

THE POWER LAW
FOR THE BUFFON NEEDLE PROBABILITY
OF THE FOUR-CORNER CANTOR SET

F. NAZAROV, Y. PERES, AND A. VOLBERG

*Dedicated to Victor Petrovich Havin
on the occasion of his 75th birthday*

ABSTRACT. Let \mathcal{C}_n be the n th generation in the construction of the middle-half Cantor set. The Cartesian square \mathcal{K}_n of \mathcal{C}_n consists of 4^n squares of side length 4^{-n} . The chance that a long needle thrown at random in the unit square will meet \mathcal{K}_n is essentially the average length of the projections of \mathcal{K}_n , also known as the Favard length of \mathcal{K}_n . A classical theorem of Besicovitch implies that the Favard length of \mathcal{K}_n tends to zero. It is still an open problem to determine its exact rate of decay. Until recently, the only explicit upper bound was $\exp(-c \log_* n)$, due to Peres and Solomyak ($\log_* n$ is the number of times one needs to take the log to obtain a number less than 1, starting from n). In the paper, a power law bound is obtained by combining analytic and combinatorial ideas.

§1. INTRODUCTION

The four-corner Cantor set \mathcal{K} is constructed by replacing the unit square by four subsquares of side length $1/4$ at its corners, and iterating this operation in a self-similar manner in each subsquare. More formally, consider the set \mathcal{C}_n that is the union of 2^n segments:

$$\mathcal{C}_n = \bigcup_{\substack{a_j \in \{0,3\} \\ j=1,\dots,n}} \left[\sum_{j=1}^n a_j 4^{-j}, \sum_{j=1}^n a_j 4^{-j} + 4^{-n} \right],$$

and let the *middle-half Cantor set* be

$$\mathcal{C} := \bigcap_{n=1}^{\infty} \mathcal{C}_n.$$

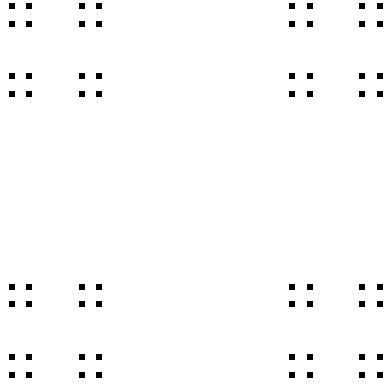
It can also be written as $\mathcal{C} = \left\{ \sum_{n=1}^{\infty} a_n 4^{-n} : a_n \in \{0, 3\} \right\}$. The four-corner Cantor set \mathcal{K} is the Cartesian square $\mathcal{C} \times \mathcal{C}$.

Since the one-dimensional Hausdorff measure of \mathcal{K} satisfies $0 < \mathcal{H}^1(\mathcal{K}) < \infty$ and the projections of \mathcal{K} in two distinct directions have zero length, a theorem of Besicovitch (see [3, Theorem 6.13]) shows that the projection of \mathcal{K} to almost every line through the origin has zero length. This is equivalent to saying that the Favard length of \mathcal{K} equals

2010 *Mathematics Subject Classification.* Primary 28A80; Secondary 28A75, 60D05, 28A78.

Key words and phrases. Favard length, four-corner Cantor set, Buffon's needle.

The research of the authors was supported in part by NSF grants DMS-0501067 (Nazarov and Volberg) and DMS-0605166 (Peres).

FIGURE 1. \mathcal{K}_3 , the third stage of the construction of \mathcal{K} .

zero. Recall (see [1, p. 357]) that the *Favard length* of a planar set E is defined by

$$(1.1) \quad \text{Fav}(E) = \frac{1}{\pi} \int_0^\pi |\text{Proj } \mathcal{R}_\theta E| d\theta,$$

where Proj denotes the orthogonal projection from \mathbb{R}^2 to the horizontal axis, \mathcal{R}_θ is the counterclockwise rotation by angle θ , and $|A|$ denotes the Lebesgue measure of a measurable set $A \subset \mathbb{R}$. The Favard length of a set E in the unit square has a probabilistic interpretation: up to a constant factor, it is the probability that “Buffon’s needle”, a long line segment dropped at random, hits E (more precisely, suppose the needle’s length is infinite, pick its direction uniformly at random, and then locate the needle in a uniformly chosen position in that direction, at distance at most $\sqrt{2}$ from the center of the unit square).

The set $\mathcal{K}_n = \mathcal{C}_n^2$ is a union of 4^n squares with side length 4^{-n} (see Figure 1 for a picture of \mathcal{K}_3). By the dominated convergence theorem, $\text{Fav}(\mathcal{K}) = 0$ implies that

$$\lim_{n \rightarrow \infty} \text{Fav}(\mathcal{K}_n) = 0.$$

We are interested in good estimates for $\text{Fav}(\mathcal{K}_n)$ as $n \rightarrow \infty$. A lower bound $\text{Fav}(\mathcal{K}_n) \geq \frac{c}{n}$ for some $c > 0$ follows from Mattila [8, 1.4]. Peres and Solomyak [10] proved that

$$\text{Fav}(\mathcal{K}_n) \leq C \exp[-a \log_* n] \quad \text{for all } n \in \mathbb{N},$$

where

$$\log_* n = \min \left\{ k \geq 0 : \underbrace{\log \log \dots \log}_k n \leq 1 \right\}.$$

This result can be viewed as an attempt to make a quantitative statement out of a qualitative Besicovitch projection theorem [1, 11], using this canonical example of the Besicovitch irregular set.

It is very interesting to see what are quantitative analogs of the Besicovitch theorem in general. The reader can find more of that in [11].

Now we state our main result, which improves this upper bound to a power law.

Theorem 1. *For every $\delta > 0$, there exists $C > 0$ such that*

$$\text{Fav}(\mathcal{K}_n) \leq C n^{\delta-1/6} \quad \text{for all } n \in \mathbb{N}.$$

Remarks.

