
Measure Theory and Fractal Geometry of Deterministic and Random Functions

Arnaud Durand

Laboratoire de Mathématiques d'Orsay

Université Paris-Saclay, France

arnaud.durand@universite-paris-saclay.fr

Beihang, September 2026

Abstract

These lecture notes develop measure theory and its applications to the fractal geometry of deterministic and random functions. We begin with measures and integration, including convergence theorems, product measures, and Fubini's theorem. We then study Borel and Radon measures, the Riesz representation theorem, weak convergence, and Lebesgue measure. This foundation leads to Hausdorff measures, Hausdorff and box counting dimensions, the mass distribution principle, Frostman's lemma, and capacities. We apply these tools to Hölder continuous functions and the Weierstrass function, examining regularity, nowhere differentiability, and graph dimensions. Finally, we construct Brownian motion and study its modulus of continuity, nowhere differentiability, and the fractal dimensions of its graph and images.

Contents

1	Measures	3
2	Integration	6
2.1	Measurable and simple functions	6
2.2	Integration of nonnegative functions	7
2.3	Integration of complex functions	12
3	Product measures	16
3.1	Definition and measurability properties	16
3.2	Fubini's theorem	18
4	Borel and Radon measures	20
4.1	Definitions	20
4.2	Urysohn's lemma and continuity of positive linear functionals	23
4.3	The Riesz representation theorem	25
4.4	Regularity and approximation properties of Radon measures	27
4.5	Image measure	29
4.6	Weak convergence	31
5	Metric measures	34
5.1	Definition and Borel property	34
5.2	The Lebesgue measure	36
5.3	Tail integrals and integral inequalities	38
6	Hausdorff measures, dimensions, and capacities	41
6.1	Hausdorff measures and contents	41
6.2	Hausdorff dimension	42
6.3	Box counting dimensions	44
6.4	Mass distribution principle and Frostman's lemma	45
6.5	Capacities	48
7	Hölder continuous functions and their graphs	49
7.1	General results	49
7.2	The Weierstrass function	53
7.2.1	Hölder continuity and modulus of continuity	54
7.2.2	Nowhere differentiability	54
7.2.3	Lower Hölder estimate	60
7.2.4	Graph dimension	61
8	Fractal properties of Brownian motion	62
8.1	Basics on probability spaces and random variables	62
8.2	Gaussian random variables	64
8.3	Lévy's construction of Brownian motion	68
8.4	Hölder continuity and modulus of continuity	71
8.5	Nowhere differentiability	72
8.6	Dimensions of graph and images	73

Prerequisites and external results. We assume elementary real analysis, linear algebra, and the basic language of metric and topological spaces, including compactness, continuity, and countable bases. The measure-theoretic arguments are developed below. A few background results from topology, approximation theory, and probability are used without proof; their precise forms are stated before use and references are supplied. The multivariable change-of-variables formula used for Gaussian integrals is also recalled before those calculations.

1 Measures

Let X be a set. The *power set* of X , denoted by 2^X , consists of all subsets of X . Following [4, 8, 9], we define a *measure* on X as a function $\mu: 2^X \rightarrow [0, \infty]$, and subsequently define *measurability* for sets with respect to a measure. While such functions are typically called *outer measures* in the literature, we simplify terminology by referring to them exclusively as measures.

Definition 1 (Measure, measurable set). Let X be a set. A *measure* on X is a function $\mu: 2^X \rightarrow [0, \infty]$ satisfying the following properties:

- (i). $\mu(\emptyset) = 0$;
- (ii). If $A \subseteq B \subseteq X$, then $\mu(A) \leq \mu(B)$;
- (iii). *Countable subadditivity*: for any sequence $(A_i)_{i \geq 1}$ of subsets of X , then

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

A measure μ on X is called a *probability measure* if it satisfies $\mu(X) = 1$. Moreover, given a measure μ on X , a set $A \subseteq X$ is called μ -*measurable* if for any set $E \subseteq X$,

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

The collection of all μ -measurable sets in X is denoted by $\mathcal{M}_\mu$.

Furthermore, the collection of all μ -measurable sets satisfies special stability properties that are described by the following definition.

Definition 2 (σ -algebra). Let X be a set. A σ -*algebra* on X is a collection $\mathcal{A} \subseteq 2^X$ satisfying the following properties:

- (i). $\emptyset \in \mathcal{A}$;
- (ii). If $A \in \mathcal{A}$, then $X \setminus A \in \mathcal{A}$;
- (iii). If $(A_i)_{i \geq 1}$ is a sequence of sets in $\mathcal{A}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

An important example of σ -algebra is the *Borel σ -algebra* induced by a topological structure on a set. A *topology* on a set X is a collection of sets in X , called *open sets*, such that any union of open sets is an open set, and any finite intersection of open sets is an open set. If a set X is endowed with a topology, this set is called a *topological space*. The Borel σ -algebra on X , denoted by $\mathcal{B}(X)$, is then defined as the smallest σ -algebra containing all open sets of X .

The following result shows that the collection $\mathcal{M}_\mu$ of μ -measurable sets is a σ -algebra, and establishes the countable additivity of the measure μ on this collection. We use the symbol $\sqcup$ to indicate that the sets forming the union are pairwise disjoint.

Theorem 1 (σ -algebra of measurable sets, countable additivity on measurable sets). *Let X be a set and let μ be a measure on X . The collection $\mathcal{M}_\mu$ of μ -measurable sets in X is a σ -algebra on X . Moreover, the following properties hold:*

- (i). *The null sets are measurable, that is, if a set $N \subseteq X$ satisfies $\mu(N) = 0$, then $N \in \mathcal{M}_\mu$;*
- (ii). *For any set $E \subseteq X$ and for any sequence $(A_i)_{i \geq 1}$ of pairwise disjoint sets in $\mathcal{M}_\mu$,*

$$\mu\left(E \cap \bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(E \cap A_i);$$

For $E = X$, this entails that the measure μ , restricted to $\mathcal{M}_\mu$, is countably additive.

Proof. The collection $\mathcal{M}_\mu$ is closed under complements, due to the symmetry of the definition of measurability. To show that null sets are measurable, consider $N \subseteq X$ with $\mu(N) = 0$, and $E \subseteq X$. Then,

$$\mu(E \setminus N) \leq \mu(E) \leq \mu(E \cap N) + \mu(E \setminus N) \leq \mu(N) + \mu(E \setminus N) = \mu(E \setminus N).$$

We thus have equality throughout, which shows that $N \in \mathcal{M}_\mu$.

To prove closure under countable unions, first take $A, B \in \mathcal{M}_\mu$ and $E \subseteq X$. We have

$$\begin{aligned} \mu(E \cap (A \cup B)) + \mu(E \setminus (A \cup B)) &= (\mu(E \cap (A \cup B) \cap A) + \mu(E \cap (A \cup B) \setminus A)) + \mu(E \setminus (A \cup B)) \\ &= \mu(E \cap A) + \mu((E \setminus A) \cap B) + \mu((E \setminus A) \setminus B) \\ &= \mu(E \cap A) + \mu(E \setminus A) = \mu(E), \end{aligned}$$

from which we conclude that $A \cup B \in \mathcal{M}_\mu$. Furthermore, let us consider a sequence $(A_i)_{i \geq 1}$ of pairwise disjoint sets in $\mathcal{M}_\mu$ with union $A = \bigsqcup_{i=1}^{\infty} A_i$, and a set $E \subseteq X$. Then, for any $n \geq 1$,

$$\begin{aligned} \mu(E \cap (A_1 \sqcup \cdots \sqcup A_n)) &= \mu(E \cap (A_1 \sqcup \cdots \sqcup A_n) \cap A_n) + \mu(E \cap (A_1 \sqcup \cdots \sqcup A_n) \setminus A_n) \\ &= \mu(E \cap A_n) + \mu(E \cap (A_1 \sqcup \cdots \sqcup A_{n-1})) \\ &= \cdots = \mu(E \cap A_n) + \cdots + \mu(E \cap A_1). \end{aligned}$$

Moreover, the previous argument shows that $A_1 \sqcup \cdots \sqcup A_n \in \mathcal{M}_\mu$, so that

$$\begin{aligned} \mu(E) &= \mu(E \cap (A_1 \sqcup \cdots \sqcup A_n)) + \mu(E \setminus (A_1 \sqcup \cdots \sqcup A_n)) \\ &= \mu(E \cap A_1) + \cdots + \mu(E \cap A_n) + \mu(E \setminus (A_1 \sqcup \cdots \sqcup A_n)) \\ &\geq \mu(E \cap A_1) + \cdots + \mu(E \cap A_n) + \mu(E \setminus A). \end{aligned}$$

Since this bound holds for all $n \geq 1$, we deduce that

$$\mu(E) \geq \left(\sum_{i=1}^{\infty} \mu(E \cap A_i) \right) + \mu(E \setminus A) \geq \mu(E \cap A) + \mu(E \setminus A) \geq \mu(E).$$

We thus have equality throughout, which shows that $A \in \mathcal{M}_\mu$. In addition, replacing E by $E \cap A$, we get

$$\mu(E \cap A) = \left(\sum_{i=1}^{\infty} \mu(E \cap A_i) \right) + \mu((E \cap A) \setminus A) = \sum_{i=1}^{\infty} \mu(E \cap A_i).$$

Finally, if the sets A_i are not pairwise disjoint, the sets $A'_i = A_i \setminus (A_1 \cup \cdots \cup A_{i-1})$ are pairwise disjoint, belong to $\mathcal{M}_\mu$, and their union satisfies $\bigsqcup_{i=1}^{\infty} A'_i = \bigcup_{i=1}^{\infty} A_i$. We can therefore apply the previous argument to conclude that this union is also in $\mathcal{M}_\mu$. $\square$

By Definition 1 and Theorem 1, a measure is a nonnegative set function defined on 2^X that vanishes on the empty set and is nondecreasing, countably subadditive, and countably additive on the σ -algebra of its measurable sets. Conversely, any nonnegative set function defined on a σ -algebra that vanishes on the empty set and is countably additive can be extended to a measure on X .

Theorem 2 (Extension of a countably additive set function). *Let X be a set, let $\mathcal{A}$ be a σ -algebra on X , and let $\nu: \mathcal{A} \rightarrow [0, \infty]$ be a countably additive set function satisfying $\nu(\emptyset) = 0$. (Hence, $\nu(\bigsqcup_i A_i) = \sum_i \nu(A_i)$ for any sequence $(A_i)_{i \geq 1}$ of pairwise disjoint sets in $\mathcal{A}$.) Then, the formula*

$$\mu(A) = \inf\{\nu(B), B \in \mathcal{A} \text{ with } B \supseteq A\},$$

where the infimum is attained, defines a measure on X that extends ν and verifies that all sets in $\mathcal{A}$ are μ -measurable. Specifically, for any subset A of X , the following properties hold:

- (i). *There exists a set $B \in \mathcal{A}$ such that $B \supseteq A$ and $\nu(B) = \mu(A)$;*
- (ii). *If $A \in \mathcal{A}$, then $\mu(A) = \nu(A)$ and $A \in \mathcal{M}_\mu$.*

Proof. It is clear from the formula that $\mu(A) = \nu(A)$ for any set $A \in \mathcal{A}$. In particular, $\mu(\emptyset) = 0$. Moreover, if $A_1 \subseteq A_2 \subseteq X$, then any set in $\mathcal{A}$ that contains A_2 also contains A_1 , so that $\mu(A_1) \leq \mu(A_2)$.

To prove that the infimum is attained, let us consider a set $A \subseteq X$ such that $\mu(A) < \infty$, and observe that for any integer $n \geq 1$, there exists a set $B_n \in \mathcal{A}$ such that $B_n \supseteq A$ and $\nu(B_n) \leq \mu(A) + 1/n$. Then, the intersection $B = \bigcap_{n=1}^{\infty} B_n$ is in the σ -algebra $\mathcal{A}$ and satisfies $B \supseteq A$. In addition $\nu(B) = \mu(A)$, since

$$\mu(A) \leq \nu(B) \leq \nu(B) + \nu(B_n \setminus B) = \nu(B_n) \leq \mu(A) + 1/n \xrightarrow{n \rightarrow \infty} \mu(A).$$

If $\mu(A) = \infty$, the infimum is attained at $B = X$ since $X \in \mathcal{A}$ and $\nu(X) = \mu(X) \geq \mu(A) = \infty$.

Now, for any sequence $(A_i)_{i \geq 1}$ of subsets of X , there exists a sequence $(B_i)_{i \geq 1}$ of sets in $\mathcal{A}$ such that $B_i \supseteq A_i$ and $\nu(B_i) = \mu(A_i)$. The union $\bigcup_{i=1}^{\infty} B_i$ is in $\mathcal{A}$, and contains the union $\bigcup_{i=1}^{\infty} A_i$. Thus,

$$\begin{aligned} \mu\left(\bigcup_{i=1}^{\infty} A_i\right) &\leq \nu\left(\bigcup_{i=1}^{\infty} B_i\right) = \nu\left(\bigsqcup_{i=1}^{\infty} B_i \setminus (B_1 \cup \dots \cup B_{i-1})\right) = \sum_{i=1}^{\infty} \nu(B_i \setminus (B_1 \cup \dots \cup B_{i-1})) \\ &\leq \sum_{i=1}^{\infty} (\nu(B_i \setminus (B_1 \cup \dots \cup B_{i-1})) + \nu(B_i \cap (B_1 \cup \dots \cup B_{i-1}))) = \sum_{i=1}^{\infty} \nu(B_i) = \sum_{i=1}^{\infty} \mu(A_i). \end{aligned} \tag{1}$$

It remains to show that $\mathcal{A} \subseteq \mathcal{M}_\mu$. Let us consider a set $A \in \mathcal{A}$, a set $E \subseteq X$, and a set $B \in \mathcal{A}$ such that $B \supseteq E$ and $\nu(B) = \mu(E)$. Since $B \cap A$ and $B \setminus A$ are disjoint and in $\mathcal{A}$, we obtain

$$\mu(E) \leq \mu(E \cap A) + \mu(E \setminus A) \leq \nu(B \cap A) + \nu(B \setminus A) = \nu(B) = \mu(E).$$

We have equality throughout, so we can conclude that $A \in \mathcal{M}_\mu$. $\square$

Theorems 1 and 2 allow us to pass between our framework and the standard σ -algebra approach: restriction gives a countably additive set function on the measurable sets, and extension preserves its values there. These operations need not recover the original outer measure on nonmeasurable sets. We now characterize measure behavior under monotone set sequences.

Theorem 3 (Measure of monotone unions and intersections). *Let X be a set, let μ be a measure on X , let $(A_i)_{i \geq 1}$ be a sequence of sets in $\mathcal{M}_\mu$, and let E be a subset of X . Then, the following properties hold:*

- (i). *If $A_1 \subseteq A_2 \subseteq \dots$, then $\mu\left(E \cap \bigcup_{i=1}^{\infty} A_i\right) = \sup_{i \geq 1} \mu(E \cap A_i)$;*
- (ii). *If $A_1 \supseteq A_2 \supseteq \dots$ and $\mu(E \cap A_i) < \infty$ for some i , then $\mu\left(E \cap \bigcap_{i=1}^{\infty} A_i\right) = \inf_{i \geq 1} \mu(E \cap A_i)$.*

Moreover, these extrema can be written as limits, because the sequences therein are monotone.

Proof. For the first property, let us define $B_1 = A_1$ and $B_i = A_i \setminus A_{i-1}$ for $i \geq 2$. The sets B_i are pairwise disjoint, belong to $\mathcal{M}_\mu$, and their union is equal to $\bigcup_{i=1}^{\infty} A_i$. Theorem 1 then yields

$$\mu\left(E \cap \bigcup_{i=1}^{\infty} A_i\right) = \sum_{j=1}^{\infty} \mu(E \cap B_j) = \sup_{i \geq 1} \sum_{j=1}^i \mu(E \cap B_j) = \sup_{i \geq 1} \mu\left(E \cap \bigcup_{j=1}^i B_j\right) = \sup_{i \geq 1} \mu(E \cap A_i).$$

For the second property, let us suppose that $\mu(E \cap A_{i_0}) < \infty$ and define $C_i = A_{i_0} \setminus A_{i_0+i}$ for $i \geq 1$. The sets $C_1 \subseteq C_2 \subseteq \dots$ belong to $\mathcal{M}_\mu$, so the first property yields

$$\mu\left((E \cap A_{i_0}) \setminus \bigcap_{i=1}^{\infty} A_{i_0+i}\right) = \mu\left(E \cap \bigcup_{i=1}^{\infty} C_i\right) = \sup_{i \geq 1} \mu(E \cap C_i) = \sup_{i \geq 1} \mu(E \cap (A_{i_0} \setminus A_{i_0+i})).$$

Moreover, the measurability of the sets A_{i_0+i} implies that $\mu(E \cap A_{i_0})$, which is finite, is equal to both

$$\begin{aligned} & \mu\left((E \cap A_{i_0}) \cap \bigcap_{i=1}^{\infty} A_{i_0+i}\right) + \mu\left((E \cap A_{i_0}) \setminus \bigcap_{i=1}^{\infty} A_{i_0+i}\right) \\ & \mu((E \cap A_{i_0}) \cap A_{i_0+i}) + \mu((E \cap A_{i_0}) \setminus A_{i_0+i}). \end{aligned}$$

In particular, all these terms are finite, and we deduce that

$$\begin{aligned} \mu(E \cap A_{i_0}) - \mu\left(E \cap \bigcap_{i=1}^{\infty} A_{i_0+i}\right) &= \mu\left((E \cap A_{i_0}) \setminus \bigcap_{i=1}^{\infty} A_{i_0+i}\right) = \sup_{i \geq 1} \mu(E \cap (A_{i_0} \setminus A_{i_0+i})) \\ &= \sup_{i \geq 1} (\mu(E \cap A_{i_0}) - \mu(E \cap A_{i_0+i})) \\ &= \mu(E \cap A_{i_0}) - \inf_{i \geq 1} \mu(E \cap A_{i_0+i}). \end{aligned}$$

This proves the desired equality, since $\bigcap_{i=1}^{\infty} A_{i_0+i} = \bigcap_{i=1}^{\infty} A_i$. □

2 Integration

2.1 Measurable and simple functions

Our study of integration theory begins with the concepts of *measurable* and *simple* functions.

Definition 3 (Measurable function, simple function). Let X be a set, let μ be a measure on X , and let f be a function defined on X and taking values in a topological space Y , endowed with its Borel σ -algebra $\mathcal{B}(Y)$. The function f is called:

- (i). μ -measurable if the inverse image $f^{-1}(B)$ is in $\mathcal{M}_\mu$ for any Borel set $B \in \mathcal{B}(Y)$;
- (ii). simple if its image $f(X)$ is a finite set;
- (iii). Borel-measurable if X is a topological space and the inverse image $f^{-1}(B)$ is in $\mathcal{B}(X)$ for any Borel set $B \in \mathcal{B}(Y)$.

The set of all μ -measurable functions from X to $\mathbb{C}$, and to $[0, \infty]$, are denoted by $\mathcal{L}_\mu$, and by $\mathcal{L}_\mu^+$, respectively. The set of all simple μ -measurable functions from X to $[0, \infty)$ is denoted by $\mathcal{S}_\mu^+$.

The Borel σ -algebra being generated by open sets, it suffices to check the measurability conditions for inverse images of open sets. Thus, it is easy to see that continuous functions are Borel-measurable. In addition, if μ is a Borel measure, meaning that all Borel sets are μ -measurable, any Borel-measurable function is μ -measurable. We refer to Definition 9 below for the definition of Borel measures.

2.2 Integration of nonnegative functions

We first integrate simple functions, then nonnegative measurable functions.

Definition 4 (Integral of a nonnegative simple measurable function). Let X be a set and let μ be a measure on X . The *integral* of a function $f \in \mathcal{S}_\mu^+$ with respect to μ is defined, adopting the convention that $0 \cdot \infty = 0$, by the formula

$$\int_X f \, d\mu = \sum_{y \in f(X)} y \, \mu(f^{-1}(\{y\})).$$

The above sum that defines the integral of $f \in \mathcal{S}_\mu^+$ has a finite number of terms, because the image $f(X)$ is finite. These terms belong to $[0, \infty]$. By convention, if $0 \in f(X)$, the corresponding term $0 \mu(f^{-1}(\{0\}))$ in the sum is equal to zero, even when the measure $\mu(f^{-1}(\{0\}))$ is infinite. The integral of a nonnegative simple measurable function is thus well-defined, and belongs to $[0, \infty]$. Furthermore, the integral of a function $f \in \mathcal{S}_\mu^+$ over a set $A \in \mathcal{M}_\mu$ is defined by

$$\int_A f \, d\mu = \int_X \mathbf{1}_A f \, d\mu, \tag{2}$$

where $\mathbf{1}_A$ is the indicator function of A . This integral is well-defined, because $\mathbf{1}_A f \in \mathcal{S}_\mu^+$. Finally, nonnegative simple measurable functions have integrals satisfying the properties listed below. The proof of these properties reveals why Definition 4 requires measurability, even though the formula therein remains well-defined without this condition.

Proposition 4 (Properties of integrals of nonnegative simple measurable functions). *Let X be a set and let μ be a measure on X . The following properties hold for any $f, g \in \mathcal{S}_\mu^+$:*

- (i). *If $c \in [0, \infty)$, then $\int_X c f \, d\mu = c \int_X f \, d\mu$;*
- (ii). *$\int_X (f + g) \, d\mu = \int_X f \, d\mu + \int_X g \, d\mu$;*
- (iii). *If $g \leq f$, then $\int_X g \, d\mu \leq \int_X f \, d\mu$;*
- (iv). *The mapping $\nu: A \mapsto \inf\{\int_B f \, d\mu, B \in \mathcal{M}_\mu \text{ with } A \subseteq B\}$ is a measure on X , and $\mathcal{M}_\mu \subseteq \mathcal{M}_\nu$.*

Proof. The first property is obvious for $c = 0$, so we assume that $c > 0$. The function $c f$ is simple, its image is $\{c y, y \in f(X)\}$, and the inverse image of $\{c y\}$ under $c f$ is equal to $f^{-1}(\{y\})$. Therefore,

$$\int_X c f \, d\mu = \sum_{y \in f(X)} c y \mu((c f)^{-1}(\{c y\})) = c \sum_{y \in f(X)} y \mu(f^{-1}(\{y\})) = c \int_X f \, d\mu.$$

To prove the second property, we first note that the function $f + g$ is nonnegative, simple, and μ -measurable. The measurability of the sum comes from the continuity of the addition operation. We then use the countable additivity of μ on its measurable sets, together with the fact that the sets $f^{-1}(\{y\})$, for $y \in f(X)$, yield a measurable partition of X . Hence, we have

$$\int_X g \, d\mu = \sum_{z \in g(X)} z \mu(g^{-1}(\{z\})) = \sum_{z \in g(X)} \sum_{y \in f(X)} z \mu(f^{-1}(\{y\}) \cap g^{-1}(\{z\})),$$

and a similar expression for $\int_X f \, d\mu$, exchanging the roles of f and g . Adding these expressions and using the partition into sets $(f + g)^{-1}(\{w\})$, for $w \in (f + g)(X)$, we obtain

$$\begin{aligned} \int_X f \, d\mu + \int_X g \, d\mu &= \sum_{y \in f(X)} \sum_{z \in g(X)} (y + z) \mu(f^{-1}(\{y\}) \cap g^{-1}(\{z\})) \\ &= \sum_{y \in f(X)} \sum_{z \in g(X)} \sum_{w \in (f+g)(X)} (y + z) \mu(f^{-1}(\{y\}) \cap g^{-1}(\{z\}) \cap (f + g)^{-1}(\{w\})). \end{aligned}$$

The latter intersection above is empty unless $w = y + z$. Since the sets $f^{-1}(\{y\}) \cap g^{-1}(\{z\})$ that are nonempty form a partition of X , the right-hand side above can be rewritten as

$$\sum_{w \in (f+g)(X)} w \mu((f + g)^{-1}(\{w\})) = \int_X (f + g) \, d\mu.$$

The third property follows from the second property and the observation that if $g \leq f$, then $f - g$ is nonnegative, simple, and μ -measurable. Consequently,

$$\int_X f \, d\mu = \int_X (g + (f - g)) \, d\mu = \int_X g \, d\mu + \int_X (f - g) \, d\mu \geq \int_X g \, d\mu.$$

The fourth property follows from Theorem 2. Indeed, ν is the extension as a measure on X of the nonnegative set function $B \mapsto \int_B f \, d\mu$ defined on $\mathcal{M}_\mu$. To check that this set function vanishes on the empty set and is countably additive, it suffices to observe that for any set $B \in \mathcal{M}_\mu$,

$$\int_B f \, d\mu = \sum_{y \in f(X)} y \mu(f^{-1}(\{y\}) \cap B),$$

and to use the countable additivity of μ on $\mathcal{M}_\mu$. □

The integral of a nonnegative measurable function is then defined as the supremum of the integrals of all simple functions that are less than or equal to this function.

Definition 5 (Integral of a nonnegative measurable function). Let X be a set and let μ be a measure on X . The *integral* of a function $f \in \mathcal{L}_\mu^+$ with respect to μ is defined by the formula

$$\int_X f \, d\mu = \sup \left\{ \int_X g \, d\mu, g \in \mathcal{S}_\mu^+ \text{ with } g \leq f \right\},$$

and then its integral over a set $A \in \mathcal{M}_\mu$ is defined by the formula (2).

The integral of a nonnegative measurable function is thus well-defined, and belongs to $[0, \infty]$. By Proposition 4(iii), the above definition is coherent with Definition 4 for simple functions. Moreover, Propositions 4(i) and 4(iii) can directly be extended to nonnegative measurable functions. We now prove a fundamental convergence theorem.

Theorem 5 (The monotone convergence theorem). Let X be a set and let μ be a measure on X . Let $(f_j)_{j \geq 1}$ be a sequence of functions in $\mathcal{L}_\mu^+$, such that $f_j \leq f_{j+1}$ for all $j \geq 1$. If f denotes the pointwise limit of f_j as $j \rightarrow \infty$, then $f \in \mathcal{L}_\mu^+$ and

$$\int_X f \, d\mu = \lim_{j \rightarrow \infty} \int_X f_j \, d\mu.$$

Proof. It is easy to check using Definition 3 that any pointwise limit of functions in $\mathcal{L}_\mu^+$ is in $\mathcal{L}_\mu^+$ too. The sequence $(\int_X f_j \, d\mu)_{j \geq 1}$ is nondecreasing and bounded above by $\int_X f \, d\mu$. Therefore, it converges to a limit $\ell \in [0, \infty]$ such that $\ell \leq \int_X f \, d\mu$. To prove the reverse inequality, let us consider $\alpha \in (0, 1)$ and $g \in \mathcal{S}_\mu^+$ such that $g \leq f$. The sets $A_j = \{x \in X \mid f_j(x) \geq \alpha g(x)\}$ are in $\mathcal{M}_\mu$, are nondecreasing in j , and their union is X . Moreover, $\int_X f_j \, d\mu \geq \int_{A_j} f_j \, d\mu \geq \alpha \int_{A_j} g \, d\mu$. Theorem 3 and Proposition 4(iv) give

$$\ell = \lim_{j \rightarrow \infty} \int_X f_j \, d\mu \geq \lim_{j \rightarrow \infty} \alpha \int_{A_j} g \, d\mu = \alpha \int_X g \, d\mu.$$

We conclude by letting $\alpha \rightarrow 1$ and taking the supremum over all $g \in \mathcal{S}_\mu^+$ such that $g \leq f$. $\square$

By Definition 5, the integral of a nonnegative measurable function is defined as a supremum of integrals of simple functions. Thanks to the monotone convergence theorem, we can also express it as a monotone limit of integrals of simple functions. The following construction gives a nondecreasing sequence of simple functions approximating any nonnegative measurable function.

Lemma 6 (Approximation of nonnegative measurable functions by simple functions). Let X be a set and let μ be a measure on X . For any function $f \in \mathcal{L}_\mu^+$, there exists a sequence $(g_j)_{j \geq 1}$ of functions in $\mathcal{S}_\mu^+$ such that $g_1 \leq g_2 \leq \dots \leq f$, and $g_j \rightarrow f$ pointwise as $j \rightarrow \infty$. Moreover, the convergence is uniform on any set where f is bounded.

Proof. For any integer $j \geq 1$, let us consider the nonnegative simple Borel-measurable function on $[0, \infty]$

$$\phi_j = 2^j \mathbf{1}_{[2^j, \infty)} + \sum_{k=0}^{2^j-1} k 2^{-j} \mathbf{1}_{(k 2^{-j}, (k+1) 2^{-j}]}.$$

For all $y \in [0, \infty]$, note that $0 \leq \phi_j(y) \leq y$, and that $\phi_j(y) \geq y - 2^{-j}$ when $y \leq 2^j$. Moreover, we can check that $\phi_{j+1}(y) \geq \phi_j(y)$ for all $y \in [0, \infty]$, by inspecting all the possible cases:

- If $y \leq 0$, then $\phi_{j+1}(y) = 0 = \phi_j(y)$;
- If $y > 2^{j+1}$, then $\phi_{j+1}(y) = 2^{j+1} \geq 2^j = \phi_j(y)$;
- If $\ell 2^{-(j+1)} < y \leq (\ell + 1)2^{-(j+1)}$ with $\ell \in \{2^{2j+1}, \dots, 2^{2(j+1)} - 1\}$, then $\phi_{j+1}(y) = \ell 2^{-(j+1)} \geq 2^j = \phi_j(y)$;
- If $\ell 2^{-(j+1)} < y \leq (\ell + 1)2^{-(j+1)}$ with $\ell \in \{0, \dots, 2^{2j+1} - 1\}$, then $\phi_{j+1}(y) = \ell 2^{-(j+1)} \geq \lfloor \frac{\ell}{2} \rfloor 2^{-j} = \phi_j(y)$.

Hence, the sequence formed by the functions $g_j = \phi_j \circ f \in \mathcal{S}_\mu^+$ fulfills the desired properties. $\square$

As a first application of the monotone convergence theorem, that is, Theorem 5, we observe that we can exchange sums and integrals for sequences of nonnegative measurable functions.

Theorem 7 (Integral of a sum of nonnegative measurable functions). *Let X be a set and let μ be a measure on X . Let $(f_j)_{j \geq 1}$ be a sequence of functions in $\mathcal{L}_\mu^+$. Then, the function $\sum_{j=1}^\infty f_j$ is in $\mathcal{L}_\mu^+$, and the following equality holds:*

$$\int_X \sum_{j=1}^\infty f_j \, d\mu = \sum_{j=1}^\infty \int_X f_j \, d\mu.$$

Proof. Note that the result encompasses the case of a finite sum, because we can consider some of the functions f_j to be zero. The result is actually equivalent to the case where the sum is finite, by virtue of Theorem 5. Therefore, by induction, it suffices to prove the result for a sum of two functions $f_1, f_2 \in \mathcal{L}_\mu^+$. By Lemma 6, there are two sequences of functions $g_{1,j}, g_{2,j} \in \mathcal{S}_\mu^+$ such that $g_{1,1} \leq g_{1,2} \leq \dots \leq f_1$ and $g_{2,1} \leq g_{2,2} \leq \dots \leq f_2$, and such that $g_{1,j} \rightarrow f_1$ and $g_{2,j} \rightarrow f_2$ pointwise as $j \rightarrow \infty$. Then, $g_{1,j} + g_{2,j} \in \mathcal{S}_\mu^+$, the sequence $(g_{1,j} + g_{2,j})_{j \geq 1}$ is nondecreasing, and its pointwise limit is $f_1 + f_2$. We conclude from Proposition 4(ii) and Theorem 5 that

$$\int_X (f_1 + f_2) \, d\mu = \lim_{j \rightarrow \infty} \int_X (g_{1,j} + g_{2,j}) \, d\mu = \lim_{j \rightarrow \infty} \left(\int_X g_{1,j} \, d\mu + \int_X g_{2,j} \, d\mu \right) = \int_X f_1 \, d\mu + \int_X f_2 \, d\mu,$$

and the result follows directly. $\square$

Another important consequence of the monotone convergence theorem is Fatou's lemma.

Theorem 8 (Fatou's lemma). *Let X be a set and let μ be a measure on X . Let $(f_j)_{j \geq 1}$ be a sequence of functions in $\mathcal{L}_\mu^+$. Then, the function $\liminf_{j \rightarrow \infty} f_j$ is in $\mathcal{L}_\mu^+$, and the following inequality holds:*

$$\int_X \liminf_{j \rightarrow \infty} f_j \, d\mu \leq \liminf_{j \rightarrow \infty} \int_X f_j \, d\mu.$$

Proof. The fact that $\liminf_{j \rightarrow \infty} f_j$ belongs to $\mathcal{L}_\mu^+$ follows directly from Definition 3. Furthermore, for any integer $j_1 \geq j_0$, we have $\inf_{j \geq j_0} f_j \leq f_{j_1}$, so that $\int_X \inf_{j \geq j_0} f_j \, d\mu \leq \int_X f_{j_1} \, d\mu$. We deduce that $\int_X \inf_{j \geq j_0} f_j \, d\mu \leq \inf_{j \geq j_0} \int_X f_j \, d\mu$. Applying Theorem 5 to the nondecreasing sequence of functions $\inf_{j \geq j_0} f_j \in \mathcal{L}_\mu^+$, indexed by $j_0 \geq 1$, we obtain

$$\int_X \liminf_{j \rightarrow \infty} f_j \, d\mu = \lim_{j_0 \rightarrow \infty} \int_X \inf_{j \geq j_0} f_j \, d\mu \leq \lim_{j_0 \rightarrow \infty} \inf_{j \geq j_0} \int_X f_j \, d\mu = \liminf_{j \rightarrow \infty} \int_X f_j \, d\mu,$$

which concludes the proof. $\square$

Recall that a property depending on $x \in X$ holds for μ -almost every $x \in X$ if the set of $x \in X$ such that the property does not hold is μ -null. Then, this property is also said to hold μ -almost everywhere. We write a.e. as a shorthand for both ‘almost every’ and ‘almost everywhere’. The next result shows that the nonnegative measurable functions with a vanishing integral are μ -a.e. equal to zero.

Proposition 9 (Functions with a null integral). *Let X be a set and let μ be a measure on X . Then, for any function $f \in \mathcal{L}_\mu^+$, it holds that*

$$\int_X f \, d\mu = 0 \quad \Longleftrightarrow \quad \mu\text{-a.e. } f = 0.$$

Proof. The result is clear when $f \in \mathcal{S}_\mu^+$, since it follows from Definition 4 that

$$\int_X f \, d\mu = 0 \quad \Longleftrightarrow \quad [\forall y \in f(X) \setminus \{0\} \quad \mu(f^{-1}(\{y\})) = 0] \quad \Longleftrightarrow \quad \mu(X \setminus f^{-1}(\{0\})) = 0.$$

In the general case, let us assume that a function $f \in \mathcal{L}_\mu^+$ is equal to zero μ -a.e. and consider a function $g \in \mathcal{S}_\mu^+$ such that $g \leq f$. Then, g is also equal to zero μ -a.e. and the above equivalence ensures that $\int_X g \, d\mu = 0$. By Definition 5, we necessarily have $\int_X f \, d\mu = 0$. Conversely, for any integer $j \in \mathbb{Z}$,

$$\int_X f \, d\mu \geq \int_X \mathbf{1}_{A_j} f \, d\mu \geq 2^{-j} \mu(A_j) \quad \text{with} \quad A_j = \{x \in X \mid f(x) \geq 2^{-j}\} \in \mathcal{M}_\mu. \quad (3)$$

If the integral in the left-hand side vanishes, then all the sets A_j , for $j \geq 0$, are μ -null. By Theorem 3, their increasing union $\{x \in X \mid f(x) > 0\}$ is μ -null as well. $\square$

Definition 6 (σ -finiteness). Let X be a set and let μ be a measure on X .

- (i). A subset A of X is σ -finite with respect to μ if $A = \bigcup_{i=1}^\infty A_i$ for some sequence of sets $A_i \in \mathcal{M}_\mu$ such that $\mu(A_i) < \infty$ for all $i \geq 1$;
- (ii). The measure μ is σ -finite if X is σ -finite with respect to μ ;
- (iii). A function f defined on X and taking values in $[0, \infty]$ or $\mathbb{C}$ is σ -finite with respect to μ if f is μ -measurable and the set $\{x \in X \mid f(x) \neq 0\}$ is σ -finite with respect to μ .

Remark 1 (Functions with a finite integral). For any function $f \in \mathcal{L}_\mu^+$ with $\int_X f \, d\mu < \infty$, the integral in the left-hand side of (3) is finite, so that $\mu(A_j) < \infty$ for all $j \geq 0$. Hence, the set $\{x \in X \mid f(x) > 0\}$ is a countable union of measurable sets with finite measure. This means that the function f is σ -finite in the sense of Definition 6. Moreover, the function f is finite μ -a.e., because (3) implies that

$$\mu(\{x \in X \mid f(x) = \infty\}) \leq \mu(A_j) \leq 2^j \int_X f \, d\mu \xrightarrow{j \rightarrow -\infty} 0.$$

Thanks to Proposition 9, we can establish versions of Fatou’s lemma and the monotone convergence theorem for a.e. convergence. Specifically, let us assume that a sequence of functions $f_j \in \mathcal{L}_\mu^+$ tends μ -a.e. to a function $f \in \mathcal{L}_\mu^+$. There exists a set $N \in \mathcal{M}_\mu$ with $\mu(N) = 0$, such that $f_j(x) \rightarrow f(x)$ as $j \rightarrow \infty$, for all $x \in X \setminus N$. By Theorem 7 and Proposition 9, we have

$$\int_X f \, d\mu = \int_X \mathbf{1}_{X \setminus N} f \, d\mu + \int_X \mathbf{1}_N f \, d\mu = \int_X \mathbf{1}_{X \setminus N} f \, d\mu, \quad (4)$$

and the same holds for the integrals of the functions f_j . Moreover, $\mathbf{1}_{X \setminus N} f_j \rightarrow \mathbf{1}_{X \setminus N} f$ pointwise as $j \rightarrow \infty$, everywhere. Applying Theorem 8, we deduce that

$$\int_X f \, d\mu = \int_X \mathbf{1}_{X \setminus N} f \, d\mu \leq \liminf_{j \rightarrow \infty} \int_X \mathbf{1}_{X \setminus N} f_j \, d\mu = \liminf_{j \rightarrow \infty} \int_X f_j \, d\mu.$$

If we assume in addition that the sequence $(f_j(x))_{j \geq 1}$ is nondecreasing for μ -a.e. $x \in X$, there is a set $N' \in \mathcal{M}_\mu$ with $\mu(N') = 0$, such that $f_j(x) \leq f_{j+1}(x)$ for all $x \in X \setminus N'$. Then, $\mathbf{1}_{X \setminus (N \cup N')} f_j \rightarrow \mathbf{1}_{X \setminus (N \cup N')} f$ pointwise and monotonically as $j \rightarrow \infty$, everywhere. Since (4) still holds with $N \cup N'$ instead of N , for the functions f and f_j , we can apply Theorem 5, which gives

$$\int_X f \, d\mu = \int_X \mathbf{1}_{X \setminus (N \cup N')} f \, d\mu = \lim_{j \rightarrow \infty} \int_X \mathbf{1}_{X \setminus (N \cup N')} f_j \, d\mu = \lim_{j \rightarrow \infty} \int_X f_j \, d\mu.$$

These methods often generalize integral results to almost everywhere conditions.

