Roth theorem via analytical approach

Martin Leshko

July 29, 2022

1 Tools

1.1 Arithmetic progressions and Fourier coefficients

We start from the trilinear functional

$$\mathrm{AP}(f,g,h) = \mathop{\mathbf{E}}_{n,d\in\mathbf{Z}_N} f(n)g(n+d)h(n+2d) \tag{1}$$

We may rewrite any function in terms of Fourier coefficients[1]

$$f(d) = \sum_{x\in\mathbf{Z}_N} \hat{f}(x)\omega^{dx} \tag{2}$$

And we substitute (2) into (1) and obtain

$$\mathrm{AP}(f,g,h) = \mathop{\mathbf{E}}_{n,d\in\mathbf{Z}_N}\left(\sum_{x\in\mathbf{Z}_N}\hat{f}(x)\omega^{nx}\cdot\sum_{x\in\mathbf{Z}_N}\hat{g}(x)\omega^{(n+d)x}\cdot\sum_{x\in\mathbf{Z}_N}\hat{h}(x)\omega^{(n+2d)x}\right)$$

We extract the sum sign

$$\mathrm{AP}(f,g,h) = \sum_{x,y,z\in\mathbf{Z}_N}\left(\hat{f}(x)\hat{g}(y)\hat{h}(z)\mathop{\mathbf{E}}_{n,d\in\mathbf{Z}_N}\omega^{nx+(n+d)y+(n+2d)z}\right)$$

And we can split the expectation into two parts

$$\mathop{\mathbf{E}}_{n,d\in\mathbf{Z}_N}\omega^{nx+(n+d)y+(n+2d)z} = \left(\mathop{\mathbf{E}}_{n\in\mathbf{Z}_N}\omega^{(x+y+z)n}\right)\left(\mathop{\mathbf{E}}_{d\in\mathbf{Z}_N}\omega^{(y+2z)d}\right)$$

[1] In fact, in the following sum we have to sum over $\hat{\mathbf{Z}}_N$ (dual group). Yet one may prove that $\hat{\mathbf{Z}}_N \cong \mathbf{Z}_N$.

Since we are summing up the powers of the primitive root we have that the expectation is equal to

$$\frac{1}{N^2}[x+y+z=0]N[y+2z=0]N = \begin{cases} 1, & (x,y,z) = (\alpha,-2\alpha,\alpha) \\ 0, & \text{otherwise} \end{cases}$$

Hence we can rewrite (1) in the following way

$$\mathrm{AP}(f,g,h) = \sum_{x\in\mathbf{Z}_N} \hat{f}(x)\hat{g}(-2x)\hat{h}(x)$$

We can extract the term at $x=0$ and obtain

$$\mathrm{AP}(f,g,h) = \hat{f}(0)\hat{g}(0)\hat{h}(0) + \sum_{\substack{x\in\mathbf{Z}_N \\ x\neq 0}} \hat{f}(x)\hat{g}(-2x)\hat{h}(x) \tag{3}$$

We bound the ‘error term’

$$\left|\sum_{x\neq 0} \hat{f}(x)\hat{g}(-2x)\hat{h}(x)\right| \overset{\infty,1}{\leq} \sup_{x\neq 0}|\hat{f}(x)| \sum_{x\neq 0} |\hat{g}(-2x)\hat{h}(x)| \overset{\frac{1}{2},\frac{1}{2}}{\leq}$$

$$\overset{\frac{1}{2},\frac{1}{2}}{\leq} \left(\sup_{x\neq 0}|\hat{f}(x)|\right)\left(\sum_{x\neq 0}|\hat{g}(x)|^2\right)^{\frac{1}{2}}\left(\sum_{x\neq 0}|\hat{h}(x)|^2\right)^{\frac{1}{2}} \leq \left(\sup_{x\neq 0}|\hat{f}(x)|\right)\|\hat{g}\|_2\|\hat{h}\|_2 \tag{4}$$

In particular, if $f = 1_A$, $g = 1_B$ and $h = 1_C$ are indicators then (4) transforms to

$$\left|\sum_{x\neq 0} \hat{1}_A(x)\hat{1}_B(-2x)\hat{1}_C(x)\right| \leq \left(\sup_{x\neq 0}|\hat{1}_A(x)|\right)\|1_B\|_2\|1_C\|_2 = \left(\sup_{x\neq 0}|\hat{1}_A(x)|\right)\sqrt{\beta\gamma} \tag{5}$$

since Plancherel’s identity and $\|1_X\|_2 = \sqrt{\mathrm{density}(X)}$.