- The $1/6$ in the exponent is certainly not optimal, and, indeed, can be improved slightly with the methods of this paper. However, a bound decaying faster than $O(n^{-1/4})$ would require new ideas.
- In [10, Theorem 2.2] a random analog of the Cantor set \mathcal{K} was analyzed, and it was shown that, with high probability, the Favard length of the n th stage in the construction has upper and lower bounds that are constant multiples of n^{-1} . However, it is not clear to us whether $\text{Fav}(\mathcal{K}_n)$ also decays at this rate.
- From the results of Kenyon [5] and Lagarias and Wang [6] it follows that $|\text{Proj } \mathcal{R}_\theta \mathcal{K}| = 0$ for all θ such that $\tan \theta$ is irrational. As was noted in [10], this information does not seem to help obtain an upper bound for $\text{Fav}(\mathcal{K}_n)$.
- The set \mathcal{K} was one of the first examples of sets of positive length and zero analytic capacity; see [2] for a survey. The asymptotic behavior of the analytic capacity of \mathcal{K}_n was determined in 2003 by Mateu, Tolsa, and Verdera [7]; it is equivalent to $\frac{1}{\sqrt{n}}$.

It will be convenient to translate \mathcal{K}_n so that its convex hull will be the unit square centered at the origin. Due to the symmetries of the square, we can average over $\theta \in (0, \frac{\pi}{4})$ in the definition (1.1) of $\text{Fav}(\mathcal{K}_n)$. After translation, the projection of $\mathcal{R}_\theta(\mathcal{K}_n - (\frac{1}{2}, \frac{1}{2}))$ to the horizontal axis is the union of 4^n intervals of length $4^{-n}(\cos \theta + \sin \theta)$ centered at the points $\sum_{k=0}^{n-1} 4^{-k} \xi_k$, where

$$\xi_k \in \left\{ \pm \frac{3\sqrt{2}}{8} \cos\left(\frac{\pi}{4} - \theta\right), \pm \frac{3\sqrt{2}}{8} \sin\left(\frac{\pi}{4} - \theta\right) \right\}.$$

Now, let $t = \tan(\frac{\pi}{4} - \theta) \in [0, 1]$. Since $\frac{\sqrt{2}}{2} \leq \cos(\frac{\pi}{4} - \theta) \leq 1$ on $(0, \frac{\pi}{4})$, the length of the projection $\text{Proj } \mathcal{R}_\theta(\mathcal{K}_n)$ is comparable to the length of the union of 4^n intervals of length $4^{-n}\rho$ centered at the points $\sum_{k=0}^{n-1} 4^{-k} \xi_k$ with $\xi_k \in \{\pm 1, \pm t\}$, where $\rho = \rho(\theta) = \frac{8}{3\sqrt{2}}(1 + \tan(\frac{\pi}{4} - \theta))$. The exact value of $\rho(\theta)$ is of no importance; the only thing that matters is that it is separated away from both 0 and $+\infty$. We shall also need the function f_n that is the product of $\frac{1}{\rho}$ and the sum of the characteristic functions of these intervals. In other words,

$$f_n = \nu^{(n)} * \frac{4^n}{\rho} \chi_{[-\frac{\rho}{2}4^{-n}, \frac{\rho}{2}4^{-n}]},$$

where

$$\nu^{(n)} = *_{k=0}^{n-1} \nu_k \quad \text{and} \quad \nu_k = \frac{1}{4} [\delta_{-4^{-k}} + \delta_{-4^{-k}t} + \delta_{4^{-k}t} + \delta_{4^{-k}}].$$

Geometrically, up to minor rescaling, f_n is the number of squares whose projections contain a given point. Finally, since $|\frac{dt}{d\theta}| = \frac{1}{\cos^2(\frac{\pi}{4} - \theta)}$ is between 1 and 2 for all $\theta \in [0, \frac{\pi}{4})$, we can replace averaging over θ with that over t .

§2. FOURIER-ANALYTIC PART

In what follows, we will use \asymp and \lesssim, \gtrsim to denote, respectively, equality or the corresponding inequality up to some positive multiplicative constant. Let K, S be large positive numbers. Our first aim is to show that there exists a power $p > 0$ (we shall see that any $p > 4$ fits) such that the measure of the set

$$E = \left\{ t \in [0, 1] : \max_{1 \leq n \leq (KS)^p} \int_{\mathbb{R}} f_n^2 \leq K \right\}$$

is at most $\frac{1}{5}$. Suppose not. Let N be the smallest even integer exceeding $\frac{1}{2}(KS)^p$. For every $t \in E$, we must have

$$K \geq \int_{\mathbb{R}} |f_N(x)|^2 dx \asymp \int_{\mathbb{R}} |\widehat{f}_N(y)|^2 dy \gtrsim \int_1^{4^{N/2}} |\widehat{\nu}^{(N)}(y)|^2 dy,$$

because $\psi = \frac{4^N}{\rho} \chi_{[-\frac{\rho}{2}4^{-N}, \frac{\rho}{2}4^{-N}]}$ satisfies $\widehat{\psi}(y) \gtrsim 1$ for all $|y| < 4^{N/2}$ if N is sufficiently large. Thus,

$$\frac{1}{|E|} \int_E \left[\sum_{n=1}^{N/2} \int_{4^{n-1}}^{4^n} |\widehat{\nu}^{(N)}(y)|^2 dy \right] dt \leq K,$$

and for each $m \leq N/2$, there exists $n \leq N/2$ satisfying

$$\frac{1}{|E|} \int_E \left[\int_{4^{n-m}}^{4^n} |\widehat{\nu}^{(N)}(y)|^2 dy \right] dt \leq \frac{4Km}{N}.$$

