Roth's thm on arithmetic progressions

Thm 1 (Szemerédi, 1975) Let $A \subseteq \mathbb{N}$ be a subset with positive upper density

$$\bar{d}(A) = \limsup_{n \to \infty} \frac{|A \cap \{1, \cdots, n\}|}{n} > 0.$$

Then A contains a k-term AP for every $k \in \mathbb{N}$.

Cor 2 (Van der Waerden, 1927) Let $l \in \mathbb{N}$. Then for any partition $A_1, \cdots, A_l$ of $\mathbb{N}$, there exists $1 \leq i \leq l$ such that A_i contains a k-term AP for every $k \in \mathbb{N}$.

Thm 3 (Szemerédi, 1975) $\forall\, k \in \mathbb{N}$ and $\delta \in (0, 1]$, $\exists\, N(k, \delta) \in \mathbb{N}$

such that $\forall\ n \geq N(k,\delta)$ and $\forall\ A \subseteq \{1, \cdots, n\}$ with $|A| \geq \delta n$, A

contains a k-term AP.

Fact : Thm 1 $\Leftrightarrow$ Thm 3.

Roth (1953) : $k=3$.

Fourier-analytic proof strategy : Fourier analysis + density increment

Notation : $[n] := \{0, 1, \cdots, n-1\}$.

$$\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}.$$

$$e_N(x) := e^{\frac{2\pi i x}{N}}.$$

1_A : characteristic function of A.

Proposition 3 (Density Increment Lemma) Let $\delta > 0$, and let $A \subseteq [N]$ be a subset with $|A| \geq \delta N$, where $N \geq 8\delta^{-2}$. Then one of the following holds:

(i) A contains a 3-term AP; or

(ii) there exists an AP in $\mathbb{N}$, say P, of length $|P| \geq \frac{1}{256}\delta^2\sqrt{N}$ such that $|A \cap P| \geq (\delta + \frac{1}{64}\delta^2)|P|$.

Prop 3 $\Rightarrow$ Roth's thm (Thm 2, case $k=3$):

$\delta_0 = \inf\{\delta > 0 : \exists N(\delta)$ s.t. every $A \subseteq [n]$ with $|A| = \delta n$ contains a 3-term AP whenever $n \geq N(\delta)\}$.

Assume Roth's thm fails. Then $\delta_0 > 0$. Let $\delta \in (0, \delta_0)$. Then for arbitrarily large N, $\exists A \subseteq [N]$ with $|A| = \delta N$ such that A contains no 3-term APs. By Prop 3, there exists an AP in $\mathbb{N}$, say P, of length $|P| \geq \frac{1}{256}\delta^2\sqrt{N}$ such that $|A \cap P| \geq (\delta + \frac{1}{64}\delta^2)|P|$.

Write $P = \{x_1, \cdots, x_{|P|}\}$ with $x_1 < \cdots < x_{|P|}$. Then $P \cong [|P|]$ and $A \cap P \cong \{i_1, \ldots, i_r\} \subseteq [|P|]$. Choose $\delta \in (0, \delta_0)$ so that $\delta + \frac{1}{64}\delta^2 > \delta_0$.

Then $A \cap P$ contains a 3-term AP by definition of δ_0, a contradiction.

The rest of the talk will be devoted to sketching the proof of Prop 3.

1. Fourier Analysis on $\mathbb{Z}_N$.

Given $f: \mathbb{Z}_N \to \mathbb{C}$, the Fourier transform of f is defined by

$$\hat{f}(k) := \sum_{x \in \mathbb{Z}_N} f(x)\, e_N(-kx).$$

Properties: 1. $\hat{f}(0) = \sum_{x \in \mathbb{Z}_N} f(x)$.

$$\hat{1}_A(0) = |A|, \text{ where } A \subseteq [N].$$

2. $|\hat{f}(k)| \leq \sum_{x \in \mathbb{Z}_N} |f(x)|$

3. Define the convolution of $f, g: \mathbb{Z}_N \to \mathbb{C}$ by

$$f * g(k) = \sum_{x \in \mathbb{Z}_N} f(x)\, g(k-x).$$

Then $\widehat{f*g}(k) = \hat{f}(k)\, \hat{g}(k)$.

