Roth's Theorem for 3-term APs in $\mathbb{Z}$

Prahladh Harsha

Ramprasad Saptharishi[*]

TIFR, Mumbai

December 1st, 2017

Contents

In this class we shall look at a different application of Fourier analysis, not in the Boolean setting but over groups such as $\mathbb{Z}/N\mathbb{Z}$. The main result in this class would be Roth's theorem about 3-term APs inside sets of large density. These notes are based on Ryan O'Donnell's lecture notes for the course Analysis of Boolean Functions [O'D07, Lecture 27] and Adam Lott's notes on Roth's Theorem [Lot17].

1 Some history

The story begins with a result of Van der Warden that proved the following:

Theorem (Van der Warden 1927)**.** *Consider any colouring of $\mathbb{Z}_{\geq 0}$ consisting of r colours. Then, for any $k \geq 0$, there would be a monochromatic AP of length k.*

[*]ramprasad@tifr.res.in

Subsequently, Erdős and Turan conjectured that this is really a question about the *density* of AP-free sets.

Conjecture (Erdős -Turan 1936). *Let* $A \subseteq \mathbb{Z}_{\geq 0}$ *with* $\limsup_N \frac{|A \cap [N]|}{N} = \delta > 0$. *Then, there are arbitrarily long APs in* A.

Let $r_k(N)$ refer to the size of the largest k-AP-free subset in $[N]$. Can we get upper and lower bounds for this? Behrend gave a cool lower bound for such sets for the case of $k = 3$.

Theorem (Behrend 1946). $r_3(N) \geq \frac{n}{2^{\sqrt{O(\log n)}}}$.

For $k = 3$, Roth proved an upper bound for $r_3(N)$.

Theorem (Roth 1956). $r_3(N) \leq O\left(\frac{N}{\log \log N}\right)$.

And finally, for all k, Szemerédi proved an upper bound.

Theorem (Szemerédi1975). $r_k(N) = o_k(N)$.

Gowers then presented a more explicit growth of $r_k(N)$ and showed the following.

Theorem (Gowers 2001).

$$r_k(N) = O\left(\frac{N}{(\log \log N)^{c_k}}\right) \quad , \quad \text{where } c_k = \frac{1}{2^{2^{k+9}}}.$$

Erdős and Turan, in their paper, also made the following conjecture.

Conjecture (Erdős-Turan 1936). *Let* A *be any subset of positive integers such that* $\sum_{a \in A} \frac{1}{a}$ *diverges. Then, there must be arbitrarily long APs in* A.

This conjecture is open even in the case of $k = 3$. But an important special case, when A is the set of primes, was solved by Green and Tao.

Theorem (Green-Tao 2005). *There are arbitrarily long APs in the set of primes.*

The best upper bound for $k = 3$ today is due to Bloom.

Theorem (Bloom 2016). $r_3(N) \leq O\left(N \frac{(\log \log N)^4}{\log N}\right)$.

This improved on previous bounds of $O\left(N\sqrt{\frac{\log \log N}{\log N}}\right)$ (Bourgain 1999), $O\left(N \frac{(\log \log N)^2}{(\log N)^{2/3}}\right)$ (Bourgain 2008), and $O\left(N \frac{(\log \log N)^6}{\log N}\right)$ (Sanders 2012).

There is still a huge gap between Behrend and Bloom and this is considered an important open question.

One can also ask the question about APs in other domains. For example, what is the size of the largest set in $\mathbb{F}_3^n$ that is 3-AP-free? From Roth's proof, one can obtain that $r_3(\mathbb{F}_3^n) \leq 3^n/n$ (this was first observed by Meshulam (1995)). A lower bound was given by Edel (2004): $r_3(\mathbb{F}_3^n) \geq (2.2174)^n$. Very recently, Croot-Lev-Pach (2016) and Ellenberg-Gijswit (2016), using the polynomial method and slice rank, proved that $r_3(\mathbb{F}_3^n) \leq (2.756)^n$.

So much for the history of the problem and related results. In this class we shall see a proof of Roth's theorem.

Roth's Theorem. *Let $\delta > 0$. Then, there exists an N large enough such that any subset $A \subseteq [N]$ with $|A| \geq \delta N$ will have a 3-AP in it.*

To solve this problem, it would be useful to think of A as a subset of $\mathbb{Z}/N\mathbb{Z}$, the additive group of integers modulo N. There are some technicalities about wrap-arounds that we need to handle but we'll figure that out when we get to that point. For now, let us take a detour to understand Fourier analysis over $\mathbb{Z}/N\mathbb{Z}$.

