Roth's Theorem in $\mathbb{F}_p^n$ and $\mathbb{Z}_N$

Stephen Zhou

November 2025

Abstract

Roth's Theorem states that any subset of the natural numbers with positive density contains a three term progression. We introduce the method Fourier analysis on finite groups, and use it to prove a quantitative versions of Roth's theorem in $\mathbb{F}_p^n$ and $\mathbb{Z}_N$, using the density increment method.

1 Introduction

Additive combinatorics is the subfield of combinatorics that asks questions about the additive properties of numbers, instead of their multiplicative properties, such as factoring. An example theorem in additive combinatorics is Van Der Waerden's Theorem.

Theorem 1. *Let $A_1, \dots, A_r$ be subsets of $\mathbb{N}$ with $A_1 \cup \cdots \cup A_r = \mathbb{N}$. Then for any value of k, there exists an A_i that has a length k arithmetic progression.*

A length k arithmetic progression is also known as a k-AP. Equivalently, the partition into $A_1, \dots, A_r$ can be viewed as coloring the integers with r colors. Van der Waerden's Theorem was strengthened into Szemeredi's Theorem.

Definition 1. *The density of $A \subseteq \mathbb{N}$ is*

$$\lim_{N \to \infty} \frac{A \cap [N]}{N} \tag{1}$$

if it exists.

Theorem 2 (Szemeredi's Theorem)**.** *Let $A \in \mathbb{N}$ have positive density. Then for any $k \in \mathbb{N}$, A has an arithmetic progression of length $\mathbb{N}$.*

Notice that any coloring the integers with r colors must have one color with positive density. Szemeredi's theorem remained as a conjecture for a long time, before Szemeredi proved it using graph theory in 1975. Klaus Roth proved the $k = 3$ case of this theorem in 1956 using Fourier analysis. It is this theorem we will be focusing on.

Theorem 3 (Roth's Theorem)**.** *Let $A \in \mathbb{Z}$ have positive density. Then A has a 3-AP.*

Actually, Roth found a quantitative bound.

Theorem 4 (Roth's Theorem, quantitative version)**.** *Let $A \in [N]$ be 3-AP free. Then the maximal size of $|A|$ is $O(N/\log\log N)$.*

We will prove this version of his Roth's theorem, as well as an analogue in $\mathbb{F}_p^n$ with a simpler proof and a bound of $O(p^n/n)$.

2 Fourier analysis on $\mathbb{F}_p^n$

In $\mathbb{R}$, Fourier series allow us to write periodic functions $f : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$ as sums of multiples of the "frequencies" $e^{2\pi inx}$. We can do a similar thing for functions $f : \mathbb{F}_p^n \to \mathbb{C}$. Define $\omega = e^{\frac{2\pi i}{p}}$.

Definition 2. *The Fourier transform of $f : \mathbb{F}_p^n \to \mathbb{C}$ is the function $\widehat{f} : \mathbb{F}_p^n \to \mathbb{C}$ defined by*

$$\widehat{f}(r) = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} f(x)\gamma_{-r}(x). \tag{2}$$

Notice that $\widehat{f}(0) = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} f(x)$ is the average of $f(x)$. This coefficient will often be distinguished from the others. The reason anyone cares about Fourier transforms is

Theorem 5 (Fourier inversion)**.** *Let $f : \mathbb{F}_p^n \to \mathbb{C}$. Then*

$$f(x) = \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r)\gamma_r(x). \tag{3}$$

Notice that

$$\sum_{x \in \mathbb{F}_p^n} \gamma_r(x) = \begin{cases} 1 & \text{if } r = 0 \\ 0 & \text{if } r \neq 0 \end{cases}. \tag{4}$$

This allows us to prove Fourier inversion by just plugging in the definition of a Fourier transform into the formula.

