MTH392A Report
Arithmetic Progressions in Sets of Integers

Anurag Sahay
11141

April 11, 2014

Abstract

We briefly survey the field exploring the connections between the "size" of a set of integers with the occurrence of arithmetic progressions in that set. We then examine a proof of Roth's theorem, given in [3]

Supervised by:

Prof. Shobha Madan
Dept. of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Contents

1 Introduction

Let $\mathbb{Z}^+$ denote the set of natural numbers. Let $A \subset \mathbb{Z}^+$ be a subset of positive integers. Further, suppose

$$[x, y] = \{a \in \mathbb{Z} : x \leq a \leq y\}$$

Then the ratio of the size of $A \cap [1, n]$ with n is roughly a measure of how "dense" the set is in the positive integers. In particular for every A, define

$$\delta(A) = \limsup_n \frac{|A \cap [1, n]|}{n}$$

to be the "upper asymptotic density" of the set A. Clearly, $\delta(A) \geq 0$. We will refer to this as the density of the set for the time being.

Suppose that A is an infinite length arithmetic progression. That is, there exist $a, d \in \mathbb{Z}^+$ such that

$$A = \{a + dx : x \in \mathbb{Z}^+\}$$

It is easy to verify that $\delta(A) = 1/d > 0$[1].

[1] To see this, note that we can add finitely many integers to any set without changing its density. Further, note that if we extend the arithmetic progression to the left until it either reaches 0, then the augmented set we obtain, A' will satisfy $|A' \cap [1, an]| = n$,

The goal of this reading project was to explore the moral converse of this fact: that is, given that $\delta(A) > 0$, what can we say about the presence of arithmetic progressions in A? In particular, this density is a rough measure of how packed the set A is among the integers, or, to take it another way, how large the set A is. The expectation is that if the set A is too large, then it should contain at least one arithmetic progression of a given length k.

In 1936, Erdős and Turán formulated this as a precise statement: they conjectured that if A is a set of positive upper density (ie, $\delta(A) > 0$), then for any $k \in \mathbb{Z}^+$, there exists a k-length arithmetic progression in A. Since removing this progression from A will not change the density of A (as this is a k-size, and hence finite subset of A), this in fact implies the existence of infinitely many non-trivial arithmetic progressions of arbitrary length in A. This statement (known erstwhile as the Erdő-Turán conjecture) was proven by Endre Szemerédi in 1975, and now goes by the name Szemerédi's theorem.

It is easy to see that the cases $k = 1$ and $k = 2$ are trivially true given the hypothesis of positive upper asymptotic density. The first non-trivial case ($k = 3$), however, was first established by Klaus Roth through fourier analytic methods in 1953. In this project I first tried to understand this proof of Roth's theorem, with the eventual hope to move on to $k = 4$, and then perhaps to the general case. In this report, we will document one proof of Roth's theorem, as provided in [3], due to the insight it provides into the nature of the original proof. During the course of this project, I also referred to Roth's original paper [1] as well as an expository note by Alex Iosevich [2], and presented this version of the proof to my supervisor. I later read the proof from [3] (an exposition by Neil Lyall), which is what is presented in this report in a little expanded format. Lyall's exposition handles the big picture much better than the other two, in terms of explaining exactly what is going on. We assume basic familiarity with the discrete Fourier transform in the report.

since there is exactly one element from A' in every consecutive sequence of a integers. Furthermore, it is quite easy to see that choosing $[1, N]$ with a different N than a multiple of a will result in the quotient of $|A' \cap [1, N]|$ with N being slightly less than $1/a$. Hence, it is easily seen that

$$\delta(A) = \limsup_{N} \frac{|A' \cap [1, N]|}{N} = \frac{1}{a}$$

1.1 Roth's theorem

We first state the version of Roth's theorem that was originally proved.

Theorem 1.1 (Roth's Original Theorem). *Let $A \subset \mathbb{Z}^+$ be a set such that*

$$\delta(A) = \limsup_n \frac{|A \cap [1,n]|}{n} > 0$$

then A contains at least one arithmetic progression of length 3.

