Strong Bounds for 3-Progressions

Zander Kelley (UIUC), Raghu Meka (UCLA)
3-Term Arithmetic Progressions

- Triple \((x, z, y)\) with \(x + y = 2z\)

- "trivial" when \(x = y = z\)
3-Term Arithmetic Progressions

**Theorem (Roth ’53)**

If $A \subseteq \{1, 2, \ldots, N\}$ is dense enough*, where density $\delta := \frac{|A|}{N}$, then $A$ must have a (nontrivial) 3-progression.

* (density threshold $\delta \approx \frac{1}{\log \log N}$)
History \((A \subseteq [N], A \geq \delta N \Rightarrow 3\text{-progression})\)

<table>
<thead>
<tr>
<th>(\delta \approx 1/\log \log N)</th>
<th>(Roth ‘53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\delta \approx 1/\log(N)^c, c &gt; 0)</td>
<td>(Heath-Brown ‘87)</td>
</tr>
<tr>
<td></td>
<td>(Szemerédi ‘90)</td>
</tr>
<tr>
<td>(\delta \approx 1/\log(N)^{2/3})</td>
<td>(Bourgain ‘08)</td>
</tr>
<tr>
<td>(\delta \approx (\log \log N)^{O(1)}/\log(N))</td>
<td>(Sanders ‘12)</td>
</tr>
<tr>
<td>(\delta \approx 1/\log(N)^{1+c}, c &gt; 0)</td>
<td>(Bloom-Sisask ‘20)</td>
</tr>
</tbody>
</table>
Our Result

**Theorem (K-Meka ‘23)**

If \( A \subseteq [N] \) is dense enough*, then \( A \) must have a (nontrivial) 3-progression.

* (density threshold \( \delta \approx 2^{-\log(N)^{1/12}} \))

- Compare to lower bound, \( \delta \approx 2^{-\log(N)^{1/2}} \)
Dense sets have many 3-progressions

**Theorem (K-Meka ’23)**

If \( A \subseteq [N], |A| \geq 2^{-d}N \) then \( A \) has \( \sim 2^{-d^{12}}N^2 \) solutions to \( x + y = 2z \)

\( \Rightarrow \) (At most \( |A| \leq N \) trivial solutions)
3-Progression over finite abelian $G$

- If $A \subseteq G$, we can ask if $A$ must have many solutions to $x + y = 2z$ (in $G$).
  - $(A \subseteq [N], |A| \geq 2^{-d}N \Rightarrow 2^{-d^{12}}N^2$ solutions.)
  - $A \subseteq [N], |A| \geq 2^{-d}||F_q^n|| \Rightarrow 2^{-d^9}||F_q^n||^2$ solutions.
  - $A \subseteq G, |A| \geq 2^{-d}|G| \Rightarrow 2^{-d^{12}}|G|^2$ solutions. [BS ‘23]

$(G = \mathbb{Z}_n$ is roughly equivalent to $[N])$
The “Analytic” Approach ($A \subseteq G$)

- Find $A' \subseteq A$, with $\approx \frac{|A'|^3}{|G|}$ solutions to $x + y = 2z$.
- (Want $A'$ large)
- E.g. try $A' = A \cap V$, $V$ structured:
  - $V =$ translate of some approximate subgroup:
    - Subgroup
    - Bohr set
    - Generalized Arithmetic Progression
The “Analytic” approach

- **$V$** = structured set.
- **$A'$** = $A \cap V$ has the "right" number of solutions to $x + y = 2z$

\[(= (1 \pm \epsilon) \frac{|A'|^3}{|V|}.)\]

($\epsilon$ is some small constant, like 1/10)
Approximate Subgroups

- Example: $I = [-m, m] \subseteq \mathbb{Z}$.
- For generic sets $S \subseteq \mathbb{Z}$, we expect $|S + S| \approx |S|^2$
- In contrast, $|I + I| = 2|I|$: “approximately closed under addition”
<table>
<thead>
<tr>
<th>$A \subseteq \mathbb{F}_q^n$, $V = \text{subgroup}$</th>
<th>$A \subseteq [N]$, $V \approx \text{subgroup}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta = 1/\log(N)$ (Roth)</td>
<td>$\delta = 1/\log \log(N)$</td>
</tr>
<tr>
<td>$\delta = 1/\log(N)^{1+c}$ (BK '12)</td>
<td>$\delta = 1/\log(N)^{1+c}$ (BS '20)</td>
</tr>
<tr>
<td>$\delta = 2^{-\log(N)^{1/9}}$ (KM '23)</td>
<td>$\delta = 2^{-\log(N)^{1/12}}$ (KM '23)</td>
</tr>
<tr>
<td></td>
<td>$\delta = 1/\log(N)^{2/3}$</td>
</tr>
<tr>
<td></td>
<td>$\delta = \log \log(N)^{\Theta(1)}/\log(N)$</td>
</tr>
</tbody>
</table>
The “Analytic” approach

\[ V = \text{structured set.} \]

\[ A' = A \cap V \text{ has the "right" number of solutions to } x + y = 2z \]

\[ A' \text{ is "pseudorandom".} \]
Notion of Pseudorandomness \((A \subseteq G)\)

