University of Oxford

Part B extended essay

3-TERM ARITHMETIC PROGRESSIONS

Rattana Pukdee

Master of mathematics

Hilary 2019

Contents

1 Introduction

This paper will discuss Szemerédi's theorem, a result concerning arithmetic progressions in subsets of the integers. The theorem proves the Erdös and Turán conjecture that all sets of natural numbers of positive upper density contain arbitrarily long arithmetic progressions. We will look at the special case when the length of arithmetic progressions is equal to 3. This case was originally proved by Klaus Roth [Rot53] via a Fourier analysis method. After Roth, there were many versions of proof of the theorem, each containing different ideas. For example, one proof uses a graph theoretic idea from Szemerédi's regularity lemma. One may also study the finite field analogue of the problem. This was originally done by Brown and Buhler [BB82] and recently a breakthrough was achieved by Croot Lev and Pach [CLP17]. We will explore the main ideas of various proofs of this special case and try to come up with a self-contain and reader-friendly version for undergraduate students.

2 Preliminary

Roth's and Meshulam's approach are based on Fourier analysis so this chapter is about Fourier transforms on $\mathbb{Z}$ and a finite abelian group G. We will begin with the definition of characters.

Definition 2.1 (Character of G)**.** *Let G be a finite abelian group. A character of G is a map $\chi : G \to \mathbb{C}\backslash 0$ that*

$$\chi(a+b) = \chi(a)\chi(b) \quad \textit{for all } a, b \in G.$$

We called χ_0, a trivial character of G if $\chi_0(a) = 1$ for all $a \in G$. The definition is the same for characters of $\mathbb{Z}$. We will explore some basic properties of the character of $\mathbb{Z}$ first. By substituting $b = 0$, we have

$$\chi(a) = \chi(a)\chi(0)$$

and because $\chi(a) \neq 0$ we must have $\chi(0) = 1$. For a positive integer m, we also have

$$\chi(m) = \chi(m-1)\chi(1) = \cdots = \chi(1)^m.$$

and $\chi(0) = \chi(m)\chi(-m)$ implies $\chi(-m) = \chi(1)^{-m}$. Therefore, a character in $\mathbb{Z}$ is uniquely determined by the value of $\chi(1)$. Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ be a set of real number quotients by a set of integer. Let $e(x) := e^{2\pi i x}$. We can check that $e(\theta x) : x \to e^{2\pi i \theta x}$ is a character of $\mathbb{Z}$ for any $\theta \in \mathbb{T}$.

On the other hand, a character of a group G yields a similar result. Consider a group $(\mathbb{Z}/3\mathbb{Z})^n$, for example, any element of this can be written as a sum of basis element $(1, 0, \ldots, 0), \ldots, (0, 0, \ldots, 1)$ which we write $b_1, b_2, \ldots, b_n$ for convenience. Similarly, a character of this group is uniquely determined by the value of $\chi(b_1), \ldots, \chi(b_n)$ as for $x = (x_1, x_2, \ldots, x_n)$ we have

$$\chi(x) = \chi((x_1, x_2, \ldots, x_n)) = \prod_{i=1}^{n} \chi(b_i)^{x_i}.$$

Moreover, for any basis b_i

$$\chi(b_i)^3 = \chi(3b_i) = \chi(0) = 1$$

so $\chi(b_i)$ must be 3^{rd} root of unity which has 3 possible values, $1, e(\frac{1}{3}), e(\frac{2}{3})$. If $\chi(b_i) = e(\frac{r_i}{3})$ for each i, then

$$\prod_{i=1}^{n} \chi(b_i)^{x_i} = \prod_{i=1}^{n} e(\frac{r_i x_i}{3}) = e(\frac{1}{3}(r_1 x_1 + \ldots r_n x_n))$$

Therefore, any character of $(\mathbb{Z}/3\mathbb{Z})^n$ can be written in the form

$$\chi(x) : x \to e(\frac{1}{3}(r_1 x_1 + \ldots r_n x_n))$$

where $r = (r_1, \ldots, r_n) \in (\mathbb{Z}/3\mathbb{Z})^n$.

In fact, characters of a group G form an abelian group under multiplication which we called it the character group $\hat{G}$ and there is an isomorphism from $\hat{G}$ to G. We won't give a formal proof here, but from above, the reader might get an idea on why there would be an isomorphism.

Definition 2.2 (Fourier transform)**.** *Let G be a finite abelian group and $f : G \to \mathbb{C}$, the Fourier transform of f is defined by*

$$\hat{f}(\chi) := \sum_{x \in G} f(x)\chi(-x)$$

and for $f : \mathbb{Z} \to \mathbb{C}$, the Fourier transform of f is

$$\hat{f}(\theta) := \sum_{x \in \mathbb{Z}} f(x)e(-x\theta)$$

Proposition 2.1. *Let G be a finite abelian group, then*

$$\sum_{\chi\in\hat{G}}\chi(x)=\begin{cases}|G| & \text{if } x=0,\\ 0 & \text{if } x\neq 0.\end{cases}$$

Proof. Although, this is true for any finite abelian group G, we will give a proof for the case $G=(\mathbb{Z}/3\mathbb{Z})^n$ as it is the only group we are interested in for this essay. It is clear that when $x=0$ the result holds. Let $x=(x_1,\ldots,x_n)\neq 0$ then there is $i\in\{1,\ldots,n\}$ that $x_i\neq 0$. Without loss of generality let $x_1\neq 0$. From above, we know that any character of $(\mathbb{Z}/3\mathbb{Z})^n$ could be written in the form

$$\chi(x):x\to e(\frac{1}{3}(r_1x_1+\ldots r_nx_n))$$

where $r=(r_1,\ldots,r_n)\in(\mathbb{Z}/3\mathbb{Z})^n$. We have

$$\begin{aligned}\sum_{\chi\in\hat{G}}\chi(x)&=\sum_{r\in(\mathbb{Z}/3\mathbb{Z})^n}e(\frac{1}{3}(r_1x_1+\ldots r_nx_n))\\&=\sum_{(r_2,\ldots,r_n)\in(\mathbb{Z}/3\mathbb{Z})^{n-1}}e(\frac{1}{3}(r_2x_2+\ldots r_nx_n))(1+e(\frac{1}{3}x_1)+e(\frac{2}{3}x_1))\\&=0\end{aligned}$$

The last equality is true because $x_1\neq 0$ so $1+e(\frac{1}{3}x_1)+e(\frac{2}{3}x_1)=0$. □

Proposition 2.2. *Let $e:\mathbb{Z}\to\mathbb{C}$ is defined as $e(\theta x)=e^{2\pi i\theta x}$, then for any integer m,*

$$\int_{\theta\in\mathbb{T}}e(m\theta)d\theta=\begin{cases}1 & \text{if } m=0,\\ 0 & \text{if } m\neq 0.\end{cases}$$

Proof. It is clear that when $m=0$ the result holds. Let $m\neq 0$ we have

$$\begin{aligned}\int_{\theta\in\mathbb{T}}e(m\theta)d\theta&=\int_{\theta\in\mathbb{T}}e^{2\pi mi\theta}d\theta\\&=\frac{1}{2\pi mi}[e^{2\pi mi\theta}]_{\theta=0}^{\theta=1}\\&=0\end{aligned}$$

□

Theorem 2.1 (Parvesal's identity)**.** *Let G be a finite abelian group and $f:G\to\mathbb{C}$, then*

$$\frac{1}{|G|}\sum_{\chi\in\hat{G}}|\hat{f}(\chi)|^2=\sum_{x\in G}|f(x)|^2$$

Proof. Expand the term directly, here we use the fact that $\overline{\chi(-x)} = \chi(x)$,

$$\begin{aligned}\sum_{\chi\in\hat{G}} |\hat{f}(\chi)|^2 &= \sum_{\chi\in\hat{G}} |\sum_{x\in G} f(x)\chi(-x)|^2 \\ &= \sum_{\chi\in\hat{G}} (\sum_{x\in G} f(x)\chi(-x))(\sum_{y\in G} \overline{f(y)}\chi(y)) \\ &= \sum_{\chi\in\hat{G}} (\sum_{x,y\in G} f(x)\overline{f(y)}\chi(-x+y) \\ &= \sum_{x,y\in G} f(x)\overline{f(y)}(\sum_{\chi\in\hat{G}} \chi(-x+y))\end{aligned}$$

From the proposition (2.1), $\sum_{\chi\in\hat{G}} \chi(-x+y) = 0$ when $x \neq y$. We have

$$\begin{aligned}\sum_{\chi\in\hat{G}} |\hat{f}(\chi)|^2 &= \sum_{x\in G} f(x)\overline{f(x)}|G| \\ &= |G| \sum_{x\in G} |f(x)|^2\end{aligned}$$

□

Theorem 2.2 (Parseval's identity)**.** *For $f : \mathbb{Z} \to \mathbb{C}$ with finite support, we have*

$$\int_{\theta\in\mathbb{T}} |\hat{f}(\theta)|^2 d\theta = \sum_{x\in\mathbb{Z}} |f(x)|^2$$

Proof. Because f has a finite support, we can interchange the integration with a sum,

$$\begin{aligned}\int_{\theta\in\mathbb{T}} |\hat{f}(\theta)|^2 d\theta &= \int_{\theta\in\mathbb{T}} |\sum_{n\in\mathbb{Z}} f(n)e(-n\theta)|^2 d\theta \\ &= \int_{\theta\in\mathbb{T}} \sum_{n,m\in\mathbb{Z}} f(n)\overline{f(m)}e((m-n)\theta)d\theta \\ &= \sum_{n,m\in\mathbb{Z}} f(n)\overline{f(m)} \int_{\theta\in\mathbb{T}} e((m-n)\theta)d\theta\end{aligned}$$

From the proposition (2.2), we know that $\int_{\theta\in\mathbb{T}} e((m-n)\theta)d\theta = 0$ if $m \neq n$ and equals to 1 otherwise, so

$$\begin{aligned}\int_{\theta\in\mathbb{T}} |\hat{f}(\theta)|^2 d\theta &= \sum_{n\in\mathbb{Z}} f(n)\overline{f(n)} \\ &= \sum_{x\in\mathbb{Z}} |f(x)|^2\end{aligned}$$

□

3 Roth's theorem on $\mathbb{Z}$

From here on in this essay, we write $[N]$ for a set $\{1, 2, \ldots, N\}$. We will go through the definition of density and upper density first.

Definition 3.1 (Density)**.** *For any set A, P, define the density of A in P as*

$$d_p(A) = \frac{|A \cap P|}{|P|}$$

Definition 3.2 (Upper density)**.** *For a subset $A \subset \mathbb{Z}$, define an upper density of A as*

$$\limsup_{N \to \infty} d_{[N]}(A)$$

The Erdos and Turan conjecture on arithmetic progression states that all sets of natural numbers with positive upper density contain k terms arithmetic progressions for any positive integer k (a progression in the form $(a, a+d, a+2d, \ldots, a+(k-1)d)$ when $d \neq 0$). If $k = 3$, we have Roth's theorem.

It is clear from the definition that any finite subset of natural numbers has zero upper density so any subset with a positive upper density must be an infinite set. Working with an infinite set might be hard so we consider an equivalent statement instead which states that for a fixed density $\delta > 0$, when N is large enough, any subset $A \subset [N]$ with a density δ must contain a 3-term arithmetic progression. We can restate the Roth's theorem in the following way.

Theorem 3.1 (Roth's theorem)**.** *For any $\delta > 0$, there exists an integer N_0 such that for every $N \geq N_0$, if $A \subset [N]$ with $|A| \geq \delta N$ then A must contain a 3-term arithmetic progression.*

3.1 Proof of Roth's theorem

We will explore what can be implied if $A \subset [N]$ does not contain a 3-term arithmetic progression. One clear result is that any subset of A must not contain a 3-term arithmetic progression. This property turns out to be a key idea in this chapter and if we combine it with the next proposition, we can prove Roth's theorem.

Proposition 3.1. *Suppose that $0 < \delta < 1$. Suppose that $P \subset \mathbb{Z}$ is an arithmetic progression of length M. If $A \subset P$ with a density $d_P(A) = \delta$ does not contain a 3-term arithmetic progression and $M \geq (\frac{18}{\delta})^6$, there is an arithmetic progression $P' \subset \mathbb{Z}$ of length M' and $A' \subset P'$ such that*

1. *$d_{P'}(A') \geq \delta + \frac{\delta^2}{16}$*
2. *A' is a subset of A*
3. *$M' \geq M^{1/3}$*

As mentioned above, a subset A' must contain no 3-term arithmetic progression either and we can apply the proposition (3.1) again to $A' \subset P'$ if the condition is satisfied and so on. After each step, we increased the density $d_{P'}(A')$ and decreased the length M'. Note that a density can't be greater than 1 so if we start with a density δ, we can apply the proposition (3.1) a finite number of times which depends on δ. For example,

if $\delta = 0.9$, we know that $0.9 + \frac{0.9^2}{16} \sim 0.951$ and $0.95 + \frac{0.95^2}{16} \sim 1.01$ so we can apply the proposition (3.1) at most 1 time. However, the only condition that prevents us from apply the proposition (3.1) is that the length of P is not long enough. Therefore, if P is long enough so that we can apply the proposition (3.1) as many times as we need to have a density greater than 1, there must be no such $A \subset P$ that does not contain a 3-term arithmetic progression in the first place. This is called a density increment technique and the idea of the proof is adapted from Ben Green's lecture notes [Gre].

