Math. Res. Lett. **15** (2008), *no.* 1, 73–81

KAKEYA SETS IN CANTOR DIRECTIONS

Michael Bateman and Nets Hawk Katz

1. Introduction

In this paper, we prove the following.

Theorem 1.1. *For any $N = 3^n$, there is a union of N parallelograms $P_1, \dots P_N$ in $\mathbb{R}^2$ of eccentricity $\sim N$ and area $\sim \frac{1}{N}$ so that the slopes of the long sides of $P_1, \dots, P_N$ are all contained in the standard middle-thirds Cantor set, so that*

$$|\bigcup_{j=1}^{N} P_j| \lesssim \frac{1}{\log N}, \tag{1.1}$$

but so that, if we let $2P_j$ be the double of the parallelogram we have

$$|\bigcup_{j=1}^{N} 2P_j| \gtrsim \frac{\log \log N}{\log N}.$$

In the statement of the theorem as in the rest of the paper, we use the convention that when S is a subset of the plane $\mathbb{R}^2$, we denote by $|S|$ the Lebesgue measure of S. Further when A_N and B_N are numbers depending on N, and we write $A_N \lesssim B_N$, we mean there is a constant C independent of N so that

$$A_N \le C B_N.$$

The proof of our theorem is by a probabilistic construction. The estimate which allows us to prove inequality (1.1) is a now fairly standard estimate on percolation on trees following the work of Russ Lyons ([5], [6]). As far as we know, this idea has not appeared in the study of Kakeya sets before. The moral of the story is that if we define (loosely) a Kakeya set in the plane as a "1 dimensional" family of unit line segments whose union has measure 0 then while it is true that the random family of line segments is not a Kakeya set, it is the case that the random, *sticky*, set of line segments is a Kakeya set. Here we use the term sticky as in [8].

If we let $\mathcal{S}$ be the set of line segments in the plane whose slope is in the standard Cantor set, and we define for $s \in \mathcal{S}$, the expression $av_s(f)$ to be the average of a function f on s, where f must be locally integrable on lines, we may define a maximal operator

$$\mathcal{M}f(x) = \sup_{x \in s \in \mathcal{S}} av_s|f|.$$

An immediate consequence of our theorem is

Corollary 1.2. *The maximal operator $\mathcal{M}$ is unbounded on any $L^p(\mathbb{R}^2)$ with $p \neq \infty$.*

Received by the editors January 15, 2007.

This was proved for $p \leq 2$ in [7]. Previously, the operator had been explicitly studied in [2] and [10]. The boundedness of this operator had been known in the folklore as an open problem for more than a decade previously. Since the publication of [7], work by Hare and Rönning ([3], [4]) attempted to address the relationship between various notions of dimensionality (e.g. Hausdorff and box dimension) and the boundedness of directional maximal operators. They managed to construct Perron-type trees for a certain class of sets of directions, but their results do not extend to the traditional Cantor set.

An early result due to Nagel, Stein, and Wainger [9] showed the boundedness of the directional maximal operator on L^p when the set of directions if lacunary, i.e., approaching a limit at least geometrically. On the other hand, the unit interval, for example, is a set that yields an unbounded directional maximal operator on L^p. This paper settles the question in the intermediate case of the Cantor set. Recently, the first author in [1] extended the methods of this paper to achieve a full classification of the boundedness of directional maximal operators in $\mathbb{R}^2$.

