Kakeya Sets: The Paper: The Talk

1 Outline

- I aim to cover roughly the following things:
  1. Brief history of the Kakeya problem in analysis
  2. The finite-field Kakeya problem >> Dvir's solution
  3. Kakeya over non-Archimedean local rings >> my and Marci's work

2 Kakeya in Analysis

- A Kakeya (needle) set is a subset of the plane $\mathbb{R}^2$ inside which it is possible to rotate a needle of length 1 completely around.
  - Obvious example: A circle of diameter 1, area $\frac{\pi}{4}$.
  - Less obvious example: A deltoid, area $\frac{\pi}{8}$.
- Question ("Kakeya needle problem"): How small (in measure) can a Kakeya needle set be?
  - Originally asked by Soichi Kakeya in 1917, for a convex set. (That problem was also solved, although I don't know what the answer is.)
  - Answer (Besicovitch, 1919): Arbitrarily small measure.
  - Construction is fairly geometric: general idea is to divide up the region into smaller pieces, then translate the pieces so they overlap a lot. End result is a sort of spiky-triangle shape.
  - It is also possible to adjust the construction so as to keep the Kakeya set simply-connected and still have arbitrarily small measure.
- If we reformulate the problem slightly, to say a Kakeya set is one containing a unit line segment in every possible direction, then one can construct a Kakeya set of measure zero.
- Followup question 1: Are there Kakeya sets of measure zero in $\mathbb{R}^n$? (yes: just take the product of a 2-dimensional set with a unit segment.)
- Followup question 2 ("Kakeya conjectures"): "How small" of a measure zero set can a Kakeya set be? In other words, must a Kakeya set in $\mathbb{R}^n$ actually be n-dimensional (Hausdorff or Minkowski), or could its dimension be less than n?
  - Answer: In the plane, a Kakeya set must be 2-dimensional. (This is a not easy theorem.)
  - For higher dimensions, only lower bounds are known. Example: Terry Tao and (Nets) Katz showed that a Kakeya set in $\mathbb{R}^n$ must have dimension at least $(2-\sqrt{2})(n-4)+3$.
- Some other remarks:
  - Kakeya sets have some uses in harmonic analysis, and are of interest to analysts. There is also a related Kakeya maximal function, which I don't know anything about, that is also interesting.
  - Example: A theorem of Fefferman uses Kakeya sets to show that certain truncated Fourier integrals need not converge in L^p norm for $p \neq 2$. Barry Mazur told me that Fefferman's result partly explains "why L^2 is the best L^p space".

3 Finite-Field Kakeya

- In 1999, Thomas Wolff posed a reformulation of the Kakeya problem in a finite field setting.
- Wolff's "finite-field Kakeya problem" asks: For $\mathbb{F}_q$ a field, we define a Kakeya set in $(\mathbb{F}_q)^n$ to be a (finite) set which contains a line in every possible direction.
  - Here, a line through x in the direction of y is the set of points $x + t\,y$ as t runs through the elements of $\mathbb{F}_q$.
  - So in this context, a Kakeya set is a set K such that, for every $y \in \mathbb{F}_q^n$ there exists $x \in \mathbb{F}_q^n$ such that $x + ty \in K$ for all $t \in \mathbb{F}_q$.
- Wolff conjectured that a Kakeya set in $\mathbb{F}_q^n$ must contain at least $c_n \cdot |\mathbb{F}_q|^n$ points, for some constant c_n depending only on n (and not on q).
- Several authors proved lower bounds on the order of $q^{4n/7}$ using fairly difficult methods in additive number theory, before Zeev Dvir proved the full conjecture in roughly one page in 2008 using very elementary algebraic geometry.
- <u>Theorem</u> (Dvir): If K is a Kakeya set in $\mathbb{F}_q^n$, then $|K| \geq \left(\begin{array}{c} n+q-2 \\ n \end{array} \right) \geq \frac{1}{n!} \cdot q^n$.
- <u>Proof:</u>
  - Suppose otherwise, and take K to be a Kakeya set of size less than $\left(\begin{array}{c} n+q-2 \\ n \end{array} \right)$. Then the collection of polynomials in $\mathbb{F}_q[x_1, \cdots, x_n]$ of degree at most $q-1$ is a vector space of dimension $\left(\begin{array}{c} n+q-2 \\ n \end{array} \right) > |K|$, so there exists a nonzero polynomial $P \in \mathbb{F}_q[x_1, \cdots, x_n]$ of degree at most $q-1$ such that $P(x) = 0$ for all $x \in K$.
  - Write $P = \sum_{i=0}^{q-1} P_i$, where P_i is homogeneous of degree i, and fix $y \in \mathbb{F}_q^n$.
  - Since K is a Kakeya set, for any $y \in \mathbb{F}_q^n$ there exists $b \in \mathbb{F}_q^n$ for which $P(b + ty) = 0$ for all $t \in \mathbb{F}$. For fixed b and y, this is a polynomial of degree $q-1$ in the variable t which vanishes for q different values of t. Hence this polynomial in t is identically zero, and so in particular its coefficient of t^{q-1} is zero. Expanding shows this is just $P_{q-1}(y)$.
  - Therefore, $P_{q-1}(y) = 0$ for every $y \in \mathbb{F}_q^n$. However, this can only happen if P_{q-1} is the zero polynomial – to see this, just use the division algorithm with respect to x_1, then x_2, up through x_n, and the fact that P_{q-1} has degree less than q.
  - Now repeat the above argument for each of $P_{q-2}, \ldots, P_1$ to see that P must be a constant, hence the zero polynomial. Contradiction.
  - Hence $|K| \geq \left(\begin{array}{c} n+q-2 \\ n \end{array} \right) = \frac{(q+n-2)(q+n-3)\cdots(q-1)}{n!} \geq \frac{1}{n!} \cdot q^n$ for $n \geq 4$.