Proof of Roth's theorem: We will prove Roth's theorem via induction on a density $\delta = \frac{2}{n}$. We will show that for any positive integer n, there is an integer N_0 such that for any $N > N_0$, any subset $A \subset [N]$ with density $d_{[N]}(A) > \frac{2}{n}$ must contain a 3-term arithmetic progression. Because $\frac{2}{n} \to 0$ as $n \to \infty$, the result implies Roth's theorem.

Cases $n = 1, 2$ are trivial. Consider any density $\delta > \frac{2}{3}$. There is an integer N_0 large enough so that for all $N > N_0$,

$$\delta N > \frac{2}{3}N + 3.$$

We partition $[N]$ into classes $\{1, 2, 3\}, \{4, 5, 6\}, \ldots$, where each class contains 3 consecutive numbers and the last class is the leftover of length ≤ 2. There are at most $\lfloor \frac{N}{3} \rfloor + 1$ classes, but $|A| = \delta N > \frac{2}{3}N + 3 \geq 2\lfloor \frac{N}{3} \rfloor + 3$. By the pigeonhole principle, there must be a class that contains 3 members of A and which implies a 3-term arithmetic progression in A and so the base case is true.

Now, assume that for a positive integer n, there exists an integer $N_0 > 0$ such that for any $N \geq N_0$, any subset $A \subset [N]$ with $d_{[N]}(A) > \frac{2}{n}$ must contain a 3-term arithmetic progression. Consider a subset $A \subset [N]$ with $d_{[N]}(A) > \frac{2}{n+1}$ that does not contain a 3-term arithmetic progression. Note that $[N]$ is also an arithmetic progression so apply the proposition (3.1) to $A \subset [N]$. There is an arithmetic progression P_1 of length at least $N^{1/3}$ and a subset $A_1 \subset P_1$ such that

$$\begin{aligned} d_{P_1}(A_1) &\geq d_{[N]}(A) + \frac{d_{[N]}(A)^2}{16} \\ &> \frac{2}{n+1} + \frac{1}{4(n+1)^2}. \end{aligned}$$

We also know that $A_1 \subset A$ does not contain a 3-term arithmetic progression and we could apply the proposition (3.1) again. Assuming that N is large enough, we apply the proposition (3.1) k times and get an arithmetic progression P_k of length at least $N^{(1/3)^k}$ and a subset $A_k \subset P_k$ that

$$d_{P_k}(A_k) > \frac{2}{n+1} + \frac{k}{4(n+1)^2}$$

If $k > 16$,

$$\frac{k}{4(n+1)^2} \geq \frac{4}{(n+1)^2} \geq \frac{2}{n(n+1)}.$$

This implies

$$d_{P_k}(A_k) > \frac{2}{n+1} + \frac{2}{n(n+1)} = \frac{2}{n}.$$

Applying a linear transformation to the induction hypothesis, we know that there is an integer $N_k > 0$ that for any arithmetic progression P_k with length more than N_k, any subset $A_k \subset P_k$ with $d_{P_k}(A_k) > \frac{2}{n}$ must contain an arithmetic progression. Therefore,

if $N^{(1/3)^k} > N_k$, there is no subset $A_k \subset P_k$ that does not contain a 3-term arithmetic progression which implies there is no subset $A \subset [N]$ with $d_{[N]}(A) > \frac{2}{n+1}$ that does not contain a 3-term arithmetic progression. This proves an inductive step and completes the proof of the theorem. □

One may be interested in a qualitative bound for the size of the largest subset $A \subset [N]$ that does not contain a 3-term arithmetic progression says $r_3(N)$. We will show next that the upper bound that can be derived from our proof is

$$r_3(N) < \frac{CN}{\log\log N}$$

Observe that if $\frac{2}{k} < \delta < \frac{2}{k-1}$ and there exists $A \subset P$ with density δ that does not contain a 3-term arithmetic progression, we need to apply the proposition (3.1) at most 16 times to increase the density to be in $(\frac{2}{k-1}, \frac{2}{k-2})$ and at most $16k$ times to be in $(\frac{2}{3}, 1)$ which guarantees a 3-term arithmetic progression and leads to a contradiction. If we know that we can apply the proposition (3.1) at least $16k$ times, we are certain that we won't a subset $A \subset [N]$ with density δ that does not contain a 3-term arithmetic progression.

In order to apply the proposition (3.1) successfully, the length of P must be greater than $(\frac{18}{\delta})^6$. Each time we apply the proposition, a new density is increased so that $(\frac{18}{\delta})^6$ is decreased and the constraint on the length of P is weaker while at the same time the length of a new arithmetic progression P' is decreased to at most a cube root of the old length. Therefore,

$$N^{(\frac{1}{3})^{16k}} > (\frac{18}{\delta})^6$$

guarantees that we can apply the proposition (3.1) at least $16k$ times. This is equivalent to

$$(\frac{1}{3})^{16k} \log(N) \geq 6 \log(\frac{18}{\delta})$$
$$-16k \log(3) + \log\log(N) > \log(6 \log(\frac{18}{\delta}))$$

Since $\frac{2}{k} < \delta < \frac{2}{k-1}$, we have $k < \frac{2}{\delta} + 1$ and the above can be implied by

$$\log\log(N) > \frac{32 \log(3)}{\delta} + 16 \log(3) + \log(6 \log(\frac{18}{\delta}))$$
$$\delta \sim O(\frac{C}{\log\log N})$$

Historically, the bound has been improved by many mathematicians such as Bourgain, Heath-Brown, Szemeredi, Sander and Bloom who holds the current record as

$$r_3(N) \leq \frac{C(\log\log N)^4 N}{\log N}.$$

Now, we will prove the proposition (3.1), we split the proof into 2 parts

1. Lack of arithmetic progression implies a large Fourier coefficient
2. With an appropriate partition of $[N]$, a large Fourier coefficient implies a density increment

3.2 A large Fourier coefficient

Let A be a subset of $[N]$ such that $|A| = \delta N$. Define a balance function of A as

$$f_A : 1_A - \delta 1_{[N]}.$$

It is a function that average the density of A through out $[N]$ as

$$\sum_{x\in[N]} f_A(x) = \sum_{x\in[N]} 1_A - \delta 1_{[N]} = |A| - \delta N = 0.$$

If a sum of f_A on a subset $P \subset [N]$ is large then A must have a large density in P. On extreme case,

$$\sum_{x\in A} f_A(x) = |A| - \delta|A| = |A|(1-\delta).$$

We will show that if A does not contain a 3-term arithmetic progression then f_A has a large Fourier coefficient.

Proposition 3.2. *Suppose that $0 < \delta < 1$ and $N \geq 4/\delta^2$. Suppose that $A \subset [N]$ with $|A| = \delta N$ and does not contain a 3-term arithmetic progression then there is $\theta \in \mathbb{T}$ such that*

$$|\hat{f}_A(\theta)| \geq \frac{\delta^2 N}{4}.$$

Proof. For $f_1, f_2, f_3 : \mathbb{Z} \to \mathbb{R}$ that have supports on $[N]$, define a trilinear operator

$$T(f_1, f_2, f_3) = \sum_{x,d\in\mathbb{Z}} f_1(x) f_2(x+d) f_3(x+2d).$$

It is clear that $T(1_A, 1_A, 1_A)$ counts the number of 3-term arithmetic progressions in A including trivial progressions (a, a, a). Since A does not contain a non-trivial 3-term arithmetic progression, we have

$$T(1_A, 1_A, 1_A) = |A| \tag{1}$$

On the other hand,

$$T(1_A, 1_{[N]}, 1_A) = \sum_{x,d\in\mathbb{Z}} 1_A(x) 1_{[N]}(x+d) 1_A(x+2d)$$

This counts the number of pairs $(a, b) \in A \times A$ that $\frac{a+b}{2} \in [N]$. We know that any $a, b \in A$, $\frac{a+b}{2} \leq N$. The statement $\frac{a+b}{2} \in [N]$ is equivalent to $a + b$ being divisible by 2 that is a, b have the same parity. If A has x odd elements then it has $|A| - x$ even elements and the number of pairs $(a, b) \in A \times A$ with the same parity is

$$x^2 + (|A| - x)^2 \geq \frac{|A|^2}{2}$$

Therefore,

$$T(1_A, 1_{[N]}, 1_A) \geq \frac{|A|^2}{2} \tag{2}$$

From (1), (2) we have

$$T(1_A, f_A, 1_A) = T(1_A, 1 - \delta 1_{[N]}, 1_A) \leq |A| - \frac{\delta|A|^2}{2} \tag{3}$$

Recall that $N \geq 4/\delta^2$, this implies

$$\begin{aligned} |A| - \frac{\delta|A|^2}{2} &= |A|(1 - \frac{\delta|A|}{2}) \\ &= |A|(1 - \frac{\delta^2 N}{2}) \\ &< 0 \end{aligned}$$

Next, the key point of this chapter that links us from lacking of a 3-term arithmetic progression to Fourier analysis is that we observe that

$$\begin{aligned} \int_{\theta\in\mathbb{T}} \hat{f}_1(\theta)\hat{f}_2(-2\theta)\hat{f}_3(\theta)d\theta &= \int_{\theta\in\mathbb{T}} \sum_{n,m,l\in\mathbb{Z}} f_1(n)f_2(m)f_3(l)e(-n\theta + 2m\theta - l\theta)d\theta \\ &= \sum_{n,m,l\in\mathbb{Z}} f_1(n)f_2(m)f_3(l) \int_{\theta\in\mathbb{T}} e((2m-n-l)\theta)d\theta \end{aligned}$$

From the proposition (2.2), we know that

$$\int_{\theta\in\mathbb{T}} e((2m-n-l)\theta)d\theta = 0$$

if $2m - n - l \neq 0$ and is equal to 1 when (n, m, l) is a 3-term arithmetic progression. We have

$$\begin{aligned} \int_{\theta\in\mathbb{T}} \hat{f}_1(\theta)\hat{f}_2(-2\theta)\hat{f}_3(\theta)d\theta &= \sum_{n+l=2m} f_1(n)f_2(m)f_3(l) \\ &= \sum_{x,d\in\mathbb{Z}} f_1(x)f_2(x+d)f_3(x+2d) \\ &= T(f_1, f_2, f_3) \end{aligned}$$

Because $T(1_A, f_A, 1_A)$ is negative, take a modulus, we have

$$\begin{aligned} \frac{\delta|A|^2}{2} - |A| &\leq |T((1_A, f_A, 1_A)| \\ &= |\int_{\theta\in\mathbb{T}} \hat{1}_A(\theta)\hat{f}_A(-2\theta)\hat{1}_A(\theta)d\theta| \\ &\leq \sup_{\theta\in\mathbb{T}} |\hat{f}_A(-2\theta)| \int_{\theta\in\mathbb{T}} |\hat{1}_A(\theta)|^2 d\theta \\ &= \sup_{\theta\in\mathbb{T}} |\hat{f}_A(\theta)| \int_{\theta\in\mathbb{T}} |\hat{1}_A(\theta)|^2 d\theta \end{aligned}$$

The last equality holds as the space $\mathbb{T}$ are the same as $\frac{-1}{2}\mathbb{T}$. By Parseval's identity

$$\int_{\theta\in\mathbb{T}} |\hat{1}_A(\theta)|^2 = \sum_{x\in\mathbb{Z}} |1_A(x)|^2 = |A|$$

Substituting this into the inequality above we have

$$\sup_{\theta\in\mathbb{T}} |\hat{f}_A(\theta)| \geq \frac{\delta|A|}{2} - 1 \geq \frac{\delta^2 N}{4}$$

□

3.3 Density increment arguments

Next, we will try to extract information about $f_A(x)$ from $\hat{f}_A(\theta)$. Recall that

$$\hat{f}_A(\theta) = \sum_{x\in\mathbb{Z}} f_A(x)e(-\theta x)$$

If there is $P \subset \mathbb{Z}$ that for any $x \in P$, $e(\theta x)$ are close to each others. We could factorise out $e(-\theta x')$ for some $x' \in P$ and get

$$\begin{aligned}\sum_{x\in P} f_A(x)e(-\theta x) &= e(-\theta x')\sum_{x\in P} f_A(x)\frac{e(-\theta x)}{e(-\theta x')}\\ &\sim e(-\theta x')\sum_{x\in P} f_A(x)\end{aligned}$$

and could get some information about $f_A(x)$. For $F : \mathbb{Z} \to \mathbb{R}$, we define a diameter of F on a set S as

$$\mathrm{diam}_S(F) = \sup_{x,x'\in S} |F(x) - F(x')|.$$

This measures how clustered an image of S is, after applying F. We will show that we could partition $[N]$ into arithmetic progressions so that $e(-\theta x)$ has a small diameter on each progression.

