Cambridge University Press & Assessment
978-1-107-13411-9 — Fractals in Probability and Analysis
Christopher J. Bishop, Yuval Peres
Excerpt
More Information

1

Minkowski and Hausdorff dimensions

In this chapter we will define the Minkowski and Hausdorff dimensions of a set and will compute each in a few basic examples. We will then prove Billingsley's Lemma and the Law of Large Numbers. These allow us to deal with more sophisticated examples: sets defined in terms of digit frequencies, random slices of the Sierpiński gasket, and intersections of random translates of the middle thirds Cantor set with itself. Both Minkowski and Hausdorff dimensions measure how efficiently a set K can be covered by balls. Minkowski dimension requires that the covering be by balls all of the same radius. This makes it easy to compute, but it lacks certain desirable properties. In the definition of Hausdorff dimension we will allow coverings by balls of different radii. This gives a better behaved notion of dimension, but (as we shall see) it is usually more difficult to compute.

1.1 Minkowski dimension

A subset K of a metric space is called **totally bounded** if for any $\varepsilon > 0$, it can be covered by a finite number of balls of diameter ε. For Euclidean space, this is the same as being a bounded set. For a totally bounded set K, let $N(K,\varepsilon)$ denote the minimal number of sets of diameter at most ε needed to cover K. We define the **upper Minkowski dimension** as

$$\overline{\dim}_{\mathscr{M}}(K) = \limsup_{\varepsilon\to 0} \frac{\log N(K,\varepsilon)}{\log 1/\varepsilon},$$

and the **lower Minkowski dimension**

$$\underline{\dim}_{\mathscr{M}}(K) = \liminf_{\varepsilon\to 0} \frac{\log N(K,\varepsilon)}{\log 1/\varepsilon}.$$

If the two values agree, the common value is simply called the **Minkowski dimension** of K and denoted by $\dim_{\mathscr{M}}(K)$. When the Minkowski dimension of a set K exists, the number of sets of diameter ε needed to cover K grows like $\varepsilon^{-\dim_{\mathscr{M}}(K)+o(1)}$ as $\varepsilon \to 0$.

We get the same values of $\overline{\dim}_{\mathscr{M}}(K)$ and $\underline{\dim}_{\mathscr{M}}(K)$ if we replace $N(K,\varepsilon)$ by $N_B(K,\varepsilon)$, which is the number of closed balls of radius ε needed to cover K. This is because $N_B(K,\varepsilon) \le N(K,\varepsilon) \le N(K,\varepsilon/2)$ (any set is contained in a ball of at most twice the diameter and any ball of radius $\varepsilon/2$ has diameter at most ε; strict inequality could hold in a metric space). For subsets of Euclidean space we can also count the number of axis-parallel squares of side length ε needed to cover K, or the number of such squares taken from a grid. Both possibilities give the same values for upper and lower Minkowski dimension, and for this reason Minkowski dimension is sometimes called the **box counting dimension**. It is also easy to see that a bounded set A and its closure $\overline{A}$ satisfy $\overline{\dim}_{\mathscr{M}}(A) = \overline{\dim}_{\mathscr{M}}(\overline{A})$ and $\underline{\dim}_{\mathscr{M}}(A) = \underline{\dim}_{\mathscr{M}}(\overline{A})$.

If X is a set and $x,y \in X$ implies $|x-y| \ge \varepsilon$, we say X is **ε-separated**. Let $N_{\text{sep}}(K,\varepsilon)$ be the number of elements in a maximal ε-separated subset X of K. Clearly, any set of diameter $\varepsilon/2$ can contain at most one point of an ε-separated set X, so $N_{\text{sep}}(K,\varepsilon) \le N(K,\varepsilon/2)$. On the other hand, every point of K is within ε of a maximal ε-separated subset X (otherwise add that point to X). Thus $N(K,\varepsilon) \le N_{\text{sep}}(K,\varepsilon)$. Therefore replacing $N(K,\varepsilon)$ by $N_{\text{sep}}(K,\varepsilon)$ in the definition of upper and lower Minkowski dimension gives the same values (and it is often easier to give a lower bound in terms of separated sets).

Example 1.1.1 Suppose that K is a finite set. Then $N(K,\varepsilon)$ is bounded and $\dim_{\mathscr{M}}(K)$ exists and equals 0.

