x)^n$ for $n \in \mathbb{N}$
3. $f_3 : x \mapsto P(\cos(x))$ where $P \in \mathbb{R}[X]$ (polynomial)

Proposition 18. Continuity and measurability.
Continuous functions defined on measurable sets are measurable.

Proof.
Let f be a continuous function on a measurable set $E \subset \mathbb{R}$. Our goal is to show that

$$\forall a \in \mathbb{R}, A = \{x \in E | f(x) > a\} \quad \text{is measurable.}$$

Let $a \in \mathbb{R}$ and A the subset defined above.

1. If $A = \emptyset$, then we already know that the empty set is measurable.
2. If $A \neq \emptyset$, then let $x \in A$. We know that $f(x) > a$ by definition (strict inequality, a is not on the frontier), and since f is continuous at x this means that there exists $\varepsilon_x > 0$ such that $X =]x - \varepsilon_x; x + \varepsilon_x[\cap E$ and $\forall y \in X, f(y) > a$.

   Hence, we can "construct" A thanks to its elements as follows

   $$A = \bigcup_{x \in A} (]x - \varepsilon_x; x + \varepsilon_x[\cap E) = \left(\bigcup_{x \in A}]x - \varepsilon_x; x + \varepsilon_x[\right) \cap E,$$

   where the ε_x are defined for each $x \in A$. A is then the intersection of a measurable set E and of a (non-necessarily countable) union of measurable sets (the intervals $]x-\varepsilon_x; x+\varepsilon_x[$), so A is measurable by stability of the measurable sets.

□

3.4 Concluding Remarks

Previously, in chapter 2, we have built the Lebesgue measure in $\mathbb{R}^n$, with the conclusion that measurable sets satisfy Carathéodory's measurability criteria. Examples of such sets are: intervals, cuboids, compacts, and all their unions, intersections, and so on. In order to define more rigorously the set of measurable sets, we introduced in this chapter (chapter 3) the notion of Borelians. From this, we saw how to define Lebesgue measurable sets, and how to approximate open and closed sets by their exteriors and interiors, respectively. The Lebesgue measure on $\mathbb{R}$ allows us to define the two very important notions of *almost everywhere* and of *measurable functions*. We will build upon these the integration theory.

This chapter concludes our discussion on the Lebesgue measure in itself, so it is worth reminding that, when measuring a set, one has to be careful with the "intuitive" results! As seen in chapter 2, the *naive* definition of lengths and measures does not work (mainly due to the countable unions). At the beginning of chapter 3, we saw that there exist non-measurable sets in $\mathbb{R}^n$, and that sometimes the measure of a set is counter-intuitive (with the example of the Cantor set).

CHAPTER 4

Integration

The integral we are all familiar with is the *Riemann integral*, which is obtained by an approximation of the integrable function by a staircase function over a fixed subdivision of the set of definition. This definition is very intuitive and easy to apply to a given function, but with this "simplicity" comes several issues, such as numerous non-integrable functions, convergence theorems, and so on. For example, what can we say about the Dirichlet integral

$$\int_0^{+\infty} \frac{\sin t}{t} \mathrm{d}t,$$

that is non-integrable? And yet, can be computed (and is equal to $\pi/2$)! Or, if we define

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

what would be

$$\int_0^1 f(x)\mathrm{d}x \ ?$$

We can say a few things, like it is bounded by 0 and 1, but not much more. In order to address these issues, Lebesgue proposed to redefine the Riemann integral thans to a change of point of view, counting the "volume under a curve" differently: based on his measure theory, this is the *Lebesgue integral*.

4.1 Riemann Integral vs. Lebesgue Integral

Let $n \in \mathbb{N}$ be a "discretisation" parameter, i.e. how much the space is discretised, or how well the integrated function is approximated (figure 4.1). We write the Riemann sum as

$$R_n(f) = \sum_{k=0}^{n} f(x_k)\delta x.$$

Such a sum corresponds to a sum of rectangles approximating the area or volume "under the curve of the function f" that can be obtained thanks to a staircase function. Note that, depending on the definition of the Riemann sum, the value of f for each rectangle starting at

x_k can be taken either in x_k or in x_{k+1}; similarly, it is also possible to define this sum using Darboux's sums of staircase functions always larger or smaller than f (we will come back to this later). The Riemann integral is then obtained by taking the limit n goes to infinity of this sum, i.e. taking the limit δx goes to 0. We then write the Riemann integral of f as

$$R_\infty(f) = \sum_{k=0}^{+\infty} f(x_k)\delta x \equiv \int_0^1 f(x)\mathrm{d}x.$$

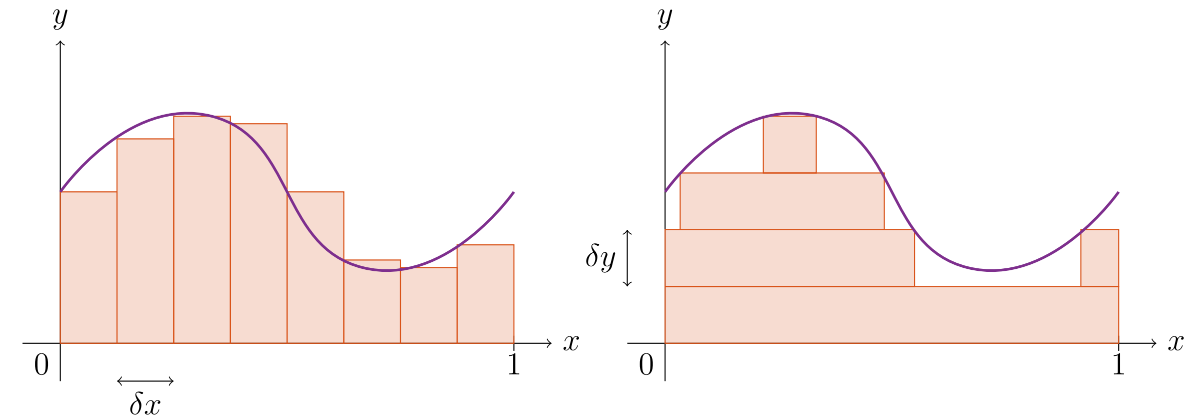


Figure 4.1: Basic Principle of the Riemann integral (left) and Lebesgue integral (right).

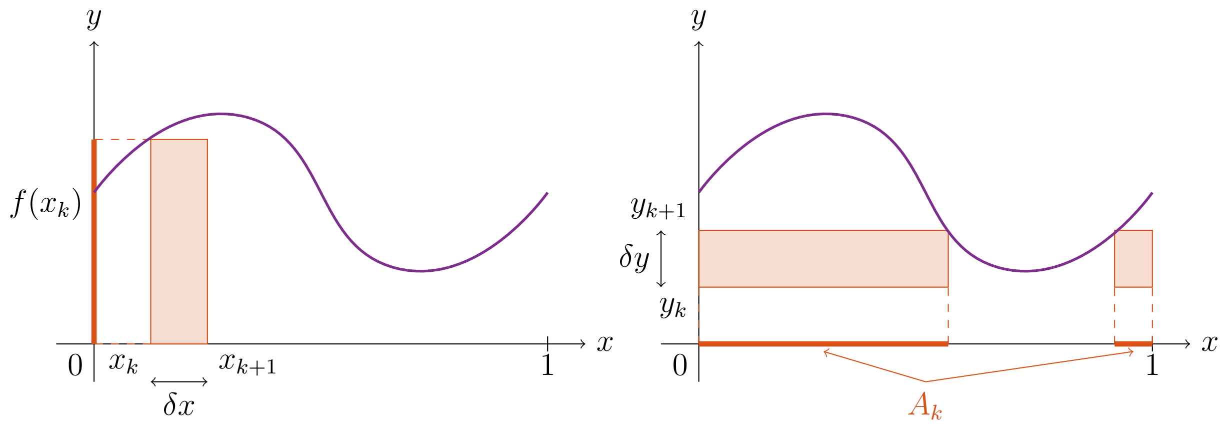


Figure 4.2: Elementary sets for the Riemann integral (left) and Lebesgue integral (right).

As shown in figure 4.1, instead of discretising the space on which the function is defined, we can discretise the set of images and write the Lebesgue sum as

$$L_n(f) = \sum_{k=0}^{n} \mu(A_k)\delta y,$$

by measuring the size of the A_k. Hence, we have the correspondances $f(x_k) \leftrightarrow \delta y$ and $\delta x \leftrightarrow \mu(A_k)$ (see figure 4.2). The idea of the Lebesgue integral is therefore to proceed as for the Riemann integral: taking the limit n goes to infinity of this sum, this time by taking the limit δy goes to 0. We want to write the Lebesgue integral of f as

$$L_\infty(f) = \sum_{k=0}^{+\infty} \mu(A_k)\delta y \equiv \int_0^1 f(x)\mathrm{d}\mu(x).$$

In the next section, we will see how such an integral can be defined rigorously, by considering the integrals of simple functions.

4.2 Simple Functions

4.2.1 Lebesgue Integration of Simple Functions

DEFINITION 62. Simple functions.
Let f be a function on a measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. We say that f is a simple function if there exists $n \in \mathbb{N}$ such that there is a finite sequence of numbers in $\{a_k\}_{k \in [\![1;\, n]\!]} \in \mathbb{R}^n$ and a finite sequence of mutually disjoint sets $\{E_k\}_{k \in [\![1;\, n]\!]} \in \mathcal{B}(\mathbb{R})^n$ such that

$$f = \sum_{k=1}^{n} a_k \chi_{E_k},$$

with χ_{E_k} the characteristic function associated to the set E_k.

EXAMPLE 11.

1. Let $E \subset \mathbb{R}$. The characteristic function of the set, χ_E, is a trivial simple function.
2. The sign function σ defined by

$$\begin{aligned} f : \mathbb{R} &\rightarrow \{-1, +1\} \\ x &\mapsto \begin{cases} +1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0, \end{cases} \end{aligned}$$

is a simple function.

3. The staircase function f defined by

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{N} \\ x &\mapsto n \in \mathbb{N}, n \leq x < n+1, \end{aligned}$$

is not a simple function, since it cannot be described by a finite set of real numbers associated to a finite set of Borelians. In fact, we can write

$$f = \sum_{k=-\infty}^{+\infty} k \chi_{E_k} \qquad \text{with} \qquad \forall k \in \mathbb{Z}, E_k = [k;\, k+1[.$$

4. Any restriction of the staircase function f to a bounded interval I, f_I, defined by

$$\begin{aligned} f_I : \mathbb{R} &\rightarrow \mathbb{N} \\ x &\mapsto \begin{cases} f(x) & \text{if } x \in I \\ 0 & \text{if } x \notin I, \end{cases} \end{aligned}$$

is a simple function.

DEFINITION 63. Integrability of a simple function.
Let f be an simple function, defined with a finite sequence of numbers in $\{a_k\}_{k \in [\![1;\, n]\!]} \in \mathbb{R}^n$ and a finite sequence of mutually disjoint sets $\{E_k\}_{k \in [\![1;\, n]\!]} \in \mathcal{B}(\mathbb{R})^n$. f is said to be *integrable* if for all k in $[\![1;\, n]\!]$, if $a_k \neq 0$, then $\mu(E_k) < +\infty$.

Definition 64. Integral of a simple function.
Let f be an integrable simple function, defined with a finite sequence of numbers in $\{a_k\}_{k\in[\![1;n]\!]} \in \mathbb{R}^n$ and a finite sequence of mutually disjoint sets $\{E_k\}_{k\in[\![1;n]\!]} \in \mathcal{B}(\mathbb{R})^n$. The *integral* of f, written

$$\int f \mathrm{d}\mu = \int_{\mathbb{R}} f(x)\mathrm{d}\mu(x),$$

is defined by

$$\int f \mathrm{d}\mu = \sum_{k=1}^{n} a_k \mu(E_k).$$

Proposition 19. Invariance of the integral.
Let f be an integrable simple function. The integral of f is independent of the representation chosen. In other words, if we have two representations of f such as $\{a_k\}_{k\in[\![1;n]\!]} \in \mathbb{R}^n$ and $\{E_k\}_{k\in[\![1;n]\!]} \in \mathcal{B}(\mathbb{R})^n$, and $\{b_k\}_{k\in[\![1;m]\!]} \in \mathbb{R}^m$ and $\{F_k\}_{k\in[\![1;m]\!]} \in \mathcal{B}(\mathbb{R})^m$, such that

$$f = \sum_{k=1}^{n} a_k \chi_{E_k} = \sum_{k=1}^{m} b_k \chi_{F_k},$$

then the integral is unambiguously defined

$$\int f \mathrm{d}\mu = \sum_{k=1}^{n} a_k \mu(E_k) = \sum_{k=1}^{m} b_k \mu(F_k).$$

Exercise 41. ★ Invariance of the integral.
Prove the proposition on the invariance of the integral with respect to the definition of a simple function.

Definition 65. Lebesgue integral.
Let f be an integrable simple function, defined on $X \subset \mathbb{R}$. The *Lebesgue integral* of f on $E \subset X$ is defined as

$$\int_E f \mathrm{d}\mu = \int_E f(x)\mathrm{d}\mu(x) = \int_{\mathbb{R}} \chi_E(x) f(x)\mathrm{d}\mu(x) = \int \chi_E f \mathrm{d}\mu.$$

Example 12.

1. The characteristic function of a set E of finite measure is integrable and

$$\int \chi_E \mathrm{d}\mu = \int_{\mathbb{R}} \chi_E(x)\mathrm{d}\mu(x) = \int_E \mathrm{d}\mu = \mu(E)$$

2. The sign function is not integrable.

3. Any restriction to a compact of the staircase function is integrable.

Exercise 42. ★ Removal of a countable subset.
We consider $X = [-1; 1]$ and $f : x \mapsto 1/2$. If f integrable on X? What is the value of the integral of f on X? We now consider $E = X \setminus \mathbb{Q}$. Is f integrable on E? If so, what is the value of the integral of f on E? If they are defined, can you comment on the values of these two integrals?

DEFINITION 66. Sets of measurable simple functions.
We will note $\mathcal{E}$ the set of measurable simple functions, and $\mathcal{E}^+$ the set of positive measurable simple functions.

4.2.2 Key Theorems on Simple Function Integrals

THEOREM 22. Linearity of the integral.
If f and g are integrable simple functions, and if $(a, b) \in \mathbb{R}^2$, then

$$\int (af + bg)\mathrm{d}\mu = a \int f\mathrm{d}\mu + b \int g\mathrm{d}\mu.$$

THEOREM 23. Positivity of the integral.
If f is an integrable simple function, positive almost everywhere, then

$$\int f\mathrm{d}\mu \geq 0.$$

THEOREM 24. Monotonicity of the integral.
If f and g are integrable simple function such that $f \geq g$ almost everywhere, then

$$\int f\mathrm{d}\mu \geq \int g\mathrm{d}\mu.$$

THEOREM 25. Triangle inequality.
If f and g are integrable simple functions, then

$$\int |f + g|\mathrm{d}\mu \leq \int |f|\mathrm{d}\mu + \int |g|\mathrm{d}\mu.$$

THEOREM 26. Absolute value inequality.
If f is an integrable simple function, then

$$\left|\int f\mathrm{d}\mu\right| \leq \int |f|\mathrm{d}\mu.$$

EXERCISE 43. ★★ Key theorems on simple function integrals.
Prove the five theorems on simple function integrals.

4.2.3 Indefinite Integrals of Simple Functions

DEFINITION 67. Indefinite integral.
The indefinite integral of an integrable function f is the set function $\mathcal{I}$ defined for every measurable set E by

$$\mathcal{I}(E) = \int_E f\mathrm{d}\mu.$$

Theorem 27. Monotonicity.
If f is an integrable function, non-negative almost everywhere, then its indefinite integral is monotone, i.e. if $E \subset F$ then

$$\int_E f \mathrm{d}\mu \leq \int_F f \mathrm{d}\mu.$$

Proof.
If $E \subset F$ are measurable subsets and if f is an integrable function, non-negative almost everywhere, then $\chi_E f \leq \chi_F f$ and the result follows from the monotonicity of the integral. □

Theorem 28. Absolute continuity.
The indefinite integral of an integrable function is absolutely continuous.

Proof.
Let c be any positive number greater than all the values of $|f|$ on the set of integration E. Then, we can write

$$\left|\int_E f \mathrm{d}\mu\right| \leq c\mu(E).$$

□

Theorem 29. Countable additivity.
The indefinite integral of an integrable function is countably additive.

Proof.
This is a direct consequence of the countable additivity of μ. □

Exercise 44. ★ Distance between two functions.
Let f and g be two integrable functions. We define the application ρ by

$$\rho(f,g) = \int |f-g| \mathrm{d}\mu.$$

Show that $\rho(f,f) = 0$, $\rho(f,g) = \rho(g,f)$, and $\rho(f,g) \leq \rho(g,h) + \rho(h,g)$. Is ρ a distance?

4.3 Lebesgue Integration

4.3.1 Integration of Positive Measurable Functions

Lemma 3. Approximation by a staircase function.
Let $f : \mathbb{R} \to [0; +\infty]$ be a positive measurable function. There exists an increasing sequence of measurable staircase functions converging pointwise to f.

Proof.
For all $n \in \mathbb{N}$ we define ϕ_n by

$$\begin{aligned} \phi_n : \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto \begin{cases} 2^{-n}E(2^n x) & \text{if } x < n \\ n & \text{if } x \geq n, \end{cases}, \end{aligned}$$

with $x \mapsto E(x)$ is the integer part. ϕ_n is a positive staircase function and $\forall t \in \mathbb{R}^+, 0 \leq \phi_n(t) \leq \phi_{n+1}(t)$. We define, for $n \in \mathbb{N}$, the function $f_n = \phi_n \circ f$. Then, $\{f_n\}_{n\in\mathbb{N}}$ is an increasing sequence of measurable staircase functions converging pointwise to f.

□

Definition 68. Lebesgue integral of a positive measurable function.
Let $f : \mathbb{R} \to [0; +\infty]$ be a positive measurable function. The *integral of f* is defined by

$$\int f \mathrm{d}\mu = \sup \left\{ \int h \mathrm{d}\mu \middle| h \in \mathcal{E}^+ \quad \text{and} \quad h \leq f \right\}.$$

And, if E is a measurable subset of $\mathbb{R}$, we note

$$\int_E f \mathrm{d}\mu = \int \chi_E f \mathrm{d}\mu.$$

Theorem 30. Properties of the Lebesgue integral.
The Lebesgue integral of positive measurable functions is linear, positive, monotonic, and satisfies the triangle inequality and the absolute value inequality.

Exercise 45. ★ Vanishing property.
Let f be a continuous positive measurable function. Prove that

$$\int f \mathrm{d}\mu = 0 \Leftrightarrow \{x \in \mathbb{R} | f(x) > 0\} = \emptyset.$$

What can we say if f is not continuous?

Exercise 46. ★ Finiteness property.
Let f be a continuous positive measurable function. Prove that

$$\int f \mathrm{d}\mu < +\infty \Leftrightarrow \{x \in \mathbb{R} | f(x) = +\infty\} = \emptyset.$$

What can we say if f is not continuous?

Proposition 20. Domination of a measurable set by an integral.
Let f be a positive measurable function defined on a subset $X \subset \mathbb{R}$. Then

$$\forall a \in X, \mu\left(\{x \in X | f(x) \geq a\}\right) \leq \frac{1}{a} \int f \mathrm{d}\mu.$$

Proof.
Let $a \in X$ and $A = \{x \in X | f(x) \geq a\}$. This set is measurable (inverse image of a measurable set by a measurable function) and $f \geq a\chi_A$, so by monotonicity

$$\int f \mathrm{d}\mu \geq a\mu(A).$$

□

Proposition 21. Integral equality *almost everywhere*.
Let f be a positive measurable function defined on a subset $X \subset \mathbb{R}$. Then

$$\int f \mathrm{d}\mu = 0 \Leftrightarrow f = 0 \text{ almost everywhere.}$$

Proof.
Let assume that $f = 0$ almost everywhere. Then, if we pick $h \in \mathcal{E}^+$ such that $h \leq f$, we have $h = 0$ almost everywhere. Therefore, the integral of h is equal to zero and this does not change when taking the sup, so

$$f = 0 \text{ almost everywhere} \Rightarrow \int f \mathrm{d}\mu = 0.$$

Now, let assume that the integral of f is equal to zero. For all $n \in \mathbb{N}^\star$ we define

$$A_n = \left\{ x \in X \,\middle|\, f(x) \geq \frac{1}{n} \right\}.$$

Then, A_n is measurable, $A_n \subset A_{n+1}$ and the countable union of the A_n is $A = \{x \in X \,|\, f(x) > 0\}$. Using the domination of a measurable set by an integral proposition, we have

$$\forall n \in \mathbb{N}^\star, \mu(A_n) \leq n \int f \mathrm{d}\mu.$$

Therefore,

$$\mu(A) = \lim_{n \to +\infty} \mu(A_n) = 0,$$

and

$$\int f \mathrm{d}\mu = 0 \Rightarrow f = 0 \text{ almost everywhere.}$$

□

Corollary 5. Integral equality *almost everywhere.*
Let f and g be two positive measurable functions defined on a subset $X \subset \mathbb{R}$. Then

$$f = g \text{ almost everywhere } \Rightarrow \int f \mathrm{d}\mu = \int g \mathrm{d}\mu.$$

Exercise 47. ★ Integral equality *almost everywhere.*
Prove this corollary, and explain why the equivalence is not true.

Theorem 31. Monotone convergence (Beppo-Levi).
Let $\{f_n\}_{n \in \mathbb{N}}$ be an increasing sequence of positive measurable functions defined on a measurable set X, and f be the pointwise limit of this sequence. Then, f is measurable and

$$\int f \mathrm{d}\mu = \lim_{n \to +\infty} \int f_n \mathrm{d}\mu.$$

Proof.
We first note that, since $\{f_n\}_{n \in \mathbb{N}}$ is an increasing sequence of functions, the monotonicity of the integral gives

$$\forall n \in \mathbb{N}, \int f_n \mathrm{d}\mu \leq \int f_{n+1} \mathrm{d}\mu,$$

and therefore the limit

$$a = \lim_{n \to +\infty} \int f_n \mathrm{d}\mu$$

exists (in $\overline{\mathbb{R}}^+$). As the sequence $\{f_n\}_{n \in \mathbb{N}}$ is converging pointwise to f, the function f is measurable, and since $\forall n \in \mathbb{N}, f_n \leq f$, we obtain from the monotonicity of the integral

$$a \leq \int f \mathrm{d}\mu.$$

Let $h \in \mathcal{E}^+$ such that $h \leq f$ and let $c \in]0;1[$. For all $n \in \mathbb{N}$, we define the subsets

$$A_n = \{x \in X | f_n(x) \geq c\, h(x)\}.$$

These subsets are measurable, we have $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$, and the countable union of the A_n forms the initial set X. Yet, we can write

$$\int f \mathrm{d}\mu \geq \int_{A_n} f_n \mathrm{d}\mu \geq \int_{A_n} ch \mathrm{d}\mu = c \int_{A_n} h \mathrm{d}\mu.$$

If we now write h as a simple function

$$h = \sum_{k=1}^{m} \alpha_k B_k,$$

then

$$\int_{A_n} h \mathrm{d}\mu = \sum_{k=1}^{m} \alpha_k \mu(B_k \cap A_n).$$

Here we can take the limit n goes to $+\infty$ and we obtain

$$\lim_{n \to +\infty} \int_{A_n} h \mathrm{d}\mu = \int h \mathrm{d}\mu = \sum_{k=1}^{m} \alpha_k \mu(B_k).$$

Hence, we deduce that

$$\forall c \in]0;1[, \forall h \in \mathcal{E}^+, h \leq f, a \geq c \int h \mathrm{d}\mu.$$

By taking the sup on $c \in]0;1[$ and then the sup on $h \in \mathcal{E}^+$ with $h \leq f$, we find that

$$a \geq \int f \mathrm{d}\mu.$$

Thus, we conclude that

$$a = \int f \mathrm{d}\mu.$$

□

Corollary 6. Monotone convergence (decreasing sequence).
Let $\{f_n\}_{n\in\mathbb{N}}$ be a decreasing sequence of positive measurable functions defined on a measurable set X, and f be the pointwise limit of this sequence. Then, if the integral of f_0 is finite, f is measurable and

$$\int f \mathrm{d}\mu = \lim_{n \to +\infty} \int f_n \mathrm{d}\mu.$$

Proof.
By application of the previous theorem to the increasing sequence of positive measurable functions $\{g_n\}_{n\in\mathbb{N}}$ defined by $\forall n \in \mathbb{N}, g_n = f_0 - f_n$. □

Corollary 7. Countable additivity.
Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of positive measurable functions defined on a measurable set X, and f be defined by

$$f = \sum_{n\in\mathbb{N}} f_n.$$

Then f is measurable and

$$\int f \mathrm{d}\mu = \sum_{n\in\mathbb{N}} \int f_n \mathrm{d}\mu.$$

Proof.
We define, for $N \in \mathbb{N}$, the function

$$g_N = \sum_{n=0}^{N} f_n.$$

Then, $\{g_n\}_{n\in\mathbb{N}}$ is an increasing sequence of positive measurable functions and, for all $N \in \mathbb{N}$, we have

$$\int g_n \mathrm{d}\mu = \sum_{n=0}^{N} \int f_n \mathrm{d}\mu.$$

The corollary follows by taking the limit N goes to $+\infty$ and by applying Beppo-Levi's theorem. □

Lemma 4. Fatou's lemma.
Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of positive measurable functions. Then

$$\int \liminf_{n\to+\infty} f_n \mathrm{d}\mu \leq \liminf_{n\to+\infty} \int f_n \mathrm{d}\mu.$$

Proof.
Since $\liminf\limits_{n\to+\infty} f_n = \lim\limits_{k\to+\infty} \left(\inf\limits_{n\geq k} f_n\right)$, this is the limit of an increasing sequence of functions, therefore Beppo-Levi's theorem gives

$$\int \liminf_{n\to+\infty} f_n \mathrm{d}\mu = \lim_{k\to+\infty} \int \left(\inf_{n\geq k} f_n\right) \mathrm{d}\mu.$$

In addition, if we let $p \geq k$, we have $\inf\limits_{n\geq k} f_n \leq f_p$, so

$$\int \left(\inf_{n\geq k} f_n\right) \mathrm{d}\mu \leq \int f_p \mathrm{d}\mu,$$

and therefore

$$\int \left(\inf_{n\geq k} f_n\right) \mathrm{d}\mu \leq \inf_{p\geq k} \int f_p \mathrm{d}\mu.$$

Hence, we have

$$\int \liminf_{n\to+\infty} f_n \mathrm{d}\mu \leq \lim_{k\to+\infty} \inf_{p\geq k} \int f_p \mathrm{d}\mu = \liminf_{n\to+\infty} \int f_n \mathrm{d}\mu.$$

□

Remark 30.
Fatou's lemma is actually a corollary of Beppo-Levi's monotone convergence theorem.

4.3.2 Integration of Measurable Functions

Definition 69. Lebesgue integral of a measurable function (in $\mathbb{R}$).
Let f be a measurable function with values in $\mathbb{R}$. f is integrable with respect to the measure μ if

$$\int |f| \mathrm{d}\mu < +\infty.$$

The *integral of f* is then defined by

$$\int f \mathrm{d}\mu = \int f^+ \mathrm{d}\mu - \int f^- \mathrm{d}\mu,$$

where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$.

Definition 70. $\mathcal{L}^1$ space.
We note $\mathcal{L}^1(X, \mathcal{B}(X), \mu)$ the space of integrable functions defined on X, where measurable sets are the borelians of X, and the measure considered is the Lebesgue measure μ. In general, we will omit to mention the Borelian set and even the measure, when we are unambiguously using the Lebesgue measure. We then note $\mathcal{L}^1(X)$.

Theorem 32. Properties of the Lebesgue integral (in $\mathbb{R}$).
The Lebesgue integral of measurable functions is linear, positive, monotonic, and satisfies the triangle inequality and the absolute value inequality.

Exercise 48. ★★ Properties of the Lebesgue integral.
Prove that the Lebesgue integral of measurable functions is linear, positive, monotonic, and satisfies the triangle inequality and the absolute value inequality.

Proposition 22. Domination.
If f is integrable and g is a measurable function such that $|g| \leq |f|$, then g is integrable.

Proof.
This is trivial using the monotonicity of the integral on functions in $\mathcal{E}^+$ approximating $|g|$. □

Proposition 23. Bounded product.
If f is integrable and g is an essentially bounded measurable function (i.e. almost everywhere), then fg is integrable.

Proof.
Let $c \in \mathbb{R}$ such that $|g| \leq c$ almost everywhere. Then, $|fg| \leq c|f|$ almost everywhere, and the result follows from the domination proposition. □

Proposition 24. Bounded function.
If f is an essentially bounded measurable function, and if E is a set of finite measure, then f is integrable over E.

Proof.
Let f be an essentially bounded measurable function and $g = \chi_E$. Then g is integrable over E since E has a finite measure, and the previous proposition on bounded product applies to fg. □

Definition 71. Lebesgue integral of a measurable function (in $\mathbb{C}$).
Let f be a measurable function with values in $\mathbb{C}$. f is integrable with respect to the measure μ if

$$\int |f| \mathrm{d}\mu < +\infty.$$

The *integral of* f is then defined by

$$\int f \mathrm{d}\mu = \int \Re(f) \mathrm{d}\mu + i \int \Im(f) \mathrm{d}\mu.$$

Remark 31.
If the imaginary part of f is zero, we recover the previous definition.

Theorem 33. Properties of the Lebesgue integral (in $\mathbb{C}$).
Again, the Lebesgue integral of measurable functions is linear, positive, monotonic, and satisfies the triangle inequality and the absolute value inequality.

Exercise 49. ★ Integrability of f and $|f|$.
Let f be a measurable function. Show that $f \in \mathcal{L}^1(\mathbb{R}) \Leftrightarrow |f| \in \mathcal{L}^1(\mathbb{R})$. Then, prove the absolute value inequality

$$\left|\int f \mathrm{d}\mu\right| \leq \int |f| \mathrm{d}\mu.$$

Exercise 50. ★ Domination.
Let f and g be two measurable functions. We assume that $g \in \mathcal{L}^1(\mathbb{R})$ and that $|f| \leq |g|$. Show that $f \in \mathcal{L}^1(\mathbb{R})$.

4.3.3 Negligeable Sets

Proposition 25. Integration on a set of measure 0.
If f is an integrable function defined on X and if $E \subset X$ is a set of measure 0, then

$$\int_E f \mathrm{d}\mu = 0.$$

Proof.
We write

$$\int_E f \mathrm{d}\mu = \int \chi_E f \mathrm{d}\mu.$$

Then, the function $x \mapsto \chi_E(x) f(x)$ is zero almost everywhere since $\mu(E) = 0$. Yet, we have seen that if a function is equal to *zero* almost everywhere, then its integral is equal to zero. □

Proposition 26. Integration on a strictly positive function.
If f is an integrable function defined on X, positive almost everywhere, and if $E \subset X$, then

$$\int_E f \mathrm{d}\mu = 0 \Rightarrow \mu(E) = 0.$$

Proof.
We define the sequence of sets $(F_n)_{n\in\mathbb{N}}$ by

$$F_0 = \{x \in E | f(x) > 0\} \qquad \text{and} \qquad \forall n \in \mathbb{N}^\star, F_n = \left\{x \in E \middle| f(x) \geq \frac{1}{n}\right\}.$$

E and F_0 are measurable, so we can write (Carathéodory)

$$\mu(E) = \mu(E \cap F_0) + \mu(E \setminus F_0).$$

Since f is positive and the integral of f over E is zero, we have $\mu(E \setminus F_0) = 0$. Now, we note that

$$\forall n \in \mathbb{N}^\star, 0 = \int_{E\cap F_n} f \mathrm{d}\mu \geq \frac{1}{n}\mu(E \cap F_n) \geq 0,$$

and that

$$F_0 = \bigcup_{n=1}^{+\infty} F_n,$$

so we have the following inequalities

$$0 \leq \mu(E \cap F_0) \leq \sum_{n=1}^{+\infty} \mu(E \cap F_n) = 0.$$

Hence, $\mu(E \cap F_0) = 0$ and the result follows. □

PROPOSITION 27. Integration measurable sets.
If f is an integrable function defined on X, then if

$$\forall E \subset X, \int_E f \mathrm{d}\mu = 0,$$

with E measurable, then $f = 0$ almost everywhere.

PROOF.
We define $A = \{x \in X | f(x) > 0\}$. By hypothesis, since $A \subset X$, we have

$$\int_A f \mathrm{d}\mu = 0.$$

By applying the proposition on the integration of a strictly positive function, we deduce that A is a set of measure 0. We can apply the same reasoning to the function $-f$ and show that $B = \{x \in X | f(x) < 0\}$ has measure 0. Hence, $A \cup B = \{x \in X | f(x) \neq 0\}$ has measure 0. □

4.3.4 Integral Calculus

THEOREM 34. Chasles.
Let f be an integrable function on $A \subset \mathbb{R}$, and $[a; b] \subset A$. Then, for all $c \in [a; b]$, we have

$$\int_a^b f \mathrm{d}\mu = \int_a^c f \mathrm{d}\mu + \int_c^b f \mathrm{d}\mu.$$

PROOF.
Trivial using the additivity property of the Lebesgue measure. □

DEFINITION 72. Convergence – divergence.
Let $a \in \mathbb{R}$ and let f be a measurable function defined on $[a; +\infty[$ with values in $\mathbb{R}$. We say that the integral of f over $[a; +\infty[$ is *converging* if the function

$$x \mapsto \int_a^x f \mathrm{d}\mu,$$

has a finite limit in $+\infty$. If not, the integral is says to be *diverging*.

REMARK 32.

1. These are also often called *proper* and *improper* integrals.

2. This definition is an extension of the definition of the integral; however, it does not change the definition of integrability. In fact, an integrable function has a finite integral.

3. A diverging integral is not necessarily equal to $\pm\infty$. It is possible that the limit does not exist at all.

THEOREM 35. Integrability.
Let f be a measurable function defined on $A \subset \mathbb{R}$. f is integrable if and only if its integral on A is converging.