2 Fourier analysis over $\mathbb{Z}/N\mathbb{Z}$

We are interested in functions $f : \mathcal{G} \to \mathbb{C}$ where $\mathcal{G} = \mathbb{Z}/N\mathbb{Z}$ and we want a nice orthonormal basis for the set of functions. The inner-product we would be working with is

$$\langle f, g \rangle = \mathop{\mathbb{E}}_{x \in \mathcal{G}} \left[f(x)\overline{g(x)} \right].$$

If $\omega = e^{2\pi \iota/N}$, the N-th primitive root of unity, then the following forms set of characters of this group:

$$\{ \chi_a : x \mapsto \omega^{ax} \mid a \in \mathcal{G} \}.$$

Note that this is well-defined as ω is an N-th root of unity and hence we only need the exponent modulo N. The following facts are trivial to verify.

Proposition 2.1. *Let $\chi_a : \mathcal{G} \to \mathbb{C}$ be the characters defined above.*

- $\|\chi_a\|_2^2 = \langle \chi_a, \chi_a \rangle = 1$ *for any $a \in \mathcal{G}$.*
- $\langle \chi_a, \chi_b \rangle = \delta_{(a=b)}$, *i.e. it is 1 when $a = b$ and 0 otherwise.*

Therefore, any function $f : \mathcal{G} \to \mathbb{C}$ can be written as $f(x) = \sum_{a \in G} \widehat{f}_a \cdot \chi_a(x)$ and $\widehat{f}_a = \langle f, \chi_a \rangle$ is the a-th Fourier coefficient of f.

Proposition 2.2 (Parseval)**.**

$$\mathop{\mathbb{E}}_{x} \left[|f(x)|^2 \right] = \langle f, f \rangle = \sum_{\alpha} \widehat{f}_\alpha^{\,2}.$$

Observation 2.3. *For any $a \in \mathcal{G}$ and $f : \mathcal{G} \to \mathbb{C}$, we have*

$$\left|\widehat{f}_a\right| \leq \mathop{\mathbb{E}}_x[|f(x)|].$$

2.1 *Convolution of two functions*

If $f, g : \mathcal{G} \to \mathbb{C}$, define the function $(f * g)$ as

$$(f * g)(x) = \mathop{\mathbb{E}}_y [f(y)g(x-y)]]$$

Observation 2.4. *For any $a \in \mathcal{G}$, we have $\widehat{(f * g)}_a = \widehat{f}_a \cdot \widehat{g}_a$.*

Proof.

$$\begin{aligned}
\widehat{(f * g)}_a &= \mathop{\mathbb{E}}_x \left[\mathop{\mathbb{E}}_y [f(y)g(x-y)]\, \overline{\chi_a(x)}\right] \\
&= \mathop{\mathbb{E}}_{x,y} \left[f(y)\overline{\chi_a(y)} \cdot g(x-y)\overline{\chi_a(x-y)}\right] \\
&= \mathop{\mathbb{E}}_y \left[f(y)\overline{\chi_a(y)}\right] \cdot \mathop{\mathbb{E}}_x \left[g(x)\overline{\chi_a(x)}\right] \\
&= \widehat{f}_a \cdot \widehat{g}_a.
\end{aligned}$$

□

A consequence of this and Observation 2.3 is the following fact.

Corollary 2.5. *For any $a \in \mathcal{G}$, and $f, g : \mathcal{G} \to \mathbb{C}$, we have*

$$\left|\widehat{f}_a\right| \cdot |\widehat{g}_a| \leq \mathop{\mathbb{E}}_x [|(f * g)(x)|] = \mathop{\mathbb{E}}_x \left[\left|\mathop{\mathbb{E}}_y [f(y)g(x-y)]\right|\right].$$

If f, g are characteristic functions of some sets S, T respectively, then the RHS gives the expected density of the intersection of S with a *random translate* of T. We will use this at a later point.

3 Structure of the proof of Roth's theorem

The proof of Roth's theorem follows the strategy of "randomness vs structure". We will show that if A is *random* in some sense, then we should have lots of 3-term APs inside A. On the other hand, if A is not *random*, then we will somehow show that A is structured in a precise sense and exploit that somehow. This notion of A being random will be captured by the fact that all non-zero Fourier coefficients of A are small.