The theorem we shall prove will be the following:

Theorem 1.2 (Roth's Theorem; finitary). *Suppose $1 \geq \delta > 0$. Then there exists a sufficiently large integer N_0 such that for all $N \geq N_0$, if $A \subset [0, N-1]$ and $|A| = \delta N$, then A necessarily contains a non-trivial arithmetic progression.*

To see that the finitary version of Roth's theorem implies Roth's original version, pick any set $A \subset \mathbb{Z}^+$ of positive density. Fix a sufficiently small ϵ such that $\delta = \delta(A) - \epsilon$ is still positive. Thus, for sufficiently large N_0, a subset A' of $[0, N_0 - 1]$ of size δN_0 shall contain an arithmetic progression. Now, since the density of A is $\delta(A)$, for every N_0 and every ϵ, there exists an $N \geq N_0$ such that

$$\frac{|A \cap [0, N-1]|}{N} \geq \delta(A) - \epsilon = \delta > 0$$

Hence, setting $A' = A \cap [0, N-1]$, we see that A' (and hence, A) must have an arithmetic progression.

1.2 Fourier transform

In this section, we recall some properties of the Fourier transform over $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$, the cyclic group of order N. For any function $f : \mathbb{Z}_N \to \mathbb{C}$, we define its Fourier transform as

$$\hat{f}(k) = \sum_{[x] \in \mathbb{Z}_N} f(x) e^{-\frac{2\pi i x k}{N}} = \sum_{x=0}^{N-1} f(x) e^{-\frac{2\pi i x k}{N}}$$

where the last equality holds because the summand does not depend on which representative of $[x] \in \mathbb{Z}_N$ we choose. Then, clearly,

$$\hat{f}(0) = \sum_x f(x)$$

and, by the triangle inequality,

$$\hat{f}(k) \leq \sum_x |f(x)|$$

Now, let

$$f * g(x) = \sum_y f(y)\overline{g(y-x)}$$

then we know that

$$f \hat{*} g(k) = \hat{f}(k)\overline{\hat{g}(k)}$$

Hence we have

$$|\hat{f}(k)||\hat{g}(k)| \leq \sum_x \left| \sum_y f(y)\overline{g(y-x)} \right|$$

It is easy to see also, the orthogonality relation

Theorem 1.3 (Orthogonality relation)**.**

$$\frac{1}{N}\sum_{k=0}^{N-1} e^{\frac{2\pi i}{N}xk} = \begin{cases} 1 & x \equiv 0 \pmod{N} \\ 0 & \textit{otherwise} \end{cases}$$

Proof. To see this, note that for $x \equiv 0 \pmod{N}$, the summand is 1, hence the sum will totally be $|\mathbb{Z}_N| = N$. For any other case, we see that it is the sum of a geometric progression, and hence equal to

$$\frac{(e^{\frac{2\pi i}{N}xk})^N - 1}{e^{\frac{2\pi i}{N}xk} - 1}$$

Clearly, the numerator is 0. □

With this, we can prove Plancherel's identity in this setting,

Theorem 1.4 (Plancherel's identity). *For any function $f : \mathbb{Z}_N \to \mathbb{C}$,*

$$\sum_x |f(x)|^2 = \frac{1}{N} \sum_k |\hat{f}(k)|^2$$

Proof. Note that $|\hat{f}(k)|^2 = \hat{f}(k)\overline{\hat{f}(k)}$. Substituting the definition of $\hat{f}(k)$

$$|\hat{f}(k)|^2 = \left(\sum_y f(y) e^{-\frac{2\pi i y k}{N}} \right) \left(\sum_z \overline{f(z)} e^{\frac{2\pi i z k}{N}} \right)$$

Hence,

$$|\hat{f}(k)|^2 = \sum_{y,z} f(y)\overline{f(z)} e^{\frac{2\pi i (z-y)k}{N}}$$

Thus, summing it over all k,

$$\sum_k |\hat{f}(k)|^2 = \sum_k \sum_{y,z} f(y)\overline{f(z)} e^{\frac{2\pi i (z-y)k}{N}}$$

We now interchange the sums, since all sums are finite to get

$$\sum_k |\hat{f}(k)|^2 = \sum_{y,z} f(y)\overline{f(z)} \left(\sum_k e^{\frac{2\pi i (z-y)k}{N}} \right)$$

By the orthogonality relation, the inner sum will vanish except when $z \equiv y \pmod{N}$, which in our domain is the same as asking $y = z$. In this case, the inner sum will be N.