2.3 Integration of complex functions

The positive and negative parts of a real-valued function f are $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$, respectively. If f is measurable, then both f^+ and f^- are nonnegative and measurable. To ensure linearity, it is then natural to define the integral of f with respect to μ by the formula

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$

The integral of f is well-defined, and valued in $[-\infty, \infty]$, when at least one of the integrals in the right-hand side is finite. Since $|f| = f^+ + f^-$, both of these integrals are finite if and only if $\int_X |f| \, d\mu < \infty$. This entails that the integral of f is finite. Furthermore, if f is complex-valued, we can define its real part $\Re f$ and its imaginary part $\Im f$ by the decomposition $f = \Re f + i \Im f$. If f is measurable, then both $\Re f$ and $\Im f$ are real-valued and measurable. Again, to ensure linearity, it is natural to define

$$\int_X f \, d\mu = \int_X \Re f \, d\mu + i \int_X \Im f \, d\mu.$$

The integral of f is well-defined if the two integrals in the right-hand side are finite, which is guaranteed if both $\int_X |\Re f| \, d\mu$ and $\int_X |\Im f| \, d\mu$ are finite. Since $|f| \leq |\Re f| + |\Im f| \leq 2|f|$, this condition is equivalent to the finiteness of $\int_X |f| \, d\mu$. We therefore make the following definition.

Definition 7 (Integral of a complex-valued measurable function). Let X be a set and let μ be a measure on X . A function $f: X \rightarrow \mathbb{C}$ is μ -integrable if $f \in \mathcal{L}_\mu$ and $\int_X |f| \, d\mu < \infty$. The collection of all μ -integrable functions in $\mathcal{L}_\mu$ is denoted by $\mathcal{L}_\mu^1$. Moreover, the *integral* of a function $f \in \mathcal{L}_\mu^1$ with respect to μ is defined by the formula

$$\int_X f \, d\mu = \left(\int_X (\Re f)^+ \, d\mu - \int_X (\Re f)^- \, d\mu \right) + i \left(\int_X (\Im f)^+ \, d\mu - \int_X (\Im f)^- \, d\mu \right),$$

and then its integral over a set $A \in \mathcal{M}_\mu$ is defined by the formula (2).

The integral of a complex-valued measurable function was defined above with linearity in mind. Unsurprisingly, we thus have the following properties.

Proposition 10 (Linearity of integrals). *The set of real-valued μ -integrable functions is a vector space over $\mathbb{R}$. The set $\mathcal{L}_\mu^1$ of complex-valued μ -integrable functions is a vector space over $\mathbb{C}$. Moreover, the integral is a linear functional on these vector spaces. More precisely, for any $f, g \in \mathcal{L}_\mu^1$ and any $\lambda \in \mathbb{C}$, it holds that*

$$\int_X (\lambda f + g) d\mu = \lambda \int_X f d\mu + \int_X g d\mu.$$

Proof. The fact that $\mathcal{L}_\mu^1$ is a vector space over $\mathbb{C}$, and that its restriction to real-valued functions is a vector space over $\mathbb{R}$, follow from the observation that $|\lambda f + g| \leq |\lambda||f| + |g|$, combined with Theorem 7 and the extension of Proposition 4(i) to nonnegative measurable functions.

To prove the linearity of the integral, let us first assume that $\lambda \in \mathbb{R}$, and that f and g are real-valued. Letting $\lambda^+ = \max\{\lambda, 0\}$ and $\lambda^- = \max\{-\lambda, 0\}$, we have $(\lambda f)^+ = \lambda^+ f^+ + \lambda^- f^-$ and $(\lambda f)^- = \lambda^- f^+ + \lambda^+ f^-$. By Theorem 7 and the extension of Proposition 4(i) to nonnegative measurable functions, we get

$$\begin{aligned} \int_X \lambda f d\mu &= \int_X (\lambda^+ f^+ + \lambda^- f^-) d\mu - \int_X (\lambda^- f^+ + \lambda^+ f^-) d\mu \\ &= \lambda^+ \int_X f^+ d\mu + \lambda^- \int_X f^- d\mu - \lambda^- \int_X f^+ d\mu - \lambda^+ \int_X f^- d\mu \\ &= (\lambda^+ - \lambda^-) \left(\int_X f^+ d\mu - \int_X f^- d\mu \right) = \lambda \int_X f d\mu. \end{aligned}$$

Moreover, the function $h = f + g$ satisfies $h^+ + f^- + g^- = h^- + f^+ + g^+$. Theorem 7 yields

$$\int_X h^+ d\mu + \int_X f^- d\mu + \int_X g^- d\mu = \int_X h^- d\mu + \int_X f^+ d\mu + \int_X g^+ d\mu.$$

Arranging the terms in this equation, we directly obtain

$$\begin{aligned} \int_X h d\mu &= \int_X h^+ d\mu - \int_X h^- d\mu \\ &= \left(\int_X f^+ d\mu - \int_X f^- d\mu \right) + \left(\int_X g^+ d\mu - \int_X g^- d\mu \right) = \int_X f d\mu + \int_X g d\mu. \end{aligned}$$

To deal with the integral of λf in the complex case, we observe that $\Re(\lambda f) = (\Re \lambda)(\Re f) - (\Im \lambda)(\Im f)$ and $\Im(\lambda f) = (\Re \lambda)(\Im f) + (\Im \lambda)(\Re f)$. The above calculations imply that

$$\begin{aligned} \int_X \lambda f d\mu &= \int_X ((\Re \lambda)(\Re f) - (\Im \lambda)(\Im f)) d\mu + i \int_X ((\Re \lambda)(\Im f) + (\Im \lambda)(\Re f)) d\mu \\ &= \left(\Re \lambda \int_X \Re f d\mu - \Im \lambda \int_X \Im f d\mu \right) + i \left(\Re \lambda \int_X \Im f d\mu + \Im \lambda \int_X \Re f d\mu \right) \\ &= (\Re \lambda + i \Im \lambda) \left(\int_X \Re f d\mu + i \int_X \Im f d\mu \right) = \lambda \int_X f d\mu. \end{aligned}$$

Concerning the integral of $f + g$, we observe that $\Re(f + g) = \Re f + \Re g$ and $\Im(f + g) = \Im f + \Im g$, and we use again the previous calculations to obtain

$$\begin{aligned} \int_X (f + g) d\mu &= \int_X (\Re f + \Re g) d\mu + i \int_X (\Im f + \Im g) d\mu \\ &= \left(\int_X \Re f d\mu + \int_X \Re g d\mu \right) + i \left(\int_X \Im f d\mu + \int_X \Im g d\mu \right) \\ &= \left(\int_X \Re f d\mu + i \int_X \Im f d\mu \right) + \left(\int_X \Re g d\mu + i \int_X \Im g d\mu \right) \end{aligned}$$

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

The linearity is thus established for all functions f and g in $\mathcal{L}_\mu^1$, and for all $\lambda \in \mathbb{C}$. $\square$

The linearity of the integral implies its monotonicity for real-valued integrable functions. More precisely, if $f, g \in \mathcal{L}_\mu^1$ are real-valued and satisfy $f \leq g$, then Proposition 10 ensures that

$$\int_X g \, d\mu = \int_X (f + (g - f)) \, d\mu = \int_X f \, d\mu + \int_X (g - f) \, d\mu \geq \int_X f \, d\mu.$$

This property persists when the inequality $f \leq g$ holds μ -a.e., as justified by Proposition 9 and the subsequent comments. Now, let us establish the triangle inequality for integrals.

Proposition 11 (Triangle inequality for integrals). *Let X be a set and let μ be a measure on X . Then, for any function $f \in \mathcal{L}_\mu^1$, it holds that*

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu.$$

Proof. Assume that $\int_X f \, d\mu \neq 0$ and set $\alpha = |\int_X f \, d\mu| / \int_X f \, d\mu$. We have

$$\left| \int_X f \, d\mu \right| = \int_X \alpha f \, d\mu = \Re \left(\int_X \alpha f \, d\mu \right) = \int_X \Re(\alpha f) \, d\mu \leq \int_X |\alpha f| \, d\mu = \int_X |f| \, d\mu,$$

which concludes the proof. $\square$

Proposition 9 and its subsequent discussion demonstrate that null-set modifications of nonnegative measurable functions do not affect their integrals. The following proposition extends this invariance property to complex-valued functions.

Proposition 12 (Invariance of integrals under null modifications). *Let X be a set, let μ be a measure on X , and let $f, g \in \mathcal{L}_\mu^1$. Then, the following equivalence holds:*

$$\left[\forall A \in \mathcal{M}_\mu \quad \int_A f \, d\mu = \int_A g \, d\mu \right] \iff \int_X |f - g| \, d\mu = 0 \iff \mu\text{-a.e. } f = g.$$

Proof. The second equivalence follows from applying Proposition 9 to the function $|f - g| \in \mathcal{L}_\mu^+$. In addition, Propositions 10 and 11 ensure that for any set $A \in \mathcal{M}_\mu$,

$$\left| \int_A f \, d\mu - \int_A g \, d\mu \right| = \left| \int_X \mathbf{1}_A (f - g) \, d\mu \right| \leq \int_X |\mathbf{1}_A (f - g)| \, d\mu \leq \int_X |f - g| \, d\mu.$$

This proves one direction of the first equivalence. Finally, if it is false that f and g coincide μ -a.e., then at least one function among $(\Re(f - g))^+$, $(\Re(f - g))^-$, $(\Im(f - g))^+$, and $(\Im(f - g))^-$ is positive on a set of positive measure. To fix ideas, if $A = \{x \in X \mid (\Re(f - g))^+(x) > 0\}$ satisfies $\mu(A) > 0$, then

$$\Re \left(\int_A f \, d\mu - \int_A g \, d\mu \right) = \int_A \Re(f - g) \, d\mu = \int_X \mathbf{1}_A (\Re(f - g))^+ \, d\mu > 0.$$

The last equality holds because $(\Re(f - g))^- = 0$ on A , and Proposition 9 gives the strict positivity. The other cases follow in the same way. $\square$

Proposition 12 establishes that integrable functions may be modified on μ -null sets without affecting their integrals. For integration purposes, it thus makes no difference to consider, instead of $\mathcal{L}_\mu^1$, the collection of all equivalence classes of functions defined on X and μ -a.e. equal to some function in $\mathcal{L}_\mu^1$, for the equivalence relation corresponding to the μ -a.e. equality. This collection of equivalence classes is denoted by L_μ^1 . For any set $A \in \mathcal{M}_\mu$, writing $\int_A f \, d\mu$ with $f \in L_\mu^1$ implicitly means that we consider the integral of an arbitrary function in $\mathcal{L}_\mu^1$ that coincides μ -a.e. with f . It follows from Proposition 12 that the value of the integral does not depend on the choice of the representative function in $\mathcal{L}_\mu^1$. Finally, Proposition 10 pertains to the equivalence classes. Indeed, L_μ^1 is a vector space over $\mathbb{C}$, and the integral is a linear functional on this vector space. Likewise, the μ -a.e. real-valued functions in L_μ^1 form a vector space over $\mathbb{R}$. These vector spaces can be endowed with the L_μ^1 -norm

$$\|f\|_{L_\mu^1} = \int_X |f| \, d\mu.$$

This induces a metric $(f, g) \mapsto \|f - g\|_{L_\mu^1}$ on L_μ^1 . The corresponding convergence is the L_μ^1 -convergence. For all functions $f, f_1, f_2, \dots \in L_\mu^1$, the norm definition and Propositions 10 and 11 ensure that

$$f_j \xrightarrow{j \rightarrow \infty} f \iff \int_X |f_j - f| \, d\mu \xrightarrow{j \rightarrow \infty} 0 \implies \int_X f_j \, d\mu \xrightarrow{j \rightarrow \infty} \int_X f \, d\mu. \quad (5)$$

The following important result gives a sufficient condition for L_μ^1 -convergence.

Theorem 13 (The dominated convergence theorem). *Let X be a set, let μ be a measure on X , and let $(f_j)_{j \geq 1}$ be a sequence of functions in L_μ^1 . Assume that there exist a function f defined on X and a function $g \in L_\mu^1$ such that for μ -a.e. $x \in X$,*

$$f_j(x) \xrightarrow{j \rightarrow \infty} f(x) \quad \text{and} \quad \sup_{j \geq 1} |f_j(x)| \leq g(x).$$

Then, the following properties hold:

$$f_j \xrightarrow{j \rightarrow \infty} f \in L_\mu^1 \quad \text{and so} \quad \int_X f_j \, d\mu \xrightarrow{j \rightarrow \infty} \int_X f \, d\mu.$$

Proof. Up to modifying the functions f, g and f_j on a μ -null set, we can assume that the function g and all the functions f_j belong to $\mathcal{L}_\mu^1$, and that the convergence and the domination assumptions hold everywhere. Hence, f is a pointwise limit of measurable functions and $|f| \leq g$ on X , so that $f \in \mathcal{L}_\mu^1$. Since $|f_j - f| \leq |f_j| + |f| \leq 2g$ by the triangle inequality, we can apply Fatou's lemma, that is, Theorem 8 to the functions $2g - |f_j - f| \in \mathcal{L}_\mu^+$. Together with Proposition 10, this gives

$$\begin{aligned} \int_X 2g \, d\mu &= \int_X \liminf_{j \rightarrow \infty} (2g - |f_j - f|) \, d\mu \\ &\leq \liminf_{j \rightarrow \infty} \int_X (2g - |f_j - f|) \, d\mu \leq \int_X 2g \, d\mu - \limsup_{j \rightarrow \infty} \int_X |f_j - f| \, d\mu. \end{aligned}$$

We conclude with the help of (5). □

Thanks to the dominated convergence theorem, we can establish the following convergence result for series of integrable functions, which can be compared with Theorem 7.

Theorem 14 (Series of integrable functions). *Let X be a set, let μ be a measure on X , and let $(f_j)_{j \geq 1}$ be a sequence of functions in L^1_μ such that $\sum_j \int_X |f_j| d\mu < \infty$. Then, the series $\sum_j f_j$ converges μ -a.e. to a function in L^1_μ , and*

$$\int_X \sum_{j=1}^{\infty} f_j d\mu = \sum_{j=1}^{\infty} \int_X f_j d\mu.$$

Proof. The functions $|f_j|$ are in $\mathcal{L}^+_\mu$. By Theorem 7, their sum $g = \sum_{j=1}^{\infty} |f_j|$ is also in $\mathcal{L}^+_\mu$ and satisfies $\int_X g d\mu = \sum_{j=1}^{\infty} \int_X |f_j| d\mu < \infty$. Therefore, g is in L^1_μ . By Remark 1, the function g is finite μ -a.e., so the series $f = \sum_j f_j$ is μ -a.e. absolutely convergent. The partial sums belong to L^1_μ and satisfy $|\sum_{j=1}^n f_j| \leq g$, by the triangle inequality. We conclude thanks to Theorem 13. $\square$

3 Product measures

3.1 Definition and measurability properties

The product of two measures is the measure on the product set that is defined as follows.

Definition 8 (Product measure). Let X and Y be sets, and let μ and ν be measures on X and Y , respectively. The *product measure* of μ and ν is the set function $\mu \otimes \nu: 2^{X \times Y} \rightarrow [0, \infty]$ defined for each set $E \subseteq X \times Y$ by the formula

$$\mu \otimes \nu(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i), E \subseteq \bigcup_{i=1}^{\infty} (A_i \times B_i) \text{ with } (A_i, B_i) \in \mathcal{M}_\mu \times \mathcal{M}_\nu \right\}.$$

The following result shows that the product measure $\mu \otimes \nu$ is a measure on $X \times Y$, in the sense of Definition 1, and discusses its measurability properties. According to our notations, the σ -algebra of measurable sets with respect to the product measure $\mu \otimes \nu$ is denoted by $\mathcal{M}_{\mu \otimes \nu}$.

For any subset E of a product set $X \times Y$, we define the *cross sections* of E as the sets

$$E_x = \{y \in Y \mid (x, y) \in E\} \quad \text{and} \quad E_y = \{x \in X \mid (x, y) \in E\}.$$

Proposition 15 (Measurability with respect to the product measure). *Let X and Y be sets, and let μ and ν be measures on X and Y , respectively. Then, the product measure $\mu \otimes \nu$ is a measure on $X \times Y$, and the following properties hold:*

- (i). If $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$, then $A \times B \in \mathcal{M}_{\mu \otimes \nu}$ and $\mu \otimes \nu(A \times B) = \mu(A) \nu(B)$;
- (ii). For any set $E \subseteq X \times Y$, it holds that $\mu \otimes \nu(E) = \inf \{ \mu \otimes \nu(F), F \in \mathcal{M}_{\mu \otimes \nu} \text{ with } F \supseteq E \}$;
- (iii). The above infimum is attained by a set F that is a countable intersection of countable unions of sets of the form $A \times B$, where $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$.

Proof. It is clear that $\mu \otimes \nu(\emptyset) = 0$, and that $\mu \otimes \nu(E) \leq \mu \otimes \nu(F)$ when $E \subseteq F$, because any covering of F is also a covering of E . To prove the countable subadditivity property, let us consider a sequence of sets $E_j \subseteq X \times Y$. We can assume that the sum $\sum_{j=1}^{\infty} \mu \otimes \nu(E_j)$ is finite. For each $\varepsilon > 0$ and $j \geq 1$, choose sequences $A_{j,i} \in \mathcal{M}_\mu$ and $B_{j,i} \in \mathcal{M}_\nu$ such that

$$E_j \subseteq \bigcup_{i=1}^{\infty} (A_{j,i} \times B_{j,i}) \quad \text{and} \quad \sum_{i=1}^{\infty} \mu(A_{j,i}) \nu(B_{j,i}) \leq \mu \otimes \nu(E_j) + \varepsilon 2^{-j}.$$

Enumerating $\{1, 2, \dots\}^2$ as $(j_n, i_n)_{n \geq 1}$, we see that $\bigcup_n (A_{j_n, i_n} \times B_{j_n, i_n})$ covers $\bigcup_j E_j$, hence

$$\mu \otimes \nu \left(\bigcup_{j=1}^{\infty} E_j \right) \leq \sum_{n=1}^{\infty} \mu(A_{j_n, i_n}) \nu(B_{j_n, i_n}) \leq \varepsilon + \sum_{j=1}^{\infty} \mu \otimes \nu(E_j) \xrightarrow{\varepsilon \rightarrow 0} \sum_{j=1}^{\infty} \mu \otimes \nu(E_j).$$

Let $\mathcal{E}$ denote the collection of all sets $E \subseteq X \times Y$ such that: for any $y \in Y$, the function $x \mapsto \mathbf{1}_E(x, y)$ is in $\mathcal{L}_\mu^+$; and the function $y \mapsto \int_X \mathbf{1}_E(x, y) d\mu(x)$ is in $\mathcal{L}_\nu^+$. For $E \in \mathcal{E}$, we can define

$$\rho(E) = \int_Y \left(\int_X \mathbf{1}_E(x, y) d\mu(x) \right) d\nu(y).$$

The set function ρ is nonnegative and nondecreasing on $\mathcal{E}$. Moreover, Theorem 7 ensures that the collection $\mathcal{E}$ is closed by taking countable unions of disjoint sets, that ρ is countably additive on $\mathcal{E}$, and that $\rho(\bigcup_i E_i) \leq \sum_i \rho(E_i)$ for any sequence of sets $E_i \in \mathcal{E}$ such that $\bigcup_i E_i \in \mathcal{E}$.

In addition, let $\mathcal{P}$ denote the collection of all product sets $A \times B$, where $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$. Note that $\mathcal{P} \subseteq \mathcal{E}$ and $\rho(A \times B) = \mu(A)\nu(B)$ for any $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$. We then define $\mathcal{P}_\sigma$ as the collection of all countable unions of sets in $\mathcal{P}$. Any set in $\mathcal{P}_\sigma$ can be written as a countable union of *disjoint* sets in $\mathcal{P}$. This follows from the observation that for any sets $A_1 \times B_1$ and $A_2 \times B_2$ in $\mathcal{P}$, we have

$$\begin{cases} (A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2) \\ (A_1 \times B_1) \setminus (A_2 \times B_2) = ((A_1 \setminus A_2) \times B_1) \sqcup ((A_1 \cap A_2) \times (B_1 \setminus B_2)), \end{cases} \quad (6)$$

where the sets in both right-hand side belong to $\mathcal{P}$. We can thus proceed as in (1) to transform the countable union of sets in $\mathcal{P}_\sigma$ into a countable union of disjoint sets in $\mathcal{P}$. We deduce that $\mathcal{P}_\sigma \subseteq \mathcal{E}$. On top of that, we define $\mathcal{P}_{\sigma\delta}$ as the collection of all countable intersections of sets in $\mathcal{P}_\sigma$.

Let us now prove that for any set $E \subseteq X \times Y$,

$$\mu \otimes \nu(E) = \inf\{\rho(P), P \in \mathcal{P}_\sigma \text{ with } P \supseteq E\}. \quad (7)$$

For any covering $E \subseteq \bigcup_i (A_i \times B_i)$, where $A_i \in \mathcal{M}_\mu$ and $B_i \in \mathcal{M}_\nu$, note that $A_i \times B_i \in \mathcal{P} \subseteq \mathcal{E}$ and $P = \bigcup_i (A_i \times B_i) \in \mathcal{P}_\sigma \subseteq \mathcal{E}$, so that $\rho(P) \leq \sum_i \rho(A_i \times B_i) = \sum_i \mu(A_i)\nu(B_i)$. Thus, $\mu \otimes \nu(E)$ is bounded below by the above infimum. Conversely, any set $P \in \mathcal{P}_\sigma$ such that $P \supseteq E$ can be written as a countable union of disjoint sets $A_i \times B_i \in \mathcal{P} \subseteq \mathcal{E}$. Then, $\rho(P) = \sum_i \rho(A_i \times B_i) = \sum_i \mu(A_i)\nu(B_i) \geq \mu \otimes \nu(E)$, because the sets $A_i \times B_i$ cover E . We thus obtain (7).

We are in position to establish the announced properties. Let $A \in \mathcal{M}_\mu$ and $B \in \mathcal{M}_\nu$. The set $A \times B$ is in $\mathcal{P}$, and satisfies $\mu \otimes \nu(A \times B) \leq \mu(A)\nu(B) = \rho(A \times B) \leq \rho(P)$ for any $P \in \mathcal{P}_\sigma$ with $P \supseteq A \times B$. It follows from (7) that $\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$. To show that $A \times B \in \mathcal{M}_{\mu \otimes \nu}$, let us consider a set $E \subseteq X \times Y$, and a set $P \in \mathcal{P}_\sigma$ with $P \supseteq E$. By (6), the sets $P \cap (A \times B)$ and $P \setminus (A \times B)$ are in $\mathcal{P}_\sigma$. By (7), we get

$$\mu \otimes \nu(E \cap (A \times B)) + \mu \otimes \nu(E \setminus (A \times B)) \leq \rho(P \cap (A \times B)) + \rho(P \setminus (A \times B)) = \rho(P),$$

and we conclude by taking the infimum over all sets $P \in \mathcal{P}_\sigma$ with $P \supseteq E$ in the right-hand side.

Finally, let us consider a set $E \subseteq X \times Y$. It is clear that $\mu \otimes \nu(E)$ is bounded above by the infimum in the statement. To prove the reverse inequality, and that the infimum is attained, it suffices to exhibit a set $F \in \mathcal{M}_{\mu \otimes \nu}$ such that $F \supseteq E$ and $\mu \otimes \nu(F) = \mu \otimes \nu(E)$. If $\mu \otimes \nu(E) = \infty$, we can take $F = X \times Y \in \mathcal{M}_{\mu \otimes \nu}$, since $\mu \otimes \nu(F) \geq \mu \otimes \nu(E) = \infty$. Otherwise, (7) yields a sequence of sets $P_j \in \mathcal{P}_\sigma$ such that $P_j \supseteq E$ and $\rho(P_j) \leq \mu \otimes \nu(E) + 1/j$, and we take $F = \bigcap_j P_j$. Indeed,

the set F contains E , belongs to $\mathcal{P}_{\sigma\delta}$, and thus to $\mathcal{M}_{\mu\otimes\nu}$, because this σ -algebra contains $\mathcal{P}$. In addition, (7) ensures that

$$\mu \otimes \nu(E) \leq \mu \otimes \nu(F) \leq \rho(P_j) \leq \mu \otimes \nu(E) + 1/j \xrightarrow{j \rightarrow \infty} \mu \otimes \nu(E).$$

The statement is thus proved. In the finite-measure case just considered, let us show additionally that $F \in \mathcal{E}$ and $\mu \otimes \nu(F) = \rho(F)$. Put $G_j = P_1 \cap \cdots \cap P_j$ and $h_j(y) = \mu((G_j)_y)$. Since $\mathcal{P}_\sigma$ is closed under finite intersections, $G_j \in \mathcal{P}_\sigma \subseteq \mathcal{E}$. Thus h_j is measurable, $h_j \leq h_1$, and $\int_Y h_1 d\nu = \rho(P_1) < \infty$. The measurable set $N = \{y \in Y \mid h_1(y) = \infty\}$ is therefore ν -null. For every $y \notin N$, the functions $x \mapsto \mathbf{1}_{G_j}(x, y)$ decrease to $\mathbf{1}_F(x, y)$ and are dominated by the integrable function $x \mapsto \mathbf{1}_{P_1}(x, y)$. Theorem 13 gives, for $y \notin N$,

$$h_j(y) \xrightarrow{j \rightarrow \infty} h(y) := \mu(F_y).$$

Every section F_y is measurable, being a countable intersection of measurable sections. The function h agrees outside N with the measurable function $\lim_{j \rightarrow \infty} h_j$; since every subset of N is ν -measurable by Theorem 1, the function h is measurable as well. Hence $F \in \mathcal{E}$. Replacing h_j , h_1 and h by zero on N gives finite-valued measurable functions dominated by an integrable function on Y . A second application of dominated convergence yields

$$\rho(F) = \int_Y h d\nu = \lim_{j \rightarrow \infty} \int_Y h_j d\nu = \lim_{j \rightarrow \infty} \rho(P_1 \cap \cdots \cap P_j).$$

Finally, using (7), we obtain

$$\mu \otimes \nu(F) \leq \rho(P_1 \cap \cdots \cap P_j) \leq \rho(P_j) \leq \mu \otimes \nu(E) + 1/j \leq \mu \otimes \nu(F) + 1/j,$$

and we conclude by letting $j \rightarrow \infty$ that $\mu \otimes \nu(F) = \rho(F)$. $\square$

3.2 Fubini's theorem

An important result about product measures is Fubini's theorem, which states that a product measure can be recovered by integrating the measures of the cross sections of the product set. This result is valid under the σ -finiteness assumption of Definition 6.

It follows from Remark 1 that any function in $\mathcal{L}_\mu^1$ is σ -finite with respect to μ . We shall split the statement and the proof of Fubini's theorem into two parts, one for sets and one for functions. We begin with Fubini's theorem for sets.

Theorem 16 (Fubini's theorem for sets). *Let X and Y be sets, and let μ and ν be measures on X and Y , respectively. Then, for any subset E of $X \times Y$ that is σ -finite with respect to $\mu \otimes \nu$, the following properties hold:*

- (i). $E_y \in \mathcal{M}_\mu$ for ν -a.e. $y \in Y$, and $[y \mapsto \mu(E_y)] \in \mathcal{L}_\nu^+$;
- (ii). $E_x \in \mathcal{M}_\nu$ for μ -a.e. $x \in X$, and $[x \mapsto \nu(E_x)] \in \mathcal{L}_\mu^+$;
- (iii). $\mu \otimes \nu(E) = \int_Y \mu(E_y) d\nu(y) = \int_X \nu(E_x) d\mu(x)$.

Proof. Let us consider a set $E \in \mathcal{M}_{\mu\otimes\nu}$ such that $\mu \otimes \nu(E) < \infty$. As established in the proof of Proposition 15, there exists a set $F \in \mathcal{M}_{\mu\otimes\nu} \cap \mathcal{E}$ such that $F \supseteq E$ and $\mu \otimes \nu(F) = \rho(F) = \mu \otimes \nu(E)$. Thus,

$$\mu \otimes \nu(F \setminus E) = \mu \otimes \nu(F) - \mu \otimes \nu(E) = 0.$$

Likewise, there exists a set $N \in \mathcal{M}_{\mu \otimes \nu} \cap \mathcal{E}$ such that $N \supseteq F \setminus E$ and $\mu \otimes \nu(N) = \rho(N) = 0$. Observe that $\rho(N) = \int_Y \mu(N_y) d\nu(y)$. Proposition 9 ensures that $\mu(N_y) = 0$ for ν -a.e. $y \in Y$, which implies that $\mu((F \setminus E)_y) = 0$. Since negligible sets are measurable, we have $(F \setminus E)_y \in \mathcal{M}_\mu$. Moreover, since $F \in \mathcal{E}$, we have $F_y \in \mathcal{M}_\mu$ for all $y \in Y$. We deduce that $E_y = F_y \setminus (F \setminus E)_y \in \mathcal{M}_\mu$ for ν -a.e. $y \in Y$. In addition, the function $y \mapsto \mu(F_y)$ is in $\mathcal{L}_\nu^+$, and coincides ν -a.e. with the function $y \mapsto \mu(E_y)$. Hence, the latter function is in $\mathcal{L}_\nu^+$ as well. Finally, Proposition 12 implies that

$$\mu \otimes \nu(E) = \rho(F) = \int_Y \mu(F_y) d\nu(y) = \int_Y \mu(E_y) d\nu(y).$$

To obtain the whole statement, it suffices to observe that any σ -finite set can be written as a countable union of disjoint measurable sets of finite measure, and that we can exchange the roles of μ and ν above and in the proof of Proposition 15. $\square$

To transfer Fubini's theorem from sets to functions, we establish the following result.

Lemma 17 (Decomposition of nonnegative measurable functions). *Let X be a set and let μ be a measure on X . For any function $f \in \mathcal{L}_\mu^+$, there exists a sequence of sets $A_j \in \mathcal{M}_\mu$ such that*

$$f = \sum_{j=1}^{\infty} \frac{\mathbf{1}_{A_j}}{j}.$$

Proof. We define the sets A_j inductively. To be specific, we set $A_1 = \{x \in X \mid f(x) \geq 1\}$, and for $j \geq 2$,

$$A_j = \left\{ x \in X \mid f(x) \geq \frac{1}{j} + \sum_{k=1}^{j-1} \frac{\mathbf{1}_{A_k}(x)}{k} \right\}.$$

It is clear that $A_j \in \mathcal{M}_\mu$. Moreover, $f \geq \mathbf{1}_{A_1}$, and for all $J \geq 2$, the following implication holds:

$$f \geq \sum_{j=1}^{J-1} \frac{\mathbf{1}_{A_j}}{j} \implies f = \mathbf{1}_{A_J} f + \mathbf{1}_{X \setminus A_J} f \geq \mathbf{1}_{A_J} \left(\frac{1}{J} + \sum_{j=1}^{J-1} \frac{\mathbf{1}_{A_j}}{j} \right) + \mathbf{1}_{X \setminus A_J} \sum_{j=1}^{J-1} \frac{\mathbf{1}_{A_j}}{j} = \sum_{j=1}^J \frac{\mathbf{1}_{A_j}}{j}.$$

By induction, we conclude that $f \geq \sum_{j=1}^{\infty} \mathbf{1}_{A_j}/j$. Finally, if $f(x) = \infty$, then x belongs to all the sets A_j , so that $\sum_{j=1}^{\infty} \mathbf{1}_{A_j}(x)/j = \infty$. Otherwise, if $f(x) < \infty$, then $x \notin A_{j_n}$ for infinitely many j_n , and thus

$$\sum_{j=1}^{\infty} \frac{\mathbf{1}_{A_j}(x)}{j} \leq f(x) < \frac{1}{j_n} + \sum_{k=1}^{j_n-1} \frac{\mathbf{1}_{A_k}(x)}{k}.$$

We conclude by letting $n \rightarrow \infty$. $\square$

We can finally state and prove Fubini's theorem for functions.

Theorem 18 (Fubini's theorem for functions). *Let X and Y be sets, and let μ and ν be measures on X and Y , respectively. For any function $f \in \mathcal{L}_{\mu \otimes \nu}^+$ that is σ -finite with respect to $\mu \otimes \nu$, the following properties hold:*

- (i). $x \mapsto f(x, y)$ is μ -measurable for ν -a.e. $y \in Y$, and $y \mapsto \int_X f(x, y) d\mu(x)$ is in $\mathcal{L}_\nu^+$;
- (ii). $y \mapsto f(x, y)$ is ν -measurable for μ -a.e. $x \in X$, and $x \mapsto \int_Y f(x, y) d\nu(y)$ is in $\mathcal{L}_\mu^+$;
- (iii). $\int_{X \times Y} f d(\mu \otimes \nu) = \int_Y \left(\int_X f(x, y) d\mu(x) \right) d\nu(y) = \int_X \left(\int_Y f(x, y) d\nu(y) \right) d\mu(x)$.

In these formulas, the inner integral is defined to be zero on the exceptional null set where the section is not measurable. For any function $f \in \mathcal{L}_{\mu \otimes \nu}^1$, almost every section is integrable; defining the inner integral to be zero on the exceptional null set where the section is not integrable, the same properties hold with $\mathcal{L}_\nu^+$ and $\mathcal{L}_\mu^+$ replaced by $\mathcal{L}_\nu^1$ and $\mathcal{L}_\mu^1$.

Proof. The statement is exactly that of Theorem 16 when f is the indicator function of a set E that is σ -finite with respect to $\mu \otimes \nu$. For a general function $f \in \mathcal{L}_{\mu \otimes \nu}^+$ that is σ -finite with respect to $\mu \otimes \nu$, we can apply Lemma 17 to write $f = \sum_{j=1}^\infty \mathbf{1}_{A_j}/j$, where the sets A_j are in $\mathcal{M}_{\mu \otimes \nu}$. These sets are σ -finite with respect to $\mu \otimes \nu$, because their union is the set of points where f is nonzero, which is σ -finite with respect to $\mu \otimes \nu$. We conclude thanks to Theorems 7 and 16.

If $f \in \mathcal{L}_{\mu \otimes \nu}^1$, we begin by observing that f is σ -finite with respect to $\mu \otimes \nu$, in view of Remark 1. The function f can be decomposed as $f = ((\Re f)^+ - (\Re f)^-) + i((\Im f)^+ - (\Im f)^-)$, where $(\Re f)^+$, $(\Re f)^-$, $(\Im f)^+$ and $(\Im f)^-$ are in $\mathcal{L}_{\mu \otimes \nu}^+$ and σ -finite with respect to $\mu \otimes \nu$. The first part of the proof ensures that the properties hold for these functions. We deduce the desired properties for f . In particular,

$$\int_Y \left(\int_X (\Re f)^+(x, y) d\mu(x) \right) d\nu(y) = \int_{X \times Y} (\Re f)^+ d(\mu \otimes \nu) \leq \int_{X \times Y} |f| d(\mu \otimes \nu) < \infty,$$

so that the function $y \mapsto \int_X (\Re f)^+(x, y) d\mu(x)$ is in $\mathcal{L}_\nu^1$. The same is true for the other components of f , so we can replace $\mathcal{L}_\nu^+$ and $\mathcal{L}_\mu^+$ by $\mathcal{L}_\nu^1$ and $\mathcal{L}_\mu^1$ in the statement. $\square$

4 Borel and Radon measures

4.1 Definitions

Let us now explore the properties of measures that are specific to the context of topological spaces. We shall work on *Hausdorff spaces*, where two distinct points have disjoint neighborhoods. The collections of all compact sets and all open sets in a Hausdorff space X are denoted by $\mathcal{K}(X)$ and $\mathcal{G}(X)$, respectively. Let us introduce several properties that combine measures and topology.

Definition 9 (Borel measure, regularity, local finiteness). Let X be a Hausdorff space and let μ be a measure on X .

- (i). μ is a *Borel measure* if all Borel sets in X are μ -measurable, that is, $\mathcal{B}(X) \subseteq \mathcal{M}_\mu$;
- (ii). μ is *Borel-regular* if μ is Borel and for any set $A \subseteq X$, the following equality holds and the infimum is attained: $\mu(A) = \inf\{\mu(B), B \in \mathcal{B}(X) \text{ with } B \supseteq A\}$;
- (iii). μ is *inner regular* on a set $A \subseteq X$ if $\mu(A) = \sup\{\mu(K), K \in \mathcal{K}(X) \text{ with } K \subseteq A\}$;
- (iv). μ is *outer regular* on a set $A \subseteq X$ if $\mu(A) = \inf\{\mu(U), U \in \mathcal{G}(X) \text{ with } U \supseteq A\}$;
- (v). μ is *locally finite* if any point $x \in X$ has a neighborhood V_x such that $\mu(V_x) < \infty$.

For any countably additive set function $\nu: \mathcal{B}(X) \rightarrow [0, \infty]$ such that $\nu(\emptyset) = 0$, the extension μ given by Theorem 2 is a Borel-regular measure on the space X . More precisely, μ is the unique Borel-regular measure on X that coincides with ν on all Borel sets.

Given any Borel measure, one can easily construct a Borel-regular measure that agrees with the original measure on all Borel sets. This is the purpose of the following result.

Proposition 19 (Regularization of a Borel measure). Let X be a Hausdorff space and let μ be a Borel measure on X . Then, the formula

$$\tilde{\mu}(A) = \inf\{\mu(B), B \in \mathcal{B}(X) \text{ with } B \supseteq A\},$$

where the infimum is attained, defines the unique Borel-regular measure $\tilde{\mu}$ on X that satisfies $\tilde{\mu}(A) = \mu(A)$ for all Borel sets $A \in \mathcal{B}(X)$.

