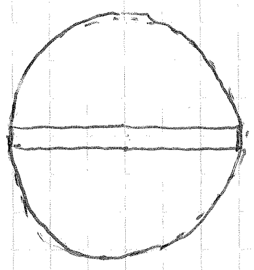
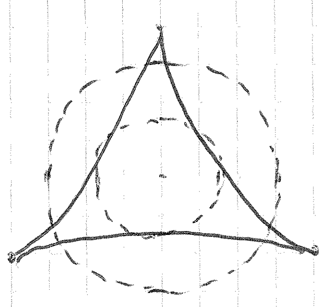
Kakeya 1917:
What is the smallest area required to rotate a unit line segment (a "needle") by 180 degrees in the plane?

<u>Kakeya needle problem</u>



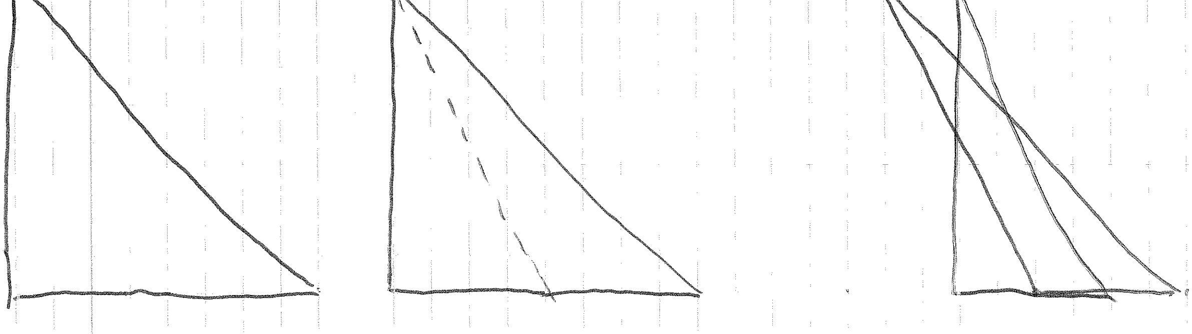
Rotate around midpoint:
Area $\pi r^2 = \pi/4$



"Three-point U-turn" or deltoid
$\pi/8$

[Besicovitch '27]: Area can be <u>arbitrarily small</u>

Can construct BESICOVITCH SETS e.g. by starting with triangle, bisect, shift together to



decrease area, and repeat...

