Discrete geometry is one of my favorite fields for two reasons. First, it is a field where one can ask simple questions (ones that any middle schooler could understand), while coming up with answers can be extremely difficult. Second, although the questions are often presented in the simple terms of Euclidean geometry (points, lines, circles, distances, etc.), the problems often turn out to have deep connections to other fields of math; in this talk, we'll see connections to combinatorics, number theory, topology, harmonic analysis, geometric measure theory, and algebraic geometry. There are also many important connections to computer science, which we won't cover.

1 Sylvester–Gallai and friends

In 1893, Sylvester asked the following question.

Question. *Is it possible to place n non-collinear points in the plane so that whenever a line passes through two of them, it also passes through a third?*

Note that the non-collinearity is important, since $n \geq 3$ points on a line certainly satisfy this property. Also note that if we allow infinitely many points, then the answer is certainly yes, since the lattice $\mathbb{Z}^2$ works (among many other examples).

It's worth trying to place n points in the plane to satisfy this property; in my experience, it always feels like you're getting close, but you never quite make it. This is not a coincidence, as shown by the next theorem.

Theorem ("The Sylvester–Gallai Theorem"; due to Melchior, 1941)**.** *Any n non-collinear points in the plane define a line containing exactly two of them.*

Proof (due to Kelly). Suppose for contradiction that this is false, and let P be the set of points, and let L be the set of lines they define. Consider the set of pairs

$$\{(p, \ell) \in P \times L : p \notin \ell\}.$$

By the assumption of non-collinearity, this set is non-empty. So pick p, ℓ with $p \notin \ell$ so that the distance between p and ℓ is minimized. Drop a perpendicular from p to ℓ. By assumption, ℓ contains at least three points of P, say a, b, c, and say they appear in this order. By the pigeonhole principle, two of these points must be on the same side of the perpendicular from p. Say that these are b, c. Then the pair of points $p, c \in P$ define another line $\ell' \in L$, and observe that the distance from b to ℓ' is strictly smaller than the distance from p to ℓ (this is intuitively clear from the picture, and can be proven by observing that we have two similar triangles, one of which is contained in the other). This contradicts the minimality of the pair (p, ℓ). □

Observe that this proof uses various properties of Euclidean geometry, or more precisely of the field $\mathbb{R}$. Namely, we use the fact that $\mathbb{R}$ is an ordered field to order the points a, b, c, and to talk about "being on one side" of the perpendicular, and we use the fact that $\mathbb{R}^2$ has a well-defined (Euclidean) distance. Nevertheless, Sylvester's question is purely a question of points and lines, so we can ask it inside $\mathbb{F}^2$, where $\mathbb{F}$ is any field.

When we do this, we discover that the properties of $\mathbb{R}$ we used aren't an artifact of the proof, since the Sylvester–Gallai theorem is false in many other fields. For instance, if $\mathbb{F}$ is a finite field $\mathbb{F}_q$ with $q > 2$, then we can take our set of points $P \subseteq \mathbb{F}_q^2$ to simply be *all* of $\mathbb{F}_q^2$, at which point Sylvester's property is satisfied (since every line contains $q \geq 3$ points). As another example, suppose $\mathbb{F} = \mathbb{C}$. Then it turns out we can again construct a set of points n points so that every line they define contains at least three of them. To do this, let E be an elliptic curve. Then it turns out that the set of $\mathbb{C}$-points of E defines a torus $S^1 \times S^1$. Moreover, it turns out that this set of points can be naturally associated with a group structure, also isomoprhic to the group $S^1 \times S^1$ (where $S^1 \cong SO(2)$ is the circle group), where the group law satisfies the property that $x + y + z = 0$ if and only if x, y, z are collinear. This fact implies that if we take the unique subgroup of E isomoprhic to $C_3 \times C_3$ (the 3-torsion points of E), then any line passing through two of them will pass through a third, and these points will not be collinear. (Technically, one of the points in this configuration will be the point at infinity in the complex projective plane, but we can apply a projective transformation to move this point into the affine plane $\mathbb{C}^2$.)