Proof. Clearly, $\tilde{\mu}(A) = \mu(A)$ for every Borel set A . In particular, $\tilde{\mu}(\emptyset) = 0$. Moreover, if $A_1 \subseteq A_2 \subseteq X$, then every Borel set that contains A_2 also contains A_1 , so that $\tilde{\mu}(A_1) \leq \tilde{\mu}(A_2)$.

To prove that the infimum is attained, let us consider a set $A \subseteq X$ such that $\tilde{\mu}(A) < \infty$, and observe that for every integer $n \geq 1$, there exists a Borel set $B_n \supseteq A$ such that $\mu(B_n) \leq \tilde{\mu}(A) + 1/n$. Then, the intersection $B = \bigcap_{n=1}^{\infty} B_n$ is a Borel set such that $B \supseteq A$. In addition $\mu(B) = \tilde{\mu}(A)$, since

$$\tilde{\mu}(A) \leq \mu(B) \leq \mu(B) + \mu(B_n \setminus B) = \mu(B_n) \leq \tilde{\mu}(A) + 1/n \xrightarrow{n \rightarrow \infty} \tilde{\mu}(A).$$

If $\tilde{\mu}(A) = \infty$, the infimum is attained at the Borel set X , because $\mu(X) = \tilde{\mu}(X) \geq \tilde{\mu}(A) = \infty$.

Now, for every sequence $(A_i)_{i \geq 1}$ of subsets of X , there exists a sequence $(B_i)_{i \geq 1}$ of Borel sets such that $B_i \supseteq A_i$ and $\mu(B_i) = \tilde{\mu}(A_i)$. The union $\bigcup_{i=1}^{\infty} B_i$ is a Borel set that contains the union $\bigcup_{i=1}^{\infty} A_i$. Thus,

$$\begin{aligned} \tilde{\mu}\left(\bigcup_{i=1}^{\infty} A_i\right) &\leq \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \mu\left(\bigsqcup_{i=1}^{\infty} B_i \setminus (B_1 \cup \cdots \cup B_{i-1})\right) = \sum_{i=1}^{\infty} \mu(B_i \setminus (B_1 \cup \cdots \cup B_{i-1})) \\ &\leq \sum_{i=1}^{\infty} (\mu(B_i \setminus (B_1 \cup \cdots \cup B_{i-1})) + \mu(B_i \cap (B_1 \cup \cdots \cup B_{i-1}))) = \sum_{i=1}^{\infty} \mu(B_i) = \sum_{i=1}^{\infty} \tilde{\mu}(A_i). \end{aligned} \tag{8}$$

To prove that all Borel sets are $\tilde{\mu}$ -measurable, let us consider a Borel set $A \subseteq X$, a set $E \subseteq X$, and a Borel set $B \supseteq E$ such that $\mu(B) = \tilde{\mu}(E)$. The Borel sets $B \cap A$ and $B \setminus A$ are disjoint, so that

$$\tilde{\mu}(E) \leq \tilde{\mu}(E \cap A) + \tilde{\mu}(E \setminus A) \leq \mu(B \cap A) + \mu(B \setminus A) = \mu(B) = \tilde{\mu}(E).$$

We have equality throughout, so we can conclude that the Borel set A is $\tilde{\mu}$ -measurable. We thus have shown that $\tilde{\mu}$ is a Borel-regular measure that coincides with μ on all Borel sets.

It remains to prove uniqueness. Let ν be another Borel-regular measure on X such that $\nu(A) = \mu(A)$ for all Borel sets $A \subseteq X$. For every $A \subseteq X$, Borel regularity of ν gives

$$\nu(A) = \inf\{\nu(B), B \in \mathcal{B}(X) \text{ with } B \supseteq A\} = \inf\{\mu(B), B \in \mathcal{B}(X) \text{ with } B \supseteq A\} = \tilde{\mu}(A).$$

We conclude that the measures ν and $\tilde{\mu}$ coincide on all sets. $\square$

Locally finite measures are finite on compact sets. In *locally compact* spaces, measures that are finite on compact sets are locally finite. These spaces are defined by the condition that every point has a compact neighborhood. The Euclidean space $\mathbb{R}^d$ is locally compact and Hausdorff, because all closed balls are compact. Metric spaces are Hausdorff, but not always locally compact.

The inner and the outer regularity properties characterize the approximation of a measure by compact and open sets, respectively. A first consequence of the inner regularity property is that a measure is null outside of its *support*, which we now define.

Definition 10 (Support of a measure). Let X be a Hausdorff space and let μ be a measure on X . The *support* of μ is the complement of the union of all μ -null open subsets of X , specifically, $\text{supp } \mu = X \setminus \bigcup\{U \in \mathcal{G}(X) \mid \mu(U) = 0\}$.

From this definition, it follows immediately that the support of a measure is closed. While its complement is a union of null sets, this union may be uncountable, and consequently need not be null itself, unless additional assumptions are made on the measure or the space. The following result shows that the complement of the support is null if the measure is inner regular on open sets.

Proposition 20 (Null complement of the support). *For any measure μ on a Hausdorff space X that is inner regular on all open sets, $\mu(X \setminus \text{supp } \mu) = 0$.*

Proof. Any compact set K contained in the open set $X \setminus \text{supp } \mu$ is covered by a finite union of μ -null open sets $U_1, \dots, U_n$. By the countable subadditivity of μ , we deduce that $\mu(K) \leq \mu(U_1) + \dots + \mu(U_n) = 0$. The result follows directly from the inner regularity of μ on all open sets, hence on $X \setminus \text{supp } \mu$. $\square$

Radon measures, defined below, are crucial for studying the positive linear functionals on the space of compactly supported continuous functions. The next definition establishes our framework. The closure of a set E is denoted by $\bar{E}$.

Definition 11 (Compactly supported function, positive linear functional, Radon measure). Let X be a Hausdorff space.

- (i). The *support* of a function $f: X \rightarrow \mathbb{C}$ is the closed set $\text{supp } f = \overline{\{x \in X \mid f(x) \neq 0\}}$.
- (ii). The set of continuous functions $X \rightarrow \mathbb{C}$ with compact support is a vector space over $\mathbb{C}$, denoted by $C_c(X)$.
- (iii). A linear functional I on $C_c(X)$ is *positive* if $I(f) \geq 0$ whenever $f \geq 0$.
- (iv). A *Radon measure* on X is a locally finite Borel measure on X that is inner regular on all open subsets of X , and outer regular on all subsets of X .

According to Definitions 9 and 11, a Radon measure is necessarily Borel-regular and outer regular on Borel sets. Conversely, these two properties ensure that a measure is outer regular on all sets.

4.2 Urysohn's lemma and continuity of positive linear functionals

Definition 11 does not mention any continuity requirement on the linear functional I . However, the positivity assumption entails a strong continuity property. To see this, we need a version of *Urysohn's lemma* on the existence of continuous functions that can separate two disjoint closed sets. Such a result requires additional assumptions on the space. For this reason, we work from now on in a locally compact Hausdorff space X . For any compact set $K \in \mathcal{K}(X)$, any open set $U \in \mathcal{G}(X)$, and any function $f \in C_c(X)$, we write $K < f < U$ to indicate that

$$0 \leq f \leq 1, \quad f = 1 \text{ on } K, \quad \text{supp } f \subseteq U.$$

The second assertion below is Urysohn's lemma for locally compact Hausdorff spaces.

Lemma 21 (Compact closed separation, Urysohn's lemma, partition of unity). *Let X be a locally compact Hausdorff space, and let $K \in \mathcal{K}(X)$ and $U, U_1, \dots, U_n \in \mathcal{G}(X)$ such that $K \subseteq U$ and $K \subseteq U_1 \cup \dots \cup U_n$. Then, the following properties hold:*

- (i). *There exists a relatively compact open set $V \in \mathcal{G}(X)$ such that $K \subseteq V \subseteq \overline{V} \subseteq U$;*
- (ii). *There exists a function $f \in C_c(X)$ such that $K < f < U$;*
- (iii). *There exist functions $f_1, \dots, f_n \in C_c(X)$ with $\emptyset < f_j < U_j$ for all j , and $f_1 + \dots + f_n = 1$ on K .*

Proof. To prove (i), let us consider, for any $x \in K$, an open neighborhood V_x of x with compact closure. By compactness, K is covered by the union W of finitely many sets $V_{x_1}, \dots, V_{x_m}$. The closure of W is contained in the union of the closures of the sets V_{x_j} , which are compact. It follows that W has compact closure. If $U = X$, we can thus choose $V = W$.

Let us assume otherwise that $F = X \setminus U \neq \emptyset$. For any $f \in F$, there exists an open set W_f such that $K \subseteq W_f$ and $f \notin \overline{W_f}$. Indeed, for any $y \in K$, the Hausdorff property ensures that there exist two disjoint open neighborhoods A_y and B_y of f and y , respectively. By compactness, K is covered by a finite union $W_f = B_{y_1} \cup \dots \cup B_{y_m}$, while f is in $A_{y_1} \cap \dots \cap A_{y_m}$, an open set disjoint from W_f , and thus from $\overline{W_f}$. It follows that $\bigcap_{f \in F} (F \cap \overline{W} \cap \overline{W_f})$ is an empty intersection of compact sets, from which we can extract finitely many compact sets $F \cap \overline{W} \cap \overline{W_{f_1}}, \dots, F \cap \overline{W} \cap \overline{W_{f_p}}$ with an empty intersection. We can finally take $V = W \cap W_{f_1} \cap \dots \cap W_{f_p}$. To extract an empty finite intersection from an arbitrary empty intersection of compact sets, observe that if $K_0 \cap \bigcap_{\alpha} K_{\alpha} = \emptyset$, then $K_0 \subseteq (X \setminus K_{\alpha_1}) \cup \dots \cup (X \setminus K_{\alpha_n})$.

To prove (ii), let $(r_j)_{j \geq 1}$ denote an enumeration of the rational numbers in $[0, 1]$, with $r_1 = 0$ and $r_2 = 1$. Then, (i) yields two relatively compact open sets V_0 and V_1 such that

$$K \subseteq V_1 \subseteq \overline{V_1} \subseteq V_0 \subseteq \overline{V_0} \subseteq U.$$

For $J \geq 2$, assume that we have relatively compact open sets $V_{r_1}, \dots, V_{r_J}$ such that for all $j, j' \in \{1, \dots, J\}$,

$$K \subseteq V_{r_j} \subseteq \overline{V_{r_j}} \subseteq U \quad \text{and} \quad [r_j < r_{j'} \implies \overline{V_{r_{j'}}} \subseteq V_{r_j}]. \quad (9)$$

We choose in $\{1, \dots, J\}$ two indices j and j' such that $r_j < r_{j+1} < r_{j'}$ and $r_{j'} - r_j$ is minimal. Then, (i) yields a relatively compact open set $V_{r_{j+1}}$ such that

$$\overline{V_{r_{j'}}} \subseteq V_{r_{j+1}} \subseteq \overline{V_{r_{j+1}}} \subseteq V_{r_j}.$$

By induction, we obtain a sequence of relatively compact open sets V_{r_j} such that (9) holds for all $j, j' \geq 1$. Let us now define the functions

$$\begin{cases} f = \sup_{r \in [0,1] \cap \mathbb{Q}} f_r \\ g = \inf_{s \in [0,1] \cap \mathbb{Q}} g_s \end{cases} \quad \text{with} \quad \begin{cases} f_r = r \mathbf{1}_{V_r} \\ g_s = s + (1-s) \mathbf{1}_{\overline{V_s}}. \end{cases}$$

The function f is lower semi-continuous, meaning that the set $\{x \in X \mid f(x) > \alpha\}$ is open for any α . Likewise, g is upper semi-continuous, meaning that $\{x \in X \mid g(x) < \alpha\}$ is open for any α . Moreover, $0 \leq f \leq 1$, with $f = 1$ on $V_1 \supseteq K$ and $\text{supp } f \subseteq \overline{V_0} \subseteq U$. To show that f is continuous, it suffices to show that $f = g$. To proceed, note that $f_r(x) > g_s(x)$ implies that $r > s$ and $x \in V_r \setminus \overline{V_s}$. This is impossible since $V_r \subseteq V_s$ when $r > s$. We deduce that $f_r \leq g_s$ for all r and s , so that $f \leq g$. For the reverse inequality, note that $f(x) < g(x)$ implies that $f(x) < r < s < g(x)$ for some rational numbers r and s , and thus that $x \in \overline{V_s} \setminus V_r$, which is again impossible. Therefore, $f = g$ and the proof of (ii) is complete.

Let us prove (iii). Any point $x \in K$ belongs to some open set U_i , and (i) ensures that there exists a relatively compact open set V_x such that $\{x\} \subseteq V_x \subseteq \overline{V_x} \subseteq U_i$. By compactness, K is covered by the union of finitely many sets $V_{x_1}, \dots, V_{x_m}$. For $i \in \{1, \dots, n\}$, let L_i be the union of the closures $\overline{V_{x_j}}$, with $1 \leq j \leq m$, that are contained in U_i . Then, (ii) yields a function $g_i \in C_c(X)$ such that $L_i < g_i < U_i$. The functions defined by

$$f_1 = g_1, \quad f_2 = (1 - g_1)g_2, \quad \dots \quad f_n = (1 - g_1) \cdots (1 - g_{n-1})g_n$$

are in $C_c(X)$, and satisfy $\emptyset < f_i < U_i$. Moreover,

$$\begin{aligned} f_1 + \dots + f_i &= 1 - (1 - g_1) \cdots (1 - g_i) \\ \implies f_1 + \dots + f_{i+1} &= 1 - (1 - g_1) \cdots (1 - g_i) + (1 - g_1) \cdots (1 - g_i)g_{i+1} \\ &= 1 - (1 - g_1) \cdots (1 - g_{i+1}). \end{aligned}$$

Therefore, by induction, we have $f_1 + \dots + f_n = 1 - (1 - g_1) \cdots (1 - g_n)$. Since K is covered by the union of the sets L_i , the product in the right-hand side is zero on K . We conclude that $f_1 + \dots + f_n = 1$ on K . $\square$

We can now show that a positive linear functional on $C_c(X)$ satisfies a continuity property. This vector space is equipped with the uniform norm, defined for any function $f \in C_c(X)$ by

$$\|f\|_\infty = \sup_{x \in X} |f(x)|.$$

We then have the following result.

Proposition 22 (Continuity of positive linear functionals). *Let X be a locally compact Hausdorff space, and let I be a positive linear functional on $C_c(X)$. Then, for any compact set $K \in \mathcal{K}(X)$, there exists a constant $C_K > 0$ such that for all $f \in C_c(X)$,*

$$\text{supp } f \subseteq K \implies |I(f)| \leq C_K \|f\|_\infty.$$

Proof. First consider real-valued functions $f \in C_c(X)$. Positivity implies that $I(f)$ is real, since $f = f^+ - f^-$. Urysohn's lemma, that is, Lemma 21(ii) yields a function $\phi \in C_c(X)$ such that $K < \phi < X$. If $\text{supp } f \subseteq K$, then $|f| \leq \|f\|_\infty \phi$. It follows that $\|f\|_\infty \phi \pm f \geq 0$, and thus $I(\|f\|_\infty \phi \pm f) \geq 0$. Hence, $|I(f)| \leq I(\phi) \|f\|_\infty$.

For complex-valued f with $I(f) \neq 0$, choose $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$ and $\lambda I(f) = |I(f)|$. Complex linearity and the preceding real-valued case give

$$|I(f)| = \Re I(\lambda f) = I(\Re(\lambda f)) \leq I(\phi) \|\Re(\lambda f)\|_\infty \leq I(\phi) \|f\|_\infty.$$

The bound is immediate if $I(f) = 0$. Taking $C_K = 1 + I(\phi) > 0$ proves the statement. $\square$

4.3 The Riesz representation theorem

Let μ denote a locally finite Borel measure on a locally compact Hausdorff space X . Any function $f \in C_c(X)$ is Borel-measurable, and thus μ -measurable. Moreover,

$$\int_X |f| d\mu \leq \|f\|_\infty \mu(\text{supp } f) < \infty. \quad (10)$$

It follows that $C_c(X)$ is contained in L^1_μ . In addition, the mapping $f \mapsto \int_X f d\mu$ is a linear functional on $C_c(X)$, by Proposition 10. It is easy to check that this functional is positive. The main result of this section is that all positive linear functionals on $C_c(X)$ are of this form, and can thus be represented by a locally finite Borel measure μ on X . In fact, μ can be chosen to be a Radon measure, which then ensures its uniqueness. This result is the *Riesz representation theorem*.

Theorem 23 (Riesz representation theorem). *Let X be a locally compact Hausdorff space. For any positive linear functional I on $C_c(X)$, there exists a unique Radon measure μ on X such that for all $f \in C_c(X)$,*

$$I(f) = \int_X f d\mu.$$

Moreover, the measure μ satisfies the following properties:

- (i). $\mu(U) = \sup\{I(f), f \in C_c(X) \text{ with } 0 < f < U\}$ for all open sets $U \in \mathcal{G}(X)$;
- (ii). $\mu(K) = \inf\{I(f), f \in C_c(X) \text{ with } K < f < X\}$ for all compact sets $K \in \mathcal{K}(X)$.

Proof. We begin by proving uniqueness. Let μ denote a Radon measure such that $I(f) = \int_X f d\mu$ for all $f \in C_c(X)$. Then, for any open set $U \in \mathcal{G}(X)$ and any function $f \in C_c(X)$ with $0 < f < U$,

$$I(f) = \int_X f d\mu \leq \int_X \mathbf{1}_U d\mu = \mu(U).$$

Moreover, for any compact set $K \in \mathcal{K}(X)$ such that $K \subseteq U$, Urysohn's lemma, that is, Lemma 21(ii) yields a function $g \in C_c(X)$ such that $K < g < U$. We have

$$I(g) = \int_X g d\mu \geq \int_X \mathbf{1}_K d\mu = \mu(K).$$

Since μ is inner regular on open sets, the supremum formula (i) holds. The Radon measure μ is thus determined by I on open sets. By outer regularity, μ is actually determined by I on all sets.

To prove existence, we begin by defining, for any open set $U \in \mathcal{G}(X)$,

$$\mu(U) = \sup\{I(f), f \in C_c(X) \text{ with } 0 < f < U\}. \quad (11)$$

We then extend the definition of μ to any subset A of X by setting

$$\mu(A) = \inf\{\mu(U), U \in \mathcal{G}(X) \text{ with } U \supseteq A\}. \quad (12)$$

We thus obtain a measure on X . Clearly, $\mu(\emptyset) = I(0) = 0$, and $\mu(A) \leq \mu(B)$ for any $A \subseteq B$ because any open set that contains B also contains A . To prove the countable subadditivity of μ , let us consider a sequence of sets A_i in X . For any $\varepsilon > 0$ and $i \geq 1$, there exists an open set $U_i \in \mathcal{G}(X)$ such that $A_i \subseteq U_i$ and $\mu(U_i) \leq \mu(A_i) + \varepsilon 2^{-i}$. Let us consider a function $f \in C_c(X)$ such that $0 < f < U = \bigcup_i U_i$. The support K of f is compact, and thus contained in a finite

union $U_{i_1} \cup \cdots \cup U_{i_n}$. By Lemma 21(iii), choose $f_1, \dots, f_n \in C_c(X)$ such that $0 < f_j < U_{i_j}$ for all j , and $f_1 + \cdots + f_n = 1$ on K . Since $f = \sum_{j=1}^n f_j f$ and $0 < f_j f < U_{i_j}$, we obtain

$$I(f) = \sum_{j=1}^n I(f_j f) \leq \sum_{j=1}^n \mu(U_{i_j}) \leq \sum_{i=1}^{\infty} \mu(U_i) \leq \varepsilon + \sum_{i=1}^{\infty} \mu(A_i).$$

For this fixed ε and the corresponding open set U , taking the supremum over all functions $f \in C_c(X)$ such that $0 < f < U$, we obtain

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \mu(U) \leq \varepsilon + \sum_{i=1}^{\infty} \mu(A_i).$$

Letting $\varepsilon \rightarrow 0$ proves countable subadditivity, so μ is a measure on X . Moreover, (11) and (12) ensure respectively that μ satisfies the supremum formula (i) and is outer regular on all sets.

The next step is to show that any open set $U \in \mathcal{G}(X)$ is μ -measurable. Let E be a subset of X such that $\mu(E) < \infty$. For any $\varepsilon > 0$, there exists an open set $V \in \mathcal{G}(X)$ such that $E \subseteq V$ and $\mu(V) \leq \mu(E) + \varepsilon$. Since $U \cap V$ is open, there exists a function $f \in C_c(X)$ with $0 < f < U \cap V$ and $I(f) \geq \mu(U \cap V) - \varepsilon$. Likewise, there exists a function $g \in C_c(X)$ with $0 < g < (V \setminus \text{supp } f)$ and $I(g) \geq \mu(V \setminus \text{supp } f) - \varepsilon$. Note that $0 < f + g < V$, so that

$$\mu(E) + \varepsilon \geq \mu(V) \geq I(f + g) = I(f) + I(g) \geq \mu(U \cap V) + \mu(V \setminus \text{supp } f) - 2\varepsilon \geq \mu(E \cap U) + \mu(E \setminus U) - 2\varepsilon.$$

We conclude by letting $\varepsilon \rightarrow 0$ that U is μ -measurable. Hence, $\mathcal{M}_\mu$ contains all open sets in X and is a σ -algebra by Theorem 1. Thus, all Borel sets are in $\mathcal{M}_\mu$, implying that μ is a Borel measure.

Let us now show that μ also satisfies the infimum formula (ii). Let $K \in \mathcal{K}(X)$, and let $f \in C_c(X)$ such that $K < f < X$. For $0 < \varepsilon < 1$, let us consider the open set $U_\varepsilon = \{x \in X \mid f(x) > 1 - \varepsilon\} \supseteq K$ and a function $g \in C_c(X)$ such that $0 < g < U_\varepsilon$. We have $(1 - \varepsilon)g \leq f$, so that $I(g) \leq (1 - \varepsilon)^{-1}I(f)$. Taking the supremum over all the functions g , we deduce that

$$\mu(K) \leq \mu(U_\varepsilon) \leq (1 - \varepsilon)^{-1}I(f) \xrightarrow{\varepsilon \rightarrow 0} I(f).$$

Conversely, for any open set $U \supseteq K$, Urysohn's lemma, that is, Lemma 21(ii) yields a function $f \in C_c(X)$ with $K < f < U$. Consequently, $\inf\{I(g), g \in C_c(X) \text{ with } K < g < X\} \leq I(f) \leq \mu(U)$. Taking the infimum over the open sets $U \supseteq K$ and using (12) gives the reverse inequality required for the infimum formula.

We thus established the infimum formula (ii). This implies that μ is finite on compact sets, and hence locally finite, because X is locally compact. Indeed, for any compact set K , Urysohn's lemma, that is, Lemma 21(ii) yields a function $f \in C_c(X)$ such that $K < f < X$. We have $\mu(K) \leq I(f) < \infty$.

To prove the inner regularity of μ , let us consider an open set U and $\alpha < \mu(U)$. By (11), there exists a function $f \in C_c(X)$ such that $0 < f < U$ and $I(f) > \alpha$. Let K denote the support of f . For any function $g \in C_c(X)$ such that $K < g < X$, we have $g \geq f$, so that $I(g) \geq I(f) > \alpha$. By the infimum formula (ii), we deduce that $\mu(K) > \alpha$. This proves that μ is inner regular on open sets.

It remains to establish that $I(f) = \int_X f \, d\mu$ for all $f \in C_c(X)$. We can assume by linearity that f takes values in $[0, 1]$. Let us define $K_0 = \text{supp } f$ and, for all integers $1 \leq j \leq n$,

$$K_j = \left\{x \in X \mid f(x) \geq \frac{j}{n}\right\} \quad \text{and} \quad f_j = \min\left\{\left(f - \frac{j-1}{n}\right)^+, \frac{1}{n}\right\}.$$

Note that $\mathbf{1}_{K_j} \leq n f_j \leq \mathbf{1}_{K_{j-1}}$, so that for any open set $U \supseteq K_{j-1}$, we have $K_j < n f_j < U$. Thus,

$$\mu(K_j) \leq n \int_X f_j d\mu \leq \mu(K_{j-1}) \quad \text{and} \quad \mu(K_j) \leq n I(f_j) \leq \mu(U).$$

We used here the infimum formula (ii) and the fact that K_j is a compact set, as a closed subset of the support of f . Moreover, taking the infimum over all open sets $U \supseteq K_{j-1}$, and using the outer regularity of μ , we can replace U by K_{j-1} in the last inequality. Finally, observing that $f = f_1 + \cdots + f_n$, we get

$$\frac{1}{n} \sum_{j=1}^n \mu(K_j) \leq \int_X f d\mu \leq \frac{1}{n} \sum_{j=0}^{n-1} \mu(K_j) \quad \text{and} \quad \frac{1}{n} \sum_{j=1}^n \mu(K_j) \leq I(f) \leq \frac{1}{n} \sum_{j=0}^{n-1} \mu(K_j).$$

As a result, since the compact set K_0 has finite measure, we have

$$\left| I(f) - \int_X f d\mu \right| \leq \frac{\mu(K_0) - \mu(K_n)}{n} \leq \frac{\mu(K_0)}{n} \xrightarrow{n \rightarrow \infty} 0.$$

This finishes the proof of the Riesz representation theorem. $\square$

4.4 Regularity and approximation properties of Radon measures

We now study in more detail the regularity and approximation properties of Radon measures. We begin by strengthening the inner regularity properties of Radon measures.

A topological space is called σ -compact if it is a countable union of compact subsets.

Proposition 24 (Inner regularity on σ -finite sets). *Let X be a locally compact Hausdorff space, and let μ be a Radon measure on X . The following properties hold:*

- (i). μ is inner regular on all its σ -finite sets;
- (ii). If X is also σ -compact, then μ is inner regular on all its measurable sets.

Proof. Let A be a σ -finite set with respect to μ . In particular, A is μ -measurable. Let us assume that $\mu(A) < \infty$. Since μ is outer regular on all sets, for any $\varepsilon > 0$, there exists an open set $U \supseteq A$ such that

$$\mu(U \setminus A) = \mu(U) - \mu(A) < \varepsilon.$$

Likewise, there exists an open set $V \supseteq U \setminus A$ with $\mu(V) < \varepsilon$. In addition, since μ is inner regular on open sets, there exists a compact set $K \subseteq U$ such that $\mu(K) > \mu(U) - \varepsilon$. The set $L = K \setminus V$ is compact, as a closed subset of K , and satisfies $L \subseteq A$. Moreover, we have

$$\mu(L) = \mu(K) - \mu(K \cap V) > (\mu(U) - \varepsilon) - \mu(V) > \mu(A) - 2\varepsilon.$$

This shows that the measure μ is inner regular on the set A . If otherwise $\mu(A) = \infty$, then A is the union of a nondecreasing sequence of μ -measurable sets A_i such that $\mu(A_i) < \infty$. By Theorem 3, we have $\mu(A_i) \rightarrow \infty$ as $i \rightarrow \infty$. For any $\ell \in \mathbb{R}$, we have $\mu(A_i) > \ell$ for some i and, by the previous argument, there exists a compact set $L \subseteq A_i \subseteq A$ such that $\mu(L) > \ell$. We conclude that μ is inner regular on A .

For the last part of the statement, we just observe that if $X = \bigcup_i K_i$ with K_i compact, and if $A \in \mathcal{M}_\mu$, then $A = \bigcup_i (A \cap K_i)$ with $A \cap K_i \in \mathcal{M}_\mu$ and $\mu(A \cap K_i) \leq \mu(K_i) < \infty$, so that A is σ -finite. It follows from the first part of the statement that μ is inner regular on A . $\square$

Proposition 24 ensures that all σ -finite Radon measures are both outer and inner regular on all measurable sets. The following result discusses the approximation properties of σ -finite Radon measures for measurable sets. Recall that a G_δ -set is a countable intersection of open sets, and that an F_σ -set is a countable union of closed sets.

Proposition 25 (Measure approximation of Borel sets). *Let X be a locally compact Hausdorff space and let μ be a σ -finite Radon measure on X . Then, for any μ -measurable subset E of X , the following properties hold:*

- (i). *For any $\varepsilon > 0$, there exist a closed set F and an open set U with $F \subseteq E \subseteq U$ and $\mu(U \setminus F) < \varepsilon$;*
- (ii). *There exist an F_σ -set A and a G_δ -set B with $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$.*

Proof. Since μ is σ -finite, we can write $X = \bigcup_i A_i$ with $A_i \in \mathcal{M}_\mu$ and $\mu(A_i) < \infty$. By outer regularity, for any $\varepsilon > 0$ and any $i \geq 1$, there exists an open set $U_i \supseteq E \cap A_i$ with $\mu(U_i) < \mu(E \cap A_i) + \varepsilon 2^{-(i+1)}$. The set $U = \bigcup_i U_i$ is open, contains E , and satisfies

$$\mu(U \setminus E) \leq \sum_{i=1}^{\infty} \mu(U_i \setminus (E \cap A_i)) = \sum_{i=1}^{\infty} (\mu(U_i) - \mu(E \cap A_i)) < \frac{\varepsilon}{2}.$$

The same arguments applied to the set $X \setminus E$ yield an open set $V \supseteq X \setminus E$ such that $\mu(V \cap E) < \varepsilon/2$. Then, the set $F = X \setminus V$ is closed and contained in E . Finally, we have

$$\mu(U \setminus F) = \mu(U \setminus E) + \mu(E \setminus F) = \mu(U \setminus E) + \mu(E \cap V) < \varepsilon.$$

We thus established (i). In order to prove (ii), we apply (i) with $\varepsilon = 1/i$, thereby obtaining a sequence of closed sets F_i and a sequence of open sets U_i such that $F_i \subseteq E \subseteq U_i$ and $\mu(U_i \setminus F_i) < 1/i$ for any $i \geq 1$. The set $A = \bigcup_i F_i$ is an F_σ -set, the set $B = \bigcap_i U_i$ is a G_δ -set, and $A \subseteq E \subseteq B$. Finally, for all $i \geq 1$,

$$\mu(B \setminus A) \leq \mu(U_i \setminus F_i) \leq 1/i \xrightarrow{i \rightarrow \infty} 0.$$

This achieves the proof of (ii). □

We finish our treatment of Radon measures by showing that, under a σ -compactness assumption on the space, any locally finite Borel-regular measure is actually a Radon measure.

Theorem 26 (Sufficiency of local finiteness). *Let X be a locally compact Hausdorff space, and let μ be a locally finite Borel measure on X . Assume that any open subset of X is σ -compact. The following properties hold:*

- (i). *μ is outer and inner regular on all Borel sets;*
- (ii). *If in addition μ is Borel-regular, then it is a Radon measure.*

Proof. Recall that $C_c(X)$ is contained in L^1_μ , as a result of (10). Moreover, the mapping $f \mapsto \int_X f \, d\mu$ is a positive linear functional on $C_c(X)$, thanks to Proposition 10. Theorem 23 yields a Radon measure ν that corresponds to this positive linear functional. Let us consider an open subset U of X and assume that $U = \bigcup_i K_i$, where the sets K_i are compact. By Lemma 21(ii), there exists a function $f_1 \in C_c(X)$ such that $K_1 < f_1 < U$, and by induction, there exists a sequence of functions $f_i \in C_c(X)$ such that $(K_1 \cup \dots \cup K_i \cup \text{supp } f_1 \cup \dots \cup \text{supp } f_{i-1}) < f_i < U$ for all $i \geq 2$. This sequence of functions is nondecreasing in i and converges to $\mathbf{1}_U$, so the monotone convergence theorem, that is, Theorem 5, yields

$$\mu(U) = \int_X \mathbf{1}_U \, d\mu = \lim_{i \rightarrow \infty} \int_X f_i \, d\mu = \lim_{i \rightarrow \infty} \int_X f_i \, d\nu = \int_X \mathbf{1}_U \, d\nu = \nu(U).$$

The measure ν is σ -finite, because X is σ -compact and ν is locally finite. Moreover, any Borel subset E of X is ν -measurable. Proposition 25(i) yields a closed set F and an open set U such that $F \subseteq E \subseteq U$ and $\nu(U \setminus F) < \varepsilon$. Since $U \setminus F$ is open, we have $\mu(U \setminus F) = \nu(U \setminus F) < \varepsilon$.

Hence, $\mu(U) \leq \mu(E) + \varepsilon$, which proves outer regularity when $\mu(E) < \infty$; when $\mu(E) = \infty$, outer regularity is immediate by monotonicity. Moreover, F is closed in the σ -compact space X , so there are increasing compact sets K_j with union F . By Theorem 3, $\mu(K_j) \rightarrow \mu(F)$. If $\mu(E) < \infty$, then $\mu(F) \geq \mu(E) - \varepsilon$, and some K_j satisfies $\mu(K_j) > \mu(E) - 2\varepsilon$. If $\mu(E) = \infty$, the inequality $\mu(E \setminus F) < \varepsilon$ gives $\mu(F) = \infty$, so the compact masses $\mu(K_j)$ are unbounded. This proves inner regularity in both cases.

For the last part of the statement, we use the fact that if μ is Borel-regular and outer regular on Borel sets, it is actually outer regular on all sets. Since μ and ν coincide on open sets, they coincide on all sets. Hence, μ is the Radon measure ν . $\square$

Remark 2 (Open sets are σ -compact in second-countable locally compact Hausdorff spaces). A topological space is called *second-countable* if there exists a countable collection of open sets such that any open set in the space can be expressed as a union of open sets from this collection. If a locally compact Hausdorff space X is second-countable, then any open subset of X is σ -compact, and so the assumption of Theorem 26 is verified. To see this, let us consider an open subset U of X . By Lemma 21(i), for any point $x \in U$, there exists a relatively compact open set V_x such that $x \in V_x \subseteq \overline{V_x} \subseteq U$. Moreover, there exists an open set U_x from the countable base of X such that $x \in U_x \subseteq V_x$. Hence, $U = \bigcup_{x \in U} \overline{U_x}$, where the sets $\overline{U_x}$ are compact, and only countably many are distinct, since the base is countable.

4.5 Image measure

Let us now study the notion of image measure. We are particularly interested in the preservation of the Borel and Radon properties of measures when taking their image.

Definition 12 (Image measure). Let X and Y be two sets, let μ be a measure on X , and let $f: X \rightarrow Y$ be a function. The *image measure* of μ under f , also called the *pushforward measure* of μ by f , is the measure ν on Y defined for all subsets A of Y by the formula

$$\nu(A) = \mu(f^{-1}(A)).$$

The image measure ν is denoted by $f_{\#}\mu$, or $\mu \circ f^{-1}$.

This definition does not make any assumption on the measure μ or the function f . When Y is a topological space and f is μ -measurable, we obtain additional measurability properties.

Proposition 27 (Borel image measure). Let X be a set and let μ be a measure on X . Then, for any μ -measurable function $f: X \rightarrow Y$, where Y is a topological space, $f_{\#}\mu$ is a Borel measure on Y . This is the case in particular for f Borel-measurable, when X is a topological space and μ is a Borel measure.

Proof. Let us consider a Borel set $A \in \mathcal{B}(Y)$. If f is μ -measurable, then $f^{-1}(A)$ is μ -measurable, and thus A is ν -measurable, because for any subset E of Y , we have

$$\begin{aligned} \nu(E \cap A) + \nu(E \setminus A) &= \mu(f^{-1}(E \cap A)) + \mu(f^{-1}(E \setminus A)) \\ &= \mu(f^{-1}(E) \cap f^{-1}(A)) + \mu(f^{-1}(E) \setminus f^{-1}(A)) \\ &= \mu(f^{-1}(E)) = \nu(E). \end{aligned}$$

When X is a topological space and μ is a Borel measure, any Borel-measurable function is necessarily μ -measurable, so the above arguments apply. $\square$

Furthermore, the Radon property of a measure is also conveniently transferred to its image under a continuous function, when the measure has compact support.

Theorem 28 (Radon image measure). *Let X and Y be locally compact Hausdorff spaces, let μ be a compactly supported Radon measure on X , and let $f: X \rightarrow Y$ be a continuous function. Then, $f_{\#}\mu$ is a Radon measure on Y . Moreover,*

$$\text{supp } f_{\#}\mu = f(\text{supp } \mu).$$

Proof. Since μ is a Borel measure and f is Borel-measurable, as a continuous function, Proposition 27 shows that $f_{\#}\mu$ is a Borel measure. Moreover, μ has compact support $S = \text{supp } \mu$ and is finite on compact sets, because it is locally finite, so $\mu(S) < \infty$. Since μ is inner regular on all open sets, we also know from Proposition 20 that $\mu(X \setminus S) = 0$. Therefore, for any compact subset L of Y ,

$$f_{\#}\mu(L) = \mu(f^{-1}(L)) \leq \mu(S) < \infty.$$

Given that the space Y is locally compact, this ensures that $f_{\#}\mu$ is locally finite.

Let us prove that $f_{\#}\mu$ is inner regular on any open subset U of Y . Since f is continuous, $f^{-1}(U)$ is open, so for any $\varepsilon > 0$, the inner regularity of μ implies that $\mu(K) > \mu(f^{-1}(U)) - \varepsilon$ for some compact set $K \subseteq f^{-1}(U)$. The set $L = f(K)$ is compact and contained in U , and satisfies

$$f_{\#}\mu(L) = \mu(f^{-1}(f(K))) \geq \mu(K) > \mu(f^{-1}(U)) - \varepsilon = f_{\#}\mu(U) - \varepsilon.$$

In order to establish that $f_{\#}\mu$ is outer regular on an arbitrary subset A of Y , we consider $\varepsilon > 0$ and we use the outer regularity of μ to obtain an open set $U \supseteq f^{-1}(A) \cap S$ such that

$$\mu(U) < \mu(f^{-1}(A) \cap S) + \varepsilon \leq f_{\#}\mu(A) + \varepsilon.$$

Note that $V = Y \setminus f(S \setminus U)$ is an open set, as the complement of a continuous image of a compact set. Moreover, $A \cap f(S \setminus U) = \emptyset$, so that $A \subseteq V$. In addition, we have

$$S \cap f^{-1}(V) = S \setminus f^{-1}(f(S \setminus U)) \subseteq S \setminus (S \setminus U) \subseteq U,$$

from which we directly deduce that

$$f_{\#}\mu(V) = \mu(f^{-1}(V)) \leq \mu(X \setminus S) + \mu(S \cap f^{-1}(V)) \leq \mu(U) \leq f_{\#}\mu(A) + \varepsilon.$$

We have thus checked all the requirements to confirm that $f_{\#}\mu$ is a Radon measure on Y . It remains to show that the support of $f_{\#}\mu$ is the image of that of μ . For any open subset U of Y , we have

$$\begin{aligned} U \subseteq Y \setminus \text{supp } f_{\#}\mu &\iff f_{\#}\mu(U) = 0 \iff \mu(f^{-1}(U)) = 0 \\ &\iff f^{-1}(U) \subseteq X \setminus \text{supp } \mu \\ &\iff f^{-1}(U) \cap \text{supp } \mu = \emptyset \iff U \cap f(\text{supp } \mu) = \emptyset \\ &\iff U \subseteq Y \setminus f(\text{supp } \mu). \end{aligned}$$

The result follows from the fact that $\text{supp } f_{\#}\mu$ and $f(\text{supp } \mu)$ are two closed sets. $\square$

Finally, we also have the following useful integration formula for image measures.