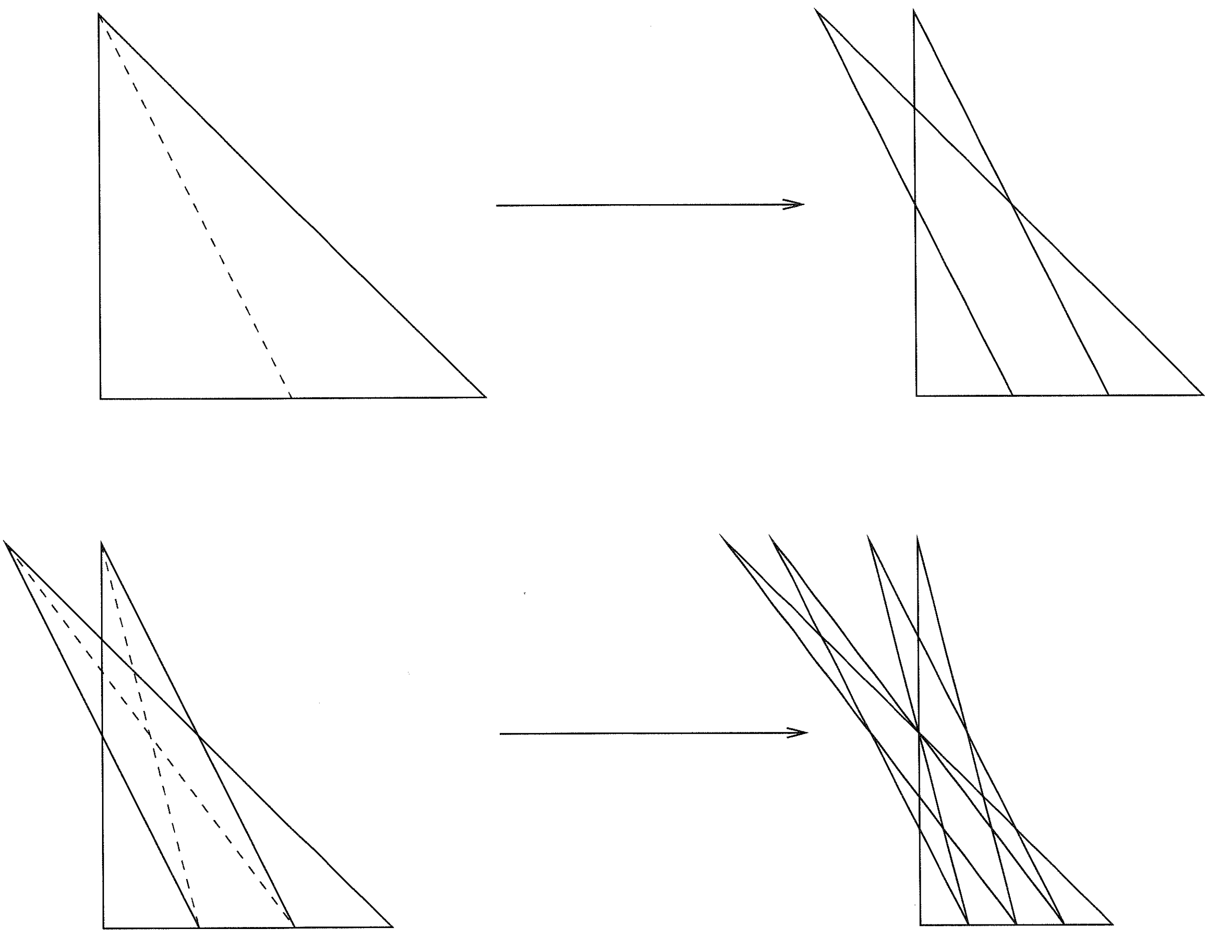


Figure 2: The iterative construction of a Besicovitch set. Each stage consists of the union of triangles. To pass to the next stage, the triangles are bisected and shifted together to decrease their area.

Just a mathematical curiosity? Not quite.
In last few decades connections discovered between this type of problem and:

- number theory,
- geometric combinatorics
- arithmetic combinatorics
- oscillatory integrals,
- analysis of dispersive and wave equations

Generalizing Kakeya's question:

For any $n \geq 2$, define a KAKEYA (or BESICOVITCH) SET to be a subset of $\mathbb{R}^n$ containing a unit length segment in every direction.

Besicovitch showed such sets can have arbitrarily small measure in any dimension.

More refined question: Consider space to be discretized and ask about the HAUSDORFF or MINKOWSKI DIMENSION of a Kakeya set. (Not important what this is)

KAKEYA CONJECTURE: Kakeya sets must have Hausdorff or Minkowski dimension n.

Well-known open problem
Resolved for $n=2$ in 1971
Open in larger dimensions.
How to make progress? Maybe gain insights by studying similar problem in different model?

Maybe not for $\mathbb{R}$ but for other fields?
As usual, $\mathbb{F}_q$ denotes finite ~~set~~ field with q elements

<u>DEF 1</u> A KAKEYA SET (or BESICOVITCH SET) is a set $K \subseteq \mathbb{F}_q^n$ s.t. K contains a line in every direction. That is, for every $x \in \mathbb{F}_q^n$ there exists a $y \in \mathbb{F}_q^n$ s.t.

$$L_{y,x} = \{ y + ax \mid a \in \mathbb{F}_q \} \subseteq K.$$

Wolff '99:
- Suggested to study finite field Kakeya sets
- Asked if $|K| \geq C_n \cdot q^n$ (for C_n depending only on n)
- Proved $|K| \geq C_n \cdot q^{\frac{n+2}{2}}$

$\Longrightarrow$ Do warm-up $|K| \geq q^{n/2}$

Later works improved this to $|K| \geq C_n q^{4n/7}$

Contributions by math superstars such as Bourgain and Tao.
Heavy-duty machinery from additive number theory. [Studies questions like Given $A, B \subseteq \mathbb{F}_q^n$, how large is set $A + B = \{ a+b \mid a \in A, b \in B \}$]

Idea to use additive number theory for Kakeya sets from Bourgain '99.