Another way to generalize Sylvester's question is by asking about higher dimensions. Namely, can we put n points in $\mathbb{R}^m$ so that any time a line contains two of them, it contains a third? The answer is again no, as shown by the following simple argument. Suppose we had such a configuration, and pick a uniformly random $\mathbb{R}^2 \subset \mathbb{R}^m$. If we project the entire configuration onto this plane, then with probability 1, no two points will collide, and no point will project onto a line it was not previously on. Thus, we will obtain a configuration in the plane contradicting the Sylvester–Gallai theorem.

Similarly, we can ask the higher-dimensional question over other fields. Here, the most interesting result is due to Kelly, who proved that any Sylvester configuration in $\mathbb{C}^m$ must actually lie in some two-dimensional plane. In other words, in order to make the Sylvester–Gallai theorem true over $\mathbb{C}$, we need only require that the points are non-collinear *and* non-coplanar. On the other hand, no such modification is possible over $\mathbb{F}_q$, since if we take all the points in $\mathbb{F}_q^m$ for any m, this will yield a set of points contradicting Sylvester–Gallai.

A closely related result is the following theorem, due to de Bruijn and Erdős in 1948.

Theorem. *Any n non-collinear points in the plane define at least n lines.*

Proof. We proceed by induction on n. The base case is $n = 3$; in this case, 3 non-collinear points define a triangle, and thus three lines. For the inductive case, suppose we have a set P of n points in the plane. By the Sylvester–Gallai theorem, there are two points, say $a, b \in P$, whose line contains no other point of P. Delete a from the configuration; if $P \setminus \{a\}$ is non-collinear, then by induction it spans at least $n - 1$ lines, which don't include the line containing a, b. Adding a and this line back gives us at least n lines. On the other hand, if $P \setminus \{a\}$ is collinear, then a must be off this line by assumption; then it defines a distinct line with each point on $P \setminus \{a\}$, and together with the collinear line, this gives the desired n lines. □

Since this proof uses the Sylvester–Gallai theorem, you might think that it is also only true in $\mathbb{R}^2$. However, this is not at all the case; in fact, this theorem is true over every field, which follows from the fact that de Bruijn and Erdős found a purely combinatorial proof of this fact.

Proof. Let the points be P and the lines they define L, with $|P| = n, |L| = m$. We wish to prove that $m \geq n$. Let $a_1, \ldots, a_n$ be the number of lines containing the points $p_1, \ldots, p_n$, respectively, and let $b_1, \ldots, b_m$ denote the number of points on the lines $\ell_1, \ldots, \ell_m$. Then observe that

$$\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} b_j,$$

since both these sums just count the number of incidences between P and L. Moreover, observe that if p_i is not on the line ℓ_j, then

$$a_i \geq b_j.$$

This is because, for every point on the line ℓ_j, it defines a distinct line through the point p_i.

Now, suppose WLOG that a_n is minimized among all the a_i, and set $x = a_n$; we have that $x \geq 2$ since we assumed the points were non-collinear. Let the lines through a_n be $\ell_1, \ldots, \ell_x$; for every such line, there is some other point on it, say $p_1, \ldots, p_x$. Note that these points are all distinct (otherwise two lines would intersect at two points), so we can conclude that

$$a_1 \geq b_2 \qquad a_2 \geq b_3 \qquad \cdots \qquad a_{x-1} \geq b_x \qquad a_x \geq b_1.$$

Thus,

$$\sum_{i=1}^{x} a_i \geq \sum_{j=1}^{x} b_j.$$

Additionally, for every $i > x$, we have by minimality that $a_i \geq a_n$. Moreover, since a_n is not on the line ℓ_j for $j > x$, we also have that $a_n \geq b_j$ for $j > x$. Therefore,

$$\sum_{i=1}^{n} a_i \geq \sum_{i=1}^{x} a_i + \sum_{i=x+1}^{n} a_n \geq \sum_{j=1}^{x} b_j + \sum_{j=x+1}^{n} b_j = \sum_{j=1}^{n} b_j$$

Thus, in order to get the full sums to be equal, we must have $m \geq n$. □

Observe that the only property of points and lines we used in this proof is the fact that any two points define a line, and any two lines intersect at most once. Thus, this theorem is true over any field, but even in greater generality: for any set of objects which we call points, and any set of their subsets which we call lines, as long as this property is satisfied, the de Bruijn–Erdős theorem is true.

