Excluding affine configurations over a finite field

Abel Prize Laureates Lectures

Dion Gijswijt
Delft University of Technology
Consider a homogeneous balanced system of linear equations:

\[
\begin{align*}
  a_{11}x_1 + \cdots + a_{1k}x_k &= 0 \\
  \vdots \notag \\
  a_{m1}x_1 + \cdots + a_{mk}x_k &= 0
\end{align*}
\]  

(\star)

Balanced: \( a_{i1} + \cdots + a_{ik} = 0 \) for all \( i \).
Consider a homogeneous balanced system of linear equations:

\[
\begin{align*}
    a_{11}x_1 + \cdots + a_{1k}x_k &= 0 \\
    & \vdots \\
    a_{m1}x_1 + \cdots + a_{mk}x_k &= 0
\end{align*}
\]  

(\star)

Balanced: \( a_{i1} + \cdots + a_{ik} = 0 \) for all \( i \).

Coefficients \( a_{ij} \in \mathbb{F}_q \). Variables \( x_j \in \mathbb{F}_q^n \) are vectors.

\[ x_j = (x_{j1}, \ldots, x_{jn}) \]
Consider a **homogeneous balanced** system of linear equations:

\[
a_{11}x_1 + \cdots + a_{1k}x_k = 0 \\
\vdots \\
a_{m1}x_1 + \cdots + a_{mk}x_k = 0
\]

\((\star)\)

**Balanced:** \(a_{i1} + \cdots + a_{ik} = 0\) for all \(i\).

Coefficients \(a_{ij} \in \mathbb{F}_q\). Variables \(x_j \in \mathbb{F}_q^n\) are vectors. \(x_j = (x_{j1}, \ldots, x_{jn})\)

**Trivial solutions:** \(x_1 = \cdots = x_k\).
Consider a **homogeneous balanced** system of linear equations:

\[
\begin{align*}
    a_{11}x_1 + \cdots + a_{1k}x_k &= 0 \\
    \vdots \\
    a_{m1}x_1 + \cdots + a_{mk}x_k &= 0
\end{align*}
\]  

\((\star)\)

**Balanced:** \(a_{i1} + \cdots + a_{ik} = 0\) for all \(i\).

Coefficients \(a_{ij} \in \mathbb{F}_q\). Variables \(x_j \in \mathbb{F}_q^n\) are vectors. \(x_j = (x_{j1}, \ldots, x_{jn})\)

**Trivial solutions:** \(x_1 = \cdots = x_k\).

**Problem**

How large must \(S \subseteq \mathbb{F}_q^n\) be to ensure a **non-trivial** solution \(x = (x_1, \ldots, x_k)\) with \(x_1, \ldots, x_k \in S\)?
A **cap set**: subset $S \subseteq \mathbb{F}_3^n$ containing no non-trivial solution to $x_1 - 2x_2 + x_3 = 0$. Equivalently: no (non-trivial) 3-term arithmetic progression (3AP).

**Cap set problem**

What is the **asymptotic growth** of maximum size of a cap set in $\mathbb{F}_3^n$?
## Online Encyclopedia of Integer Sequences: A090245

Pellegrino cap (1971)

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>max cap size</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td><strong>20</strong></td>
<td>45</td>
<td>112</td>
<td><strong>236 – 291</strong></td>
</tr>
<tr>
<td>$3^n$</td>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
<td>243</td>
<td>729</td>
<td>2187</td>
</tr>
</tbody>
</table>
Online Encyclopedia of Integer Sequences: A090245

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>max cap size</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>20</td>
<td>45</td>
<td>112</td>
<td>236 – 291</td>
</tr>
<tr>
<td>( 3^n )</td>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
<td>243</td>
<td>729</td>
<td>2187</td>
</tr>
</tbody>
</table>

- \( f(n) = O\left(\frac{3^n}{n}\right) \) [Meshulam, 1995]
- \( O\left(\frac{3^n}{n^{1+c}}\right) \) [Bateman-Katz, 2012]
Online Encyclopedia of Integer Sequences: A090245

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>max cap size</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td><strong>20</strong></td>
<td>45</td>
<td>112</td>
<td><strong>236 – 291</strong></td>
</tr>
<tr>
<td>$3^n$</td>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
<td>243</td>
<td>729</td>
<td>2187</td>
</tr>
</tbody>
</table>