Additive number theory has turned out to be very useful in several areas of TCS — could be good topic for future PhD course.

WARM-UP

If K is a Kakeya set, then for every $x \in \mathbb{F}_q^n$ there is a $y \in \mathbb{F}_q^n$ such that

$$\ell_{y,x} = \{ y + ax \mid a \in \mathbb{F}_q \} \subseteq K$$

So K "contains lines in q^n different directions"

$\ell_{y,x}$ determined by pair $(y, y+x)$ and K contains a full line for each such pair - q^n distinct pairs, since $(y+x)-y$ determines x

The points in K can yield at most $|K|^2$ such pairs, so

$$|K|^2 \geq q^n$$

or

$$|K| \geq q^{n/2}$$

But not needed for Wolff's question whether $|K| \geq C_n \cdot q^n$ – this can be resolved by elementary means.

<u>Rough idea:</u>

Fix any Kakeya set $K \subseteq \mathbb{F}_q^n$

Then any $g \in \mathbb{F}_q[x_1, \ldots, x_n]$, $\deg(g) = q-2$, g homogeneous (i.e., all monomials have same total degree) can be reconstructed from its values ~~of~~ on K.

But then $|K|$ had better be as large as the dimension of this space of polynomials, which is roughly q^{n-1}

Proof by Zeev Dvir in 2008

Strikingly simple and came "out of the blue"

Article in Journal of the AMS 4½ pages, out of which 1 page intro & 1 page references.

Main theorem proves robust version of lower bounds for "somewhat Kakeya" sets

<u>DEF 2</u> $K \subseteq \mathbb{F}_q^n$ is a (δ, γ)-KAKEYA SET if exists set $\mathcal{L} \subseteq \mathbb{F}_q^n$ (of directions) of size $|\mathcal{L}| \geq \delta \cdot q^n$ 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, i.e., $\exists y$ s.t.

$$|\mathcal{L}_{y,x} \cap K| = \gamma q.$$

Note that a Kakeya set is a $(1,1)$-Kakeya set.

Dvir's main theorem is as follows.

<u>THM 3</u> If $K \subseteq \mathbb{F}_q^n$ is a (δ,γ)-Kakeya set, then

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

where

$$d = \lfloor q \cdot \min(\delta,\gamma) \rfloor - 2.$$

Let us postpone the proof of Thm 3 and instead see how to use it.

<u>THM 4</u> If $K \subseteq \mathbb{F}_q^n$ is a Kakeya set, then

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

where C_n depends only on n (and not on q).

<u>Proof</u> By Thm 3, the $(1,1)$-Kakeya set K must have size $|K| \geq \binom{q-2+n-1}{n-1}$.

Using $\binom{n}{k} \geq \left(\frac{n}{k}\right)^k$ we get

$$|K| \geq \left(\frac{q+n-3}{n-1}\right)^{n-1} \geq \left(\frac{1}{n-1}\right)^{n-1} \cdot q^{n-1} \qquad \blacksquare$$

[Assuming $n \geq 3$; also true for $n=2$ with special fix any the case of 1-dimensional Kakeya sets is left as an exercise]

This is still a factor q off from Wolff's question. Is it possible to do better?

Suppose that $K \subseteq \mathbb{F}_q^n$ Kakeya set. Then the Cartesian product $K^r = \{(y_1, y_2, \dots, y_r) \mid y_i \in K\}$ is a Kakeya set in $\mathbb{F}_q^{nr}$.

[Given a direction $x \in \mathbb{F}_q^{nr}$, split into $x_1, \dots, x_r$; $x_i \in \mathbb{F}_q^n$. For every such x_i there is a point y_i s.t. $L_{y_i, x_i} \subseteq K$.
Then
$$L_{(y_1, \dots, y_r), (x_1, \dots, x_r)} \subseteq K^r \quad]$$

This yields the following corollary

<u>COROLLARY 5</u>

For every $n \in \mathbb{N}^+$ and every $\varepsilon > 0$ there exists a constant $C_{n,\varepsilon}$, depending only on n and ε, such that any Kakeya set ~~in~~ $K \subseteq \mathbb{F}_q^n$ satisfies
$$|K| \geq C_{n,\varepsilon} \cdot q^{n-\varepsilon}$$

<u>Proof</u>: Use Thm 4 on the Kakeya set $K^r \subseteq \mathbb{F}_q^n$. We get
$$|K^r| = |K|^r \geq C_{n \cdot r} \cdot q^{n \cdot r - 1}$$
or
$$|K| \geq \sqrt[r]{C_{n,r}} \cdot q^{n - 1/r} \qquad \blacksquare$$

This sort of suggests that the right answer should really be $|K| \geq c_n \cdot q^n$.

