Contents

CHAPTER 1

Frostman's theory and capacity.

In Chapter 2 we proved the mass distribution principle and Billingsley's lemma. These showed that estimates for the Hausdorff dimension of a set could be given by finding appropriate measures on the set. In this chapter we show that finding the Hausdorff dimension is equivalent to finding a "smooth" measure on the set and use this powerful idea to derive a variety of results including the behavior of dimension under products, projections and slices.

1. Frostman's Lemma.

Although the definition of Hausdorff dimension involves only coverings of a set by balls, we have seen that computing the dimension of a set usually involves constructing a measure on the set. In particular, the mass distribution principle, Lemma **??** of Chapter **??** says that if a compact set K supports a measure μ which satisfies

$$\mu(B) \leq C|B|^{\alpha},$$

for every ball B then $\mathcal{H}^{\alpha}_{\infty}(K) \geq \frac{1}{C}\mu(K)$. In this section we will prove a very sharp converse of this.

Lemma 1.1 (Frostman's lemma). *Let $K \subset \mathbb{R}^d$ be a compact set with positive Hausdorff content, $\mathcal{H}^{\varphi}_{\infty}(K) > 0$. Then there is a positive measure μ on K satisfying*

$$\mu(B) \leq \varphi(|B|),$$

for all balls B and $\mu(K) \geq C_d \mathcal{H}^{\varphi}_{\infty}(K)$

Here C_d is a positive constant depending only on d. This was proved by Frostman in 1935. To prove it we will use the Maxflow-Mincut theorem. This is a result from graph theory about general networks, but we will only need it in the case of trees.

Let Γ be a rooted tree of depth n and with vertex set V and edges E. We let σ_0 denote the root vertex and let $|\sigma|$ denote the depth of σ, i.e., its distance from σ_0. For each vertex σ (except the root) we let σ' denote the unique vertex adjacent to σ and closer to the root. We sometimes call σ' the "parent" of σ. To each edge we associate a

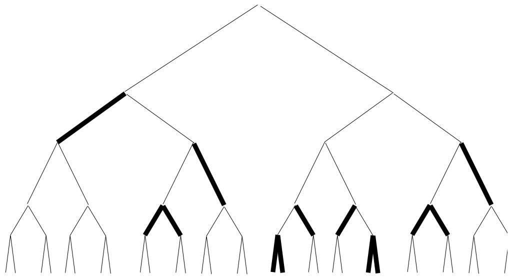

FIGURE 1.1. Cutset in dyadic tree

positive number $C(e), e \in E$, called the conductance of the edge. If we think of the tree as a network of pipes the conductance of any edge represents the maximum amount of material that can flow through it. A *flow* on Γ is a non-negative function f on Γ which satisfies

$$f(\sigma'\sigma) = \sum_{\tau:\sigma'=\tau} f(\sigma\tau),$$

for all vertices σ of depth from 1 to n. A *legal flow* must satisfy

$$f(\sigma'\sigma) \leq C(\sigma'\sigma)$$

for all vertices σ. The *norm* of a flow f is the sum

$$|\, f \,| = \sum_{|\sigma|=1} f(\sigma_0\sigma).$$

A *cutset* of Γ is a set of edges Π which disconnects the root from the nth level vertices. For example, taking all the edges from kth to $(k+1)$st level vertices for some $0 \leq k < n$ form a cutset. A *minimal cutset* is a cutset which has no proper subsets which are cutsets (i.e., the removal of any edge allows a path from the root to some nth level vertex). Because a cutset corresponds to a covering of the compact set $\partial\Gamma$ by open sets, a minimal cutset is finite. Given a cutset we define Γ_Π to be the connected component of $\Gamma \setminus \Pi$ containing the root σ_0. Vertices in Γ_Π are also said to be "above" Π.

Cutsets give a way to bound the norm of a flow, for given any cutset Π, we have

$$|\, f \,| \leq \sum_{e\in\Pi} f(e).$$

Moreover, if Π is a minimal cutset then one can write

$$|\, f \,| = \sum_{|\sigma|=1} f(\sigma_0\sigma) + \sum_{\sigma\in\Gamma_\Pi\setminus\{\sigma_0\}} [-f(\sigma'\sigma) + \sum_{\tau'=\sigma} f(\sigma\tau)] = \sum_{e\in\Pi} f(e).$$

If f is a legal flow then the right hand side is bounded by

$$\sum_{e\in\Pi} C(e) \equiv C(\Pi).$$

This number, $C(\Pi)$, is called the *cut* corresponding to the cutset Π. It is now clear that

$$\max_{\text{legal flows}} |\, f\,| \le \min_{\text{cutsets}} C(\Pi).$$

The Maxflow-Mincut theorem says that we actually have equality.

THEOREM 1.2 (Ford-Fulkerson). *For a finite rooted tree of depth n,*

$$\max_{\textit{legal flows}} |\, f\,| = \min_{\textit{cutsets}} C(\Pi).$$

PROOF. The set of legal flows is compact, so there exists legal flow f^* attaining the maximum possible norm. Let Π_0 be the set of all saturated edges for f^*, i.e, $\Pi_0 = \{e \in E : f(e) = C(e)\}$. If there is a path from the root to level n which avoids Π_0 then we can add ϵ to f^* on each of the edges in the path and obtain another legal flow, contradicting f^*'s maximality. Therefore there is not such path and we deduce Π_0 is a cutset. Now let $\Pi \subset \Pi_0$ be a minimal cutset. But then

$$|\, f^*\,| = \sum_{e\in\Pi} f^*(e) = \sum_{e\in\Pi} C(e) = C(\Pi),$$

which proves

$$\max_{\text{legal flows}} |\, f\,| = |\, f^*\,| \ge \min_{\text{cutsets}} C(\Pi).$$

Since the opposite inequality always holds, we have proven the result. $\square$

COROLLARY 1.3. *If Γ is an infinite tree,*

$$\max_{\textit{legal flows}} |\, f\,| = \min_{\textit{cutsets}} C(\Pi).$$

PROOF. We first note that a maximal flow will still exist. This is because for each n there is a flow f_n^* which is maximal with respect to the first n levels of Γ and a corresponding minimal cutset Π_n so that $|\, f_n^*\,| = C(\Pi_n)$. Now take a pointwise limit f^* by a Cantor diagonal argument to obtain a flow which is good for all levels. Clearly

$$|\, f^*\,| = \lim_{n\to\infty} |\, f_n^*\,| = \lim_{n\to\infty} C(\Pi_n),$$

so that $|\, f^*\,| = \inf_{\Pi} C(\Pi)$. $\square$

LEMMA 1.1. Fix an integer $b > 1$ (say $b = 2$) and construct the b-adic tree corresponding to K (vertices correspond to b-adic intervals which hit K and so that each b-adic interval is connected to its "parent"). Define a conductance by assigning to each edge $\sigma'\sigma$, $|\sigma| = n$ the number $C(\sigma'\sigma) = \varphi(b^{-n})$. Let f be the maximal flow given by the theorem and note that the definition of a flow allows one to define a measure μ on K by

$$\mu(I_\sigma) = f(\sigma'\sigma).$$

Since the flow is legal we have

$$\mu(I_\sigma) \leq \varphi(\sigma) = \varphi(|I_\sigma|),$$

for every b-adic interval I_σ. Any interval J can be covered by C_d b-adic intervals of smaller size, and so since φ is increasing

$$\mu(J) \leq C_d \varphi(|J|).$$

Now each b-adic cover of K corresponds to a cutset of the tree Γ and vice versa, so

$$\inf_{\Pi} C(\Pi) = \hat{\mathcal{H}}^\varphi_\infty(K) \geq \mathcal{H}^\varphi_{(\infty)}(K),$$

where $\hat{\mathcal{H}}^\varphi_\infty$ is the b-adic Hausdorff content of K. By the Maxflow-Mincut theorem,

$$|\, f\,| = \mu(K) \geq \mathcal{H}^\varphi_\infty(K),$$

and so μ satisfies all the required conditions. ☐

2. The dimension of product sets

THEOREM 2.1. *If $A, B \subset \mathbb{R}^d$, then*

$$dim(A) + dim(B) \leq dim(A \times B) \leq dim(A) + Pdim(B).$$

The left hand inequality is due to Marstrand [?] and the right hand side to Tricot [?], refining an early result of Besicovitch and Moran [?].

PROOF. We start with the left hand inequality. Suppose $\alpha < \dim(A)$ and $\beta < \dim(B)$. Then there are Frostman probability measures μ_A, μ_B supported on A and B respectively such that for all D, F,

$$\mu_A(D) \leq C_A |D|^\alpha,$$
$$\mu_B(D) \leq C_B |F|^\beta.$$

Suppose $|D| = |F|$ and take the "square" $D \times F$. Then

$$\mu_A \times \mu_B(D \times F) \leq C_A C_B |D|^{\alpha+\beta},$$

so it is a Frostman measure for $A \times B$. Thus $\dim(A \times B) \geq \alpha + \beta$ and taking $\alpha \to \dim(A)$, $\beta \to \dim(B)$ gives the result.

Now for the right hand inequality. First we prove

$$\dim(A \times B) \leq \dim(A) + \overline{\dim}_M(B). \tag{2.1}$$

Choose $\alpha > \dim(A)$ and $\beta > \overline{\dim}_M(B)$. Then we can find ϵ_0 such that $N(\epsilon, B) \leq \epsilon^{-\beta}$ for any $\epsilon < \epsilon_0$. We can also find a cover $\{A_j\}$ of A satisfying $|A_j| \leq \epsilon_0$ for all j and $\sum_j |A_j|^\alpha < \infty$.

For each j cover B by a collection $\{B_{jk}\}_{k=1}^{L_j}$ of balls with $|B_{jk}| = |A_j|$. We can do this with $L_j \leq |A_j|^{-\beta}$. The family $\{A_j \times B_{j,k}\}_{j,k}$ covers $A \times B$ and

$$\begin{aligned} \sum_j \sum_{k=1}^{L_j} |A_j \times B_{jk}|^{\alpha+\beta} &\leq C \sum_j L_j |A_j|^{\alpha+\beta} \\ &\leq C \sum_j |A_j|^\alpha \\ &< \infty. \end{aligned}$$

Thus $\dim(A \times B) \leq \alpha + \beta$. Taking limits as $\alpha \to \dim(A)$ and $\beta \to \overline{\dim}_M(B)$ gives (2.1).

To finish the proof of the theorem we use the definition of packing dimenion to cover B by closed sets $\{B_n\}$ such that $\sup_n \overline{\dim}_M(B_n) \leq \mathrm{Pdim}(B) + \epsilon$. Thus

$$\dim(A \times B) \leq \sup_n \dim(A \times B_b) \leq \dim(A) + \mathrm{Pdim}(B) + \epsilon.$$

Taking $\epsilon \to$ gives the theorem. □

EXAMPLE 2.2. In Chapter 2 we computed the dimension of the Cantor middle thirds set $E_{1/3}$ and the product $E_{1/3} \times E_{1/3}$ (see examples **??** and **??**) and found

$$\dim(E_{1/3} \times E_{1/3}) = \log_3 4 = 2\log_3 2 = 2\dim(E_{1/3}).$$

Thus equality holds in Marstrand's theorem (as required since $E_{1/3}$ is self-similar and hence its Minkowski and Hausdorff dimensions agree).

EXAMPLE 2.3. Recall Examples **??** and **??** of Chapter 2: suppose $S \subset \mathbb{N}$, and define $A_S = \{x = \sum_{k=1}^\infty x_k 2^{-k}\}$ where

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

We showed that

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

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

Now let

$$S = \cup_{j=1}^{\infty}[(2j-1)!, (2j)!-1],$$
$$S^c = \cup_{j=1}^{\infty}[(2j)!, (2j+1)!-1].$$

In this case it is easy to check that

$$\liminf_{N\to\infty} \frac{|S \cap \{1,\dots,N\}|}{N} = \liminf_{N\to\infty} \frac{|S^c \cap \{1,\dots,N\}|}{N} = 0,$$

so that $\dim(A_S) = \dim(A_{S^c}) = 0$. On the other hand, if

$$E = A_S \times A_{S^c} = \{(x,y) : x \in A_S, y \in A_{S^c}\},$$

then we claim $\dim(E) \geq 1$. To see this, note that

$$[0,1] \subset A_S + A_{S^c} = \{x+y : x \in A_S, y \in A_{S^c}\}$$

since we can partition the binary expansion of any real number in $[0,1]$ appropriately. Thus the map

$$(x,y) \to x+y,$$

defines a Lipschitz map of E onto $[0,1]$. Since it is easy to see that a Lipschitz map cannot increase Hausdorff dimension, we deduce $\dim(E) \geq 1$.