Note that the de Bruijn–Erdős theorem is sharp, as shown by the example where all but one point are collinear, which defines exactly n lines. In fact, de Bruijn and Erdős also proved that the only configurations where this theorem is tight are that one and a *finite* *projective plane*, namely a configuration where all lines contain the same number of points, and all lines intersect exactly once. Finite projective planes include all projective planes over finite fields, but there are others as well, and in general understanding which finite projective planes exist is a huge and very difficult open problem.

2 The Erdős unit distance problem

In 1946, Erdős asked the following innocuous-seeming problem, which is one of my favorites in all of math. Let $u(n)$ denote the maximal number of unit distances that can be defined by n points in the plane, namely

$$u(n) = \max_{\substack{P \subseteq \mathbb{R}^2 \\ |P| = n}} |\{a, b \in P : \|a - b\| = 1\}|.$$

Erdős wanted to know how fast $u(n)$ grows as a function of n.

Let's begin by thinking about lower bounds. A pretty dumb thing to do is to put all the n points on a line, spaced 1 apart from each other. Then this will define $n-1$ unit distances, so we see that $u(n)$ grows at least linearly in n. This configuration is dumb, however, since it doesn't use the fact that we're in two dimensions. Another natural thing to do is to place the n points on a $\sqrt{n} \times \sqrt{n}$ unit grid. Then (almost) every point will be at unit distance from four others, so this configuration will define roughly $2n$ unit distances (in reality, it'll be a bit less because the points on the boundary of the grid will contribute a bit less, but it's roughly correct). We can improve this to roughly $3n$ by taking a triangular grid instead of a square one, but it's not clear how to push past linear growth. However, Erdős came up with the following trick. Suppose that instead of putting our points on a $\sqrt{n} \times \sqrt{n}$ *unit* grid, we instead made the distances between adjacent points $1/5$. Then each point will be at unit distance from 12 others, instead of just 4, namely the four points five steps away horizontally or vertically, plus the eight points that are $(4, 3)$ away. Thus, this configuration will define roughly $6n$ unit distances. Even better, if we make the adjacent distances $1/\sqrt{65}$, then every point will be at unit distance from 16 others, namely those defined by the distances $(8, 1)$ and $(7, 4)$, and thus we will get roughly $8n$ unit distances. More generally, if an integer x can be expressed as a sum of squares in r ways, then if we rescale the grid by $1/\sqrt{x}$, every point will be at unit distance from at least r others, and we will define roughly $(r/2)n$ unit distances.

So now we have a number theory problem; among the integers less than n, which one can be expressed as a sum of squares in the largest number of ways? Using results due to Fermat and Lagrange, Erdős was able to show that there is always an $x \leq n$ that can be expressed as a sum of squares in at least $n^{c/\log\log n}$ distinct ways, for some $c > 0$. Therefore, Erdős proved that

$$u(n) \geq n^{1+c/\log\log n}.$$

He conjectured that this was close to correct; usually, his conjecture is stated as follows.

Conjecture (Erdős, 1946). *$u(n) = n^{1+o(1)}$, namely $u(n) = O(n^{1+\varepsilon})$ for all $\varepsilon > 0$.*

For the upper bound, we can first observe that $u(n) \le \binom{n}{2} = O(n^2)$, since there are only $\binom{n}{2}$ pairs of points among n points. This bound is obviously pretty bad, since it doesn't use the geometry at all. Erdős was able to do better.

Theorem (Erdős, 1946)**.** $u(n) = O(n^{3/2})$.