2. Geometric Constructions

We denote by T_n the set of all n-digit strings $.a_1 a_2 \dots a_n$ with each a_j taking on the value $0, 1$ or 2. Here we consider T_0 to be the singleton set containing ".", the empty decimal. We define the maps

$$\pi_j : T_n \longrightarrow \{0, 1, 2\},$$

by

$$\pi_j(.a_1 a_2 \dots a_n) = a_j,$$

and for $j < n$, we define

$$\pi^j : T_n \longrightarrow T_j,$$

by

$$\pi^j(.a_1 a_2 \dots a_n) = .a_1 a_2 \dots a_j.$$

We define

$$T_n^* = \bigcup_{j=0}^{n} T_j.$$

We may view T_n^* as a rooted ternary tree with an edge between $s \in T_j$ and $t \in T_{j-1}$ whenever $\pi^{j-1}(s) = t$. We denote this edge by $e_{t,\pi_j(s)}$ and say that s is the $\pi_j(s)$th child of t. We identify the tree T_n^* with the triadic intervals of length greater than 3^{-n}, by the map

$$I(s) = [s, s + \frac{1}{3^j}],$$

when $s \in T_j$ and $s = .a_1 \dots a_j$ is identified with the triadic rational

$$\frac{a_1}{3} + \frac{a_2}{9} + \dots + \frac{a_j}{3^j}.$$

Whenever $s, t \in T_n^*$, and $I(s) \subset I(t)$, we say that t is an ancestor of s, or s is a descendant of t.

We denote by $C_n \subset T_n$ the set of all n-digit strings $.a_1a_2\dots a_n$ so that each a_j takes on the value either 0 or 2. Then $I(C_n)$ is the nth stage of the construction of the standard Cantor set. We will say that a map

$$\sigma : T_n \longrightarrow C_n,$$

is *sticky* provided that for any $s \in T_n$, the value of $\pi_j(\sigma(s))$ depends only on $\pi^j(s)$. We shall define a random variable σ_n which takes values in sticky maps from T_n to C_n. This random variable shall in fact be evenly distributed among such maps, but we define its components more explicitly.

To each edge $e_{t,a}$ of T_n^*, we define a random variable $r_{t,a}$. The variables $r_{t,a}$ are independent and take on the values 0 and 2 with probability $\frac{1}{2}$ each. Now we define

$$\sigma_n(s) = s',$$

where $\pi_j(s') = r_{\pi^{j-1}(s),\pi_j(s)}$.

Following [7], we assign a "Kakeya set" to every possible value of the random variable σ_n. (In [7], this was actually when $r_{t,a} = 0$ for $a = 0, 1$ and $r_{t,a} = 2$ for $a = 2$, independently of t.) Given a sticky map

$$\sigma : T_n \longrightarrow C_n,$$

we define for each $s \in T_n$, a parallelogram in $\mathbb{R}^2$ which we will denote by $P_{\sigma,s}$. The parallelogram $P_{\sigma,s}$ has as its corners the points $(0, \frac{s}{3}), (0, \frac{s}{3} + \frac{1}{3^{n+1}}), (1, \frac{s}{3} + \sigma(s))$, and $(1, \frac{s}{3} + \frac{1}{3^{n+1}} + \sigma(s))$. (Here we again identify s and $\sigma(s)$ as real numbers by the ternary expansion.) We think of $P_{\sigma,s}$ as a tube with eccentricity approximately $\frac{1}{3^{n+1}}$ which begins at $(0, \frac{s}{3})$ and has slope $\sigma(s)$. Then we define a "Kakeya set" by

$$K_\sigma = \bigcup_{s\in T_n} P_{\sigma,s}.$$

Our first goal is to prove

Lemma 2.1. *For any choice of a sticky map*

$$\sigma : T_n \longrightarrow C_n,$$

we have that

$$|K_\sigma| \gtrsim \frac{\log n}{n}.$$

Notice that Lemma 2.1 is a generalization of ([7],Lemma 2.3).

To prove this, we first establish the following elementary uniformity inequality in measure theory.

Proposition 2.2. *Suppose $(X, \mathcal{N}, \mu)$ is a measure space and $A_1, \dots, A_K$ are sets with $\mu(A_j) = \alpha$. Let $m > 0$. Suppose that*

$$\sum_{j=1}^{K}\sum_{k=1}^{K} \mu(A_j \cap A_k) \le Km\alpha,$$

then

$$\mu(\bigcup_{j=1}^{K} A_j) \ge \frac{K\alpha}{16m}.$$

(The 16 in the denominator is unnecessary, but simplifies the proof slightly.)