Proposition 3.3. *Suppose that $\theta \in \mathbb{T}$ and $N \geq (\frac{18}{\delta})^6$. We can partition $[N]$ into arithmetic progressions P_i each of length at least $N^{\frac{1}{3}}$ such that, for all i, we have*

$$diam_{P_i}(e(\theta x)) \leq \frac{\delta^2}{8}.$$

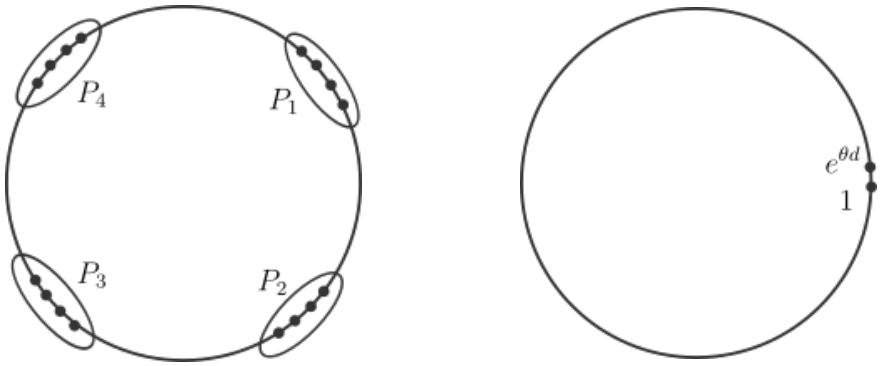
Proof. The idea is that we will try to find d that $e(d\theta)$ is closed to 1 as shown in the figure (1b) and build our progression based on that. For any $x \in \mathbb{R}$, let $||x||$ be the minimum distance from x to an integer. We observe that

$$\begin{aligned}|e(\theta d) - 1| &= |e^{2\pi i\theta d} - 1|\\ &= |e^{\pi i\theta d}||e^{\pi i\theta d} - e^{-\pi i\theta d}|\\ &= 2|sin(\pi d\theta)|\\ &= 2|sin(\pi||d\theta||)|\\ &\leq 2\pi||d\theta||\end{aligned}$$

so the closer $d\theta$ to an integer, the closer $e(d\theta)$ to 1.

Let $Q := \lceil \frac{1}{2}N^{\frac{2}{3}} \rceil$. Consider $\theta, 2\theta, \ldots, Q\theta$ as elements of $\mathbb{T}$. For example, if $\theta = 0.7$ then $2\theta = 0.4, 3\theta = 0.1$. There are Q elements in $[0, 1)$. Splits $[0, 1)$ into Q parts, $[0, \frac{1}{Q}), [\frac{1}{Q}, \frac{2}{Q}), \ldots, [\frac{Q-1}{Q}, 1)$. If there are distinct $j_1, j_2 \leq Q$ such that $j_1\theta, j_2\theta$ lie in the same part then $j = |j_1 - j_2| \leq Q$ is a positive integer that $|j\theta| = |j_1\theta - j_2\theta| < \frac{1}{Q}$. Otherwise, $\theta, 2\theta, \ldots, Q\theta$ must lie in different parts and there must be $j \leq Q$ that $j\theta$ lies in $[0, \frac{1}{Q})$ so that $||j\theta|| < \frac{1}{Q}$. In one way or another, there exists a positive integer $d \leq Q$ such that $||d\theta|| < \frac{1}{Q}$.

We partitions $[N]$ as follows



(a) $e(P_i\theta)$ on a unit circle

(b) 1 and $e^{\theta d}$ on a unit circle

Figure 1

1. We naturally partition $[N]$ into arithmetic progressions with a common difference d
$$\{1, 1+d, 1+2d, \dots\}, \{2, 2+d, 2+2d, \dots\}, \dots$$
so each progression has length around
$$\frac{N}{d} \geq \frac{N}{\lceil \frac{1}{2} N^{\frac{2}{3}} \rceil} > N^{\frac{1}{3}}.$$

2. For each progression above, partition along from the left to the right into progression of length $\lceil N^{\frac{1}{3}} \rceil$ and merge the left over partition of length $< \lceil N^{\frac{1}{3}} \rceil$ with the preceding progression.

We have partitioned $[N]$ into arithmetic progressions P_i with length at least $N^{\frac{1}{3}}$ and at most $2\lceil N^{\frac{1}{3}} \rceil - 1 \leq 3N^{\frac{1}{3}}$. Observe that

$$\begin{aligned}
|e(\theta(x+kd)) - e(\theta x)| &= |e(\theta x)||e(\theta kd) - 1| \\
&= |e(\theta kd) - e(\theta(k-1)d) + e(\theta(k-1)d) - \cdots + e(\theta d) - 1| \\
&\leq \sum_{n=1}^{k} |e(\theta nd) - e(\theta(n-1)d)| \\
&= k|e(\theta d) - 1| \\
&< 2\pi k||d\theta||
\end{aligned}$$

Because $||d\theta|| < \frac{1}{Q}$,

$$|e(\theta(x+kd)) - e(\theta x)| < \frac{2\pi k}{Q} \tag{4}$$

For our arithmetic progression P_i with a common difference d with length less than $3N^{\frac{1}{3}}$. Because P_i is finite, there exists $x_1, x_2 \in P_i$ that

$$\operatorname{diam}_{P_i}(e(\theta x)) = |e(\theta x_1) - e(\theta x_2)|$$

From (4), we have

$$\begin{aligned}\mathrm{diam}_{P_i}(e(\theta x)) &\leq |\frac{x_1 - x_2}{d}|\frac{2\pi}{Q} \\ &\leq 2\pi\frac{3N^{\frac{1}{3}}}{\lceil\frac{1}{2}N^{\frac{2}{3}}\rceil} \\ &\leq 12\pi N^{-\frac{1}{3}} \\ &\leq \frac{\delta^2}{8}.\end{aligned}$$

as required. □

We note that the proposition (3.3) implies we can partition $[N]$ so that the diameter of $e(\theta x)$ on each progression is arbitrary small, given that N is large enough. (4) illustrates a trade off between an arithmetic progression length and the size of diameter on it. Now, we are ready to prove the proposition (3.1)

Proposition 3.4. *Suppose that $|\hat{f}_A(\theta)| \geq \frac{\delta^2 N}{4}$ and $N \geq (\frac{18}{\delta})^6$. Let $[N] = \bigcup_i P_i$ be a partition as in the proposition (3.3). Then there exists some i such that*

$$d_{P_i}(A) \geq \delta + \frac{\delta^2}{16}$$

Proof. Because $f_A(x)$ has support on $[N]$ and by the triangle inequality

$$\begin{aligned}|\hat{f}_A(\theta)| &= |\sum_{x\in\mathbb{Z}} f_A(x)e(-\theta x)| \\ &\leq \sum_i |\sum_{x\in P_i} f_A(x)e(-\theta x)| \\ &\leq \sum_i |\sum_{x\in P_i} f_A(x)| + \sum_i \sum_{x\in P_i} |f_A(x)||e(-\theta x) - 1|\end{aligned}$$

Because $|f_A(x)| \leq 1$ and $|e(-\theta x) - 1| \leq \mathrm{diam}_{P_i}(e(\theta x)) \leq \frac{\delta^2}{8}$, we have

$$\begin{aligned}|\hat{f}_A(\theta)| &\leq \sum_i |\sum_{x\in P_i} f_A(x)| + \sum_i |P_i|\mathrm{diam}_{P_i}(e(\theta x)) \\ &\leq \sum_i |\sum_{x\in P_i} f_A(x)| + \frac{\delta^2 N}{8}\end{aligned}$$

Since $|\hat{f}_A(\theta)| \geq \frac{\delta^2 N}{4}$, we have

$$\sum_i |\sum_{x\in P_i} f_A(x)| \geq |\hat{f}_A(\theta)| - \frac{\delta^2 N}{8} \geq \frac{\delta^2 N}{8}$$

Because $\sum_i \sum_{x\in P_i} f_A(x) = |A| - \delta N = 0$ and $\sum_i |P_i| = N$,

$$\sum_i |\sum_{x\in P_i} f_A(x)| + \sum_i \sum_{x\in P_i} f_A(x) \geq \frac{\delta^2 N}{8} = \frac{\delta^2}{8}\sum_i |P_i|$$

There must be some i such that

$$|\sum_{x\in P_i} f_A(x)| + \sum_{x\in P_i} f_A(x) \geq \frac{\delta^2}{8}|P_i|$$

If $\sum_{x\in P_i} f_A(x) < 0$ then the LHS above is zero which is impossible, so we must have $|\sum_{x\in P_i} f_A(x)| = \sum_{x\in P_i} f_A(x)$ and

$$\begin{aligned}\sum_{x\in P_i} f_A(x) &\geq \frac{\delta^2}{16}|P_i| \\ |A\cap P_i| - \delta|P_i| &\geq \frac{\delta^2}{16}|P_i| \\ d_{P_i}(A) = \frac{|A\cap P_i|}{|P_i|} &\geq \delta + \frac{\delta^2}{16}\end{aligned}$$

□

Let $P' = P_i$ and $A' = A \cap P_i$, we have proved a special case of the proposition (3.1) when $P = [N]$. We could extend this to any arithmetic progression P by applying a linear transformation. For $P = \{a, a+d, \ldots, a+(M-1)d\}$ and $A \subset P$, we apply a linear transformation $T(x) = \frac{x-a}{d} + 1$ so that $T(P) = \{1, 2, \ldots, M\}$. We note that if A does not contain a 3-term arithmetic progression then so does a linear transformation $T(A)$ so we can apply the special case of the proposition (3.1) to $T(A) \subset T(P)$. We can then transform the result back with $T^{-1}(x) = a + d(x-1)$ which proved the proposition (3.1).

4 Graph-theoretic method

In the last chapter, our approach is based on the property that A does not contain a 3-term arithmetic progression. This chapter we will consider what is going to happen to a subset $A \subset [N]$ when N gets large via a graph-theoretic method. We will show that we can construct a graph G such that a 3-term arithmetic progression in A is represented by a triangle in G (will see below). Then, we will show that if $[N]$ is large enough, the corresponding graph G will contain a triangle that implies a non-trivial 3-term arithmetic progression. Our main proposition is the triangle removal lemma.

Proposition 4.1 (Triangle removal lemma)**.** *For any $\varepsilon > 0$, there exists $\delta > 0$ such that for a large enough n, for any graph G with n vertices with at most δn^3 triangles, it is possible to remove at most εn^2 edges from G to make it triangle-free.*

4.1 Proof of Roth's theorem

We will use the triangle removal lemma to prove Roth's theorem first and then we will state and prove the Szemerédi regularity lemma, the triangle counting lemma and then use them to prove the triangle removal lemma.

Proof of Roth's theorem: we adapted an idea of this proof from Choongbum Lee's lecture notes [Lee]. Let $A \subset [N]$ such that $|A| \geq \delta N$. Consider a 3-partite graph $G = U \cup V \cup W$ with $3N$ vertices in each part so that G has $9N$ vertices. Label vertices in U, V, W by u_i, v_j, w_k for $i, j, k = 1, \ldots, 3N$ respectively. Define edges in G as

follows, $u_iv_j \in e(G)$ or $v_jw_k \in e(G)$ if and only if $j - i \in A$ or $k - j \in A$ respectively, $u_iw_k \in e(G)$ if and only if $\frac{1}{2}(k-i) \in A$. Note that $(j-i)+(k-j) = 2(\frac{k-i}{2})$ so that $(j-i, \frac{1}{2}(k-i), k-j)$ is a 3-term arithmetic progression. Therefore, a triangle $u_iv_jw_k$ in $U \times V \times W$ is equivalent to a 3-term arithmetic progression in A, $(j-i, \frac{1}{2}(k-i), k-j)$. We call a triangle $u_iv_jw_k$ in G trivial if $(j-i) = \frac{1}{2}(k-i) = (k-j) = a \in A$ that is it represents a trivial arithmetic progression (a, a, a). One arithmetic progression can be represented by many triangles. For example, $u_1v_3w_5$ and $u_2v_4w_6$ both represent a progression $(2, 2, 2)$.

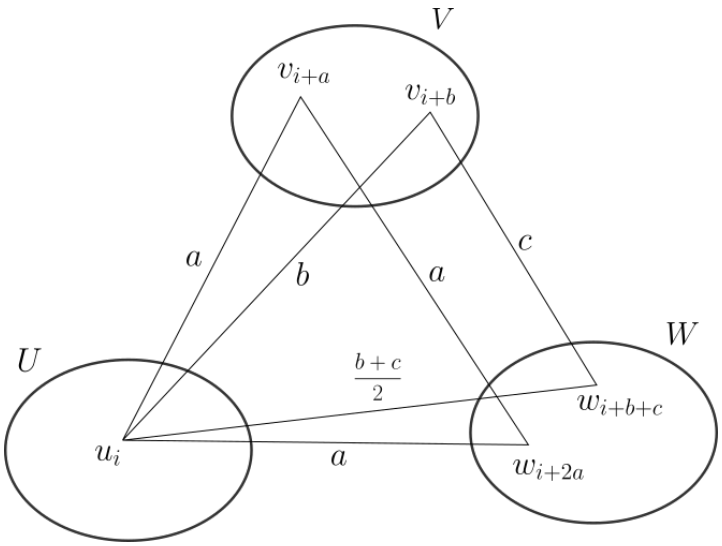


Figure 2: A diagram illustrates triangles in G

Because one vertex of a trivial triangle, $u_i \in U$ and a common difference $a \in A$ uniquely determines the trivial triangle, it is impossible for 2 different trivial triangles to share an edge. Therefore, trivial triangles are edge-disjoint. Since there are at most $3N$ ways to choose a vertex and at most N ways to choose a common difference in A, there are at most $3N^2$ trivial triangles in G.