Proof. Fix a set of n points in the plane, and draw a unit circle around each one; let P be the set of points and C the set of circles. Denote by $I(P, C)$ the number of incidences between a point of P and a circle of C; then the number of unit distances defined by P is just $I(P, C)/2$, so it suffices to upper-bound $I(P, C)$. We can write

$$I(P,C) = \sum_{p\in P}\sum_{c\in C} \mathbf{1}_{p\in c},$$

where $\mathbf{1}_{p\in c}$ is the indicator function that is 1 if $p \in c$ and 0 otherwise. By Cauchy–Schwarz, this is at most

$$\begin{aligned} I(P,C) &= \sum_{p\in P}\sum_{c\in C} \mathbf{1}_{p\in c} \\ &\le \left(\sum_{p\in P}\left(\sum_{c\in C}\mathbf{1}_{p\in c}\right)^2\right)^{1/2}\left(\sum_{p\in P} 1^2\right)^{1/2} \\ &= \sqrt{n}\left(\sum_{p\in P}\sum_{c,c'\in C}\mathbf{1}_{p\in c}\mathbf{1}_{p\in c'}\right)^{1/2} \\ &\le \sqrt{n}\left(\sum_{c,c'\in C} 2\right)^{1/2} \\ &= \sqrt{2}\cdot n^{3/2} \end{aligned}$$

where we use the fact that any two fixed circles c, c' intersect at most twice, and in particular both contain at most two distinct points of P. □

Observe that this argument only uses the fact that two distinct circles intersect at most twice. This property is true not just in $\mathbb{R}^2$, but also in $\mathbb{F}_q^2$, if we interpret a unit circle around (x_0, y_0) to be the solution set of $(x - x_0)^2 + (y - y_0)^2 = 1$. Thus, the $O(n^{3/2})$ upper bound is also true for the unit distance problem over $\mathbb{F}_q$; moreover, there it is tight. Indeed, if we take P to be all the points in $\mathbb{F}_q^2$, then $n = q^2$, and every point is a unit distance away from roughly q others (since the circle is a one-dimensional variety, so its size is roughly q). Thus, there will be roughly $q^2/2$ unit distances, which is $Cn^{3/2}$ for some constant C.

Therefore, to improve the upper bound, we will need to use further properties of Euclidean geometry, in addition to the fact that circles intersect at most twice. The next major improvement on the upper bound was the following theorem.

Theorem (Spencer–Szemerédi–Trotter, 1984)**.** $u(n) = O(n^{4/3})$.

You can see that the sequence of exponents goes $2/1, 3/2, 4/3$, so it's natural to assume the next result is $O(n^{5/4})$. However, we're still very far from proving that. Instead, the next two improvements were

Theorem (Székely, 1997)**.** *$u(n) = O(n^{4/3})$.*

Theorem (Pach–Tardos, 2006)**.** *$u(n) = O(n^{4/3})$.*

The astonishing thing is that all three of these proofs use very different techniques. The original proof of Spencer–Szemerédi–Trotter used a planar partitioning argument; these arguments, which come up all over discrete geometry, say that it is possible to partition the plane into regions that behave nicely with respect to our point configuration. Székely's proof uses the *crossing number lemma*, a beautiful tool at the intersection of topology and graph theory. Finally, Pach and Tardos used a "purely combinatorial" argument, based on the Turán problem for ordered graphs (of course, it's not truly purely combinatorial, as the geometric structure of $\mathbb{R}^2$ must come into play at some point, but its entry in this argument is very minimal).

Székely's proof. For any graph G, its crossing number $\operatorname{cr}(G)$ is defined to be the minimal number of edge crossings among all drawings of G in the plane. Thus, G is planar if and only if $\operatorname{cr}(G) = 0$. Our main tool is the following.

Lemma (Ajtai–Chvátal–Newborn–Szeméredi, 1982; Leighton, 1983)**.** *If G is a graph on n vertices and $m \geq 4n$ edges, then*

$$\operatorname{cr}(G) \geq \frac{m^3}{64n^2}.$$

Proof. A simple consequence of Euler's formula $V - E + F = 2$ implies that a planar with n vertices and m edges has $m \leq 3n$. Therefore, in any graph G with n vertices and m edges, we have

$$\operatorname{cr}(G) \geq m - 3n,$$

since we may repeatedly delete an edge involved in a crossing, and we will need to do this at least $m - 3n$ times to get a planar graph. Now, let G be the graph we are interested in. For some $p \in (0, 1)$, we define a random subgraph G_p by keeping each vertex independently with probability p, and keeping every edge if both its endpoints are kept. Let $n(G_p), m(G_p)$ denote the number of vertices and edges of G_p. Then observe that

