Contents

CHAPTER 1

Minkowski and Hausdorff dimension

In this chapter we will define the Minkowski and Hausdorff dimensions of a set and will compute each in a few easy examples. We will then prove Billingsley's lemma and state 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 Sierpinski gasket and random intersections of the middle thirds Cantor set with itself.

1. Definitions and mass distribution principle.

Both Minkowski and Hausdorff dimension measure how efficiently a set K can be covered by balls. Minkowski dimension requires that the covering be by balls all of the same size. 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 sizes. This gives a better behaved notion of dimension, but (as we shall see) is usually much more difficult to compute.

Suppose K is a bounded set in $\mathbb{R}^d$ (or any metric space for that matter) and let $N(K, \epsilon)$ be the minimal number of ϵ balls needed to cover K. We define the *upper Minkowski dimension* as

$$\overline{\dim}_M(K) = \limsup_{\epsilon \to 0} \frac{\log N(K, \epsilon)}{\log 1/\epsilon},$$

and the *lower Minkowski dimension*

$$\underline{\dim}_M(K) = \liminf_{\epsilon \to 0} \frac{\log N(K, \epsilon)}{\log 1/\epsilon}.$$

If the two values agree, the common value is simply called the Minkowski dimension of K. The Minkowski dimension says the the number of balls of size ϵ needed to cover K grows like $\epsilon^{-\dim_M(K)}$. Minkowski dimension is sometimes called the *box counting dimension.*

EXAMPLE 1.1. Suppose $K = [0, 1]$. The at least $1/\epsilon$ balls of radius ϵ are needed and clearly $\epsilon^{-1} + 1$ suffice. Thus $\dim_M(K)$ exists and equals 1.

EXAMPLE 1.2. Let $E_{1/3}$ be the usual "middle thirds" Cantor set obtained as follows. Let $E_{1/3}^0 = [0,1]$ define $E_{1/3}^1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1] \subset E_{1/3}^0$ by removing the center interval of length $\frac{1}{3}$. In general, $E_{1/3}^n$ is a union of 2^n intervals of length 3^{-n} and $E_{1/3}^{n+1}$ is obtained be removing the center third of each. This gives a decreasing nested sequence of compact sets and we set $E_{1/3} = \cap_n E_{1/3}^n$.

The construction gives a covering of $E_{1/3}$ that uses 2^n intervals of length 3^{-n}. Thus for $3^{-n} \le \epsilon \le 3^{-n+1}$ we have

$$N(E_{1/3}, \epsilon) \le 2^n,$$

and hence

$$\overline{\dim}_M(E_{1/3}) \le \frac{\log 2}{\log 3}.$$

Conversely, any interval of length 3^{-n} can hit $E \cap I$ for at most two of the 2^n nth generation intervals I. Hence $N(E, 3^{-n}) \ge 2^{n-1}$. Thus

$$\underline{\dim}_M(E_\alpha) \ge \frac{\log 2}{\log 3}.$$

Therefore the Minkowski dimension exists and equals this common value, $\log_3 2$. If at each stage we remove the α (instead of the middle 1/3) we get a Cantor set E_α with Minkowski dimension,

$$\dim_M(E_\alpha) = \frac{\log 2}{\log 2 + \log \frac{1}{1-\alpha}}.$$

The Minkowski dimension has several drawbacks. For example, it need not exist for a general set (see Example 2.2). Hausdorff dimension is another notion of dimension that has the advantage of always being well defined, but is a little more cumbersome to define and calculate.

Given a set K we define the α-dimensional Hausdorff content as

$$\mathcal{H}^\alpha_{(\infty)}(K) = \inf\{\sum |U_i|^\alpha\},$$

where $\{U_i\}$ is a countable cover of K and $|E|$ denotes the diameter of a set $|E|$. More generally we define

$$\mathcal{H}^\alpha_{(\epsilon)}(K) = \inf\{\sum |U_i|^\alpha : E \subset \cup_i U_i, |U_i| \le \epsilon\},$$

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

$$\mathcal{H}^\alpha(K) = \lim_{\epsilon \to 0} \mathcal{H}^\alpha_{(\epsilon)}(K).$$

If K and K' are separated, i.e.,

$$d(K, K') \ge \delta > 0,$$

then

$$\mathcal{H}^\alpha(K \cup K') = \mathcal{H}^\alpha(K) + \mathcal{H}^\alpha(K').$$

In general we have $\leq$. Thus if A is arbitrary and K is closed,

$$\mathcal{H}^\alpha(A) = \mathcal{H}^\alpha(A \cap K) + \mathcal{H}^\alpha(A \cap K^c),$$

which implies closed sets are measurable for $\mathcal{H}^\alpha$. Therefore the Borel sets are also measurable for $\mathcal{H}^\alpha$.

Note that $\mathcal{H}^\alpha_{(\infty)}(K) = 0$ if and only if $\mathcal{H}^\alpha(K) = 0$. The construction of Hausdorff measure can be made a little more general by introducing 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

$$\mathcal{H}^\varphi_{(\infty)}(K) = \inf\{\sum \varphi(|U_i|)\},$$

just as above. 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 for many applications.

Lemma 1.3. *If $\mathcal{H}^\alpha(K) < \infty$ then $\mathcal{H}^\beta(K) = 0$ for any $\beta > \alpha$. Similarly, if $\mathcal{H}^\alpha(K) > 0$ then $\mathcal{H}^\beta(K) = \infty$ for all $\beta < \alpha$.*

Proof. It follows from the definition of $\mathcal{H}^\alpha_{(\epsilon)}$ that

$$\mathcal{H}^\beta_{(\epsilon)}(K) \leq \epsilon^{(\beta-\alpha)} \mathcal{H}^\alpha_{(\epsilon)}(K),$$

which gives the desired result as $\epsilon \to 0$. The second claim is proved in a similar manner. $\square$

Thus if we think of $\mathcal{H}^\alpha(K)$ as a function of α it must be either 0 or ∞ except possible for one value of α. The value of α where the measure jumps is the Hausdorff dimension of the set.

Definition 1.4. The *Hausdorff dimension* of K is defined to be

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

The following relationship to Minkowski dimension is clear

$$\dim(K) \leq \underline{\dim}_M(K) \leq \overline{\dim}_M(K). \tag{1.1}$$

Strict inequality is possible.

Example 1.5. Consider the countable set $K = \{0\} \cup \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\}$. The distance between $\frac{1}{n}$ and $\frac{1}{n+1}$ is

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

So if we choose n so that $(n+1)^{-2} < \epsilon \leq n^{-2}$, then $n \approx \epsilon^{-1/2}$ and n distinct intervals of length ϵ are need to cover the points $1, \frac{1}{2}, \dots \frac{1}{n}$. The remaining interval $[0, \frac{1}{n}]$ can be covered by n more intervals of length ϵ. Thus

$$\frac{1}{2}\epsilon^{-1/2} \leq N(K,\epsilon) \leq 2\epsilon^{-1/2}.$$

Thus $\dim_M(K) = 1/2$. However, any countable set has Hausdorff dimension 0, for if enumerate the points $\{x_1, x_2, \dots\}$ and cover the nth point by a ball of radius $r_n = \epsilon e^{-n}$ we can make $\sum_n r_n^\alpha$ as small as we wish for any $\alpha > 0$. Thus K is compact set for which the Minkowski dimension exists, but is different from the Hausdorff dimension.

EXAMPLE 1.6. Consider the Cantor set obtained by replacing the unit square in the plane by four equal sub-squares of side length α. See Figure 1.1. It is is easy to see that we can cover the set by 4^n disk of size $2 \cdot \alpha^n$. Thus

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

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

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

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

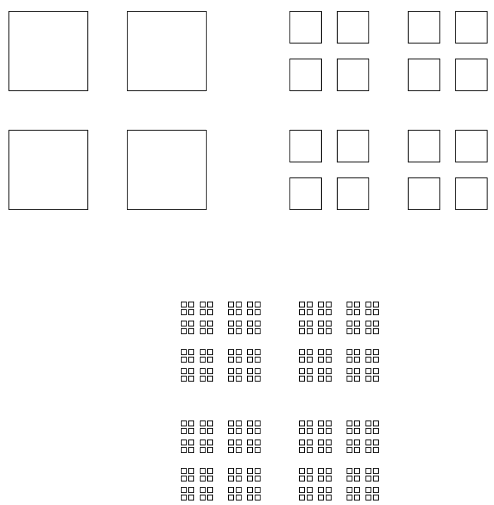

FIGURE 1.1. Constructing a Cantor set

Thus we automatically get the Hausdorff dimension estimate $\dim(E) \leq d(\alpha)$. To prove the equality requires a little more work. The general procedure to prove a lower bound on Hausdorff dimension is to construct an appropriate measure on the set (we will see this technique many times).

In this case, we consider the probability measure μ defined on the Cantor set which gives each of the square of the nth generation the same mass (namely 4^{-n}. One then shows that this measure satisfies the estimate

$$\mu(B(x,r)) \leq Cr^{d(\alpha)},$$

for all disks and some uniform $C < \infty$. 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

$$\mu(B \cap E) \leq 4 \cdot 4^{-n} \leq 4\alpha^{nd_\alpha} \leq 4\alpha^{-d_\alpha} r^{d_\alpha}.$$

Using this, we see that given any covering of the Cantor set by disks, $\{D(x_j, r_j)\}$, we have

$$1 \leq \sum_j \mu(D(x_j, r_j)) \leq C^{-1} r_j^{d(\alpha)},$$

which shows the dimension is $\geq d(\alpha)$.

This argument proves the following

Lemma 1.7 (Mass Distribution Principle). *If E supports a strictly positive measure μ which satisfies*

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

for every ball $B(x,r)$, then $\mathcal{H}^d(E) \geq \mu(E)/C$. In particular, $dim_H(E) \geq d$.

This is one of the major ideas used throughout the book; upper bounds for Hausdorff dimension come from finding explicit coverings of the set, but lower bounds are always 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 3.1) and will generalize it even further in later chapters.

Example 1.8. Another simple set for which the two dimensions agree and is easy to compute is the von Koch snowflake. To construct this we start with an equilateral triangle. At each stage we we continue by adding 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 for the first four iterations of this process. The boundary of this region is a curve with dimension $\log 4/\log 3$. 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.

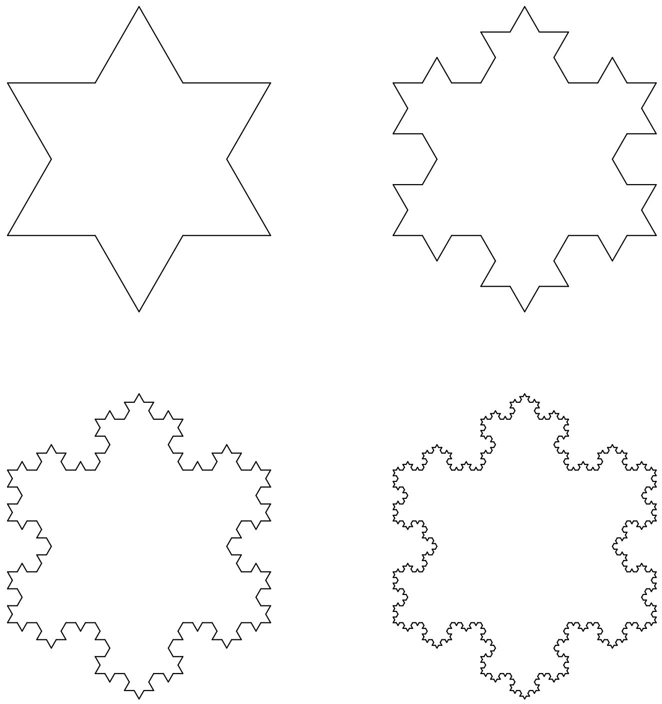

FIGURE 1.2. Creating the snowflake

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

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

where b is an integer larger than 1 and $0 < \alpha, 1$. It is conjectured that the Hausdorff dimension of its graph is $2 - \alpha$, but this has not been proven (see Chapter 6 for some related results).