Let us consider (3) when f,g,h are indicators. It is easy to see that $\hat{1}_X(0) = \mathrm{density}(X)$, consequently

$$\mathrm{AP}(1_A,1_B,1_C) = \alpha\beta\gamma + \sum_{\substack{x\in\mathbf{Z}_n \\ x\neq 0}} \hat{1}_A(x)\hat{1}_B(-2x)\hat{1}_C(x) \tag{6}$$

Yet not all elements from $A\times B\times C$ form AP, hence

$$\mathrm{AP}(1_A,1_B,1_C) \leq \mathop{\mathbf{E}}_{\substack{a\in A \\ b\in B \\ c\in C}} 1_A(a)1_B(b)1_C(c) = \alpha\beta\gamma.$$

Consequently, by (5) and (6) we obtain

$$\mathrm{AP}(1_A,1_B,1_C)=\alpha\beta\gamma-\left|\sum_{\substack{x\in\mathbf{Z}_n\\ x\neq 0}}\hat{1}_A(x)\hat{1}_B(-2x)\hat{1}_C(x)\right|\geq\sqrt{\beta\gamma}\left(\alpha\sqrt{\beta\gamma}-\sup_{x\neq 0}|\hat{1}_A(x)|\right). \quad (7)$$

1.2 Small Fourier coefficients

1.2.1

We denote $A_{\max}:=\sup_{x\neq 0}|\hat{1}_A(x)|$.
Since (7), it is natural to consider two cases:

1. All the coefficients are small: $A_{\max}\leq\alpha\sqrt{\beta\gamma}/2$.

2. There is a large coefficient: $A_{\max}>\alpha\sqrt{\beta\gamma}/2$.

1.2.2

We consider the first case. By (7) we obtain

$$\mathrm{AP}(1_A,1_B,1_C)\geq\sqrt{\beta\gamma}\cdot\frac{\alpha\sqrt{\beta\gamma}}{2}=\frac{\alpha\beta\gamma}{2}$$

and if we ask for $\alpha\beta\gamma/2>1/N$ (or $N>2/\alpha\beta\gamma$) we obtain

$$\mathrm{AP}(1_A,1_B,1_C)\geq\frac{\alpha\beta\gamma}{2}>\frac{1}{N} \quad (8)$$

Let us assume that there is no 3-AP in $A\times B\times C$. Then we can bound AP:

$$\mathrm{AP}(1_A,1_B,1_C)=\mathop{\mathbf{E}}_{n,d\in\mathbf{Z}_N}1_A(n)1_B(n+d)1_C(n+2d)=\frac{\#\mathrm{AP}}{N^2}\leq\frac{N}{N^2}=\frac{1}{N} \quad (9)$$

since $A\times B\times C$ contains only degenerate 3-APs and, maybe, not all of them.
Yet (8) contradicts (9), hence in the first case (and $N>2/\alpha\beta\gamma$) we have 3-AP.
Otherwise, we have $r\in\mathbf{Z}_N$ such that $|\hat{1}_A(r)|\geq\alpha\sqrt{\beta\gamma}/2$.
In other words, we have established the following result.

Proposition 1. *Let N be an odd integer and $A,B,C\subseteq\mathbf{Z}_N$ with densities α,β,γ respectively. Assume that $N\geq 2/\alpha\beta\gamma$. Then either there exists non-degenerate 3-AP $(x,x+d,x+2d)\in A\times B\times C$, or there exists non-zero r such that $|\hat{1}_A(r)|\geq\alpha\sqrt{\beta\gamma}/2$.*

1.3 Splitting into dense progressions

1.3.1

We start from the unit circle $|z| = 1$. For any positive integer r we have that for any $0 \le u < v \le r$ holds

$$|\omega^{\phi u} - \omega^{\phi v}| \le \frac{2\pi}{r}.$$

and, denoting $d = u - v$,

$$|1 - \omega^{\phi d}| \le \frac{2\pi}{r}.$$

By triangle identity,

$$|1 - \omega^{\phi t d}| \le \sum_{s=1}^{t} |\omega^{\phi d s} - \omega^{\phi d(s-1)}| \le \frac{2\pi t}{r}.$$

If we ask for t to be at most $\varepsilon r/2\pi$ we obtain that

$$\operatorname{diam} \omega^{\phi P} \le \varepsilon$$

for P be a 3-AP with common difference d.