Thus,

$$E_* = \left\{ t \in E : \int_{4^{n-m}}^{4^n} |\widehat{\nu}^{(N)}(y)|^2 dy \leq \frac{8Km}{N} \right\}$$

satisfies $|E_*| \geq |E|/2$. Our assumption on E implies that $|E_*| \geq \frac{1}{25}$. Now for $y \in [4^{n-m}, 4^n]$, we have

$$|\widehat{\nu}^{(N)}(y)|^2 \asymp \prod_{k=0}^n \left| \frac{\cos 4^{-k}y + \cos 4^{-k}ty}{2} \right|^2,$$

because the remaining terms (that correspond to $k \in [n+1, N]$) in the product converge geometrically to 1. Making the change of variable $y \mapsto 4^n y$, we get

$$\int_{4^{n-m}}^{4^n} |\widehat{\nu}^{(N)}(y)|^2 dy \asymp 4^n \int_{4^{-m}}^1 \left| \prod_{k=0}^n \frac{\cos 4^k y + \cos 4^k ty}{2} \right|^2 dy.$$

Now we split the last product into

$$P_1(y) = \prod_{k=0}^m \frac{\cos 4^k y + \cos 4^k ty}{2} \quad \text{and} \quad P_2(y) = \prod_{k=m+1}^n \frac{\cos 4^k y + \cos 4^k ty}{2}.$$

First, consider the integral

$$\int_{4^{-m}}^1 |P_2(y)|^2 dy.$$

Writing the cosines as sums of exponentials, we have

$$P_2(y) = 4^{m-n} \sum_{j=1}^{4^{n-m}} e^{i\lambda_j y},$$

where $\{\lambda_j\}_{j=1}^{4^{n-m}}$ are the sums of all subsets of $\{\pm 4^k, \pm 4^k t : k \in [m+1, n]\}$. For $t \in E_* \subset E$, the definition of E shows that

$$\int_{\mathbb{R}} \left(\sum_j \chi_{[\lambda_j - \frac{\rho}{2}4^m, \lambda_j + \frac{\rho}{2}4^m]} \right)^2 \leq K \cdot 4^n$$

(this is equivalent to $\int_{\mathbb{R}} f_{n-m}^2 \leq K$). The last inequality can be viewed as a separation condition on the spectrum, so one can hope that a variation of Salem's trick should allow us to conclude that

$$\int_{4^{-m}}^1 |P_2(y)|^2 dy \gtrsim 4^{m-n},$$

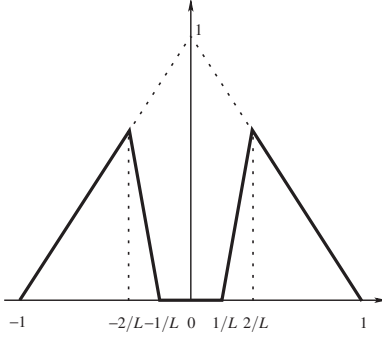


FIGURE 2. Triangle kernel function.

provided that $L = 4^m$ is chosen appropriately. We shall choose m such that $4^m = L$ is a large constant multiple of K . Since $|P_2(-y)| = |P_2(y)|$, we can integrate over $[-1, 1] \setminus [-L^{-1}, L^{-1}]$. Consider the function g given by

$$g(y) = (1 - |y|)_+ - 2(1 - L^{-1})(1 - \frac{L}{2}|y|)_+ + (1 - 2L^{-1})(1 - L|y|)_+.$$

Note that g is even, $0 \leq g \leq 1$, $\text{supp } g \subset [-1, 1] \setminus [-L^{-1}, L^{-1}]$, and $\int_{-1}^1 g \geq \frac{1}{2}$ if L is not too small. Now, let h denote “the triangle function” that is 1 at 0, vanishes on $\mathbb{R} \setminus (-1, 1)$, and is linear on $[-1, 0]$ and on $[0, 1]$. Then

$$g(y) = h(y) - 2(1 - L^{-1})h(\frac{L}{2}y) + (1 - 2L^{-1})h(Ly).$$

Since $\widehat{h}(\lambda) = 2 \frac{1 - \cos \lambda}{\lambda^2} \in [0, \frac{C}{1 + \lambda^2}]$, we obtain

$$\widehat{g}(\lambda) \geq -\frac{C}{L} \cdot \frac{1}{1 + (\lambda/L)^2}.$$

So, we have obtained the estimate

$$\widehat{g}(\lambda) \geq -C \frac{L}{\lambda^2 + L^2}$$

with some numerical constant C .

Denote $M = 4^{n-m}$. Let us call $k \in \{1, \dots, M\}$ *good* if

$$C \sum_j \frac{L}{L^2 + (\lambda_j - \lambda_k)^2} \leq \frac{1}{8}.$$

Then

$$\begin{aligned} \int_{[-1, 1] \setminus [-L^{-1}, L^{-1}]} \left| \sum_k e^{i\lambda_k y} \right|^2 dy &\geq \int_{\mathbb{R}} g(y) \left| \sum_k e^{i\lambda_k y} \right|^2 dy \\ &\geq \sum_{\{k : k \text{ is good}\}} \frac{1}{2} + \int_{\mathbb{R}} g(y) \left| \sum_{\{k : k \text{ is bad}\}} e^{i\lambda_k y} \right|^2 dy - 2 \sum_{\{k : k \text{ is good}\}} C \sum_j \frac{L}{L^2 + (\lambda_j - \lambda_k)^2} \\ &\geq \frac{1}{4} \#\{k : k \text{ is good}\}. \end{aligned}$$

Now we only need to show that the number of good indices is comparable to M . For this, note that we have the condition

$$\int_{\mathbb{R}} \left(\sum_j \chi_{[\lambda_j - \frac{\theta}{2}L, \lambda_j + \frac{\theta}{2}L]}(\lambda) \right)^2 d\lambda \leq MLK.$$