2. Sets defined by digit restrictions

In this section we will consider some more complicated sets whose Minkowski dimensions are easy to compute, but whose Hausdorff dimensions are not 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 introducing b-adic grids and net measures.

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

$$x = \sum_{n=0}^{\infty} x_n b^{-n}.$$

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. It has been observed (by Frostman and Besicovitch) that we can restrict the infimum in the definition of Hausdorff measure to coverings of the the set which only involve intervals of the form $[\frac{k-1}{b^n}, \frac{k}{b^n}]$ (for subsets of the real line). Such intervals are called b-adic (dyadic if $b=2$). The advantage of dealing with these intervals is that they are nested; i.e., if I and J are two such then either they are disjoint or one is contained in the other. In particular any covering by b-adic intervals always contains a disjoint subcover. Furthermore, the b-adic intervals can be given the structure of a tree, an observation that we will use extensively in later chapters.

We define the *grid Hausdorff measures.* Define

$$\tilde{\mathcal{H}}^{\alpha}_{(\epsilon)}(A) = \inf\{\sum_i |J_i|^{\alpha}\}$$

$$\tilde{\mathcal{H}}^{\alpha}(A) = \lim_{\epsilon \to 0} \mathcal{H}^{\alpha}_{(\epsilon)}(A),$$

where the infimum is over all coverings of A by collections of $\{J_i\}$ of b-adic intervals. Then clearly

$$\tilde{\mathcal{H}}^{\alpha}(A) \geq \mathcal{H}^{\alpha}(A),$$

since we are taking the infimum over a smaller set of coverings. However, the two sides are almost the same size, i.e.,

$$\tilde{\mathcal{H}}^{\alpha}(A) \leq (b+1)\mathcal{H}^{\alpha}(A),$$

since any interval I can be covered by at most $(b+1)$ b-adic intervals (just take the smallest n so that $b^{-n} \leq |I|$ and take all b-adic intervals of length b^{-n} which hit I). Thus $\tilde{\mathcal{H}}^{\alpha}(A) = 0$ if and only if $\mathcal{H}^{\alpha}(A) = 0$ and so we can use $\tilde{\mathcal{H}}^{\alpha}$ to find Hausdorff or Minkowski dimensions of sets.

Definition 2.1. For $S \subset \mathbb{N}$ the *upper density* of S is

$$\overline{\mathrm{d}}(S) = \limsup_{N\to\infty} \frac{|S \cap \{1,\dots,N\}|}{N}.$$

The *lower density* is of course

$$\underline{\mathrm{d}}(S) = \liminf_{N\to\infty} \frac{|S \cap \{1,\dots,N\}|}{N},$$

and if the two values agree the limit exists and is called $d(S)$, the *density* of S.

Now for some examples.

Example 2.2. Suppose $S \subset \mathbb{N}$, and define $A_S = \{x = \sum_{k=1}^{\infty} x_k 2^{-k}\}$ where

$$x_k \in \begin{cases} \{0,1\}, & k \in S \\ \{0\}, & k \notin S \end{cases}.$$

The set S hits exactly

$$2^{\sum_{k=1}^{n} \chi_S(k)},$$

dyadic intervals of generation n, so

$$\log_2 \tilde{N}(S, 2^{-n}) = \sum_{k=1}^{n} \chi_S(k).$$

Thus

$$\frac{\log \tilde{N}(S, 2^{-n})}{\log 2^{-n}} = \frac{1}{n}\sum_{k=1}^{n} \chi_S(k),$$

whcih implies

$$\overline{\dim}_M(A_S) = \overline{\mathrm{d}}(S),$$
$$\underline{\dim}_M(A_S) = \underline{\mathrm{d}}(S).$$

It is easy to construct sets S where the liminf and limsup differ and hence we get compact sets where the Minkowski dimension does not exist. See Exercise 2.

The Hausdorff dimension of A_S is equal to the lower Hausdorff dimension. We shall prove this in the next section as an application of Billingsley's lemma.

We can also construct A_S as follows. Start with the interval $[0,1]$ and subdivide it into two equal length subintervals $[0,1/2]$ and $[1/2,1]$. If $1 \in S$ then keep both intervals and if $1 \notin S$ then keep only the leftmost, $[0,1/2]$. Cut each of the remaining intervals in half, keeping both subintervals if $2 \in S$ and only keeping the left interval otherwise. In general, at the nth step we have a set $A_n^S \subset A_{n-1}^S$ which is a finite union of intervals of length 2^n (some may be adjoining). We cut each

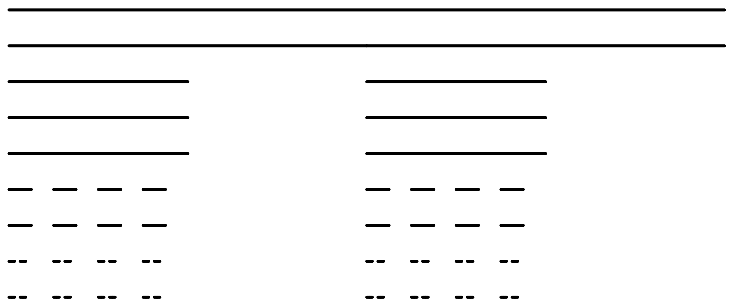

FIGURE 2.1. First 8 generations of A_S for $S = \{1, 3, 4, 6, 8, 10, \dots\}$.

of intervals in half, keeping both subintervals if $n \in S$ and throwing away the right hand one if $n \notin S$. The limiting set is $A_S = \lim A_n^S$. In Figure 2.1 we have drawn the first few generations of the construction corresponding to $S = \{1, 3, 4, 6, 8, 10, \dots\}$.

EXAMPLE 2.3. The *shifts of finite type* are defined by restricting what digits can follow other digits. Let $A = (a_{ij})$, $0 \le i, j < b$ be a $b \times b$ matrix of 0's and 1's and define

$$X_A = \{\sum_{n=1}^{\infty} x_n b^{-n} : A_{x_n x_{n+1}} = 1 \text{ for all } n \ge 1\}.$$

This condition on the $\{x_n\}$ is clearly shift invariant and is the intersection of countable many "closed" conditions, so the set X_A is compact and T_b invariant. For example, if

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

then $x \in X_A$ if and only if the binary representation of x has no two consecutive 1's. Any such matrix can be represented by a directed graph were the vertices represent the numbers $\{0, \dots, b-1\}$ and a directed edge connects i to j if $A_{ij} = 1$. For example, the 2×2 matrix A above is represented by Figure 2.2. The first ten generations in the construction of the Cantor set X_A are shown in Figure 2.3.

An element of $x = \sum_n x_n b^{-n} \in X_A$ corresponds in a natural way to infinite path $\{x_1, x_2, \dots\}$ on this graph and a b-adic intervals hitting X_A correspond to a finite paths. Conversely, any finite a path of length n with vertices $\sigma_1, \dots, \sigma_n$ corresponds to a b-adic interval of length b^{-n} hitting X_A. Thus $N_n(X_A)$ is the number of distinct paths of length n is the graph. It is an easy fact from graph theory (actually a fact about matrix multiplication) that the number of paths of length n from i to j

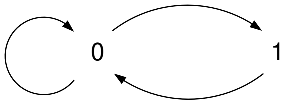

FIGURE 2.2. The directed graph representing A

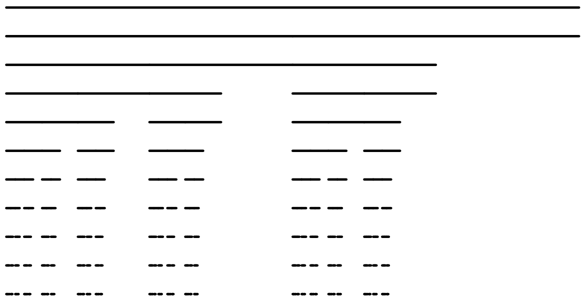

FIGURE 2.3. Approximations of X_A

is the $(A^n)_{ij}$, (i.e., the (ith, jth) entry in the matrix A^n). See Exercise **??**. Thus the total number of length n paths is

$$N_n(X_A) = \mid A^n \mid,$$

where the norm of a matrix is defined as

$$\mid B \mid = \sum_{i,j} |B_{ij}|.$$

Thus

$$\dim_M(X_A) = \lim_{n\to\infty} \frac{\log N_n(X_A)}{n \log b} = \lim_{n\to\infty} \frac{\log[N_n(X_A)]^{1/n}}{\log b} = \frac{\log \rho(A)}{\log b},$$

where

$$\rho(A) = \lim_{n\to\infty} \mid A^n \mid^{1/n},$$

is, by definition, the *spectral radius* of the matrix A. To see why the limit must exist set $a_n = \log \mid A^n \mid$ and note that the resulting sequence satisfies

$$a_{nm} \leq a_n + a_m.$$

Now apply Exercise 9. For the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

the spectral radius is $(1+\sqrt{5})/2$ (Exercise 12), so the Minkowski dimension X_A is

$$\log \rho(A) = \frac{(\log(1+\sqrt{5}) - \log 2)}{\log 2}.$$

We shall see in Example **??** of Chapter 3, that for these examples the Hausdorff and Minkowski dimensions agree.

If $A = (a_{ij})$ has real, non-negative entries then the Perron-Frobenius theorem says that $\rho(A)$ is an eigenvalue of A of maximal modulus. Moreover, if A is primitive, then all the other eigenvalues are strictly smaller in modulus. *Primitive* means that some power A^n of A has all positive entries. Thus the dimension of X_A can (in theory) be computed from A. Parry first gave this formula for $\dim_M(X_A)$ in 1964, and Furstenberg proved the equality with Hausdorff dimension in 1967. We will give the proof in Section **??** of Chapter 3.

A trivial case where it is easy to compute $\dim_M(X_A)$ is

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

The corresponding graph is drawn in Figure 2.4. X_A is easily seen to have only four elements $\{0\bar{1}, \bar{1}, 2\bar{1}, 20\bar{1}\}$, where the $\bar{1}$ denotes an infinite sequences of 1's. Thus $\dim(X_A) = 0$.

