1

Geometric Measure Theory

1.1 Measures, Integrals, and Measure Spaces

This section has an introductory character. It collects a minimum of knowledge from abstract measure theory needed in subsequent chapters of the book. Most, commonly well known, theorems are brought up without proofs. A full account of measure theory can be found in many books, e.g., [**Coh**], [**Fr**], [**RF**].

Definition 1.1.1 A family $\mathfrak{F}$ of subsets of a set X is said to be a σ-algebra if and only if the following conditions are satisfied:

$$X \in \mathfrak{F}, \tag{1.1}$$

$$A \in \mathfrak{F} \Rightarrow A^c, \in \mathfrak{F}, \tag{1.2}$$

$$\{A_i\}_{i=1}^{\infty} \subseteq \mathfrak{F} \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}. \tag{1.3}$$

It follows from this definition that $\emptyset \in \mathfrak{F}$, i.e., that the σ-algebra $\mathfrak{F}$ is closed under countable intersections and under subtractions of sets. If (1.3) is assumed only for finite subfamilies of $\mathfrak{F}$, then $\mathcal{F}$ is called an algebra. The elements of the σ-algebra $\mathfrak{F}$ are frequently called measurable sets.

Definition 1.1.2 For any family $\mathfrak{F}$ of subsets of X, we denote by $\sigma(\mathfrak{F})$ the least σ-algebra that contains $\mathfrak{F}$, and we call it the σ-algebra generated by $\mathfrak{F}$.

Definition 1.1.3 A function on a σ-algebra $\mathfrak{F}$, $\mu\colon \mathfrak{F} \to [0, +\infty]$, is said to be σ-additive or countably additive if, for any countable subfamily $\{A_i\}_{i=1}^{\infty}$ of $\mathfrak{F}$ consisting of mutually disjoint sets, we have that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i). \tag{1.4}$$

We say then that μ is a measure.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

If we consider in (1.4) only finite families of sets, we say that μ is additive. The two notions of additivity and of σ-additivity make sense for a σ-algebra as well as for an algebra, provided that, in the case of an algebra, one considers only families $\{A_i\}_{i=1}^{\infty} \subseteq \mathfrak{F}$ such that $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{F}$. The simplest consequences of the definition of measure are the following:

$$\mu(\emptyset) = 0. \tag{1.5}$$

$$\text{If } \ A, B \in \mathfrak{F} \text{ and } \ A \subseteq B, \ \text{ then } \mu(A) \leq \mu(B). \tag{1.6}$$

$$\text{If } A_1 \subseteq A_2 \subseteq \cdots \text{ and } \{A_i\}_{i=1}^{\infty} \subseteq \mathfrak{F}, \text{ then } \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sup_i \mu(A_i) = \lim_{i\to\infty} \mu(A_i). \tag{1.7}$$

Definition 1.1.4 We say that the triple $(X, \mathfrak{F}, \mu)$ with a σ-algebra $\mathfrak{F}$ and μ, a measure on $\mathfrak{F}$, is a measure space. If $\mu(X)$ =1, the triple $(X, \mathfrak{F}, \mu)$ is called a probability space and μ is a probability measure.

Definition 1.1.5 We say that $\varphi\colon X \to \mathbb{R}$ is a measurable function if $\varphi^{-1}(J) \in \mathfrak{F}$ for every interval $J \subseteq \mathbb{R}$, equivalently for every Borel set $J \subseteq \mathbb{R}$.

Throughout the book, for any set $A \subseteq X$, we denote by $1\!\!1_A$ the characteristic function of the set A:

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

A step function is a linear combination of (finitely many) characteristic functions. It is easy to see that any nonnegative measurable function $\varphi\colon X \to \mathbb{R}$ can be represented as the pointwise limit of a monotone increasing sequence of nonnegative step functions, say

$$\varphi = \lim_{n\to\infty} \varphi_n.$$

The integral of φ against the measure μ is then defined as:

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

It is easy to see that this definition is independent of the choice of a sequence $(\varphi)_{n=1}^{\infty}$ of monotone increasing nonnegative step functions. Writing any (not necessarily nonnegative) measurable function $\varphi\colon X \to \mathbb{R}$ in its canonical form

$$\varphi = \varphi_+ - \varphi_-,$$

where

$$\varphi_+ := \max\{\varphi, 0\} \text{ and } \varphi_- := -\min\{\varphi, 0\},$$

we say that the function φ is μ integrable if

$$\int_X \varphi_+ \, d\mu < +\infty \text{ and } \int_X \varphi_- \, d\mu < +\infty.$$

We then define the integral of φ against the measure μ to be

$$\int_X \varphi \, d\mu := \int_X \varphi_+ \, d\mu - \int_X \varphi_- \, d\mu.$$

The integral of φ is also frequently denoted by

$$\mu(\varphi).$$

Since $|\varphi| = \varphi_+ - \varphi_-$, we see that φ is integrable if and only if $|\varphi|$ is, i.e., if $\int_X |\varphi| d\mu < \infty$. We then write $\varphi \in L^1(\mu)$. We now bring up two fundmental properties of integrals – theose that make integrals such powerful and convenient tools.

Theorem 1.1.6 (Lebesgue Monotone Convergence Theorem) *Suppose that $(\varphi)_{n=1}^{\infty}$ is a monotone-increasing sequence of integrable, real-valued functions on a probability space $(X, \mathfrak{F}, \mu)$. Denote its limit by φ. Then*

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

In particular, the above limit exists. As a matter of fact, it is enough to assume only that the sequence $(\varphi)_{n=1}^{\infty}$ is monotone-increasing on a measurable set whose complement is of measure zero.

Theorem 1.1.7 (Lebesgue Dominated Convergence Theorem) *Suppose that $(\phi_n)_{n=1}^{\infty}$ is a sequence of measurable, real-valued functions on a probability space $(X, \mathfrak{F}, \mu)$, that $|\phi_n| \leq g$ for an integrable function g, and that the sequence $(\phi_n)_{n=1}^{\infty}$ converges μ-a.e. to a function $\varphi \colon X \to \mathbb{R}$. Then the function ϕ is μ-integrable and*

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

More generally than $L^1(\mu)$, for every $1 \leq p < \infty$, we write

$$||\varphi||_p := \left(\int_X |\varphi|^p d\mu \right)^{\frac{1}{p}}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

and we say that φ belongs to $L^p(\mu) = L^p(X, \mathfrak{F}, \mu)$. If

$$\inf_{\mu(E)=0}\left\{\sup_{X\setminus E}|\varphi|\right\} < \infty,$$

then we denote the latter expression by $||\varphi||_\infty$, we say that the function φ is essentially bounded, and we write that $\varphi \in L^\infty$. The numbers $||\varphi||_p$, $1 \le p < \infty$, are called L^p-norms of φ. The vector spaces $L^p(X, \mathfrak{F}, \mu)$ become Banach spaces when endowed with respective norms $||\cdot||_p$.

Definition 1.1.8 A measure space $(X, \mathfrak{F}, \mu)$ and the measure μ are called

- finite if $\mu(X) < +\infty$,
- probability if $\mu(X) = 1$,
- infinite if $\mu(X) = +\infty$,
- σ-finite if the space X can be expressed as a countable union of measurable sets with finite measure μ.

Given two measures μ and ν on the same measurable space $(X, \mathfrak{F})$, we say that μ is absolutely continuous with respect to ν if, for any set A in $\mathfrak{F}$, $\nu(A) = 0$ entails $\mu(A) = 0$. The famous Radon–Nikodym Theorem gives the following.

Theorem 1.1.9 *Let $(X, \mathfrak{F})$ be a measurable space. Let μ and ν be two σ-finite measures on $(X, \mathfrak{F})$. Then the following statements are equivalent.*

(a) *μ is absolutely continuous with respect to ν ($\nu(A) = 0$ entails $\mu(A) = 0$).*
(b) $\forall_{\varepsilon>0}\,\exists_{\delta>0}\forall_{A\in\mathfrak{F}}\,[\nu(A) < \delta \Rightarrow \mu(A) < \varepsilon]$.
(c) *There exists a unique (up to sets of measure zero) measurable function $\rho \colon X \to [0, +\infty)$ such that*

$$\mu(A) = \int_A \rho\, d\nu$$

for every $A \in \mathfrak{F}$.

We then write

$$\mu \prec \nu$$

in order to indicate that a measure μ is absolutely continuous with respect to ν. The unique function $\rho \colon X \longrightarrow [0, +\infty)$ appearing in item (c) is denoted by $d\mu/d\nu$ and is called the Radon–Nikodym derivative of μ with respect to ν.

We say that two measures μ and ν on the same measurable space $(X, \mathfrak{F})$ are equivalent if each one is absolutely continuous with respect to the other. To denote this fact, we frequently write

$$\mu \asymp \nu.$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

On the other hand, there is a concept that is somehow opposite to equivalence or even to absolute continuity of measures. Namely, we say that two measures μ and ν on $(X, \mathfrak{F})$ are (mutually) singular if there exists a set $Y \in \mathfrak{F}$ such that

$$\mu(X \backslash Y) = 0 \ \text{ while } \ \nu(Y) = 0.$$

We then write that

$$\mu \perp \nu.$$

1.2 Measures on Metric Spaces: (Metric) Outer Measures and Weak* Convergence

In this section, we will show how to construct measures starting with functions of sets that are required to satisfy much weaker conditions than those defining a measure. These are called outer measures. At the end of the section, we also deal with the weak* topology of measures and Riesz Representation Theorem. Again, we refer, for example, to [**Coh**], [**Fr**], [**RF**] for complete accounts.

Definition 1.2.1 An outer measure on a set X is a function μ defined on all subsets of X taking values in $[0, \infty]$ such that

$$\mu(\emptyset) = 0; \tag{1.8}$$

$$\text{if } \ A \subseteq B, \ \text{ then } \ \mu(A) \leq \mu(B); \tag{1.9}$$

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n) \tag{1.10}$$

for any countable family $\{A_n\}_{n=1}^{\infty}$ of subsets of X.

A subset A of X is called μ-measurable or simply measurable with respect to the outer μ if and only if

$$\mu(B) \geq \mu(A \cap B) + \mu(B \backslash A) \tag{1.11}$$

for all sets $B \subseteq X$. The opposite inequality follows immediately from (1.10). One can immediately check that if $\mu(A) = 0$, then A is μ-measurable.

Theorem 1.2.2 *If μ is an outer measure on X, then the family $\mathfrak{F}$ of all μ-measurable sets is a σ-algebra and restriction of μ to $\mathfrak{F}$ is a measure.*

Proof Obviously, $X \in \mathfrak{F}$. By symmetry (1.11), $A \in \mathfrak{F}$ if and only if $A^c \in \mathcal{F}$. So the conditions (1.1) and (1.2) of the definition of σ-algebra are satisfied.

To check the condition (1.3) that $\mathfrak{F}$ is closed under a countable union, suppose that $A_1, A_2, \dots, \in \mathfrak{F}$ and let $B \subseteq X$ be any set. Applying (1.11) in turn to $A_1, A_2, \dots$, we get, for all $k \geq 1$,

$$\begin{aligned}\mu(B) &\geq \mu(B \cap A_1) + \mu(B \backslash A_1) \\ &\geq \mu(B \cap A_1) + \mu((B \backslash A_1) \cap A_2) + \mu(B \backslash A_1 \backslash A_2) \\ &\geq \dots \\ &\geq \sum_{j=1}^{k} \mu\left(\left(B \backslash \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \mu\left(B \backslash \bigcup_{j=1}^{k} A_j\right) \\ &\geq \sum_{j=1}^{k} \mu\left(\left(B \backslash \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \mu\left(B \backslash \bigcup_{j=1}^{\infty} A_j\right);\end{aligned}$$

therefore,

$$\mu(B) \geq \sum_{j=1}^{k} \mu\left(\left(B \backslash \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \mu\left(B \backslash \bigcup_{j=1}^{\infty} A_j\right). \tag{1.12}$$

Since

$$B \cap \bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} \left(B \backslash \bigcup_{i=1}^{j-1} A_i\right) \cap A_j,$$

using (1.10) we, thus, get

$$\mu(B) \geq \mu\left(\bigcup_{j=1}^{\infty} \left(B \backslash \bigcup_{i=1}^{j-1} A_i\right) \cap A_j\right) + \mu\left(B \backslash \bigcup_{j=1}^{\infty} A_j\right).$$

Hence, condition (1.3) is also satisfied and $\mathfrak{F}$ is a σ-algebra. To see that μ is a measure on $\mathcal{F}$, meaning that condition (1.4) is satisfied, consider mutually disjoint sets $A_1, A_2, \dots, \in \mathfrak{F}$ and apply (1.12) to $B = \bigcup_{j=1}^{\infty} A_j$. We get

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

Combining this with (1.10), we conclude that μ is a measure on $\mathfrak{F}$. ■

Definition 1.2.3 Let (X, ρ) be a metric space. An outer measure μ on X is said to be a metric outer measure if

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{1.13}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

for all positively separated sets $A, B \subseteq X$, i.e, those satisfying the following condition:

$$\rho(A,B) = \inf\{\rho(x,y):\ x \in A,\ y \in B\} > 0.$$

We assume the convention that $\rho(A,\emptyset) = \rho(A,\emptyset) = \infty$.

Recall that the Borel σ-algebra on X is the σ-algebra generated by open, or equivalently closed, sets. We want to show that if μ is a metric outer measure, then the family of all μ-measurable sets contains this σ-algebra. The proof is based on the following lemma.

Lemma 1.2.4 *Let μ be a metric outer measure on (X,ρ). Let $\{A_n\}_{n=1}^{\infty}$ be an ascending sequence of subsets of X. Denote $A := \bigcup_{n=1}^{\infty} A_n$. If $\rho(A_n, A\backslash A_{n+1}) > 0$ for all $n \geq 1$, then*

$$\mu(A) = \lim_{n\to\infty} \mu(A_n).$$

Proof By (1.9) it is sufficient to show that

$$\mu(A) \leq \lim_{n\to\infty} \mu(A_n). \tag{1.14}$$

If $\lim_{n\to\infty} \mu(A_n) = \infty$, there is nothing to prove. So, suppose that

$$\lim_{n\to\infty} \mu(A_n) = \sup_{n\to\infty} \mu(A_n) < \infty. \tag{1.15}$$

Let $B_1 = A_1$ and $B_n = A_n \backslash A_{n-1}$ for $n \geq 2$. If $n \geq m+2$, then $B_m \subseteq A_m$ and $B_n \subseteq A\backslash A_{n-1} \subseteq A\backslash A_{m+1}$. Thus, B_m and B_n are positively separated, and applying (1.13) we get, for every $j \geq 1$,

$$\mu\left(\bigcup_{i=1}^{j} B_{2i-1}\right) = \sum_{i=1}^{j} \mu(B_{2i-1}) \quad \text{and} \quad \mu\left(\bigcup_{i=1}^{j} B_{2i}\right) = \sum_{i=1}^{j} \mu(B_{2i}). \tag{1.16}$$

We also have, for every $n \geq 1$, that

$$\begin{aligned} \mu(A) =& \mu\left(\bigcup_{k=n}^{\infty} A_k\right) = \mu\left(A_n \cup \bigcup_{k=n+1}^{\infty} B_k\right) \\ &\leq \mu(A_n) + \sum_{k=n+1}^{\infty} \mu(B_k) \\ &\leq \lim_{l\to\infty} \mu(A_l) + \sum_{k=n+1}^{\infty} \mu(B_k). \end{aligned} \tag{1.17}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Since the sets $\bigcup_{i=1}^{j} B_{2i-1}$ and $\bigcup_{i=1}^{j} B_{2i}$ appearing in (1.16) are both contained in A_{2j}, it follows from (1.15) and (1.16) that the series $\sum_{k=1}^{\infty} \mu(B_k)$ converges. Therefore, (1.14) follows immediately from (1.17). The proof is complete. ■

Theorem 1.2.5 *If μ is a metric outer measure on (X, ρ), then all Borel subsets of X are μ-measurable.*

Proof Since the Borel sets form the least σ-algebra containing all closed subsets of X, it follows from Theorem 1.2.2 that it is enough to check (1.11) for every nonempty closed set $A \subseteq X$ and every $B \subseteq X$. For all $n \geq 1$, let $B_n = \{x \in B \backslash A \colon \rho(x, A) \geq 1/n\}$. Then $\rho(B \cap A, B_n) \geq 1/n$ and by (1.13)

$$\mu(B \cap A) + \mu(B_n) = \mu\big((B \cap A) \cup B_n\big) \leq \mu(B). \tag{1.18}$$

The sequence $\{B_n\}_{n=1}^{\infty}$ is ascending and, since A is closed, $B \backslash A = \bigcup_{n=1}^{\infty} B_n$. In order to apply Lemma 1.2.4, we shall now show that

$$\rho\big(B_n, (B \backslash A) \backslash B_{n+1}\big) > 0$$

for all $n \geq 1$. And, indeed, if $x \in (B \backslash A) \backslash B_{n+1}$, then there exists $z \in A$ with $\rho(x, z) < 1/(n+1)$. Thus, if $y \in B_n$, then

$$\rho(x, y) \geq \rho(y, z) - \rho(x, z) > \frac{1}{n} - \frac{1}{n(n+1)} > 0.$$

Applying now Lemma 1.2.4 with $A_n = B$ shows that $\mu(A \backslash B) = \lim_{n \to \infty} \mu(B_n)$. Thus, (1.11) follows from (1.18). The proof is complete. ■

This theorem, as well as many other reasons disseminated over mathematics, many of which we will encounter in this book, justifies the following definition.