Hence,

$$\frac{1}{N} \sum_k |\hat{f}(k)|^2 = \sum_y f(y)\overline{f(y)} = \sum_x |f(x)|^2$$

□

With this we can now begin our proof of Roth's theorem.

2 Outline: Randomness and Structure

Problems such as Roth's theorem are often considered in a relatively modern field called additive combinatorics, which asks questions like these about sets having some sort of additive structure. A general theme that occurs in problems from this field is of the dichotomy between "structure" and "randomness". In each problem, the solution lies in identifying what a suitable notion of "randomness" is for the problem at hand. Then, the problem is divided into two cases - one where the set we are considering is random in the above sense, and the other where it is more structured. Different methods are used to solve both cases, and then a hybrid method is used to solve the general case which is neither too random nor too structured.

First note the concept of "being in arithmetic progression" is invariant under affine transformations. That is, let A be a set of integers and for $x, y \in \mathbb{Z}^+$, let

$$x + y \cdot A = \{x + ya : a \in A\}$$

Then A is clearly an arithmetic progression if and only $x + y \cdot A$ is an arithmetic progression. In particular, rather than considering subsets of $[0, N-1]$, we can consider any subset of an arbitrary progression P of length N. If P has the property that $A \subset P$, $|A| = \delta|P| = \delta N$ implies that A has an arithmetic progression, then $[0, N-1]$ has the same property (by using the affine transformation $t \mapsto a + dt$, where a and d are the first term and common difference of P).

We can now given an overview of how the randomness-structure ideology will work out in this case. We will define a notion of randomness (which we call ϵ-randomness), such that if A is "random", we will immediately have many arithmetic progressions of length 3. If A is not "random", we will show that there exists a substructure of $[0, N-1]$ (that is, a subprogression of $[0, N-1]$ such that A is concentrated higher on that subprogression that in the larger progression. In other words, for some $\epsilon_0 > 0$ depending only on δ, we will show that there exists a progression $P \subset [0, N-1]$ such that A has density at least $\delta + \epsilon_0$ on P. That is,

$$|A \cap P| \geq (\delta + \epsilon_0)|P|$$

We then replace A by A_1 and replace $[0, N-1]$ with $[0, N_1 - 1]$, where $[0, N_1 - 1]$ is the progression obtained by applying an affine transformation to P with the goal of making the first term of the progression 0, and A_1 is the corresponding set obtained by applying the same affine transformation on $A \cap P$. We then examine A_1 in the above sense of "randomness". If it isn't random, we do the same thing to obtain an A_2 and N_2, and a new progression P_2 and keep iterating this until we get

$$|A_k \cap P_k| \geq (\delta + M(k)\epsilon_0)|P_k| \geq |P_k|$$

That is, the density of A_k on P_k exceeds 1. This is clearly absurd, and hence we must conclude the theorem. To be sure that this argument works, we will explicitly need to show that for the k for which $\delta + M(k)\epsilon_0$ exceeds 1, the progression P_k is not empty.

2.1 Progressions in $\mathbb{Z}$ and $\mathbb{Z}_N$

One of the advantages of considering only 3-progressions is that there is a simple algebraic formula defining them. We know that x, y, z are in arithmetic progression if and only if

$$x + y = 2z$$

Thus, instead of explicitly searching for 3-progressions in A, we can restrict ourselves to finding solutions to the equation $x + y - 2z = 0$. This solution will have at least $|A| = \delta N$ solutions given by $x = y = z$. These are the trivial APs.