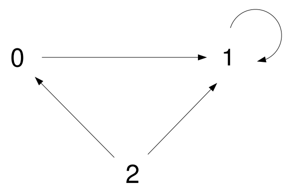

FIGURE 2.4. All paths end at 1

EXAMPLE 2.4. Now we consider some sets in the plane. Suppose A is a $b \times b$ matrix of 0's and 1's and let

$$Y_A = \{(x, y) : A_{x_n, y_n} = 1 \forall n\},$$

where $\{x_n\}, \{y_n\}$ are the b-ary expansions of x and y. For example, if

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix},$$

the resulting set is the Sierpinski Gasket. See Figure 2.5.

If

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

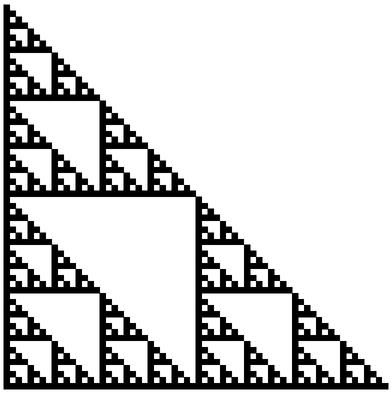

FIGURE 2.5. The Sierpinski gasket (6 generations).

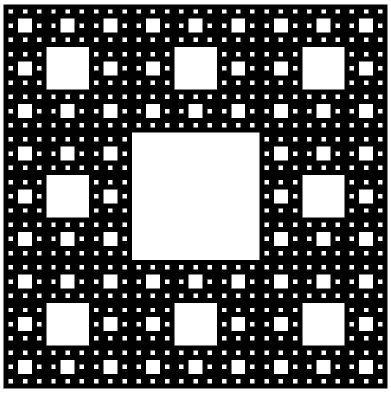

FIGURE 2.6. Sierpinski carpet (4 generations)

Then we get $E_{1/3} \times E_{1/3}$, the product of the middle thirds Cantor set with itself. For

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

we get the Sierpinski carpet.

If the matrix A has r 1's in it, then we can cover Y_A by r^n squares of size b^{-n} and it is easy to see this many are needed. Thus $\dim_M(Y_A) = \log r / \log b \log_b r$. Defining a measure on Y_A which gives equal mass to each of the nth generational squares and using the Mass Distribution Principle shows that the Hausdorff dimension of the set is also $\log_b r$.

The Sierpinski gasket (as well as the other examples so far) are "self-similar" in the sense that they are invariant under certain collection of maps which are constructed from isometries and dilations. In some sense these sets are "easy" to understand. The basic results are covered in Section ?? of Chapter 3.

If we replace the square matrices of the previous example by rectangular ones (i.e., use different base expansions for x and y) we get a much more difficult class of sets which are invariant under affine maps. These sets will be the studied in a later chapter, but we give an example here.

EXAMPLE 2.5. We modify the previous example by taking a non-square matrix. Suppose $m < n$ and suppose A is $m \times n$ matrix of 0's and 1's. Let

$$Y_A = \{(x,y) : A_{y_k,x_k} = 1 \forall k\},$$

where $\{x_k\}$ is the n-ary expansion of x and $\{y_k\}$ is the m-ary expansion of y. For example, if

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

we obtain the set on the right of Figure 2.5

This set can also be described as follows. We start with a box Q and divide it into 6 sub-rectangles by making on horizontal and 3 vertical subdivisions. Choose one of the squares on the top row and two squares from the bottom row. Now subdivide the chosen rectangles in the same way and select the corresponding sub-rectangles. Figure 2.5 illustrates some choice of boxes and the resulting set after 5 iterations.

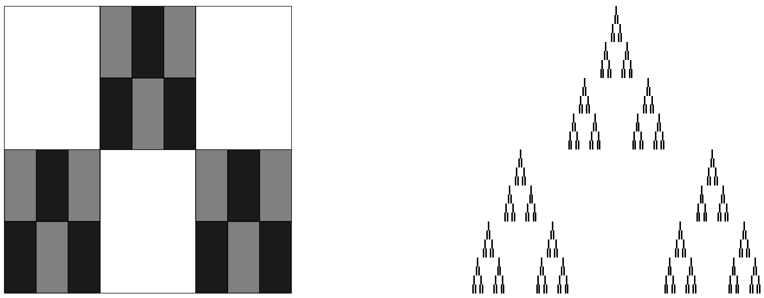

FIGURE 2.7. The McMullen set

The Minkowski dimension of this set is $1 + \log_3 \frac{3}{2} \approx 1.58496$ (see Theorem ?? in Chapter 5). We shall eventually prove C. McMullen's result that the Hausdorff dimension of this set is $\log_2(2^{\log_3 2} + 1) \approx 1.34968$ which is strictly less than the Minkowski dimension. We shall prove later in this chapter that its Hausdorff dimension is strictly larger than 1.

3. Billingsley's lemma

In this section we prove a more refined version of the mass distribution principle. An even more general result (the Frostman-Rodgers-Taylor density theorem) will be proved later in Section **??** of Chapter 4. Let $b \geq 2$ be an integer and for let $I_n(x)$ denote the nth generation b-adic interval containing x.

LEMMA 3.1 (Billingsley's lemma). *Let $A \subset [0,1]$ be closed and let μ be a finite measure on $[0,1]$. Suppose $\mu(A) > 0$. If*

$$\alpha_1 \leq \liminf_{n\to\infty} \frac{\log \mu(I_n(x))^{-1}}{\log |I_n(x)|^{-1}} \leq \alpha_2,$$

for all $x \in A$, then $\alpha_1 \leq dim(A) \leq \alpha_2$.

PROOF. Manipulating the right hand inequality, we get

$$\limsup_{n\to\infty} \frac{\mu(I_n(x))}{|I_n(x)|^{\alpha_2}} \geq 1, \tag{3.1}$$

and the left yields

$$\limsup_{n\to\infty} \frac{\mu(I_n(x))}{|I_n(x)|^{\alpha_1}} \leq 1. \tag{3.2}$$

First we will show that (3.1) implies $\mathcal{H}^\alpha(A) \leq \tilde{\mathcal{H}}^\alpha(A) \leq C^{-1}\mu(A)$. Then we will show that (3.2) implies $\tilde{\mathcal{H}}^\alpha(A) \geq C^{-1}\mu(A)$. The lemma follows from these two claims.

We start with the first assertion. Fix $\epsilon > 0$ and choose an open set $V \supset A$ with $\mu(V)$ close to $\mu(A)$. For every $x \in A$ we can find integers n as large as we wish and such that

$$\frac{\mu(I_n(x))}{|I_n(x)|^\alpha} > C.$$

Take $n(x)$ to be the minimal integer satisfying this condition and such that $b^{-n} \leq \epsilon$ and $I_n(x) \subset V$.

Now $\{I_{n(x)}(x)\}$ is a cover of A, and suppose $\{J_k\}$ is a disjoint subcover. This covering of A has the property that that $|J_k| \leq \epsilon$ for all k (by our choice of $n(x)$) and

$$\sum_k |J_k|^\alpha \leq \sum_k \mu(J_k) \leq \mu(V).$$

This implies

$$\tilde{\mathcal{H}}^\alpha_\epsilon(A) \leq C^{-1}\mu(A).$$

Therefore $\tilde{\mathcal{H}}^\alpha(A) \leq C^{-1}\mu(A)$.

Next we prove the second assertion. Let

$$A_m = \{x \in A : \mu(I_n(x)) < C|I_n(x)| \text{ for all } n > m\}.$$

A trick of Frostman is to look at the set A_m instead of A. Since we have $A = \cup_m A_m$ and $A_{m+1} \supset A_m$ we have $\mu(A) = \lim_{m\to\infty} \mu(A_m)$. Thus it suffices to prove the estimate for A_m.

Fix $\epsilon < b^{-m}$ and consider any cover of A by b-adic intervals $\{J_k\}$ with $|J_k| \leq \epsilon$. With out loss of generality we assume every interval in the covering intersects A. Then

$$\sum_k |J_k|^\alpha \quad \geq \sum |J_k|^\alpha \geq \sum \mu(J_k) \geq C^{-1}\mu(A_m).$$

This shows $\tilde{\mathcal{H}}^\alpha_{(\epsilon)}(A) \geq C^{-1}\mu(A_m)$. Taking $\epsilon \to 0$ and $m \to \infty$ gives the desired result. $\square$

EXAMPLE 3.2. For $S \subset \mathbb{N}$, define $A_S = \{x = \sum_{k=1}^\infty x_k 2^{-k}\}$ where

$$x_k \in \begin{cases} \{0,1\}, & k \in S \\ \{0\}, & k \notin S \end{cases}.$$

We claimed in Example 2.2 that

$$\dim(A_S) = \liminf_{N\to\infty} \frac{|S \cap \{1,\dots,N\}|}{N}.$$

To prove this, let μ_n be Lebesgue measure on A_n^S, normalized to have mass 1, and let $\mu \in \Pr(A_S)$ be the limiting measure. For any $x \in A_S$,

$$\frac{\log 1/\mu I_n(x)}{\log 1/|I_n(x)|} = \frac{\log 2^{-|S\cap\{1,\dots,n\}|}}{\log 2^{-n}} = \frac{|S \cap \{1,\dots,n\}|}{n}.$$

Thus the liminf of the left hand side is the liminf of the right hand side. By Billingsley's lemma, this proves the claim.

It is convenient to introduce the dimension of a measure.

DEFINITION 3.3. If μ is a measure on $\mathbb{R}^n$ we define

$$\dim(\mu) = \inf\{\dim(A) : \mu(A^c) = 0, \emptyset \neq A \subset \mathbb{R}^n\}.$$

Observe that the infimum is really a minimum. If $\dim(A_n) \to \dim(\mu)$ and $\mu(A_n^c) = 0$ for all n then $A = \cap A_n$ satisfies $\dim(A) = \dim(\mu)$. An equivalent definition is to write

$$\dim(\mu) = \inf\{\alpha : \mu \perp \mathcal{H}^\alpha\}$$

where $\mu \perp \nu$ means the two measure are mutually singular, i.e., there is a set $A \subset \mathbb{R}^n$ such that $\mu(A) = \nu(A^c) = 0$.

As before, let b be a positive integer and given $x \in [0,1]$ let $I_n(x)$ denote the b-adic interval of the form $[\frac{j-1}{b^n}, \frac{j}{b^n})$ containing x where $j \in \{1, \dots b^n\}$. Let

$$\alpha_\mu = \text{ess sup}\{\liminf_{n\to\infty} \frac{\log 1/\mu(I_n(x))}{\log 1/|I_n(x)|}\},$$

where the essential supremum is taken with respect to μ.

LEMMA 3.4. *$dim(\mu) = \alpha_\mu$.*

PROOF. To see this, first take

$$A = \{x : \liminf_{n\to\infty} \frac{\log 1/\mu(I_n(x))}{\log 1/|I_n(x)|} < \alpha\},$$

for $\alpha > \alpha_\mu$. Because $\mu(A^c) = 0$ (definition of essential supremum) we have $\dim(\mu) \le \dim(A)$ and $\dim(A) \le \alpha$ by Billingsley's lemma. Thus taking $\alpha \to \alpha_\mu$ gives $\dim(\mu) \le \alpha_\mu$.

