Contents

CHAPTER 1

When do Minkowski and Hausdorff dimension agree?

We saw in the last section that the Minkowski dimension of a general set need not exist, and that when it does exist it need not agree with the Hausdorff dimension. In this chapter we will consider various conditions on a compact set K which insure that the Minkowski dimension of K exists and equals the Hausdorff dimension. We start with the theory of "self-similar"sets and then consider more general rescaling conditions.

1. Self-similar sets.

In this section we will consider the simplest examples of sets with fractional dimension, the so called "self-similar" sets. These are sets which are invariant under some finite set of similarities of $\mathbb{R}^n$. Our first observation is

LEMMA 1.1. *Suppose $S_1, \dots, S_r$ are contractions of a complete metric space X and suppose*

$$d(S_j(x), S_j(y)) \leq \lambda_j d(x, y),$$

for some $\lambda_j < 1$, $j = 1, \dots r$. Then there exists a unique compact set K such that

$$K = S(K) \equiv \cup_{j=1}^r S_j(K).$$

For example, the usual middle $\frac{1}{3}$ Cantor set $E \subset [0, 1]$ is the set associated to the two contractions $K_1(x) = x/3$ and $K_2(x) = \frac{2}{3} + x/3$.

[**Hut81**]. Let comp(X) denote the space of compact subsets of X with the Hausdorff metric

$$d_H(K_1, K_2) = \inf\{\epsilon : K_2 \subset K_1^{(\epsilon)}, K_1 \subset K_2^{(\epsilon)}\},$$

where $K^{(\epsilon)} = \{x : d(x, K) < \epsilon$ is the ϵ neighborhood of K. Then comp(X) is a complete metric space and

$$S(K) = \cup_{j=1}^r S_j(K),$$

defines a contraction map of comp(X) to itself, since

$$\begin{aligned} d_H(S(K_1), S(K_2)) &\leq \max_j d_H(S_j(K_1), S_j(K_2)) \\ &\leq (\max_j \lambda_j) d_H(K_1, K_2). \end{aligned}$$

Therefore by the Banach contraction principle, there is a unique fixed point of S, which is the desired set K. $\square$

The set K can be associated to a sequence space using the maps $\{S_j\}$. Consider the space $\{1, \ldots, r\}^{\mathbb{N}}$. given an element $\eta = \{i_1, i_2, i_3 \ldots\}$ of this space we have a nested sequence of compact sets,

$$K \supset S_{i_1}(K) \supset S_{i_1} S_{i_2}(K) \supset \ldots .$$

By the Cantor intersection theorem the intersection is nonempty and since the maps are contractions, the intersection must be a single point, which we call $\Psi(\eta)$. Thus $\Psi : \{1, \ldots r\}^{\mathbb{N}} \to K$ defines a map. It is easy to check that it is surjective, though it need not be injective (consider constructing the snowflake in this way). Give $\{1, \ldots, r\}^{\mathbb{N}}$ the product topology and the metric

$$d(\eta, \zeta) = e^{-|\eta \wedge \zeta|},$$

where $|\eta \wedge \zeta|$ denotes the length of the longest initial segment on which the two sequences agree. The mapping Ψ is continuous, indeed, it is Hölder continuous,

$$d(\Psi(\eta), \Psi(\zeta)) \leq |K| (\lambda_{\max})^{|\eta \wedge \zeta|} \leq |K| d(\eta, \zeta)^{\log 1/(\lambda_{\max})}.$$

Surjectivity can be derived from the fact that $S(\text{Image}(\Psi)) = \text{Image}(\Psi)$ which is a compact set. Indeed, if $\tilde{S}_j$ is the map from $\{1, \ldots r\}^{\mathbb{N}}$ defined by a right shift and inserting j in the first place, then $\Psi \circ \tilde{S}_j = S_j \circ \Psi$.

For any probability vector $(p_1, \ldots, p_r)$ there is a natural product measure on the sequence space. Consider $(p_1, \ldots, p_r) = (\lambda_1^\alpha, \ldots, \lambda_r^\alpha)$ and the corresponding product measure $\tilde{\mu}$. The mapping Ψ allows us to map $\tilde{\mu}$ to a measure μ on K, by $\mu(A) = \tilde{\mu}(\Psi^{-1}(A))$, for $A \subset K$ Borel. The measure μ can also be constructed as follows. It is defined as a solution of the integral equation

$$\int \varphi d\mu = \sum_{j=1}^{r} \lambda_j^\alpha \int \varphi \circ S_j d\mu$$

for every bounded measurable function $\varphi : K \to \mathbb{R}$. To show that there is a unique probability measure on K satisfying this equation we define a distance between measures on K by

$$D(\mu,\nu)=\sum_{\varphi}|\int\varphi d\mu-\int\varphi d\nu|$$

where the supremum is taken over all Lipschitz functions on K, i.e., all $\varphi: K\to\mathbb{R}$ such that $|\varphi(x)-\varphi(y)|\leq d(x,y)$ for all pairs $x,y\in K$. For example, if δ_x denotes the Dirac mass at the point x then

$$D(\delta_x,\delta_y)=d(x,y).$$

Now consider the operator on the $\Pr(K)$, the space of probability measures on K given by:

$$A\mu=\sum_{j=1}^{r}\lambda_j^{\alpha}\mu\circ S_j^{-1}.$$

Since for any two measures μ_1,μ_2 we can easily verify

$$D(\mu_i\circ S_j^{-1},\mu_2\circ S_j^{-1})\leq\lambda_j D(\mu_1,\mu_2),$$

we see that A is a contraction map with coefficient $\max_j\lambda_j$. The space $\Pr(K)$ of probability measures on K is a compact and hence complete, so by the Banach principle there is a unique fixed point of A, i.e., a unique probability measure μ on K so that $A\mu=\mu$. Moreover, this measure can be obtained as the limit $\lim_{n\to\infty}A^n\mu_0$ for any choice of $\mu_0\in\Pr(K)$.

If one considers the Markov process with state space K in which a point x maps to $S_j(x)$ with probability λ_j^{α} then this solution μ will be the limiting distribution of the process.

DEFINITION 1.2. We say $K\subset\mathbb{R}^d$ is self-similar if it is the attractor of a collection of maps $\{S_1,\dots,S_r\}$ which are Euclidean similarities, i.e., $|S_ix-S_iy|=\lambda_i|x-y|$, $i=1,\dots,r$.

LEMMA 1.3. *For a self-similar set* $\mathcal{H}^{\alpha}(K)=\mathcal{H}^{\alpha}_{\infty}(K)$.

PROOF. Since for any set we have $\mathcal{H}^{\alpha}(K)\geq\mathcal{H}^{\alpha}_{\infty}(K)$ we only have to show the opposite inequality, i.e, given any finite cover of K by open balls we have to find a covering by small balls with at most the same sum. To see this suppose $\sigma=\{i_1,\dots,i_n\}$ is a sequence of length $|\sigma|=n$ and let $\lambda_\sigma=\{\lambda_{i_1},\dots,\lambda_{i_n}\}$. Then the collection of sets $\{S_\sigma(K) : |\sigma| = n\}$ is a covering of K,

$$\begin{aligned}\sum_{|\sigma|=n} |S_\sigma(K)|^\alpha &= \sum_{|\sigma|=n} |K|^\alpha \lambda_\sigma^\alpha \\ &= |K|^\alpha (\sum_{i=1}^r \lambda_i^\alpha)^n \\ &= |K|^\alpha\end{aligned}$$

Thus if $\mathcal{B} = \{B_j\}$ is a finite covering of K by open balls then $\{S_\sigma(B) : |\sigma| = n, B \in \mathcal{B}\}$ is a covering of K by small balls whose sum is the same as the original covering. □

We are now ready to show that for self-similar sets the Hausdorff dimension is what we think it should be; the same as the similarity dimension.

DEFINITION 1.4. The set K and maps $\{S_i\}$ are said to satisfy the open set condition if there is a bounded open set V such that $S_i(V) \subset V$, $i = 1, \dots r$ and the $\{S_1(V), \dots, S_r(V)\}$ are disjoint.

Note that we do not require that V contains K, but it will follow from the proof below that $K \subset \overline{V}$.