The notion of arithmetic progression is not particular to the ring $\mathbb{Z}$ and can easily be generalized to $\mathbb{Z}_N$ as well. We say that a set of integers x, y, z is a $\mathbb{Z}_N$-progression if

$$x + y \equiv 2z \pmod{N}$$

Using the orthogonality relation, we have very nice way through which we can count the $\mathbb{Z}_N$-progressions in any set A. Replacing x by $x+y-2z$, we see that

$$\sum_{k=0}^{N-1} e^{\frac{2\pi i k}{N}(x+y-2z)} = \begin{cases} N & x+y \equiv 2z \pmod{N} \\ 0 & \text{otherwise} \end{cases}$$

Thus,

$$\frac{1}{N} \sum_{x,y,z\in A} \sum_{k=0}^{N-1} e^{\frac{2\pi i k}{N}(x+y-2z)} = \#\{(x,y,z) \in A^3 : x+y \equiv 2z \pmod{N}\}$$

We let $\mathfrak{N}_0$ denote the number of solutions in A to the given congruence. Then, rearranging the sums and separating out the parts that depend on x, y and z respectively,

$$\mathfrak{N}_0 = \sum_{k=0}^{N-1} \left(\sum_{x\in A} e^{-\frac{2\pi i}{N}x}\right) \left(\sum_{y\in A} e^{-\frac{2\pi i}{N}y}\right) \left(\sum_{z\in A} e^{-\frac{2\pi i}{N}(-2z)}\right)$$

Now, if $1_A(x)$ is the indicator function of A, that is

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$$

Then, using the definition of Fourier transform, it is easily obtained that

$$\mathfrak{N}_0 = \frac{1}{N} \sum_{k=0}^{N-1} \widehat{1_A}(k)^2 \widehat{1_A}(-2k)$$

This idea of using the orthogonality relation to count progressions is central to the entire proof.

Now, we know that $\widehat{1_A}(0) = \sum_x 1_A(x) = |A| = \delta N$. Hence, separating the $k=0$ term from the sum, we see that

$$\mathfrak{N}_0 = \frac{\widehat{1_A}(0)^3}{N} + \frac{1}{N}\sum_{k=1}^{N-1} \widehat{1_A}(k)^2\widehat{1_A}(-2k)$$

That is,

$$\mathfrak{N}_0 = \delta^3 N^2 + \frac{1}{N}\sum_{k=1}^{N-1} \widehat{1_A}(k)^2\widehat{1_A}(-2k)$$

The leading term here helps us in formulating a suitable notion of randomness. If the set A truly were random, then suppose we choose $x, z \in A$. There are clearly $|A|^2 = \delta^2 N^2$ of doing this. If A is "truly random", the probability that $x - 2z \in A$ is should be $|A|/N = \delta$. Meaning that there should, roughly speaking, be $\delta^3 N^2$ many solutions to the congruence.

In other words, the $k = 0$ term contributes what we would expect in the random case. Thus, if we bound the other sum in a particular manner, we can come up with a suitable notion of randomness. We thus, define the notion as follows:

We say the set A is ϵ-uniform or ϵ-random if, for all $k = 1, 2, \cdots, N-1$,

$$|\widehat{1_A}(k)| \leq \epsilon N$$

Now note that

$$\left|\sum_{k=1}^{N-1} \widehat{1_A}(k)^2\widehat{1_A}(-2k)\right| \leq \max_{k\neq 0} |\widehat{1_A}(-2k)| \sum_{k=0}^{N-1} |\widehat{1_A}(k)|^2$$

Here we use Plancherel's identity, and assume that A is random,

$$\left|\sum_{k=1}^{N-1} \widehat{1_A}(k)^2\widehat{1_A}(-2k)\right| \leq \epsilon N^2 \sum_{x=0}^{N-1} |1_A(x)|^2 = \epsilon\delta N^3$$

Replacing this in our earlier counting identity,

$$\mathfrak{N}_0 \geq \delta^3 N^2 - \epsilon\delta N^2$$

Thus, the definition we have made measures in some sense how close A is to being random. If ϵ is sufficiently small, we have that A is sufficiently random, and can use the identity to prove the existence of non-trivial $\mathbb{Z}_N$-progressions. We will later adapt this argument to count regular $\mathbb{Z}$-progressions instead.