THEOREM 36. Integration by parts.
Let $(a, b) \in \overline{\mathbb{R}}^2$ and $(f, g) \in \mathcal{C}^1(]a; b[)$. We assume that the limits of fg in a and b exist and are finite. If $f'g$ is integrable, then so if $g'f$ and we have

$$\int_a^b f'g\mathrm{d}\mu = \lim_{x\to b} f(x)g(x) - \lim_{x\to a} f(x)g(x) - \int_a^b fg'\mathrm{d}\mu.$$

THEOREM 37. Change of variables.
Let f be integrable on $]a; b[\subset \mathbb{R}$ and $\phi :]\alpha; \beta[\to]a; b[$ a function in $\mathcal{C}^1(]\alpha; \beta[)$, that is increasingly monotonic and bijective. We have

$$\int_a^b f\mathrm{d}\mu = \int_\alpha^\beta (f \circ \phi)\phi'\mathrm{d}\mu.$$

REMARK 33.
We will see, in the next chapter, how to properly formulate a change of variables.

PROPOSITION 28. Integration of asymptotics: “domination”.
Let (f, g) be two positive measurable functions on $[a; b[\subset \mathbb{R}$. We assume that $f \underset{x\to b}{O}(g)$. Then

1. If $\int_a^b f\mathrm{d}\mu$ is diverging, then $\int_a^b g\mathrm{d}\mu$ is diverging and $\int_a^x f\mathrm{d}\mu \underset{x\to b}{O}\left(\int_a^x g\mathrm{d}\mu\right)$.

2. If $\int_a^b g\mathrm{d}\mu$ is converging, then $\int_a^b f\mathrm{d}\mu$ is converging and $\int_x^b f\mathrm{d}\mu \underset{x\to b}{O}\left(\int_x^b g\mathrm{d}\mu\right)$.

PROPOSITION 29. Integration of asymptotics: “preponderance”.
Let (f, g) be two positive measurable functions on $[a; b[\subset \mathbb{R}$. We assume that $f \underset{x\to b}{o}(g)$. Then

1. If $\int_a^b f\mathrm{d}\mu$ is diverging, then $\int_a^b g\mathrm{d}\mu$ is diverging and $\int_a^x f\mathrm{d}\mu \underset{x\to b}{o}\left(\int_a^x g\mathrm{d}\mu\right)$.

2. If $\int_a^b g\mathrm{d}\mu$ is converging, then $\int_a^b f\mathrm{d}\mu$ is converging and $\int_x^b g\mathrm{d}\mu \underset{x\to b}{o}\left(\int_x^b g\mathrm{d}\mu\right)$.

Proposition 30. Integration of asymptotics: "equivalence".
Let (f, g) be two positive measurable functions on $[a; b[\subset \mathbb{R}$. We assume that $f \underset{x\to b}{\sim} g$. Then

1. If $\int_a^b f \mathrm{d}\mu$ is diverging, then $\int_a^b g \mathrm{d}\mu$ is diverging and $\int_a^x f \mathrm{d}\mu \underset{x\to b}{\sim} \int_a^x g \mathrm{d}\mu$.
2. If $\int_a^b f \mathrm{d}\mu$ is converging, then $\int_a^b g \mathrm{d}\mu$ is converging and $\int_x^b f \mathrm{d}\mu \underset{x\to b}{\sim} \int_x^b g \mathrm{d}\mu$.

4.3.5 Scaling Functions

Definition 73. Reference functions.
A family of functions $(\phi_i)_{i\in I}$ defined on a subset $A \subset \mathbb{R}$ is a family of *reference functions* in the neighbourhood of $a \in A$ if and only if

$$\forall (i, j) \in I^2, i \neq j, \phi_i = o(\phi_j) \quad \text{or} \quad \phi_j = o(\phi_i).$$

Proposition 31. Canonical reference functions on $\mathbb{R}$.
The following functions are reference functions:

1. In a neighbourhood of 0, $(x^\alpha)_{\alpha\in\mathbb{R}}$ and $\left(x^\alpha|\ln(x)|^\beta\right)_{(\alpha,\beta)\in\mathbb{R}^2}$;
2. In a neighbourhood of $+\infty$, $(x^\alpha)_{\alpha\in\mathbb{R}}$, $\left(x^\alpha|\ln(x)|^\beta\right)_{(\alpha,\beta)\in\mathbb{R}^2}$, and $(e^{\alpha x})_{\alpha\in\mathbb{R}}$;
3. In a neighbourhood of $a \in \mathbb{R}$, $(|x-a|^\alpha)_{\alpha\in\mathbb{R}}$.

Exercise 51. ★ Reference functions.
Using the definition of o in a neighbourhood from chapter 1, prove the proposition on canonical reference functions on $\mathbb{R}$.

Definition 74. Expansion on reference functions.
Let f be a function defined in a neighbourhood of $a \in A \subset \mathbb{R}$, and $(\phi_i)_{i\in I}$ a family of reference functions in the neighbourhood of a. f is said to have an asymptotic development on $(\phi_i)_{i\in I}$ of order p in a if there exists $p \in \mathbb{N}$, a p-uplet $(i_1, ..., i_p) \in I^p$, and a p-uplet $(c_1, ..., c_p) \in \mathbb{R}^p$, and $j \in I$ such that

$$\forall k \in [\![1; p]\!], \phi_{i_{k+1}} = o(\phi_{i_k}),$$

and

$$f - \sum_{k=1}^{p} c_{i_k}\phi_{i_k} = O(\phi_j).$$

We also write

$$f = \sum_{k=1}^{p} c_{i_k}\phi_{i_k} + O(\phi_j).$$

Remark 34.

1. This expansion is unique.

2. If the first coefficient is non-zero (i.e. $c_1 \neq 0$) then $c_1\phi_1$ is the dominant part of f in a.

PROPOSITION 32. Integrability of reference functions.

1. $x \mapsto x^\alpha$ is integrable on $[0;\, 1]$ for $\alpha > -1$.
2. $x \mapsto x^\alpha$ is integrable on $[1;\, +\infty[$ for $\alpha < -1$.
3. $x \mapsto |x-a|^\alpha$ is integrable on $[1;\, +\infty[$ for $\alpha > -1$.
4. $x \mapsto e^{\alpha x}$ is integrable on $[1;\, +\infty[$ for $\alpha < 0$.

REMARK 35.
These asymptotics are useful in both Riemann and Lebesgue integration theory, since they ensure the same integrability.

EXERCISE 52. ★ Integrability and domination.

1. Show that $x \mapsto x^3 e^{-x}\chi_{\mathbb{R}\setminus\mathbb{Q}}(x)$ is Lebesgue integrable on $\mathbb{R}^{+\star}$.
2. What can you say about $x \mapsto |x-2|^3 e^{-x}\chi_{\mathbb{R}\setminus\mathbb{Q}}(x)$ on $\mathbb{R}^{+\star}$?
3. What about $x \mapsto x^3 e^{-|x|}\chi_{\mathbb{R}\setminus\mathbb{Q}}(x)$ on $\mathbb{R}$?
4. Let f be polynomial. Can you comment on the Lebesgue integrability of the function $x \mapsto f(x)e^{-x}\chi_{\mathbb{R}\setminus\mathbb{Q}}(x)$ on $\mathbb{R}^{+\star}$?

4.4 Convergence Theorems

4.4.1 Dominated Convergence

THEOREM 38. Dominated convergence.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions. If

1. the sequence converges pointwise to a function f (i.e. the limit exists),
2. there exists a positive integrable function g such that $\forall n \in \mathbb{N}, |f_n| \leq g$,

then f is integrable and

$$\lim_{n\to+\infty} \int f_n \mathrm{d}\mu = \int f \mathrm{d}\mu \qquad \text{and} \qquad \lim_{n\to+\infty} \int |f_n - f| \mathrm{d}\mu = 0.$$

PROOF.
With the pointwise convergence, we know that f is measurable and, since $|f| \leq g$, we also know that f is integrable. Noting that $\forall n \in \mathbb{N}, |f - f_n| \leq 2g$, we can apply Fatou's lemma to the sequence $(2g - |f - f_n|)_{n\in\mathbb{N}}$ (constituted of positive terms), and

$$\begin{aligned}
\int 2g\mathrm{d}\mu = \int \liminf_{n\to+\infty}(2g - |f-f_n|)\mathrm{d}\mu &\leq \liminf_{n\to+\infty} \int 2g - |f - f_n|\mathrm{d}\mu \\
&= \liminf_{n\to+\infty} \left(\int 2g\mathrm{d}\mu - \int |f-f_n|\mathrm{d}\mu\right) \\
&= \int 2g\mathrm{d}\mu - \limsup_{n\to+\infty} \int |f - f_n|\mathrm{d}\mu.
\end{aligned}$$

By substracting, we deduce that

$$\limsup_{n\to+\infty}\int|f-f_n|\mathrm{d}\mu\leq 0,$$

hence

$$\lim_{n\to+\infty}\int|f-f_n|\mathrm{d}\mu=0.$$

Besides, since

$$\left|\int f_n\mathrm{d}\mu-\int f\mathrm{d}\mu\right|\leq\int|f-f_n|\mathrm{d}\mu,$$

we conclude that

$$\lim_{n\to+\infty}\int f_n\mathrm{d}\mu=\int f\mathrm{d}\mu.$$

□

THEOREM 39. Dominated convergence (almost everywhere).
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions. If

1. the sequence converges pointwise almost everywhere in X to a function f (i.e. the limit exists),

2. there exists a positive integrable function g such that $\forall n\in\mathbb{N}, |f_n|\leq g$ almost everywere in X,

then f is integrable and

$$\lim_{n\to+\infty}\int f_n\mathrm{d}\mu=\int f\mathrm{d}\mu \qquad\text{and}\qquad \lim_{n\to+\infty}\int|f_n-f|\mathrm{d}\mu=0.$$

PROOF.
We introduce

$$A=\left\{x\in X\,\middle|\,\lim_{n\to+\infty}f_n\text{ exists}\right\},$$

and

$$\forall n\in\mathbb{N}, B_n=\{x\in X\,||f_n(x)|\leq g(x)\}.$$

We note that we have $\mu(A^c)=0$ and $\forall n\in\mathbb{N},\mu(B_n^c)=0$. Let E be the set defined by

$$E=A\cap\left(\bigcap_{n=0}^{+\infty}B_n\right).$$

We also have $\mu(E^c)=0$ (using set operators and Morgan's laws).

Now, introducing for all $n\in\mathbb{N}$ the functions $\tilde{f}_n=\chi_E f_n$, $\tilde{f}=\chi_E f$, and $\tilde{g}=\chi_E g$, we can apply the dominated convergence theorem to the functions $\tilde{f}_n$, $\tilde{f}$, and $\tilde{g}$. Since

$$\int\tilde{f}_n\mathrm{d}\mu=\int f_n\mathrm{d}\mu \qquad\text{and}\qquad \int\tilde{f}\mathrm{d}\mu=\int f\mathrm{d}\mu,$$

the conclusion follows.

□

Corollary 8. Convergent series.
Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions on X such that

$$\sum_{n=0}^{+\infty}\int |f_n|\mathrm{d}\mu < +\infty,$$

then the series of the $(f_n)_{n\in\mathbb{N}}$ converges (absolutely) almost everywhere in X to an integrable function f and we can write

$$\int f\mathrm{d}\mu = \sum_{n=0}^{+\infty}\int f_n\mathrm{d}\mu.$$

Proof.
Let ϕ the function defined on X by

$$\forall x\in X, \phi(x) = \sum_{n=0}^{+\infty}|f_n(x)|.$$

Using the monotone convergence theorem (countable additivity version), we know that

$$\int \phi\mathrm{d}\mu = \sum_{n=0}^{+\infty}\int |f_n|\mathrm{d}\mu < +\infty.$$

Hence, ϕ is integrable and there exists a measurable subset $E\subset X$ such that $\mu(E^c)=0$ and $\forall x\in E, \phi(x) < +\infty$. Therefore the series based on the f_n converges absolutely and we can define its limit f by

$$\begin{array}{rcl} f : X & \to & \mathbb{R} \\ x & \mapsto & \begin{cases} \sum_{n=0}^{+\infty} f_n(x) & \text{if } x\in E, \\ 0 & \text{if } x\notin E. \end{cases} \end{array}$$

By applying the dominated convergence theorem to the sequence $\left(\sum_{n=0}^{k} f_n\right)_{k\in\mathbb{N}}$, dominated by the characteristic function of the set E, we have

$$\int f\mathrm{d}\mu = \lim_{k\to+\infty}\int\sum_{n=0}^{k} f_n\mathrm{d}\mu = \lim_{k\to+\infty}\sum_{n=0}^{k}\int f_n\mathrm{d}\mu = \sum_{n=0}^{+\infty}\int f_n\mathrm{d}\mu. \qquad \square$$

Remark 36.
These results mean that, under the right assumptions (not too restrictive), the limit and the integral can be switched, and the sum and the integral as well.

Exercise 53. ★★ Application of the dominated convergence theorem.
We consider the function f defined on $\mathbb{R}^+$ by

$$f : x\mapsto e^{-x^2}\cos x.$$

1. Show that f can be written as the limit of a series of functions as follows

$$f(x) = \lim_{n\to+\infty} f_n(x) = \lim_{n\to+\infty}\sum_{k=0}^{n}\frac{(-1)^k}{(2k!)}x^{2k}e^{-x^2}.$$

2. Is f integrable?

3. Using the dominated convergence theorem, show that

$$\int_{\mathbb{R}^+} f(x)\mathrm{d}\mu(x) = \frac{\sqrt{\pi}}{2}e^{-1/4}.$$

4.4.2 Parametrised Integrals

DEFINITION 75. Parametrised integral.
Let X be a measured space and Λ a metric space (e.g. $\mathbb{N}$ or $\mathbb{R}$). Let $f : X \times \Lambda \to \mathbb{C}$ be an integrable function on X. For all $\lambda \in \Lambda$, we define the function $F : \Lambda \to \mathbb{C}$, or *parametrised integral*, by

$$\forall \lambda \in \Lambda, F(\lambda) = \int_X f(x, \lambda) \mathrm{d}\mu(x).$$

THEOREM 40. Continuity.
Let $F : \Lambda \to \mathbb{C}$ be a parametrised function defined from $f : X \times \Lambda \to \mathbb{C}$. If

1. $\forall \lambda \in \Lambda, x \mapsto f(x, \lambda)$ is integrable on X
2. For μ-almost everywhere in X, the function $\lambda \to f(x, \lambda)$ is continuous on Λ
3. There exists $g : X \to \mathbb{R}^+$ integrable such that, for μ-almost everywhere in X and for all $\lambda \in \Lambda$, $|f(x, \lambda)| \leq g(x)$,

then $F : \Lambda \to \mathbb{C}$ in continuous on Λ.

REMARK 37.

1. For the first condition, we actually do not need the integrability of f but only its measurability, since the third condition implies the integrability.
2. In the second condition, if we require that $\lambda \to f(x, \lambda)$ is continuous only on a point λ_0, then we obtain F continuous only on λ_0.

PROOF.
We note that assumption 1 ensures that the function F is actually well defined.
Let $\lambda_0 \in \Lambda$ and $(\lambda_n)_{n \in \mathbb{N}}$ a sequence in Λ converging to λ_0 (using the distance on Λ). Then, thanks to assumption 2 we have that, μ-almost everywhere in X,

$$\lim_{n \to +\infty} f(x, \lambda_n) = f(x, \lambda_0).$$

With assumption 3 we also have that, μ-almost everywhere in X,

$$\forall n \in \mathbb{N}, |f(x, \lambda_n)| \leq g(x),$$

therefore applying the dominated convergence theorem we deduce that

$$\lim_{n \to +\infty} F(\lambda_n) = \lim_{n \to +\infty} \int_X f(x, \lambda_n) \mathrm{d}\mu(x) = \int_X f(x, \lambda_0) \mathrm{d}\mu(x) = F(\lambda_0).$$

This result being true for any arbitrary λ_0, the theorem is proved.

□

THEOREM 41. Derivability.
Let $F : \Lambda \to \mathbb{C}$ be a parametrised function defined from $f : X \times \Lambda \to \mathbb{C}$, with $\Lambda \subset \mathbb{R}$ non-empty. If

1. $\forall \lambda \in \Lambda, x \mapsto f(x, \lambda)$ is integrable on X
2. For μ-almost everywhere in X, the function $\lambda \to f(x, \lambda)$ is derivable on Λ
3. There exists $g : X \to \mathbb{R}^+$ integrable such that, for μ-almost everywhere in X and for all $\lambda \in \Lambda$, $|\partial_\lambda f(x, \lambda)| \leq g(x)$,

then $F : \Lambda \to \mathbb{C}$ is derivable on Λ and

$$F'(\lambda) = \int_X \frac{\partial}{\partial \lambda} f(x, \lambda) \mathrm{d}\mu(x).$$

REMARK 38.

1. $\partial_\lambda f$ is defined μ-almost everywhere in X, and can be set to zero where it is not defined.
2. Even if we only want to prove a pointwise derivability, we still need to assume that condition 3 is fulfilled (i.e. $|\partial_\lambda f|$ bounded on a larger scale).

PROOF.
Let $\lambda_0 \in \Lambda$ and $(\lambda_n)_{n \in \mathbb{N}}$ a sequence in Λ converging to λ_0 with $\forall n \in \mathbb{N}, \lambda_n \neq \lambda_0$. We define for all $x \in X$ and for all $n \in \mathbb{N}$ the sequence of functions

$$h_n(x) = \frac{f(x, \lambda_n) - f(x, \lambda_0)}{\lambda_n - \lambda_0}.$$

Thanks to assumption 2 we know that there exists a set $A \subset X$ such that $\mu(A^c) = 0$ and $\lambda \mapsto f(x, \lambda)$ is derivable for all $x \in A$. Hence, we have

$$\forall x \in A, \lim_{n \to +\infty} h_n(x) = \frac{\partial}{\partial \lambda} f(x, \lambda_0).$$

We define the function h by

$$\begin{array}{rcl} h : X & \to & \mathbb{C} \\ x & \mapsto & \begin{cases} \dfrac{\partial}{\partial \lambda} f(x, \lambda_0) & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \end{array}$$

Using assumption 3, there exists a measurable set $B \subset A$ with $\mu(B^c) = 0$ and

$$\forall n \in \mathbb{N}, \forall x \in B, |h_n(x)| \leq \sup_{\lambda \in \Lambda} \left| \frac{\partial}{\partial \lambda} f(x, \lambda) \right| \leq g(x).$$

Thus, using the dominated convergence theorem, we deduce that

$$\lim_{n \to +\infty} \frac{F(\lambda_n) - F(\lambda)}{\lambda_n - \lambda_0} = \lim_{n \to +\infty} \int_X h_n(x) \mathrm{d}\mu(x) = \int_X h(x) \mathrm{d}\mu(x).$$

Since the sequence $(\lambda_n)_{n \in \mathbb{N}}$ is arbitrary, we finally have F derivable on λ_0 and

$$F'(\lambda) = \int_X \frac{\partial}{\partial \lambda} f(x, \lambda) \mathrm{d}\mu(x).$$

□

REMARK 39.
These theorems are similar to the parametrised integral theorems with the Riemann integral: do we really have something more? The answer is yes, mostly for two reasons.

1. These theorems apply to functions that do not behave nicely with the Riemann integral (see example of the function defined on $[0;\,1] \setminus \mathbb{Q}$).

2. These theorems only require a validity μ-almost everywhere on X, meaning that isolated discontinuities, for example, do not matter.

EXERCISE 54. ★★ Parametrised cosine integral.
Show that

$$\int_{\mathbb{R}^{+\star}} e^{-x^2}\cos(xt)\mathrm{d}\mu(x) = \frac{\sqrt{2}}{\pi}e^{-t^2/4}.$$

4.4.3 Euler Function

DEFINITION 76. Euler function.
For $\alpha \in \mathbb{R}^{+\star}$ we define the Euler function, noted Γ, as the parametrised integral

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x.$$

PROPOSITION 33. Continuity.
For all $\alpha \in \mathbb{R}^{+\star}$, Γ is defined, has values in $\mathbb{R}^{+\star}$, and is continuous.

PROOF.
First of all, we note that for all $\alpha \in \mathbb{R}^{+\star}$, the function $f : x \mapsto x^{\alpha-1}e^{-x}$ is continuous on $\mathbb{R}^{+\star}$, thus measurable, and has positive values. We can define the integral

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x,$$

in the Lebesgue sense, which belongs to $\overline{\mathbb{R}}^+$; but the question is to know if this integral is finite or not for all values of $\alpha \in \mathbb{R}^{+\star}$, i.e. if it is integrable or not.

We cut the integral in two parts, and will consider the integration on $[1;+\infty[$ and on $]0;\,1]$ to conclude. We fix an arbitrary $\alpha \in \mathbb{R}^{+\star}$ to proceed.

Part 1: For all $N \in \mathbb{N}$ we define $f_N : x \mapsto x^{\alpha-1}e^{-x}\chi_{[1;\,N]}(x)$. The sequence $(f_N)_{N\in\mathbb{N}}$ is an increasing sequence of positive measurable functions converging pointwise to f, so using Beppo-Levi's monotone convergence theorem we have

$$\int_1^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x = \lim_{N\to+\infty}\int_1^{+\infty} x^{\alpha-1}e^{-x}\chi_{[1;\,N]}(x)\mathrm{d}x \equiv \lim_{N\to+\infty}\int_1^{N} x^{\alpha-1}e^{-x}\mathrm{d}x.$$

We now write

$$\int_1^{N} x^{\alpha-1}e^{-x}\mathrm{d}x = \int_1^{N} x^{\alpha-1}e^{-x/2}e^{-x/2}\mathrm{d}x.$$

Since $g : x \mapsto x^{\alpha-1}e^{-x/2}$ is a continuous function on $[1;+\infty[$ that has for limit 0 in $+\infty$, g is bounded and there exists a (positive) constant c_α such that $\forall x \in [1;+\infty[, g(x) \leq c_\alpha$ (this constant depends, a priori, on α). Thus

$$\int_1^{N} x^{\alpha-1}e^{-x}\mathrm{d}x \leq \int_1^{N} c_\alpha e^{-x/2}\mathrm{d}x = 2c_\alpha(e^{-1/2} - e^{-N/2}) \leq 2c_\alpha < +\infty.$$

Therefore, taking the limit in N, we have

$$\forall \alpha \in \mathbb{R}^{+\star}, \int_1^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x < +\infty.$$

Part 2: The same reasoning gives us

$$\int_0^1 x^{\alpha-1}e^{-x}\mathrm{d}x = \lim_{\varepsilon\to 0}\int_\varepsilon^1 x^{\alpha-1}e^{-x}\mathrm{d}x.$$

Yet, we have

$$\int_\varepsilon^1 x^{\alpha-1}e^{-x}\mathrm{d}x \le \int_\varepsilon^1 x^{\alpha-1}\mathrm{d}x = \frac{1}{\alpha}(1-\varepsilon^\alpha) \le \frac{1}{\alpha} < +\infty.$$

Then, taking the limit in ε, we have

$$\forall \alpha \in \mathbb{R}^{+\star}, \int_0^1 x^{\alpha-1}e^{-x}\mathrm{d}x < +\infty.$$

Finally, using the linearity of the integral, we conclude that

$$\forall \alpha \in \mathbb{R}^{+\star}, \int_0^{+\infty} x^{\alpha-1}e^{-x}\mathrm{d}x < +\infty.$$

Note that, since $f : x \mapsto x^{\alpha-1}e^{-x}$ is positive, having $\Gamma(\alpha) < +\infty$ for all $\alpha \in \mathbb{R}^{+\star}$ means that f is integrable. This integrability, however, cannot be proved using a domination of f over $\mathbb{R}^+$ at once! To show the continuity of Γ, we thus need to use the two dominations we just found.

Let β and γ such that $0 < \beta < \gamma < +\infty$, and $A = [\beta;\gamma]$. For all $\alpha \in A$, $x \mapsto x^{\alpha-1}e^{-x}$ is integrable on $\mathbb{R}^{+\star}$ and for all $x \in \mathbb{R}^{+\star}$, $\alpha \mapsto x^{\alpha-1}e^{-x}$ is continuous. Moreover, if we define h as

$$h : x \mapsto \begin{cases} c_\gamma e^{-x/2} & \text{if } x \ge 1, \\ x^{\beta-1} & \text{if } x \le 1, \end{cases}$$

then

$$\forall \alpha \in A, \forall x \in \mathbb{R}^{+\star}, 0 \le x^{\alpha-1}e^{-x} \le h(x),$$

and h is integrable on $\mathbb{R}^{+\star}$.

The theorem on continuity of parametrised integrals gives the continuity of Γ on $A = [\beta;\gamma]$ and, since β and γ are arbitrary, we obtain the continuity of Γ on $\mathbb{R}^{+\star}$. $\square$

Proposition 34. Derivatives.
$\Gamma \in \mathcal{C}^\infty(]0;+\infty[)$ and

$$\forall \alpha \in \mathbb{R}^{+\star}, \forall k \in \mathbb{N}, \Gamma^{(k)}(\alpha) = \int_0^{+\infty} (\ln(x))^k x^{\alpha-1}e^{-x}\mathrm{d}x.$$

Exercise 55. ★★★ Derivatives of Γ.
Prove that $\Gamma \in \mathcal{C}^\infty(]0;+\infty[)$ and

$$\forall \alpha \in \mathbb{R}^{+\star}, \forall k \in \mathbb{N}, \Gamma^{(k)}(\alpha) = \int_0^{+\infty} (\ln(x))^k x^{\alpha-1}e^{-x}\mathrm{d}x.$$

Corollary 9. Convexity.
The Γ function is strictly convex.

PROOF.
We have
$$\forall \alpha \in \mathbb{R}^{+\star}, \Gamma''(\alpha) = \int_0^{+\infty} (\ln(x))^2 x^{\alpha-1} e^{-x} \mathrm{d}x > 0.$$
□

PROPOSITION 35. Iterations.

$$\forall \alpha \in \mathbb{R}^{+\star}, \Gamma(\alpha+1) = \alpha\Gamma(\alpha).$$

EXERCISE 56. ★★ Iterations of Γ.
Prove that $\forall \alpha \in \mathbb{R}^{+\star}, \Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.

COROLLARY 10. Iterations on integers.

$$\forall n \in \mathbb{N}, \Gamma(n+1) = n!$$

PROOF.
Can be shown by induction, since it is a consequence of the iteration proposion on $\mathbb{R}^{+\star}$. □

EXERCISE 57. ★★ Values of Γ on half-integers.
Show that $\Gamma(1/2) = \sqrt{\pi}$. What is the value of $\Gamma(3/2)$? And of $\Gamma(n/2)$ for $n \in \mathbb{N}^\star$?

PROPOSITION 36. Asymptotic behaviours.

1. When α goes to zero, we have
$$\Gamma(\alpha) \sim \frac{1}{\alpha}.$$
2. When α goes to $+\infty$, we have
$$\Gamma(\alpha+1) = \sqrt{\alpha} e^{-\alpha} \alpha^\alpha f(\alpha),$$
where
$$f(\alpha) = \int_{-\sqrt{\alpha}}^{+\infty} \left(1 + \frac{x}{\sqrt{\alpha}}\right)^\alpha e^{-x\sqrt{\alpha}} \mathrm{d}x,$$
and
$$\lim_{\alpha \to +\infty} f(\alpha) = \sqrt{2\pi}.$$
Hence, we have
$$\Gamma(\alpha+1) \sim \sqrt{2\pi\alpha}\,\alpha^\alpha e^{-\alpha}.$$

COROLLARY 11. Stirling.
We have an symptotic approximation for large values of $n \in \mathbb{N}$ as

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n}.$$

EXERCISE 58. ★★ Beta function.
For $(a, b) \in (\mathbb{R}^{+\star})^2$, we define the Beta function by

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} \mathrm{d}x.$$

1. Show that $\forall (a,b) \in (\mathbb{R}^{+\star})^2, B(a,b) = B(b,a)$.

2. Show that, for $(a,b) \in (\mathbb{R}^{+\star})^2$, we can write

$$B(a,b) = 2\int_0^{\pi/2} (\sin\theta)^{2a-1} (\cos\theta)^{2b-1} \theta \mathrm{d}\theta.$$

3. Show that $B(1/2, 1/2) = \pi$ and that $B(a,1) = 1/a$.

4. Prove that, for $(a,b) \in (\mathbb{R}^{+\star})^2$, we have

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

5. Use these results to prove that $\forall a \in \mathbb{R}^{+\star}, \Gamma(a+1) = a\Gamma(a)$.

4.5 Comparison to Riemann Integral

Definition 77. Riemman integration of a staircase function.
Let $A \subset \mathbb{R}$ and $(I_k)_{k\in[\![0;\,n]\!]}$ a finite collection of mutually disjoint intervals covering A. Let f be the *staircase function* defined on A by

$$f : x \mapsto \sum_{k=0}^{n} a_k \chi_{I_k},$$

where for all $k \in [\![0;\,n]\!]$, $a_k \in \mathbb{R}$ and χ_{I_k} is the characteristic function of the interval I_k. The *Riemann integral* of f is defined by

$$\int_A f(x)\mathrm{d}x = \sum_{k=0}^{n} a_k l(I_k),$$

with l the length application.

Remark 40.
An important difference with the Lebesgue integration of simple functions is that, in order to define the Riemann integral, we consider mutually disjoint intervals and not simply mutually disjoint sets.

Definition 78. Riemman integration (sequential).
Let f be a function defined on $A \subset \mathbb{R}$, with values in $\mathbb{R}$. f is said to be *Riemann integrable* if there exists two sequences of staircase functions $(f_n)_{n\in\mathbb{N}}$ and $(\phi_n)_{n\in\mathbb{N}}$, with $\forall n \in \mathbb{N}, \phi_n \geq 0$, such that

$$\forall n \in \mathbb{N}, |f - f_n| \leq \phi_n \qquad \text{and} \qquad \lim_{n\to+\infty} \int_A \phi_n(x)\mathrm{d}x = 0.$$

In that case, we define the *Riemann integral* of f by

$$\int_A f(x)\mathrm{d}x = \lim_{n\to+\infty} \int_A f_n(x)\mathrm{d}x.$$

Exercise 59. ★ Riemman integration (sequential).
We consider the definition on Riemann integration (sequential).

1. Explain why the limit $\lim_{n\to+\infty}\int_A f_n(x)\mathrm{d}x$ is correctly defined.

2. Show that the definition of $\int_A f(x)\mathrm{d}x$ does not depend on the choice of the sequences $(f_n)_{n\in\mathbb{N}}$ and $(\phi_n)_{n\in\mathbb{N}}$ (consider two other sequences).

Definition 79. Riemman integration (Darboux's sums).
Let f be a function defined on $A\subset\mathbb{R}$, with values in $\mathbb{R}$. f is said to be *Riemann integrable* if there exists two sequences of staircase functions $(f_n^+)_{n\in\mathbb{N}}$ and $(f_n^-)_{n\in\mathbb{N}}$ such that

$$\forall n\in\mathbb{N}, f_n^+\geq f_{n+1}^+ \quad\text{and}\quad f_n^-\leq f_{n+1}^-,$$

with

$$\forall n\in\mathbb{N}, f_n^-\leq f\leq f_n^+ \qquad\text{and}\qquad \lim_{n\to+\infty}\int_A (f_n^+(x)-f_n^-(x))\mathrm{d}x=0.$$

Then, the *Riemann integral* of f is defined as

$$\int_A f(x)\mathrm{d}x=\lim_{n\to+\infty}\int_A f_n^+(x)\mathrm{d}x=\lim_{n\to+\infty}\int_A f_n^-(x)\mathrm{d}x,$$

where the integrals of f_n^+ and of f_n^- are called Darboux's sums.

Proposition 37. Riemann-Lebesgue equivalence.
Let f be a Riemann integrable function, defined on a mesured space $(\mathbb{R},\mathcal{B}(\mathbb{R}),\mu)$. f is Lebesgue-measurable (for the Borelian tribe), Lebesgue-integrable, and

$$\underbrace{\int f(x)\mathrm{d}x}_{\text{Riemann integral}} = \underbrace{\int f(x)\mathrm{d}\mu(x)}_{\text{Lebesgue integral}}.$$

Proof.
Let f be a Riemann integrable function, defined on $\mathbb{R}$, with values in $\mathbb{R}$. Let $(f_n)_{n\in\mathbb{N}}$ and $(\phi_n)_{n\in\mathbb{N}}$, with $\forall n\in\mathbb{N},\phi_n\geq 0$, be two sequences of staircase functions (defined through a family of mutually disjoint intervals) such that

$$\forall n\in\mathbb{N}, |f-f_n|\leq\phi_n \qquad\text{and}\qquad \lim_{n\to+\infty}\int\phi_n(x)\mathrm{d}x=0.$$

We assume that (since we can always extract a subsequence)

$$\forall n\in\mathbb{N}, \int\phi_n(x)\mathrm{d}x\leq 2^{-n}.$$

The definition of a staircase function proposed here agrees with the definition of simple functions, meaning that $\forall n\in\mathbb{N},\phi_n\in\mathcal{E}^+$ and thus for all $n\in\mathbb{N}$, ϕ_n is Lebesgue integrable and

$$\forall n\in\mathbb{N}, \int\phi_n(x)\mathrm{d}x=\int\phi_n(x)\mathrm{d}\mu(x).$$

Since we assumed that, for all $n\in\mathbb{N}$, the integral of ϕ_n is smaller than 2^{-n}, we know that

$$\sum_{n=0}^{+\infty}\int\phi_n(x)\mathrm{d}\mu(x)<+\infty,$$

so

$$\sum_{n=0}^{+\infty} \phi_n(x)\mathrm{d}x < +\infty,$$

μ-almost everywhere on $\mathbb{R}$. If we define E the subset of $\mathbb{R}$ such that this inequality is satisfied, we have $\mu(E^c) = 0$. In particular, we have a convergence μ-almost everywhere

$$\forall x \in E, \lim_{n\to+\infty} \phi_n(x) = 0,$$

which implies

$$\forall x \in E, \lim_{n\to+\infty} f_n(x) = f(x).$$

Hence, f is Lebesgue measurable: let $a \in \mathbb{R}$, we have

$$\{x \in \mathbb{R} | f(x) > a\} = \{x \in E | f(x) > a\} \cup \{x \in E^c | f(x) > a\} \in \mathcal{B}(\mathbb{R}),$$

that is the union of two Lebesgue measurable sets. In addition, let $N \in \mathbb{N}$ such that f_N is Riemann integrable (such a N exists since the sequential definition of the Riemann integral ensures the convergence of the integral when n goes to $+\infty$) and is, by consequence, Lebesgue integrable. We can write $|f| = |f - f_N + f_N| \leq |f - f_N| + |f_N| \leq |\phi_N| - |f_N|$, and since $|\phi_N|$ and $|f_N|$ are both integrable, $|f|$ is Lebesgue integrable.