THEOREM 1.5. *If $K \subset \mathbb{R}^d$ is self-similar, define α by the equation $\sum \lambda_i^\alpha = 1$ (α is called the "similarity dimension"). Assume the open set condition. Then*

$$dim(K) = dim_M(K) = \alpha$$

and

$$0 < \mathcal{H}^\alpha(K) = \mathcal{H}^\alpha_\infty(K) < \infty.$$

PROOF. We start with a lemma.

LEMMA 1.6. *Let $\{W_j\}$ be disjoint sets such that $B_{ar} \subset W_j \subset B_{br}$, where B_s denotes some ball of radius s. Then any B_r ball of radius r can intersect at most $(1 + 2b)^d a^{-d}$ of the W_j.*

PROOF. Let B_r intersect q of the W_j. All of them are contained in the ball of radius $(1+2b)r$ concentric with B_r. But the volume of W_j is at least the volume of B_{ar} so their disjointness implies $q \leq (1+2b)^d a^{-d}$. □

We now return to the proof of the theorem. We use the mass distribution principle Lemma **??**. Consider the probability measure μ on K constructed above and let V be a bounded open set as described in the open set condition. By definition $S(V) \subset V$, so for the closure of V we have

$$\overline{V} \supset S(\overline{V}) \supset \dots K,$$

and hence $K \subset \overline{V}$. Denote by Π_ρ the set of all finite sequences $\sigma = \{i_1, \dots, i_s\}$ such that

$$\lambda_\sigma = \prod_{j=1}^{s} \lambda_{i_j} \le \rho,$$

but $\lambda_{\tilde\sigma} > \rho$ where $\tilde\sigma = \{i_1, \dots i_{s-1}\}$ denotes σ with the last element removed. Here $\rho < 1$ is a real number.

The space of finite sequences can be considered as a tree with the empty sequence as root with an edge connecting each pair $\{\sigma, \tilde\sigma\}$. The space of infinite sequences X corresponds to the boundary of this tree in a natural way. For each σ define $F_\sigma = \{\tau = (i_1, i_2 \dots) \in X : (i_1, \dots, i_{|\sigma|}) = \sigma\}$, i.e., F_σ is just the collection of all infinite sequences which begin with σ.

The set of finite sequences Π_ρ define a "cut" of this tree, i.e., a connected subset containing the root. The collection of sets F_σ where $\sigma \in \Pi_\rho$ is a finite cover of X. For any $\sigma \in \Pi_\rho$,

$$\tilde\mu(F_\sigma) = \lambda_\sigma^\alpha \le \rho^\alpha.$$

Now apply the lemma to the collection $W_\sigma = \{S_\sigma V\}$. By the open set condition, for every pair of distinct sequences $\sigma_1 \ne \sigma_2$ we have $W_{\sigma_1} \cap W_{\sigma_2} = \emptyset$. Furthermore, if we choose a and b so that $B_a \subset V \subset B_b$ (which we can do since V is open and bounded), then our sets satisfy

$$B_{a\lambda_\sigma} \subset W_\sigma \subset B_{b\lambda_\sigma}.$$

Therefore the lemma applies and any ball B_ρ of radius ρ intersects at most $C(1+2b)^d(a\lambda_0)^{-d}$ of the sets where $\lambda_0 = \min_j \lambda_j$. Thus for any ball B_ρ of radius ρ we have

$$\begin{aligned} \mu(B_\rho) &\le \sum_{\sigma \in \Pi_\rho} \mu(B_\rho \cap W_\sigma) \\ &\le C \max_{\sigma \in \Pi_\rho} \mu(W_\sigma) \\ &\le C \max_{\sigma \in \Pi_\rho} \tilde\mu(F_\sigma) \\ &\le C\rho^\alpha. \end{aligned}$$

So $\mathcal{H}^\alpha(k) > 0$. Since we already know $\mathcal{H}^\alpha(K) = \mathcal{H}^\alpha_\infty(K)$ and the latter is finite (since K is bounded), we have proven the theorem.

Next we will compute the Minkowski dimension of the self-similar set K. As in the previous proof we use the cover Π_ρ of X. Every set in this collection has diameter less than ρ, so by expanding each to a ball of radius ρ we see that

$$N(\rho|K|, K) \leq \#\Pi_\rho.$$

Furthermore,

$$1 = \sum_{\sigma \in \Pi_\rho} \lambda_\sigma^\alpha \geq (\lambda_0 \rho)^\alpha \#\Pi_\rho.$$

Therefore $N(\rho|K|, K) \leq (\lambda_0 \rho)^{-\alpha}$. But

$$\overline{\dim_M}(K) = \limsup_{\rho \to 0} \frac{\log N(\rho|K|, K)}{\log 1/\rho},$$

so $\overline{\dim_M}(K) \leq \alpha$. Combining this with our previous result gives

$$\alpha = \dim(K) \leq \underline{\dim_M}(K) \leq \overline{\dim_M}(K) = \alpha,$$

and hence $\dim_M(K) = \alpha$. □

We note that the open set condition is necessary in the previous result. For example, if we take the maps

$$S_1(x) = \frac{2}{3}x, \quad S_2(x) = 1 - \frac{2}{3}x,$$

on $\mathbb{R}$ then the similarity dimension satisfies $2(2/3)^\alpha = 1$ which implies $\alpha > 1$, which is impossible for the Hausdorff dimension of a subset of the line.

Even more surprisingly, it has recently been shown [] that if the open set condition fails then so does the conclusion of Theorem 1.5.

2. Homogeneous sets

In this section we will consider sets which, in some sense, look the same at all scales. In particular we will present criteria which assure that the Hausdorff and Minkowski dimensions agree. We have already seen that the self-similar sets have this property.

Using the b-adic intervals introduced earlier we define

$$N_n(K) = \#(\{j \in \{1, \ldots, b^n\} : [\frac{j-1}{b^n}, \frac{j}{b^n}] \cap K \neq \emptyset\}).$$

It is straightforward to verify (as above) that

$$\overline{\dim_M}(K) = \limsup_{n \to \infty} \frac{\log N_n(K)}{n \log b}.$$

For the integer $b > 0$ define the b-to-1 map T_b mapping $[0, 1]$ to itself by

$$T_b(x) = bx \mod 1.$$

We now suppose that $K \subset [0,1]$ is a compact set such that $T_b K = K$. For example, the usual middle thirds Cantor set is invariant under the map T_3. More generally, if $D \subset \{0,\dots,b-1\}$ and

$$K = \{\sum_{n=1}^{\infty} a_n b^{-n} : a_n \in D\},$$

then K is compact and invariant under T_b.

If this is the case, we claim $N_{n+m}(K) \leq N_n(K)N_m(K)$. To see this, suppose the interval $I = [\frac{j-1}{b^m}, \frac{j}{b^m}]$ hits K and that inside this interval there are M intervals of the form $I_k = [\frac{k-1}{b^{n+m}}, \frac{k}{b^{n+m}}]$ which hit K. Multiplying by b^m and reducing mod 1 (i.e., applying T_b^n) the interval I is mapped onto $[0,1]$ and each of the intervals I_k is mapped to an intervals of length b^{-n} which cover $K \cap I$. Thus $M \leq N_m(K)$ and so $N_{n+m}(K) \leq N_n(K)N_m(K)$.

It is a general fact about sequences of real numbers that 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. Thus for T_b invariant sets, we deduce

$$\dim_M(K) = \lim_{n\to\infty} \frac{\log N_n(K)}{n \log b},$$

exists (see Exercise **??** in Chapter 2). Next we show that the Minkowski dimension of such a set agrees with it Hausdorff dimension.

LEMMA 2.1 (Furstenberg's Lemma). *If $K \subset [0,1]$ is compact and $T_b K = K$ then $dim(K) = dim_M(K)$.*

PROOF. Since we always have $\dim(K) \leq \overline{\dim}_M(K)$ we have to prove the other direction. Given a covering of K by b-adic intervals of possibly different sizes we have to produce an equally efficient covering by b-adic intervals all of the same size. In order to do this, it is convenient to introduce some notation and consider K as a subset of a sequence space rather than as a subset of $[0,1]$.