Definition 1.2.6 Any measure on a metric space that is defined on its σ-algebra of Borel sets (or larger) is called a Borel measure.

Let us list the following well-known properties of finite Borel measures.

Theorem 1.2.7 *Any finite Borel measure μ on a metric space X is both outer and inner regular. Outer regularity means that*

$$\mu(A) = \inf\{\mu(G) \colon G \supseteq A \text{ and } G \text{ is open}\},$$

while inner regularity means that

$$\mu(A) = \sup\{\mu(F) \colon F \subseteq A \text{ and } F \text{ is closed}\}.$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

In addition, if the space X is completely metrizable, then the closed sets involved in the concept of inner regularity can be replaced by compact ones.

Given a metric space (X, ρ), we denote by $M(X)$ the collection of all Borel probability measures on X. We denote by $C(X)$ the vector space of all real-valued continuous functions on X and by $C_b(X)$ its vector subspace consisting of all bounded elements of $C(X)$. Let us record the following easy theorem.

Theorem 1.2.8 *If (X, ρ) is a metric space, the two measures μ and ν in $M(X)$ are equal if and only if*

$$\nu(g) = \mu(f)$$

for all functions $g \in C_b(X)$.

If X is compact, then $C(X)$ becomes a Banach space if endowed with the supremum metric. Denote by $C^*(X)$ the dual of $C(X)$. Endow $C^*(X)$ with the weak* topology. This means that

$$\text{a net } (F_\lambda)_{\lambda\in\Lambda} \text{ in } C^*(X) \text{ converges to an element } F \in C^*(X)$$

if and only if

$$\text{the net } (F_\lambda(g))_{\lambda\in\Lambda} \text{ converges to } F(g)$$

for every $g \in C(X)$. $M(X)$, the space of all Borel probability measures on X, can then be naturally viewed as a subset of $C^*(X)$: every measure $\mu \in M(X)$ induces the functional

$$C(X) \ni g \longmapsto \mu(g).$$

We will frequently use the following.

Theorem 1.2.9 *Let X be a compact metrizable space. Consider $C^*(X)$ with its weak* topology. Then*

(a) *$M(X)$ is a convex compact subset of $C^*(X)$.*
(b) *$M(X)$ is a metrizable space. In particular, proving continuity or convergence one can restrict oneself to sequences only (as opposed to nets).*
(c) *(Riesz Representation Theorem) Every nonnegative linear functional $F\colon C(X) \longrightarrow \mathbb{R}$ such that $F(1\!\!1) = 1$ is (uniquely) represented by an element in $M(X)$. More precisely, there exists $\mu \in M(X)$ such that*

$$F(g) = \mu(g)$$

for all $g \in C(X)$.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

It follows from item (c) of this theorem that the functional F considered therein is bounded. In fact, this is a quite elementary property whose short proof we leave for the reader as an exercise.

Definition 1.2.10 If X is a topological space and μ is a Borel measure on X, then the topological support of μ is defined as the set of all points $x \in X$ such that, for every open set G containing x, $\mu(G) > 0$. It is denoted by $\text{supp}(\mu)$.

The following proposition collects the basic properties of topological supports.

Proposition 1.2.11 *If X is a topological space and μ is a Borel measure on X, then the topological support* $\text{supp}(\mu)$ *of μ is a closed subset of X.*

If, in addition, X is a separable metrizable space, then the following hold.

(1) $\mu(X\backslash\text{supp}(\mu)) = 0$.
(2) *If $F \subseteq X$ is a closed set such that $\mu(X\backslash F) = 0$, then $F \supseteq \text{supp}(\mu)$.*
(3) *If, in addition, μ is a nonzero finite measure, then* $\text{supp}(\mu)$ *is the smallest closed subset of X such that $\mu(F) = \mu(X)$.*

Proof The fact that the topological support $\text{supp}(\mu)$ is closed is immediate from its definition. Let us prove item (1). If $x \in X\backslash\text{supp}(\mu)$, then there exists an open set $G_x \subseteq X$ containing X such that $\mu(G_x) = 0$. Since X is a separable metrizable space, so is $X\backslash\text{supp}(\mu)$. But then $X\backslash\text{supp}(\mu)$ a Lindelöf space. Therefore, the cover G_x, $x \in \text{supp}(\mu)$, of $X\backslash\text{supp}(\mu)$ has a countable subcover. This means that there exists a countable set $D \subseteq X\backslash\text{supp}(\mu)$ such that

$$\bigcup_{x\in D} G_x = X\backslash\text{supp}(\mu).$$

Hence,

$$\mu(X\backslash\text{supp}(\mu)) \leq \sum_{x\in D} \mu(G_x) = 0,$$

meaning that item (1) holds.

In order to prove item (2), note that, since $X\backslash F$ is an open set and its measure is equal to zero, it is contained in the complement of $\text{supp}(\mu)$. This means that item (2) holds.

Proving item (3), its hypotheses yield $\mu(X\backslash F) = 0$. Since the set F is also closed, item (2) implies that $F \supseteq \text{supp}(\mu)$, and we are done. ■

We end this section with the following easy fact, which will be frequently used throughout both volumes of the book.

Theorem 1.2.12 *If X is a compact metric space and μ is a finite Borel measure on X with full topological support, then, for every $r > 0$,*

$$M(\mu,r) := \inf\{\mu(B(x,r))\colon x \in X\} > 0. \tag{1.19}$$

Proof Since the space X is compact, there exists a finite set $F \subseteq X$ such that

$$\bigcup_{y\in F} B(y,r/2) = X.$$

Since $\operatorname{supp}(\mu) = X$, we have that

$$M := \min\{\mu(B(y,r/2)) > 0.$$

Now, if $x \in X$, there exists $y \in F$ such that $x \in B(y,r/2)$. But then $B(x,r) \supseteq B(y,r/2)$, and, therefore,

$$\mu(B(x,r)) \geq \mu(B(y,r/2)) \geq M > 0.$$

The proof of Theorem 1.2.12 is complete. ■

1.3 Covering Theorems: 4*r*, Besicovitch, and Vitali Type; Lebesgue Density Theorem

In this section, we prove first the $4r$ Covering Theorem. Following the arguments of [**MSzU**], we prove it for all metric spaces. If we do not insist on $4r$ but are content with $5r$ (which is virtually always the case), a shorter, less involved proof is possible. This can be found, for example, in [**Heino**]. Then, following [**Mat**], we will prove the Besicovitch Covering Theorem and, as its fairly straightforward consequence, the Vitali-Type Covering theorem. We finally deduce from the latter the Lebesgue Density Points Theorem. All these theorems are classical and can be found in many sources with extended discussions. More applications of covering theorems will appear in further sections of this chapter and throughout the entire book. For every ball $B := B(x,r)$, we put $r(B) = r$ and $c(B) = x$.

Theorem 1.3.1 ($4r$ Covering Theorem). *Suppose that (X,ρ) is a metric space and $\mathcal{B}$ is a family of open balls in X such that* $\sup\{r(B)\colon B \in \mathcal{B}\} < +\infty$. *Then there is a family $\mathcal{B}' \subseteq \mathcal{B}$ consisting of mutually disjoint balls such that $\bigcup_{B\in\mathcal{B}} B \subseteq \bigcup_{B\in\mathcal{B}'} 4B$. In addition, if the metric space X is separable, then $\mathcal{B}'$ is countable.*

Proof Fix an arbitrary $M > 0$. Suppose that there is a family $\mathcal{B}'_M \subseteq \mathcal{B}$ consisting of mutually disjoint balls such that

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

(a) $r(B) > M$ for all $B \in \mathcal{B}'_M$,
(b) $\bigcup_{B\in\mathcal{B}'_M} 4B \supseteq \bigcup\{B\colon B \in \mathcal{B} \text{ and } r(B) > M\}$.

We shall show that then there exists a family $\mathcal{B}''_M \subseteq \mathcal{B}$ with the following properties:

(c) $\mathcal{B}''_M \subseteq \mathfrak{F} := \{B \in \mathcal{B}\colon 3M/4 < r(B) \le M\}$,
(d) $\mathcal{B}'_M \cup \mathcal{B}''_M$ consists of mutually disjoint balls,
(e) $\bigcup_{B\in\mathcal{B}'_M\cup\mathcal{B}''_M} 4B \supseteq \bigcup\{B\colon B \in \mathcal{B} \text{ and } r(B) > 3M/4\}$.

Indeed, put

$$\mathcal{B}'''_M = \left\{B \in \mathfrak{F}\colon B \cap \bigcup_{D\in\mathcal{B}'_M} D = \emptyset\right\}. \tag{1.20}$$

Consider $B \in \mathfrak{F}\backslash\mathcal{B}'''_M$. Then there exists $D \in \mathcal{B}'_M$ such that $B \cap D \neq \emptyset$. Hence, $r(B) \le M < r(D)$ and, in consequence,

$$\rho(c(B),c(D)) < r(B) + r(D) \le M + r(D) < r(D) + r(D) = 2r(D)$$

and

$$B \subseteq B(c(D), r(B) + 2r(D)) \subseteq B(c(D), 3r(D)) = 3D \subseteq 4D.$$

Therefore,

$$\bigcup_{B\in\mathfrak{F}\backslash\mathcal{B}'''_M} B \subseteq \bigcup_{B\in\mathcal{B}'_M} 4B. \tag{1.21}$$

So, if $\mathcal{B}'''_M = \emptyset$, we are done with the proof by setting $\mathcal{B}''_M = \emptyset$. Otherwise, fix an arbitrary $B_0 \in \mathcal{B}'''_M$ and further, proceeding by transfinite induction, fix some $\mathcal{B}_\alpha \in \mathcal{B}'''_M$ such that

$$c(B_\alpha) \in c\big(\mathcal{B}'''_M\big)\backslash \bigcup_{\gamma<\alpha} \frac{8}{3} B_\gamma$$

for some some ordinal number $\gamma \ge 0$, as long as the difference on the right-hand side above is nonempty. This procedure terminates at some ordinal number λ. First, we claim that the balls $(B_\alpha)_{\alpha<\lambda}$ are mutually disjoint. Indeed, fix $0 \le \alpha < \beta < \lambda$. Then $c(B_\beta) \notin \frac{8}{3}B_\alpha$. So,

$$\rho\big(c(B_\beta), c(B_\alpha)\big) \ge \frac{8}{3} r(B_\alpha) > \frac{8}{3}\cdot\frac{3}{4}M = 2M$$

and

$$r(B_\beta) + r(B_\alpha) \le M + M = 2M.$$

Thus, $B_\beta \cap B_\alpha = \emptyset$. Now if $B \in \mathcal{B}'_M$ and $0 \le \alpha < \lambda$, then $B_\alpha \in \mathcal{B}'''_M$ and, by (1.20), $B_\alpha \cap B = \emptyset$. Thus, we proved item (d) with $\mathcal{B}''_M = \{B_\alpha\}_{\alpha<\lambda}$. Item (c) is obvious since $B_\alpha \in \mathcal{B}'''_M \subseteq \mathfrak{F}$ for all $0 \le \alpha < \lambda$. It remains to prove item (e). By the definition of λ, $c(\mathcal{B}'''_M) \subset \bigcup_{\gamma<\lambda} \frac{8}{3}B_\gamma = \bigcup_{B\in\mathcal{B}''_M} \frac{8}{3}B$. Hence, if $x \in B$ and $B \in \mathcal{B}'''_M$, then there exists $D \in \mathcal{B}''_M$ such that $c(B) \in \frac{8}{3}D$. Therefore,

$$\begin{aligned}\rho(x,c(D)) &\le \rho(x,c(B)) + \rho(c(B),c(D)) \le r(B) + \frac{8}{3}r(D)\\ &\le M + \frac{8}{3}r(D) < \frac{4}{3}r(D) + \frac{8}{3}r(D)\\ &= 4r(D).\end{aligned}$$

Thus, $x \in 4D$; consequently, $\bigcup \mathcal{B}'''_M \subseteq \bigcup_{D\in\mathcal{B}''_M} 4D$. Combining this and (1.21), we get that $\bigcup_{B\in\mathfrak{F}} B \subseteq \bigcup_{B\in\mathcal{B}'_M\cup\mathcal{B}''_M} 4B$. This and (b) immediately imply (e). The properties (c), (d), and (e) are established. Now take $S = \sup\{r(B)\colon B \in \mathcal{B}\} + 1 < +\infty$ and define inductively the sequence $(\mathcal{B}'_{(3/4)^n S})_{n=0}^\infty$ by declaring that $\mathcal{B}'_S = \emptyset$ and $\mathcal{B}'_{(3/4)^{n+1}S} = \mathcal{B}'_{(3/4)^n S} \cup \mathcal{B}''_{(3/4)^n S}$. Then

$$\mathcal{B}' = \bigcup_{n=0}^{\infty} \mathcal{B}'_{(3/4)^n S}.$$

It then follows directly from (d) and our inductive definition that $\mathcal{B}'$ consists of mutually disjoint balls. It follows from (e) that $\bigcup_{B\in\mathcal{B}'} 4B \supseteq \bigcup\{B \in \mathcal{B}\colon r(B) > 0\} = \bigcup \mathcal{B}$. The first part of our theorem is, thus, proved. The last part follows immediately from the fact that any family of mutually disjoint open subsets of a separable space is countable. ■

Remark 1.3.2 Assume the same as in Theorem 1.3.1 (no separability of X is required) and suppose that there exists a finite Borel measure μ on X such that $\mu(B) > 0$ for all $B \in \mathcal{B}'$. Then $\mathcal{B}'$ is countable.

We shall now prove the *Besicovitch Covering Theorem*. We consider it to be one of the most powerful geometric tools when dealing with some aspects of fractal sets. We can easily deduce from it two fundamental classical theorems: the *Vitali-Type Covering Theorem* and the *Lebesgue Density Points Theorem*. For the proof of the Besicovitch Covering Theorem, we introduce two concepts. First the following definition.

Definition 1.3.3 Let (X,ρ) be a metric space. A collection $\mathcal{B} = \{B(x_i,r_i)\}_{i=1}^\infty$ of open balls centered at a set $A \subseteq X$, meaning that $x_i \in A$ for all $i \ge 1$, is said to be a packing of A if and only if, for any pair $i \ne j$,

$$\rho(x_i,x_j) \ge r_i + r_j.$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

This property is not in general equivalent to the requirement that all the balls $B(x_i, r_i)$ be mutually disjoint. It is obviously so if X is a Euclidean space. We call the number

$$r(\mathcal{B}) := \sup\{r_i : i \geq 1\}$$

the radius of packing $\mathcal{B}$.

This notion has a far-reaching meaning. It is the key concept to define packing measures and dimensions, which will be done in Section 1.5. The other notion we need is the following.

For any $x \in \mathbb{R}^n$, any $0 < r \leq \infty$, and any $0 < \alpha < \pi$ by $\mathrm{Con}(x, \alpha, r)$, we will denote any solid central cone with vertex x, radius r, and angle α. That is, with the above data, for an arbitrary ray l emanating from x, we denote

$$\begin{aligned}\mathrm{Con}(x,\alpha,r) &= \mathrm{Con}(l,x,\alpha,r) \\ &:= \{y \in \mathbb{R}^n : 0 < |y - x| < r, \angle(y - x, l) \leq \alpha\} \cup \{x\}.\end{aligned}$$

The proof of Theorem 1.3.5 makes substantial use of the following obvious geometric observation.