Applying the dominated convergence theorem, we have

$$\int f(x)\mathrm{d}\mu(x) = \lim_{n\to+\infty} \int f_n(x)\mathrm{d}\mu(x).$$

Then, since for all $n \in \mathbb{N}$, f_n is a staircase function, its Lebesgue and Riemann integrals are the same, so

$$\lim_{n\to+\infty} \int f_n(x)\mathrm{d}\mu(x) = \lim_{n\to+\infty} \int f_n(x)\mathrm{d}x,$$

and, by definition of the Riemann integral, we have

$$\lim_{n\to+\infty} \int f_n(x)\mathrm{d}x = \int f(x)\mathrm{d}x.$$

Thus, we conclude with

$$\underbrace{\int f(x)\mathrm{d}x}_{\text{Riemann integral}} = \underbrace{\int f(x)\mathrm{d}\mu(x)}_{\text{Lebesgue integral}}.$$

□

PROPOSITION 38. Riemann-Lebesgue equivalence.
Let f be a function defined on a mesured space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. f is Riemann-integrable if and only if

1. f is bounded, and defined on a compact space; and
2. the set E of discontinuity points of f is Lebesgue-negligeable.

EXERCISE 60. ★★★ Riemann-Lebesgue equivalence.
Prove the previous proposition, i.e. for f a function defined on a mesured space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, f is Riemann-integrable if and only if

1. f is bounded, and defined on a compact space; and
2. the set E of discontinuity points of f is Lebesgue-negligeable.

Remark 41.

The Lebesgue integration solves several definition issues we could have with the Riemann integration, but the *generalised integrals* are still not integrable. If we consider the Dirichlet integral

$$\int_0^{+\infty} \frac{\sin t}{t} \mathrm{d}t,$$

then the function $x \mapsto \sin(x)/x$ is continuous on $\mathbb{R}^+$ and therefore is measurable, but

$$\int_0^{+\infty} \frac{|\sin(t)|}{t} \mathrm{d}\mu(t) = +\infty,$$

so it is still non-integrable for Lebesgue. Yet, we can directly compute

$$\lim_{n\to+\infty} \int_0^{n} \frac{\sin t}{t} \mathrm{d}\mu(t) = \frac{\pi}{2}.$$

4.6 Conclusions

Fundamentally, the theory of integration developped by Lebesgue consists in a different point of view on the integral: we could say, to simplify, that Riemann defines the integral thanks to a subdivision of the domain of a function and by measuring in its codomain, while Lebesgue defines the integral thanks to a subdivision of the codomain of a function and by measuring in its domain. When they are both defined, these integrals are the same; however, the set of Lebesgue integrable functions is larger than the set of Riemann integrable functions: for example, the characteristic function of $\mathbb{Q}$, restricted to a compact, has a meaning in the theory of Lebesgue but not in the theory of Riemann. The elementary operations are formally the same as well, but we gain in generality with Lebesgue since the notions of measurability and of almost everywhere save us from dealing with negligeable singularities and generally allows for weaker assumptions in the theorems.

CHAPTER 5

Integration and Product Spaces

In chapter 3, we have defined the Lebesgue measure (outer and inner measure) on $\mathbb{R}$ and on $\mathbb{R}^n$. As discussed in the corresponding section, the Lebesgue measure on $\mathbb{R}^n$ corresponds, in some sense, to the extension of the intuitive notion of volume in $\mathbb{R}^3$ (and higher dimensions as well). Here, we will extend the Lebesgue measure theory to product spaces in order to formulate Lebesgue integration theory on such spaces. Note that, according to the previous definitions, we already have an intuitive sense of what would be the Lebesgue integral of the characteristic function of a set: the measure of the set itself. How can we define it rigorously? How can we integrate functions defined on product spaces?

5.1 Product Spaces and Product Measure

5.1.1 Rectangles

Definition 80. Rectangle.
Let X and Y be two sets, and $A \subset X$ and $B \subset Y$. A *rectangle* R is a subset of $X \times Y$ formed by the Cartesian product of A and B, i.e. $R = A \times B$. A and B are called the *sides* of the rectangle.

Remark 42.
Note that, in the previous definition, we did not specify the nature of the sets A and B that form the rectangle. The intuitive notion of a rectangle is a bi-dimensional set that can be defined, in $\mathbb{R}^2$, by the product of two intervals, i.e. a cuboid in $\mathbb{R}^2$. Here, and throughout this chapter, a rectangle is a set in any product space (not only $\mathbb{R} \times \mathbb{R}$) formed by the product of two sets that are not necessarily intervals, and not necessarily connex either.

Theorem 42. Empty rectangle.
A rectangle is empty if and only if one of its sides is empty.

Proof.
Let $R = A \times B$ be a rectangle.
First, let assume $R \neq \emptyset$. Then, we can pick $(a, b) \in R$ with, by definition, $a \in A$ and $b \in B$, so $A \neq \emptyset$ and $B \neq \emptyset$.

Now, let assume $A \neq \emptyset$ and $B \neq \emptyset$. Then, we can pick $a \in A$ and $b \in B$ and we have, by definition of the Cartesian product, $(a,b) \in A \times B$, so $R \neq \emptyset$. $\square$

Theorem 43. Sub-rectangles.
If $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ are non-empty rectangles, then $R_1 \subset R_2$ if and only if $A_1 \subset A_2$ and $B_1 \subset B_2$.

Proof.
With the notations of the theorem, if $A_1 \subset A_2$ and $B_1 \subset B_2$ then $R_1 \subset R_2$ is obvious: any $(a,b) \in A_1 \times B_1$ is also in $A_2 \times B_2$.

Now, we assume that $R_1 \subset R_2$. Let $(a,b) \in A_1 \times B_1$, and we suppose that there exists $c \in A_1$ such that $c \notin A_2$. Then, $(c,b) \in A_1 \times B_1$ and $(c,b) \notin A_2 \times B_2$. Yet, this is contradictory since $A_1 \times B_1 \subset A_2 \times B_2$, and we conclude that $A_1 \subset A_2$. The same reasoning yields $B_1 \subset B_2$. $\square$

Theorem 44. Equal rectangles.
If $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ are non-empty rectangles, then $R_1 = R_2$ if and only if $A_1 = A_2$ and $B_1 = B_2$.

Proof.
This comes from the previous theorem: $R_1 = R_2$ if and only if $R_1 \subset R_2$ and $R_2 \subset R_1$. On one side, we have $R_1 \subset R_2$ if and only if $A_1 \subset A_2$ and $B_1 \subset B_2$. On the other side, we have $R_2 \subset R_1$ if and only if $A_2 \subset A_1$ and $B_2 \subset B_1$. The conclusion $A_1 = A_2$ and $B_1 = B_2$ follows. $\square$

Theorem 45. Disjoint union of rectangles.
Let $R = A \times B$, $R_1 = A_1 \times B_1$, and $R_2 = A_2 \times B_2$ be non-empty rectangles. Then, a necessary and sufficient condition that R is the disjoint union of R_1 and R_2 is that either A is the disjoint union of A_1 and A_2 with $B = B_1 = B_2$, or B is the disjoint union of B_1 and B_2 with $A = A_1 = A_2$.

Proof.
Showing that the condition is necessary:
We assume that R is the disjoint union of R_1 and R_2, hence we have $R_1 \subset R$ and $R_2 \subset R$. Using the previous results, we deduce that $A_1 \subset A$ and $A_2 \subset A$ so $A_1 \cup A_2 \subset A$ (same for $B_1 \cup B_2 \subset B$). Since we have

$$R_1 \cup R_2 \subset (A_1 \cup A_2) \times (B_1 \cup B_2),$$

it follows that $A \subset A_1 \cup A_2$ and $B \subset B_1 \cup B_2$. Thus, $A = A_1 \cup A_2$ and $B = B_1 \cup B_2$.

We can proceed similarly for the intersection, and using that R_1 and R_2 are disjoint we find that

$$(A_1 \cap A_2) \times (B_1 \cap B_2) \subset R_1 \cap R_2 = \emptyset.$$

This means that the rectangle $(A_1 \cap A_2) \times (B_1 \cap B_2)$ is empty, so one of its sides (at least) is empty, either $A_1 \cap A_2$ or $B_1 \cap B_2$.

We suppose that $A_1 \cap A_2 = \emptyset$ (i.e. they are disjoint). Let assume that there exists $b \in B \setminus B_1$. Then, for any $a \in A_1$ we have $(a,b) \in R$ but $(a,b) \notin R_1$ (because $b \notin B_1$) and $(a,b) \notin R_2$ (because $a \notin A_2$): this is contradictory since $R = R_1 \cup R_2$, so $B \setminus B_1 = \emptyset$. With the same reasoning on $B \setminus B_2$ we show that $B = B_1 = B_2$. Similary, we show that if $B_1 \cap B_2 = \emptyset$ we have $A = A_1 = A_2$.

Showing that the condition is sufficient:

If A is the disjoint union of A_1 and A_2, and $B = B_1 = B_2$, then: $A_1 \subset A$, $A_2 \subset A$, $B_1 \subset B$, and $B_2 \subset B$, so $R_1 \cup R_2 \subset R$. Moreover, if $(a, b) \in R$, then either $(a, b) \in R_1$ or $(a, b) \in R_2$, since either $a \in A_1$ or $a \in A_2$. Therefore, E is the disjoint union of R_1 and R_2.

□

Theorem 46. Ring of rectangles.
If $\mathcal{S}$ and $\mathcal{T}$ are rings of subsets of X and Y, respectively, then the set $\mathcal{R}$ of all finite, disjoint unions of rectangles of the form $A \times B$ with $A \in \mathcal{S}$ and $B \in \mathcal{T}$, is a ring of sets.

Proof.

In order to prove that $\mathcal{R}$ is a ring of sets, we have to prove that $\mathcal{R}$ is not empty, and that it is stable by union and by difference.

1. Since $\mathcal{S}$ and $\mathcal{T}$ are ring of sets they are not empty, thus we can find $A \in \mathcal{S}$ and $B \in \mathcal{T}$, and $A \times B \in \mathcal{R}$ so $\mathcal{R} \neq \emptyset$.

2. Let $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ be two elements of $\mathcal{R}$. We note $R = R_1 \cup R_2$. R is a subset of $X \times Y$ so we can write $R = A \times B$ with $A \in X$ and $B \in Y$, but at this stage we do not know whether A and B are in $\mathcal{S}$ and $\mathcal{T}$ or not. Then, using the previous theorems, we know that $A_1 \subset A$ and $A_2 \subset A$ so $A_1 \cup A_2 \subset A$, and similarly $B_1 \cup B_2 \subset B$. Then, writing
$$R_1 \cup R_2 \subset (A_1 \cup A_2) \times (B_1 \cup B_2),$$
we deduce that $A \subset A_1 \cup A_2$ and $B \subset B_1 \cup B_2$. Thus, we have $A = A_1 \cup A_2 \in \mathcal{S}$ and $B = B_1 \cup B_2 \in \mathcal{T}$ since $\mathcal{S}$ and $\mathcal{T}$ are rings (stable by union). Hence, R is an element of $\mathcal{R}$, and $\mathcal{R}$ is stable by union.

3. Let $R_1 = A_1 \times B_1$ and $R_2 = A_2 \times B_2$ be two elements of $\mathcal{R}$. Then
$$R_1 \setminus R_2 = [(A_1 \cap A_2) \times (B_1 \setminus B_2)] \cup [(A_1 \setminus A_2) \times B_1] .$$
Yet, $\mathcal{S}$ and $\mathcal{T}$ are ring of sets so they are stable by intersection and difference, and $A_1 \cap A_2 \in \mathcal{S}$, $A_1 \setminus A_2 \in \mathcal{S}$, and $B_1 \setminus B_2 \in \mathcal{T}$. Thus, $(A_1 \cap A_2) \times (B_1 \setminus B_2) \in \mathcal{R}$ and $(A_1 \setminus A_2) \times B_1 \in \mathcal{R}$, and we just showed that $\mathcal{R}$ is stable by union.

We conclude that $\mathcal{R}$ is a ring of sets.

□

Definition 81. Cartesian product of σ-rings.
Let X and Y be two sets,and $\mathcal{S}$ and $\mathcal{T}$ be σ-rings of sets of X and Y, respectively. Then, we note $\mathcal{S} \times \mathcal{T}$ the σ-ring of sets of $X \times Y$, generated by the class of all sets of the form $A \times B$ where $A \in \mathcal{S}$ and $B \in \mathcal{T}$.

Theorem 47. Product of measurable spaces.
If $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ are measurable spaces, then $(X \times Y, \mathcal{S} \times \mathcal{T})$ is a measurable space. It is the Cartesian product of two measurable spaces.

PROOF.
If $(x, y) \in X \times Y$, then we can find two sets A and B such that $x \in A \in \mathcal{S}$ and $y \in B \in \mathcal{T}$ Hence, $(x, y) \in A \times B \in \mathcal{S} \times \mathcal{T}$.

$\square$

DEFINITION 82. Measurable rectangle.
Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be two measurable spaces, and $(X \times Y, \mathcal{S} \times \mathcal{T})$ their Cartesian product. Rectangles of $\mathcal{S} \times \mathcal{T}$ are called *measurable rectangles*. Alternatively, this means that $A \times B$ is a measurable rectangle if $A \in \mathcal{S}$ and $B \in \mathcal{T}$.

5.1.2 Sections

DEFINITION 83. Section of a set.
Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be two measurable spaces, and $(X \times Y, \mathcal{S} \times \mathcal{T})$ their Cartesian product. Let E be a subset of $X \times Y$ and $x \in X$. We call *section of E* or *section of E determined by x* and write E_x the set

$$E_x = \{y \in Y | (x, y) \in E\}.$$

Similarly, for $y \in Y$, we call *section of E* or *section of E determined by y* and write E^y the set

$$E^y = \{x \in X | (x, y) \in E\}.$$

REMARK 43.
With the same notations, we note that E_x is not a subset of $X \times Y$, but a subset of Y (and similarly for E^y, which is a subset of X). Sometimes, we only need the fact that the section is determined by a fixed point and is a subset of the other space; if so, we will use the terminology *X-section* for a section E_x determined by a point of X, and *Y-section* for a section E^y determined by a point of Y.

DEFINITION 84. Section of an application.
Let $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ be two measurable spaces, and $(X \times Y, \mathcal{S} \times \mathcal{T})$ their Cartesian product. Let E be a subset of $X \times Y$ and f an application defined on E. If x is a point of X, we call *section of f*, or *section of f determined by x*, or X-section of f the application defined on E_x by

$$f_x : y \mapsto f(x, y).$$

Similarly, if y is a point of Y, we call *section of f*, or *section of f determined by y*, or Y-section of f the application defined on E_y by

$$f^y : x \mapsto f(x, y).$$

THEOREM 48. Measurability of the section of a set.
Every section of a measurable set is a measurable set.

PROOF.
We take $(X, \mathcal{S})$ and $(Y, \mathcal{T})$ two measurable spaces. Let $\mathcal{M}$ be the set of all subsets of $X \times Y$ with the property that each of their sections is measurable. Let $E = A \times B$ be a measurable set of $X \times Y$: this is a measurable rectangle, so A and B are measurable.

We now consider $x \in X$ and the section E_x. Using the definition of the section of a set, we can show that either $E_x = \emptyset$ or $E_x = B$. Hence, E_x is measurable. Similarly, we obtain E^y measurable, and finally $E \in \mathcal{M}$.

Then, it is easy to show that $\mathcal{M}$ is a σ-ring of sets, so $\mathcal{S} \times \mathcal{T} \subset \mathcal{M}$, and the proof is complete. □

THEOREM 49. Measurability of the section of a function.
Every section of a measurable function is a measurable function.

PROOF.
Let f is a measurable function on $X \times Y \subset \mathbb{R} \times \mathbb{R}$, $x \in X$, and $a \in \mathbb{R}$. We consider $N = \{y \in Y | f_x(y) > a\}$. Since f is measurable, by definition, $M = \{(x, y) \in X \times Y | f(x, y) > a\}$ is measurable. Using the previous theorem, we know that any section of M is also measurable and in particular $M_x = \{y \in Y | (x, y) \in M\}$ is measurable. Yet, we can re-write $M_x = \{y \in Y | f(x, y) > a\} = \{y \in Y | f_x(y) > a\} = N$, so N and therefore f_x is measurable. Similarly, we show that f^y is measurable. □

REMARK 44.
This has to be added to the list of stable operations for the class of measurable functions: again, when a function is measurable, most transformations keep the function measurable.

EXAMPLE 13. Let χ be the characteristic function of a set $E \subset X \times Y$. Then χ_x and χ^y are the characteristic functions of E_x and E^y, respectively. In particular, if $E = A \times B$ is a rectangle, then $\forall (x, y) \in E, \chi(x, y) = \chi_A(x)\chi_B(y)$.

EXERCISE 61. ★ Section of a simple function.
Using the previous example, show that every section of a simple function is a simple function.

5.1.3 Product Tribes

DEFINITION 85. Product tribe.
Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. The *product tribe* $\mathcal{M} \otimes \mathcal{N}$ is the tribe generated by the rectangles $A \times B$ where $A \in X$ and $B \in Y$.

REMARK 45.

1. In general, the set of all rectangles $A \times B$ where $A \in X$ and $B \in Y$ is not a tribe, which is why we need to define the product tribe as generated by them and not as the set of possible rectangles.

2. The product tribe gives to the product space $X \times Y$ the structure of measurable space.

DEFINITION 86. Projections.
Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. The projections π_X and π_Y are the two applications defined on $X \times Y$ such that

$$\pi_X : (x, y) \mapsto x \qquad \text{and} \qquad \pi_Y : (x, y) \mapsto y.$$

PROPOSITION 39. Product tribe and projections.
Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be two measurable spaces. The product tribe $\mathcal{M} \otimes \mathcal{N}$ is the smallest tribe for which π_X and π_Y are measurable.

Proposition 40. Associativity.
Let $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ be three tribes, then

$$(\mathcal{T}_1 \otimes \mathcal{T}_2) \otimes \mathcal{T}_3 = \mathcal{T}_1 \otimes (\mathcal{T}_2 \otimes \mathcal{T}_3) = \mathcal{T}_1 \otimes \mathcal{T}_2 \otimes \mathcal{T}_3.$$

Exercise 62. ★★ Associativity.
Prove the associativity of the product on tribes.

Proposition 41. Let X and Y be two subsets of $\mathbb{R}$, and $\mathcal{B}(X)$ and $\mathcal{B}(Y)$ their Borelian tribes. Then $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$.

Remark 46.
If X and Y are only topological spaces with their borelian tribes, then the previous proposition is not true. We would have, in general, $\mathcal{B}(X) \otimes \mathcal{B}(Y) \subset \mathcal{B}(X \times Y)$. If X and Y are separable metric spaces, then we can conclude $\mathcal{B}(X) \otimes \mathcal{B}(Y) = \mathcal{B}(X \times Y)$. In our case, we are interested in $\mathbb{R}$ (and $\mathbb{R}^n$), and these are separable metric spaces.

Proposition 42. Measurability of the section of a function.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two measurable spaces and f a function defined on $X \times Y$. If f is measurable for the product tribe $\mathcal{M} \otimes \mathcal{N}$, then for any $(x, y) \in X \times Y$, the sections of f, f_x and f^y, are measurable.

Proposition 43. Sub-spaces.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two measurable spaces, and $E \subset X \times Y$. If $E \in \mathcal{M} \otimes \mathcal{N}$, then for any $(x, y) \in X \times Y$, $E_x \in \mathcal{N}$ and $E^y \in \mathcal{M}$.

Remark 47.
These are the extension of what we have done for rectangles and to tribes and tribe product.

5.1.4 Product Measure

Definition 87. σ-finite space.
Let $(X, \mathcal{T}, \mu)$ be a measured space. We say that it is *σ-finite* if there exists an increasing sequence of sets $(E_n)_{n\in\mathbb{N}}$ such that

$$X = \bigcup_{n=0}^{+\infty} E_n \qquad \text{and} \qquad \forall n \in \mathbb{N}, \mu(E_n) < +\infty.$$

Theorem 50. Measurability of measured section.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and E be a measurable subset of $X \times Y$. The functions f and g defined on X and Y, respectively, by

$$f : x \mapsto \nu(E_x) \qquad \text{and} \qquad g : y \mapsto \mu(E^y),$$

are non-negative measurable functions such that

$$\int f \mathrm{d}\mu = \int g \mathrm{d}\nu.$$

Theorem 51. Product measure.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces. Then

1. There exists a unique measure λ on $(X \times Y, \mathcal{M} \otimes \mathcal{N})$ such that
$$\forall A \in \mathcal{M}, \forall B \in \mathcal{N}, \lambda(A \times B) = \mu(A)\nu(B).$$
2. For all $E \in \mathcal{M} \otimes \mathcal{N}$,
$$\lambda(E) = \int_X \nu(E_x)\mathrm{d}\mu(x) = \int_Y \mu(E_y)\mathrm{d}\nu(y).$$

Definition 88. Product measure.
With the notations of the previous theorem, λ is called the *product measure* of the measures μ and ν, and we write $\lambda = \mu \otimes \nu$.

5.2 Fubini's Theorems

5.2.1 Double Integral and Iterated Integrals

Definition 89. Double integral.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and f an integrable function defined on $X \times Y$. We note $\lambda = \mu \otimes \nu$ the product measure. The *double integral* of f over $X \times Y$ is the quantity
$$\int f \mathrm{d}\lambda = \int f \mathrm{d}(\mu \otimes \nu),$$
also written
$$\int f(x, y)\mathrm{d}\lambda(x, y) = \int f(x, y)\mathrm{d}(\mu \otimes \nu)(x, y).$$

Definition 90. Iterated integrals.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and f an integrable function defined on $X \times Y$. The *iterated integrals* of f over $X \times Y$ are the quantity
$$\iint f \mathrm{d}\mu \mathrm{d}\nu = \int \left(\int f \mathrm{d}\mu \right) \mathrm{d}\nu = \int_Y \left(\int_X f(x, y)\mathrm{d}\mu(x) \right) \mathrm{d}\nu(y),$$
and
$$\iint f \mathrm{d}\nu \mathrm{d}\mu = \int \left(\int f \mathrm{d}\nu \right) \mathrm{d}\mu = \int_X \left(\int_Y f(x, y)\mathrm{d}\nu(y) \right) \mathrm{d}\mu(x).$$

Remark 48.

1. The double integral and the iterated integrals are not the same quantities! As we defined it, the double integral is an integral over a product space $(X \times Y)$, whereas the iterated integrals are the integrals over two spaces (X and Y).
2. Similarly, the two iterated integrals we defined are not the same in general!
3. The point of the Fubini's theorems we will prove is to tell that (1) the two iterated integrals are the same and (2) they correspond to the double integral.

Proposition 44. Negligeable subset of $X \times Y$.
A necessary and sufficient condition that a measurable subset E of $X \times Y$ has its measure equal to zero is that almost every X-section (or almost every Y-section) has measure zero. E will be said negligeable.

Proof.
Using the definition of the product measure, we can write

$$\lambda(E) = \int \nu(E_x)\mathrm{d}\mu(x) = \int \mu(E^y)\mathrm{d}\nu(y).$$

Thus, if $\lambda(E) = 0$, at least one of these two integrals has to be zero. Since the integrands are non-negative (measures), this means that either $\nu(E_x) = 0$ or $\mu(E^y) = 0$. On the other hand, if either of the integrands is zero we directly have that $\lambda(E) = 0$. □

Theorem 52. Fubini-Tonelli (positive functions).
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and $f : X \times Y \to \mathbb{R}^+$ a measurable function for the product tribe $\mathcal{M} \otimes \mathcal{N}$. Then

1. The functions I_Y and I_X defined on X and Y, respectively, by

$$I_Y : x \mapsto \int_Y f(x,y)\mathrm{d}\nu(y) \qquad \text{and} \qquad I_X : y \mapsto \int_X f(x,y)\mathrm{d}\mu(x),$$

are measurable.

2. The following equality is satisfied

$$\int_{X\times Y} f\mathrm{d}(\mu\otimes\nu) = \int_X \left(\int_Y f(x,y)\mathrm{d}\nu(y)\right)\mathrm{d}\mu(x) = \int_Y \left(\int_X f(x,y)\mathrm{d}\mu(x)\right)\mathrm{d}\nu(y) \in \mathbb{R}^+.$$

Proof.
We proceed by steps.

1. Let $E \in \mathcal{M} \otimes \mathcal{N}$, and $f = \chi_E$. Then, we have

$$\forall x \in X, I_Y : x \mapsto \int_Y f(x,y)\mathrm{d}\nu(y) = \nu(E_x),$$

and

$$\forall y \in Y, I_X : y \mapsto \int_X f(x,y)\mathrm{d}\mu(x) = \mu(E^y)$$

and the theorem is true by definition of the product measure: I_Y and I_X are measurable, and the double integral is equal to the two iterated integrals (and is equal to $\lambda(E)$).

2. Using the linearity of the Lebesgue integral, the theorem is still true for simple functions.

3. Now, let consider $f : X \times Y \to \mathbb{R}^+$ a measurable function. There exists an increasing sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ converging to f for all $x \in X$ and all $y \in Y$. Using the monotone convergence theorem, we have

$$\int_Y f(x,y)\mathrm{d}\nu(y) = \lim_{n\to+\infty} \int_Y f_n(x,y)\mathrm{d}\nu(y),$$

so I_Y and I_X are measurable. Moreover, the equality on double and interated integrals is true for all of the f_n functions, so thanks to the increasing limit and applying twice the monotone convergence theorem we prove that the equality is also true for f.

□

THEOREM 53. Fubini-Lebesgue (integrable functions).
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and f a function defined on $X \times Y$, integrable for the product tribe $\mathcal{M} \otimes \mathcal{N}$ and product measure $\mu \otimes \nu$. Then

1. For almost every $x \in X$, the function $y \mapsto f(x, y)$ is integrable on Y.
   For almost every $y \in Y$, the function $x \mapsto f(x, y)$ is integrable on X.
2. $I_Y : x \mapsto \int_Y f(x, y) \mathrm{d}\nu(y)$, defined μ-almost everywhere, is integrable on X.
   $I_X : y \mapsto \int_X f(x, y) \mathrm{d}\mu(x)$, defined ν-almost everywhere, is integrable on Y.
3. The following equality is satisfied

$$\int_{X \times Y} f \mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f(x, y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int_Y \left(\int_X f(x, y) \mathrm{d}\mu(x) \right) \mathrm{d}\nu(y).$$

PROOF.
We consider $f : X \times Y \to \mathbb{R}$, integrable.

1. The X and Y-sections of f are measurable, so for all $x \in X$ the function $y \mapsto f(x, y)$ is measurable. Therefore, applying Fubini-Tonelli's theorem to the function $|f|$, we find that
$$\int_X \left(\int_Y |f(x, y)| \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) = \int_{X \times Y} |f| \mathrm{d}(\mu \otimes \nu) < +\infty.$$
Therefore, for almost every $x \in X$, we have
$$\int_Y |f(x, y)| \mathrm{d}\nu(y) < +\infty,$$
and the function $y \mapsto f(x, y)$ is integrable on Y. Similarly, for almost every $y \in Y$, we have
$$\int_X |f(x, y)| \mathrm{d}\mu(x) < +\infty,$$
and the function $x \mapsto f(x, y)$ is integrable on X.

2. Now, we write $f = f^+ - f^-$ and we define the function g on X by
$$g : x \mapsto \begin{cases} \int_Y f^+(x, y) \mathrm{d}\nu(y) - \int_Y f^-(x, y) \mathrm{d}\nu(y) & \text{if } \int_Y |f(x, y)| \mathrm{d}\nu(y) < +\infty, \\ 0 & \text{otherwise,} \end{cases}$$
Then, applying Fubini-Tonelli to f^+ and f^- yields g measurable and
$$\int_X |g| \mathrm{d}\mu \leq \int_X \left(\int_Y |f(x, y)| \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x) < +\infty,$$
so g is integrable, and the second proposition of the theorem is true.

3. Finally, using Fubini-Tonelli's theorem, we have
$$\int_{X \times Y} f^+ \mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f^+(x, y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x),$$
and
$$\int_{X \times Y} f^- \mathrm{d}(\mu \otimes \nu) = \int_X \left(\int_Y f^-(x, y) \mathrm{d}\nu(y) \right) \mathrm{d}\mu(x).$$
The difference of these two integrals leads to the equality of the double and iterated integrals.

□

Remark 49.
Fubini-Lebesgue's theorem applies if the integrable function has values in $\mathbb{R}$ or in $\mathbb{C}$.

Corollary 12. Fubini and factorisation.
Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be two σ-finite measured spaces, and f a function defined on $X \times Y$ that satisfies the assumptions of either Fubini-Tonelli or Fubini-Lebesgue. If f can be factorised, i.e. if $f = hg$ with h a function defined on X and g a function defined on Y, then

$$\int_{X\times Y} f \mathrm{d}(\mu \otimes \nu) = \left(\int_X g \mathrm{d}\mu\right)\left(\int_Y h \mathrm{d}\nu\right).$$

Exercise 63. ★ Cavalieri's principle (short version).
Let E and F be measurable subsets of $X \times Y$ for the product measure $\lambda = \mu \otimes \nu$, such that $\nu(E_x) = \nu(F_x)$ for almost every $x \in X$. Show that $\lambda(E) = \lambda(F)$.

Remark 50.
Why do we need two Fubini's theorems?

Usually, we would like to apply Fubini-Lebesgue's theorem to a function f, but in order to do so we need to be sure that the function f is integrable. Fubini-Tonelli's theorem is the right tool to prove this condition.

Example 14. Direct application of Fubini's theorems.
We consider the function $f : \mathbb{R}^{+\star} \times \mathbb{R}^{+\star} \to \mathbb{R}$ defined by

$$f : (x, y) \mapsto \frac{e^{-y} \sin(xy)}{x\sqrt{1 + x^2}}.$$

Here, we consider $\mathbb{R}^{+\star}$ with its Borelian tribe and the Lebesgue measure μ; for simplicity, we will use the notation $\mathrm{d}\mu(x) = \mathrm{d}x$. We will show that f is integrable and we will compute its double integral.

1. First, we note that f is continuous on $(\mathbb{R}^{+\star})^2$, thus measurable. Since f can have positive and negative values, we will use Fubini-Lebesgue's theorem, and we have to show that f is integrable in order to apply it. To prove the integrability, we will use Fubini-Tonelli's theorem with $|f|$, which is a positive measurable function.

2. Showing that f is integrable: to do so, we need to prove that the integral of $|f|$ is finite. Since $|f|$ is a positive measurable function, we can apply Fubini-Tonelli's theorem and directly bound the iterated integral.
   We recall that $\forall x \in \mathbb{R}, |\sin(x)| \leq \min(|x|, 1) \leq |x|$.

   (a) For $x \in]0; 1]$, we have
   $$|f(x, y)| \leq \frac{e^{-y}xy}{x\sqrt{1 + x^2}} \leq \frac{e^{-y}y}{\sqrt{1 + x^2}},$$
   thus
   $$\int_0^{+\infty} |f(x, y)|\mathrm{d}y \leq \frac{1}{\sqrt{1 + x^2}}.$$

   (b) For $x \in [1; +\infty[$, we have
   $$|f(x, y)| \leq \frac{e^{-y}}{x\sqrt{1 + x^2}},$$
   thus
   $$\int_0^{+\infty} |f(x, y)|\mathrm{d}y \leq \frac{1}{x\sqrt{1 + x^2}}.$$

Therefore, we write

$$\int_0^{+\infty} |f(x,y)|\mathrm{d}y \leq \chi_{]0;\,1[}(x)\frac{1}{\sqrt{1+x^2}} + \chi_{[1;\,+\infty[}(x)\frac{1}{x\sqrt{1+x^2}}.$$

Using the reference integrable functions (power laws), we see that this function of x is integrable. Actually, we have

$$\int_0^{+\infty}\left(\int_0^{+\infty}|f(x,y)|\mathrm{d}y\right)\mathrm{d}x \leq \int_0^1 \mathrm{d}x + \int_1^{+\infty}\frac{1}{x^2}\mathrm{d}x = 2 < +\infty.$$

Hence, f is integrable.

3. Computing the double integral: f being integrable, Fubini-Lebesgue's theorem apply and, noting $\lambda = \mu \otimes \mu$, we can write

$$\begin{aligned}
\int_{\mathbb{R}^{+\star}\times\mathbb{R}^{+\star}} f\mathrm{d}\lambda &= \int_{\mathbb{R}^{+\star}\times\mathbb{R}^{+\star}} \frac{e^{-y}\sin(xy)}{x\sqrt{1+x^2}}\mathrm{d}\lambda(x,y) \\
&= \int_0^{+\infty}\frac{1}{x\sqrt{1+x^2}}\left(\int_0^{+\infty} e^{-y}\sin(xy)\mathrm{d}y\right)\mathrm{d}x.
\end{aligned}$$

Now, we can compute the integral on y as

$$\begin{aligned}
\int_0^{+\infty} e^{-y}\sin(xy)\mathrm{d}y &= \Im\left(\int_0^{+\infty} e^{-y}e^{ixy}\mathrm{d}y\right) \\
&= \Im\left(\frac{1}{1-ix}\right) = \frac{x}{1+x^2}.
\end{aligned}$$

Thus

$$\begin{aligned}
\int_{\mathbb{R}^{+\star}\times\mathbb{R}^{+\star}} f\mathrm{d}\lambda &= \int_0^{+\infty}\frac{1}{x\sqrt{1+x^2}}\frac{x}{1+x^2}\mathrm{d}x \\
&= \int_0^{+\infty}\frac{1}{(1+x^2)^{3/2}}\mathrm{d}x \\
&= \int_0^{+\infty}\frac{1}{(1+\mathrm{sh}^2(t))^{3/2}}\mathrm{ch}(t)\mathrm{d}t \qquad x \to \mathrm{sh}(t) \\
&= \int_0^{+\infty}\frac{1}{\mathrm{ch}^2(t)}\mathrm{d}t \\
&= [\mathrm{th}(t)]_0^{+\infty} \\
&= 1.
\end{aligned}$$

Example 15. Importance of the integrability condition in Fubini-Lebesgue.
We consider the function $f :]0;\,1[\times]0;\,1[\to \mathbb{R}$ defined by

$$f : (x,y) \to \frac{x^2-y^2}{(x^2+y^2)^2}.$$

We note that, for a fixed $x \in]0;\,1[$, we have

$$\int_0^1 f(x,y)\mathrm{d}y = \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\mathrm{d}y = \left[\frac{y}{x^2+y^2}\right]_0^1 = \frac{1}{1+x^2}.$$

Therefore, if we compute the iterated integrals, we find that

$$\int_0^1\left(\int_0^1 f(x,y)\mathrm{d}y\right)\mathrm{d}x = \int_0^1 \frac{1}{1+x^2}\mathrm{d}x = [\arctan(x)]_0^1 = \frac{\pi}{4},$$

and, since f is antisymmetric when permuting x and y,

$$\int_0^1 \left(\int_0^1 f(x,y)\mathrm{d}x\right)\mathrm{d}y = \int_0^1 -\frac{1}{1+y^2}\mathrm{d}y = [-\arctan(y)]_0^1 = -\frac{\pi}{4}.$$

The two iterated integrals are different! Here, Fubini-Lebesgue's theorem does not apply, because f is not an integrable function. One should always check the integrability of f before applying Fubini-Lebesgue's theorem!