We can get a much more conceptual explanation by viewing the set of all functions $f : \mathbb{F}_p^n \to \mathbb{C}$ as a p^n-dimensional vector space. We can define an inner product on this vector space as

$$\langle f, g \rangle = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} \overline{f(x)} g(x), \tag{5}$$

and a norm

$$||f||_2 = \langle f, f \rangle^{1/2}. \tag{6}$$

It is simple to check that $\langle \cdot, \cdot \rangle$ is an inner product. Then

$$\langle \gamma_r, \gamma_s \rangle = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} \gamma_{s-r}(x) = \begin{cases} 1 & \text{if } r = s \\ 0 & \text{if } r \neq s \end{cases} \tag{7}$$

so the characters γ_r are mutually orthogonal unit vectors. Since there are p^n different characters, and the vector space is p^n-dimensional, the characters form an orthonormal basis. We call this basis the *fourier basis*. Thus, for any $f : \mathbb{F}_p^n \to \mathbb{C}$, we can write

$$f(x) = \sum_{r \in \mathbb{F}_p^n} \langle \gamma_r, f \rangle \gamma_r. \tag{8}$$

Since $\langle \gamma_r, f \rangle = \widehat{f}(r)$, we have proven Theorem 5.

Next, we will prove Parseval's identity, which for $\mathbb{F}_p^n$, basically converts the inner product for functions $f(x)$ in "physical space" into an inner product for their Fourier transforms $\widehat{f}(r)$ in "frequency space". This new inner product isn't all that different.

Definition 3. *Let the inner product $\langle \cdot, \cdot \rangle_{\ell^2}$ be defined as*

$$\langle \widehat{f}, \widehat{g} \rangle_{\ell^2} = \sum_{r \in \mathbb{F}_{\text{!`}}^{\times}} \overline{\widehat{f}(r)} \widehat{g}(r). \tag{9}$$

and the norm be $||f||_{\ell^2} = \langle f, f \rangle_{\ell^2}^{1/2}$.

It is clear that $\langle \cdot, \cdot \rangle_{\ell^2}$ is an inner product.

Theorem 6 (Parseval's identity)**.** *For any $f, g : \mathbb{F}_p^n \to \mathbb{C}$,$\langle f, g \rangle = \langle \widehat{f}, \widehat{g} \rangle_{\ell^2}$, or*

$$\frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} \overline{f(x)} g(x) = \sum_{r \in \mathbb{F}_p^n} \overline{\widehat{f}(r)} \widehat{g}(r). \tag{10}$$

Specifically, in the case $f = g$, we have that $||f||_2 = ||\widehat{f}||_{\ell^2}$, or

$$\frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} |f(x)|^2 = \sum_{r \in \mathbb{F}_p^n} |\widehat{f}(r)|^2. \tag{11}$$

Proof. By the linearity of the inner product and the orthogonality of the Fourier basis, we have

$$\langle f, g \rangle = \sum_{r \in \mathbb{F}_p^n} \overline{\widehat{f}(r)} \langle \gamma_r, g \rangle = \sum_{r \in \mathbb{F}_p^n} \sum_{s \in \mathbb{F}_p^n} \overline{\widehat{f}(r)} g(r) \langle \gamma_r, \gamma_s \rangle = \sum_{r \in \mathbb{F}_p^n} \overline{\widehat{f}(r)} g(r). \tag{12}$$

□

Now we define the convolution operation.

Definition 4. *Let $f, g : \mathbb{F}_p^n \to \mathbb{C}$. Then the convolution $f * g : \mathbb{F}_p^n \to \mathbb{C}$ is defined as*

$$(f * g)(x) = \frac{1}{p^n} \sum_{y \in \mathbb{F}_p^n} f(y)g(x - y). \tag{13}$$

That is, $f * g$ is the average of $f(y)g(z)$ over y, z such that $x = y + z$.

The reason we care about convolutions is that the Fourier transform turns convolutions into multiplication.

