3-term Arithmetic Progressions

Siddharth Iyer

September 2022

1. Roth's Theorem

Theorem 1. *For n sufficiently large, any subset $A \subseteq \{1, \dots, n\}$ with density $\delta = \Omega(1/\log\log n)$ must contain a 3-term arithmetic progression.*

Suppose $A \subseteq \{1, \dots, n\}$ of density δ, that does not contain 3-term arithmetic progressions. It will be convenient to interpret A as a subset of $\mathbb{Z}_p$, where p is a prime between $2n$ and $4n$, and addition being mod p.

Fourier Analysis over $\mathbb{Z}_N$. The Fourier transform of a function $f : \mathbb{Z}_N \to \mathbb{C}$ is given as

$$f(x) = \sum_r \widehat{f}(r)\omega^{rx},$$

where ω is the n-th root of unity and $\widehat{f}(r) = \mathbf{E}_y[f(y)\omega^{-ry}]$. The following fact known as Plancherel's theorem, follows from the above definitions of the Fourier coefficients

$$\sum_r \left|\widehat{f}(r)\right|^2 = \mathop{\mathbf{E}}_x[|f(x)|^2] = \|f\|_2^2.$$

Counting the number of 3-term APs. For a function $f : \mathbb{Z}_p \to \mathbb{C}$, define

$$\Lambda_3(f) := \sum_{x,y\in\mathbb{Z}_p} f(x-y)f(x)f(x+y).$$

Let 1_A be the indicator function corresponding to the subset A.

Observation 2. $\Lambda_3(1_A) = \delta n$.

Proof. The summands corresponding to $y = 0$ contribute to $\Lambda_3(1_A)$ whenever $x \in A$. Hence, $\Lambda_3(1_A) \geq \delta n$.

When $y \neq 0$, we can assume without loss of generality that $y \in \{1, \dots, (p-1)/2\}$. Otherwise, we have $-y \in \{1, \dots, (p-1)/2\}$, and

$$1_A(x-y)1_A(x)1_A(x+y) = 1_A(x-(-y))1_A(x)1_A(x+(-y)).$$

Assume to a contradiction that $1_A(x-y) = 1_A(x) = 1_A(x+y) = 1$. In particular, $x \in \{1, \dots, n\} \subseteq \{1, \dots, (p-1)/2\}$. Moreover, $x+y \in \{2, \dots, p-1\}$, and hence $x + y \bmod p = x + y$. Finally, $x - y \in A$ implies that $y \le x$, for otherwise, $x - y \bmod p \in \{(p+1)/2, \dots, p-1\}$. Therefore, the integers $x - y, x, x + y$, are distinct elements of A and form an arithmetic progression, which is impossible. $\square$

$1_A - \delta 1_{[n]}$ **has a large Fourier coefficient.** For three functions $f, g, h : \mathbb{Z}_p \to \mathbb{C}$, define

$$\Lambda(f, g, h) := \sum_{x,y \in \mathbb{Z}_p} f(x - y) g(x) h(x + y).$$

Note that $\Lambda(f, f, f) = \Lambda_3(f)$. Using the Fourier transforms of f, g and h, we have

$$\Lambda(f, g, h) = p^2 \sum_r \widehat{f}(r) \widehat{g}(-2r) \widehat{h}(r). \tag{1}$$

Let us denote $1_A - \delta 1_{[n]} := g$.

Claim 3. *If $\delta^2 \ge 8/n$, then there exists $r \ne 0$ such that $|\widehat{g}(r)| \ge C\delta^2$, for some absolute constant $C > 0$.*

Proof. Note that $\Lambda_3(1_A) = \Lambda_3(g + \delta 1_{[n]})$. Using the Λ function, we can write

$$\begin{aligned}\Lambda_3(g + \delta 1_{[n]}) = &\Lambda_3(g) + \delta\Lambda(g, 1_{[n]}, g) + 2\delta\Lambda(g, g, 1_{[n]}) + \\ &\delta^3 \Lambda_3(1_{[n]}) + \delta^2 \Lambda(1_{[n]}, g, 1_{[n]}) + 2\delta^2 \Lambda(1_{[n]}, 1_{[n]}, g).\end{aligned}$$