Let $\Omega = \{0,\dots,b-1\}$ let $\Omega^\infty = \Omega^{\mathbb{N}}$ be the sequence space of Ω. There is a natural identification of Ω^∞ with $[0,1]$ via the b-ary representations of real numbers $x = \sum_{n=1}^{\infty} x_n b^{-n}$. Using this we see that the map T_p can be written as

$$T_b(x) = \sum_{n=1}^{\infty} x_{n+1} b^{-n},$$

so that the induced map on on Ω^∞ is the left shift map T. The identification map from Ω^∞ to $[0,1]$ is continuous, so we may consider K as a compact, T invariant subset of Ω^∞.

We also define Ω^* as the space of finite sequences in Ω. It is a semi-group under the operation of concatenation

$$(a_1, \dots, a_n)(b_1, \dots, b_m) = (a_1, \dots, a_n, b_1, \dots, b_m).$$

we define the length of an element $\sigma \in \Omega^*$ as the number of entries and denote it by $|\sigma|$.

We label the b-adic intervals in $[0, 1]$ by elements of Ω^* as follows. For the interval $[\frac{j-1}{b^n}, \frac{j}{b^n}]$ we write the base b expansion of $j - 1$, as

$$j - 1 = \sigma_{n-1} b^{n-1} + \dots + \sigma_1 b + \sigma_0,$$

where $0 \leq \sigma_j < b$ and label this interval by the sequence

$$\sigma = (\sigma_{n-1}, \dots, \sigma_0) \in \Omega^*.$$

Let Π be a cover of K by b-adic intervals (not necessarily all the same size). Because K is compact we may take Π finite, so $L(\Pi) = \max_{\sigma \in \Pi} |\sigma|$ is well defined ($L(\Pi)$ measures the shortest interval used in Π). Let S be all the elements of Ω^* obtained by truncating some element of K so S is T invariant (since K is invariant).

Let S_Π be the subset of S with lengths bigger than $L(\Pi)$.

Note that any element of S_Π must have an initial segment belonging to Π. Therefore any $\sigma \in S_\Pi$ can be written as $\tau_1 \sigma_1$ with $\tau_1 \in \Pi$. But because S is T invariant , we must have $\sigma_1 \in S$. If $|\sigma_1| \leq L(\Pi)$ we stop, otherwise an initial segment of σ_1 must belong to Π as well. Now be induction we can write any $\sigma \in S$ as

$$\sigma = \tau_1 \tau_2 \dots \tau_r \sigma',$$

where $\tau_j \in \Pi$ and $|\sigma'| < L(\Pi)$. There are at most $b^{L(\Pi)}$ distinct possible values for σ'.

Suppose that for some α we have

$$\sum_{\tau \in \Pi} b^{-\alpha|\tau|} = q < 1.$$

Then

$$\sum_{\tau-1, \dots, \tau_r \in \Pi} b^{-\alpha|\tau_1 \cdots \tau_r|} = q^r < 1,$$

and

$$\sum_{r=1}^{\infty} \sum_{\tau-1, \dots, \tau_r \in \Pi} b^{-\alpha|\tau_1 \cdots \tau_r|} = \frac{q}{1-q}.$$

Piling up all the above we get

$$\sum_{\sigma \in S} b^{-\alpha|\sigma|} < \frac{b^{L(\Pi)}}{1-q} < \infty,$$

but since any $\sigma \in S$ corresponds to some $\frac{j-1}{b^n}$ for some n and $j \in \{1, \dots, b^n\}$, we find that

$$\sum_{\sigma \in S} b^{-\alpha|\sigma|} = \sum_{n=1}^{\infty} \sum_{\sigma \in S, |\sigma|=n} b^{-\alpha n} = \sum_{n=1}^{\infty} N_n(K) b^{-\alpha n} < 1.$$

Therefore, eventually $N_n(K)b^{-\alpha n} < 1$ and hence

$$\frac{\log N_n(K)}{n \log b} < \alpha,$$

and as $n \to \infty$ we obtain $\overline{\dim_M}(K) < \alpha$.

For any $\alpha > \dim(K)$ we have a covering Π with the desired condition $\sum_{\Pi} b^{-\alpha|\tau|} < 1$, so we deduce $\overline{\dim_M}(K) \leq \dim(K)$. Since we always have the opposite inequality

$$\dim(K) \leq \underline{\dim_M}(K) \leq \overline{\dim_M}(K),$$

we get equality, proving the lemma. □

EXAMPLE 2.2. We have already seen the example

$$K = \{\sum_{n=1}^{\infty} a_n b^{-n} : a_n \in D\},$$

where $D \subset \{0, \dots, b-1\}$, satisfies the hypothesis of Furstenberg's lemma, and its easy to check that $N_n(K) = |D|^n$, so we can deduce

$$\dim(K) = \dim_M(K) = \frac{\log D}{\log b}.$$

EXAMPLE 2.3. Another collection of compact sets which are T_b invariant are the *shifts of finite type* defined in Example **??**: let $A = (a_{ij})$, $0 \leq i, j < b$ be a $b \times b$ matrix of 0's and 1's and define

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

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. Since we computed the Minkowski dimension of this set in Chapter 2 we can now deduce

$$\dim(X_A) = \dim_M(X_A) = \rho(A) = \lim_{n \to \infty} \mid A^n \mid^{1/n}.$$

3. Microsets

DEFINITION 3.1. Suppose X is a compact metric space and A, B are compact subsets of X. We define the *Hausdorff distance* between these sets as

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

This makes $K(X)$, the set of compact subsets of X into a compact metric space itself.

A microset of $K \subset \mathbb{R}^d$ is a way of quantifying what we see as we "zoom in" closer and closer to K. Suppose $K \subset [0,1]$ is compact and $\{S_n\}$ is a sequence of expanding similarities, i.e.,

$$S_n(x) = \lambda x + a, \quad |\lambda| > 1.$$

DEFINITION 3.2. A set $\tilde{K} \subset [0,1]$ is called a *microset* of K if it is a limit point in the Hausdorff topology (see Section 1) of $S_n(K) \cap [0,1]$ for some sequence of expanding similarities $S_n(x) = \lambda x + a_n$, $|\lambda_n| \nearrow \infty$.

Note that a microset is something which is seen at arbitrarily small scales of K, but is not necessarily a subset of K (or necessarily seen at all scales). For example, if C is the usual middle 1/3 Cantor set then C is a microset of itself, but a copy of C on $[\frac{1}{3}, \frac{2}{3}]$ will also be a microset of C which is not contained in C.

DEFINITION 3.3. A set K is called *Furstenberg homogeneous* if for all microsets $\tilde{K}$ of K, $\dim(\tilde{K}) \le \dim(K)$.

DEFINITION 3.4. $\tilde{K}$ is called a *b-microset* of K when if is a limit using similarities of the form $S_n(x) = b^{l_n}x - a_n$ and $0 \le a_n < b^{l_n}$ is an integer (i.e., S_n corresponds to "blowing up" a b-adic interval to $[0,1]$.)

It is fairly easy to see that the set $F = \{\frac{1}{2}, \frac{1}{3}, \dots\} \cup \{0\}$, has microset $\tilde{K} = [0,1]$, so that it is not Furstenberg homogeneous. In particular, this gives an example of a zero dimensional set with microset of dimension 1. We leave it as an exercise to verify that in the definition of Furstenberg homogeneous, its suffices to only requires $\dim(\tilde{K}) \le \dim(K)$ for all b-microsets. Using this fact, we note that any T_b invariant sets (as described in the previous section) are Furstenberg homogeneous since any b-microset must be a subset of K.

