Three announcements

1. 3 problem sets (not 4)
   - Set 2 was posted on Friday

2. You have to provide at least one solution on Piazza to pass a set. This holds also for set 1
   - (Plus you need at least a C on the set itself to pass, but local TCS students usually get an A)

3. Time to start thinking about a research paper to present

---

LAST TIME

Saw an application of the POLYNOMIAL METHOD

1. Given subset $K$ satisfying some algebraic conditions, construct nonzero low-degree polynomial $g$ that vanishes on all $x \in K$

2. Use algebraic conditions on $K$ to show that in fact $g$ (or some relative $g^*$) vanishes on a set $K'$ such that $|K'| \gg |K|

3. Conclude that $g$ (or $g^*$) has to be identically zero. Contradiction. Hence properties of $K$ "impossibly good"
Recall Kakeya set $K \subseteq \mathbb{F}_q^n$ 'contains a line in every direction'

$$\forall x \in \mathbb{F}_q^n \exists y \in \mathbb{F}_q^n \quad \text{such that} \quad L_{y,x} = \{ y + ax \mid a \in \mathbb{F}_q \} \subseteq K$$

Also considered "somewhat" Kakeya sets or "statistical" Kakeya sets

$K \subseteq \mathbb{F}_q^n$ $(\delta, \chi)$-Kakeya set if

$\exists L \subseteq \mathbb{F}_q^n$ (good directions), $|L| \geq \delta \cdot q^n$

$\forall x \in L \exists y \in \mathbb{F}_q^n$

$|L_{y,x} \cap K| \geq \delta \cdot q$

First proved lower bound

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

using statistical Kakeya set

1. Construct low-degree homogeneous $g \in \mathbb{F}_q[C]$
   s.t. $g(x) = 0 \quad \forall x \in K$

2. By homogeneity $g(y) = 0 \quad \forall y$ on lines through 0 and $y \in K$

3. For direction $x \in L$, can find line $L = \{ a + t \overline{a} \mid t \in \mathbb{F}_q \}$ such that
   (a) $x \in L$
   (b) $g(t)$ has many zeros (and is univariate polynomial in $t$).
Hence \( g(x) = 0 \ \forall x \in L \)
but this set is too large!

contradicts Schwartz-Zippel lemma \( \Rightarrow \)
g zero after all \( \Rightarrow \)
contradicts that \( K \) can be small.

Then proved lower bound

\[ |K| \geq C_n \cdot q^n \]

let \( K \) real Kakeya set (not statistical).

construct low-degree poly that vanishes on \( K \).

write \( g = g_h + g * \)

\( g_h \) homogeneous part of max degree \( (d=0) \)

For any \( z \in \mathbb{F}_q^n \), \( \exists \ y \in \mathbb{F}_q^n \) s.t.

\[ \{ y + a z \mid a \in \mathbb{F}_q \} \subseteq K \]

then \( h(t) = g_h(y + t z) \) has

many zeros and must be identically

zero \( \Rightarrow h = \sum_{i=0}^{d} a_i t^i \) with all \( a_i = 0 \)

But coefficient \( a_0 \) in \( h(t) \) is

\( a_0 = \overline{g}(x) \), so \( \overline{g}(x) = 0 \ \forall x \in \mathbb{F}_q^n \)

contradicts Schwartz-Zippel, so

\( K \) cannot be this small.

OK, so now we know

\[ |K| \geq C_n \cdot q^n \]

and question is settled, right?
Not quite... What is $C_n$?

$C_n \leq 1$ since $F^n_n$ is a Kakeya set.

In fact, Din showed there are Kakeya sets for

$$C_n \times \frac{1}{2^{n-1}}$$

Lower bound form last lecture yields

$$C_n \geq 1/n!$$

[Saraf & Sudan]:

$$C_n \geq 1/(2.6)^n$$

**TODAY AND NEXT LECTURE**

Show result from [Din, Kopparty, Saraf & Sudan '13] that

$$C_n \geq 1/2^n$$

Effectively closes the gap (except for factor $2$).

Improves state-of-the-art randomness extractors (we won't have time to talk about this).
(1) Given set $K$ satisfying algebraic properties, construct nonzero low-degree $g \in \mathbb{F}[x]$ that vanishes on $K$ with high multiplicity.

(2) Use algebraic conditions on $K$ to show that $g$ (or relative $g^*$) vanishes on proper sets with high multiplicity.