4 Non-Archimedean Kakeya, interlude

- In 2010, Jordan along with Richard Oberlin and Terry Tao wrote a paper reviewing the Kakeya problem over finite fields. Others had shown that the constant in Dvir's estimate is not sharp – it can be improved to $\left(\frac{1}{2} + o(1)\right)^n$, which EOT thought likely to be optimal, perhaps up to removing the $o(1)$.
- If we take the probability measure on $\mathbb{F}_q^n$ (so that the whole space has measure 1) then Dvir's result says that the measure of a Kakeya set in $\mathbb{F}_q^n$ is at least $\frac{1}{n!}$, which is positive. (This is in opposition to the case of Kakeya sets over the reals, which can have zero measure.)

- In section 4.20 of their paper, EOT discuss the analogy between the finite-field Kakeya problem and the real one. Part of the reason that the finite-field results might not capture useful information about the problem over $\mathbb{R}$ is the lack of "multiple scales": over $\mathbb{F}_q^n$ there is no useful idea of a distance between points and lines – either two points are the the same, or they are not – while over $\mathbb{R}$ there is a very strong notion of distance.
- EOT suggested considering Kakeya sets over other rings which do have some notion of "distance", such as the finite rings $\mathbb{Z}/p^n\mathbb{Z}$ and $\mathbb{F}_q[t]/t^n$. Over these rings we can pose the Kakeya conjecture since we have a notion of dimension.
- Even better would be to consider the infinite rings $\mathbb{Z}_p$ and $\mathbb{F}_q[[t]]$, which are complete and thus seem much closer to $\mathbb{R}$.
- EOT also observed that although the measure of a Kakeya set in $\mathbb{F}_q^n$ is always positive, the best known bound nevertheless goes to zero as n goes to ∞. They therefore asked whether, perhaps, there might be a Kakeya set of measure zero lurking in the completion $\mathbb{F}_q[[t]]$.

5 My actual stuff

- Let R be an infinite ring admitting a Haar measure μ such that $\mu(R)$ is finite.
- A **line** with direction vector $v \in R^n$ through the point $x \in R^n$ consists of the elements in R^n of the form $x + tv$ as t runs through the elements of R.
- A **Kakeya set** in R^n is a subset of R^n which contains (all the points on) at least one line with each possible direction vector.
- For a finite ring R, the <u>Minkowski dimension</u> of a set $E \subset R^n$ is defined as $\dfrac{\log|E|}{\log|R|}$. The natural analogue of the Minkowski dimension of a compact subset $E \subset \mathbb{F}_q[[t]]^n$ is $\lim_{k\to\infty} \dfrac{\log|E_k|}{\log|\mathbb{F}_q|^k}$, where E_k is the image of the projection of E onto $\mathbb{F}_q[[t]]/t^k$. (Similarly, for E inside $\mathbb{Z}_p^n$.)
- <u>Theorem</u> 1 (–, Habliscek): For all $n > 1$, there exists a Kakeya set $E \subset \mathbb{F}_q[[t]]^n$ of measure 0.
- <u>Theorem</u> 2 (–, Habliscek): The Minkowski dimension of any Kakeya set in $\mathbb{F}_q[[t]]^2$ or $\mathbb{Z}_p^2$ is 2.