To get the other direction we use the other part of Billingsley's lemma. For $\alpha < \alpha_\mu$ look at the set

$$B = \{x : \liminf_{n\to\infty} \frac{\log 1/\mu(I_n(x))}{\log 1/|I_n(x)|} > \alpha\}.$$

By the definition of α_μ B has positive μ measure and If A is any set with full μ measure then $0 < \mu(B) = \mu(A \cap B)$ and

$$\frac{\log 1/\mu(I_n(x))}{\log 1/|I_n(x)|} > \alpha,$$

on $A \cap B$, so applying Billingsley's lemma to $A \cap B$ shows $\dim(A) \ge \dim(A \cap B) > \alpha$. Thus $\dim(\mu) \ge \alpha$ for all $\alpha < \alpha_\mu$, and we deduce $\dim(\mu) = \alpha_\mu$. □

EXAMPLE 3.5. In the first section we constructed a measure μ on the middle third Cantor set $E_{1/3}$ by giving equal mass to each nth generation interval in the construction. If we let $b = 3$ and consider a 3-adic interval I of length 3^{-n} then

$$\mu(I) = 2^{-n} = |I|^{n \log_3 2},$$

if I hits $E_{1/3}$ and is 0 otherwise. Thus

$$\lim_{n\to\infty} \frac{\log 1/\mu(I_n(x))}{\log 1/|I_n(x)|} = \log_3 2,$$

for all $x \in E_{1/3}$, (and hence for μ almost every x), so $\dim(\mu) = \log_3 2$.

4. Sets defined by digit frequency

We previously considered sets with restrictions on what the nth digit of the b-ary expansion could be. In this section we not restrict particular digits, but will require that each digit occurs with a certain frequency. The resulting sets are not closed (in general they will be dense), so we will only consider the Hausdorff dimension of such sets.

EXAMPLE 4.1. $A_p = \{x = \sum_{n=1}^{\infty} x_n 2^{-n} : \lim_{j\to\infty} \frac{1}{j}\sum_{k=1}^{j} x_k = p\}$.

Thus A_p is the set of real numbers in $[0,1]$ so that a 1 occurs in the binary expansion about a pth of the time. For "typical" real numbers we expect a 1 to occur about half the time, and indeed, $A_{1/2}$ is a set of full measure in $[0,1]$ (see below).

In general we will show

$$\dim(A_p) = H_2(p) = -p\log_2 p - (1-p)\log_2(1-p).$$

This quantity is called the entropy of p and is strictly less than 1 except for $p = 1/2$. It is a measure of the uncertainty associated to the probability (if $p = 0$ or 1 the entropy is 0; it is maximized when $p = 1/2$). In addition to Billingsley's lemma we will need

THEOREM 4.2 (Law of Large Numbers). *Let $\{f_n\}$, $n = 1, 2\dots$ be a sequence of orthogonal functions on a probability space $(X, d\nu)$ and suppose $E(f^2) = \int |f|^2 d\nu \le 1$. Then*

$$\frac{1}{n}S_n = \frac{1}{n}\sum_{k=1}^{n} f_k \to 0,$$

a.e. (with respect to ν) as $n \to \infty$.

We will give a very simple proof of this in Appendix B and discuss some of the history of the result there. This version is sometimes called the strong law of large numbers because it gives almost everywhere convergence, as opposed to the weak version which only says that $\frac{1}{n}S_n$ converges to 0 in L^2.

Define a measure probability μ_p on $[0,1]$ as follows. It suffices to define μ on dyadic intervals. Let

$$\mu_p([\frac{j-1}{2^n}, \frac{j}{2^n}) = p^{k(j)}(1-p)^{n-k(j)},$$

where $k(j)$ is the number of 1's in the binary expression of k. We can also write μ_p as

$$\mu_p(I_n(x)) = p^{\sum_{k=1}^{n} x_k}(1-p)^{n-\sum_{k=1}^{n} x_k},$$

where $x = \sum x_k 2^{-k}$.

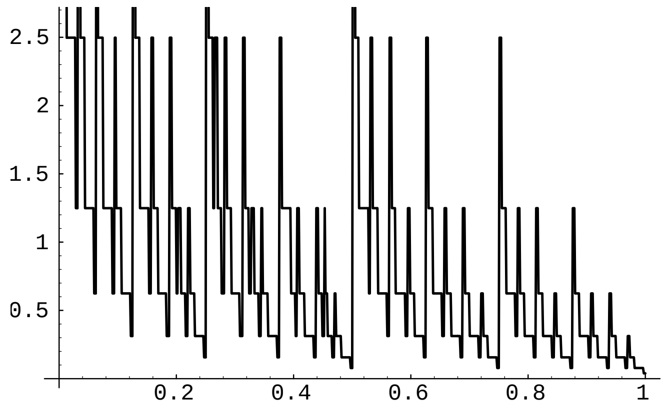

FIGURE 4.1. Graph of the density of $\mu_{1/3}$.

An alternate way of describing μ is that it gives the whole interval $[0,1]$ mass 1, and gives the two subintervals $[0,1/2]$, $[1/2,1]$ mass $1-p$ and p respectively. In general if a dyadic interval I has measure $\mu(I)$ the left half has measure $(1-p)\mu(I)$ and the right half has measure $p\mu(I)$. Observe that if $p = 1/2$ then this is exactly Lebesgue measure on $[0,1]$. In Figure 4.1 we have graphed the density for $p = 1/3$ for intervals of down to size 2^{-7}.

LEMMA 4.3. *$dim(A_p) = dim(\mu) = h_2(p)$*

PROOF. We start by proving $\mu_p(A_p) = 1$. To see this let $f_n(x) = (x_n - p)$ (where $x = \sum x_k 2^{-k}$ is the binary expansion of x). If $S_n = \sum_{k=1}^n f_n$, then unwinding definitions shows that A_p is exactly the set where $\frac{1}{n} S_n \to 0$, so we will be done if we can apply the LLN to $\{f_n\}$ and the measure μ_p. The first thing we need to check is that $\int f_n^2 d\mu_p \leq 1$. This is easy since $|f_n| \leq 1$ and μ_p is a probability measure. To check the orthogonality, note

$$\begin{aligned}\int (x_n - p)(x_m - p) d\mu_p &= (1-p)^2(\mu_p(E_{11}) \\ &\quad -p(1-p)(\mu_p(E_{10}) + \mu(E_{01})) \\ &\quad +p^2(\mu_p(E_{00}),\end{aligned}$$

where $E_{ij} = \{x : x_n = i, x_m = j\}$. If $n \neq m$ then since μ is a product measure,

$$\mu_p(E_{11}) = p^2, \quad \mu_p(E_{10}) = \mu(E_{01}) = p(1-p), \quad \mu_p(E_{00}) = (1-p)^2.$$

Substituting into the expression above gives $\int f_n f_m d\mu_p = 0$ for $n \neq m$, as desired. Thus $\mu(A_p) = 1$.

It is now an easy matter to compute the dimension of the measure μ_p. Note

$$\frac{\log 1/\mu_p(I_n(x))}{\log 1/|I_n(x)|} = \frac{1}{\log 2}((\frac{1}{n}\sum_{k=1}^{n} x_k)\log\frac{1}{p} + (\frac{1}{n}\sum_{k=1}^{n}(1-x_k)\log\frac{1}{1-p})).$$

Since μ_p almost every point is in A_p, we see that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} x_k = p,$$

for μ_p a.e. x. Therefore,

$$\frac{\log 1/\mu_p(I_n(x))}{\log 1/|I_n(x)|} \to \frac{p\log\frac{1}{p} + (1-p)\log\frac{1}{1-p}}{\log 2} = h_2(p),$$

for a.e. x with respect to μ_p. Thus by Lemma 3.4 $\dim(\mu_p) = h_2(p)$. □

Since we have already checked that $\mu_p(A_p) = 1$ we deduce $\dim(A_p) \geq h_2(p)$ (by the definition of $\dim(\mu_p)$). However, since

$$A_p \subset \{x \in [0,1] : \liminf_{n\to\infty}\frac{\log 1/\mu_p(I_n(x))}{\log 1/|I_n(x)|} \leq h_2(p)\},$$

and $\mu_p(A_p) = 1$, Billingsley's lemma implies $\dim(A_p) \leq h_2(p)$, hence equality.

Example 4.4. Consider the related sets

$$\tilde{A}_p = \{x \in [0,1] : \limsup_{n\to\infty}\sum_{k=1}^{n} x_k \leq p\},$$

$$\hat{A}_p = \{x \in [0,1] : \liminf_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} x_k \leq p\}.$$

Clearly we have $A_p \subset \tilde{A}_p \subset \hat{A}_p$ and for $q \leq p$ we have $A_q \subset \tilde{A}_p$. Since $\dim(A_{1/2}) = 1$ we have $\dim(\tilde{A}_p) = \dim(\hat{A}_p) = 1$ if $p \geq 1/2$.

If $p < 1/2$ clearly

$$h_2(p) = \dim(A_p) \leq \dim(\tilde{A}_p) \leq \dim(\hat{A}_p).$$

On the other hand, if $x \in \hat{A}_p$ and taking any $q < p$ our previous computations imply

$$\liminf_{n\to\infty}\frac{\log 1/\mu_p(I_n(x))}{\log 1/|I_n(x)|} \leq q\log_2\frac{1}{p} + (1-q)\log_2\frac{1}{1-p},$$

for a.e. x with respect to μ_q. Since μ_q gives full measure to A_q and $A_q \subset \hat{A}_p$, it gives full mass to $\hat{A}_p$ as well. Therefore Billingsley's lemma implies

$$\dim(\hat{A}_p) \leq q \log_2 \frac{1}{p} + (1-q) \log_2 \frac{1}{1-p}.$$

Since $p < 1/2$, this is an increasing function of q and thus

$$q \log_2 \frac{1}{p} + (1-q) \log_2 \frac{1}{1-p} \leq p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p} = h_2(p).$$

Thus $\dim(\tilde{A}_p) = \dim(\hat{A}_p) = h_2(p)$.

Example 4.5. The same argument that we just used to compute the dimension of A_p as works in a more general setting. Suppose $\mathbf{p} = (p_0, \dots, p_{b-1})$ is a probability vector, i.e., $\sum_{k=0}^{b-1} p_k = 1$ and define a measure $\mu_{\mathbf{p}}$ by

$$\mu_{\mathbf{p}}(I_n(x)) = \prod_{k=1}^{n} p_{x_n},$$

where $\{x_n\}$ is the b-ary expansion of x. Then repeating the proof of the lemma shows

$$\dim(\mu_{\mathbf{p}}) = h_b(\mathbf{p}) = \sum_{k=0}^{b-1} p_k \log \frac{1}{p_k}.$$

Similarly, the set of x's in $[0,1]$ such that

$$\lim_{N\to\infty} \frac{|\{n \in [0,N] : x_n = k\}|}{N} = p_k,$$

for each $k = 0, \dots b-1$ has dimension $h_b(\mathbf{p})$.

The following is a special case we will need in a later section. Fix an integer $b \geq 2$ and a set $E \subset \{0, 1, \dots b-1\}$. Let $\{x_n\}$ be the b-ary expansion of the real number x and let $e_n = 1$ if $x_n \in E$ and 0 otherwise.