1.3.2

We split $\{1, \ldots, N\}$:

$$\{1, \ldots, N\} \xrightarrow{\quad \bmod d \quad} \{A_0, \ldots, A_{d-1}\}.$$

Firstly, $d \le r$.

Secondly, we have that for any i it holds that $|A_i| \ge \lfloor n/d \rfloor \ge \lfloor n/r \rfloor$. Consequently, if A_i can be partitioned into pieces of size between $t/2$ and t (we can straightforwardly obtain such partition by the simple greedy algorithm) then $\lfloor n/r \rfloor \ge 2t$. Note that such 'consecutive' pieces forms AP in the original sequence $\{1, \ldots, N\}$. Well, then, we can safely ignore the integer part and obtain $n/r \ge 2t$.

On the other hand, we have that $t \le \varepsilon r/2\pi$. Combining these two bounds we obtain

$$\frac{n}{r} \ge 2t \ge \frac{\varepsilon r}{\pi}.$$

In particular,

$$\frac{n}{r} \ge \frac{\varepsilon r}{\pi} \qquad \text{hence} \qquad r \le \sqrt{\frac{\pi n}{\varepsilon}}.$$

It is easy to see that

$$t = \sqrt{\frac{\varepsilon n}{4\pi}} \qquad \text{and} \qquad r = \sqrt{\frac{\pi n}{\varepsilon}}$$

satisfies all the bounds. And the shortest possible length is $t/2 = \sqrt{\varepsilon n/8\pi}$.
In other words, we have proved that for any $\phi \in \mathbf{R}$ and $\varepsilon > 0$ there exists N sufficiently large such that $\{1,\dots,N\}$ can be partitioned into several disjoint AP $\{P_1,\dots,P_m\}$ such that for every i the following statements holds:

1. The length is big enough $|P_i| \geq \sqrt{\varepsilon n/8\pi}$ and
2. the diameter is artitrarily small $\operatorname{diam} \omega^{\phi P_i} = \max_{x,y} |\omega^{\phi P_{i,x}} - \omega^{\phi P_{i,y}}| \leq \varepsilon$.

1.4 Long and dense progression

1.4.1

First, we introduce a balanced indicator as

$$\mathfrak{I}_A := 1_A - \underbrace{\operatorname{density}(A)}_{\alpha}$$

The crucial and easy-to-establish properties of balanced indicators are

$$\mathop{\mathbf{E}}_{x \in \mathbf{Z}_N} \mathfrak{I}_A(x) = 0; \tag{10}$$

$$\forall x \neq 0, \quad \hat{\mathfrak{I}}_A(x) = \hat{1}_A(x). \tag{11}$$

Particularly, as in the subsubsection 1.2.2,

$$|\mathfrak{I}_A(r)| = |1_A(r)| \geq \frac{\alpha\sqrt{\beta\gamma}}{2} =: \mathfrak{X}.$$

1.4.2

By subsection 1.3 with $\varepsilon = \mathfrak{X}/2$ and $\phi = r/N$ we obtain that it is possible to partition $\{1,\dots,N\}$ into several arithmetical progressions $P_1,\dots,P_m$ of least length $\sqrt{\mathfrak{X} n/16\pi}$ and such that their diameters are small: $\operatorname{diam} \omega^{rP_i/N} \leq \mathfrak{X}/2$.

Our current goal is to establish a bound for $\mathcal{X}$. For instance,

$$\mathcal{X} \leq |\hat{\mathcal{I}}_A(r)| = \left|\frac{1}{N}\sum_{x\in\mathbf{Z}_N}\mathcal{I}_A(x)\omega^{-\frac{rx}{N}}\right| \leq \frac{1}{N}\sum_{i=1}^{m}\left|\sum_{x\in P_i}\mathcal{I}_A(x)\omega^{-\frac{rx}{N}}\right| =$$

