ON THE SIZE OF KAKEYA SETS IN FINITE FIELDS

ZEEV DVIR

ABSTRACT. A Kakeya set is a subset of $\mathbb{F}^n$, where $\mathbb{F}$ is a finite field of q elements, that contains a line in every direction. In this paper we show that the size of every Kakeya set is at least $C_n \cdot q^n$, where C_n depends only on n. This answers a question of Wolff [Wol99].

1. Introduction

Let $\mathbb{F}$ denote a finite field of q elements. A Kakeya set (also called a Besicovitch set) in $\mathbb{F}^n$ is a set $K \subset \mathbb{F}^n$ such that K contains a line in every direction. More formally, K is a Kakeya set if for every $x \in \mathbb{F}^n$ there exists a point $y \in \mathbb{F}^n$ such that the line

$$L_{y,x} \triangleq \{y + a \cdot x | a \in \mathbb{F}\}$$

is contained in K.

The motivation for studying Kakeya sets over finite fields is to try and understand better the more complicated questions regarding Kakeya sets in $\mathbb{R}^n$. A Kakeya set $K \subset \mathbb{R}^n$ is a compact set containing a line segment of unit length in every direction. The famous Kakeya Conjecture states that such sets must have Hausdorff (or Minkowski) dimension equal to n. The importance of this conjecture is partially due to the connections it has to many problems in harmonic analysis, number theory and PDE. This conjecture was proved for $n = 2$ [Dav71] and is open for larger values of n (we refer the reader to the survey papers [Wol99, Bou00, Tao01] for more information)

It was first suggested by Wolff [Wol99] to study finite field Kakeya sets. It was asked in [Wol99] whether there exists a lower bound of the form $C_n \cdot q^n$ on the size of such sets in $\mathbb{F}^n$. The lower bound appearing in [Wol99] was of the form $C_n \cdot q^{(n+2)/2}$. This bound was further improved in [Rog01, BKT04, MT04, Tao08] both for general n and for specific small values of n (e.g for $n = 3, 4$). For general n, the currently best lower bound is the one obtained in [Rog01, MT04] (based on results from [KT99]) of $C_n \cdot q^{4n/7}$. The main technique used to show this bound is an additive number theoretic lemma relating the sizes of different sum sets of the form $A + r \cdot B$ where A and B are fixed sets in $\mathbb{F}^n$ and r ranges over several different values in $\mathbb{F}$ (the idea to use additive number theory in the context of Kakeya sets is due to Bourgain [Bou99]).

The next theorem, proven in Section 2, gives a near-optimal bound on the size of Kakeya sets. Roughly speaking, the proof follows by observing that any degree $q - 2$ homogenous polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ can be 'reconstructed' from its value

Date: April 2008.
2000 *Mathematics Subject Classification.* Primary 52C17; Secondary 05B25.
Key words and phrases. Kakeya, Finite Fields, Polynomial Method.
Research supported by Binational Science Foundation (BSF) Grant.

on any Kakeya set $K \subset \mathbb{F}^n$. This implies that the size of K is at least the dimension of the space of polynomials of degree $q-2$, which is $\approx q^{n-1}$ (when q is large).

Theorem 1.1. *Let $K \subset \mathbb{F}^n$ be a Kakeya set. Then*

$$|K| \geq C_n \cdot q^{n-1},$$

where C_n depends only on n.

The result of Theorem 1.1 can be made into an even better bound using the simple observation that a product of Kakeya sets is also a Kakeya set.

Corollary 1.2. *For every integer n and every $\epsilon > 0$ there exists a constant $C_{n,\epsilon}$, depending only on n and ϵ such that any Kakeya set $K \subset \mathbb{F}^n$ satisfies*

$$|K| \geq C_{n,\epsilon} \cdot q^{n-\epsilon},$$

Proof. Observe that, for every integer $r > 0$, the Cartesian product $K^r \subset \mathbb{F}^{n \cdot r}$ is also a Kakeya set. Using Theorem 1.1 on this set gives

$$|K|^r \geq C_{n \cdot r} \cdot q^{n \cdot r - 1},$$

which translates into a bound of $C_{n,r} \cdot q^{n-1/r}$ on the size of K. □

We derive Theorem 1.1 from a stronger theorem that gives a bound on the size of sets that contain only 'many' points on 'many' lines. Before stating the theorem we formally define these sets.