It is useful to keep in mind the *tree description* of a set $K \subset [0,1]$. Recall that the collection of all b-adic intervals in $[0,1]$ can be viewed as the vertices of a tree T where edges connect each interval I to the b subintervals of length $|I|/b$. Moreover we can visualize this tree drawn

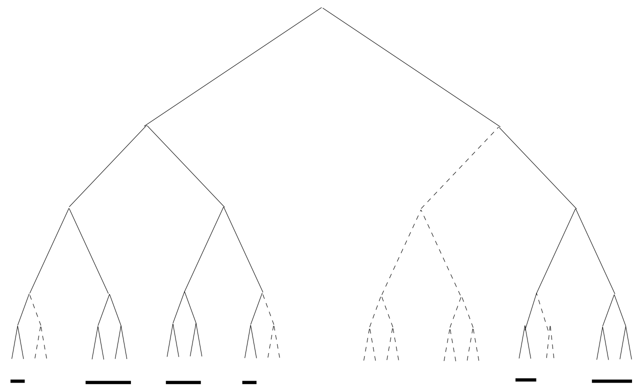

FIGURE 3.1. A set $K \subset [0,1]$ and the corresponding tree

in the plane with the root at the top and the edges ordered left to right in the same order as the corresponding edges. See Figure **??**.

The set of intervals which hit a nonempty set $K \subset [0,1]$ then form a subtree $T_k \subset T$. A "blow up" of K by a similarity $S_n(x) = b^{l_n} = a_n$ corresponds to choosing a vertex of the full tree and taking the subtree of T_K separated from the root ($[0,1]$) by this vertex, and making the chosen vertex the new root. The b-microsets correspond to taking all possible limits of such subtrees. A sequence of rooted trees $\{T_n\}$ is said to converge to a tree T if for any k there is an m so that $n \geq m$ implies the kth generation truncations of T and T_n agree. With this definition of convergence, the set of rooted trees of degree at most b becomes a compact space.

For example the full tree T corresponds to the set $K = [0,1]$, and an infinite (non self-intersecting) path corresponds to a point. The sets A_S, $S \subset \mathbb{N}$, described in Example **??** can easily be viewed in terms of the corresponding tree. Start by placing the root in the tree. If $1 \in S$ then but both adjacent vertices in the tree. Otherwise just but in the left vertex. In general, at level n we add both branches to the $(n+1)$st level if $1 \in S$ and only the left branch if $1 \notin S$.

EXAMPLE 3.5. If we take $S = \cup_{n=1}^{\infty}[(2n)!, (2n+1)!-1]$, we already know the set A_S has dimension 0, but there are long intervals were the corresponding tree T_S has full branching. Thus the full tree T is a limit of subtrees of T_S. Thus A_S has $[0,1]$ as a microset.

In the previous section we proved that compact sets invariant under T_b satisfy $\dim(K) = \dim_M(K)$. We have also noted that such sets are Furstenberg homogeneous, and now we will show that this property suffices:

LEMMA 3.6 (Extended Furstenberg lemma). *Any $K \subset [0,1]$ has a microset $\tilde{K}$ with dim$(\tilde{K}) \geq \overline{\text{dim}_M}(K)$.*

COROLLARY 3.7. *For any Furstenberg homogeneous set, dim$(K) =$ dim$_M(K)$.*

PROOF. The proof of the corollary is clear, so we only proof the lemma. To do that, it is enough to show that for any fixed $n > 1$ and $\alpha < \dim_M(K)$, there is a magnification of K

$$K_\alpha^n = S(K) \cap [0,1],$$

and a probability measure μ_α^n supported on K_α^n with the property that

$$\mu_\alpha^n([\frac{j-1}{b^r}, \frac{j}{b^r}]) \leq b^{-r\alpha},$$

for every $r = 0, 1, \dots, n$. If this is so, take a sequence $\{\alpha_k\}$ tending to $\dim_M(K)$ and get a sequence of sets and measures $\{\mu_{\alpha_k}^n\}$. Taking a weak limit of $\{\mu_{\alpha_n}^n\}$ gives a measure μ_α supported on a microset $\tilde{K}_\alpha$ of K which satisfies

$$\mu_\alpha([\frac{j-1}{b^r}, \frac{j}{b^r}]) \leq b^{-r\alpha},$$

for every $r = 0, 1, \dots$. The mass distribution principle Lemma **??** then implies $\dim(\tilde{K}) \geq \alpha$.

We are now reduced to proving the existence of K_α^n and μ_α^n. By definition

$$\alpha < \overline{\dim}_M(K) = \limsup_{n\to\infty} \frac{\log N_n(K)}{n \log b},$$

so there exists an $\epsilon > 0$ and arbitrarily large integers M such that

$$N_M(K) > b^{M(\alpha+\epsilon)}.$$

Define a probability measure ν_α on K by putting equal mass on each of the $N_M(K)$ b-adic intervals of size b^{-M}.

We claim that there is a a b-adic interval I and a M so that $\nu_\alpha/\nu_\alpha(I)$ restricted to I works. If not, then in particular the claim fails for $I = [0,1]$ so there exists a b-adic subinterval I_1 of generation $r \leq n$ so that

$$\nu_\alpha(I_1) > |I_1|^\alpha.$$

Since the claim also fails for I_1, there is a b-adic subinterval $I_1 \subset I_1$ such that

$$\frac{\nu_\alpha(I_2)}{\nu_\alpha(I_1)} > \frac{|I_2|^\alpha}{|I_1|^\alpha}.$$

Continuing in this way we obtain a sequence $I_1 \supset I_2 \subset \dots$, so that

$$\frac{\nu_\alpha(I_{j+1})}{\nu_\alpha(I_j)} > \frac{|I_{j+1}|^\alpha}{|I_j|^\alpha}.$$

Multiplying out and canceling, we get

$$\nu_\alpha(I_N) > |I_N|^\alpha \geq b^{-M\alpha}.$$

In the construction of ν_α, our sequence of intervals skips at most n generations at a time, so I_N belongs to some level between $M-n$ and M. Thus,

$$\nu_\alpha(I_N) \leq \frac{b^n}{N_M(K)},$$

since I_N contains at most b^M intervals of generation M. Since $N_M(K) > b^{M(\alpha+\epsilon)}$, we get

$$b^{-M\alpha} \leq \nu_\alpha(I_N) \leq b^n b^{-M(\alpha+\epsilon)}.$$

If M is large enough, this gives a contradiction, so the claim must hold. □

4. Poincaré sets

Let $H \subset \mathbb{N}$ and consider the set

$$X_H = \{\sum_{n=1}^{\infty} x_n 2^{-n} : x_n x_{n+h} \neq 1 \text{ for } h \in H\}.$$

This set is compact and is clearly shift invariant so by Furstenberg's lemma its Hausdorff and Minkowski dimensions agree.

For example if $H = \{1\}$, then this says that in x's binary expansion, a 1 always be followed by a 0. This is the same as the shift of finite type set X_A defined in Section 2 corresponding to the 2 matrix

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

The spectral radius of this matrix is $(1+\sqrt{5})/2$, so the dimension X_A is

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

For a general finite $H \subset \{1, \dots, N\}$ the set X_H, the values of the binary coefficients $x_{kN}, \dots, x_{(k+1)N-1}$ determine which values of $x_{(k+1)N}, \dots, x_{(k+2)N-1}$ are allowed. Thus X_H can be written as a shift of finite type with a $2^N \times 2^N$ matrix A, and the dimension of X_H can be computed from the spectral radius of this matrix

$$\dim(X_H) = \frac{\log \rho(A)}{N \log 2}.$$

Note that for $H' \subset H$ we have $X_H \subset X_{H'}$. For infinite sets H one could try to compute $\dim(H')$ using the spectral radius for all $H' \subset H$ finite, and take a limit, but this does not provide any useful formula. Rather than try to compute the dimension of X_H for general infinite sets H, we will concentrate on a more limited question: for which infinite sets H do we get $\dim(X_H) = 0$? The answer turns out to depend on some interesting combinatorial properties of the set H.

Recall that for $S \subset \mathbb{N}$ the upper density of S is

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

DEFINITION 4.1. $H \subset \mathbb{N}$ is called a Poincaré set if for every $S \subset \mathbb{N}$ with $\overline{\mathrm{d}}(S) > 0$ we have $(S - S) \cap H \neq \emptyset$ (i.e., there is an $h \in H$ such that $(S + h) \cap S \neq \emptyset$).

