Roth's theorem on arithmetic progressions, or there and back again

Jing Guo (dev.guoj@gmail.com)
University of Regensburg, ALGANT

April 3, 2024

Contents

1 History

In 1927, van der Waerden published the celebrated theorem, which states that if the positive integers are partitioned into finitely many classes, then at least one of these classes contains arbitrarily long arithmetic progressions (AP), or formally:

Theorem 1.1 (van der Waerden, 1927). *For any positive integers r and k, an integer N exists such that, if the integers in the set $[N] = \{1, 2, \ldots, N\}$ are each colored with one of r distinct colors, then there will be at least k integers forming an AP where all elements share the same color.*

It is considered as one of the fundamental results of Ramsey theory. A strengthening of this theorem was conjectured by Erdős and Turán. Before stating the conjecture, we need the following notion:

Definition 1.2 (Positive upper density). A subset A of $\mathbb{N}$ is said to have *positive upper density* if

$$\limsup_{n\to\infty} \frac{|A \cap [N]|}{N} > 0.$$

Conjecture 1.3 (Erdős–Turán, 1936)**.** Every set of integers with positive upper density contains a k-term AP, for every positive integer k.

The cases $k = 1$ and $k = 2$ are trivial to prove. The case $k = 3$, known as *Roth's theorem*, was proved by Klaus Roth [Rot53] in 1953 with Fourier analysis, which will be the focus of this talk. The case $k = 4$ was proved by Szemerédi [Sze69] in 1969. Roth [Rot72] gave a second proof in 1972.

The general case was completely settled in 1975 by Szemerédi:

Theorem 1.4 (Szemerédi, 1975 [Sze75])**.** *Any subset of* $\mathbb{N}$ *with positive upper density contains infinitely many* k*-term APs.*

Remark 1.5. The original proof was very combinatorial, called "a masterpiece of combinatorial reasoning" by Erdős.

Many other proofs now exist. The second proof (considered by many to be "the most important") was given by Furstenberg [Fur77; FKO82] in 1977, with ergodic theory. In 2001, William Timothy Gowers "invented" higher-order Fourier analysis [Gow01] to prove the theorem, for which (along with other work on functional analysis) he was awarded the Fields Medal in 2002.

Terence Tao called the various proofs of Szemerédi's theorem a "Rosetta stone" for connecting many fields of mathematics.

One of the most exciting developments in additive combinatorics at the beginning of this century is the following celebrated theorem proved by Ben Green and Terence Tao:

Theorem 1.6 (Green–Tao, 2004 [GT08])**.** *The prime numbers contain arbitrarily long arithmetic progressions.*

Remark 1.7. This result is not implied by Szemerédi's theorem since the primes are of density 0 in the natural numbers. Green and Tao introduced a "relative" version of Szemerédi's theorem which applies to some subsets of the integers that satisfy certain pseudo-randomness conditions.

Interested readers are referred to the excellent exposition [CFZ14] for a proof of the Green–Tao theorem, written by David Conlon, Jacob Fox, and Yufei Zhao.

Before closing this section, we mention the following well-known Erdős–Turán conjecture. If true, it would imply both Szemerédi's and Green–Tao theorems:

Conjecture 1.8 (Erdős–Turán, \$5000)**.** If the sum of reciprocals of a set of integers diverges, then that set contains arbitrarily long arithmetic progressions.

2 Bounds of Roth's theorem

The infinite version of Roth's theorem is often stated as the following:

Theorem 2.1 (Roth, 1953)**.** *A subset of* $\mathbb{N}$ *with positive upper density contains a* 3*-term AP.*

However, often people are more interested in the quantitative version: We say that A is 3*-AP-free* if there are no $x, x + y, x + 2y \in A$, with $y \neq 0$. A 3-AP is *trivial* if $y = 0$.

We can retrieve the infinite version from the finite version if we let N tend to infinity, A becomes negligible compared to N, implying that the density of A approaches zero.

Theorem 2.2 (Roth)**.** *Let* $A \subseteq [N]$ *be* 3*-AP-free. Then* $|A| = o(N)$.