We will show that the dominant term in the above sum is due to $\delta^3 \Lambda_3(1_{[n]})$. First, it is easy to see that $\Lambda_3(1_{[n]}) \ge n^2/4$. Indeed, for any $x \in [n]$, we have $1_{[n]}(x - y) 1_{[n]}(x) 1_{[n]}(x + y) = 1$, when $y \in \{0\} \cup ([x - 1] \cap [n - x])$. Hence,

$$\Lambda_3(1_{[n]}) \ge \sum_{x=1}^{n} \min\{x, n - x + 1\} \ge 2 \sum_{x=1}^{n/2} x \ge n^2/4.$$

Next, we upper bound the remaining terms using the Fourier transforms of the respective functions and Equation (1). For example,

$$\begin{aligned}\Lambda(g, g, 1_{[n]}) &= p^2 \Big| \sum_r \widehat{g}(r) \widehat{g}(-2r) \widehat{1_{[n]}}(r) \Big| \\ &\le \max_r |\widehat{g}(r)| \sum_r |\widehat{g}(-2r)| \cdot \left| \widehat{1_{[n]}}(r) \right| && \text{(triangle inequality)} \\ &\le \max_r |\widehat{g}(r)| \sqrt{\sum_r |\widehat{g}(r)|^2} \sqrt{\sum_r \left| \widehat{1_{[n]}}(r) \right|^2} && \text{(Cauchy-Schwarz)} \\ &= \max_{r \ne 0} |\widehat{g}(r)| \cdot \|g\|_2 \cdot \|1_{[n]}\|_2, && \text{(Plancherel)}\end{aligned}$$

where in the last line we also use the fact that $\widehat{g}(0) = 0$. The same analysis also shows that $\Lambda(g, 1_{[n]}, g) \le \max_{r \ne 0} |\widehat{g}(r)| \cdot \|g\|_2 \cdot \|1_{[n]}\|_2$. Similarly,

$$\Lambda(g, 1_{[n]}, 1_{[n]}), \Lambda(1_{[n]}, g, 1_{[n]}) \le p^2 \max_{r \ne 0} |\widehat{g}(r)| \cdot \|1_{[n]}\|_2^2,$$

and $\Lambda_3(g) \leq p^2 \max_{r\neq 0} |\widehat{g}(r)| \cdot \|g\|_2^2$. Therefore,

$$\begin{aligned}\Lambda_3(g+\delta 1_{[n]}) &\geq \delta^3 n^2/4 - p^2 \max_{r\neq 0} |\widehat{g}(r)| \left(3\delta\|g\|_2 \cdot \|1_{[n]}\|_2 + 3\delta^2 \|1_{[n]}\|_2^2 + \|g\|_2^2\right) \\ &\geq \delta^3 n^2/4 - 8\delta p^2 \max_{r\neq 0} |\widehat{g}(r)|\,,\end{aligned}$$

which follows by noting that $\|g\|_2^2 \leq \|g\|_1 \leq 2\delta$.

Recalling that $\Lambda_3(g+\delta 1_{[n]}) = \Lambda_3(1_A) = \delta n$ and rearranging, we get

$$8\delta p^2 \max_{r\neq 0} |\widehat{g}(r)| \geq \delta^3 n^2/4 - \delta n \geq \delta^3 n^2/8,$$

which implies $\max_{r\neq 0} |\widehat{g}(r)| \geq C\delta^2$. □

Density increment. We will now find a progression of the form

$$P_y = \{y, y+\beta, \ldots, y+(k-1)\beta\},$$

for some y, β and $k \in \mathbb{Z}_p$, restricted to which, A is denser. Assuming $P_y \subseteq [n]$, we have

$$\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| = \left|\frac{|A\cap P_y| - \delta|P_y \cap [n]|}{|P_y|}\right| = \left|\frac{|A\cap P_y|}{|P_y|} - \delta\right|,$$

which shows that $\left|\mathbf{E}_{z\in P_y}[g(z)]\right|$ is a measure of how much the density $|A \cap P_y|/|P_y|$ differs from δ. In order to reason about $\left|\mathbf{E}_{z\in P_y}[g(z)]\right|$ and choose an appropriate progression P_y we will use the concept of convolution. For two functions $f_1, f_2 : \mathbb{Z}_p \to \mathbb{C}$, the convolution $f_1 * f_2$ is a function defined as