$$= \frac{1}{N}\sum_{i=1}^{m}\left(\left|\sum_{x\in P_i}\mathcal{I}_A(x)\omega^{-\frac{rx_i}{N}}\right| + \left|\sum_{x\in P_i}\mathcal{I}_A(x)(\omega^{-rx/N}-\omega^{-rx_i/N})\right|\right) \overset{*}{\leq} \frac{1}{N}\sum_{i=1}^{m}\left|\sum_{x\in P_i}\mathcal{I}_A(x)\right| + \frac{\mathcal{X}}{2}. \quad (12)$$

where x_i is an element which belongs to the progression P_i. In particular, the starred transition is true because we extracted the constant from the sum:

$$\left|\sum_{x\in P_i}\mathcal{I}_A(x)\omega^{-\frac{rx_i}{N}}\right| = \left|\omega^{-\frac{rx_i}{N}}\sum_{x\in P_i}\mathcal{I}_A(x)\right| \leq \left|\omega^{-\frac{rx_i}{N}}\right|\left|\sum_{x\in P_i}\mathcal{I}_A(x)\right| = \left|\sum_{x\in P_i}\mathcal{I}_A(x)\right|.$$

It follows from (12) and $N = |P_1| + \cdots + |P_m|$ that

$$\sum_{i=1}^{m}\left|\sum_{x\in P_i}\mathcal{I}_A(x)\right| \geq \frac{\mathcal{X}}{2}\sum_{i=1}^{m}|P_i|.$$

From (10) we know that

$$\sum_{i=1}^{m}\sum_{x\in P_i}\mathcal{I}_A(x) = 0$$

We sum up the last inequality and identity to obtain

$$\sum_{i=1}^{m}\left(\left|\sum_{x\in P_i}\mathcal{I}_A(x)\right| + \sum_{x\in P_i}\mathcal{I}_A(x) - \frac{\mathcal{X}}{2}|P_i|\right) \geq 0$$

1.4.3

Therefore, there exists at least one i such that the sum in the parentheses is non-negative. It means that

$$\left|\sum_{x\in P_i}\mathcal{I}_A(x)\right| + \sum_{x\in P_i}\mathcal{I}_A(x) \geq \frac{\mathcal{X}}{2}|P_i| > 0$$

that implies that the sum is positive (otherwise absolute value + actual value will be zero)

$$\sum_{x\in P_i}\mathcal{I}_A(x) \geq \frac{\mathcal{X}|P_i|}{4}.$$

By definition,

$$\sum_{x\in P_i}\mathcal{I}_A(x) = \sum_{x ß P_i}(1_A(x)-\alpha) = |A\cap P_i| - \alpha|P_i|.$$

Consequently, we have found the long, specifically of the length at least $\sqrt{\mathcal{X}n/16\pi}$, arithmetic progression, namely P_i, such that most of its elements lie in A:

$$|A\cap P_i| \geq (\alpha+\mathcal{X}/4)|P_i|$$

where $\mathcal{X} = \alpha\sqrt{\beta\gamma}/2$.
The following proposition gives a brief description of the current subsection.

Proposition 2. *Let A be a α-density subset of $\mathbf{Z}_N$ and there exists r such that $|\hat{1}_A(r)| \geq \mathcal{X}$. Then there is an arithmetic progression $P \subset \{1,\dots,N\}$ such that*

1. *P is long enough ($|P| \geq \sqrt{\mathcal{X}N/16\pi}$) and*
2. *P has a large density in A, namely $|A\cap P| \geq (\alpha+\mathcal{X}/4)|P|$.*

2 Proof of Roth's theorem

2.1 Density increment argument

2.1.1

Let A_0 be a subset of $\{1,\dots,N_0\}$ with density at least δ_0, i.e. $|A_0| \geq \delta_0 N_0$. In the case if N_0 is odd, we set A be A_0 and N be N_0.
In case if N_0 is even, we split $\{1,\dots,N_0\}$ into two pieces

$$\{1,\dots,N_0\} \xrightarrow{\text{split}} \{1,\dots,L_0\} \sqcup \{L_0+1,\dots,N_0\}.$$

such that each piece‘s length is at least $N_0/4$ and odd. Well, then, either

$$|A_0\cap\{1,\dots,L_0\}| \geq \delta_0 L_0 \qquad \text{or} \qquad |A_0\cap\{L_0,\dots,N_0\}| \geq \delta_0(N_0-L_0).$$

Indeed, otherwise, the density of A_0 should be less than δ, and we reach a contradiction.
In the first case, we set A be $A_0\cap\{1,\dots,L_0\}$ and $N = L_0$. In the opposite one, we set A be $(A_0-L_0)\cap\{1,\dots,N_0-L_0\}$ and $N = N_0-L_0$.