After Dvir's initial result, Noga Alon and Terence Tao independently observed that Dvir's technique can be tweaked to obtain:

<u>THM 6</u> If $K \subseteq \mathbb{F}_q^n$ is a Kakeya set, then $|K| \geq C_n \cdot q^n$ where C_n depends only on n.

Our plan today:

1. First prove Thm 3
2. Then see how to modify argument to get Thm 6.

Main technical tool: Schwartz-Zippel

<u>LEM 7</u> For a non-zero $f \in \mathbb{F}_q[x_1, \ldots, x_n]$ with $\deg(f) \leq d$ it holds that

$$\left| \{ x \in \mathbb{F}_q^n : f(x) = 0 \} \right| \leq \frac{d}{q} |\mathbb{F}_q^n| = d \cdot q^{n-1}$$

Let us now prove Thm 3

Suppose (in order to derive contradiction) that K is a (δ, γ)-Kakeya set with

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

<u>Claim A</u> The number of monomials of total degree exactly d is $\binom{d+n-1}{n-1}$

<u>Proof</u> Consider $d+n-1$ balls

○ ○ ○ ○ ○ ○ ○ ○ $\quad n=5$, $d=4$

Choose $n-1$ balls to colour black

○ ● ● ○ ○ ● ○ ●

Let distance (# white balls) between $(i-1)$st and ith ball $=$ degree of x_i

$$x_1 x_3^2 x_4$$

d white balls left $\Rightarrow$ degree exactly d

Given any degree-d monomial, get valid colouring as above — map bijective $\boxed{}$

<u>Claim B</u> There is a non-zero homogeneous degree-d polynomial $g \in \mathbb{F}_q[x_1, \dots, x_n]$ such that

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

<u>Proof</u> Fix $x \in K$. We want to find

$$g(y) = \sum_{\substack{i_1+\dots+i_n=d \\ i_j \geq 0}} c_{i_1,\dots,i_n} y_1^{i_1} \cdots y_n^{i_n}$$

such that for this $x \in K$

$$\sum_{i_1+\dots+i_n=d} c_{i_1,\dots,i_n} x_1^{i_1} \cdots x_n^{i_n} = 0$$

This is a linear equation in variables $c_{i_1, \dots, i_n}$, which in addition is homogeneous

We have:

- $\binom{d+n-1}{n-1}$ variables
- less than $\binom{d+n-1}{n-1}$ equations (one for each $x \in K$).

Hence there is a non-trivial solution g.

[We have equation system

$$\begin{array}{cc} & n \\ m & \boxed{\quad A \quad} \end{array} \; \vec{x} = \vec{A_1} x_1 + \vec{A_2} x_2 + \dots + \vec{A_n} x_n = \vec{0}$$

and n vectors in a space of dimension $m < n$ can't be linearly independent. Hence $\exists$ non-trivial linear combination summing to zero]

Since K is a (δ, γ)-Kakeya set, there exists $\mathcal{L} \subseteq \mathbb{F}_q^n$, $|\mathcal{L}| \geq \delta \cdot q^n$, such that for every $z \in \mathcal{L}$ there exists a line with direction z that intersects K in at least $\gamma \cdot q$ points. That is, for every $z \in \mathcal{L}$ can find y such that for

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

we have

$$| L_{y,z} \cap K | \geq \gamma \cdot q$$

We want to show that g has too many zeros to be non-zero. In particular $g(z)$ for all $z \in \mathcal{L}$

<u>Claim C (main)</u>

For every $z \in \mathcal{L}$ we have $g(z) = 0$

Before proving Claim C, let's see why we're now done.

g evaluates to zero on $|\mathcal{L}| \geq \delta \cdot q^n$ points

g is a non-zero polynomial of degree d.