Example 1.1.2 Suppose $K = [0,1]$. Then at least $1/\varepsilon$ intervals of length ε are needed to cover K and clearly $\varepsilon^{-1}+1$ suffice. Thus $\dim_{\mathscr{M}}(K)$ exists and equals 1. Similarly, any bounded set in $\mathbb{R}^d$ with interior has Minkowski dimension d.

Example 1.1.3 Let $\mathbf{C}$ be the usual middle thirds Cantor set obtained as follows. Let $\mathbf{C}^0 = [0,1]$ and define $\mathbf{C}^1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1] \subset \mathbf{C}^0$ by removing the central interval of length $\frac{1}{3}$. In general, $\mathbf{C}^n$ is a union of 2^n intervals of length 3^{-n} and $\mathbf{C}^{n+1}$ is obtained by removing the central third of each. This gives a decreasing nested sequence of compact sets whose intersection is the desired set $\mathbf{C}$.

The construction gives a covering of $\mathbf{C}$ that uses 2^n intervals of length 3^{-n}. Thus for $3^{-n} \le \varepsilon < 3^{-n+1}$ we have

$$N(\mathbf{C},\varepsilon) \le 2^n,$$

Figure 1.1.1 The Cantor middle thirds construction.

and hence

$$\overline{\dim}_{\mathscr{M}}(\mathbf{C}) \le \frac{\log 2}{\log 3}.$$

Conversely, the centers of the nth generation intervals form a 3^{-n}-separated set of size 2^n, so $N_{\text{sep}}(\mathbf{C}, 3^{-n}) \ge 2^n$. Thus

$$\underline{\dim}_{\mathscr{M}}(\mathbf{C}) \ge \frac{\log 2}{\log 3} = \log_3 2.$$

Therefore the Minkowski dimension exists and equals this common value. If at each stage we remove the middle α $(0 < \alpha < 1)$ we get a Cantor set $\mathbf{C}_\alpha$ with Minkowski dimension $\log 2 / (\log 2 + \log \frac{1}{1-\alpha})$.

Example 1.1.4 Consider $K = \{0\} \cup \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. Observe that

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)} > \frac{1}{n^2}.$$

So, for $\varepsilon > 0$, if we choose n so that $(n+1)^{-2} < \varepsilon \le n^{-2}$, then $n \le \varepsilon^{-1/2}$ and n distinct intervals of length ε are needed to cover the points $1, \frac{1}{2}, \dots, \frac{1}{n}$. The interval $[0, \frac{1}{n+1}]$ can be covered by $n+1$ additional intervals of length ε. Thus

$$\varepsilon^{-1/2} \le N(K, \varepsilon) \le 2\varepsilon^{-1/2} + 1.$$

Hence $\dim_{\mathscr{M}}(K) = 1/2$.

This example illustrates a drawback of Minkowski dimension: finite sets have dimension zero, but countable sets can have positive dimension. In particular, it is not true that $\dim_{\mathscr{M}}(\bigcup_n E_n) = \sup_n \dim_{\mathscr{M}}(E_n)$, a useful property for a dimension to have. In the next section, we will introduce Hausdorff dimension, which does have this property (Exercise 1.6). In the next chapter, we will introduce packing dimension, which is a version of upper Minkowski dimension forced to have this property.

1.2 Hausdorff dimension and the Mass Distribution Principle

Given any set K in a metric space, we define the **α-dimensional Hausdorff content** as

$$\mathscr{H}^{\alpha}_{\infty}(K) = \inf\left\{ \sum_i |U_i|^{\alpha} : K \subset \bigcup_i U_i \right\},$$

where $\{U_i\}$ is a countable cover of K by any sets and $|E|$ denotes the diameter of a set E.

Definition 1.2.1 The **Hausdorff dimension** of K is defined to be

$$\dim(K) = \inf\{\alpha : \mathscr{H}^{\alpha}_{\infty}(K) = 0\}.$$

More generally we define

$$\mathscr{H}^{\alpha}_{\varepsilon}(K) = \inf\left\{ \sum_i |U_i|^{\alpha} : K \subset \bigcup_i U_i, |U_i| < \varepsilon \right\},$$

where each U_i is now required to have diameter less than ε. The **α-dimensional Hausdorff measure** of K is defined as

$$\mathscr{H}^{\alpha}(K) = \lim_{\varepsilon \to 0} \mathscr{H}^{\alpha}_{\varepsilon}(K).$$

This is an outer measure; an **outer measure** on a non-empty set X is a function μ^* from the family of subsets of X to $[0,\infty]$ that satisfies

- $\mu^*(\emptyset) = 0$,
- $\mu^*(A) \le \mu^*(B)$ if $A \subset B$,
- $\mu^*(\bigcup_{j=1}^{\infty} A_j) \le \sum_{j=1}^{\infty} \mu^*(A_j)$.