This may seem unwieldy at first, but turns out to be a natural concept. The original definition looks different, but is completely equivalent and may be easier to motivate. Given a a dynamical system, e.g., a space X and a mapping T from X to itself, it is natural to ask if iterates of a subset $A \subset X$ ever intersect itself. In finite measure preserving systems this must occur and we can ask at what times an iterate must return. A Poincaré sequence is one along which returns must occur, namely

DEFINITION 4.2. $H \subset \mathbb{N}$ is a Poincaré set if and only if for any probability measure μ on a measure space $(X, \mathcal{B})$, any measure preserving map $T : X \to X$ and any $A \subset X$ with $\mu(A) > 0$ we have $\mu(A \cap T^{-h}(A)) > 0$ for some $h \in H$.

See Furstenberg's book [**Fur81**] for the proof of the equivalence of the two definitions. It may not be clear from the second definition that many Poincaré sets exist, but its easy to check the first definition in some cases. For example, if $\overline{\mathrm{d}}(S) > 0$ then S certainly contains at least 3 distinct numbers a, b, c and if $a - b, b - c$ are both odd then $a - c$ is even. Thus the even numbers are a Poincaré set. The same argument shows $k\mathbb{Z}$ is Poincaré for any integer k. On the other hand, taking $S = 2\mathbb{Z}$ has positive density and only even differences so the odd numbers do not form a Poincaré set.

The square of an integer is always either 0 or 1 $\mod 3$, so $n^2 + 1 \mod 3$ must be either 1 or 2. Thus taking $S = 3\mathbb{N}$ shows $H = \{n^2 + 1 : n \in \mathbb{N}\}$ is not Poincaré. However, $H = \{n^2 : n \in \mathbb{N}\}$ is Poincaré (see Exercise 18.

DEFINITION 4.3. Suppose $H = \{h_j\}$ with $H_1 < h_2 < \ldots$. H is *lacunary* if there exists a $q > 1$ such that $h_{j+1} \geq q h_j$ for all j.

LEMMA 4.4. *A lacunary set is never Poincaré.*

PROOF. This is not too unreasonable since Poincaré sets should be "large" and lacunary sets are very "thin" (but note that $\{n^2\}$ is both Poincaré and fairly thin). The proof depends on the fact that given a lacunary H we have the following Diophantine approximation property: there is an irrational α and an $\epsilon > 0$ so that $\text{dist}(h\alpha, \mathbb{Z}) \geq \epsilon$ for every $h \in H$. This is left as an exercise (see Exercise 29).

If we assume this property then the proof goes as follows. Let

$$S = \{s \in \mathbb{N} : d(s\alpha, \mathbb{Z}) < \frac{\epsilon}{2}\}.$$

since α is irrational the multiples of α are equidistributed $\mod 1$ in $[0, 1]$ and therefore

$$\overline{\mathrm{d}}(S) = d(S) = \epsilon.$$

However, if $a, b \in S$ then $\text{dist}((a-b)\alpha, \mathbb{Z}) \leq \epsilon$, so $(S - S) \cap H = \emptyset$. Thus H is not Poincaré. ☐

THEOREM 4.5 (Furstenberg). *With X_h defined as above, dim$(X_H) =$ 0 if and only if H is Poincaré.*

PROOF. First assume H is Poincaré. It's enough to show that for

$$x = \sum_{n=1}^{\infty} x_n 2^{-n} \in X_H,$$

we have

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

for then Billingsley's lemma implies $\dim(X_H) = 0$. If this is not true then there is a bad $x \in X_H$ such that

$$\frac{1}{n}\sum_{k=1}^{n} x_k > \delta > 0.$$

Let $S = \{n \in \mathbb{N} : x_n = 1$, and note that $\overline{\mathrm{d}}(S) \geq \delta$. On the other hand, for any two elements $n, n+k \in S$, we have $x_n = x_{n+k} = 1$, but since $x \in H$, $x_n x_{n+h} \neq 1$ for all $h \in H$. Thus $k \notin H$, and so H is not Poincare', a contradiction. Therefore, the claim is true and $\dim)X_H) = 0$.

Suppose H is not Poincaré and suppose $S \subset \mathbb{N}$ has positive upper density and $(S - S) \cap H = \emptyset$. Define a real number x by

$$x = \sum_{n \in S} 2^{-n}.$$

Let $\{x_n\}$ be the binary expansion of this x (i.e., the characteristic function of S) and define

$$A = \{y : y_n \leq x_n\}.$$

Note

$$\overline{\dim_M}(A) \quad = \quad \limsup_{n\to\infty} \frac{\log 2^{|S\cap[1,n]|}}{n \log 2} = \overline{\mathrm{d}}(S) > 0$$

If $n \notin S$ then $x_n = 0$. If $n \in S$ then for any $h \in H$, we have $n + h \notin S$ which implies $x_{n+h} = 0$. Thus in either case $x_n x_{n+h} = 0 \neq 1$ for every This implies $A \subset X_H$, and we saw above that $\overline{\dim_M}(A) > 0$. Since X_H is invariant, Furstenberg's lemma implies

$$\dim(X_H) = \overline{\dim_M}(X_H) \geq \overline{\dim_M}(A) > 0.$$

$\square$

How can we check whether a set is Poincaré? One sufficient (but not necessary) condition that is easier to check in practice is the following.

DEFINITION 4.6. A set $H \subset \mathbb{N}$ is called a *Van der Corput* set (VDC set) if for any positive measure μ supported on the circle $\mathbb{R}/\mathbb{Z} = [0, 1)$, $\hat{\mu}(h) = 0$ for all $h \in H$, implies $\mu(\{0\}) = 0$.

Here $\hat{\mu}$ denotes the Fourier transform of μ

$$\hat{\mu}(k) = \int_0^1 e^{2\pi i k\theta} d\mu(\theta).$$

The examples we considered earlier in the section are VDC exactly when they are Poincaré. For example, let us show

LEMMA 4.7. *$7\mathbb{Z}$ is Van der Corput.*

PROOF. Given a positive measure μ on $[0, 1]$ assume $\hat{\mu}(7n) =$ for all n. The

$$0 = \frac{1}{n}\sum_{k=1}^{n} \hat{\mu}(7k) = \int_0^1 \frac{1}{n}\sum_{k=1}^{n} e^{-2\pi i 7k\theta} d\mu(\theta) \to \int_0^1 f(\theta) d\mu(\theta),$$

as $n \to \infty$ where

$$f(\theta) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} e^{-2\pi i 7k\theta} = \begin{cases} 0, & \theta \notin \mathbb{Z}/7 \\ 1, & \text{otherwise} \end{cases}$$

Thus

$$0 = \int_0^1 f(\theta) d\mu(\theta) = \mu(0) + \mu(\frac{1}{7}) + \cdots + \mu(\frac{6}{7}).$$

Since μ is positive each of the above is zero, in particular $\mu(\{0\}) = 0$. $\square$

LEMMA 4.8. *$H = 7\mathbb{N} + 1$ is not Van der Corput.*

PROOF. Let μ have equidistributed mass on $\{\frac{k}{7}\}$ for $k = \{0, \ldots, 6\}$. The $\mu(\{0\}) \neq 0$, but

$$\hat{\mu}(7k+1) = \int_0^1 e^{-2\pi i(7k+1)x} d\mu(x) = \frac{1}{7}\sum_{k=1}^{6} e^{-2\pi ik/7} = 0.$$

□

LEMMA 4.9. *If $S \subset \mathbb{N}$ is infinite, then $S - S$ is Van der Corput.*

PROOF. Suppose $\hat{\mu}(s_i - s_j) = 0$ for all $i, j \in S$, $i \neq j$. Let $e_m(x) = e^{2\pi imx}$ and note

$$\langle e_m, e_n \rangle_{L^2(\mu)} = \delta_{mn} \int e_0(x) d\mu(x).$$

Thus

$$\begin{aligned} N^2\mu(\{0\}) &\leq \int |\sum_{k=1}^{n} e_{s_k}(x)|^2 d\mu(x) \\ &= \int \sum_{1 \leq i,j \leq N} e_{s_i - s_j}(x) d\mu(x) \\ &= \int \sum_{1 \leq i \leq N} e_0(x) d\mu(x) \\ &= N \,|\, \mu \,|, \end{aligned}$$

which implies that

$$\mu(\{0\}) \leq \frac{|\,\mu\,|}{N} \to 0.$$

□

A set $H \subset \mathbb{N}$ is called thick if it contains arbitrarily long intervals. It is easy to see that any thick set contains an infinite difference set: given H choose $s_1 = 1$ and s_2 to be the right endpoint of an interval in H of length at least 2. Thus $s_2 - s_1 = s_2 - 1$ is in H. If we have chosen $s_1 < s_2 < \cdots < s_n$, choose s_{n+1} to be the right endpoint of an interval in H of length at least s_n. Thus all the differences $s_{n+1} - s_j$, $1 \leq j \leq n$ are in H and the other differences are in H by the induction hypothesis. Thus the previous lemma implies

COROLLARY 4.10. *Any thick set is Van der Corput.*

THEOREM 4.11. *If H is Van der Corput then H is Poincaré.*

Bourgain proved the converse is not true by constructing a Poincaré sequence which is not Van der Corput.