- Draw \(a, a' \sim A\) (uniformly) at random
- Say that \(A\) is pseudorandom if:

\[ a + a' \text{ is near-uniform over } G. \]
Notion of Pseudorandomness \((a, a' \sim A)\)

- Let \(D(x) = \text{PDF}(a + a')\)

\[\implies \text{for any } C \subseteq G, \quad \#\text{sol}(a + a' = c) = (1 \pm \epsilon) \frac{|A|^2 |C|}{|G|}\]

\[\implies (\text{e.g. } C = \{2z | z \in A\})\]
Definition of “near-uniform”

\[ |A| = 2^{-d} |G|, \]
\[ D(x) = \text{PDF}(a + a') \]
\[ C = \{2z \mid z \in A \} \]

If \( p := d + 1 \), then
\[ \#\text{sol}(a + a' = c) \geq \frac{1}{4} \frac{|A|^3}{|G|} \]

\[ (\|D - 1\|_p \leq \epsilon \equiv) \frac{|S|}{|G|} \leq 2^{-p} \]
Notation For (Min)-Entropy Deficit

Write

\[ \Delta(A) = d \]

iff

\[ |A| = 2^{-d} |G| \]
Main Lemma (for general $G$)

- Let $A \subseteq G$, $\Delta(A) \leq d$.
- Either
  1. $PDF(a + a')$ is near-uniform, or
  2. $\frac{|A \cap V|}{|V|} \geq (1 + \epsilon) \frac{|A|}{|G|}$, for some approximate subgroup $V$, $\Delta(V) \leq poly(d, p)$. 
Plan for (II): Zoom in on $A' = A \cap V$ until it looks like (I)
Main Lemma (for \( G = \mathbb{F}_q^n \))

- Let \( A \subseteq \mathbb{F}_q^n, \Delta(A) \leq d \).
- Either
  1. PDF\((a + a')\) is near-uniform, or
  2. \( \frac{|A \cap V|}{|V|} \geq (1 + \epsilon) \frac{|A|}{|\mathbb{F}_q^n|} \),

  for some affine subspace \( V \),
  \( \text{Codim}(V) \leq O(d^4p^4) \).
Density Increments

- Initialize $A_0 = A$, $V_0 = \mathbb{F}_q^n$.
- If $A_i$ is not pseudorandom, pass to some $A_{i+1} := A_i \cap V_{i+1}$, \[
\frac{|A_{i+1}|}{|V_{i+1}|} \geq (1 + \epsilon) \frac{|A_i|}{|V_i|}.
\]
- If $\frac{|A_t|}{|V_t|} \geq (1 + \epsilon)^t \frac{|A|}{|\mathbb{F}_q^n|} \geq 2^{\epsilon t - d}$, then $t \leq d/\epsilon$, and $\Delta(A_t) \leq O(td^8) = O(d^9)$. 

Proof of Main Lemma: Setup

- Let $D(x) = \text{PDF}(a + a')$.
- Assume $D$ is not near-uniform: $\|D - 1\|_p \geq \epsilon$.
- We want to find a large $V$, $\mathbb{E}_V[\mathbf{1}_A] \geq (1 + \epsilon)\mathbb{E}_{F_q^n}[\mathbf{1}_A]$.
- Actually, we will find a “density increment”

$$\mathbb{E}_V[D] \geq 1 + \epsilon$$
Main Idea #1: Spectral Positivity

Let $D = \text{PDF}(a + a')$, $F = \text{PDF}(a - a')$.

- $\|D - 1\|_p \leq \|F - 1\|_p$.
- $\|(F - 1)_-\|_p \leq \|(F - 1)_+\|_p$.

because $F(x - y) \geq 0$. 

![Graphs showing the comparison between D and F with respect to spectral positivity.](image-url)
Main Idea #2: Sifting

- Hard case: $A$ is mostly pseudorandom, but with a “planted” (strong but rare) structured part.
- Suppose $A = V \cup R$, for some subspace $V$ and a random set $R$. How to find $V$?

$$A \rightarrow A \cap (A+s_1) \cap (A+s_2)$$
Main Idea #2: Sifting

- Let $F(x) = \text{PDF}(a - a')$ and assume $\|F\|_p \geq 1 + \epsilon$.
- We use sifting to find a set $B = \bigcap_{i=1}^{p} (A + s_i)$,
  - of size roughly $|B| \geq 2^{-dp}|A|$,
  - witnessing

$$\mathbb{E}_{b,b' \in B}[F(b - b')] \geq 1 + \epsilon/2$$

$A \quad \rightarrow \quad A \cap (A + s_1) \cap (A + s_2)$
Rough Proof Outline

\[ \mathbb{E}_C[|F - 1|] \geq \epsilon, \]
\[ \Rightarrow \mathbb{E}_S[F] \geq 1 + \epsilon/2, \]
\[ \Rightarrow \mathbb{E}_{b,b' \in B}[F(b - b')] \geq 1 + \epsilon/4, \quad \Delta(B) \leq O(pd) \]
\[ \Rightarrow \mathbb{E}_V[F] \geq 1 + \epsilon/8, \quad \Delta(V) \leq O(p^4 d^4). \]

\[ \Delta(C) \leq p \]
\[ \Delta(S) \leq O(p) \]
Thanks for listening!