This gives an example of inequality in Marstrand's theorem.

3. The product theorem is sharp

The theorem proven in the previous section gives upper and lower bounds for the Hausdorff dimension of a product $A \times B$ as

$$\dim(A) + \dim(B) \leq \dim(A \times B) \leq \mathrm{Pdim}(A) + \dim(B).$$

We saw examples where both inequalities where sharp, but there is a more subtle question: for any A are both inequalities sharp? In other words, given A, do there exist sets B_1 and B_2 so that

$$\dim(A \times B_1) = \dim(A) + \dim(B_1),$$
$$\dim(A \times B_2) = \mathrm{Pdim}(A) + \dim(B_2)?$$

The first is easy since any set B_1 where $\dim(B_1) = \mathrm{Pdim}(B_1)$ will work, say $B_1 = [0,1]$, or any self-similar set. The second question is harder, but we will prove in this section that it is true by constructing such a B_2. This is more technical than most of the other results in the book, but it illustrates the advantages of the tree representation of sets to clarify an argument and is elementary in the sense that we only use the definitions and mass distribution principle. The proof we give here is from the authors paper [**?**]. A similar result for sets in $\mathbb{R}$ was given independently by Y. Xiao in [?]. We shall only give the proof for compact sets. By quoting results of Joyce and Preiss [?] this implies it is true for the much larger class of analytic sets (which contain all Borel sets as a proper subclass).

THEOREM 3.1. *Given a compact set $A \subset \mathbb{R}^d$ and $\epsilon > 0$ there is a compact $B \subset \mathbb{R}^d$ so that $dim(B) \le d - Pdim(A) + \epsilon$ and $dim(A \times B) \ge d - \epsilon$.*

Thus
$$\text{Pdim}(A) = \sup_B[\dim(A \times B) - \dim(B)],$$
where the supremum is over all compact subsets of $\mathbb{R}^d$. It is a simple matter to modify the proof to take $\epsilon = 0$, but we will leave this as an exercise for the reader. Before starting the construction, it is convenient to replace A by a subset which is more regular in the following sense.

LEMMA 3.2. *Suppose E is a subset of a separable metric space and $Pdim(A) > \alpha$. Then there is a non-empty subset $F \subset E$ such that $Pdim(F \cap V) > \alpha$ for every open set V such that $F \cap V \neq \emptyset$.*

PROOF. Define F to be the set of points x in E such that $\text{Pdim}(E \cap U) > \alpha$ for every neighborhood U of x. It is easy to check that F is closed (if $\{x_n\} \subset F$ converges to$x$ then any neighborhood W of x contains some neiborhood U of some x_n and hence $\text{Pdim}(W \cap E) \ge \text{Pdim}(U \cap A) > \alpha$). If F were empty then E could be covered by a countable subcover (by separability) of sets of packing dimension $\le \alpha$, a contradiction. Fianlly we have to check F has the desired regularity. If V is an open set hitting F then $V \cap (E \setminus F)$ has a countable cover by sets of packing dimension $\le \alpha$ and hence $\text{Pdim}(V \cap (E \setminus F)) \le \alpha$. On the other hand, since V is a neighborhood of some point of F, $\text{Pdim}(E \cap V) > \alpha$. Since
$$V \cap E = (V \cap F) \cup (V \cap (E \setminus F),$$
we deduce $\text{Pdim}(V \cap F) > \alpha$. □

For the rest of the section we will assume A satisfies the conclusion of the lemma (and absord the difference into the ϵ).

Before beginning the proof we also observe how to find the tree associated to the product of two sets. Given two compact sets A, B in $[0,1]^d$, we can construct a tree associated to the Cartesian product $A \times B \subset [0,1]^{2d}$ from trees $T(A)$ and $T(B)$. In general, given two trees T^1, T^2, the product tree $T^1 * T^2$ has the vertex set
$$\{(v_1, v_2) \ : \ v_i \in T^i \ \text{ for } \ i = 1,2 \ \text{ and } \ |v_1| = |v_2|\},$$

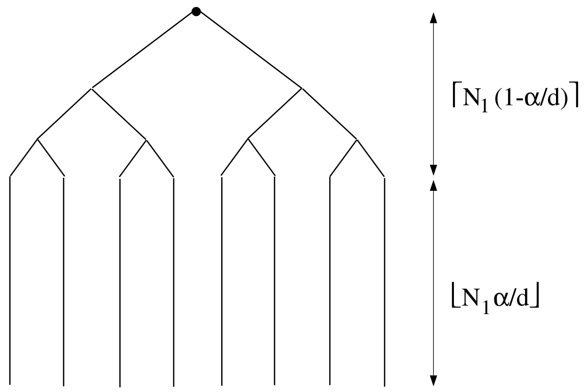

FIGURE 3.1. Maximal branching followed by no branching

the root $(\mathrm{root}(T^1), \mathrm{root}(T^2))$, and the adjacency relation: *the vertex* (v_1, v_2) *is adjacent to* (u_1, u_2) *iff* v_i *is adjacent to* u_i *for* $i = 1, 2$.
It is easy to see that $T(A) * T(B)$ is the tree associated with $A \times B$.

THEOREM 3.1. Assume $A \subset [0,1]^d$ is compact set satisfying the conclusion of Lemma 3.2 and let $\mathrm{Pdim}(A) - \epsilon < \alpha < \mathrm{Pdim}(A)$. his implies that for any vertex v of the tree $T(A)$, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log_2 |T_n(A, v)| \geq \alpha \,, \tag{3.1}$$

where $|T_n(A, v)|$ is the number of descendants w of v such that $|w| = n$.

We will consider truncations $T^k(A)$ of the tree $T(A)$ to certain levels N_k, and use them to inductively construct a nested increasing sequence of trees Γ^k, whose limit Γ is associated to the promised set B. Simultaneously, we will describe directed unit flows from the root to the boundaries of the product trees $T^k(A) * \Gamma^k$ which satisfy uniform Hölder estimates of orders d_k close to d. This will suffice to establish the result.

First step: By (3.1), there is a level N_1 such that $T(A)$ has more than $2^{\alpha N_1}$ vertices at level N_1. Construct Γ^1 of uniform depth N_1, where every vertex in the first $\lceil N_1(1-\alpha/d)\rceil$ levels has 2^d children, and every vertex in the remaining $\lfloor N_1\alpha/d\rfloor$ levels has a single child (see Figure 1). The leaves of Γ^1 form a cut-set Π_1 of cardinality at most $2^{N_1(d-\alpha)+d}$, so the corresponding $d-\alpha$ dimensional cut-set sum satisfies

$$\sum_{v\in\Pi_1} 2^{(\alpha-d)|v|} \leq 2^d \,.$$

Now consider the uniform unit flow μ on the product tree $T^1(A)*\Gamma^1$ which assigns each vertex at level N_1 the same mass.

Let (v, w) be a vertex at level $j \leq N_1$ of this product tree. Observe that v has at most $2^{d(N_1-j)}$ descendants at level N_1 of $T(A)$. If $j > \lceil N_1(1-\alpha/d)\rceil$ then w has only one descendant at level N_1 of Γ^1, so that

$$\mu(v,w) \leq 2^{d(N_1-j)} \cdot 1 \cdot 2^{-\alpha N_1} \cdot 2^{-dN_1(1-\alpha/d)} = 2^{-dj+(\alpha-\alpha)N_1} \,. \tag{3.2}$$

Denoting

$$\tilde{d} := d - \frac{\alpha - \alpha}{1-\alpha/d} \,, \tag{3.3}$$

we deduce from (3.2) that the Hölder estimate

$$\mu(v,w) \leq 2^{-\tilde{d}\,|(v,w)|} \tag{3.4}$$

holds if $j = |(v,w)| > \lceil N_1(1-\alpha/d)\rceil$. On the other hand, an improved Hölder estimate (with d in place of $\tilde{d}$) clearly holds for smaller values of j, since Γ_1 has full branching at these levels. We conclude that (3.4) is valid for all vertices (v, w) of $T^1(A) * \Gamma^1$. For use in the induction below, we define $N_0 = 0$ and $d_0 = d_1 = \tilde{d}$.

The inductive step: We assume that we have constructed a tree Γ^k of uniform depth N_k, in which every vertex has at most 2^d children. Also, we assume there is a unit flow μ on the product tree $T^k(A) * \Gamma^k$ which satisfies $\mu(v,w) \leq 2^{-d_{k-1}j}$ for any $N_{k-1} \leq j \leq N_k$ and any vertex (v, w) at level j of the product tree; for <u>leaves</u> (v, w) of Γ^k, we assume the estimate $\mu(v,w) \leq 2^{-d_k N_k}$ holds.

The construction of Γ^{k+1} from Γ^k is done in four stages:

1. To each leaf of Γ^k attach a full 2^d-ary tree of depth kN_k to obtain a tree $\Gamma^{k,1}$ of uniform depth $(k+1)N_k$.
2. To each leaf of $\Gamma^{k,1}$ attach a copy of $T^k(A)$. This yields a tree $\Gamma^{k,2}$ of uniform depth $(k+2)N_k$, where each leaf of $\Gamma^{k,2}$ corresponds to some leaf of $T^k(A)$ (this is a many-to-one correspondence).
3. Now we mimic the first step of the induction. For each leaf u of $T^k(A)$, choose an integer $n(u)$ such that u has at least $2^{n(u)\alpha}$ descendants at level $|u| + n(u)$, and $n(u) \geq (k^2+k+1)N_k$. This is possible by (3.1). For future reference, we denote by $T^{k,3}$ the subtree of $T(A)$ obtained by appending to every vertex u at level N_k its descendants in the next $n(u)$ levels. To each vertex of $\Gamma^{k,2}$ corresponding to u, append a tree of uniform depth $m(u) = n(u) - (k+1)N_k$, where every vertex in the first $\lceil m(u)(1-\alpha/d)\rceil$ levels has 2^d children, and every vertex in the remaining $\lfloor m(u)\alpha/d\rfloor$ levels has a single child. This defines $\Gamma^{k,3}$.

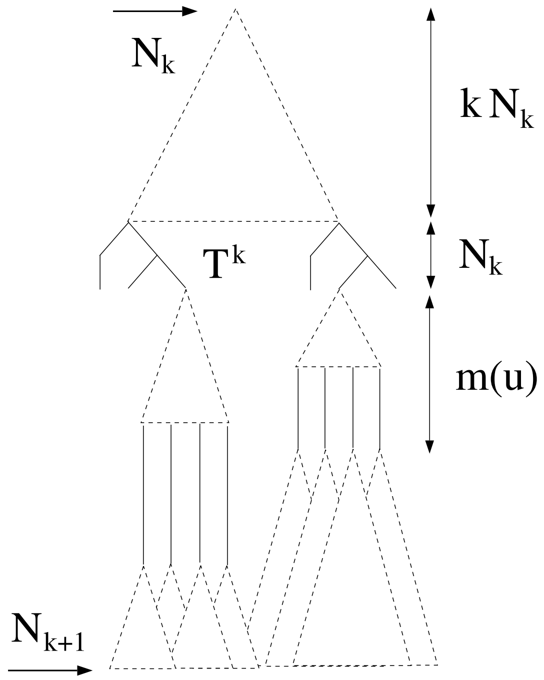

FIGURE 3.2. The four stages of the construction

The leaves of $\Gamma^{k,3}$ form a cut-set Π_{k+1} such that the corresponding β-dimensional cut-set sum satisfies

$$\sum_{w\in\Pi_{k+1}} 2^{-\beta|w|} \le 2^{d(k+1)N_k} \cdot \sum_{u\in\partial T^k(A)} 2^{m(u)(d-\alpha)+d} \cdot 2^{-\beta(N_k+n(u))}, \tag{3.5}$$

where the first factor $2^{d(k+1)N_k}$ is a bound on the number of leaves of $\Gamma^{k,1}$. The number of leaves of $T^k(A)$ is at most 2^{dN_k}. Using this and the inequality
$m(u) \le n(u)$, we see that the cut-set sum (3.5) is bounded by

$$2^{d(k+1)N_k} \cdot \sum_{u\in\partial T^k(A)} 2^d \cdot 2^{n(u)(d-\alpha-\beta)} \le 2^{d(k+2)N_k+d} \cdot 2^{k^2N_k(d-\alpha-\beta)}. \tag{3.6}$$

where the last inequality used $n(u) > k^2N_k$. For any fixed $\beta > d-\alpha$, these bounds tend to zero as $k\to\infty$.