5.2.2 Multiple Integrals

Definition 91. Multiple integrals.
Let $N \in \mathbb{N}$ and $(A_n)_{n\in[\![0;N]\!]}$ a family of subsets of $\mathbb{R}$, to which we associate the measures $(\mu_n)_{n\in[\![0;N]\!]}$ and the Borelian tribes $(\mathcal{B}(A_n))_{n\in[\![0;N]\!]}$. We define the measured space $(E,\mathcal{B}(E),\mu)$ by

$$(E,\mathcal{B}(E),\mu) = \left(\prod_{n=0}^N A_n, \bigotimes_{n=0}^N \mathcal{B}(A_n), \bigotimes_{n=0}^N \mu_n\right),$$

where we define the (finite) Cartesian product

$$\prod_{n=0}^N A_n = A_0 \times A_2 \times \ldots \times A_N,$$

the (finite) product tribe

$$\bigotimes_{n=0}^N \mathcal{B}(A_n) = \mathcal{B}(A_0) \otimes \mathcal{B}(A_1) \otimes \ldots \otimes \mathcal{B}(A_N),$$

and the (finite) product measure

$$\bigotimes_{n=0}^N \mu_n = \mu_0 \otimes \mu_1 \otimes \ldots \otimes \mu_N.$$

Remark 51.

As previously, it is important to note that $\mathcal{B}(E) \neq \bigotimes_{n=0}^N \mathcal{B}(A_n)$ in general. It is true, however, in the case of subsets of $\mathbb{R}$.

Definition 92. Multiple integrals.
Let $N \in \mathbb{N}$ and $(A_n)_{n\in[\![0;N]\!]}$ a family of subsets of $\mathbb{R}$, to which we associate the measures $(\mu_n)_{n\in[\![0;N]\!]}$ and the Borelian tribes $(\mathcal{B}(A_n))_{n\in[\![0;N]\!]}$. We consider the measured space $(E,\mathcal{B}(E),\mu)$ as defined previously. Then, if f is an integrable function on E, we define its multiple integral by

$$\int_E f\mathrm{d}\mu = \int_E f\mathrm{d}\left(\bigotimes_{n=0}^N \mu_n\right).$$

Definition 93. Iterated integrals.
Let $N \in \mathbb{N}$ and $(A_n)_{n\in[\![0;\,N]\!]}$ a family of subsets of $\mathbb{R}$, to which we associate the measures $(\mu_n)_{n\in[\![0;\,N]\!]}$ and the Borelian tribes $(\mathcal{B}(A_n))_{n\in[\![0;\,N]\!]}$. Let $\sigma : [\![0;\,N]\!] \to [\![0;\,N]\!]$ a bijective ordering function (i.e. a permutation). Then, the iterated integrals of f are the integrals defined by

$$\int\int ... \int f \mathrm{d}\mu_{\sigma(0)}\mathrm{d}\mu_{\sigma(1)}...\mathrm{d}\mu_{\sigma(N)} = \int \left(... \left(\int \left(\int f \mathrm{d}\mu_{\sigma(0)} \right) \mathrm{d}\mu_{\sigma(1)} \right) ... \right) \mathrm{d}\mu_{\sigma(N)}.$$

Theorem 54. Fubini's theorems.
Fubini's theorems can be extended to such multiple integrals.

5.2.3 Applications

Gauss's Integral

Let $f : \mathbb{R}^{+\star} \times \mathbb{R}^{+\star} \to \mathbb{R}^+$ the function defined by

$$f : (x, y) \mapsto e^{-x(1+y^2)}.$$

The function f is measurable (because it is continuous) and positive, so we can apply Fubini-Tonelli's theorem. Performing the integration over x yields

$$\int_0^{+\infty} e^{-x(1+y^2)}\mathrm{d}x = \frac{1}{1+y^2},$$

then

$$\int_0^{+\infty} \left(\int_0^{+\infty} e^{-x(1+y^2)}\mathrm{d}x \right) \mathrm{d}y = \int_0^{+\infty} \frac{1}{1+y^2}\mathrm{d}y = \frac{\pi}{2}.$$

On the other hand, integrating over y first yields

$$\begin{aligned}
\int_0^{+\infty} e^{-x(1+y^2)}\mathrm{d}y &= e^{-x} \int_0^{+\infty} e^{-xy^2}\mathrm{d}y \\
&= e^{-x} \int_0^{+\infty} \frac{e^{-t^2}}{\sqrt{x}}\mathrm{d}t \qquad y \to \frac{t}{\sqrt{x}} \\
&= \frac{e^{-x}}{\sqrt{x}} \left(\int_0^{+\infty} e^{-t^2}\mathrm{d}t \right),
\end{aligned}$$

then

$$\begin{aligned}
\int_0^{+\infty} \left(\int_0^{+\infty} e^{-x(1+y^2)}\mathrm{d}y \right) \mathrm{d}x &= \left(\int_0^{+\infty} \frac{e^{-x}}{\sqrt{x}}\mathrm{d}x \right) \left(\int_0^{+\infty} e^{-t^2}\mathrm{d}t \right) \\
&= 2 \left(\int_0^{+\infty} e^{-t^2}\mathrm{d}t \right)^2 \qquad x \to t^2.
\end{aligned}$$

Combining these two results, we find that

$$\int_0^{+\infty} e^{-t^2}\mathrm{d}t = \frac{\sqrt{\pi}}{2},$$

and we deduce the value of Gauss's integral

$$\int_{\mathbb{R}} e^{-t^2}\mathrm{d}t = \sqrt{\pi}.$$

Open Ball in $\mathbb{R}^n$

We work in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda)$ and we define

$$\omega_n = \lambda\left(\{x \in \mathbb{R}^n |\ ||x|| \leq 1\}\right).$$

This quantity ω_n corresponds to the volume of an open ball in $\mathbb{R}^n$, of radius 1. We know that

1. for $n = 1$, we have $\omega_1 = 2$ which is the length of the interval $]-1; 1[$;
2. for $n = 2$, we have $\omega_2 = \pi$ which is the area of a unit disc; and
3. for $n = 3$, we have $\omega_3 = 4\pi/3$ which is the volume of a unit sphere.

Can we derive a general result?

Let $n \in \mathbb{N}$. We consider the product space $\mathbb{R}^n \times \mathbb{R}$ and, for $x \in \mathbb{R}^n$, we will write $x = (x_1, ..., x_n)$. Let $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ defined by

$$f : (x, t) \mapsto e^{-t}\chi_{\{||x||^2<t\}}.$$

The function f is measurable for the Borel tribe $\mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R})$ and f is positive, so we can apply Fubini-Tonelli's theorem. Performing the integration on t, we have

$$\int_{\mathbb{R}} f(x,t)\mathrm{d}t = \int_{||x||^2}^{+\infty} e^{-t}\mathrm{d}t = e^{-||x||^2},$$

hence

$$\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}} f(x,t)\mathrm{d}t\right)\mathrm{d}x = \int_{\mathbb{R}^n} e^{-(x_1^2+...+x_n^2)}\mathrm{d}x = \prod_{i=1}^{n}\left(\int_{\mathbb{R}} e^{-x_i^2}\mathrm{d}x_i\right) = \pi^{n/2}.$$

Now, if we first integrate over x, we obtain

$$\int_{\mathbb{R}^n} f(x,t)\mathrm{d}x = e^{-t}\lambda\left(\left\{x \in \mathbb{R}^n |\ ||x|| \leq \sqrt{t}\right\}\right) = e^{-t}\omega_n t^{n/2}.$$

Thus

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}^n} f(x,t)\mathrm{d}x\right)\mathrm{d}t = \int_{\mathbb{R}} e^{-t}\omega_n t^{n/2}\mathrm{d}t = \omega_n\Gamma\left(\frac{n}{2}+1\right),$$

with Γ the Euler function.

With Fubini-Tonelli's theorem, we know that these two iterated integrals are equal, so we deduce

$$\omega_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

We check that we find the expected results for $n = 1$, 2, and 3; now, if we consider $n = 4$, we obtain that the volume of the unit ball in $\mathbb{R}^4$ is

$$\omega_4 = \frac{\pi^2}{\Gamma(3)} = \frac{\pi^2}{2}.$$

Exercise 64. ★★ Dirichlet integral.

1. Integrating $x \mapsto \sin x e^{-xy}$ on $]0; a[\times]0; +\infty[$, show that

$$\int_0^a \frac{\sin x}{x}\mathrm{d}x = \frac{\pi}{2} - \cos a\int_0^{+\infty}\frac{e^{-ay}}{1+y^2}\mathrm{d}y - \sin a\int_0^{+\infty}\frac{ye^{-ay}}{1+y^2}\mathrm{d}y.$$

2. Use this result to prove that
$$\int_0^{+\infty} \frac{\sin x}{x} \mathrm{d}x = \frac{\pi}{2}.$$
3. Now, let $c \in \mathbb{R}$. Show that
$$I = \int_0^{+\infty} \frac{\sin(cx)}{x} \mathrm{d}x,$$
is equal to $\pi/2$ if $c > 0$, to 0 if $c = 0$, and to $-\pi/2$ if $c < 0$.

EXERCISE 65. ★ Logarithms.

By integrating $(x, y) \mapsto e^{-xy}$ over an appropriate region, show that, for $(a, b) \in (\mathbb{R}^{+\star})^2$, we have
$$\int_0^{+\infty} \frac{e^{-ax} - e^{-bx}}{x} \mathrm{d}x = \ln\left(\frac{b}{a}\right).$$

EXERCISE 66. ★ Changing sign (1/2).

Compute the iterated integrals on $[0;\,1]^2$ of
$$f : (x, y) \mapsto \frac{x - y}{(x + y)^3}.$$
Can we apply Fubini-Lebesgue's theorem?

EXERCISE 67. ★ Changing sign (2/2).

Compute the iterated integrals on $[0;\,1]^2$ of
$$f : (x, y) \mapsto \begin{cases} x^{-2} & \text{if } y < x < 1, \\ -y^{-2} & \text{if } x < y < 1. \end{cases}$$
Can we apply Fubini-Lebesgue's theorem?

EXERCISE 68. ★ Cosine integral.

We define $K = \{(x, y) \in \mathbb{R}^2 | x \geq 0, y \geq 0, x + y \leq 1\}$. Compute the following integral
$$\iint_K \cos\left(\frac{x - y}{x + y}\right) \mathrm{d}(x, y).$$

EXERCISE 69. ★★ Another function of two variables.

We define the function f on $\mathbb{R}^{+\star} \times \mathbb{R}^{+\star}$ by
$$f : (x, y) \mapsto \frac{1}{(1 + y)(1 + x^2 y)}.$$

1. If f Lebesgue-integrable?
2. Compute
$$\iint_{\mathbb{R}^{+\star} \times \mathbb{R}^{+\star}} f(x, y) \mathrm{d}x\mathrm{d}y.$$
3. Use this result to compute
$$\int_0^{+\infty} \frac{\ln(x)}{x^2 - 1} \mathrm{d}x.$$

EXERCISE 70. ★★ Yet another function of two variables.

We define the function f on $[0;\,1] \times \mathbb{R}^{+\star}$ by
$$f : (x, y) \mapsto e^{-y} \sin(2xy).$$

1. Prove that f is Lebesgue-integrable.
2. By integrating f, compute
$$\int_0^{+\infty} \frac{1}{t} \sin^2(t) e^{-t} \mathrm{d}t.$$

5.3 Change of variables

5.3.1 Partial Derivatives and Jacobian

DEFINITION 94. Partial derivatives.
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ a $\mathcal{C}^1$ function of several variables, and $u \in \mathbb{R}^n$. We note $f = (f_1, f_2, ..., f_n)$ and $u = (u_1, u_2, ...u_n)$. The partial derivative of the i^{th} component of f with respect to the j^{th} variable, noted

$$\frac{\partial f_i}{\partial u_j},$$

is obtained by looking at the derivative of f_i with respect to u_j when keeping all other variables constant.

DEFINITION 95. Jacobian.
Let $f : \mathbb{R}^n \to \mathbb{R}^n$ a $\mathcal{C}^1$ function of several variables, and $u \in \mathbb{R}^n$. We note $f = (f_1, f_2, ..., f_n)$ and $u = (u_1, u_2, ...u_n)$. We define the matrice of the partial derivatives of f evaluated in u as

$$(Df)_{ij}(u) = \left(\frac{\partial f_i}{\partial u_j}(u)\right)_{ij}.$$

The determinant of this matrice is called *Jacobian of f in u* and is noted $J_f(u)$.

5.3.2 Diffeomorphism

DEFINITION 96. Diffeomorphism.
Let U and V be two non-empty sets of $\mathbb{R}^n$. The application $\phi : U \to V$ is a *diffeomorphism* if

1. ϕ is bijective;
2. ϕ is $\mathcal{C}^1$ on U; and
3. ϕ^{-1} is $\mathcal{C}^1$ on V.

EXAMPLE 16.
If $U = V = \mathbb{R}$, then $x \mapsto x$ and $x \mapsto -x$ are diffeomorphisms.

EXERCISE 71. ★ Diffeomorphisms.

1. If $U = V = \mathbb{R}$, is $\phi : U \to V$ defined by $x \mapsto x^2$ a diffeomorphism? Same question if $U = V = \mathbb{R}^+$.
2. If $U =]0; +\infty[\times]0; \pi/2[$ and $V =]0; +\infty[^2$, is $\phi : U \to V$ defined by $(x, y) \mapsto (x\cos(y), x\sin(y))$ a diffeomorphism?

EXERCISE 72. ★ Jacobian of a diffeomorphism.
Let $\phi : U \to V$ be a diffeomorphism and $x \mapsto J_\phi(x)$ the function associated to the Jacobian matrix of ϕ. Then, show that for all $x \in U$, $J_{\phi^{-1}}(\phi(x)) = (J_\phi(x))^{-1}$.

Lemma 5. Measure invariant by translation.
We consider the measured space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R})^n, \mu)$, where μ is the Lebesgue measure. If $\nu : \mathcal{B}(\mathbb{R}^n) \to \overline{\mathbb{R}}^{+\star}$ is a measure and is invariant by translation, finite on bounded sets, then there exists $c \in \mathbb{R}^+$ such that $\mu = c\nu$.

Proof.
Let $c = \nu([0;\,1[^n) \geq 0$. For all $k \in \mathbb{N}^\star$, the set $[0;\,1[^n$ can be written as a disjoint union of k^n cuboids that are all translated of the cuboid $[0;\,1/k[^n$. Therefore, since ν is invariant by translation, we have

$$\forall k \in \mathbb{N}^\star, \nu([0;\,1/k[^n) = \frac{c}{k^n}.$$

Let P the subset of $\mathbb{R}^n$ defined by

$$P = \prod_{i=1}^{n} [0;\,a_i[,$$

where $\forall i \in [\![1;\,n]\!], a_i \in \mathbb{R}^+$. If we note E the integer part function, then we have for all $k \in \mathbb{N}^\star$,

$$\prod_{i=1}^{n} \left[0;\,\frac{E(ka_i)}{k}\right[\subset P \subset \prod_{i=1}^{n} \left[0;\,\frac{E(ka_i)+1}{k}\right[.$$

Hence, using the previous result and the invariance by translation of ν, we obtain that

$$\forall k \in \mathbb{N}^\star, c\prod_{i=1}^{n} \frac{E(ka_i)}{k} \leq \nu(P) \leq c\prod_{i=1}^{n} \frac{E(ka_i)+1}{k}.$$

Taking the limit $k \to +\infty$, we obtain that

$$\nu(P) = c\prod_{i=1}^{n} a_i = c\mu(P).$$

By looking at countable unions and intersections of translated of such cuboids, we find that for all open (and for all closed) cuboid P, we have that $\nu(P) = c\mu(P)$. Since the Borelian tribe of $\mathbb{R}^n$ is generated by such cuboids, and as the Lebesgue measure is unique, we have $\nu = c\mu$ for all set of $\mathcal{B}(\mathbb{R}^n)$.

□

Proposition 45. Measure: linear diffeomorphism.
Let $M \in GL_n(\mathbb{R})$ ($n \times n$ inversible matrix) and $b \in \mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}^n$ the function defined by $x \mapsto Mx + b$. Then,

$$A \in \mathcal{B}(\mathbb{R}^n) \Rightarrow f(A) \in \mathcal{B}(\mathbb{R}^n) \quad \text{and} \quad \mu(f(A)) = |\det(M)|\mu(A).$$

Remark 52.
This proposition actually corresponds to the invariance by translation, rotation, and dilatation of the Lebesgue measure on $\mathbb{R}^n$.

Proof.
With the notations of the proposition, we consider the functions f and g defined by

$$f : x \mapsto Mx + b \qquad \text{and} \qquad g : y \mapsto M^{-1}(y - b) = f^{-1}(y).$$

If $A \in \mathcal{B}(\mathbb{R}^n)$, then $f(A) = g^{-1}(A) \in \mathcal{B}(\mathbb{R}^n)$ since $\chi_{g^{-1}(A)} = \chi_A \circ g$ is measurable.

Now, we want to compute $\mu(f(A))$. The Lebesgue measure being invariant by translation, we can assume for simplicity that $b = 0$. If we define the set application ν such that $\forall A \in \mathcal{B}(\mathbb{R}^n), \nu(A) = \mu(f(A))$, then ν is a measure (same properties as μ), ν is invariant by translation (since f is linear), and ν has finite values on bounded sets. Hence, there exists a positive constant $c(M)$ such that $\nu = c(M)\mu$. We only have to prove that this constant is the determinant of M.

There are different cases:

1. If M is orthogonal (i.e. if $M^{-1} = {}^tM$), then we note $B = \{x \in \mathbb{R}^n | \ ||x|| \leq 1\}$. We have $f(B) = \{Mx | x \in B\} = B$. Hence, $\nu(B) = \mu(f(B)) = \mu(B)$ and $\nu = 1\mu$, so $c(M) = 1 = \det(M)$ (M is orthogonal).

2. If M is real, symmetric, and definite, then there exists an orthogonal matrix O and a (positive) diagonal matrix $D = \mathrm{diag}(d_1, d_2, ...d_n)$ such that we can write
$$ {}^tOMO = D. $$
Let $P = [0;\, 1]^n$ and A the set defined by $A = OP$, then
$$ \mu(A) = \mu(OP) = \mu(P) = 1, $$
and
$$ \nu(A) = \mu(f(A)) = \mu(MA) = \mu(ODP) = \mu\left(\prod_{i=1}^{n} [0;\, d_i]\right) = \det(M). $$
In this case, $\nu = c(M)\mu$ with $c(M) = \det(M)$.

3. In the general case, we write the polar decomposition of M as $M = RS$ with
$$ S = ({}^tMM)^{1/2} \qquad \text{and} \qquad R = MS^{-1}, $$
where S is real, symmetric, and positive, and R is orthogonal. Then, for all Borelian set A we can write
$$ \nu(A) = \mu(RSA) = \mu(SA) = \det(S)\mu(A), $$
and therefore $\nu = c(M)\mu$ with $c(M) = \det(S) = |\det(M)|$.

□

EXERCISE 73. ★ Translation, rotation, and dilatation of a set in $\mathbb{R}^2$.
Let A be a Borelian set of $\mathbb{R}^2$. Using this proposition, show that

1. If $u \in \mathbb{R}^2$, then $\mu(A + u) = \mu(A)$.

2. If $\theta \in [0;\, 2\pi[$ and A_θ is obtained by rotatation of angle θ of the set A, then $\mu(A_\theta) = \mu(A)$.

3. If $k \in \mathbb{R}^{+\star}$ and if kA is the set obtained by dilatation of A, then $\mu(kA) = \mu(A)$.

How does this translate in $\mathbb{R}$? And in $\mathbb{R}^n$?

PROPOSITION 46. Measure: general diffeomorphism.
Let $\phi : U \to V$ be a diffeomorphism. If $A \in \mathcal{B}(\mathbb{R}^n)$, $A \subset U$, then $\phi(A) \in \mathcal{B}(\mathbb{R}^n)$ and
$$ \mu(\phi(A)) = \int_A |J_\phi| \mathrm{d}\mu = \int_A |J_\phi(u)| \mathrm{d}\mu(u). $$

5.3.3 Change of Variables

THEOREM 55. Change of variables in $\mathbb{R}^n$ (positive measurable functions).
Let U and V be two open sets of $\mathbb{R}^n$ and $\phi : U \to V$ a $\mathcal{C}^1$ diffeomorphism. Then if $f : V \to [0; +\infty[$ is positive and measurable, then $(f \circ \phi)|J_\phi| : U \to [0; +\infty[$ is measurable, and

$$\int_V f(v)\mathrm{d}v = \int_U f(\phi(u))|J_\phi(u)|\mathrm{d}u.$$

THEOREM 56. Change of variables in $\mathbb{R}^n$ (integrable functions).
Let U and V be two open sets of $\mathbb{R}^n$ and $\phi : U \to V$ a $\mathcal{C}^1$ diffeomorphism. Then if $f : V \to \mathbb{R}$ is integrable, then $(f \circ \phi)|J_\phi| : U \to \mathbb{R}$ is integrable, and

$$\int_V f(v)\mathrm{d}v = \int_U f(\phi(u))|J_\phi(u)|\mathrm{d}u.$$

PROOF.
We will prove the two theorems at once.
Let $B \subset V$ be a Borelian, and let $A = \phi^{-1}(B) \subset U$.
If we consider $f = \chi_B$, then $f \circ \phi = \chi_B \circ \phi = \chi_{\phi^{-1}(B)} = \chi_A$, and we obtain

$$\int_B f(v)\mathrm{d}v = \int_B \mathrm{d}v = \mu(B),$$

but also

$$\int_A f(\phi(u))|J_\phi(u)|\mathrm{d}u = \mu(\phi(A)) = \mu(B),$$

thanks to the proposition on a diffeomorphism (general case). Hence, by linearity of the integral, the first theorem is true for positive simple functions and, by taking the limit (see definition of the Lebesgue integral), for all positive measurable function.

Now, if f is integrable (and real), we can apply this theorem to f^+ and f^- and we obtain that the second theorem is true. We can also prove it if f is integrable and complex, by applying the results to $\Re(f)$ and $\Im(f)$.

□

COROLLARY 13. Change of variables in $\mathbb{R}$.
Let f be integrable on $]a; b[\subset \mathbb{R}$ and $\phi :]\alpha; \beta[\to]a; b[$ a function in $\mathcal{C}^1(]\alpha; \beta[)$, that is increasingly monotonic and bijective. We have

$$\int_a^b f\mathrm{d}\mu = \int_\alpha^\beta (f \circ \phi)\phi'\mathrm{d}\mu.$$

EXERCISE 74. ★ Change of variables in $\mathbb{R}$.
Using the previous theorems, prove the corollary on the change of variables in $\mathbb{R}$

EXAMPLE 17.
On $\mathbb{R}^n$, let $u \in \mathbb{R}^n$ and $\phi : x \to x + u$. Then, ϕ is diffeomorphism and $|J_\phi| = 1$. Then, if $f : \mathbb{R} \to \mathbb{R}$ is integrable, we have

$$\int_{\mathbb{R}^n} f(x)\mathrm{d}x = \int_{\mathbb{R}^n} f(x+u)\mathrm{d}x.$$

5.3.4 Applications to $\mathbb{R}^2$ and $\mathbb{R}^3$

PROPOSITION 47. Polar coordinates in $\mathbb{R}^2$.
Let $U = \{(r,\theta) | r \in \mathbb{R}^{+\star} \text{ and } \theta \in]-\pi;\pi[\}$ and $V = \mathbb{R}^2 \setminus \{(x,0) | x \in \mathbb{R}^-\}$. We consider $\phi : U \to V$ defined by

$$\phi : (r,\theta) \mapsto (r\cos\theta, r\sin\theta).$$

Then ϕ is a $\mathcal{C}^\infty$ diffeomorphism, and

$$J_\phi(r,\theta) = \det \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r > 0.$$

If f is a positive measurable function, or an integrable function, defined on $\mathbb{R}^2$, we have

$$\int_{\mathbb{R}^2} f(x,y)\mathrm{d}x\mathrm{d}y = \int_U f(r\cos\theta, r\sin\theta) r\mathrm{d}r\mathrm{d}\theta.$$

COROLLARY 14. Invariance by rotation in $\mathbb{R}^2$.
Applying the previous theorem to a function f invariant by rotation, i.e. there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that $\forall (x,y) \in \mathbb{R}^2, f(x,y) = g(\sqrt{x^2+y^2})$, then Fubini yields

$$\int_{\mathbb{R}^2} f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^{+\infty} 2\pi r g(r)\mathrm{d}r.$$

EXAMPLE 18. Gauss's integral.
Using Fubini-Tonelli's theorem, we can write that

$$\int_0^{+\infty} e^{-x^2}\mathrm{d}x = \left(\int_0^{+\infty} e^{-x^2}\mathrm{d}x \times \int_0^{+\infty} e^{-y^2}\mathrm{d}y\right)^{1/2} = \left(\int_{\mathbb{R}^{+\star}\times\mathbb{R}^{+\star}} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y\right)^{1/2}.$$

In polar coordinates, we have

$$\int_{\mathbb{R}^{+\star}\times\mathbb{R}^{+\star}} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y = \int_0^{\pi/2}\int_0^{+\infty} e^{-r^2} r\mathrm{d}r\mathrm{d}\theta = \frac{\pi}{2}\int_0^{+\infty} re^{-r^2}\mathrm{d}r = \frac{\pi}{4}.$$

We deduce the value of Gauss's integral

$$\int_{\mathbb{R}} e^{-x^2}\mathrm{d}x = \sqrt{\pi}.$$

EXERCISE 75. ★ Integrability and polar coordinates.
Let $\alpha \in \mathbb{R}^{+\star}$ and $f : \mathbb{R}^2 \to \mathbb{R}^+$ defined by

$$f : (x,y,z) \mapsto \frac{1}{(x^2+y^2)^{\alpha/2}}.$$

Let $\mathcal{B} = \{(x,y,z) \in \mathbb{R}^3 | x^2+y^2+z^2 \leq 1\}$. Is f integrable on $\mathcal{B}$? Is f integrable on $\mathcal{B}^c$? Compute the integral of f on $\mathcal{B}$ and $\mathcal{B}^c$ when it is possible to do so.

Corollary 15. Cylindrical coordinates in $\mathbb{R}^3$.
Let $U = \{(r,\theta,z) | r \in \mathbb{R}^{+\star},\ \theta \in]-\pi;\pi[, \text{ and } z \in \mathbb{R}\}$ and $V = \mathbb{R}^3 \setminus \{(x,0,0) | x \in \mathbb{R}^-\}$. We consider $\phi : U \to V$ defined by

$$\phi : (r,\theta,z) \mapsto (r\cos\theta, r\sin\theta, z).$$

Then ϕ is a $\mathcal{C}^\infty$ diffeomorphism, and

$$J_\phi(r,\theta,z) = \det\begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = r > 0.$$

If f is a positive measurable function, or an integrable function, defined on $\mathbb{R}^3$, we have

$$\int_{\mathbb{R}^3} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_U f(r\cos\theta, r\sin\theta, z) r\mathrm{d}r\mathrm{d}\theta\mathrm{d}z.$$

Proposition 48. Spherical coordinates in $\mathbb{R}^3$.
We consider the two following sets: $U = \{(r,\theta,\varphi) | r \in \mathbb{R}^{+\star},\ \theta \in]-\pi;\pi[, \text{ and } \varphi \in]0;\pi[\}$ and $V = \mathbb{R}^3 \setminus \{(x,0,z) | x \in \mathbb{R}^- \text{ and } z \in \mathbb{R}\}$. Let $\phi : U \to V$ defined by

$$\phi : (r,\theta,\varphi) \mapsto (r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi).$$

Then ϕ is a $\mathcal{C}^\infty$ diffeomorphism, and

$$J_\phi(r,\theta,\varphi) = \det\begin{pmatrix} \cos\theta\sin\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\sin\theta\cos\varphi & r\cos\theta\sin\varphi \\ \cos\varphi & -r\sin\varphi & 0 \end{pmatrix} = r^2\sin\varphi > 0.$$

If f is a positive measurable function, or an integrable function, defined on $\mathbb{R}^3$, we have

$$\int_{\mathbb{R}^3} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_U f(r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi) r^2\sin\varphi\mathrm{d}r\mathrm{d}\theta\mathrm{d}\varphi.$$

Corollary 16. Invariance by rotation in $\mathbb{R}^3$.
Applying the previous theorem to a function f invariant by rotation, i.e. there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that $\forall (x,y,z) \in \mathbb{R}^3, f(x,y,z) = g(\sqrt{x^2+y^2+z^2})$, then Fubini yields

$$\int_{\mathbb{R}^3} f(x,y,z)\mathrm{d}x\mathrm{d}y = \int_0^{+\infty} 4\pi r^2 g(r)\mathrm{d}r.$$

Example 19.
Let $\alpha \in \mathbb{R}^{+\star}$ and $f : \mathbb{R}^3 \to \mathbb{R}^+$ defined by

$$f : (x,y,z) \mapsto \frac{1}{(x^2+y^2+z^2)^{\alpha/2}}.$$

The function f is positive, measurable. Is f integrable on $\mathbb{R}^3$?
Let $\mathcal{B} = \{(x,y,z) \in \mathbb{R}^3 | x^2+y^2+z^2 \leq 1\}$. We have

1. f integrable on $\mathcal{B}$ if and only if $\alpha < 3$; and
2. f integrable on $\mathcal{B}^c$ if and only if $\alpha > 3$.

This can be shown by computing the integral of f on $\mathcal{B}$ and on $\mathcal{B}^c$.

$$\int_{\mathcal{B}} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = 4\pi \int_0^1 r^{2-\alpha}\mathrm{d}r = \begin{cases} \frac{4\pi}{3-\alpha} & \text{if } \alpha < 3 \\ +\infty & \text{if } \alpha \geq 3 \end{cases},$$

and

$$\int_{\mathcal{B}^c} f(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z = 4\pi \int_1^{+\infty} r^{2-\alpha}\mathrm{d}r = \begin{cases} \frac{4\pi}{\alpha-3} & \text{if } \alpha > 3 \\ +\infty & \text{if } \alpha \leq 3 \end{cases}.$$

5.3.5 Generalisation to $\mathbb{R}^n$

Definition 97. Euclidean surface measure.
Let $n \in \mathbb{N}^\star$. We define $S^{n-1} = \{x \in \mathbb{R}^n | \ ||x|| = 1\}$. For all $A \in S^{n-1}$, we define the portion $\gamma(A)$ by

$$\gamma(A) = \{rx \in \mathbb{R}^n | x \in A \text{ and } 0 < r \leq 1\}.$$

Then, we call *Euclidean surface measure* the Borelian measure defined on S^{n-1} by

$$\forall A \in \mathcal{B}(S^{n-1}), \sigma_n(A) = n\mu(\gamma(A)).$$

Example 20.
We recover the measure of the area of a surface (that has dimension $n-1$ in a space of dimension n). For example, $\sigma_1(S^0) = 2$ (i.e. counting the bounds of the unit interval), $\sigma_2(S^1) = 2\pi$ (perimeter of the unit circle in a plane), $\sigma_3(S^2) = 4\pi$ (area of the unit sphere).

Proposition 49. Polar coordinates in $\mathbb{R}^n$.
If f is a positive measurable, or integrable, function defined on $\mathbb{R}^n$, then

$$\int_{\mathbb{R}^n} f\mathrm{d}\mu = \int_0^{+\infty}\int_{S^{n-1}} f(r,\omega)r^{d-1}\mathrm{d}\mu(r)\mathrm{d}\sigma_n(\omega).$$

Corollary 17. Invariance by rotation in $\mathbb{R}^n$.
Applying the previous theorem to a function f invariant by rotation, i.e. there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that $\forall x \in \mathbb{R}^n, f(x) = g(||x||)$, then Fubini yields

$$\int_{\mathbb{R}^n} f\mathrm{d}\mu = \alpha_n \int_0^{+\infty} g(r)r^{n-1}\mathrm{d}r \quad \text{with} \quad \alpha_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}.$$

5.4 Conclusions

In this chapter we focused on integration on product spaces, i.e. integration on spaces defined through a Cartesian product. In order to do so, we defined the Cartesian product of two measured spaces, formed by the Cartesian produt of the two spaces, the product tribe (in the case of subsets of $\mathbb{R}$, the product of the Borelian tribes is the Borelian tribe of the Cartesian product space), and the product measure (in our case, the product of two Lebesgue measures). This leads to the definition of iterated integrals and of double integrals. Iterated integrals are the integrals performed on one space, then on the other one, and so on, whereas double integrals (or multiple integrals) are integrals performed over the Cartesian product space, at once. In general, these integrals are not the same. Based on this construction, we have seen three important results: Fubini-Tonelli, Fubini-Lebesgue, and the change of variable theorem. These three theorems are not difficult to use, but one should be very careful with the assumptions. As for the permutation theorems (limits-integrals, sums-integrals, etc), permuting the integration signs gives wrong, or contradictory, results when Fubini's assumptions are not satisfied.

CHAPTER 6

Integration Calculus Techniques

After having constructed the Lebesgue measure and the Lebesgue integration theory, we will now spend some time on actually computing integrals of real functions. As such, this chapter simply consists in an overview of some integration techniques that can be useful to compute integrals. Of course, these are not exclusive to Lebesgue integration and can also be applied to Riemann integrals.