PROOF. Suppose $S \subset \mathbb{N}$ has positive upper density and let $S_N = S \cap [0, N]$. Let

$$\mu_N = \frac{1}{|S_N|} |\sum_{s \in S_n} e_s(x)|^2 dx,$$

where $e_s(x) = e^{2\pi isx}$ is as above. These are probability measures, so there is a weakly convergent subsequence with limit μ. Fix a $\delta > 0$ and choose N so large that $\delta \leq (10N)^{-1}$. Let $I_N = [\frac{-1}{10N}, \frac{1}{10N}]$. Then

$$\begin{aligned}
\mu_N(-\delta, \delta) &\geq \mu_N(I_N) \\
&= \int_{I_N} \frac{1}{|S_N|} |\sum_{s \in S_N} e_s(x)|^2 dx \\
&= \int_{I_N} \frac{1}{|S_N|} (\sum_{s \in S_N} \mathrm{Re}(e_s(x)))^2 dx
\end{aligned}$$

since $|z|^2 \geq (\mathrm{Re} z)^2$. Now observe that for for $x \in I_N$ and $s < n$

$$\mathrm{Re}(e_s(x) = \cos(2\pi sx) \geq \cos(2\pi \frac{N}{10N}) \geq \cos(\frac{\pi}{5}).$$

Therefore

$$\begin{aligned}
\mu_N(-\delta, \delta) &\geq \frac{1}{|S_N|} |S_n|^2 \inf s \in S_n, x \in I_N [\mathrm{Re}(e_s(x))]^2 \int_{I_N} dx \\
&\geq \frac{2\cos(\frac{\pi}{5})}{10N} |S_N| \\
&\geq C \frac{|S_N|}{N} \\
&\geq \overline{\mathrm{d}}(S) \\
&> 0.
\end{aligned}$$

Thus $\mu(-\delta, \delta) > \overline{\mathrm{d}}(S)$ for all $\delta > 0$ and so μ has a positive atom at 0.

Suppose H is Van der Corput. Then the Fourier transform of μ cannot vanish on H since it has an atom at 0. If we can show that $\hat{\mu}$ is supported in $S - S$ then we deduce that $H \cap (S - S) \neq \emptyset$. Thus H is Poincaré.

Thus we need only show that for $v \notin S - S$, we have $\hat{\mu}(v) = 0$. So take $v \notin S - S$ and write

$$\begin{aligned}
\hat{\mu}(v) &= \int_0^1 e_v(x) \frac{1}{|S_N|} |\sum_{s \in S_N} e_s(x)|^2 dx \\
&= \frac{1}{|S_N|} \int_0^1 e_v(x) (\sum_{s_1 \in S_N} e_{s_1}(x)) (\sum_{s_2 \in S_N} \bar{e}_{s_2}(x)) dx \\
&= \frac{1}{|S_N|} \sum_{s_1 - s_2 \in S_N - S_N} \int_0^1 (e_v(x) e_{s_1 - s_2}(x)) dx \\
&= \frac{1}{|S_N|} \sum_{S_N - S_N} \langle e_v, e_{s_1 - s_2} \rangle \\
&= 0,
\end{aligned}$$

since $v \neq s_1 - s_2$. This completes the proof. $\square$

5. Alternative definitions of Minkowski dimension

There are several equivalent definitions for Minkowski dimension. For instance, a packing of size ϵ of K is a collection of disjoint balls of size ϵ centered at points of K. Let $N_{\text{pack}}(K, \epsilon)$ be the maximum possible number of balls in a ϵ-packing of K and replace N_{cov} by N_{pack} in the definition above. This gives the same result. To see this, note that if $\{B_\epsilon(x_1), \ldots, B_\epsilon(x_n)\}$, $n = N_{\text{pack}}(K, \epsilon)$, is a maximal packing and x is any point of K then there is a x_j so that $|x - x_j| < 2\epsilon$ (otherwise we could insert a ball around x into the packing, contradicting maximality). Thus

$$N_{\text{cov}}(K, \epsilon) \leq N_{\text{pack}}(K, 2\epsilon).$$

On the other hand, if $B_\epsilon(y)$ belongs to a minimal ϵ-covering then it contains at most one center of a maximal ϵ-packing, so

$$N_{\text{pack}}(K, \epsilon) \leq N_{\text{cov}}(K, \epsilon),$$

and the equivalence of the two definitions follows.

Minkowski's original definition was different. If $K \subset \mathbb{R}^d$ he considered a "thickening" of the set K by ϵ, $K_\epsilon = \{x : \text{dist}(x, K) < \epsilon\}$ and defined

$$\overline{\dim_M}(K) = d + \limsup_{\epsilon \to 0} \frac{\log \text{Vol}(K_\epsilon)}{\log 1/\epsilon} = \limsup_{\epsilon \to 0} \frac{\log \text{Vol}((K_\epsilon)/\text{Vol}(B_\epsilon))}{\log 1/\epsilon},$$

where Vol is the d-dimensional volume. The equivalence of this the definitions above can be verified by checking the two inequalities

$$\text{Vol}(K_\epsilon) \leq N_{\text{cov}}(K, \epsilon) \text{Vol}(B_{2\epsilon}),$$

$$\text{Vol}(K_\epsilon) \geq N_{\text{pack}}(K, \epsilon)\text{Vol}(B_\epsilon),$$

which are both trivial.

Another version of Minkowski dimension that we will need later (in the Chapter on Kleinian groups) is due to Besicovitch and Taylor (on the line) and involves the idea of a Whitney decomposition. We will explain for domains in the plane (which is where we will use these results later).

Given a compact set K in the plane let $\Omega = \mathbb{R}^2 \setminus K$ be its complement. A Whitney decomposition of Ω is a collection of squares $\{Q_j\}$ which are disjoint, except along their boundaries, and such that

$$\frac{1}{10}\text{dist}(Q_j, \partial\Omega) \leq |Q_j| \leq 10\text{dist}(Q_j, \partial\Omega).$$

The existence of Whitney decomposition for any open set is a standard fact in real analysis (e.g., [**Ste71**]). The squares $\{Q_j\}$ may also be taken to be dyadic.

For any compact set $K \subset \mathbb{C}$ we can define an exponent of convergence

$$\kappa = \kappa(K) = \inf\{\alpha : \sum_j |Q_j|^\alpha < \infty\},$$

where the sum is taken over all squares in Whitney decomposition of $\Omega = K^c$ which are within distance 1 of K (we have to drop the "far away" squares or the series might not converge). It is easy to check that this does not depend on the particular choice of Whitney decomposition. This number has been rediscovered many times in the literature, but seem to have been first used by Besicovitch and Taylor in [**BT54**]. See also Tricot's paper[**Tri81**] where he shows that it agrees with the upper Minkowski dimension.

LEMMA 5.1. *For any compact set K, $\kappa \leq \overline{\dim}_M(K)$. If K also has zero area then $\kappa = \overline{\dim}_M(K)$.*

PROOF. We start with the easy direction, $\kappa \leq \overline{\dim}_M(K)$. Let $D = \overline{\dim}_M(K)$ and choose $\epsilon > 0$. Then by the definition of $\overline{\dim}_M(K)$ we have that

$$N(K, 2^{-n}) \leq C2^{D+\epsilon}.$$

Let $\mathcal{S}_n$ be a covering of K by fewer than $C2^{D+\epsilon}$ squares of size 2^{-n}. If Q is a Whitney square with $2^{-n-1} \leq |Q| < 2^{-n}$, then choose a point $x \in K$ with $\text{dist}(x, Q)| \leq |Q|)$. Let $S(Q) \in \mathcal{S}_n$ be the square containing Q. Since $S(Q)$ and Q have comparable sizes and there distance apart is at most $|Q|$, we easily see that each $S \in \mathcal{S}_n$ can only be associated to a uniformly bounded number of Q's in the Whitney decomposition, say A.