$$f_1 * f_2(y) = \mathop{\mathbf{E}}_{z}[f_1(z) f_2(y-z)].$$

For a subset S, if we set f_2 to be its characteristic function, 1_S, then

$$f_1 * 1_S(y) = \frac{|S|}{p} \mathop{\mathbf{E}}_{z\in y-S}[f_1(z)],$$

where $x - S$ is simply the reflected set $-S$, shifted by y. Therefore, for any y, the quantity $\left|\mathbf{E}_{z\in P_y}[g(z)]\right|$ is proportional to $|g * 1_{-P_0}(y)|$. The convolution $f_1 * f_2$ additional enjoys the property that its Fourier coefficients are simply products of those of f_1 and f_2. In particular, we have

$$\left|\widehat{g * 1_{-P_0}}(r)\right| = |\widehat{g}(r)| \cdot \left|\widehat{1_{-P_0}}(r)\right| = |\widehat{g}(r)| \cdot \left|\widehat{1_{P_0}}(r)\right|.$$

From the previous section, we know that $|\widehat{g}(r)| = \Omega(\delta^2)$. Now, we will choose the parameters β and k to ensure that $\left|\widehat{1_{P_0}}(r)\right|$ is also large. This will imply that the convolution has a large Fourier coefficient. A lower bound on the Fourier coefficient in turn implies that $g * 1_{-P_0}(y)$ cannot be small everywhere, which will allow us to find a progression P_y where the density increases.

Claim 4.

$$\left|\widehat{1_{P_0}}(r)\right| \geq \frac{k}{p}\left(1 - 2\pi k \left\|\frac{r\beta}{p}\right\|\right),$$

where $\|x\|$ denotes the distance of x to the closest integer.

Proof. We have

$$\begin{aligned}
\left|\widehat{1_{P_0}}(r)\right| &= \frac{k}{p}\Big|\mathop{\mathbf{E}}_{z\in P_0}[\omega^{-rz}]\Big| \\
&\geq \frac{k}{p}\left(1 - \Big|\mathop{\mathbf{E}}_{z\in P_0}[1-\omega^{-rz}]\Big|\right) \\
&\geq \frac{k}{p}\left(1 - \max_{\ell\in[k]}\left|1-\omega^{\ell\beta r}\right|\right) \geq \frac{k}{p}\left(1 - k\left|1-\omega^{\beta r}\right|\right).
\end{aligned}$$

Moreover, $\left|1-\omega^{\beta r}\right| = 2\left|\sin(\pi\beta r/p)\right| = 2\left|\sin(\pi\|r\beta/p\|)\right| \leq 2\pi\|r\beta/p\|$. ☐

In order to use the previous claim, we must impose certain constraints on the common difference β to control the error in the above approximation while simultaneously ensuring that P_0 is a legitimate progression in $[n]$. Choosing $\beta = r^{-1}$ would ensure that the error in the approximation is sufficiently small, however, r^{-1} could be quite large, which might result in several points in P_0 wrapping around. If instead, we set $\beta = 1$, then it easy to ensure P_0 is a valid progression but the error could be quite large. It turns out that there is a choice of β that is not too large and also ensures a small error in the approximation.

Observation 5. *Let $t < p$. There exists $\beta \in [t]$, such that $r\beta \bmod p \leq p/t$.*

Proof. Indeed, if we partition $\{0,\ldots,p-1\}$ into t consecutive disjoint intervals, each of length at most p/t, it follows by the pigeonhole principle that for some two distinct values $a > b \in \{1,\ldots,t\}$, both $ar \bmod p$ and $br \bmod p$ land in the same interval. The claim follows by setting $\beta = a - b$. ☐

For a parameter $t \in \mathbb{N}$ that we shall set later, let β be as promised by the previous observation. Then, the common difference in P_0 is bounded by t. If in addition, we set the length k so that $kt \ll n$, then P_y is a legitimate progression in $[n]$ as long as $y \in [n-kt]$. Let us denote $E_1 := \{n+1,\ldots,p-kt\}$ and $E_2 := \{0,\ldots,p-1\} \setminus ([n-kt] \cup E_1)$. Note that

$$|g * 1_{-P_0}(y)| = \frac{k}{p}\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| = \begin{cases} 0, & \text{if } y \in E_1 \\ k/p, & \text{if } y \in E_2, \end{cases}$$

where the first case follows by observing that $g(z) = 0$ for all $z \in P_y$ since $P_y \cap [n] = \emptyset$, and the second case uses the fact that g is bounded in $[-1,1]$.