The main progress measure would be a density increase. Formally, this is the main lemma that we would be proving.

Lemma 3.1. *Let $A \subseteq [N]$ with $|A| = \delta N$ and $\delta > \sqrt{\frac{50}{N}}$. Then, one of the following must be true:*

1. *A has non-trivial 3-APs over $\mathbb{Z}$.*
2. *There is an arithmetic progression $\tilde{P}$ over $\mathbb{Z}$, with $|\tilde{P}| \geq \left(\frac{\delta^2}{8000}\right)\sqrt{N}$ such that*

$$\left|A \cap \tilde{P}\right| \geq \left(\delta + \frac{\delta^2}{800}\right)\left|\tilde{P}\right|.$$

Notice that since APs are invariant under affine transformations, the above lemma basically allows an inductive argument. Of course, the density can't keep increasing as δ can't go beyond 1 and hence we must encounter non-trivial APs soon enough.

Let us quickly finish the proof assuming this lemma.

Remark. *There would be large numbers thrown (8000, 800 etc.); absolutely no attempt has been made to optimize these constants. The numbers were just made large enough (often exorbitantly) to show the validity of the proof.* ♢

Proof of Roth's Theorem assuming Lemma 3.1. Fix $\delta_0 = \delta$ and let us start with $N_0 = N$ large enough (we'll shortly figure out how large). Suppose $A_0 = A$, whose size is $\delta_0 N_0$ does not contain any 3-APs in it. Then, by Lemma 3.1, there must be an arithmetic progression $\tilde{P}_0$ of size $N_1 \geq \left(\frac{\delta^2}{8000}\right)\sqrt{N_0}$ such that $A_0 \cap \tilde{P}_0$ has density at least $\delta_1 = \left(\delta_0 + \frac{\delta_0^2}{800}\right)$ in $\tilde{P}_0$. Since arithmetic progressions are invariant under shifts and dilations, we have a set $A_1 \subseteq [N_1]$ (A_1 is just the set $A_0 \cap \tilde{P}_0$ appropriately transformed) of density δ_1.

The recurrence we are working with is $\delta_{i+1} = \delta_i\left(1 + \frac{\delta_i}{800}\right)$ and $N_{i+1} \geq \left(\frac{\delta_i^2}{8000}\right)\sqrt{N_i}$. Since the density is increasing at every step by a factor of at least $(1 + \delta/800)$, after $\left(\frac{800}{\delta}\right)$ steps the density would have increased from δ to 2δ. In another $\left(\frac{800}{2\delta}\right)$ steps, it would increase to 8δ. Since the density can never be greater than 1, it follows that this density increase cannot continue for more than $t = \left(\frac{800}{\delta}\right)\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) = \left(\frac{1600}{\delta}\right)$. Lemma 3.1 states that if the density increase doesn't happen, there must be non-trivial APs inside the set, as long as we can ensure $\delta_i > \sqrt{50/N_i}$.

Thus, we want to ensure that N_t is large enough so that $\delta_i > \sqrt{50/N_i}$ throughout; simplest to just make sure $\delta > \sqrt{50/N_t}$ which is implied by $N_t \geq 100/\delta^2$. This can be done by choosing N large enough so that

$$N_t \geq \left(\frac{\delta^2}{800}\right)^t N^{1/2^t} \geq \frac{100}{\delta^2}$$

Suffices to take $N = \exp\left(\exp\left(O\left(\frac{1}{\delta}\right)\right)\right)$. This completes the proof of Roth's theorem. $\square$

Now we need to prove Lemma 3.1. We shall divide this into two cases depending on the Fourier coefficients of the characteristic function of A. We shall abuse notation to use A to refer to the characteristic function $A : \mathcal{G} \to \{0,1\}$ (i.e., $A(x) = 1$ if and only if $x \in A$).

Clearly, $\widehat{A_0} = \delta$ as this is just the density of the set. How large are the other Fourier coefficients of A? We'll first handle the easy case when we know that $\left|\widehat{A_\alpha}\right|$ is small for all $\alpha \neq 0$.

3.1 *When A looks random (small non-trivial Fourier coefficients)*

Lemma 3.2. *Let $A \subseteq [N]$ with $|A| = \delta N$ and $\left|\widehat{A_\alpha}\right| \leq \frac{\delta^2}{100}$ for all $\alpha \neq 0$. Suppose $\delta > \sqrt{\frac{50}{N}}$.*

Then, either A has non-trivial 3-APs in it, or one of the intersections $A \cap [1, N/3)$ or $A \cap [2N/3, N]$ has size at least $\frac{N}{3}\left(\delta + \frac{\delta}{6}\right)$.