5.1 Sketch of Theorem 1 (Measure Zero)

- I construct a Kakeya set K of measure zero in $\mathbb{F}_q[[t]]^2$; to get one in a higher dimension just take $K \times \mathbb{F}_q[[t]]^{n-2}$.
- I refer to a nonzero direction vector $v = (a, b)$ in $\mathbb{F}_q[[t]]^2$ as **nonreduced** if t divides both a and b, and as **reduced** otherwise. It is obvious that any line with nonreduced direction vector v passing through (x, y) is contained in the line with direction vector v/t through (x, y); thus, we need only consider reduced direction vectors.
- Every nonzero reduced direction vector is of the form $(1, b)$ or $(b, 1)$. So we need only find a set E of measure zero containing a line with direction vector $(1, b)$ for each $b \in \mathbb{F}_q[[t]]$; then $K = \{(x, y) : (x, y) \text{ or } (y, x) \in E\}$ still has measure zero, and is Kakeya.
- Notation:
  - For any $a \in \mathbb{F}_q[[t]]$, let a_i denote the coefficient of t^i.
  - For any $a \in \mathbb{F}_q[[t]]$, define the element $a^* \in \mathbb{F}_q[[t]]$ by

$$a_i^* = \begin{cases} 0 & \text{if } i = 2^k - 2 \text{ for some natural number } k, \\ a_{i+1} & \text{otherwise.} \end{cases}$$

- Here is the construction: define

$$E = \{(x, y) \in \mathbb{F}_q[[t]]^2 : ax + y = a^* \text{ for some } a \in \mathbb{F}_q[[t]]\}.$$

  - To see that this contains the required lines, observe that, for any $b \in \mathbb{F}_q[[t]]$, the points $(x, y) = (0, -b^*) + s(1, b)$ are contained in E, as s ranges over $\mathbb{F}_q[[t]]$, since $(-b)s + (-b^* + bs) = (-b)^*$.
  - It only remains to prove that E has measure zero (which is, obviously, the hard part).

- We can see that $(x, y) \in E$ if and only if there exist $a_i \in \mathbb{F}_q$ for all $i \geq 0$ such that the coefficients of x and y satisfy the following infinite system:

$$\begin{aligned}
&a_0x_0 + y_0 = 0, && [0]\\
&a_1x_0 + a_0x_1 + y_1 = a_2, && [1]\\
&a_2x_0 + a_1x_1 + a_0x_2 + y_2 = 0, && [2]\\
&\quad\vdots\\
&a_nx_0 + a_{n-1}x_1 + \cdots + a_0x_n + y_n = a_n^*, && [n]\\
&\quad\vdots
\end{aligned}$$

- For an arbitrary element $(x, y) \in \mathbb{F}_q[[t]]^2$, define $s_n(x, y)$ to be the number of tuples $(a_0, \ldots, a_n) \in \mathbb{F}_q^{n+1}$ satisfying the first n equations; observe that $s_n(x, y)$ only depends on $\{x_0, y_0, \ldots, x_n, y_n\}$. Clearly if $s_n(x, y) = 0$ for any integer n, then $(x, y) \notin E$, and $s_k(x, y) = 0$ for all $k > n$, so $\mu(\{(x, y)|s_n(x, y) = 0\})$ is non-decreasing as $n \to \infty$. We show this measure tends to 1 as $n \to \infty$.

- Observe that the equations at any stage are linear in the a_i. Moreover, for i not of the form $2^k - 2$ for some integer k, Equation $[i]$ states $a_{i+1} = a_ix_0 + a_{i-1}x_1 + ... + a_0x_i + y_i$, and so we may reduce our system of equations by eliminating a_{i+1}. Basic linear algebra then implies that $s_{2^k-2}(x, y)$ is either zero or q^l for some integer $l \leq k$.