Improving upper and lower bounds on $|A|$ is a very much active research problem. We state the important progress below:

The original bound given in Roth's proof:

$$|A| = O\left(\frac{N}{\log\log N}\right)$$

Over the years, this bound has been continually lowered by Szemerédi, Heath-Brown, Bourgain, and Sanders. The previous best bound "breaking the logarithmic barrier" was due to Bloom and Sisask [BS20], who showed the existence of a constant $c > 0$, such that

$$|A| \leq \frac{N}{(\log N)^{1+c}}.$$

In February 2023, Kelley and Meka [KM23] gave the following remarkable result:

$$|A| \leq \frac{N}{2^{O((\log N)^c)}}$$

Four days later, Bloom and Sisask [BS23b] simplified the result with a little improvement to $|A| \leq N/\exp\bigl(O((\log N)^{1/11})\bigr)$. Several months later, Bloom and Sisask [BS23a] obtained a further improvement to $|A| \leq N/\exp\bigl(O((\log N)^{1/9})\bigr)$, and stated (without proof) that their techniques can be used to show $|A| \leq N/\exp\bigl(O((\log N)^{5/41})\bigr)$.

There has also been work on the other direction: Constructing the largest set with no 3-AP. However, the best construction has barely seen improvement since 1946, when Behrend [Beh46] improved on the initial construction by Salem and Spencer and showed

$$|A| \geq \frac{N}{\exp\bigl(O(\sqrt{\log N})\bigr)}.$$

Due to lack of improvements in over 70 years, it is conjectured that Behrend's construction is asymptotically close to the best [BS20]. If correct, the Kelley–Meka bound will prove this conjecture.

3 Fourier analysis in the integers

In this section, we will introduce Fourier analysis in the integers, which is a crucial tool in the proof of Roth's theorem (also known as the Hardy–Littlewood circle method in this context). It allows us to detect non-randomness in the distribution of integers by examining the magnitude of Fourier coefficients.

Before going into the proof, we will review some basic notions of Fourier analysis on the integers. We will denote by $\mathbb{R}/\mathbb{Z}$ the set of real numbers mod 1, and assume the Lebesgue measure.

Definition 3.1 (Fourier transform in $\mathbb{Z}$)**.** Given a finitely supported function $f : \mathbb{Z} \to \mathbb{C}$, define $\widehat{f} : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ by setting, for all $\theta \in \mathbb{R}$,

$$\widehat{f}(\theta) = \sum_{x \in \mathbb{Z}} f(x) e(-x\theta),$$

where $e(t) = \exp(2\pi i t)$, for $t \in \mathbb{R}$.

Note: $\widehat{f}(\theta) = \widehat{f}(\theta + n)$, for all $n \in \mathbb{Z}$.

Theorem 3.2 (Fourier inversion formula)**.** *Given finitely supported $f : \mathbb{Z} \to \mathbb{C}$, for all $x \in \mathbb{Z}$,*

$$f(x) = \int_0^1 \widehat{f}(\theta) e(x\theta)\ d\theta$$

Theorem 3.3 (Parseval)**.** *Given finitely supported $f, g : \mathbb{Z} \to \mathbb{C}$,*

$$\sum_{x \in \mathbb{Z}} \overline{f(x)} g(x) = \int_0^1 \overline{\widehat{f}(\theta)} \widehat{g}(\theta)\ d\theta$$

In particular, when $f = g$, we have

$$\sum_{x \in \mathbb{Z}} |f(x)|^2 = \int_0^1 \left|\widehat{f}(\theta)\right|^2\ d\theta$$

Definition 3.4 (Convolution)**.** Given finitely supported $f : \mathbb{Z} \to \mathbb{C}$, define $f * g : \mathbb{Z} \to \mathbb{C}$ by

$$(f * g)(x) = \sum_{y \in \mathbb{Z}} f(y) g(x - y).$$

Theorem 3.5 (Convolution identity)**.** *Let functions f and g defined as before. For all $x \in \mathbb{R}/\mathbb{Z}$,*

$$\widehat{f * g}(x) = \widehat{f}(x)\ \widehat{g}(x).$$

Remark 3.6. Interested readers are encouraged to watch the excellent videos on Fourier transform [1] and convolution [2] produced by 3Blue1Brown.

Given finitely supported $f, g, h : \mathbb{Z} \to \mathbb{C}$, define

$$\Lambda(f, g, h) = \sum_{x, y \in \mathbb{Z}} f(x) g(x + y) h(x + 2y)$$
$$\Lambda_3(f) = \Lambda(f, f, f)$$

Then for any finite set A of integers,

$$\Lambda_3(A) = \Lambda_3(1_A) = |\{(x, y) : x, x + y, x + 2y \in A\}|$$

counts the number of 3-APs in A, where each non-trivial 3-AP is counted twice, and each trivial 3-AP is counted once.