For background on real analysis see Folland (1999). The α-dimensional Hausdorff measure is even a Borel measure in $\mathbb{R}^d$; see Theorem 1.2.4 below. When $\alpha = d \in \mathbb{N}$, then $\mathscr{H}^{\alpha}$ is a constant multiple of $\mathscr{L}_d$, d-dimensional Lebesgue measure.

If we admit only open sets in the covers of K, then the value of $\mathscr{H}^{\alpha}_{\varepsilon}(K)$ does not change. This is also true if we only use closed sets or only use convex sets. Using only balls might increase $\mathscr{H}^{\alpha}_{\varepsilon}$ by at most a factor of 2^{α}, since any set K is contained in a ball of at most twice the diameter. Still, the values for which $\mathscr{H}^{\alpha}(K) = 0$ are the same whether we allow covers by arbitrary sets or only covers by balls.

Definition 1.2.2 Let μ^* be an outer measure on X. A set K in X is **μ^*-measurable**, if for Every set $A \subset X$ we have

$$\mathscr{H}^{\alpha}(A) = \mu^*(A \cap K) + \mu^*(A \cap K^c).$$

Definition 1.2.3 Let (Ω, d) be a metric space. An outer measure μ on Ω is called a **metric outer measure** if $\operatorname{dist}(A,B) > 0 \implies \mu(A \cup B) = \mu(A) + \mu(B)$, where A and B are two subsets of Ω.

Since Hausdorff measure $\mathscr{H}^\alpha$ is clearly a metric outer measure, the following theorem shows that all Borel sets are $\mathscr{H}^\alpha$-measurable. This implies that $\mathscr{H}^\alpha$ is a Borel measure (see Folland (1999)).

Theorem 1.2.4 *Let μ be a metric outer measure. Then all Borel sets are μ-measurable.*

Proof It suffices to show any closed set K is μ-measurable, since the measurable sets form a σ-algebra. So, let K be a closed set. We must show for any set $A \subset \Omega$ with $\mu(A) < \infty$,

$$\mu(A) \geq \mu(A \cap K) + \mu(A \cap K^c). \tag{1.2.1}$$

Let $B_0 = \emptyset$ and for $n \geq 1$ define $B_n = \{x \in A : \operatorname{dist}(x,K) > \frac{1}{n}\}$, so that

$$\bigcup_{n=1}^{\infty} B_n = A \cap K^c$$

(since K is closed). Since μ is a metric outer measure and $B_n \subset A \setminus K$,

$$\mu(A) \geq \mu[(A \cap K) \cup B_n] = \mu(A \cap K) + \mu(B_n). \tag{1.2.2}$$

For all $m \in \mathbb{N}$, the sets $D_n = B_n \setminus B_{n-1}$ satisfy

$$\sum_{j=1}^{m} \mu(D_{2j}) = \mu\left(\bigcup_{j=1}^{m} D_{2j}\right) \leq \mu(A), \forall m,$$

since if $x \in B_n$, and $y \in D_{n+2}$, then

$$\operatorname{dist}(x,y) \geq \operatorname{dist}(x,K) - \operatorname{dist}(y,K) \geq \frac{1}{n} - \frac{1}{n+1}.$$

Similarly $\sum_{j=1}^{m} \mu(D_{2j-1}) \leq \mu(A)$. So $\sum_{j=1}^{\infty} \mu(D_j) < \infty$. The inequality

$$\mu(B_n) \leq \mu(A \cap K^c) \leq \mu(B_n) + \sum_{j=n+1}^{\infty} \mu(D_j)$$

implies that $\mu(B_n) \to \mu(A \cap K^c)$ as $n \to \infty$. Thus letting $n \to \infty$ in (1.2.2) gives (1.2.1). $\square$

The construction of Hausdorff measure can be made a little more general by considering a positive, increasing function φ on $[0,\infty)$ with $\varphi(0)=0$. This is called a **gauge function** and we may associate to it the Hausdorff content

$$\mathscr{H}_\infty^\varphi(K) = \inf\Big\{\sum_i \varphi(|U_i|) : K \subset \bigcup_i U_i\Big\};$$

then $\mathscr{H}_\varepsilon^\varphi(K)$, and $\mathscr{H}^\varphi(K) = \lim_{\varepsilon\to 0}\mathscr{H}_\varepsilon^\varphi(K)$ are defined as before. The case $\varphi(t)=t^\alpha$ is just the case considered above. We will not use other gauge functions in the first few chapters, but they are important in many applications, e.g., see Exercise 1.59 and the Notes for Chapter 6.