2.1.2

If we have a small intersection with the middle third $(N/3;2N/3)$, id est

$$\left|A \cap \left(\frac{N}{3};\frac{2N}{3}\right]\right| < \frac{\delta N}{5}$$

Then, by similar argument as in subsubsection 2.1.1 we obtain that either

$$\left|A \cap \left[1;\frac{N}{3}\right]\right| > \frac{2\delta N}{5} \quad \text{or} \quad \left|A \cap \left(\frac{2N}{3};N\right]\right| > \frac{2\delta N}{5}. \tag{13}$$

Never mind which third is dense in A (id est $> 2\delta N/5$) we can consider it as an arithmetic progression, say P. The length of P is at most $N/3$. We can colour all the elements in $A \cap P$ in red and perform an affine transformation to send P into $\{1,\dots,|P|\}$. Then all the red elements will lie in $\{1,\dots,|P|\}$. Then, as in (13)

$$|A \cap P| \geq \frac{2\delta N}{5} = \frac{6\delta}{5} \cdot \frac{N}{3} \geq \frac{6\delta}{5}|P|.$$

Then we repeat our density increment argument for $|A \cap P| \subseteq \mathbf{Z}_{|P|}$.

2.1.3

We assume that $|A \cap (N/3;2N/3]| \geq \delta N/5$. Let $B = C = A \cap [N/3;2N/3)$. Then $|B|,|C| \geq (\delta/5)N$. In addition, it is easy to see that if P is an arithmetic progression in $A \times B \times C$ (regarded as subset of $\mathbf{Z}_N$) then P is arithmetic progression in $\{1,\dots,N\}$.
Let α,β be the densities of sets A and $B = C$. Then $\alpha \geq \delta$ and $\beta \geq \delta/5$.
Let N be sufficiently large, namely

$$N \geq \frac{2}{\alpha\beta^2} = \frac{50}{\delta^3}.$$

in order to apply proposition 1. We are extremely happy if we found 3-AP. Otherwise, we have found a large Fourier coefficient, id est

$$\exists r \neq 0: \quad |\hat{1}_A(x)| \geq \frac{\alpha\beta}{2} \geq \frac{\delta^2}{10}.$$

We apply proposition 2 with $\mathcal{X} = \delta^2/10$ to obtain that there exists a long arithmetic progression

P of size at least
$$|P| \geq \sqrt{\frac{\delta^2 N}{160\pi}} = \frac{\delta\sqrt{N}}{\sqrt{160\pi}} \geq \frac{\delta\sqrt{N}}{50}.$$
such that its density in A is quite large
$$|A \cap P| \geq \left(\delta + \frac{\delta^2}{40}\right)|P|.$$

We mark all the elements of $|A \cap P|$ in red colour. Then we apply affine map to send P (which currently a subset of $\{1,\ldots,N\}$) to $\{1,\ldots,|P|\}$. In particular, all red elements (in fact, they are common elements of the progression and the original set) will be there. Well, then, considering $A \cap P$ as subset in $\{1,\ldots,|P|\}$, we have augmented the density from δ to at least $\delta + \delta^2/40$. One may repeat this argument until the density is greater than one.

2.1.4

Let $d(k)$ be a lower bound on the density. Then,
$$\begin{cases} d(k+1) = d(k) + \frac{d(k)^2}{40}, \\ d(0) = \delta. \end{cases}$$

Note that
$$d(k+1) - d(k) = \frac{d(k)^2}{40}.$$
We consider the following telescoping sum
$$d(k) - d(0) = \sum_{i=0}^{k}(d(i+1) - d(i)) = \frac{1}{40}\sum_{i=0}^{k} d(i)^2$$

Obviously, $d(a+1) > d(a)$ then
$$d(k+1) = d(0) + \frac{1}{40}\sum_{i=0}^{k} d(i)^2 > d(0) + \frac{k+1}{40}d(0)^2 = \delta + \frac{(k+1)\delta^2}{40}.$$

Hence, for the sufficiently large number of iterations, namely k density will be greater than one, and we reach a contradiction.