Definition 1.3 ((δ, γ)-Kakeya Set)**.** *A set $K \subset \mathbb{F}^n$ is a (δ, γ)-*Kakeya Set *if there exists a set $\mathcal{L} \subset \mathbb{F}^n$ of size at least $\delta \cdot q^n$ such that for every $x \in \mathcal{L}$ there is a line in direction x that intersects K in at least $\gamma \cdot q$ points.*

The next theorem, proven in Section 2, gives a lower bound on the size of (δ, γ)-Kakeya sets. Theorem 1.1 will follow by setting $\delta = \gamma = 1$.

Theorem 1.4. *Let $K \subset \mathbb{F}^n$ be a (δ, γ)-Kakeya Set. Then*

$$|K| \geq \binom{d+n-1}{n-1},$$

where

$$d = \lfloor q \cdot \min\{\delta, \gamma\} \rfloor - 2.$$

Notice that, in order to get a bound of $\approx q^{n(1-\epsilon)}$ on the size of K, Theorem 1.4 allows δ and γ to be as small as $q^{-\epsilon}$.

1.1. **Improving the bound to $\approx q^n$.** Following the initial publication of this work, Noga Alon and Terence Tao [AT08] independently observed that it is possible to turn the proof of Theorem 1.1 into a proof that gives a bound of $C_n \cdot q^n$, thus achieving an optimal bound. A proof of the following theorem appears in Section 3

Theorem 1.5. *Let $K \subset \mathbb{F}^n$ be a Kakeya set. Then*

$$|K| \geq C_n \cdot q^n,$$

where C_n depends only on n.

2. Proof of Theorem 1.4

We will use the following bound on the number of zeros of a degree d polynomial proven by Schwartz and Zippel [Sch80, Zip79].

Lemma 2.1 (Schwartz-Zippel)**.** *Let* $f \in \mathbb{F}[x_1, \ldots, x_n]$ *be a non zero polynomial with* $\deg(f) \le d$. *Then*

$$|\{x \in \mathbb{F}^n \,|\, f(x) = 0\}| \le d \cdot q^{n-1}.$$

Proof of Theorem 1.4. Suppose in contradiction that

$$|K| < \binom{d+n-1}{n-1}.$$

Then, the number of monomials in $\mathbb{F}[x_1, \ldots, x_n]$ of degree d is larger than the size of K. Therefore, there exists a homogenous degree d polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ such that g is not the zero polynomial and

$$\forall x \in K, \quad g(x) = 0$$

(this follows by solving a system of linear equations, one for each point in K, where the unknowns are the coefficients of g). Our plan is to show that g has too many zeros and therefore must be identically zero (which is a contradiction).

Consider the set

$$K' \triangleq \{c \cdot x \,|\, x \in K, c \in \mathbb{F}\}$$

containing all lines that pass through zero and intersect K at some point. Since g is homogenous we have

$$g(c \cdot x) = c^d \cdot g(x)$$

and so

$$\forall x \in K', \quad g(x) = 0.$$

Since K is a (δ, γ)-Kakeya set, there exists a set $\mathcal{L} \subset \mathbb{F}^n$ of size at least $\delta \cdot q^n$ such that for every $y \in \mathcal{L}$ there exists a line with direction y that intersects K in at least $\gamma \cdot q$ points.

Claim 2.2. *For every* $y \in \mathcal{L}$ *we have* $g(y) = 0$.

Proof. Let $y \in \mathcal{L}$ be some non zero vector (if $y = 0$ then $g(y) = 0$ since g is homogenous). Then, there exists a point $z \in \mathbb{F}^n$ such that the line

$$L_{z,y} = \{z + a \cdot y | a \in \mathbb{F}\}$$

intersects K in at least $\gamma \cdot q$ points. Therefore, since $d + 2 \le \gamma \cdot q$, there exist $d + 2$ distinct field elements $a_1, \ldots, a_{d+2} \in \mathbb{F}$ such that

$$\forall i \in [d+2], \; z + a_i \cdot y \in K.$$

If there exists i such that $a_i = 0$ we can remove this element from our set of $d + 2$ points and so we are left with at least $d+1$ distinct *non-zero* field elements (w.l.o.g $a_1, \ldots, a_{d+1}$) such that

$$\forall i \in [d+1], \quad z + a_i \cdot y \in K \quad \text{and } a_i \ne 0$$

Let $b_i = a_i^{-1}$ where $i \in [d + 1]$. The $d + 1$ points

$$w_i \triangleq b_i \cdot z + y, \; i \in [d+1]$$

are all in the set K' and so

$$g(w_i) = 0, \; i \in [d+1].$$

If $z = 0$ then we have $w_i = y$ for all $i \in [d+1]$ and so $g(y) = 0$. We can thus assume that $z \neq 0$ which implies that $w_1, \ldots, w_{d+1}$ are $d+1$ *distinct* points belonging to the same line (the line through y with direction z). The restriction of $g(x)$ to this line is a degree $\leq d$ univariate polynomial and so, since it has $d+1$ zeros (at the points w_i), it must be zero on the entire line. We therefore get that $g(y) = 0$ and so the claim is proven. □