Lemma 1.2.5 *If $\mathscr{H}^\alpha(K)<\infty$ then $\mathscr{H}^\beta(K)=0$ for any $\beta>\alpha$.*

Proof It follows from the definition of $\mathscr{H}_\varepsilon^\alpha$ that

$$\mathscr{H}_\varepsilon^\beta(K) \le \varepsilon^{\beta-\alpha}\mathscr{H}_\varepsilon^\alpha(K),$$

which gives the desired result as $\varepsilon\to 0$. □

Thus if we think of $\mathscr{H}^\alpha(K)$ as a function of α, the graph of $\mathscr{H}^\alpha(K)$ versus α shows that there is a critical value of α where $\mathscr{H}^\alpha(K)$ jumps from ∞ to 0. This critical value is equal to the Hausdorff dimension of the set. More generally we have:

Proposition 1.2.6 *For every metric space E we have*

$$\mathscr{H}^\alpha(E)=0 \quad\Leftrightarrow\quad \mathscr{H}_\infty^\alpha(E)=0$$

and therefore

$$\begin{aligned}\dim E &= \inf\{\alpha\colon \mathscr{H}^\alpha(E)=0\} = \inf\{\alpha\colon \mathscr{H}^\alpha(E)<\infty\}\\ &= \sup\{\alpha\colon \mathscr{H}^\alpha(E)>0\} = \sup\{\alpha\colon \mathscr{H}^\alpha(E)=\infty\}.\end{aligned}$$

Proof Since $\mathscr{H}^\alpha(E)\ge\mathscr{H}_\infty^\alpha(E)$, it suffices to prove "$\Leftarrow$". If $\mathscr{H}_\infty^\alpha(E)=0$, then for every $\delta>0$ there is a covering of E by sets $\{E_k\}$ with $\sum_{k=1}^\infty |E_k|^\alpha<\delta$. These sets have diameter less than $\delta^{1/\alpha}$, hence $\mathscr{H}_{\delta^{1/\alpha}}^\alpha(E)<\delta$. Letting $\delta\downarrow 0$ yields $\mathscr{H}^\alpha(E)=0$, proving the claimed equivalence. The equivalence readily implies that $\dim E=\inf\{\alpha\colon \mathscr{H}^\alpha(E)=0\}=\sup\{\alpha\colon \mathscr{H}^\alpha(E)>0\}$. The other conclusions follow from Lemma 1.2.5. □

The following relationship to Minkowski dimension is clear

$$\dim(K) \le \underline{\dim}_{\mathscr{M}}(K) \le \overline{\dim}_{\mathscr{M}}(K). \tag{1.2.3}$$

Indeed, if $B_i = B(x_i, \varepsilon/2)$ are $N(K,\varepsilon)$ balls of radius $\varepsilon/2$ and centers x_i in K that cover K, then consider the sum

$$S_\varepsilon = \sum_{i=1}^{N(K,\varepsilon)} |B_i|^\alpha = N(K,\varepsilon)\varepsilon^\alpha = \varepsilon^{\alpha - R_\varepsilon},$$

where $R_\varepsilon = \frac{\log N(K,\varepsilon)}{\log(1/\varepsilon)}$. If $\alpha > \liminf_{\varepsilon\to 0} R_\varepsilon = \underline{\dim}_{\mathscr{M}}(K)$ then $\inf_{\varepsilon>0} S_\varepsilon = 0$. Strict inequalities in (1.2.3) are possible.

Example 1.2.7 Example 1.1.4 showed that $K = \{0\} \bigcup_n \{\frac{1}{n}\}$ has Minkowski dimension $\frac{1}{2}$. However, any countable set has Hausdorff dimension 0, for if we enumerate the points $\{x_1, x_2, \dots\}$ and cover the nth point by a ball of diameter $\delta_n = \varepsilon 2^{-n}$ we can make $\sum_n \delta_n^\alpha$ as small as we wish for any $\alpha > 0$. Thus K is a compact set for which the Minkowski dimension exists, but is different from the Hausdorff dimension.