Observation 1.3.4 Let $n \geq 1$ be an integer. Then there exists $\alpha(n) > 0$ so small that the following holds. If $x \in \mathbb{R}^n$, $0 < r < \infty$, if $z \in B(x,r)\backslash B(x,r/3)$ and if $x \in \mathrm{Con}(z, \alpha(n), \infty)$, then the set $\mathrm{Con}(z, \alpha(n), \infty)\backslash B(x, r/3)$ consists of two connected components: one of z and one of ∞. The one containing z is contained in $B(x, r)$.

Theorem 1.3.5 (Besicovitch Covering Theorem) *Let $n \geq 1$ be an integer. Then there exists an integer constant $b(n) \geq 1$ such that the following holds.*

If A is a bounded subset of $\mathbb{R}^n$, then, for any function $r: A \to (0, \infty)$, there exists $\{x_k\}_{k=1}^{\infty}$, a countable subset of A, such that the collection

$$\mathcal{B}(A, r) := \{B(x_k, r(x_k)) : k \geq 1\}$$

covers A and can be decomposed into $b(n)$ packings of A.

Proof We will construct the sequence $\{x_k : k = 1, 2, \ldots\}$ inductively. Let

$$a_0 := \sup\{r(x) : x \in A\}.$$

If $a_0 = \infty$, then one can find $x \in A$ with $r(x)$ so large that $B(x, r(x)) \supseteq A$ and the proof is finished.

If $a_0 < \infty$, choose $x_1 \in A$ so that $r(x_1) > a_0/2$. Fix $k \geq 1$. Assume that the points $x_1, x_2, \ldots, x_k$ have already been chosen. If $A \subseteq B(x_1, r(x_1)) \cup \cdots \cup B(x_k, r(x_k))$, then the selection process is finished. Otherwise, put

$$a_k := \sup \left\{ r(x) \colon x \in A \setminus \left(B(x_1, r(x_1)) \cup \cdots \cup B(x_k, r(x_k)) \right) \right\}$$

and take

$$x_{k+1} \in A \setminus \left(B(x_1, r(x_1)) \cup \cdots \cup B(x_k, r(x_k)) \right) \tag{1.22}$$

such that

$$r(x_{k+1}) > a_k/2. \tag{1.23}$$

In order to shorten notation from now on, throughout this proof we will write r_k for $r(x_k)$. By (1.22), we have that $x_l \notin B(x_k, r_k)$ for all pairs k, l with $k < l$. Hence,

$$\|x_k - x_l\| \geq r(x_k). \tag{1.24}$$

It follows from the construction of the sequence (x_k) that

$$r_k > a_{k-1}/2 \geq r_l/2. \tag{1.25}$$

therefore, $r_k/3 + r_l/3 < r_k/3 + 2r_k/3 = r_k$. By combining this and (1.24) we obtain that

$$B(x_k, r_k/3) \cap B(x_l, r_l/3) = \emptyset \tag{1.26}$$

for all pairs k, l with $k \neq l$ since then either $k < l$ or $l < k$.

Now we shall show that the balls $\{B(x_k, r_k) \colon k \geq 1\}$ cover A. Indeed, if the selection process stops after finitely many steps this claim is obvious. Otherwise, it follows from (1.26) that $\lim_{k\to\infty} r_k = 0$ and if $x \notin \bigcup_{k=1}^{\infty} B(x_k, r_k)$ for some $x \in A$, then by construction $r_k > a_{k-1}/2 \geq r(x)/2$ for every $k \geq 1$. The contradiction obtained proves that $\bigcup_{k=1}^{\infty} B(x_k, r_k) \supseteq A$.

The main step of the proof is given by the following.

Claim 1°. For every $z \in \mathbb{R}^n$ and any cone $\mathrm{Con}(z, \alpha(n), \infty)$ ($\alpha(n)$ given by Observation 1.3.4), we have that

$$\#\left\{ k \geq 1 \colon z \in B(x_k, r_k) \setminus B(x_k, r_k/3) \text{ and } x_k \in \mathrm{Con}(z, \alpha(n), \infty) \right\} \leq (12)^n.$$

Proof Denote the above set by Q. Our task is to estimate its cardinality from above. If $Q = \emptyset$, there is nothing to prove. Otherwise, let $i = \min Q$. If $k \in Q$ and $k \neq i$, then $k > i$ and, therefore, $x_k \notin B(x_i, r_i)$. Therefore, by Observation 1.3.4 applied with $x = x_i$, $r = r_i$, and by the the definition of Q, we get that $\|z - x_k\| \geq 2r_i/3$. Hence,

$$r_k \geq \|z - x_k\| \geq 2r_i/3. \tag{1.27}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

On the other hand, by (1.25), we have that $r_k < 2r_i$; therefore, $B(x_k, r_k/3) \subseteq B(z, 4r_k/3) \subseteq B(z, 8r_i/3)$. Thus, using (1.26), (1.27) and the fact that the n-dimensional volume of balls in $\mathbb{R}^n$ is proportional to the nth power of radii, we obtain that, $\#Q \leq (8r_i/3)^n/(2r_i/9)^n = 12^n$. The proof of the claim is finished. ■

Clearly, there exists an integer $c(n) \geq 1$ such that, for every $z \in \mathbb{R}^n$, the space $\mathbb{R}^n$ can be covered by at most $c(n)$ cones of the form $\mathrm{Con}(z, \alpha(n), \infty)$. Therefore, it follows from the above claim that, for every $z \in \mathbb{R}^n$,

$$\#\{k \geq 1 \colon z \in B(x_k, r_k) \backslash B(x_k, r_k/3)\} \leq c(n)(12)^n.$$

Thus, applying (1.26),

$$\#\{k \geq 1 \colon z \in B(x_k, r_k) \leq 1 + c(n)(12)^n. \tag{1.28}$$

Since the ball $\overline{B}(0, 3/2)$ is compact, it contains a finite subset P such that

$$\bigcup_{x \in P} B(x, 1/2) \supseteq \overline{B}(0, 3/2).$$

Now, for every $k \geq 1$, consider the composition of the map $\mathbb{R}^n \ni x \longmapsto r_k x \in \mathbb{R}^n$ and the translation determined by the vector from 0 to x_k. Call the image of P under this affine map P_k. Then, $\#P_k = \#P$, $P_k \subseteq \overline{B}(x_k, 3r_k/2)$, and

$$\bigcup_{x \in P_k} B(x, r_k/2) \supseteq \overline{B}(0, 3r_k/2). \tag{1.29}$$

Consider now two integers $1 \leq k < l$ such that

$$B(x_k, r_k) \cap B(x_l, r_l) \neq \emptyset. \tag{1.30}$$

Let $y \in \mathbb{R}^n$ be the only point lying on the interval joining x_l and x_k at the distance $r_k - r_l/2$ from x_k. As $x_l \notin B(x_k, r_k)$, by (1.30) we have that $\|y - x_l\| \leq r_l + r_l/2 = 3r_l/2$ and, therefore, by (1.29) there exists $z \in P_l$ such that $\|z - y\| < r_l/2$. Consequently, $z \in B(x_k, r_l/2 + r_k - r_l/2) = B(x_k, r_k)$. Thus, applying (1.28), with z being the elements of P_l, we obtain the following:

$$\#\big\{1 \leq k \leq l-1 \colon B(x_k, r_k) \cap B(x_l, r_l) \neq \emptyset\big\} \leq \#P(1 + c(n)(1)2^n) \tag{1.31}$$

for every $l \geq 1$. Putting

$$b(n) := \#P(1 + c(n)(12)^n) + 1,$$

this property allows us to decompose the set $\mathbb{N}$ of positive integers into $b(n)$ subsets $\mathbb{N}_1, \mathbb{N}_2, \ldots, \mathbb{N}_{b(n)}$ in the following inductive way. For every $k = 1, 2, \ldots, b(n)$, set $\mathbb{N}_k(b(n)) = \{k\}$ and suppose that, for every

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

$k = 1,2,\ldots,b(n)$ and some $j \geq b(n)$, the mutually disjoint families $\mathbb{N}_k(j)$ have already been defined so that

$$\mathbb{N}_1(j) \cup \cdots \cup \mathbb{N}_{b(n)}(j) = \{1,2,\ldots,j\}.$$

Then, by (1.31), there exists at least one $1 \leq k \leq b(n)$ such that $B(x_{j+1}, r_{j+1}) \cap B(x_i, r_i) = \emptyset$ for every $i \in \mathbb{N}_k(j)$. We set

$$\mathbb{N}_k(j+1) := \mathbb{N}_k(j) \cup \{j+1\}$$

and

$$\mathbb{N}_l(j+1) = \mathbb{N}_l(j)$$

for all $l \in \{1,2,\ldots,b(n)\} \backslash \{k\}$. Putting now, for every $k = 1,2,\ldots,b(n)$,

$$\mathbb{N}_k := \mathbb{N}_k(b(n)) \cup \mathbb{N}_k(b(n)+1) \cup \cdots,$$

we see from the inductive construction that these sets are mutually disjoint, that they cover $\mathbb{N}$, and that for every $k = 1,2,\ldots,b(n)$ the families of balls $\{B(x_l, r_l) \colon l \in \mathbb{N}_k\}$ are also mutually disjoint. The proof of the Besicovitch Covering Theorem is finished. ■

We would like to emphasize here that the same statement remains true, if open balls, are replaced by closed ones. It also remains true if, instead of balls, one considers n-dimensional cubes. And, in this latter case, it is even better: namely, the proof based on the same idea is technically considerably easier. There are further, frequently useful, generalizations, especially a theorem of Morse. The reader is advised to consult the book [**Gu**] by Guzman on such topics.

As we have already mentioned, we can easily deduce from the Besicovitch Covering Theorem some other fundamental facts.

Theorem 1.3.6 (Vitali-Type Covering Theorem) *Let μ be a probability Borel measure on $\mathbb{R}^n$, let $A \subset \mathbb{R}^n$ be a Borel set, and let $\mathcal{B}$ be a family of closed balls such that each point of A is the center of arbitrarily small balls of $\mathcal{B}$, i.e.,*

$$\inf\{r \colon B(x,r) \in \mathcal{B}\} = 0$$

for all $x \in A$. Then there exists a countable (finite or infinite) collection $\mathcal{B}(A)$ of mutually disjoint balls in $\mathcal{B}$ such that

$$\mu\left(A \backslash \bigcup \{B \in \mathcal{B}(A)\}\right) = 0.$$

Proof We assume that A is bounded, leaving the unbounded case to the reader. We may assume that $\mu(A) > 0$. The measure μ restricted to a compact

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

ball $B(0,R)$ such that $A \subset B(0,R/2)$ is Borel, hence regular. Hence, there exists an open set $U \subset \mathbb{R}^n$ containing A such that

$$\mu(U) \leq \big(1 + (4b(n))^{-1}\big)\mu(A),$$

where $b(n)$ is as in the Besicovitch Covering Theorem 1.3.5. By that theorem applied for closed balls we can decompose $\mathcal{B}$ in packings $\mathcal{B}_1, \ldots, \mathcal{B}_{b(n)}$ of A contained in U, i.e., each $\mathcal{B}_i$ consists of disjoint balls and

$$A \subset \bigcup_{i=1}^{b(n)} \bigcup \mathcal{B}_i \subset U.$$

Then $\mu(A) \leq \sum_{i=1}^{b(n)} \mu\big(\bigcup \mathcal{B}_i\big)$; consequently, there exists an i such that

$$\mu(A) \leq b(n)\mu\left(\bigcup \mathcal{B}_i\right).$$

Further, for some finite subfamily $\mathcal{B}'_i$ of $\mathcal{B}_i$,

$$\mu(A) \leq 2b(n)\mu\left(\bigcup \mathcal{B}'_i\right).$$

Letting $A_1 = A \backslash \big(\bigcup \mathcal{B}'_i\big)$, we get

$$\begin{aligned}\mu(A_1) &\leq \mu\left(U \backslash \bigcup B'_i\right) = \mu(U) - \mu\left(\bigcup B'_i\right)\\ &\leq \left(1 + \frac{1}{4}(b(n))^{-1} - \frac{1}{2}(b(n))^{-1}\right)\mu(A)\\ &= u\mu(A)\end{aligned}$$

with $u := 1 - \frac{1}{4}(b(n))^{-1} < 1$. Now consider A_1 in the role of A before. Since

$$A_1 \subset \mathbb{R}^n \backslash \left(\bigcup \mathcal{B}'_i\right),$$

which is open, we find a packing, playing the role of B'_i contained in

$$\mathbb{R}^n \backslash \left(\bigcup \mathcal{B}'_i\right),$$

so disjoint from $\bigcup \mathcal{B}'_i$. We then get the measure of a noncovered remnant bounded above by $u\mu(A_1) \leq u^2\mu(A)$. We can continue, consecutively constructing packings that exhaust the whole set A except at most a set of measure 0. The proof is complete. ■

Now we shall prove two quite straightforward consequences of the Besicovitch Covering Theorem (Theorem 1.3.5), the first one being the celebrated, and to some extent counter-intuitive, Density Points Theorem. It in fact follows from the Vitali-Type Covering Theorem (Theorem 1.3.6), which itself is a consequence of the Besicovitch Covering Theorem.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Theorem 1.3.7 (Lebesgue Density Theorem) *Let μ be a probability Borel measure on $\mathbb{R}^n$, $n \in \mathbb{N}$, and $A \subset \mathbb{R}^n$ be a Borel set. Then the limit*

$$\lim_{r \to 0} \frac{\mu(A \cap B(x,r))}{\mu(B(x,r))}$$

exists and is equal to 1 for μ-almost every point $x \in A$.

Proof First of all, for every Borel set $B \subseteq \mathbb{R}^n$ and every $x \in \mathbb{R}^n$, we obviously have that

$$\lim_{s \nearrow r} \mu(B \cap B(x,s)) = \mu(B \cap B(x,r))$$

and

$$\lim_{s \searrow r} \mu(B \cap B(x,s) \geq \mu(B \cap B(x,r))).$$

Therefore, the function

$$\mathbb{R}^n \ni x \longmapsto \mu(B \cap B(x,r)) \in \mathbb{R}$$

is lower semi-continuous, thus Borel measurable. Hence, the function

$$\mathbb{R}^n \ni x \longmapsto \frac{\mu(A \cap B(x,r))}{\mu(B(x,r))} \in \mathbb{R}$$

is also Borel measurable. Furthermore, since

$$\lim_{\mathbb{Q} \ni s \nearrow r} \mu(B \cap B(x,s)) = \mu(B \cap B(x,r)),$$

it follows that the set of points $x \in \mathbb{R}^n$ for which the limit

$$\lim_{r \to 0} \frac{\mu(A \cap B(x,r))}{\mu(B(x,r))} \tag{1.32}$$

exists is the same as the set of points $x \in \mathbb{R}^n$ for which the limit

$$\lim_{\mathbb{Q} \ni r \to 0} \frac{\mu(A \cap B(x,r))}{\mu(B(x,r))}$$

exists. Since the set $\mathbb{Q}$ of rational numbers is countable, we, thus, conclude that the set of points in X for which in (1.32) exists is Borel measurable.

Seeking contradiction, suppose now that the set of points in A where this limit is either not equal to 1 or does not exist has positive measure μ. Then there exists $0 \leq a < 1$ and Borel $A' \subset A$ of positive measure μ such that, for every $x \in A'$, there exists a sequence $(r_i(x))_{i=1}^\infty$ of positive radii converging to 0 such that

$$\frac{\mu\big(A' \cap B(x, r_i(x))\big)}{\mu\big(B(x, r_i(x))\big)} < a$$

for all $i \geq 1$. Given an open set U containing A, let

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

$$\mathcal{B}_U := \left\{B(x, r_i(x)) \colon x \in A',\ B(x, r_i(x)) \subseteq U\right\}.$$

Then let $\mathcal{B}_U(A')$ be the corresponding collection of balls whose existence is asserted in the Vitali-Type Covering Theorem (Theorem 1.3.6). Then

$$\mu(A') = \sum_{B \in \mathcal{B}_U(A')} \mu(A' \cap B) \le a \sum_{B \in \mathcal{B}_U(A')} \mu(B) \le a\mu(U).$$

Since measure μ is regular, this yields $\mu(A') \le a\mu(A')$. This contradiction finishes the proof. ■

Every point in the set A for which the assertion of the Lebesgue Density Theorem holds will be called a Lebesgue density point of A with respect to the measure μ.

The second consequence of the Besicovitch Covering Theorem (Theorem 1.3.5), which we have mentioned above, is the following technical, but very useful and frequently applied, lemma, which is suitable for proving that one given measure is absolutely continuous with respect to the other. We follow the proof from [**DU2**].