4. Finally, denote by N_{k+1} the maximum depth of $\Gamma^{k,3}$, and to each leaf w of $\Gamma^{k,3}$ append a tree of uniform depth $N_{k+1}-|w|$ with 2^d children at every level. This yields Γ^{k+1}.

The construction is illustrated in Figure 2. The dashed boxes represent subtrees with full growth (2^d children at each vertex).

Now that we have constructed the tree Γ^{k+1} we want to construct a flow μ on the product tree $T^{k+1}(A) * \Gamma^{k+1}$ which extends the flow μ constructed previously on $T^k(A) * \Gamma^k$. (Rather than call the flow constructed at each step μ_k, we simply refer to them all as μ since the definitions are consistent.) It is most convenient to do this in four stages following the construction of the tree.

1. Suppose (v, w) is a leaf of $T^k(A) * \Gamma^k$. By our induction hypothesis the flow satisfies the Hölder estimate
$$\mu(v, w) \leq 2^{-d_k N_k}.$$
Let $T^{k,1}$ be the truncation of $T(A)$ to level $(k+1)N_k$ (so it has the same depth as $\Gamma^{k,1}$). We define the flow on $T^{k,1} * \Gamma^{k,1}$ by dividing the mass assigned by μ to each vertex of the smaller tree $T^k(A) * \Gamma^k$ equally among its descendants at level $(k+1)N_k$ of the extended tree $T^{k,1} * \Gamma^{k,1}$. Since $\Gamma^{k,1}$ has full branching at these levels (i.e., each vertex has 2^d children) it is clear that the resulting flow satisfies
$$\mu(v, w) \leq 2^{-d_k j},$$
for vertices in level j such that $N_k \leq j \leq (k+1)N_k$.
2. *This is the only stage where we do not divide the mass equally among descendants.* Let $T^{k,2}$ be the truncation of $T(A)$ to level $(k+2)N_k$ (so it has the same depth as $\Gamma^{k,2}$). For each leaf (v, w) of $T^{k,1} * \Gamma^{k,1}$, let u be the leaf of $T^k(A)$ which is an ancestor of v. Divide the mass of (v, w) evenly among the leaves (v', w') of $T^{k,2} * \Gamma^{k,2}$ which satisfy: *v' is a descendant of v and w' is the unique descendant of w corresponding to u.*
At worst, we have concentrated all the mass of (v, w) onto a single descendant, so for leaves (v', w') of $T^{k,2} * \Gamma^{k,2}$ we get
$$\mu(v', w') \leq \mu(v, w) \leq 2^{-dkN_k} \leq 2^{-d_{k+1}(k+2)N_k},$$
where we define
$$d_{k+1} := \min\{\frac{dk}{k+2}, \tilde{d}\}\,. \tag{3.7}$$
Thus at level $(k+2)N_k$, the flow μ satisfies a Hölder estimate of order d_{k+1}.
3. Next, we extend the definition of μ to the product tree $T^{k,3} * \Gamma^{k,3}$. The two trees in the product do not have uniform depth. Let v be a leaf of $T^{k,2}$ which is a descendant of a leaf u of T^k, and let w be a leaf of $\Gamma^{k,2}$ which corresponds to u. Divide the mass $\mu(v, w)$ equally among the descendants of (v, w) at level $N_k + n(u)$. Since the flow μ is Hölder of order $d_{k+1} \leq \tilde{d}$ at (v, w), the argument in

step 1 of the induction implies that the same estimate holds for all descendants of (v, w) down to level $N_k + n(u)$.

4. Finally, to define the flow on $T^{k+1} * \Gamma^{k+1}$, use equal division again; since Γ^{k+1} has full branching here, the Hölder estimate of order d_{k+1} established at the previous stage persists.

This completes the inductive step.

The limit of the trees Γ^k is a tree Γ which is associated to a compact set B in the cube $[0,1]^d$. Each cut-set of Γ corresponds to a cover of B by dyadic cubes, so the cut-set sum estimates (3.6) in stage 3 of the inductive step imply that

$$\dim(B) \leq d - \alpha. \tag{3.8}$$

The flow μ extends to the product tree $T(A) * \Gamma$ and corresponds to a measure μ' on $A \times B$ which satisfies the estimate $\mu'(Q) \leq \operatorname{diam}(Q)^{d_k}$ for any dyadic cube Q of side-length at most 2^{-N_k}. Since $d_k = \tilde{d}$ for all large k by (3.7), the mass distribution principle (see, e. g., Proposition 4.2 in Falconer (1990)) implies that

$$\dim(A \times B) \geq \tilde{d}. \tag{3.9}$$

Under our standing assumption $\alpha = \operatorname{Pdim}(A) < d$, we can get $\tilde{d}$ arbitrarily close to d by choosing α close to α and using (3.3). Thus (3.8) and (3.9), combined with Tricot's upper bound (**??**), prove the variational formula (**??**) in this case.

□

4. More about slicing

We now return to a topic we first discussed in Chapter 2; intersecting a set with subspaces of $\mathbb{R}^n$, i.e., slicing the set. Our earlier result (Theorem **??** of Chapter **??**) said that if $A \subset \mathbb{R}^2$ then for almost every $x \in \mathbb{R}$ the intersection of A with the vertical line L_x passing though x has dimension $\leq \dim(A) - 1$. The generalization we consider here is based on the following question: Suppose $F \subset \mathbb{R}^2$. Given a set $A \subset [0,1]$ with $\dim(A) \geq \alpha$ can we find an $x \in A$ with

$$\dim(F_x) \leq \dim(F) - \alpha?$$

We cannot expect to make the right hand side any smaller because of the following example: take sets A, B so that $\dim(B) = \dim_M(B)$. The Marstrand product theorem (Theorem 2.1) says that for $F = A \times B$, $\dim(F_x) = \dim(B) = \dim(F) - \dim(A)$ for all $x \in A$. However, question has an affirmative answer

THEOREM 4.1 (Generalized Marstrand slicing theorem). *Suppose* μ *is a probability measure on* $A \subset [0,1]$ *that satisfies* $\mu(I) \leq C|I|^\alpha$, *for all intervals* $I \subset [0,1]$. *If* $F \subset [0,1]^2$ *has dimension* $\geq \alpha$, *then*

$$dim(F_x) \leq dim(F) - \alpha,$$

for μ *almost every* x.

PROOF. This proof follows the one in Chapter **??** . The idea is to choose $\gamma > \dim(F)$ and show that

$$\int \mathcal{H}^{\gamma-\alpha}(F_x) d\mu(x) < \infty,$$

and hence that $\dim(F_x) \leq \gamma - \alpha$ at μ almost every x. Cover F by squares $\{Q_j\} = \{A_j \times B_j\}$ with $\sum_j |Q_j|^\gamma < \epsilon$. Let

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

Then with μ as in the hypothesis,

$$\int \int_{\mathbb{R}} f(x,y) dy d\mu(x) = \sum_j |Q_j|^{\gamma-1-\alpha} |A_j| \mu(B_j).$$

However, $\mu(B_j) \leq C|B_j|^\gamma \leq C|Q_j|^\alpha$, and $|A_j| \leq |Q_j|$ so

$$\int \int_{\mathbb{R}} f(x,y) dy d\mu(x) \leq C \sum_j |A_j|^\alpha < C\epsilon.$$

Now define

$$Q_j^x = \begin{cases} A_j, & x \in B_j \\ \emptyset, & x \notin B_j \end{cases}.$$

The $\{Q_j^x\}$ form a cover of F_x. By Fubini,

$$\begin{aligned} \int \int_{\mathbb{R}} f(x,y) dy d\mu(x) &= \int [\int_{\mathbb{R}} f(x,y) dy] d\mu(x) \\ &= \int \sum_j |Q_x|^{\gamma-\alpha} d\mu(x) \\ &\geq \int \mathcal{H}^{\gamma-\alpha}_{(\epsilon)}(F_x) d\mu(x). \end{aligned}$$

Thus

$$0 \leq \int \mathcal{H}^{\gamma-\alpha}_{(\epsilon)}(F_x) d\mu(x) \leq C\epsilon.$$

Now let $\epsilon \searrow 0$ and we get

$$0 = \int \mathcal{H}^{\gamma-\alpha}(F_x) d\mu(x),$$

so $\mathcal{H}^{\gamma-\alpha}(F_x) = 0$ μ almost everywhere. Taking limits on γ and α gives the result. $\square$

COROLLARY 4.2. *If $E = \{x : dim(F_x) > dim(F) - \alpha\}$ then $dim(E) \leq \alpha$.*

PROOF. If $\dim(E) > \alpha$ then there is a Frostman measure μ on E with $\mu(E) = 1$ and $\mu(I) \leq C|I|^\alpha$ for all intervals. Then the theorem says that $\dim(F_x) \leq \dim(F) - \alpha$ for μ almost every $x \in E$. This contradiction give the result. $\square$

EXAMPLE 4.3. Back in Chapter 2 we saw that the if G is the Sierpinski gasket and G_x is the vertical slice at x then $G_x = A_S$ where $S = \{x : x_n = 1\}$. Thus

$$\dim(G_x) = \liminf_{n\to\infty} \frac{1}{N} \sum_{n=1}^{N} (1 - x_n),$$

where $\{x_n\}$ is the binary expansion of x. Thus

$$\begin{aligned} \dim(\{x : \dim(G_x) \geq p\}) &= \dim(\{x : \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} x_n \leq 1 - p\}) \\ &= \dim(\tilde{A}_{1-p}), \end{aligned}$$

where $\tilde{A}_p$ is the set defined in Example **??** of Chapter 2. The argument of Section **??**, Chapter **??** shows that the dimension of this set is

$$\dim(\tilde{A}_{1-p}) = \begin{cases} h_2(p), & p \geq 1/2 \\ 1, & p \leq 1/2 \end{cases}.$$

Thus

$$\begin{aligned} &\dim(\{x : \dim(G_x) \geq \dim(G) - \alpha\}) \\ &\qquad = \dim(\{x : \dim(G_x) \geq \log_2 3 - \alpha\}) \\ &\qquad = \begin{cases} 0, & \alpha \leq \log_2 3 - 1 \\ h_2(\log_2 3 - \alpha), & \log_2 3 - 1 \leq \alpha \leq \log_2 3 - 1/2 \\ 1, & \alpha \geq \log_2 3 - 1/2 \end{cases}. \end{aligned}$$

Since $\log_2 3 - \frac{1}{2} \approx 1.08496 > 1$ the third case is consistent with the corollary. A little more work shows

$$h_2(p) \leq \log_2 3 - p,$$

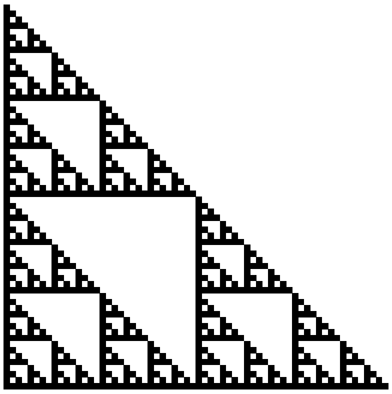

FIGURE 4.1. The Sierpinski gasket (6 generations).

on $[1/2, 1]$, hence taking $p = \log_2 3 - \alpha$,

$$h_2(\log_2 3 - \alpha) \leq \alpha,$$

on $[\log_2 3 - 1, \log_2 3 - 1/2]$, again as required by the corollary.