We now get a contradiction since

$$d/q < \delta$$

and, using Lemma 2.1, a polynomial of degree d can be zero on at most a d/q fraction of $\mathbb{F}^n$. □

3. Proof of Theorem 1.5

Suppose, in contradiction, that $K \subset \mathbb{F}^n$ is a Kakeya set such that

$$|K| < \binom{q+n-1}{n}.$$

Then, as is explained in the proof of Theorem 1.1, there exists a nonzero polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ of degree $d \leq q-1$ so that $g(x) = 0$ for all $x \in K$ (notice that g is not necessarily homogeneous). Let $\bar{g} \in \mathbb{F}[x_1, \ldots, x_n]$ be the homogeneous part of degree d of g so that $\bar{g}$ is non-zero and homogenous. Fix some $y \in \mathbb{F}^n$. Then there exists $z \in \mathbb{F}^n$ so that the line $\{z + t \cdot y \,|\, t \in \mathbb{F}\}$ is contained in K. Therefore,

$$P_{y,z}(t) \triangleq g(z + t \cdot y) = 0$$

for all $t \in \mathbb{F}$. Since $P_{y,z}(t)$ is a univariate polynomial of degree $d \leq q-1$ this means that $P_{y,z}(t)$ is identically zero, and hence all its coefficients are zero. In particular, the coefficient of t^d is zero, but it is easy to see that this is exactly $\bar{g}(y)$. Since y was arbitrary it follows that the polynomial $\bar{g}$ is identically zero – a contradiction. This concludes the proof. □

4. Acknowledgments

I am grateful to Avi Wigderson for encouraging me to work on this problem and for many helpful discussions. I thank my advisers Ran Raz and Amir Shpilka for their continuous support. I thank Noga Alon, Richard Oberlin and Terence Tao for pointing out the improvements to Theorem 1.1.

References

AT08. N. Alon and T. Tao. Private communication. 2008.

BKT04. J. Bourgain, N. Katz, and T. Tao. A sum-product estimate in finite fields, and applications. *GAFA*, 14(1):27–57, 2004.

Bou99. J. Bourgain. On the dimension of Kakeya sets and related maximal inequalities. *Geom. Funct. Anal.*, (9):256–282, 1999.

Bou00. J. Bourgain. Harmonic analysis and combinatorics: How much may they contribute to each other? *IMU/Amer. Math. Soc.*, pages 13–32, 2000.

Dav71. R. Davies. Some remarks on the Kakeya problem. *Proc. Cambridge Philos. Soc.*, (69):417–421, 1971.

KT99. N. Katz and T. Tao. Bounds on arithmetic projections, and applications to the Kakeya conjecture. *Math. Res. Letters*, 6:625–630, 1999.

MT04. G. Mockenhaupt and T. Tao. Restriction and Kakeya phenomena for finite fields. *Duke Math. J.*, 121:35–74, 2004.

Rog01. K.M Rogers. The finite field Kakeya problem. *Amer. Math. Monthly 108*, (8):756–759, 2001.

Sch80. J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. ACM*, 27(4):701–717, 1980.

Tao01. T. Tao. From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and pde. *Notices Amer. Math. Soc.*, 48(3):294–303, 2001.

Tao08. T. Tao. A new bound for finite field besicovitch sets in four dimensions. *Pacific J. Math (to appear)*, 2008.

Wol99. T. Wolff. Recent work connected with the Kakeya problem. *Prospects in mathematics (Princeton, NJ, 1996).* pages 129–162, 1999.

Zip79. R. Zippel. Probabilistic algorithms for sparse polynomials. In *Proceedings of the International Symposiumon on Symbolic and Algebraic Computation*, pages 216–226. Springer-Verlag, 1979.

Department of Computer Science, Weizmann Institute of Science, Rehovot, Israel
E-mail address: zeev.dvir@weizmann.ac.il