Proof. It must be that there is $S \subset \{1, \dots K\}$ with $\#(S) \geq \frac{K}{2}$ so that we have

$$\sum_{j=1} \mu(A_j \cap A_k) \leq 2m\alpha,$$

whenever $k \in S$. For any such k, there must be a measurable set $B_k \subset A_k$ so that

$$\sum_{j=1}^{K} \chi_{A_j}(x) \leq 4m,$$

for any $x \in B_k$, and so that $\mu(B_k) \geq \frac{\alpha}{2}$. Then

$$\int \sum_{k \in S} \chi_{B_k} \geq \frac{K\alpha}{4},$$

but

$$\sum_{k \in S} \chi_{B_k}(x) \leq \sum_{j=1}^{K} \chi_{A_j}(x) \leq 4m,$$

for $x \in \cup_{k \in S} B_k$. Thus by Chebychev's inequality, we have

$$\frac{K\alpha}{16m} \leq \mu(\bigcup_{k \in S} B_k) \leq \mu(\bigcup_{j=1}^{K} A_j),$$

which was to be shown. □

Proof. We will show that for $0 \leq j < \log n$, with $S_j = [3^{-j}, 3^{1-j}] \times \mathbb{R}$ we have the estimate

$$|K_\sigma \cap S_j| \gtrsim \frac{1}{n}.$$

We see that

$$K_\sigma \cap S_j = \bigcup_{s \in T_n} P_{\sigma,s,j},$$

where

$$P_{\sigma,s,j} = P_{\sigma,s} \cap S_j.$$

Since for each value of s, we have

$$|P_{\sigma,s,j}| = \frac{2}{3^{j+n}},$$

it suffices to show, in light of Proposition 2.2 that

$$\sum_{s_1 \in T_n} \sum_{s_2 \in T_n} |P_{\sigma,s_1,j} \cap P_{\sigma,s_2,j}| \lesssim \frac{n}{3^{2j}}. \tag{2.1}$$

(Note that the inequality fails for $j \geq \log n$ because of the diagonal part of the sum.) Between any s_1 and s_2 we define the triadic distance $d(s_1, s_2)$ to be 3^{-k} where k is the largest number for which $\pi^k(s_1) = \pi^k(s_2)$. Note that for any $s_1 \neq s_2$, we have that

$$P_{\sigma,s_1,j} \cap P_{\sigma,s_2,j} \neq \emptyset$$

implies that $d(s_1, s_2) \gtrsim 3^j |s_1 - s_2|$, where again we have identified s_1 and s_2 as numbers. Further, we always have the estimate

$$|P_{\sigma,s_1,j} \cap P_{\sigma,s_2,j}| \lesssim \frac{1}{3^{2n+j}|s_1 - s_2|}, \tag{2.2}$$

because $3^j|s_1 - s_2|$, bounds below the difference in the slopes $\sigma(s_1)$ and $\sigma(s_2)$. We divide up the sum in (2.1) according to the approximate value of $3^j|s_1 - s_2|$. For $k = -n, ..., -1, 0$, let $A_{k,j}$ be the number of pairs (s_1, s_2) for which $d(s_1, s_2) \geq 3^j|s_1 - s_2|$ and $3^j|s_1 - s_2| \sim 3^k$, and observe that

$$A_{k,j} \lesssim 3^{2n+k-2j}. \tag{2.3}$$

Combining (2.2) and (2.3) and summing over k proves the estimate (2.1). □

For the remainder of this section, we fix a point $(t, y) \in \mathbb{R}^2$ with $\frac{1}{3} < t < 1$. We investigate the probability $P_n(t, y)$ of the event that $(t, y) \in K_{\sigma_n}$.