5. Capacity and dimension.

Motivated by Frostman's lemma we make the following definition.

DEFINITION 5.1. For a measure μ in $\mathbb{R}^d$ and $\alpha > 0$ we define the *α-dimensional energy* of μ to be

$$\mathcal{E}_\alpha(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{d\mu(x)d\mu(y)}{|x-y|^\alpha},$$

and for a set $K \subset \mathbb{R}^d$ we define the *α-dimensional capacity* of K to be

$$\mathrm{Cap}_\alpha(K) = \sup\{\mu(K) : \mu > 0, \mathcal{E}_\alpha(\mu) \leq 1\}.$$

Capacity is not additive, but is monotone and if $K_1 \subset K_2 \subset \dots$, we have

$$\mathrm{Cap}_\alpha(K) = \lim_n \mathrm{Cap}_\alpha(K_n).$$

Furthermore, the union of two (hence countable many) sets of zero capacity also has zero capacity (otherwise the restriction of a finite energy measure to at least one of the subsets would also be a positive, finite energy measure). Note that k has positive α-dimensional capacity if and only if K supports a measure of finite α-dimensional energy. Of course there is an analogous definition when $|x-y|^\alpha$ is replaced by a more general kernel, $\varphi(|x-y|)$.

Pôlya and Szegö defined the critical α for the vanishing of $\mathrm{Cap}_\alpha(K)$ to be the *capacitary dimension* of K, but this name is no longer used because of the following result.

THEOREM 5.2 (Frostman). *Suppose $K \subset \mathbb{R}^d$ is compact.*

1. *If $\mathcal{H}^\alpha(K) > 0$ then $Cap_\beta(K) > 0$ for all $\beta < \alpha$.*
2. *If $Cap_\alpha(K) > 0$ then $\mathcal{H}^\alpha(K) = \infty$ and is not even σ-finite (i.e., can not be written as a countable union of finite measure sets).*

In particular, $dim(K) = \inf\{\alpha : Cap_\alpha(K) = 0\}$.

PROOF. We start with (1). Suppose $\mathcal{H}^\alpha(K) > 0$. Then the α-content is positive, so by Frostman's lemma there is a positive measure μ on K so that

$$\mu(B) \leq C|B|^\alpha,$$

for all balls B. Take a $\beta < \alpha$ and fix a point $x \in K$. Then if $\mathrm{diam}(K) \leq 2^r$,

$$\begin{aligned}
\int_K \frac{d\mu(y)}{|x-y|^\beta} &= \sum_{n=-r}^{\infty} \int_{2^{-n-1} \leq |y-x| \leq 2^{-n}} \frac{d\mu(y)}{|x-y|^\beta} \\
&\leq \sum_{n=-r}^{\infty} \mu(B(x, 2^{-n})) 2^{\beta(n+1)} \\
&\leq C \sum_n 2^{-n\alpha} 2^{n\beta} \\
&\leq M \\
&< \infty
\end{aligned}$$

where C and M are independent of x. Thus $\mathcal{E}_\beta(\mu) \leq M\,|\,\mu\,| < \infty$ and so $\mathrm{Cap}_\beta(K) > 0$.

Now for part (2). Since $\mathrm{Cap}_\alpha(K) > 0$ there is a measure μ on K with $\mathcal{E}_\alpha(\mu) < \infty$, so for sufficiently large M the set

$$K_1 = \{x \in K : \int_K \frac{d\mu(y)}{|x-y|^\alpha} \leq M\},$$

has positive μ measure. One can write this integral as a sum and use summation by parts (e.g., Exercise 9) to get,

$$\begin{aligned}\int_K \frac{d\mu(y)}{|x-y|^\beta} &\geq \sum_{n=-r}^{\infty} \mu(\{2^{-n-1} \leq |y-x| \leq 2^{-n}\})2^{n\alpha} \\ &= \sum_n [\mu(B(x,2^{-n-1})) - \mu(B(x,2^{-n}))]2^{n\alpha} \\ &= C_1 + \sum_n (2^{n\alpha} - 2^{(n-1)\alpha}\mu(B(x,2^{-n})) \\ &= C_1 + C_2 \sum_n 2^{n\alpha}\mu(B(x,2^{-n})).\end{aligned}$$

If $x \in K_1$, then the integral, and thus the sum, is finite, so

$$\lim_{n\to\infty} \frac{\mu(B(x,2^{-n}))}{2^{-n\alpha}} = 0.$$

The mass distribution principle now implies $\mathcal{H}^\alpha(K_1) = \infty$, hence $\mathcal{H}^\alpha(K) = \infty$.

To see that K cannot even be σ-finite, suppose $K = \cup_n K_n$ is a union of sets of finite measure. Then by the argument above, $\mathrm{Cap}_\alpha(K_n) = 0$, which implies $\mathrm{Cap}_\alpha(K) = 0$, a contradiction. □

The capacities defined above with respect to the kernels $K(x,y) = |x-y|^{-\alpha}$ can be generalized to other kernels, just as Hausdorff measures generalize to other gauge functions besides t^α. One kernel that will be very important for our later study of Brownian motion and harmonic measure in $\mathbb{R}^2$ is the logarithmic kernel,

$$K(x,y) = \log \frac{1}{|x-y|}.$$

As before we define the energy of a measure as

$$\mathcal{E}(\mu) = \int\int \log \frac{1}{|x-y|} d\mu(x)d\mu(y),$$

and define

$$\mathrm{Cap}_{\log}(K) = (\sup_{\mu\in \Pr(K)} \mathcal{E}(\mu))^{-1}.$$

This is also called the Robin's constant and is sometimes denoted $\gamma(K)$ or $r(K)$. However, since the logarithmic kernel is not positive this is not a monotone set function. It is sometimes more convenient to deal with

$$R(K) = \sup_{\mu\in \Pr(K)} e^{-\mathcal{E}(\mu)} = \exp(-1/\mathrm{Cap}_{\log}(K)).$$

This is a monotone set function and if K is a disk then $R(K)$ equals its radius. $C(K)$ is called the logarithmic capacity of K and $\mathrm{Cap}_{\log}(K)$ is called the Robin's constant of K and is sometimes denoted γ_K.

Some basic facts about logarithmic capacity are recorded in Appendix A.

6. Marstrand Projection theorem.

Let $K \subset \mathbb{R}^2$ be a compact set and let P_θ be the projection of $\mathbb{R}^2$ onto the line through the origin in the direction $(\cos\theta, \sin\theta)$, for $\theta \in [0, \pi)$. Since P_θ is Lipschitz and $P_\theta K \subset \mathbb{R}$,

$$\dim(P_\theta K) \leq \min\{\dim(K), 1\}.$$

Marstrand's theorem states that for almost all directions the equality holds. We will prove the following version

THEOREM 6.1 (Marstrand, []). *If $Cap_\alpha(K) > 0$ for some $0 < \alpha < 1$ then for almost every θ $Cap_\alpha(P_\theta K) > 0$.*

This has the following corollary

COROLLARY 6.2. *If $dim(K) \leq 1$ then $dim(P_\theta K) = dim(K)$ for almost every θ.*

PROOF. Take $\alpha < \dim(K)$. By Theorem 5.2 $\mathrm{Cap}_\alpha(K) > 0$, so by the theorem $\mathrm{Cap}_\alpha(P_\theta K) > 0$ for almost every θ. By the other part of Frostman's theorem $\dim(P_\theta K) \geq \alpha$ for almost every θ. $\square$

THEOREM 6.1. Let μ be a measure on K with $\mathcal{E}_\alpha(\mu) < \infty$, $\alpha < 1$ and consider the projected measure $\mu_\theta = P_\theta \mu$ on $\mathbb{R}$ defined by

$$\int_{\mathbb{R}} f(x) d\mu_\theta(x) = \int_{\mathbb{R}^2} f(P_\theta(y)) d\mu(y).$$

It's enough to show that μ_θ has finite energy for a.e. θ which will follow from the stronger fact

$$\int_0^\pi \mathcal{E}_\alpha(\mu_\theta) d\theta < \infty.$$

Let's write this integral explicitly as

$$\begin{aligned}
\int_0^\pi \mathcal{E}_\alpha(\mu_\theta)d\theta &= \int_0^\pi \iint \frac{d\mu_\theta(t)d\mu_\theta(s)}{|t-s|^\alpha} \\
&= \int_0^\pi \int_K \int_K \frac{d\mu(x)d\mu(y)}{|P_\theta(x-y)|^\alpha} \\
&= \int_K \int_K \int_0^\pi \frac{d\theta}{|P_\theta(x-y)|^\alpha} d\mu(x)d\mu(y) \\
&= \int_K \int_K \frac{1}{|x-y|^\alpha} \int_0^\pi \frac{d\theta}{|P_\theta(u)|^\alpha} d\mu(x)d\mu(y),
\end{aligned}$$

where $u = (x-y)/|x-y|$ is the unit vector in the direction $(x-y)$. Since the inner integral doesn't depend on u and converges, we can write it as a C_α and pull it out of the integral. Thus

$$\int_0^\pi \mathcal{E}_\alpha(\mu_\theta)d\theta = C_\alpha \int_K \int_K \frac{\mu_\theta(t)d\mu_\theta(s)}{|x-y|^\alpha} = C_\alpha \mathcal{E}_\alpha(\mu) < \infty,$$

by hypothesis. □

If K has dimension 1 then Marstrand's theorem says its projectioonto almost every line has dimension 1, but it need not have positive 1 dimensional measure, even if K does. For example, the standard 1/4 Cantor set in the plane has projections of zero length in almost all directions. However,

THEOREM 6.3 (Kaufman, [**Kau68**]). *If $Cap_1(K) > 0$ then $\mathcal{L}^1(P_\theta K) > 0$ for almost every θ.*

Another version of Theorem 6.3 is due to Mattila (1989). It involves the notion of the Favard length of a set

$$\mathrm{Fav}(K) = \int_0^\pi \mathcal{L}^1(P_\theta K)d\theta.$$

THEOREM 6.4 (Mattila, 1989). *$Fav(K) \geq \pi^2 Cap_1(K)$.*

The constant π^2 is sharp; equality is obtained when K is a disk. No such estimate is true in the opposite direction since a line segment has positive Favard length but zero 1 dimensional capacity (since it has finite 1 dimensional measure). It is not entirely obvious that Mattila's theorem implies Kaufman's since positive Favard length only implies $\mathcal{L}^1(P_\theta K) > 0$ for a positive measure set of directions and not almost all directions. However, as a consequence of the proof we will obtain

LEMMA 6.5. *$\int_0^\pi \mathcal{L}^1(P_\theta(K))^{-1} \leq \mathcal{E}_1(K)$. for every probability measure μ supported on K.*

This implies Kaufman's theorem since the integral would be infinite if $\mathcal{L}^1(P_\theta K) = 0$ on a set of positive measure. The lemma says that $|P_\theta(K)|^{-1} \in L^1(0,\pi)$. It is not known whether it lies in $L^{1+\epsilon}$ for any $\epsilon > 0$.

We will now sketch how Kaufman proved Theorem 6.3 and then give a complete proof of Theorem 6.4. Given a measure μ on K with $\mathcal{E}_1(\mu) < \infty$, Kaufman used Fourier transforms and elementary facts about Bessel's functions to show that

$$\int_0^\pi \int_{\mathbb{R}} |\hat{\mu}_\theta(s)|^2 d\theta < \infty.$$

Thus for a.e. θ $\hat{\mu}$ in is L^2 and so by the Plancherel formula $d\mu_\theta(x) = f(x)dx$ for some L^2 function f. In particular μ is absolutely continuous with respect to Lebesgue measure and so $\mathrm{supp}(\mu_\theta) \subset P_\theta(K)$ must have positive length for a.e. θ.