Let W_n be the number of Whitney squares with size $2^{-n-1} \leq |Q| < 2^{-n}$. Then

$$\begin{aligned}
\sum_j |Q_j|^{D+2\epsilon} &\leq \sum_{n=0}^{\infty} W_n 2^{-n(D+2\epsilon)} \\
&\leq \sum_{n=0}^{\infty} CN(K, 2^{-n}) 2^{-n(D+2\epsilon)} \\
&\leq C \sum_{n=0}^{\infty} 2^{n(D+\epsilon)} 2^{-n(D+2\epsilon)} \\
&\leq C \sum_{n=0}^{\infty} 2^{-n\epsilon} \\
&< \infty,
\end{aligned}$$

which proves $\kappa \leq D + 2\epsilon$. Taking $\epsilon \to 0$ gives $\kappa \leq D = \overline{\dim}_M(K)$.

Now we assume K has zero area and will prove $\kappa \geq D = \overline{\dim}_M(K)$. As above, let $\epsilon > 0$ and suppose $\{Q_j\}$ is a Whitney decomposition of $\Omega = K^c$. By the definition of $\overline{\dim}_M(K)$ we have

$$N(K, 2^{-n}) \geq 2^{n(D-\epsilon)},$$

for infinitely many n. Suppose n_0 is a value where this occurs and let $\mathcal{S} = \{S_k\}$ be a covering of K with dyadic squares of size 2^{-n_0}. Let for each $S_k \in \mathcal{S}$, $\mathcal{C}_k = \{Q_{jk}\}$ be the collection of Whitney squares which intersect S_k. If we assume the Q_j are dyadic, then every square hitting S_k is contained in S_k. Since the area of K is zero, this gives

$$2^{-2n_0} = \text{area}(S_k) = \text{area}(S_k \setminus K) = \text{area}(S_k \cap \Omega) = \sum_{\mathcal{C}_k} |Q_{jk}|^2.$$

Therefore,

$$\begin{aligned}
\sum_{\mathcal{C}_k} |Q_{jk}|^{D-2\epsilon} &= \sum_{\mathcal{C}_k} |Q_{jk}|^2 |Q_{jk}|^{-2+D-2\epsilon} \\
&\geq \sum_{\mathcal{C}_k} |Q_{jk}|^2 |S_k|^{-2+D-2\epsilon} \\
&= |S_k|^{-2+D-2\epsilon} \sum_{\mathcal{C}_k} |Q_{jk}|^2 \\
&= |S_k|^{-2+D-2\epsilon} |S_k|^2 \\
&= |S_k|^{D-2\epsilon} \\
&= 2^{-n_0(D-2\epsilon)}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\sum_j |Q_j|^{D-2\epsilon} &\geq \sum_k \sum_{\mathcal{C}_k} |Q_{jk}|^{D-2\epsilon} \\
&\geq \sum_k |S_k|^{D-2\epsilon} \\
&\geq N(K, 2^{-n_0}) 2^{-n_0(D-2\epsilon)} \\
&\geq 2^{n_0(D-\epsilon)} 2^{-n_0(D-2\epsilon)} \\
&= 2^{n_0\epsilon}.
\end{aligned}$$

Taking $n_0 \to \infty$, we get $\sum_j |Q_j|^{D-2\epsilon} = \infty$, and hence $\kappa \geq D - 2\epsilon$. Taking $\epsilon \to 0$ gives the desired result. □

6. Packing measures and dimension.

Tricot (1979,1982) introduced packing dimension, which is dual to Hausdorff dimension in several senses and does come with associated measures.

For any increasing function $\varphi : [0, \infty) \to \mathbf{R}$ such that $\varphi(0) = 0$ and any set E in a metric space, define first the *packing premeasure* (in gauge φ) by

$$\tilde{\mathcal{P}}_\varphi(E) = \lim_{\epsilon \downarrow 0} \sup \sum_{j=1}^{\infty} \varphi(\text{diam } B_j),$$

where the supremum is over all collections of disjoint closed balls $\{B_j\}_{j=1}^{\infty}$ with centers in E and diameters $\text{diam}(B_j) < \epsilon$. Then define the *packing measure* in gauge φ:

$$\mathcal{P}_\varphi(E) = \inf \left\{ \sum_{i=1}^{\infty} \tilde{\mathcal{P}}_\varphi(E_i) \mid E \subset \bigcup_{i=1}^{\infty} E_i \right\}.$$

This is indeed a Borel measure; when $\varphi(t) = t^\theta$ we write $\mathcal{P}_\theta$ for $\mathcal{P}_\varphi$ ($\mathcal{P}_\theta$ is called θ-dimensional packing measure).

Finally, define the *packing dimension* of E:

$$\text{Pdim}(E) = \inf \{\theta : \mathcal{P}_\theta(E) = 0\}. \tag{6.1}$$

We always have

$$\dim_H(E) \leq \text{Pdim}(E) \leq \overline{\dim}_M(E).$$

Packing measures are studied in detail in Taylor and Tricot (1985) and in Saint-Raymond and Tricot (1988); here we only mention the general properties we need.

The packing dimension of any set F in a metric space may be expressed in terms of upper Minkowski dimensions:

$$\text{(6.2)} \quad \mathrm{Pdim}(F) = \inf \left\{ \sup_{j \geq 1} \overline{\dim_M}(F_j) \mid F \subset \bigcup j = 1^{\infty} F_j \right\}$$

(See Tricot (1982) or Falconer (1990)).

Another useful fact is that the packing premeasure (in any gauge) assigns the same value to a set and to its closure. Thus in the definition (6.1) of packing measure, we can restrict attention to covers by closed sets. Using (6.2) and a Baire Category argument yields the following.

THEOREM 6.1. (Tricot (1982); see also Falconer (1990)). *Let K be a compact set in a metric space which satisfies*

$$\overline{dim_M}(K \cap V) = \overline{dim_M}(K)$$

for any open set V which intersects K. Then

$$Pdim(K) = \overline{dim_M}(K).$$

Next, let $\{f_j\}_{j=1}^{\ell}$ be contracting self maps of a complete metric space. Hutchinson [**Hut81**] showed there is a unique compact set K (the attractor) which satisfies

$$\text{(6.3)} \qquad K = \bigcup j = 1^{\ell} f_j(K).$$

THEOREM 6.2. *Assume $\{f_1, \ldots, f_\ell\}$ are contracting Bi-Lipschitz maps of a complete metric space, i.e.*

$$\epsilon_j d(x,y) \leq d(f_j(x), f_j(y)) \leq r_j d(x,y)$$

for all $1 \leq j \leq \ell$ and any x, y, where

$$0 < \epsilon_j \leq r_j < 1.$$

Denote by K the compact attractor satisfying (6.3). Then

1. *$Pdim(K) = \overline{dim_M}(K)$ and*
2. *for any gauge function φ such that K is σ-finite for $\mathcal{P}_\varphi$ we have*

$$\tilde{\mathcal{P}}_\varphi(K) < \infty.$$

PROOF. (1) Let V be an open set which intersects K and let $x \in K \cap V$. For any $m \geq 1$ there are $j_1, j_2, \ldots, j_m$ such that $x \in f_{j_1} \circ f_{j_2} \circ \cdots \circ f_{j_m}(K) \subset K$. Since $r_j < 1$, for m large enough we have $f_{j_1} \circ f_{j_2} \circ \cdots \circ f_{j_m}(K) \subset V$. The set on the left hand side is bi-Lipshitz equivalent to K. Consequently

$$\overline{\dim_M}(K \cap V) = \overline{\dim_M}(K)$$

and Theorem 6.1 applies.