$$\mathbb{E}[n(G_p)] = pn \qquad \mathbb{E}[m(G_p)] = p^2 m \qquad \mathbb{E}[\operatorname{cr}(G_p)] \leq p^4 \operatorname{cr}(G),$$

by linearity of expectation. Observe that we only get an inequality for $\mathbb{E}[\operatorname{cr}(G_p)]$; indeed, if we fix a drawing of G, then every crossing will survive with probability p^4, but there might be an even better drawing of this graph, leading to the inequality. Therefore, we find that

$$p^4 \operatorname{cr}(G) \geq p^2 m - 3pn,$$

for any $p \in (0, 1)$. Plugging in $p = 4n/m$ gives the desired result. □

Now, suppose we have a configuration of n points in the plane. Draw a unit circle around each one. We now define a graph G on these n points by connecting two of them if they appear sequentially on one of these unit circles. Then G has m edges, where m is the number of incidences between points and circles, which is in turn twice the number of unit distances defined by these n points, so it suffices to upper-bound m. By the crossing number lemma, we have that

$$\operatorname{cr}(G) \geq \frac{m^3}{64n^2}.$$

On the other hand, any two circles can intersect at most twice, and we have n circles, so this drawing of G yields at most $2\binom{n}{2} \leq n^2$ crossings. Thus, $\operatorname{cr}(G) \leq n^2$. Plugging this in, we find that

$$m^3 \leq 64n^2 \operatorname{cr}(G) \leq 64n^4$$

which yields the bound $m = O(n^{4/3})$, as desired. □

It's natural to ask why 4/3 appears to be such a sticking point. One good reason is that all three of the proofs cited above actually work in further generality; namely, they apply when the metric on $\mathbb{R}^2$ is given by any norm, not just by the Euclidean one. This can be seen above in Székely's proof, which only used the fact that two circles intersect at most twice, which is true for any convex curve, and thus for the unit circle of any norm on $\mathbb{R}^2$. Moreover, Valtr observed that if we define our norm to have as its unit circle the curve $|y| = 1 - x^2$, then the grid

$$\left\{ \left(\frac{i}{k}, \frac{j}{k^2} \right) : |i| \leq k, |j| \leq k^2 \right\}$$

will have $4k^3$ points and roughly k^4 unit distances in this metric (since every point will be at unit distance from roughly k others). Setting $n = 4k^3$ shows that in this metric, the $n^{4/3}$ bound is tight. So in order to improve the upper bound further, one would need to find an argument that is able to distinguish the Euclidean metric from other similar metrics such as this one; alternatively, Erdős's conjecture might be false, and there may be a set of n points in the plane defining $Cn^{4/3}$ unit Euclidean distances.

As a final remark on related problems, note that we may ask the unit distance problem in any dimension. In $\mathbb{R}^3$, Erdős used a similar number-theoretic argument (rescaling the three-dimensional $n^{1/3} \times n^{1/3} \times n^{1/3}$ grid) to find a set of points that define $Cn^{4/3} \log \log n$ unit distances. The best upper bound is $O(n^{3/2})$, so there is still a large gap here. However, for dimensions $d \geq 4$, the problem becomes pretty uninteresting. Indeed, in $\mathbb{R}^4$, one may find two orthogonal circles of radius $1/\sqrt{2}$. Note that by the Pythagorean theorem, any point on one circle is at unit distance from any point on the other. So by putting $n/2$ points on each circle, we may produce $n^2/4$ unit distances, meaning that the trivial quadratic upper bound is correct, up to the constant. Moreover, using a bit more work, Erdős was actually able to pin down the precise constant in all dimensions above 3.

3 The Kakeya conjecture

In 1917, Sōichi Kakeya asked the following question.

Question (Kakeya)**.** *What is the smallest area of a subset of* $\mathbb{R}^2$ *in which you can turn a segment of length* 1 *through a full* 360° *rotation?*

We call such a set a *Kakeya set*. Kakeya conjectured that the optimal shape was the so-called "three-pointed deltoid", which has area $\pi/8$. However, this guess was not only wrong, but roughly as wrong as can be.