Before beginning the proof of Mattila's theorem we record a lemma on the differentiation of measures that we will need.

LEMMA 6.6. *Suppose ν is a measure on $\mathbb{R}$ satisfying*

$$\liminf_{\delta \to 0} \frac{\nu(t-\delta, t+\delta)}{2\delta} < \infty,$$

for a.e. $[\nu]$ t. Then $\nu \ll dt$, i.e., ν is absolutely continuous with respect to Lebesgue measure.

PROOF. Let $A \subset \mathbb{R}$ with $\mathcal{L}^1(A) = 0$. We wish to show $\nu(A) = 0$. Given an $\epsilon > 0$ there is a m so that we can write $A = A_m \cup A'_m$ where $\nu(A') < \epsilon$ and

$$\liminf_{\delta \to 0} \frac{\nu(t-\delta, t+\delta)}{2\delta} \leq m,$$

for $x \in A_m$. Since $A_m \subset A$ has zero Lebesgue measure we can take an open set U containing A_m with $\mathcal{L}^1(U) \leq \epsilon/m$. For each $x \in A_m$ take a dyadic interval J_x with $x \in J_x \subset U$ and

$$\frac{\nu(J)}{|J|} \leq 4m.$$

There is such an interval J because x is the center of of an interval I such that $\nu(I) \leq 2m|I|$. If we let J_x be the dyadic interval of maximal length contained in I and containing x then $|J| \geq |I|/2$ and $\nu(J) \leq \nu(I)$, as desired. Thus

$$\nu(A_m) \leq \sum_{x \in A_m} \nu(J_x) \leq 4m \sum_{x \in A_m} |J_x| \leq 4m|U| \leq 4\epsilon.$$

Letting $\epsilon \to 0$ proves $\nu(A) = 0$ as desired. $\square$

THEOREM 6.4. Let $K \subset \mathbb{R}^2$ and μ be a probability measure on K with $\mathcal{E}_1(\mu) < \infty$. Put $\mu_\theta = P_\theta \mu$. Further, let

$$\phi_\theta(t) = \liminf_{\delta \to 0} \frac{\mu_\theta(t-\delta, t+\delta)}{2\delta}.$$

We claim that

$$\int_0^\pi \int_{\mathbb{R}} \phi_\theta(t) d\mu_\theta(t) d\theta \leq \mathcal{E}_1(\mu).$$

To prove this we start with Fatou's lemma which states that

$$\liminf_n (\int f_n d\mu) \geq \int (\liminf f_n) d\mu.$$

This implies that the left hand side is bounded above by

$$\begin{aligned} &\liminf_{\delta \to 0} \frac{1}{2\delta} \int_0^\pi \int_{\mathbb{R}} \mu_\theta(t-\delta, t+\delta) d\mu_\theta(t) d\theta \\ &\quad = \liminf_{\delta \to 0} \frac{1}{2\delta} \int_0^\pi \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbf{1}_{(t-\delta,t+\delta)}(s) d\mu_\theta(s) d\mu_\theta(t) d\theta \end{aligned}$$

Now $\mathbf{1}_{(t-\delta,t+\delta)}(s) = 1$ iff $|s-t| < \delta$ (and is zero otherwise). If $s = P_\theta(x)$ and $t = P_\theta(y)$ then this occurs iff $|P_\theta(x-y)| = |P_theta(x) - P_\theta(y)| < \delta$. Thus the integral becomes,

$$\begin{aligned} &= \liminf_{\delta \to 0} \frac{1}{2\delta} \int_0^\pi \int_K \int_K \mathbf{1}_{P_\theta(x-y)<\delta}(x,y) d\mu(x) d\mu(y) d\theta \\ &= \liminf_{\delta \to 0} \frac{1}{2\delta} \int_K \int_K \int_0^{pi} [\mathbf{1}_{P_\theta(x-y)<\delta}(x,y) d\theta] d\mu(x) d\mu(y). \end{aligned}$$

It is fairly easy to check that given x and y, that $\{\theta : P_\theta(x-y) < \delta\}$ is an interval of length $2\arcsin(\delta/|x-y|)$. Thus the integral bcomes

$$= \liminf_{\delta \to 0} \frac{1}{2\delta} \int_{\mathbb{K}} \int_{\mathbb{K}} 2\arcsin \frac{\delta}{|x-y|} d\mu(x) d\mu(y) \tag{6.1}$$

Note that since $\arcsin(0)$, the definition of deriviative implies

$$\lim_{\delta \to 0} \frac{1}{\delta} \arcsin \frac{\delta}{|x-y|} = \arcsin'(\frac{t}{|x-y|})|_{t=0} = \frac{1}{|x-y|},$$

and if it is easy to check that

$$\sup_{\delta \to 0} \frac{1}{\delta} \arcsin \frac{\delta}{|x-y|} \leq frac\pi|x-y|,$$

which by assumption is in $L^1(\mu)$. Hence we can invoke the Lebegue dominated convergence theorem to bring the limit inside the integral and (6.1 becomes

$$\begin{aligned} &= \int_{\mathbb{K}}\int_{\mathbb{K}} \liminf_{\delta\to 0} \frac{1}{\delta} \arcsin \frac{\delta}{|x-y|} d\mu(x)d\mu(y). \\ &= \int_{\mathbb{K}}\int_{\mathbb{K}} \frac{1}{|x-y|} d\mu(x)d\mu(y). \end{aligned}$$

This is the claim we wished to prove.

Using the claim we now know that for a.e. θ

$$\int_{\mathbb{R}} \phi_\theta(t)d\mu_\theta(t) < \infty,$$

which implies that $\phi_\theta(t) < \infty$ for a.e. $[\mu_\theta]$ t. By the previous lemma this implies μ_θ is absolutely continuous for a.e. θ and that

$$\phi_t = \frac{d\mu_\theta}{dt}.$$

Using this, the claim becomes

$$\int_0^\pi \int_{\mathbb{R}} (\frac{d\mu_\theta}{dt})^2 dt d\theta \leq \mathcal{E}_1(\mu).$$

Now use the Cauchy-Schwarz inequality

$$1 \quad = \quad [\mu_\theta(P_\theta(K))]^2 = [\int \mathbf{1}_{P_\theta(K)} \frac{d\mu_\theta}{dt} dt]^2 \leq \mathcal{L}^1(P_\theta(K)) \int (\frac{d\mu_\theta}{dt})^2 dt.$$

Thus

$$|P_\theta(K)|^{-1} \leq \int (\frac{d\mu_\theta}{dt})^2 dt.$$

Integrating $d\theta$ over $[0,\pi]$ gives

$$\int_0^\pi \mathcal{L}^{-1}(P_\theta(K))^{-1} \leq \mathcal{E}_1(\mu).$$

(Note that this is Lemma 6.5.)

To finish the proof of Mattila's theorem we use the Cauchy-Schwarz inequality again and apply the above inequality

$$\begin{aligned} \pi^2 &= [\int_0^\pi \mathcal{L}^1(P_\theta(K))^{1/2}\mathcal{L}^1(P_\theta(K))^{-1/2}d\theta]^2 \\ &\leq \int_0^\pi \mathcal{L}^1(P_\theta(K))d\theta \int_0^\pi \mathcal{L}^1(P_\theta(\mu))^{-1}d\theta \\ &\leq \operatorname{Fav}(K)\mathcal{E}_1(\mu). \end{aligned}$$

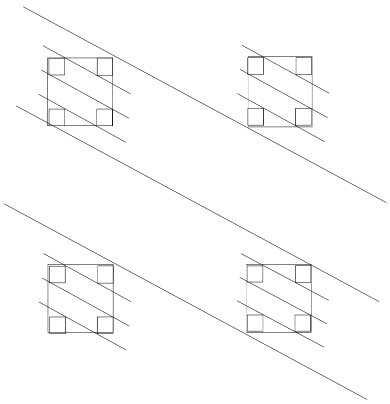

FIGURE 6.1. A large projection of the 1/4 Cantor set.

Thus

$$\mathrm{Fav}(K) \geq \frac{\pi^2}{\mathcal{E}_1(\mu)}.$$

Taking the supremum over all probability measures supported on K gives the result. $\square$

REMARK . Much more can be said about projections of sets K which have finite $\mathcal{H}^1$ measure. By results of Besicovitch such sets can always be written as a union of a regular and iregular set (and a set of zero $\mathcal{H}^1$ measure). A set is regular if it lies on a countable union of rectifiable arc and is irregular if its intersection with any rectifiable arc has zero length. Regular sets of positive $\mathcal{H}^1$ measure project to positive length in almost all directions. Irregular sets project to zero length in almost all directions. See Falconer's book [?] for details. For example, it is easy to use this definition to show that the standard 1/2 Cantor set in the plane is an irrgular set (Exercise 27. Thus it projects to zero length in almost all directions, although Figure 6.1 shows there is a least one direction whose projection contains an interval.

7. Mapping a tree to Euclidean space preserves capacity

We have already seen that it is very useful to think of b-adic cubes in $\mathbb{R}^d$ as forming a tree and associating subsets of $\mathbb{R}^d$ with subtrees of this tree. This is a theme which will also be important later in the book. In this section we record a useful result which allows us to compare the capacity of a subset of $\mathbb{R}^d$ with the capacity of the associated tree.

In order to interpret theorems proved on trees as theorems about sets in Euclidean space, we employ the canonical mapping $\mathcal{R}$ from the boundary of a b^d-ary tree $\Gamma^{(b^d)}$ (every vertex has b^d children) to the cube $[0,1]^d$. Formally, label the edges from each vertex to its children in a one-to-one manner with the vectors in $\Omega = \{0, 1, \dots, b-1\}^d$. Then the boundary $\partial\Gamma^{(b^d)}$ is identified with the sequence space $\Omega^{\mathbf{Z}^+}$ and we define $\mathcal{R} : \Omega^{\mathbf{Z}^+} \to [0,1]^d$ by

$$\mathcal{R}(\omega_1, \omega_2, \dots) = \sum_{n=1}^{\infty} b^{-n}\omega_n. \tag{7.1}$$

Similarly, a vertex σ of $\Gamma^{(b^d)}$ is identified with a finite sequence $(\omega_1, \dots, \omega_k) \in \Omega^k$ if $|\sigma| = k$, and we write $\mathcal{R}(\sigma)$ for the cube of side b^{-k} obtained as the image under $\mathcal{R}$ of all sequences in $\Omega^{\mathbf{Z}^+}$ with prefix $(\omega_1, \dots, \omega_k)$.

The notions of energy and capacity are meaningful on any compact metric space. Given a decreasing function $g : (0,\infty) \to (0,\infty)$ such that $g(0+) = \infty$, define the energy of a Borel measure μ by

$$\mathcal{E}(\mu) = \int\int g(|x-y|)\, d\mu(x)\, d\mu(y)$$

and the capacity of a set Λ by

$$\mathrm{cap}_g(\Lambda) = \left[\inf_{\mu(\Lambda)=1} \mathcal{E}_g(\mu)\right]^{-1}.$$

THEOREM 7.1. *With the notation above, let T be a subtree of the b^d-ary tree $\Gamma^{(b^d)}$, so we may take $\partial T \subseteq \Omega^{\mathbf{Z}^+}$. Given a decreasing function $g : (0,\infty) \to (0,\infty)$ define $f(n) = g(b^{-n})$. Then for any finite measure μ on ∂T we have*

$$\mathcal{E}_f(\mu) < \infty \Leftrightarrow \mathcal{E}_g(\mu\mathcal{R}^{-1}) < \infty \tag{7.2}$$

and in fact the ratio is bounded between positive constants depending only on the dimension d. (Note that $\mathcal{E}_f$ denotes energy on the tree and $\mathcal{E}_g$ is energy in Euclidean space.) It follows that

$$\mathit{cap}_f(T) > 0 \Leftrightarrow \mathit{cap}_g(\mathcal{R}(\partial T)) > 0.$$

Remark 7: For $g(t) = \log(1/t)$ and $f(n) = n \log b$ this is proved in Benjamini and Peres (1992). As noted there, the potentials may become infinite when passing from the tree to Euclidean space.