Looking at the statement of Thm 3, we have

$$d = \lfloor q \cdot \min(\delta, \gamma) \rfloor - 2$$

and hence in particular

$$d \leq q \cdot \delta - 2$$

$$\frac{d}{q} < \delta$$

So now we have a non-zero degree-d polynomial $g \in \mathbb{F}_q[x_1, \ldots, x_n]$ which is zero on $\delta \cdot q^n > d\, q^{n-1}$ points, i.e., on a fraction $\delta > \frac{d}{q}$ of $\mathbb{F}_q^n$. This violates Schwartz-Zippel, and Thm 3 follows.

Hence, all that remains is to prove Claim C.

Proof of Claim C ($\forall z \in \mathcal{L}\ g(z)=0$)

Consider set $K' = \{ c \cdot x \mid x \in K, c \in \mathbb{F}_q \}$

containing all lines that
- pass through zero
- intersect K at some point

Since g is homogeneous we have

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

Hence, for all $x \in K'$ we have

$$\forall x \in K'\ g(x) = 0 \qquad (\dagger)$$

As above, let $\mathcal{L} \subseteq \mathbb{F}_q^n$ be such that

$$|\mathcal{L}| \geq \delta \cdot q^n$$

Can assume $z \neq 0$ otherwise $g(z) = g(0) = 0$ by homogeneity

and

$$\forall z \in \mathcal{L}\ \exists y \in \mathbb{F}_q^n\ |\mathcal{L}_{y,z} \cap K| \geq \gamma \cdot q \qquad (*)$$

~~Since~~ Fix z and y s.t. $(*)$ holds. Since

$$d \leq \gamma q - 2$$

there exist $d+2$ distinct $a_1, a_2, \dots, a_{d+2} \in \mathbb{F}_q$ such that

$$\forall i \in [d+2] \quad y + a_i \cdot z \in K$$

Furthermore, for at least $d+1$ a_i's (say $a_1, \dots, a_{d+1}$) it holds that

$$a_i \neq 0$$

and so we have

$\forall i \in [d+1] \quad y + a_i \cdot z \in K \quad a_i \neq 0$

For $i \in [d+1]$ let

$$b_i = a_i^{-1}$$

Then the $d+1$ points

$$w_i = b_i(y + a_i z) = b_i y + z \qquad (**)$$

are all in the set K' and so by $(\dagger)$ it holds that

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

Look at y in $(**)$. If $y = 0$, then $w_i = z$ $\forall i$ and hence $g(z) = 0$ by $(\ddagger)$ and we are done.

If $y \neq 0$, then all w_i in $(**)$ are <u>distinct</u>

Furthermore, they lie on the same line, namely the line through z with direction y, a.k.a. $L_{z,y} = \{ z + ty \mid t \in \mathbb{F}_q \}$

Now consider the <u>univariate</u> polynomial

$$h(t) = g(z + ty)$$

We have $\deg(h) \leq \deg(g) \leq d$

By $(**)$ and $(\ddagger)$ h has $d+1$ (distinct) zeros

$b_1, b_2, \ldots, b_{d+1}$

By the Degree Mantra, $h(t)$ must be identically zero. Thus, in particular, it holds that

$$h(0) = g(z) = 0$$

and Claim C follows $\boxed{4}$

This completes the proof of Thm 3

So now we know that we can get

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

by Thm 4 and even

$$|K| \geq c_{n,\varepsilon} \, q^{n-\varepsilon}$$

by Cor 5. But how to get all the way to

$$|K| \geq c_n \, q^n$$

in Thm 6?

We will just (essentially) rerun the proof of Thm 3 again, but with a neat twist that actually shortens the proof and makes in simpler

Let us start by the follow variants of Claims A and B.