Lemma 1.2.8 (Mass Distribution Principle) *If E supports a strictly positive Borel measure μ that satisfies*

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

for some constant $0 < C < \infty$ and for every ball $B(x,r)$, then $\mathscr{H}^\alpha(E) \ge \mathscr{H}^\alpha_\infty(E) \ge \mu(E)/C$. In particular, $\dim(E) \ge \alpha$.

Proof Let $\{U_i\}$ be a cover of E. For $\{r_i\}$, where $r_i > |U_i|$, we look at the following cover: choose x_i in each U_i, and take open balls $B(x_i, r_i)$. By assumption,

$$\mu(U_i) \le \mu(B(x_i, r_i)) \le C r_i^\alpha.$$

We deduce that $\mu(U_i) \le C|U_i|^\alpha$, whence

$$\sum_i |U_i|^\alpha \ge \sum_i \frac{\mu(U_i)}{C} \ge \frac{\mu(E)}{C}.$$

Thus $\mathscr{H}^\alpha(E) \ge \mathscr{H}^\alpha_\infty(E) \ge \mu(E)/C$. □

We note that upper bounds for Hausdorff dimension usually come from finding explicit coverings of the set, but lower bounds are proven by constructing an appropriate measure supported on the set. Later in this chapter we will generalize the Mass Distribution Principle by proving Billingsley's Lemma (Theorem 1.4.1) and will generalize it even further in later chapters. As a special case of the Mass Distribution Principle, note that if $A \subseteq \mathbb{R}^d$ has positive Lebesgue d-measure then $\dim(A) = d$.

Example 1.2.9 Consider the Cantor set E obtained by replacing the unit square in the plane by four congruent sub-squares of side length $\alpha < 1/2$ and continuing similarly. See Figure 1.2.1. We can cover the set by 4^n squares of diameter $\sqrt{2} \cdot \alpha^n$. Thus

$$\overline{\dim}_{\mathscr{M}}(E) \leq \lim_{n\to\infty} \frac{\log 4^n}{-\log(\sqrt{2}\alpha^n)} = \frac{\log 4}{-\log \alpha}.$$

On the other hand, it is also easy to check that at least 4^n sets of diameter α^n are needed, so

$$\underline{\dim}_{\mathscr{M}}(E) \geq \frac{\log 4}{-\log \alpha}.$$

Thus the Minkowski dimension of this set equals $\beta = -\log 4/\log \alpha$.

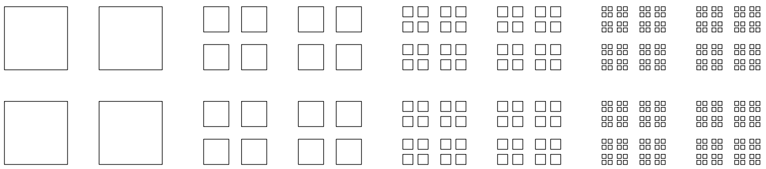

Figure 1.2.1 Four generations of a Cantor set.

We automatically get $\dim(E) \leq \beta$ and we will prove the equality using Lemma 1.2.8. Let μ be the probability measure defined on E that gives each nth generation square the same mass (namely 4^{-n}). We claim that

$$\mu(B(x,r)) \leq Cr^\beta,$$

for all disks and some $0 < C < \infty$. To prove this, suppose $B = B(x,r)$ is some disk hitting E and choose n so that $\alpha^{n+1} \leq r < \alpha^n$. Then B can hit at most 4 of the nth generation squares and so, since $\alpha^\beta = 1/4$,

$$\mu(B \cap E) \leq 4 \cdot 4^{-n} = 4\alpha^{n\beta} \leq 16 r^\beta.$$

Example 1.2.10 Another simple set for which the two dimensions agree and are easy to compute is the von Koch snowflake. To construct this we start with an equilateral triangle. At each stage we add to each edge an equilateral triangle pointing outward of side length 1/3 the size of the current edges and centered on the edge. See Figure 1.2.2 for the first four iterations of this process. The boundary of this region is a curve with dimension $\log 4/\log 3$ (see Theorem 2.2.2). We can also think of this as a replacement construction, in which at each stage, a line segment is replaced by an appropriately scaled copy of a polygonal curve.