6.1 Usual Primitives

Definition 98. Primitives.
Let $I = [a; b] \subset \mathbb{R}$ and let $f : I \to \mathbb{R}$ be an integrable function on I. The function $F : I \to \mathbb{R}$, defined by

$$\forall x \in I, F(x) = \int_a^x f(t)\mathrm{d}t,$$

is called a *primitive of f*.

Theorem 57. Existence of primitives.
Let f be a continuous function on an interval $I \subset \mathbb{R}$. This function has primitives.

Remark 53.

1. The primitive of a function f on an interval I are defined with respect to one constant.
2. With these notations, $F : x \mapsto \displaystyle\int_a^x f(t)\mathrm{d}t$ is the unique primitive of f such that $F(a) = 0$.
3. Here, primitives are defined in the Riemann sense; we will come back to it in the chapter on derivation to look at them in the Lebesgue theory.

6.2 Integration Techniques and Change of Variables

This section provides a list of integration techniques, such as relevant integration by parts and change of variables, that can be used to quickly compute integrals.

Function	Primitive	Interval
$x^n, n \in \mathbb{N}$	$\frac{x^{n+1}}{n+1}$	$\mathbb{R}$
$x^n, n \in \mathbb{Z}, n < -1$	$\frac{x^{n+1}}{n+1}$	$\mathbb{R}^{+\star}$ or $\mathbb{R}^{-\star}$
$x^\alpha, \alpha \in \mathbb{R} \setminus \{-1\}$	$\frac{x^{\alpha+1}}{\alpha+1}$	$\mathbb{R}^{+\star}$
$\frac{1}{x}$	$\ln\|x\|$	$\mathbb{R}^{+\star}$ or $\mathbb{R}^{-\star}$
$\ln\|x\|$	$x\ln\|x\| - x$	$\mathbb{R}^{+\star}$ or $\mathbb{R}^{-\star}$
e^x	e^x	$\mathbb{R}$
$\cos x$	$\sin x$	$\mathbb{R}$
$\sin x$	$-\cos x$	$\mathbb{R}$
$\mathrm{ch}x$	$\mathrm{sh}x$	$\mathbb{R}$
$\mathrm{sh}x$	$\mathrm{ch}x$	$\mathbb{R}$
$\tan x$	$-\ln\|\cos x\|$	$]-\pi/2 + k\pi;\ \pi/2 + k\pi[, k \in \mathbb{Z}$
$\cot x$	$\ln\|\sin x\|$	$]k\pi;\ (k+1)\pi[, k \in \mathbb{Z}$
$\mathrm{th}x$	$\ln\|\mathrm{ch}x\|$	$\mathbb{R}$
$\frac{1}{\cos^2 x}$	$\tan x$	$]-\pi/2 + k\pi;\ \pi/2 + k\pi[, k \in \mathbb{Z}$
$\frac{1}{\sin^2 x}$	$-\cot x$	$]k\pi;\ (k+1)\pi[, k \in \mathbb{Z}$
$\frac{1}{\mathrm{ch}^2 x}$	$\mathrm{th}x\|$	$\mathbb{R}$
$\frac{1}{\mathrm{sh}^2 x}$	$-\frac{1}{\mathrm{th}x}$	$\mathbb{R}^{+\star}$ or $\mathbb{R}^{-\star}$
$\frac{1}{1+x^2}$	$\arctan x$	$\mathbb{R}$
$\frac{1}{1-x^2}$	$\frac{1}{2}\ln\left\|\frac{1+x}{1-x}\right\|$	$]-\infty;\ -1[$ or $]-1;\ 1[$ or $]1;\ +\infty[$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$]-1;\ 1[$
$\frac{1}{\sqrt{h+x^2}}, h \in \mathbb{R}^\star$	$\ln\|x + \sqrt{x^2+h}\|$	$\mathbb{R}$ if $h > 0$, $]-\infty;\ -\sqrt{-h}[\cup]\sqrt{-h};+\infty[$ otherwise

Table 6.1: Table of usual primitives.

6.2.1 Polynomial Functions

METHOD:
Use the linearity of the integral, and perform the integral on each of the $x \mapsto x^n$ terms.

6.2.2 Polynomial Functions × Log, Exp, Sine and Cosine

METHOD:

1. If the function can be written $x \mapsto P(x)e^{\omega x}$ with P a polynomial function, then do an integration by parts by derivating P. We recover the same problem, but the polynomial function has now been reduced of 1 degree. If P is initially of degree n, simply do n integrations by parts. Indeed
$$\int^t P(x)e^{\omega x}\mathrm{d}x = \frac{1}{\omega}\left(P(t)e^{\omega t} - \int^t P'(x)e^{\omega x}\mathrm{d}x\right).$$
2. If the function can be written $x \mapsto P(x)\cos(\omega x)$ or $x \mapsto P(x)\sin(\omega x)$ with P a polynomial function, the do exactly as for $x \mapsto P(x)e^{\omega x}$.
3. If the function can be written $x \mapsto P(x)\ln(\omega x)$, then do one integration by parts by integrating P to recover a simple polynomial integration.

EXERCISE 76. ★ Polynomial functions.
Compute the primitives of $f : x \mapsto (x^2 - 3x + 2)\cos x$ and $g : x \mapsto (x^3 + 2x^2 - 4)\ln x$.

6.2.3 Frational Functions

METHOD:

1. Use the decomposition in simple elements in $\mathbb{C}[X]$, using the singular roots of the fractional function. Then, the problem is reduced to the integration of polynomial functions and of rational functions of the form $x \mapsto 1/(x-\alpha)^n$.
2. Or, use the decomposition in simple elements in $\mathbb{R}[X]$. Then the problem is reduced to the integration of polynomial functions, and to functions of the form $x \mapsto 1/(x-\alpha)^n$ and $x \mapsto (ax+b)/(x^2 - \beta x + \gamma)^n$.

EXERCISE 77. ★ Fractional functions.
Compute the primitive of $f : x \mapsto \dfrac{1}{x^3 - 1}$.

6.2.4 Fractional Cosine and Sine Functions

METHOD:

1. Use the change of variable $u = \tan(x/2)$ to recover a fractional function in u. This works all the time, but requires some care. We have, in particular
$$\mathrm{d}x = \frac{2\mathrm{d}u}{1+u^2}, \qquad \cos x = \frac{1-u^2}{1+u^2} \qquad \text{and} \qquad \sin x = \frac{2u}{1+u^2}.$$
2. Or, sometimes, you can use Bioche's rules and an appropriate change of variable.

PROPOSITION 50. Bioche's rules.
Since we want to compute $\int f(x)\mathrm{d}x$, we note $\omega(x) = f(x)\mathrm{d}x$. Then

1. If $\omega(-x) = \omega(x)$, use $u = \cos x$.
2. If $\omega(\pi - x) = \omega(x)$, use $u = \sin x$.
3. If $\omega(\pi + x) = \omega(x)$, use $u = \tan x$.

EXERCISE 78. ★ Fractional cosine and sine functions.
Compute the primitives of $f : x \mapsto \dfrac{1}{2 + \cos x}$ and $g : x \mapsto \dfrac{2 + \sin^2 x}{\cos x}$.

6.2.5 Fractional Functions in Exp, Hyperbolic Sine and Cosine

METHOD:

1. Use the change of variable $u = e^x$ to recover a fractional function in u. This works all the time, but requires some care. We have, in particular $\mathrm{d}u = u\mathrm{d}x$.
2. Or, sometimes, you can use another version of Bioche's rules and an appropriate change of variable.

PROPOSITION 51. Bioche's rules.
Since we want to compute $\int f(x)\mathrm{d}x$, we note $\omega(x) = f(x)\mathrm{d}x$ in which we replace the hyperbolic functions ch, sh, and th, by cos, sin, and tan, respectively. Then

1. If $\omega(-x) = \omega(x)$, use $u = \mathrm{ch}x$.
2. If $\omega(\pi - x) = \omega(x)$, use $u = \mathrm{sh}x$.
3. If $\omega(\pi + x) = \omega(x)$, use $u = \mathrm{th}x$.

EXERCISE 79. ★ Fractional function in exponential, hyperbolic sine and cosine.
Compute the primitive of $f : x \mapsto \dfrac{1}{\mathrm{ch}x}$.

6.2.6 Fractional Function with a Square Root Argument

METHOD:

1. If the argument is $\sqrt{ax^2 + bx + c}$, then
   (a) If $a = 0$, it is a fractional function in $\sqrt{bx + c}$, then use the change of variable $u = \sqrt{bx + c}$.
   (b) If $a \neq 0$, then factorise by $\sqrt{|a|}$ and factorise the 2nd order polynomial. After that, use a linear change of variable to obtain $\sqrt{u^2 + \alpha^2}$ (then write $u = \alpha\mathrm{sh}t$), $\sqrt{u^2 - \alpha^2}$ (then write $u = \alpha\mathrm{ch}t$), or $\sqrt{-u^2 + \alpha^2}$ (then write $u = \alpha \sin t$).
2. If the argument is $\sqrt{\dfrac{ax + b}{cx + d}}$ then use the change of variable $u = \sqrt{\dfrac{ax + b}{cx + d}}$.

EXERCISE 80. ★ Fractional function with a square root argument.
Compute the primitive of $f : x \mapsto \dfrac{x}{\sqrt{x^2 + x + 1}}$.

6.2.7 Multi-variable Sine and Cosine Functions

Method:

When the functions are separable, it is often (but not always) possible to use the complex form of sine and cosine functions in order to do the calculus.

Example 21.

$$\int_0^{+\infty} e^{-y}\sin(xy)\mathrm{d}y = \Im\left(\int_0^{+\infty} e^{-y}e^{ixy}\mathrm{d}y\right) = \Im\left(\frac{1}{1-ix}\right) = \frac{x}{1+x^2}.$$

6.3 Conclusion

This chapter provides a table of the usual primitives as well as a few calculus tips that are useful to compute integrals. When confronted to an integration problem, part of the work is to prove that integrals are well defined, or that permutation theorems apply, but part of the work is also often to compute the integrals in order to get a final result. Most of the integrals on simple functions can be dealt with using these techniques.

CHAPTER 7

$\mathcal{L}^p$ and L^p Spaces

The previous chapters focused the construction of Lebesgue integration theory, with its measure and its integral. While discussing this, we always considered the functions independently of each other when mentioning their properties: measurability, integrability, etc. Our goal, now, is to look at them at a larger scale: is there a “space” of Lebesgue integrable functions, and what are its properties? We already had a glance at it, when we defined the $\mathcal{L}^1$ space of integrable functions; in this chapter, we will study it more carefully.

7.1 Functional Spaces

7.1.1 $\mathcal{L}^p$ Space

Definition 99. $\mathcal{L}^p$ space.
Let $p \in [1; +\infty[$. We define

$$\mathcal{L}^p(X, \mathcal{M}, \mu) = \left\{ f : X \to \mathbb{K} \text{ measurable} \,\middle|\, \int_X |f|^p \mathrm{d}\mu < +\infty \right\}.$$

Definition 100. $\mathcal{L}^p$-norm.
Let $p \in [1; +\infty[$. For $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, we define

$$||f||_p = \left(\int_X |f|^p \mathrm{d}\mu \right)^{\frac{1}{p}}.$$

Example 22.

1. Case $p = 1$: $\mathcal{L}^1(X, \mathcal{M}, \mu)$ is the set of Lebesgue integrable functions.

2. Case $p = 2$: $\mathcal{L}^2(X, \mathcal{M}, \mu)$ is the set of functions f such that f^2 is Lebesgue integrable.

Exercise 81. ★ Which space?
Let $\alpha \in \mathbb{R}$. We define $f_\alpha : x \mapsto x^\alpha$ and $g_\alpha : x \mapsto e^{\alpha x}$. In which space(s) $\mathcal{L}^p(]0;\ 1], \mathcal{B}(]0;\ 1]), \mu)$ are f_α and g_α? Same question with $\mathcal{L}^p([1; +\infty[, \mathcal{B}([1; +\infty[), \mu)$ and with $\mathcal{L}^p(\mathbb{R}^{+\star}, \mathcal{B}(\mathbb{R}^{+\star}), \mu)$.

EXERCISE 82. ★ p-norm and exponent.
Let $p \in [1; +\infty[$ and $a \in \mathbb{R}^{+\star}$. Show that

$$|| \, |f|^a ||_p = ||f||_{ap}^a.$$

DEFINITION 101. Essentially bounded.
Let $f : X \to \mathbb{K}$. f is said to be *essentially bounded* if there exists $a \in \mathbb{R}^+$ such that $|f| \leq a$ μ-almost everywhere. In terms of "measuring a set", this means

$$\mu \left(\{x \in X | \; |f(x)| > a\} \right) = 0.$$

We note $\mathcal{L}^\infty(X, \mathcal{M}, \mu)$ the set of essentially bounded functions on X.

DEFINITION 102. Essentially bounded.
Let $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$. We note

$$||f||_\infty = \sup_{x \in X} \text{ess} |f(x)| = \inf \left\{ a \in \mathbb{R}^+ | \; |f| \leq a \quad \mu - \text{almost everywhere} \right\}.$$

EXERCISE 83. ★ Essentially bounded and sup.
Let f and g be two functions such that $f = g$ μ-almost everywhere. Prove that $||f||_\infty$ is the smallest of all numbers of the form $\sup\limits_{x \in X} \{|g(x)|\}$.

DEFINITION 103. Conjugated exponents.
Let $(p, q) \in [1; +\infty]^2$. p and q are *conjugated exponents* if

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

LEMMA 6. Conjugated exponents.
Let $(p, q) \in]1; +\infty[^2$ be two conjugated exponents, and $(a, b) \in (\mathbb{R}^+)^2$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

and the equality is reached if and only if $a^p = b^q$.

PROOF.
We first note that, if either $a = 0$ or $b = 0$ (or both), then the lemma is true. Thus, we assume that $(a, b) \in (\mathbb{R}^{+\star})^2$ and we define

$$x = p \ln(a) \qquad \text{and} \qquad y = q \ln(b),$$

hence

$$a = e^{x/p} \qquad \text{and} \qquad b = e^{y/q}.$$

Since the exponential function is strictly convex, we deduce that

$$ab = \exp\left(\frac{x}{p} + \frac{y}{q} \right) \leq \frac{1}{p} e^x + \frac{1}{q} e^y = \frac{a^p}{p} + \frac{b^q}{q},$$

and the equality is reached if and only if $x = y$, i.e. $a^p = b^q$.

□

Proposition 52. Hölder inequality.
Let $(p, q) \in [1; +\infty]^2$ two conjugated exponents. If f and g are two measurable functions defined on X, then

$$||fg||_1 \leq ||f||_p ||g||_q.$$

Proof.
Let f and g be two measurable functions. Then, fg is also measurable and the quantities in the inequality are well defined.

We note that, if $p = 1$ and $q = +\infty$ (and reciprocally), the result is trivial using domination theorems on the Lebesgue integral. We therefore assume that $(p, q) \in]1; +\infty[^2$.

We identify two trivial cases: if $f = 0$ almost everywhere (or if $g = 0$ almost everywhere), then $fg = 0$ almost everywhere and both sides of the inequality are equal to zero, the proposition is true; if $||f||_p = +\infty$ (or if $||g||_q = +\infty$), then the right-hand side of the inequality is equal to $+\infty$ and the proposition is true.

Now, we assume that $(p, q) \in]1; +\infty[^2$ and that $||f||_p < +\infty$ and $||g||_q < +\infty$. We define the following functions

$$F : x \mapsto \frac{|f(x)|}{||f||_p} \qquad \text{and} \qquad G : x \mapsto \frac{|g(x)|}{||g||_q}.$$

Then, $F \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, $G \in \mathcal{L}^q(X, \mathcal{M}, \mu)$, and we have $||F||_p = ||G||_q = 1$. The lemma on conjugated exponents gives, for all $x \in X$, the inequality

$$F(x)G(x) \leq \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q.$$

By integrating over X, we find

$$\int_X FG\mathrm{d}\mu \leq \frac{1}{p}||F||_p^p + \frac{1}{q}||G||_q^q = \frac{1}{p} + \frac{1}{q} = 1,$$

hence

$$||fg||_1 \leq ||f||_p ||g||_q.$$

□

Remark 54.

1. If $(p, q) \in]1; +\infty[^2$, then the inequality writes

$$\int_X |fg|\mathrm{d}\mu \leq \left(\int_X |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}} \left(\int_X |g|^q \mathrm{d}\mu\right)^{\frac{1}{q}}.$$

2. If $p = 1$ and $q = +\infty$, then the inequality writes

$$\int_X |fg|\mathrm{d}\mu \leq \left(\int_X |f|\mathrm{d}\mu\right) ||g||_\infty.$$

Exercise 84. ★★ A variation of Hölder inequality.
Let $(p, q, r) \in [1; +\infty[^k$ such that

$$\frac{1}{r} = \frac{1}{q} + \frac{1}{p}.$$

For f and g two integrable functions, and assuming that the norms are defined, show that

$$||fg||_r \leq ||f||_p ||g||_q.$$

EXERCISE 85. ★★ A more general Hölder inequality.
Let $k \in \mathbb{N}$ and $(p_1, ..., p_k) \in [1; +\infty[^k$ such that

$$\sum_{i=1}^{k} \frac{1}{p_i} = 1.$$

We consider a family of integrable functions f_1, ..., f_k. Assuming that all the norms are defined, show that

$$\left| \int_X \prod_{i=1}^{k} f_i \mathrm{d}\mu \right| \leq \prod_{i=1}^{k} ||f_i||_{p_i}.$$

EXERCISE 86. ★★ Hölder equality.
Taking the Hölder inequality with positive functions $f \geq 0$ and $g \geq 0$, and

$$\int_X fg \mathrm{d}\mu = ||f||_p ||g||_q.$$

Prove that $f^p = g^q$ μ-almost everywhere, to within a multiplicative constant.

PROPOSITION 53. Minkowski inequality.
Let $p \in [1; +\infty]$, and let f and g two measurable functions defined on X. Then

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

PROOF.
First of all, we note that the result is trivial for $p = 1$ and $p = +\infty$, since $|f + g| \leq |f| + |g|$. Thus, we assume that $p \in]1; +\infty[$. Since we can replace f and g by $|f|$ and $|g|$, we also assume that f and g are positive functions.
The function $x \mapsto x^p$ is convex for $x \in [0; +\infty[$, therefore

$$\forall (a, b) \in (\mathbb{R}^+)^2, \left(\frac{a}{2} + \frac{b}{2}\right)^p \leq \frac{a^p}{2} + \frac{b^p}{2},$$

hence

$$\forall (a, b) \in (\mathbb{R}^+)^2, (a + b)^p \leq 2^{p-1}(a^p + b^p).$$

Applying this to the functions f and g, we have

$$\forall x \in X, (f(x) + g(x))^p \leq 2^{p-1}(f^p(x) + g^p(x)),$$

and $f + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. Now, we introduce the positive function $h = f + g$. This yields

$$h^p = (f + g)h^{p-1} = fh^{p-1} + gh^{p-1}.$$

Let introduce q the conjugated exponent of p, i.e. $q = p/(p - 1)$. Then, we can write

$$||h^{p-1}||_q = \left(\int_X h^p \mathrm{d}\mu\right)^{\frac{1}{q}} = ||h||_p^{p/q} = ||h||_p^{p-1} < +\infty.$$

And, using Hölder inequality of $h^p = fh^{p-1} + gh^{p-1}$, we obtain

$$||h||_p^p \leq ||f||_p ||h^{p-1}||_q + ||g||_p ||h^{p-1}||_q = ||h||_p^{p-1}(||f||_p + ||g||_p||).$$

If $||h||_p = 0$ then f and g are 0 μ-almost everywhere, so the Minkowski inequality is true. If $||h||_p \neq 0$, then we can divide by $||h||_p^{p-1}$ and we obtain the result. □

Exercise 87. ★★ p-norm and q-norm.
Let $(p,q) \in [1;+\infty]^2$ with $p<q$. We assume that $\mu(X) < +\infty$. Let f be a function defined on X, and we assume that $f \in \mathcal{L}^p(X, \mathcal{B}(X), \mu)$ and $f \in \mathcal{L}^q(X, \mathcal{B}(X), \mu)$. Show that

$$||f||_p \leq (\mu(X))^{\frac{1}{p}-\frac{1}{q}} \, ||f||_q.$$

Exercise 88. ★★ p-norm, q-norm, and ∞-norm.
Let $(p,q) \in [1;+\infty]^2$ with $p<q$.

1. Prove that $\mathcal{L}^p(X, \mathcal{B}(X), \mu) \cap \mathcal{L}^\infty(X, \mathcal{B}(X), \mu) \subset \mathcal{L}^q(X, \mathcal{B}(X), \mu)$.

2. Let $f \in \mathcal{L}^p(X, \mathcal{B}(X), \mu) \cap \mathcal{L}^\infty(X, \mathcal{B}(X), \mu)$. Show that

$$||f||_q \leq ||f||_p^{p/q} ||f||_\infty^{1-p/q}.$$

Proposition 54. Vector space $\mathcal{L}^p$.
Let $p \in [1;+\infty]$. The space $\mathcal{L}^p(X, \mathcal{M}, \mu)$ is a vector space.

Proof.

1. It is a sub-space of the vector space $\mathcal{C}(X)$.

2. $0 \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ (the constant function equal to 0^p is integrable).

3. If $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ and $\lambda \in \mathbb{K}$, then $\lambda f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$.

4. If f and g are two elements of $\mathcal{L}^p(X, \mathcal{M}, \mu)$, then using Minkowski we show that $|f+g|^p$ is integrable, so $f+g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$.

□

Proposition 55. Semi-norm $f \mapsto ||f||_p$.
Let $p \in [1;+\infty]$. $f \mapsto ||f||_p$ is a semi-norm on $\mathcal{L}^p(X, \mathcal{M}, \mu)$.

Proof.

1. $\forall f \in \mathcal{L}^p(X, \mathcal{M}, \mu), ||f||_p \geq 0$

2. $\forall f \in \mathcal{L}^p(X, \mathcal{M}, \mu), \forall \lambda \in \mathbb{K}, ||\lambda f||_p = |\lambda| \, ||f||_p$

3. $\forall (f,g) \in (\mathcal{L}^p(X, \mathcal{M}, \mu))^2, ||f+g||_p \leq ||f||_p + ||g||_p$

4. $||f||_p = 0 \Leftrightarrow f = 0$ almost everywhere

This last property prevents $f \mapsto ||f||_p$ from being a norm: because of the negligeable sets (for the Lebesgue measure), the function f will only be 0 almost everywhere and not necessarily everywhere.

□

7.1.2 L^p Space

Definition 104. Equality almost everywhere.
Let $p \in [1; +\infty]$. We define the equivalence relation of *equality almost everywhere* on $\mathcal{L}^p(X, \mathcal{M}, \mu)$, noted $\diamond$, by

$$f \diamond g \Leftrightarrow f = g \quad \mu \text{ almost everywhere.}$$

Definition 105. Class of equality almost everywhere.
Let $p \in [1; +\infty]$ and $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. The equivalent class of f associated to the equivalence relation $\diamond$ is noted $f^\diamond$ and is defined by

$$f^\diamond = \{g \in \mathcal{L}^p(X, \mathcal{M}, \mu) | g = f \ \mu \text{ almost everywhere}\}.$$

Exercise 89. ★ Equality almost everywhere.
Justify that the relation $\diamond$ we defined is an equivalence relation, i.e. it is reflexive, symmetric, and transitive.

Definition 106. L^p space.
Let $p \in [1; +\infty]$. We define $L^p(X, \mathcal{M}, \mu)$ as the quotient set of $\mathcal{L}^p(X, \mathcal{M}, \mu)$ by $\diamond$, i.e.

$$L^p(X, \mathcal{M}, \mu) = \{f^\diamond | f \in \mathcal{L}^p(X, \mathcal{M}, \mu)\}.$$

Remark 55.
Elements of $L^p(X, \mathcal{M}, \mu)$ are not functions, but classes of equivalence of a function. In particular, we cannot evaluate an element of $L^p(X, \mathcal{M}, \mu)$ in a given point of X, since two functions from the same equivalence class can differ on negligeable sets. We will often, however, identify the equivalence class with its representant.

Proposition 56. Vector space L^p.
For all $p \in [1; +\infty]$, $L^p(X, \mathcal{M}, \mu)$ with $f \mapsto ||f||_p$ is a normed vector space.

Proof.

1. $L^p(X, \mathcal{M}, \mu)$ is a vector space for the same reasons $\mathcal{L}^p(X, \mathcal{M}, \mu)$ is a vector space (subspace, 0, linearity of the integral, and Minkowski inequality).
2. $f \mapsto ||f||_p$ already has the properties of the semi-norm, but now we have the additional property $||f^\diamond||_p = 0 \Leftrightarrow f^\diamond = 0$ since we are working on equivalent classes.

□

Theorem 58. Dominated convergence in $L^p(X, \mathcal{M}, \mu)$.
Let $p \in [1; +\infty[$ and $(f_n)_{n\in\mathbb{N}}$ a sequence of functions in $L^p(X, \mathcal{M}, \mu)$, converging μ-almost everywhere to f. We assume that there exists a function $g \in L^p(X, \mathcal{M}, \mu)$ such that $\forall n \in \mathbb{N}, |f_n| \leq g$ μ-almost everywhere. Then, $f \in L^p(X, \mathcal{M}, \mu)$ and $(f_n)_{n\in\mathbb{N}}$ converges to f with the $||\cdot||_p$ norm

$$\lim_{n\to+\infty} \int_X |f_n - f|^p \mathrm{d}\mu = 0.$$

Proof.

First of all, we deduce from the hypothesis that $|f| \leq g$ μ-almost everywhere. Hence, $f \in L^p(X, \mathcal{M}, \mu)$.

Then, we can apply the "regular" dominated convergence theorem to the sequence defined by, for $n \in \mathbb{N}, h_n = |f_n - f|^p$. For all $n \in \mathbb{N}$, h_n is bounded by the integrable function $|2g|^p$. Since the family of functions $(h_n)_{n\in\mathbb{N}}$ converges to 0 μ-almost everywhere, its integral converges to 0, and the proof is complete.

□

Theorem 59. Riesz-Fisher.
For all $p \in [1; +\infty]$, $L^p(X, \mathcal{M}, \mu)$ with $f \mapsto ||f||_p$ is a complete normed vector space (i.e. $L^p(X, \mathcal{M}, \mu)$ is a Banach space).

Remark 56.

This is a very important theorem in terms of functional spaces, and in the next section we will see why, by comparing this result to the case of Riemann integrable functions.

Proof.

We already know that for all $p \in [1; +\infty]$, $L^p(X, \mathcal{M}, \mu)$ with $f \mapsto ||f||_p$ is a normed vector space. Thus, we have to prove that this is a complete space, i.e. that any Cauchy sequence of $L^p(X, \mathcal{M}, \mu)$ converges in that space. In other words, since we are working with classes of equality, we have to show that if $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence of functions in $L^p(X, \mathcal{M}, \mu)$, then choosing representants of this sequence (in $\mathcal{L}^p(X, \mathcal{M}, \mu)$), there exists a function $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ such that $\lim\limits_{n\to+\infty} ||f_n - f||_p = 0$.

Case $p \in [1; +\infty[$:

We consider a Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $L^p(X, \mathcal{M}, \mu)$. By induction, we can extract a sub-sequence $(f_{n_k})_{k\in\mathbb{N}}$ such that

$$\forall k \in \mathbb{N}, ||f_{n_{k+1}} - f_{n_k}||_p \leq 2^{-k}.$$

For all $N \in \mathbb{N}$, we then define

$$\forall x \in X, g_N(x) = \sum_{k=0}^{N} |f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

This function g_N is in $\mathcal{L}^p(X, \mathcal{M}, \mu)$ and using Minkowski inequality we have

$$\forall N \in \mathbb{N}, ||g_N||_p \leq \sum_{k=0}^{N} ||f_{n_{k+1}} - f_{n_k}||_p \leq \sum_{k=0}^{N} 2^{-k} < 2.$$

Hence, the sequence $(g_N^p)_{N\in\mathbb{N}}$ is an increasing sequence of positive integrable functions, such that

$$\sup_{N\in\mathbb{N}} \int_X g_N^p \mathrm{d}\mu \leq 2^p < +\infty.$$

Therefore, using the monotone convergence theorem, we obtain that

$$\sup_{N\in\mathbb{N}} g_N(x) < +\infty,$$

μ-almost everywhere on X. Again, we can split X in two spaces E and $X \setminus E$ of $\mathcal{M}$, with $\mu(E) = 0$ and

$$\forall x \in X \setminus E, \sum_{k=0}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| < +\infty.$$

Based on these results, we define the function g as

$$g : x \mapsto \begin{cases} \sum_{k=0}^{+\infty} |f_{n_{k+1}}(x) - f_{n_k}(x)| & \text{if } x \in X \setminus E \\ 0 & \text{if } x \in E \end{cases}.$$

We note that $g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, and that $||g||_p \leq 2$.

Now, for all $k \in \mathbb{N}$ we rewrite f_{n_k} as

$$\forall x \in X, f_{n_k}(x) = f_{n_0}(x) + \sum_{j=0}^{k-1}(f_{n_{j+1}}(x) - f_{n_j}(x)),$$

and we define the function f as

$$f : x \mapsto \begin{cases} \lim\limits_{k\to+\infty} f_{n_k}(x) & \text{if } x \in X \setminus E \\ 0 & \text{if } x \in E \end{cases}.$$

Since $(f_{n_k})_{k\in\mathbb{N}}$ converges to f μ-almost everywhere (by construction) and since we have that $\forall k \in \mathbb{N}, |f_{n_k}| \leq |f_{n_0}| + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, we deduce from the dominated convergence theorem that $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. Moreover, we note that for μ-almost every $x \in X$,

$$\lim_{k\to+\infty} |f_{n_k}(x) - f(x)| = 0,$$

and that $|f_{n_k} - f| \leq |f_{n_0} - f| + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. Hence, this shows using the dominated convergence theorem that

$$\lim_{k\to+\infty} ||f_{n_k} - f||_p = 0,$$

and since $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, the limit is unique when it exists, so

$$\lim_{n\to+\infty} ||f_n - f||_p = 0.$$

Thus, every Cauchy sequence of functions in $\mathcal{L}^p(X, \mathcal{M}, \mu)$ converges μ-almost everywhere to a function in $\mathcal{L}^p(X, \mathcal{M}, \mu)$. Since we are working with representants of classes of equality, this means that every Cauchy sequence of functions in $L^p(X, \mathcal{M}, \mu)$ converges to a function in $L^p(X, \mathcal{M}, \mu)$. The space $L^p(X, \mathcal{M}, \mu)$ is complete.

Case $p = +\infty$:

We consider a Cauchy sequence $(f_n)_{n\in\mathbb{N}}$ of functions in $L^\infty(X, \mathcal{M}, \mu)$. For $(n, m) \in \mathbb{N}^2$ we define

$$A_n = \{x \in X |\ |f_n(x)| > ||f_n||_\infty\},$$

$$B_{n,m} = \{x \in X |\ |f_n(x) - f_m(x)| > ||f_n - f_m||_\infty\},$$

$$E = \left(\bigcup_{n\in\mathbb{N}} A_n\right) \cup \left(\bigcup_{(n,m)\in\mathbb{N}^2} B_{n,m}\right).$$

Since we are working in $L^\infty(X, \mathcal{M}, \mu)$, we note that $\forall n \in \mathbb{N}, \mu(A_n) = 0$, and $\forall (n, m) \in \mathbb{N}^2, \mu(B_{n,m}) = 0$, so $\mu(E) = 0$. By construction, we also have that

$$\forall x \in X \setminus E, \lim_{(n,m)\to+\infty} |f_n(x) - f_m(x)| \leq \lim_{(n,m)\to+\infty} ||f_n - f_m||_\infty = 0.$$

We define the function f as

$$f : x \mapsto \begin{cases} \lim\limits_{n\to+\infty} f_n(x) & \text{if } x \in X \setminus E \\ 0 & \text{if } x \in E \end{cases}.$$

Then, f is measurable and for all $x \in X$ we have $|f(x)| \leq \liminf_{n \to +\infty} ||f_n||_\infty$, so $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$ and $||f||_\infty < +\infty$. Similarly, we have for all $n \in \mathbb{N}$,

$$\forall x \in X \setminus E, |f_n(x) - f(x)| \leq \liminf_{m \to +\infty} ||f_n - f_m||_\infty,$$

so

$$\lim_{n \to +\infty} ||f_n - f||_\infty \leq \lim_{n \to +\infty} \liminf_{m \to +\infty} ||f_n - f_m||_\infty = 0.$$

We conclude that $(f_n)_{n \in \mathbb{N}}$ converges μ-almost everywhere to f in $\mathcal{L}^\infty(X, \mathcal{M}, \mu)$. The same conclusion as in the general case applies: every Cauchy sequence of functions in $L^\infty(X, \mathcal{M}, \mu)$ converges to a function in $L^\infty(X, \mathcal{M}, \mu)$. The space $L^\infty(X, \mathcal{M}, \mu)$ is complete.

□

Corollary 18. Hilbert space $L^2(X, \mathcal{M}, \mu)$.
$L^2(X, \mathcal{M}, \mu)$ is a Hilbert space with the scalar product defined by

$$\langle f | g \rangle = \int_X f \overline{g} \mathrm{d}\mu.$$

Proof.
We already know that $L^2(X, \mathcal{M}, \mu)$ is a vector space. We simply have to show that the application defined by

$$\langle f | g \rangle = \int_X f \overline{g} \mathrm{d}\mu,$$

is a scalar product, i.e. a bi-linear form, symmetric, definite and positive (if $X \subset \mathbb{R}$), or a sesqui-linear form, hermitian, definite, and positive (if $X \subset \mathbb{C}$). We proceed on $\mathbb{C}$, since the result on $\mathbb{R}$ will naturally follow.