(3) Conclude that $g$ (or $g^*$) has to be identically zero, using beefed-up version of Schwartz-Zippel. Contradicts "too good" properties of $K$.

To warm up, look at univariate polynomials over, say, $\mathbb{R}$ or $\mathbb{Q}$.

$p(a) = 0$ if $(x-a) \mid p(x)$

A multiple zero if $(x-a)^2 \mid p(x)$

$p(x) = (x-a)^2 \cdot q(x)$

$p'(x) = 2(x-a)q(x) + (x-a)^2q'(x)$

So $p$ and derivative $p'$ share zero $a$. 
\[ p(x) = x^3 - 3x^2 + 3x - 1 \]

\[ p(x) = (x-1)^3 \]

1 zero of multiplicity 3

\[ p'(x) = 3(x-1)^2 \]  also have 1 as zero

\[ p''(x) = 6(x-1) \]

Now look at \((x-1)^3\) over, say, \(\mathbb{F}_3\)

\[ p(x) = (x-1)^3 = x^3 - 1 \]

\[ p'(x) = 0 \]

Lost all info. Can't distinguish between, e.g.,

\((x-1)^3\) and identically zero polynomial.

Not nice.

Standard derivative not so helpful in positive characteristic.

Define instead \textit{Hasse derivative}

For \( p(x) = \sum a_k x^k \), the \(i\)th Hasse derivative is

\[ p^{(i)}(x) = \sum \binom{k}{i} a_k x^{k-i} \]

Need to generalize this to multivariate case.
From now on, write
\[ \vec{z} = (i_1, \ldots, i_n) \]
\[ \vec{x} = (x_1, \ldots, x_n) \]

etc (perhaps sometimes even without vector arrow)

\[ \vec{z} \] will denote vector of (non-negative) exponents

**Weight** \( \text{wt}(\vec{z}) = \sum_{j=1}^{n} z_j \)

For compactness of notation

\[ \vec{x}^\vec{z} = \prod_{j=1}^{n} x_j^{z_j} \]

Note that \( \text{deg}(\vec{x}^\vec{z}) = \text{wt}(\vec{z}) \)

Also use notation

\[ \left( \begin{array}{c} \vec{z} \\ \vec{f} \end{array} \right) = \prod_{h=1}^{n} \left( \begin{array}{c} i_k \\ f_h \end{array} \right) \]

**Lemma 1** Let \( \vec{z} = (z_1, \ldots, z_n) \) and \( \vec{w} = (w_1, \ldots, w_n) \) and \( \vec{r} = (r_1, \ldots, r_n) \) and consider

\[ (\vec{z} + \vec{w})^\vec{r} = (z_1 + w_1)^{r_1} \cdot \cdots \cdot (z_n + w_n)^{r_n} \quad (*) \]

Then the coefficient of \( z^{\vec{z}} w^{\vec{w}} \)
m in \( (*) \) equals \( \left( \begin{array}{c} \vec{r} \\ \vec{s} \end{array} \right) \)

**Proof** Exercise.
DEF 2 For $p(x) \in \mathbb{F}[x]$ and $\mathbf{\alpha} = (\alpha_1, \ldots, \alpha_n)$ non-negative, the $\mathbf{\alpha}$th HAIRE DERIVATIVE of $p$, denoted $p^{(\mathbf{\alpha})}(x)$, is the coefficient of $x^{\mathbf{\alpha}}$ in the polynomial

$$p^{(\mathbf{\alpha})}(x, z) := p(x + z) \in \mathbb{F}[x, z].$$

In particular

$$p^{(\mathbf{\alpha})}(x + z) = \sum_{\mathbf{\alpha}} p^{(\mathbf{\alpha})}(x) z^{\mathbf{\alpha}} \quad (\dagger)$$