PROOF: Using summation by parts, the energy on ∂T can be rewritten:

$$\begin{aligned} \mathcal{E}_f(\mu) &= \sum_{n=1}^{\infty} \int\int_{|\xi\wedge\eta|=n} f(n)\, d\mu(\xi)\, d\mu(\eta) \\ &= \sum_{k=0}^{\infty} h(k)(\mu\times\mu)\left\{|\xi\wedge\eta| \ge k\right\}, \end{aligned}$$

where $h(k) = f(k) - f(k-1)$ and by convention $f(-1) = 0$. Two points ξ and η in the boundary of the tree correspond to infinite paths in the tree, and we let $\xi \wedge \eta$ denote the maximum subpath they have in common and $|\xi \wedge \eta|$ the length of this path. Moreover, if σ is a vertex of the tree we write $\xi \geq \sigma$ to denote that σ is on the path corresponding to ξ (i.e., ξ is "below" σ in the tree). Since

$$\{|\xi \wedge \eta| \geq k\} = \bigcup_{|\sigma|=k} \{\xi \wedge \eta \geq \sigma\} = \bigcup_{|\sigma|=k} [\{\xi \geq \sigma\} \cap \{\eta \geq \sigma\}],$$

we obtain

$$\mathcal{E}_f(\mu) = \sum_{k=0}^{\infty} h(k) \sum_{|\sigma|=k} \mu(\sigma)^2 = \sum_{k=0}^{\infty} h(k) S_k \tag{7.3}$$

where $S_k = S_k(\mu) = \sum_{|\sigma|=k} \mu(\sigma)^2$. Now we wish to adapt this calculation to the set $\mathcal{R}(\partial T)$ in the cube $[0,1]^d$. First observe that the same argument yields

$$\begin{aligned} \mathcal{E}_g(\mu\mathcal{R}^{-1}) &\leq \sum_{n=0}^{\infty} g(b^{-n})(\mu\mathcal{R}^{-1} \times \mu\mathcal{R}^{-1}) \left\{(x,y) : b^{-n} < |x-y| \leq b^{1-n}\right\} \\ &= \sum_{k=0}^{\infty} h(k)(\mu\mathcal{R}^{-1} \times \mu\mathcal{R}^{-1}) \left\{(x,y) : |x-y| \leq b^{1-k}\right\} , \qquad (7.4) \end{aligned}$$

where we have implicitly assumed that μ has no atoms (otherwise the energies are automatically infinite).

For vertices σ, τ of T we write $\sigma \sim \tau$ if $\mathcal{R}(\sigma)$ and $\mathcal{R}(\tau)$ intersect (this is not an equivalence relation!). If $x, y \in \mathcal{R}(\partial T)$ satisfy $|x - y| \leq b^{1-k}$ then there exist vertices σ, τ of T with $|\sigma| = |\tau| = k-1$ and $\sigma \sim \tau$ satisfying $x \in \mathcal{R}(\sigma)$ and $y \in \mathcal{R}(\tau)$. Therefore

$$(\mu\mathcal{R}^{-1} \times \mu\mathcal{R}^{-1}) \left\{(x,y) : |x-y| \leq b^{1-k}\right\} \leq \sum_{|\sigma|=|\tau|=k-1} \mathbf{1}_{\{\sigma \sim \tau\}} \mu(\sigma)\mu(\tau).$$

Now use the inequality

$$\mu(\sigma)\mu(\tau) \leq \frac{\mu(\sigma)^2 + \mu(\tau)^2}{2}$$

and the key observation that

$$\#\{\tau \in T : |\tau| = |\sigma| \text{ and } \tau \sim \sigma\} \leq 3^d \quad \text{for all } \sigma \in T$$

to conclude that

$$(\mu\mathcal{R}^{-1} \times \mu\mathcal{R}^{-1})\left\{(x,y) : |x-y|b^{1-k}\right\} \le \frac{1}{2}\Big(\sum_{\sigma}\sum_{\tau\sim\sigma}\mu(\sigma)^2 + \sum_{\tau}\sum_{\sigma\sim\tau}\mu(\tau)^2\Big) \tag{7.5}$$

$$\le \frac{1}{2}\Big(\sum_{\sigma}\mu(\sigma)^2 3^d + \sum_{\tau}\mu(\tau)^2 3^d\Big) \tag{7.6}$$

$$\le 3^d S_{k-1}. \tag{7.7}$$

It is easy to compare S_{k-1} to S_k: clearly $|\sigma| = k-1$ implies that

$$\mu(\sigma)^2 = \left(\sum_{\tau\ge\sigma\,;\,|\tau|=k}\mu(\tau)\right)^2 \le b^d \sum_{\tau\ge\sigma\,;\,|\tau|=k}\mu(\tau)^2$$

and therefore

$$S_{k-1} \le b^d S_k. \tag{7.8}$$

Combining this with (7.3), (7.4) and (7.5) yields

$$\mathcal{E}_g(\mu\mathcal{R}^{-1}) \le (3b)^d \sum_{k=0}^{\infty} h(k) S_k = (3b)^d \mathcal{E}_f(\mu).$$

This proves the direction ($\Rightarrow$) in (7.2).

The other direction is immediate in dimension $d = 1$ and easy in general:

$$\begin{aligned}
\mathcal{E}_g(\mu\mathcal{R}^{-1}) &\ge \sum_{k=0}^{\infty} g(b^{-k})(\mu\mathcal{R}^{-1}\times\mu\mathcal{R}^{-1})\left\{(x,y) : b^{-k-1} < |x-y| \le b^{-k}\right\} \\
&= \sum_{n=0}^{\infty} h(n)(\mu\mathcal{R}^{-1}\times\mu\mathcal{R}^{-1})\left\{(x,y) : |x-y| \le b^{-n}\right\} \\
&\ge \sum_{n=0}^{\infty} h(n) S_{n+l},
\end{aligned}$$

where l is chosen to satisfy $b^l \ge d^{1/2}$ and therefore

$$\left\{(x,y) : |x-y| \le b^{-n}\right\} \supseteq \bigcup_{|\sigma|=n+l} [\mathcal{R}(\sigma)\times\mathcal{R}(\sigma)].$$

Invoking (7.8) we get

$$\mathcal{E}_g(\mu\mathcal{R}^{-1}) \ge b^{-dl}\sum_{n=0}^{\infty} h(n) S_n = b^{-dl}\mathcal{E}_f(\mu)$$

which completes the proof of (7.2).

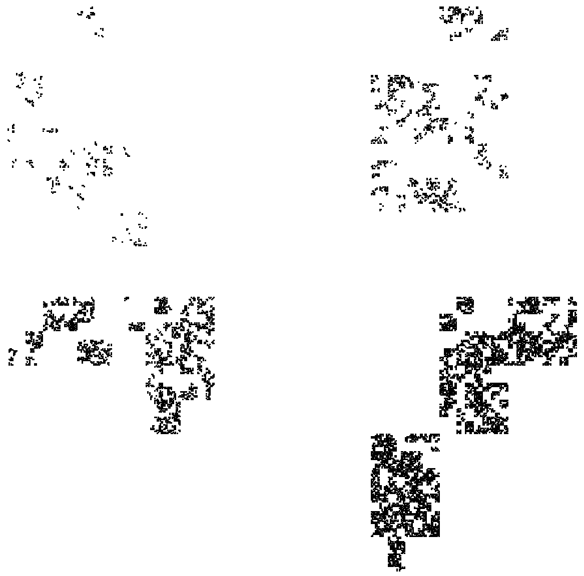

FIGURE 8.1. Random Cantor sets for $p = .5, .6, .7, .8, .9, .99$

The capacity assertion of the theorem follows, since any measure ν on $\mathcal{R}(\partial T) \subseteq [0,1]^d$ can be written as $\mu\mathcal{R}^{-1}$ for an appropriate measure μ on ∂T. $\square$

8. Dimension of random Cantor sets

Fix a number $0 < p < 1$ and define a random set by dividing the unit cube $[0,1]^d$ of $\mathbb{R}^d$ into b^d b-adic cubes in the usual way. Each cube is kept with probability p, and the kept cubes are further subdivided and again each subcube is kept with probability p. The result is a random Cantor set, or more precisely, a measure on the space of compact subsets of $[0,1]^d$ which we will denote μ_p in this section. If a property holds on a set of full μ_p measure we will say it holds almost surely.

A few examples of such sets are illustrated in Figure 8. These were drawn using 8 generations and the values $d = b = 2$ and $p \in \{.5, .6., .7, .8, .9, .99\}$.

THEOREM 8.1. *If $p \leq b^{-d}$ then μ_p gives full measure to the empty set. If $p > b^{-d}$, then μ_p almost every set is either empty or has Hausdorf dimension $d + \log_b p$. Moreover, the latter occurs with positive probability and almost surely the set is dimension regular (i.e., the Minkowski dimension exists and equals the Hausdorff dimension).*

This was first proven by Hawkes in 1981 [**J.H81**]. We will actually give another proof of this later in the book, using Lyon's beautiful theorem on percolation on trees (see Section ??). However, the direct proof is simple and is a nice illustration of the ideas introduced in this chapter.

In terms of the corresponding trees, these random Cantor sets have as simple interpretation. We start with the usual rooted tree $\Gamma = \Gamma^{b^d}$ where each vertex has b^d children and we select edges with probability p and consider the connected component of the chosen edges which contains the root. Our random Cantor sets then correspond to the boundary of these random subtrees. Because of the equivalence of dimension and capacity under the cannonical mapping from trees to sets, it suffices to do all our computations in the tree setting.

The first thing we want to do is compute the expected number of vertices at level n. Given a vertex we can easily compute the probability of having exactly k children as

$$q_k = p^k(1-p)^{N-k}\binom{N}{k},$$

where we set $N = b^d$ for the remainder of this section. Thus the expected number of vertices in the first generation is

$$pN = \sum_{k=0}^{\infty} kq_k.$$

(In our case, this sum is finite since $q_k = 0$ for $k > N$, but many of our calculations are valid for branching processes given by the data $\{q_k\}$ only assuming finite mean and variance.)

Let Γ denote the full N-ary tree and $Q = Q(\Gamma)$ denote the set of rooted subtrees of Γ (with the same root as Γ). This is the space on which we think of μ_p as living. A basis for its topology is given by "cylinder sets" $Q(T)$, determined by fixing a finite rooted T and considering all extensions. The μ_p measure of such a cyliner set is

$$\mu_p(Q(T)) = \prod_{\sigma \in T \setminus \partial T} q_{k(\sigma)},$$

where $k(\sigma)$ is the number of children of the vertex σ. Let $Z_n(T)$ denote the number of nth generation vertices of T. This is a random variable on the probability space (Q, μ_p) and as usual we define its expectation by

$$\mathbb{E}Z_n = \int_Q Z_n(T)\mu_p(T).$$

Also in keeping with probabilistic notation, if $A, B \subset Q$, we write,

$$\mathbb{P}(A) = \mu_p(A),$$

and let

$$\mathbb{P}(A|B) = \mu_p(AcapB)/\mu_p(B),$$

be the probability of A conditioned on B.

LEMMA 8.2. $\mathbb{E}Z_n = m^n$.

PROOF. This is easy by induction. We have already observed $\mathbb{E}Z_1 = m$ and in general,

$$\mathbb{E}Z_n = \sum_{k=0}^{\infty} \mathbb{E}(Z_n|Z_{n-1} = k)\mathbb{P}(Z_{n-1} = k) + \sum_k mk\mathbb{P}(Z_{n-1} = k) = m\mathbb{E}Z_{n-1} = m^n.$$

$\square$

Already, this is enough to give the correct upper estimate for the dimension of our random Cantor sets. Recall that $N(A,\epsilon)$ is the minimum number of ϵ cubes needed to cover A.