1. The hermitian part is a consequence of the linearity of the integral

$$\forall (f, g) \in \left(L^2(X, \mathcal{M}, \mu)\right)^2, \langle f | g \rangle = \overline{\langle g | f \rangle}$$

2. Using the linearity of the Lebesgue integral, we also directly have the sesqui-linearity

$$\forall (f, g, h) \in \left(L^2(X, \mathcal{M}, \mu)\right)^3, \forall \lambda \in \mathbb{C}, \langle f + \lambda h | g \rangle = \langle f | g \rangle + \lambda \langle h | g \rangle$$

$$\forall (f, g, h) \in \left(L^2(X, \mathcal{M}, \mu)\right)^3, \forall \lambda \in \mathbb{C}, \langle f | g + \lambda h \rangle = \langle f | g \rangle + \overline{\lambda} \langle f | h \rangle$$

3. It is clearly positive, since

$$\forall f \in L^2(X, \mathcal{M}, \mu), \langle f | f \rangle = \int_X |f|^2 \mathrm{d}\mu.$$

4. Is is also definite, since if we have

$$\langle f | f \rangle = \int_X |f|^2 \mathrm{d}\mu = 0,$$

then the function $|f|$, and by consequent f, is equal to the function $x \mapsto 0$ μ-almost everywhere, and therefore is equal to it in the sense of "class of equality". Note that this is true only in the L^2 space and not in $\mathcal{L}^2$.

Moreover, we easily verify that the scalar product do not depend on the representant of the class $f^\diamond$ since two functions of the same class only differ on a μ negligeable set, and the integral over this set is 0. The Riesz-Fiser theorem also gives us that $L^2(X, \mathcal{M}, \mu)$ is a complete space. $\square$

Remark 57.
All the other L^p spaces, for $p \neq 2$, are not Hilbert spaces.

Corollary 19. Cauchy-Schwarz-Buniakovski inequality.
If $(f, g) \in (L^2(X, \mathcal{M}, \mu))^2$, then

$$\left|\int_X fg \mathrm{d}\mu\right| \leq \left(\int_X |f|^2 \mathrm{d}\mu\right)^{\frac{1}{2}} \left(\int_X |g|^2 \mathrm{d}\mu\right)^{\frac{1}{2}}.$$

Proof.
This is simply an application of the Hölder inequality for $p = q = 1/2$, with the scalar product of f and $\overline{g}$, $\langle f|\overline{g}\rangle$. $\square$

7.1.3 Riemann Integrable vs. Lebesgue Integrable

Definition 107. Riemann integrable space $R^p(X)$.
Let $p \in [1; +\infty[$. We note $R^p(X)$ the space defined by

$$R^p(X) = \left\{ f : X \to \mathbb{K} \text{ Riemann integrable} \,\middle|\, \int_X |f(x)|^p \mathrm{d}x < +\infty. \right\}.$$

Example 23.

1. For $p = 1$, $R^1(X)$ is the set of the functions that are Riemann integrable on X.
2. For $p = 2$, $R^2(X)$ is the set of the functions f such that f^2 is Riemann integrable on X.

Remark 58.

1. In this definition, it really is the Riemann integral, and not the Lebesgue integral.
2. In general, for $p \in \mathbb{N}$, the set $R^p(X)$ is the set of functions whose p^{th} power is Riemann integrable. As mentioned, the two common examples are $R^1(X)$ and $R^2(X)$.
3. These sets are often noted $L^p(X)$ in the literature, when discussing the Riemann integration. Here, to use a different notation than the Lebesgue spaces L^p defined previously, we will note them R^p.

Proposition 57. Riemann vector space $R^p(X)$.
Let $p \in [1; +\infty[$. The space $R^p(X)$ is a vector space.

Proof.

1. It is a sub-space of the vector space $\mathcal{C}(X)$.
2. $0 \in R^p(X)$ (the constant function equal to 0^p is Riemann integrable).

3. If $f \in R^p(X)$ and $\lambda \in \mathbb{K}$, then $\lambda f \in R^p(X)$.
4. If f and g are two elements of $R^p(X)$, then using Minkowski (the norm p works here too) we show that $|f+g|^p$ is Riemann integrable, so $f+g \in R^p(X)$.

□

Theorem 60. Non-completeness of $R^1(X)$.
The space $R^1(X)$ is not complete.

Proof.
We consider $R^1(I)$ where $I = [0;\, 1]$ is an interval, and the usual norm defined by

$$\forall f \in R^1(I), ||f||_1 = \int_I |f(x)|\mathrm{d}x.$$

If it is complete, a vectorial space is closed. We will therefore show that this is not the case, by constructing a sequence of functions for which each term belongs to $R^1(I)$, but that converges to $f \notin R^1(I)$.

For all $n \in \mathbb{N}$, we introduce f_n as

$$f_n : x \mapsto \begin{cases} 0 & \text{if } x \in [0;\, 1/2[\\ (x-1/2)\times 2^n & \text{if } x \in [1/2;\, 1/2+1/2^{n+1}] \\ 1 & \text{if } x \in]1/2+1/2^{n+1};\, 1] \end{cases}.$$

The sequence of functions $(f_n)_{n\in\mathbb{N}}$ converges to f defined by

$$f : x \mapsto \begin{cases} 0 & \text{if } x \in [0;\, 1/2[\\ 1/2 & \text{if } x = 1/2 \\ 1 & \text{if } x \in]1/2;\, 1] \end{cases}.$$

In fact, we note that we have

$$\forall n \in \mathbb{N}, \int_0^1 |f(x) - f_n(x)|\mathrm{d}x \leq \frac{1}{2^{n+1}}.$$

We now show that $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence:
Let $(p,q) \in \mathbb{N}^2$ and $n_0 \in \mathbb{N}$ such that $q \geq p \geq n_0$. Then we have

$$||f_p - f_q||_1 \leq \max\left(\frac{1}{2^p}, \frac{1}{2^q}\right) = \frac{1}{2^p} \leq \frac{1}{2^{n_0}}.$$

Yet,

$$\lim_{n_0\to+\infty} \frac{1}{2^{n_0}} = 0.$$

So,

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}, \forall (p,q) \in \mathbb{N}^2, p > q \geq n_0, ||f_p - f_q||_1 < \varepsilon.$$

Hence, $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

We now study its convergence:
Let assume that $(f_n)_{n\in\mathbb{N}}$ converges to $g \in R^1(I)$. Then we automatically have

$$\lim_{n\to+\infty} ||g - f_n||_1 = 0.$$

We can write, using the definition of the functions f_n, that

$$\forall n \in \mathbb{N}, \int_0^{1/2} |g(x) - f_n(x)|\mathrm{d}x \leq ||g - f_n||_1,$$

then

$$\forall n \in \mathbb{N}, \int_0^{1/2} |g(x)|\mathrm{d}x \leq ||g - f_n||_1.$$

So, looking at the limit $n \to +\infty$, we obtain that

$$\int_0^{1/2} |g(x)|\mathrm{d}x = 0,$$

and since g is continuous on $[0;\, 1/2]$ (because it is a function in $R^1(I)$), then we conclude that $\forall x \in [0;\, 1/2], g(x) = 0$.

Now, let $a > 1/2$. There exists $n_0 \in \mathbb{N}$ such that $\forall n \in \mathbb{N}, n \geq n_0, 1/2 + 1/2^{n+1} \leq a$. Then, we know that for all $n \in \mathbb{N}$ such that $n \geq n_0$, we have $\forall x \in [a;\, 1], f_n(x) = 1$. Moreover, we note that

$$\forall n \geq n_0, \int_a^1 |g(x) - f_n(x)|\mathrm{d}x = \int_a^1 |g(x) - 1|\mathrm{d}x \leq ||g - f_n||_1.$$

Again, looking at the limit $n \to +\infty$, we obtain that

$$\forall a > 1/2, \int_a^1 |g(x) - 1|\mathrm{d}x = 0,$$

so $\forall a > 1/2, \forall x \geq a, g(x) = 1$. This means that $\forall x \geq 1/2, g(x) = 1$.

Finally, we note that

$$\lim_{x \to \frac{1}{2}^+} g(x) = 1 \neq \lim_{x \to \frac{1}{2}^-} g(x) = 0.$$

Therefore, g is not continuous at $1/2$ and is not in $R^1(I)$. This conclude the proof: $R^1(I)$ is not complete.

□

THEOREM 61. Non-completeness of $R^2(X)$.
The space $R^2(X)$ is not complete.

EXERCISE 90. ★★★ Non-completeness of $R^2(X)$.
Drawing from the proof on the non-completeness of $R^1(X)$, show that $R^2(X)$ is not complete.

THEOREM 62. Non-completeness of $R^p(X)$ spaces.
Let $p \in [1; +\infty[$. In general, $R^p(X)$ is not complete.

REMARK 59.

This is a crucial difference between the Riemann and the Lebesgue integration. We have already seen that the set of functions that are Lebesgue integrable is larger than the set of functions that are Riemann integrable: Riemann integrable functions are also Lebesgue integrable, but there are functions integrable in the Lebesgue theory and not in the Riemann theory. This theorem is, in a way, a consequence of this difference between the sets of integrable functions. Sequences of Riemann integrable functions may converge towards a well defined limit that is not Riemann integable. Sequences of Lebesgue integrable functions, however, have their pointwise limit (if it exists) still Lebesgue integrable.

7.2 Inclusions of Functional Spaces

LEMMA 7. Inclusions of functional spaces.
There is no general rule for the inclusion of functional spaces R^p, L^p, and $\mathcal{L}^p$.

REMARK 60.
This is not a "theorem" or a "proposition", so to speak, but simply something to keep in mind. We will see, in this section, that the functional spaces are sometimes comparable, but almost always with underlying assumptions. Inclusions of functional spaces of integrable functions are false in general...

7.2.1 Riemann Spaces

PROPOSITION 58. Inclusion of $R^p(X)$ spaces.
Let I be a bounded interval of $\mathbb{R}$. Let $(p, q) \in [1; +\infty]^2$. If $p \leq q$, then we have the following inclusion

$$R^q(I) \subset R^p(I).$$

EXAMPLE 24.
For example, if I is an interval of $\mathbb{R}$, then $R^2(I) \subset R^1(I)$: for any (continuous) function f such that f^2 is Riemann integrable on I, we also have that f is Riemann integrable on I.

REMARK 61.
One should be very careful with this proposition. In general, we only have an inclusion. Sometimes, however, we have an equality.

- If $I = [0; 1]$, then $\mathcal{C}(I) = R^1(I) = R^2(I) = R^\infty(I)$. Why? Because on a compact, a continuous function f is bounded so there exists a constant $M \in \mathbb{R}$ such that $\forall x \in I, f(x) \leq M$. Then, a constant function is always Riemann integrable, so f is in all the Riemann integrable spaces.

- If $I =]0; 1]$, then $\mathcal{C}(I) \supsetneq R^1(I) \supsetneq R^2(I) \supsetneq R^\infty(I)$. Why? For example, if we consider $f : x \mapsto 1/\sqrt{x}$, then $f \in R^1(I)$ (integrable in 0); but $f^2 : x \mapsto 1/x$, and $f^2 \notin R^2(I)$ (not integrable in 0).

- If the interval is not bounded, we do not have these inclusions. For example, consider the interval $I = [1; +\infty[$ and $f : x \mapsto 1/x$. Then, $f^2 : x \mapsto 1/x^2$ and $f^2 \in R^2(I)$ (integrable in $+\infty$) but $f \notin R^1(I)$ (not integrable in $+\infty$).

PROPOSITION 59. Inclusion of $R^p(X)$ in $\mathcal{L}^p(, \mathcal{M}, \mu)$.
Let I be an interval of $\mathbb{R}$. Let $p \in [1; +\infty]$. We have

$$R^p(I) \subset \mathcal{L}^p(I, \mathcal{B}(I), \mu).$$

PROOF.
This is simply a consequence of the fact that a Riemann integrable function $f \in R^p(I)$ is also Lebesgue integrable, and therefore $f \in \mathcal{L}^p(I, \mathcal{B}(I), \mu)$.

□

7.2.2 Lebesgue Spaces

Proposition 60. Inclusion of $\mathcal{L}^p(X, \mathcal{M}, \mu)$ spaces.
Let X be a non-negligeable set of finite measure, $\mu(X) < +\infty$. Let $(p, q) \in [1; +\infty]^2$. If $p \leq q$, then we have the following inclusion

$$\mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu).$$

Proof.
With the notation of the proposition, we assume that $f \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ we apply Hölder inequality to the functions $|f|^p$ and 1, which yields

$$\begin{aligned} \int_X |f|^p \mathrm{d}\mu &= \int_X 1|f|^p \mathrm{d}\mu \\ &\leq \left(\int_X |f|^{pq/p} \mathrm{d}\mu\right)^{p/q} \left(\int 1 \mathrm{d}\mu\right)^{1-p/q} \\ &\leq \left(\int_X |f|^q \mathrm{d}\mu\right)^{p/q} \mu(X)^{1-p/q}. \end{aligned}$$

Therefore, we obtain

$$||f||_p \leq ||f||_q \mu(x)^{\frac{1}{p} - \frac{1}{q}},$$

and since $\mu(x) < +\infty$ we deduce that $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$, so $\mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu)$. □

Proposition 61. Limit of the p-norm.
Let $r \in [1; +\infty[$ and $f \in \mathcal{L}^r(X, \mathcal{M}, \mu)$. Then

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

Proof.
We first note that if $||f||_\infty = 0$ then the result is trivial, since the essential bound of f is equal to 0. We therefore assume that $||f||_\infty > 0$.

Let $t \in \mathbb{R}$ such that $0 \leq t < ||f||_\infty$. Thanks to the definition of the ∞-norm, we know that the set

$$A = \{x \in X | \ |f(x)| \geq t\},$$

is not negligeable: $\mu(A) > 0$. A is a subset of X so we can write

$$\begin{aligned} ||f||_p &\geq \left(\int_A |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}} \\ &\geq (t^p \mu(A))^{1/p} \\ &\geq t\mu(A)^{1/p}. \end{aligned}$$

If $0 < \mu(A) < \infty$, then $\lim\limits_{p \to +\infty} \mu(A)^{1/p} = 1$, and if $\mu(A) = +\infty$, then $\lim\limits_{p \to +\infty} \mu(A)^{1/p} = +\infty$, but both cases lead to the same result

$$\liminf_{p \to +\infty} ||f||_p \geq t.$$

The choice of t being arbitrary but such that $0 \leq t < ||f||_\infty$, we deduce

$$\liminf_{p \to +\infty} ||f||_p \geq ||f||_\infty.$$

Let $p > r$ (we assumed that $f \in \mathcal{L}^r(X, \mathcal{M}, \mu)$). Then, we can show that (cf. exercise 88)

$$||f||_p \leq ||f||_r^{r/p} ||f||_\infty^{1-r/p}.$$

The r-norm of f is finite, which implies

$$\limsup_{p \to +\infty} ||f||_p \leq ||f||_\infty.$$

Hence, we have

$$\limsup_{p \to +\infty} ||f||_p \leq ||f||_\infty \leq \liminf_{p \to +\infty} ||f||_p,$$

and therefore

$$\lim_{p \to +\infty} ||f||_p = ||f||_\infty. \qquad \square$$

PROPOSITION 62. Reverse inclusion of $\mathcal{L}^p(X, \mathcal{M}, \mu)$ spaces.
Let assume that $\mu_0(X) > 0$, where μ_0 is defined by

$$\mu_0(X) = \inf\left(\{\mu(E) | E \in \mathcal{M}, \mu(E) > 0\}\right).$$

Let $(p, q) \in [1; +\infty]^2$ with $p \leq q$. Then we have the following inclusion

$$\mathcal{L}^p(X, \mathcal{M}, \mu) \subset \mathcal{L}^q(X, \mathcal{M}, \mu).$$

REMARK 62.

These propositions show that, if $(p, q) \in [1; +\infty]^2$ with $p \leq q$, depending on the assumptions on the set X we can either have $\mathcal{L}^p(X, \mathcal{M}, \mu) \subset \mathcal{L}^q(X, \mathcal{M}, \mu)$ or $\mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu)$...

PROOF.

With the notations of the proposition, let $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ with $p < +\infty$. If there exists a constant $A > 0$ and a set $E \in \mathcal{M}$ with $\mu(E) > 0$ such that $|f(x)| \geq A$ μ-almost everywhere on E, then

$$\int_X |f|^p \mathrm{d}\mu \geq \int_E A^p \mathrm{d}\mu = A^p \mu(E) \geq A^p \mu_0(X).$$

Therefore, we have

$$A \leq ||f||_p \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{p}},$$

and $f \in \mathcal{L}^\infty(X, \mathcal{M}, \mu)$ (essentially bounded) and

$$||f||_\infty \leq ||f||_p \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{p}}.$$

Now, let $q \geq p$. We have

$$\int_X |f|^q \mathrm{d}\mu = \int_X |f|^{q-p} |f|^p \mathrm{d}\mu \leq ||f||_\infty^{q-p} ||f||_p^p \leq ||f||_p^q \left(\frac{1}{\mu_0(X)}\right)^{\frac{q-p}{p}}.$$

Hence

$$||f||_q \leq ||f||_p \left(\frac{1}{\mu_0(X)}\right)^{\frac{1}{p} - \frac{1}{q}},$$

from which we deduce the inclusion

$$\mathcal{L}^p(X, \mathcal{M}, \mu) \subset \mathcal{L}^q(X, \mathcal{M}, \mu). \qquad \square$$

EXERCISE 91. ★ x^α functions.

Let $X \subset \mathbb{R}$. Based on the previous discussions, in which $\mathcal{L}^p(X, \mathcal{M}, \mu)$ spaces are the functions $x \mapsto x^\alpha$, with $\alpha \in \mathbb{R}$?

7.2.3 Local Lebesgue Spaces

Definition 108. Local $\mathcal{L}^p_{\text{loc}}$ spaces.
Let $p \in [1;+\infty[$, and G be an open set of $\mathbb{R}^n$. The *local* $\mathcal{L}^p$ *space* on G is the set of all the measurable functions defined μ-almost everywhere on G such that for every compact $K \subset G$, the function $(f\chi_K)^p$ is integrable (i.e. $f\chi_K$ has a finite p-norm, or is essentially bounded in the case $p = +\infty$. We note this set

$$\mathcal{L}^p_{\text{loc}}(G, \mathcal{M}, \mu).$$

Proposition 63. Inclusion of the local $\mathcal{L}^p_{\text{loc}}$ spaces.
Let $(p,q) \in [1;+\infty[^2$, with $p < q$. We have

$$L^\infty_{\text{loc}}(G) \subset L^q_{\text{loc}}(G) \subset L^p_{\text{loc}}(G) \subset L^1_{\text{loc}}(G).$$

7.2.4 Convexity Relations

Proposition 64. Convexity.
Let $(p,q,r) \in [1;+\infty[$, with $p < r < q$. Let $f \in \mathcal{L}^p \cap \mathcal{L}^q$. Then, $f \in \mathcal{L}^r$ and

$$\log(||f||_r) \leq \frac{\frac{1}{r} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{q}} \log(||f||_p) + \frac{\frac{1}{p} - \frac{1}{r}}{\frac{1}{p} - \frac{1}{q}} \log(||f||_q).$$

Proof.
With the notations of the proposition, given $p < r < q$ we have $q^{-1} < r^{-1} < p^{-1}$ and there exists a unique s such that

$$\frac{1}{r} = \frac{s}{p} + \frac{1-s}{q}.$$

with $0 < s < 1$. We can compute this number and we find that

$$s = \frac{\frac{1}{r} - \frac{1}{q}}{\frac{1}{p} - \frac{1}{q}} \qquad \text{and} \qquad 1 - s = \frac{\frac{1}{p} - \frac{1}{r}}{\frac{1}{p} - \frac{1}{q}}.$$

We note that

$$\frac{rs}{p} + \frac{r(1-s)}{p} = 1,$$

so rs/p and $r(1-s)/p$ are Hölder conjugates. We can therefore use Hölder inequality, assuming that f is positive, and we find

$$\begin{aligned}
||f||_r &= ||f^s f^{1-s}||_r &(7.1)\\
&= ||f^{rs} f^{r(1-s)}||_1^{1/r} &(7.2)\\
&\leq \left(||f^{rs}||_{p/rs} ||f^{r(1-s)}||_{q/r(1-s)}\right)^{1/r} &(7.3)\\
&\leq \left(||f||_p^{rs} ||f||_q^{r(1-s)}\right)^{1/r} &(7.4)\\
&\leq ||f||_p^s ||f||_q^{1-s}. &(7.5)
\end{aligned}$$

Taking the log of this expression, we obtain the result. □

7.3 Density

Definition 109. Dense space.
Let X be a set, and $A \subset X$. A is said to be *dense* in X if any non-empty open subset of X contains at least one element of A.

Remark 63.
This definition is equivalent to other definitions, notably: A dense in X if and only if $\overline{A} = X$.

Example 25.
$\mathbb{Q}$ is dense in $\mathbb{R}$.

Proposition 65. Sequential characterization of the density.
Let X be a set, and $A \subset X$. A dense in X if any element $x \in X$ can be written as a limit of a sequence $(a_n)_{n\in\mathbb{N}}$ of elements in A.

Exercise 92. ★ Density of $\mathbb{Q}$ in $\mathbb{R}$.
Prove that $\mathbb{Q}$ is dense in $\mathbb{R}$.

Lemma 8. Sequence of simple functions.
Let f be a positive measurable function defined on X. There exists a monotone increasing sequence of positive simple functions $(f_n)_{n\in\mathbb{N}}$ converging pointwise to f. Moreover, if f is bounded, the convergence is uniform.

Proof.
Let $n \in \mathbb{N}$. We write the interval $[0; 2^n]$ as the union of 2^{2n} sub-intervals of length 2^{-n}, defined by

$$\forall k \in [\![0;\, 2^{2n}-1]\!], I_{k,n} =]k2^{-n};\, (k+1)2^{-n}].$$

and we define $J_n = [2^n; +\infty[$. Then, we introduce the sets

$$A_{k,n} = f^{-1}(I_{k,n}) \qquad \text{and} \qquad B_n = f^{-1}(J_n).$$

We can now write a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ satisfying the criteria of the lemma, as

$$\forall n \in \mathbb{N}, f_n = \sum_{k=1}^{2^n-1} k2^{-n}\chi_{A_{k,n}} + 2^n\chi_{B_n}.$$

□

Proposition 66. Density of simple functions.
Let $p \in [1; +\infty]$. The set of simple functions that belong to $L^p(X, \mathcal{M}, \mu)$ is dense in $L^p(X, \mathcal{M}, \mu)$.

Remark 64.
Why is this result important? In previous proofs, we already used it without explicitly formulating it. The density of simple functions in L^p means that, with the right convergence theorems, it may be only necessary to work with simple functions and then extend the results thanks to their density.

Proof.

We note that, since a function can be expressed in terms of its positive and negative part, we can prove this proposition only for positive functions (and then extend it naturally to any function). We will therefore prove that we can approximate any positive function f by simple functions.

Case $p \in [1; +\infty[$:

Let $f \in L^p(X, \mathcal{M}, \mu)$. The lemma tells us that there exists an increasing sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ converging pointwise to f. These simple functions belong to $L^p(X, \mathcal{M}, \mu)$ by definition. We have, for all $n \in \mathbb{N}$, that

$$|f - f_n|^p \leq |f|^p,$$

so, using the dominated convergence theorem, we have

$$\lim_{n\to+\infty} \int_X |f - f_n|^p \mathrm{d}\mu = 0,$$

and f is the limit in $L^p(X, \mathcal{M}, \mu)$ of a sequence of simple functions.

Case $p = +\infty$:

Let $f \in L^\infty(X, \mathcal{M}, \mu)$, and let choose a representant of f that is bounded everywhere on X. Then, using the previous lemma, there exists a sequence of simple functions that converges uniformly to f, and belonging to $L^\infty(X, \mathcal{M}, \mu)$.

□

Remark 65.

This is not the only result on density with the Lebesgue spaces. For example, under certain assumptions, the Lipschitzian functions are dense in the Lebesgue spaces. Here, we will see another important density result.

Definition 110. Continuous functions with compact support.
We note $\mathcal{C}_c(X)$ the space of continuous functions with a compact support, i.e. the space of functions f such that there exists a compact $K \subset X$ such that $f = 0$ on $X \setminus K$.
For $n \in \mathbb{N}$, we note $\mathcal{C}_c^n(X)$ the space of continuous functions with a compact support whose n^{th} derivative is continuous.

Exercise 93. ★ Vector space $\mathcal{C}_c(X)$.
Prove that $\mathcal{C}_c(X)$ is a vector space. Is it a normed vector space?

Proposition 67. Density of continuous functions with compact support.
Let $p \in [1; +\infty[$. Then for all $n \in \mathbb{N}$, $\mathcal{C}_c^n(X)$ is dense in $L^p(X, \mathcal{M}, \mu)$.

Remark 66.

There are different ways to prove this theorem. Here, I will follow an approach that uses only the notions we have covered; but first, we need two lemmas.

Lemma 9. Regularity of the Lebesgue measure.
Let A be a Borelian of $\mathbb{R}^n$ such that $\mu(A) < +\infty$. Then, for all $\varepsilon > 0$, there exists an open set U and a compact K such that

$$K \subset A \subset U \qquad \text{and} \qquad \mu(U \setminus K) < \varepsilon.$$

Proof.
This is simply a consequence of theorem 18 on the exterior and interior approximations of a measurable set, applied in $\mathbb{R}^n$.
□

Exercise 94. ★★ Regularity of the Lebesgue measure.
Prove the previous lemma, using the results from theorem 18.

Lemma 10. Approximation by a continuous function with compact support.
Let U be an open set of $X \subset \mathbb{R}^n$ and K a compact such that $K \subset U$. Then, we can find a function $f \in \mathcal{C}_c(X)$ such that for all $x \in K$, $f(x) = 1$ and for all $x \in X \setminus U$, $f(x) = 0$, with $0 \leq f \leq 1$.

Proof.
Since K is compact and U open, we can find a bounded open set W such that

$$K \subset W \qquad \text{and} \qquad \overline{W} \subset U.$$

We can take, for example $W = \{x \in \mathbb{R}^n | d(x, K) < \varepsilon\}$ for $\varepsilon > 0$ small enough. Then, we have

$$\forall x \in \mathbb{R}^n, d(x, K) + d(x, \mathbb{R}^n \setminus W) > 0.$$

This can be justified as follows: K and $\mathbb{R}^n \setminus W$ are closed sets, so for $x \in \mathbb{R}^n$, either $x \notin K$ or $x \notin \mathbb{R}^n \setminus W$, so at least one of the two distances is non-zero.
Thus, we can define $f : X \to \mathbb{R}$ by

$$f : x \mapsto \frac{d(x, \mathbb{R}^n \setminus W)}{d(x, K) + d(x, \mathbb{R}^n \setminus W)}.$$

This function f is continuous, and satisfies that for all $x \in K$, $f(x) = 1$ and for all $x \in X \setminus U$, $f(x) = 0$, with $0 \leq f \leq 1$. Moreover, f is equal to zero if $x \notin \overline{W}$, which is a closed and bounded set, hence a compact, contained in X.
□

Proof.
We will do the proof of the proposition only for $\mathcal{C}_c(X)$, but it works similarly for the other spaces.
We first note that, since the space of simple functions is dense in $L^p(X, \mathcal{M}, \mu)$, then we only need to show that the closure of $\mathcal{C}_c(X)$ contains all the simple functions in $L^p(X, \mathcal{M}, \mu)$. Moreover, since $\mathcal{C}_c(X)$ is a vector space, its closure $\overline{\mathcal{C}_c(X)}$ is also a vector space. An easy way to proceed now is to show that $\overline{\mathcal{C}_c(X)}$ contains all the characteristic functions of $L^p(X, \mathcal{M}, \mu)$, i.e. all the functions of the form χ_A with $A \subset X$ and $\mu(A) < +\infty$ (then, by linearity, we will have the result).
Let $A \subset X$ with $\mu(A) < +\infty$, and $\varepsilon > 0$. We have to prove that there exists a function $f \in \mathcal{C}_c(X)$ such that $||f - \chi_A||_p < \varepsilon$. We will use the two lemmas we have just discussed.
The first lemma tells us that we can find a compact K and an open set U such that $K \subset A \subset U$ and $\mu(U \setminus K) < \varepsilon$.
The second lemma tells us that there exists a function $f \in \mathcal{C}_c(X)$ such that $f = 1$ on K and $f = 0$ on $X \setminus U$, with $0 \leq f \leq 1$.
By definition of f, we have that $\chi_K \leq f \leq \chi_U$. Moreover, by definition of the sets K and U, we also know that $\chi_K \leq \chi_A \leq \chi_U$. Therefore, we have

$$|f - \chi_A| \leq \chi_U - \chi_K = \chi_{U \setminus K}.$$

Hence

$$\begin{aligned}||f-\chi_A||_p^p &= \int_X |f-\chi_A|^p \mathrm{d}\mu \\ &\leq \int_X \chi_{U\setminus K}^p \mathrm{d}\mu \\ &\leq \mu(U\setminus K) \\ &< \varepsilon,\end{aligned}$$

and since ε is arbitrary, this proves the proposition.

□

Exercise 95. ★★★ Non-density in L^∞.
Show that $\mathcal{C}_c(X)$ is not dense in $L^\infty(X,\mathcal{M},\mu)$ (you can show that the constant function equal to 1 is not in the closure of $\mathcal{C}_c(X)$ for the ∞-norm).

7.4 L^p Spaces for $p<1$

Definition 111. $\mathcal{L}^p$ space.
Let $p\in]0;+\infty[$. We define

$$\mathcal{L}^p(X,\mathcal{M},\mu)=\left\{f:X\to\mathbb{K} \text{ measurable}\,\middle|\int_X |f|^p\mathrm{d}\mu<+\infty\right\}.$$

Definition 112. $\mathcal{L}^p$-norm.
Let $p\in]0;+\infty[$. For $f\in\mathcal{L}^p(X,\mathcal{M},\mu)$, we define

$$||f||_p=\left(\int_X |f|^p\mathrm{d}\mu\right)^{\frac{1}{p}}.$$

Remark 67.
The definition of the L^p spaces follows naturally by working on equivalence classes.

Proposition 68. Additive inequality.
Let $p\in]0;1[$. If f and g are two functions in $\mathcal{L}^p(X,\mathcal{M},\mu)$, then $f+g\in\mathcal{L}^p(X,\mathcal{M},\mu)$ and

$$||f+g||_p^p\leq ||f||_p^p+||g||_p^p.$$

Proof.
Let $y\in\mathbb{R}^{+\star}$. We define the function

$$\begin{aligned}h:\mathbb{R}^{+\star} &\to \mathbb{R}\\ x &\mapsto (x+y)^p-a^p.\end{aligned}$$

We have $\forall x\in\mathbb{R}^{+\star}, h'(x)=p\left((x+y)^{p-1}-x^{p-1}\right)$ and, since $p<1$, $x>0$ and $y>0$, we have $\forall x\in\mathbb{R}^{+\star}, h'(x)<0$. Thus, the function h is a decreasing function of x and is bounded by its limit when x goes to zero, which is y^p. Therefore, we have

$$\forall (x,y)\in(\mathbb{R}^{+\star})^2, (x+y)^p\leq x^p+y^p.$$

For all $x \in X$, we can apply this result to $f(x)$ and $g(x)$, so we obtain

$$\forall x \in X, |f(x) + g(x)|^p \leq |f(x)|^p + |g(x)|^p.$$

Note that the absolute value takes care of the signs of $f(x)$ and $g(x)$.

Using the monotonicity of the Lebesgue integral, we can therefore write the inequality

$$\int_X |f+g|^p \mathrm{d}\mu \leq \int_X |f|^p \mathrm{d}\mu + \int_X |g|^p \mathrm{d}\mu.$$

□

Proposition 69. p-metric.
Let $p \in]0;1[$. The $\mathcal{L}^p$-norm defines a metric on $L^p(X, \mathcal{M}, \mu)$, i.e. it defines a distance d: if f and g are two functions in $\mathcal{L}^p(X, \mathcal{M}, \mu)$, then

$$d(f,g) = ||f-g||_p^p.$$

Remark 68.
The $\mathcal{L}^p$-norm is not a norm on the $L^p(X, \mathcal{M}, \mu)$ spaces, since it does not satisfy all the properties of a norm. It can be used, however, to define a distance (and a metric).

Exercise 96. ★ p-distance.
Prove that $d(f,g) = ||f-g||_p^p$ defines a distance on $L^p(X, \mathcal{M}, \mu)$.

Exercise 97. ★ Triangle inequality.
Prove that the triangle inequality does not hold for $||f||_p$ if $0 < p < 1$.

Proposition 70. Simili-Hölder inequality.
Let f and g be positive measurable functions defined on X, and p and q such that

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

Then we have

$$\int_X fg \mathrm{d}\mu \geq \left(\int f^p \mathrm{d}\mu\right)^{\frac{1}{p}} \left(\int g^q \mathrm{d}\mu\right)^{\frac{1}{q}}.$$

Proof.
We first note that the case $f = 0$ or $g = 0$ are trivial, as well as the case in which the product fg is not bounded. We therefore assume that f and g are strictly positive and bounded.

Now, we note that, if

$$\frac{1}{p} + \frac{1}{q} = 1,$$

then q is defined by $q = p/(p-1)$.

We will apply Hölder inequality for $1/p$ and its conjugated exponent, i.e. $1/(1-p)$.

We write

$$\begin{aligned}
\int_X f^p \mathrm{d}\mu &= \int_X (fg)^p g^{-p} \mathrm{d}\mu \\
. \leq & \left(\int_X fg \mathrm{d}\mu\right)^p \left(\int_X g^{-p/(1-p)} \mathrm{d}\mu\right)^{1-p} && \text{using Hölder inequality} \\
&\leq \left(\int_X fg \mathrm{d}\mu\right)^p \left(\int_X g^q \mathrm{d}\mu\right)^{-p/q}.
\end{aligned}$$

Now, taking the power $1/p$ on each side, we obtain

$$\int_X fg \mathrm{d}\mu \geq \left(\int f^p \mathrm{d}\mu\right)^{\frac{1}{p}} \left(\int g^q \mathrm{d}\mu\right)^{\frac{1}{q}}.$$

□

Exercise 98. ★★ Reverse Minkowski inequality.
Let $p \in]0;1[$, and f and g be two positive functions in $\mathcal{L}^p(X, \mathcal{M}, \mu)$. Prove that

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

Exercise 99. ★★ Limit of the p-norm (v1).
Let f be a positive integrable function with $\mu(\{f>0\}) < 1$. Using Hölder inequality, show that

$$\lim_{p\to 0} ||f||_p = 0.$$

Exercise 100. ★★★ Limit of the p-norm (v2).
Let $p_0 \in]0; +\infty[$, and $f \in \mathcal{L}^{p_0}(X, \mathcal{M}, \mu)$. Prove that

$$\lim_{p\to 0} ||f||_p^p = \lim_{p\to 0} \int_X |f|^p \mathrm{d}\mu = \mu\left(\{x \in X | f(x) \neq 0\}\right).$$

7.5 Conclusions

We conclude now our study of the functional spaces of integrable functions. We have discussed the definition of such spaces (Riemann integrable, Lebesgue integrable, and the quotient space) and we have seen some of their properties. A fundamental theorem here is the Riesz-Fischer theorem: the L^p spaces are complete (which is not the case of the R^p spaces). This result in itself shows the strength of Lebesgue integration theory and the stability of the Lebesgue integrable functions.