EXAMPLE A Let $p(x^2) = 2x_1^3 + x_2 x_3^2$

$$p(x^2 + z) = 2(x_1 + x_1)^3 + (x_2 + x_2)(x_3 + x_3)^2$$

$$= 2x_1^3 + 6x_1 x_1 + 6x_1 x_2 + 2x_2^3$$

$$+ x_2 x_3^2 + 2x_2 x_3 x_3 + x_2 x_3^2$$

$$+ x_3 x_3 + 2x_3 x_2 x_3 + x_2 x_3^2$$

$$= (2x_1^3 + x_2 x_3^2) + (6x_1) \cdot z^{(0,0,1)} + (6x_1) \cdot z^{(2,0,0)} + (2x_2^3) \cdot z^{(3,0,0)} + (x_2 x_3^2) \cdot z^{(1,0,0)} + (2x_2 x_3) \cdot z^{(2,0,0)} + (x_2 x_3) \cdot z^{(0,0,2)}$$

$$\begin{align*}
\rho^{(0,0)} &= 2x_1^3 + x_2 x_3^2 \\
\rho^{(1,0,0)} &= 6x_1 \\
\rho^{(2,0,0)} &= 6x_1 \\
\rho^{(2,0,1)} &= 6x_1 \\
\rho^{(3,0,0)} &= 2 \\
\rho^{(0,0,1)} &= x_2 \\
\rho^{(0,0,2)} &= x_2 \\
\rho^{(0,1,1)} &= 2x_2 \\
\rho^{(0,1,2)} &= 1
\end{align*}$$
PROPOSITION 3 (Basic properties of Hafez derivatives)

1. \( p^{(2)}(x) + q^{(2)}(x) = (p+q)^{(2)}(x) \)

2. If \( p(x) \) is homogeneous of degree \( d \), then either \( p^{(2)}(x) \) is homogeneous of degree \( d - \text{wt}(p) \) or \( p^{(2)}(x) = 0 \)

3. Let \( H_q(x) \) denote the homogeneous part of \( q \) of highest total degree. Then either \( (H_p)^{(2)}(x) = H_{p^{(2)}}(x) \) or \( (H_p)^{(2)}(x) = 0 \)

4. \( (p^{(2)})(x) = \left( \frac{x^2}{2} \right) p^{(2)}(x) \)

Note that (1) - (3) are the same as for standard derivatives but (4) is not!

However (4) says that if \( (2+\frac{3}{2}) \) th derivative is zero, then \( p \) th derivative of \( 2 \) th derivative is also zero, which will be critical

Proof of Prop 3

Exercise.
DEF 4 (MULTIPLICITY)

For \( p(\bar{x}) \in \mathbb{F}[\bar{x}] \) and \( \bar{a} \in \mathbb{F}^n \), the 
MULTIPLICITY of \( p \) AT \( \bar{a} \), denoted \( \text{mult}(p, \bar{a}) \),
is the largest integer \( M \) such that for every nonnegative \( \bar{i} \) with \( \text{wt}(\bar{i}) < M \)
it holds that \( p^{(\bar{i})}(\bar{a}) = 0 \)

If \( M \) can be taken arbitrarily large, we set \( \text{mult}(p, \bar{a}) = \infty \).

Note that:

1. \( \text{mult}(p, \bar{a}) \geq 0 \) for every \( \bar{a} \) (vacuously)

2. \( p(\bar{a}) = 0 \) if and only if \( \text{mult}(p, \bar{a}) \geq 1 \) (since \( p^{(\bar{i})}(\bar{a}) = p(\bar{a}) \))

3. Condition \( p^{(\bar{i})}(\bar{a}) = 0 \) imposes homogeneous linear constraint on coefficients of \( p \)

Suppose weight limit \( M \)

\[
\# \; \bar{i} \; \text{s.t.} \; \text{wt}(\bar{i}) < M \\
= \# \; \bar{i} \; \text{s.t.} \; i_j > 0, \; \sum_{j = 1}^n i_j \leq M - 1 \\
= \binom{M + n - 1}{n}
\]

So \( \text{mult}(p, \bar{a}) \geq M \) imposes at most \( \binom{M + n - 1}{n} \) homogeneous linear constraints of coefficients of \( p(\bar{x}) \).
In a further twist, we will need to extend this definition to tuples of polynomials

\[ \vec{p} = (p_1, \ldots, p_m) \in \mathbb{F}[\mathbb{R}]^m \]

Define

\[ \vec{p}^{(\vec{a})} = (p_1^{(a)}, \ldots, p_m^{(a)}) \]

then

\[ \text{mult} (\vec{p}, \vec{a}) = \min \left( \text{mult}(p_i, a_i) \right) \]

We need to translate some of the properties of lattice derivations into properties of multiplicities.