Theorem 29 (Integration formula). *Let X and Y be topological spaces, let μ be a Borel measure on X , and let $f: X \rightarrow Y$ be a Borel-measurable function. Then, for any nonnegative Borel-measurable function g defined on Y ,*

$$\int_Y g \, d(f_{\#}\mu) = \int_X g \circ f \, d\mu.$$

This formula remains valid for g valued in $\mathbb{C}$, provided that either of the integrals is finite when the integrand is replaced by its modulus.

Proof. We first remark that g and $g \circ f$ are Borel-measurable functions, and that $f_{\#}\mu$ and μ are Borel measures, in view of Proposition 27. Hence, g is $f_{\#}\mu$ -measurable and $g \circ f$ is μ -measurable, so the integrals are well-defined. Let us assume that g is simple, so that $g \circ f$ is also simple. By Definition 4,

$$\int_X g \circ f \, d\mu = \sum_{z \in (g \circ f)(X)} z \mu((g \circ f)^{-1}(\{z\})) = \sum_{z \in g(f(X))} z \mu(f^{-1}(g^{-1}(\{z\}))).$$

For $z \in g(Y) \setminus g(f(X))$, we have $f^{-1}(g^{-1}(\{z\})) = \emptyset$, so $f_{\#}\mu(g^{-1}(\{z\})) = 0$. Hence,

$$\int_Y g \, d(f_{\#}\mu) = \sum_{z \in g(Y)} z f_{\#}\mu(g^{-1}(\{z\})) = \sum_{z \in g(f(X))} z \mu(f^{-1}(g^{-1}(\{z\}))).$$

It follows that the formula is valid when g is simple. To deduce the result in the general case, we use the sequence of nonnegative simple Borel-measurable functions ϕ_j built in the proof of Lemma 6, and we conclude with the help of the monotone convergence theorem, that is, Theorem 5.

To deal with the case where g is valued in $\mathbb{C}$, it suffices to apply the formula for nonnegative functions to the positive and the negative parts of the real and the imaginary parts of g . $\square$

4.6 Weak convergence

Let us now study the convergence properties of Radon measures, in the following dual sense.

Definition 13 (Weak convergence of Radon measures). *Let X be a locally compact Hausdorff space, and let μ_j and μ be Radon measures on X . The sequence $(\mu_j)_{j \geq 1}$ converges weakly to μ if for any function $f \in C_c(X)$,*

$$\int_X f \, d\mu_j \xrightarrow{j \rightarrow \infty} \int_X f \, d\mu.$$

This property is denoted by $\mu_j \rightharpoonup \mu$ as $j \rightarrow \infty$.

This use of $C_c(X)$ test functions is also called *vague convergence*; probability texts often reserve weak convergence for convergence against all bounded continuous functions.

On a second-countable locally compact Hausdorff space, any uniformly locally finite sequence of Radon measures admits a weakly convergent subsequence. To prove this result, we shall use the Riesz representation theorem, and the separability of the space $C_c(X)$ when X is compact and metrizable, which is addressed in the following lemma.

We use the following form of the *Stone-Weierstrass theorem*: if K is a compact metrizable space and $\mathcal{A} \subseteq C(K; \mathbb{C})$ is a complex subalgebra containing the constants, separating points, and closed under complex conjugation, then $\mathcal{A}$ is dense in $C(K; \mathbb{C})$ for the uniform norm. This approximation theorem is admitted here; see Theorem 7.33, p. 165, in [10].

Lemma 30 (Separability of $C_c(X)$). *If X is compact and metrizable, then $C_c(X)$ is separable.*

Proof. Let d be a compatible metric on X . Since X is compact, it is separable, so that there exists a countable dense subset $\{x_n, n \geq 1\}$ of X . Let us consider the functions defined by $f_n(x) = d(x, x_n)$, and the family $\mathcal{P}$ consisting of the functions $P(f_1, \dots, f_n)$, where P is a polynomial in n variables with complex coefficients. The family $\mathcal{P}$ is an algebra over $\mathbb{C}$ of complex continuous functions on X that is self-adjoint (stable under complex conjugation), that vanishes at no point of X (because $\mathcal{P}$ contains the constant function equal to 1), and that separates points on X (because if $x \neq y$, there exists $n \geq 1$ such that $d(x_n, x) \neq d(x_n, y)$). The Stone-Weierstrass theorem ensures that the algebra $\mathcal{P}$ is dense in $C_c(X)$. We refer to Theorem 7.33, p. 165, in [10] for the statement of the Stone-Weierstrass theorem that we use here. However, the functions f_n are uniformly bounded on X by the diameter of X , and the set of Gaussian rationals $\mathbb{Q}(i)$ is dense in $\mathbb{C}$, so restricting the coefficients of the polynomials to $\mathbb{Q}(i)$, we establish that the algebra $\mathcal{P}$ is separable, and thus that $C_c(X)$ is separable. $\square$

We also use the following topological fact: every second-countable locally compact Hausdorff space admits a compatible metric and is a countable union of compact sets. This is the metrization and σ -compactness result recalled in the corollary to Proposition 16 in Chapter IX, § 2, no. 9, pp. 158–159, of [2].

Combining Lemma 30 with Theorem 23, we can now establish the following result.

Theorem 31 (Uniformly locally finite sequences of Radon measures). *Let X be a second-countable locally compact Hausdorff space, and let $(\mu_j)_{j \geq 1}$ be a sequence of Radon measures on X that is uniformly locally finite, meaning that for any compact set $K \subseteq X$,*

$$C_K = \sup_{j \geq 1} \mu_j(K) < \infty.$$

Then, there exists a subsequence $(\mu_{j_k})_{k \geq 1}$ and a Radon measure μ such that $\mu_{j_k} \rightarrow \mu$ as $k \rightarrow \infty$.

Proof. Since the space X is second-countable and locally compact, it is metrizable and σ -compact, see the corollary to Proposition 16 in Chapter IX, § 2, no. 9, pp. 158–159, of [2]. It is also exhaustible by compact sets. Indeed, let us assume that $X = \bigcup_m L_m$ with L_m compact. By Lemma 21(i), there exists a relatively compact open set $U_1 \supseteq L_1$. For $m \geq 2$, we can thus define inductively a relatively compact open set $U_m \supseteq \overline{U_{m-1}} \cup L_m$. Letting $K_m = \overline{U_m}$, we see that $X = \bigcup_m K_m$, with K_m compact and $K_m \subseteq \text{int } K_{m+1}$. Every compact K_m is metrizable, so that $C_c(K_m)$ is separable, by Lemma 30. Let E_m be the subspace of $C_c(X)$ consisting of functions with support contained in K_m . Restriction to K_m identifies E_m isometrically with a subspace of the separable metric space $C_c(K_m)$. Every subspace of a separable metric space is separable: intersect a countable base with the subspace and choose one point from each nonempty intersection. Thus, let $\{\phi_p^{(m)}, p \geq 1\}$ denote a countable dense subset of E_m . Up to adding a countable number of functions, we can assume that for any $p \geq 1$, there exists $p' \geq 1$ such that $\phi_{p'}^{(m)} = (\Re \phi_p^{(m)})^+$. Since $z \mapsto (\Re z)^+$ is 1-Lipschitz continuous, this ensures that any nonnegative function in E_m can be approximated by nonnegative functions from the dense subset.

Let $\psi_1, \psi_2, \dots$ be an enumeration of the set $\{\phi_p^{(m)}, m, p \geq 1\}$. The uniform local finiteness yields

$$\left| \int_X \psi_k d\mu_j \right| \leq \int_X |\psi_k| d\mu_j \leq C_{\text{supp } \psi_k} \|\psi_k\|_\infty.$$

The sequence in the left-hand side is bounded in j , so we can extract a convergent subsequence. Specifically, we obtain an extraction $\eta_1(j)$ such that $\int_X \psi_1 d\mu_{\eta_1(j)}$ converges. By induction, we further obtain extractions $(\eta_1 \circ \dots \circ \eta_k)(j)$ such that $\int_X \psi_k d\mu_{(\eta_1 \circ \dots \circ \eta_k)(j)}$ converges. Finally, the diagonal extraction defined by $j_k = (\eta_1 \circ \dots \circ \eta_k)(k)$ satisfies for all $m, p \geq 1$,

$$\int_X \phi_p^{(m)} d\mu_{j_k} \xrightarrow{k \rightarrow \infty} J(\phi_p^{(m)}) \in \mathbb{C}.$$

Let us consider a function $f \in C_c(X)$. We have $\text{supp } f \subseteq K_m$ for m large enough, and there exists a sequence of functions $\phi_{p_n}^{(m)}$ that converges uniformly to f . By the uniform local finiteness, we have

$$\left| \int_X \phi_{p_n}^{(m)} d\mu_{j_k} - \int_X \phi_{p_{n'}}^{(m)} d\mu_{j_k} \right| \leq \int_X |\phi_{p_n}^{(m)} - \phi_{p_{n'}}^{(m)}| d\mu_{j_k} \leq C_{K_m} \|\phi_{p_n}^{(m)} - \phi_{p_{n'}}^{(m)}\|_\infty.$$

Letting $k \rightarrow \infty$, we deduce that

$$|J(\phi_{p_n}^{(m)}) - J(\phi_{p_{n'}}^{(m)})| \leq C_{K_m} \|\phi_{p_n}^{(m)} - \phi_{p_{n'}}^{(m)}\|_\infty.$$

The sequence $(J(\phi_{p_n}^{(m)}))_{n \geq 1}$ is Cauchy, and thus converges to some limit $I_m(f) \in \mathbb{C}$. The same arguments imply that the value $I_m(f)$ does not depend on the choice of the approximating functions $\phi_{p_n}^{(m)}$, and does not depend on m provided that $\text{supp } f \subseteq K_m$. Hence, $I_m(f) = I(f) \in \mathbb{C}$ when $\text{supp } f \subseteq K_m$.

To show that I is linear, let us consider $\alpha \in \mathbb{C}$ and $f, g \in C_c(X)$. For m large enough, the supports of f, g and $\alpha f + g$ are contained in K_m . There are sequences of functions $\phi_{p_n}^{(m)}, \phi_{p'_n}^{(m)}$ and $\phi_{p''_n}^{(m)}$ that converge uniformly to f, g and $\alpha f + g$, respectively. The uniform local finiteness yields

$$\begin{aligned} & \left| \int_X \phi_{p''_n}^{(m)} d\mu_{j_k} - \int_X (\alpha \phi_{p_n}^{(m)} + \phi_{p'_n}^{(m)}) d\mu_{j_k} \right| \\ & \leq \int_X |\phi_{p''_n}^{(m)} - (\alpha \phi_{p_n}^{(m)} + \phi_{p'_n}^{(m)})| d\mu_{j_k} \leq C_{K_m} \|\phi_{p''_n}^{(m)} - (\alpha \phi_{p_n}^{(m)} + \phi_{p'_n}^{(m)})\|_\infty. \end{aligned}$$

Using the linearity of the integral and letting $k \rightarrow \infty$ in the left-hand side, we obtain

$$|J(\phi_{p''_n}^{(m)}) - \alpha J(\phi_{p_n}^{(m)}) - J(\phi_{p'_n}^{(m)})| \leq C_{K_m} \|\phi_{p''_n}^{(m)} - (\alpha \phi_{p_n}^{(m)} + \phi_{p'_n}^{(m)})\|_\infty.$$

Letting $n \rightarrow \infty$ with m fixed, we deduce that $I(\alpha f + g) = \alpha I(f) + I(g)$. Hence, I is a linear functional on $C_c(X)$. To show that I is positive, let us consider a nonnegative function $f \in C_c(X)$. Again, for m large enough, the support of f is contained in K_m , and there exists a sequence of functions $\phi_{p_n}^{(m)}$ that converges uniformly to f . However, we have another sequence $\phi_{p'_n}^{(m)} = (\Re \phi_{p_n}^{(m)})^+ \geq 0$ that converges uniformly to $(\Re f)^+ = f$. Note that $J(\phi_{p'_n}^{(m)}) \geq 0$, so that $I_m(f) \geq 0$, and finally $I(f) \geq 0$.

Since I is a positive linear functional on $C_c(X)$, the Riesz representation theorem, that is, Theorem 23, ensures that there exists a unique Radon measure μ on X such that $I(f) = \int_X f d\mu$ for all $f \in C_c(X)$. It remains to prove that $\mu_{j_k} \rightarrow \mu$ as $k \rightarrow \infty$. Let us consider a function $f \in C_c(X)$ such that $\text{supp } f \subseteq K_m$, and a sequence of functions $\phi_{p_n}^{(m)}$ that converges uniformly to f . We have

$$\left| \int_X f d\mu - \int_X f d\mu_{j_k} \right| \leq |I_m(f) - J(\phi_{p_n}^{(m)})| + \left| J(\phi_{p_n}^{(m)}) - \int_X \phi_{p_n}^{(m)} d\mu_{j_k} \right|$$

$$\begin{aligned}
& + \left| \int_X \phi_{p_n}^{(m)} d\mu_{j_k} - \int_X f d\mu_{j_k} \right| \\
& \leq |I_m(f) - J(\phi_{p_n}^{(m)})| + \left| J(\phi_{p_n}^{(m)}) - \int_X \phi_{p_n}^{(m)} d\mu_{j_k} \right| \\
& + C_{K_m} \|\phi_{p_n}^{(m)} - f\|_\infty.
\end{aligned}$$

Letting $k \rightarrow \infty$, we obtain

$$\limsup_{k \rightarrow \infty} \left| \int_X f d\mu - \int_X f d\mu_{j_k} \right| \leq |I_m(f) - J(\phi_{p_n}^{(m)})| + C_{K_m} \|\phi_{p_n}^{(m)} - f\|_\infty \xrightarrow{n \rightarrow \infty} 0,$$

which implies that $\int_X f d\mu_{j_k} \rightarrow \int_X f d\mu$ as $k \rightarrow \infty$, thereby achieving the proof. $\square$

The simple example of a Dirac mass δ_{x_j} on $\mathbb{R}$ that escapes to infinity when $x_j \rightarrow \infty$ shows that the weak convergence of Radon measures does not imply the convergence of the measure of a particular set. Indeed, $\delta_{x_j} \rightarrow 0$ as $j \rightarrow \infty$, but for instance $\delta_{x_j}(\mathbb{R}) = 1$ for all j . However, we have the following semicontinuity result for the measures of open and compact sets, which is a direct consequence of the Riesz representation theorem.

Theorem 32 (Portmanteau for Radon Measures). *Let X be a locally compact Hausdorff space, and let μ and μ_j be Radon measures on X such that $\mu_j \rightarrow \mu$ as $j \rightarrow \infty$. Then:*

- (i). *For any open set $U \in \mathcal{G}(X)$, it holds that $\mu(U) \leq \liminf_{j \rightarrow \infty} \mu_j(U)$;*
- (ii). *For any compact set $K \in \mathcal{K}(X)$, it holds that $\mu(K) \geq \limsup_{j \rightarrow \infty} \mu_j(K)$.*

Proof. Let us consider the positive linear functionals defined by $I_j(f) = \int_X f d\mu_j$ and $I(f) = \int_X f d\mu$ for $f \in C_c(X)$. The formulas (i) and (ii) of Theorem 23 hold for these functionals and the corresponding Radon measures. Let us consider an open set U , and a function $f \in C_c(X)$ such that $0 < f < U$. Then,

$$I(f) = \lim_{j \rightarrow \infty} I_j(f) \leq \liminf_{j \rightarrow \infty} \mu_j(U),$$

as a result of (i), that we subsequently use again to take the supremum over f in the left-hand side. Similarly, let us consider a compact set K , and a function $f \in C_c(X)$ such that $K < f < X$. Then,

$$I(f) = \lim_{j \rightarrow \infty} I_j(f) \geq \limsup_{j \rightarrow \infty} \mu_j(K),$$

as a result of (ii), that we use once more to take the infimum over f in the left-hand side. $\square$

5 Metric measures

5.1 Definition and Borel property

We now study the particular case of metric spaces, and consider important examples of measures related to the size and the geometry of sets. These measures fall in the category of *metric measures*.

Definition 14 (Metric measure). Let (X, d) be a metric space. A measure μ on X is *metric* if for all sets $A, B \subseteq X$,

$$d(A, B) > 0 \implies \mu(A \cup B) = \mu(A) + \mu(B),$$

where $d(A, B) = \inf\{d(a, b), a \in A, b \in B\}$ is the distance between A and B .

Metric measures are Borel measures. To prove this, we need the following lemma, which is Lemma 1.4, pp. 5–6, in [5].

Lemma 33 (Carathéodory's lemma). Let (X, d) be a metric space. For any metric measure μ on X , and for any nondecreasing sequence of subsets A_j of X with union $A = \bigcup_j A_j$, the following implication holds:

$$[\forall j \geq 1 \quad d(A_j, A \setminus A_{j+1}) > 0] \implies \mu(A) = \sup_{j \geq 1} \mu(A_j) = \lim_{j \rightarrow \infty} \mu(A_j).$$

Proof. Define $B_1 = A_1$ and $B_j = A_j \setminus A_{j-1}$ for all $j \geq 2$. For $i \geq j + 2$, we have $B_i \subseteq A \setminus A_{i-1} \subseteq A \setminus A_{j+1}$, so that $d(A_j, B_i) \geq d(A_j, A \setminus A_{j+1}) > 0$. Hence, for any $m \geq 2$, we have

$$d(B_1 \cup B_3 \cup \cdots \cup B_{2m-3}, B_{2m-1}) \geq d(A_{2m-3}, B_{2m-1}) > 0,$$

and a similar property for sets with even indices. By induction, using Definition 14, we get

$$\mu\left(\bigcup_{k=1}^m B_{2k-1}\right) = \sum_{k=1}^m \mu(B_{2k-1}) \quad \text{and} \quad \mu\left(\bigcup_{k=1}^m B_{2k}\right) = \sum_{k=1}^m \mu(B_{2k}).$$

If at least one of these series diverges, we have $\mu(A) \geq \mu(A_j) \rightarrow \infty$ as $j \rightarrow \infty$, because the unions are both contained in A_{2m} . We may therefore assume that $\sum_k \mu(B_k) < \infty$, which gives

$$\mu(A) = \mu\left(A_j \cup \bigcup_{k=j+1}^{\infty} B_k\right) \leq \mu(A_j) + \sum_{k=j+1}^{\infty} \mu(B_k) \xrightarrow{j \rightarrow \infty} \lim_{j \rightarrow \infty} \mu(A_j) \leq \mu(A),$$

which achieves the proof. $\square$

We can now establish that Borel sets are measurable with respect to metric measures.

Theorem 34 (Metric measures are Borel). Let X be a metric space, and let μ be a metric measure on X . Then, μ is a Borel measure.

Proof. It suffices to show that any closed set A is μ -measurable. For all $E \subseteq X$ and $j \geq 1$, let us define $E_j = \{x \in E \setminus A \mid d(\{x\}, A) \geq 1/j\}$. We have $d(E_j, E \cap A) \geq d(E_j, A) \geq 1/j > 0$. Hence, by Definition 14,

$$\mu(E_j) + \mu(E \cap A) = \mu(E_j \cup (E \cap A)) \leq \mu((E \setminus A) \cup (E \cap A)) = \mu(E).$$

Moreover, the sets E_j are nondecreasing in j , and their union is equal to $E \setminus A$, because A is closed. In addition, they satisfy the assumption of Lemma 33, specifically, $d(E_j, (E \setminus A) \setminus E_{j+1}) > 0$ for all $j \geq 1$. To see this, consider a point $x \in (E \setminus A) \setminus E_{j+1}$. Then, $d(x, a) < 1/(j+1)$ for some $a \in A$. For $y \in E_j$, we have $d(y, a) \geq d(\{y\}, A) \geq 1/j$, and the triangle inequality implies that $d(x, y) > 1/j - 1/(j+1)$, which ensures that $d(E_j, (E \setminus A) \setminus E_{j+1}) \geq 1/(j(j+1)) > 0$. Lemma 33 now gives $\mu(E \setminus A) + \mu(E \cap A) \leq \mu(E)$ for all $E \subseteq X$, proving the measurability of A . $\square$

5.2 The Lebesgue measure

The principal example of a metric measure is Lebesgue measure, defined below.

Definition 15 (*N-dimensional Lebesgue measure*). For any integer $N \geq 1$, the *N-dimensional Lebesgue measure* is the measure on $\mathbb{R}^N$ defined by

$$\mathcal{L}^N(A) = \inf \left\{ \sum_{i=1}^{\infty} (2r_i)^N, A \subseteq \bigcup_{i=1}^{\infty} Q(x_i, r_i) \right\}.$$

The infimum is over all coverings of the set $A \subseteq \mathbb{R}^N$ by open cubes centered at $x \in \mathbb{R}^N$ with side length $2r \geq 0$, namely, by cubes $Q(x, r) = \{y \in \mathbb{R}^N \mid \max_n |x_n - y_n| < r\}$. A cube with $r = 0$ is empty.

Theorem 34 and the results of Section 4.4 imply the following important property.

Theorem 35 (*Metric and Radon properties of the Lebesgue measure*). For any $N \geq 1$, the *N-dimensional Lebesgue measure* $\mathcal{L}^N$ is a metric Radon measure on $\mathbb{R}^N$.

Proof. Let us check that $\mathcal{L}^N$ is a measure. First, the empty set $\emptyset$ is covered by cubes with side length 0, so that $\mathcal{L}^N(\emptyset) = 0$. Furthermore, if $A \subseteq B$, any covering of B is a covering of A , so that $\mathcal{L}^N(A) \leq \mathcal{L}^N(B)$. Finally, for any sequence of sets A_j and any $\varepsilon > 0$, there exists a covering of A_j by cubes $Q(x_{j,i}, r_{j,i})$ such that $\sum_{i=1}^{\infty} (2r_{j,i})^N \leq \mathcal{L}^N(A_j) + \varepsilon 2^{-j}$. The union of the cubes $Q(x_{j,i}, r_{j,i})$ is a covering of the set $A = \bigcup_j A_j$. The associated sum $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} (2r_{j,i})^N$ is bounded above by $\varepsilon + \sum_{j=1}^{\infty} \mathcal{L}^N(A_j)$, and below by $\mathcal{L}^N(A)$. We deduce the countable subadditivity by letting $\varepsilon \rightarrow 0$.

In order to show that the Lebesgue measure is a metric measure, we observe that the infimum that defines $\mathcal{L}^N(A)$ is unchanged when the coverings are restricted to cubes of arbitrarily small side length. Specifically, for any $\delta > 0$,

$$\mathcal{L}^N(A) = \inf \left\{ \sum_{i=1}^{\infty} (2r_i)^N, A \subseteq \bigcup_{i=1}^{\infty} Q(x_i, r_i) \text{ with } r_i < \delta \right\}. \quad (13)$$

To prove this, let us consider a covering $A \subseteq \bigcup_i Q(x_i, r_i)$. For any $\varepsilon > 0$ and $m_i \geq 1$, the open cube $Q(x_i, r_i)$ can be covered by m_i^N open cubes with side length $2(1 + \varepsilon)r_i/m_i$, which is less than δ for m_i large enough. We obtain a covering of A with sum $\sum_i m_i^N (2(1 + \varepsilon)r_i/m_i)^N = (1 + \varepsilon)^N \sum_i (2r_i)^N$. By taking the infimum over all coverings, we deduce that the right-hand side of (13) is bounded above by $(1 + \varepsilon)^N \mathcal{L}^N(A)$, while being clearly bounded below by $\mathcal{L}^N(A)$. We conclude by letting $\varepsilon \rightarrow 0$ that (13) holds. Now, let us consider two sets $A, B \subseteq \mathbb{R}^N$ such that $d(A, B) > 0$, and a covering of $A \cup B$ by cubes $Q(x_i, r_i)$ with $r_i < \delta$. Define I_A as the set of indices i such that $A \cap Q(x_i, r_i) \neq \emptyset$, and define I_B similarly. The sets I_A and I_B are disjoint if δ is small enough. Moreover, the set A is covered by the cubes $Q(x_i, r_i)$ with $i \in I_A$, and a similar property holds for B . It follows that $\mathcal{L}^N(A \cup B) \geq \mathcal{L}^N(A) + \mathcal{L}^N(B)$.

We just proved that $\mathcal{L}^N$ is a metric measure, and Theorem 34 implies that $\mathcal{L}^N$ is a Borel measure. Furthermore, for every set $A \subseteq \mathbb{R}^N$ and $m \geq 1$, there exists a covering $A \subseteq B_m = \bigcup_{i=1}^{\infty} Q(x_{m,i}, r_{m,i})$ such that $\mathcal{L}^N(B_m) \leq \sum_{i=1}^{\infty} (2r_{m,i})^N \leq \mathcal{L}^N(A) + 1/m$. The set $B = \bigcap_m B_m$ is a G_δ -set containing A and satisfying $\mathcal{L}^N(B) \leq \mathcal{L}^N(A)$. Therefore, $\mathcal{L}^N$ is a Borel-regular measure.

By Theorem 26, only local finiteness remains to be checked. Every point has an open ball of finite measure, since $\mathcal{L}^N(B(x, r)) \leq (2r)^N < \infty$. $\square$

Proposition 36 (Lebesgue measure of rectangles). *The Lebesgue measure of a rectangle in $\mathbb{R}^N$ with sides parallel to the coordinate axes is given by*

$$\mathcal{L}^N\left(\prod_{i=1}^N [a_i, b_i]\right) = \prod_{i=1}^N (b_i - a_i), \quad (14)$$

where $-\infty < a_i \leq b_i < \infty$ for all i . Moreover, the above formula is also valid for open rectangles.

Proof. An open rectangle with a zero side length is empty and has measure zero. Otherwise, since open rectangles can be sandwiched between two arbitrarily close closed rectangles, the formula (14) for closed rectangles implies the formula for open rectangles. To prove (14), note from Definition 15 that the Lebesgue measure is translation-invariant and satisfies the scaling property $\mathcal{L}^N([0, \delta]^N) = \delta^N \mathcal{L}^N([0, 1]^N)$ for every $\delta > 0$. In addition, the rectangle $\prod_{i=1}^N [a_i, b_i]$ is sandwiched between the union of $\prod_{i=1}^N \lfloor \frac{b_i - a_i}{\delta} \rfloor$ disjoint closed cubes with side length $\delta' \in (0, \delta)$, and the union of $\prod_{i=1}^N \max\left\{1, \lceil \frac{b_i - a_i}{\delta} \rceil\right\}$ closed cubes with side length δ , where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ denote the floor and the ceiling function, respectively. Therefore,

$$\mathcal{L}^N([0, 1]^N) \delta'^N \prod_{i=1}^N \left\lfloor \frac{b_i - a_i}{\delta} \right\rfloor \leq \mathcal{L}^N\left(\prod_{i=1}^N [a_i, b_i]\right) \leq \mathcal{L}^N([0, 1]^N) \delta^N \prod_{i=1}^N \max\left\{1, \left\lceil \frac{b_i - a_i}{\delta} \right\rceil\right\}.$$

Letting $\delta' \rightarrow \delta$ and then $\delta \rightarrow 0$, we deduce that

$$\mathcal{L}^N\left(\prod_{i=1}^N [a_i, b_i]\right) = \mathcal{L}^N([0, 1]^N) \prod_{i=1}^N (b_i - a_i).$$

It now remains to prove (14) for the unit cube $[0, 1]^N$. For every $\delta > 0$, the unit cube is contained in an open cube with side length $1 + \delta$. It follows from Definition 15 that $\mathcal{L}^N([0, 1]^N) \leq (1 + \delta)^N$. Letting $\delta \rightarrow 0$, we obtain $\mathcal{L}^N([0, 1]^N) \leq 1$. To prove the reverse inequality, assume that the unit cube is covered by open cubes of the form $Q(x_i, r_i)$. By the countable subadditivity, the translation invariance and the scaling property of the Lebesgue measure, we have

$$\mathcal{L}^N([0, 1]^N) \leq \sum_{i=1}^{\infty} \mathcal{L}^N(Q(x_i, r_i)) \leq \mathcal{L}^N([0, 1]^N) \sum_{i=1}^{\infty} (2r_i)^N.$$

We deduce from Definition 15 that $\mathcal{L}^N([0, 1]^N) \leq \mathcal{L}^N([0, 1]^N)^2$, and thus $\mathcal{L}^N([0, 1]^N) \in \{0, 1\}$.

To prove that $\mathcal{L}^N([0, 1]^N) = 1$, we extract by compactness a finite subcover of the unit cube, and thus assume that $[0, 1]^N \subseteq \bigcup_{i=1}^n Q(x_i, r_i)$ with $r_i > 0$. If one of these cubes has side length $2r_i \geq 1$, then $1 \leq \sum_i (2r_i)^N$, which already gives the required lower bound for this covering. Otherwise, all $2r_i < 1$, and we argue as follows. Recall that a dyadic cube is a set of the form $2^{-j}(k + [0, 1]^N)$, for $j \in \{0, 1, \dots\}$ and $k \in \mathbb{Z}^N$. Each set $[0, 1]^N \cap Q(x_i, r_i)$ is covered by at most 2^N dyadic subcubes of $[0, 1]^N$ with side length 2^{-j_i} , where j_i is chosen such that $2^{-j_i-1} \leq 2r_i < 2^{-j_i}$. Each of these subcubes contains $2^{N(j-j_i)}$ dyadic subcubes with side length 2^{-j} , where $J = \max\{j_1, \dots, j_n\}$. The total number of the latter subcubes is equal to 2^{NJ} , because their union is equal to $[0, 1]^N$. (However, each of them can belong to several of the former subcubes.) Hence, $2^{NJ} \leq 2^N \sum_{i=1}^n 2^{N(j-j_i)}$, which implies that $1 \leq 2^N \sum_{i=1}^n 2^{-Nj_i} \leq 2^{2N} \sum_{i=1}^{\infty} (2r_i)^N$. We conclude from Definition 15 that $\mathcal{L}^N([0, 1]^N) \geq 2^{-2N} > 0$, and finally that $\mathcal{L}^N([0, 1]^N) = 1$. $\square$

The next result shows that the N -dimensional Lebesgue measure can also be obtained as a consequence of the Riesz representation theorem and the properties of product measures.

Proposition 37 (Alternative construction of the Lebesgue measure). *The mapping $I: f \mapsto \sum_{j=-\infty}^{\infty} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f\left(j + \frac{k}{n}\right)$ is a positive linear functional on $C_c(\mathbb{R})$, represented by a unique Radon measure μ on $\mathbb{R}$. Then, the measure μ_N defined inductively by $\mu_1 = \mu$ and $\mu_N = \mu_{N-1} \otimes \mu$ is the N -dimensional Lebesgue measure $\mathcal{L}^N$ on $\mathbb{R}^N$.*

Proof. It is easy to see that I is a positive linear functional on $C_c(\mathbb{R})$. Indeed, there are finitely many nonvanishing terms in the series, and those terms are well-defined as limits of Riemann sums. The Riesz representation theorem, that is, Theorem 23 implies that there exists a unique Radon measure μ on $\mathbb{R}$ such that $I(f) = \int_{\mathbb{R}} f \, d\mu$. For $a < b$ and $0 < \varepsilon < (b - a)/2$, let f_ε denote the continuous function that is constant equal to zero outside of (a, b) , constant equal to one on $[a + \varepsilon, b - \varepsilon]$, and affine on $[a, a + \varepsilon]$ and $[b - \varepsilon, b]$. Similarly, let g_ε denote the continuous function that is constant equal to zero outside of $(a - \varepsilon, b + \varepsilon)$, constant equal to one on $[a, b]$, and affine on $[a - \varepsilon, a]$ and $[b, b + \varepsilon]$. We have $f_\varepsilon \leq \mathbf{1}_{(a,b)} \leq g_\varepsilon$, so that

$$b - a - \varepsilon = I(f_\varepsilon) \leq \mu((a, b)) \leq I(g_\varepsilon) = b - a + \varepsilon.$$

It follows that $\mu((a, b)) = b - a$. Furthermore, let us consider a covering of a set $A \subseteq \mathbb{R}$ by open intervals $(x_i - r_i, x_i + r_i)$. By countable subadditivity, $\mu(A) \leq \sum_i (2r_i)$. Taking the infimum over all coverings, we obtain $\mu(A) \leq \mathcal{L}^1(A)$ for any set $A \subseteq \mathbb{R}$. Moreover, since $\mathcal{L}^1$ and μ are both Borel-regular measures, there exists a Borel set $B \supseteq A$ such that $\mathcal{L}^1(B) = \mathcal{L}^1(A)$ and $\mu(B) = \mu(A)$. Then, for any $r > 0$,

$$\begin{aligned} 0 = \mathcal{L}^1((-r, r)) - \mu((-r, r)) &= (\mathcal{L}^1((-r, r) \cap B) - \mu((-r, r) \cap B)) \\ &\quad + (\mathcal{L}^1((-r, r) \setminus B) - \mu((-r, r) \setminus B)), \end{aligned}$$

where the two terms in parentheses are nonnegative. These two terms are thus equal to zero, and Theorem 3 implies that $\mu(B) = \mathcal{L}^1(B)$, so that $\mu(A) = \mathcal{L}^1(A)$. The Radon measures μ and $\mathcal{L}^1$ coincide.

Using Proposition 15, we can show by induction that $\mu_N(Q(x, r)) = (2r)^N$ for any open cube $Q(x, r)$ in $\mathbb{R}^N$, since a cube in dimension N is the product of a cube in dimension $N - 1$ and an interval. Then, for any covering of a set $A \subseteq \mathbb{R}^N$ by open cubes $Q(x_i, r_i)$, we have by countable subadditivity $\mu_N(A) \leq \sum_i (2r_i)^N$. Taking the infimum over all coverings, we obtain $\mu_N(A) \leq \mathcal{L}^N(A)$. To justify Borel regularity of μ_N , assume inductively that both factors are Borel-regular. Each measurable rectangle in a covering can be enlarged to a Borel rectangle with the same cost, by taking Borel hulls of its two factors. Intersecting countable unions of these rectangles from coverings whose costs tend to $\mu_N(A)$ gives a Borel hull of A with the same measure when $\mu_N(A) < \infty$; otherwise $\mathbb{R}^N$ is such a hull. Moreover, every open set in $\mathbb{R}^N$ is a countable union of open rectangles, which are μ_N -measurable by Proposition 15. Hence μ_N is a Borel-regular measure. The preceding comparison argument, with expanding open cubes in place of intervals, now gives $\mu_N = \mathcal{L}^N$. $\square$

5.3 Tail integrals and integral inequalities

Considering the product of a measure with the one-dimensional Lebesgue measure, we can express the integral of any nonnegative function as an integral of its tail distribution. In these integrals, we may write dt for $d\mathcal{L}^1(t)$.

Proposition 38 (Tail formula for integrals). *Let X be a set and let μ be a measure on X . Then, for any function $f \in \mathcal{L}_\mu^+$ that is σ -finite with respect to μ , the following equality holds:*

$$\int_X f \, d\mu = \int_0^\infty \mu(\{x \in X \mid f(x) > t\}) \, dt.$$

Proof. The set under the graph is measurable, including when f takes the value ∞ , because

$$A_f = \{(t, x) \in \mathbb{R} \times X \mid 0 < t < f(x)\} = \bigcup_{q \in \mathbb{Q}, q > 0} (0, q) \times \{x \in X \mid f(x) > q\}.$$

Each rectangle is measurable by Proposition 15, and the union is countable. In addition, the set A_f is σ -finite with respect to $\mathcal{L}^1 \otimes \mu$. This follows from Proposition 15 and the following observations: the function f is σ -finite with respect to μ , the Lebesgue measure is σ -finite, and $A_f \subseteq \mathbb{R} \times \{x \in X \mid f(x) \neq 0\}$. Therefore, Theorem 16 yields

$$\mathcal{L}^1 \otimes \mu(A_f) = \int_X \mathcal{L}^1(\{t \in \mathbb{R} \mid 0 < t < f(x)\}) \, d\mu(x) = \int_{\mathbb{R}} \mu(\{x \in X \mid 0 < t < f(x)\}) \, d\mathcal{L}^1(t),$$

and the result follows directly. $\square$

Proposition 38 can be used to derive *Kolmogorov's inequality* for functions in weak Lebesgue spaces. To be more precise, if a set X is endowed with a measure μ , then for any $p \in (1, \infty)$, the Lebesgue space L_μ^p consists of equivalence classes, modulo equality μ -almost everywhere, of μ -measurable functions $f: X \rightarrow \mathbb{C}$ such that $|f|^p \in L_\mu^1$. It is known that L_μ^p is a vector space, on which the mapping $f \mapsto \|f\|_{L_\mu^p} = (\int_X |f|^p \, d\mu)^{1/p}$ is a norm, see [11]. Then, the weak Lebesgue space $L_\mu^{p,\infty}$ consists of equivalence classes, under the same identification, of μ -measurable functions $f: X \rightarrow \mathbb{C}$ such that

$$\|f\|_{L_\mu^{p,\infty}} = \sup_{t>0} t \, \mu(\{x \in X \mid |f(x)| > t\})^{1/p} < \infty.$$

The mapping $f \mapsto \|f\|_{L_\mu^{p,\infty}}$ fails to be a norm, but $L_\mu^{p,\infty}$ is a topological vector space, see [7]. In addition, note that $\|f\|_{L_\mu^{p,\infty}} \leq \|f\|_{L_\mu^p}$, so that $L_\mu^p \subseteq L_\mu^{p,\infty}$. Indeed, if $|f(x)| > t$, then $t^{-p}|f(x)|^p > 1$, and we have

$$t \, \mu(\{x \in X \mid |f(x)| > t\})^{1/p} \leq t \left(\int_X t^{-p} |f(x)|^p \, d\mu(x) \right)^{1/p} = \|f\|_{L_\mu^p}.$$

The value of $\|f\|_{L_\mu^{p,\infty}}$ yields an upper bound on the integral of $|f|$ over measurable sets of finite measure. This is the purpose of the following result, which is a direct application of Proposition 38.