2.2 Main Theorem

We will now state the main theorem of [3] and then show how this theorem implies the finitary version of Roth's theorem.
Theorem 2.1 (Iteration)**.** *Let $A \subset [0, N-1]$, $\delta > 0$, with $A = \delta N$, and $N \geq 8\delta^{-2}$. Then, either A contains a non-trivial 3-AP, or there exists a $\mathbb{Z}$-progression P with length $\geq \frac{\delta^2\sqrt{N}}{256}$, such that*

$$|A \cap P| \geq \left(\delta + \frac{\delta^2}{64}\right)|P|$$

We will prove this theorem in the rest of the report. For now, note that this is the exact estimate that was roughly outlined earlier in the randomness-structure discussion. With this, we can establish Roth's theorem.

Suppose that under the hypothesis of the theorem, we have not yet chosen an N, but we have some $A \subset [0, N-1]$ containing no non-trivial progressions. By the Iteration theorem, we have some progression (say P_1). We now apply an affine transformation on it to get the progression $[0, N_1 - 1]$. Furthermore, we identify A_1 as the set $A \cap P_1$ under this same affine transformation. We then know by our earlier theorems, and by the Iteration theorem, that $|A_1| = \delta_1 N_1$ with $N_1 \geq \frac{\delta^2\sqrt{N}}{256}$ and $\delta_1 \geq (\delta + \frac{\delta^2}{64})$. Now, since A has not non-trivial 3-APs, neither does A_1. Thus, we can iterate the argument to obtain P_k, A_k, δ_k and N_k. At $k = 64/\delta$, we will have density

$$\delta_k \geq \delta + \frac{k\delta^2}{64} = 2\delta$$

By the same argument, after another $64/2\delta$ steps, the density will double again, to 4δ. After another $64/4\delta$ steps to 8δ, and so on. Hence, after

$$\frac{64}{\delta}(1 + \frac{1}{2} + \cdots + \frac{1}{2^{l-1}}) \leq \frac{128}{\delta}$$

steps we have density $2^l\delta$. Hence, by the $128/\delta$-th step, the density will have far exceeded 1.

Now, we can easily see by induction that $N_k \geq \frac{\delta^4 N^{1/2^k}}{256^2}$ (both the base case and the induction are trivial). Thus, we need to ensure that $\frac{\delta^4 N^{1/2^k}}{256^2} \geq 1$ for $k = 128/\delta$ for the argument to work as stated. Clearly, this can be done by choosing N large enough in the beginning.

Hence, if we establish the Iteration theorem, Roth's theorem follows.

3 Proof of the Theorem

In this section, we move onto proving the main theorem of the previous section. We must first deal with a way to detect genuine progressions rather than detecting progressions modulo an integer. Let $\mathfrak{N}$ denote the number of true 3-APs in the set A. Now note that if $x, z \in A \cap [N/3, 2N/3)$ (ie, the middle third of A), then

$$N > 2z - x \geq 0$$

and hence $y \in A \cap [0, N-1]$. Thus, in this particular case, we actually detect genuine progressions rather than $\mathbb{Z}_N$-progressions. Let $M = A \cap [N/3, 2N/3)$. Then, we can change $\mathfrak{N}_0$ to $\mathfrak{N}$ if we change the subscripts to M in two cases, and replace the equality sign by a greater than sign (since not all genuine progressions may be of the form above). Thus, we get

$$\mathfrak{N} \geq \frac{1}{N} \sum_{k=0}^{N-1} \widehat{1_M}(k)\widehat{1_A}(k)\widehat{1_M}(-2k)$$