Of course, there are many other results on the Lebesgue spaces. For example, we have derived several inequalities with the p-norms, which are the most common, but there are many other inequalities more or less related to these ones. As mentioned, there are also other density results than the main one on simple functions. Another key result is the duality of the Lebesgue spaces and the relations between them

CHAPTER 8

Product of Convolution

Based on the Lebesgue integration theory and on the definition of the Lebesgue spaces L^p, we can define a very interesting operation on functions: the convolution. Formally, it constitutes a product on functional spaces, with an absorbant ($x \mapsto 0$) and a neutral element. Given its behaviour, this operation is widely used to solve differential equations. This chapter aims at introducing the convolution of two functions, as well as some of its basic properties.

8.1 Translations

DEFINITION 113. Translation.
Let $f : \mathbb{R}^n \to \mathbb{K}$, and $y \in \mathbb{R}^n$. We define the *translation application* τ_y as

$$\forall x \in \mathbb{R}^n, (\tau_y f)(x) = f(x - y).$$

The function $\tau_y f$ is called the *translated of f by y*.

PROPOSITION 71. Continuity of translations.
Let $y \in \mathbb{R}^n$ and $p \in [1; +\infty[$. The translation τ_y is continuous in $L^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ and, if $f \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, then

$$\lim_{y \to 0} ||\tau_y f - f||_p = 0.$$

REMARK 69.
This result is not true in L^∞! For example, if we consider the Heaviside function $H : \mathbb{R} \to \mathbb{R}$ defined by

$$H : x \mapsto \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}$$

then for all $y \neq 0$ we have $||\tau_y H - H||_\infty = 1$.

PROOF.
Let $y \in \mathbb{R}^n$. We first prove the continuity at 0 of the function $y \mapsto \tau_y g$ for any function g continuous with compact support. To do so, let g such a function and $M = \sup |g|$ (well defined, since g is continuous with compact support), and $R > 0$ such that the support of g is a subset of $E_R = \{x \in \mathbb{R}^n | \ ||x|| < R\}$.

Then, if $y \in \mathbb{R}^n$ is such that $||y|| < 1$, we have $|\tau_y g - g| \leq 2M\chi_{E_{R+1}} \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$. The dominated convergence theorem yields

$$\lim_{y\to 0} \int |\tau_y g - g|^p \mathrm{d}\mu = \lim_{y\to 0} \int |g(x-y) - g(x)|^p \mathrm{d}x = 0.$$

Hence,

$$\lim_{y\to 0} ||\tau_y g - g||_p = 0.$$

Now, let $f \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ and $\varepsilon > 0$. Since the continuous functions with compact support are dense in L^p, there exists a function g continuous with compact suport such that $||f-g||_p \leq \varepsilon/3$. The previous demonstration shows that there exists $\delta > 0$ such that if $||y|| \leq \delta$, then $||\tau_y g - g||_p \leq \varepsilon/3$. Therefore,

$$\begin{aligned} ||f - \tau_y f||_p &\leq ||f - g||_p + ||g - \tau_y g||_p + ||\tau_y g - \tau_y f||_p \\ &\leq 2||f - g||_p + ||g - \tau_y g||_p \\ &\leq \varepsilon, \end{aligned}$$

since $||\tau_y g - \tau_y f||_p = ||g - f||_p$ by invariance of the Lebesgue measure by translation. We deduce that, for all $y_0 \in \mathbb{R}^n$,

$$\lim_{y\to 0} ||\tau_{y_0+y} f - \tau_{y_0} f||_p = \lim_{y\to 0} ||\tau_y f - f||_p = 0,$$

so the translation is continuous at all $y_0 \in \mathbb{R}^n$. □

8.2 Product of Convolution

8.2.1 Definition and Properties

Definition 114. Product of convolution.
Let f and g be two measurable functions defined on $\mathbb{R}^n$, with values in $\mathbb{K}$. The *product of convolution*, noted $*$, is defined by

$$(f * g)(x) = \int f(x-y)g(y)\mathrm{d}y.$$

Proposition 72. Symmetry of the product of convolution.
Let f and g be two measurable functions defined on $\mathbb{R}^n$, with values in $\mathbb{K}$. The product of convolution of f and g is symmetric, i.e.

$$(f * g)(x) = (g * f)(x).$$

Proof.
This is simply a change of variables in the integral. □

Proposition 73. Associativity of the product of convolution.
Let f, g, and h be three measurable functions defined on $\mathbb{R}^n$, with values in $\mathbb{K}$. The product of convolution of f, g, and h is associative, i.e.

$$((f * g) * h)(x) = (f * (g * h))(x).$$

Exercise 101. ★ Associativity.
Prove that the product of convolution is associative, i.e. that $(f*g)*h = f*(g*h)$.

Exercise 102. ★ Convolution of characteristic functions.
Let $(a,b) \in (\mathbb{R}^+)^2$ with $a \geq b$. We define $f : x \mapsto \chi_{[-a;a]}$ and $g : x \mapsto \chi_{[-b;b]}$. Show that

$$f*g : x \mapsto \begin{cases} 2b & \text{if } |x| \leq a-b, \\ a+b-|x| & \text{if } a-b \leq |x| \leq a+b, \\ 0 & \text{if } a+b \leq |x|. \end{cases}$$

Definition 115. Product of convolution and function of 2 variables.
Let f and g be two measurable functions defined on $\mathbb{R}^n$, with values in $\mathbb{K}$. We define the function $h : (x,y) \mapsto f(x)g(y)$. The product of convolution of f and g can be re-written as

$$(f*g)(x) = \int h(x-y,y)\mathrm{d}y.$$

Exercise 103. ★ Convolution of rational functions.
Let $a \in \mathbb{R}^{+\star}$ and $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f : x \mapsto \frac{1}{a^2+x^2}.$$

Show that

$$\forall x \in \mathbb{R}, (f*f)(x) = \frac{2\pi}{a(4a^2+x^2)}.$$

Exercise 104. ★ Convolutions of complex functions.
Let f and g two functions from $\mathbb{R}$ to $\mathbb{C}$ defined by

$$f : x \mapsto \frac{1}{x+i} \qquad \text{and} \qquad g : x \mapsto \frac{1}{x-i}.$$

Show that

$$\forall x \in \mathbb{R}, (f*f)(x) = -\frac{2\pi i}{x+2i},$$

and

$$\forall x \in \mathbb{R}, (f*g)(x) = 0,$$

and

$$\forall x \in \mathbb{R}, (f^2*g^2)(x) = 0,$$

Exercise 105. ★ Convolution of exponentials.
Let $(a,b) \in (\mathbb{R}^+)^2$ with $a \neq b$. We define $f : x \mapsto e^{-|ax|}$ and $g : x \mapsto e^{-|bx|}$. Compute the convolution of f and g.

Exercise 106. ★ Convolution of Gaussian functions.
Let $(a,b) \in (\mathbb{R}^+)^2$. We define $f : x \mapsto e^{-ax^2}$ and $g : x \mapsto e^{-bx^2}$. Show that, if $a+b > 0$, then

$$\forall x \in \mathbb{R}, (f*g)(x) = \sqrt{\frac{\pi}{a+b}}e^{-\frac{ab}{a+b}x^2}.$$

8.2.2 Young's Theorem

Theorem 63. Specific case of Young's theorem.
If $f \in \mathcal{L}^1(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ and $g \in \mathcal{L}^1(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, then the product of convolution h of f and g defined by

$$h : x \mapsto (f * g)(x) = \int f(x-y)g(y)\mathrm{d}y,$$

is well defined for all $x \in \mathbb{R}^n$ and we have $h \in \mathcal{L}^1(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, with $||h||_1 \leq ||f||_1 ||g||_1$.

Proof.
We first assume that f and g are positive functions defined on $\mathbb{R}^n$. In that case, the function $h : (x, y) \mapsto f(x-y)g(y)$ is a positive measurable function defined on $(\mathbb{R}^n)^2$. Therefore, we can apply Fubini-Tonelli theorem and we obtain

$$\int \left(\int f(x-y)g(y)\mathrm{d}y \right) \mathrm{d}x = \int \left(\int f(x-y)g(y)\mathrm{d}x \right) \mathrm{d}y.$$

We can re-write the left-hand side as

$$\int \left(\int f(x-y)g(y)\mathrm{d}y \right) \mathrm{d}x = \int (f * g)(x)\mathrm{d}x = ||f * g||_1 = ||h||_1,$$

and the right-hand side as

$$\int \left(\int f(x-y)g(y)\mathrm{d}x \right) \mathrm{d}y = \left(\int f(x)\mathrm{d}x \right) \left(\int g(y)\mathrm{d}y \right) = ||f||_1 ||g||_1.$$

Hence, the theorem is true.

In the general case, we shall extend this proof to all integrable functions. The proof on positive functions shows that $|f| * |g|$ exists μ-almost everywhere, so the function $y \mapsto |f(x-y)g(y)|$ is integrable for almost every $x \in \mathbb{R}^n$. By definition of the Lebesgue integral, this means that $h : (x, y) \mapsto f(x-y)g(y)$ is integrable on $(\mathbb{R}^n)^2$. Hence, the product of convolution exists μ-almost everywhere and since $|f * g| \leq |f| * |g|$, we obtain that

$$||h||_1 \leq ||f||_1 ||g||_1.$$

□

Theorem 64. General Young's theorem.
Let $(p, q, r) \in [1; +\infty]^3$ such that

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}.$$

If $f \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ and $g \in \mathcal{L}^q(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, then the product of convolution h of f and g defined by

$$h : x \mapsto (f * g)(x) = \int f(x-y)g(y)\mathrm{d}y,$$

is well defined for all $x \in \mathbb{R}^n$ and we have $h \in \mathcal{L}^r(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, with $||h||_r \leq ||f||_p ||g||_q$. Besides

1. If $r = +\infty$, h is uniformly continuous; and
2. If $r = +\infty$, and $1 < p < q < +\infty$, then $\lim_{x \to +\infty} h(x) = 0$.

Proof.

If either f or g is equal to zero μ-almost everywhere then the result is trivial. We therefore assume that it is not the case, and that we can therefore re-normalize f and g such that $||f||_p = ||g||_q = 1$. As in he previous proof, we will show the result for positive functions, and then the general result will naturally follow.

Assuming f and g positive, we apply Hölder inequality and find that

$$\begin{aligned}(f*g)(x) &= \int \left(f^{p/r}(x-y)g^{q/r}(y)\right) f^{1-p/r}(x-y)g^{1-q/r}(y)\mathrm{d}y \\ &\leq \left(\int f^p(x-y)g^q(y)\mathrm{d}y\right)^{\frac{1}{r}} \left(\int f^{q'(1-p/r)}(x-y)\mathrm{d}y\right)^{\frac{1}{q'}} \left(\int g^{p'(1-q/r)}(y)\mathrm{d}y\right)^{\frac{1}{p'}},\end{aligned}$$

where p' and q' are the Hölder conjugates of p and q, so that they satisfy

$$\frac{1}{r} + \frac{1}{q'} + \frac{1}{p'} = \frac{1}{r} + \left(1 - \frac{1}{q}\right) + \left(1 - \frac{1}{p}\right) = 1,$$

and

$$\begin{aligned}q'\left(1 - \frac{p}{r}\right) &= q'p\left(\frac{1}{p} - \frac{1}{r}\right) = q'p\left(1 - \frac{1}{q}\right) = p, \\ p'\left(1 - \frac{q}{r}\right) &= p'q\left(\frac{1}{q} - \frac{1}{r}\right) = p'q\left(1 - \frac{1}{p}\right) = q.\end{aligned}$$

Therefore, we deduce that

$$\begin{aligned}(f*g)(x) &\leq \left(\int f^p(x-y)g^q(y)\mathrm{d}y\right)^{\frac{1}{r}} \left(f^p(x-y)\mathrm{d}y\right)^{\frac{1}{q'}} \left(g^q(y)\mathrm{d}y\right)^{\frac{1}{p'}} \\ &\leq \left(\int f^p(x-y)g^q(y)\mathrm{d}y\right)^{\frac{1}{r}} ||f||_p^{\frac{p}{q'}} ||g||_q^{\frac{q}{p'}} \\ &\leq \left(\int f^p(x-y)g^q(y)\mathrm{d}y\right)^{\frac{1}{r}},\end{aligned}$$

so

$$(f*g)^r(x) \leq \int f^p(x-y)g^q(y)\mathrm{d}y,$$

that we can re-write

$$(f*g)^r \leq f^p * g^q.$$

We can now use the previous theorem with the 1-norms and we obtain

$$\begin{aligned}\int (f*g)^r \mathrm{d}x &\leq ||f^p * g^q||_1 \\ &\leq ||f^p||_1 ||g^q||_1 \\ &\leq ||f||_p^p ||g||_q^q \\ &\leq 1,\end{aligned}$$

which proves the first part of the theorem. This result can be easily extended to the general case in which functions are not assumed to be positive.

Now, to prove the second part of the theorem, we assume that $r = +\infty$. In this case, p and q are Hölder conjugates. Since we have shown that translations are continuous, we know that for all $\varepsilon > 0$ there exists $\delta > 0$ such that, if $|y| \leq \delta$, then $||\tau_y g - g||_q \leq \varepsilon$. Therefore, if $|x - x'| \leq \delta$, we have

$$||\tau_x g - \tau_{x'} g||_q = ||\tau_{x-x'} g - g||_q \leq \varepsilon.$$

Using Hölder inequality, we obtain

$$\begin{aligned}|(f*g)(x)-(f*g)(x')| &\leq \int |f(y)||g(x-y)-g(x'-y)|\mathrm{d}y \\ &\leq \int |f(-y)||g(x+y)-g(x'+y)|\mathrm{d}y \\ &\leq ||f||_p||\tau_x g-\tau_{x'}g||_q \\ &\leq ||f||_p\varepsilon.\end{aligned}$$

This inequality ensures that $f*g$ is uniformly continuous. Now, we assume that both p and q are finite. Since the continuous functions with compact support are dense in the Lebesgue spaces, we pick two sequences of such functions, $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$, converging to f in L^p and to g in L^q, respectively. We can show (but not here) that for all $n\in\mathbb{N}$, f_n*g_n is also continuous with compact support. Then, we write that, for all $n\in\mathbb{N}$, we have

$$\begin{aligned}||f_n*g_n-f*g||_\infty &\leq ||f_n*(g_n-g)||_\infty+||(f_n-f)*g||_\infty \\ &\leq ||f_n||_p||g_n-g||_q+||f_n-f||_p||g||_q,\end{aligned}$$

from which we can easily see that

$$\lim_{n\to+\infty}||f_n*g_n-f*g||_\infty=0.$$

The sequence $(f_n*g_n)_{n\in\mathbb{N}}$ therefore converges uniformly to $f*g$. We deduce from this result that $f*g$ goes to zero as $|x|$ goes to $+\infty$. □

Exercise 107. ★★ Special case of Young's theorem.
In Young's theorem, if $q=1$ and $r=p$, we have

$$||f*g||_p\leq||f||_p||g||_1.$$

Prove this inequality through the use of Minkowski inequality.

8.3 Regularization

Definition 116. ε-rescaling.
Let $\phi:\mathbb{R}^n\to\mathbb{R}$ such that $\displaystyle\int\phi\mathrm{d}\mu=1$. Then, for all $\varepsilon>0$, we define the *ε-rescaling* of ϕ as

$$\phi_\varepsilon(x)=\frac{1}{\varepsilon^n}\phi\left(\frac{x}{\varepsilon}\right).$$

Exercise 108. ★ From ϕ to ϕ_ε.
Show that

$$\forall\varepsilon>0,\int\phi_\varepsilon\mathrm{d}\mu=1.$$

Proposition 74. Regularization by convolution.
Let $\phi:\mathbb{R}^n\to\mathbb{R}$. Let $p\in[1;+\infty[$ and $f\in\mathcal{L}^p(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n),\mu)$. We have

$$\lim_{\varepsilon\to 0}||f*\phi_\varepsilon-f||_p=0.$$

Remark 70.
This is called a "regularization by convolution" of the function ϕ.

Proof.
We already know that, for all $\varepsilon > 0$, the convolution $f * \phi_\varepsilon \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, and that $||f * \phi_\varepsilon||_p \leq ||f||_p$. In addition, μ-almost everywhere on $\mathbb{R}^n$, we have

$$\begin{aligned}(f * \phi_\varepsilon)(x) - f(x) &= \int f(x-y)\phi_\varepsilon(y)\mathrm{d}y - f(x)\int \phi_\varepsilon(y)\mathrm{d}y, \qquad \text{since } \int \phi_\varepsilon(y)\mathrm{d}y = 1 \\ &= \int (f(x-y) - f(x))\phi_\varepsilon(y)\mathrm{d}y \\ &= \frac{1}{\varepsilon^n}\int (f(x-y) - f(x))\phi_\varepsilon\left(\frac{y}{\varepsilon}\right)\mathrm{d}y \\ &= \int (f(x-\varepsilon z) - f(x))\phi(z)\mathrm{d}z.\end{aligned}$$

Now, we write

$$|f*\phi_\varepsilon - f|(x) \leq \left(\int |f(x-\varepsilon z) - f(x)|^p \phi(z)\mathrm{d}z\right)^{\frac{1}{p}}\left(\int \phi(z)\mathrm{d}z\right)^{1-\frac{1}{p}} = \left(\int |f(x-\varepsilon z) - f(x)|^p\phi(z)\mathrm{d}z\right)^{\frac{1}{p}},$$

meaning that

$$||f * \phi_\varepsilon - f||_p^p \leq \int \phi(z)||\tau_{\varepsilon z} f - f||_p^p \mathrm{d}z.$$

Since for all $\varepsilon > 0$, $||\tau_{\varepsilon z} f - f||_p^p$ converges pointwise to 0 and is dominated by an integrable function. Using the dominated convergence theorem, the right-hand side goes to 0 when ε goes to 0 and so the convergence of the left-hand side is proven. □

Lemma 11. Unit Gaussian integral.
There exists a $\mathcal{C}^\infty$ application $\phi : \mathbb{R}^n \to \mathbb{R}^+$ such that

$$\int \phi \mathrm{d}\mu = 1,$$

and $\phi(x) = 0$ if and only if $|x| \geq 1$.

Proof.
Let $\psi : \mathbb{R}^n \to \mathbb{R}^+$ the application defined by

$$\psi : x \mapsto \begin{cases} \exp\left(-\frac{1}{1-||x||^2}\right) & \text{if } ||x|| < 1, \\ 0 & \text{otherwise,}\end{cases}$$

then this function is integrable and is $\mathcal{C}^\infty$. We can now define ϕ as

$$\phi : x \mapsto \frac{\psi(x)}{\int \psi(y)\mathrm{d}y}.$$

□

Exercise 109. ★★ Continuity of ϕ.
Let ϕ defined as in the lemma. We assume that f is locally integrable, i.e. on all compacts. Show that $g = \phi * f$ is $\mathcal{C}^\infty$.

Corollary 20. Density of $\mathcal{C}_c^\infty$.
The space of $\mathcal{C}^\infty$ functions with compact support is dense in $\mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$ for $p \in [1; +\infty[$.

Proof.

Let $f \in \mathcal{L}^p(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu)$, and $\delta > 0$. Then, there exists a continuous function with compact support g such that $||g - f||_p \leq \delta/2$ (density result). Then, let ϕ as in the lemma, and let $\varepsilon > 0$ such that $||g - g * \phi_\varepsilon||_p \leq \delta/2$. Then, $h = g * \phi_\varepsilon$ is $\mathcal{C}^\infty$ with compact support, and

$$||f - h||_p \leq ||f - g||_p + ||g - h||_p \leq \delta,$$

and since this is true for all $\delta > 0$, it proves the corollary.

□

8.4 A Brief Extension: Distributions

There are different approaches to the notion of *distributions*. I will simply briefly discuss two of them:

- A "mathematical" one, motivated by the lack of some functions and by the fact that some functions are not well-defined. An example of that is, actually, the product of convolution: what could be the neutral element of this product?

- A "physical" one, or "applied maths" one, motivated by modeling issues: for a flow, or a field, how can we model sharp discontinuities or point sources?

The rigorous way to answer these issues is to define *generalised functions*, also called *distributions*.

Remark 71.

This section is a mere introduction to the notion of distributions. This is a whole analytical domain that would require much more than a section within a chapter, so I encourage you to look at it in other books if you are interested.

8.4.1 Definition

Definition 117. Test space.
The *test space*, noted $\mathcal{D}(\mathbb{R}^n)$, is the vector space of $\mathcal{C}^\infty$ functions defined on $\mathbb{R}^n$ and with values in $\mathbb{K}$, that have a bounded support.

Definition 118. Test function.
A test function is any function φ defined on the test space $\mathcal{D}(\mathbb{R}^n)$.

Definition 119. Distribution.
A *distribution* on $\mathbb{R}^n$ is any continuous linear functional defined on $\mathcal{D}(\mathbb{R}^n)$. If T is a distribution and φ is a test function, then we note $\langle T, \varphi \rangle \in \mathbb{K}$ the value of T at φ.

Remark 72.

A distribution is therefore an application whose argument is a function (hence a *functional*) with values in $\mathbb{K}$. The space of distributions defined on $\mathcal{D}(\mathbb{R}^n)$ forms a vector space which is the topological dual of $\mathcal{D}(\mathbb{R}^n)$.

8.4.2 Elementary Distributions

Definition 120. Dirac distribution.
We defined the *Dirac distribution* as the particular distribution δ which, to $\varphi \in \mathcal{D}(\mathbb{R}^n)$, associates its value at 0, i.e.
$$\langle \delta, \varphi \rangle = \varphi(0).$$

Definition 121. Shifted Dirac distribution.
Let $a \in \mathbb{R}^n$. We defined the *shifted Dirac distribution* as the particular distribution δ_a which, to $\varphi \in \mathcal{D}(\mathbb{R}^n)$, associates its value at a, i.e.
$$\langle \delta_a, \varphi \rangle = \varphi(a).$$

Remark 73.
This is usually called the *Dirac function* or the *delta function* while it is formally a distribution, not a function. One should prefer the name *Dirac distribution*.

Definition 122. Dirac notation.
Let $a \in \mathbb{R}^n$. The shifted Dirac distribution is generally identified to a function defined on $\mathbb{R}^n$ and is noted $\delta(x-a)$.

Remark 74.
There are many justifications to this, but we will not go into details here. An issue with the distributions is that, in most cases, they are not analytical on $\mathbb{R}^n$, which prevents from such a writing. In the case of the Dirac distribution, however, it can be expressed as the limit of a sequence of Gaussian functions, which justifies a formal writing as proposed in the definition.

Definition 123. Heaviside function.
The *Heaviside function* $H : \mathbb{R} \to \mathbb{R}$ is defined by
$$H : x \mapsto \begin{cases} 0 & \text{if } x < 0, \\ 1/2 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Definition 124. Heaviside distribution.
The *Heaviside distibution* H is the distribution associated to the Heaviside function, i.e. it is defined by
$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n), \langle H, \varphi \rangle = \int_0^{+\infty} \varphi(x)\mathrm{d}x.$$

8.4.3 Distributions and Convolutions

Proposition 75. Convolution and Heaviside distribution.
Let $a \in \mathbb{R}$. The Heaviside distribution shifted by a, applied on $\varphi \in \mathcal{D}(\mathbb{R})$, can be written as
$$(H * \varphi)(a) = \int H(x-a)\varphi(x)\mathrm{d}x.$$

Remark 75.
If $a = 0$, we recover the expression in the defintion of the non-shifted Heaviside distribution, i.e.

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n), \langle H, \varphi \rangle = \int_0^{+\infty} \varphi(x)\mathrm{d}x.$$

Proposition 76. Convolution with the Dirac distribution.
Let $a \in \mathbb{R}^n$. The Dirac distribution shifted by a, applied on $\varphi \in \mathcal{D}(\mathbb{R})$, can be written as a convolution

$$(\delta * \varphi)(a) = \int \delta(x - a)\varphi(x)\mathrm{d}x = \varphi(a).$$

Exercise 110. ★ "Product" of convolution.
Prove that the product of convolution formally defines a product on the test space $\mathcal{D}(\mathbb{R}^n)$, with an absorbant element $x \mapsto 0$ and a neutral element $x \mapsto \delta(x)$, i.e.

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n), \forall x \in \mathbb{R}^n, (0 * \varphi)(x) = 0,$$

and

$$\forall \varphi \in \mathcal{D}(\mathbb{R}^n), \forall x \in \mathbb{R}^n, (\delta * \varphi)(x) = \varphi(x).$$

8.5 Conclusions

In this chapter, we discussed a particular operation on integrable functions: the product of convolution. We have seen that, through Lebesgue theory and using the notion of distributions, we can formally define it as an associative and commutative product on the Lebesgue functional space. There are various reasons why the product of convolution is such an interesting mathematical tool. As mentioned here, it can be used to regularize functions, and to bound integrals using the p-norm (Young's theorems). Yet, one powerful use of the product of convolution really is in the notion of distributions: most of the basic operations one can think about on functions –such as translation, dilatation, picking a value at a given location, taking into account only half of a domain, etc– can be expressed in terms of distributions, and of a product of convolution. This has many implications while solving ordinary or partial differential equations, notably using Green functions.

CHAPTER 9

Differentiation and Primitives

We already know that the Riemann integral allows to define primitives F of a function f. Moreover, assuming f continuous on an interval $[a;\, b]$, the primitive F defined by

$$\begin{array}{rcl} F : [a;\, b] & \to & \mathbb{R} \\ x & \mapsto & \displaystyle\int_a^x f(y)\mathrm{d}y, \end{array}$$

is differentiable on $[a;\, b]$, and its derivative is, for all $x \in [a;\, b], F'(x) = f(x)$. This result, although powerful, is the best one can get with Riemann integration theory. It relies on the assumption that the function is continuous, and one can easily see why issues arise if f is not continuous: the equality between F' and f, assuming that F' is well defined, is not necessarily true at the points of discontinuities. With the Lebesgue integration theory, through classes of equivalence and the notion of μ-almost everywhere, we are able to obtain a more general result than with Riemann, in which the function f only has to be integrable.

9.1 Lebesgue's Differentiation Theorem

Theorem 65. Lebesgue's theorem.
Let $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mathcal{M}, \mu)$. Then, μ-almost everywhere in $\mathbb{R}^n$,

$$\lim_{r\to 0} \frac{1}{\mu\left(\mathcal{B}(x,r)\right)} \int_{\mathcal{B}(x,r)} |f(y) - f(x)|\mathrm{d}y = 0,$$

and, μ-almost everywhere in $\mathbb{R}^n$,

$$\lim_{r\to 0} \frac{1}{\mu\left(\mathcal{B}(x,r)\right)} \int_{\mathcal{B}(x,r)} f(y)\mathrm{d}y = f(x).$$

Definition 125. Lebesgue set of a function.
Let $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mathcal{M}, \mu)$ and $x \in \mathbb{R}^n$. Then x is a point in the *Lebesgue set of f* if there exists $A \in \mathbb{R}$ such that

$$\lim_{r\to 0} \frac{1}{\mu\left(\mathcal{B}(x,r)\right)} \int_{\mathcal{B}(x,r)} |f(y) - A|\mathrm{d}y = 0.$$

Definition 126. Regularly convergent.
A sequence of measurable sets $(E_n)_{n\in\mathbb{N}}$ is *regularly convergent*, or *converges regularly*, to x if there exist a positive constant c and a sequence of positive numbers $(r_n)_{n\in\mathbb{N}}$ such that

$$\forall n \in \mathbb{N}, E_n \subset \mathcal{B}(x, r_n),$$

with

$$\lim_{n\to+\infty} r_n = 0,$$

and

$$\forall n \in \mathbb{N}, \mu\left(\mathcal{B}(x, r_n)\right) \leq c\mu(E_n).$$

Theorem 66. Regular convergence.
Let $f \in L^1_{\text{loc}}(\mathbb{R}^n, \mathcal{M}, \mu)$ and $x \in \mathbb{R}^n$ in the Lebesgue set of f. Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of sets that converges regularly to x. Then we have

$$f(x) = \lim_{n\to+\infty} \frac{1}{\mu(E_n)} \int_{E_n} f(y)\mathrm{d}y.$$

Proof.
Let $n \in \mathbb{N}$. With the notations of the theorem, for all x in the domain of definition of f we write

$$\begin{aligned}
\left|\frac{1}{\mu\left(E_n\right)} \int_{E_n} f(y)\mathrm{d}y - f(x)\right| &= \left|\frac{1}{\mu\left(E_n\right)} \int_{E_n} f(y)\mathrm{d}y - f(x)\frac{1}{\mu\left(E_n\right)} \int_{E_n} \mathrm{d}y\right| \\
&= \left|\frac{1}{\mu\left(E_n\right)} \int_{E_n} (f(y) - f(x))\mathrm{d}y\right| \\
&\leq \frac{1}{\mu\left(E_n\right)} \int_{E_n} |f(y) - f(x)|\mathrm{d}y \\
&\leq \frac{c}{\mu\left(\mathcal{B}(x, r_n)\right)} \int_{\mathcal{B}(x,r_n)} |f(y) - f(x)|\mathrm{d}y,
\end{aligned}$$

and if we know take the limit n goes to $+\infty$ we find that

$$\lim_{n\to+\infty} \left|\frac{1}{\mu\left(E_n\right)} \int_{E_n} f(y)\mathrm{d}y - f(x)\right| = 0.$$

□

9.2 Primitives

Definition 127. Primitive.
Let $f \in L^1_{\text{loc}}(\mathbb{R}, \mathcal{M}, \mu)$ and $a \in \mathbb{R}$. We define F the *primitive of* f by

$$F : x \mapsto \int_a^x f(y)\mathrm{d}y.$$

Remark 76.
Contrary to the definition of the primitive of a function in the Riemann sense, in which the function has to be continuous, with Lebesgue the function only has to be locally integrable. Note that we are, again, working on class of equivalences, so the functions may differ on a negligeable set.

THEOREM 67. Differentiability.
Let $f \in L^1_{\text{loc}}(\mathbb{R}, \mathcal{M}, \mu)$, and F a primitive of f. Then F is differentiable μ-almost everywhere and

$$F' = f.$$

REMARK 77.
More precisely: the class of equivalence of F' and the class of equivalence of f are the same.

PROOF.
Since $f \in L^1_{\text{loc}}(\mathbb{R}, \mathcal{M}, \mu)$, we have that almost every $x \in \mathbb{R}$ is in the Lebesgue set of f, satifying the conclusion of the previous theorem on regular convergence. Hence, we only have to show that for all x in the Lebesgue set of f, we have $F'(x) = f(x)$.

Let $(r_n)_{n\in\mathbb{N}}$ be a sequence of positive numbers converging to 0. For all $n \in \mathbb{N}$, we define $E_n = [x; x + r_n]$. The sequence $(E_n)_{n\in\mathbb{N}}$ is a sequence of sets that converges regularly to x, so that we have

$$\lim_{n\to+\infty} \frac{1}{\mu(E_n)} \int_{E_n} f(y)\mathrm{d}y = f(x),$$

or, better said,

$$\lim_{n\to+\infty} \frac{1}{r_n} \int_x^{x+r_n} f(y)\mathrm{d}y = f(x).$$

This can be rewritten, using the primitive of f, as

$$\lim_{n\to+\infty} \frac{F(x + r_n) - F(x)}{r_n} = f(x).$$

The sequence $(r_n)_{n\in\mathbb{N}}$ being arbitrary, we conclude that

$$\lim_{h\to 0^+} \frac{F(x + h) - F(x)}{h} = f(x).$$

If we now consider the sequence $(E'_n)_{n\in\mathbb{N}}$ where $\forall n \in \mathbb{N}, E'_n = [x - r_n; x]$, then a similar reasoning yields

$$\lim_{h\to 0^-} \frac{F(x + h) - F(x)}{h} = f(x).$$

Finally, we conclude that

$$\lim_{h\to 0} \frac{F(x + h) - F(x)}{h} = f(x),$$

thus μ-almost everywhere on $\mathbb{R}$, $F'(x) = f(x)$ (i.e. they have the same class of equivalence). $\square$

9.3 Conclusions

This (brief) chapter introduces one last fundamental result of Lebesgue integration theory: the primitives of a function and their derivability. If we recall Riemann's result, we have that, if f is continuous on an interval $[a; b]$, then we can define its primitive F on $[a; b]$ and F is differentiable with

$$\forall x \in [a; b], F'(x) = f(x).$$

Now, Lebesgue's result goes one step further by stating that, if f is integrable on a subset of $\mathbb{R}$, then we can define its primitive F and F is differentiable. In particular, in the case in which the set is an interval $[a; b]$, then

$$\mu \text{ almost everywhere on } [a; b], F'(x) = f(x).$$

Formally, the functions F' and f are no longer equal (as it was the case in Riemann's theory), but their classes of equivalence, in Lebesgue spaces, are equal, i.e. the functions are equal μ-almost everywhere.

CHAPTER 10

Fourier Series and Fourier Transform

The Fourier analysis, i.e. Fourier series and transform, constitutes a very powerful tool to study harmonic signals. Historically, they have been introduced by Joseph Fourier in his 1822 publication *Théorie analytique de la chaleur*: unable to solve the heat equation he just derived

$$\frac{\partial T}{\partial t} - D\nabla^2 T = \rho(\mathbf{x}, t),$$

he decided to decompose the temperature field T over trigonometric functions, whose convergence can be easily proven. Such a decomposition for periodic signals has been later called *Fourier series* (discrete) and extended to non-periodic functions as *Fourier transform* (continuous).

In this chapter, we will present these mathematical tools and some of their properties. For simplicity, we will only consider the case of functions defined on $\mathbb{R}$, but all results can be readily extended to $\mathbb{R}^n$.