LEMMA 8.3. *Suppose A is a random set in $[0,1]^d$ so that*

$$\mathbb{E}N(A, b^{-nd}) \leq Cb^{\alpha n}.$$

Then $\overline{dim_M}(A) \leq \alpha$.

PROOF. Fix $\epsilon > 0$ and note that by the monotone convergence theorem

$$\mathbb{E}\sum_{n=0}^{\infty} \frac{N(A, b^{-nd})}{2^{n(\alpha+\epsilon)}} = \sum_{n=0}^{\infty} \frac{\mathbb{E}N(A, b^{-nd})}{2^{n(\alpha+\epsilon)}} = \sum_{n=0}^{\infty} \frac{C2^{\alpha n}}{2^{n(\alpha+\epsilon)}} < \infty.$$

Thus $\sum_{n=0}^{\infty} \frac{N(A,b^{-nd})}{2^{n(\alpha+\epsilon}}$ converges for almost every A and hence $N(A, b^{-nd}) \leq C2^{n(\alpha+\epsilon)}$ for almost every A. Letting $\epsilon \to$ through a countable sequence gives the lemma. $\square$

Combining the last two lemmas shows that our random Cantor sets has dimension at most $d + \log_b p$, since

$$m^n = b^{n\log_b m} = b^{n\log_b b^d p} = b^{n(d+\log_b p)}.$$

To prove that this is also a lower bound (at least when we condition on the event that the set is non-empty) is harder and will occupy the rest of this section. The main point is to show that for almost every $T \in Q(\Gamma)$, either $Z_n = 0$ for large enough n or $Z_n \geq Cm^n$ for some positive constant (which may depend on T). To do this, define the random variable on Q,

$$W_n = m^{-n}Z_n,$$

The main result we need is that

THEOREM 8.4. *Assume* $\sum_k k^2 q_k < \infty$ *(which is automatic for random Cantor sets). There is a function* $W \in L^2(Q, \mu_p)$ *so that* $W_n \to W$ *pointwise almost everywhere and in* L^2*. Moreover, the sets* $\{T : W(T) = 0\}$ *and* $\{T : Z_n(T) = 0$ *for some* $n\}$ *are the same up to a set of* μ_p *measure zero.*

We will prove this later in the section. For the present, we simply use this result to compute the capacity of our sets using the kernel $f(n) = b^{-\alpha n}$ on the tree. Given a tree T so that $W_n(T)$ converges to $W(T) > 0$, we define a flow on T by

$$W(\sigma) = \lim_{n\to\infty} m^{-n} Z_n(T, \sigma),$$

where $Z_n(T, \sigma)$ denote the number of nth level vertices of T which are descendants of σ. Its easy to check that this is a flow on T with total mass $W(T)$. Moreover, $W(\sigma)$ is a random variable on Q with the same distribution as $m^{-|\sigma|}W$ (assuming σ is in the tree; otherwise it is zero). In particular, $\mathbb{E}W(\sigma)^2 = m^{-2|\sigma|}\mathbb{E}W < \infty$ if $\sigma \in T$ by Theorem 8.4.

According to formula 7.3, the f-energy of this measure is given by

$$\mathcal{E}_f(W) = \sum_{n=0}^{\infty} \sum_{\sigma\in\Gamma_n} [f(n) - f(n-1)] W(\sigma)^2,$$

where Γ_n denotes the nth level vertices of the full tree Γ. Thus the expected enegry is

$$\begin{aligned}
\mathbb{E}\mathcal{E}_f(\nu) &= \sum_{n=0}^{\infty} \sum_{\sigma\in\Gamma_n} [f(n) - f(n-1)]\mathbb{E}(W(\sigma)^2) \\
&= \sum_{n=0}^{\infty} \sum_{\sigma\in\Gamma_n} [f(n) - f(n-1)] m^{-2n} \mathbb{E}(W^2)\, \mathbb{P}(\sigma \in T)
\end{aligned}$$

Now the probability that $\sigma \in T$ is the probability that all the edges connecting it to the root are in T. Each edge is chosen independently with probability p, so this probability is p^n. Thus

$$\begin{aligned}\mathbb{E}\mathcal{E}_f(\nu) &= \mathbb{E}(W^2)\sum_{n=0}^{\infty}\sum_{\sigma\in\Gamma_n}(b^{\alpha n} - b^{\alpha(n-1)})m^{-2n}p^n \\ &= \mathbb{E}(W^2)(1-b^{-\alpha})\sum_{n=0}^{\infty} b^{\alpha n}N^n m^{-2n}p^n \\ &= \mathbb{E}(W^2)(1-b^{-\alpha})\sum_{n=0}^{\infty} b^{\alpha n}N^n (Np)^{-2n}p^n \\ &= \mathbb{E}(W^2)(1-b^{-\alpha})\sum_{n=0}^{\infty} b^{n(\alpha-\log_b m)}.\end{aligned}$$

This sum is clearly finite if $\alpha < \log_b m = d + \log_b p$. Thus if $W(T) > 0$, the corresponding set has positive α-capacity for every $\alpha < d + \log_b p$ and hence dimension $\geq d + \log_b p$. This completes the proof of Theorem 8.1 except for the proof of Theorem 8.4.

THEOREM 8.4. Our first goal is to show

$$\mathbb{E}(|W_n - W_{n+1}|^2) = ||W_n - W_{n+1}||_2^2 < Cm^{-n}$$

Clearly,

$$\mathbb{E}(|W_n - W_{n+1}|^2) = m^{-2n-2}\mathbb{E}(|mZ_n - Z_{n+1}|^2)$$

so it is enough to show

$$\mathbb{E}(|Z_{n+1} - mZ_n|^2) \leq Cm^n \tag{8.1}$$

To prove this, we will condition on Z_n and write $Z_{n+1} - mZ_n$ as a sum of Z_n orthogonal functions $\{X_j\}$ on the space Q of random trees. To do this we will actually condition on the first n levels of the tree, i.e., fix some finite n-level tree T with a fixed number Z_n of nth-level leaves and let $Q(T)$ be the associated cylinder set in Q (i.e., all the infinite trees which agree with T in the first n levels). Let

$$\mathbb{E}_T(X) = \frac{1}{\mu_p(Q(T))}\int_{Q(T)} X\,d\mu_p,$$

denote the expectation of x over $Q(T)$. We will order the Z_n leaves of T in some way and let

$$X_j = [\sum_k (k-m)\mathbf{1}_{Q(j,k)}],$$

where $\mathbf{1}_X$ denotes the indicator function of X and $Q(j,k)$ is the set of trees in $Q(T)$ so that the jth leaf of T has k children. Note that $\mu_p(Q(j,k)) = q_k\mu(Q(T))$.

Note that $\sum_{j=1}^{Z_n} X_j$ for a particular tree $T' \in Q(T)$ is the total number of $(n+1)$st level vertices in T' minus mZ_n. Thus

$$Z_{n+1}(T') - mZ_n(T') = \sum_j X_j(T').$$

We next claim that the mean value of X_j is zero.

$$\mathbb{E}_T(X_j) \;=\; \sum_k (k-m)q_k = 0,$$

since $\sum_k kq_k = m$. Next we want to see that X_i and X_j are orthogonal if $i \neq j$. Now choose $i < j$. The sets $Q(i,l)$ and $Q(j,k)$ are independent, so

$$\mu(Q(i,l) \cap Q(j,k)) = \mu(Q(i,l)\mu(Q(j,k))$$

and hence X_i and X_j are independent. Thus $\mathbb{E}(X_iX_j) = \mathbb{E}(X_i)\mathbb{E}(X_j) = 0$. Finally we need the L^2 norm of the X_j's to be bounded uniformly. Note,

$$\mathbb{E}_T(|X_j|^2) \;=\; \sum_k |k-m|^2 q_k \equiv \sigma^2,$$

which we assume is bounded.

Thus if we take expected values over $Q(T)$ we get

$$\mathbb{E}(|Z_{n+1} - mZ_n|^2) = \mathbb{E}(|\sum_{j=1}^{Z_n} X_j|^2) = \sum_{j=1}^{Z_n} \mathbb{E}(|X_j|^2) = \sum_{j=1}^{Z_n} \sigma^2 = \sigma^2 Z_n(T),$$

as desired. Finally we have to write Q as a finite disjoint union of sets of the form $Q(T)$. Since the expected value over $Q(T)$ is $\sigma^2 Z_n(T)$ and the expected value (over all nth level trees T) of $Z_n(T)$ is m^n, we get that

$$\mathbb{E}(|Z_{n+1} - mZ_n|^2) \sum_T \mathbb{E}_T(|Z_{n+1} - mZ_n|^2) = \sigma^2 \sum_T Z_n \mathbb{P}(T) = \sigma^2 m^n,$$

as desired.

We have now proven that $|\, W_n - W_{n-1} \,|_2 \leq Cm^{-n/2}$. This easily implies that the sequence $\{W_n\}$ is Cauchy in L^2 and hence converges in the L^2 norm to a function W. Moreover, $W \geq 0$ since the same is true for every W_n. Moreover,

$$\mathbb{E}W = \lim_n \mathbb{E}W_n = \lim_n m^{-n}\mathbb{E}Z_n = 1,$$

and so $\{T : W(T) > 0\}$ must have positive measure.

To see that W_n converges pointwise almost everywhere to W, note that using the Cauchy-Schwarz inequality we get

$$\begin{aligned}\mathbb{E}\sum_n |W_n - W_{n-1}| &\leq \sum_n \mathbb{E}|W_n - W_{n-1}| \\ &\leq \sum_n ||W_n - W_{n-1}||_2 \\ &\leq \sum_n Cm^{-n/2},\end{aligned}$$

which converges if $m > 1$. Since the left hand side must be finite almost everywhere and hence the series converges absolutely for almost every T. Thus $W_n \to W$ almost everywhere, as desired.

Since $Z_n(T) = 0$ for some n implies $W(T) = 0$, the final claim of the theorem will follow is we show the two sets have the same measure. This can be done by computing each of them using a generating function argument. Let

$$f(x) = \sum_{k=0}^{\infty} q_k x^k,$$

be the generating function assoicated to the sequence $\{q_k\}$ (as noted before, in the case of random Cantor sets this is a finite sum and f is a polynomial, but the argument that follows is more general). Since f is a positive sum of the increasing, convex up functions $\{x_k\}$, it is also increasing and convex up on $[0, 1]$. Furthermore, $f(1) = 1$, $f(0) = q_0$ and $f'(1) = \sum_k kq_k = m$. If $m > 1$ this means that f has a unique fixed point $x_0 < 1$ which is zero if q_0 and is $q_0 < x_0 < 1$ if $q_0 > 0$. The plot in Figure 8 shows the generating functions for $d = b = 2$ and the values $p = .75$ (supercritical), $p = .25$ (critical) and $p = .1$ (subcritical).

We claim that $\mathbb{P}(Z_n = 0)$ and $\mathbb{P}(W = 0)$ can be computed in terms of the generating function f. Let $\{q_{k,j}\}$ be the probability of $Z_1 = j$ given that $Z_0 = k$. Each of the k vertices has children with distribution $\{q_k\}$ independent of the other $k - 1$ vertices. Thus the probability of choosing exactly j overall is the coefficient of x^j in

$$(q_0 + q_1 x + q_2 x^2 + \dots)^k,$$

or in other words, the generating function of $\{q_{k,j}\}$ is

$$\sum_j q_{k,j} s^j = [f(s)]^k.$$

Now let $q^n_{k,j} = \mathbb{P}(Z_n = j | Z_0 = k)$ be the n-step transition probabilities, What is F_n, the corresponding generating function? We claim that

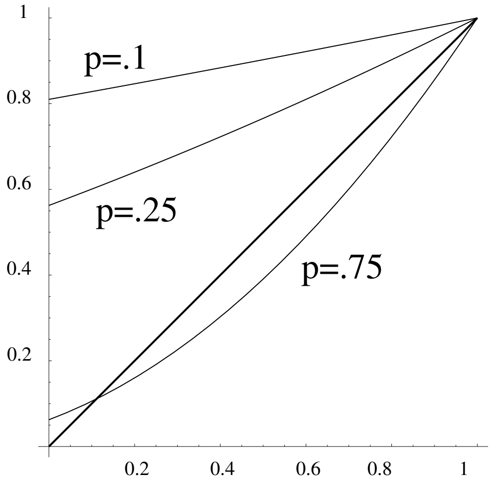

FIGURE 8.2. Generating functions for $p = .1, .25, .75$

$F_n = f^{(n)}$ where $f^{(n)} = f \circ \cdots \circ f$ denotes the n fold composition of f with itself. We prove this by induction, the case $n =$ being obvious. Computing $q_{1,j}^{n+1}$ from $q_{k,j}^n$ gives

$$q_{1,j}^{n+1} = \sum_k q_{1,k}^n q_{k,j},$$

so in terms of generating functions,

$$\sum_j q_{1,j}^{n+1} x^j = \sum_k \sum_j q_{1,j}^n q_{k,j} x^j = \sum_k q_{1,k}^n [f(x)]^j = F_n(f(x)) = f^{(n)}(f(x)) = f^{(n+1)}(x),$$

as desired.