Theorem (Besicovitch, 1928)**.** *For every* $\varepsilon > 0$*, there is a Kakeya set* $K \subset \mathbb{R}^2$ *with area at most* ε*.*

Moreover, it turns out that this theorem cannot be improved, in the sense that every Kakeya set has positive area. However, we can weaken our requirements; let's call a subset of $\mathbb{R}^2$ a *Besicovitch set* if it contains a unit line segment in every direction, without requiring the ability to turn this segment. With this definition, we actually can do better.

Theorem (Besicovitch, 1919)**.** *There is a Besicovitch set* $B \subset \mathbb{R}^2$ *with measure zero.*

This seems basically as good as we can do. However, there are different types of measure-zero sets in the plane, and a useful way to distinguish measure-zero sets is by their dimension; for our purposes, we will focus on the Minkowski (or box-counting) dimension. For a bounded set $S \subseteq \mathbb{R}^n$ and $\varepsilon > 0$, let $N(S, \varepsilon)$ denote the number of boxes S intersects when we partition $\mathbb{R}^n$ into a grid of boxes of side length ε. Then the Minkowski dimension of S is defined by

$$\dim_M(S) = \lim_{\varepsilon \to 0} \frac{\log N(S,\varepsilon)}{\log \frac{1}{\varepsilon}},$$

assuming this limit exists. The intuition is that an d-dimensional set S should intersect roughly $C(1/\varepsilon)^d$ boxes, for some constant C, so this expression will recover the d. With this definition, we can state the modern form of the Kakeya conjecture; similarly to before, a Besicovitch set in $\mathbb{R}^n$ will be a subset of $\mathbb{R}^n$ containing a unit line segment in every direction.

Conjecture. *Every Besicovitch set in* $\mathbb{R}^n$ *has (Minkowski) dimension* n*.*

In other words, the Kakeya conjecture says that, although Besicovitch sets can have measure zero, they still need to be "as big as possible" among all measure-zero sets. It turns out that this conjecture also has many close connections to harmonic analysis, and much of the interest in it today comes from analysts.

There is also a natural definition of Besicovitch sets over finite fields.

Definition. A set $B \subseteq \mathbb{F}_q^n$ is called a *Besicovitch* set if it contains a line in every direction. Namely, for all $m \in \mathbb{F}_q^n \setminus \{0\}$, there is some $b \in \mathbb{F}_q^n$ so that $b + tm \in B$ for all $t \in \mathbb{F}_q$.

Suppose we partition the interval $[0,1]$ into q subintervals of equal length, and use this to define a partition of $[0,1]^n$ into q^n boxes. If we put a point in the center of each box, then we can pretend that $\mathbb{F}_q^n$ is a discrete approximation to $[0,1]^n$, and we can guess that this approximation gets better and better as q gets larger. In particular, we might hope that as q gets larger, a Besicovitch set in $\mathbb{F}_q^n$ looks more and more like a Besicovitch set in $[0,1]^n$.

If we think of a set $S \subseteq [0,1]^n$ as coming from a subset of $\mathbb{F}_q^n$, again imagined as living inside $[0,1]^n$, then we have that $N(S,1/q) = |S|$, since the number of boxes intersecting S is precisely the number of elements of $\mathbb{F}_q^n$ in S. Therefore, if we believe the Kakeya conjecture, that $\dim_M B = n$ for any Besicovitch set $B \subseteq \mathbb{R}^n$, we might hope that something like the following holds.

$$n = \dim_M B = \lim_{q\to\infty} \frac{\log |B_q|}{\log q},$$

where $B_q \subseteq \mathbb{F}_q^n$ is a Besicovitch set in $\mathbb{F}_q^n$. Rearranging this gives us the following guess.

Conjecture (Finite Field Kakeya Conjecture). *For every n, there is some constant C_n so that for any q and any Besicovitch set $B_q \subseteq \mathbb{F}_q^n$, we have*

$$|B_q| \geq C_n q^n$$

Note that if this conjecture is true, then we indeed have that

$$\lim_{q\to\infty} \frac{\log |B_q|}{\log q} \geq \lim_{q\to\infty} \left(\frac{\log C_n}{\log q} + \frac{n \log q}{\log q}\right) = n,$$

as our heuristic argument above suggested. This Finite Field Kakeya Conjecture was first conjectured by Wolff in 1999, and his idea was that it might serve as another regime where ideas for the real Kakeya conjecture could be tested. There is no formal reduction from one conjecture to the other (all our arguments above were purely heuristic, and cannot be turned into proofs), but the hope was that understanding one problem would help us understand the other. For about a decade, any time someone made an advance towards solving either the Kakeya Conjecture or the Finite Field Kakeya Conjecture, some work very quickly followed that got the same result for the other conjecture. Moreover, the relationship between these two conjectures allowed new ideas to come into play. But this changed when Zeev Dvir shocked everyone by proving the Finite Field Kakeya Conjecture, using a proof technique that seems impossible to adapt to the real case.