We are now ready a progression P_y where the density increases. We have,

$$\begin{aligned}\left|\widehat{g * 1_{-P_0}}(r)\right| &= \left|\widehat{g * 1_{-P_0}}(r)\right| + \widehat{g * 1_{-P_0}}(0)\\ &\leq \mathop{\mathbf{E}}_{y}[\left|g * 1_{-P_0}(y)\omega^{-ry}\right| + g * 1_{-P_0}(y)]\\ &\leq \frac{k}{p}\left(\max_{y\in[n-kt]}\left(\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| + \mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right) + \frac{2|E_2|}{p}\right).\end{aligned}$$

In the first equality, above we use the fact that $\widehat{g * 1_{-P_0}}(0) = 0$ since $\widehat{g}(0) = 0$.

Using the fact that $|\widehat{g}(r)| \geq \Omega(\delta^2)$ and Claim 4, we have

$$\begin{aligned}&\max_{y\in[n-kt]}\left(\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| + \mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right) + \frac{2|E_2|}{p} \geq C\delta^2\left(1 - \frac{2\pi k}{t}\right)\\ &\implies \max_{y\in[n-kt]}\left(\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| + \mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right) \geq C\delta^2\left(1 - \frac{2\pi k}{t}\right) - \frac{4kt}{p}.\end{aligned}$$

Setting $kt = Cp\delta^2/16$ and $k/t = 1/4\pi$, we see that

$$\max_{y\in[n-kt]}\left(\left|\mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right| + \mathop{\mathbf{E}}_{z\in P_y}[g(z)]\right) \geq C\delta^2/4,$$

which implies that $\mathbf{E}_{z\in P_y}[g(z)] \geq C\delta^2/8$, since $\left|\mathbf{E}_{z\in P_y}[g(z)]\right| + \mathbf{E}_{z\in P_y}[g(z)] \geq 0$. By our choice parameters β, k and y, we have a progression $P_y \subseteq [n]$ of length at least $C'\delta\sqrt{p}$, where $|A \cap P_y| \geq |P_y|(\delta + C'\delta^2)$, for an absolute constant $C' > 0$.

Calculations for the density constraint. Starting with a set $A \subseteq [n]$ of density δ, that does not contain 3-term APs, we can repeatedly restrict A to a progression and increase the density so long as $\delta^2 \geq 8/n$. In every iteration, the size of the ambient set shrinks to $C'\sqrt{n}\delta$ while the density increases by additive $C'\delta^2$, for some absolute constant C'. This process must terminate after $m := 8C'/\delta + \log(1/C'\delta)$ steps because then the density would exceed one. However, even after these number of steps the size of the ambient set is at least $n^{2^{-m}}(\delta^2 C')^m > 8/\delta^2$, if $\delta = \Omega(1/\log\log n)$.

2. Improved Density Estimate

Again let $A \subseteq [n]$ of density δ. In this section, we shall think of A being a subset of $\mathbb{Z}_n$ with addition being mod n. Throughout this section, we will denote $J := \{-n/12 + 1, \ldots, n/12\}$.

Claim 6. *If* $\left|A \cap \left(\frac{n}{2} + J\right)\right| \geq \frac{\delta|J|}{2}$ *and* $\delta^2 > \frac{1}{1200n}$ *then*

$$\sum_{r\neq 0}\left|\widehat{1_A}(r)\right|^3 \geq 10^{-6}\delta^3. \tag{2}$$

Note that if $\left|A \cap \left(\frac{n}{2}+J\right)\right| < \frac{\delta|J|}{2}$ then

$$\left|A \cap \left[\frac{5n}{12}\right]\right| + \left|A \cap \left(\frac{7n}{12} + \left[\frac{5n}{12}\right]\right)\right| > \delta n - \frac{\delta n}{12} = \frac{11\delta n}{12} = \frac{11\delta n}{10} \cdot \frac{10n}{12}.$$

This implies that the density of A restricted to a shift of $[5n/12]$ increases by $\delta/10$.