Again, separating out the $k = 0$ term, we will get a contribution of $|M|$ from the first and third term, and a contribution of δN from in between. Thus,

$$\mathfrak{N} \geq= \delta|M|^2 + \frac{1}{N} \sum_{k=1}^{N-1} \widehat{1_M}(k)\widehat{1_A}(k)\widehat{1_M}(-2k)$$

We now bound away the sum under the assumption of ϵ-randomness for some ϵ. We get

$$\sum_{k=1}^{N-1}\left|\widehat{1_M}(k)\widehat{1_A}(k)\widehat{1_M}(-2k)\right| \leq \max_{k\neq 0}|1_A(k)|\sum_{k=0}^{N-1}|\widehat{1_M}(k)\widehat{1_M}(-2k)|$$

Applying Cauchy-Schwarz and the randomness hypothesis,

$$\sum_{k=1}^{N-1}\left|\widehat{1_M}(k)\widehat{1_A}(k)\widehat{1_M}(-2k)\right| \leq \epsilon N\sum_{k}|\widehat{1_M}(k)|^2 = \epsilon N^2\sum_{x}1_M(x) = \epsilon N^2|M|$$

where we have use Plancherel's identity.

We also assume that $M \geq \frac{\delta N}{4}$. Now if the set is random enough (in particular if the set is ϵ-uniform for some $\epsilon < \delta^2/8$, we see that

$$\mathfrak{N} \geq \frac{\delta|M|^2}{2} \geq \frac{\delta^3N^2}{32}$$

The above discussions yield that A being ϵ-random for $\epsilon < \delta^2/8$, along with $|M| \geq \frac{\delta N}{4}$ yields the existence of many progressions (which easily exceed the trivial APs). Hence, if $|M| < \frac{\delta N}{4}$, we know that one of the first and last third of $[0, N)$ must contain at least $3\delta N/8$ terms of A, otherwise A cannot have δN elements. Hence, if we take P to be this set, we note that P is progression such that $|P| \geq N/3$ such that

$$|A \cap P| = \frac{3\delta N}{8} = \frac{9\delta}{8}\frac{N/3}{=}\left(\delta + \frac{\delta}{8}\right)|P|$$

Hence, on P, A has increased density by at least $\delta/8$.

We have thus established the following theorem, by taking contrapositives to the above, which is proposition 1 of [3]

Theorem 3.1 (Base Iteration)**.** *Let $A \subset [0, N-1]$ with $|A| = \delta N$. If A contains no non-trivial AP, then either*

1. *$N \leq 8\delta^{-2}$.*
2. *There exists a progression P, $|P| \geq N/3$ such that $|A \cap P| \geq (\delta + \delta/8)|P|$.*
3. *A is not ϵ-random for $\epsilon \leq \delta^2/8$.*

Clearly the first point can be avoided by choosing N large enough. We will now examine the third case - that is, we will examine the case where are set is not random enough. In this non-random case, we shall be able to prove that the largeness of the non-zero Fourier coefficients $\widehat{1_A}(k)$ can be used to conclude that there is a large progression on which the set A is strongly concentrated.

3.1 Non-random Sets

We follow [3] and consider the non-random sets case by proving a sequence of theorems that add up to give the main theorem stated and used above.

The first step is to consider what we call non-overlapping $\mathbb{Z}_N$-progressions: we say such a progression is non-overlapping if its common difference d and length L satisfy $dL < N$. We know that if we do this, then at most one multiple of N can occur between the smallest and largest term in the non-overlapping $\mathbb{Z}_N$-AP. We can thus separate this into two genuine APs: one before this multiple of N and one after this multiple of N.