Theorem 3.7 (Fourier and 3-AP)**.** *Given finitely supported $f, g, h : \mathbb{Z} \to \mathbb{C}$,*

$$\Lambda(f, g, h) = \int_0^1 \widehat{f}(\theta) \widehat{g}(-2\theta) \widehat{h}(\theta)\ d\theta$$

[1] `https://www.youtube.com/watch?v=spUNpyF58BY`

[2] `https://www.youtube.com/watch?v=KuXjwB4LzSA`

Proof. We will start by expanding the right-hand side:

$$\begin{aligned}\int_0^1 \widehat{f}(\theta)\widehat{g}(-2\theta)\widehat{h}(\theta)\ d\theta &= \int_0^1 \left(\sum_{x\in\mathbb{Z}} f(x)e(-x\theta)\right)\left(\sum_{y\in\mathbb{Z}} g(y)e(2y\theta)\right)\left(\sum_{z\in\mathbb{Z}} h(z)e(-z\theta)\right)\ d\theta \\ &= \sum_{x,y,z\in\mathbb{Z}} f(x)g(y)h(z)\int_0^1 e((-x+2y-z)\theta)\ d\theta\end{aligned}$$

We notice that the integral only evaluates to 1 when $-x+2y-z=0$, and 0 otherwise. Therefore,

$$\begin{aligned}\int_0^1 \widehat{f}(\theta)\widehat{g}(-2\theta)\widehat{h}(\theta)\ d\theta &= \sum_{\substack{x,y,z\in\mathbb{Z}\\ x+z=2y}} f(x)g(y)h(z) \\ &= \sum_{x,y\in\mathbb{Z}} f(x)g(x+y)h(x+2y) \\ &= \Lambda(f,g,h).\end{aligned}$$

We obtain the second equality by the substitution $y \mapsto x+y$. □

4 Proof of Roth's theorem

There are essentially two ways to prove Roth's theorem: Roth originally used Fourier analysis to prove the theorem, and subsequent improvements on the bounds rely on this approach as well. The second "much simpler" approach needs Szemerédi's regularity lemma (provided that you understand the regularity lemma and its derived triangle counting lemma). However, it only gives a very weak bound $o(N)$ (compared to Roth's $O(N/\log\log N)$). This proof will not be covered here, but interested readers are invited to read Yufei Zhao's excellent "Graph Theory and Additive Combinatorics" [Zha23].

Before going into the proof, we will introduce some notations:

$$\left\|\widehat{f}\right\|_\infty = \sup_\theta \left|\widehat{f}(\theta)\right| \quad \text{and} \quad \|f\|_2 = \left(\sum_{x\in\mathbb{Z}} |f(x)|^2\right)^{1/2}$$

The following proposition says that if f and g are "Fourier-close", then they should have similar 3-AP counts:

Proposition 4.1 (3-AP counting lemma)**.** *Let $f, g : \mathbb{Z} \to \mathbb{C}$ be finitely supported functions. Then*

$$|\Lambda_3(f) - \Lambda_3(g)| \leq 3\left\|\widehat{f-g}\right\|_\infty \max\{\|f\|_2^2, \|g\|_2^2\}.$$

Proof. We notice

$$\Lambda_3(f) - \Lambda_3(g) = \Lambda(f-g,f,f) + \Lambda(g,f-g,f) + \Lambda(g,g,f-g).$$

We then can bound the first term in the following way:

$$\begin{aligned}
|\Lambda(f-g,f,f)| &= \left|\int_0^1 \widehat{(f-g)}(\theta)\widehat{f}(-2\theta)\widehat{f}(\theta)\ d\theta\right| \\
&\leq \left\|\widehat{f-g}\right\|_\infty \left|\widehat{f}(-2\theta)\widehat{f}(\theta)\ d\theta\right| \\
&\leq \left\|\widehat{f-g}\right\|_\infty \left(\int_0^1 \left|\widehat{f}(-2\theta)d\theta\right|^2\right)^{1/2} \left(\int_0^1 \left|\widehat{f}(\theta)\right|^2 d\theta\right)^{1/2} \\
&\leq \left\|\widehat{f-g}\right\|_\infty \|f\|_2^2.
\end{aligned}$$

The first inequality is obtained by the triangle inequality, the second by Cauchy-Schwarz, and the third by Parseval.