Now, for any $i \leq N, a \in A$, we have $i + a \leq 3N$ and $i + 2a \leq 3N$ so that if $j = i + a$ and $k = i + 2a$ then $u_iv_jw_k$ is a trivial triangle. Thus there are at least $N|A| \geq \delta N^2$ trivial triangles in G. Because these trivial triangles are edge-disjoint, we need to remove at least δN^2 edges from G to make it triangle free.

By the triangle removal lemma, for $\varepsilon = \frac{\delta}{81}$, there exists δ' such that for N large enough, for any graph G with $9N$ vertices with at most $\delta'(9N)^3$ triangles, it is possible to remove at most $\epsilon(9N)^2 = \delta N^2$ edges from G to make it triangle free. From above, we can't remove δN^2 edges from G to make it triangle free which means there are at least $\delta'(9N)^3$ triangles in G. For a sufficiently large N, $\delta'(9N)^3 > 3N^2$. Since there are at most $3N^2$ trivial triangles in G, there must be a non-trivial triangle $u_iv_jw_k$ in G which implies a non-trivial arithmetic progression in A. □

4.2 The regularity lemma

In order to prove the triangle removal lemma, we require Szemerédi's regularity lemma. The lemma can also be used to prove the generalisation of the triangle removal lemma, the graph removal lemma. In fact, Szemerédi first introduced the earlier version of the regularity lemma in a part of the proof of the Erdös and Turán conjecture or so called Szemerédi's theorem that we are discussing its special case when $k = 3$ in this essay. We will introduce some notations to understand this lemma. The proofs and definitions here are adapted from David Colon, Jacob Fox [CF13], János Komlós, Miklos Simonovits [KS96], Endre Szemerédi [Sze75] and Choongbum Lee [Lee].

Let G be a graph, for X, Y disjoint sets of vertices in G, define $e(X, Y)$ as the number of edges in G between X and Y. Define $d(X, Y) = \frac{e(X,Y)}{|X||Y|}$ as the density between X and Y.

Definition 4.1 (ε-regular pair)**.** *Let G be a graph. A disjoint pair (X, Y) of sets of vertices in G is ε-regular if for every $X' \subset X$ and $Y' \subset Y$ such that $|X'| \geq \varepsilon|X|$ and $|Y'| \geq \varepsilon|Y|$,*

$$|d(X, Y) - d(X', Y')| \leq \varepsilon$$

If a pair (X, Y) is ε-regular then edges between them are highly uniform or so called random-like so that for any large enough subsets of vertices X', Y', the density of X', Y' is about the same as the density of the pair X, Y. If (X, Y) is not ε-regular, we call it an ε-irregular pair.

For example, we know that for a complete bipartite graph (X, Y) , a pair (X, Y) is ε-regular for any $0 < \varepsilon \leq 1$ because for any pair $X' \subset X$, $Y' \subset Y$, we always have a density $d(X', Y') = d(X, Y) = 1$.

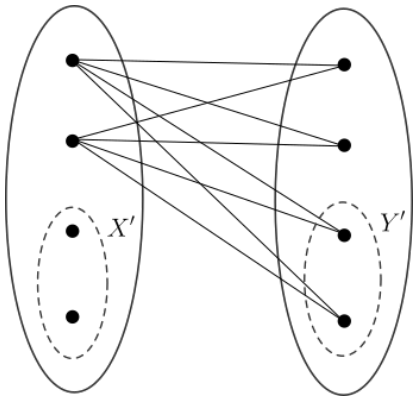


Figure 3: an ε-irregular pair for $0 < \varepsilon < \frac{1}{2}$

The pair (X, Y) that is shown in the figure (3) is not ε-regular for $0 < \varepsilon < \frac{1}{2}$ because we can pick a subset $X' \subset X$ of size up to 2 that has no edge adjacent to it. For any $Y' \subset Y$, we have

$$|d(X, Y) - d(X', Y')| = |d(X, Y)| = \frac{1}{2}$$

On the other hand, consider a random bipartite graph (X, Y) with n vertices in each part where there is an edge between any 2 vertices with a probability $\frac{1}{2}$. This means

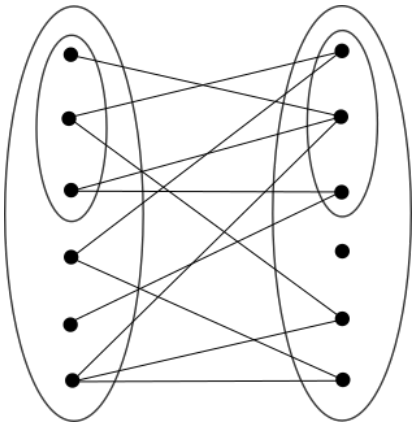


Figure 4: A random bipartite graph

when n is large we would expect the density $d(X, Y)$ to be close to $\frac{1}{2}$. We would also expect a density of subsets $X' \subset X, Y' \subset Y$ to be close to $\frac{1}{2}$ because edges between X', Y' are random with the probability $\frac{1}{2}$ too. A pair (X, Y) is an example of a ε-regular pair for a small value of ε. To make a long story short, we would say ε-regular measures how similar our pair is to a random-graph, the smaller the ε, the more randomness within the pair.

Definition 4.2 (ε-regular partition)**.** *Let G be a graph with n vertices. A partition $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ of a vertex set V of G is an ε-regular partition if $|V_0| < \varepsilon n$ and all but at most εk^2 pair (V_i, V_j) for $1 \leq i < j \leq k$ are ε-regular.*

An ε-regular partition is a partition where almost every part are pairwise ε-regular so that edges between any 2 parts are almost random.

Definition 4.3 (Equitable partition)**.** *Let G be a graph with n vertices. A partition $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ of a vertex set V of G is an equitable partition if $|V_i| = |V_j|$ for all distinct $1 \leq i, j \leq k$ and V_0 can have different size. We said V_0 is an exceptional set of P.*

The Regularity lemma states that every large enough graph can be partitioned into parts with the same size and a not too large remainder so that almost every pair of parts are ε-regular, that is edges between any part are almost random. The nice thing about this lemma is it applies to every graph!

Theorem 4.1 (Regularity lemma)**.** *For every $\varepsilon > 0$ and $m > 0$ there exist M, N that depends on ε, m such that for every graph G with $n \geq N$ vertices, there is a partition of the vertex set $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ such that*

1. *$m \leq k \leq M$*
2. *P is equitable*
3. *P is ε-regular*

The proof of this lemma is also based on density increment arguments. The idea is if our partition doesn't satisfy the condition of the lemma, we can find a refinement of that partition so that a certain index value of the refinement increased by a fixed amount,

we know that our index value is bounded so the process can't go on forever and it must stop at some point. We must be left with the partition that satisfies the condition. We will begin with a series of definitions and lemmas that will be used in the proof.

Definition 4.4 (Refinement)**.** *Let P be a partition, we say that a partition Q is a refinement of P if every part of Q is a subset of a part of P. For any partitions $P_1, \ldots, P_n$ we said P_c is a common refinement of $P_1, \ldots, P_n$ if any part of a partition P_i is a union of some parts of the P_c. The smallest common refinement is the common refinement that has minimum number of parts.*

For example, If $V = \{1, 2, 3, 4\}$ is a set of vertices. Let $P_1 : V = \{1, 2, 3\} \cup \{4\}$ and $P_2 : V = \{1\} \cup \{2, 3, 4\}$ then $\{1\} \cup \{2, 3\} \cup \{4\}$ and $\{1\} \cup \{2\} \cup \{3\} \cup \{4\}$ are common refinements of P_1 and P_2 but $\{1\} \cup \{2, 3\} \cup \{4\}$ is the smallest common refinement.

Now, we will try to define an index that could measure ε-regularity of our graph G. For vertex sets V_i, V_j, we define

$$f(V_i, V_j) = |V_i||V_j|d(V_i, V_j)^2.$$

On disjoint partitions of a vertex set $P : V = V_0 \cup V_1 \cup \ldots V_k$, $Q : W = W_0 \cup W_1 \cup \ldots W_l$, we define

$$\begin{aligned} f(P, Q) &= f(V_0 \cup V_1 \cup \ldots V_k, W_0 \cup W_1 \cup \ldots W_l) \\ &= \sum_{0 \le i \le k} \sum_{0 \le j \le l} f(V_i, W_j) \end{aligned}$$

Proposition 4.2. *For any disjoint vertex sets X, Y if we have $X = X_1 \cup X_2$ then*

$$f(X_1 \cup X_2, Y) \ge f(X, Y)$$

Proof. Because $e(X_1, Y) + e(X_2, Y) = e(X, Y)$. We apply the Cauchy-Schwarz inequality,

$$\begin{aligned} (|X_1| + |X_2|)(\frac{e(X_1, Y)^2}{|X_1|} + \frac{e(X_2, Y)^2}{|X_2|}) &\ge (e(X_1, Y) + e(X_2, Y))^2 \\ \frac{e(X_1, Y)^2}{|X_1|} + \frac{e(X_2, Y)^2}{|X_2|} &\ge \frac{e(X, Y)^2}{|X|} \\ |X_1||Y|d(X_1, Y)^2 + |X_2||Y|d(X_2, Y)^2 &\ge |X||Y|d(X, Y)^2 \\ f(X_1 \cup X_2, Y) &\ge f(X, Y) \end{aligned}$$

□

So f is not decreasing when we partitions a vertex set to smaller parts. We will show that if (X, Y) is not ε-regular, there is a partition $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ that $f(X_1 \cup X_2, Y_2 \cup Y_2) > f(X, Y)$

Proposition 4.3. *For any ε-irregular pairs X, Y, there is a partition $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ that*

$$f(X_1 \cup X_2, Y_2 \cup Y_2) \ge f(X, Y) + \varepsilon^4 |X||Y|.$$

Proof. Because (X, Y) is ε-irregular, there exists $X' \subset X, Y' \subset Y$ that $|X'| \geq \varepsilon|X|, |Y'| \geq \varepsilon|Y|$ and

$$|d(X', Y') - d(X, Y)| \geq \varepsilon.$$

Let $X_1 = X', Y_1 = Y'$ and partitions $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$. Observe that

$$\begin{aligned}\sum_{i,j=1}^{2} |X_i||Y_j|(d(X_i, Y_j) - d(X, Y))^2 &\geq |X_1||Y_1|(d(X_1, Y_1) - d(X, Y))^2 \\ &\geq \varepsilon|X| \cdot \varepsilon|Y|\varepsilon^2 \\ &\geq \varepsilon^4|X||Y|.\end{aligned}$$

On the other hand

$$\begin{aligned}&\sum_{i,j=1}^{2} |X_i||Y_j|(d(X_i, Y_j) - d(X, Y))^2 \\ &= \sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j)^2 - 2d(X, Y)\sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j) + d(X, Y)^2 \sum_{i,j=1}^{2} |X_i||Y_j| \\ &= \sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j)^2 - |X||Y|d(X, Y)^2\end{aligned}$$

The last equality is from the fact that $\sum_{i,j=1,2}^{2} |X_i||Y_i|d(X_i, Y_i) = |X||Y|d(X, Y)$. Putting inequalities together we have

$$\begin{aligned}\sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j)^2 - |X||Y|d(X, Y)^2 &\geq \varepsilon^4|X||Y| \\ \sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j)^2 &\geq |X||Y|d(X, Y)^2 + \varepsilon^4|X||Y| \\ f(X_1 \cup X_2, Y_1 \cup Y_2) &\geq f(X, Y) + \varepsilon^4|X||Y|.\end{aligned}$$

□

Definition 4.5 (Index)**.** *Let G be a graph with n vertices. We define an index of a partition $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ of a vertex set V of G as*

$$\begin{aligned}ind(P) &= f(V_1 \cup \ldots V_k, V_1 \cup \ldots V_k) \\ &= \frac{1}{n^2} \sum_{1 \leq i < j \leq k} f(V_i, V_j) \\ &= \frac{1}{n^2} \sum_{1 \leq i < j \leq k} |V_i||V_j|d(V_i, V_j)^2\end{aligned}$$

Note that, here we don't take the exceptional set V_0 into account.

Since $0 \leq d(V_i, V_j) \leq 1$,

$$\text{ind}(P) \leq \frac{1}{n^2} \sum_{1 \leq i < j \leq k} |V_i||V_j|$$

We also know that

$$2\sum_{1\le i<j\le k}|V_i||V_j| \le (\sum_{1\le i\le k}|V_i|)^2 \le n^2.$$

so

$$\text{ind}(P) \le \frac{1}{2}$$

Proposition 4.4. *Let G be a graph with n vertices. Let P be a partition of a vertex set of G and Q be a refinement of P then*

$$ind(Q) \ge ind(P)$$

Proof. The proposition (4.2) implies that for a partition P, if we split one part into two, the index value is not decrease. For any partition P, we can get a refinement Q of P from continuously splitting one part of P to two at a time. Apply the proposition (4.2) repeatedly we have $\text{ind}(Q) \ge \text{ind}(P)$. □

Recall that if a partition P is not ε-regular, there are at least εk^2 ε-irregular pairs in P. The proposition (4.3) implies that we can find a refinement of P that increases the value of index significantly.