Proposition 39 (Kolmogorov's inequality). *Let X be a set and let μ be a measure on X . Then, for any $p \in (1, \infty)$, any function $f \in \mathcal{L}_\mu$ and any set $A \in \mathcal{M}_\mu$, the following inequality holds:*

$$\int_A |f| \, d\mu \leq \frac{\mu(A)^{1-1/p}}{1-1/p} \|f\|_{L_\mu^{p,\infty}}.$$

Proof. If $\mu(A) = 0$, the integral vanishes. If $\|f\|_{L_\mu^{p,\infty}} = 0$, then $\mu(\{x \in X \mid |f(x)| > 1/n\}) = 0$ for every $n \geq 1$, so that $f = 0$ μ -a.e. and the integral again vanishes. These cases agree with the convention $0 \cdot \infty = 0$. If neither factor is zero and one is infinite, the right-hand side is

infinite and the result is immediate. We may therefore assume that both $\mu(A)$ and $\|f\|_{L_\mu^{p,\infty}}$ are positive and finite. In particular, $\mathbf{1}_A |f| \in \mathcal{L}_\mu^+$ is σ -finite with respect to μ . By Proposition 38, we have

$$\int_A |f| d\mu = \int_0^\infty \mu(\{x \in A \mid |f(x)| > t\}) dt \leq \int_0^\infty \min\{\mu(A), t^{-p} \|f\|_{L_\mu^{p,\infty}}^p\} dt.$$

Comparing the two terms in the minimum, we deduce that

$$\begin{aligned} \int_A |f| d\mu &\leq \mu(A)^{1-1/p} \|f\|_{L_\mu^{p,\infty}} + \|f\|_{L_\mu^{p,\infty}}^p \int_{\mu(A)^{-1/p} \|f\|_{L_\mu^{p,\infty}}}^\infty \frac{dt}{t^p} \\ &= \mu(A)^{1-1/p} \|f\|_{L_\mu^{p,\infty}} + \frac{\|f\|_{L_\mu^{p,\infty}}^p}{p-1} (\mu(A)^{-1/p} \|f\|_{L_\mu^{p,\infty}})^{1-p} = \frac{\mu(A)^{1-1/p}}{1-1/p} \|f\|_{L_\mu^{p,\infty}}, \end{aligned}$$

which proves the inequality. $\square$

We now state and prove Jensen's inequality.

Proposition 40 (Jensen's inequality). *Let X be a set and let μ be a measure on X such that $\mu(X) = 1$. Then, for any $-\infty \leq a < b \leq \infty$, any convex function $\phi: (a, b) \rightarrow \mathbb{R}$, and any μ -integrable function $f: X \rightarrow (a, b)$, it holds that*

$$\phi\left(\int_X f d\mu\right) \leq \int_X \phi(f) d\mu.$$

The integral on the right is well-defined in $(-\infty, \infty]$: it means $\int_X (\phi \circ f)^+ d\mu - \int_X (\phi \circ f)^- d\mu$, and the second integral is finite.

Proof. Put $m = \int_X f d\mu$, which is finite by integrability. If a is finite, then $f - a > 0$ everywhere, so Proposition 9 gives $m - a = \int_X (f - a) d\mu > 0$. Similarly, if b is finite, then $b - m > 0$. Hence, $m \in (a, b)$.

Since ϕ is convex, it is continuous on (a, b) , and its finite left-hand and right-hand derivatives satisfy, for $a < u < v < b$,

$$\phi'_+(u) \leq \frac{\phi(v) - \phi(u)}{v - u} \leq \phi'_-(v) \leq \phi'_+(v).$$

It follows, considering separately $t < m$ and $t > m$, that for every $t \in (a, b)$,

$$\phi(t) \geq \phi(m) + \phi'_+(m)(t - m).$$

Thus, the measurable function $\phi \circ f$ is bounded below by the integrable function $h = \phi(m) + \phi'_+(m)(f - m)$. In particular, $(\phi \circ f)^- \leq h^-$ is integrable, so the asserted extended integral is well-defined. If its positive part has infinite integral, the desired inequality is immediate. Otherwise, $\phi \circ f$ is integrable, and integrating the supporting-line inequality gives

$$\int_X \phi(f) d\mu \geq \int_X h d\mu = \phi(m) + \phi'_+(m) \left(\int_X f d\mu - m \right) = \phi(m),$$

as required. $\square$

In comparison with Proposition 39, *Jensen's inequality* yields another bound on the integral of $|f|$ over measurable sets, in terms of $\|f\|_{L_\mu^p}$. When $0 < \mu(A) < \infty$ and $\|f\|_{L_\mu^p} < \infty$, Proposition 39 and the bound $\|f\|_{L_\mu^{p,\infty}} \leq \|f\|_{L_\mu^p}$ ensure that $|f|$ is integrable on A . Applying

Jensen's inequality to the probability measure $\mu(A)^{-1}\mu|_A$ on A and the convex function $\phi: t \in \mathbb{R} \mapsto |t|^p$, we obtain

$$\left(\frac{1}{\mu(A)} \int_A |f| d\mu \right)^p \leq \frac{1}{\mu(A)} \int_A |f|^p d\mu \quad \text{and so} \quad \int_A |f| d\mu \leq \mu(A)^{1-1/p} \|f\|_{L_\mu^p}.$$

If $\|f\|_{L_\mu^p} = \infty$, the last inequality is immediate.

6 Hausdorff measures, dimensions, and capacities

6.1 Hausdorff measures and contents

The definition of the Lebesgue measure given in Definition 15 can be generalized in several natural ways: for instance, we can consider coverings by arbitrary sets in a metric space instead of cubes in $\mathbb{R}^N$, or we can raise the diameter of the covering sets in the sum to a power $s \geq 0$ instead of N . The diameter of a set A in a metric space (X, d) is defined by $|A| = \sup\{d(x, y), x, y \in A\}$. We use the convention $|\emptyset| = 0$. The previous ideas are put together in the following definition.

Definition 16 (Hausdorff content and Hausdorff measure). Let (X, d) be a metric space, and let $s \in [0, \infty)$. For all $A \subseteq X$ and $\delta \in (0, \infty]$, define

$$\mathcal{H}_\delta^s(A) = \inf \left\{ \sum_{i=1}^{\infty} |A_i|^s, A \subseteq \bigcup_{i=1}^{\infty} A_i \text{ with } |A_i| < \delta \right\},$$

where the infimum is over all coverings of the set $A \subseteq X$ by arbitrary sets $A_i \subseteq X$ with diameter less than δ . (We adopt the convention that $|A_i|^0 = \mathbf{1}_{\{A_i \neq \emptyset\}}$, i.e., we do not count the empty sets in the sum when $s = 0$.) Then, the s -dimensional Hausdorff content of A is defined as the value $\mathcal{H}_\infty^s(A)$, and the s -dimensional Hausdorff measure of A is defined as the monotone limit

$$\mathcal{H}^s(A) = \sup_{\delta > 0} \mathcal{H}_\delta^s(A) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(A).$$

We then have the following analog of Theorem 35 for Hausdorff measures.

Theorem 41 (Metric and Radon properties of Hausdorff measures and contents). Let X be a metric space and let $s \in [0, \infty)$. Then, the following properties hold:

- (i). $\mathcal{H}_\delta^s$ is a measure on X for every $\delta \in (0, \infty]$, so the Hausdorff content $\mathcal{H}_\infty^s$ is a measure on X ;
- (ii). The Hausdorff measure $\mathcal{H}^s$ is a Borel-regular metric measure on X ;
- (iii). for every locally compact subspace $A \subseteq X$ such that $\mathcal{H}^s(A) < \infty$, the restriction $\mathcal{H}^s|_A$ is a Radon measure on A .

Proof. For every $\delta \in (0, \infty]$, we can show as in the beginning of the proof of Theorem 35 that $\mathcal{H}_\delta^s$ is a measure. We can deduce that $\mathcal{H}_\infty^s$ and $\mathcal{H}^s$ are measures by letting $\delta = \infty$ and $\delta \rightarrow 0$, respectively. To check that $\mathcal{H}^s$ is a metric measure, we can use the same argument as in the proof of Theorem 35 to observe that if $d(A, B) > 0$, then $\mathcal{H}_\delta^s(A \cup B) = \mathcal{H}_\delta^s(A) + \mathcal{H}_\delta^s(B)$ for δ small enough, and we conclude by letting $\delta \rightarrow 0$. To show that $\mathcal{H}^s$ is Borel-regular, first note that if $\mathcal{H}^s(A) = \infty$, the Borel set X is a hull of the same measure. We may therefore assume $\mathcal{H}^s(A) < \infty$. Then, for every $m \geq 1$, there is a covering $A \subseteq \bigcup_{i=1}^{\infty} A_{m,i}$ such that $\sum_i |A_{m,i}|^s \leq \mathcal{H}_{1/m}^s(A) + 1/m$ and $|A_{m,i}| < 1/m$. The closures $\overline{A_{m,i}}$ also cover A and have the

same diameter as $A_{m,i}$, so yield the same sum. The set $B = \bigcap_m \bigcup_{i=1}^{\infty} \overline{A_{m,i}}$ is thus a Borel set containing A and satisfying $\mathcal{H}_{1/m}^s(B) \leq \mathcal{H}_{1/m}^s(A) + 1/m$. We conclude that $\mathcal{H}^s(B) = \mathcal{H}^s(A)$ by letting $m \rightarrow \infty$. Finally, if $\mathcal{H}^s(A) < \infty$, then $\mathcal{H}^s(A \cap B(x, r)) < \infty$ for every open ball centered at $x \in A$ with radius $r > 0$. Thus, $\mathcal{H}^s|_A$ is locally finite on A . Moreover, A is separable, hence second-countable, as it can be covered by countably many arbitrarily small sets. If A is also locally compact, we conclude that $\mathcal{H}^s|_A$ is a Radon measure on A , thanks to Theorem 26 and the fact that open sets are σ -compact in second-countable locally compact Hausdorff spaces (see Remark 2). $\square$

We now compare s -dimensional Hausdorff measures and contents for different values of s .

Proposition 42 (Hausdorff measures and contents comparison). *Let X be a metric space, let $t > s \geq 0$, and let $A \subseteq X$. Then, the following implications hold:*

$$\begin{cases} \mathcal{H}_{\infty}^s(A) = 0 & \iff & \mathcal{H}^s(A) = 0 \\ \mathcal{H}^s(A) < \infty & \implies & \mathcal{H}^t(A) = 0 \\ \mathcal{H}^t(A) > 0 & \implies & \mathcal{H}^s(A) = \infty. \end{cases}$$

Proof. Since $\mathcal{H}_{\infty}^s(A) \leq \mathcal{H}^s(A)$, vanishing of the Hausdorff measure implies vanishing of the content. For the converse when $s = 0$, any covering of a nonempty set has at least one nonempty member and therefore has total cost at least one. Thus, $\mathcal{H}_{\infty}^0(A) = 0$ forces $A = \emptyset$, and hence $\mathcal{H}^0(A) = 0$. Now let $s > 0$ and assume that $\mathcal{H}_{\infty}^s(A) = 0$. For every $\varepsilon > 0$, there is a covering $\bigcup_i A_i \supseteq A$ such that $\sum_i |A_i|^s < \varepsilon$. Hence, $|A_i| < \varepsilon^{1/s}$ for every nonempty A_i , so that $\mathcal{H}_{\varepsilon^{1/s}}^s(A) < \varepsilon$. Letting $\varepsilon \rightarrow 0$, we get $\mathcal{H}^s(A) = 0$.

For every $t > s \geq 0$ and every covering $\bigcup_i A_i \supseteq A$ with $|A_i| < \delta$, omitting empty members and using the convention in Definition 16 when $s = 0$, we have

$$\mathcal{H}_{\delta}^t(A) \leq \sum_i |A_i|^t \leq \delta^{t-s} \sum_i |A_i|^s.$$

Taking the infimum over all coverings, we obtain $\mathcal{H}_{\delta}^t(A) \leq \delta^{t-s} \mathcal{H}_{\delta}^s(A)$. If $\mathcal{H}^s(A) < \infty$, then $\mathcal{H}_{\delta}^s(A) \leq \mathcal{H}^s(A)$, and letting $\delta \rightarrow 0$ implies that $\mathcal{H}^t(A) = 0$. The last implication is the contrapositive: if $\mathcal{H}^t(A) > 0$, then $\mathcal{H}^s(A) = \infty$. $\square$

6.2 Hausdorff dimension

Proposition 42 shows that there is a cutoff at which the Hausdorff measure drops from infinity to zero. This cutoff is called the *Hausdorff dimension* of the set, and is formalized as follows. In the definition below, we adopt the convention that $\sup \emptyset = 0$ and $\inf \emptyset = \infty$.

Definition 17 (Hausdorff dimension). *Let X be a metric space. The Hausdorff dimension of a nonempty set $A \subseteq X$ is defined as*

$$\dim_{\text{H}} A = \sup\{s \geq 0 \mid \mathcal{H}^s(A) = \infty\} = \inf\{s \geq 0 \mid \mathcal{H}^s(A) = 0\}.$$

The Hausdorff dimension of the empty set is defined to be $-\infty$.

Recall that a mapping f defined on a metric space (X, d_X) and taking values in a metric space (Y, d_Y) is *bilipschitz* if there exists a constant $\ell \geq 1$ such that $\ell^{-1} d_X(x, x') \leq d_Y(f(x), f(x')) \leq \ell d_X(x, x')$ for all $x, x' \in X$. Hausdorff dimension satisfies the following important properties. Here, countable sets include both finite and countably infinite sets.

Proposition 43 (Hausdorff dimension properties). *Let X and Y be metric spaces. Then, the following properties hold:*

- (i). *Monotonicity: If $A \subseteq B \subseteq X$, then $\dim_H A \leq \dim_H B$;*
- (ii). *Finite or countable stability: If $A_1, A_2, \dots \subseteq X$, then $\dim_H \bigcup_i A_i = \sup_i \dim_H A_i$;*
- (iii). *Bilipschitz invariance: If $A \subseteq X$ and $f: X \rightarrow Y$ is bilipschitz, then $\dim_H f(A) = \dim_H A$;*
- (iv). *Nonempty countable sets: If $A \subseteq X$ is nonempty and countable, then $\dim_H A = 0$;*

Proof. *Monotonicity:* The case $A = \emptyset$ follows from our convention, and the case $\dim_H B = \infty$ is immediate. Otherwise, observe that if $A \subseteq B \subseteq X$ and $s > \dim_H B$, then $\mathcal{H}^s(B) = 0$, and thus $\mathcal{H}^s(A) = 0$ as well. This shows that $\dim_H A \leq s$, and the result follows from letting $s \rightarrow \dim_H B$.

Finite or countable stability: If all the sets are empty, the equality follows from our convention. If $\sup_i \dim_H A_i = \infty$, monotonicity gives the equality. In the remaining case, we may discard empty sets and assume that the supremum is finite and nonnegative. Note by monotonicity that $\sup_i \dim_H A_i \leq \dim_H \bigcup_i A_i$. Conversely, if $s > \sup_i \dim_H A_i$, then $\mathcal{H}^s(A_i) = 0$ for all i . By countable subadditivity, $\mathcal{H}^s(\bigcup_i A_i) = 0$. It follows that $s \geq \dim_H \bigcup_i A_i$, and we conclude by letting $s \rightarrow \sup_i \dim_H A_i$.

Bilipschitz invariance: The case $A = \emptyset$ is immediate. If $f: X \rightarrow Y$ is bilipschitz with constant ℓ and a set $A \subseteq X$ is covered by sets $A_i \subseteq X$ with $|A_i| < \delta$, then its image $f(A)$ is covered by the sets $f(A_i) \subseteq Y$ with $|f(A_i)| \leq \ell|A_i| < \ell\delta$. For every $s \geq 0$, we thus have $\mathcal{H}_{\ell\delta}^s(f(A)) \leq \ell^s \mathcal{H}_\delta^s(A)$. Letting $\delta \rightarrow 0$, we obtain $\mathcal{H}^s(f(A)) \leq \ell^s \mathcal{H}^s(A)$. This shows that $\dim_H f(A) \leq \dim_H A$. Conversely, if $f(A)$ is covered by sets $B_i \subseteq Y$ with $|B_i| < \delta$, then A is covered by the sets $f^{-1}(B_i) \subseteq X$ with $|f^{-1}(B_i)| \leq \ell|B_i| < \ell\delta$. It follows that $\dim_H A \leq \dim_H f(A)$.

Nonempty countable sets: For every $x \in X$ and $s, \delta > 0$, we have $\mathcal{H}_\delta^s(\{x\}) \leq |\{x\}|^s = 0$. Letting $\delta \rightarrow 0$, we obtain $\mathcal{H}^s(\{x\}) = 0$. If A is a nonempty countable set, it follows from the countable subadditivity that $\mathcal{H}^s(A) = 0$. This being true for all $s > 0$, and the dimension of a nonempty set being nonnegative, we conclude that $\dim_H A = 0$. $\square$

In Euclidean spaces, Hausdorff measures and dimension enjoy the following additional properties in relation to Lebesgue measure.

Proposition 44 (Comparison between the N -dimensional Hausdorff and Lebesgue measures). *For every integer $N \geq 1$ and every set $A \subseteq \mathbb{R}^N$, it holds that*

$$2^{-N} \mathcal{L}^N(A) \leq \mathcal{H}^N(A) \leq N^{N/2} \mathcal{L}^N(A). \quad (15)$$

It follows that $\dim_H A \leq N$, with equality if A has positive Lebesgue measure, hence in particular if A is open and nonempty.

Proof. For any covering of the set A by open cubes $Q(x_i, r_i)$ with diameter less than δ , that is, such that $2r_i N^{1/2} < \delta$, we have $\mathcal{H}_\delta^N(A) \leq \sum_i (2r_i N^{1/2})^N = N^{N/2} \sum_i (2r_i)^N$. In view of (13), taking the infimum over all such coverings and then letting $\delta \rightarrow 0$, we obtain $\mathcal{H}^N(A) \leq N^{N/2} \mathcal{L}^N(A)$. Conversely, for any covering of the set A by sets A_i with diameter less than δ , we can cover each set A_i of positive diameter by an open cube with side length $2(1 + \varepsilon)|A_i|$, for every $\varepsilon > 0$. The nonempty sets of diameter zero are singletons; for any $\tau > 0$, cover the i -th such member by an open cube of volume at most $\tau 2^{-i}$. Empty members can be omitted. Hence, $\mathcal{L}^N(A) \leq \sum_i (2(1 + \varepsilon)|A_i|)^N + \tau$. Letting $\tau \rightarrow 0$, taking the infimum over all such coverings, and then letting $\varepsilon \rightarrow 0$ and $\delta \rightarrow 0$, we obtain $\mathcal{L}^N(A) \leq 2^N \mathcal{H}^N(A)$.

If the set A is bounded, it has finite Lebesgue measure, hence finite N -dimensional Hausdorff measure, so that $\dim_{\mathcal{H}} A \leq N$. If A is unbounded, it is covered by countably many bounded sets, so that $\dim_{\mathcal{H}} A \leq N$ by the countable stability of Hausdorff dimension established in Proposition 43. If A has positive Lebesgue measure, then it has positive N -dimensional Hausdorff measure, so that $\dim_{\mathcal{H}} A \geq N$. This holds in particular if A is open and nonempty, as it then contains a closed cube with positive side length, and thus positive Lebesgue measure by Proposition 36. $\square$

For the following remark only, we use *uniqueness of Haar measure*: on a locally compact Hausdorff group, any two nonzero left-translation-invariant Radon measures are proportional on Borel sets, provided they are positive on nonempty open sets. In our framework, their Borel-regular extensions are therefore proportional on all subsets. This theorem is admitted here; see [7].

Remark 3 (Proportionality between the N -dimensional Hausdorff and Lebesgue measures). Beyond the inequalities in (15), it is true that for some constant $c_N > 0$ and for every set $A \subseteq \mathbb{R}^N$,

$$\mathcal{H}^N(A) = c_N \mathcal{L}^N(A).$$

This follows from the fact that both $\mathcal{H}^N$ and $\mathcal{L}^N$ are *Haar measures* on the locally compact additive group $\mathbb{R}^N$, specifically, they are translation-invariant Borel-regular Radon measures on $\mathbb{R}^N$ that are not identically zero. The uniqueness of Haar measures up to a multiplicative constant (see [7]) then implies the existence of the constant c_N .

6.3 Box counting dimensions

Recall that a subset A of a metric space X is *totally bounded* if for every $\varepsilon > 0$, the set A can be covered by a finite number of balls of radius ε . Box counting dimensions are closely related to this concept.

Definition 18 (Box counting dimensions). Let X be a metric space and let $A \subseteq X$. If A is nonempty and totally bounded, the *lower box counting dimension* and the *upper box counting dimension* of A are defined respectively by

$$\underline{\dim}_{\mathcal{B}} A = \liminf_{\varepsilon \rightarrow 0} \frac{\log N_{\varepsilon}(A)}{-\log \varepsilon} \quad \text{and} \quad \overline{\dim}_{\mathcal{B}} A = \limsup_{\varepsilon \rightarrow 0} \frac{\log N_{\varepsilon}(A)}{-\log \varepsilon},$$

where $N_{\varepsilon}(A)$ denotes the minimal number of open balls of radius ε needed to cover A . For $A = \emptyset$, we set both dimensions equal to $-\infty$. Moreover, if A is not totally bounded, then $\underline{\dim}_{\mathcal{B}} A = \overline{\dim}_{\mathcal{B}} A = \infty$. Finally, if $\underline{\dim}_{\mathcal{B}} A$ and $\overline{\dim}_{\mathcal{B}} A$ coincide, their common value is called the *box counting dimension* of A and is denoted by $\dim_{\mathcal{B}} A$.

The box counting dimensions provide a simple way of measuring the size of a set in a metric space, taking into account its geometry. They can be compared to the Hausdorff dimension as follows.

Proposition 45 (Comparison of dimensions). *Let X be a metric space. Then, for every set $A \subseteq X$, we have*

$$\dim_H A \leq \underline{\dim}_B A \leq \overline{\dim}_B A.$$

Proof. The inequality between the lower and upper box counting dimensions is trivial. To prove the inequality between the Hausdorff dimension and the lower box counting dimension, we may assume that A is nonempty and totally bounded. (Otherwise, the Hausdorff dimension is $-\infty$ or the lower box counting dimension is ∞ , so the inequality is trivial.) For $\varepsilon > 0$, the set A can be covered by $N_\varepsilon(A)$ open balls of radius ε . Hence, for every $s \geq 0$, we have $\mathcal{H}_{3\varepsilon}^s(A) \leq N_\varepsilon(A) (2\varepsilon)^s$. If $s > \underline{\dim}_B A$, then

$$\mathcal{H}^s(A) = \lim_{\varepsilon \rightarrow 0} \mathcal{H}_{3\varepsilon}^s(A) \leq \liminf_{\varepsilon \rightarrow 0} N_\varepsilon(A) (2\varepsilon)^s = 0,$$

so that $\dim_H A \leq s$. The result follows from letting $s \rightarrow \underline{\dim}_B A$. $\square$

6.4 Mass distribution principle and Frostman's lemma

There is an important relationship between the Hausdorff dimension of a set and the existence of measures with support in that set. This can be formalized in the two directions. First, if a set supports a measure with certain scaling properties, then we obtain a lower bound on its Hausdorff dimension. This is the purpose of the following *mass distribution principle*.

Lemma 46 (Mass distribution principle). *Let X be a metric space, let $s \in [0, \infty)$, and let $A \subseteq X$. Assume that there exists a measure μ on X such that $\mu(B(x, r)) \leq r^s$ for all $x \in A$ and $0 < r < r_0$, for some $r_0 > 0$. Then, $\mathcal{H}^s(A) \geq \mu(A)$. If in addition $\mu(A) > 0$, then $\dim_H A \geq s$.*

Proof. Assume that $s > 0$. The condition on μ implies that $\mu(\{x\}) = 0$ for all $x \in A$. Let us consider a covering $\bigcup_i A_i \supseteq A$ with $|A_i| < \delta < r_0$, and a real number $\varepsilon > 0$ such that $(1 + \varepsilon)\delta < r_0$. If $x_i \in A \cap A_i$ and $|A_i| > 0$, then $A_i \subseteq B(x_i, (1 + \varepsilon)|A_i|)$, so that $\mu(A_i) \leq (1 + \varepsilon)^s |A_i|^s$. If $|A_i| = 0$, then $A_i = \{x_i\}$, so that $\mu(A_i) = 0 \leq (1 + \varepsilon)^s |A_i|^s$. Therefore, we obtain

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A \cap A_i) \leq \sum_{A \cap A_i \neq \emptyset} \mu(A_i) \leq (1 + \varepsilon)^s \sum_{i=1}^{\infty} |A_i|^s.$$

We conclude by letting $\varepsilon \rightarrow 0$, taking the infimum over all coverings, and finally letting $\delta \rightarrow 0$.

If $s = 0$, then $\mathcal{H}^0(A)$ is the number of points in A . We can assume that A is finite and nonempty, say $A = \{x_1, \dots, x_n\}$. The condition on μ implies that $\mu(\{x_i\}) \leq 1$ for all i , so that $\mu(A) \leq n = \mathcal{H}^0(A)$. $\square$

Example 1 (The middle-third Cantor set). Let $\mathbb{K}_0$ be the closed unit interval $[0, 1]$ and, for all integers $j \geq 1$, let $\mathbb{K}_j$ be the union over all possible j -tuples $(a_1, \dots, a_j) \in \{0, 1\}^j$ of the closed intervals $2a_1 3^{-1} + \dots + 2a_j 3^{-j} + [0, 3^{-j}]$. Each set $\mathbb{K}_j$ is the union of 2^j closed intervals, each of length 3^{-j} . The sets $\mathbb{K}_j$ are nonincreasing in j , nonempty and compact. Hence, their intersection $\mathbb{K} = \bigcap_{j=0}^{\infty} \mathbb{K}_j$ is nonempty and compact, by Cantor's intersection theorem. The set $\mathbb{K}$ is called the *middle-third Cantor set*. Let us show that

$$\dim_H \mathbb{K} = \frac{\log 2}{\log 3}. \quad (16)$$

Let $s = \log 2 / \log 3$ and $\delta > 0$. Choose an integer $j \geq 0$ such that $3^{-j} < \delta$. Since $\mathbb{K} \subseteq \mathbb{K}_j$, we have $\mathcal{H}_\delta^s(\mathbb{K}) \leq 2^j (3^{-j})^s = 1$. Letting $\delta \rightarrow 0$, we get $\mathcal{H}^s(\mathbb{K}) \leq 1$, so that $\dim_H \mathbb{K} \leq s$. To show

the reverse inequality, consider for each j the probability measure on $\mathbb{R}$ derived from the restriction to the set $\mathbb{K}_j$ of the Lebesgue measure $\mathcal{L}^1$. To be precise, for any set $A \subseteq \mathbb{R}$, define

$$\mu_j(A) = (3/2)^j \mathcal{L}^1(\mathbb{K}_j \cap A).$$

By Theorem 31, there exists a subsequence $(\mu_{j_k})_{k \geq 1}$ converging weakly to a Radon measure μ on $\mathbb{R}$. The open-set and compact-set inequalities of Theorem 32 apply to this subsequence. For any integer $j_0 \geq 0$, the set $\mathbb{R} \setminus \mathbb{K}_{j_0}$ is open and has μ_{j_0} -measure zero for all $j \geq j_0$. It follows from Theorem 32 that $\mu(\mathbb{R} \setminus \mathbb{K}_{j_0}) = 0$. Theorem 3 then ensures that $\mu(\mathbb{R} \setminus \mathbb{K}) = 0$. Moreover, the closed unit interval $[0, 1]$ is compact and has μ_j -measure one for all $j \geq j_0$. The compact-set inequality gives $\mu([0, 1]) \geq 1$, whereas the open-set inequality applied to $\mathbb{R}$ gives $\mu(\mathbb{R}) \leq \liminf_{k \rightarrow \infty} \mu_{j_k}(\mathbb{R}) = 1$. Hence $\mu([0, 1]) = \mu(\mathbb{R}) = 1$. Thus, μ is a probability measure supported in $\mathbb{K}$. In addition, consider $x \in \mathbb{R}$ and $r \in (0, 1/2]$, and choose an integer $j_0 \geq 0$ such that $3^{-j_0-1} < 2r \leq 3^{-j_0}$. The open ball $B(x, r) = (x - r, x + r)$ intersects at most one interval among the union $\mathbb{K}_{j_0}$, hence at most 2^{j-j_0} intervals among the union $\mathbb{K}_j$ if $j \geq j_0$. Then, $\mu_j(B(x, r)) \leq (3/2)^j 2^{j-j_0} 3^{-j} = 2^{-j_0} = (3^{-j_0})^s < (6r)^s$, hence

$$\mu(B(x, r)) \leq (6r)^s = 2^{s+1} r^s \leq 4 r^s, \quad (17)$$

as a result of Theorem 32. By the mass distribution principle (Lemma 46), we conclude that $\mathcal{H}^s(\mathbb{K}) \geq \mu(\mathbb{K})/4 = 1/4 > 0$, which implies that $\dim_H \mathbb{K} \geq s$. This finishes the proof of (16).

The mass distribution principle (Lemma 46) provides one direction of the relationship between measures and Hausdorff dimension. The other direction is subtler. It states that if a set has positive s -dimensional Hausdorff measure, then it supports a measure that scales as in the assumption of the mass distribution principle. The existence of such a measure, formalized in *Frostman's lemma*, is crucial in various applications, particularly in fractal geometry and the study of sets with specific dimensional properties. Though this result is valid in more general settings, we only state and prove it for F_σ -sets, that is, for countable unions of closed sets in the Euclidean space $\mathbb{R}^N$.

Lemma 47 (Frostman's lemma). *Let $N \geq 1$ and $s \geq 0$, and an F_σ -set $A \subseteq \mathbb{R}^N$. If $\mathcal{H}^s(A) > 0$, then there exists a nonzero finite Radon measure μ on $\mathbb{R}^N$ with compact support $\text{supp } \mu \subseteq A$, such that for every $x \in \mathbb{R}^N$ and $r > 0$,*

$$\mu(B(x, r)) \leq r^s.$$

Proof. We use the Euclidean distance throughout. If $s = 0$, the assumption implies that A is nonempty. A unit point mass at any point of A is a nonzero finite Radon measure with compact support in A and satisfies $\mu(B(x, r)) \leq 1 = r^0$ for every x and $r > 0$. We may therefore assume that $s > 0$.

An F_σ -set in $\mathbb{R}^N$ can be written as a countable union of compact sets, each contained in a closed cube of the form $k + [0, 1]^N$ for some $k \in \mathbb{Z}^N$. By countable subadditivity of Hausdorff measures, if the F_σ -set has positive measure, then at least one of the compact sets must have positive measure. We can thus assume that A is compact and contained in some cube of the previous form. Up to translating A , we can assume that it is contained in the cube $Q_0 = [0, 1]^N$. For every integer $m \geq 1$, this cube is the union of 2^{Nm} disjoint cubes with side length 2^{-m} , specifically, the cubes $Q_{m,k} = \prod_{n=1}^N [k_n 2^{-m}, (k_n + 1) 2^{-m}]$, for $k_n \in \{0, 1, \dots, 2^m - 1\}$, where $\rangle$ means that the right endpoint is excluded unless it is equal to 1. We put $\mathcal{Q}_0 = \{Q_0\}$ and define $\mathcal{Q}_m$ as the collection of all cubes $Q_{m,k}$, for $k \in \{0, 1, \dots, 2^m - 1\}^N$. If $m \geq 1$, any cube $Q \in \mathcal{Q}_m$ is contained in a unique cube that belongs to $\mathcal{Q}_{m-1}$, is denoted by $Q^\uparrow$, and is called the *parent* of Q . Hence Q is a *child* of $Q^\uparrow$.

We define by induction a family of weights α_Q , for $Q \in \mathcal{Q}_m$ and $m \geq 0$. Set $\alpha_{Q_0} = \mathcal{H}_\infty^s(A)$, which is positive by Proposition 42 and at most $|Q_0|^s = N^{s/2}$. For each cube $P \in \mathcal{Q}_{m-1}$, put

$$S_P = \sum_{R \uparrow = P} \mathcal{H}_\infty^s(A \cap R).$$

For each child Q of P , define

$$\alpha_Q = \begin{cases} \frac{\mathcal{H}_\infty^s(A \cap Q)}{S_P} \alpha_P, & S_P > 0, \\ 0, & S_P = 0. \end{cases}$$

We check simultaneously by induction that $\alpha_Q \leq \mathcal{H}_\infty^s(A \cap Q)$ and that $\sum_{Q \uparrow = P} \alpha_Q = \alpha_P$. The bound holds at the root. If it holds for P , countable subadditivity gives

$$\alpha_P \leq \mathcal{H}_\infty^s(A \cap P) \leq S_P.$$

If $S_P = 0$, then $\alpha_P = 0$, and both assertions follow from the zero-weight convention. If $S_P > 0$, the formula for α_Q proves the bound for each child, and summing that formula proves conservation of mass. Thus, for every $m' \geq m$ and $Q \in \mathcal{Q}_m$, the weights of all level- m' descendants of Q sum to α_Q .

This enables us to define a sequence of measures μ_m on $\mathbb{R}^N$, with support contained in Q_0 and with constant total mass $\mu_m(Q_0) = \mathcal{H}_\infty^s(A)$, by the formula

$$\mu_m = \sum_{Q \in \mathcal{Q}_m} \alpha_Q \frac{\mathcal{L}^N|_Q}{\mathcal{L}^N(Q)}.$$

These measures are Radon measures on $\mathbb{R}^N$, because of Theorem 26 and the fact that they are Borel-regular and locally finite, as finite linear combinations of restrictions of the Lebesgue measure to Borel sets. By Theorems 31 and 32, there exists a subsequence $(\mu_{m_k})_{k \geq 1}$ and a Radon measure μ on $\mathbb{R}^N$ such that for every open set $U \subseteq \mathbb{R}^N$ and for every compact set $K \subseteq \mathbb{R}^N$,

$$\mu(U) \leq \liminf_{k \rightarrow \infty} \mu_{m_k}(U) \quad \text{and} \quad \mu(K) \geq \limsup_{k \rightarrow \infty} \mu_{m_k}(K). \quad (18)$$

Since $\mathbb{R}^N \setminus Q_0$ is open, we have in particular $\mu(\mathbb{R}^N \setminus Q_0) = 0$, so that the support of μ satisfies $\text{supp } \mu \subseteq Q_0$. Moreover, since Q_0 is compact and $\mathbb{R}^N$ is open, we also have $\mathcal{H}_\infty^s(A) \leq \mu(Q_0) \leq \mu(\mathbb{R}^N) \leq \mathcal{H}_\infty^s(A)$. Hence, the measure μ has total mass equal to $\mathcal{H}_\infty^s(A)$, which is both nonzero and finite.

The dyadic partitions are nested and their boundaries have zero Lebesgue measure. Conservation of the weights therefore gives $\mu_{m'}(Q) = \alpha_Q = \mu_m(Q)$ for every $Q \in \mathcal{Q}_m$ and $m' \geq m$. In particular, if $Q \cap A = \emptyset$, then $\alpha_Q = 0$ and $\mu_{m'}(Q) = 0$. Define the open set

$$U_m = \{x \in \mathbb{R}^N \mid d(\{x\}, A) > \sqrt{N} 2^{-m}\}.$$

Since every cube in $\mathcal{Q}_m$ has Euclidean diameter $\sqrt{N} 2^{-m}$, we have

$$U_m \cap Q_0 \subseteq \bigcup_{\substack{Q \in \mathcal{Q}_m \\ Q \cap A = \emptyset}} Q.$$

Every $\mu_{m'}$ is concentrated on Q_0 , so $\mu_{m'}(U_m) = 0$ for $m' \geq m$. Applying the open-set inequality in (18) along the subsequence yields $\mu(U_m) = 0$. As A is compact, $\mathbb{R}^N \setminus A = \bigcup_{m \geq 0} U_m$, whence $\mu(\mathbb{R}^N \setminus A) = 0$ and $\text{supp } \mu \subseteq A$.

Finally, let $0 < r < 1$ and choose $m \geq 0$ such that $2^{-m-1} \leq r < 2^{-m}$. Each coordinate projection of the Euclidean ball $B(x, r)$ is an interval of length $2r < 2 \cdot 2^{-m}$, so the ball intersects at most 3^N cubes of $\mathcal{Q}_m$. For all $m' \geq m$, we obtain

$$\begin{aligned} \mu_{m'}(B(x, r)) &\leq \sum_{\substack{Q \in \mathcal{Q}_m \\ Q \cap B(x, r) \neq \emptyset}} \mu_{m'}(Q) = \sum_{\substack{Q \in \mathcal{Q}_m \\ Q \cap B(x, r) \neq \emptyset}} \alpha_Q \\ &\leq \sum_{\substack{Q \in \mathcal{Q}_m \\ Q \cap B(x, r) \neq \emptyset}} \mathcal{H}_\infty^s(A \cap Q) \leq 3^N N^{s/2} 2^{-sm} \leq 3^N (2\sqrt{N})^s r^s. \end{aligned}$$

The open-set inequality in (18) gives $\mu(B(x, r)) \leq 3^N (2\sqrt{N})^s r^s$. For $r \geq 1$, the total-mass bound gives

$$\mu(B(x, r)) \leq \mu(\mathbb{R}^N) = \mathcal{H}_\infty^s(A) \leq N^{s/2} \leq 3^N (2\sqrt{N})^s r^s.$$

Dividing μ by $3^N (2\sqrt{N})^s$ proves the asserted estimate for every $r > 0$, while preserving its positive finite mass and compact support in A . $\square$

6.5 Capacities

The scaling behavior of measures in the mass distribution principle and in Frostman's lemma is conveniently encoded by *dimensional energies*, which in turn lead to the notion of *capacity*.

Definition 19 (Energies and capacities). Let (X, d) be a separable and locally compact metric space and let $s \in (0, \infty)$. The s -dimensional energy of a Borel measure μ on X is defined by the formula

$$\mathcal{E}_s(\mu) = \int_X \int_X \frac{\mu(dx) \mu(dx')}{d(x, x')^s}.$$

The s -dimensional capacity of a set $A \subseteq X$ is defined by the relationship

$$\text{Cap}_s(A)^{-1} = \inf_{\substack{\mu(X)=1 \\ \text{supp } \mu \subseteq A}} \mathcal{E}_s(\mu),$$

where the infimum is over all Borel probability measures μ on X with support contained in A . The energy kernel is assigned the value ∞ when $x = x'$. We use the conventions $\inf \emptyset = \infty$, $1/\infty = 0$ and $1/0 = \infty$; in particular, $\text{Cap}_s(\emptyset) = 0$.