**Lemma 5** If \( \text{mult} (p, \vec{a}) = s \), then \( \text{mult} (p^{(\vec{a})}, \vec{a}) \geq s - \text{wt} (\vec{a}) \)

**Proof** By assumption, for any \( \vec{p} \) with \( \text{wt} (\vec{p}) < s \) we have \( p^{(\vec{a})} (\vec{a}) = 0 \). Take any \( \vec{f} \) s.t. \( \text{wt} (\vec{f}) = s - \text{wt} (\vec{a}) \).

We have

\[ (p^{(\vec{a})})(\vec{f})(\vec{a}) = \left( \vec{f}^{(\vec{a})} \right) p^{(\vec{a})} (\vec{a}) \]

Since \( \text{wt} (\vec{p} + \vec{f}) = \text{wt} (\vec{p}) + \text{wt} (\vec{f}) < s \), \( p^{(\vec{a})} (\vec{a}) = 0 \) and hence \( \left( p^{(\vec{a})} \right) (\vec{a}) = 0 \).

Thus \( \text{mult} (p^{(\vec{a})}, \vec{a}) \geq s - \text{wt} (\vec{a}) \).
Now let us see what happens when we compose polynomial tuples.

Let 
\[ \mathbf{x} = (x_1, \ldots, x_n) \]
\[ \mathbf{y} = (y_1, \ldots, y_n) \]

\[ \mathbf{p} = (p_1(x), \ldots, p_m(x)) \in \mathbb{F}^{m}[x] \]
\[ \mathbf{q} = (q_1(y), \ldots, q_n(y)) \in \mathbb{F}^{n}[y] \]

Define the polynomial tuple 
\[ \mathbf{p} \circ \mathbf{q} (\mathbf{y}) \in \mathbb{F}^{m}[y] \]

by
\[ \mathbf{p} \circ \mathbf{q} (\mathbf{y}) = 
\left( p_1 (q_1 (y), \ldots, q_n (y)), \ldots, p_m (q_1 (y), \ldots, q_n (y)) \right) \]

**Proposition 11.4.6**

For any \( \mathbf{a} \in \mathbb{F}^n \), it holds that

\[ \text{mult} (\mathbf{p} \circ \mathbf{q}, \mathbf{a}) \geq \text{mult} (\mathbf{p}, \mathbf{q}(\mathbf{a})) \cdot \text{mult} (\mathbf{q} - \mathbf{q}(\mathbf{a}), \mathbf{a}) \]

In particular, since \( \text{mult} (\mathbf{q} - \mathbf{q}(\mathbf{a})) = 1 \), we have

\[ \text{mult} (\mathbf{p} \circ \mathbf{q}, \mathbf{a}) \geq \text{mult} (\mathbf{p}, \mathbf{q}(\mathbf{a})) \]
Proof

Let \( m_1 = \text{mult} ( \overrightarrow{p}, \overrightarrow{\varphi (\overrightarrow{a})} ) \)
\( m_2 = \text{mult} ( \overrightarrow{\varphi (\overrightarrow{a})}, \overrightarrow{a} ) > 0 \)

If \( m_2 = 0 \) there is nothing to prove

Assume \( m_1 > 0 \) (so that \( \overrightarrow{p} (\overrightarrow{\varphi (\overrightarrow{a})}) = 0 \))

By literal use of (T) we get
\[
\overrightarrow{p} (\overrightarrow{\varphi (\overrightarrow{a}) + \overrightarrow{z}}) =
\]
\[
= \overrightarrow{p} \left( \overrightarrow{\varphi (\overrightarrow{a})} + \sum_{\overrightarrow{z} \neq 0} \overrightarrow{\varphi (\overrightarrow{a})} (a) \overrightarrow{z} \right) \quad \text{[by (T)]}
\]
\[
= \overrightarrow{p} \left( \overrightarrow{\varphi (\overrightarrow{a})} + \sum_{\overrightarrow{z} \neq 0} \overrightarrow{\varphi (\overrightarrow{a})} (a) \overrightarrow{z} \right) \quad \text{[smaller derivatives of \( \overrightarrow{\varphi} = 0 \) since \( \text{mult} (\overrightarrow{\varphi (\overrightarrow{a})}, \overrightarrow{a}) = m_2 \)]}
\]
\[
= \overrightarrow{p} \left( \overrightarrow{\varphi (\overrightarrow{a})} + \overrightarrow{h (\overrightarrow{z})} \right)
\]
\[
\overrightarrow{h (\overrightarrow{z})} = \sum_{\text{wt}(\overrightarrow{z}) = m_2} \overrightarrow{\varphi (\overrightarrow{a})} (a) \overrightarrow{z} \quad \text{[by (T)]}
\]
\[
= \overrightarrow{p} \left( \overrightarrow{\varphi (\overrightarrow{a})} \right) + \sum_{\overrightarrow{z} \neq 0} \overrightarrow{p} (\overrightarrow{z}) (\overrightarrow{\varphi (\overrightarrow{a})}) \overrightarrow{h (\overrightarrow{z})}^2
\]
\[
\text{[by (T)]}
\]
= \sum_{\text{wt}(\overline{p}) \geq m_1} p(\overline{q}) \cdot q(\overline{\alpha}) \cdot \overline{\eta}(\overline{z})