Lemma 1.3.8 *Let μ and ν be Borel probability measures on X, a bounded subset of a Euclidean space $\mathbb{R}^d$, $d \ge 1$. Suppose that there is a constant $M > 0$ and, for every point $x \in Y$, there is a converging to zero sequence $\left(r_j(x)\right)_{i=0}^{\infty}$ of positive radii such that, for all $j \ge 1$ and all $x \in X$,*

$$\mu(B(x, r_j(x)) \le M\nu(B(x, r_j(x)).$$

Then the measure μ is absolutely continuous with respect to ν and the Radon–Nikodym derivative satisfies

$$d\mu/d\nu \le Mb(d),$$

where $b(d)$ is the constant coming from the Besicovitch Covering Theorem, i.e., Theorem 1.3.5.

Proof Consider an arbitrary Borel set $E \subseteq X$ and fix $\varepsilon > 0$. Since $\lim_{j\to\infty} r_j(x) = 0$ and measure ν is regular, for every $x \in E$ there exists a radius $r(x)$ of the form $r_j(x)$ such that

$$\nu\left(\bigcup_{x\in E} B_e(x, r(x)) \setminus E\right).$$

Now, by the Besicovitch Covering Theorem (Theorem 1.3.5) we can choose a countable subcover $\{B(x_i, r_i(x))\}_{i=1}^{\infty}$ from the cover $\{B(x_i, r_i(x))\}_{x\in E}$ of E, of multiplicity bounded above by $b(d)$. Therefore, we obtain that

$$\mu(E) \leq \sum_{i=1}^{\infty} \mu(B(x_i, r_i(x))) \leq M \sum_{i=1}^{\infty} \nu(B(x_i, r_i(x)))$$
$$\leq Mb(d)\nu\left(\bigcup_{i=1}^{\infty} B(x_i, r_i(x))\right)$$
$$\leq Mb(d)(\varepsilon + \nu(E)).$$

Letting $\varepsilon \searrow 0$, we, thus, obtain that $\mu(E) \leq Mb(d)\nu(E)$. Therefore, μ is absolutely continuous with respect to ν with the Radon–Nikodym derivative bounded above by $Mb(d)$. ■

1.4 Conditional Expectations and Martingale Theorems

The content of this section belongs to probability theory rather than to classical measure theory. Its culmination (for us), i.e., Theorem 1.4.11, is, however, similar to the Lebesgue Density Theorem, i.e., Theorem 1.3.7, so is natural to place it here. This chapter is about conditional expectations and martingales and is closely modeled on a chapter form Billingsley's book [**Bil2**].

We start with conditional expectations. Let $(X, \mathfrak{F}, \mu)$ be a probability space. Let $\mathfrak{D}$ be a sub-σ-algebra of $\mathfrak{F}$. Let

$$\phi : X \longrightarrow \mathbb{R}$$

be a measurable function, integrable with respect to the measure μ. We denote by

$$E(\phi|\mathfrak{D}) = E_\mu(\phi|\mathfrak{D})$$

the (conditional) expected value of ϕ with respect to the σ-algebra $\mathfrak{D}$. This is the only function (up to sets of measure zero) that is measurable with respect to the σ-algebra $\mathfrak{D}$ such that

$$\int_D E_\mu(\phi|\mathfrak{D}) d\mu = \int_D \phi \, d\mu$$

for every set $D \in \mathfrak{D}$. Its existence for nonnegative integrable functions is a straightforward consequence of the Radon–Nikodym Theorem. In the general case, one sets

$$E_\mu(\phi|\mathfrak{D}) := E_\mu(\phi_+|\mathfrak{D}) - E_\mu(\phi_-|\mathfrak{D}).$$

Uniqueness is obvious.

Conditional expectations exhibit several natural properties. We list the most basic ones below. Their proofs are straightforward and are omitted.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Proposition 1.4.1 *Let $(X, \mathcal{A}, \mu)$ be a probability space, $\mathcal{B}$ and $\mathcal{C}$ denote some sub-σ-algebras of $\mathcal{A}$ and $\varphi \in L^1(X, \mathcal{A}, \mu)$. Then the following hold.*

(a) *If $\varphi \geq 0$ μ-a.e., then*

$$E(\varphi|\mathcal{B}) \geq 0 \quad \mu\text{-}a.e.$$

(b) *If $\varphi_1 \geq \varphi_2$ μ-a.e., then*

$$E(\varphi_1|\mathcal{B}) \geq E(\varphi_2|\mathcal{B}) \ \ \mu\text{-}a.e.$$

(c) $\left|E(\varphi|\mathcal{B})\right| \leq E\big(|\varphi| \,\big|\, \mathcal{B}\big)$.

(d) *The functional $E(\cdot|\mathcal{B})$ is linear. In other words, for any $c_1, c_2 \in \mathbb{R}$ and $\varphi_1, \varphi_2 \in L^1(X, \mathcal{A}, \mu)$, we have that*

$$E(c_1\varphi_1 + c_2\varphi_2|\mathcal{B}) = c_1\, E(\varphi_1|\mathcal{B}) + c_2\, E(\varphi_2|\mathcal{B}).$$

(e) *If φ is already $\mathcal{B}$-measurable, then $E(\varphi|\mathcal{B}) = \varphi$. In particular, we have that*

$$E\big(E(\varphi|\mathcal{B})\big|\mathcal{B}\big) = E(\varphi|\mathcal{B}).$$

Also, if $\varphi = c \in \mathbb{R}$ is a constant function, then $E(\varphi|\mathcal{B}) = \varphi = c$.

(f) *If $\mathcal{B} \supseteq \mathcal{C}$, then*

$$E\big(E(\varphi|\mathcal{B})\big|\mathcal{C}\big) = E(\varphi|\mathcal{C}).$$

We will now determine the conditional expectations of an arbitrary integrable function φ with respect to various sub-σ-algebras that are of particular interest and are simple enough.

Example 1.4.2 Let $(X, \mathcal{A}, \mu)$ be a probability space. The family $\mathcal{B}$ of all measurable sets that are either of null or of full measure constitutes a sub-σ-algebra of $\mathcal{A}$. Let $\varphi \in L^1(X, \mathcal{A}, \mu)$. Since $E(\varphi|\mathcal{B})$ is $\mathcal{B}$-measurable,

$$E(\varphi|\mathcal{B})^{-1}(\{t\}) \in \mathcal{B}$$

for each $t \in \mathbb{R}$, meaning that the set $E(\varphi|\mathcal{B})^{-1}(\{t\})$ is either of measure zero or of measure 1. Also bear in mind that

$$X = E(\varphi|\mathcal{B})^{-1}(\mathbb{R}) = \bigcup_{t \in \mathbb{R}} E(\varphi|\mathcal{B})^{-1}(\{t\}).$$

Since the above union consists of mutually disjoint sets of measure zero and 1, it follows that only one of these sets can be of measure 1. In other words, there exists a unique $t \in \mathbb{R}$ such that

$$E(\varphi|\mathcal{B})^{-1}(\{t\}) = A$$

for some $A \in \mathcal{A}$ with $\mu(A) = 1$. Because the function $E(\varphi|\mathcal{B})$ is unique up to a set of measure zero, we may assume without loss of generality that $A = X$. Hence, $E(\varphi|\mathcal{B})$ is a constant function. Therefore,

$$E(\varphi|\mathcal{B}) = \int_X E(\varphi|\mathcal{B})d\mu = \int_X \varphi\, d\mu.$$

Example 1.4.3 Let $(X, \mathcal{A})$ be a measurable space and α be a countable measurable partition of X. The sub-σ-algebra $\sigma(\alpha)$ of $\mathcal{A}$ generated by α is the family of all sets which can be represented as a union of elements of α. When α is finite, so is $\sigma(\alpha)$. When α is countably infinite, $\sigma(\alpha)$ is uncountable; in fact, it is of cardinality continuum. Let μ be a probability measure on $(X, \mathcal{A})$. Let $\varphi \in L^1(X, \mathcal{A}, \mu)$. Since $E(\varphi|\sigma(\alpha))$ is $\mathcal{B}$-measurable,

$$E(\varphi|\mathcal{B})^{-1}(\{t\}) \in \sigma(\alpha)$$

for each $t \in \mathbb{R}$, i.e.,

$$E\big(\varphi|\sigma(\alpha)\big)^{-1}(\{t\}) \in \sigma(\alpha).$$

This means that the set $E\big(\varphi|\sigma(\alpha)\big)^{-1}(\{t\})$ is a union of elements of α. This further means that the conditional expectation function $E(\varphi|\sigma(\alpha))$ is constant on each element of α. Let $A \in \alpha$. If $\mu(A) = 0$, then $E(\varphi|\sigma(\alpha))|_A = 0$. Otherwise,

$$E(\varphi|\sigma(\alpha))|_A = \frac{1}{\mu(A)} \int_A E(\varphi|\sigma(\alpha))d\mu = \frac{1}{\mu(A)} \int_{A_n} \varphi\, d\mu. \tag{1.33}$$

In summary, the conditional expectation $E(\varphi|\mathcal{B})$ of a function φ with respect to a sub-σ-algebra generated by a countable measurable partition is constant on each element of that partition. More precisely, on any given element of the partition, $E(\varphi|\mathcal{B})$ is equal to the mean value of φ on that element. In particular, if α is a trivial partition, i.e., consisting of sets of measure zero and 1 only, then

$$E(\varphi|\sigma(\alpha)) = \int_X \varphi\, d\mu \quad \mu\text{-a.e.} \tag{1.34}$$

The next result is a special case of a theorem originally due to Doob and is commonly called the Martingale Convergence Theorem. In order to discuss it, we first define the martingale itself.

Definition 1.4.4 Let $(X, \mathcal{A}, \mu)$ be a probability space. Let $(\mathcal{A}_n)_{n=1}^{\infty}$ be a sequence of sub-σ-algebras of $\mathcal{A}$. Let also $(\varphi_n : X \longrightarrow \mathbb{R})_{n=1}^{\infty}$ be a sequence of random variables, i.e., a sequence of $\mathcal{A}$-measurable functions. The sequence

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

$$\big((\varphi_n, \mathcal{A}_n)\big)_{n=1}^{\infty}$$

is called a *martingale* if and only if the following conditions are satisfied:

(a) $(\mathcal{A}_n)_{n=1}^{\infty}$ is an ascending sequence, i.e., $\mathcal{A}_{n+1} \supseteq \mathcal{A}_n$ for all $n \in \mathbb{N}$.
(b) φ_n is $\mathcal{A}_n$-measurable for all $n \in \mathbb{N}$.
(c) $\varphi_n \in L^1(\mu)$ for all $n \in \mathbb{N}$.
(d) $E(\varphi_{n+1}|\mathcal{A}_n) = \varphi_n$ μ-a.e. for all $n \in \mathbb{N}$.

The main, and frequently referred to as the simplest, convergence theorem concerning martingales is this.

Theorem 1.4.5 (Martingale Convergence Theorem) *Let $(X, \mathcal{A}, \mu)$ be a probability space. If $((\varphi_n, \mathcal{A}_n))_{n=1}^{\infty}$ is a martingale such that*

$$\sup\{\|\varphi_n\|_1 : n \in \mathbb{N}\} < +\infty,$$

then there exists $\widehat{\varphi} \in L^1(X, \mathcal{A}, \mu)$ such that

$$\lim_{n\to\infty} \varphi_n(x) = \widehat{\varphi}(x) \quad \textit{for } \mu\textit{-a.e. } x \in X$$

and

$$\|\widehat{\varphi}\|_1 \le \sup\{\|\varphi_n\|_1 : n \in \mathbb{N}\} < +\infty.$$

This is a special case of Theorem 35.5 in Billingsley's book [**Bil2**], proved therein. Its proof is just too long and too involved to be reproduced here. We omit it.

One natural martingale is formed by the conditional expectations of a function with respect to an ascending sequence of sub-σ-algebras.

Proposition 1.4.6 *Let $(X, \mathcal{A}, \mu)$ be a probability space and let $(\mathcal{A}_n)_{n=1}^{\infty}$ be an ascending sequence of sub-σ-algebras of $\mathcal{A}$. For any $\varphi \in L^1(X, \mathcal{A}, \mu)$, the sequence*

$$\big((E(\varphi|\mathcal{A}_n), \mathcal{A}_n)\big)_{n=1}^{\infty}$$

is a martingale.

Proof Indeed, set

$$\varphi_n := E(\varphi|\mathcal{A}_n)$$

for all $n \in \mathbb{N}$. Condition (a) in Definition 1.4.4 is automatically fulfilled. Conditions (b) and (c) follow from the very definition of the conditional

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

expectation function. Regarding condition (d), a straightforward application of Proposition 1.4.1(f) gives

$$E(\varphi_{n+1}|\mathcal{A}_n) = E\big(E(\varphi|\mathcal{A}_{n+1})|\mathcal{A}_n\big) = E(\varphi|\mathcal{A}_n) = \varphi_n$$

μ-a.e. for all $n \in \mathbb{N}$. So $\big((E(\varphi|\mathcal{A}_n), \mathcal{A}_n)\big)_{n=1}^{\infty}$ is a martingale. ■

With the hypotheses of this proposition, by using Proposition 1.4.1(c), we see that

$$\sup_{n\in\mathbb{N}} \|\varphi_n\|_1 = \sup_{n\in\mathbb{N}} \int_X \big|E(\varphi|\mathcal{A}_n)\big| d\mu \le \sup_{n\in\mathbb{N}} \int_X E\big(|\varphi|\big|\mathcal{A}_n\big) d\mu = \int_X |\varphi|\, d\mu < \infty.$$

According to Theorem 1.4.5, there, thus, exists $\widehat{\varphi} \in L^1(X, \mathcal{A}, \mu)$ such that

$$\lim_{n\to\infty} E(\varphi|\mathcal{A}_n)(x) = \widehat{\varphi}(x) \text{ for } \mu\text{-a.e. } x \in X \qquad \text{and} \qquad \|\widehat{\varphi}\|_1 \le \|\varphi\|_1.$$

What is $\widehat{\varphi}$? This is the question we will address now. For this we need the concept of uniform integrability and the convergence theorem that it entails.

Definition 1.4.7 Let $(X, \mathcal{A}, \mu)$ be a measure space. A sequence of measurable functions $(f_n)_{n=1}^{\infty}$ is called uniformly integrable if and only if

$$\lim_{M\to\infty} \sup_{n\in\mathbb{N}} \int_{\{|f_n|\ge M\}} |f_n|\, d\mu = 0.$$

The following theorem is classical in measure theory. It is proved, for example, as Theorem 16.14 in Billingsley's book [**Bil2**].

Theorem 1.4.8 *Let $(X, \mathcal{A}, \mu)$ be a finite measure space and $(f_n)_{n=1}^{\infty}$ a sequence of measurable functions that converges pointwise μ-a.e. to a function f.*

(a) *If $(f_n)_{n=1}^{\infty}$ is uniformly integrable, then $f_n \in L^1(\mu)$ for all $n \in \mathbb{N}$ and $f \in L^1(\mu)$. Moreover,*

$$\lim_{n\to\infty} \|f_n - f\|_1 = 0 \quad \textit{and} \quad \lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu.$$

(b) *If $f, f_n \in L^1(\mu)$ and $f_n \ge 0$ μ-a.e. for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} \int_X f_n\, d\mu = \int_X f\, d\mu$ implies that $(f_n)_{n=1}^{\infty}$ is uniformly integrable.*

We shall now prove the uniform integrability of the martingale appearing in Proposition 1.4.6.