Theorem 7. *For any $f, g : \mathbb{F}_p^n \to \mathbb{C}$,*

$$\widehat{f * g}(r) = \widehat{f}(r)\widehat{g}(r). \tag{14}$$

Proof. Expanding, we get

$$\widehat{f * g}(r) = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} (f * g)(x)\omega^{-r \cdot x} = \frac{1}{p^n} \sum_{x \in \mathbb{F}_p^n} \frac{1}{p^n} \sum_{y \in \mathbb{F}_p^n} f(y)g(x - y)\omega^{-r \cdot x} = \frac{1}{p^{2n}} \sum_{x \in \mathbb{F}_p^n} \sum_{y \in \mathbb{F}_p^n} f(y)g(x - y)\omega^{-r \cdot x}. \tag{15}$$

Swapping the order of summation and substituting in $x = z + y$, we have

$$\frac{1}{p^{2n}} \sum_{y \in \mathbb{F}_p^n} \sum_{z \in \mathbb{F}_p^n} f(y)g(z)\omega^{-r \cdot (y+z)} = \left(\frac{1}{p^n} \sum_{y \in \mathbb{F}_p^n} f(y)\omega^{-r \cdot y} \right) \left(\frac{1}{p^n} \sum_{z \in \mathbb{F}_p^n} f(z)\omega^{-r \cdot z} \right) = \widehat{f}(r)\widehat{g}(r). \tag{16}$$

Thus $\widehat{f * g}(r) = \widehat{f}(r)\widehat{g}(r)$. □

3 Roth's Theorem for $\mathbb{F}_p^n$

We now define a functional that helps us count the number of 3-APs in a subset $A \subseteq \mathbb{F}_p^n$.

Definition 5. *For any $f, g, hf : \mathbb{F}_p^n \to \mathbb{C}$, we define*

$$\Lambda(f, g, h) = \frac{1}{p^{2n}} \sum_{x,y \in \mathbb{F}_p^n} f(x)g(x + y)h(x + 2y). \tag{17}$$

and

$$\Lambda_3(f) = \Lambda(f, f, f). \tag{18}$$

Notice that if f is an indicator function $1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$, then $\Lambda_3(1_A)$ is the probability that a randomly chosen(possibly trivial) 3-AP is in A.

Our proof of Roths Theorem has three steps.

1. Prove that if A does not contain a nontrivial 3-AP, then 1_A has a large Fourier coefficient $\widehat{1_A}(r)$.
2. If 1_A has a large Fourier coefficient $\widehat{1_A}(r)$, then there exists a hyperplane $P \in \mathbb{F}_p^n$ such that the density of A in P is large.
3. Repeat the density increment.

To begin, we express Λ_3 in terms of $\widehat{f}$.

Lemma 1. *Let $f : \mathbb{F}_p^n \to \mathbb{C}$. Then*

$$\Lambda(f,g,h) = \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r)\widehat{g}(-2r)\widehat{h}(r). \tag{19}$$

Proof. Let $a = x, b = x + y, c = x + 2y$ and $g_1(b) = g(-b/2)$. Applying Fourier inversion and the convolution identity to Λ, we get

$$\begin{aligned}
\Lambda(x,y,z) &= \sum_{x,y \in \mathbb{F}_p^n} f(x)g(x+y)h(x+2y) \\
&= \sum_{a-2b+c=0} f(a)g(b)h(b) \\
&= \sum_{a+b+c=0} f(a)g_1(b)h(b) \\
&= f * g_1 * h(0) \\
&= \sum_{r \in \mathbb{F}_p^n} \widehat{f * g_1 * h}(r) \\
&= \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r)\widehat{g_1}(r)\widehat{h}(r) \\
&= \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r)\widehat{g}(-2r)\widehat{h}(r). \qquad (20)
\end{aligned}$$

□

In particular, we have $\Lambda_3(f) = \sum_{r \in \mathbb{F}_p^n} \widehat{f}(r)^2 \widehat{f}(-2r)$.

The following lemma formalizes our intuition that a subset of A with small Fourier coefficients should have about as many 3-APs as a random set.