Observe that a set has zero s -dimensional capacity precisely when it supports no Borel probability measure with finite s -dimensional energy. The following theorem relates capacities to Hausdorff measures and yields a capacity characterization of Hausdorff dimension. The second part of this statement relies on Frostman's lemma and thus inherits the assumptions of Lemma 47, although the conclusion holds in greater generality. In the statement below, we adopt the convention that $\sup \emptyset = 0$ and $\inf \emptyset = \infty$.

Theorem 48 (Relationship between capacities and Hausdorff measures and dimension). *Let X be a separable and locally compact metric space, let $A \subseteq X$ and let $s \in (0, \infty)$. The following properties hold:*

- (i). *If $\text{Cap}_s(A) > 0$, then $\mathcal{H}^s(A) > 0$;*
- (ii). *If $X = \mathbb{R}^N$ and A is an F_σ -set such that $\mathcal{H}^s(A) > 0$, then $\text{Cap}_t(A) > 0$ for every $t \in (0, s)$;*
- (iii). *If $X = \mathbb{R}^N$ and A is a nonempty F_σ -set, then*

$$\dim_H A = \sup\{s > 0 \mid \text{Cap}_s(A) > 0\} = \inf\{s > 0 \mid \text{Cap}_s(A) = 0\}.$$

Proof. Let us first assume that $\text{Cap}_s(A) > 0$. Then, there exists a Borel probability measure μ on X such that $\text{supp } \mu \subseteq A$ and $\mathcal{E}_s(\mu) < \infty$. Since X is a separable metric space, it has a countable base; the complement of $\text{supp } \mu$ is therefore a countable union of μ -null open sets. Thus, $\mu(X \setminus \text{supp } \mu) = 0$, and $\mu(A \cap E) = \mu(E)$ for every set $E \subseteq X$. For every integer $i \geq 1$, define the set

$$A_i = \left\{ x \in X \mid \int_X \frac{\mu(dx')}{d(x, x')^s} \leq i \right\}.$$

The sets A_i are μ -measurable, because the function $x \mapsto \int_X \frac{\mu(dx')}{d(x, x')^s}$ is μ -measurable, as a consequence of Fubini's theorem (Theorem 18). The union over all $i \geq 1$ of the sets A_i is equal to the set of all $x \in X$ such that $\int_X \frac{\mu(dx')}{d(x, x')^s} < \infty$. This set has μ -measure equal to one, because $\mathcal{E}_s(\mu) < \infty$. Then,

$$\sup_{i \geq 1} \mu(A_i) = \mu\left(\left\{ x \in X \mid \int_X \frac{\mu(dx')}{d(x, x')^s} < \infty \right\}\right) = 1,$$

as a result of Theorem 3, so that $\mu(A_i) > 0$ for some $i \geq 1$. Moreover, for all $x, x' \in X$ and $r > 0$, note that $\mathbf{1}_{B(x, r)}(x') d(x, x')^s < r^s$. Consequently, for every $x \in A \cap A_i$ and $r > 0$, we have

$$\mu(B(x, r)) = \int_X \mathbf{1}_{B(x, r)}(x') \mu(dx') \leq \int_X \frac{r^s}{d(x, x')^s} \mu(dx') \leq i r^s.$$

By the mass distribution principle (Lemma 46), we obtain $\mathcal{H}^s(A) \geq \mathcal{H}^s(A \cap A_i) \geq \mu(A \cap A_i)/i > 0$.

Let us now assume that $X = \mathbb{R}^N$ and that A is an F_σ -set with $\mathcal{H}^s(A) > 0$. By Frostman's lemma (Lemma 47), there exists a nonzero finite Radon measure μ on $\mathbb{R}^N$ with compact support $K = \text{supp } \mu \subseteq A$ and such that $\mu(B(x, r)) \leq r^s$ for every $x \in \mathbb{R}^N$ and $r > 0$. For all $t \in (0, s)$ and $x \in K$, we then have

$$\begin{aligned} \int_K \frac{\mu(dx')}{\|x' - x\|^t} &= \sum_{j=-r}^{\infty} \int_K \mathbf{1}_{\{2^{-j-1} \leq \|x' - x\| < 2^{-j}\}} \frac{\mu(dx')}{\|x' - x\|^t} \\ &\leq \sum_{j=-r}^{\infty} 2^{t(j+1)} \mu(B(x, 2^{-j})) \leq \sum_{j=-r}^{\infty} 2^{t(j+1)} 2^{-js} < \infty, \end{aligned}$$

where r is an integer such that $|K| < 2^r$. The first equality follows from the observation that $\mu(\{x\}) = 0$ and Theorem 7. Since the upper bound is independent of $x \in K$, we deduce that $\mathcal{E}_t(\mu) < \infty$. The measure $\mu/\mu(\mathbb{R}^N)$ is thus a Borel probability measure with support contained in A and with finite t -dimensional energy, so that $\text{Cap}_t(A) > 0$. Finally, put $d = \dim_H A$. If $u > d$, then $\mathcal{H}^u(A) = 0$, so the first assertion implies $\text{Cap}_u(A) = 0$. If $0 < u < d$, choose v with $u < v < d$; Proposition 42 gives $\mathcal{H}^v(A) > 0$, and the second assertion gives $\text{Cap}_u(A) > 0$. These two observations prove both formulas for d , including $d = 0$ with the stated empty-supremum convention. $\square$

7 Hölder continuous functions and their graphs

7.1 General results

Recall that the *graph* of a function f defined on an interval $I \subseteq \mathbb{R}$ and valued in $\mathbb{R}$ is the set

$$\text{gph } f = \{(x, f(x)), x \in I\} \subseteq \mathbb{R}^2.$$

If the function f is continuous and the interval I is compact, then the graph $\text{gph } f$ is a compact subset of $\mathbb{R}^2$, because it is the image of the compact interval I under the continuous

mapping $x \mapsto (x, f(x))$. Additional properties on $\text{gph } f$ can be derived from the Hölder continuity of f , which we now define.

Definition 20 (Hölder continuity, Lipschitz continuity, and modulus of continuity). Let $\alpha \in (0, 1]$. A real-valued function f defined on a nonempty interval $I \subseteq \mathbb{R}$ is called α -Hölder continuous if there exists a constant $C > 0$ such that for all $x, x' \in I$,

$$|f(x) - f(x')| \leq C |x - x'|^\alpha. \quad (19)$$

If $\alpha = 1$, then the function f is also called *Lipschitz continuous*. More generally, if η is a continuous function defined on the interval $[0, \infty)$ such that $\eta(0) = 0$, then the function f is said to have *modulus of continuity* η if there exists a constant $C > 0$ such that for all $x, x' \in I$,

$$|f(x) - f(x')| \leq C \eta(|x - x'|).$$

We assume in Definition 20 that $\alpha \in (0, 1]$, since the only functions that satisfy (19) for $\alpha > 1$ are the constant ones. Indeed, for all $x, x' \in I$, and for $x_{n,k} = x + k(x' - x)/n$, the triangle inequality gives

$$\begin{aligned} |f(x) - f(x')| &= \left| \sum_{k=0}^{n-1} (f(x_{n,k}) - f(x_{n,k+1})) \right| \leq \sum_{k=0}^{n-1} |f(x_{n,k}) - f(x_{n,k+1})| \\ &\leq C \sum_{k=0}^{n-1} |x_{n,k} - x_{n,k+1}|^\alpha = C |x' - x|^\alpha n^{1-\alpha} \xrightarrow{n \rightarrow \infty} 0, \end{aligned}$$

which implies that the function f is constant on the interval I . The following result derives an upper bound on the upper box counting dimension of the graph $\text{gph } f$ from the Hölder continuity of f .

Proposition 49 (Graph dimension upper bound for Hölder continuous functions). Let $\alpha \in (0, 1]$ and let f be a real-valued α -Hölder continuous function defined on a nonempty bounded interval $I \subseteq \mathbb{R}$. Then, the upper box counting dimension of the graph $\text{gph } f$ satisfies

$$\overline{\dim}_B \text{gph } f \leq 2 - \alpha.$$

Hence, the Hausdorff dimension of the graph $\text{gph } f$ satisfies

$$\dim_H \text{gph } f \leq 2 - \alpha.$$

Proof. If I is a singleton, its graph has dimension zero and the conclusion is immediate. Otherwise put $L = |I| > 0$. For $0 < \varepsilon < \min\{L, 1\}$, partition I into at most $L/\varepsilon + 2$ intervals of length at most ε . On each interval the oscillation of f is at most $C\varepsilon^\alpha$, so its graph is covered by at most $C\varepsilon^{\alpha-1} + 2$ closed squares of side length ε . Each square is contained in an open Euclidean ball of radius 2ε . Consequently,

$$N_{2\varepsilon}(\text{gph } f) \leq (L/\varepsilon + 2)(C\varepsilon^{\alpha-1} + 2) \leq 3L(C + 2)\varepsilon^{\alpha-2}.$$

The upper box dimension bound follows from Definition 18, and the Hausdorff bound from Proposition 45. $\square$

The next result explains how Hausdorff dimension behaves under Hölder continuous mappings. Recall from Proposition 43 the bilipschitz invariance of Hausdorff dimension.

Proposition 50 (Hausdorff dimension upper bound on Hölder continuous images). *Let $\alpha \in (0, 1]$ and let f be a real-valued α -Hölder continuous function defined on a nonempty interval $I \subseteq \mathbb{R}$. Then, for any set $A \subseteq I$, it holds that*

$$\dim_{\mathbb{H}} f(A) \leq \frac{1}{\alpha} \dim_{\mathbb{H}} A.$$

Proof. If $A = \emptyset$, the conclusion follows from $f(A) = \emptyset$. Assume henceforth that $A \neq \emptyset$. Let us consider two real numbers $s > \dim_{\mathbb{H}} A$ and $\delta > 0$, and a covering $\bigcup_i A_i \supseteq A$ by sets $A_i \subseteq I$ (intersect any covering with I) such that $|A_i| < \delta$ for every $i \geq 1$. Then, the set $f(A)$ is covered by the sets $f(A_i)$, each of which has diameter at most $C|A_i|^\alpha$ by (19). Therefore, $\sum_i |f(A_i)|^{s/\alpha} \leq C^{s/\alpha} \sum_i |A_i|^s$. Taking the infimum over all such coverings, we obtain $\mathcal{H}_{C\delta^\alpha}^{s/\alpha}(f(A)) \leq C^{s/\alpha} \mathcal{H}_\delta^s(A)$. Letting $\delta \rightarrow 0$, we deduce that $\mathcal{H}^{s/\alpha}(f(A)) \leq C^{s/\alpha} \mathcal{H}^s(A) = 0$, which implies that $\dim_{\mathbb{H}} f(A) \leq s/\alpha$. The result follows by letting $s \rightarrow \dim_{\mathbb{H}} A$. $\square$

Example 2 (The Cantor function). Let $\mathbb{K}$ be the middle-third Cantor set and let μ be the associated probability measure, as defined in Example 1. Define the *Cantor function* $f: [0, 1] \rightarrow [0, 1]$ by the formula $f(x) = \mu([0, x])$. It follows from (17) that for $0 \leq x < x' \leq 1$ and $\varepsilon > 0$,

$$|f(x) - f(x')| = \mu((x, x']) \leq \mu((x, x' + \varepsilon)) \leq 2|x' + \varepsilon - x|^s \xrightarrow{\varepsilon \rightarrow 0} 2|x' - x|^s,$$

where $s = \log 2 / \log 3$. Therefore, the function f is nondecreasing and s -Hölder continuous on the unit interval $[0, 1]$. Moreover, it is clear that $f(\mathbb{K}) \subseteq [0, 1]$. To prove the reverse inclusion, consider $y \in [0, 1]$, write $y = \sum_{j=1}^{\infty} a_j 2^{-j}$ with $a_j \in \{0, 1\}$, and define $x = \sum_{j=1}^{\infty} 2a_j 3^{-j}$. We assert that $x \in \mathbb{K}$ and $f(x) = y$. Indeed, for every integer $j_0 \geq 1$, we have

$$x = \sum_{j=1}^{\infty} 2a_j 3^{-j} = \sum_{j \leq j_0} 2a_j 3^{-j} + \sum_{j > j_0} 2a_j 3^{-j} \in 2a_1 3^{-1} + \cdots + 2a_{j_0} 3^{-j_0} + [0, 3^{-j_0}] \subseteq \mathbb{K}_{j_0}.$$

In addition, the interval that contains x among the union $\mathbb{K}_{j_0}$ is the $(2^{j_0-1}a_1 + \cdots + a_{j_0} + 1)$ -th, considering the order of the intervals within $[0, 1]$. Since single points have μ_{j_0} -measure equal to zero, it follows that

$$(2^{j_0-1}a_1 + \cdots + a_{j_0})2^{-j_0} \leq \mu_{j_0}((0, x)) = \mu_{j_0}([0, x]) \leq (2^{j_0-1}a_1 + \cdots + a_{j_0} + 1)2^{-j_0}.$$

The interval $(0, x)$ is open and the interval $[0, x]$ is compact. Hence, letting $j_0 \rightarrow \infty$ along a subsequence to which Theorem 32 applies, we obtain

$$\mu((0, x)) \leq \lim_{j_0 \rightarrow \infty} (2^{j_0-1}a_1 + \cdots + a_{j_0} + 1)2^{-j_0} = y \quad \text{and} \quad \mu([0, x]) \geq \lim_{j_0 \rightarrow \infty} (2^{j_0-1}a_1 + \cdots + a_{j_0})2^{-j_0} = y.$$

By (17), single points also have μ -measure equal to zero. We conclude that $f(x) = y$, and then that $f(\mathbb{K}) = [0, 1]$. Therefore, $\dim_{\mathbb{H}} f(\mathbb{K}) = 1 = (1/s) \dim_{\mathbb{H}} \mathbb{K}$, which shows that the upper bound given by Proposition 50 is attained in this case, proving that the proposition is sharp.

A converse to Proposition 49 requires a *lower Hölder estimate*.

Definition 21 (Lower Hölder estimate). Let $\alpha > 0$. A real-valued function f defined on a nonempty interval $I \subseteq \mathbb{R}$ satisfies a *lower Hölder estimate* of order α if there exists a constant $C > 0$ such that for every nonempty subinterval $J \subseteq I$, there exist $x, x' \in J$ such that

$$|f(x) - f(x')| \geq C|J|^\alpha. \quad (20)$$

Under this condition, we can derive the following converse result to Proposition 49.

Proposition 51 (Graph box dimension lower bound under a lower Hölder estimate). *Let $\alpha > 0$ and let f be a real-valued continuous function defined on a bounded interval of positive length $I \subseteq \mathbb{R}$ and satisfying a lower Hölder estimate of order α . Then, the lower box counting dimension of the graph $\text{gph } f$ satisfies*

$$\underline{\dim}_B \text{gph } f \geq 2 - \alpha.$$

Proof. If the graph is not totally bounded, its lower box dimension is infinite and the conclusion is immediate. Otherwise choose a closed subinterval $[a, b] \subseteq I$ of length $L = b - a > 0$. For $0 < \varepsilon < L/2$, it contains $M = \lfloor L/\varepsilon \rfloor \geq L/(2\varepsilon)$ consecutive closed intervals of length ε . On each such interval J , continuity and (20) imply that $f(J)$ contains an interval of length at least $C\varepsilon^\alpha$. If N open balls of radius ε cover the graph, let n_J count those meeting its graph over J . Their vertical projections cover $f(J)$, so $2\varepsilon n_J \geq C\varepsilon^\alpha$. The horizontal projection of any one ball has length 2ε and meets at most four of the consecutive intervals J . Summing gives

$$4N \geq \sum_J n_J \geq \frac{MC}{2} \varepsilon^{\alpha-1} \geq \frac{CL}{4} \varepsilon^{\alpha-2}.$$

Therefore $N_\varepsilon(\text{gph } f) \geq (CL/16)\varepsilon^{-(2-\alpha)}$, which yields the conclusion. $\square$

Since the inequalities in Proposition 45 are not equalities in general, the previous result does not imply a lower bound on the Hausdorff dimension of the graph. However, for functions that are α -Hölder continuous and satisfy a lower Hölder estimate of the same order $\alpha \in (0, 1)$, the graph has Hausdorff dimension larger than one, as the next proposition shows.

Proposition 52 (Hausdorff dimension lower bound under a lower Hölder estimate). *Let $\alpha \in (0, 1)$ and let f be a real-valued α -Hölder continuous function defined on a bounded interval of positive length $I \subseteq \mathbb{R}$ and satisfying a lower Hölder estimate of the same order α . Then, the Hausdorff dimension of the graph $\text{gph } f$ satisfies*

$$\dim_H \text{gph } f > 1.$$

Proof. It suffices to prove the bound for the restriction of f to a closed subinterval of positive length: its graph is contained in $\text{gph } f$, and both Hölder estimates are preserved. We may thus assume that I is a compact interval of positive length.

By (19) and (20), there exist two constants $0 < C_1 \leq C_2$ such that the oscillation of the function f over any interval $J \subseteq I$ is between $C_1|J|^\alpha$ and $C_2|J|^\alpha$. For every $\lambda \in \mathbb{R} \setminus \{0\}$, the graph of the function λf is the image of the graph of f under the bilipschitz mapping $(x, y) \mapsto (x, \lambda y)$. Hence, by Proposition 43, these two graphs have the same Hausdorff dimension. Therefore, up to dividing f by C_2 , we can assume that $C_2 = 1$. Up to reducing C_1 , we can also assume that $C_1 = 4\sqrt{2}b^{-\alpha}$ for some integer $b \geq 2$.

Let μ be the measure obtained by lifting the Lebesgue measure on the interval I to the graph of the function f . More precisely, for every set $A \subseteq \mathbb{R}^2$, define

$$\mu(A) = \mathcal{L}^1(\{x \in I \mid (x, f(x)) \in A\}).$$

We aim at applying the mass distribution principle (Lemma 46) to the measure μ , to derive the required lower bound on the Hausdorff dimension of $\text{gph } f$. Note that $\mu(\text{gph } f) = \mathcal{L}^1(I) > 0$.

Consider a square Q of side length less than one, and assume that Q intersects $\text{gph } f$ and that the projection J of Q onto the first coordinate axis is contained in I . (The square Q can include all, none, or some of its boundary sides.) Then, there exist $x_0, x_1, x_2 \in J$ such that $(x_0, f(x_0)) \in Q$ and $|f(x_1) - f(x_2)| \geq C_1|J|^\alpha$, in view of (20). By the triangle inequality, we

have $|f(x_i) - f(x_0)| \geq C_1 |J|^\alpha / 2$ for some $i \in \{1, 2\}$. Then, for any point $(u, v) \in Q$, the triangle inequality also implies that

$$\begin{aligned} \|(x_i, f(x_i)) - (u, v)\| &\geq \|(x_i, f(x_i)) - (x_0, f(x_0))\| - \|(x_0, f(x_0)) - (u, v)\| \\ &\geq \frac{C_1}{2} |J|^\alpha - |Q| \geq \frac{C_1}{2} |J|^\alpha - |J| \sqrt{2}. \end{aligned}$$

Now, divide J into b disjoint subintervals $J_1, \dots, J_b$ with equal length, and assume that $x_i \in J_k$ for some $k \in \{1, \dots, b\}$. If Q intersects the graph of f over the interval J_k , then there exists a point $(x'_0, f(x'_0)) \in Q$ with $x'_0 \in J_k$. Using (19) with constant $C_2 = 1$, the previous inequalities with $(u, v) = (x'_0, f(x'_0))$, and the fact that $\alpha \in (0, 1)$ and $C_1 = 4\sqrt{2} b^{-\alpha}$, we obtain

$$\begin{aligned} \frac{4\sqrt{2} b^{-\alpha}}{2} |J|^\alpha - |J| \sqrt{2} &\leq \|(x_i, f(x_i)) - (x'_0, f(x'_0))\| = ((x_i - x'_0)^2 + (f(x_i) - f(x'_0))^2)^{1/2} \\ &\leq (|J_k|^2 + |J_k|^{2\alpha})^{1/2} \leq \left(\frac{|J|}{b}\right)^\alpha \sqrt{2}, \end{aligned}$$

which is equivalent to $2b^{-\alpha} |J|^\alpha - |J| \leq (|J|/b)^\alpha$, that is, $|J| \geq b^{-\alpha/(1-\alpha)}$. The upshot is that

$$|J| < b^{-\alpha/(1-\alpha)} \implies \exists k \in \{1, \dots, b\} \quad Q \cap \{(x, f(x)), x \in J_k\} = \emptyset.$$

We apply the same reasoning to the other intervals J_j with $j \neq k$. Each of these intervals contains a subinterval of length $|J|/b^2$ over which the graph of f does not intersect Q , provided that the condition $2b^{-\alpha} (|J|/b)^\alpha - |J| \leq (|J|/b^2)^\alpha$ does not hold, meaning that $|J| < b^{-2\alpha/(1-\alpha)}$. Iterating this argument, we deduce that if $|J| < b^{-n\alpha/(1-\alpha)}$ for some integer $n \geq 1$, then there exist disjoint subintervals of J , one of length $|J|/b$, then $b-1$ of length $|J|/b^2$, and so on, up to $(b-1)^{n-1}$ of length $|J|/b^n$, over which the graph of f does not intersect Q . As a consequence,

$$\mu(Q) \leq |J| - \left(\frac{|J|}{b} + (b-1) \frac{|J|}{b^2} + \dots + (b-1)^{n-1} \frac{|J|}{b^n} \right) = |J| - \frac{|J|}{b} \frac{1 - (1-1/b)^n}{1 - (1-1/b)} = \left(1 - \frac{1}{b}\right)^n |J|.$$

Choosing $\varepsilon > 0$ such that $1-1/b \leq b^{-\varepsilon\alpha/(1-\alpha)}$ and assuming that $b^{-(n+1)\alpha/(1-\alpha)} \leq |J| < b^{-n\alpha/(1-\alpha)}$ for some integer $n \geq 1$, we conclude that

$$\mu(Q) \leq \left(1 - \frac{1}{b}\right)^n |J| \leq b^{-n\varepsilon\alpha/(1-\alpha)} |J| \leq b^{\varepsilon\alpha/(1-\alpha)} |J|^{1+\varepsilon}.$$

Finally, for any point $z \in \text{gph } f$ and for any positive radius r such that $r < |I|/2$ and $r < b^{-\alpha/(1-\alpha)}/2$, the set $\text{gph } f \cap B(z, r)$ is contained in a square Q of side length $2r$ that intersects $\text{gph } f$ and whose projection onto the first coordinate axis is contained in I . Therefore,

$$\mu(B(z, r)) \leq \mu(Q) \leq 2^{1+\varepsilon} b^{\varepsilon\alpha/(1-\alpha)} r^{1+\varepsilon}.$$

Hence, the mass distribution principle (Lemma 46) implies that $\dim_H \text{gph } f \geq 1 + \varepsilon > 1$. $\square$

7.2 The Weierstrass function

The *Weierstrass function* is a classic example of a continuous function that is nowhere differentiable. It is also Hölder continuous and satisfies a lower Hölder estimate, which makes it a perfect illustration of the previous results. This function is defined as follows.

Definition 22 (Weierstrass function). The Weierstrass function with parameters a real number $\alpha \in (0, 1]$ and an integer $b > 1$ is defined for all $x \in \mathbb{R}$ by the formula

$$W_{\alpha,b}(x) = \sum_{n=1}^{\infty} b^{-n\alpha} \cos(b^n x).$$

7.2.1 Hölder continuity and modulus of continuity

We begin with the Hölder continuity and the modulus of continuity of the Weierstrass function. In the following statement, $\log_b^+$ denotes the positive part of the base b logarithm, defined by the formula $\log_b^+ t = \max\{0, \log t / \log b\}$ for every real number $t > 0$.

Proposition 53 (Hölder continuity and modulus of continuity of the Weierstrass function). *Let $\alpha \in (0, 1]$ and let $b > 1$ be an integer. Then, the following properties hold:*

- If $\alpha < 1$, then the Weierstrass function $W_{\alpha,b}$ is α -Hölder continuous;
- If $\alpha = 1$, then $W_{\alpha,b}$ has modulus of continuity $\eta(0) = 0$ and $\eta(t) = t(1 + \log_b^+(1/t))$ for $t > 0$.

Proof. For all $x, x' \in \mathbb{R}$ such that $|x - x'| \geq 1$, we have

$$|W_{\alpha,b}(x) - W_{\alpha,b}(x')| \leq \sum_{n=1}^{\infty} b^{-n\alpha} |\cos(b^n x) - \cos(b^n x')| \leq 2 \sum_{n=1}^{\infty} b^{-n\alpha} \leq \frac{2b^{-\alpha}}{1 - b^{-\alpha}} |x - x'|^{\alpha},$$

which yields the desired estimates. Now, consider $x, x' \in \mathbb{R}$ such that $0 < |x - x'| < 1$, and then the integer $N \geq 1$ such that $b^{-N} \leq |x - x'| < b^{1-N}$. The same remainder estimate gives

$$\left| \sum_{n=N+1}^{\infty} b^{-n\alpha} \cos(b^n x) - \sum_{n=N+1}^{\infty} b^{-n\alpha} \cos(b^n x') \right| \leq 2 \sum_{n=N+1}^{\infty} b^{-n\alpha} = \frac{2b^{-(N+1)\alpha}}{1 - b^{-\alpha}} \leq \frac{2b^{-\alpha}}{1 - b^{-\alpha}} |x - x'|^{\alpha}.$$

In addition, the mean value inequality implies that $|\cos(b^n x) - \cos(b^n x')| \leq b^n |x - x'|$ for all $n \geq 1$. Hence, for $\alpha < 1$, the N -th partial sum satisfies

$$\begin{aligned} \left| \sum_{n=1}^N b^{-n\alpha} \cos(b^n x) - \sum_{n=1}^N b^{-n\alpha} \cos(b^n x') \right| &\leq \sum_{n=1}^N b^{-n\alpha} b^n |x - x'| = b^{1-\alpha} \frac{b^{(1-\alpha)N} - 1}{b^{1-\alpha} - 1} |x - x'| \\ &\leq \frac{b^{2(1-\alpha)}}{b^{1-\alpha} - 1} (b^{1-N})^{\alpha-1} |x - x'| \leq \frac{b^{2(1-\alpha)}}{b^{1-\alpha} - 1} |x - x'|^{\alpha}. \end{aligned}$$

This shows that the Weierstrass function $W_{\alpha,b}$ is α -Hölder continuous. For $\alpha = 1$, we have

$$\begin{aligned} \left| \sum_{n=1}^N b^{-n\alpha} \cos(b^n x) - \sum_{n=1}^N b^{-n\alpha} \cos(b^n x') \right| &\leq \sum_{n=1}^N b^{-n} b^n |x - x'| \\ &= N |x - x'| \leq |x - x'| \left(1 + \log_b \frac{1}{|x - x'|} \right). \end{aligned}$$

We thus obtain the desired modulus of continuity on the Weierstrass function $W_{\alpha,b}$. $\square$

7.2.2 Nowhere differentiability

We now show that the Weierstrass function is nowhere differentiable. To this purpose, we place the Weierstrass function within the general framework of *Fourier series* of the form

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx}, \quad (21)$$

where $(a_n)_{n \in \mathbb{Z}}$ is a sequence of complex numbers such that $\sum_n |a_n| < \infty$. The partial sums of the above series are continuous and 2π -periodic functions, and they converge uniformly on $\mathbb{R}$ to a complex-valued function f which is then continuous and 2π -periodic as well. In addition, the Fourier series satisfy the following basic properties. The first one can be interpreted as an inversion of (21).

Proposition 54 (Basic properties of Fourier series). *Let $(a_n)_{n \in \mathbb{Z}}$ and $(b_n)_{n \in \mathbb{Z}}$ be sequences of complex numbers such that $\sum_n |a_n| < \infty$ and $\sum_n |b_n| < \infty$, and let f and g be the corresponding Fourier series as in (21). Then:*

(i). *For every integer $n \in \mathbb{Z}$, it holds that*

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx;$$

(ii). *For every real number $x_0 \in \mathbb{R}$, the function $x \mapsto f(x - x_0)$ is the Fourier series associated with the sequence $(e^{-inx_0} a_n)_{n \in \mathbb{Z}}$;*

(iii). *The convolution $f * g$ is the Fourier series associated with the sequence $(a_n b_n)_{n \in \mathbb{Z}}$, that is,*

$$f * g(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) g(y) dy = \sum_{n=-\infty}^{\infty} a_n b_n e^{inx}.$$

Proof. The first property can be proven using Fubini's theorem (Theorem 18). To be specific, for any $n_0 \in \mathbb{Z}$, we have

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) e^{-in_0 x} dx &= \int_{-\pi}^{\pi} \sum_{n=-\infty}^{\infty} a_n e^{inx} e^{-in_0 x} dx \\ &= \sum_{n=-\infty}^{\infty} a_n \int_{-\pi}^{\pi} e^{i(n-n_0)x} dx = \sum_{n=-\infty}^{\infty} 2\pi \mathbf{1}_{\{n=n_0\}} a_n = 2\pi a_{n_0}. \end{aligned}$$

In order to prove the second property, we just write

$$f(x - x_0) = \sum_{n=-\infty}^{\infty} a_n e^{in(x-x_0)} = \sum_{n=-\infty}^{\infty} e^{-inx_0} a_n e^{inx}.$$

Finally, using Fubini's theorem again, we obtain

$$\begin{aligned} f * g(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x - y) g(y) dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m=-\infty}^{\infty} a_m e^{im(x-y)} \sum_{n=-\infty}^{\infty} b_n e^{iny} dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_m e^{imx} b_n e^{i(n-m)y} dy = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_m b_n e^{imx} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n-m)y} dy \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_m b_n e^{imx} \mathbf{1}_{\{m=n\}} = \sum_{n=-\infty}^{\infty} a_n b_n e^{inx}, \end{aligned}$$

which proves the third property. $\square$

As mentioned above, the Weierstrass function $W_{\alpha,b}$ can be expressed as a Fourier series. To be specific, let sgn denote the sign function, defined by $\text{sgn } x = x/|x|$ for $x \neq 0$, and by $\text{sgn } 0 = 0$. Then,

$$W_{\alpha,b}(x) = -\frac{1}{2} + \sum_{n=-\infty}^{\infty} \frac{b^{-|n|\alpha}}{2} e^{i \text{sgn}(n) b^{|n|} x}. \quad (22)$$

More precisely, the Weierstrass function is a *lacunary* Fourier series, in the following sense.

Definition 23 (Lacunary sequences and lacunary Fourier series). (i). A sequence of positive integers $(n_k)_{k \geq 1}$ is *lacunary* if there exists a real number $q > 1$ such that $n_{k+1} \geq q n_k$ for all integers $k \geq 1$.
(ii). A Fourier series as in (21) is *lacunary* if there exists a lacunary sequence $(n_k)_{k \geq 1}$ such that $a_n = 0$ for all integers $n \in \mathbb{Z}$ that are not of the form $\pm n_k$ with $k \geq 1$.

The first property of Proposition 54 can be strengthened into the following statement, which is particularly relevant for lacunary Fourier series.

Lemma 55 (Inverse formula with trigonometric polynomials). *Let $(n_k)_{k \in \mathbb{Z}}$ be a strictly increasing sequence of integers, let $(a_n)_{n \in \mathbb{Z}}$ be a sequence of complex numbers such that $\{n \in \mathbb{Z} \mid a_n \neq 0\} \subseteq \{n_k, k \in \mathbb{Z}\}$ and $\sum_n |a_n| < \infty$, and let f be the corresponding Fourier series as in (21). Then, for every $k \in \mathbb{Z}$ and every trigonometric polynomial $T_k(x) = \sum_{n=-\ell_k}^{\ell_k} c_n e^{inx}$ with $c_0 \neq 0$ and degree at most $\ell_k < \min\{n_{k+1} - n_k, n_k - n_{k-1}\}$, it holds that*

$$a_{n_k} = \frac{\int_{-\pi}^{\pi} f(x) T_k(x) e^{-in_k x} dx}{\int_{-\pi}^{\pi} T_k(x) dx}.$$

Proof. For every $k \in \mathbb{Z}$, define $g(x) = T_k(-x) e^{in_k x}$. We have

$$f(x) = \sum_{n=-\infty}^{\infty} a_n e^{inx} \quad \text{and} \quad g(x) = \sum_{n=-\ell_k}^{\ell_k} c_n e^{i(n_k-n)x} = \sum_{n=n_k-\ell_k}^{n_k+\ell_k} c_{n_k-n} e^{inx}.$$

The third property of Proposition 54 gives

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) T_k(x) e^{-in_k x} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) g(-x) dx = \sum_{n=-\infty}^{\infty} a_n c_{n_k-n} \mathbf{1}_{\{n_k-\ell_k \leq n \leq n_k+\ell_k\}} = a_{n_k} c_0.$$

The last equality follows from the following observations: if $n_k - \ell_k \leq n \leq n_k + \ell_k$, then $n_{k-1} < n < n_{k+1}$; hence, if $a_n \neq 0$, then $n = n_k$ and $c_{n_k-n} = c_0$. To conclude, we observe that $\frac{1}{2\pi} \int_{-\pi}^{\pi} T_k(x) dx = c_0$, as a consequence of the first property of Proposition 54. $\square$

Lemma 55 can be used to estimate the coefficients of lacunary Fourier series. For such series, it is convenient to apply this lemma to the trigonometric polynomials obtained as the squares of the Fejér kernels. For every integer $N \geq 1$, the Fejér kernel F_N is defined by the formula

$$F_N(x) = \sum_{n=-N+1}^{N-1} \left(1 - \frac{|n|}{N}\right) e^{inx} = 1 + 2 \sum_{n=1}^{N-1} \left(1 - \frac{n}{N}\right) \cos(nx),$$

for all $x \in \mathbb{R}$, and satisfies the following properties.

Proposition 56 (Basic estimates on the Fejér kernel). *For every integer $N \geq 1$, the Fejér kernel F_N satisfies the following properties (the bound involving $1/x^2$ is interpreted as ∞ at $x = 0$):*

- (i). $F_N(x) = \frac{1}{N} \left(\frac{\sin(Nx/2)}{\sin(x/2)} \right)^2$ for $x \notin 2\pi\mathbb{Z}$, with continuous extension $F_N(2\pi k) = N$;
- (ii). $F_N(x) \leq \min \left\{ N, \frac{\pi^2}{Nx^2} \right\}$ for all $x \in [-\pi, \pi]$;
- (iii). $F_N(x) \geq \frac{4N}{\pi^2}$ for all $x \in [-\pi/N, \pi/N]$.

Proof. To prove the first property, we first observe that for all $x \in \mathbb{R}$,

$$N F_N(x) = \sum_{|n| < N} (N - |n|) e^{inx} = \sum_{|n| < N} \sum_{|n| \leq m < N} e^{inx} = \sum_{m=0}^{N-1} \sum_{|n| \leq m} e^{inx}. \quad (23)$$

The formula at $x \in 2\pi\mathbb{Z}$ follows directly from the defining finite sum; its continuity also gives the claimed extension. For every integer $m \geq 0$ and $x \notin 2\pi\mathbb{Z}$, we have

$$\sum_{|n| \leq m} e^{inx} = \sum_{p=0}^{2m} e^{i(p-m)x} = e^{-imx} \sum_{p=0}^{2m} e^{ipx} = e^{-imx} \frac{1 - e^{i(2m+1)x}}{1 - e^{ix}} = \frac{e^{-imx} - e^{i(m+1)x}}{1 - e^{ix}}.$$

Therefore, we obtain

$$\begin{aligned} (1 - e^{ix}) N F_N(x) &= \sum_{m=0}^{N-1} (e^{-imx} - e^{i(m+1)x}) = \sum_{m=0}^{N-1} e^{-imx} - \sum_{m=0}^{N-1} e^{i(m+1)x} = \frac{1 - e^{-iNx}}{1 - e^{-ix}} - e^{ix} \frac{1 - e^{iNx}}{1 - e^{ix}} \\ &= \frac{1 - e^{-iNx}}{1 - e^{-ix}} - \frac{e^{iNx} - 1}{1 - e^{-ix}} = \frac{2 - e^{iNx} - e^{-iNx}}{1 - e^{-ix}} = -\frac{(e^{iNx/2} - e^{-iNx/2})^2}{1 - e^{-ix}}. \end{aligned}$$

This concludes the proof of the first property, since we now have

$$F_N(x) = -\frac{(e^{iNx/2} - e^{-iNx/2})^2}{N(1 - e^{ix})(1 - e^{-ix})} = \frac{(e^{iNx/2} - e^{-iNx/2})^2}{N(e^{ix/2} - e^{-ix/2})^2} = \frac{1}{N} \left(\frac{\sin(Nx/2)}{\sin(x/2)} \right)^2.$$

To prove the second property, we infer from (23) and the triangle inequality that for all $x \in \mathbb{R}$,

$$|F_N(x)| \leq \frac{1}{N} \sum_{m=0}^{N-1} \sum_{|n| \leq m} |e^{inx}| = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{|n| \leq m} 1 = \frac{1}{N} \sum_{m=0}^{N-1} (2m+1) = \frac{1}{N} \left(2 \frac{(N-1)N}{2} + N \right) = N.$$

By the first property, F_N is a nonnegative function, so we have $F_N(x) \leq N$. Moreover, $\sin t \geq 2t/\pi$ for all $t \in [0, \pi/2]$. Since F_N is an even function, the first property then implies that for all $x \in [-\pi, \pi] \setminus \{0\}$,

$$F_N(x) = \frac{1}{N} \left(\frac{\sin(Nx/2)}{\sin(x/2)} \right)^2 \leq \frac{1}{N} \frac{1}{(\sin(x/2))^2} \leq \frac{1}{N} \left(\frac{\pi}{x} \right)^2 = \frac{\pi^2}{Nx^2}.$$

Finally, we use the fact that $2t/\pi \leq \sin t \leq t$ for all $t \in [0, \pi/2]$, the first property, and the fact that F_N is an even function to obtain that for all $x \in [-\pi/N, \pi/N] \setminus \{0\}$,

$$F_N(x) = \frac{1}{N} \left(\frac{\sin(Nx/2)}{\sin(x/2)} \right)^2 \geq \frac{1}{N} \left(\frac{Nx/\pi}{x/2} \right)^2 = \frac{4N}{\pi^2},$$

which proves the third property. The identity $F_N(0) = N$ gives the remaining bounds at zero. $\square$

As mentioned above, we shall apply Lemma 55 to the trigonometric polynomial $T_N = F_N^2$, that is, the square of the Fejér kernel F_N . We shall also use the following estimates on T_N .