Convolving with the Poisson kernel $\mathcal{P}_L(\lambda) = \frac{1}{\pi} \frac{L}{L^2 + \lambda^2}$ and taking into account the inequality

$$\chi_{[\lambda_j - \frac{\rho}{2}L, \lambda_j + \frac{\rho}{2}L]} * \mathcal{P}_L \geq cL\mathcal{P}_L(\cdot - \lambda_j)$$

with $c > 0$ (here we use the fact that ρ stays bounded away from 0 and $+\infty$), we get

$$L^2 \int_{\mathbb{R}} \left[\sum_j \mathcal{P}_L(\lambda - \lambda_j) \right]^2 d\lambda \leq C' MLK,$$

but

$$\int_{\mathbb{R}} \mathcal{P}_L(\lambda - \lambda_j) \mathcal{P}_L(\lambda - \lambda_k) d\lambda \geq c' \mathcal{P}_L(\lambda_j - \lambda_k).$$

Thus,

$$c' \sum_{j,k} \mathcal{P}_L(\lambda_j - \lambda_k) \leq C' MKL^{-1}$$

and

$$\#\{k : k \text{ is bad}\} \leq \frac{8C\pi C'}{c'} (KL^{-1})M \leq \frac{M}{2},$$

provided that $L \geq \frac{16C\pi C'}{c'} K$. Therefore, indeed, $\int_{4^{-m}}^1 |P_2(y)|^2 dy \gtrsim 4^{m-n}$.

The danger is that this large integral can be completely killed when the integrand is multiplied by $|P_1|^2$. Note that

$$\frac{\cos 4^k y + \cos 4^k ty}{2} = \cos 2^{-1} 4^k (y + ty) \cos 2^{-1} 4^k (y - ty),$$

so that

$$P_1(y) = \prod_{k=0}^m \cos 2^{-1} 4^k (y + ty) \cos 2^{-1} 4^k (y - ty).$$

Using the formula

$$2 \cdot 4^m \sin\left(\frac{u}{2}\right) \prod_{\ell=0}^{2m} \cos 2^{\ell-1} u = \sin 4^m u,$$

we conclude that

$$|P_1(y)| \gtrsim 4^{-2m} |\sin 4^m (y + ty)| \cdot |\sin 4^m (y - ty)|.$$

This can be small only if $\sin 4^m (y + ty)$ or $\sin 4^m (y - ty)$ is small. For $\delta \in (0, 1)$, denote by \mathcal{I}_δ the union of intervals of length $4^{-m}\delta$ centered at the points $\frac{\pi\ell}{4^m}$, $\ell \in \mathbb{Z}$. Define $\omega(t; \delta)$ by

$$\omega(t; \delta) = \{y \in (4^{-m}, 1) : y + ty \in \mathcal{I}_\delta \text{ or } y - ty \in \mathcal{I}_\delta\}.$$

We would like to estimate $\int_{\omega(t; \delta)} |P_2(y)|^2 dy$ from above. This may be a hard task for an individual $t \in E_*$, but we can bound the average fairly easily. We have

$$\begin{aligned} \frac{1}{|E_*|} \int_{E_*} \left(\int_{\omega(t; \delta)} |P_2(y)|^2 dy \right) dt &\leq 2S \int_0^1 \left(\int_{\omega(t; \delta)} |P_2(y)|^2 dy \right) dt \\ &\lesssim 2S \int_{[4^{-m}, 1] \cap \mathcal{I}_\delta} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k u \right) \frac{du}{u+v} \int_{[0, 1]} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k v \right) dv \\ &\quad + 2S \int_{[4^{-m}, 1]} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k u \right) \frac{du}{u+v} \int_{[0, 1] \cap \mathcal{I}_\delta} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k v \right) dv, \end{aligned}$$

where $u = y + ty$ and $v = y - ty$.

Using the formula $\cos^2 \alpha = \frac{1}{2}(1 + \cos 2\alpha)$ and the inequality $\frac{1}{u+v} \leq L$, we can estimate the last expression by

$$\mathbf{E} := CSL \cdot 4^{m-n} \left[\int_{\mathcal{I}} \prod_{k=m+1}^n (1 + \cos 4^k u) du \right] \cdot \left[\int_{[0,1]} \prod_{k=m+1}^n (1 + \cos 4^k v) dv \right].$$

Above we observed that

$$\prod_{k=m+1}^n (\cos^2 2^{-1} 4^k u) = 2^{m-n} \prod_{k=m+1}^n (1 + \cos 4^k u) =: 2^{m-n} R(u).$$

Now we want to show that

$$\mathbf{E} \leq cSL \cdot 4^{m-n} \sqrt{\delta}.$$

For this, we notice that the Riesz product $R(u)$ is $\frac{\pi}{4^m}$ -periodic.

Note also that, for any interval \mathcal{J} of length $4^{-m} \frac{\pi}{4^j}$ ($j \in \mathbb{Z}_+$), we have

$$\int_{\mathcal{J}} R(u) du = \int_{\mathcal{J}} R_1(u) R_2(u) du,$$

where $R_1(u) = \prod_{k=m+1}^{m+j} (1 + \cos 4^k u)$ and $R_2(u) = \prod_{k=m+j+1}^n (1 + \cos 4^k u)$. Observe that $R_1(u) \leq 2^j$ for all u and $R_2(u)$ is $\frac{\pi}{4^{m+j}}$ -periodic, whence

$$\int_{\mathcal{J}} R_2(u) du = \frac{1}{4^{m+j}} \int_0^\pi R_2(u) du = \frac{\pi}{4^{m+j}}.$$

Thus, $\int_{\mathcal{J}} R(u) du \leq \frac{\pi}{2^j} 4^{-m}$. We choose j so that $\delta \asymp 4^{-j}$.