Lemma 2. *Let $A \in \mathbb{F}_p^n$ and $\alpha = |A|/p^n$. Then*

$$|\Lambda_3(1_A) - \alpha^3| \leq \max_{r \neq 0} |\widehat{1_A}(r)| ||f||_2^2. \tag{21}$$

Proof. By Lemma 1, we get that

$$\Lambda_3(1_A) = \sum_r \widehat{1_A}(r)^2\widehat{1_A}(-2r) = \widehat{1_A}(0)^3 + \sum_{r\neq 0} \widehat{1_A}(r)^2\widehat{1_A}(-2r). \tag{22}$$

Since $\widehat{1_A}(0) = \alpha$,

$$|\Lambda_3(1_A) - \alpha^3| \leq \sum_{r\neq 0} |\widehat{1_A}(r)|^2|\widehat{1_A}(-2r)| \leq \max_{r\neq 0}|\widehat{1_A}(r)| \sum_{r\neq 0}|\widehat{1_A}(r)|^2 = \max_{r\neq 0}|1_A|\|1_A\|_2^2.$$

□

Now we can complete the first step.

Lemma 3. *Let $A \in \mathbb{F}_p^n$ and $\alpha = |A|/p^n$. Then if A is 3-AP-free and $\alpha^2 \geq \frac{2}{p^n}$, there exists $r \neq 0$ such that $|\widehat{1_A}(r)| \geq \alpha^2/2$.*

Proof. Since A is 3-AP-free, all 3-APs are trivial, so $\Lambda_3(1_A) = \frac{\alpha^3}{p^n}$. By Lemma 2

$$\alpha^3 - \frac{\alpha}{p^n} \leq \max_{r\neq 0}|\widehat{1}_A(r)|\|\widehat{1}_A\| = \max_{r\neq 0}|\widehat{1}_A(r)|\alpha. \tag{23}$$

Since $\alpha^2 \geq \frac{2}{p^n}$, we get that $\alpha^3 - \frac{\alpha}{p^n} \geq \frac{\alpha^3}{2}$. So $\max_{r\neq 0}|\widehat{1_A}(r)| \geq \alpha^2/2$. □

Now we prove that a large Fourier coefficient implies a density increment.

Lemma 4. *Let $A \in \mathbb{F}_p^n$ and $\alpha = |A|/p^n$. If there exists $r \in \mathbb{F}_p^n$ such that $|\widehat{1_A}(r)| \geq \delta$, then A has density at least $\alpha + \delta/2$ in some hyperplane.*

Proof. Let $r^\perp$ be the set of vectors in $\mathbb{F}_p^n$ orthogonal to r. Viewing $r^\perp$ as a subspace of $\mathbb{F}_p^n$, $r^\perp$ has cosets $R_0 = r^\perp, R_1, \ldots, R_{p-1}$. For any coset R_i, $v \cdot r$ is constant for all $v \in R_i$. So we will assume that $v \cdot r = i$ for all $v \in R_i$. Let $\alpha_i = |A \cap R_i|/p^{n-1}$ be the density of A in R_i. Then we have

$$\widehat{1_A}(r) = \frac{1}{p^n}\sum_{x\in\mathbb{F}_p^n} 1_A(x)\omega^{-r\cdot x} = \frac{1}{p}\sum_{i=0}^{p-1}\alpha_i\omega^i. \tag{24}$$

The last equality follows by grouping the xs by coset. Notice that $p\alpha = \sum_i \alpha_i$. By the triangle inequality,

$$p\delta \leq \left|\sum_{i=0}^{p-1}\alpha_i\omega^i\right| = \left|\sum_{i=0}^{p-1}(\alpha_i - \alpha)\omega^i\right| \leq \sum_{i=0}^{p-1}|(\alpha_i - \alpha)| = \sum_{i=0}^{p-1}(|(\alpha_i - \alpha)| + \alpha_i - \alpha).$$

So there exists some j such that $|(\alpha_j - \alpha)| + \alpha_j - \alpha \geq \delta$. This is equivalent to $\alpha_j - \alpha \geq \delta/2$. □

Combining the previous two lemmas, we get

Lemma 5 (Density Increment)**.** *Let $A \in \mathbb{F}_p^n$ and $\alpha = |A|/p^n$. If A is 3-AP-free and $\alpha^2 > 2/p^n$, then A has density at least $\alpha + \alpha^2/4$ in some hyperplane.*

By repeating this, we prove Roth's Theorem for $\mathbb{F}_p^n$.