10.1 Fourier Series

10.1.1 Periodicity

Definition 128. Periodic function.
Let f be a function defined on $\mathbb{R}$, and $T \in \mathbb{R}$. The function f is said to be *periodic of period T*, or *T-periodic*, if

$$\forall x \in \mathbb{R}, f(x+T) = f(x).$$

Example 26. For $n \in \mathbb{N}$, we can define the T-periodic functions $e_n : \mathbb{R} \to \mathbb{C}$

$$e_n : t \mapsto \exp\left(\frac{2i\pi nt}{T}\right).$$

Remark 78.
T is called the *period* of a function, and is often recasted to be equal to 2π or to 1.

Definition 129. Periodic domain.
Let $T \in \mathbb{R}$ and f a T-periodic function. We note $\mathbb{T} = [0;T]$ the periodic domain of definition of f. Moreover, we note $\mathcal{D}_T$ the T-periodic functions continuous by steps on $\mathbb{T}$, and $\mathcal{C}_T$ the T-periodic functions continuous on $\mathbb{T}$.

Lemma 12. Integration on $\mathbb{T}$.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. We have

$$\forall a \in \mathbb{R}, \int_a^{T+a} f(t)\mathrm{d}t = \int_0^T f(t)\mathrm{d}t = \int_{-T/2}^{T/2} f(t)\mathrm{d}t.$$

Exercise 111. ★ Integration on $\mathbb{T}$.
Prove that, if $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$, we have

$$\forall a \in \mathbb{R}, \int_a^{T+a} f(t)\mathrm{d}t = \int_0^T f(t)\mathrm{d}t = \int_{-T/2}^{T/2} f(t)\mathrm{d}t.$$

Theorem 68. Scalar product on $\mathcal{C}_T$.
Let $T \in \mathbb{R}$. The application

$$\begin{array}{rcl} \mathcal{C}_T \times \mathcal{C}_T & \to & \mathbb{C} \\ (f,g) & \mapsto & \langle f,g\rangle = \dfrac{1}{T}\displaystyle\int_0^T \overline{f(t)}g(t)\mathrm{d}t, \end{array}$$

defines a scalar product on $\mathcal{C}_T$.

Proof.
We already proved it when discussing the Hilbert space L^2. □

10.1.2 Fourier Coefficients and Fourier Sums

Theorem 69. Orthonormality of the $(e_n)_{n\in\mathbb{N}}$ functions.
Let $T \in \mathbb{R}$. The family of T-periodic functions $(e_n)_{n\in\mathbb{N}}$, defined from $\mathbb{R}$ to $\mathbb{C}$ by

$$e_n : t \mapsto \exp\left(\frac{2i\pi nt}{T}\right),$$

forms an orthonormal basis of $\mathcal{C}_T$, i.e.

$$\forall (n,m) \in \mathbb{N}^2, \langle e_n, e_m\rangle = \delta_{nm}.$$

Proof.
We note that

$$\forall (n,m) \in \mathbb{N}^2, \langle e_n, e_m\rangle = \frac{1}{T}\int_0^T \exp\left(-\frac{2i\pi(n-m)t}{T}\right)\mathrm{d}t.$$

Therefore, if $n = m$ we have

$$\langle e_n, e_m\rangle = \frac{1}{T}\int_0^T 1\mathrm{d}t = 1,$$

and if $n \neq m$ we have

$$\langle e_n, e_m \rangle = \left[\frac{1}{T} \frac{T}{-2i\pi(n-m)T} \exp\left(-\frac{2i\pi(n-m)t}{T} \right) \right]_0^T = 0,$$

because of the T-periodicity.

□

Definition 130. Fourier coefficients.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. Let $n \in \mathbb{Z}$. We call *Fourier coefficient of order n of f* the quantity

$$c_n(f) = \langle e_n, f \rangle = \frac{1}{T} \int_0^T f(t) e^{-2i\pi nt/T} \mathrm{d}t.$$

Definition 131. Fourier sum.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. Let $n \in \mathbb{N}$. We call *Fourier sum of order n* the quantity

$$S_n(f) = \sum_{k=-n}^{n} c_k(f) e_k.$$

Remark 79.
The Fourier sum will help us define a series but it is relevant to make this distinction: in some cases, the Fourier sum does not converge when $n \to +\infty$.

Proposition 77. Fourier sum.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. Then,

$$\forall t \in \mathbb{R}, \forall n \in \mathbb{N}^\star, (S_n(f))(t) = \frac{a_0(f)}{2} + \sum_{k=1}^{n} \left[a_k(f) \cos\left(\frac{2\pi kt}{T}\right) + b_k(f) \sin\left(\frac{2\pi kt}{T}\right) \right],$$

with

$$\forall k \in \mathbb{N}, a_k(f) = \frac{2}{T} \int_0^T f(t) \cos\left(\frac{2\pi kt}{T}\right) \mathrm{d}t \qquad \text{and} \qquad b_k(f) = \frac{2}{T} \int_0^T f(t) \sin\left(\frac{2\pi kt}{T}\right) \mathrm{d}t.$$

Proposition 78. Fourier coefficients (case $T = 2\pi$).
Let $T = 2\pi$ and $f \in \mathcal{D}_T$. Then,

$$\forall k \in \mathbb{Z}, c_k(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} \mathrm{d}t,$$

and

$$\forall k \in \mathbb{N}, a_k(f) = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(kt)\, \mathrm{d}t \qquad \text{and} \qquad b_k(f) = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(kt)\, \mathrm{d}t.$$

PROPOSITION 79. Parity.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. Then,

- If $\forall x \in \mathbb{R}, f(x) = f(-x)$, then

$$\forall k \in \mathbb{N}, b_k(f) = 0 \qquad \text{and} \qquad a_k(f) = \frac{4}{T}\int_0^{T/2} f(t)\cos\left(\frac{2\pi kt}{T}\right)\mathrm{d}t;$$

- If $\forall x \in \mathbb{R}, f(x) = -f(-x)$, then

$$\forall k \in \mathbb{N}, a_k(f) = 0 \qquad \text{and} \qquad b_k(f) = \frac{4}{T}\int_0^{T/2} f(t)\sin\left(\frac{2\pi kt}{T}\right)\mathrm{d}t.$$

10.1.3 Fourier Series

DEFINITION 132. Fourier series.
Let $T \in \mathbb{R}$ and $f \in \mathcal{D}_T$. The sequence $(S_n(f))_{n\in\mathbb{N}}$ of the Fourier sums is called the *Fourier series of f*.

THEOREM 70. Convergence and Fourier series.
Let $T \in \mathbb{R}$ and $f \in \mathcal{C}_T$, $\mathcal{C}^1$ by steps. Then, the sequence of the Fourier sums $(S_n(f))_{n\in\mathbb{N}}$ of f converges uniformly to f in $(\mathcal{C}_T, ||\cdot||_2)$. By extension, if it exists, the limit of this sequence is called the *Fourier series of f* and we have

$$\forall t \in \mathbb{R}, f(t) = \sum_{n=-\infty}^{+\infty} c_n(f)e_n.$$

EXAMPLE 27.

1. One of the most simple examples we can give is the case where $\forall n \in \mathbb{N}, a_n = 0$, $\forall n \in \mathbb{N}\setminus\{1\}, b_n = 0$, and $b_1 \neq 0$. Then, the function defined by these Fourier coefficients is simply $f : t \mapsto b_1 \sin(2\pi t/T)$.

2. A signal with a slow varying enveloppe can be written as $f : t \mapsto \cos(\omega t)\cos(\Omega t)$, with ω and Ω two pulsations. Then, using trigonometry, it is easy to translate it into a sum of cosines: $f : t \mapsto [\cos((\omega+\Omega)t) + \cos(\omega-\Omega)t)]/2$, whose Fourier series is straightforward to write.

3. The square signal, defined on $[-\pi; \pi]$ by

$$f : t \mapsto \begin{cases} 0 & \text{if } t \in [-\pi; 0] \\ 1 & \text{if } t \in [0; \pi], \end{cases}$$

can be computed easily. First, the function is anti-symmetric so $\forall n \in \mathbb{N}^\star, a_n = 0$. Then,

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\mathrm{d}t = \frac{1}{\pi}\int_0^{\pi} 1\mathrm{d}t = 1,$$

and similarly

$$\forall n \in \mathbb{N}^\star, b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)\mathrm{d}t = \frac{1}{\pi}\int_0^{\pi}\sin(nt)\mathrm{d}t = \frac{1-\cos(n\pi)}{n\pi} = \frac{1-(-1)^n}{\pi n}.$$

Therefore we write the Fourier series

$$f : t \mapsto \frac{1}{2} + \sum_{n=1}^{+\infty} \frac{1-(-1)^n}{\pi n} \sin(nt).$$

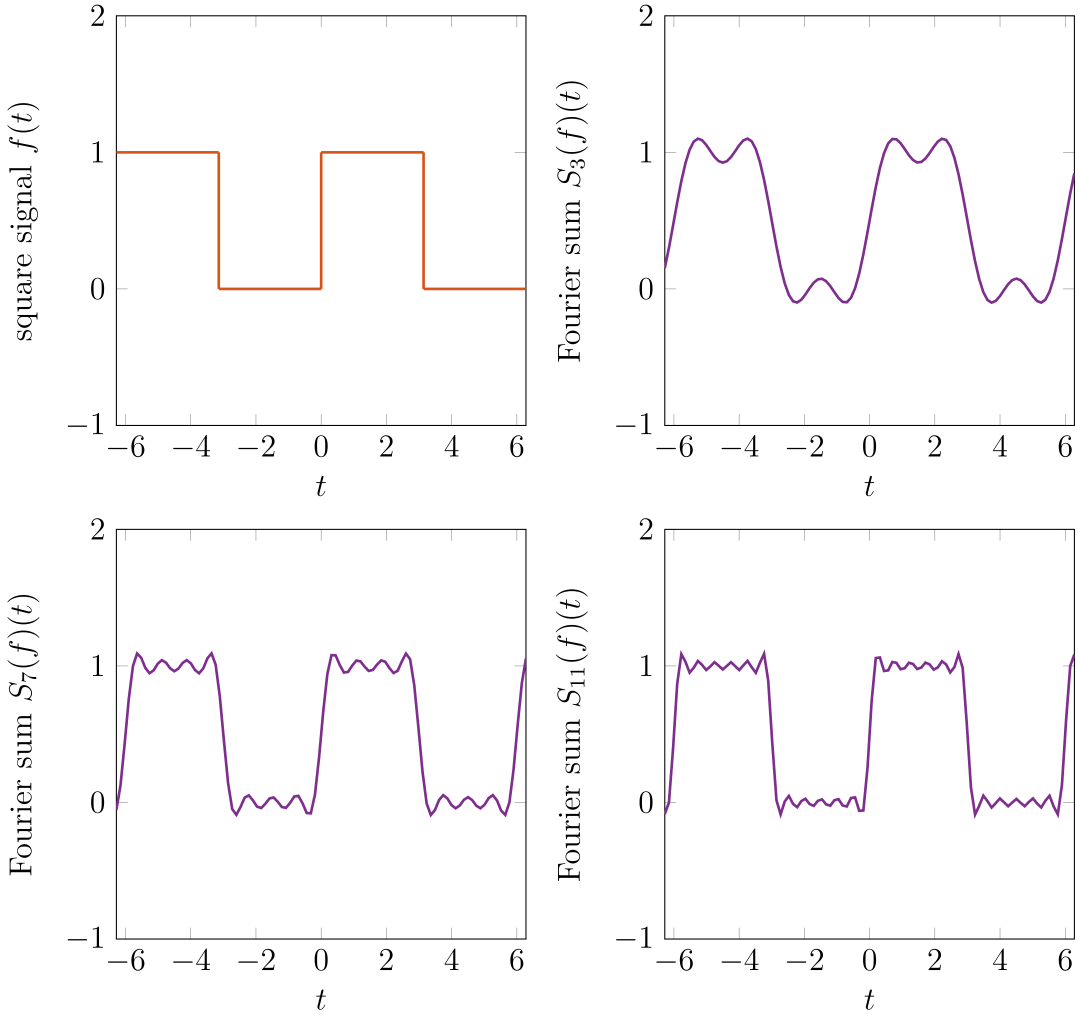


Figure 10.1: Square signal f and Fourier sums, converging to f.

Exercise 112. ★ Triangle function.
Compute the Fourier series of the triangle function f, defined on $[-\pi;\, \pi]$ by

$$f : t \mapsto \begin{cases} \pi/2 + t & \text{if } t \in [-\pi;\, 0] \\ \pi/2 - t & \text{if } t \in [0;\, \pi]. \end{cases}$$

Exercise 113. ★ Square functions.
Compute the Fourier series of the function f, defined on $[-\pi;\, \pi]$ by $f : t \mapsto t^2$. Then, compute the Fourier series of the function g, defined on $[0;\, 2\pi]$ by $g : t \mapsto t^2$.

COROLLARY 21. Parseval equality.
Let $T \in \mathbb{R}$ and $f \in \mathcal{C}_T$, $\mathcal{C}^1$ by steps. Then, the series

$$\sum_{n\geq 0} |c_n(f)|^2 \qquad \text{and} \qquad \sum_{n\geq 0} |c_{-n}(f)|^2,$$

are convergent, and

$$||f||_2^2 = \sum_{n=-\infty}^{+\infty} |c_n(f)|^2.$$

DEFINITION 133. Regularised function.
Let f a function defined on $\mathbb{R}$, continuous by steps. The *regularised function* associated to f, noted $\tilde{f}$, is defined by

$$\tilde{f} : x \mapsto \frac{1}{2}\left(\lim_{y\to x^+} f(x) + \lim_{y\to x^-} f(x)\right).$$

THEOREM 71. Dirichlet theorem.
Let $T \in \mathbb{R}$ and $f \in \mathcal{C}_T$, $\mathcal{C}^1$ by steps. Then, the sequence of the Fourier sums $(S_n(f))_{n\in\mathbb{N}}$ of f converges pointwise to $\tilde{f}$.

10.2 Fourier Transform

10.2.1 Definition

DEFINITION 134. Fourier transform.
Let f be a function defined on $\mathbb{R}$. The *Fourier transform of* f, noted $\mathcal{F}[f]$ or $\hat{f}$, is the function defined on $\mathbb{R}$ with values in $\mathbb{C}$, by

$$\mathcal{F}[f] = \hat{f} : k \mapsto \int f(t)e^{-2i\pi kt}\mathrm{d}t.$$

DEFINITION 135. Conjugate Fourier transform.
Let f be a function defined on $\mathbb{R}$. The *conjugate Fourier transform of* f, noted $\overline{\mathcal{F}}[f]$, is the function defined on $\mathbb{R}$ with values in $\mathbb{C}$, by

$$\overline{\mathcal{F}}[f] : k \mapsto \int f(t)e^{2i\pi kt}\mathrm{d}t.$$

REMARK 80.
Unfortunately, there are different conventions to define the Fourier transform. This one has the advantage to be symmetric and does not have normalisation prefactor, but you can also find

$$\mathcal{F}[f] : k \mapsto \frac{1}{\sqrt{2\pi}}\int f(t)e^{-ikt}\mathrm{d}t \qquad \text{and} \qquad \mathcal{F}[f] : t \mapsto \frac{1}{\sqrt{2\pi}}\int \hat{f}(k)e^{ikt}\mathrm{d}k,$$

or

$$\mathcal{F}[f] : k \mapsto \int f(t)e^{-ikt}\mathrm{d}t \qquad \text{and} \qquad \mathcal{F}[f] : t \mapsto \frac{1}{2\pi}\int \hat{f}(k)e^{ikt}\mathrm{d}k.$$

THEOREM 72. Fourier transform and integrable functions.
Let $f \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. Then, the Fourier transform of f, $k \mapsto \hat{f}(k)$, is defined for all $k \in \mathbb{R}$.

EXAMPLE 28.

1. The Fourier transform of the Gaussian function

$$f : x \mapsto e^{-\pi x^2}$$

   is also a Gaussian function: $\forall k \in \mathbb{R}, \mathcal{F}[f](k) = e^{-\pi k^2}$.

2. Let $a \in \mathbb{R}^{+\star}$. We define the Lorentzian function f by

$$f : x \mapsto \frac{2a}{a^2 + 4\pi^2 x^2}.$$

   The Fourier transform of f writes: $\forall k \in \mathbb{R}, \mathcal{F}[f](k) = e^{-a|k|}$.

PROPOSITION 80. Fourier transform in $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$.
Two integrable functions that are equal μ-almost everywhere (i.e. they have the same class of equivalence in $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$) have the same Fourier transform.

PROOF.
This is straightforward: two functions from the same class of equivalence are equal, except on a negligeable set, but in the Lebesgue integration theory the integral on a negligeable set is equal to zero.
□

PROPOSITION 81. Properties of the Fourier transform application.
The application $\mathcal{F}$ defined on $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ that associates to a function its Fourier transform, is linear and continuous.

LEMMA 13. Riemann-Lebesgue lemma.
Let $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. Its Fourier transform $\hat{f}$ is continuous and satisfies

$$\lim_{k \to \pm\infty} \hat{f}(k) = 0.$$

EXERCISE 114. ★★★ Riemann-Lebesgue lemma.
Prove Riemann-Lebesgue lemma.

10.2.2 Inversion

THEOREM 73. Inverse Fourier transform.
Let $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. We assume that its Fourier transform $\hat{f}$ is integrable. Then, we have that

$$\mu \text{ almost everywhere, } \overline{\mathcal{F}}[\hat{f}](x) = f(x).$$

This is true, in particular, for all x at which f is continuous.

COROLLARY 22. Inverse Fourier transform of a continuous function.
Let $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ be a continuous function. We assume that its Fourier transform $\hat{f}$ is integrable. Then, we have that

$$\forall x \in \mathbb{R}, \overline{\mathcal{F}}[\hat{f}](x) = f(x).$$

REMARK 81. Since $\mathcal{F}$ and $\overline{\mathcal{F}}$ are related, it is sometimes easier to compute the inverse Fourier transform of known functions instead of computing the Fourier transform of complicated functions.

10.2.3 Properties of the Fourier Transform

PROPOSITION 82. Symmetry and translation.
Let $a \in \mathbb{R}$, $k_0 \in \mathbb{R}$, and $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. We define

$$\overline{f} : x \mapsto \overline{f(x)}, \quad u : x \mapsto f(-x), \quad v : x \mapsto f(x-a), \quad \text{and} \quad w : x \mapsto f(x)e^{2i\pi k_0 x}.$$

Then we have

$$\hat{\overline{f}} : k \mapsto \overline{\hat{f}(-k)}, \quad \hat{u} : k \mapsto \hat{f}(-k), \quad \hat{v} : k \mapsto e^{-2i\pi ka}\hat{f}(k), \quad \text{and} \quad \hat{w} : k \mapsto \hat{f}(k-k_0).$$

PROPOSITION 83. Dilatation.
Let $a \in \mathbb{R}^\star$ and $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. We define g the dilated function function obtained by a change of scale as $g : x \mapsto f(ax)$. Then we have

$$\hat{g} : k \mapsto \frac{1}{|a|}\hat{f}\left(\frac{k}{a}\right).$$

PROPOSITION 84. Derivatives.
Let $p \in \mathbb{N}$ and $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ a function such that $g_p : x \mapsto (-2i\pi x)^p f(x)$ is still integrable. Then, $\hat{f}$ can be differentiated up to p times and

$$\forall q \in [\![0;\, p]\!], \hat{g_q} : k \mapsto \hat{f}^{(q)}(k).$$

Moreover, if $f \in \mathcal{C}^p$ such that all its derivatives $h_p : x \mapsto f^{(p)}(x)$ are integrable, then

$$\forall q \in [\![0;\, p]\!], \hat{h_q} : k \mapsto (2i\pi k)^q \hat{f}(k).$$

EXERCISE 115. ★ Symmetry, translation, dilatation and derivatives.
Prove the corresponding propositions.

REMARK 82.
In particular, assuming that the correct assumptions are satisfied, we deduce that

$$\begin{aligned} \mathcal{F}[x \mapsto f'(x)] : k &\mapsto 2i\pi k\hat{f}(k), \\ \mathcal{F}[x \mapsto -2i\pi x f(x)] : k &\mapsto \frac{\mathrm{d}}{\mathrm{d}k}\hat{f}(k). \end{aligned}$$

Corollary 23. $\mathcal{C}^\infty$ transform.
If $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, with a bounded support, then its Fourier transform $\hat{f}$ is $\mathcal{C}^\infty$.

Corollary 24. Iterated derivatives.
Let $p \in \mathbb{N}$. We assume that $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ is a $\mathcal{C}^p$ function, and that all of these derivatives are integrable. Then, for any $k \in \mathbb{R}$, we have

$$\left|\hat{f}(k)\right| \leq \frac{1}{|2\pi k|^p} \int \left|f^{(p)}(t)\right| \mathrm{d}t,$$

and

$$\left|\hat{f}^{(p)}(k)\right| \leq \int |2\pi t|^p \, |f(t)| \, \mathrm{d}t.$$

Remark 83.
These inequalities give bounds on the decay of the Fourier transform $\hat{f}$ at $\pm\infty$. For example, if $f^{(p)}$ is integrable, then $\hat{f}$ decays at least as fast as $1/k^p$ at $\pm\infty$. Conversely, the same procedure applies with the inverse Fourier transform. Therefore, the regularity of f (respectively $\hat{f}$) is linked to the decay of $\hat{f}$ (respectively f) at $\pm\infty$.

Example 29.
Solving a linear ODE (or PDE) can be very straightforward using these properties. It is especially true for periodic signals with the Fourier series. Let consider the following ordinary differential equation

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) + l^2 f(x) = g(x),$$

then its Fourier transform is

$$(k^2 + l^2)\hat{f}(k) = \hat{g}(k),$$

and we deduce that

$$f = \mathcal{F}^{-1}\left[k \mapsto \frac{\hat{g}(k)}{k^2 + l^2}\right].$$

In particular, to solve the homogeneous ODE, we see that the solution is either $f = 0$, or an oscillatory signal with wave length $k = l$.

Exercise 116. ★ A forced damped harmonic oscillator.
Using the Fourier transform, solve the following ODE

$$f''(t) + 2\gamma f'(t) + \omega_0^2 f(t) = a\cos(\omega t),$$

where $\gamma, \omega_0, a, \omega$ are positive real constants.

10.2.4 Usual Fourier Transforms

Proposition 85. Fourier transform of a Dirac.
We have

$$\mathcal{F}[x \mapsto 1] : k \mapsto \delta(k) \qquad \text{and} \qquad \mathcal{F}[x \mapsto \delta(x)] : k \mapsto 1.$$

Remark 84.

Many other results can be deduced from this one, notably

$$\begin{aligned}
\mathcal{F}[x \mapsto \delta(x - x_0)] : k &\mapsto e^{-2i\pi k x_0} \\
\mathcal{F}[x \mapsto \sin(2\pi k_0 x)] : k &\mapsto \frac{1}{2i}\left[\delta(k - k_0) - \delta(k + k_0)\right] \\
\mathcal{F}[x \mapsto \cos(2\pi k_0 x)] : k &\mapsto \frac{1}{2}\left[\delta(k - k_0) + \delta(k + k_0)\right]
\end{aligned}$$

EXERCISE 117. ★ Fourier transform of trigonometric functions.
Prove the previous results.

PROPOSITION 86. Fourier transform of a finite window.
Let $\Pi : x \mapsto \chi_{[-1/2;1/2]}(x)$. We have

$$\hat{\Pi} : k \mapsto \frac{\sin(\pi k)}{\pi k} = \mathrm{sinc}(\pi k).$$

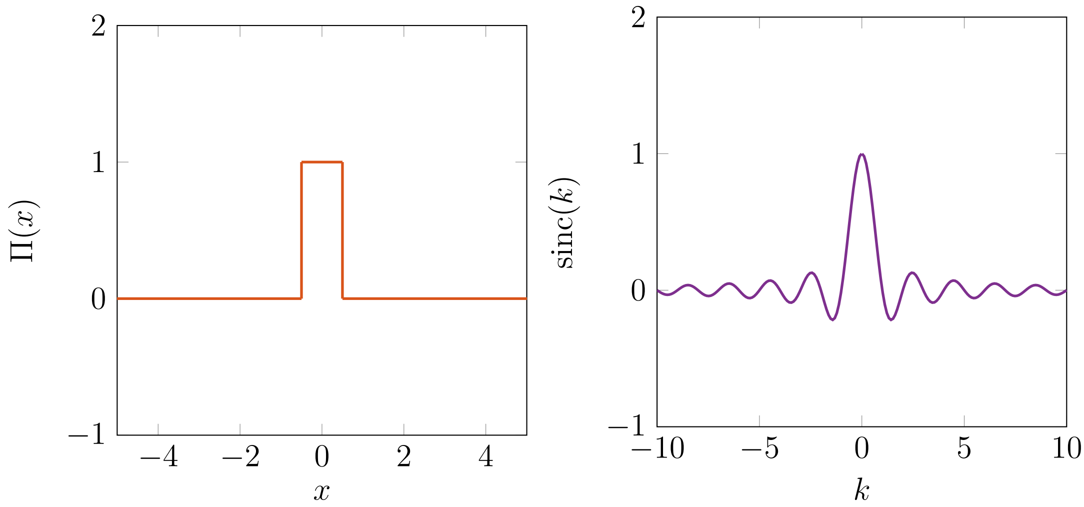


Figure 10.2: Window function (left) and its Fourier transform, a sine cardinal function (right).

PROOF.
We compute

$$\begin{aligned}
\hat{\Pi}(k) &= \int_{-\infty}^{+\infty} \Pi(x) e^{-2i\pi k x} \mathrm{d}x \\
&= \int_{-1/2}^{+1/2} e^{-2i\pi k x} \mathrm{d}x \\
&= \left[-\frac{1}{2i\pi k} e^{-2i\pi k x}\right]_{-1/2}^{+1/2} \\
&= \frac{\sin(\pi k)}{\pi k} = \mathrm{sinc}(\pi k).
\end{aligned}$$

□

EXERCISE 118. ★ Fourier transform of a Gaussian.
Let $\sigma \in \mathbb{R}^{+\star}$ and f a Gaussian function defined by $f : x \mapsto e^{-\sigma x^2}$. Show that

$$\hat{f} : k \mapsto \sqrt{\frac{\pi}{\sigma}} e^{-\pi^2 k^2/\sigma}.$$

EXERCISE 119. ★ Fourier transform of a characteristic function.
Let $(a, b) \in \mathbb{R}^2$ with $b > a$. We define f by

$$f : x \mapsto \frac{1}{b-a}\chi_{[a;b]}(x).$$

Show that the Fourier transform of f is

$$\hat{f} : k \mapsto \frac{\sin(\pi(b-a)k)}{\pi k} e^{-i\pi(a+b)k}.$$

10.3 Schwartz Space

10.3.1 Definition and Relation to the Fourier Transform

DEFINITION 136. Schwartz space.
The *Schwartz space* $\mathcal{S}$ is the space of $\mathcal{C}^\infty$ functions that are rapidly decaying in $\pm\infty$, as well as all of their derivatives.

PROPOSITION 87. Fourier transform on Schwartz space.
The Fourier transform is a continuous linear operator $\mathcal{F} : \mathcal{S} \to \mathcal{S}$.

REMARK 85.
This is a fundamental result of stability *via* Fourier transform: if $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$ and if a sequence of functions $(f_n)_{n\in\mathbb{N}}$ goes to 0 then $\left(\hat{f}_n\right)_{n\in\mathbb{N}}$ also goes to zero.

PROPOSITION 88. Inverse Fourier transform on Schwartz space.
The Fourier transform is an isomorphism of $\mathcal{S}$, i.e. it is bijective, and we have that $\mathcal{F}^{-1} = \overline{\mathcal{F}}$ everywhere on $\mathcal{S}$.

DEFINITION 137. Schwartz convergence.
A sequence of functions $(f_n)_{n\in\mathbb{N}}$ in the Schwartz space converges in $\mathcal{S}$ to 0 if, for any $(p, q) \in \mathbb{N}^2$, we have

$$\lim_{n\to+\infty} \sup_{x\in\mathbb{R}} \left| x^p f_n^{(q)}(x) \right| = 0.$$

10.3.2 Fourier Transform in L^2

PROPOSITION 89. Density of the Schwartz space.
The Schwartz space $\mathcal{S}$ is a dense subspace of $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$.

REMARK 86.
This shows, again, the powerful structure of the L^2 space: it is a normed vector space, it is complete, it is a Hilbert space with an integral scalar product, the Schwartz space is dense in it, it can be identified to its topological dual...

Theorem 74. Parseval-Plancherel.
The Fourier transform is an isometry of $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$. For two functions f and g in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, the Fourier transforms $\hat{f}$ and $\hat{g}$ are also in $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ and we have

$$\int \overline{f}(x) g(x) \mathrm{d}x = \int \overline{\hat{f}(k)} \hat{g}(k) \mathrm{d}k.$$

In particular, if $f = g$, we have

$$\int |f(x)|^2 \,\mathrm{d}x = \int \left|\hat{f}(k)\right|^2 \mathrm{d}k, \qquad \text{i.e.} \qquad ||f||_2 = ||\hat{f}||_2.$$

Proposition 90. Inverse Fourier transform in L^2.
For any $f \in L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, we have μ-almost everywhere

$$\overline{\mathcal{F}}\left[\mathcal{F}[f]\right] = \mathcal{F}\left[\overline{\mathcal{F}}[f]\right] = f,$$

i.e. μ-almost everywhere $\mathcal{F}^{-1} = \overline{\mathcal{F}}$.

Remark 87.
At this point, one may ask the following question: why did we need to do all this? These results simply define a "nice" space for the Fourier transform, i.e. a space of functions with the appropriate properties to apply the Fourier transform and its inverse transform. In particular, physically, functions are almost always in L^2 and in $\mathcal{S}$: physical signals are linked to energetic quantities, that are finite quantities computed by integrating the squared quantity (squared velocities, intensities, etc).

10.4 Fourier Transform and Convolutions

Theorem 75. Fourier Transform and Convolutions (Faltung theoreom).
Let f and g two functions such that their Fourier transform $\mathcal{F}[f]$ and $\mathcal{F}[g]$ exist, and that the product of convolution $f * g$ is defined and integrable. Then we have

$$\mathcal{F}[f * g] = \mathcal{F}[f] \cdot \mathcal{F}[g],$$

and

$$\mathcal{F}[f \cdot g] = \mathcal{F}[f] * \mathcal{F}[g].$$

Remark 88.
This result shows the strong link between the product of convolution, the translations, and the Fourier transform. It is actually, again, strongly related to the theory of distributions and the formalism of Green functions.

Example 30.
If we consider Laplace equation

$$\nabla^2 f(\mathbf{x}) = \rho(\mathbf{x}),$$

with ρ an arbitrary function and $\nabla^2 = \partial_x^2 + \partial_y^2 + \partial_z^2$ in 3D: such an equation is not easy to solve. We consider the following problem

$$\nabla^2 G(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}'),$$

where G is called the *Green function* of the Laplace equation. Since the differential operator is invariant by translation, we can write $G : (\mathbf{x}, \mathbf{x}') \mapsto G(\mathbf{x} - \mathbf{x}')$, so we want to solve

$$\nabla^2 G(\mathbf{x} - \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}').$$

This yields, after a 3D Fourier transform and since $\hat{\delta} = 1$, to

$$\hat{G}(\mathbf{k}) = \frac{1}{\mathbf{k}^2}.$$

We now return to the direct space with an inverse Fourier transform

$$\begin{aligned}
G(\mathbf{x}) &= \frac{1}{(2\pi)^3} \iiint_{\mathbb{R}^3} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{\mathbf{k}^2} \mathrm{d}\mathbf{k} \\
&= \frac{1}{(2\pi)^3} \int_0^{2\pi} \int_0^{\pi} \int_0^{+\infty} \frac{e^{ikx\cos\theta}}{k^2} k^2 \sin\theta \mathrm{d}k \mathrm{d}\theta \mathrm{d}\phi \qquad \text{(spherical coordinates)} \\
&= \frac{1}{2\pi^2} \int_0^{+\infty} \frac{\sin kx}{kx} \mathrm{d}k \\
&= \frac{1}{2\pi^2 x} \int_0^{+\infty} \frac{\sin u}{u} \mathrm{d}u \\
&= \frac{1}{4\pi x} \qquad \left(\text{because } \int_0^{+\infty} \frac{\sin u}{u} \mathrm{d}u = \frac{\pi}{2}\right).
\end{aligned}$$

Therefore, the Green functions G of this problem are given by $G : x \mapsto (4\pi x)^{-1} + \phi$ where ϕ is a function satisfying $\nabla^2 \phi = 0$. Then, we note that

$$(\delta * \rho)(\mathbf{x}) = \rho(\mathbf{x}),$$

so the solutions of the original 3D Laplace equation are given by

$$f : \mathbf{x} \mapsto \int G(\mathbf{x} - \mathbf{x}') \rho(\mathbf{x}') \mathrm{d}\mathbf{x}' = \int \frac{\rho(\mathbf{x}')}{4\pi |\mathbf{x} - \mathbf{x}'|} \mathrm{d}\mathbf{x}'.$$

EXERCISE 120. *★★ 2D Green operator.*
Compute the Green operator of the 2D Laplace equation, and write the general solution for the 2D Laplace equation with a source term $\rho(\mathbf{x})$.

10.5 Conclusions

This chapter concludes this class on math analysis by providing an overview of the harmonic analysis introduced by Fourier. This is a broad and dense domain, which is why we skipped many proofs, but also many intriguing propositions and theorems. I encourage you to look at these in more detail if this part of analysis is of interest to you. Today, the Fourier analysis is widely used to solve ODE and PDE problems, to provide convergence proofs on solutions of differential equations (or estimates before a solution breaks down, or before singularities or shock emerge), to analyse data (from physical sensors, or in computer science), and many other applications. Here, we have mostly seen the mathematical aspect of the Fourier transform, but there are many aspects that are very interesting in applied maths and in physics. For example, most of the wave theories are based on Fourier formalism, and in optics diffraction and interference figures are directly linked to Fourier transforms (it is event possible to perform and visualise the optic Fourier transform).

It is important to keep in mind that other transforms exist and we can cite, among others: the Laplace transform (formally, the Fourier and Laplace transform are linked, but have different meanings: the Fourier transform is designed to describe periodic processes, while the Laplace transform is used to work with causal processes), the Hilbert transform (that also allows to recover the direction in which a signal propagates), or the Hankel transform (based on Bessel functions to describe radially-dependent functions in polar coordinates instead of using Cartesian coordinates).