We can bound the second and third terms in a similar way. Therefore,

$$\begin{aligned}
|\Lambda_3(g,f-g,f)| &\leq \left\|\widehat{f-g}\right\|_\infty \|f\|_2\|g\|_2 \\
|\Lambda_3(g,g,f-g)| &\leq \left\|\widehat{f-g}\right\|_\infty \|g\|_2^2.
\end{aligned}$$

Combining the three terms, we obtain the desired inequality. □

We recall that we will be proving the following version of Roth's theorem:

Theorem 4.2 (Roth)**.** *Every 3-AP-free subset of $[N]$ has size $O(N/\log\log N)$.*

Let A be a 3-AP-free subset of $[N]$. The proof will proceed in three steps:

1. Show that A admits a large Fourier coefficient.

2. Show that a large Fourier coefficient implies density increment on a sub-AP.

3. Iterate the density increment.

By "density increment" on a sub-AP, we refer to the method that if A does not have a 3-AP, then we can find a subset of A with a higher density.

4.1 Step 1: A 3-AP-free set has a large Fourier coefficient

Lemma 4.3 (3-AP-free implies a large Fourier coefficient)**.** *Let $A \subseteq [N]$ be a 3-AP-free set with density $\alpha = |A|/N$. If $N \geq 5\alpha^{-2}$, then there exists $\theta \in \mathbb{R}/\mathbb{Z}$, satisfying*

$$\left|\sum_{x=1}^{N}(1_A-\alpha)(x)\ e(x\theta)\right| \geq \frac{\alpha^2}{10}N.$$

Proof. Since A is 3-AP-free, we have

$$\Lambda_3(1_A) = |A| = \alpha N$$

due to the fact that $\Lambda_3(1_A)$ only counts the trivial 3-APs.

On the other hand, in $[N]$ we can count 3-APs by only considering pairs of integers with same parity to form first and thid elements of the 3-AP:

$$\Lambda_3(1_N) \geq \frac{N^2}{2}$$

Now we apply the counting lemma (Proposition 4.1) with $f = 1_A$ and $g = \alpha 1_N$: First, we have

$$\|1_A\|_2^2 = |A| = \alpha N \quad \text{and} \quad \left\|\alpha 1_{[N]}\right\|_2^2 = \alpha^2 N$$

So we obtain

$$\frac{\alpha^3 N^2}{2} - \alpha N \leq \alpha^3 \Lambda_3(1_{[N]}) - \Lambda_3(1_A) = \left|\Lambda_3(1_A) - \Lambda_3(\alpha 1_{[N]})\right| \leq 3\left\|\widehat{(1_A - \alpha 1_{[N]})}\right\|_\infty \cdot \alpha N$$

We now consider the first and last terms in the inequality:

$$\left\|\widehat{(1_A - \alpha 1_{[N]})}\right\|_\infty \geq \frac{\frac{\alpha^3 N^2}{2} - \alpha N}{3\alpha N} = \frac{1}{6}\alpha^2 N - \frac{1}{3} \geq \frac{1}{10}\alpha^2 N$$

where the last inequality follows from the assumption $N \geq 5\alpha^{-2}$.

The L^∞ norm being large implies the existence of a large Fourier coefficient, therefore, there exists some $\theta \in \mathbb{R}/\mathbb{Z}$ with

$$\left|\sum_{x=1}^{N}(1_A - \alpha)(x) \cdot e(\theta x)\right| = \widehat{(1_A - \alpha 1_{[N]})}(\theta) \geq \frac{1}{10}\alpha^2 N.$$

□

4.2 Step 2: A large Fourier coefficient implies density increment on a sub-AP

Lemma 4.4 (Dirichlet's lemma)**.** *We let $\|\theta\|_{\mathbb{R}/\mathbb{Z}}$ be the distance from θ to the nearest integer.*

Let $\theta \in \mathbb{R}$ and $0 < \delta < 1$. Then there exists a positive integer $d \leq 1/\delta$, such that $\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leq \delta$.