Lemma 4.6. *The set X of x's so that*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{\infty} e_n = p,$$

has dimension

$$p \log_b(\frac{p}{|E|}) + (1-p) \log_b(\frac{1-p}{b-|E|}).$$

PROOF. Define a measure by

$$\mu(I_n(x)) = (\frac{p}{|E|})^{\sum_{k\le n} e_k}(\frac{1-p}{b-|E|})^{\sum_{k\le n} 1-e_k}.$$

The proof of Lemma 4.3 repeated here shows that

$$\dim(X) = \dim(\mu) = p\log_b(\frac{p}{|E|}) + (1-p)\log_b(\frac{1-p}{b-|E|}).$$

□

We will use this formulation later in the chapter when we consider intersections of random translates of Cantor sets (see proof of Proposition 6.8).

5. Slices

If $A \subset \mathbb{R}^2$ has dimension α what should we expect the dimension of $A \cap L$ to be, where L is a "random" line? For the present, let us consider only vertical lines and set

$$A_x = \{y : (x,y) \in A\}.$$

THEOREM 5.1 (Marstrand Slicing theorem). *Let $A \subset \mathbb{R}^2$ and suppose $dim(A) \ge 1$. Then*

$$dim(A_x) \le dim(A) - 1,$$

for (Lebesgue) almost every x.

If $\dim(A) < 1$ then A_x is empty for almost every x, indeed is empty except possible for a set of dimension $= \dim(A)$ (See Exercise 7). On the other hand it is possible that $\dim(A_x) = \dim(A)$ for some values of d, e.g., if A is a vertical line segment.

The inequality can be strict for every x, e.g., there are real valued functions on $\mathbb{R}$ whose graphs have dimension strictly larger than one, even though all the vertical slices are singletons. See Chapter 6. For example, we will prove the graph of a 1 dimensional Brownian motion has dimension 3/2 almost surely (see Theorem **??**, Chapter 8). We shall give several other examples of strict inequality later in this section.

PROOF. We start with the following claim: for $1 \le \alpha \le 2$

$$\mathcal{H}^\alpha(A) \ge c_2 \int \mathcal{H}^{\alpha-1}(A_x)dx.$$

(We can take $c_2 = 1$. In higher dimensions the analogous constant is $c_n > 1$). To see that it suffices to prove this claim, take $\alpha > \dim(A)$. Then

$$0 = \mathcal{H}^\alpha(A) \ge \int \mathcal{H}^{\alpha-1}(A_x)dx,$$

implies $\mathcal{H}^{\alpha-1}(A) = 0$ for almost every x, i.e., $\dim(A_x) \leq \alpha - 1$ for a.e. x.

Now to prove the claim. Fix $\epsilon > 0$ and $\delta > 0$ and let $\{D_j\}$ be a cover of A with $|D_j| < \epsilon$ and

$$\sum_j |D_j|^\alpha < \mathcal{H}^\alpha_{(\epsilon)}(A) + \delta.$$

Enclose each D_j in a square Q_j with sides parallel to the axes and with side length $s(Q_j) \leq |D_j|$. Let I_j be the projection of Q_j onto the x axis. Define the function

$$f(x, y) = \sum_j \mathbf{1}_{Q_j}(x, y) s(Q_j)^{\alpha-2}.$$

For each x, the slices $\{(Q_j)_x\}$ form a cover of A_x and have length

$$|(Q_j)_x| = \begin{cases} s(Q_j), & x \in I_j \\ 0, & x \notin I_j \end{cases}.$$

(In the higher dimensional case, this is the point where the constant $c_n > 1$ appears. For example, if $n = 3$, then a slice of a cube Q_j is either $\sqrt{2}s(Q_j)$ or 0.)

So we now have an ϵ cover of A_x and

$$\mathcal{H}^{\alpha-1}_{(\epsilon)}(A_x) \leq \sum_j |(Q_j)_x|^{\alpha-1} \leq \sum_{j:x\in I_j} s(Q_j)^{\alpha-1}.$$

For a fixed x,

$$\int_{\mathbb{R}} f(x, y)dy = \int_{\mathbb{R}} \sum_{j:x\in I_j} \mathbf{1}_{Q_j}(x, y)s(Q_j)^{\alpha-2}dxdy = \sum_{j:x\in I_j} s(Q_j)^{\alpha-2}.$$

This implies,

$$\iint f(x, y)dxdy = \int_{\mathbb{R}} \sum j s(Q_j)^{\alpha-1}.$$

Therefore

$$\begin{aligned}\int \mathcal{H}^{(\epsilon)}_{\alpha-1}(A_x)dx &\leq \int_{\mathbb{R}} \sum_j s(Q_j)^{\alpha-1} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y)dxdy \\ &= \sum_j s(Q_j)^{\alpha-2} s(Q_j)^2 \\ &= \sum_j s(Q_j)^{\alpha} \\ &\leq \mathcal{H}^{\alpha}_{(\epsilon)}(A_x) + \delta\end{aligned}$$

Taking $\delta \to 0$ gives

$$\int \mathcal{H}^{(\epsilon)}_{\alpha-1}(A_x)dx \leq \mathcal{H}^{(\epsilon)}_{\alpha}(A).$$

As $\epsilon \to 0$, $\mathcal{H}^{(\epsilon)}_{\alpha-1}(A_x) \nearrow \mathcal{H}_{\alpha-1}(A_x)$, so the Monotone Convergence Theorem implies

$$\int \mathcal{H}_{\alpha-1}(A_x)dx \leq \mathcal{H}_{\alpha}(A).$$

□

There is generalization of the slicing theorem for more general codimension.

THEOREM 5.2. *Let $A \subset \mathbb{R}^n$ be a Borel set with dim$(A) > n-m$ and let E_m be an m-dimensional subspace. Then for almost every $x \in E^{\perp}$ (or equivalently, almost every $x \in \mathbb{R}^n$)*

$$dim(A \cap (E_m + x)) \leq dim(A) - (n-m),$$

and

$$\mathcal{H}^{\alpha}(A) \geq c_n \int_{E_m^{\perp}} \mathcal{H}^{\alpha-(n-m)}(A \cap (E_m + x))dx.$$

The proof is the same as before, so we omit it. Also as before, the equality can be strict. There are also generalizations which replace the "almost everywhere with respect to Lebesgue measure on "$\mathbb{R}^n$" with a more general measure. This will be considered in Chapter 4 (Theorem **??**).

EXAMPLE 5.3. If $A = E_{1/3} \times E_{1/3}$, is the product of the middle thirds Cantor set with itself, we saw in Example 2.4 that $\dim(A) = \log_3 4$. The vertical slices of A are empty almost surely, so $\dim(A_x) = 0 < (\log_3 4) - 1$ almost surely. This gives strict inequality for almost all slices.

EXAMPLE 5.4. A set with more interesting slices is the Sierpinski gasket (see Example 2.4). We claim that almost every (Lebesgue) vertical slice G_x of G has dimension 1/2 (which is strictly smaller than $\dim(G) - 1 = \log_2 3 - 1 = .58496\ldots$).

PROOF. For each $x \in [0, 1]$, let

$$S(x) = \{n : x_n = 1\},$$

where $\{x_n\}$ is the binary expansion of x. Then its easy to see

$$G_x = A_{S(x)},$$

where A_S is the set defined in Example 2.2 following Billingsley's lemma. Since the law of large numbers says almost every x is normal, we deduce $\dim(G_x) = 1/2$ for almost every x. □

More generally, $\dim(G_x) = p$ if and only if $x \in \hat{A}_p$, the set discussed in Example 4.1. Thus

$$\dim(\{x : \dim(G_x) = \alpha\}) = h_2(\alpha).$$

EXAMPLE 5.5. Consider the self-affine sets of Example 2.5. More precisely, consider the McMullen set X obtained by taking

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

The intersection of X with any vertical line is a single point. On the other hand the intersection of the set with a horizontal line of height y is a set described as follows: at the nth generation replace each interval by its middle third if $y_n = 1$ but remove the middle third if $y_n = 0$. The dimension of such a set is

$$(\log_3 2) \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} (1 - y_n).$$

Since for almost every $y \in [0, 1]$ the limit exists and equals $\frac{1}{2}$, we see that almost every horizontal cross section of the McMullen set has dimension $\frac{1}{2}\log_3 2$. The dimension of the set of y's where the horizontal cross section is exactly α can also be worked out as above.

It is interesting to note that we can compute the dimension of the cross sections, but that we have not yet computed the dimension of the set itself. By Marstrand's projection theorem we can deduce that its dimension is at least $1 + \frac{1}{2}\log_3 2 \approx 1.3146$ (we shall use the law of large numbers in Chapter 5 to prove that its dimension is exactly $\log_2(1 + 2^{\log_3 2}) \approx 1.34968$).

6. Intersecting translates of Cantor sets

In the previous section we considered intersecting a set in the plane with random lines. In this section we will intersect a set in the line with random translates of itself. It turns out that this is easiest in the case when the set has a particular arithmetic structure, such as the middle thirds Cantor set. We include the discussion here, because this special cases behaves very much like the "digit restriction sets" considered earlier. This section also serves to illustrate a principle that will be much more apparent later in the book: it is often easier to compute the expected dimension of a random family of sets than to compute it for any particular member of the family.

In 1973 J. Hawkes proved the following result:

THEOREM 6.1. *Let $\boldsymbol{C}$ be the classical ternary Cantor set. Then*

$$dim[\boldsymbol{C} \cap \boldsymbol{C} + t)] = \frac{1}{3}\frac{\log 2}{\log 3}$$

for Lebesgue–a.e. $t \in [-1, 1]$.

It is natural to expect that the method used to prove this theorem should extend to other Cantor sets defined by digit restrictions, but this extension, given in Kenyon and Peres (1991), turns out to depend on the theory of random matrix products, which are beyond the scope of this volume. The dimensions which arise in general also seem different, as they do not appear to be ratios of logarithms of rational numbers. For example,

EXAMPLE 6.2. If $\Lambda = \{\sum_1^\infty d_n 4^{-n} | d_n \in \{0, 1, 2\}\}$ then for a.e. $t \in [-1, 1]$,

(6.1)
$$\dim[(\Lambda + t) \cap \Lambda] = \frac{1}{6}\log_4 \frac{2}{3} + \sum_{k=0}^{\infty} 4^{-k-1} \log_4 \frac{(3 \cdot 2^k)!}{(2^{k+1})!} \approx 0.575228.$$

Next we discuss the cases to which (essentially) Hawkes' method extends. We start by sketching a simplified proof of Hawkes' theorem (Proposition 6.1). Hawkes' original proof is more elaborate, as he proves a result about general Hausdorff gauge functions and derives Proposition 6.1 as a corollary.