Proposition 4.5. *Let G be a graph with n vertices. Let $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ be an equitable partition that is not ε-regular. There is a refinement Q of P into at most $k2^{k-1}+1$ parts for which*

$$ind(Q) \ge ind(P) + (1 - \frac{|V_0|}{n})^2\varepsilon^5$$

Proof. Define partitions on vertex set V_i, V_j of P for $1 \le i < j \le k$ as follows. If (V_i, V_j) is ε-irregular, there must be subsets $V_{ij} \subset V_i$ and $V_{ji} \subset V_j$ such that $|V_{ij}| \ge \varepsilon|V_i|$ and $|V_{ji}| \ge \varepsilon|V_j|$ but $|d(V_{ij}, V_{ji}) - d(V_i, V_j)| \ge \varepsilon$, we define partitions $P_{ij} : V_i = V_{ij} \cup V_i \backslash V_{ij}$, $P_{ji} : V_j = V_{ji} \cup V_j \backslash V_{ji}$. Otherwise, if (V_i, V_j) is ε-regular then P_{ij}, P_{ji} does nothing to V_i, V_j.

Let P_i be the smallest common refinement of partitions P_{ij} for $j \ne i$. Let Q be the union of the refinements of $P_1, \ldots, P_k$ and V_0. Each partition P_{ij} partition V_i into at most 2 parts so P_i partitions V_i into at most 2^{k-1} parts. This means Q has at most $1 + k2^{k-1}$ parts.

For any V_i, V_j that is not ε-regular, apply the proposition (4.4), (4.3) on $V_i = V_{ij} \cup V_i \backslash V_{ij}$ and $V_j = V_{ji} \cup V_j \backslash V_{ji}$

$$\begin{aligned} f(V_i, V_j) + \varepsilon^4|V_i||V_j| &\le f(V_{ij} \cup V_i \backslash V_{ij}, V_{ji} \cup V_j \backslash V_{ji}) \\ &= f(P_{ij}, P_{ji}) \\ &\le f(P_i, P_j) \end{aligned}$$

The second inequality is from P_i, P_j is a refinement of P_{ij}, P_{ji} respectively. Because $Q = V_0 \cup P_1 \cup \cdots \cup P_k$ and there are at least εk^2 ε-irregular pairs, applying the inequality above we have

$$\begin{aligned}\mathrm{ind}(Q) &= \frac{1}{n^2}\sum_{1\le i<j\le k} f(P_i,P_j)\\ &\ge \frac{1}{n^2}\sum_{1\le i<j\le k} f(V_i,V_j) + \varepsilon k^2\frac{|V_i|^2}{n^2}\varepsilon^4\\ &= \mathrm{ind}(P) + (1-\frac{|V_0|}{n})^2\varepsilon^5\end{aligned}$$

The last equality follows from P is an equitable partition so each V_i has cardinality $\frac{n-|V_0|}{k}$. Therefore, Q is the desired partition. □

Note that the refinement in the proposition (4.5) may not be an equitable partition. We will show that we could re-partition Q into an equitable partition P' without losing much information on the index value.

Proposition 4.6. *Let G be a graph with n vertices. Let $Q : V = V_0 \cup V_1 \cup \cdots \cup V_l$ be a vertex partition. There is an equitable partition P' of V with at most $t+1$ parts such that*

1. *$ind(P') \ge ind(Q) - \frac{l}{t}$*
2. *The exceptional set of P' has size $|W_0| < |V_0| + \frac{ln}{t}$*

Proof. Let $n_0 = n - |V_0|$ be the number of vertices not in V_0. We arbitrarily partition every part of Q except V_0 into parts of size $\lceil \frac{n_0}{t} \rceil$. For $1 \le i \le l$, we write $W_i = W_i^{(1)} \cup \cdots \cup W_i^{(k_i)} \cup W_i^*$ where $W_i^{(1)}, \ldots, W_i^{(k_i)}$ are parts that have equal size $\lceil \frac{n_0}{t} \rceil$ and W_i^* is the remainder. This partition is a refinement of Q and we call it Q'. Define a partition P' to be the same as Q' but put together every remainder W_i^* into the set V_0, we called the new exceptional set W_0. Because we partition $V \backslash V_0$ into parts with size

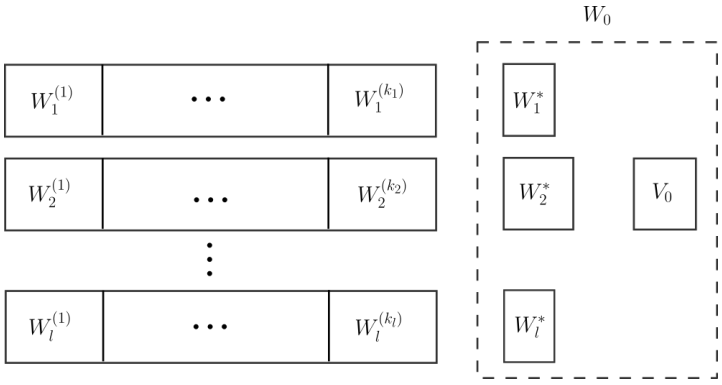


Figure 5: each part of a partition P'

$\lceil \frac{n_0}{t} \rceil$, P' has at most $t+1$ parts. Because the size of each remainder $|W_i^*|$ can't be larger than $\lceil \frac{n_0}{t} \rceil - 1$, we have

$$|W_0| \le |V_0| + l(\lceil \frac{n_0}{t} \rceil - 1) \le |V_0| + \frac{ln}{t}.$$

The difference between the index of Q' and the index of P' are terms that only involve the remainders W_i^* so

$$\mathrm{ind}(Q') - \mathrm{ind}(P') = \frac{1}{n^2} \sum_{1 \le i,j \le l} f(W_i^*, W_j^{(1)} \cup W_j^{(2)} \cup \cdots \cup W_j^{(k_j)})$$

Recall that $f(V_i, V_j) = |V_i||V_j|d(V_i, V_j)^2 \le |V_i||V_j|$, we have

$$\begin{aligned}
\mathrm{ind}(Q') - \mathrm{ind}(P') &\le \frac{1}{n^2} \sum_{1 \le i,j \le l} |W_i^*||W_j^{(1)} \cup W_j^{(2)} \cup \cdots \cup W_j^{(k_j)}| \\
&= \frac{1}{n^2} \sum_{1 \le i,j \le l} |W_i^*||W_j \backslash W_j^*| \\
&= \frac{1}{n^2} (\sum_{1 \le i \le l} |W_i^*|)(\sum_{1 \le j \le l} |W_j \backslash W_j^*|) \\
&= \frac{n_0}{n^2} (\sum_{1 \le i \le l} |W_i^*|) \\
&\le \frac{l(\lceil \frac{n_0}{t} \rceil - 1)}{n^2} n_0 \\
&\le \frac{l}{t}
\end{aligned}$$

The second inequality follows from W_i^* is a remainder so $|W_i^*| \le \lceil \frac{n_0}{t} \rceil - 1$. Because Q' is a refinement of Q, apply the proposition (4.4),

$$\mathrm{ind}(P') \ge \mathrm{ind}(Q') - \frac{l}{t} \ge \mathrm{ind}(Q) - \frac{l}{t}.$$

Thus, P' is the desired partition. □

Now, we are ready to prove the regularity lemma. Since our choice of partition will depend on the value of ε and a large ε doesn't give us meaningful information about the graph, for simplicity we will prove the case when $\varepsilon \le 1/5$.

Proof of the regularity lemma: Let G be a graph with n vertices and $\varepsilon \le 1/5$. First, we begin with an arbitrary equitable partition of vertices of G, $P : V = V_0 \cup V_1 \cup \cdots \cup V_k$ for k small enough so that $|V_0| < k < \frac{\varepsilon n}{2}$. Applying the proposition (4.5), there is a refinement $Q^1 : V = V_0 \cup V_1 \cup \cdots \cup V_l$ of P into at most $k2^{k-1} + 1$ parts for which $\mathrm{ind}(Q^1) \ge \mathrm{ind}(P) + (1 - \frac{|V_0|}{n})^2 \varepsilon^5$. When $\varepsilon \le 1/5$, we have

$$(1 - \varepsilon)^2 \ge 1 - 2\varepsilon \ge \frac{1}{2} + \frac{\varepsilon}{2}.$$

Because $|V_0| < k < \frac{\varepsilon n}{2}$,

$$(1 - \frac{|V_0|}{n})^2 \ge (1 - \varepsilon)^2 \ge \frac{1}{2} + \frac{\varepsilon}{2}.$$

Therefore,

$$\begin{aligned}
\mathrm{ind}(Q^1) &\ge \mathrm{ind}(P) + (1 - \frac{|V_0|}{n})^2 \varepsilon^5 \\
&\ge \mathrm{ind}(P) + (\frac{1}{2} + \frac{\varepsilon}{2}) \varepsilon^5
\end{aligned}$$

Next, apply the proposition (4.6), there is an equitable partition $P^1 : V = W_0 \cup W_1 \cup \cdots \cup W_t$ such that $\text{ind}(P^1) \geq \text{ind}(Q^1) - \frac{l}{t}$ and $|W_0| < |V_0| + \frac{ln}{t}$. We choose $t = \lceil 2l\varepsilon^{-6} \rceil$ so that $t \geq 2l\varepsilon^{-6}$ and $\frac{l}{t} \leq \frac{\varepsilon^6}{2}$. We get an equitable partition P^1 such that

$$\begin{aligned}\text{ind}(P^1) &\geq \text{ind}(Q^1) - \frac{l}{t} \\ &\geq \text{ind}(P) + (\frac{1}{2} + \frac{\varepsilon}{2})\varepsilon^5 - \frac{l}{t} \\ &= \text{ind}(P) + \frac{1}{2}\varepsilon^5.\end{aligned}$$

and

$$\begin{aligned}|W_0| &\leq |V_0| + \frac{nl}{t} \\ &\leq |V_0| + \frac{n\varepsilon^6}{2}\end{aligned}$$

We apply the proposition (4.5),(4.6) again and again. We can see that after each round, the index of the new partition increased by $\varepsilon^5/2$. Since the index is bounded by $1/2$, the process can be done at most ε^{-5} times. Moreover, after at most ε^{-5} rounds, the exceptional set of the partition has size most $|V_0| + \varepsilon^{-5}(\varepsilon^6 n)/2 < \varepsilon n$. We also know that an index increment of at least $\varepsilon^5/2$ is valid throughout the process because we know that our exceptional set after k steps,

$$|W_k| < \varepsilon n$$

so that

$$(1 - \frac{|W_k|}{n})^2 \geq (1 - \varepsilon)^2 \geq \frac{1}{2} + \frac{\varepsilon}{2}$$

Therefore, when the process is finished, we are left with an equitable, ε-regular partitions of vertices of G. This proves the regularity lemma. Note that each step, the number of parts increases from k to at most $k2^k\varepsilon^{-6}$ which is called a tower-type bound. The density in this chapter is increased by a fixed amount while in the chapter 3, the increment depends on the value of the density at each step.

□

4.3 The triangle removal lemma

For a vertex partition of a graph G, $P : V = P_1 \cup \cdots \cup P_k$, we could zoom out and see each part of the partition as a vertex P_i' of a new graph G' where a vertex P_i' is adjacent to P_j' in G' if and only if the corresponding part (P_i, P_j) is ε-regular. G' can give us some information about G. For example, if there is a triangle (P_i', P_j', P_k') in G', pairs $(P_i, P_j), (P_j, P_k), (P_k, P_i)$ must be ε-regular which means edges between them are almost random. If densities $d(P_i, P_j), d(P_j, P_k), d(P_k, P_i)$ are not too small and we are not unlucky, there must be a triangle $V_iV_jV_k$ in G such that $V_i \in P_i, V_j \in P_j, V_k \in P_k$. The next proposition counts the number of triangles we can guarantee in G if we have a triangle in G'. The idea in the proof is adapted from Stephanie Bell and Will Grodzicki [BG].

Proposition 4.7 (Triangle counting lemma)**.** *Let X, Y, Z be a graph. Suppose that $(X, Y), (Y, Z), (Z, X)$ are ε-regular pairs with density at least 2ε then the number of triangles formed by $X \times Y \times Z$ is at least*

$$(1 - 2\varepsilon)\varepsilon^3|X||Y||Z|.$$

Proof. For any $x \in X$, let Y_x, Z_x be sets of neighbors of x in Y, Z respectively. We know that an edge between Y_x and Z_x generates a triangle in $X \times Y \times Z$. If $|Y_x| \geq \varepsilon|Y|$ and $|Z_x| \geq \varepsilon|Z|$, ε-regularity between Y, Z implies

$$\begin{aligned}
&|d(Y_x, Z_x) - d(Y, Z)| < \varepsilon \\
&d(Y_x, Z_x) > d(Y, Z) - \varepsilon > \varepsilon \\
&e(Y_x, Z_x) > \varepsilon|Y_x||Z_x| > \varepsilon^3|Y||Z|
\end{aligned}$$

There are at least $\varepsilon^3|Y||Z|$ triangles in $X \times Y \times Z$ that contain the vertex x.