Lemma 1.4.9 *Let $(X, \mathcal{A}, \mu)$ be a probability space and $(\mathcal{A}_n)_{n=1}^{\infty}$ be a sequence of sub-σ-algebras of $\mathcal{A}$. Then, for every $\varphi \in L^1(X, \mathcal{A}, \mu)$, the sequence $(E(\varphi|\mathcal{A}_n))_{n=1}^{\infty}$ is uniformly integrable.*

Proof Without loss of generality, we may assume that $\varphi \geq 0$. Let $\varepsilon > 0$. Since the measure ν on $(X, \mathcal{A})$ given by

$$\nu(A) := \int_A \varphi \, d\mu$$

is absolutely continuous with respect to μ, it follows from the Radon–Nikodym Theorem (Theorem 1.1.9) that there exists $\delta > 0$ such that

$$A \in \mathcal{A}, \ \mu(A) < \delta \quad \Longrightarrow \quad \int_A \varphi \, d\mu < \varepsilon. \tag{1.35}$$

Consider any

$$M > \frac{1}{\delta} \int_X \varphi \, d\mu.$$

For each $n \in \mathbb{N}$, let

$$X_n(M) := \big\{ x \in X : E(\varphi|\mathcal{A}_n)(x) \geq M \big\}.$$

Observe that $X_n(M) \in \mathcal{A}_n$ since $E(\varphi|\mathcal{A}_n)$ is $\mathcal{A}_n$-measurable. Therefore, by Tchebyschev's Inequality, we get that

$$\mu(X_n(M)) \leq \frac{1}{M} \int_{X_n(M)} E(\varphi|\mathcal{A}_n) d\mu = \frac{1}{M} \int_{X_n(M)} \varphi \, d\mu \leq \frac{1}{M} \int_X \varphi \, d\mu < \delta$$

for all $n \in \mathbb{N}$. Consequently, by (1.35),

$$\int_{X_n(M)} E(\varphi|\mathcal{A}_n) d\mu = \int_{X_n(M)} \varphi \, d\mu < \varepsilon$$

for all $n \in \mathbb{N}$. Thus,

$$\sup_{n \in \mathbb{N}} \int_{\{E(\varphi|\mathcal{A}_n) \geq M\}} E(\varphi|\mathcal{A}_n) d\mu \leq \varepsilon.$$

Therefore,

$$\lim_{M \to \infty} \sup_{n \in \mathbb{N}} \int_{\{E(\varphi|\mathcal{A}_n) \geq M\}} E(\varphi|\mathcal{A}_n) d\mu = 0,$$

i.e., $(E(\varphi|\mathcal{A}_n))_{n=1}^{\infty}$ is uniformly integrable. ■

Theorem 1.4.10 (Martingale Convergence Theorem for Conditional Expectations) *Let $(X, \mathcal{A}, \mu)$ be a probability space and $\varphi \in L^1(X, \mathcal{A}, \mu)$. Let $(\mathcal{A}_n)_{n=1}^{\infty}$ be an ascending sequence of sub-σ-algebras of $\mathcal{A}$ and*

$$\mathcal{A}_\infty := \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{A}_n \right).$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Then

$$\lim_{n\to\infty} E(\varphi|\mathcal{A}_n) = E(\varphi|\mathcal{A}_\infty) \;\; \mu\text{-}a.e. \; on \; X$$

and

$$\lim_{n\to\infty} \left\| E(\varphi|\mathcal{A}_n) - E(\varphi|\mathcal{A}_\infty) \right\|_1 = 0.$$

Proof Let

$$\varphi_n := E(\varphi|\mathcal{A}_n).$$

It follows from Proposition 1.4.6 and Lemma 1.4.9 that $((\varphi_n, \mathcal{A}_n))_{n=1}^\infty$ is a uniformly integrable martingale such that

$$\lim_{n\to\infty} \varphi_n = \widehat{\varphi} \;\; \mu\text{-a.e. on } X$$

for some $\widehat{\varphi} \in L^1(X, \mathcal{A}, \mu)$. For all $n \in \mathbb{N}$, the function φ_n is $\mathcal{A}_\infty$-measurable since it is $\mathcal{A}_n$-measurable and $\mathcal{A}_n \subseteq \mathcal{A}_\infty$. Thus, $\widehat{\varphi}$ is $\mathcal{A}_\infty$-measurable, too. Moreover, it follows from Theorem 1.4.8 that

$$\lim_{n\to\infty} \|\varphi_n - \widehat{\varphi}\|_1 = 0 \quad \text{and} \quad \lim_{n\to\infty} \int_A \varphi_n \, d\mu = \int_A \widehat{\varphi} \, d\mu$$

for all $A \in \mathcal{A}$. Therefore, it just remains to show that

$$\widehat{\varphi} = E(\varphi|\mathcal{A}_\infty).$$

Let $k \in \mathbb{N}$ and $A \in \mathcal{A}_k$. If $n \geq k$, then $A \in \mathcal{A}_n \subseteq \mathcal{A}_\infty$ and, thus,

$$\int_A \varphi_n \, d\mu = \int_A E(\varphi|\mathcal{A}_n) d\mu = \int_A \varphi \, d\mu = \int_A E(\varphi|\mathcal{A}_\infty) d\mu.$$

Letting $n \to \infty$, this yields

$$\int_A \widehat{\varphi} \, d\mu = \int_A E(\varphi|\mathcal{A}_\infty) d\mu._k.$$

Since k was arbitrary, this entails

$$\int_B \widehat{\varphi} \, d\mu = \int_B E(\varphi|\mathcal{A}_\infty) d\mu$$

for all $B \in \bigcup_{k=1}^\infty \mathcal{A}_k$. Finally, since $\bigcup_{k=1}^\infty \mathcal{A}_k$ is a π-system generating $\mathcal{A}_\infty$ and both $\widehat{\varphi}$ and $E(\varphi|\mathcal{A}_\infty)$ are $\mathcal{A}_\infty$-measurable, we conclude that

$$\widehat{\varphi} = E(\varphi|\mathcal{A}_\infty) \quad \mu\text{-a.e. in } X.$$

■

Recall that countable measurable partitions of a measurable space are defined and systematically treated in Section 6.1; they form the key concept for all of Chapter 6. As an immediate consequence of the Martingale Convergence Theorem for Conditional Expectations, i.e., Theorem 1.4.10 and (1.33), we get the following theorem, which is somewhat similar to the Lebesgue Density Theorem, i.e., Theorem 1.3.7 from the previous section.

Theorem 1.4.11 *Let $(X, \mathcal{A}, \mu)$ be a probability space and $(\alpha_n)_{n=1}^{\infty}$ be a sequence of finer and finer (more precisely, α_{n+1} is finer than α_n for all $n \geq 1$) countable measurable partitions of X which generates the σ-algebra $\mathcal{A}$, i.e., $\sigma\left(\bigcup_{n\geq 1} \alpha_n\right) = \mathcal{A}$. Then, for every set $A \in \mathcal{A}$ and for μ-a.e. $x \in A$, we have that*

$$\lim_{n\to\infty} \frac{\mu(A \cap \alpha_n(x))}{\mu(\alpha_n(x))} = 1\!\!1_A(x) = 1.$$

1.5 Hausdorff and Packing Measures: Hausdorff and Packing Dimensions

In this section, we introduce the basic geometric concepts on metric spaces. These are Hausdorff measures, Hausdorff dimensions, packing measures, and packing dimensions. We prove their fundamental properties. While Hausdorff measures and Hausdorff dimensions were introduced quite early, in 1919, by Felix Hausdorff in [**H**], it took several decades more for packing measures and packing dimensions to be defined. It was done in stages in [**Tr**], [**TT**], and [**Su6**]. There are now plenty of books on these concepts; we refer the reader to, for example, [**Fal1**], [**Fal12**], [**Fal13**], [**Mat**], and [**PU2**]. The classical book [**Ro**] by C. A. Rogers is also interesting, and not only for historical reasons, appearing for the first time in 1970. The 1998 edition is particularly interesting because of the comments by Falconer that it contains.

Let $\varphi\colon [0, +\infty) \longrightarrow [0, +\infty)$ be a function with the following properties:

- φ is nondecreasing, meaning that $s \leq t \Rightarrow \varphi(s) \leq \varphi(t)$.
- $\varphi(0) = 0$ and φ is continuous at 0.
- $\varphi((0, +\infty)) \subseteq (0, +\infty)$.

Any function $\varphi\colon [0, +\infty) \longrightarrow [0, +\infty)$ with such properties is referred to in what follows as a gauge function.

Let (X, ρ) be a metric space. For every $\delta > 0$, define

$$\mathrm{H}_{\varphi}^{\delta}(A) := \inf\left\{\sum_{i=1}^{\infty} \varphi(\mathrm{diam}(U_i))\right\}, \tag{1.36}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

where the infimum is taken over all countable covers $\{U_i\}_{i=1}^{\infty}$ of A of diameter not exceeding δ.

We shall check that, for every $\delta > 0$, $\mathrm{H}_\varphi^\delta$ is an outer measure. Conditions (1.8) and (1.9) of Definition 1.2.1, defining the concept of outer measures, are obviously satisfied with $\mu = \mathrm{H}_\varphi^\delta$. To verify (1.10), let $\{A_n\}_{n=1}^{\infty}$ be a countable family of subsets of X. Fix $\varepsilon >$. Then, for every $n \geq 1$, we can find a countable cover $\{U_i^n\}_{i=1}^{\infty}$ of A_n with diameters not exceeding δ such that

$$\sum_{i=1}^{\infty} \varphi(\mathrm{diam}(U_i)) \leq \mathrm{H}_\varphi^\delta(A_n) + \frac{\varepsilon}{2^n}.$$

Then the family $\{U_i^n : i, n \geq 1\}$ covers $\bigcup_{n=1}^{\infty} A_n$ and

$$\begin{aligned}\mathrm{H}_\varphi^\delta\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \sum_{n=1}^{\infty}\sum_{i=1}^{\infty} \varphi(\mathrm{diam}(U_i^n)) \leq \sum_{n=1}^{\infty}\left(\mathrm{H}_\varphi^\delta(A_n) + \frac{\varepsilon}{2^n}\right)\\ &= \sum_{n=1}^{\infty} \mathrm{H}_\varphi^\delta(A_n) + \varepsilon.\end{aligned}$$

Thus, letting $\varepsilon \searrow 0$, (1.10) follows, proving that $\mathrm{H}_\varphi^\delta$ is an outer measure. Define

$$\mathrm{H}_\varphi(A) := \sup_{\delta > 0}\left\{\mathrm{H}_\varphi^\delta(A)\right\} = \lim_{\delta \to 0} \mathrm{H}_\varphi^\delta(A). \tag{1.37}$$

The limit exists since $\mathrm{H}_\varphi^\delta(A)$ increases as δ decreases, though it may happen to be infinite. Since all $\mathrm{H}_\varphi^\delta$ are outer measures. It is therefore immediate that H_φ is an outer measure too. Moreover, H_φ is a metric measure, since if A and B are two positively separated sets in X, then no set of diameter less than $\rho(A, B)$ can intersect both A and B. Consequently,

$$\mathrm{H}_\varphi^\delta(A \cup B) = \mathrm{H}_\varphi^\delta(A) + \mathrm{H}_\varphi^\delta(B)$$

for all $\delta < \rho(A, B)$. Letting $\delta \searrow 0$, we get the same formula for H_φ, which is just (1.13) with $\mu = \mathrm{H}_\varphi$. The metric outer measure H_φ is called the Hausdorff outer measure associated with the gauge function φ. Its restriction to the σ-algebra of H_φ-measurable sets, which by Theorem 1.2.5 includes all the Borel sets, is called the Hausdorff measure associated with the function φ. We should add that even if $E \subseteq X$ is not a Borel set, nor even H_φ-measurable, we nevertheless commonly refer to $\mathrm{H}_\varphi(E)$ as the Hausdorff measure of E rather than the Hausdorff outer measure of E.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

As an immediate consequence of the definition of the Hausdorff measure and the properties of the function φ, we get the following.

Proposition 1.5.1 *For any gauge function φ, the Hausdorff measure H_φ is atomless.*

A particularly important role is played by the gauge functions of the form

$$\varphi_t(r) = r^t$$

for $t > 0$. In this case, the corresponding outer Hausdorff measure is denoted by H^t. So, H^t can be briefly defined as follows.

Definition 1.5.2 Given that $t \geq 0$, the t-dimensional outer Hausdorff measure $\mathrm{H}^t(A)$ of the set A is equal to

$$\mathrm{H}^t(A) = \sup_{\delta>0} \inf \left\{ \sum_{i=1}^{\infty} \mathrm{diam}^t(A_i) \right\},$$

where the infimum is taken over all countable covers $\{A_i\}_{i=1}^{\infty}$ of A by the sets with diameters $\leq \delta$.

Remark 1.5.3 Since $\mathrm{diam}(\overline{A}) = \mathrm{diam}(A)$ for every set $A \subseteq X$, we may, in Definition 1.5.2, restrict ourselves to closed sets A_i only.

Having defined Hausdorff measures, we now move on to define the dual concept, i.e., that of packing measures. As mentioned at the beginning of this section, while Hausdorff measures were introduced quite early, in 1919 by Felix Hausdorff in [**H**], it took several decades more for packing measures to be defined. It was done in stages in [**Tr**], [**TT**], and [**Su6**]. We do it again here now. We recall that, in Definition 1.3.3, we introduced the concept of packing. We will also use it now. For every $A \subseteq X$ and every $\delta > 0$, let

$$\Pi_\varphi^{*\delta}(A) := \sup \left\{ \sum_{i=1}^{\infty} \varphi(\mathrm{diam}(r_i)) \right\}, \tag{1.38}$$

where the supremum is taken over all packings $\{B(x_i, r_i)\}_{i=1}^{\infty}$ of A with radii not exceeding δ. Let

$$\Pi_\varphi^*(A) := \inf_{\delta>0} \left\{ \Pi_\varphi^{*\delta}(A) \right\} = \lim_{\delta\to 0} \Pi_\varphi^{*\delta}(A). \tag{1.39}$$

The limit exists since $\Pi_\varphi^{*\delta}(A)$ decreases as δ decreases. Although the function Π_φ^* satisfies condition (1.9) of outer measures, albeit in contrast to the case of Hausdorff measures, this function does not need to be subadditive, i.e., conditions (1.10) in general fail. In order to obtain an outer measure, we take one more step and we put

$$\Pi_{\varphi}(A) := \inf\left\{\sum_{i=1}^{\infty} \Pi_{\varphi}^{*\delta}(A_i)\right\}, \tag{1.40}$$

where the infimum is taken over all countable covers $\{A_i\}_{i=1}^{\infty}$ of A. Analogously, as in the case of Hausdorff measures, one checks, with similar arguments, that Π_{φ} is already an outer measure. Furthermore, it is a metric outer measure on X. It will be called the outer packing measure, associated with the gauge function φ. Its restriction to the σ-algebra of Π_{φ}-measurable sets, which by Theorem 1.2.5 includes all Borel sets, will be called the packing measure associated with the gauge function φ.

In the case of gauge functions,

$$\varphi_t(r) = r^t,$$

where $t > 0$, the definition of the outer packing measure takes the following form.

Definition 1.5.4 The t-dimensional outer packing measure $\Pi^t(A)$ of a set $A \subseteq X$ is given by

$$\Pi^t(A) = \inf_{\cup A_i = A}\left\{\sum_i \Pi_*^t(A_i)\right\}$$

(A_i are arbitrary subsets of A), where

$$\Pi_*^t(A) = \sup_{\delta>0}\sup\left\{\sum_{i=1}^{\infty} r_i^t\right\}.$$

Here, the second supremum is taken over all packings $\{B(x_i, r_i)\}_{i=1}^{\infty}$ of the set A by open balls centered at A with radii which do not exceed δ.

From now on, in order to obtain more meaningful geometric consequences, we assume that, for a given gauge function $\phi\colon [0, +\infty) \to [0, +\infty)$, there exists a function $C_\phi\colon (0,\infty) \to (0,\infty)$ such that, for every $a \in (0,\infty)$ and every $t > 0$ sufficiently small (depending on a),

$$C_\phi(a)^{-1}\phi(t) \le \phi(at) \le C_\phi(a)\phi(t). \tag{1.41}$$

We frequently refer to such gauge functions as evenly varying. Since $(at)^r = a^r t^r$, all the gauge functions ϕ of the form $r \mapsto r^t$ satisfy (1.41) with $C_\phi(a) = a^t$.

We shall now establish a simple, but crucial for geometric consequences, relation between Hausdorff and packing measures.

Proposition 1.5.5 *For every set $A \subseteq X$, it holds that* $\mathrm{H}_\varphi(A) \le C_\varphi(2)\Pi_\varphi(A)$.