Proof of Proposition 6.1 (Sketch): The key observation is that the digits $\{-2, 0, 2\}$ can be used to represent any $t \in [-1, 1]$ in base 3, and the representation is unique if t is not a ternary rational. Fix t which is not a ternary rational, and let $\boldsymbol{C}$denote the classical Cantor set. Represent

$$t = \sum_{n=1}^{\infty} t_n 3^{-n},$$

with $t_n \in \{-2, 0, 2\}$. We have

$$\begin{aligned}(\boldsymbol{C}+t) \quad \cap \quad \boldsymbol{C} &= \{y \in \boldsymbol{C} | \exists x \in \boldsymbol{C}, y - x = t\} \\ &= \left\{ \sum_1^{\infty} y_n 3^{-n} | \exists x = \sum_1^{\infty} x_n 3^{-n}, \text{ with } y_n, x_n \in \{0,2\}, y_n - x_n = t_n \right\} \\ &= \left\{ \sum_1^{\infty} y_n 3^{-n} | y_n \in \{0,2\}, y_n = 2 \text{ if } t_n = 2, \text{ and } y_n = 0 \text{ if } t_n = -2 \right\}.\end{aligned}$$

Thus for $y \in (\boldsymbol{C}+t) \cap \boldsymbol{C}$ the n'th ternary digit is determined unless $t_n = 0$, in which case it has two possibilities. For random $t \in [-1, 1]$ with respect to Lebesgue measure, $\{n | t_n = 0\}$ has density $1/3$, and so we expect to need $2^{n/3}$ intervals of size 3^{-n} to cover the set. This gives

$$\dim_M[(\boldsymbol{C}+t) \cap \boldsymbol{C}] = \frac{\log 2^{\frac{1}{3}}}{\log 3} = \frac{1}{3} \log_3 2,$$

for such t. Extending the result to Hausdorff dimension is now easy; a rigorous discussion is contained in the proof of the next theorem.

Definition 6.3. Fix an integer $b > 1$. For any finite set of integers D set

$$\Lambda(D, b) = \left\{ \sum_{n=1}^{\infty} d_n b^{-n} | d_n \in D \right\}.$$

Say that two Cantor sets $\Lambda_1, \Lambda_2 \subset \mathbf{R}$ satisfy the *difference set condition* if $\Lambda_i = \Lambda(D, b)$ for $i = 1, 2$ and the difference set $D_2 - D_1 \subset \mathbf{Z}$ is contained in an arithmetic progression of length b.

Theorem 6.4. *Let $\Lambda_1 = \Lambda(D_1, b)$ and $\Lambda_2 = \Lambda(D_2, b)$ be two Cantor sets, satisfying the difference set condition. For each $i \in D = D_2 - D_1$, let $M_i = |(D_1 + i) \cap D_2|$.*

(i) *For all $t \in \mathbf{R}$, $dim[(\Lambda_1 + t) \cap \Lambda_2] = dim_{\mathrm{M}}[(\Lambda_1 + t) \cap \Lambda_2]$.*

(ii) *If $|D| = b$ then*

$$dim[(\Lambda_1 + t) \cap \Lambda_2] = \frac{1}{b} \sum_{i \in D} \log_b M_i,$$

for almost all t in the interval $\Lambda(D, b) = \Lambda_2 - \Lambda_1$.

In the examples below, we show how Theorem 6.4 implies Hawkes' theorem, as well as Theorem 6.8.

The proof of Theorem 6.4 depends on the following lemma.

LEMMA 6.5. *Assume* $\Lambda_1 = \Lambda(D_1, b)$ *and* $\Lambda_2 = \Lambda(D_2, b)$ *satisfy the difference-set condition. Denote* $D = D_2 - D_1$ *and for each* $i \in D$, $M_i = |(D_1 + i) \cap D_2|$. *Let* $t \in \Lambda(D, b)$. *If* t *has a unique representation* $t = \sum_{n=1}^{\infty} t_n b^{-n}$ *with* $t_n \in D$, *define*

$$\varphi_1(t) = \liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \log_b M_{t_n}.$$

If t *has two such representations, or* $t \notin \Lambda(D, b)$, *define* $\varphi_1(t) = 0$. *Then*

$$dim[(\Lambda_1 + t) \cap \Lambda_2] = dim_{\mathrm{M}}[(\Lambda_1 + t) \cap \Lambda_2] = \varphi_1(t), \tag{6.2}$$

for all $t \in \mathbf{R}$.

REMARK . Note the representation in base b with digits from D "behaves" like the standard representation in base b: We claim that the representation is unique except at countably many points. This follows from the difference-set condition which guarantees that for some a_0, $d_0 \in \mathbf{Z}$ the inclusion

$$D \subset \{a_0 + j d_0 | 0 \leq j < b\}$$

holds. Let $s, t \in \mathbf{R}$ with $t = a_0/(b-1) + s d_0$. Then $s \in \Lambda(\{0, 1, \ldots, b-1\}, b)$ if and only if $t \in \Lambda(D, b)$, and this equivalence establishes the claim above.

Proof of Lemma 6.5: Assume first that t has a unique representation

$$t = \sum_{n=1}^{\infty} t_n b^{-n}$$

with $t_n \in D$. Then

$$(\Lambda_1 + t) \cap \Lambda_2 = \left\{ \sum_{n=1}^{\infty} d_n b^{-n} | d_n \in (D_1 + t_n) \cap D_2 \text{ for all } n \right\}$$

and it follows immediately from the definition of Minkowski dimension and φ_1 that

$$\dim_{\mathrm{M}}[(\Lambda_1 + t) \cap \Lambda_2] = \varphi_1(t).$$

To compute the Hausdorff dimension, define a probability measure μ_t supported on $(\Lambda_1 + t) \cap \Lambda_2$ as follows. For any sequence $d_1^*, d_2^*, \ldots, d_N^*$ with $d_j^* \in (D_1 + t_j) \cap D_2$ let

$$\mu_t \left\{ \sum_{j=1}^{N} d_j^* b^{-j} + \sum_{j=N+1}^{\infty} d_j b^{-j} | d_j \in (D_1 + t_j) \cap D_2 \text{ for all } j > N \right\} = \prod_{j=1}^{N} M_{t_j}^{-1}.$$

Fix $y \in (\Lambda_1+t)\cap\Lambda_2$, and let $C_N(y)$ denote the subset of $(\Lambda_1+t)\cap\Lambda_2$ consisting of points for which the first N digits in the Λ_2 representation agree with the first N digits of y. Then

$$\liminf_{N\to\infty} \frac{-\log \mu_t[C_n(y)]}{N\log b} = \varphi(t)$$

so by Billingsley's lemma, $\dim[(\Lambda_1+t)\cap\Lambda_2] = \varphi_1(t)$.

Similar considerations appear in Cajar's book [Caj][**?**]

If $t \in \Lambda(D,b)$ has two distinct representations

$$t = \sum_{n=1}^{\infty} t_n b^{-n} = \sum_{n=1}^{\infty} t'_n b^{-n}$$

then by the remark above for some $n_0 \in \mathbf{Z}^+$, for all $n \geq n_0$ we have $t_n = \max D$ and $t'_n = \min D$ (or vice versa). It follows that for such t, the set $(\Lambda_1+t)\cap\Lambda_2$ is finite.

Finally, for $t \notin \Lambda(D,b)$ the set $(\Lambda_1+t)\cap\Lambda_2$ is empty. $\square$

Proof of Theorem 6.4 Statement (i) is contained in Lemma 6.5. Next we prove (ii). By the remark above, when representing numbers in $\Lambda(D,b)$, the digits are independent symmetric random variables with respect to normalized Lebesgue measure on the closed interval $\Lambda(D,b)$. Thus, by the strong law of large numbers,

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \log_b M_{t_n} = \frac{\int_{\Lambda(D,b)} \log_b(M_{t_1})dt}{\int_{\Lambda(D,b)}} = \frac{1}{b}\sum_{i\in D} \log_b M_i$$

for a.e. $t = \sum_{n=1}^{\infty} t_n b^{-n} \in \Lambda(D,b)$. Lemma 6.5 now gives the desired conclusion.

Example 6.6. If $b=3$ and $\Lambda_1,\Lambda_2 = C$, the classical Cantor set, then $D_1 = D_2 = \{0,2\}$, the difference set is $D = \{-2,0,2\}$ and $\{M_j\}_{j\in D} = \{1,2,1\}$. Thus Theorem 6.1 is contained in Theorem 6.4.

Example 6.7. Extending the previous example, assume D_1, D_2 are arithmetic progressions with the same difference, of length n_1, n_2 respectively, where $n_1+n_2 = b+1$. Without loss of generality, $n_2 \leq n_1$,

$$D_1 = \{1,\dots,n_1\} \text{ and } D_2 = \{1,\dots,n_2\}.$$

Then $D = D_2 - D_1 = \{1-n_1, 2-n_1, \dots, n_2-1\}$ and

$$M_j = \begin{cases} j+n_1 & -n_1 < j \leq n_2 - n_1, \\ n_2 & n_2-n_1 \leq j \leq 0, \\ n_2 - j & 0 \leq j < n_2. \end{cases}$$

Thus Theorem 6.4 (ii) shows that for *a.e.* $t \in \Lambda(D, b)$

$$\dim[(\Lambda_1 + t) \cap \Lambda_2] = \frac{1}{b} \log_b[(n_2!) n_2^{b-2n_2}].$$

Theorem 6.4 gives the dimension of intersections for typical translates; in the difference-set case, it is possible to go further and compute certain iterated dimensions, just as we did for slices of the Sierpinski gasket. Such iterated dimensions are closely related to the popular "multi-fractal analysis".

PROPOSITION 6.8 (Kenyon and Peres (1994)).

$$dim\{t | dim[(\boldsymbol{C} + t) \cap \boldsymbol{C}] = \alpha \log_3 2\} = h_3\left(\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}\right),$$

where $\boldsymbol{C}$ *is the ternary Cantor set,* $0 < \alpha < 1$ *and*

$$h_3(p_1, p_2, p_3) = -\sum_{i=1}^{3} p_i \log_3 p_i$$

is the ternary entropy function.

PROOF. From the considerations in the proofs of Lemma 6.5 and Theorem 6.4 we see that we want to find the dimension of the set of t's such that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \log_b M_{t_n} = \alpha \log_3 2,$$

where $\{t_n\}$ is the ternary expansion of t. Equivalently, (since $M_{t_n} = 2$ if $t_n = 0$ and 1 otherwise) we want the set of t's so that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} T_n = \alpha,$$

where $T_n = 1$ if $t_n = 0$ and equals 0 otherwise. Thus Lemma 4.6 gives the dimension as

$$\alpha \log_3 \alpha + (1-\alpha) \log_3 \frac{1}{2}(1-\alpha) = h_3\left(\frac{1-\alpha}{2}, \alpha, \frac{1-\alpha}{2}\right),$$

as desired. This proves Proposition 6.8. $\square$

Exercises

EXERCISE 1. Let $E = \{0\} \cup \{n^{-\alpha}\}_{n=1}^{\infty}$. Find $\dim_M(E)$.

EXERCISE 2. Let $S = \cup_{n=1}^{\infty}[(2n), (2n+1)] \subset \mathbb{N}$. Show that $\underline{\mathrm{d}}(S) = 0$ and $\overline{\mathrm{d}}(S) = 1$.