Theorem (Dvir, 2008). *For every n, q, and every Besicovitch set $B \subseteq \mathbb{F}_q^n$, we have*

$$|B| \geq \binom{q+n-1}{n} \geq \frac{1}{n!} q^n$$

Thus, the Finite Field Kakeya Conjecture is true with $C_n = 1/n!$.

Proof. Dvir's technique is the so-called *polynomial method*, whose basic mantra is "a set is small if and only if a non-zero low-degree polynomial vanishes on it". For this proof, the following precise version of this mantra will suffice.

Proposition. *Let $\mathbb{F}$ be any field and $d, n \in \mathbb{N}$, and let $T \subseteq \mathbb{F}^n$ be any set so that $|T| < \binom{d+n}{n}$. Then there is a non-zero polynomial $P(x_1, \ldots, x_n)$ of degree at most d such that P vanishes on T.*

This proposition is easy to prove with basic linear algebra. The space of polynomials of degree at most d has dimension $\binom{d+n}{n}$, and requiring a polynomial to vanish at a point is a linear condition on its coefficients; so as long as $|T| < \binom{d+n}{n}$, the space of polynomials that work has positive dimension, so in particular we can find a non-zero one.

Now, suppose for contradiction that we had a Besicovitch set $B \subseteq \mathbb{F}_q^n$ with

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

By this proposition, this implies that there is some non-zero polynomial $P(x_1, \ldots, x_n)$ with coefficients in $\mathbb{F}_q$ and degree at most $q-1$ with the property that P vanishes on B. Suppose $\deg P = d \leq q-1$, and write

$$P = \sum_{i=0}^{d} P_i$$

where each P_i is a homogeneous polynomial of degree i. Since $\deg P = d$, we know that P_d is not the zero polynomial (for otherwise the degree would be strictly smaller).

For any $0 \neq m \in \mathbb{F}_q^n$, we know that B contains a line in the direction of m, namely there is some $b \in \mathbb{F}_q^n$ so that $b + tm \in B$ for all $t \in \mathbb{F}_q$. Define a new single-variate polynomial $Q_m(t)$ by

$$Q_m(t) = P(b + tm).$$

Since we are just plugging in values to P, we find that $\deg Q_m \leq \deg P \leq q-1$. On the other hand, for any value of t, we have that $b + tm \in B$, so $P(b + tm) = 0$. Thus, $Q_m(t) = 0$ for every $t \in \mathbb{F}_q$, so Q_m has at least q roots. Since $\deg Q_m \leq q-1$, this implies that Q_m is the zero polynomial. Thus, in particular, the coefficient of t^d in $Q_m(t)$ is zero. However, the coefficient of t^d in $Q_m(t)$ is precisely the value of $P_d(m)$. So we find that $P_d(m) = 0$ for every $0 \neq m \in \mathbb{F}_q^n$. Moreover, since P_d is homogeneous of degree d, this implies that in fact, P_d vanishes on all of $\mathbb{F}_q^n$. Finally, since $d < q = |\mathbb{F}_q|$, this implies that P_d must in fact be the zero polynomial. This is a contradiction. □

References

[1] Z. Dvir, Incidence theorems and their applications, *Found. Trends Theor. Comput. Sci.*, 6(4):(2010), 257–393.

[2] J. Matoušek, *Lectures on discrete geometry*, *Graduate Texts in Mathematics*, vol. 212, Springer-Verlag, New York, 2002.

[3] E. Szemerédi, Erdős's unit distance problem, in *Open problems in mathematics*, Springer, [Cham], 2016, pp. 459–477.