For every $s \in T_k$ and every $c \in C_k$, we consider $I_{s,c,t}$ which is the set of y so that (t, y) is contained in a line whose y-intercept is in the interval $[\frac{s}{3}, \frac{s}{3} + \frac{1}{3^{k+1}}]$ and whose slope is contained in $[c, c + \frac{1}{3^k}]$. We easily see that

$$I_{s,c,t} = [\frac{s}{3} + tc, \frac{s}{3} + \frac{1+3t}{3^{k+1}} + tc].$$

We observe that for any distinct $c_1, c_2 \in C_k$, we have $|c_1 - c_2| \geq \frac{2}{3^k}$, so that since $t > \frac{1}{3}$, the collection

$$\{I_{s,c,t}\}_{c \in C_k},$$

is pairwise disjoint. Therefore for each value of s, there is at most one value of c so that $y \in I_{s,c,t}$. (There may be no such value.) If such a value c exists we denote it by $c = c_{t,y}(s)$. Otherwise, we write $c_{t,y}(s) = \infty$. Note that, by definition, if $c_{t,y}(s)$ is finite then $c_{t,y}(s')$ is finite for any ancestor s' of s. Note further that if we are given s_1 and s_2 with $c_{t,y}(s_1), c_{t,y}(s_2)$ both finite and if $I(s_2) \subset I(s_1)$ then $I(c_{t,y}(s_2)) \subset I(c_{t,y}(s_1))$. We denote by $T^*_{n,t,y}$, the set of those $s \in T^*_n$ so that $c_{t,y}(s)$ is finite. Then the collection $T^*_{n,t,y}$ is a subtree of T^*_n.

We make two observations about the tree $T^*_{n,t,y}$. The first observation is that the event $(t, y) \in K_{\sigma_n}$ occurs only if there is some $s \in T_n \cap T^*_{n,t,y}$ so that $\sigma_n(s) = c_{t,y}(s)$. This, in turn, happens if and only if for every $0 < k \leq n$ we have that

$$\pi_k(c_{t,y}(s)) = r_{\pi^{k-1}(s),\pi_k(s)}. \tag{2.4}$$

The events in (2.4) are in one to one correspondence with the edges $e_{\pi^{k-1}(s),\pi_k(s)}$, are independent of one another, and occur with probability $\frac{1}{2}$. Thus $P_n(t, y)$ is bounded by the probability that if we remove each edge of $T^*_{n,t,y}$ independently with probability $\frac{1}{2}$, that we leave in place a path from the root to the nth generation. This is called, in the probability literature, (see *e.g.* [5],[6]) the survival probability of Bernoulli($\frac{1}{2}$) percolation on the tree $T^*_{n,t,y}$. We record this observation as a Lemma.

Lemma 2.3. *With $\frac{1}{3} < t < 1$, we have that $P_n(t, y)$, the probability that (t, y) is in the random "Kakeya set" K_{σ_n} is bounded by the survival probability of Bernoulli($\frac{1}{2}$) percolation on the associated tree $T^*_{n,t,y}$.*

The second observation is that for any k, the set of $s \in T_k$ such that $y \in I_{s,c,t}$ for some c is contained in 2^k intervals of length $t3^{-k}$ which in turn is contained in $\lesssim 2^k$ triadic intervals of length $3^{-(k+1)}$. Thus, we get immediately

Lemma 2.4. *We have, for every $0 \leq k \leq n$ the estimate*

$$\#(T_k \cap T^*_{n,t,y}) \lesssim 2^k.$$

Lemmas 2.3 and 2.4 will be enough to allow us to obtain the estimate which we require for $P_n(t,y)$. We carry this out in the following section.

3. Percolation on Trees

In this section, we review part of the theory of percolation on trees. We do not claim any originality. All results are special cases of theorems of Russ Lyons (see *e.g.* [5],[6]). Pointers may be found there to a much wider literature).

We let $T' \subset T^*_n$ be a subtree. We remove each edge of T' independently with probability $\frac{1}{2}$. We denote by $P(T')$ the probability that a path remains from the root to $T_n \cap T'$.