It follows that, for each constituent interval \mathcal{J} of \mathcal{I}_δ , we have $\int_{\mathcal{J}} R(u) du \lesssim 4^{-m} \sqrt{\delta}$ and

$$\int_{[4^{-m}, 1] \cap \mathcal{I}_\delta} R(u) du \lesssim 4^m \cdot 4^{-m} \sqrt{\delta} \lesssim \sqrt{\delta}.$$

In conjunction with the estimate $\int_{[0,1] \cap \mathcal{I}_\eta} R(v) dv \lesssim 1$, finally we get

$$\mathbf{E} \leq cSL \cdot 4^{m-n} \sqrt{\delta}.$$

The resulting estimate is much less than 4^{m-n} if δ is much less than $S^{-2}L^{-2}$. Therefore, for at least one $t \in E$ we must have (recall that $L = 4^m$)

$$\int_{[L^{-1}, 1] \setminus \Omega(t)} |P_2(y)|^2 dy \geq c4^{m-n},$$

and thereby (if we remember that K was a small constant times 4^m),

$$\int_{[L^{-1}, 1]} |P_1(y)|^2 |P_2(y)|^2 dy \geq 4^{-4m} (S^{-2}L^{-2})^4 \cdot 4^{m-n} \geq cS^{-8}K^{-11}4^{-n}.$$

Thus, if $p > 12$, then our choice of N at the beginning of the proof gives $N > (KS)^{12+\varepsilon}/2 > (KS)^{12+\varepsilon/2}$; hence $\frac{2K \log K}{N}$ is much less than $S^{-8}K^{-11}$, and we get a contradiction.

However, we promised to show that $N > (KS)^{4+\varepsilon}$ already leads to a contradiction. To do this, we make our considerations more elaborate, but follow the same lines. In fact, let us consider

$$\Omega(t; \delta, \eta) = \{y \in (4^{-m}, 1) : y + ty \in \mathcal{I}_\delta \text{ and } y - ty \in \mathcal{I}_\eta\}.$$

We changed the word ‘‘or’’ in the definition of $\omega(t; \delta)$ by the word ‘‘and’’ in the definition of $\Omega(t; \delta, \eta)$. This will allow us to obtain a subtler estimate. Notice that

$$\{y : |\sin 4^m(y + ty)| \cdot |\sin 4^m(y - ty)| \leq 2^{-l}\} \subset \bigcup_{k=0}^{\ell} \Omega(t; 2^{-k}, 2^{-\ell+k+1}).$$

We would like to estimate $\int_{\Omega(t;\delta,\eta)} |P_2(y)|^2 dy$ from above. As before, we have

$$\begin{aligned} \frac{1}{|E_*|} \int_{E_*} \left(\int_{\omega(t;\delta)} |P_2(y)|^2 dy \right) dt &\leq 2S \int_0^1 \left(\int_{\Omega(t;\delta,\eta)} |P_2(y)|^2 dy \right) dt \\ &\lesssim 2S \int_{[4^{-m},1] \cap \mathcal{I}_\delta} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k u \right) \frac{du}{u} \int_{[0,1] \cap \mathcal{I}_\eta} \left(\prod_{k=m+1}^n \cos^2 2^{-1} 4^k v \right) dv, \end{aligned}$$

where $u = y + ty$ and $v = y - ty$ as before.

We already introduced $R(u) = \prod_{k=m+1}^n (1 + \cos 4^k u)$ and established the estimate

$$\int_{[0,1] \cap \mathcal{I}_\eta} R(v) dv \lesssim \sqrt{\eta}.$$

Now we can estimate

$$\int_{[4^{-m},1] \cap \mathcal{I}_\delta} R(u) \frac{du}{u} \lesssim \sum_{1 \leq j \leq \frac{1}{\pi} 4^m} \frac{4^m}{\pi j} \cdot 4^{-m} \sqrt{\delta} \leq \sqrt{\delta} m.$$

Therefore,

$$\frac{1}{|E_*|} \int_{E_*} \left(\int_{\Omega(t;\delta,\eta)} |P_2(y)|^2 dy \right) dt \lesssim S m \sqrt{\delta \eta} 4^{m-n}.$$

Denote $\Omega_\ell(t) := \bigcup_{k=0}^\ell \Omega(t; 2^{-k}, 2^{-\ell+k+1})$. Now we know that

$$\frac{1}{|E_*|} \int_{E_*} \left(\int_{\Omega_\ell(t)} |P_2(y)|^2 dy \right) dt \lesssim S m \ell \cdot 2^{-\ell/2} 4^{m-n}.$$

If $S m \ell \cdot 2^{-\ell/2}$ is a small constant (much less than 1), it follows that there exists $t \in E_*$ such that

$$\int_{[4^{-m},1] \setminus \Omega_\ell(t)} |P_2(y)|^2 dy \geq c \cdot 4^{m-n}.$$

But $\Omega_\ell(t)$ contains $\{y : |\sin 4^m(y + ty)| \cdot |\sin 4^m(y - ty)| \leq 2^{-\ell}\}$. This means that $|P_1| \gtrsim 4^{-2m} 2^{-\ell}$ on $(4^{-m}, 1) \setminus \Omega_\ell$. Thus, for the same t , we have

$$\int_{4^{-m}}^1 |P_2(y)|^2 dy \geq 4^{-4m} 2^{-2\ell} 4^{m-n}.$$

If $\frac{K m}{N}$ is much less than $4^{-4m} 2^{-2\ell} 4^m$, we get a contradiction. Since $4^m \asymp K$, we see that we get a contradiction if it is possible to find ℓ such that $S m \ell \cdot 2^{-\ell/2}$ is much less than 1 and N is much greater than $K^4 m 2^{2\ell}$ simultaneously. If $N > (KS)^{4+\gamma}$ with $\gamma > 0$, we can take $2^{\ell/2} \asymp (SK)^{\frac{\gamma}{8}} S$, thus finishing the proof of our claim with any $p > 4$.