Proof. Let $N = \lceil 1/\delta \rceil$. Consider the sequence of fractional parts $\{\theta\}, \{2\theta\}, \ldots, \{N\theta\}$, where $\{x\}$ denotes the fractional part of x. The unit interval $[0, 1)$ is partitioned into N sub-intervals of equal length $1/N$.

By the pigeonhole principle, since we have $N + 1$ fractional parts (including 0) and only N sub-intervals, there must be at least one sub-interval that contains at least two of these fractional parts. Let $\{m\theta\}$ and $\{n\theta\}$ be two such fractional parts that lie within the same sub-interval, where m and n are distinct integers, such that $1 \leq m, n \leq N$, and without loss of generality, assume $m > n$.

The distance between $\{m\theta\}$ and $\{n\theta\}$ is at most $1/N$. The fractional part of their difference $\{(m - n)\theta\}$ measures the distance of the multiple $(m - n)\theta$ from the nearest integer, which is $\|(m - n)\theta\|_{\mathbb{Z}}$.

Setting $d = m - n$, we have $1 \leq d \leq N$ because both m and n are within the range from 1 to N, and their difference is non-negative and at most N. Moreover, $\|d\theta\|_{\mathbb{Z}} \leq 1/N \leq \delta$, fulfilling the conditions of the lemma. □

Lemma 4.5 (Partition into progression level sets)**.** *Let* $0 < \varepsilon < 1$ *and* $\delta \in \mathbb{R}$*. Suppose* $N \geq (4\pi/\varepsilon)^6$*, then one can partition* $[N]$ *into sub-AP* P_i*, each with length*

$$N^{1/3} \leq |P_i| \leq 2N^{1/3},$$

such that for each i*,*

$$\sup_{x,y \in P_i} |e(x\theta) - e(y\theta)| < \varepsilon.$$

Proof. By the Dirichlet's lemma, there is a positive integer $d < \sqrt{N}$, such that $\|d\theta\|_{\mathbb{R}/\mathbb{Z}} \leq 1/\sqrt{N}$. We then partition $[N]$ greedily into sub-AP P_i, with common difference d, of length between $N^{1/3}$ and $2N^{1/3}$.

For two any two elements x, y within the same P_i, we have

$$|e(x\theta) - e(y\theta)| \leq |P_i||e(d\theta) - 1| \leq 2N^{1/3} \cdot 2\pi \cdot N^{-1/2} \leq \varepsilon$$

the second inequality follows from the fact that the length of a chord on a circle is at most the length of the corresponding arc. □

Lemma 4.6 (3-AP-free implies density increment)**.** *Let* $A \subset [N]$ *be* 3*-AP-free, with density* $\alpha = |A|/N$ *and* $N \geq (16/\alpha)^{12}$*. Then there exists a sub-AP* $P \subset [N]$ *with* $|P| \geq N^{1/3}$ *and* $|A \cap P|/|P| \geq \alpha + \alpha^2/40$*.*

Proof. By Lemma 4.3, there exists $\theta \in \mathbb{R}/\mathbb{Z}$ satisfying

$$\left|\sum_{x=1}^{N}(1_A - \alpha)(x) \cdot e(x\theta)\right| \geq \frac{\alpha^2}{10}N.$$

Next, apply Lemma 4.5 with $\varepsilon = \alpha^2/20$ to obtain a partition $P_1, \ldots, P_k$ of $[N]$ satisfying $N^{1/3} \leq |P_i| \leq 2N^{1/3}$ and

$$|e(x\theta) - e(y\theta)| \leq \frac{\alpha^2}{20}$$

for all i and $x, y \in P_i$.

So on each P_i, by the triangle inequality and Lemma 4.3, we have

$$\left|\sum_{x\in P_i}(1_A - \alpha)(x) \cdot e(x\theta)\right| \leq \left|\sum_{x\in P_i}(1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}|P_i|$$

Thus,

$$\begin{aligned}
\frac{\alpha^2}{10} &\leq \left|\sum_{x=1}^{N}(1_A - \alpha)(x)e(x\theta)\right| \\
&\leq \sum_{i=1}^{k}\left|\sum_{x\in P_i}(1_A - \alpha)(x)e(x\theta)\right| \\
&\leq \sum_{i=1}^{k}\left(\left|\sum_{x\in P_i}(1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}|P_i|\right) \\
&= \sum_{i=1}^{k}\left|\sum_{x\in P_i}(1_A - \alpha)(x)\right| + \frac{\alpha^2}{20}N
\end{aligned}$$