- $f(n) = O\left(\frac{3^n}{n}\right)$ [Meshulam, 1995]
- $O\left(\frac{3^n}{n^{1+c}}\right)$ [Bateman-Katz, 2012]
- $f(n) = \Omega(2.217^n)$ [Edel, 2004]
Motivation

Arithmetic progressions
The cap set problem is a toy model for understanding arithmetic progressions in the integers

Terence Tao: “Perhaps my favourite open question is the problem on the maximal size of a cap set”

Fast matrix multiplication
Possible schemes for fast matrix multiplication rely on large cap sets (e.g. Coppersmith-Winograd conjecture)

Related to other problems in extremal combinatorics
e.g. Erdős-Szemerédi sunflower conjecture.
Solution of the cap set problem

Theorem (2016) [Ellenberg-G.]
For every dimension $n$ we have $f(n) \leq 2.756^n$.

Consequences
- Erdős Szemerédi sunflower conjecture is true.
- Coppersmith-Winograd conjecture is false (not viable path for fast matrix multiplication)
- Proof builds upon work of Croot-Lev-Pach for 3APs in $(\mathbb{Z}/4\mathbb{Z})^n$.
- CLP lemma.
- Proof reformulated by Tao in terms of slice rank of tensors.
- Slice rank method.
Theorem (2016) [Ellenberg-G.]
For every dimension $n$ we have $f(n) \leq 2.756^n$.

Consequences
- Erdős Szemerédi sunflower conjecture is true.
- Coppersmith-Winograd conjecture is false (not viable path for fast matrix multiplication)
Solution of the cap set problem

**Theorem (2016) [Ellenberg-G.]**
For every dimension $n$ we have $f(n) \leq 2.756^n$.

**Consequences**
- Erdős Szemerédi sunflower conjecture is true.
- Coppersmith-Winograd conjecture is false
  (not viable path for fast matrix multiplication)
- Proof builds upon work of Croot-Lev-Pach for 3APs in $(\mathbb{Z}/4\mathbb{Z})^n$.
  CLP lemma.
- Proof reformulated by Tao in terms of slice rank of tensors.
  Slice rank method.
Slice rank method

\[ a_{11}x_1 + \cdots + a_{1k}x_k = 0 \]
\[ \vdots \]
\[ a_{m1}x_1 + \cdots + a_{mk}x_k = 0 \]

where \( a_{ij} \in \mathbb{F}_q \). Variable vectors \( x_j \in \mathbb{F}_q^n \).
Slice rank method

\[
a_{11}x_1 + \cdots + a_{1k}x_k = 0 \\
\vdots \\
a_{m1}x_1 + \cdots + a_{mk}x_k = 0
\]

(\ast)

where \(a_{ij} \in \mathbb{F}_q\). Variable vectors \(x_j \in \mathbb{F}_q^n\).

Theorem

Suppose that \(S \subseteq \mathbb{F}_q^n\) contains no nontrivial solutions to (\ast).

If \(k \geq 2m + 1\), then \(|S| \leq q^{(1-\delta)n}\) for some \(\delta > 0\).

Note: No (non-trivial) bound for \(k \leq 2m\).
Theorem

Suppose that \( S \subseteq \mathbb{F}_q^n \) contains no nontrivial solutions to (\(*)\). If \( k \geq 2m + 1 \) then there is a \( \delta > 0 \) such that \( |S| \leq q^{(1-\delta)n} \).

Note: No (non-trivial) bound for \( k \leq 2m \).

Open problem 4APs

Let \( p \geq 5 \) prime. Is there a \( \delta > 0 \) such that the following holds. If \( S \subseteq \mathbb{F}_p^n \) has no (non-trivial) solutions to

\[
\begin{align*}
x_1 - 2x_2 + x_3 &= 0 \\
x_2 - 2x_3 + x_4 &= 0
\end{align*}
\]

then \( |S| \leq p^{(1-\delta)n} \)?
Non-degenerate solutions

Sometimes non-trivial is still too degenerate!
Non-degenerate solutions

Sometimes non-trivial is still too degenerate!

A solution \((x_1, \ldots, x_k)\) is all-different if all \(x_j\) are distinct.
Non-degenerate solutions

Sometimes non-trivial is still too degenerate!

A solution \((x_1, \ldots, x_k)\) is all-different if all \(x_j\) are distinct.

**Erdős-Ginzburg-Ziv**

Max size of \(S \subseteq \mathbb{F}_p^n\) without all-different solution to

\[ x_1 + \cdots + x_p = 0. \]
Sometimes non-trivial is still too degenerate!