Proof of Roth's Theorem in $\mathbb{F}_p^n$. Repeatedly apply Lemma 5 to a 3-AP-free $A \in \mathbb{F}_p^n$. Say we can do this m times. Then we get a chain of subspaces $\mathbb{F}_p^n = V_0 \supseteq V_1 \supseteq \cdots \supseteq V_{m-1}$, where V_i has dimension $n - i$. Let the strictly increasing sequence $\alpha_i = |A \cap V_i|/p^{n-i}$ be the density of A in V_i. Then for $0 \leq i < m$, $2\alpha_i^{-2} \leq |V_i|$, by the conditions of Lemma 5. And since we can only apply Lemma 5 m times, we have that $2\alpha_m^{-2} \geq |V_i|$. Each round, the density increases by at most $\alpha^2/4$. So it takes at most $\lceil 4/\alpha \rceil$ turns for α_i to double. Then it takes $\lceil 1/\alpha \rceil$ turns for α_i to double again. Since $\alpha_i \leq 1$, the total number of rounds m is bounded by a geometric sequence, meaning $m = O(1/\alpha)$. But $|V_{m-1}| = p^{n-m} \leq 2\alpha_m^{-2} \leq 2\alpha^{-2}$. Taking logs, we get that $\alpha = O(1/n)$, which is exactly what we wanted. □

We have gotten a bound of $O(p^n/n)$ on cap sets. Improving this bound is an interesting problem.

4 Fourier analysis in $\mathbb{Z}_N$

Now we will modify our proof to count arithmetic progressions in $\mathbb{Z}$. The quantitative version of Roth's Theorem we will prove is

Theorem 8. *Let $A \in [N]$ be 3-AP free. Then the maximal size of A is $O(N/\log\log n)$.*

We will use Fourier analysis on $\mathbb{Z}_n = \mathbb{Z}/N\mathbb{Z}$. This is a variation of Roth's original proof, which used Fourier analysis over $\mathbb{Z}$. Fourier analysis over $\mathbb{Z}_N$ is basically analogous to Fourier analysis over $\mathbb{F}_p^n$.

Definition 6. *Let $f : \mathbb{Z}_N \to \mathbb{C}$. Then the Fourier transform of f is*

$$\widehat{f}(r) = \frac{1}{N} \sum_{x \in \mathbb{Z}_N} f(x)\gamma^{-r}(x), \tag{25}$$

where $\gamma_r(x) = e^{2\pi irx/N}$.

All the theorems proven about Fourier transforms in $\mathbb{F}_p^n$ also hold for Fourier transforms over $\mathbb{Z}_N$, but with p^n replaced by $1/N$. More generally, this type of Fourier analysis works for all finite abelian groups.

5 Roth's Theorem over $\mathbb{Z}$

From now on, we will assume N is odd, so that 2 has an inverse in $\mathbb{Z}/N\mathbb{Z}$. Our proof of Roth's Theorem for $\mathbb{Z}$ will have the same density increment as the proof for $\mathbb{F}_p^n$. Recall that the function Λ helps count the number of 3-APs in A.

Definition 7. *Let $f, g, h : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$. Define*

$$\Lambda(f, g, h) = \frac{1}{N^2} \sum_{x,y \in \mathbb{Z}/N\mathbb{Z}} f(x)g(x+y)h(x+2y). \tag{26}$$

Although $\Lambda(1_A, 1_A, 1_A)$ counts 3-AP density in $\mathbb{Z}_N$, these are not exactly the same as 3-APs in $\mathbb{Z}$ because we can "cycle back". So we need to be a bit more clever. Let $B = A \cap [N/3, 2N/3]$. Notice that if x, y, z are a 3-AP with $x + z - 2y = 0$ in $\mathbb{Z}_N$, then x, y, z is also an 3-AP in $\mathbb{Z}$ if $x, y \in B$. Thus the number of $\mathbb{Z}$ 3-AP's in A is at least $\Lambda(1_B, 1_B, 1_A)$.