(2) By the hypothesis, K can be covered by sets $\{E_i\}_{i=1}^{\infty}$ such that $\mathcal{P}_\varphi(E_i) < \infty$ for all i. Thus $E_i \subset \bigcup_{j=1}^{\infty} E_{ij}$ where $\tilde{\mathcal{P}}_\varphi(E_{ij}) < \infty$ and without loss of generality E_{ij} are closed sets. By Baire's theorem there are indices i, j and an open set V intersecting K such that $E_{ij} \supset K \cap V$. As detailed in part (i), $K \cap V$ contains a bi-Lipshitz image of K and therefore $\tilde{\mathcal{P}}_\varphi(K) < \infty$. □

If $\dim(K) = \overline{\dim_M}(K)$ the result is trivial since the packing dimension is trapped between these numbers. Non-trival examples will be given in Chapter 5 when we consider the self-affine sets (see Section 5, Chapter 5).

Finally, we quote a basic density result.

THEOREM 6.3. *Let μ be a Borel probability measure on $\mathbf{R}^d$ and $E \subset \mathbf{R}^d$ a set such that $\mu(E) > 0$.*

Fix a gauge function φ and denote by $B(z, \epsilon)$ the ball of radius ϵ centered at z. Then

1. *If $0 < c_1 \le \frac{\mu[B(z,\epsilon)]}{\varphi(2\epsilon)} \le c_2 < \infty$ for all $z \in E$ and $0 < \epsilon < 1$, then*
$$0 < \mathcal{P}_\varphi(E) < \infty.$$
2. *If $\limsup_{\epsilon \to 0} \frac{\varphi(\epsilon)}{\mu B(z,\epsilon)} = 0$ for all $z \in E$, then $\mathcal{P}_\varphi(E) = 0$.*

Exercises

EXERCISE 1. Using the notation of Lemma 1.1, show that K is the closure of the countable set of fixed points of all maps of the form $\{S_{j_1} \circ \cdots \circ 1S_{j_n}\}$.

EXERCISE 2. Using the notation of Lemma 1.1, show that K is the closure of the countable set of fixed points of all maps of the form $\{S_{j_1} \circ \cdots \circ S_{j_n}\}$.

EXERCISE 3. Suppose K_1 and K_2 are two self similar sets which are Jordan arcs and suppose that $\dim(K_1) = \dim(K_2)$. Then there is a bi-Lipschitz mapping of one to the other. (Bi-Lipschitz means that there is a $C < \infty$ such that $C^{-1}|x - y| \le |f(x) - f(y)| \le C|x - y|$.)

EXERCISE 4. If K is a self-similar Jordan arc of dimension α, prove there is a 1-1 mapping $f : [0, 1] \to K$ of Hölder class $\beta = 1/\alpha$. (Hölder class β means there is a $C < \infty$ such that $|f(x) - f(y)| \le C|x - y|^\beta$ for all x and y.

EXERCISE 5. Write down explicit linear maps which give the Sierpinski gasket as the attractor. Do the same for the von Koch snowflake.

Exercise 6. Suppose E is a self-similar set in $\mathbb{R}^n$. Is the projection of E onto a k-dimensional subspace necessarily a self-similar set in $\mathbb{R}^k$?

Exercise 7. Suppose E is a self-similar set in $\mathbb{R}^n$. Is the intersection of E with a k-dimensional subspace necessarily a self-similar set in $\mathbb{R}^k$?

Exercise 8. Suppose K is the self-similar set associated to the similarities $\{f_1, \dots, f_n\}$. Prove that the set associated to $\{f_1, \dots, f_{n-1}$ is a subset of strictly smaller dimension.

Exercise 9. Let $\{x_n\}$ be the base 5 expansion of x. Find
$$\dim(\{x : x_n x_{n+2} = 3 \bmod 5 \text{ for all } n\}).$$

Exercise 10. Let $\{x_n\}$ be the base 5 expansion of x. Find
$$\dim(\{x : (x_n + x_{n-1})^2 + x_{n-2} = 2 \bmod 5 \text{ for all } n\}).$$

Exercise 11. Let $\{x_n\}$ be the base 5 expansion of x. Let
$$X = \{x : x_n^2 = n^2 \bmod 5 \text{ for all } n\}.$$
Show X is Furstenberg homogeneous, but is not invariant under $T_5 : x \to 5x \bmod 1$.

Exercise 12. Suppose
$$f_1(x) = \frac{1}{3}x, f_2(x) = \frac{2}{3} + \frac{1}{3}x, f_3(x) = \frac{1}{6} + \frac{1}{3}x.$$
What is the dimension of the attractor of (f_1, f_2, f_3)?

Exercise 13. Suppose $f_n(x) = \frac{1}{n} + \frac{1}{n^2}x$. What is the dimension of the attractor of $\{f_n\}_{n=2}^{\infty}$?

Exercise 14. Let $\{x_n\}$ denote the binary expansion of x. Find
$$\dim(\{x : \sum_{k=1}^{n} x_n \le k/2 \text{ for all } k\}).$$

Exercise 15. Let $\{x_n\}$ denote the binary expansion of x. Find
$$\dim(\{x : \lim_{k\to\infty} \sum_{k=1}^{n} x_n \text{ does not exist }\}).$$

Exercise 16. Prove that if α is irrational then $n = \alpha \mod 1$ is dense in $[0, 1]$.

Exercise 17. Suppose $p(x) = \alpha_n x^n + \cdots + \alpha_1 x + \alpha_0$ has at least one irrational coefficient. Prove that $p(n) \mod 1$ is dense in $[0, 1]$.

EXERCISE 18. Shows that $H = \{n^2\}$ is Van der Corput (Hint: use the fact that $n^2\alpha$ is equidistributed for any irrational α.

EXERCISE 19. Construct a compact set such that the Besicovitch-Taylor index is not equal to its dimension.

EXERCISE 20. Suppose $K \subset \mathbb{R}^2$ is compact and $x \in K$. We say that x is a point of approximation (or a conical limit point) of K if there is a $C < \infty$ such that for every $\epsilon > 0$ there is a $y \notin K$ with $|x-y| \leq \epsilon$ and $\operatorname{dist}(y, K) \geq |x-y|/C$. Denote the conical limit points of K by K_c. Show that the Besicovitch-Taylor index of K is always an upper bound $\dim(K_c)$.

EXERCISE 21. Construct a set K so that $\dim(K_c) < \dim(K)$.

EXERCISE 22. Is it true that $\dim(K_c)$ equals the Besicovitch-Taylor index of K for any compact set K?

EXERCISE 23. Suppose $E = [0,1] \setminus \cup_j I_j$ where $\{I_j\}_1^\infty$ are disjoint open intervals of length r_j. Suppose $\sum_j r_j = 1$ and $\sum_j r_j^2 = 1/2$. Show that $\dim(E) \leq 1/2$. Find an example to show this is sharp.

EXERCISE 24. Suppose $K \subset \mathbb{R}^2$ is compact and has zero area. Let N_n be the number of Whitney squares in $\Omega = \mathbb{R}^2 \setminus K$ with side length $2^{-n} \leq \ell(Q) < 2^{-n+1}$. Show that the lower Minkowski dimension is

$$\liminf_{n\to\infty} \frac{\log N_n}{\log 2}.$$

EXERCISE 25. Verify that in the definition of Furstenberg homogeneous, its suffices to only requires $\dim(\tilde{K}) \leq \dim(K)$ for all b-microsets.

EXERCISE 26. Show that there is a compact set K which has every compact subset of $[0,1]$ as a microset.

EXERCISE 27. Is $K = \{0\} \cup \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ homogeneous? What are all its microsets?

EXERCISE 28. Given $K \subset [0,1]$, let $M(K)$ be the set of all microsets of K. Show M is connected with respect to the Hausdorff metric.

EXERCISE 29. Given a lacunary sequence $\{h_n\}$ there is an irrational α and an $\epsilon > 0$ so that $\operatorname{dist}(h_n\alpha, \mathbb{Z}) \geq \epsilon$ for every n.

EXERCISE 30. Is $\{e^{\lfloor\sqrt{n}\rfloor}\}$ Poincaré?

EXERCISE 31. Construct a set K so that

$$\dim(K) < \operatorname{Pdim}(K) M\dim_M(K).$$

EXERCISE 32. Prove Theorem 6.3.

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.