We introduce one other quantity associated to T'. We view T' as an electric circuit which has a battery whose positive node is connected to the root and whose negative part is connected in parallel to each vertex of $T_n \cap T'$. On each edge of T' which connects a vertex of T_{k-1} to a vertex of T_k, we place a resistor with resistance 2^k. We denote by $R(T')$, the resistance between the root of T' and the bottom $T' \cap T_n$. (For more on the mathematical theory of electrical circuits, see [6].) The following theorem is due to Lyons [5], in greater generality and with a better constant. We include the proof which follows simply to make the paper self-contained.

Theorem 3.1. *(Lyons)We have that*

$$P(T') \lesssim \frac{1}{2 + R(T')}.$$

Proof. We prove this by induction on n. Clearly it is true for constant 2, when $n = 0$. We assume up to $n-1$, we have

$$P(T') \leq \frac{12}{2 + R(T')}.$$

We observe that if T' is subtree of T containing the root, we may view T' as the root, together with up to 3 edges connected to 3 trees T_1, T_2, and T_3. (If some of these trees are empty, we assign them probabilty zero and infinite resistance.) We denote

$$P(T_j) = P_j,$$

and

$$R(T_j) = R_j.$$

Then we have the recursive formulae

$$P(T) = \frac{1}{2}(P_1 + P_2 + P_3) - \frac{1}{4}(P_1P_2 + P_1P_3 + P_2P_3) + \frac{1}{8}P_1P_2P_3 \tag{3.1}$$

and

$$\frac{1}{R(T)} = \frac{1}{2 + 2R_1} + \frac{1}{2 + 2R_2} + \frac{1}{2 + 2R_3}.$$

Now we break into two cases. In the first case, we have $\frac{12}{2+R_j} > 2$ for some j. Then we have $R_j < 4$. This implies $R(T) < 10$ which implies $\frac{12}{2+R(T)} > 1$, so that we certainly have

$$P(T) \leq \frac{12}{2+R(T)}.$$

We define

$$Q_j = \frac{12}{2+R_j}.$$

We may assume each $Q_j \leq 2$. Observe that if we define

$$F(x,y,z) = 1 - (1-\frac{1}{2}x)(1-\frac{1}{2}y)(1-\frac{1}{2}z),$$

on the domain $[0,2]\times[0,2]\times[0,2]$ then F is monotone increasing in each variable. Therefore we have that

$$\begin{aligned} P(T) &= F(P_1,P_2,P_3) \\ &= F(Q_1,Q_2,Q_3) \\ &\leq \frac{1}{2}(Q_1+Q_2+Q_3) - \frac{1}{6}(Q_1Q_2+Q_1Q_3+Q_2Q_3). \end{aligned} \tag{3.2}$$

Note that the equality is (3.1), while for the two inequalities we have used that the Q's are ≤ 2.

Now plugging the definition of the Q's into (3.2), we obtain

$$\begin{aligned} P(T) &\leq \frac{12}{2}[\frac{(R_1+2)(R_2+2)+(R_1+2)(R_3+2)+(R_2+2)(R_3+2)-4(R_1+R_2+R_3+6)}{(R_1+2)(R_2+2)(R_3+2)}] \\ &\leq \frac{12}{2}[\frac{(R_1+2)(R_2+2)+(R_1+2)(R_3+2)+(R_2+2)(R_3+2)-4(R_1+R_2+R_3+6)}{(R_1+2)(R_2+2)(R_3+2)-R_1-R_2-R_3-4}] \\ &\leq \frac{12}{2}[\frac{(R_1+1)(R_2+1)+(R_1+1)(R_3+1)+(R_2+1)(R_3+1)}{(R_1+2)(R_2+2)(R_3+2)-R_1-R_2-R_3-4}] \\ &= \frac{12}{R(T)+2}. \end{aligned}$$

Here the second inequality is by decreasing the denominator and the third inequality is by increasing the numerator. □

Next we estimate the resistance of the trees we are interested in.