[\text{since } \text{mult}(\overline{p}, q(\overline{\alpha})) = m_1, > 0]$

Now:
- Each monomial $\overline{z}^\alpha$ in $\overline{h}$ has $\text{wt}(\overline{z}) \geq m_2$.
- Also, each occurrence of $\overline{h}(\overline{z})$ is raised to $\overline{f}$ with $\text{wt}(\overline{f}) = m_1$.

It follows that (†) can be written in the form

\[ \sum_{\text{wt}(\overline{k}) \geq m_1 \cdot m_2} c_k \overline{z}^k \]

By using (†) again, it follows that $(\overline{p} \cdot \overline{q}) \cdot \overline{f}(\overline{\alpha}) = 0$ for $\text{wt}(\overline{k}) < m_1 \cdot m_2$ and the proposition follows. \(\Box\)
**Corollary**

Let \( p(x) \in \mathbb{F}[x] = \mathbb{F}[x_1, \ldots, x_n] \)

Let \( a, b \in \mathbb{F}^n \)

Let \( p \circ b(t) := p(a + t \cdot b) \in \mathbb{F}[t]^* \)

Then for any \( t \in \mathbb{F} \) it holds that

\[
\text{mult}(p \circ b(t), t) \geq \text{mult}(p, a + t \cdot b).
\]

**Proof**

Let \( q(t) = a + t \cdot b \in \mathbb{F}[t]^n \)

Apply previous proposition to \( p, q \)

and \( p \circ q = p \circ b \).

Our previous lower bounds on Kakeya sets boiled down to applications of Schwartz-Zippel. We want to extend Schwartz-Zippel to take multiplicities of zeros into account.

**Standard form of Schwartz-Zippel** says that for \( p(x) \in \mathbb{F}[x] \) of degree \( d \)

and \( S \subseteq \mathbb{F} \), probability that \( p(x) = 0 \) when a drawn uniformly at random from \( S^n \) is \( \leq d / |S| \)

*) Apologies for not warning sensitive viewers, but yes, this is suddenly a univariate, standard, non-vector polynomial
Note that
\[ \min (1, \text{mult}(p, \overline{a})) = \begin{cases} 1 & \text{if } p(\overline{a}) = 0 \\ 0 & \text{otherwise} \end{cases} \]
then we can write Schwartz–Zippel as follows.

**Schwartz–Zippel Lemma, Standard Version**

Let \( p(z) \in \mathbb{F}[z] \), \( \deg(p) \leq d \), \( p \neq 0 \), and let \( S \subseteq \mathbb{F} \). Then
\[
\sum_{a \in S} \min (1, \text{mult}(p, \overline{a})) \leq d \cdot |S|^{n-1}
\]

But what if \( p \) vanishes with higher multiplicity on some \( \overline{a} \)?

Look at univariate case.

Suppose \( \text{mult}(p, a) \geq 2 \)

Then \( p(x) = (x-a)^2 \cdot p^*(x) \)

for \( \deg(p^*) = d - 2 \)

so \( p \) can only have \( d - 2 \) other zeroes, since it is "wasted" multiplicity 2 on \( a \).

The key insight (or one of them) in [Dvir, Kopparty, Saraf, Sudan '13] is that this holds also in general.
Lemma 8 (Schwartz-Zippel with multiplicities)

Let \( p \in \mathbb{F}[x_1, \ldots, x_n] \) non-zero polynomial of total degree \( \deg(p) \leq d \).

Then for any finite \( S \subseteq \mathbb{F} \) it holds that

\[
\sum_{a \in S^n} \text{mult}(p, a) \leq d \cdot |S|^{n-1}
\]