We now show that if A is not ϵ-uniform, it must be heavily concentrated on a large non-overlapping $\mathbb{Z}_N$-progression. In particular, if A is not uniform, then at least one of its Fourier coefficients must be "large" in the sense of comparable to N in size. We have the following theorem

Theorem 3.2. *If, for some $r \neq 0$, $|\widehat{1_A}(r) \geq \epsilon N$, then there exists a non-overlapping $\mathbb{Z}_N$-progression B', $|B'| \geq \sqrt{N}/4$, such that*

$$|A \cap B'| \geq \left(\delta + \frac{\epsilon}{4}\right) |B'|$$

Proof. Before we move on, we note that we know that no matter what happens, one Fourier coefficient of A, the one for $k = 0$ is always δN, and hence quite large. We also know that displacing the function by a constant quantity does not change the higher Fourier coefficients by the orthogonality relation. Hence, if we want to concentrate on the large Fourier coefficient, we try to construct a function which has a small 0 coefficient (in fact, we just make it 0), by displacing the indicator function by a well-chose constant quantity. In particular, let

$$f_A(x) = 1_A(x) - \delta = \begin{cases} 1 - \delta & x \in A \\ -\delta & \text{otherwise} \end{cases}$$

It is easy to see that $\widehat{f_A}(0) = 0$ and that $\widehat{f_A}(k) = \widehat{1_A}(k)$ for all $k \neq 0$.

Now, write $B = B' + x$, for some x. Then $|B| = |B'|$ and,

$$|A \cap B'| - \delta|B| = \sum_y (1_{A\cap B'}(y) - 1_B(y))$$

Now, we know that $1_{A\cap B'}(y) = 1_A(y)1'_B(y)$. Also, $y \in B'$ if and only if $y - x \in B$, by definition. Hence,

$$|A \cap B'| - \delta|B| = \sum_y (1_A(y)1_B(y-x) - 1_B(y-x)) = \sum_y f_A(y)1_B(y-x)$$

Hence, the condition that

$$|A \cap B'| \geq (\delta + \epsilon/4)|B'|$$

is the same as

$$\sum_y f_A(y)1_B(y-x) = |A \cap B'| - \delta|B| \geq \frac{\epsilon|B|}{4}$$

Now we established earlier that

$$|\hat{f}(k)||\hat{g}(k)| \leq \sum_x \left| \sum_y f(y)\overline{g(y-x)} \right|$$

for function f and g. Thus, if the same Fourier coefficient of two functions are large, then at least one of the translates of the functions must have a large inner product with the other function. We can use this fact to establish the above theorem - the non-randomness implies that one Fourier coefficient (at $r \neq 0$) of f_A is already large. We try to construct a B such that the same Fourier coefficient of 1_B is also large.

First, we construct such a B of length at least $\sqrt{N}/4$. In particular, for $r \neq 0$, we have that we construct such a non-overlapping progression such that

$$|\widehat{1_B}(r)| \geq \frac{|B|}{2}$$

To see this, partition the square $[0, N-1]^2$ into less than $\lceil\sqrt{N}-1\rceil^2$ equal squares. Then by considering the points

$$\{(0,0), (1,r), \cdots (N-1, (N-1)r)\}$$

we get that there must exist integers l, k such that $l(1,r)$ and $k(1,r)$ are in the same square, and hence

$$(k-l) \leq \sqrt{N}$$

and

$$r(k-l) \leq \sqrt{N}$$

We let $d = k - l$, and consider the progression $\{\cdots, -2d, d, 0, d, 2d, \cdots\}$ equally on both sides around 0 until the size of the progression reaches $|B| = \lfloor\sqrt{N}/\pi\rfloor$. We then have

$$\left|\widehat{1_B}(r) - |B|\right| \leq \left|\sum_x 1_B(x)\left(e^{-\frac{2\pi i x r}{N}} - 1\right)\right|$$

Now, by definition of B, we can write $x = dl$ with $|l| \leq |B|/2$, and hence

$$\left|\widehat{1_B}(r) - |B|\right| \leq \left|\sum_{|l|\leq|B|/2} 1_B(x)\left(e^{-\frac{2\pi i d l r}{N}} - 1\right)\right| < \frac{1}{2}|B|\left(\frac{2\pi|B|\sqrt{N}}{2N}\right) \leq \frac{|B|}{2}$$

Hence, we have $\widehat{1_B}(r) \geq |B|/2$.