In other words, either A has a non-trivial 3-AP or A has an increased density in an arithmetic progression of size $N/3$ (much better than what was required by Lemma 3.1).

Proof. We want to see if A has 3-APs in it. To do this, we can estimate the quantity $\mathbb{E}_{x,y}[A(x)A(x+y)A(x+2y)]$. However, there is a slight catch as, once we start using the characteristic function $A : \mathcal{G} \to \{0,1\}$, the above term estimates the fraction of 3 APs in $\mathcal{G}$ and not just in $\mathbb{Z}$.

To get around this, let $B = A \cap \left[\frac{N}{3}, \frac{2N}{3}\right)$. Now notice that if $x, x+y, x+2y$ is an AP in $\mathcal{G}$ with $x, x+y \in B$, then $(x, x+y, x+2y)$ is also an AP in $\mathbb{Z}$.

Suppose B is substantially smaller than A, that is $|B| \leq |A|/5$. Then, either $\left[1, \frac{N}{3}\right)$ or $\left[\frac{2N}{3}, N\right]$ must have $2|A|/5$ elements of A. But then, the density of A inside this third is $6\delta/5 = \delta + \frac{\delta}{5}$. Therefore, from now on, we shall assume that $|B| \geq |A|/5$.

Let us estimate $\mathbb{E}_{x,y}[B(x)B(x+y)A(x+2y)]$. Expanding via the Fourier representation, we get

$$\begin{aligned}
\mathop{\mathbb{E}}_{x,y}[B(x)B(x+y)A(x+2y)] &= \sum_{\alpha,\beta,\gamma} \widehat{B_\alpha}\widehat{B_\beta}\widehat{A_\gamma} \mathop{\mathbb{E}}_{x}\left[\chi_{\alpha+\beta+\gamma}(x)\right] \mathop{\mathbb{E}}_{y}\left[\chi_{\beta+2\gamma}(x)\right]\\
&= \sum_{\alpha} \widehat{B_\alpha}\widehat{B_{-2\alpha}}\widehat{A_\alpha}\\
&= \widehat{B_0}^2\widehat{A_0} \quad + \quad \sum_{\alpha\neq 0} \widehat{B_\alpha}\widehat{B_{-2\alpha}}\widehat{A_\alpha}\\
&\geq \frac{\delta^3}{25} \quad - \quad \max_{\alpha\neq 0}\left|\widehat{A_\alpha}\right| \cdot \sum_{\alpha\neq 0} \widehat{B_\alpha}\widehat{B_{-2\alpha}}\\
\text{(by Cauchy-Schwartz)} \qquad &\geq \frac{\delta^3}{25} \quad - \quad \max_{\alpha\neq 0}\left|\widehat{A_\alpha}\right| \sqrt{\sum_\alpha \widehat{B_\alpha}^2}\sqrt{\sum_\alpha \widehat{B_{-2\alpha}}^2}\\
&\geq \frac{\delta^3}{25} \quad - \quad 2\max_{\alpha\neq 0}\left|\widehat{A_\alpha}\right| \left(\sum_\alpha \widehat{B_\alpha}^2\right)\\
&\geq \frac{\delta^3}{25} \quad - \quad \frac{\delta^3}{50} = \frac{\delta^3}{50}.
\end{aligned}$$

Hence, we have at least $\frac{\delta^3 N^2}{50}$ many 3-APs in A, and this includes δN trivial (x,x,x) APs. Thus, if $\delta > \sqrt{\frac{50}{N}}$ then this shows that there must be non-trivial 3-APs in A.

Hence, if A looks random (and the stupid corner case doesn't arise), we know that A must have non-trivial 3-APs in it. □

3.2 *When some $\widehat{A_\alpha}$ is large*

We are in the setting when $\left|\widehat{A_\alpha}\right| \geq \frac{\delta^2}{100}$ for some $\alpha \neq 0$. What does it mean to say that this Fourier coefficient is large?

$$\begin{aligned}
\widehat{A_\alpha} &= \mathop{\mathbb{E}}_{x}\left[A(x)\overline{\chi_\alpha(x)}\right]\\
&= \frac{1}{N}\left(\sum_{x\in A}\omega^{\alpha x}\right).
\end{aligned}$$

Therefore, if $\widehat{A_\alpha}$ is large, then the set $\{\omega^{\alpha x} \mid x \in A\}$ is in some sense *aligned* in a common direction. We therefore want to say that there is a large part of A that looks sort-of periodic. More formally, we wish to show that there is some arithmetic progression P (over $\mathbb{Z}$) in $[N]$ of decent size (about $O(\sqrt{N})$ size) such that

$$|A\cap P| \geq \left(\delta + \frac{\delta^2}{800}\right)|P|.$$

That is, the density of A inside a progression P (of decent size) is substantially larger than the density of A in $[N]$. This would finish the proof of Lemma 3.1.