§3. COMBINATORIAL PART

Fix the rotation angle θ and some large positive integer N . As before, let $F_n(x)$ be the number of the squares in $\mathcal{R}_\theta \mathcal{K}_n$ whose projections to the horizontal axis contain x . Define

$$F_*(x) = \max_{0 \leq n \leq N} F_n(x).$$

Our key observation is the following inequality: for any positive integers K, M , we have

$$\mu\{F_* \geq 4KM\} \leq 108K \mu\{F_* \geq K\} \mu\{F_* \geq M\},$$

where μ denotes the usual Lebesgue measure on the real line.

Proof. For each point $x \in \mathbb{R}$ where $F_*(x) \geq 2K$, choose the smallest $n = n(x)$ for which $F_n(x) \geq 2K$ and mark all the squares in $\mathcal{R}_\theta \mathcal{K}_n$ whose projections contain x . Note that the number of such squares for a given point x cannot exceed $4K$; otherwise, we would have $F_{n-1}(x) \geq 2K$, which contradicts our choice of n . Now unmark all marked squares that are contained in larger marked squares and consider the family of the remaining maximal marked squares. The desired inequality is immediately implied by the following two claims.

Claim 1. In order to reach the level $4KM$ at x , we have to reach the level M in at least one maximal marked square whose projection contains x .

Claim 2. The sum of the side lengths of all maximal marked squares does not exceed $108K\mu\{F_* \geq K\}$.

Proof of Claim 1. Obviously, in order to reach the level $4KM$ at x , one must reach the level M in at least one of the squares of generation $n(x)$ whose projection contains x (recall that there are at most $4K$ such squares!). Each such square is contained in some maximal marked square, whence the claim. \square

Proof of Claim 2. Consider all 4-adic intervals $I \subset \mathbb{R}$ such that I intersects a projection of some maximal marked square Q whose side length is at least $|I|$. Clearly, the union of all such intervals contains the projections of all maximal marked squares. Now consider the maximal intervals I with this property. Clearly, each such maximal interval I intersects the projection of some maximal marked square Q with side length $|I|$, but intersects no projection of a maximal marked square with a larger side length. \square

Now we estimate the sum of the side lengths of the maximal marked squares intersecting one such maximal 4-adic interval I . Let $\sigma = \sin \theta + \cos \theta$. Note that each maximal marked square whose projection intersects I is contained in some square of generation $\log_4 \frac{1}{|I|}$ with side length $|I|$ whose projection intersects I . Since the projection of each such square is contained in $(2\sigma + 1)I$, having more than $\frac{2\sigma+1}{\sigma} \cdot 4K$ such squares would imply the existence of a point $x \in I$ that is contained in more than $4K$ projections of squares of generation $\log_4 \frac{1}{|I|}$. But this implies that there are at least $2K$ squares of the previous generation above x , so $n(x) \leq \log_4 \frac{1}{|I|} - 1$ and there exists a marked square of side length greater than $|I|$ whose projection intersects I . The maximal marked square containing it has at least the same side length and its projection still intersects I . But this contradicts the maximality of I .

Since the maximal marked squares are disjoint, the sum of the side lengths of maximal marked squares contained in one square of generation $\log_4 \frac{1}{|I|}$ does not exceed $|I|$. Hence, the sum of the side lengths of all maximal marked squares whose projections intersect I is at most $\frac{2\sigma+1}{\sigma} \cdot 4K|I| \leq 12K|I|$. Thus, the total sum of the side lengths of all maximal marked squares is at most $12K \sum_I |I| = 12K|U_I I|$, because the maximal intervals are disjoint.

Now, let I be one of our maximal intervals, and let Q_1 be a maximal marked square with side length $|I|$ whose projection intersects I . Since Q_1 is a marked square, there exists a point x and $2K - 1$ other squares Q_2, \dots, Q_{2K} of side length $|I|$ such that $\text{Proj } Q_j \ni x$ for all $j = 1, \dots, 2K$. Choosing K such squares whose centers lie on one side of x , we see that there exists an interval J of length $\sigma/2$ containing x such that $F_* \geq F_{\log_4 \frac{1}{|I|}} \geq K$ on J . Since $\text{dist}(x, I) \leq \sigma|I|$, we have $I \subset \frac{5\sigma+4}{\sigma} J \subset 9J$. Hence, if J' is the constituent interval of the set $\{F_* \geq K\}$, containing J , we also have $I \subset 9J'$. Therefore, $|\bigcup_I I| \leq 9\mu\{F_* \geq K\}$ and we are done.

Now fix θ , K , and N . Let $\nu = \mu\{F_* \geq K\}$. By induction, we get

$$\mu\{F_* \geq (4K)^j K\} \leq [108K\nu]^j \nu, \quad j = 1, 2, \dots$$

Hence, for all $n = 0, 1, \dots, N$, we have

$$\begin{aligned} \int_{\mathbb{R}} f_n^2 &= \int_{\{f_n \leq K\}} f_n^2 + \int_{\{K \leq f_n \leq 4K^2\}} f_n^2 + \sum_{j \geq 1} \int_{\{(4K)^j K \leq f_n \leq (4K)^{j+1} K\}} f_n^2 \\ &\leq \sqrt{2}K + 16K^4\nu + \sum_{j \geq 1} [108K\nu]^j (4K)^{2j} 16K^4\nu \leq 2K, \end{aligned}$$

provided that $108 \cdot 16K^3\nu \leq \frac{1}{2}$. The Fourier-analytic part implies that the measure of all angles θ with this property is less than some absolute constant times $\frac{K}{N^{1/4-\gamma}}$ with arbitrarily small $\gamma > 0$.