4. Plancherel's identity

$$\sum_{x \in \mathbb{Z}_N} |f(x)|^2 = \frac{1}{N} \sum_{k \in \mathbb{Z}_N} |\hat{f}(k)|^2,$$

which follows from the orthogonality relation

$$1_{N\mathbb{Z}}(x) = \frac{1}{N} \sum_{k=0}^{N-1} e_N(-kx) = \frac{1}{N} \hat{1}(x).$$

5. Fourier inversion formula

$$f(x) = \sum_{k \in \mathbb{Z}_N} \hat{f}(k)\, e_N(kx).$$

2. Counting 3-term AP's

Let $A, B, C \subseteq [n]$. Then the number of solutions of $a+b=2c$

in $\mathbb{Z}_N$ with $a \in A$, $b \in B$, $c \in C$ is

$$\mathcal{N} = \sum_{a \in A} \sum_{b \in B} \sum_{c \in C} 1_{N\mathbb{Z}}(a+b-2c)$$

$$= \sum_{a \in A} \sum_{b \in B} \sum_{c \in C} \frac{1}{N} \sum_{k=0}^{N-1} e_N(-k(a+b-2c))$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \hat{1}_A(k)\, \hat{1}_B(k)\, \hat{1}_C(-2k).$$

Observation: If $x+y=2z$ holds in $\mathbb{Z}_N$ with $x, z \in M_A \cap [\frac{N}{3}, \frac{2N}{3})$

and $y \in A$, then $x+y=2z$ holds in $\mathbb{Z}$, where $|A| = \delta N$.

Thus, the number of solutions of $x+y=2z$ in A is at least

$$\frac{1}{N} \sum_{k=0}^{N-1} \hat{1}_{M_A}(k)\, \hat{1}_A(k)\, \hat{1}_{M_A}(-2k) = \delta |M_A|^2 + \frac{1}{N} \sum_{k=1}^{N-1} \hat{1}_{M_A}(k)\, \hat{1}_A(k)\, \hat{1}_{M_A}(-2k).$$

To make sure this is large, we want $|M_A|$ to be large and

$$\left| \sum_{k=1}^{N-1} \hat{1}_{M_A}(k)\, \hat{1}_A(k)\, \hat{1}_{M_A}(-2k) \right|$$

to be small.

Def^n. We say that A is ε-uniform (or suitably uniform) if

$$|\hat{1}_A(k)| \leqslant \varepsilon N.$$

Now, if $A \subseteq [N]$ is ε-uniform, then

$$\left| \sum_{k=1}^{N-1} \hat{1}_{M_A}(k)\, \hat{1}_A(k)\, \hat{1}_{M_A}(-2k) \right| \leqslant \varepsilon N \sum_{k=1}^{N-1} \left| \hat{1}_{M_A}(k)\, \hat{1}_{M_A}(-2k) \right|$$

$$\overset{C\text{-}S}{\leqslant} \varepsilon N \Big(\sum_{k \in \mathbb{Z}_N} |\hat{1}_{M_A}(k)|^2 \Big)^{\frac{1}{2}} \Big(\sum_{k \in \mathbb{Z}_N} |\hat{1}_{M_A}(-2k)|^2 \Big)^{\frac{1}{2}}$$

$$\underset{2 \nmid N}{\overset{\text{Plancherel}}{=}} \varepsilon N \Big(N \sum_{x \in \mathbb{Z}_N} 1_{M_A}(x) \Big) = \varepsilon N^2 |M_A|.$$

So the number of solutions of $x + y = 2z$ in A is at least

$$\delta |M_A|^2 - \varepsilon N^2 |M_A| \geq \frac{\delta^3 N^2}{32} > \delta N.$$

if $\varepsilon \leq \frac{\delta^2}{8}$ and $|M_A| \geq \frac{\delta}{4} N$.