Proof First, we show that, for every set $A \subseteq X$ and every $\delta > 0$,

$$\mathrm{H}_\varphi^{2\delta}(A) \le C_\varphi(2)\Pi_\varphi^{*\delta}(A). \tag{1.42}$$

Indeed, if there is no finite maximal (in the sense of inclusion) packing of the set A of the form $\{B(x_i,\delta)\}_{i=1}^\infty$, then, for every $k \ge 1$, there is a packing $\{B(x_i,\delta)\}_{i=1}^k$ of A; therefore,

$$\Pi_\varphi^{*\delta}(A) \ge \sum_{i=1}^k \varphi(\delta) = k\varphi(\delta).$$

Since $\varphi(\delta) > 0$, this yields $\Pi_\varphi^{*\delta}(A) = \infty$, and (1.42) holds. Otherwise, let $\{B(x_i,\delta)\}_{i=1}^l$ be a finite maximal packing of A. Then the collection $\{B(x_i,2\delta)\}$ covers A; therefore,

$$\mathrm{H}_\varphi^{2\delta}(A) \le \sum_{i=1}^l \varphi(2\delta) \le C_\varphi(2)l\varphi(\delta) \le C_\varphi(2)\Pi_\varphi^{*\delta}(A).$$

Hence, (1.42) is satisfied. Thus, letting $\delta \searrow 0$, we get that

$$\mathrm{H}_\varphi(A) \le C_\varphi(2)\Pi_\varphi^*(A). \tag{1.43}$$

So, if $\{A_n\}_{n\ge1}$ is a countable cover of A, then

$$\mathrm{H}_\varphi(A) \le \sum_{n=1}^\infty \mathrm{H}_\varphi(A_i) \le C_\varphi(2)\sum_{n=1}^\infty \Pi_\varphi^*(A_i).$$

Hence, applying (1.40), the lemma follows. ■

Definition 1.5.6 HD(A), the Hausdorff dimension of the set A, is defined to be

$$\mathrm{HD}(A) := \inf\{t\colon \mathrm{H}^t(A) = 0\} = \sup\{t\colon \mathrm{H}^t(A) = \infty\}. \tag{1.44}$$

Likewise, PD(A), the packing dimension of the set A, is defined to be

$$\mathrm{PD}(A) := \inf\{t\colon \Pi^t(A) = 0\} = \sup\{t\colon \Pi^t(A) = \infty\}. \tag{1.45}$$

The following theorem is the immediate consequence of the definition of Hausdorff and packing dimensions and the corresponding outer measures.

Theorem 1.5.7 *The Hausdorff and packing dimensions are monotone increasing functions of sets, i.e., if $A \subseteq B$, then*

$$\mathrm{HD}(A) \le \mathrm{HD}(B) \quad \textit{and} \quad \mathrm{PD}(A) \le \mathrm{PD}(B).$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

We shall prove the following theorem, commonly referred to as the σ-stability Hausdorff and packing dimensions.

Theorem 1.5.8 *If* $\{A_n\}_{n=1}^{\infty}$ *is a countable family of subsets of* X*, then*

$$\mathrm{HD}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sup_{n\geq 1}\{\mathrm{HD}(A_n)\}$$

and

$$\mathrm{PD}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sup_{n\geq 1}\{\mathrm{PD}(A_n)\}.$$

Proof We shall prove only the Hausdorff dimension part. The proof for the packing dimension is analogous. Inequality

$$\mathrm{HD}\left(\bigcup_{n=1}^{\infty} A_n\right) \geq \sup_{n\geq 1}\{\mathrm{HD}(A_n)\}$$

is an immediate consequence of Theorem 1.5.7. Thus, if $\sup_n\{\mathrm{HD}(A_n)\} = \infty$, there is nothing to prove. So, suppose that

$$s := \sup_{n\geq 1}\{\mathrm{HD}(A_n)\}$$

is finite and consider an arbitrary $t > s$. Then, in view of (1.44), $\mathrm{H}^t(A_n) = 0$ for every $n \geq 1$; therefore, since H^t is an outer measure,

$$\mathrm{H}^t\left(\bigcup_{n=1}^{\infty} A_n\right) = 0.$$

Hence, by (1.44) again,

$$\mathrm{HD}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq t.$$

The proof is complete. ■

As an immediate consequence of this theorem, Proposition 1.5.1, and (1.44), we obtain the following.

Proposition 1.5.9 *The Hausdorff dimension of any countable set is equal to* 0.

These are the most basic, transparent, and also probably most useful properties of Hausdorff and packing measures and dimensions. We will apply them frequently in both volumes of the book.

1.6 Hausdorff and Packing Measures: Frostman Converse-Type Theorems

In this section, we derive several geometric consequences of Theorem 1.3.1. Their meaning is to tell us when a Hausdorff measure or packing measure is positive, finite, zero, or infinity. We refer to them as the Frostman Converse-Type Theorems. Somewhat strangely, these theorems are frequently called the Mass Redistribution Principle in the fractal literature, as if to indicate that such measures would always have to appear in the process of an iterative construction. At the end of the section, we formulate the Frostman Direct Theorem and compare it with the Frostman Converse-Type Theorems. As already mentioned, the advantage of the latter theorems is that they provide tools to calculate, or at least to estimate, both Hausdorff and packing measures and dimensions. We recall that in this section, as in the entire book, we keep

$$\varphi\colon [0,+\infty) \longrightarrow [0,+\infty),$$

an evenly varying gauge function, i.e., satisfying (1.41). We start with the following.

Theorem 1.6.1 (Frostman Converse-Type Theorem for Generalized Hausdorff Measures) *Let $\varphi\colon [0,+\infty) \longrightarrow [0,+\infty)$ be a continuous evenly varying gauge function. Let (X,ρ) be an arbitrary metric space and μ a Borel probability measure on X. Fix a Borel set $A \subseteq X$. Assume that there exists a constant $c \in (0,+\infty]$ $(1/+\infty = 0)$ such that*

(1)

$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{\varphi(r)} \ge c$$

for all points $x \in A$ except for countably many perhaps. Then the Hausdorff measure H_φ, corresponding to the gauge function φ, satisfies

$$\mathrm{H}_\varphi(E) \le c^{-1} C_\varphi(8)\mu(E)$$

for every Borel set $E \subseteq A$. In particular,

$$\mathrm{H}_\varphi(A) < +\infty \;\; (\mathrm{H}_\varphi(A) = 0 \;\; \text{if} \;\; c = +\infty).$$

(2) *If, conversely,*

$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{\varphi(r)} \le c < +\infty$$

for all $x \in A$*, then*

$$\mu(E) \leq \mathrm{H}_\varphi(E)$$

for every Borel set $E \subseteq A$*. In particular,*

$$\mathrm{H}_\varphi(A) > 0$$

whenever $\mu(E) > 0$*.*

Proof Part (1). Since H_φ of any countable set is equal to 0, we may assume without loss of generality that E does not intersect the exceptional countable set. Fix $\varepsilon > 0$. Then fix $\delta > 0$. Since measure μ is regular, there exists an open set $G \supseteq E$ such that

$$\mu(G) \leq \mu(E) + \varepsilon.$$

Further, for every $x \in E$, there exists $r(x) \in (0, \delta)$ such that $B(x, r(x)) \subseteq G$ and

$$(c^{-1} + \varepsilon)\mu(B(x, r(x))) \geq \varphi(r(x)) > 0.$$

By virtue both of $4r$ Covering Theorem, i.e., Theorem 1.3.1, and of Remark 1.3.2, there exists $\{x_k\}_{k=1}^\infty$, a sequence of points in E such that

$$B(x_i, r(x_i)) \cap B(x_j, r(x_j)) = \emptyset \qquad \text{for } i \neq j$$

and

$$\bigcup_{k=1}^\infty B(x_k, 4r(x_k)) \supseteq \bigcup_{x \in E} B(x, r(x)) \supseteq E.$$

Hence,

$$\begin{aligned}
\mathrm{H}_\varphi^{2\delta}(E) &\leq \sum_{k=1}^\infty \varphi(2 \cdot 4r(x_k)) \leq \sum_{k=1}^\infty c_\varphi(8)\varphi(r(x_k)) \\
&\leq c_\varphi(8) \sum_{k=1}^\infty (c^{-1} + \varepsilon)\mu(B(x_k, r(x_k))) \\
&= c_\varphi(8)(c^{-1} + \varepsilon)\mu\left(\bigcup_{k=1}^\infty B(x_k, r(x_k))\right) \leq c_\varphi(8)(c^{-1} + \varepsilon)\mu(G) \\
&\leq c_\varphi(8)(c^{-1} + \varepsilon)(\mu(E) + \varepsilon).
\end{aligned}$$

So, letting $\delta \searrow 0$, we get

$$\mathrm{H}_\varphi(E) \leq c_\varphi(8)(c^{-1} + \varepsilon)(\mu(E) + \varepsilon)$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

and, since $\varepsilon > 0$ was arbitrary, we finally get

$$\mathrm{H}_\varphi(E) \le c_\varphi(8)c^{-1}\mu(E).$$

This finishes the first part of the proof.
Part (2). Now we deal with the second part of our theorem. Fix an arbitrary $s > c$. Note that, for every $r > 0$, the function

$$X \ni x \longmapsto \frac{\mu(B(x,r))}{\varphi(r)} \quad \text{is Borel measurable.}$$

For every $k \ge 1$, consider the function

$$X \ni x \longmapsto \varphi_k(x) := \sup\left\{\frac{\mu(B(x,r)}{\varphi(r)} : \ r \in \mathbb{Q} \cap (0,1/k]\right\},$$

where $\mathbb{Q}$ denotes the set of rational numbers. This function is Borel measurable as the supremum of countably many measurable functions. Let

$$A_k = A \cap \varphi_k^{-1}((0,s]) \quad \text{for } k \ge 1.$$

All A_k, $k \ge 1$, are then Borel subsets of X. Fix an arbitrary $r \in (0,1/k)$. Then pick $r_j \searrow r$, $r_j \in Q$. Since the function $t \mapsto \mu(B(x,t))$ is nondecreasing and the function φ is continuous, we get, for every $x \in A_k$, that

$$\frac{\mu(B(x,r))}{\varphi(r)} \le \limsup_{j\to\infty} \frac{\mu(B(x,r_j))}{\varphi(r_j)} \le s.$$

Now fix $k \ge 1$. Then fix a Borel set $F \subseteq A_k$. Our first objective is to prove the assertion of Part (2) for the set F. To do this, fix $r < 1/k$ and then $\{F_i\}_1^\infty$, a countable cover of F by subsets of F that are closed relative to F and have diameters less than $r/2$. For every $i \ge 1$, pick $x_i \in F_i$. Then $F_i \subset B(x_i, \mathrm{diam}(F_i))$. Since all the sets F_i, $i \ge 1$, are also Borel in X, we, therefore, have that

$$\sum_{i=1}^\infty \varphi(\mathrm{diam}(F_i)) \ge s^{-1}\sum_{i=1}^\infty \mu(B(x_i,\mathrm{diam}(F_i))) \ge s^{-1}\sum_{i=1}^\infty \mu(F_i) \ge s^{-1}\mu(F).$$

Hence, invoking Remark 1.5.3, we get that

$$\mathrm{H}_\varphi(F) \ge s^{-1}\mu(F). \tag{1.46}$$

Moving on, by our hypothesis, we have that

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

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Define inductively

$$B_1 := A_1 \cap A$$

and

$$B_{k+1} := A_{k+1} \cap \left(A \setminus \bigcup_{j=1}^{k} A_j \cap A \right).$$

Obviously, the family $\{B_k\}_1^\infty$ consists of mutually disjoint sets, with each set B_k contained in A_k, $k \geq 1$, and

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

Hence, if E is a Borel subset of A, then applying (1.46) for sets $F = E \cap B_k$, $k \geq 1$, we get

$$\mathrm{H}_\varphi(E) = \bigcup_{k=1}^{\infty} \mathrm{H}_\varphi(E \cap B_k) \geq s^{-1} \sum_{k=1}^{\infty} \mu(E \cap B_k) = s^{-1}\mu(E).$$

Letting $s \searrow c$ then finishes the proof. ■

Now let us prove the corresponding theorem for packing measures.

Theorem 1.6.2 (Frostman Converse-Type Theorem for Generalized Packing Measures) *Let $\varphi\colon [0,\infty) \longrightarrow [0,\infty)$ be a continuous evenly varying gauge function. Let (X,ρ) be an arbitrary metric space and μ be a Borel probability measure on X. Fix a Borel set $A \subset X$ and assume that there exists $c \in (0,+\infty]$ $(1/+\infty = 0)$ such that*

(1)

$$\liminf_{r \to 0} \frac{\mu(B(x,r))}{\varphi(r)} \leq c \quad \textit{for all } x \in A.$$

Then

$$\mu(E) \leq \Pi_\varphi(E)$$

for every Borel set $E \subseteq A$, where, we recall, Π_φ denotes the packing measure corresponding to the gauge function φ. In particular, if $\mu(E) > 0$, then

$$\Pi_\varphi(E) > 0.$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

(2) *If, conversely,*

$$\liminf_{r\to 0} \frac{\mu(B(x,r))}{\varphi(r)} \geq c \quad \textit{for all } x \in A,$$

then

$$\Pi_\varphi(E) \leq c^{-1}\mu(E)$$

for every Borel set $E \subseteq A$. In particular, if $\mu(E) < +\infty$, then

$$\Pi_\varphi(E) < +\infty.$$

Proof Part (1). Let $\varepsilon > 0$. Fix an arbitrary subset $F \subseteq A$. Define a decreasing sequence $(G_n)_{n\geq 1}$ of open sets containing F as follows. By our hypothesis, for every $x \in A$, there exists $0 < r_1(x) < 1$ such that

$$\frac{\mu\big(B(x,r_1(x))\big)}{\phi(r_1(x))} \leq c + \varepsilon.$$

Take the family of balls $\big\{B\big(x,\frac{1}{4}r_1(x)\big)\big\}_{x\in F}$. According to the $4r$ Covering Theorem (Theorem 1.3.1), there is a countable set $F_1 \subseteq F$ such that the subfamily $\big\{B\big(x,\frac{1}{4}r_1(x)\big)\big\}_{x\in F_1}$ consists of mutually disjoint balls satisfying

$$F \subseteq \bigcup_{x\in F} B\Big(x,\frac{1}{4}r_1(x)\Big) \subseteq \bigcup_{x\in F_1} B\big(x,r_1(x)\big).$$

Let $G_1 := \bigcup_{x\in F_1} B(x,r_1(x))$. For the inductive step, suppose that G_n has been defined for some $n \geq 1$. By our hypothesis again, for every $x \in A$, there exists some $0 < r_{n+1}(x) < \frac{1}{n+1}$ such that $B(x,r_{n+1}(x)) \subseteq G_n$ and

$$\frac{\mu\big(B(x,r_{n+1}(x))\big)}{\phi(r_{n+1}(x))} \leq c + \varepsilon. \tag{1.47}$$

Consider the family of balls $\big\{B\big(x,\frac{1}{4}r_{n+1}(x)\big)\big\}_{x\in F}$. According to the $4r$ Covering Theorem (Theorem 1.3.1), there exists a countable set $F_{n+1} \subseteq F$ such that the subfamily $\big\{B\big(x,\frac{1}{4}r_{n+1}(x)\big)\big\}_{x\in F_{n+1}}$ consists of mutually disjoint balls satisfying

$$F \subseteq \bigcup_{x\in F} B\left(x,\frac{1}{4}r_{n+1}(x)\right) \subseteq \bigcup_{x\in F_{n+1}} B\big(x,r_{n+1}(x)\big).$$

Let

$$G_{n+1} := \bigcup_{x\in F_{n+1}} B(x,r_{n+1}(x)).$$

It is clear that G_{n+1} is an open set and $F \subseteq G_{n+1} \subseteq G_n$. Moreover, for all pairs $x, y \in F_{n+1} \subseteq F$, we know that

$$d(x, y) \geq \frac{1}{4} \max\{r_{n+1}(x), r_{n+1}(y)\} \geq \frac{1}{8}(r_{n+1}(x) + r_{n+1}(y)).$$

Therefore, the collection $\left\{\left(x, \frac{1}{8} r_{n+1}(x)\right)\right\}_{x \in F_{n+1}}$ forms an $\left(\frac{1}{n+1}\right)$-packing of F. Using (1.47), it follows that

$$\begin{aligned}
\Pi_{\phi}^{*\frac{1}{n+1}}(F) &\geq \sum_{x \in F_{n+1}} \phi\left(\frac{1}{8} r_{n+1}(x)\right) \geq (C_\phi(8))^{-1} \sum_{x \in F_{n+1}} \phi(r_{n+1}(x)) \\
&\geq (C_\phi(8))^{-1} \sum_{x \in F_{n+1}} \frac{\mu\big(B(x, r_{n+1}(x))\big)}{c + \varepsilon} \\
&= \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \mu\left(\bigcup_{x \in F_{n+1}} B(x, r_{n+1}(x))\right) \\
&\geq \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \mu(G_{n+1}).
\end{aligned}$$

Letting n increase to infinity, we, thus, obtain that

$$\Pi_\phi^*(F) \geq \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \inf_{n \geq 1} \mu(G_n) = \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \lim_{n \to \infty} \mu(G_n) = \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \mu(G_F),$$

where $G_F := \cap_{n \geq 1} G_n$ is a G_δ set and, therefore, in particular, is a Borel set. Consequently, for every Borel set $E \subseteq A$, we have that

$$\begin{aligned}
\Pi_\phi(E) &= \inf\left\{\sum_{k=1}^{\infty} \Pi_\phi^*(A_k) \colon \{A_k\}_{k=1}^{\infty} \text{ is a cover of } E\right\} \\
&= \inf\left\{\sum_{k=1}^{\infty} \Pi_\phi^*(A_k) \colon \{A_k\}_{k=1}^{\infty} \text{ is a partition of } E\right\} \\
&\geq \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \inf\left\{\sum_{k=1}^{\infty} \mu(G_{A_k}) \colon \{A_k\}_{k=1}^{\infty} \text{ is a partition of } E\right\} \\
&\geq \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \inf\left\{\mu\left(\bigcup_{k=1}^{\infty} G_{A_k}\right) \colon \{A_k\}_{k=1}^{\infty} \text{ is a partition of } E\right\} \\
&\geq \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \inf\big\{\mu(E) \colon \{A_k\}_{k=1}^{\infty} \text{ is a partition of } E\big\} \\
&= \frac{(C_\phi(8))^{-1}}{c + \varepsilon} \mu(E).
\end{aligned}$$

Since this holds for all $\varepsilon > 0$, we deduce that $\Pi_\phi(E) \geq (C_\phi(8)c)^{-1}\mu(E)$. The proof of Part (1) is complete.