Assume that $\nu > 32^{-1} \cdot 108^{-1} \cdot K^{-3}$. For each point $x \in \mathbb{R}$ where $F_*(x) \geq K$, we choose some $n \in \{0, 1, \dots, N\}$ for which $F_n(x) \geq K$ and mark all squares of the n th generation whose projections contain x . Now, in the N th generation, we color green all squares contained in the marked squares. Let φ be the sum of the characteristic functions of projections of green squares, and let Ξ be the union of the projections of all marked squares. First, we want to show that

$$\Xi \subset \left\{ y \in \mathbb{R} : \mathcal{M}\varphi(y) \geq \frac{K}{4} \right\},$$

where $\mathcal{M}\varphi(y) = \sup_{r>0} \frac{1}{2r} \int_{y-r}^{y+r} \varphi(s) ds$ is the central Hardy–Littlewood maximal function. Indeed, if $y \in \Xi$, then the vertical line through y intersects at least one marked square Q_1 . Thus, there exists $x \in \mathbb{R}$ and $K-1$ other marked squares Q_2, \dots, Q_K of the same size as Q_1 such that $x \in \text{Proj } Q_j$ for all $j = 1, \dots, K$. Now, the interval J centered at y of length $4|\text{Proj } Q_1|$ contains all the projections of the squares Q_1, \dots, Q_K . Then the integral $\int_J \varphi$ is not less than the sum of all lengths of the projections of the green squares contained in Q_1, \dots, Q_k , which is $K|\text{Proj } Q_1|$. Hence, $\mathcal{M}\varphi(y) \geq \frac{1}{|J|} \int_J \varphi \geq \frac{K}{4}$. Using the weak type L^1 estimate for the Hardy–Littlewood maximal function, we conclude that

$$\mu(\Xi) \lesssim \frac{1}{K} \int_{\mathbb{R}} \varphi.$$

Since $F_*(x) \geq K$ implies $x \in \Xi$, we deduce that $\int_{\mathbb{R}} \varphi \gtrsim K\nu \gtrsim K^{-2}$, i.e., that there are at least $cK^{-2}4^N$ green squares.

On the other hand, Ξ contains the projections of all green squares, and, up to a constant factor, $\int_{\mathbb{R}} \varphi$ is the sum of all side lengths of all green squares. Thus, the length of the projection of the union of all green squares is at most $\frac{C}{K}$ times the sum of their side lengths. The net outcome of the previous construction is that in the N th generation rotated Cantor square $\mathcal{R}_\theta \mathcal{K}_N$, we have $\mathcal{U} \gtrsim K^{-2}4^N$ green squares, whose projections overlap a lot (more precisely, their total projection is only about $\frac{1}{K}$ times their total side length), and $4^N - \mathcal{U}$ other (white) squares about which we know nothing. This only gives the estimate $\frac{C}{K} \mathcal{U} \cdot 4^{-N} + \sqrt{2} \cdot 4^{-N} (4^N - \mathcal{U})$ for the total length of the projection of $\mathcal{R}_\theta \mathcal{K}_N$, which doesn't look very impressive. But here is where the self-similarity comes into play.

Let us repeat the construction of green squares in each of the white squares (this will bring us to the consideration of \mathcal{K}_{2N} instead of \mathcal{K}_N). Now we shall have $\mathcal{U} \cdot 4^N$ squares contained in the original green squares, which still give us the projections $\leq \frac{C}{K} \mathcal{U} \cdot 4^{-N}$, but we shall also have $(4^N - \mathcal{U})\mathcal{U}$ new small green squares and the total length of their projection will be not greater than $\frac{C}{K} \mathcal{U} (4^N - \mathcal{U}) \cdot 4^{-2N}$. Therefore, the total length of the projections of all these squares will be at most $\frac{C}{K} \mathcal{U} \cdot 4^{-N} [1 + (1 - \frac{\mathcal{U}}{4^N})]$. To this we

should add $\sqrt{2} \cdot 4^{-2N} (4^N - \mathcal{U})^2 = \sqrt{2} (1 - \frac{\mathcal{U}}{4^N})^2$, which is the trivial upper bound for the total projection of the remaining $(4^N - \mathcal{U})^2$ squares in $\mathcal{R}_\theta \mathcal{K}_{2N}$. Proceeding to \mathcal{K}_{3N} in a similar manner, we shall get

$$|\text{Proj } \mathcal{R}_\theta \mathcal{K}_{3N}| \leq \frac{C}{K} \mathcal{U} \cdot 4^{-N} \left[1 + \left(1 - \frac{\mathcal{U}}{4^N}\right) + \left(1 - \frac{\mathcal{U}}{4^N}\right)^2 \right] + \sqrt{2} \left(1 - \frac{\mathcal{U}}{4^N}\right)^3,$$

and so on. By the time we reach $\mathcal{R}_\theta \mathcal{K}_{XN}$ with a large positive integer X , we shall get

$$|\text{Proj } \mathcal{R}_\theta \mathcal{K}_{XN}| \leq \frac{C}{K} \frac{\mathcal{U}}{4^N} \sum_{\ell=0}^{X-1} \left(1 - \frac{\mathcal{U}}{4^N}\right)^\ell + \sqrt{2} \left(1 - \frac{\mathcal{U}}{4^N}\right)^X.$$

The first term does not exceed $\frac{C}{K} \frac{\mathcal{U}}{4^N} \sum_{\ell=0}^{\infty} \left(1 - \frac{\mathcal{U}}{4^N}\right)^\ell = \frac{C}{K}$, while the second is at most $\sqrt{2} e^{-4^{-N} \mathcal{U} X}$, which is less than $\frac{\sqrt{2}}{K}$ if $4^{-N} \mathcal{U} X > \log K$, i.e., if X is much greater than $K^2 \log K$. The moral of the story is that, given two positive integers K and N , we can find an exceptional set of measure $\lesssim \frac{K}{N^{1/4-\gamma}}$ such that for all θ outside this set we have $|\text{Proj } \mathcal{R}_\theta \mathcal{K}_{XN}| \leq \frac{1}{K}$ for all integers X that are much greater than $K^2 \log K$.