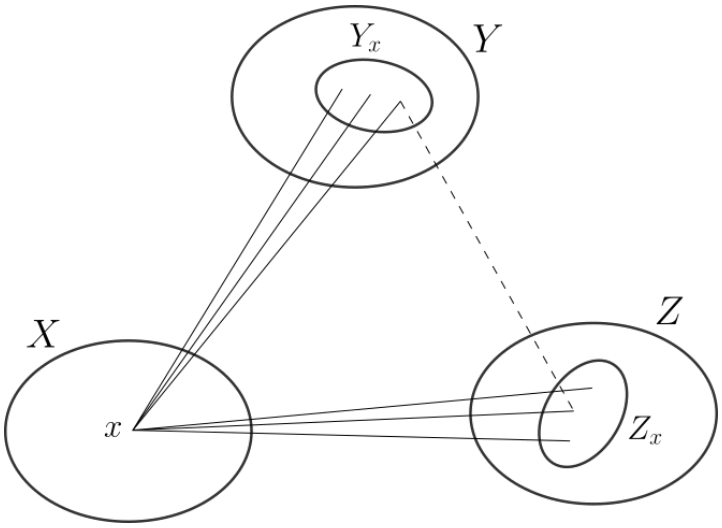


Figure 6: An edge between Y_x, Z_x generates a triangle

Now, we will find the number of x such that $|Y_x|, |Z_x|$ are large enough. Let $A := \{x \in X : |Y_x| < \varepsilon|Y|\}$ be the subset of X such that each element has number of neighbors in Y less than $\varepsilon|Y|$. If $|A| \geq \varepsilon|X|$ then by ε-regularity between X, Y

$$\begin{aligned}
&|d(A, Y) - d(X, Y)| < \varepsilon \\
&d(A, Y) > d(X, Y) - \varepsilon > \varepsilon
\end{aligned}$$

But

$$d(A, Y) = \frac{e(A, Y)}{|A||Y|} = \frac{\sum_{x \in A} |Y_x|}{|A||Y|}, < \frac{\varepsilon|A||Y|}{|A||Y|} = \varepsilon$$

This leads to contradiction so we have, $|A| < \varepsilon|X|$. Similarly, let $B := \{x \in X : |Z_x| < \varepsilon|Z|\}$ then $|B| < \varepsilon|X|$. Therefore, there are at least $|X| - |A| - |B| \geq (1 - 2\varepsilon)|X|$, x in X such that Y_x, Z_x have size greater than $\varepsilon|Y|, \varepsilon|Z|$. Each x generates at least $\varepsilon^3|Y||Z|$ triangles in $X \times Y \times Z$. Thus there are at least $(1 - 2\varepsilon)\varepsilon^3|X||Y||Z|$ triangles in $X \times Y \times Z$. □

The idea in the proposition above doesn't only work for a triangle but it also works for any graph. If there is a square in G', it's likely that there is a square in G too. Now,

we are ready to prove the triangle removal lemma.

Proof of the triangle removal lemma: Let G be a graph with n vertices. If n is large enough, we can apply the regularity lemma to G with $\frac{\varepsilon}{4}$, there is an equitable partition $P : V = V_0 \cup V_1 \cdots \cup V_k$ of the set of vertices of G such that $k \geq \frac{2}{\varepsilon}$ and P is $\frac{\varepsilon}{4}$ -regular. We aim to remove edges from G so that we are left with edges between a $\frac{\varepsilon}{4}$ -regular pair that has a density at least $\frac{\varepsilon}{2}$ so that we could apply the triangle counting lemma. Remove edges from G as follows

- remove edges incident with V_0
- remove edges inside V_i for $1 \leq i \leq k$
- remove edges between pairs (V_i, V_j) that are not $\frac{\varepsilon}{4}$ -regular
- remove edges between pairs (V_i, V_j) with density less than $\frac{\varepsilon}{2}$

With $n = |V_0| + k|V_i|$ in mind, we calculate how many edges we could remove. There are $|V_0|$ vertices in V_0 so there are at most $|V_0|n < \frac{\varepsilon n^2}{4}$ edges incident with V_0 (P is $\frac{\varepsilon}{4}$ -regular so $|V_0| < \frac{\varepsilon n}{4}$). There are at most $\binom{|V_i|}{2}$ edges inside V_i so there are at most $k\binom{|V_i|}{2} < \frac{k|V_i|^2}{2} < \frac{n^2}{2k} < \frac{\varepsilon n^2}{4}$ edges inside V_i. Because P is $\frac{\varepsilon}{4}$ -regular, there are at most $\frac{\varepsilon k^2}{4}$ $\frac{\varepsilon}{4}$-irregular pairs (V_i, V_j), each has at most $|V_i|^2$ edges between them so there are at most $\frac{\varepsilon}{4}k^2|V_i|^2 < \frac{\varepsilon n^2}{4}$ edges between $\frac{\varepsilon}{4}$ -irregular pairs. Last, there are at most $\binom{k}{2}$ pairs of (V_i, V_j) with density less than $\frac{\varepsilon}{2}$ and each pair has at most $\frac{\varepsilon}{2}|V_i|^2$ edges between them so there are at most $\binom{k}{2}\frac{\varepsilon}{2}|V_i|^2 < \frac{k^2|V_i|\varepsilon}{4} < \frac{\varepsilon n^2}{4}$ edges in the last category.

The total number of edges we may remove is at most εn^2 edges. If there is still a triangle left in the graph, It must come from a triple V_a, V_b, V_c where each pair $(V_a, V_b), (V_b, V_c), (V_c, V_a)$ is $\frac{\varepsilon}{4}$ -regular with density at least $\frac{\varepsilon}{2}$. From the triangle counting lemma, there are at least

$$\begin{aligned}(1 - 2(\frac{\varepsilon}{4}))(\frac{\varepsilon}{4})^3|V_a||V_b||V_c| &= (1 - \frac{\varepsilon}{2})(\frac{\varepsilon}{4})^3(\frac{n - |V_0|}{k})^3 \\ &\geq (1 - \frac{\varepsilon}{2})(\frac{\varepsilon}{4})^3(\frac{n - \frac{\varepsilon n}{4}}{k})^3 \\ &= (1 - \frac{\varepsilon}{2})(\frac{4\varepsilon - \varepsilon^2}{16k})^3 n^3\end{aligned}$$

triangles left in $V_a \times V_b \times V_c$. If we have δn^3 triangles at the start, where $\delta < (1 - \frac{\varepsilon}{2})(\frac{4\varepsilon - \varepsilon^2}{16k})^3$. After removing at most εn^2 edges as above, we can't have any triangle left. This proves the triangle removal lemma. □

5 Finite field analogue of the problem

We can extend the notion of Roth's theorem to a finite field. The key difference between a finite field and integers is that the finite field is closed under addition, subtraction, multiplication and division. This affects the scope of a 3-term arithmetic progression. For example, in $\mathbb{Z}/5\mathbb{Z}$, $(1, 4, 7) = (1, 4, 2)$ is a 3-term arithmetic progression while $(1, 4, 2)$ is not a 3-term arithmetic progression in $\{1, 2, 3, 4, 5\}$.

Among all of finite fields, $(\mathbb{Z}/3\mathbb{Z})^n$ is the most popular one. We are interested in how large a subset $A \subset (\mathbb{Z}/3\mathbb{Z})^n$ can be so that it does not contain a 3-term arithmetic progression. The problem is also known as the cap set problem. Brown and Buhler [BB82] are the first people who explored this problem and found that the size of A is $o(3^n)$ as n grows. Meshulam used fourier analysis arguements similar to Roth's to achieve an upper bound of $\frac{2\cdot 3^n}{n}$ in [Mes95] . There was no improvement of the bound for almost 20 years until Bateman and Katz in [BK12] as $O(3^n/n^{1+\varepsilon})$ with $\varepsilon > 0$ which was considered to be a breakthrough.

On the other hand, it's clear that 2^n is a lower bound for the size of A because a set $\{(x_1, \ldots, x_n) : x_i \in \{0, 1\}\}$ does not contain a a 3-term arithmetic progression. Moreover, Edel has shown in [Ede04] that we can construct a subset up to a size 2.2^n that does not contain a 3-term arithmetic progression. The ultimate goal of this problem is to find whether we could improve the upper bound up to the form $(3 - c)^n$ for some constant c which is called an exponential bound. Croot-Lev-Pach [CLP17] achieved a breakthrough by using a polynomial method on a relating group $(\mathbb{Z}/4\mathbb{Z})^n$. A few weeks later, Ellenberg and Gijswijt [EG17] generalised the lemma from [CLP17] to work with finite fields $(\mathbb{Z}/q\mathbb{Z})^n$ which gave us an upper bound for the cap set problem of the form $(3 - c)^n$.

5.1 Meshulam's proof

In fact, Meshulam has found an upper bound for a size of a subset A that does not contain a 3-term arithmetic progression of any finite abelian group G. However, we want to focus on presenting simple idea of proofs so we will contract his idea to find the bound for a finite field $G_n = (\mathbb{Z}/3\mathbb{Z})^n$. As mentioned that Meshulam used a similar Fourier analysis idea to chapter 3, it may be implicit in his paper [Mes95] that they are similar. We will rewrite his proof in the same pattern as in the chapter 3 and make comparison along the way.

Theorem 5.1 (Meshulam)**.** *Let $G_n = (\mathbb{Z}/3\mathbb{Z})^n$ be an abelian group. Let $d(n)$ be the maximum density of a subset $A \subset G_n$ that does not contain a 3-term arithmetic progression. We have*

$$d(n) \leq \frac{2}{n}.$$

Proof. We will prove this theorem by an induction. It's clear that the base case when $n = 1$ is true. Now, assume that $d(n-1) \leq \frac{2}{n-1}$ for an integer $n \geq 2$. Let A be a subset of G_n that does not contain a 3-term arithmetic progression. Define a balance function

$$f_A := 1_A - d(n-1)1_{G_n}.$$

We will show that the lack of 3-term arithmetic progression of A implies a large Fourier coefficient of this function.

Proposition 5.1. *Suppose that $A \subset G_n$ does not contain a 3-term arithmetic progression then there is a character $\chi \in \hat{G}_n$ such that*

$$|\hat{f}_A(\chi)| \geq d(n-1)|A| - 1$$

Proof. Consider a similar trilinear T on a function $f_i : G_n \to \mathbb{C}$

$$T(f_1, f_2, f_3) = \sum_{x,d \in G_n} f_1(x) f_2(x+d) f_3(x+2d).$$

$T(1_A, 1_A, 1_A)$ counts the number of 3-term arithmetic progression in A so we have

$$T(1_A, 1_A, 1_A) = |A| \tag{5}$$

and also

$$T(1_A, 1_{G_n}, 1_A) = \sum_{x,d \in G_n} 1_A(x) 1_{G_n}(x+d) 1_A(x+2d)$$

This counts the number of $a, b \in A$ that $\frac{a+b}{2} \in G_n$. Because G_n is a group of odd order, $\frac{1}{2}$ is well-defined in G_n so that for any $a, b \in A$, we have $\frac{a+b}{2} \in G_n$. We have

$$T(1_A, 1_{G_n}, 1_A) = |A|^2 \tag{6}$$

From (5),(6)

$$T(1_A, f_A, 1_A) = |A| - d(n-1)|A|^2 \tag{7}$$

If $|A|$ is large enough then $T(1_A, f_A, 1_A)$ is negative. The equation here is an equality compared to the inequality in the chapter 3. Next, we know that if χ is a character of G then χ^{-2} is well-defined and is a character of G as

$$\chi^{-2}(x) = \chi(-2x) = \chi(x)$$

The last equality holds because we are working on $(\mathbb{Z}/3\mathbb{Z})^n$. Now, although the definition of a Fourier transform is different in G_n to a Fourier transform in $\mathbb{Z}$, we still have a key observation that

$$\begin{aligned}
\sum_{\chi \in \hat{G}_n} \hat{f}_1(\chi)\hat{f}_2(\chi)\hat{f}_3(\chi) &= \sum_{\chi \in \hat{G}_n} \hat{f}_1(\chi)\hat{f}_2(\chi^{-2})\hat{f}_3(\chi) \\
&= \sum_{\chi \in \hat{G}_n} \sum_{n,m,l \in G_n} f_1(n) f_2(m) f_3(l) \chi(2m-n-l) \\
&= \sum_{n,m,l \in G_n} f_1(n) f_2(m) f_3(l) \Big(\sum_{\chi \in \hat{G}_n} \chi(2m-n-l) \Big)
\end{aligned}$$

From the proposition (2.1), we know that

$$\sum_{\chi \in \hat{G}_n} \chi(2m-n-l) = 0$$

if $2m-n-l \neq 0$ and equals to $|G_n|$ if (n, m, l) is a 3-term arithmetic progression.

$$\begin{aligned}
\sum_{\chi \in \hat{G}_n} \hat{f}_1(\chi)\hat{f}_2(\chi)\hat{f}_3(\chi) &= |G_n| \sum_{n+l=2m} f_1(n) f_2(m) f_3(l) \\
&= |G_n| \sum_{x,d \in G_n} f_1(x) f_2(x+d) f_3(x+2d) \\
&= |G_n| T(f_1, f_2, f_3)
\end{aligned}$$

Apply this result with (7), we have

$$\begin{aligned} d(n-1)|A|^2 - |A| &\leq |T((1_A, f_A, 1_A)| \\ &= \frac{1}{|G_n|} |\sum_{\chi \in \hat{G}_n} \hat{1}_A(\chi)\hat{f}_A(\chi)\hat{1}_A(\chi)| \\ &\leq \frac{1}{|G_n|} \sup_{\chi \in \hat{G}_n} |\hat{f}_A(\chi)| \sum_{\chi \in \hat{G}_n} |\hat{1}_A(\chi)| \end{aligned}$$

By the Parseval's identity

$$\frac{1}{|G_n|} \sum_{\chi \in \hat{G}_n} |1_A(\chi)^2| = \sum_{x \in G_n} |1_A(x)|^2 = |A|$$

Substituting back to the inequality, we have

$$\sup_{\chi \in \hat{G}_n} |\hat{f}_A(\chi)| \geq d(n-1)|A| - 1$$

□

The proposition (5.1) is similar to the proposition (3.1) in the chapter 3 but with a stronger bound thanks to the exploitation of properties of G_n. Next, we will find an appropriate partition P_i of G_n that for each i, $\text{diam}_{P_i}(\chi(x))$ is small. Surprisingly, there is a partition such that values of diameter in each part are zero!