A solution \((x_1, \ldots, x_k)\) is all-different if all \(x_j\) are distinct.

**Erdős-Ginzburg-Ziv**

Max size of \(S \subseteq \mathbb{F}_p^n\) without all-different solution to

\[
x_1 + \cdots + x_p = 0.
\]

Slice rank method does not work (for \(p > 3\))!

However, bounds \(O(p^{(1-\delta)n})\) obtained by modifying/augmenting the slice rank method

Naslund (2020), Fox-Sauermann (2018), Sauermann (2021)
For which systems is there a $\delta > 0$ such that $|S| = O(q^{(1-\delta)n})$ if $S$ has no all-different solution?

Proved for several systems. Mimura-Tokushige: 3 papers, several explicit systems and some families of systems. van Dobben de Bruyn-G.: coefficient matrix has 'many' linearly dependent columns. Sauermann: all $m \times m$ minors nonzero and $k \geq 3m$. 
For which systems is there a $\delta > 0$ such that $|S| = O(q^{(1-\delta)n})$ if $S$ has no all-different solution?

Proved for several systems.

- **Mimura-Tokushige**: 3 papers, several explicit systems and some families of systems.
For which systems is there a $\delta > 0$ such that $|S| = O(q^{(1-\delta)n})$ if $S$ has no all-different solution?

Proved for several systems.

- **Mimura-Tokushige**: 3 papers, several explicit systems and some families of systems.
- **van Dobben de Bruyn-G.**: coefficient matrix has ‘many’ linearly dependent columns.
For which systems is there a $\delta > 0$ such that

$$|S| = O(q^{(1-\delta)n})$$

if $S$ has no all-different solution?

Proved for several systems.

- **Mimura-Tokushige**: 3 papers, several explicit systems and some families of systems.
- **van Dobben de Bruyn-G.**: coefficient matrix has ‘many’ linearly dependent columns.

- **Sauermann**: all $m \times m$ minors nonzero and $k \geq 3m$. 
generic solutions

A solution to (⋆) is **generic** if it only satisfies affine relations implied by (⋆).
A solution to (⋆) is generic if it only satisfies affine relations implied by (⋆).

The affine rank of \( \{x_1, \ldots, x_k\} \) is max. number of affinely independent \( x_j \).

\[
\text{generic} \iff \text{affine rank } k - m.
\]
A solution to (⋆) is **generic** if it only satisfies affine relations implied by (⋆).

The **affine rank** of \( \{x_1, \ldots, x_k\} \) is max. number of affinely independent \( x_j \).

\[
\text{generic} \iff \text{affine rank } k - m.
\]

**Problem**

How large must \( S \subseteq F^n_q \) be to ensure a **generic** solution \( x = (x_1, \ldots, x_k) \) with \( x_1, \ldots, x_k \in S \)?
A solution to (⋆) is **generic** if it only satisfies affine relations implied by (⋆).

The **affine rank** of \( \{x_1, \ldots, x_k\} \) is max. number of affinely independent \( x_j \).

\[
generic \iff \text{affine rank } k - m.
\]

**Problem**

How large must \( S \subseteq \mathbb{F}_q^n \) be to ensure a **generic** solution \( x = (x_1, \ldots, x_k) \)
with \( x_1, \ldots, x_k \in S \)?

Note: to use the slice-rank method, we certainly need \( k \geq 2m + 1 \) and similarly for every implied system.
We call (⋆) tame if every implied system with $m'$ equalities uses $k' \geq 2m' + 1$ variables.

**Theorem**

Suppose that (⋆) is tame. Consider subsets $S \subseteq \mathbb{F}_q^n$.

There is a $\delta > 0$ such that $|S| = \Omega(q^{(1-\delta)n})$ implies generic solutions to (⋆) in $S$ (for $n$ large enough).
Proof sketch 1/5 (setup)

- Restrict to the ‘worst’ case: \( k = 2m + 1 \).
  - **Goal**: Show: \(|S| = \Omega(q^{(1-\delta)n})\) implies generic solutions in \( S \)
    generic \( \equiv \) affine rank \( m + 1 \)

- Induction on \( r \):
  - **Assume**: we get solutions of affine rank \( r < m + 1 \), but not \( r + 1 \).
  - **Goal**: obtain a contradiction.

Important tool is super saturation.

**Proposition (Super saturation)**

Let \( 0 < \delta' < \delta \). There is a constant \( c > 0 \) such that the following holds.