The last result can be restated as follows: If K, S are sufficiently large and $N \geq K^p S^q$ with $p > 6, q > 4$, then

$$\mu \left\{ \theta \in \left(0, \frac{\pi}{4}\right) : |\text{Proj } \mathcal{R}_\theta \mathcal{K}_N| \geq \frac{C}{K} \right\} \lesssim \frac{1}{S}.$$

This gives us the weak type inequality

$$\mu \left\{ \theta \in \left(0, \frac{\pi}{4}\right) : |\text{Proj } \mathcal{R}_\theta \mathcal{K}_N| \geq t \right\} \lesssim (N^{-1} t^{-p})^{1/q},$$

provided that $N^{-1} t^{-p}$ is much less than 1. Combining this with the trivial estimate $\mu \left\{ \theta \in \left(0, \frac{\pi}{4}\right) : \dots \right\} \leq \frac{\pi}{4}$ for all other t , we finally get

$$\begin{aligned} \int_0^{\frac{\pi}{4}} |\text{Proj } \mathcal{R}_\theta \mathcal{K}_N| d\theta &= \int_0^{\infty} \mu \left\{ \theta \in \left(0, \frac{\pi}{4}\right) : |\text{Proj } \mathcal{R}_\theta \mathcal{K}_N| \geq t \right\} dt \\ &\lesssim \int_0^{CN^{-1/p}} 1 dt + \int_{CN^{-1/p}}^{\infty} N^{-1/q} t^{-p/q} dt \\ &\lesssim N^{-\frac{1}{p}} + N^{-\frac{1}{q}} N^{\frac{1}{p}(\frac{p}{q}-1)} = 2N^{-\frac{1}{p}}, \end{aligned}$$

finishing the proof. \square

§4. h -HAUSDORFF MEASURES OF THE PROJECTIONS

If a function h is monotone increasing and continuous, and $h(0) = 0$, we can define the Hausdorff measure \mathcal{H}_h on compact sets by the usual procedure. When $h(t) = t$, this is exactly the Hausdorff measure \mathcal{H}^1 of dimension 1. We know that the \mathcal{H}^1 measure of almost all projections of the 1/4 Cantor set is zero, and the Hausdorff dimension of almost every projection is 1. We can get more information about these projections by measuring their \mathcal{H}_h with respect to a more refined scale of gauge functions than merely powers of t . Namely, the main result obtained in this paper readily implies the following corollary. Consider the gauge function $h(t) = t(\log \frac{1}{t})^c$ with small positive c . We have proved the following statement.

Theorem 2. *If c is sufficiently small ($c \in (0, 1/6)$), then almost every projection of the four-corner Cantor set \mathcal{K} has zero \mathcal{H}_h measure.*

REFERENCES

- [1] A. S. Besicovitch, *Tangential properties of sets and arcs of infinite linear measure*, Bull. Amer. Math. Soc. **66** (1960), 353–359. MR0120335 (22:11090)
- [2] G. David, *Analytic capacity, Calderón–Zygmund operators, and rectifiability*, Publ. Mat. **43** (1999), 3–25. MR1697514 (2000e:30044)
- [3] K. J. Falconer, *The geometry of fractal sets*, Cambridge Tracts in Math., vol. 85, Cambridge Univ. Press, Cambridge, 1986. MR0867284 (88d:28001)
- [4] P. W. Jones and T. Murai, *Positive analytic capacity but zero Buffon needle probability*, Pacific J. Math. **133** (1988), 99–114. MR0936358 (89m:30050)
- [5] R. Kenyon, *Projecting the one-dimensional Sierpinski gasket*, Israel J. Math. **97** (1997), 221–238. MR1441250 (98i:28002)
- [6] J. C. Lagarias and Y. Wang, *Tiling the line with translates of one tile*, Invent. Math. **124** (1996), 341–365. MR1369421 (96i:05040)
- [7] J. Mateu, X. Tolsa, and J. Verdera, *The planar Cantor sets of zero analytic capacity and the local $T(b)$ -theorem*, J. Amer. Math. Soc. **16** (2003), 19–28. MR1937197 (2003k:30041)
- [8] P. Mattila, *Orthogonal projections, Riesz capacities, and Minkowski content*, Indiana Univ. Math. J. **39** (1990), 185–198. MR1052016 (91d:28018)
- [9] Y. Peres, K. Simon, and B. Solomyak, *Self-similar sets of zero Hausdorff measure and positive packing measure*, Israel J. Math. **117** (2000), 353–379. MR1760599 (2001g:28017)
- [10] Y. Peres and B. Solomyak, *How likely is Buffon’s needle to fall near a planar Cantor set?* Pacific J. Math. **204** (2002), 473–496. MR1907902 (2003h:28010)
- [11] T. Tao, *A quantitative version of the Besicovitch projection theorem via multiscale analysis*, arXiv:0706.2446v1 [math. CA] 18 Jun 2007, pp. 1–28. Proc. London Math. Soc. (3) **98** (2009), 559–584. MR2500864 (2010f:28009)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN
E-mail address: `nazarov@math.wisc.edu`

MICROSOFT RESEARCH REDMOND – AND – DEPARTMENTS OF STATISTICS AND MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY
E-mail address: `peres@microsoft.com`

DEPARTMENT OF MATHEMATICS, MICHIGAN STATE UNIVERSITY – AND – THE UNIVERSITY OF EDINBURGH, UNITED KINGDOM
E-mail address: `volberg@math.msu.edu`
E-mail address: `a.volberg@ed.ac.uk`

Received 20/OCT/2008
 Originally published in English