Proposition 5.2. *Let $G_n = (\mathbb{Z}/3\mathbb{Z})^n$. For any non-trivial character $\chi \in \hat{G}_n$, let $W_\chi = ker(\chi) = \{x \in G_n : \chi(x) = 1\}$ be the kernel of χ. Then*

$$W_\chi \cong (\mathbb{Z}/3\mathbb{Z})^{n-1}$$

Proof. Recall that any character of $(\mathbb{Z}/3\mathbb{Z})^n$ could be written in the form

$$\chi(x) : x \to e(\frac{1}{3}(r_1x_1 + \dots r_nx_n))$$

where $r = (r_1, \dots, r_n) \in (\mathbb{Z}/3\mathbb{Z})^n$. If $\chi \neq \chi_0$, consider the corresponding r of χ, there exists i that $r_i \neq 0$. Without loss of generality, let $r_1 \neq 0$ Define a map $\phi : W_\chi \to (\mathbb{Z}/3\mathbb{Z})^{n-1}$ as

$$\phi((x_1, x_2, \dots, x_n)) = (x_2, x_3, \dots, x_n).$$

The map is clearly a homomorphism. Next, let $x, y \in W_\chi$ that $\phi(x) = \phi(y)$. Because x, y are in the kernel W_χ,

$$\chi(x) = \chi(y) = 1$$
$$e(\frac{1}{3}(x_1r_1 + \dots x_nr_n)) = e(\frac{1}{3}(y_1r_1 + \dots y_nr_n))$$

But $\phi(x) = \phi(y)$ implies $x_2r_2 + \dots + x_nr_n = y_2r_2 + \dots + y_nr_n$ we have

$$e(\frac{1}{3}(x_1 - y_1)r_1) = 1$$

Since $r_1 = 0$ and $x_1, y_1 \in \{0, 1, 2\}$, we have $x_1 = y_1$ which implies $x = y$. Therefore, ϕ is injective.

On the other hand, for any $(x_2, x_3, \ldots, x_n) \in (\mathbb{Z}/3\mathbb{Z})^{n-1}$, $r_1 \neq 0$ implies that there must be $x_1 \in \{0, 1, 2\}$ such that $r^T x = x_1 r_1 + \cdots + x_n r_n$ is divisible by 3 so $e(\frac{1}{3} r^T x) = 1$ and $(x_1, \ldots, x_n) \in W_\chi$. Thus ϕ is surjective. We can conclude that W_χ is isomorphic to $(\mathbb{Z}/3\mathbb{Z})^{n-1}$.

□

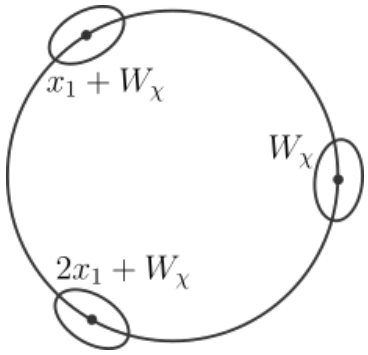


Figure 7: image of $W_\chi, x_1 + W_\chi, 2x_1 + W_\chi$ after applying χ on a unit circle

If χ is a non-trivial character, there must be $x_1 \in G_n$ such that $\chi(x_1) \neq 0$. Consider sets $W_\chi, x_1 + W_\chi, 2x_1 + W_\chi$. It is clear that they are disjoint and from the proposition (5.2) each set has size 3^{n-1}. Therefore, $W_\chi, x_1 + W_\chi, 2x_1 + W_\chi$ must be a partition of G_n. Moreover, after applying χ, each part of the partition have only one value which equals to $1, \chi(x_1), \chi(2x_1)$ respectively. We have

$$\mathrm{diam}_{W_\chi}(\chi(x)) = \mathrm{diam}_{x_1 + W_\chi}(\chi(x)) = \mathrm{diam}_{2x_1 + W_\chi}(\chi(x)) = 0.$$

We could compare this with the proposition (3.3) where we try to partition $[N]$ into parts where $e(\theta x)$ has a small diameter on them. We will use this partition to find an upper bound for $\hat{f}_A(\chi)$.

Proposition 5.3. *Suppose that $A \subset G_n$ does not contain a 3-term arithmetic progression. For any $\chi \in \hat{G}_n$ we have*

$$|\hat{f}_A(\chi)| \leq d(n-1)|G_n| - |A|$$

Proof. For any non-trivial χ, there exist a partition $W_\chi, x_1 + W_\chi, 2x_1 + W_\chi$ as above,

for convenience we write P_1, P_2, P_3 for the partition respectively. We have

$$\begin{aligned}
|\hat{f}_A(\chi)| &= |\sum_{x\in G_n} f_A(x)\chi(-x)| \\
&\leq |\sum_{x\in P_1} f_A(x)| + |\sum_{x\in P_2} f_A(x)||\chi(-x_1)| + |\sum_{x\in P_3} f_A(x)||\chi(-2x_1)| \\
&= |\sum_{x\in P_1} f_A(x)| + |\sum_{x\in P_2} f_A(x)| + |\sum_{x\in P_3} f_A(x)| \\
&= \sum_{i=1}^{3} ||P_i \cap A| - d(n-1)3^{n-1}|
\end{aligned}$$

Since A does not contain a 3-term arithmetic progression and P_i is isomorphic to G_{n-1}, we could see $P_i \cap A$ as a subset of G_{n-1} that does not contain a 3-term arithmetic progression. The size of $P_i \cap A$ must be less than or equal to $d(n-1)3^{n-1}$. Thus, each term in the modulus in the inequality above is negative so that

$$\begin{aligned}
|\hat{f}_A(\chi)| &\leq d(n-1)3^n - (|P_1 \cap A| + |P_2 \cap A| + |P_3 \cap A|) \\
&= d(n-1)3^n - |A| \\
&= d(n-1)|G_n| - |A|
\end{aligned}$$

□

Now, we are ready to prove the inductive step. By the proposition (5.1) and proposition (5.3)

$$\begin{aligned}
d(n-1)|A| - 1 &\leq d(n-1)|G_n| - |A| \\
|A|(d(n-1)+1) &\leq 1 + d(n-1)|G_n| \\
\frac{|A|}{|G_n|} &\leq \frac{|G_n|^{-1} + d(n-1)}{1 + d(n-1)} \\
d_{G_n}(A) &\leq \frac{3^{-n} + d(n-1)}{1 + d(n-1)}
\end{aligned}$$

By the induction hypothesis, $d(n-1) \leq \frac{2}{n-1}$ so

$$\begin{aligned}
d_{G_n}(A) &\leq \frac{3^{-n} + \frac{2}{n-1}}{1 + \frac{2}{n-1}} \\
&= \frac{3^{-n}(n-1) + 2}{n+1} \\
&\leq \frac{2}{n}
\end{aligned}$$

Because $d(n)$ is the maximum density of a subset $A \subset G_n$ that does not contain a 3-term arithmetic progression, we have

$$d(n) \leq \frac{2}{n}$$

This proves the inductive step and hence proves the theorem.

□

5.2 Polynomial Method

We have presented many ideas on the proof of Roth's theorem, some of them are based on the Fourier analysis, one is based on the graph-theoretic method but all of them use a density increment technique. This section provides an alternative approach to the problem that is not a density increment argument. We will use a polynomial method. The idea is first appeared in [CLP17] and was generalised in [EG17].

We begin with notations. Let $\mathbb{F}_q$ be a finite field and let n be a positive integer. Let M_n be the set of monomials in $x_1, \dots, x_n$ whose degree in each variable is at most $q-1$. For example, if $q = 3$ then $M_2 = \{1, x_1, x_2, x_1x_2, x_1^2, x_2^2\}$. Let P_n be the space of polynomials in $x_1, \dots, x_n$ whose degree in each variable is at most $q-1$ on a field $\mathbb{F}_q$ so M_n is a basis for P_n. Let M_n^d be the set of monomials in M_n with degree at most d and P_n^d be the subset of P_n spanned by M_n^d. With the same example, if $q = 3$ then $M_2^1 = \{1, x_1, x_2\}$. Roughly speaking, P_n^d is a space of polynomials of degree at most d. Write $|M_n^d| = m_d$ for the dimension of P_n^d.

Proposition 5.4 (Croot-Lev-Pach lemma)**.** *Let $\mathbb{F}_q$ be a finite field and let A be a subset of $\mathbb{F}_q$. Suppose a polynomial $P \in P_n^d$ satisfies $P(a-b) = 0$ for every pair a, b of distinct elements of A and $|A| \geq 2m_{d/2}$. Then $P(0) = 0$.*

Proposition 5.5 (Generalisation of the Croot-Lev-Pach lemma)**.** *Let $\mathbb{F}_q$ be a finite field and let A be a subset of $\mathbb{F}_q$. Let α, β be elements of $\mathbb{F}_q$. Suppose a polynomial $P \in P_n^d$ satisfies $P(\alpha a + \beta b) = 0$ for every pair a, b of distinct elements of A. Then the number of $a \in A$ for which $P((\alpha+\beta)a) \neq 0$ is at most $2m_{d/2}$.*

Proof. For any $P \in P_n^d$, we can write it as a linear combination of monomials of degree at most d,

$$\begin{aligned} P(\alpha x + \beta y) &= \sum_{m \in M_n^d} C_m m(\alpha x + \beta y) \\ &= \sum_{m, m' \in M_n^d, \deg(mm') \leq d} C_{m,m'} m(x) m'(y), \end{aligned}$$

where $C_{m,m'}$ are constants, $m(x)$ and $m(y)$ are monomials on x and y respectively. Because the total degree of P is at most d, at least one of m and m' has degree at most $d/2$. We can write (not necessarily unique)

$$\begin{aligned} P(\alpha x + \beta y) &= \sum_{\deg(m) \leq d/2} C_{m,m'} m(x) m'(y) + \sum_{\deg(m') \leq d/2} C_{m,m'} m(x) m'(y) \\ &= \sum_{m \in M_n^{d/2}} m(x) F_m(y) + \sum_{m' \in M_n^{d/2}} m'(y) G_{m'}(x) \\ &= \sum_{m \in M_n^{d/2}} m(x) F_m(y) + \sum_{m \in M_n^{d/2}} m(y) G_m(x) \end{aligned}$$

for some families of polynomials F_m, G_m indexed by $m \in M_n^d$.

We write $A = \{a_1, \dots, a_k\}$ for some finite k. Let B be a $k \times k$ matrix whose i, j entry is $P(\alpha a_i + \beta a_j)$, then

$$B_{ij} = \sum_{m \in M_n^{d/2}} m(a_i) F_m(a_j) + \sum_{m \in M_n^{d/2}} m(a_j) G_m(a_i).$$

We can write B as a sum of matrices U_m with i,j entry $m(a_i)F_m(a_j)$ and V_m with i,j entry $m(a_j)G_m(a_i)$

$$B = \sum_{m \in M_n^{d/2}} U_m + \sum_{m \in M_n^{d/2}} V_m$$

When

$$U_m = (m(a_1), \ldots, m(a_k))^T (F_m(a_1), \ldots, F_m(a_k)) \in \mathbb{F}_q^{k \times k}.$$
$$V_m = (G_m(a_1), \ldots, G_m(a_k))^T (m(a_1), \ldots, m(a_k)) \in \mathbb{F}_q^{k \times k}.$$

Recall from linear algebra that the rank of a matrix is the dimension of the vector space generated by its columns. It's clear that U_m, V_m has rank 1 because each column of the matrix is a factor of $(m(a_1), \ldots, m(a_k))^T$ and $(G_m(a_1), \ldots, G_m(a_k))^T$ respectively. Because there are at most $m_{d/2}$ monomials with degree at most $d/2$, B can be written as a sum of at most $2m_{d/2}$ matrices with rank 1. Since for matrices U, V with the same dimension

$$\text{rank}(U+V) \leq \text{rank}(U) + \text{rank}(V)$$

We have that the rank of B is at most $2m_{d/2}$.