Lemma 57 (Estimates on the square of the Fejér Kernel). *For each integer $N \geq 1$ and each $\delta \in (0, \pi)$, we have the following bounds for $T_N = F_N^2$:*

$$\int_{-\pi}^{\pi} |x| T_N(x) dx \leq \frac{100}{N} \int_{-\pi}^{\pi} T_N(x) dx \quad \text{and} \quad \int_{\delta < |x| < \pi} T_N(x) dx \leq \frac{63}{\delta^3 N^3} \int_{-\pi}^{\pi} T_N(x) dx.$$

Proof. Proposition 56 implies on the one hand that

$$\begin{aligned} \int_{-\pi}^{\pi} |x| T_N(x) dx &\leq \int_{-\pi}^{\pi} |x| \min \left\{ N^2, \frac{\pi^4}{N^2 x^4} \right\} dx \leq 2 \int_0^{1/N} x N^2 dx + 2 \int_{1/N}^{\pi} \frac{\pi^4}{N^2 x^3} dx \\ &\leq 2N^2 \left[\frac{x^2}{2} \right]_0^{1/N} + \frac{2\pi^4}{N^2} \left[-\frac{1}{2x^2} \right]_{1/N}^{\pi} = 1 + \pi^4 - \frac{\pi^2}{N^2} \leq 1 + \pi^4, \end{aligned}$$

and on the other hand that

$$\int_{-\pi}^{\pi} T_N(x) dx \geq \int_{-\pi/N}^{\pi/N} \frac{16N^2}{\pi^4} dx = \frac{32N}{\pi^3}, \quad (24)$$

from which the first desired estimate follows, specifically,

$$\int_{-\pi}^{\pi} |x| T_N(x) dx \leq 1 + \pi^4 = \frac{(1 + \pi^4)\pi^3}{32N} \frac{32N}{\pi^3} \leq \frac{100}{N} \int_{-\pi}^{\pi} T_N(x) dx.$$

In addition, Proposition 56 also implies together with (24) that

$$\begin{aligned} \int_{\delta < |x| < \pi} T_N(x) dx &\leq \int_{\delta < |x| < \pi} \frac{\pi^4}{N^2 x^4} dx = \frac{2\pi^4}{N^2} \left[-\frac{1}{3x^3} \right]_{\delta}^{\pi} = \frac{2\pi^4}{N^2} \left(-\frac{1}{3\pi^3} + \frac{1}{3\delta^3} \right) \\ &\leq \frac{2\pi^4}{3\delta^3 N^2} = \frac{\pi^7}{48\delta^3 N^3} \frac{32N}{\pi^3} \leq \frac{63}{\delta^3 N^3} \int_{-\pi}^{\pi} T_N(x) dx, \end{aligned}$$

which proves the second desired estimate. $\square$

We are now in position to establish the following connection between the coefficients of a lacunary Fourier series and its differentiability properties.

Theorem 58 (Differentiability and decay of coefficients for lacunary Fourier series). *Let f be a lacunary Fourier series with coefficients denoted by a_n for $n \in \mathbb{Z}$. Assume that the function f is differentiable at some point $x_0 \in \mathbb{R}$. Then, the coefficients satisfy $n a_n \rightarrow 0$ as $|n| \rightarrow \infty$.*

Proof. Define two coefficients $c_0 = f(x_0) + i f'(x_0)$ and $c_1 = -i f'(x_0)$, and then the function

$$g(x) = f(x_0 + x) - c_0 - c_1 e^{ix} = (a_0 - c_0) + (a_1 e^{ix_0} - c_1) e^{ix} + \sum_{n \in \mathbb{Z} \setminus \{0,1\}} a_n e^{inx_0} e^{inx},$$

which satisfies $g(0) = g'(0) = 0$. Hence, for every $\varepsilon > 0$, there exists $\delta \in (0, \pi)$ such that $|g(x)| \leq \varepsilon|x|$ for all $x \in [-\delta, \delta]$. In addition, g is bounded by some constant $C > 0$ on the interval $[-\pi, \pi]$, because f is bounded. Moreover, since f is a lacunary Fourier series, there exists a lacunary sequence of integers $(n_k)_{k \geq 1}$ such that $a_n = 0$ for all integers $n \in \mathbb{Z}$ that are not of the form $\pm n_k$ with $k \geq 1$. Also, there exists a real number $q > 1$ such that $n_{k+1} \geq q n_k$ for all integers $k \geq 1$, which implies, for $k \geq 2$, that

$$\min\{n_{k+1} - n_k, n_k - n_{k-1}\} \geq \min\{(q-1)n_k, (1-q^{-1})n_k\} = (1-q^{-1})n_k \xrightarrow[k \rightarrow \infty]{} \infty. \quad (25)$$

Let $(m_k)_{k \in \mathbb{Z}}$ be a strictly increasing enumeration of the distinct elements of $\{0, 1\} \cup \{\pm n_j, j \geq 1\}$. This set contains the support of the Fourier coefficients b_n of g . Adding the two frequencies 0 and 1 affects only finitely many gaps, so (25) gives, for all sufficiently large $|k|$,

$$\min\{m_{k+1} - m_k, m_k - m_{k-1}\} \geq (1 - q^{-1})|m_k| \xrightarrow{|k| \rightarrow \infty} \infty.$$

Choosing $|k|$ large enough and $N \geq 1$ maximal such that $2N < (1 - q^{-1})|m_k|$, and applying Lemma 55 to the function g and the trigonometric polynomial $T_N = F_N^2$, where F_N is the Fejér kernel, we obtain

$$b_{m_k} = \frac{\int_{-\pi}^{\pi} g(x) T_N(x) e^{-im_k x} dx}{\int_{-\pi}^{\pi} T_N(x) dx}.$$

Thanks to Lemma 57 and the aforementioned bounds on g , the numerator can be estimated as follows:

$$\begin{aligned} \left| \int_{-\pi}^{\pi} g(x) T_N(x) e^{-im_k x} dx \right| &\leq \int_{|x| \leq \delta} \varepsilon |x| T_N(x) dx + \int_{\delta < |x| < \pi} C T_N(x) dx \\ &\leq \left(\frac{100\varepsilon}{N} + \frac{63C}{\delta^3 N^3} \right) \int_{-\pi}^{\pi} T_N(x) dx. \end{aligned}$$

The maximality of N implies that $4N \geq 2(N+1) \geq (1 - q^{-1})|m_k|$. Therefore,

$$\begin{aligned} |m_k b_{m_k}| &\leq \frac{2(N+1)}{1 - q^{-1}} \left(\frac{100\varepsilon}{N} + \frac{63C}{\delta^3 N^3} \right) \leq \frac{4}{1 - q^{-1}} \left(100\varepsilon + \frac{63C}{\delta^3 N^2} \right) \\ &\leq \frac{4}{1 - q^{-1}} \left(100\varepsilon + \frac{1008C}{\delta^3 (1 - q^{-1})^2 |m_k|^2} \right). \end{aligned}$$

Since the Fourier coefficients of g that are not indexed by an integer m_k are equal to zero, and since the coefficients of f and g have the same modulus, except maybe those indexed by 0 and 1, we get

$$|n a_n| = |n b_n| \leq \frac{4}{1 - q^{-1}} \left(100\varepsilon + \frac{1008C}{\delta^3 (1 - q^{-1})^2 n^2} \right)$$

for all $|n|$ large enough. This directly implies that

$$\limsup_{|n| \rightarrow \infty} |n a_n| \leq \frac{400\varepsilon}{1 - q^{-1}} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

and the theorem follows. $\square$

The following corollary follows from Theorem 58, since the Weierstrass series is lacunary.

Corollary 59 (Nowhere differentiability of the Weierstrass function). *For every $\alpha \in (0, 1]$ and every integer $b > 1$, the Weierstrass function $W_{\alpha, b}$ is nowhere differentiable.*

Proof. Recall from (22) that the Weierstrass function $W_{\alpha, b}$ is a lacunary Fourier series whose coefficients a_n satisfy $a_n = b^{-|k|^\alpha}/2$ when $n = \operatorname{sgn}(k) b^{|k|} \neq 0$. In particular, for a positive integer n of the previous form, we have $a_n = n^{-\alpha}/2$, so that $n a_n = n^{1-\alpha}/2$. Since $\alpha \leq 1$, we have $n a_n \not\rightarrow 0$ as $n \rightarrow \infty$. The conclusion then follows from Theorem 58. $\square$

7.2.3 Lower Hölder estimate

Recall from Proposition 53 that the Weierstrass function $W_{\alpha,b}$ is α -Hölder continuous when $0 < \alpha < 1$. We now prove the corresponding lower Hölder estimate of order α .

Proposition 60 (Lower Hölder estimate of the Weierstrass function). *Let $\alpha \in (0, 1)$ and let $b > 1$ be an integer. Then, the Weierstrass function $W_{\alpha,b}$ satisfies a lower Hölder estimate of order α in the interval $[-\pi, \pi]$.*

Proof. Assume that $W_{\alpha,b}$ does not satisfy a lower Hölder estimate of order α in $[-\pi, \pi]$. Any interval in $[-\pi, \pi]$ can be divided into seven equal subintervals of length at most 1. Applying the negation of Definition 21 with constant $\varepsilon^3/7^\alpha$ and choosing one such subinterval shows that, for every $\varepsilon \in (0, 2^{-1/2})$, there exists an interval $I \subseteq [-\pi, \pi]$ such that $|I| \leq 1$ and for all $x, x' \in I$,

$$|W_{\alpha,b}(x) - W_{\alpha,b}(x')| < \varepsilon^3 |I|^\alpha.$$

Let x_0 be the center of the interval I , and define $J = I - x_0 = \{x - x_0, x \in I\}$, so that J is an interval centered at zero. Define also

$$f(x) = W_{\alpha,b}(x_0 + x) - W_{\alpha,b}(x_0) = -\frac{1}{2} - W_{\alpha,b}(x_0) + \sum_{n=-\infty}^{\infty} \frac{b^{-|n|\alpha}}{2} e^{i \operatorname{sgn}(n)b^{|n|x_0}} e^{i \operatorname{sgn}(n)b^{|n|x}}.$$

The latter expression follows from (22). The nonconstant frequencies of f are lacunary; its full frequency set is contained in $\{0\} \cup \{\pm b^j, j \geq 1\}$. Moreover, $f(0) = 0$ and

$$|f(x) - f(x')| < \varepsilon^3 |J|^\alpha$$

for all $x, x' \in J$. We may take ε arbitrarily small; since $|J| \leq 1$, the frequency chosen below then tends to infinity and the maximal integer N is at least 1. Now, define $r = \varepsilon|J|/2 > 0$ and observe that if $2\varepsilon r \leq |x| < r/\varepsilon$, then $x \in J$ and

$$|f(x)| = |f(x) - f(0)| < \varepsilon^3 |J|^\alpha = \varepsilon^3 (2r/\varepsilon)^\alpha = \varepsilon^{3-2\alpha} (2\varepsilon r)^\alpha \leq \varepsilon |x|^\alpha.$$

Also, recall from Proposition 53 that $W_{\alpha,b}$ is α -Hölder continuous. Hence, there exists a constant $C > 0$ such that $|f(x)| \leq C|x|^\alpha$ for all $x \in \mathbb{R}$. Note from the proof of Proposition 53 that the constant C depends on α and b only. Choose two integers $k, N \geq 1$ such that $b^{-k} \leq r < b^{1-k}$ and

$$2N < (1 - b^{-1})b^k = \min\{b^{k+1} - b^k, b^k - b^{k-1}\} \leq 2(N+1).$$

Applying Lemma 55 to the function f and the trigonometric polynomial $T_N = F_N^2$, where F_N is again the Fejér kernel, we obtain

$$\frac{b^{-k\alpha}}{2} e^{ib^k x_0} = \frac{\int_{-\pi}^{\pi} f(x) T_N(x) e^{-ib^k x} dx}{\int_{-\pi}^{\pi} T_N(x) dx}. \quad (26)$$

The numerator of the right-hand side has modulus bounded above by

$$\int_{|x| < 2\varepsilon r} C|x|^\alpha T_N(x) dx + \int_{2\varepsilon r \leq |x| < r/\varepsilon} \varepsilon |x|^\alpha T_N(x) dx + \int_{r/\varepsilon \leq |x| < \pi} C|x|^\alpha T_N(x) dx.$$

Thanks to Proposition 56, the first integral satisfies

$$\int_{|x| < 2\varepsilon r} C|x|^\alpha T_N(x) dx \leq CN^2 (2\varepsilon r)^\alpha (4\varepsilon r) = 2^{2+\alpha} CN^2 \varepsilon^{1+\alpha} r^{1+\alpha}$$

$$< 2^{2+\alpha} C N^2 \varepsilon^{1+\alpha} \left(\frac{b-1}{2N} \right)^{1+\alpha} = 2(b-1)^{1+\alpha} C N^{1-\alpha} \varepsilon^{1+\alpha},$$

the third integral satisfies

$$\begin{aligned} \int_{r/\varepsilon \leq |x| < \pi} C|x|^\alpha T_N(x) dx &\leq \int_{r/\varepsilon \leq |x| < \pi} C|x|^\alpha \frac{\pi^4}{N^2 x^4} dx = 2\pi^4 C N^{-2} \int_{r/\varepsilon}^{\pi} x^{\alpha-4} dx \\ &\leq 2\pi^4 C N^{-2} \left[\frac{x^{\alpha-3}}{\alpha-3} \right]_{r/\varepsilon}^{\pi} \leq \frac{2\pi^4 C}{3-\alpha} \varepsilon^{3-\alpha} N^{-2} r^{\alpha-3} \\ &\leq \frac{2\pi^4 C}{3-\alpha} \varepsilon^{3-\alpha} N^{-2} \left(\frac{4N}{1-b^{-1}} \right)^{3-\alpha} = \frac{2^{7-2\alpha} \pi^4 C}{(3-\alpha)(1-b^{-1})^{3-\alpha}} N^{1-\alpha} \varepsilon^{3-\alpha}, \end{aligned}$$

and the second integral satisfies

$$\begin{aligned} \int_{2\varepsilon r \leq |x| < r/\varepsilon} \varepsilon |x|^\alpha T_N(x) dx &\leq \int_{|x| < r} \varepsilon |x|^\alpha N^2 dx + \int_{|x| \geq r} \varepsilon |x|^\alpha \frac{\pi^4}{N^2 x^4} dx \\ &= 2\varepsilon N^2 \int_0^r x^\alpha dx + 2 \frac{\varepsilon \pi^4}{N^2} \int_r^\infty x^{\alpha-4} dx \\ &= 2\varepsilon N^2 \left[\frac{x^{\alpha+1}}{\alpha+1} \right]_0^r + 2 \frac{\varepsilon \pi^4}{N^2} \left[\frac{x^{\alpha-3}}{\alpha-3} \right]_r^\infty \\ &\leq \frac{2}{\alpha+1} N^2 \varepsilon r^{\alpha+1} + \frac{2\pi^4}{3-\alpha} N^{-2} \varepsilon r^{\alpha-3} \\ &\leq \frac{2}{\alpha+1} N^2 \varepsilon \left(\frac{b-1}{2N} \right)^{\alpha+1} + \frac{2\pi^4}{3-\alpha} N^{-2} \varepsilon \left(\frac{4N}{1-b^{-1}} \right)^{3-\alpha} \\ &= \left(\frac{2^{-\alpha}(b-1)^{\alpha+1}}{\alpha+1} + \frac{2^{7-2\alpha} \pi^4}{(3-\alpha)(1-b^{-1})^{3-\alpha}} \right) N^{1-\alpha} \varepsilon. \end{aligned}$$

We deduce that the numerator of the right-hand side of (26) has modulus bounded above by $N^{1-\alpha} \varepsilon$ times a constant that depends on α and b only. Moreover, the denominator is bounded below by $32N/\pi^3$, as shown in (24). Hence, there exists a constant $D > 0$ depending on α and b only such that

$$\frac{b^{-k\alpha}}{2} = \left| \frac{b^{-k\alpha}}{2} e^{ib^k x_0} \right| = \frac{\left| \int_{-\pi}^{\pi} f(x) T_N(x) e^{-ib^k x} dx \right|}{\int_{-\pi}^{\pi} T_N(x) dx} \leq D N^{-\alpha} \varepsilon \leq \left(\frac{4}{1-b^{-1}} \right)^\alpha D b^{-k\alpha} \varepsilon.$$

This yields a contradiction if ε is chosen small enough. $\square$

7.2.4 Graph dimension

Recall from Proposition 49 that the upper box counting dimension, and thus the Hausdorff dimension, of the graph of an α -Hölder continuous function defined on an interval are bounded above by $2 - \alpha$. Conversely, recall from Proposition 51 that the lower box counting dimension of the graph is bounded below by $2 - \alpha$ when the function satisfies a lower Hölder estimate of order α . These properties are satisfied by the Weierstrass function when $0 < \alpha < 1$, as shown by Propositions 53 and 60. Using also Proposition 52 which discusses Hausdorff dimension, we have the following result.

Corollary 61 (Graph dimension of the Weierstrass function). *Let $\alpha \in (0, 1)$ and let $b > 1$ be an integer. Then, the graph of the Weierstrass function $W_{\alpha,b}$ restricted to the interval $[-\pi, \pi]$ satisfies*

$$\dim_B \text{gph } W_{\alpha,b} = 2 - \alpha \quad \text{and} \quad 1 < \dim_H \text{gph } W_{\alpha,b} \leq 2 - \alpha.$$

At the endpoint $\alpha = 1$, the Hausdorff and both box counting dimensions of the same graph are equal to 1. Indeed, the logarithmic modulus in Proposition 53 implies β -Hölder continuity on this interval for every $0 < \beta < 1$. Proposition 49 therefore bounds the upper box dimension by $2 - \beta$, and letting $\beta \rightarrow 1$ gives the upper bound 1. Projection onto the first coordinate does not increase diameters of covering sets, so it cannot increase Hausdorff dimension. Since the projected graph is the interval $[-\pi, \pi]$, its Hausdorff dimension is at least 1.

8 Fractal properties of Brownian motion

8.1 Basics on probability spaces and random variables

We begin with the minimal notions of probability theory needed to define Brownian motion and study some of its fractal properties.

A *probability space* is a triple $(\Omega, \mathcal{A}, \mathbb{P})$, where Ω is a set, $\mathcal{A}$ is a σ -algebra on Ω , and $\mathbb{P}$ is a probability measure on Ω such that $\mathcal{A} \subseteq \mathcal{M}_{\mathbb{P}}$. An *event* is an element of $\mathcal{A}$. A *real-valued random variable* is a mapping $X: \Omega \rightarrow \mathbb{R}$ such that for every Borel set $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B)$ is an event. In this case, X is $\mathbb{P}$ -measurable in the sense of Definition 3. For any Borel set A , it is customary to write $\mathbb{P}(X \in A)$ instead of $\mathbb{P}(X^{-1}(A))$ for the probability of the event $X^{-1}(A)$.

A real-valued random variable is *simple* if it takes a finite number of values. The integral of a simple random variable $X: \Omega \rightarrow [0, \infty)$ with respect to the probability measure $\mathbb{P}$ is called the *expectation* of X and is denoted by $\mathbb{E}[X]$. It is defined by Definition 4 by the formula

$$\mathbb{E}[X] = \int_{\Omega} X \, d\mathbb{P} = \sum_{x \in X(\Omega)} x \, \mathbb{P}(X = x).$$

Furthermore, the expectation of a random variable $X: \Omega \rightarrow [0, \infty]$ is defined by

$$\mathbb{E}[X] = \int_{\Omega} X \, d\mathbb{P} = \sup \{ \mathbb{E}[Y], Y \text{ simple and } 0 \leq Y \leq X \}.$$

Now, if $X: \Omega \rightarrow \mathbb{R}$ is an arbitrary real-valued random variable, then its positive part $X^+ = \max\{X, 0\}$ and its negative part $X^- = \max\{-X, 0\}$ are two random variables valued in $[0, \infty)$. If they both have finite expectations, or equivalently if $|X|$ has finite expectation, then the expectation of X is defined by

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

We can also define a measure μ_X on $\mathbb{R}$ by the condition $\mu_X(A) = \mathbb{P}(X \in A)$ for any set $A \subseteq \mathbb{R}$. Then, for every continuous function $f: \mathbb{R} \rightarrow [0, \infty)$, we have the following *integration formula*:

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f \, d\mu_X. \quad (27)$$

This can be proven by approximating f by a monotone sequence of simple measurable functions, using Theorem 5 and Lemma 6, as in the proof of Theorem 29.

An *N -dimensional vector-valued random variable*, also called an *N -dimensional random vector*, is a mapping $X: \Omega \rightarrow \mathbb{R}^N$ such that every component X_i is a real-valued random variable.

Random vectors are usually considered as column vectors. Then, the *expectation* of a random vector $X = (X_1, \dots, X_N)$ is defined by $\mathbb{E}[X] = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_N])$ when all the expectations of the components exist. Random matrices and their expectations are defined similarly. Moreover, the *covariance matrix* of a random vector X is defined by

$$\text{Cov}(X) = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^\top] \quad (28)$$

provided that all the involved expectations exist. If X is a real-valued random variable, this covariance matrix has only one component, namely, the *variance* of X , which is given by

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Finally, a family $(X_i)_{i \in I}$ of real-valued random variables X_i , indexed by some finite, countably infinite, or uncountable set I , is called a *stochastic process* or a *random process*. Its components are called *independent* if for all integers $n \geq 1$, all distinct indices $i_1, \dots, i_n \in I$ and all Borel sets $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$,

$$\mathbb{P}(X_{i_1} \in B_1, \dots, X_{i_n} \in B_n) = \mathbb{P}(X_{i_1} \in B_1) \cdots \mathbb{P}(X_{i_n} \in B_n). \quad (29)$$

For distinct indices $i_1, \dots, i_n$, if these components are nonnegative, approximating them by monotone sequences of nonnegative simple functions of the corresponding components (using Theorem 5 and Lemma 6) gives, with the convention $0 \cdot \infty = 0$,

$$\mathbb{E}[X_{i_1} \cdots X_{i_n}] = \mathbb{E}[X_{i_1}] \cdots \mathbb{E}[X_{i_n}]. \quad (30)$$

If instead the components are real-valued and integrable, applying this identity first to their absolute values shows that their product is integrable; decomposing each component into its positive and negative parts then proves (30) in this case as well. Therefore, the covariance matrix of a random vector with independent square-integrable components is diagonal. In addition, the independence of random vectors of various dimensions can be defined similarly by extending the condition (29) to Borel sets of the Euclidean spaces of appropriate dimensions.

We shall use the following *countable product theorem*, admitted here: for any sequence of Borel probability distributions ν_j on $\mathbb{R}$, there exists a probability space carrying independent random variables X_j such that $\mathbb{P}(X_j \in A) = \nu_j(A)$ for every Borel set A . Equivalently, their finite-dimensional distributions are the corresponding finite products. This is the countable-product case of Kolmogorov's extension theorem; see [3]. It also permits countably many independent copies of a construction based on countably many such variables, by indexing the coordinates with pairs of positive integers.

We recall the analytic change-of-variables formula used below: for a C^1 diffeomorphism $T: U \rightarrow V$ between open subsets of $\mathbb{R}^N$ and a nonnegative Borel function g ,

$$\int_V g(y) dy = \int_U g(T(x)) |\det DT(x)| dx.$$

The same formula holds for integrable functions. Its polar-coordinate consequence for a nonnegative radial integrand is

$$\int_{\mathbb{R}^N} h(\|x\|) dx = \kappa_N \int_0^\infty h(r) r^{N-1} dr,$$

where κ_N is the surface area of the Euclidean unit sphere ($\kappa_1 = 2$). These standard integration formulas are used without proof; see [7].

8.2 Gaussian random variables

Brownian motion is a paramount example of a *Gaussian* stochastic process, meaning that every finite vector of its components is jointly Gaussian, as defined below.

Definition 24 (Gaussian random variable, Gaussian random vector, Gaussian process).

- (i). A real-valued random variable X on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is a *standard Gaussian random variable* if for every $x \in \mathbb{R}$,

$$\mathbb{P}(X > x) = \int_x^\infty \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt. \quad (31)$$

- (ii). An N -dimensional random vector $X = (X_1, \dots, X_N)$ is a *standard Gaussian random vector* if its components $X_1, \dots, X_N$ are standard Gaussian and independent random variables.
- (iii). An N -dimensional random vector X is a *Gaussian random vector* if there exist an integer $M \geq 0$, an M -dimensional standard Gaussian random vector Y on the same probability space, an $N \times M$ matrix A and an N -dimensional vector b with real entries such that

$$X = AY + b \quad \text{almost surely.} \quad (32)$$

For $M = 0$, Y is the empty vector and $AY = 0$; this includes almost surely constant vectors without requiring auxiliary randomness. Equalities in Gaussian representations below are also understood almost surely.

- (iv). A stochastic process $(X_i)_{i \in I}$ is a *Gaussian process* if for every integer $n \geq 1$ and all indices $i_1, \dots, i_n \in I$, the random vector $(X_{i_1}, \dots, X_{i_n})$ is Gaussian.

Suppose that a real-valued random variable X satisfies

$$\mathbb{P}(X > u) = \int_u^\infty \varphi(t) dt \quad (u \in \mathbb{R})$$

for a nonnegative Borel function φ . Then, for every nonnegative Borel function f , we have

$$\mathbb{E}[f(X)] = \int_{-\infty}^\infty f(t)\varphi(t) dt. \quad (33)$$

We call φ the *density* of X . Here is a justification using the preceding measure theory. Letting $u \rightarrow -\infty$ gives $\int_{\mathbb{R}} \varphi = 1$. The Borel probability measures μ_X and $\nu(B) = \int_B \varphi(t) dt$ agree on all intervals (u, ∞) . (The latter measure on Borel sets may be extended to an outer measure by the extension theorem.) Continuity of measures and of the integral shows that both measures give mass zero to each singleton, and hence agree on all open intervals. Every open subset of $\mathbb{R}$ is a countable disjoint union of open intervals: each component contains a distinct rational number. Thus they agree on open sets, and outer regularity on Borel sets, supplied by Theorem 26, gives agreement on all Borel sets. Formula (33) follows first for nonnegative simple functions and then by monotone convergence. It also holds for signed functions whose absolute value has finite integral, by taking positive and negative parts.

For a standard Gaussian variable, $\varphi(t) = (2\pi)^{-1/2}e^{-t^2/2}$. The integral of $|t|\varphi(t)$ is finite, symmetry gives mean zero, and integration by parts gives

$$\int_{\mathbb{R}} t^2 \varphi(t) dt = [-t\varphi(t)]_{-\infty}^\infty + \int_{\mathbb{R}} \varphi(t) dt = 1.$$

Consequently,

$$\mathbb{E}[X] = 0 \quad \text{and} \quad \text{Var}(X) = 1. \quad (34)$$

A plain consequence of (30) and (34) is that the covariance matrix of a standard Gaussian random vector is the identity matrix. Hence, if X is a Gaussian random vector given by (32), then its expectation and covariance matrix are given respectively by

$$\mathbb{E}[X] = b \quad \text{and} \quad \text{Cov}(X) = AA^\top. \quad (35)$$

Moreover, the following lemma shows that the image of a standard Gaussian random vector by an orthogonal transformation is still a standard Gaussian random vector. Recall that a square matrix Q is *orthogonal* if $QQ^\top = I$, where I denotes the identity matrix.

Lemma 62 (Stability by orthogonal transformations). *Let X be an N -dimensional standard Gaussian random vector, and let Q be an $N \times N$ orthogonal matrix. Then, the random vector $Y = QX$ is a standard Gaussian random vector.*

Proof. Independence and the one-dimensional density formula imply that the law of X and the measure with density $\prod_{i=1}^N (2\pi)^{-1/2} e^{-t_i^2/2}$ agree on Borel rectangles. They therefore agree on finite unions of open rectangles, by finite inclusion-exclusion. Every open subset of $\mathbb{R}^N$ is a countable union of rectangles with rational endpoints contained in it; continuity under increasing unions gives equality on open sets. Outer regularity on Borel sets then gives equality on all Borel sets. Simple approximation and monotone convergence justify integration of any nonnegative Borel function against this product density.

For every Borel function $f: \mathbb{R}^N \rightarrow [0, \infty]$, we have

$$\begin{aligned} \mathbb{E}[f(Y^\top)] &= \mathbb{E}[f((QX)^\top)] \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f((t_1, \dots, t_N)Q^\top) \frac{e^{-t_1^2/2}}{\sqrt{2\pi}} \cdots \frac{e^{-t_N^2/2}}{\sqrt{2\pi}} dt_1 \cdots dt_N \\ &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(u_1, \dots, u_N) \frac{e^{-u_1^2/2}}{\sqrt{2\pi}} \cdots \frac{e^{-u_N^2/2}}{\sqrt{2\pi}} du_1 \cdots du_N. \end{aligned}$$

The second equality follows from (33) and Fubini's theorem (Theorem 18), and the third one from the change of variable formula combined with the fact that $|\det Q| = 1$ and $\|t\| = \|tQ^\top\|$. The above computation guarantees that the components of Y are standard Gaussian and independent. (To prove that every component of Y is a standard Gaussian random variable, recover (31) by choosing f as the indicator of the interval (x, ∞) for only one component. To prove the independence of these components, choose f as the product of the indicators of Borel sets of $\mathbb{R}$ for each component.) $\square$

Two N -dimensional random vectors X_1 and X_2 are said to *have the same distribution* if

$$\forall B \in \mathcal{B}(\mathbb{R}^N) \quad \mathbb{P}(X_1 \in B) = \mathbb{P}(X_2 \in B).$$

The following result shows that two Gaussian random vectors with the same expectation and covariance matrix can both be written as in (32) with the same A and b , up to a suitable choice of the standard Gaussian random vector Y . This entails that they have the same distribution.

Lemma 63 (Characterization of Gaussian random vectors). *Let X_1 and X_2 be two N -dimensional Gaussian random vectors such that*

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] \quad \text{and} \quad \text{Cov}(X_1) = \text{Cov}(X_2).$$

Then, there exist two M -dimensional standard Gaussian random vectors Y_1 and Y_2 , an $N \times M$ matrix A and an N -dimensional vector b with real entries such that

$$X_1 = AY_1 + b \quad \text{and} \quad X_2 = AY_2 + b.$$

In particular, the random vectors X_1 and X_2 have the same distribution.

Proof. Up to replacing X_1 and X_2 by $X_1 - \mathbb{E}[X_1]$ and $X_2 - \mathbb{E}[X_2]$, we can assume that $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$. By Definition 24, there exist matrices C_1 and C_2 , and standard Gaussian vectors Z_1 and Z_2 such that

$$X_1 = C_1 Z_1 \quad \text{and} \quad X_2 = C_2 Z_2.$$

Write M_i for the dimension of Z_i and put $r = \text{rank Cov}(X_1) = \text{rank Cov}(X_2)$. The identity $\ker(C_i C_i^\top) = \ker C_i^\top$ shows that both C_i have rank r . If $r = 0$, both centered vectors vanish almost surely; take $M = 0$, let A be the matrix with no columns, and take Y_1, Y_2 to be empty vectors. This proves the claim when $r = 0$, so assume $r > 0$. Choose an $M_i \times M_i$ orthogonal matrix Q_i whose first r columns span the row space of C_i . Then $C_i Q_i = (D_i, 0)$, where D_i has r columns. By Lemma 62, the first r coordinates of $Q_i^\top Z_i$ form a standard Gaussian vector, and multiplying this vector by D_i gives X_i . We may therefore replace the two representations by these reduced representations, so both have input dimension $M = r$, without adding any random variables.

Let C_1 and C_2 denote the spans of the rows of the reduced matrices C_1 and C_2 . Up to simultaneously permuting the output coordinates of X_1, X_2 and the rows of C_1, C_2 , we can assume that the first r rows of C_1 form a basis of C_1 . We undo this common permutation on the final matrix A and vector b . There is a linear mapping $\mathfrak{Q}: C_1 \rightarrow C_2$ that sends these rows to the first r rows of C_2 , meaning that

$$\forall i \in \{1, \dots, r\} \quad \mathfrak{Q}(C_1^{(i)}) = C_2^{(i)}, \quad (36)$$

where $C_1^{(i)}$ and $C_2^{(i)}$ denote the i -th row of C_1 and C_2 , respectively. Let $u \in C_1$. There is an N -dimensional column vector λ whose components are $\lambda_1, \dots, \lambda_r$ followed by $N - r$ zeros such that $u = \lambda^\top C_1$. In view of (35), we have

$$\begin{aligned} \lambda^\top C_2 &= \lambda_1 C_2^{(1)} + \dots + \lambda_r C_2^{(r)} = \mathfrak{Q}(\lambda_1 C_1^{(1)} + \dots + \lambda_r C_1^{(r)}) \\ &= \mathfrak{Q}(\lambda^\top C_1) = \mathfrak{Q}(u), \\ \|\lambda^\top C_2\|^2 &= (\lambda^\top C_2)(\lambda^\top C_2)^\top = \lambda^\top \text{Cov}(X_2) \lambda = \lambda^\top \text{Cov}(X_1) \lambda \\ &= (\lambda^\top C_1)(\lambda^\top C_1)^\top = \|\lambda^\top C_1\|^2 = \|u\|^2. \end{aligned}$$

This implies that $\mathfrak{Q}$ is injective, and thus that the dimension of C_1 is at most that of C_2 . By symmetry, these two dimensions are equal. Hence, $\mathfrak{Q}$ is an isomorphism from C_1 to C_2 . Moreover, the isomorphism $\mathfrak{Q}$ preserves the Euclidean norm. It is therefore an orthogonal transformation from C_1 to C_2 . We can extend $\mathfrak{Q}$ to an orthogonal transformation between M -dimensional row vectors that sends the orthogonal complement of C_1 to that of C_2 . This orthogonal transformation can be represented by an $M \times M$ orthogonal matrix Q such that $\mathfrak{Q}(u) = uQ$ for every row vector u . For $1 \leq i \leq r < j \leq N$, formulas (35) and (36) give

$$\langle C_2^{(j)} - \mathfrak{Q}(C_1^{(j)}), C_2^{(i)} \rangle = \langle C_2^{(j)}, C_2^{(i)} \rangle - \langle \mathfrak{Q}(C_1^{(j)}), \mathfrak{Q}(C_1^{(i)}) \rangle = \langle C_2^{(j)}, C_2^{(i)} \rangle - \langle C_1^{(j)}, C_1^{(i)} \rangle$$

$$= (C_2 C_2^\top)_{j,i} - (C_1 C_1^\top)_{j,i} = (\text{Cov}(X_2))_{j,i} - (\text{Cov}(X_1))_{j,i} = 0.$$

Since the rows $C_2^{(1)}, \dots, C_2^{(r)}$ form a basis of C_2 , we deduce that $C_2^{(j)} - \mathfrak{L}(C_1^{(j)})$ belongs to the orthogonal complement of C_2 . Moreover, it also belongs to C_2 , so it is equal to zero. We have proved that (36) holds for every $i \in \{1, \dots, N\}$, meaning that $C_2 = C_1 Q$. We deduce that

$$X_1 = C_1 Z_1 \quad \text{and} \quad X_2 = C_2 Z_2 = C_1 Q Z_2.$$

To conclude the proof, set $A = C_1$, $Y_1 = Z_1$, $Y_2 = Q Z_2$ and $b = 0$, and observe using Lemma 62 that Y_2 is a standard Gaussian random vector since Q is orthogonal. The last claim of the lemma follows from the observation that for every Borel set $B \in \mathcal{B}(\mathbb{R}^N)$,

$$\begin{aligned} \mathbb{P}(X_1 \in B) &= \mathbb{P}(\psi_{A,b}(Y_1) \in B) = \mathbb{P}(Y_1 \in \psi_{A,b}^{-1}(B)) = \mathbb{P}(Y_2 \in \psi_{A,b}^{-1}(B)) \\ &= \mathbb{P}(\psi_{A,b}(Y_2) \in B) = \mathbb{P}(X_2 \in B), \end{aligned}$$

where $\psi_{A,b}$ denotes the continuous mapping $y \mapsto Ay + b$. The middle equality follows from the fact that Y_1 and Y_2 are both M -dimensional standard Gaussian random vectors. $\square$

Lemma 63 shows that two Gaussian random vectors with the same expectation μ and the same covariance matrix Σ have the same distribution. For this reason, we thus say that X has *distribution* $N(\mu, \Sigma)$. The symbol N means *normal*, another name for Gaussian. In particular, a standard Gaussian random vector has distribution $N(0, I)$. Lemma 63 also gives the following corollary.

Corollary 64 (Sum and difference of independent Gaussian random variables). *Let X_1 and X_2 be two independent random variables with distribution $N(0, \sigma^2)$ for some $\sigma > 0$. Then, $X_1 + X_2$ and $X_1 - X_2$ are independent random variables with distribution $N(0, 2\sigma^2)$.*

Proof. Set $Y_i = X_i/\sigma$ for $i = 1, 2$. Scaling and Lemma 63 show that each Y_i is standard Gaussian, and independence is inherited from X_1, X_2 . Thus $Y = (Y_1, Y_2)^\top$ is a standard Gaussian vector and $X_i = \sigma Y_i$. Then,

$$\frac{1}{\sigma\sqrt{2}} \begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} Y_1 + Y_2 \\ Y_1 - Y_2 \end{pmatrix} = QY \quad \text{with} \quad Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

It is easy to check that $QQ^\top = I$. Thus, the matrix Q is orthogonal. We deduce from Lemma 62 that the vector in the left-hand side above is a standard Gaussian random vector. Hence, its components are independent standard Gaussian random variables. This implies that $X_1 + X_2$ and $X_1 - X_2$ are independent random variables with distribution $N(0, 2\sigma^2)$. $\square$

We end this section on Gaussian random variables with the following tail estimates.