By combining the first and the last terms in the inequality, we obtain

$$\frac{\alpha^2}{20}N \leq \sum_{i=1}^{k} \left| \sum_{x \in P_i} (1_A - \alpha)(x) \right|$$

and thus

$$\frac{\alpha^2}{20} \sum_{i=1}^{k} |P_i| \leq \sum_{i=1}^{k} ||A \cap P_i| - \alpha |P_i||.$$

Next, we want to show that there exists some P_i, such that A has a density increment when restricted to P_i. We notice the following trick:

$$\begin{aligned} \frac{\alpha^2}{20} \sum_{i=1}^{k} |P_i| &\leq \sum_{i=1}^{k} ||A \cap P_i| - \alpha |P_i|| \\ &= \sum_{i=1}^{k} ||A \cap P_i| - \alpha |P_i|| + (|A \cap P_i| - \alpha |P_i|). \end{aligned}$$

as the newly added terms in the final step sum to zero.

Therefore, there exists some i such that

$$\frac{\alpha^2}{20}|P_i| \leq ||A \cap P_i| - \alpha |P_i|| + (|A \cap P_i| - \alpha |P_i|)$$

Further, we notice that $|t| + t$ is $2t$ for $t > 0$, and 0 otherwise, so we obtain

$$\frac{\alpha^2}{20}|P_i| \leq 2(|A \cap P_i| - \alpha |P_i|),$$

which implies

$$|A \cap P_i| \geq (\alpha + \frac{\alpha^2}{40})|P_i|.$$

□

By translation and re-scaling, we can identify P with $[N']$ with $N' = |P|$. Then $A \cap P$ becomes a subset $A' \subseteq [N']$. Since A' is 3-AP-free, we now may iterate this argument.

4.3 Step 3: Iterate the density increment

We now iterate the second step. Let α_t be the density of A after the t-th iteration, and N_t the size of our current progression after t iterations. We start with $\alpha_0 = \alpha$ and $N_0 = N$. After i iterations, we arrive at a sub-AP of length N_i, where A has density α_i. As long as $N \geq (16/\alpha_i)^{12}$, we can apply Lemma 4.6 to pass down to a sub-AP with

$$N_{i+1} \geq N_i^{1/3} \quad \text{and} \quad \alpha_{i+1} \geq \alpha_i + \frac{\alpha_i^2}{40}.$$

Notice that we double α_i from α_0 after $\leq \lceil 40/\alpha \rceil$ iterations. Once the density reaches $\geq 2\alpha$, the next doubling takes $\leq \lceil 20/\alpha \rceil$ iterations, and so on. In general, the k-th doubling requires $\leq \lceil 40 \cdot 2^{-k}/\alpha \rceil$ iterations. Since density is at most 1, there are at most $\log_2(1/\alpha)$ doublings.

Summing up everything, the total number of iterations is

$$m \leq \sum_{i=1}^{\log_2(1/\alpha)} \lceil 40 \cdot 2^{-k}/\alpha \rceil = O(1/\alpha).$$

When the process terminates, by Lemma 4.6,

$$N^{1/3^m} \leq N_m < (16/\alpha_i)^{12} \leq (16/\alpha)^{12}.$$

Rearranging gives

$$N \leq (16/\alpha)^{12 \cdot 3^m} \leq (16/\alpha)^{\exp(O(1/\alpha))}.$$

Therefore,

$$\frac{|A|}{N} = \alpha = O\left(\frac{1}{\log \log N}\right).$$

This completes the proof of Roth's theorem.

5 Roth's theorem in finite fields: The cap-set problem

An interesting variation is Roth's theorem in finite fields: A 3-AP-free subset A of $(\mathbb{Z}/3\mathbb{Z})^n$ is called a *cap set* (named after the card game "SET"). The *cap-set problem* asks to determine the size of the largest cap set in F_3^n.

In 1982, Brown and Buhler [BB82] were the first to show that $|A| = o(3^n)$. In 1995, Mesuhlam [Mes95] show that $|A| = O(3^n/n)$ with Fourier analysis. The bound was improved to $|A| = O(3^n/n^{1+\varepsilon})$ by Bateman and Katz [BK12] in 2012.