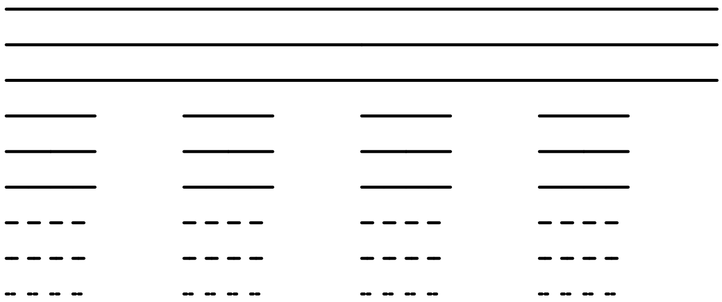

FIGURE 6.1. Set in Exercise 10

EXERCISE 3. Construct a set K so that

$$\dim(K) < \underline{\dim_M}(K) < \overline{\dim_M}(K).$$

EXERCISE 4. Construct a countable closed set so that

$$\underline{\dim_M}(K) < \overline{\dim_M}(K).$$

EXERCISE 5. Suppose $E_1 \subset E_2 \subset \dots.$ prove that

$$\dim(\cup_n E_n) = \lim_{n\to\infty} \dim(E_n).$$

Show this is false for Minkowski dimension (even for upper Minkowski dimension).

EXERCISE 6. If $E_1 \supset E_2 \supset \dots$ is it true that

$$\dim(\cap_n E_n) = \lim_{n\to\infty} \dim(E_n)?$$

EXERCISE 7. Suppose $E \subset \mathbb{R}^n$ and $P : \mathbb{R}^n toV$ is orthogoonal projection onto a subspace V. Prove that $\dim(P(E)) \leq \dim(E)$. More generally, prove this if $P : \mathbb{R}^n \to \mathbb{R}^n$ is any map satisfying the Lipschitz condition

$$|\, P(x) - P(y)\,| \leq A\,|\,x - y\,|,$$

for some $A < \infty$ and all $x, y \in \mathbb{R}^n$.

EXERCISE 8. Find the exact measure of the von Koch snowflake in dimension $\log 3/\log 2$.

EXERCISE 9. If $\{a_n\} > 0$ and $a_{n+m} \leq a_n + a_m$ then $\lim_{n\to\infty} \frac{a_n}{n} = \inf_n \frac{a_n}{n}$ exists. This was used to prove that the spectral radius of a matrix A is well defined in Section 2.

EXERCISE 10. Figure 6.1 shows approximations to a set of the form A_S for some sequence of integers S. Determine as much of S as possible.

FIGURE 6.2. Set in Exercise 16

EXERCISE 11. Suppose a finite graph is represented by the $0-1$ matrix A (with a 1 is position (i,j) if i and j are joined by an edge). Show the number of paths of length n from i to j is the $(A^n)_{ij}$, (i.e., the (ith, jth) entry in the matrix A^n).

EXERCISE 12. Show the spectral radius of $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ is $(1+\sqrt{5})/2$.

Hint: Show A^n has the form $\begin{pmatrix} F_{n+1} & F_{n-1} \\ F_{n-1} & F_n \end{pmatrix}$ where F_n is the nth Fibbonaci number.

EXERCISE 13. Compute the dimension of the set of numbers x so that 2 is isolated in the base 3 expansion of x, (e.g. $x = .33212312111...$ is allowed, but $x = .311212231131$ is not).

EXERCISE 14. Compute $\dim(X_A)$ for $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

EXERCISE 15. Compute $\dim(X_A)$ for

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

EXERCISE 16. Figure 6.2 shows the first few approximations to a linear Cantor set of finite type, corresponding to a 3×3 matrix. What is the matrix? What is the dimension of the set?

EXERCISE 17. Create a set by starting with the line segment $[0,1]$ and at each stage replacing each interval I =[x,y] by $[x,z] \cup [z,w] \cup [w,z] \cup [z,y]$ where $z = \frac{1}{2}(x+y)$ and $w = z + i\beta(y-x)$. What is the dimension of the resulting set? See Figure 6.3. This set is sometimes called the "antenna set".

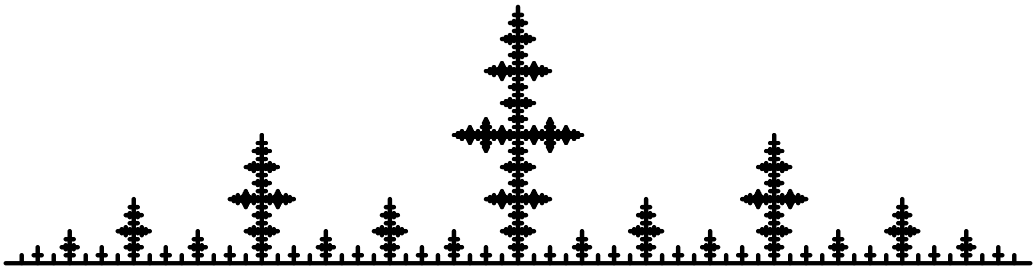

FIGURE 6.3. Set in Exercise 17

EXERCISE 18. In the previous exercise, show the set is a solid triangle when $\beta = 1/2$. Use this to construct a continuous function f on the interval which covers an open set in the plane.

EXERCISE 19. What is the dimension of almost every vertical slice of the antenna set in Exercise 17?

EXERCISE 20. Show the two definitions of $\dim(\mu)$ in Section 3 are the same.

EXERCISE 21. Show that the following two constructions define the Sierpinski Gasket G, from Example 2.4

1. $G = \{\sum_{n=1}^{\infty} a_n 2^{-n} : a_n \in \{0, 1, i\}\}$.
2. Let G_0 be the solid triangle with vertices $\{0, 1, i\}$. Find the midpoints of each edge of G_0 and remove from G_0 the triangle with these as vertices. What remains is the union of three triangles and we denote it G_1. In general, G_n is the union of 3^n right triangles with legs of length 2^{-n}. For each triangle we find the midpoints of each edge and remove the corresponding triangle. This gives a nested sequence of sets the limiting set is G.

EXERCISE 22. Suppose $S \subset \mathbb{N}$, and we are given $E, F \subset \{0, 1, 2\}$. Define $B_S = \{x = \sum_{k=1}^{\infty} x_k 2^{-k}\}$ where

$$x_k \in \begin{cases} E, & k \in S \\ F, & k \notin S \end{cases}.$$

Find $\dim(B_S)$ in terms of E, F and S.

EXERCISE 23. Consider the sets in Example 2.2. What conditions on S_1, S_2 insure that $A_{S_1} + A_{S_2} = [0, 2]$?

EXERCISE 24. When does A_S (Example 2.2) have non-zero measure in its dimension? Finite measure?

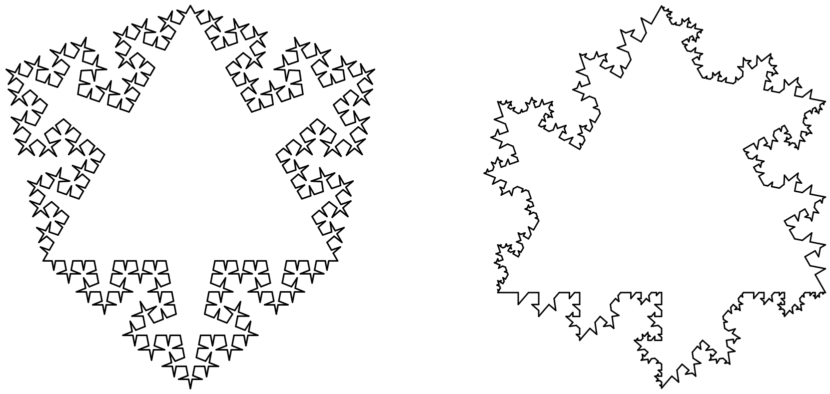

FIGURE 6.4. Curves for Exercise 25

EXERCISE 25. Estimate the dimension of the curves shown in Figure 6.4 assuming they are constructed like the von Koch snowflake by replacing intervals by the same polygonal arc each time. (You may need a ruler to do this one.)

EXERCISE 26. Characterize which directed graphs correspond to sets of dimension 0 in Example 2.3 (shifts of finite type).

EXERCISE 27. Construct a directed graph so that the number of paths of length n is $\sim n^2$.

EXERCISE 28. Given two compact sets A, B, we define the Hausdorff distance between them as

$$d(A,B) = \max_{a\in A} \operatorname{dist}(a,B) + \max_{b\in B} \operatorname{dist}(b,A).$$

Show that if X is compact then $K(X)$, the collection of compact subsets of X, is compact with this metric. (This is called the Borel selection theorem.)

EXERCISE 29. If $A_n \to A$ in the Hausdorff metric, does $\dim(A_n) \to \dim(A)$?

EXERCISE 30. A set is called first category if it is a finite union of nowhere dense sets. The complement of a first category set is called residual. Prove A_p is first category.

EXERCISE 31. If $p < 1/2$ what can be said of $\hat{A}_p + \hat{A}_p$?

EXERCISE 32. Show that if A is residual, then $A + A$ contains an interval.

EXERCISE 33. Construct a residual set of dimension 0.

EXERCISE 34. Compute the Minkowski dimension of the McMullen set.

EXERCISE 35. Construct a Cantor set which is disjoint from all of its non-zero rational translates. (By a Cantor set we mean a set which is perfect, totally disconnected and compact.)

EXERCISE 36. In Example 2.4, which matrices give connected sets?

EXERCISE 37. Let $E, F \subset \{0, 1, \dots, b-1\}$ and let $\{x_n\}$ be the b-ary expansion of $x \in [0,1]$. Let $S \subset \mathbb{N}$. What is the dimension of

$$A_{E,F,S} = \{\sum_{n=1}^{\infty} x_n b^{-n} : x_n \in \begin{cases} E, & n \in S \\ F, & n \notin S \end{cases}.$$

EXERCISE 38. Given $\{s_n\} \in \{0, 1, \dots, b-1\}^{\mathbb{N}}$ and sets $E_i \subset \{0, 1, \dots b-1\}$, $0 \le i \le b-1$, find the dimension of the set

$$\{\sum_{n=1}^{\infty} x_n b^{-n} : x_n \in E_{s_n}\}.$$

EXERCISE 39. Let C be the Sierpinski carpet (see Example 2.4). Find $\dim(\{x : \dim(C_x) = \alpha\})$.

EXERCISE 40. Let G be the Sierpinski gasket. Let L_c be the line in the plane $\{(x,y) : x + y = c\}$. For almost every $c \in [0,1]$ find $\dim(L_c \cap G)$.

EXERCISE 41. Let S denote the McMullen set of Example 2.5 and let S_y denote the intersection of this set with the horizontal line at height y. Prove

$$\dim(\{y : \dim(S_y) = \alpha\}) = H_2(\frac{\alpha}{\log_3 2}),$$

where H_2 is the entropy function of Section 4 .

EXERCISE 42. Let $X = E_{1/4} \times E_{1/4}$ and let L_c be as in previous exercise. Find $\dim(c : \dim(L_c \cap X) = \alpha)$.

EXERCISE 43. Let t_n denote the ternary expansion of $t \in [0,1]$. For $0 \le p \le 2$, what is the dimension of the set of t's such that

$$\lim_{N\to\infty} \sum_{n=1}^{\infty} t_n = p?$$

EXERCISE 44. Let $\{x_n\}$ be the b-ary expansion of x and suppose f is a function on $\{0, 1 \ldots, b-1\}$ to the reals. Show

$$\dim(\{x : \lim_{N\to\infty} \sum_{n=1}^{N} f(x_n) = \alpha\}) = \max_{\mathbf{p}:\sum_j p_i f(i)=\alpha} h_b(\mathbf{p}).$$

See [**Egg49**] and [**Col70**].