Lemma 3.3. *Let $A \subseteq [N]$ with $|A| = \delta N$ and suppose $\left|\widehat{A_\alpha}\right| \geq \varepsilon$ for some $\alpha \neq 0$. Then, there exists an arithmetic progression $\tilde{P} \subseteq [N]$ (over $\mathbb{Z}$) with $|\tilde{P}| \geq \left(\frac{\varepsilon}{80}\right)\sqrt{N}$ such that*

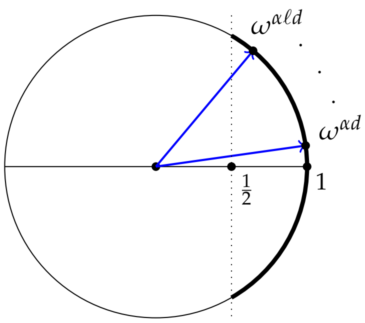
$$\left|A \cap \tilde{P}\right| \geq \left(\delta + \frac{\varepsilon}{8}\right)|\tilde{P}|.$$

This will proceed in three steps.

1. We will find a arithmetic progression P, over $\mathcal{G}$, such that its characteristic function has the property that $|P_\alpha|$ is also large.
2. Since A and P have a large common Fourier coefficient, we will show that A has large intersection with P', a *translate* of P.
3. From P', which is an arithmetic progression over $\mathcal{G}$, we will finally get a progression $\tilde{P}$ over $\mathbb{Z}$ such that A has large intersection with $\tilde{P}$.

Finding a P such that $\widehat{P_\alpha}$ is large

Suppose $P = \{d, 2d, 3d, \cdots, \ell d\}$. Then, $\left|\widehat{P_\alpha}\right| = \frac{1}{N}\left|\omega^{\alpha d} + \cdots + \omega^{\alpha \ell d}\right|$. One way to make sure that all these $\omega^{\alpha j d}$'s are aligned is two ensure that $\omega^{\alpha d} = e^{\iota\theta}$ where the angle θ is really small.



We will set $\ell = \sqrt{N}/10$ and we will find a d such that $d \leq \sqrt{N}$ and "$rd \leq \sqrt{N} \bmod N$". What we mean by the second statement is that there is some $k \in [-\sqrt{N}, \sqrt{N}]$ such that $rd - k = 0 \bmod N$.

Why would such a d exist? Well break up the set $[N]$ in to $\sqrt{N}$ blocks of size $\sqrt{N}$ each and look at the residue of $\alpha, 2\alpha, \ldots, (\sqrt{N}+1)\alpha$. Two of these residues, say $i\alpha$ and $j\alpha$ must fall in the same block. Then $d = i - j$ (say $i > j$) is what we are looking for.

With this choice of d, we know that all of $\left\{\omega^{\alpha d}, \ldots, \omega^{\ell \alpha d}\right\}$ are all within an arc that's at most 1/10-of the circle and would each contribute at least 1/2 to the component along the x-axis. Hence,

$$\widehat{P_\alpha} = \frac{1}{N}\left|\omega^{\alpha d} + \cdots + \omega^{\alpha \ell d}\right| \geq \frac{|P|}{2N}.$$

Thus, we have found a progression P (in fact over $\mathbb{Z}$, as both $d, \ell \leq \sqrt{N}$) such that $\left|\widehat{P}(\alpha)\right|$ is large.

Showing A has large density inside a translate of P

It would be convenient to work with the *balanced* version of A, which we'll call $A' : \mathcal{G} \to \mathbb{C}$ given by $A'(x) = A(x) - \delta$ for all x. Clearly, $\widehat{A'_0} = 0$, and $\widehat{A'_\beta} = \widehat{A_\beta}$ for all non-zero α. In particular, we know from our hypothesis that $\left|\widehat{A'_\alpha}\right| \geq \varepsilon$.