Lemma 3.2. *Let $T^*_{n,t,y}$ be as in section 1. Then*

$$R(T^*_{n,t,y}) \gtrsim n.$$

Proof. We use the basic physical principle, that the resistance of any circuit may be reduced by shortcircuiting it with perfect conductors. We identify all vertices in each T_k, thus reducing the resistance. Then by Lemma 2.4, we have that T_{k-1} and T_k are connected by $\lesssim 2^k$ resistors of resistance 2^k connected in parallel. Thus the resistance between T_{k-1} and T_k is $\gtrsim 1$. Thus the total resistance is $\gtrsim n$. □

Corollary 3.3. *Let $\frac{1}{3} < t \leq 1$. Then with $P_n(t,y)$, the probability that $(t,y) \in K_{\sigma_n}$, we have that*

$$P_n(t,y) \lesssim \frac{1}{n}.$$

Proof. We combine Lemma 2.3, Theorem 3.1, and Lemma 3.2. □

4. Proof of the main theorem

Proof. We observe that in order for a point (t, y) to be in any set K_σ, it must be that $0 \leq y \leq \frac{4}{3}$. Thus E, the expected measure of $K_{\sigma_n} \cap ([\frac{1}{3}, 1] \times \mathbb{R})$ is given by

$$E = \int (\int_{\frac{1}{3}}^{1} \int_{0}^{\frac{4}{3}} \chi_{K_\sigma}(t, y) dy dt) d\sigma,$$

where the outside integral takes place on a finite probability space. Interchanging the integrals, we see that

$$E = \int_{\frac{1}{3}}^{1} \int_{0}^{\frac{4}{3}} P_n(t, y) dy dt \lesssim \frac{1}{n}.$$

Therefore there is a choice of σ for which

$$|K_\sigma \cap ([\frac{1}{3}, 1] \times \mathbb{R})| \lesssim \frac{1}{n}.$$

On the other hand

$$|K_\sigma| \gtrsim \frac{\log n}{n}.$$

Thus K_σ is the desired example. □

Acknowledgements

The first author was supported by NSF grant DMS-FRG-0139874; The second author was supported by NSF grant DMS 0432237. We thank Russ Lyons for helpful discussions.

References

[1] Bateman, M. *Kakeya Sets and Directional Maximal Operators in the plane* preprint.
[2] Duoandikoetxea, J. and Vargas, A. *Directional Operators and radial functions on the plane* Ark. Mat. **33** (1995) 281-291.
[3] Hare, K. and Rönning, J-O. *Applications of generalized Perron trees to maximal functions and density bases* J. Fourier Anal. and Appl. **4** (1998) 215-227.
[4] Hare, K. and Rönning, J-O. *Size of Max(p) sets and density bases* J. Fourier Anal. and Appl. **8** (2002) 259-268.
[5] Lyons, R. *Random walks, Capacity, and Percolation on trees* Ann. Probab. **20** (1992) 2043-2088.
[6] Lyons, R. and Peres, Y. *Probability on Trees and Networks,* in preparation, http://mypage.iu.edu/ rdlyons/prbtree/prbtree.html.
[7] Katz, N.H. *A counterexample for maximal operators over a Cantor set of directions* Mat Res. Let. **3** (1996) 527–536.
[8] Katz, N.H, Laba, I., and Tao, T. *An improved bound on the Minkowski dimension of Besicovitch sets* Ann. Math. **152** (2000) 383-446.
[9] Nagel, A., Stein, E.M., and Wainger, S. *Differentiation in Lacunary Directions* Proc. Nat. Acad. Sci. **75** (1978) 1060-1062.
[10] Vargas, A. *A remark on a maximal function over a Cantor set of directions* Rend. Circ. Mat. Palermo **44** (1995) 273–282.

Department of Mathematics, Indiana University, Bloomington, Indiana 47405-7106
E-mail address: mdbatema@indiana.edu

Department of Mathematics, Indiana University, Bloomington, Indiana 47405-7106
E-mail address: nhkatz@indiana.edu