Claim A' The number of monomials of degree at most d over n variables is $\binom{d+n}{n}$

Proof Consider $d+n$ balls

○ ○ ○ ○ ○ ○ ○ ○ ○ $\quad n=5$, $d=4$

and colour n black

$$○ ● ● ○ ○ ● ● ● \underbrace{○}_{\text{left-over}}$$

Let # white balls between $(i-1)$st and ith black ball = degree of x_i

$x_1 \, x_3^2$

White balls to the right of rightmost black ball = "left-over degree"

Can verify this yields bijection $\boxed{///}$

We now claim that a Kakeya set $K \subseteq \mathbb{F}_q^n$ satisfies

$$|K| \geq \binom{q+n-1}{n} \geq C_n \cdot q^n.$$

Suppose towards contradiction that on the contrary

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

Claim B' There is a non-zero polynomial $g \in \mathbb{F}_q[x_1, \ldots, x_n]$ of degree $d \le q-1$ (not necessarily homogeneous) such that

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

Proof Pattern matching on the proof of Claim B.

We have $\binom{(q-1)+n}{n}$ monomials and less than $\binom{q-1+n}{n}$ equations restricting the coefficients of these monomials, and the equations are homogeneous. Thus there exists a non-zero polynomial g of degree $\le q-1$ such that $g(x) = 0 \ \forall x \in K$ $\boxed{\text{Th}}$

Let $d = \deg(g) \le q-1$.

Then we can write

$$g(x) = \sum_{\substack{i_1 + \ldots + i_n = d \\ i_j \ge 0 \, \forall j}} c_{i_1, \ldots, i_n} x_1^{i_1} \ldots x_n^{i_n} + \sum_{\substack{i_1 + \ldots + i_n < d \\ i_j \ge 0 \, \forall j}} c_{i_1, \ldots, i_n} x_1^{i_1} \ldots x_n^{i_n} \qquad (***)$$

$$= \bar{g}(x) + g'(x)$$

where $\bar{g} \neq 0$ is the HOMOGENEOUS PART OF DEGREE d of g. (Note that we might have $g'(x) = 0$ but we know by construction that $\bar{g}(x) \neq 0$.)

Now we have an even more dramatic final claim:

<u>Claim C'</u> For every $z \in \mathbb{F}_q^n$ it holds that $\bar{g}(z) = 0$.

Accepting this claim on faith for now, we turn to Schwartz–Zippel. S-Z says that $\bar{g}$ is zero on at most $(q-1)q^{n-1} < q^n$ points in $\mathbb{F}_q^n$ or else $\bar{g}$ is the identically zero polynomial.

But $\bar{g}$ was constructed specifically to be non-zero. Contradiction. Thm 6 follows.

<u>Proof of Claim C'</u>: Let $z \in \mathbb{F}_q^n$ be arbitrary. Since K Kakeya set, $\exists y \in \mathbb{F}_q^n$ s.t.

$$L_{y,z} = \{ y + tz \mid t \in \mathbb{F}_q \} \subseteq K$$

Let $h(t) = g(y + tz)$. Then h is a polynomial of degree $\leq d \leq q-1$ and

$$h(t) = 0 \quad \forall t \in \mathbb{F}_q ,$$

i.e., h must be identically zero by the Degree Mantra, meaning that if we write

$$h(t) = \sum_{i=0}^{d} a_i t^i \qquad a_i \in \mathbb{F}_q$$

it holds that $a_i = 0 \quad \forall i$. In particular $a_d = 0$.

Now look at the coefficient a_d (from another angle). Using (***) and writing

$$h(t) = \bar{g}(y+tz) + g'(y+tz)$$
$$= \bar{h}(t) + h'(t)$$

we see that $\deg(h'(t)) < d$.
For $\bar{h}(t)$ we get

$$\bar{h}(t) = \sum_{i_1+\dots+i_n=d} c_{i_1,\dots,i_n} (y_1+tz_1)^{i_1}(y_2+tz_2)^{i_2}\cdots(y_n+tz_n)^{i_n}$$

Eyeballing this, we observe that the only way to get a degree-d term is to pick up tz_j in every factor for which $i_j \neq 0$ and avoid y_j's.

What this means is that

$$a_d = \sum_{i_1+\dots+i_n=d} c_{i_1,\dots,i_d}\, z_1^{i_1} \cdot z_2^{i_2} \cdots z_n^{i_n}$$
$$= \bar{g}(z)$$

But we just concluded that $a_d = 0$, and hence for an arbitrary $z \in \mathbb{F}_q^n$ it holds that $\bar{g}(z) = 0$, proving Claim C. $\boxed{}$

This establishes the bound in Thm 6.