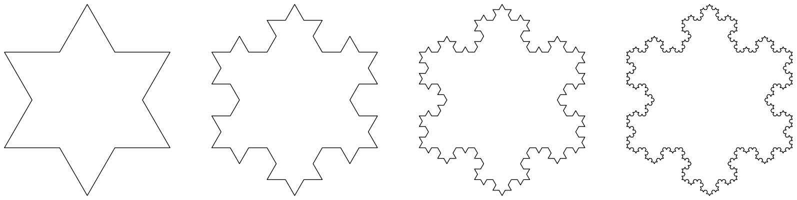

Figure 1.2.2 Four generations of the von Koch snowflake.

Even for some relatively simple sets the Hausdorff dimension is still unknown. Consider the Weierstrass function (Figure 1.2.3)

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

where $b > 1$ is real and $0 < \alpha < 1$. It is conjectured that the Hausdorff dimension of its graph is $2 - \alpha$, and this has been proven when b is an integer; see the discussion in Example 5.1.7. On the other hand, some sets that are more difficult to define, such as the graph of Brownian motion (Figure 1.2.4), will turn out to have easier dimensions to compute ($3/2$ by Theorem 6.4.3).

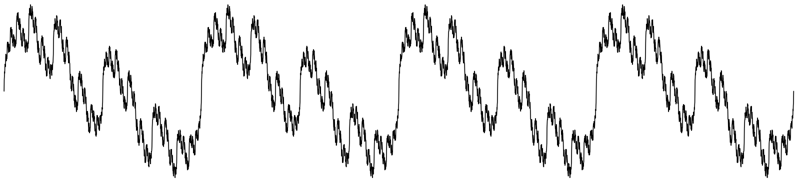

Figure 1.2.3 The Weierstrass function with $b = 2$, $\alpha = 1/2$. This graph has Minkowski dimension $3/2$ and is conjectured to have the same Hausdorff dimension.

1.3 Sets defined by digit restrictions

In this section we will consider some more complicated sets for which the Minkowski dimension is easy to compute, but the Hausdorff dimension is not

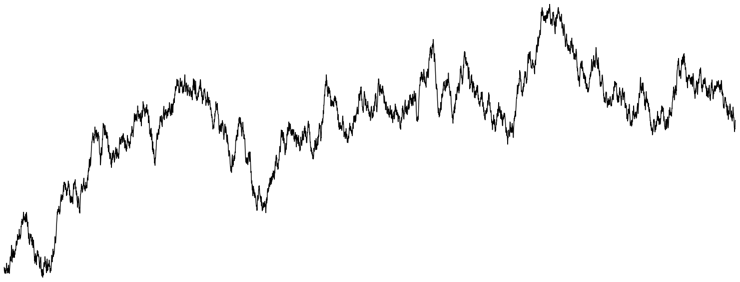

Figure 1.2.4 1-dimensional Brownian motion. This graph has dimension 3/2 almost surely.

so obvious, and will be left to later sections. These subsets of $[0,1]$ will be defined by restricting which digits can occur at a certain position of a number's b-ary expansion. In a later section we will consider sets defined by the asymptotic distribution of the digits. We start by adapting Hausdorff measures to b-adic grids.

Let $b \geq 2$ be an integer and consider b-adic expansions of real numbers, i.e., to each sequence $\{x_n\} \in \{0,1,\dots,b-1\}^{\mathbb{N}}$ we associate the real number

$$x = \sum_{n=1}^{\infty} x_n b^{-n} \in [0,1].$$

b-adic expansions give rise to Cantor sets by restricting the digits we are allowed to use. For example, if we set $b=3$ and require $x_n \in \{0,2\}$ for all n we get the middle thirds Cantor set $\mathbf{C}$.

For each integer n let $I_n(x)$ denote the unique half-open interval of the form $[\frac{k-1}{b^n}, \frac{k}{b^n})$ containing x. Such intervals are called **b-adic intervals of generation n** (**dyadic** if $b=2$).

It has been observed (by Frostman (1935) and Besicovitch (1952)) that we can restrict the infimum in the definition of Hausdorff measure to coverings of the set that involve only b-adic intervals and only change the value by a bounded factor. The advantage of dealing with these intervals is that they are nested, i.e., two such intervals either are disjoint or one is contained in the other. In particular, any covering by b-adic intervals always contains a subcover by disjoint intervals (just take the maximal intervals). Furthermore, the b-adic intervals can be given the structure of a tree, an observation that we will use extensively in later chapters.