Proof. (of Claim 6) Denote $u = 1_A \cdot 1_{n/2+2J}$. We claim that

$$\Lambda(1_A, u, u) = \sum_{x,y} 1_A(x-y)u(x)u(x+y) \leq |A|.$$

Indeed, if $u(x) = u(x+y) = 1$ then both $x, x+y \in \frac{n}{2} + 2J = \{\frac{n}{3}+2, \dots, \frac{2n}{3}\}$. This implies that $y \in \{-\frac{n}{3}+2, \dots, \frac{n}{3}-2\} = -2+4J$. Since this set is symmetric, it contains $-y$, which implies

$$x - y \in \left(\frac{n}{2} - 2\right) + 6J = \left(\frac{n}{2} - 2\right) + \left\{-\frac{n}{2} + 6, \dots, \frac{n}{2}\right\} = \{4, \dots, n-2\}.$$

Then, $x-y, x$ and $x+y$ form a 3-term AP in $[n]$, which is impossible, unless $y = 0$, in which case, the above sum is at most $|A|$.

Now, we will lower bound $\Lambda(1_A, u, u)$. Denote $v = 1_A \cdot (1_{\frac{n}{2}+J} * 1_{\frac{n}{2}+J})$. Since for every x, $u(x) \geq v(x)$, it follows that $\Lambda(1_A, v, v) \leq \Lambda(1_A, u, u) \leq \delta n$. Applying Equation (1) for $\Lambda(1_A, v, v)$, we have

$$\delta n \geq n^2 \sum_r \widehat{1_A}(r) \cdot \widehat{v}(r) \cdot \widehat{v}(-2r) \geq n^2 \left(\frac{\delta^3}{600} - \sum_{r \neq 0} \left|\widehat{1_A}(r)\right| \cdot |\widehat{v}(r)| \cdot |\widehat{v}(-2r)|\right),$$

where in the last step, we used the fact that $|\widehat{v}(0)| \geq \delta/24$. This follows by observing that $v(x) \geq \frac{1_A(x)}{2}$, for every $x \in \frac{n}{2} + J$, and $|A \cap (\frac{n}{2} + J)| \geq \frac{\delta n}{12}$. Rearranging, we get

$$\sum_{r \neq 0} \left|\widehat{1_A}(r)\right| \cdot |\widehat{v}(r)| \cdot |\widehat{v}(-2r)| \geq \frac{\delta^3}{600} - \frac{\delta}{n} > \frac{\delta^3}{1200}.$$

We can upper bound the above sum as follows

$$\begin{aligned}
\sum_{r \neq 0} \left|\widehat{1_A}(r)\right| \cdot |\widehat{v}(r)| \cdot |\widehat{v}(-2r)| &\leq \sqrt{\sum_{r \neq 0} \left|\widehat{1_A}(r)\right| \cdot |\widehat{v}(r)|^2} \cdot \sqrt{\sum_{r \neq 0} \left|\widehat{1_A}(r)\right| \cdot |\widehat{v}(-2r)|^2} \\
&\leq \left(\sum_{r \neq 0} \left|\widehat{1_A}(r)\right|^3\right)^{1/3} \cdot \left(\sum_r |\widehat{v}(r)|^3\right)^{2/3},
\end{aligned}$$

where we use Cauchy-Schwarz inequality in the first step and Hölder's inequality in the second. We can bound the norm of the Fourier spectrum of the function

$f * g$ in terms of appropriate norms of the Fourier spectrum of f and g. We have,

$$\|\widehat{f * g}\|_p \leq \|\widehat{f}\|_p \cdot \|\widehat{g}\|_1. \tag{3}$$