In 2017, Croot, Lev, and Pach [CLP17] achieved a breakthrough result by applying the *polynomial method* to Roth-type problems in the finite field model. Less than two weeks after their paper was made public, Ellenberg and Gijswijt [EG17] adapted their argument to prove the current best bound:

$$|A| = O(2.756^n).$$

It is interesting to note that both papers were published in the *Annals* and were surprisingly short – just 6 and 4 pages long, respectively.

The best lower bound is 2.2202^n [Rom+24], obtained by a group of Google DeepMind researchers with a large language model (LLM) in 2023.

References

[BB82] Tom C Brown and Joe P Buhler. "A density version of a geometric Ramsey theorem". In: *Journal of Combinatorial Theory, Series A* 32.1 (1982), pp. 20–34.

[Beh46] Felix A Behrend. "On sets of integers which contain no three terms in arithmetical progression". In: *Proceedings of the National Academy of Sciences* 32.12 (1946), pp. 331–332.

[BK12] Michael Bateman and Nets Katz. "New bounds on cap sets". In: *Journal of the American Mathematical Society* 25.2 (2012), pp. 585–613.

[BS20] Thomas F Bloom and Olof Sisask. “Breaking the logarithmic barrier in Roth’s theorem on arithmetic progressions”. In: *arXiv preprint arXiv:2007.03528* (2020).

[BS23a] Thomas F Bloom and Olof Sisask. “An improvement to the Kelley-Meka bounds on three-term arithmetic progressions”. In: *arXiv preprint arXiv:2309.02353* (2023).

[BS23b] Thomas F Bloom and Olof Sisask. “The Kelley–Meka bounds for sets free of three-term arithmetic progressions”. In: *arXiv preprint arXiv:2302.07211* (2023).

[CFZ14] David Conlon, Jacob Fox, and Yufei Zhao. “The Green-Tao theorem: an exposition”. In: *EMS Surveys in Mathematical Sciences* 1.2 (2014), pp. 249–282.

[CLP17] Ernie Croot, Vsevolod F Lev, and Péter Pál Pach. “Progression-free sets in are exponentially small”. In: *Annals of Mathematics* (2017), pp. 331–337.

[EG17] Jordan S Ellenberg and Dion Gijswijt. “On large subsets of with no three-term arithmetic progression”. In: *Annals of Mathematics* (2017), pp. 339–343.

[FKO82] Hillel Furstenberg, Yitzhak Katznelson, and Donald Ornstein. “The ergodic theoretical proof of Szemerédi’s theorem”. In: (1982).

[Fur77] Harry Furstenberg. “Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions”. In: *Journal d’Analyse Mathématique* 31.1 (1977), pp. 204–256.

[Gow01] William T Gowers. “A new proof of Szemerédi’s theorem”. In: *Geometric & Functional Analysis GAFA* 11.3 (2001), pp. 465–588.

[GT08] Ben Green and Terence Tao. “The primes contain arbitrarily long arithmetic progressions”. In: *Annals of mathematics* (2008), pp. 481–547.

[KM23] Zander Kelley and Raghu Meka. “Strong bounds for 3-progressions”. In: *arXiv preprint arXiv:2302.05537* (2023).

[Mes95] Roy Meshulam. “On subsets of finite abelian groups with no 3-term arithmetic progressions”. In: *Journal of Combinatorial Theory, Series A* 71.1 (1995), pp. 168–172.

[Rom+24] Bernardino Romera-Paredes et al. “Mathematical discoveries from program search with large language models”. In: *Nature* 625.7995 (2024), pp. 468–475.

[Rot53] Klaus F Roth. “On certain sets of integers”. In: *J. London Math. Soc* 28.104-109 (1953), p. 3.

[Rot72] Klaus F Roth. “Irregularities of sequences relative to arithmetic progressions, IV”. In: *Periodica Mathematica Hungarica* 2.1 (Mar. 1972), pp. 301–326. DOI: 10.1007/BF02018670.

[Sze69] Endre Szemerédi. “On sets of integers containing no four elements in arithmetic progression”. In: *Acta Mathematica Hungarica* 20.1-2 (1969), pp. 89–104.

[Sze75] Endre Szemerédi. “On sets of integers containing no k elements in arithmetic progression”. In: *Acta Arith* 27.199-245 (1975), p. 2.

[Zha23] Yufei Zhao. *Graph Theory and Additive Combinatorics: Exploring Structure and Randomness.* Cambridge University Press, 2023.