Part (2). The sequence of functions $(\psi_k)_{k=1}^{\infty}$, where

$$X \ni x \longmapsto \psi_k(x) := \inf\left\{\frac{\mu(B(x,r))}{\phi(r)} : r \in \mathbb{Q} \cap \left(0, \frac{1}{k}\right]\right\},$$

forms an increasing sequence of measurable functions. Let $0 < s < c$. For each $k \geq 1$, let

$$A_k := \psi_k^{-1}([s, +\infty)).$$

As $(\psi_k)_{k=1}^{\infty}$ is increasing, so is the sequence $(A_k)_{k=1}^{\infty}$. Moreover, since $s < c$, it follows from our hypothesis that

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

Furthermore, since $[s, +\infty)$ is a Borel subset of $\mathbb{R}$, the measurability of ψ_k ensures that A_k is a Borel subset of X. Fix $k \geq 1$. Choose some arbitrary $r \in (0, 1/k]$ and pick a sequence $(r_j)_{j\geq 1} \in \mathbb{Q}$ such that r_j increases to r. Since μ is a measure and ϕ is continuous, we deduce that, for all $x \in A_k$,

$$\frac{\mu(B(x,r))}{\phi(r)} = \lim_{j\to\infty} \frac{\mu(B(x,r_j))}{\phi(r_j)} \geq \psi_k(x) \geq s.$$

Thus, if $x \in A_k$, then

$$\inf\left\{\frac{\mu(B(x,r))}{\phi(r)} : r \in (0, 1/k]\right\} \geq s.$$

Now fix any set $F \subseteq A_k$ and any $r \in \left(0, \frac{1}{k}\right]$. Let $\{(x_i, t_i)\}_{i\geq 1}$ be an r-packing of F. Then

$$\sum_{i=1}^{\infty} \phi(t_i) \leq s^{-1} \sum_{i=1}^{\infty} \mu(B(x_i, t_i)) = s^{-1}\mu\left(\bigcup_{i=1}^{\infty} B(x_i, t_i)\right) \leq s^{-1}\mu(F_r),$$

where F_r denotes the open r-neighborhood of F. Taking the supremum over all r-packings yields

$$\Pi_\phi(F) \leq \Pi_\phi^*(F) \leq \Pi_\phi^{*r}(F) \leq s^{-1}\mu(F_r).$$

Thus, we have that $P_\phi(F) \leq s^{-1}\mu(F_r)$ for all $r \in (0, 1/k]$ and each subset $F \subseteq A_k$. Consequently, $\Pi_\phi(F) \leq s^{-1}\mu(F_0) = s^{-1}\mu(\overline{F})$ for all $F \subseteq A_k$. In particular, if C is a closed subset of E, then

$$\Pi_\phi(C \cap A_k) \leq s^{-1}\mu(\overline{C \cap A_k}) \leq s^{-1}\mu(C) \leq s^{-1}\mu(E).$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

As this holds for all integers $k \geq 1$ and closed sets $C \subseteq E \subseteq A = \cup_{k\geq 1} A_k$, we deduce that $\Pi_\phi(C) \leq s^{-1}\mu(E)$. By the regularity of μ, on taking the supremum over all closed sets C contained in E, we conclude that

$$\Pi_\phi(E) \leq s^{-1}\mu(E).$$

Letting s increase to c finishes the proof. ■

Replacing $\varphi(r)$ by r^t and H_φ by H^t in Theorem 1.6.1 and Π_φ by Π^t in Theorem 1.6.2, we immediately get the following two results.

Theorem 1.6.3 (Frostman Converse-Type Theorem for Hausdorff Measures) *Fix $t > 0$ arbitrary. Let (X, ρ) be a metric space and μ a Borel probability measure on X. Fix a Borel set $A \subseteq X$. Assume that there exists a constant $c \in (0, +\infty]$ $(1/+\infty = 0)$ such that*

(1)
$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{r^t} \geq c$$

for all points $x \in A$ except for countably many perhaps. Then the Hausdorff measure H_t satisfies

$$H_t(E) \leq c^{-1}8^t\mu(E)$$

for every Borel set $E \subseteq A$. In particular,

$$H_t(A) < +\infty \quad (H_t(A) = 0 \text{ if } c = +\infty).$$

(2) *If, conversely,*

$$\limsup_{r\to 0} \frac{\mu(B(x,r))}{r^t} \leq c < +\infty$$

for all $x \in A$, then

$$\mu(E) \leq H_t(E)$$

for every Borel set $E \subseteq A$. In particular,

$$H_t(A) > 0$$

whenever $\mu(E) > 0$.

Theorem 1.6.4 (Frostman Converse-Type Theorem for Packing Measures) *Fix $t > 0$ arbitrary. Let (X, ρ) be a metric space and μ a Borel probability measure on X. Fix a Borel set $A \subset X$ and assume that there exists $c \in (0, +\infty]$ $(1/+\infty = 0)$ such that*

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

(1)

$$\liminf_{r\to 0}\frac{\mu(B(x,r))}{r^t}\le c \quad \textit{for all } x\in A.$$

Then

$$\mu(E)\le \Pi_t(E)$$

for every Borel set $E\subseteq A$. *In particular, if* $\mu(E)>0$, *then*

$$\Pi_t(E)>0.$$

(2) *If conversely,*

$$\liminf_{r\to 0}\frac{\mu(B(x,r))}{r^t}\ge c \quad \textit{for all } x\in A,$$

then

$$\Pi_t(E)\le c^{-1}\mu(E)$$

for every Borel set $E\subseteq A$. *In particular, if* $\mu(E)<+\infty$, *then*

$$\Pi_t(E)<+\infty.$$

In the opposite direction to Frostman Converse Theorems, there is the following well-known theorem.

Theorem 1.6.5 (Frostman Direct Lemma) *Let* X *be either a Borel subset of a Euclidean space* $\mathbb{R}^d$, $d\ge 1$, *or an arbitrary compact metric space. If* $t>0$ *and* $\mathrm{H}_t(X)>0$, *then there exists a Borel probability measure* μ *on* X *such that*

$$\mu(B(x,r))\le r^t$$

for every point $x\in X$ *and all radii* $r>0$.

This is a very interesting theorem, although Frostman Converse Theorems seem to be more suitable for estimating and calculating Hausdorff and packing measures and dimensions.

1.7 Hausdorff and Packing Dimensions of Measures

In this section, we define the concepts of dimensions, both Hausdorff and packing, of Borel measures. We then provide tools to calculate and estimate them. We also establish some relations between them. The dimensions of measures play an important role in both fractal geometry and dynamical systems.

We start this section with the following simple but crucial consequence of Theorem 1.6.3.

Theorem 1.7.1 (Volume Lemma for Hausdorff Measures) *Suppose that μ is a Borel probability measure on a metric space (X, ρ) and that A is a bounded Borel subset of $\mathbb{R}^n$. Then*

(a) *If $\mu(A) > 0$ and there exists θ_1 such that, for every $x \in A$,*

$$\liminf_{r \to 0} \frac{\log \mu(B(x,r))}{\log r} \geq \theta_1,$$

then $\mathrm{HD}(A) \geq \theta_1$.

(b) *If there exists θ_2 such that, for every $x \in A$,*

$$\liminf_{r \to 0} \frac{\log \mu(B(x,r))}{\log r} \leq \theta_2,$$

then $\mathrm{HD}(A) \leq \theta_2$.

Proof (a) Take any $0 < \theta < \theta_1$. Then, by assumption,

$$\limsup_{r \to 0} \mu(B(x,r))/r^{\theta} = 0.$$

It, therefore, follows from Theorem 1.6.3(2) that $\mathrm{H}^{\theta}(A) = +\infty$. Hence, $\mathrm{HD}(A) \geq \theta$. Consequently, $\mathrm{HD}(A) \geq \theta_1$.

(b) Now take an arbitrary $\theta > \theta_2$. Then, by assumption,

$$\limsup_{r \to 0} \mu(B(x,r))/r^{\theta} = +\infty.$$

Therefore, applying Theorem 1.6.3(1), we obtain that $\mathrm{H}^{\theta}(A) = 0$. Thus, $\mathrm{HD}(A) \leq \theta$ and, consequently, $\mathrm{HD}(A) \leq \theta_2$. The proof is finished. ■

Similarly, one proves the following consequence of Theorem 1.6.4.

Theorem 1.7.2 (Volume Lemma for Packing Measures) *Suppose that μ is a Borel probability measure on $\mathbb{R}^n$, $n \geq 1$, and A is a bounded Borel subset of $\mathbb{R}^n$.*

(a) *If $\mu(A) > 0$ and there exists θ_1 such that, for every $x \in A$,*

$$\limsup_{r \to 0} \frac{\log \mu(B(x,r))}{\log r} \geq \theta_1,$$

then $\mathrm{PD}(A) \geq \theta_1$.

(b) *If there exists θ_2 such that, for every $x \in A$,*

$$\limsup_{r \to 0} \frac{\log \mu(B(x,r))}{\log r} \leq \theta_2,$$

then $\mathrm{PD}(A) \leq \theta_2$.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

We will now apply Theorem 1.7.1(a) to get quite a general lower bound for the Hausdorff dimension, the one that is a generalization of a result due to McMullen [**McM1**], the proof of which is taken from [**U1**]. Although this result is usually applied in a dynamical context, it really does not require any dynamics to formulate and to prove it.

As always in this section, let (X,ρ) be a metric space and μ be a Borel probability upper Ahlfors measure on X, meaning that there exist constants $h > 0$ and $C \geq 1$ such that, for every $x \in X$ and $r > 0$,

$$\mu(B(x,r)) \leq Cr^h. \tag{1.48}$$

We then call h the exponent of μ. For any integer $k \geq 1$, let E_k be a finite collection of compact subsets of X, each element of which has positive measure μ. We denote:

$$K := \bigcup_{F \in E_1} F. \tag{1.49}$$

We assume the following.

$$\text{If } k \geq 1,\ F, G \in E_k, \text{ and } F \neq G, \text{ then } \mu(F \cap G) = 0. \tag{1.50}$$

$$\text{Every set } F \in E_{k+1} \text{ is contained in a unique element } G \in E_k. \tag{1.51}$$

For every integer $k \geq 1$ and every set $F \in E_k$, define

$$\text{density}\left(\bigcup_{D \in E_{k+1}} D, F\right) := \frac{\mu\left(D \cap \bigcup_{D \in E_{k+1}} D\right)}{\mu(F)} \tag{1.52}$$

and assume that

$$\Delta_k := \inf\left\{\text{density}\left(\bigcup_{D \in E_{k+1}} D, F\right) :\ F \in E_k\right\} > 0 \tag{1.53}$$

for every $k \geq 1$. Put also

$$d_k := \sup\left\{\text{diam}(F) :\ F \in E_k\right\}.$$

Suppose that $d_k < 1$ for every $k \geq 1$ and that

$$\lim_{k \to \infty} d_k = 0. \tag{1.54}$$

We then call the collection

$$\{E_k\}_{k=1}^{\infty}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

a McMullen sequence of sets. Let

$$E_\infty := \bigcap_{k=1}^{\infty} \bigcup_{F \in E_k} F.$$

We shall prove the following generalization of a McMullen result from [**McM1**], the proof of which is taken from [**U1**].

Proposition 1.7.3 *If* $\{E_k\}_{k=1}^{\infty}$ *is a McMullen sequence of subsets of a metric space* (X, ρ) *endowed with a Borel probability upper Ahlfors measure* μ *having exponent h, then*

$$\mathrm{HD}(E_\infty) \geq h - \limsup_{k \to \infty} \frac{\sum_{j=1}^{k-1} \log \Delta_j}{\log d_k}.$$

Proof We construct inductively a sequence of Borel probability measures $\{\nu_k\}_{k=1}^{\infty}$ on K as follows.

Put $\nu_1 := \mu$ and define ν_{k+1} by putting, for each Borel set $A \subseteq K$,

$$\nu_{k+1}(A) := \sum_{F \in E_k} \frac{\mu\left(A \cap F \cap \bigcup_{D \in E_{k+1}} D\right)}{\mu\left(F \cap \bigcup_{D \in E_{k+1}} D\right)} \nu_k(F). \tag{1.55}$$

This definition makes sense since, by (1.52) and (1.53), we see that $\mu\left(F \cap \bigcup_{D \in E_{k+1}} D\right) > 0$. By induction, we get for every $k \geq 1$, that

$$\nu\left(\bigcup_{D \in E_k} D\right) = 1, \tag{1.56}$$

and it follows from properties (1.49)–(1.51) that ν_{k+1} is a Borel probability measure indeed. In view of (1.55) and (1.50), we have that $\nu_{k+1}(F) = \nu_k(F)$ for each $F \in E_k$. Hence, using (1.50) and (1.51), we conclude by induction that $\nu_n(F) = \nu_k(F)$ for every $n \geq k$. Since $\lim_{k \to \infty} d_k = 0$, we, therefore, obtain a unique probability measure ν on K (being the weak limit of measures ν_k) such that

$$\nu(F) = \nu_k(F) \tag{1.57}$$

for every $F \in E_k$. Looking now at (1.56) and the definition of the set E, one gets

$$\nu(E_\infty) = 1. \tag{1.58}$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Making use of (1.55) and (1.57), one easily estimates, for every $F \in E_k$, that

$$\nu(F) \leq \frac{\mu(F)}{\Delta_{k-1}, \ldots, \Delta_1}. \tag{1.59}$$

In view of Theorem 1.7.1, the Volume Lemma for Hausdorff Measures, in order to prove that $\mathrm{HD}(E) \geq \delta$ for some $\delta \geq 0$ it is enough to show that

$$\liminf_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} \geq \delta \tag{1.60}$$

for ν-a.e. $x \in E$.

Now consider $x \in E_\infty$ and $0 < r < \sup_{k\geq 1}(d_k)$ arbitrary. Then there exists an integer $k = k(r) \geq 1$ such that $d_{k+1} \leq r \leq d_k$. Let $\tilde{B}(x,r)$ be the union of all sets in E_{k+1} which meet $B(x,r)$. Then $\tilde{B}(x,r) \subseteq B(x,2r)$ and, using (1.59) and (1.48), we get

$$\begin{aligned}\frac{\log \nu(B(x,r))}{\log r} &\geq \frac{\log \mu(\tilde{B}(x,r)) - \sum_{j=1}^{k-1} \log \Delta_j}{\log r} \\ &\geq \frac{\log C + h\log 2 + h\log r}{\log r} - \frac{\sum_{j=1}^{k-1} \log \Delta_j}{\log d_j}.\end{aligned}$$

Since $\lim_{r\to 0} k(r) = \infty$, we, therefore, obtain that

$$\liminf_{r\to 0} \frac{\log \nu(B(x,r))}{\log r} \geq h - \limsup_{k\to\infty} \frac{\sum_{j=1}^{k-1} \log \Delta_j}{\log d_k}.$$

In view of (1.58) and by applying Theorem 1.7.1(a), this finishes the proof. ■

Now we define the following main concepts of this section.