Using this, we have

$$\begin{aligned}
\|\widehat{v}\|_3 &\leq \|\widehat{1_A}\|_3 \cdot \|\widehat{1_{\frac{n}{2}+J} * 1_{\frac{n}{2}+J}}\|_1 \\
&= \|\widehat{1_A}\|_3 \cdot \left(\sum_r \left|\widehat{1_{\frac{n}{2}+J}}(r)\right|^2\right) = \|\widehat{1_A}\|_3 \cdot \|1_{\frac{n}{2}+J}\|_2^2 = \|\widehat{1_A}\|_3/6.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
\frac{\delta^3}{1200} &\leq \left(\sum_{r\neq 0} \left|\widehat{1_A}(r)\right|^3\right)^{1/3} \cdot \frac{\|\widehat{1_A}\|_3^2}{36} \\
&\leq \frac{1}{36}\left(\sum_{r\neq 0} \left|\widehat{1_A}(r)\right|^3\right)^{1/3} \cdot \left(\delta^3 + \sum_{r\neq 0} \left|\widehat{1_A}(r)\right|^3\right)^{2/3},
\end{aligned}$$

which implies that $\sum_{r\neq 0} \left|\widehat{1_A}(r)\right|^3 \geq 10^{-6}\delta^3$. □

Density Increment. In the previous section, we found a progression of length $\approx \delta\sqrt{n}$, restricted to which, the density of A increased by $\approx \delta^2$. In order to obtain this we used the fact that some non trivial Fourier coefficient of $1_A - \delta 1_{[n]}$ had magnitude $\approx \delta^2$.

We will follow the same density increment template. First, let us order the elements in $\mathbb{Z}_{[n]} \setminus \{0\}$ as $r_1, \ldots, r_{n-1}$ so that $\left|\widehat{1_A}(r_1)\right| \geq \ldots \geq \left|\widehat{1_A}(r_{n-1})\right|$. Our starting point is Equation (2)[1], and using this, we will show the following.

1. Assuming A is sufficiently ($\delta \geq (\log n)^{-O(1)}$), there exists $q \leq (\log n)^{O(1)}$ such that $\sum_{i=1}^q \left|\widehat{1_A}(r_i)\right|^2 = \Omega(q^{1/3}\delta^2)$, and
2. there is a progression $P_y = \{y, y+\beta, \ldots, y+(k-1)\beta\}$ such that $k\beta < n$, $|P_y| = \Omega(n^{1/(q+1)})$, and $\mathbf{E}_{z\in P_y}[1_A(z)]^2 = \delta^2 + \Omega\left(\sum_{i=1}^q \left|\widehat{1_A}(r_i)\right|^2\right)$.

The assumption that $k\beta < n$ is useful because it implies, P_y can be partitioned into two sub-progressions that are legitimate progressions in $[n]$.

Claim 7. *For any $\varepsilon > 0$, assuming Equation (2), there exists $q \leq \delta^{-3}$ and a constant C (depending on ε) such that*

$$\sum_{i=1}^q \left|\widehat{1_A}(r_i)\right|^2 \geq C\delta^2 q^{1/3-2\varepsilon} \tag{4}$$

[1]Note that this equation already implies that some non-trivial Fourier coefficient of 1_A has magnitude at least $\approx \delta^2$.

Proof. Denote $t = \delta^{-3}$. We shall prove the stronger statement that for some $q \le t$ it holds that $\left|\widehat{1_A(r_q)}\right| \ge C\delta q^{-(1/3+\varepsilon)}$. Indeed, this implies,

$$\sum_{i=1}^{q}\left|\widehat{1_A(r_i)}\right|^2 \ge q\left|\widehat{1_A(r_q)}\right|^2 \ge C^2\delta^2 q^{1/3-2\varepsilon}.$$

Therefore, assume to a contradiction that $\left|\widehat{1_A(r_q)}\right| < C\delta q^{-(1/3+\varepsilon)}$ for all $q \le t$. This immediately implies,

$$\sum_{i>t}\left|\widehat{1_A(r_i)}\right|^3 \le \delta\cdot\left|\widehat{1_A(r_t)}\right| \le C\delta^2 t^{-1/3} = C\delta^3.$$