We now prove an analogue to Lemma 2 for $\Lambda(1_b, 1_A, 1_A)$.

Lemma 6. *Let $A \in [N]$, $B = A \cap [\frac{N}{3}, \frac{2N}{3}]$ and define $\alpha = |A|/N, \beta = |B|/N$. Then*

$$|\Lambda(1_B, 1_B, 1_A) - \alpha^2\beta| \leq \max_{r \neq 0} |\widehat{1_A}(r)|\beta. \tag{27}$$

Proof. By Lemma 1, we have $\Lambda(1_B, 1_B, 1_A) = \sum_r \widehat{1_B}(r)^2\widehat{1_A}(r) = \alpha^2\beta + \sum_{r\neq 0} \widehat{1_B}(r)^2\widehat{1_A}(r)$. Thus,

$$\begin{aligned}
|\Lambda(1_B, 1_B, 1_A) - \alpha^2\beta| &= \left| \sum_{r\neq 0} \widehat{1_B}(r)^2\widehat{1_A}(r) \right| \\
&\leq \sum_{r\neq 0} |\widehat{1_B}(r)^2\widehat{1_A}(r)| \leq \max_{r\neq 0} |\widehat{1_A}(r)| \sum_{r\neq 0} \widehat{1_B}(r)^2 \\
&= \max_{r\neq 0} |\widehat{1_A}(r)| \|1_B\|_2^2 = \max_{r\neq 0} |\widehat{1_A}(r)|\beta N.
\end{aligned}$$

□

We conclude that a a 3-AP-free set must have a large Fourier coefficient.

Lemma 7. *Let $A \in [N], B = (A \cap [\frac{N}{3}, \frac{2N}{3}])$ and $\alpha = |A|/N, \beta = |B|/N$. If A is $\mathbb{Z}$ 3-AP-free and $N \geq 32\alpha^{-2}$, then either $\widehat{1_A}(k) \geq \frac{\alpha^2}{8}$ for some $k \neq 0$, or $\beta \leq \alpha/4$.*

Proof. If A is 3-AP-free, then $\Lambda(\widehat{1_B}, \widehat{1_B}, \widehat{1_A}) = \beta/N$. Now assume that $\widehat{1_A}(k) < \frac{\alpha^2}{8}$ for all $k \neq 0$ and $\beta > \alpha/4$.

Then, by the previous lemma, we get

$$\Lambda(\widehat{1_B}, \widehat{1_B}, \widehat{1_A}) \geq \alpha\beta^2 - \frac{\alpha^2}{8}\beta. \tag{28}$$

The bounds turn this into

$$\Lambda(\widehat{1_B}, \widehat{1_B}, \widehat{1_A}) \geq \frac{\alpha^3}{32}.$$

But $N \geq 32\alpha^{-2}$ and $\beta \leq \alpha/4$, so this contradicts $\Lambda(\widehat{1_B}, \widehat{1_B}, \widehat{1_A}) = \beta/N$. □

Now we try to find a density increment. In $\mathbb{F}_p^n$, we found a subspace A had higher density in. But $\mathbb{Z}$ doesn't have subspaces, so we need to use another kind of substructure. It turns out to be appropriate to restrict to arithmetic subprogressions.

To construct the arithmetic progression we want, we need to use the pigeonhole principle.

Lemma 8. *Let $l = N/\lfloor\sqrt{N}\rfloor$. Then for any r, there exists $d \leq l$ such that*

$$rd \leq l \pmod{N}. \tag{29}$$

Proof. Divide the equivalence classes modulo N into $\lfloor N \rfloor$ intervals of size l each. Then by the pigeonhole principle, there exist p, q such that $pr - qr \leq l \pmod{N}$. Then $d = p - q$ works. □

Now we can prove the density increment.