**Suppose**: \(|S| = \Omega(q^{(1-\delta)n})\) implies solutions of affine rank \( \geq r \) (for \( n \) large)

**Then**: \(|S| = \Omega(q^{(1-\delta')n})\) implies \( \Omega(q^{nr-c\delta'n}) \) solutions of affine rank \( \geq r \)
The solutions to (⋆) can be modeled by a low-degree polynomial. Let \( f : S \times \cdots \times S \rightarrow \{0, 1\} \subseteq \mathbb{F}_q \) be the indicator function of the solution set.

Then
\[
f(x_1, \ldots, x_k) = \prod_{i=1}^{m} \prod_{\ell=1}^{n} \left[ 1 - (a_{i1}x_{1\ell} + \cdots + a_{ik}x_{k\ell})^{q-1} \right],
\]
a polynomial of degree \( mn(q - 1) \).

Note: \( \deg(f) = \frac{m}{2m+1} \cdot \text{maximum possible degree} \) (recall that \( k = 2m + 1 \)).
Proof sketch 3/5 (Using tameness)

Tameness of $(\star)$ implies (by matroid union theorem):

If $(x_1, \ldots, x_{2m+1})$ is a solution of affine rank $r$, there exist disjoint $I, J \subseteq \{1, \ldots, 2m + 1\}$ of size $r$ such that $\{x_i : i \in I\}$ and $\{x_i : i \in J\}$ are affinely independent.

Assume:
- all solutions have affine rank $r$
- can always take $I = \{1, \ldots, r\}$ and $J = \{r + 1, \ldots, 2r\}$.

Rename:
- $x = (x_1, \ldots, x_r)$
- $y = (x_{r+1}, \ldots, x_{2r})$
- $z = (x_{2r+1}, \ldots, x_{2m+1})$
Proof sketch 4/5 (constructing low rank matrix, CLP lemma)

Let \( g : S^{2m+1-2r} \rightarrow \mathbb{F}_q \) be random function such that

\[
\sum_{z \in S} g(z) z^\alpha = 0 \quad \text{for all monomials } z^\alpha \text{ of degree } |\alpha| \leq (q - 1)n \cdot (2m + 1 - 2r) \cdot \frac{m}{2m + 1}
\]

Compress \( f \) to a function \( M : S^{2r} \rightarrow \mathbb{F}_q \):

\[
M(x, y) = \sum_z f(x, y, z) g(z)
\]
Proof sketch 4/5 (constructing low rank matrix, CLP lemma)

Let $g : S^{2m+1-2r} \to \mathbb{F}_q$ be random function such that

$$
\sum_{z \in S} g(z)z^\alpha = 0 \quad \text{for all monomials } z^\alpha \text{ of degree } |\alpha| \leq (q - 1)n \cdot (2m + 1 - 2r) \cdot \frac{m}{2m + 1}
$$

Compress $f$ to a function $M : S^{2r} \to \mathbb{F}_q$:

$$
M(x, y) = \sum_z f(x, y, z)g(z)
$$

Then $M$ has low degree: $\deg(M) \leq (q - 1)n \cdot 2r \cdot \frac{m}{2m + 1}$.

Can view $M$ as a $|S|^r \times |S|^r$-matrix.

Croot-Lev-Pach lemma: $M$ has small rank.
Matrix $M$ satisfies:

- Bounded number of non-zeroes in each row/column.
- Total number of non-zeroes is $\Omega(q^{nr-\epsilon n})$ (by supersaturation).

Conclusion: $M$ has high rank ($\Omega(q^{nr-\epsilon n})$). Contradiction!
Thank you!
CLP lemma

Let $f \in \mathbb{F}_q[x_1, \ldots, x_n, y_1, \ldots, y_n]$ be a polynomial of degree $d$. Then the $q^n \times q^n$-matrix

$$M_{a,b} = f(a_1, \ldots, a_n, b_1, \ldots, b_n)$$

has rank $\leq 2 \times$ the number of monomials $x^\alpha$, where $\alpha \in \{0, \ldots, q-1\}^n$ and $|\alpha| := \alpha_1 + \cdots + \alpha_n \leq d/2$.

Proof.

Write

$$f = \sum_{|\alpha| \leq d/2} x^\alpha f_\alpha(y) + \sum_{|\beta| \leq d/2} y^\beta g_\beta(x)$$

for certain $f_\alpha$ and $g_\beta$. Each term $x^\alpha f_\alpha(y)$ and each term $y^\beta g_\beta(x)$ corresponds to a rank 1 matrix (outer product of two vectors). □