Moreover,

$$\begin{aligned}\sum_{i\ge 1}\left|\widehat{1_A(r_i)}\right|^3 &\le C^3\delta^3\sum_{q=1}^{t} q^{-(1+3\varepsilon)} + C\delta^3\\ &\le C^3\delta^3\left(1+\int_2^q (x-1)^{-(1+3\varepsilon)}\right) + C\delta^3 < C\delta^3\left(2+\frac{1}{3\varepsilon}\right),\end{aligned}$$

which contradicts Claim 6, if $C = \varepsilon 10^{-6}$. □

Claim 8. *For every $q > 0$, there exists a progression $P = \{0, \beta, \ldots, (k-1)\beta\}$ and $y \in \mathbb{Z}_n$ such that $k\beta < n$ and $k \ge (n^{1/(q+1)})/4\pi$, and*

$$\mathop{\mathbf{E}}_{z\in y+P}[1_A(z)]^2 = \delta^2 + \frac{1}{4}\sum_{i=1}^{q}\left|\widehat{1_A}(r_i)\right|^2.$$

Proof. First, we choose a common difference β so that for the progression $P = \{0, \beta, \ldots, (k-1)\beta\}$, we have

$$\left|\widehat{1_P}(r_i)\right| \ge \frac{k}{2n}, \text{ for every } i \in [q].$$

For a parameter $t > 0$, let us partition $\{0, \ldots, n-1\}$ into $\ell = t^{1/q}$ disjoint consecutive intervals, $I_1, \ldots, I_\ell$. These sets induce a partition of the product set $\{0, \ldots, n-1\}^q$ into cells, where each cell is a Cartesian product of q intervals. For each $j \in \{1, \ldots, t\}$, we can identify the tuple $(jr_1, \ldots, jr_q)$ by the cell that contains it. Since the number of cells is $\ell^q \le t$, the pigeonhole principle implies that there must be distinct j, j' such that the corresponding tuples are contained in the same cell. Setting $\beta = j - j'$, we see that $\beta r_i \in \{-n/\ell, \ldots, n/\ell\}$. Now, we have

$$\begin{aligned}\left|\widehat{1_P(r_i)}\right| &= \frac{k}{n}\left|\mathop{\mathbf{E}}_{z\in P}[\omega^{-zr_i}]\right|\\ &\ge \frac{k}{n}\left(1-\left|\mathop{\mathbf{E}}_{z\in P}[1-\omega^{-zr_i}]\right|\right)\\ &\ge \frac{k}{n}\left(1-\max_{j\in[k]}\left|1-\omega^{-r_ik\beta}\right|\right) \ge \frac{k}{n}\left(1-k\left|1-\omega^{-r_i\beta}\right|\right)\end{aligned}$$

Using the fact that

$$\left|1-\omega^{-r_i\beta}\right| = 2\sin\left(\frac{\pi r_i\beta}{n}\right) \leq 2\pi\left\|\frac{r_i\beta}{n}\right\| \leq \frac{2\pi}{\ell},$$

and setting $k=\ell/4\pi$, we have

$$\left|\widehat{1_P(r_i)}\right| \geq \frac{k}{n}\left(1-\frac{2\pi k}{\ell}\right) \geq \frac{k}{2n}.$$

Note that $k\beta \leq \ell t/4\pi \leq t^{(q+1)/q}/4\pi$. Setting $t=n^{q/(q+1)}$, we can ensure that $k\beta < n$. This gives us a progression P of length $k=\ell/4\pi = n^{1/(q+1)}/4\pi$, as desired. It remains to find a shift of P where the density of A increases. Since $|1_A * 1_{-P}(y)| = (k/n)\,\mathbf{E}_{z\in y+P}[1_A(z)]$, it suffices to lower bound $\|1_A * 1_{-P}\|_\infty^2$. We have

$$\begin{aligned}
\|1_A * 1_{-P}\|_\infty^2 &\geq \|1_A * 1_{-P}\|_2^2 \\
&= \sum_r \left|\widehat{1_A}(r)\right|^2 \left|\widehat{1_{-P}}(r)\right|^2 \\
&\geq \delta^2\left|\widehat{1_P}(0)\right|^2 + \sum_{i=1}^{q}\left|\widehat{1_A}(r)\right|^2\left|\widehat{1_P}(r)\right|^2 \\
&= \left(\frac{k}{n}\right)^2\left(\delta^2 + \frac{1}{4}\sum_{i=1}^{q}\left|\widehat{1_A}(r)\right|^2\right),
\end{aligned}$$

which implies the claim. $\square$