Lemma 65 (Gaussian tail bounds). *Let X be a standard Gaussian random variable. Then, for every $x > 0$,*

$$\frac{e^{-x^2/2}}{(x + 1/x)\sqrt{2\pi}} \leq \mathbb{P}(X > x) \leq \frac{e^{-x^2/2}}{x\sqrt{2\pi}}.$$

Proof. On the one hand, the formula (31) yields

$$\mathbb{P}(X > x) = \int_x^\infty \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \leq \int_x^\infty \frac{t e^{-t^2/2}}{x \sqrt{2\pi}} dt = \frac{1}{x} \left[\frac{-e^{-t^2/2}}{\sqrt{2\pi}} \right]_x^\infty = \frac{e^{-x^2/2}}{x\sqrt{2\pi}}.$$

On the other hand, the function f defined by

$$f(x) = x e^{-x^2/2} - (x^2 + 1) \int_x^\infty e^{-t^2/2} dt$$

satisfies $f(0) < 0$ and $f(x) \rightarrow 0$ as $x \rightarrow \infty$, by the previous bound. Its derivative is

$$f'(x) = e^{-x^2/2} - x^2 e^{-x^2/2} - 2x \int_x^\infty e^{-t^2/2} dt + (x^2 + 1) e^{-x^2/2} = -2x \left(\int_x^\infty e^{-t^2/2} dt - \frac{1}{x} e^{-x^2/2} \right).$$

The previous bound implies that $f'(x) \geq 0$ for every $x > 0$. Hence, $f(x) \leq 0$ for every $x \geq 0$. $\square$

8.3 Lévy's construction of Brownian motion

Brownian motion is a Gaussian process indexed by an interval $[0, a]$ with $a > 0$, or by $[0, \infty)$, with stationary independent increments and continuous trajectories. More precisely:

Definition 25 (Standard Brownian motion). Let $I = [0, a]$ with $a > 0$, or $I = [0, \infty)$. A real-valued stochastic process $(B_t)_{t \in I}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is a *standard Brownian motion* if the following conditions are satisfied:

- (i). $B_0 = 0$;
- (ii). for all integers $n \geq 3$ and all numbers $t_1 < \dots < t_n$ in I , the increments $B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ are independent random variables;
- (iii). for all numbers $s, t \in I$ with $s < t$, the increment $B_t - B_s$ is a Gaussian random variable with distribution $N(0, t - s)$;
- (iv). $\mathbb{P}(t \mapsto B_t \text{ is continuous on } I) = 1$.

For pathwise arguments we may work with a version having continuous paths everywhere: choose the measurable event of probability one in the definition, retain the process on that event, and set the whole path equal to zero on its complement. All finite-dimensional distributions are unchanged. This version is jointly measurable in (ω, t) . Indeed, let $q_m(t) = 2^{-m} \lfloor 2^m t \rfloor$. For every m , the function $(\omega, t) \mapsto B_{q_m(t)}(\omega)$ is measurable for the σ -algebra generated by measurable rectangles, because q_m has countably many values with Borel level sets. Continuity gives $B_{q_m(t)}(\omega) \rightarrow B_t(\omega)$ pointwise. This also proves joint measurability of functions of (ω, t, t') used in the energy integrals below, and justifies their use in Fubini's theorem. Any almost sure path property established for this version holds for the original process on the same event of continuous paths.

In addition to (28), recall that the covariance of two real-valued random variables X and Y is

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

A plain consequence of Definition 25 is that for all $s, t \in I$, the covariance of B_s and B_t is given by

$$\text{Cov}(B_s, B_t) = \min\{s, t\}.$$

Indeed, if for instance $s \leq t$, then $B_t = B_s + (B_t - B_s)$, where B_s and $B_t - B_s$ are independent random variables with distribution $N(0, s)$ and $N(0, t - s)$, respectively. Hence, using (30) and (35), we obtain

$$\text{Cov}(B_s, B_t) = \text{Cov}(B_s, B_s) + \text{Cov}(B_s, B_t - B_s) = \text{Var}(B_s) + 0 = s.$$

Existence of Brownian motion was first established by Wiener in 1923. An explicit construction as a uniform limit of piecewise affine interpolations at dyadic time values was then given by Lévy in 1948. This construction relies on the following convergence part of the Borel-Cantelli lemma.

Lemma 66 (Borel-Cantelli lemma (convergence part)). *Let $(A_n)_{n \geq 1}$ be a sequence of events on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The event that A_n occurs for infinitely many indices n is given by*

$$\{A_n \text{ i.o.}\} = \limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m,$$

where i.o. stands for infinitely often. Moreover, if $\sum_n \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(A_n \text{ i.o.}) = 0$.

Proof. For every integer $n \geq 1$, the event $\bigcup_{m=n}^{\infty} A_m$ is the event that A_m occurs for some index $m \geq n$. Hence, the event $\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$ is the event that for every integer $n \geq 1$, there exists an index $m \geq n$ such that A_m occurs. This is precisely the event that A_n occurs for infinitely many indices n . This proves the first claim of the lemma. In addition, the monotonicity and the countable subadditivity of the probability measure, and the convergence of the series imply that

$$\mathbb{P}(A_n \text{ i.o.}) \leq \mathbb{P}\left(\bigcup_{m=n}^{\infty} A_m\right) \leq \sum_{m=n}^{\infty} \mathbb{P}(A_m) \xrightarrow{n \rightarrow \infty} 0,$$

which proves the second claim. $\square$

We can now present Lévy's construction of Brownian motion.

Theorem 67 (Existence of Brownian motion). *Standard Brownian motion exists on the interval $[0, \infty)$, and thus on any interval $[0, a]$ with $a > 0$.*

Proof. We first construct the process on $[0, 1]$ using the dyadic numbers

$$D = \bigcup_{j=0}^{\infty} D_j \quad \text{with} \quad D_j = \{k2^{-j}, k \in \{0, \dots, 2^j\}\},$$

and then a collection of independent standard Gaussian random variables X_s , for $s \in D$, on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. (Existence follows from the countable product theorem stated above.) We define inductively a collection of random variables B_s that give the values of the Brownian motion at dyadic time values $s \in D$ in such a way that for every integer $j \geq 0$, the following conditions are satisfied:

- (i). For $n \geq 3$ and $s_1 < \dots < s_n$ in D_j , the increments $B_{s_2} - B_{s_1}, \dots, B_{s_n} - B_{s_{n-1}}$ are independent;
- (ii). For all $s_1 < s_2$ in D_j , the increment $B_{s_2} - B_{s_1}$ has distribution $N(0, s_2 - s_1)$;
- (iii). The families of random variables $(B_s)_{s \in D_j}$ and $(X_s)_{s \in D \setminus D_j}$ are independent.

These conditions are satisfied for $j = 0$ by setting $B_0 = 0$ and $B_1 = X_1$. Assume that they are satisfied for some $j \geq 0$. For every $s \in D_{j+1} \setminus D_j$, consider $s^- = s - 2^{-(j+1)}$ and $s^+ = s + 2^{-(j+1)}$ in D_j , and define

$$B_s = \frac{B_{s^-} + B_{s^+}}{2} + \frac{X_s}{2^{(j+2)/2}}. \quad (37)$$

By induction, the random variables $(B_{s^+} - B_{s^-})/2$ and $X_s/2^{(j+2)/2}$ are independent and both have distribution $N(0, 2^{-(j+2)})$. By Corollary 64, their sum $B_s - B_{s^-}$ and their difference $B_{s^+} - B_s$ are independent and have distribution $N(0, 2^{-(j+1)})$. This implies, again by induction, that the above three conditions are satisfied for $j + 1$. These conditions are thus satisfied for every integer $j \geq 0$.

We now interpolate between the dyadic time values to define a sequence of piecewise affine random functions. Specifically, set $Y_{0,t} = tX_1$ for every $t \in [0, 1]$, and then for every $j \geq 1$, define $t \mapsto Y_{j,t}$ as the piecewise affine function on $[0, 1]$ that interpolates between the points $(s, 2^{-(j+1)/2} X_s \mathbf{1}_{\{s \notin D_{j-1}\}})$ for $s \in D_j$. These functions are continuous on $[0, 1]$. Moreover, for all $j \geq 0$ and $s \in D_j$, we have

$$B_s = \sum_{i=0}^j Y_{i,s} = \sum_{i=0}^{\infty} Y_{i,s}. \quad (38)$$

The second equality holds because $Y_{i,s} = 0$ for every $i > j$. The first equality can be proven by induction on j . It is satisfied for $j = 0$ since $B_0 = 0$ and $B_1 = X_1$. Assume that it is satisfied for some $j \geq 0$. For every $s \in D_{j+1} \setminus D_j$ and every $i \in \{0, \dots, j\}$, the function $t \mapsto Y_{i,t}$ is affine between the dyadic time values s^- and s^+ which are consecutive in D_j , so between two consecutive dyadic time values in D_j . The value of $Y_{i,s}$ at the midpoint s of s^- and s^+ is thus the average of the values Y_{i,s^-} and Y_{i,s^+} . Hence,

$$\sum_{i=0}^{j+1} Y_{i,s} = \sum_{i=0}^j \frac{Y_{i,s^-} + Y_{i,s^+}}{2} + Y_{j+1,s} = \frac{1}{2} \left(\sum_{i=0}^j Y_{i,s^-} + \sum_{i=0}^j Y_{i,s^+} \right) + Y_{j+1,s} = \frac{B_{s^-} + B_{s^+}}{2} + \frac{X_s}{2^{(j+2)/2}} = B_s,$$

where the last equality follows from (37). By induction, this proves (38) for all $j \geq 0$ and $s \in D_j$.

Let us prove that the series in (38) converges uniformly on the whole interval $[0, 1]$ with probability one. For every integer $j \geq 1$, the function $t \mapsto |Y_{j,t}|$ is bounded above on the interval $[0, 1]$ by the maximal value of $2^{-(j+1)/2} |X_s|$ for $s \in D_j$. Using Lemma 65, we infer that for all $c > 0$,

$$\mathbb{P} \left(\max_{s \in D_j} |X_s| > (2cj)^{1/2} \right) \leq \sum_{s \in D_j} \mathbb{P} \left(|X_s| > (2cj)^{1/2} \right) \leq \frac{2(2^j + 1)e^{-cj}}{(4c\pi j)^{1/2}}.$$

If $c > \log 2$, then the sum of the right-hand side above over $j \geq 1$ is finite. By the convergence part of the Borel-Cantelli lemma (Lemma 66), we deduce that

$$\mathbb{P} \left(\sum_{j=0}^{\infty} \sup_{t \in [0,1]} |Y_{j,t}| = \infty \right) \leq \mathbb{P} \left(\sup_{t \in [0,1]} |Y_{j,t}| > (cj)^{1/2} 2^{-j/2} \text{ i.o.} \right) \leq \mathbb{P} \left(\max_{s \in D_j} |X_s| > (2cj)^{1/2} \text{ i.o.} \right) = 0. \quad (39)$$

This means that on an event of probability one, the series $\sum_j Y_{j,t}$ converges uniformly on $[0, 1]$ to a continuous function. By (38), the sum is equal to B_s for every $s \in D$. The sum of the series is thus the unique continuous extension of the mapping $s \mapsto B_s$ defined on the dense subset D of $[0, 1]$. We use the notation $t \mapsto B_t$ for this extension on the measurable event of uniform convergence, and set $B_t = 0$ for every t on its null complement. Each B_t is then a random variable, and every path of this version is continuous. Its values at dyadic times agree almost surely with the previously defined variables.

Let us prove that this process is a standard Brownian motion. The first and fourth conditions of Definition 25 hold by construction. Fix $n \geq 2$ and $t_1 < \dots < t_n$ in $[0, 1]$, and choose $s_{j,p} \in D_j$ tending to t_p . For sufficiently large j these times are strictly increasing. Put

$$Z_j = (B_{s_{j,p+1}} - B_{s_{j,p}})_{p=1}^{n-1}, \quad Z = (B_{t_{p+1}} - B_{t_p})_{p=1}^{n-1}.$$

Continuity implies $Z_j \rightarrow Z$ almost surely. Let $G_1, \dots, G_{n-1}$ be independent standard Gaussian variables and put $a_{j,p} = \sqrt{s_{j,p+1} - s_{j,p}}$ and $a_p = \sqrt{t_{p+1} - t_p}$. The dyadic construction says

that Z_j has the same distribution as $(a_{j,p}G_p)_{p=1}^{n-1}$. For every bounded continuous function $\phi: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$, dominated convergence on the two probability spaces gives

$$\mathbb{E}[\phi(Z)] = \lim_{j \rightarrow \infty} \mathbb{E}[\phi(Z_j)] = \lim_{j \rightarrow \infty} \mathbb{E}[\phi((a_{j,p}G_p)_p)] = \mathbb{E}[\phi((a_pG_p)_p)].$$

These equalities determine the Borel distributions. Indeed, regularize either distribution by Proposition 19; the resulting finite Borel-regular measures on Euclidean space are Radon by Theorem 26. Equality of their integrals on $C_c(\mathbb{R}^{n-1})$ implies equality of these measures by Theorem 23, and hence equality of the original distributions on Borel sets. Thus the components of Z are independent Gaussians with variances $t_{p+1} - t_p$. This proves the second and third conditions, including the single-increment case $n = 2$.

Finally, to extend the construction to the interval $[0, \infty)$, consider a sequence of independent standard Brownian motions $(B_{m,t})_{t \in [0,1]}$ on $[0, 1]$ for $m \geq 0$, and then define for every $t \geq 0$,

$$B_t = B_{[t], t - [t]} + \sum_{i=0}^{[t]-1} B_{i,1},$$

where $[t]$ denotes the integer part of t . It is now easy to check that the stochastic process $(B_t)_{t \in [0, \infty)}$ is a standard Brownian motion on the interval $[0, \infty)$. $\square$

8.4 Hölder continuity and modulus of continuity

The construction developed in the proof of Theorem 67 shows that Brownian motion has continuous trajectories. It actually gives a modulus of continuity, and thus proves that Brownian motion is Hölder continuous with exponent less than $1/2$. We highlight this in the following statement.

Proposition 68 (Hölder continuity and modulus of continuity of Brownian motion). *Let $(B_t)_{t \in [0,1]}$ be a standard Brownian motion on the interval $[0, 1]$. Then, there exists an event of probability one on which the function $t \mapsto B_t$ has modulus of continuity $\eta(0) = 0$ and $\eta(t) = (t(1 + \max\{0, \log(1/t)\}))^{1/2}$ for $t > 0$. Hence, on this event, the function $t \mapsto B_t$ is α -Hölder continuous for every $\alpha \in (0, 1/2)$.*

Proof. First consider the Brownian motion constructed in the proof of Theorem 67. For all $j \geq 1$ and $t, t' \in [0, 1]$, we have

$$\begin{aligned} |Y_{j,t} - Y_{j,t'}| &\leq \min \left\{ |t - t'| \max_{s \in D_j} \frac{2^{-(j+1)/2} |X_s|}{2^{-j}}, 2 \max_{s \in D_j} 2^{-(j+1)/2} |X_s| \right\} \\ &= 2^{-(j+1)/2} \min \{ 2^j |t - t'|, 2 \} \max_{s \in D_j} |X_s|. \end{aligned}$$

In addition, we know from (39) that if $c > \log 2$, then on an event of probability one, the maximal value of $|X_s|$ for $s \in D_j$ is at most $(2cj)^{1/2}$ for all but finitely many indices j . Thus, on an event of probability one, there exists a (random) constant $Z \in (0, \infty)$ such that for every $j \geq 1$,

$$\max_{s \in D_j} |X_s| \leq Z j^{1/2}.$$

For the initial term, $|Y_{0,t} - Y_{0,t'}| = |X_1||t - t'|$. Since B_t is the sum of the series $\sum_j Y_{j,t}$, we deduce that on this event,

$$|B_t - B_{t'}| \leq \sum_{j=0}^{\infty} |Y_{j,t} - Y_{j,t'}| \leq |X_1||t - t'| + Z \sum_{j=1}^{\infty} 2^{-(j+1)/2} \min \{ 2^j |t - t'|, 2 \} j^{1/2}$$

$$\leq |X_1| |t - t'| + \frac{Z}{\sqrt{2}} |t - t'| \sum_{j=1}^{j_0-1} 2^{j/2} j^{1/2} + Z\sqrt{2} \sum_{j=j_0}^{\infty} 2^{-j/2} j^{1/2},$$

for any integer $j_0 \geq 1$. In addition, we have

$$\sum_{j=1}^{j_0-1} 2^{j/2} j^{1/2} \leq j_0^{1/2} \sum_{j=1}^{j_0-1} 2^{j/2} \leq \frac{2^{j_0/2} j_0^{1/2}}{\sqrt{2} - 1} \quad \text{and} \quad \sum_{j=j_0}^{\infty} 2^{-j/2} j^{1/2} \leq 2^{-j_0/2} j_0^{1/2} \sum_{j=0}^{\infty} 2^{-j/2} (1+j)^{1/2},$$

so there are two numerical constants $C_1, C_2 \in (0, \infty)$ such that

$$|B_t - B_{t'}| \leq |X_1| |t - t'| + C_1 Z |t - t'| 2^{j_0/2} j_0^{1/2} + C_2 Z 2^{-j_0/2} j_0^{1/2}.$$

Assuming that $t \neq t'$ and choosing the integer $j_0 \geq 1$ such that $2^{-j_0} < |t - t'| \leq 2^{1-j_0}$, we obtain

$$|B_t - B_{t'}| \leq |X_1| |t - t'| + (C_1 Z \sqrt{2} + C_2 Z) |t - t'|^{1/2} (1 + \log_2(1/|t - t'|))^{1/2} \leq C \eta(|t - t'|)$$

for some (random) constant $C \in (0, \infty)$. This proves the modulus bound for the constructed process, including $t = t'$ by continuity of η at zero. To transfer it to any standard Brownian motion, let D be the dyadic times in $[0, 1]$ and consider the event

$$E = \bigcup_{M=1}^{\infty} \bigcap_{u, v \in D} \{|B_u - B_v| \leq M \eta(|u - v|)\}.$$

Its probability is determined by the finite-dimensional distributions: enumerate the countably many pairs, use continuity from above for each fixed M , and then continuity from below in M . All standard Brownian motions have the same finite-dimensional distributions, since their values are sums of independent Gaussian increments. Hence $\mathbb{P}(E) = 1$ for every standard Brownian motion. Intersecting E with its event of continuous paths extends the estimate from dyadic pairs to all pairs by continuity. This proves the first claim. The second claim follows from the observation that for every $\alpha \in (0, 1/2)$, there exists a numerical constant $C_\alpha \in (0, \infty)$ such that $\eta(t) \leq C_\alpha t^\alpha$ for every $t \in (0, 1]$. $\square$

8.5 Nowhere differentiability

In contrast with the Hölder continuity property established in Proposition 68, Brownian motion is almost surely nowhere differentiable. This is stated in the next proposition.

Proposition 69 (Nowhere differentiability of Brownian motion). *Let $(B_t)_{t \in [0,1]}$ be a standard Brownian motion on the interval $[0, 1]$. Then, there exists an event of probability one on which the function $t \mapsto B_t$ is nowhere differentiable. Furthermore, on this event, for every $t \in [0, 1]$, either $D_* B_t = -\infty$ or $D^* B_t = \infty$, where the limits are taken through $h \neq 0$ with $t + h \in [0, 1]$ (thus one-sided at the endpoints):*

$$D_* B_t = \liminf_{h \rightarrow 0} \frac{B_{t+h} - B_t}{h} \quad \text{and} \quad D^* B_t = \limsup_{h \rightarrow 0} \frac{B_{t+h} - B_t}{h}.$$

Proof. We use the continuous-path version described above and work on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and define subsets of Ω by the formulas

$$\Omega' = \{\exists t_0 \in [0, 1] : -\infty < D_* B_{t_0} \leq D^* B_{t_0} < \infty\}$$

$$\Omega_M = \left\{ \exists t_0 \in [0, 1] : \sup_{t \in [0,1] \setminus \{t_0\}} \left| \frac{B_t - B_{t_0}}{t - t_0} \right| \leq M \right\}.$$

If both Dini derivatives at t_0 are finite, the difference quotients are bounded near t_0 ; continuity on $[0, 1]$ bounds them away from t_0 as well. Thus the set Ω' is contained in the union of the sets Ω_M over all integers $M \geq 1$. Hence, it suffices to prove that $\mathbb{P}(\Omega_M) = 0$ for every $M \geq 1$. If a time value $t_0 \in [0, 1]$ satisfies the condition in the definition of Ω_M and belongs to the dyadic interval $[k_0 2^{-j}, (k_0 + 1) 2^{-j}]$ for $j \geq 0$ and $k_0 \in \{0, \dots, 2^j - 1\}$, then for every $k \in \{0, \dots, 2^j - 1\}$, the triangle inequality implies that

$$|B_{(k+1)2^{-j}} - B_{k2^{-j}}| \leq |B_{(k+1)2^{-j}} - B_{t_0}| + |B_{k2^{-j}} - B_{t_0}| \quad (40)$$

$$\leq M|(k+1)2^{-j} - t_0| + M|k2^{-j} - t_0| \leq 2M(|k - k_0| + 1)2^{-j}. \quad (41)$$

Consequently,

$$\Omega_M \subseteq \bigcap_{j=0}^{\infty} \bigcup_{k_0=0}^{2^j-1} \Omega_{M,j,k_0} \quad \text{with} \quad \Omega_{M,j,k_0} = \bigcap_{k=0}^{2^j-1} \{|B_{(k+1)2^{-j}} - B_{k2^{-j}}| \leq M(|k - k_0| + 1)2^{1-j}\}.$$

The increments $B_{(k+1)2^{-j}} - B_{k2^{-j}}$ for $k \in \{0, \dots, 2^j - 1\}$ are independent Gaussian random variables with distribution $N(0, 2^{-j})$. Hence, letting X denote a standard Gaussian random variable and observing from (31) that $\mathbb{P}(|X| \leq x) \leq (2/\pi)^{1/2} x \leq x$, we obtain

$$\begin{aligned} \mathbb{P}(\Omega_{M,j,k_0}) &= \prod_{k=0}^{2^j-1} \mathbb{P}(2^{j/2} |B_{(k+1)2^{-j}} - B_{k2^{-j}}| \leq M(|k - k_0| + 1)2^{1-j/2}) \\ &= \prod_{k=0}^{2^j-1} \mathbb{P}(|X| \leq M(|k - k_0| + 1)2^{1-j/2}) \\ &\leq \prod_{k=0}^{2^j-1} \min\{1, M(|k - k_0| + 1)2^{1-j/2}\} \leq (6M)^3 2^{-3j/2}. \end{aligned}$$

For the last inequality, we assume that $j \geq 2$ and just retain three factors in the product, namely for three values of k such that $|k - k_0| \leq 2$. We finally deduce that

$$\mathbb{P}(\Omega_M) \leq \inf_{j \geq 2} \sum_{k_0=0}^{2^j-1} \mathbb{P}(\Omega_{M,j,k_0}) \leq \inf_{j \geq 2} 2^j (6M)^3 2^{-3j/2} = 0,$$

so that $\mathbb{P}(\Omega') = 0$. More explicitly, the measurable set $E_M = \bigcap_{j \geq 2} \bigcup_{k_0=0}^{2^j-1} \Omega_{M,j,k_0}$ contains Ω_M and has probability zero by the same estimate. The complement of $\bigcup_{M \geq 1} E_M$ is therefore a measurable event of probability one on which the conclusion holds. This avoids needing a separate measurability assertion for the existentially defined sets Ω_M . $\square$

Remark 4 (Nowhere α -Hölder regularity of Brownian motion for $\alpha > 1/2$). The proof of Proposition 69 can be adapted to show that for every $\alpha > 1/2$, there exists an event of probability one on which the function $t \mapsto B_t$ is nowhere α -Hölder regular, in the sense that for every $t \in [0, 1]$, either $\liminf_{h \rightarrow 0} (B_{t+h} - B_t)/|h|^\alpha = -\infty$ or $\limsup_{h \rightarrow 0} (B_{t+h} - B_t)/|h|^\alpha = \infty$. The limits again use $h \neq 0$ with $t + h \in [0, 1]$. The proof is essentially the same, except that the upper bound (40) is replaced by $2M(|k - k_0| + 1)^\alpha 2^{-\alpha j}$. It is then necessary to assume j sufficiently large to retain more than $(\alpha - 1/2)^{-1}$ factors in the product, so as to obtain a bound that tends to zero as $j \rightarrow \infty$.

8.6 Dimensions of graph and images

Recall from Section 7.1 that the Hölder regularity of a function has an impact on the Hausdorff and box-counting dimensions of its graph and images. We now consider Brownian motion.

Corollary 70 (Dimension upper bounds for the graph and images of Brownian motion). *Let $(B_t)_{t \in [0,1]}$ be a standard Brownian motion on the interval $[0, 1]$. Then, there exists an event of probability one on which the following properties hold:*

(i). *The Hausdorff and box-counting dimensions of the graph of the function $B: t \mapsto B_t$ satisfy*

$$\dim_H \text{gph } B \leq \underline{\dim}_B \text{gph } B \leq \overline{\dim}_B \text{gph } B \leq 3/2;$$

(ii). *For every set $A \subseteq [0, 1]$, it holds that*

$$\dim_H B(A) \leq \min\{1, 2 \dim_H A\}.$$

Proof. Proposition 68 ensures that on some event of probability one, the function $t \mapsto B_t$ is α -Hölder continuous for every $\alpha \in (0, 1/2)$. Propositions 45, 49 and 50 then ensure that on this event, for every $\alpha \in (0, 1/2)$ and every set $A \subseteq [0, 1]$, we have

$$\dim_H \text{gph } B \leq \underline{\dim}_B \text{gph } B \leq \overline{\dim}_B \text{gph } B \leq 2 - \alpha \quad \text{and} \quad \dim_H B(A) \leq \frac{1}{\alpha} \dim_H A.$$

The conclusion follows by letting α tend to $1/2$. $\square$

The upper bound of $3/2$ for the Hausdorff and box-counting dimensions of the graph of Brownian motion is actually an equality. This is the content of the following statement, due to Taylor in 1953.

Theorem 71 (Graph dimensions for Brownian motion). *Let $(B_t)_{t \in [0,1]}$ be a standard Brownian motion on the interval $[0, 1]$. Then, there exists an event of probability one on which the Hausdorff and box-counting dimensions of the graph of the function $B: t \mapsto B_t$ satisfy*

$$\dim_H \text{gph } B = \underline{\dim}_B \text{gph } B = \overline{\dim}_B \text{gph } B = 3/2.$$

Proof. The upper bounds are given by Corollary 70. Hence, we need to prove that $\dim_H \text{gph } B \geq 3/2$ on some event of probability one. In view of Theorem 48, it suffices to prove that for every $s \in (1, 3/2)$, there exists an event of probability one on which

$$\text{Cap}_s(\text{gph } B) > 0. \quad (42)$$

(Taking a sequence of values of s that increases to $3/2$ and intersecting the corresponding events of probability one then gives the conclusion.) For every set $A \subseteq \mathbb{R}^2$, define

$$\mu(A) = \mathcal{L}^1(\{t \in [0, 1] \mid (t, B_t) \in A\}).$$

If we work on an event of probability one on which the function $t \mapsto B_t$ is continuous, then μ is a Borel measure with support contained in the graph of B . The s -dimensional energy of μ satisfies

$$\mathcal{E}_s(\mu) = \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \frac{\mu(dx) \mu(dx')}{\|x - x'\|^s} = \int_0^1 \int_0^1 \frac{dt' dt}{(|t - t'|^2 + |B_t - B_{t'}|^2)^{s/2}}.$$

This follows from the integration formula given by Theorem 29. Applying Fubini's theorem (Theorem 18) to the product of probability measure and the Lebesgue measure, we obtain

$$\mathbb{E}[\mathcal{E}_s(\mu)] = \int_0^1 \int_0^1 \mathbb{E}\left[(|t - t'|^2 + |B_t - B_{t'}|^2)^{-s/2} \right] dt' dt.$$

Since the integrand is symmetric in t and t' , we can restrict the integration domain to the triangle $\{(t, t') \in [0, 1]^2 \mid t' \leq t\}$ and then multiply the integral by 2. For $0 \leq t' \leq t \leq 1$, the

increment $B_t - B_{t'}$ has distribution $N(0, t - t')$, so it has the same distribution as $X\sqrt{t - t'}$, where X is a standard Gaussian random variable. Performing the change of variable $u = t - t'$, we thus obtain

$$\begin{aligned}\mathbb{E}[\mathcal{E}_s(\mu)] &= 2 \int_0^1 \int_0^t \mathbb{E}\left[\left((t - t')^2 + (t - t')X^2\right)^{-s/2}\right] dt' dt \\ &= 2 \int_0^1 \int_0^t \mathbb{E}\left[(u^2 + u X^2)^{-s/2}\right] du dt \\ &\leq 2 \int_0^1 \int_0^1 \mathbb{E}\left[(u^2 + u X^2)^{-s/2}\right] du dt = 2 \int_0^1 \mathbb{E}\left[(u^2 + u X^2)^{-s/2}\right] du.\end{aligned}$$

Thanks to (31) and (33), the expectation in the integrand satisfies

$$\mathbb{E}\left[(u^2 + u X^2)^{-s/2}\right] = \int_{-\infty}^{\infty} (u^2 + u x^2)^{-s/2} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} (u^2 + u x^2)^{-s/2} e^{-x^2/2} dx.$$

To estimate the last integral, we split the integration interval at the values $\sqrt{u}$ and 1. Accordingly, we consider the following three integrals, recalling that $s \in (1, 3/2)$ when estimating them:

$$\begin{aligned}\text{(i). } & \int_0^{\sqrt{u}} (u^2 + u x^2)^{-s/2} e^{-x^2/2} dx \leq \int_0^{\sqrt{u}} (u^2)^{-s/2} dx = u^{1/2-s}; \\ \text{(ii). } & \int_{\sqrt{u}}^1 (u^2 + u x^2)^{-s/2} e^{-x^2/2} dx \leq \int_{\sqrt{u}}^1 (u x^2)^{-s/2} dx = \left[\frac{u^{-s/2} x^{1-s}}{1-s} \right]_{\sqrt{u}}^1 = u^{-s/2} \frac{u^{(1-s)/2} - 1}{s-1} \leq \frac{u^{1/2-s}}{s-1}; \\ \text{(iii). } & \int_1^{\infty} (u^2 + u x^2)^{-s/2} e^{-x^2/2} dx \leq u^{-s/2} \int_1^{\infty} e^{-x^2/2} dx \leq u^{-s/2} \sqrt{\frac{\pi}{2}}.\end{aligned}$$

Summing these three bounds, we finally obtain

$$\mathbb{E}[\mathcal{E}_s(\mu)] \leq 2\sqrt{\frac{2}{\pi}} \int_0^1 \left(\frac{s}{s-1} u^{1/2-s} + u^{-s/2} \sqrt{\frac{\pi}{2}} \right) du < \infty.$$

This shows that on some event of probability one, $\mathcal{E}_s(\mu)$ is finite. Dividing the measure μ by its total mass, we obtain a Borel probability measure with finite s -dimensional energy and support contained in the graph of B . This proves that (42) holds and concludes the proof. $\square$

Similarly, the upper bound on the Hausdorff dimension of the images of Brownian motion given in Corollary 70 is an equality. This is even true for N -dimensional standard Brownian motions, namely, stochastic processes $B_t = (B_t^{(1)}, \dots, B_t^{(N)})$ taking values in $\mathbb{R}^N$, whose components $B_t^{(i)}$ are independent standard Brownian motions. We have indeed the following statement, due to McKean in 1955.

Theorem 72 (Hausdorff dimension of the images of N -dimensional Brownian motion). *Let $N \geq 1$ be an integer, let $(B_t)_{t \in [0,1]}$ be a N -dimensional standard Brownian motion on the interval $[0, 1]$, and let A be an F_σ -set in $[0, 1]$ such that $\dim_H A \leq N/2$. Then, there exists an event of probability one on which the function $B: t \mapsto B_t$ satisfies*

$$\dim_H B(A) = 2 \dim_H A.$$

Proof. Proposition 50 can easily be adapted to functions f defined on a nonempty interval $I \subseteq \mathbb{R}$ and taking values in $\mathbb{R}^N$ that are α -Hölder continuous, meaning that the absolute value in the left-hand side of (19) is replaced by the Euclidean norm in $\mathbb{R}^N$. Proposition 68 shows that, on some event of probability one, all the components of the N -dimensional Brownian motion are α -Hölder continuous for every $\alpha \in (0, 1/2)$. The N -dimensional Brownian motion $B: t \mapsto B_t$ is thus α -Hölder continuous for every such α on this event. This implies that for any set $A \subseteq [0, 1]$,

$$\dim_H B(A) \leq \min \left\{ N, \frac{1}{\alpha} \dim_H A \right\} \xrightarrow{\alpha \rightarrow 1/2} \min \{N, 2 \dim_H A\}.$$

If $A = \emptyset$, the conclusion follows from the convention for the dimension of the empty set. If $A \neq \emptyset$ and $\dim_H A = 0$, the upper bound and nonnegativity of dimension give the conclusion. Otherwise put $d = \dim_H A > 0$ and choose a sequence $0 < s_n < d$ increasing to d . Theorem 48 gives $\text{Cap}_{s_n}(A) > 0$. Also, A is a countable union of compact subsets of $[0, 1]$, so continuity makes $B(A)$ a countable union of compact sets, hence an F_σ -set.

To establish the reverse inequality when A is an F_σ -set such that $\dim_H A \leq N/2$, we note from Theorem 48 that it suffices to prove that for every $s \in (0, N/2)$ such that $\text{Cap}_s(A) > 0$, there exists an event of probability one on which $\text{Cap}_{2s}(B(A)) > 0$. Hence, let us consider such a value of s and a Borel probability measure μ with support contained in A and finite s -dimensional energy. Define

$$\mu_B(E) = \mu(\{t \in [0, 1] \mid B_t \in E\})$$

for every set $E \subseteq \mathbb{R}^N$. If we work on an event of probability one on which the function B is continuous, then μ_B is a Borel probability measure with support contained in $B(A)$. Indeed, $K = \text{supp } \mu$ is closed and contained in $[0, 1]$, hence compact, and $\mu(K) = 1$ (the complement of the support has a countable cover by open sets of zero mass). The compact set $B(K) \subseteq B(A)$ has full μ_B -mass and contains its support. The integration formula given by Theorem 29 ensures that the $2s$ -dimensional energy of μ_B satisfies

$$\mathcal{E}_{2s}(\mu_B) = \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \frac{\mu_B(dx) \mu_B(dx')}{\|x - x'\|^{2s}} = \int_0^1 \int_0^1 \frac{\mu(dt) \mu(dt')}{\|B_t - B_{t'}\|^{2s}}.$$

Finite s -dimensional energy implies that μ has no atoms: an atom of mass $a > 0$ would give the diagonal point product mass a^2 and infinite energy. By Fubini's theorem, the diagonal in $[0, 1]^2$ therefore has $\mu \otimes \mu$ -measure zero. We may omit it when using the Gaussian scaling formula below. Applying Fubini's theorem (Theorem 18) to probability measure and $\mu \otimes \mu$ gives

$$\begin{aligned} \mathbb{E}[\mathcal{E}_{2s}(\mu_B)] &= \int_0^1 \int_0^1 \mathbb{E}[\|B_t - B_{t'}\|^{-2s}] \mu(dt) \mu(dt') = \int_0^1 \int_0^1 \mathbb{E}[|t - t'|^{-s} \|X\|^{-2s}] \mu(dt) \mu(dt') \\ &= \mathbb{E}[\|X\|^{-2s}] \int_0^1 \int_0^1 \frac{\mu(dt) \mu(dt')}{|t - t'|^s} = \mathbb{E}[\|X\|^{-2s}] \mathcal{E}_s(\mu). \end{aligned}$$

The second equality follows from the fact that the increment $B_t - B_{t'}$ has the same distribution as the random vector $|t - t'|^{1/2} X$, where X is an N -dimensional standard Gaussian random vector. To check that $\mathbb{E}[\|X\|^{-2s}]$ is finite, we combine Fubini's theorem with (31) and (33), and a spherical change of variable. Accordingly, letting κ_N denote the surface of the unit sphere in $\mathbb{R}^N$, we obtain

$$\begin{aligned} \mathbb{E}[\|X\|^{-2s}] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1^2 + \cdots + x_N^2)^{-s} \frac{e^{-(x_1^2 + \cdots + x_N^2)/2}}{(2\pi)^{N/2}} dx_1 \cdots dx_N \\ &= \frac{\kappa_N}{(2\pi)^{N/2}} \int_0^{\infty} r^{-2s} e^{-r^2/2} r^{N-1} dr < \infty. \end{aligned}$$

This shows that on some event of probability one, $\mathcal{E}_{2s}(\mu_B)$ is finite, and thus $\text{Cap}_{2s}(B(A))$ is positive. Intersecting the probability-one events for the countable sequence s_n and applying Theorem 48 to $B(A)$ gives $\dim_H B(A) \geq 2s_n$ for every n , hence $\dim_H B(A) \geq 2d$. $\square$

Applying Theorem 72 with $A = [0, 1]$, we see that the range of an N -dimensional standard Brownian motion has Hausdorff dimension 2 when $N \geq 2$. In dimension $N = 1$, the range clearly has Hausdorff dimension 1: It is an interval by continuity and is nondegenerate almost surely, since $\mathbb{P}(B_1 = 0) = 0$. As a further result, which we do not prove here, in dimension $N \geq 2$, Theorem 72 can be strengthened in the sense that the event of probability one on which the conclusion holds does not depend on the set A . This uniform dimension doubling property was proven by Kaufman in 1969. It does not hold in dimension $N = 1$: the zero set has Hausdorff dimension $1/2$ almost surely, whereas its image is $\{0\}$. These additional facts are discussed in Theorem 6.4.5, p. 173, in [1], and in the discussion on p. 174.

References

- [1] C.J. Bishop and Y. Peres, *Fractals in probability and analysis*, Cambridge University Press, 2017.
- [2] N. Bourbaki, *Elements of mathematics. General topology. Part 2*, Hermann; Addison-Wesley Publishing Co., 1966.
- [3] R. Durrett, *Probability: theory and examples*, Duxbury Press, 1996.
- [4] L.C. Evans and R.F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Taylor & Francis Group, 2015.
- [5] K.J. Falconer, *The geometry of fractal sets*, Cambridge University Press, 1986.
- [6] K.J. Falconer, *Fractal geometry*, John Wiley & Sons, Ltd., 2014.
- [7] G.B. Folland, *Real analysis*, John Wiley & Sons, Inc., 1999.
- [8] P. Mattila, *Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability*, Cambridge University Press, 1995.
- [9] C. Rogers, *Hausdorff Measures*, Cambridge University Press, 1970.
- [10] W. Rudin, *Principles of mathematical analysis*, third edition, McGraw-Hill Book Co., 1976.
- [11] W. Rudin, *Real and complex analysis*, McGraw-Hill Book Co., 1987.