EXERCISE 45. Given real number x let

$$\begin{aligned} \mathbf{p}_N(x) &= \{p_0^n(x), \ldots, p_{b-1}^N(x)\} \\ &= \frac{1}{N}(\#\{n \le N : x_n = 0\}, \ldots, \#\{n \le N : x_n = b-1\}). \end{aligned}$$

Given a closed connected set E of probability vectors show that the dimension of the set of x's such that $\{\mathbf{p}_N(x)\}$ accumulates only on E is

$$\dim(X) = \max_{\mathbf{p}\in E} h_b(p).$$

EXERCISE 46. Given a string of length N on the alphabet $\{0, \ldots, b-1\}$ let n_i be the number of elements in the string equal to i. Show that the number of strings with $P_N \in U$ is at most

$$n^g \frac{n}{n_0 \ldots n_{b-1}}.$$

Use Stirling's formula to show this is less than

$$g^{n(h_b(p^*)+\epsilon+o(1))}.$$

EXERCISE 47. Let $\mathbf{p}_N(x)$ be defined as in the previous exercise. Let $V(x)$ be the accumulation set of the sequence $\{\mathbf{p}_N(x)\}$. Given a subset V_0 of the space of probability vectors show

$$\dim(\{x : V(x) \subset V_0\}) = \max_{p\in V_0} h_b(p).$$

This is from [**Vol58**]. (Hint: The $\ge$ direction is easy. To prove the other direction, let p^* be the point of V_0 where the max is attained and let U be the set of probability vectors with $h_b(p) \le h_b(p^*) + \epsilon$. Now use the previous exercise.

Volkmann's paper also contains the following formulas

1. $\dim(\{x : V(x) = V_0\}) = \min_{p\in V_0} h_b(p)$.
2. $\dim(\{x : V(x) \supset V_0\}) \min_{p\in V_0^*} h_b(p)$.
3. $\dim(\{x : V_0 \subset V(x) \subset V_1\}) = \min_{p\in V_0^*} h_b(p)$,

where V_0^* is the intersection of all continua that contain V_0.)

EXERCISE 48. Let C be the middle thirds Cantor set. What is

$$\dim(C \cap (C+t) \cap (C+s)),$$

for almost every $(s,t) \in [0,1]^2$?

EXERCISE 49. Consider the self-affine type set X_A corresponding to the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

What is $\dim(X_A \cap (X_A + t))$ for almost every $t \in [0,1]^2$?

EXERCISE 50. Find $\dim(\{x \in [0,1] : \sin(2^n x) \to 0\})$.

EXERCISE 51. Suppose $S \subset \mathbb{N}$ is a set of density α. Let $x \in [0,1]$ be the collection of numbers whoose binary expansions satisfy

$$\frac{1}{|[0,n] \cap S|} \sum_{n \in S} x_n = \beta,$$

$$\frac{1}{|[0,n] \setminus S|} \sum_{n \notin S} x_n = \gamma.$$

What is $\dim(E)$ in terms of α, β and γ?

Bibliography

[BP94] I. Benjamini and Y. Peres. Tree-indexed random walks on groups and first passage percolation. *Probab. Th. rel. Fields*, 98:91–112, 1994.

[BPP94] I. Benjamini, R. Pemantle, and Y. Peres. Martin capacity and Markov chains. pages –, 1994. preprint.

[BS93] C.J. Bishop and T. Steger. Representation theoretic rigidity in PSL$(2, R)$. *Acta Math.*, 170:121–149, 1993.

[BT54] A.S. Besicovitch and S.J. Taylor. On the complementary intervals of closed sets of zero Lebesgue measure. *J. London Math. Soc.*, 29:449–459, 1954.

[BT92] M. Barlow and S.J. Taylor. Defining fractal subsets of $\mathbb{Z}^d$. *Proc. London Math. Soc.*, 64:125–152, 1992.

[CCD88] J Chayes, L. Chayes, and R. Durrett. Connectivity properties of Mandelbrot's percolation process. *Probab. Thy Rel. Fields*, 78:307–324, 1988.

[Col70] C.M. Colebrook. Hausdorff dimension of certain sets of non-normal numbers. *Mich. Math. J.*, 17:103–116, 1970.

[DEK61] A. Dvoretsky, P. Erdös, and S. Kakutani. Nonincreasing everywhere of the Brownian motion process. In *Proceedings of the fourth Berkeley Symposium, vol. II*, pages 103–116, 1961.

[DG88] M. Dekking and G.R. Grimmett. Superbranching process and projections of random Cantor sets. *Probab. Thy. Rel. Fields*, 78:335–355, 1988.

[Egg49] H.G. Eggleston. The fractional dimesnion of a set defined by decical properties. *Quart. J. Math. Oxford*, 20:31–35, 1949.

[Erd61] P. Erdös. A problem about prime numbers and the random walk ii. *Illinois J. Math.*, 5:352–353, 1961.

[Eva92] S. Evans. Polar and nonpolar sets for a tree-indexed process. *Ann. Prob.*, 20:579–590, 1992.

[Fal88] K.J. Falconer. Projections of random Cantor sets. *J. Theoretical Prob.*, 2:65–70, 1988.

[Fre62] G. Freud. Über trigonometrische approximation und Fouriersche reihen. *Math. Z.*, 78:252–262, 1962.

[Fur81] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton University Press, 1981.

[GL92] D. Gatzouras and S. Lalley. Hausdorff and box dimensions of certain self-affine fractals. *Indiana Univ. Math. J.*, 41:533–568, 1992.

[GMW88] S. Graaf, R.D. Mauldin, and S.C. Williams. *The exact Hausdorff dimension in random recursive constructions.* American Math. Soc., 1988.

[Har16] G.H. Hardy. Weierstrass's nondifferentiable function. *Trans. Amer. Math. Soc.*, 17:301–325, 1916.

[Hoe63] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.

[Hou94] E.A. Housworth. Escape rate for 2-dimesnional Brownian motion conditioned to be transient and application to Zygmund functions. *Trans. Amer. Math. Soc.*, 343:843–, 1994.

[Hut81] J.E. Hutchinson. Fractals and self similarity. *Indiana Univ. Math. J.*, 30:713–747, 1981.

[IIK64] M. Izumi, S.I. Izumi, and J.-P. Kahane. Théorèmes élèmentaaires sur les séries de Fourier lacinaires. *J. d'Analyse Math.*, ??:235–246, 1964?

[J.H81] J.Hawkes. Trees generated by a simple branching process. *J. London Math. Soc.*, 24:373–384, 1981.

[Kah64] J.P. Kahane. Lacunary Taylor and Fourier series. *Bull. Amer. Math. Soc.*, 70:199–213, 1964.

[Kak44] S. Kakutani. Two dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo*, 21:706–714, 1944.

[Kau68] R. Kaufman. On Hausdorff dimension of projections. *Mathematika*, 15:153–155, 1968.

[Kho94] D. Khoshnevisan. A discrete fractal in $\mathbb{Z}_+^1$. *Proc. Amer. MAth. Soc.*, 120:577–584, 1994.

[KP76] J.P. Kahane and J. Peyiere. Sur cetaines martingales de B. mandlebrot. *Advances in Math.*, 22:131–145, 1976.

[KS64] S. Kochen and C. Stone. A note on the borel-cantelli lemma. *Illinois J. Math.*, 8:248–251, 1964.

[KWW63] J.-P. Kahane, G. Weiss, and M. Weiss. On lacunary power series. *Ark. Mat.*, 5:1–26, 1963.

[Lam63] J. Lamperti. Wiener's test and Markov chains. *J. math. Anal. Appl.*, 6:58–66, 1963.

[Law91] G. Lawler. *Intesections of random walks.* Birkhäuser, 1991.

[LP92] R. Lyons and R. Pemantle. Random walk in a random enviroment and first-passage percolation on trees. *Ann. Probab.*, 20:125–136, 1992.

[Lyo89] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Comm. Math. Phys.*, 125:337–353, 1989.

[Lyo90] R. Lyons. Random walks and percolation on trees. *Ann. Probab.*, 18:931–958, 1990.

[Lyo92] R. Lyons. Random walks, capacity and percolation on trees. *Ann. Probab.*, 20:2043–2088, 1992.

[Man74] B.B. Mandelbrot. Intermittent turbulence in self-similar cascades: diverence of high moments and and dimension of the carrier. *J. Fluid Mechanics*, 62:331–358, 1974.

[McK61] H.P. McKean. A problem about prime numbers and the random walk i. *Illinois J. Math.*, 5:351, 1961.

[McM84] C. McMullen. The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.*, 96:1–9, 1984.

[MW] R.D. Maudlin and S.C. Williams. Graphs of real Zygmund class functions have σ-finite measure. *Trans. Amer. Math. Soc.*, 1986(298):793–803.

[Per94a] Y. Peres. Intersection-equivalence of Brownian paths nd certain branching processes. pages –, 1994. Preprint.

[Per94b] Y. Peres. The packing measure of self-affine carpets. *Math. Proc. Camb. Phil. Soc.*, 115:437–450, 1994.

[PP94] R. Pemantle and Y. Peres. Domination between trees and appliation to an explosion problem. *Ann. Probab.*, pages –, 1994. To appear.

[Roh] S. Rohde. On conformal welding and quasicircles. *Michigan Math. J.*, pages –.

[RS82] I.Z. Ruzsa and G.J. Széely. Intersections of traces of random walks with fixed sets. *Ann. Prob.*, 10:132–136, 1982.

[Sel50] A. Selberg. The general sieve-method and its place in prime number theory. In *Proceedings of the International Congress of Mathematicians, Vol. I*, pages 286–292, 1950.

[Spi64] F. Spitzer. *Principles of random walk.* Van Nostrand, 1964.

[Ste71] E.M. Stein. *Singular integrals and differentiablility properties of functions.* Princeton University Press, 1971.

[SZ45] R. Salem and A. Zygmund. Launary power series and Peano curves. *Duke Math. J.*, 12:569–578, 1945.

[Tri81] C. Tricot. Douze definitions de la densité logarithmique. *Comptes Rendus Acad. Sci. Paris*, 293:549–552, 1981.

[Tuk89] P. Tukia. Hausdorff dimension and quasisymmetric mappings. *Math. Scand.*, 65:152–160, 1989.

[Urb90] M. Urbanski. The Hausdorff dimension of the graphs of continuous self-affine functions. *Proc. Amer. Math. Soc.*, 108:921–930, 1990.

[Vol58] B. Volkmann. Über Hausdorffsche dimensionen von mengen die durch zifferneigenschaften charakterisiert sind. vi. *Math. Z.*, 68:439–449, 1958.

[Wei72] K. Weierstrass. Über contiuirliche functionen eines reellen arguments, die für keinen werth des letzeren einen bestimmten differentialquotienten besitzen. *Königl. Akad. Wiss.*, 3:71–74, 1872. Mathematische Werke II.