Lemma 9. *Say that $A \in [N]$ is such that $\widehat{1_A}(r) \geq \delta$ for some $r \neq 0$. Then there exists some subprogression B in $\mathbb{Z}_N$ with length at least $\sqrt{N}/8$ such that A has density at least $\alpha + \delta/8$ in B.*

Proof. Let d and l be as in the previous lemma. Let B_0 be the arithmetic progression in $\mathbb{Z}_N$ of length $\sqrt{N}/2\pi$ given by

$$-2d, -d, 0, d, 2d, 3d, \ldots \tag{30}$$

Notice that B_0 is the union of two arithmetic progressions of $[N]$. Let $\beta_0 = |B_0|/N$. Then we calculate that

$$\begin{aligned}\left|\widehat{1_{B_0}}(r) - \beta_0\right| &= \left|\sum_{x\in\mathbb{Z}_N} 1_{B_0}(x)\left(e^{-\frac{2\pi i}{N}rx} - 1\right)\right| \\ &= \left|\sum e^{-\frac{2\pi i}{N}rdx} - 1\right| \leq \sum \left|e^{-\frac{2\pi i}{N}rdx} - 1\right| \\ &\leq 2\sum \frac{2\pi}{N} rdx \leq \frac{4\pi l}{N}\sum x \leq \frac{\beta_0}{2}.\end{aligned}$$

This implies that $\left|\widehat{1_{B_0}}(r)\right| \geq \frac{\beta_0}{2}$. The desired subprogression will be $B_c = B_0 + c$ for some $c \in \mathbb{Z}_N$.

Now define the balanced indicator of A to be $f_A = 1_A - \alpha$. f_a is called balanced because its mean is 0. Notice that for any $B \subseteq \mathbb{Z}_N$, we have that

$$\sum_{x=0}^{N-1} f_A(x)1_B(x) = \sum_{x\in B} f_A(x) = (1-\alpha)|A\cap B| - \alpha(|B| - |A\cap B|) = |A\cap B| - \alpha|B|,$$

so

$$\sum_{x=0}^{N-1} f_A(x)1_B(x) \qquad\qquad \geq \qquad\qquad \gamma|B|$$

is equivalent to $|A \cap B| \geq |\alpha + \gamma||B|$. So we need to prove that there exists c such that

$$G(c) = \sum_{x=0}^{N-1} f_A(x) 1_{B_c}(x) = \sum_{x=0}^{N-1} f_A(x) 1_{B_0}(x-c) \geq \frac{1}{4}\gamma|B_0|.$$

We calculate the Fourier transform of $G(c)$ to be $\widehat{G}(r) = N\widehat{f_A}(r)\widehat{1_{B_0}}(r)$. Since

$$\sum |G(c)| \geq |\widehat{G}(r)| \geq \frac{1}{2}\gamma N|B_0|$$

and

$$\sum G(c) = 0,$$

we get that there exists c such that $|G(c)| + G(c) \geq \frac{1}{2}\gamma|B_0|$, or $G(c) \geq \frac{1}{4}\gamma|B_0|$ as desired. Since B_c is the union of two arithmetic progressions $\mathbb{Z}_N$, there is a arithmetic progression in $\mathbb{Z}_N$ where A has density $\alpha + \frac{\delta}{8}$. □

Finally, we have the density increment by combining the previous three lemmas.

Lemma 10. *If $A \in [N]$ does not have a 3-AP, and $N \geq 32\alpha^{-2}$, there exists a subprogression of size $O(\sqrt{N})$ such that the density of N on the subprogression is at least $\alpha + \alpha^2/64$.*

Repeatedly applying this lemma to A and analyzing the doubling time gives us the bound $O(N/\log\log N)$.

Proof of Roth's Theorem for $\mathbb{Z}$. As in the proof for $\mathbb{F}_p^n$, we find that there are at most $O(1/a)$ density increments by looking at the doubling time. Say the process ends after m steps and density α_m. The final subprogression has size $O(N^{1/2^m})$ and this is less α^{-2}, up to a constant factor. Taking logarithms twice gives that $\alpha = O(1/\log\log N)$. □