Definition 1.7.4 Let μ be a Borel measure on a metric space (X,ρ). We write

$$\mathrm{HD}_\star(\mu) := \inf\{\mathrm{HD}(Y) : \mu(Y) > 0\} \text{ and } \mathrm{HD}^\star(\mu) = \inf\{\mathrm{HD}(Y) : \mu(X \setminus Y) = 0\}.$$

Of course,

$$\mathrm{HD}_\star(\mu) \leq \mathrm{HD}^\star(\mu),$$

and, in the case when $\mathrm{HD}_\star(\mu) = \mathrm{HD}^\star(\mu)$, we call this common value the *Hausdorff dimension* of the measure μ and we denote it by $\mathrm{HD}(\mu)$.

An analogous definition can be formulated for packing dimensions, with respective notation $\mathrm{PD}_\star(\mu)$, $\mathrm{PD}^\star(\mu)$, and $\mathrm{PD}(\mu)$, and the name *packing dimension* of the measure μ.

The next definition introduces concepts that are effective tools to calculate the dimensions introduced above.

Definition 1.7.5 Let μ be a Borel probability measure on a metric space (X,ρ). For every point $x \in X$, we define the *lower and upper pointwise dimension*of the measure μ at the point $x \in X$, respectively, as

$$\underline{d}_\mu(x) := \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \quad \text{and} \quad \overline{d}_\mu(x) := \limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}.$$

In the case when both numbers $\underline{d}_\mu(x)$ and $\overline{d}_\mu(x)$ are equal, we denote their common value by $d_\mu(x)$. We then obviously have that

$$d_\mu(x) = \lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r},$$

and we call $d_\mu(x)$ the pointwise dimension of the measure μ at the point $x \in X$.

The following theorem about Hausdorff and packing dimensions of a Borel measure μ follows easily from Theorems 1.7.1 and 1.7.2.

Theorem 1.7.6 *If μ is a Borel probability measure on a metric space (X,ρ), then*

$$\mathrm{HD}_\star(\mu) = \operatorname{ess\,inf} \underline{d}_\mu, \quad \mathrm{HD}^\star(\mu) = \operatorname{ess\,sup} \underline{d}_\mu$$

and

$$\mathrm{PD}_\star(\mu) = \operatorname{ess\,inf} \overline{d}_\mu, \quad \mathrm{PD}^\star(\mu) = \operatorname{ess\,sup} \overline{d}_\mu.$$

Proof Recall that the μ-essential infimum ess inf of a measurable function ϕ and the μ-essential supremum ess sup of this function are, respectively, defined by

$$\operatorname{ess\,inf}(\phi) := \sup_{\mu(N)=0} \inf_{x\in X\setminus N} \phi(x) \quad \text{and} \quad \operatorname{ess\,sup}(\phi) := \inf_{\mu(N)=0} \sup_{x\in X\setminus N} \phi(x).$$

Put $\phi_* ::= \operatorname{ess\,inf} \phi$. We shall prove that

$$\mu(\phi^{-1}((0,\phi_*))) = 0 \quad \text{and} \quad \mu(\phi^{-1}((0,\theta))) > 0 \tag{1.61}$$

for all $\theta > \phi_*$. Indeed, if we had $\mu(\phi^{-1}((0,\phi_*))) > 0$, then there would exist $\theta < \phi_*$ with $\mu(\phi^{-1}((0,\theta])) > 0$. Hence, for every measurable set $N \subseteq X$ with $\mu(N) = 0$, we would have that $\inf_{X\setminus N} \phi \leq \theta$. Thus, ess inf $\phi \leq \theta$, which is a contradiction, and the first part of (1.61) is proved.

For the second part, proceeding also by way of contradiction, assume that there exists $\theta > \phi_*$ with

$$\mu(\phi^{-1}((0,\theta)) = 0.$$

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Then for $N := \phi^{-1}((0,\theta))$, we would have that $\inf_{X\setminus N}(\phi) \geq \theta$. Hence, $\operatorname{ess\,inf}\phi \geq \theta$, which is a contradiction, and this finishes the proof of (1.61).

This formula, applied to the function $\phi := \underline{d}_\mu$, tells us that, for every Borel set $A \subseteq X$, with $\mu(A) > 0$, there exists a Borel set $A' \subseteq A$ with $\mu(A') = \mu(A) > 0$ such that, for every $x \in A'$, we have that $\underline{d}_\mu(x) \geq d_{\mu*}$. Hence,

$$\mathrm{HD}(A) \geq \mathrm{HD}(A') \geq d_{\mu*}$$

by Theorem 1.7.1(a). Thus,

$$\mathrm{HD}_\star(\mu) \geq d_{\mu*}. \tag{1.62}$$

On the other hand, for every $\theta > \theta_1$, we have that $\mu\big(\{x \in X : \underline{d}_\mu(x) < \theta\}\big) > 0$. Hence, by Theorem 1.7.1(b),

$$\mathrm{HD}(\{x : \underline{d}_\mu(x) < \theta\}) \leq \theta.$$

Therefore, $\mathrm{HD}_\star(\mu) \leq \theta$. So, letting $\theta \searrow \theta_1$, we get

$$\mathrm{HD}_\star(\mu) \leq d_{\mu*}.$$

Along with (1.62), we, thus, conclude that $\mathrm{HD}_\star(\mu) = d_{\mu*}$.

One should proceed similarly to prove that $\mathrm{HD}^\star(\mu) = \operatorname{ess\,sup}\underline{d}_\mu(x)$ and to obtain corresponding results for packing dimensions. For the latter, one should refer to Theorem 1.7.2 instead of Theorem 1.7.1. ■

Definition 1.7.7 A Borel probability measure μ on a metric space (X,ρ) is called dimensional exact if and only if, for μ-a.e. $x \in X$, $d_\mu(x)$, the pointwise dimension of the measure μ at x exists and is μ-a.e. constant.

As an immediate consequence of Theorem 1.7.6, we get the following.

Proposition 1.7.8 *If μ is a Borel probability dimensional exact measure on a metric space (X,ρ), then both $\mathrm{HD}(\mu)$ and $\mathrm{PD}(mu)$ exist; moreover*

$$\mathrm{HD}(\mu) = \mathrm{PD}(\mu) = d_\mu,$$

where d_μ is the μ-a.e. constant value of the pointwise dimension of μ.

1.8 Box-Counting Dimensions

We shall now examine a slightly different type of dimension; namely, the box-counting dimension. This dimension, as we will shortly see, is not given by means of any outer measure. It behaves worse: it is not σ-stable and a set and its closure have the same box-counting dimension. Its definition is, however,

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

substantially simpler than those of Hausdorff and packing dimensions and is frequently easier to calculate or to estimate; it also frequently agrees with Hausdorff and packing dimensions and is widely used in the physics literature.

Definition 1.8.1 Let $0 < r < 1$ and $A \subseteq X$ be a bounded set. Define $N(A,r)$ to be the minimum number of balls of radius at most r with centers in A needed to cover A. Then the *upper and lower box-counting* (or, more simply, *box*) dimensions of A are, respectively, defined to be

$$\overline{\mathrm{BD}}(A) := \limsup_{r\to 0} \frac{\log N(A,r)}{-\log r}$$

and

$$\underline{\mathrm{BD}}(A) := \liminf_{r\to 0} \frac{\log N(A,r)}{-\log r}.$$

If these two quantities are equal, their common value is called the *box-counting dimension*, or simply the *box dimension*, of A, and we denote it by $\mathrm{BD}(A)$.

As said, the box-counting dimensions do not share all the congenial properties of the Hausdorff dimensions. In particular, they are not σ-stable. To see this, observe that

$$\mathrm{BD}(\mathbb{Q}\cap[0,1]) = 1 \neq 0 = \sup\{\mathrm{BD}(\{q\})\colon q \in \mathbb{Q}\cap[0,1]\}.$$

The box-counting dimension is, however, easily seen to be finitely stable; see Proposition 1.8.2.

Proposition 1.8.2 *If (X,ρ) is a metric space and F_1, $F_2,\dots,F_n$ is a finite collection of subsets of X, then*

$$\overline{\mathrm{BD}}\big(F_1 \cup F_2 \cup \cdots \cup F_n\big) = \max\big\{\overline{\mathrm{BD}}(F_1),\overline{\mathrm{BD}}(F_2),\dots,\overline{\mathrm{BD}}(F_n)\big\}$$

and the same formula holds for the lower box-counting dimension.

The terminology "box counting" comes from the fact that in Euclidean spaces we may use boxes from a lattice rather than balls to cover the set under scrutiny. Indeed, let $n \geq 1$, $X = \mathbb{R}^n$, and $\mathcal{L}(r)$ be any lattice in $\mathbb{R}^n$ consisting of cubes (boxes) with edges of length r. For any $A \subseteq X$, define

$$L(A,r) = \mathrm{card}\{C \in \mathcal{L}(r)\colon C \cap A \neq \emptyset\}.$$

Proposition 1.8.3 *If A is a bounded subset of $\mathbb{R}^n$, then*

$$\overline{\mathrm{BD}}(A) = \limsup_{r\to 0} \frac{\log L(A,r)}{-\log r}$$

and

$$\underline{\mathrm{BD}}(A) = \liminf_{r\to 0} \frac{\log L(A,r)}{-\log r}.$$

Proof Without loss of generality, let $0 < r < 1$. Select points $x_i \in A$ so that

$$A \subseteq \bigcup_{i=1}^{N(A,r)} B(x_i, r).$$

Fix $1 \le i \le N(A,r)$ momentarily. If $C \in \mathcal{L}(r)$ is such that $d(C, x_i) < r$, where d denotes the standard Euclidean metric on $\mathbb{R}^n$, one immediately verifies that $C \subseteq B(x_i, r + r\sqrt{n}) = B(x_i, (1+\sqrt{n})r)$. Thus, for any given $1 \le i \le N(A,r)$, we have that

$$\begin{aligned}\#\{C \in \mathcal{L}(r) \colon d(C, x_i) < r\} &= \frac{\lambda(B(x_i, (1+\sqrt{n})r))}{\lambda(\text{cube of side } r)} \le \frac{c_n\left[(1+\sqrt{n})r\right]^n}{r^n} \\ &= c_n(1+\sqrt{n})^n,\end{aligned}$$

where λ denotes the Lebesgue measure on $\mathbb{R}^n$ and c_n denotes the volume of the unit ball in $\mathbb{R}^n$. Since every $C \in L(A,r)$ admits at least one number $1 \le i \le N(A,r)$ such that $d(C, x_i) < r$, we deduce that $L(A,r) \le N(A,r)c_n(1+\sqrt{n})^n$. Therefore,

$$\log L(A,r) \le \log\bigl(c_n(1+\sqrt{n})^n\bigr) + \log N(A,r).$$

Hence,

$$\frac{\log L(A,r)}{-\log r} \le \frac{\log\bigl(c_n(1+\sqrt{n})^n\bigr)}{-\log r} + \frac{\log N(A,r)}{-\log r}.$$

So,

$$\limsup_{r\to 0} \frac{\log L(A,r)}{-\log r} \le \overline{\mathrm{BD}}(A) \quad\text{and}\quad \liminf_{r\to 0} \frac{\log L(A,r)}{-\log r} \le \underline{\mathrm{BD}}(A).$$

For the opposite inequality, again let $0 < r < 1$ and, for each $C \in L(A,r)$, choose $x_C \in C \cap A$. Then $C \cap A \subseteq B(x_C, r\sqrt{n})$. Thus, the family of balls

$$\bigl\{B(x_C, r\sqrt{n}) \colon C \in L(A,r)\bigr\}$$

covers A. Therefore, $N(A, r\sqrt{n}) \le L(A,r)$. It then follows that

$$\overline{\mathrm{BD}}(A) \le \limsup_{r\to 0} \frac{\log L(A,r)}{-\log r} \quad\text{and}\quad \underline{\mathrm{BD}}(A)(A) \le \liminf_{r\to 0} \frac{\log L(A,r)}{-\log r}.$$

■

Now we return to our general setting. Let (X, ρ) be again a metric space, $A \subseteq X$, and $r > 0$. Further, define $P(A, r)$ to be the supremum of the cardinalities of all packings of A of the form $\{B(x_i, r)\}_{i=1}^{\infty}$, so

$$P(A, r) := \sup \left\{ \#\{B(x_i, r)\}_{i=1}^{\infty} \right\}.$$

Such packings will be called r-packings of A in what follows. We shall prove the following technical, though interesting in itself, fact.

Lemma 1.8.4 *If $A \subseteq X$ and $r > 0$, then $N(A, 2r) \leq P(A, r) \leq N(A, r)$.*

Proof The first inequality certainly holds if $P(A, r) = \infty$. So, assume that this is not the case and let $\{(x_i, r)\}_{i=1}^{k}$ be an r-packing of A that is maximal in the sense of inclusion. Then $\{B(x_i, 2r)\}_{i=1}^{k}$ is a cover of A and, consequently, $N(A, 2r) \leq P(A, r)$. For the second inequality, there is nothing to prove if $N(A, r) = \infty$. So, again, let $\{(x_i, r)\}_{i=1}^{k}$ be a finite r-packing of A and assume that

$$\{B(y_j, r)\}_{j=1}^{\ell}$$

is a finite cover of A with centers in A. Then, for each $1 \leq i \leq k$, there exists $1 \leq j(i) \leq \ell$ such that

$$x_i \in B\big(y_{j(i)}, r\big).$$

We will show that $k \leq \ell$. In order to do this, it is enough to show that the function $i \mapsto j(i)$ is injective. But, for each $1 \leq j \leq \ell$, the cardinality of the set

$$\left\{ \{x_i\}_{i=1}^{k} \cap B(y_j, r) \right\}$$

is at most 1 (otherwise, $\{(x_i, r)\}_{i=1}^{k}$ would not be an r-packing), and so the function $i \mapsto j(i)$ is injective, as required. Thus, $P(A, r) \leq N(A, r)$. ■

These inequalities have the following immediate implications.

Corollary 1.8.5 *If X is a metric space and $A \subseteq X$, then*

$$\overline{\mathrm{BD}}(A) = \limsup_{r \to 0} \frac{\log P(A, r)}{-\log r}$$

and

$$\underline{\mathrm{BD}}(A) = \liminf_{r \to 0} \frac{\log P(A, r)}{-\log r}.$$

As an immediate consequence of this corollary and of the second part of Definition 1.8.1, we obtain the following.

https://doi.org/10.1017/9781009215930.005 Published online by Cambridge University Press

Corollary 1.8.6 *If X is a metric space and $A \subseteq X$, then*

$$\mathrm{HD}(A) \le \underline{\mathrm{BD}}(A) \le \mathrm{PD}(A) \le \overline{\mathrm{BD}}(A).$$

We end this section with the following.

Proposition 1.8.7 *Let (X,ρ) be a metric space endowed with a finite Borel measure μ such that*

$$\mu(B(x,r)) \ge Cr^t$$

for some constant $C > 0$, all $x \in X$, and all radii $0 \le r \le 1$. Then

$$\overline{\mathrm{BD}}(X) \le t.$$

If, on the other hand, $\mu(X) > 0$ and

$$\mu(B(x,r)) \le Cr^t$$

for some constant $C < +\infty$, all $x \in X$, and all radii $0 \le r \le 1$, then

$$\underline{\mathrm{BD}}(X) \ge \mathrm{HD}(X) \ge t.$$

Finally, if $\mu(X) > 0$ and

$$C^{-1}r^t \le \mu(B(x,r)) \le Cr^t$$

for some constant $C \in [1, +\infty)$, all $x \in X$, and all radii $0 \le r \le 1$, then

$$\mathrm{BD}(X) = \mathrm{PD}(X) = \mathrm{HD}(X) = t.$$

Proof We start with the first inequality. Let $\{B(x,r)\}_{i=1}^k$ be an r-packing of X. Then

$$kr^t \le C^{-1}\sum_{i=1}^{k} \mu(B(x_i,r)) \le C^{-1}.$$

Hence, $k \le C^{-1}r^{-t}$. Therefore, $P(X,r) \le C^{-1}r^{-t}$. Consequently,

$$\log P(X,r) \le -\log C - t\log r.$$

In conjunction with the first formula of Corollary 1.8.5, this yields

$$\overline{\mathrm{BD}}(X) \le t.$$

The second assertion of our proposition directly follows from the first inequality of Corollary 1.8.6 and from item (2) of the Frostman Converse-Type Theorem for Hausdorff Measures (Theorem 1.6.3).

The last assertion of our proposition is now an immediate consequence of the two first assertions and Corollary 1.8.6. ∎