Lemma 4. Let $\delta > 0$, and let $A \subseteq [n]$ be a subset with $|A| = \delta N$, where $N \geq 8\delta^{-2}$. Then one of the following holds:

(i) A contains a 3-term AP; or

(ii) There exits an AP in $\mathbb{N}$, say P, of length $|P| \geq \frac{N}{3}$ such that $|A \cap P| \geq (\delta + \frac{\delta}{8}) |P|$; or

(iii) A is not ε-uniform for any $\varepsilon \leq \frac{\delta^2}{8}$.

Pf. If $|M_A| \geq \frac{\delta}{4} N$, can take $P = [\frac{N}{3}, \frac{2N}{3}) \cap \mathbb{N}$. Otherwise,

$$\max\left(\left|A\cap\left[0,\tfrac{N}{3}\right)\right|,\ \left|A\cap\left[\tfrac{2N}{3},N\right)\right|\right) \geq \frac{1}{2}\left(\delta-\frac{\delta}{4}\right)N = \frac{3}{8}\delta N = \frac{9\delta}{8}\cdot\frac{N}{3},$$

so that $P=\left[0,\frac{N}{3}\right)$ or $P=\left[\frac{2N}{3},N\right)$ works.

3. Sets Which Are Not suitaly Random

An AP in $\mathbb{Z}_N$ of length L and common difference d is said to be

non-overlapping if $dL<N$, so that it is a disjoint union of ≤ 2 $\mathbb{Z}$-APs.

Fact: If P_0 is a non-overlapping AP in $\mathbb{Z}_N$ with $|A\cap P_0| = (\delta+\varepsilon)|P_0|$, (in $\mathbb{Z}_N$)

then there exists an AP in $\mathbb{Z}$, say P, of length $|P|\geq\frac{1}{2}\varepsilon|B|$ with

$|A\cap P|\geq(\delta+\frac{1}{2}\varepsilon)|P|$.

Lemma 5. If $|\hat{1}_A(k)| \geq \varepsilon N$ for some $1 \leq k < N$, then there exists a non-overlapping AP in $\mathbb{Z}_N$, say B, of length $|B| \geq \frac{\sqrt{N}}{4}$ such that $|A \cap B| \geq (\delta + \frac{\varepsilon}{4})|B|$.

Lemma 5 + Fact handles Lemma 4 (iii):

Suppose that A is not ε-uniform for any $\varepsilon \leq \frac{\delta^2}{8}$. Then Lemma 5 $\Rightarrow$ $\exists$ AP in $\mathbb{Z}_N$, say B, with $|B| \geq \frac{\sqrt{N}}{4}$ and $|A \cap B| \geq (\delta + \frac{\varepsilon}{4})|B|$. By Fact, $\exists$ AP in $\mathbb{Z}$, say P, with $|P| \geq \frac{\varepsilon}{32}\sqrt{N}$ and $|A \cap P| \geq (\delta + \frac{\varepsilon}{8})|P|$. May take $\varepsilon = \frac{\delta^2}{8}$. Hence Prop 3 holds.

4. APs in Sparse Subsets

What about subsets $A \subseteq \mathbb{N}$ with zero upper density ?

Ex. $A = \{n^2 : n \in \mathbb{N}\}$ contains infinitely many 3-term APs.

$$(a^2-b^2-2ab)^2+(a^2-b^2+2ab)^2=2(a^2+b^2)^2$$

Thm 6 (Green-Tao, 2004) The set of primes contains infinitely many k-term APs for every $k \in \mathbb{N}$.

Twin Prime Conjecture $\Rightarrow$ infinitely many 2-term APs with common differece 2.

Conjecture (Erdős, 1976) If $A \subseteq \mathbb{N}$ satisfies $\sum_{n \in A} \frac{1}{n} = \infty$, then A contains infinitely many k-term APs for every $k \in \mathbb{N}$.

Erdős' conjecture implies both Szemerédi's thm and the Green-Tao thm.