- Now we do a somewhat unconscionable amount of polynomial arithmetic to show the following:

- <u>Lemma</u>: If for a given $\{x_0, x_1, \cdots, x_{2^n-2}\}$ and $\{y_0, y_1, \cdots, y_{2^n-2}\}$ we have $s_{2^n-2}(x, y) = q^l$, and $\{x_{2^n-1}, y_{2^n-1}, ..., x_{2^{n+1}-2}, y_2$ are randomly and uniformly chosen from $\mathbb{F}_q$, then

$$s_{2^{n+1}-2}(x, y) = \begin{cases} 0 & \text{with probability } \frac{q-1}{q^{l+2}}, \\ q^l & \text{with probability } 1 - \frac{1}{q^{l+1}}, \\ q^{l+1} & \text{with probability } \frac{1}{q^{l+2}}. \end{cases}$$

- What this says is that we have a Markov chain [DRAW PICTURE!] on the points $0, 1, q, q^2, \cdots$ such that a positive proportion of measure, independent of time, at each nonzero point is sent to 0. One can also easily check that the expected value does not change over time. Since the expected value at time zero is 1, and eventually (by a trivial induction) the measure at q^l points goes to zero, we see that the measure concentrated at 0 must go to 1 as $n \to \infty$.

- <u>Remark</u>: Using a more intricate analysis (which is surprisingly difficult despite the fact that everything is totally explicit!), one can prove that at time t, the measure away from 0 is $\ll \dfrac{\ln(t)}{t}$, for an explicitly computable constant that depends only on q.

5.2 Sketch of Theorem 2 (Minkowski Dimension)

- <u>Theorem</u> 2 (–, Habliscek): The Minkowski dimension of any Kakeya set in $\mathbb{F}_q[[t]]^2$ or $\mathbb{Z}_p^2$ is 2.

- <u>Proposition</u>: Let E be a Kakeya set in R^2 where $R = \mathbb{F}_q[t]/t^k$ or $\mathbb{Z}/p^k\mathbb{Z}$. Then $|E| \geq \frac{|R|^2}{2k}$.

- Enumerate the lines by their "coefficients" – take L_i to be any line with equation $\alpha_i x + y = b_i$ where $\alpha_i \equiv i \bmod p^k$. (Do the same thing for $R = \mathbb{F}_q[t]/t^k$ by reading polynomials in base q.)
- We prove the proposition essentially by a counting argument. Given two lines in R^2 with equations $L_i : \alpha_i x + y = b_i$ and $L_j : \alpha_j x + y = b_j$ with $i \neq j$, we see that if $(x_0, y_0) \in L_i \cap L_j$ then $(\alpha_i - \alpha_j)x_0 = b_i - b_j$. Hence if $v(\alpha_i - \alpha_j) = l$, then the number of possible values for x_0 cannot exceed $|\mathfrak{m}|^l$, where v is the $\mathfrak{m}$-adic valuation. In particular, since the value of x_0 determines the value of y_0, the size of the intersection $|L_i \cap L_j|$ is at most $|\mathfrak{m}|^l = |\mathfrak{m}|^{v(\alpha_i - \alpha_j)}$.

- Lemma: For the function $f(u) := \sum_{i=1}^{u} \mathfrak{m}^{v(\alpha_i)}$, we have $f(u) \le u \cdot \left\lceil \log_{|\mathfrak{m}|}(u) \right\rceil$. In particular, for $u \le |\mathfrak{m}|^k/k = |R|/k$, we have $f(u) \le |R|$.
  - Proof: Just count the number of terms for which $v(\alpha_i) = w$.

- Proof of Prop: Apply inclusion-exclusion by writing our Kakeya set E as the union of a bunch of lines, and subtracting away their possible intersections.
  - For $l = \left\lfloor \frac{|R|}{k} \right\rfloor$ we have

$$|E| \ge |\cup_{j=1}^{l+1} L_j| \ge \sum_{j=1}^{l+1} \left(|L_j| - \sum_{i=1}^{j-1} |L_i \cap L_j| \right)$$

  - Now look at the terms in the sum. We have $|L_j| - \sum_{i=1}^{j-1} |L_i \cap L_j| = |R| - \sum_{i=1}^{j-1} \mathfrak{m}^{v(\alpha_j - \alpha_i)} = |R| - f(j-1)$.
  - Now just sum and use the upper bound on f from the lemma. The result follows.

- Proof of Theorem:
  - Suppose E is a Kakeya set in $\mathbb{F}_q[[t]]$ or $\mathbb{Z}_p$. Then its projection E_k to $R = \mathbb{F}_q[t]/t^k$ or $\mathbb{Z}/p^k\mathbb{Z}$ is also Kakeya.
  - By the proposition, we have $|E_k| \ge \dfrac{|R|}{2k}$.
  - Then the Minkowski dimension of E_k is at least $2 - \dfrac{\log(2k)}{k \log(|\mathfrak{m}|)}$.
  - So for fixed p or q, as $k \to \infty$ this bound goes to 2. Hence result.