Now, let

$$G(x) = \sum_y f_A(y) 1_B(y-x)$$

Thus,

$$|\widehat{G}(r)| = |\widehat{f_A}(r)||\widehat{1_B}(r)| \geq \epsilon N|B|/2$$

Thus using facts about the Fourier transform shown earlier,

$$\sum_x |G(x)| \geq |\widehat{G}(r) \geq \epsilon N|B|/2$$

Now this of course does not imply that any of the $G(x)$ is large and positive, it only implies that they are large in magnitude. However, we know the $\sum_x G(x) = 0$. Thus, we can use that fact to get that

$$\sum_x |G(x)| + G(x) \geq \epsilon N|B|/2$$

Since the sum is non-negative, at least for some x we must have

$$|G(x)| + G(x) \geq \epsilon|B|/2$$

as otherwise the inequality above would not be satisfied. Now, this implies that the given $G(x)$ is positive (or else the left hand side would be 0). Hence we get

$$G(x) \geq \frac{\epsilon|B|}{4}$$

which is what we needed to show.

□

We now show how to go from non-overlapping $\mathbb{Z}_N$-progressions to genuine $\mathbb{Z}$ progressions.

Theorem 3.3. *If, for some $r \neq 0$, $|\widehat{1_A}(r) \geq \epsilon N$, then there exists a $\mathbb{Z}$-progression P, $|P| \geq \epsilon\sqrt{N}/32$, such that*

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

Proof. We note that the non-overlapping B' from the previous theorem is the union of two genuine progressions P_1 and P_2, $|P_1| \leq |P_2|$. If $|P_1| \leq \epsilon|B|/8$, then

$$|A \cap P_2| = |A \cap B'| - |A \cap P_1| \geq |A \cap B'| - |P_1|$$

Thus,

$$|A \cap P_2| \geq (\delta + \epsilon/4)|B| - \epsilon|B|/8 = (\delta + \epsilon/8)|B| \geq (\delta + \epsilon/8)|P_2|$$

If not, then both P_1 and P_2 are of length at least $\epsilon|B|/8$, and hence A must have density $\delta + \epsilon/4$ on at least one of them (as otherwise, A cannot have that density on their union - $|A \cap B'| = |A \cap P_1| + |A \cap P_2|$ is less than expected.

Hence we have the theorem.

□

Now, by the earlier theorem, we consider the case where A is not ϵ-random for any $\epsilon \leq \delta^2/8$. In particular, we can replace ϵ in the previous result with $\delta^2/8$. This establishes the following theorem:
Theorem 3.4 (Iteration)**.** *Let $A \subset [0, N-1]$, $\delta > 0$, with $A = \delta N$, and $N \geq 8\delta^{-2}$. Then, either A contains a non-trivial 3-AP, or there exists a $\mathbb{Z}$-progression P with length $\geq \frac{\delta^2\sqrt{N}}{256}$, such that*

$$|A \cap P| \geq \left(\delta + \frac{\delta^2}{64}\right)|P|$$

Proof. We note by our Base Iteration, if we choose $N \geq 8\delta^{-2}$, then either we have the second case or the third. In the second case, the same progression P obtained in that theorem suffices - it is in fact, much longer and denser than the progression required. In the third case, we know that A is not $\delta^2/8$-uniform. Hence, by the just proved theorem, there is a progression satisfying exactly the conclusion of this theorem (as we obtain by replacing ϵ by $\delta^2/8$. Hence, the theorem is established.

□

This theorem with the discussion in the previous section establishes Roth's theorem.

References

[1] K. Roth,*On certain sets of integers* (1953), J. London Math Soc. **28**

[2] A. Iosevich, *Roth's theorem on arithmetic progressions* (2003), `http://www.cs.umd.edu/~gasarch/TOPICS/vdw/notes-roth3ap.pdf`

[3] N. Lyall, *Roth's theorem: the Fourier analytic approach* `http://www.math.uga.edu/~lyall/REU/Roth.pdf`