Thus $\mathbb{P}(Z_n = 0) = f^{(n)}(0)$. Since f is a strictly increasing function on $[0,1]$, the iterates of 0 increase to the first fixed point of f. If $m > 1$ this is the point x_0 described above. If $m \leq 1$, then the first fixed point is 1. Thus if $m > 1$, Z_n remains positive with positive probability and if $m \leq 1$ it is eventially equal to zero with probability equal to 1.

To compute the probability of the set $\{W = 0\}$, let $r = \mathbb{P}(W = 0)$ and note that

$$r = \sum_k \mathbb{P}(W = 0 | Z_1 = k)\mathbb{P}(Z_1 = k) = \sum_k r^k q_k = f(r).$$

Thus r is a fixed point of f. If $m > 1$, then we saw above that $\mathbb{E}W = 1$, so we must have $r < 1$ and so $r = x_0 = \mathbb{P}(Z_n = 0 \text{ for some } n)$, as desired. $\square$

REMARK . The arguments of this section prove that given a kernel f that almost surely our random Cantor sets have positive capacity with respect to f iff

$$\sum_{n=1}^{\infty} m^{-n} f(n) < \infty \tag{8.2}$$

The set of full measure depends on f however. Pemantle and Peres proved that there is a set of full μ_p measure such that ∂T has positive capacity for the kernel f iff (8.2) holds (with the set independent of the kernel). This is given in [**?**].

REMARK . The sets $Q_d(b,p)$ were described by Mandelbrot in [**Man74**] and analyized by Kahane and Peyrier [**KP76**]. The exact measure of more general random sets was found by Graf, nauldin and Williams [**GMW88**]. Chayes, Chayes and Durrett [**CCD88**] noted that the dimension could be easily infereed from the work in [**KP76**] and proved the remarkable result that $Q_d(2,p)$ contains connected components for p close enough to 1.

Exercises

EXERCISE 1. Give an example of sets $E, F \subset \mathbb{R}$ with

$$\dim(E) + \dim(F) < \dim(E \times F) < \dim(E) + \dim(F) + 1.$$

EXERCISE 2. Show that we can take $\epsilon = 0$ in Theorem 3.1. (Hint: let $B = \cup_n B_n$ where B_n corresponds to $\epsilon = 1/n$ and the B_n's are chosen to have a compact union.)

EXERCISE 3. Prove that $\mathrm{Pdim}(E \times F) \geq \underline{\mathrm{Pdim}}(E) + \mathrm{Pdim}(F)$. This is from [**?**] and improves a result of Tricot [**?**]

EXERCISE 4. Show that strict inequality is possible in the previous problem, i.e., there is a compact E so that

$$\mathrm{Pdim}(E \times F) - \mathrm{Pdim}(F) > \underline{\mathrm{Pdim}}(E),$$

for every compact F. See [**?**]. (Hint: take E_i so that the corresponding tree has 1 child at levels $(4k+2i) \leq n < (4k+2i+i)$ and 2 children otherwise.)

EXERCISE 5. Use the result of Joyce and Preiss to prove Theorem 3.1 is true for all analytic sets.

EXERCISE 6. If E is closed and $\overline{\dim}_M(E \cap V) \geq \alpha$ for every open set such that $E \cap V \neq \emptyset$, prove $\mathrm{Pdim}(E) \geq \alpha$.

EXERCISE 7. Show that there are compact sets $A \subset \mathbb{R}^d$ so that $\dim(A \times B) = \mathrm{Pdim}(A) + \dim(B)$ iff $\dim(B) = d - \mathrm{Pdim}(A)$.

EXERCISE 8. Show that $\dim(A) \leq \underline{\mathrm{Pdim}}(A)$.

EXERCISE 9. Prove Abel's summation by parts formula,

$$\sum_{k=1}^{n} a_k b_k = A_n b_n + \sum_{k=1}^{n-1} A_k (b_k - b_{k+1}),$$

where $A_k = sum_{j=1}^{k} a_k$.

EXERCISE 10. Construct a set $E \subset \mathbb{R}^2$ so that for every $0 \leq \alpha \leq 1$ there is a direction such that $\dim(P_\theta(E)) = \alpha$.

EXERCISE 11. Let $E = E_{1/3} \times E_{1/3}$. Show that for almost every direction θ, $P_\theta(E)$ contains an interval.

EXERCISE 12. Show $|P_\theta(E)| = 0$ for almost every θ, where $E = E_{1/4} \times E_{1/4}$?

EXERCISE 13. If the projection of $E \subset \mathbb{R}^2$ in direction θ contains an interval must its projection also contain an interval for all directions close enough to θ? For any directions other than θ?

EXERCISE 14. Suppose $E \subset \mathbb{R}^2$ and $\dim(E) \geq 1$. For $0 < \alpha \leq \dim(E)$, what can we say about

$$\dim(\{\theta : \dim(P_\theta(E)) = \dim(E) - \alpha\})?$$

EXERCISE 15. Parameterize the set of lines in the plane by (r, θ) where r is the distance of a line L to the origin and θ is the normal direction to the line. Define a measure on the set of lines by $d\mu = dr d\theta$. Show this measure is invariant under rigid motions of the plane.

EXERCISE 16. Using the measure in the previous exercise, show that the measure of the set of lines which hit a convex set K in the plane is equal to the length of the boundary of K.

EXERCISE 17. Use the previous exercise to solve Buffon's problem, i.e., show that if a needle of length $d < 1$ is dropped at random onto a collection of parallel lines distance 1 apart, the probability it hits one of the lines is $2d/\pi$.

EXERCISE 18. Find the minimal cutset for the tree in Figure **??**.

EXERCISE 19. Construct an infinite tree with a legal flow so that no cutset attains the minimum cut.

EXERCISE 20. Give an example to show that α capacity is not additive.

EXERCISE 21. Show that α capacity is monotone, i.e., if $K_1 \subset K_2 \subset \dots$, then $\cap_\alpha(\cup_n K_n) = \lim_n \cap_\alpha(K_n)$.

EXERCISE 22. Prove $\mathrm{Cap}_\alpha(E \cup F) \leq \mathrm{Cap}_\alpha(E) + \mathrm{Cap}_\alpha(F)$.

EXERCISE 23. Prove that E has σ-finite $\mathcal{H}^\varphi$ measure if and only if $\mathcal{H}^\psi(E) = 0$ for every ψ such that

$$\lim_{t\to 0} \frac{\psi(t)}{\varphi(t)} = 0.$$

This is due to Besicovitch [].

EXERCISE 24. Given a gauge function φ, construct a set E so that $0 < \mathcal{H}^\varphi(E) < \infty$.

EXERCISE 25. Construct a Jordan curve of σ-finite $\mathcal{H}^1$ measure, but which has no subarc of finite length.

EXERCISE 26. Show the π in Theorem 6.4 is sharp.

EXERCISE 27. Show that the 1/4 Cantor set in the plane is irregular, i.e., its intersection with any rectifiable curve has zero length. (Hint: consider a family of disjoint square aanuli which surround each generational square in the standard construction. A curve hitting E must cross many of these annuli and hence have infinite length.)

EXERCISE 28. Modify thf of Lemma 8.3 to show that almost surely $\mathcal{H}^\alpha(A) < \infty$.

EXERCISE 29. Suppose we generate random subsets of $[0, 1]$ by choosing dyadic subintervals with probability p. Compute the probability of getting the empty set (as a function of p).

EXERCISE 30. Use generating functions to show that for $m \neq 1$,

$$\mathrm{var} Z_n = \mathbb{E}(|Z_n - \mathbb{E}Z_n|^2) = \sigma^2 m^{n-1}(m^n - 1)/(m - 1),$$

where

$$\sigma^2 = \mathrm{var} Z_1 = \sum_k (k - m)^2 q_k.$$

If $m = 1$ show the varieance of Z_n is $n\sigma^2$. (See [**?**].) (Hint: the variance of Z_n can be given in terms of $f_n'(1)$ and $f_n''(1)$, which can be computed by induction and the chain rule.)

EXERCISE 31. If $m < 1$, show that $\mathbb{P}(Z_n > 0)$ decays like $O(m^n)$.

EXERCISE 32. If $m = 1$, estimate $\mathbb{P}(Z_n > 0)$ in terms of $f''(1)$.

EXERCISE 33. Suppose we define a random subset of $[0, 1]$ iteratively by dividing an interval I into two disjoint subintervals with lengths chosen uniformly from $[0, \frac{1}{2}|I|]$. What is the dimension of the resulting set almost surely?

FIGURE 8.3. Set in Exercise 38

EXERCISE 34. (**) Suppose A is a random set in $\mathbb{R}^2$ generated with b-adic squares chosen with probability p (for brevity we shall call this a (b,p) random set in $\mathbb{R}^2$). What is the almost sure dimension of the vertical projection of A onto the real axis?

EXERCISE 35. If $p > 1/b$ show that the vertical projection has positive length almost surely (if it is non empty).

EXERCISE 36. (**) If $p > 1/b$ does the vertical projection contain an interval almost surely (if it is non empty)?

EXERCISE 37. With A as above, what is the almost sure dimension of an intersection with a random horizontal line?

EXERCISE 38. What is the value of p for the random set shown in Figure 38?

EXERCISE 39. Suppose A_1, A_2 are random sets generated from b-adic cubes in $\mathbb{R}^d$ with probabilities p_1 and p_2. Give necessary and sufficient conditions on p_1 and p_2 for $A_1 \cap A_2 \neq \emptyset$ with positive probability.

EXERCISE 40. Suppose $K \subset \mathbb{R}$ is a fixed compact set such which hits the random $(2,p)$-set with positive probability. Prove that $\dim(K) \geq -\log_2 p$. (Hint: define a measure on K by $\mu(I) = \mathbb{P}(I \cap K \cap A \neq \emptyset)$.)

EXERCISE 41. Prove that a (b,p) random set is totally disconnected almost surely if $p < 1/b$.

EXERCISE 42. Prove the same if $p = 1/b$.

EXERCISE 43. (**) Is there an $\epsilon > 0$ so that the set is almost surely disconnected if $p < (1+\epsilon)/b$?

EXERCISE 44. Consider random sets contructed by prelacing p by a sequence $\{p_n\}$ (so that edges at level n are chosen indenpently with probobility p_n. Compute the dimension in terms of $\{p_n\}$.

EXERCISE 45. With sets as in the previous exercise, characterize the sequences which give sets of positive Lebesgue measure in $\mathbb{R}^d$.

EXERCISE 46. Suppose A is a random Cantor set with $Np = m$. Show that there is a set of full μ_p measure so that A has zero capacity for any kernel f such that $\sum_{n=1}^{\infty} m^{-n} = \infty$. See [?]. (Hint: take the set of T such that $\sup_n m^{-n} Z_n > 0$.)

EXERCISE 47. (**) Suppose we make random self-affine sets by using n by m grids and choosing children independently with probability p. What is the Minkowski dimension of the resulting sets? What is the Hausdorff dimension?

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.