Let $H = A' * P$, where P is the progression we found above. Note that $\mathbb{E}_x[H(x)] = \widehat{H_0} = 0$, as $\widehat{A'_0} = 0$, and we also know that $\left|\widehat{H_\alpha}\right| \geq \varepsilon \cdot \frac{|P|}{2N}$. By Observation 2.3, we have

$$\varepsilon \cdot \frac{|P|}{2N} \leq \mathop{\mathbb{E}}_x\left[|H(x)|\right] = \mathop{\mathbb{E}}_x\left[|H(x)| + H(x)\right] \qquad \text{as } \mathop{\mathbb{E}}_x[H(x)] = 0.$$

This implies that there exists an $x \in [N]$ such that $H(x) + |H(x)| \geq \frac{\varepsilon|P|}{2N}$, which forces $H(x) \geq \frac{\varepsilon|P|}{4N}$. Therefore,

$$\begin{aligned}
\frac{\varepsilon|P|}{4N} &\leq \mathop{\mathbb{E}}_y\left[A'(y)P(y-x)\right] \\
&= \mathop{\mathbb{E}}_y\left[A(y)P(y-x)\right] - \delta \mathop{\mathbb{E}}_y\left[P(y-x)\right] \\
&= \frac{|A \cap (P+x)|}{N} - \frac{\delta|P|}{N} \\
\implies \frac{|A \cap P'|}{|P'|} &\geq \left(\delta + \frac{\varepsilon}{4}\right) \qquad \text{where } P' = (P+x) = \{a + x \mid a \in P\}.
\end{aligned}$$

Note that any translate P' of P is also an arithmetic progression (in $\mathcal{G}$). Hence, we have found a progression P' (over $\mathcal{G}$) of size $\sqrt{N}/10$ such that the density of A in P is substantially bigger than its density in $[N]$. We are almost done, except that the progression P' is a progression over $\mathcal{G}$ and not necessarily over $\mathbb{Z}$. This is the last step.

Finding a progression $\tilde{P}$ over $\mathbb{Z}$ in which A has large density

We started of with P that was indeed a progression over $\mathbb{Z}$. However, when we move to a translate P', it may no longer be a progression over $\mathbb{Z}$. Nevertheless, since $d \leq \sqrt{N}$ and $\ell = \sqrt{N}/10$, there can be at most one wrap-around. Hence, we can write $P' = P_1 \cup P_2$ where P_1 and P_2 are arithmetic progressions over $\mathbb{Z}$. Let's assume that $|P_1| \leq |P_2|$.

Since A has density at least $\delta + \frac{\varepsilon}{4}$ in P', it follows that A must have density $\delta + \frac{\varepsilon}{4}$ in either P_1 or P_2. The issue is that we want a decently sized $\tilde{P}$ in which A has large density and A could have large density in P_1 but P_1 could be really small. Hence we need to deal with two cases depending on whether P_1 is size-able or not.

Case 1: $|P_1| \geq \left(\frac{\varepsilon}{8}\right)|P|$.

Then, both P_1 and P_2 are size-able and hence can choose $\tilde{P}$ to be the P_i in which A has larger density.

Case 2: $|P_1| \leq \left(\frac{\varepsilon}{8}\right)|P|$.

In this case,

$$\begin{aligned} |A \cap P_2| &\geq \left(\delta + \frac{\varepsilon}{4}\right)|P'| - |P_1| \\ &\geq \left(\delta + \frac{\varepsilon}{8}\right)|P'| \\ &\geq \left(\delta + \frac{\varepsilon}{8}\right)|P_2| \end{aligned}$$

Hence $\tilde{P} = P_2$ satisfies our requirements.

In either case, we have found an arithmetic progression $\tilde{P}$ (over $\mathbb{Z}$) with $|\tilde{P}| \geq \left(\frac{\varepsilon}{80}\right)\sqrt{N}$ such that

$$\left|A \cap \tilde{P}\right| \geq \left(\delta + \frac{\varepsilon}{8}\right)\left|\tilde{P}\right|.$$

This completes the proof of Lemma 3.3. Combining this with Lemma 3.2, we complete the proof of Lemma 3.1. ☐

References

[Lot17] Adam Lott. *Roth's theorem on arithmetic progressions*. BA honor's thesis, University of Rochester, 2017.

[O'D07] Ryan O'Donnell. *15-859s 2007: Analysis of boolean functions*, 2007. (A course on Boolean Function Analysis at CMU, Spring 2007).