On the other hand, our hypothesis on P such that $P(\alpha a + \beta b) = 0$ for every pair a, b of distinct elements of A implies that B is a diagonal matrix. In fact, a rank of a diagonal matrix is the number of non-zero diagonal entries so at most $2m_{d/2}$ diagonal entries of B are non zero. Since a diagonal entry is in the form $P(\alpha a + \beta a)$. The number of $a \in A$ for which $P((\alpha+\beta)a) \neq 0$ is at most $2m_{d/2}$.

□

The Croot-Lev-Pach lemma is a special case of the proposition (5.5) when $(\alpha, \beta) = (1, -1)$.

Theorem 5.2 (Ellenberg-Gijswijt)**.** *Let A be a subset of $(\mathbb{Z}/3\mathbb{Z})^n$ such that A does not contain a 3-term arithmetic progression. Let m_d be the number of monomials in $x_1, \ldots, x_n$ with total degree at most d and in which each variable appear with the degree at most 2. Then*

$$|A| \leq 3m_{2n/3}$$

Proof. Write $S(A)$ for the set of all element of $(\mathbb{Z}/3\mathbb{Z})^n$ of the form $a+b$, with a, b distinct elements of A. We note that (a, b, c) is a 3-term arithmetic progression if and only if $a + c = 2b$ and the progression is not trivial if $a \neq c$. Because A does not contain a non-trivial 3-term arithmetic progression, $S(A)$ and $2A$ must be disjoint. This implies $S(A) \subset (\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$. We will consider a space of polynomials $V \subset P_n^d$ that vanishes on $(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$. Note that a polynomial P that vanishes on $(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$ also vanishes on $S(A)$ too.

Now, we will use the proposition (5.5) to find an upper bound of the dimension of V. Consider an element $P^* \in V$ that has maximum support. Let $\Sigma := \{a \in (\mathbb{Z}/3\mathbb{Z})^n : P^*(a) \neq 0\}$ be the support of P^*. Because P^* vanishes on $(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$, the support $\Sigma \subset 2A$. We will show that $\dim(V) \leq |\Sigma|$ by a contradiction. Suppose that $\dim(V) > |\Sigma|$. Let $\dim(V) = k$ and $M_1, \ldots, M_k$ be a basis for V. Any element $Q \in V$ can be written in the form

$$Q = \sum_{1 \leq i \leq k} c_i M_i,$$

for some constant c_i. We will show that there is a polynomial $Q \in V$ that vanishes on Σ. To vanish on Σ, coefficients c_i must satisfy

$$Q(a) = \sum_{1 \le i \le k} c_i M_i(a) = 0.$$

for all $a \in \Sigma$. We have $|\Sigma|$ equations and $k > |\Sigma|$ variables so we can solve for a set of coefficients $\{c_1, \ldots, c_k\}$ that satisfy this system of linear equation. Thus, there exists a non-zero element $Q \in V$ that vanishes on Σ. We have $P^* + Q \neq 0$ on Σ. Because $Q \neq 0$, there must be a point $x \in (\mathbb{Z}/3\mathbb{Z})^n \backslash \Sigma$ such that $Q(x) \neq 0$. Such x is not in the support of P^* so $P^*(x) = 0$ and therefore $P^*(x) + Q(x) \neq 0$. This implies the support of $P^* + Q$ strictly contains Σ which contradicts the fact that P^* has maximum support. Therefore,

$$|\Sigma| \ge \dim(V).$$

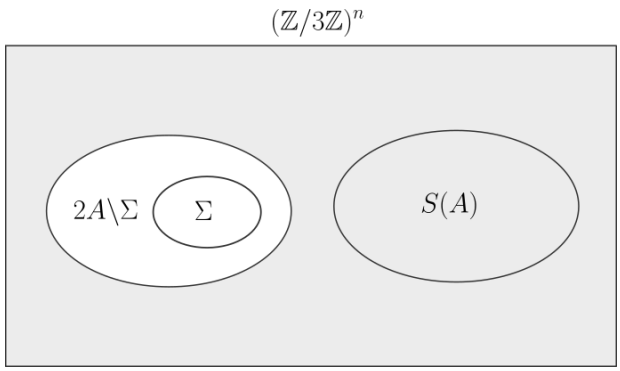


Figure 8: $P \in V$ vanishes on the shaded area

Since P^* vanishes on $(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$, it also vanishes on $S(A)$ that is $P^*(a+b) = 0$ for every pair a, b of distinct elements of A. The Proposition (5.5) implies that there are at most $2m_{d/2}$ element $a \in A$ that $P(2a) \neq 0$. Therefore, the support of P^* has the cardinality at most $2m_{d/2}$. Combine with the result above,

$$2m_{d/2} \ge |\Sigma| \ge \dim(V) \tag{8}$$

Now, we will find a lower bound for $\dim(V)$.

Proposition 5.6. *Let $T \subset (\mathbb{Z}/3\mathbb{Z})^n$ and P_n^d is a space of polynomials of degree at most d with n variables $x_1, \ldots, x_n$. If $V_T \subset P_n^d$ is a set of polynomials that vanish on T then*

$$\dim(V_T) \ge m_d - |T|.$$

Proof. For any polynomial $P \in V_T$, we can write P as a sum of monomials

$$P = \sum_{m \in M_n^d} C_m m$$

Since P vanishes on T, the coefficients C_m for each monomial m must satisfies the system of linear equations

$$\sum_{m \in M_n^d} C_m m(t) = 0 \tag{9}$$

for all $t \in T$. We have $|T|$ linear equations, so the degree of freedom for the choice of C_m is decreased by at most $|T|$ degree. This means we need to have at least $m_d - |T|$ coefficients to uniquely determine the set of coefficients C_m that satisfy (9). Therefore, $\dim(V_T) \geq m_d - |T|$.

□

V is a set of polynomials that vanishes on $(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)$, the proposition (5.6) implies

$$\dim(V) \geq m_d - |(\mathbb{Z}/3\mathbb{Z})^n \backslash (2A)| = m_d - (3^n - |A|). \tag{10}$$

It follows from (8), (10) that

$$\begin{aligned} 2m_{d/2} \geq \dim(V) &\geq m_d - (3^n - |A|) \\ |A| &\leq 2m_{d/2} + (3^n - m_d) \end{aligned}$$

We note that $3^n - m_d$ is the number of monomials in $x_1, \ldots, x_n$ with degree greater than d. By symmetry, the number of such monomials is equal to the number of monomials whose degree is at most $2n - d$ which is m_{2n-d} so

$$|A| \leq 2m_{d/2} + m_{2n-d}$$

Take $d = 4n/3$, we have

$$|A| \leq 3m_{2n/3}$$

as claimed. □

Next, we will show that $m_{2n/3}$ is exponentially small. Let $X_1, \ldots X_n$ be independent identical random variables which take value $0, 1, 2$ with probability $1/3$, then $\frac{1}{3^n} m_{2n/3}$ is equivalent to the probability that $X_1 + \cdots + X_n \leq 2n/3$. By symmetry, this is equal to the probability that $X_1 + \cdots + X_n \geq 4n/3$. We will make use of Markov's inequality.

Proposition 5.7 (Markov's inequality)**.** *Let Z be a non-negative random variable. Then for all $t \geq 0$,*

$$\mathbb{P}(Z \geq t) \leq \frac{\mathbb{E}[Z]}{t}$$

Let $Y = \sum_{i=1}^n X_i$. For any $t \geq 0$, we have $Y \geq a$ if and only if $e^{tY} \geq e^a$. Apply Markov's inequality

$$\mathbb{P}(Y \geq 4n/3) = \mathbb{P}(e^{tY} \geq e^{4nt/3}) \leq \frac{\mathbb{E}[e^{tY}]}{e^{4nt/3}}$$

Because X_i are identically independent random variables,

$$\mathbb{E}[e^{tY}] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}] = (\mathbb{E}[e^{tX_1}])^n = \frac{1}{3^n}(1 + e^t + e^{2t})^n$$

We have

$$\begin{aligned} \mathbb{P}(Y \geq 4n/3) &\leq \frac{(1 + e^t + e^{2t})^n}{3^n e^{4nt/3}} \\ &= \frac{1}{3^n}(e^{-4t/3} + e^{-t/3} + e^{2t/3})^n \end{aligned}$$

Because the bound is true for any $t \geq 0$,

$$\mathbb{P}(Y \geq 4n/3) \leq \min_{t \geq 0} \frac{1}{3^n}(e^{-4t/3} + e^{-t/3} + e^{2t/3})^n.$$

We minimise that term by taking derivatives and setting to zero,

$$n(e^{-4t/3} + e^{-t/3} + e^{2t/3})^{n-1} e^{-7t/3}(2e^{2t} - e^t - 4) = 0.$$

Because $e^t > 0$ for any real number t, the equation implies $(2e^{2t} - e^t - 4) = 0$ that is $e^t = (1 + \sqrt{33})/4$. Substituting into the equation, we have $(e^{-4t/3} + e^{-t/3} + e^{2t/3}) \sim 2.7551 < 2.756$. Therefore,

$$\frac{1}{3^n} m_{2n/3} = \mathbb{P}(Y \geq 4n/3) < \min_{t \geq 0} \frac{1}{3^n}(e^{-4t/3} + e^{-t/3} + e^{2t/3})^n < \frac{1}{3^n} 2.756^n.$$

We have $m_{2n/3} < 2.756^n$. Because $|A| \leq 3m_{2n/3}$, we can conclude that $|A| \sim O((3-c)^n)$ for some $c > 0$.

6 Conclusion

There are many nice ideas that the author want to point out again. First, we observe a connection between Fourier transforms and a trilinear that we introduce to count the number of 3-term arithmetic progression in a subset A.

$$\int_{\theta \in \mathbb{T}} \hat{f}_1(\theta)\hat{f}_2(-2\theta)\hat{f}_3(\theta)d\theta = T(f_1, f_2, f_3).$$

We introduced a balance function f_A which averages the density of A through out $[N]$. We found a relationship between lacking of 3-term arithmetic progression and a large Fourier coefficient of a balance function. We extracted information of $f_A(x)$ from $\hat{f}_A(\theta)$ by finding an appropriate partition of A. These gave us the main proposition which states that if A does not contain a 3-term arithmetic progression, under a certain condition, we can find a subset of A that has a higher density in some sense. We note that a subset of A also does not contain a 3-term arithmetic progression and we apply the proposition again and again. However, a density can't be greater that 1 so some condition must be violated eventually. This technique is called a density increment argument.

Next, we presented a different approach to the problem. We formulated the problem into a graph-theoretic problem where a triangle represents a 3-term arithmetic progression in A. The construction as in chapter 4 implies that triangles which represent trivial arithmetic progressions are edge disjoint. We combined this property with the triangle removal lemma to prove Roth's theorem.

In order to prove the triangle removal lemma, we explored ε-regularity that measures how closed is a graph to a random graph. We gave a constructive proof of the Szemerédi regularity lemma via a density increment argument. We showed that we can consider a large graph by considering groups of its vertices instead and with this idea we proved the triangle counting lemma and the triangle removal lemma.

In the last chapter, we studied an extension of Roth's theorem from $\mathbb{Z}$ to a finite field $(\mathbb{Z}/3\mathbb{Z})^n$. We used the same technique as in chapter 3 to prove Roth's theorem on

$(\mathbb{Z}/3\mathbb{Z})^n$. The property that a finite field is closed under addition, subtraction, multiplication and division increased our scope of 3-term arithmetic progressions and simplified many parts of the proof.

So far, we have used only density increment techniques, so we introduced an alternative proof using polynomial method. The breakthrough lemma of Croot Lev Pach states that if P is a polynomial on $(\mathbb{Z}/3\mathbb{Z})^n$ that $P(a+b)=0$ for any distinct $a,b\in A$, then there are at most $2m_{d/2}$ elements $a\in A$ that $P(2a)\neq 0$ where m_d be the number of monomials in $x_1,\ldots,x_n$ with total degree at most d. Since there is a connection between a 3-term arithmetic progression in a set A and a common element of $S(A)$ and $2A$, we can apply this lemma to help find an upper bound of A. Then, the problem of finding m_d can be seen as a probability problem which we can use Markov's inequality to deal with. Although, the polynomial method does not require an advanced knowledge of mathematics to understand, it gave us a really nice result. For example, the upper bound of a subset $A\subset(\mathbb{Z}/3\mathbb{Z})^n$ that does not contain a 3-term arithmetic progression from Meshulam's is $\frac{2\cdot 3^n}{n}$ while one from the polynomial method is $O(2.76^n)$.

To conclude, we have discussed 4 proofs of Roth's theorem in this essay, they involve Number Theory, Combinatorics, Fourier analysis, Linear Algebra and Probability. We would say that Roth's theorem is a wonderful problem that connects seemingly unrelated fields of mathematics.

References

[BB82] Tom Brown and Joe Buhler. A density version of a geometric ramsey theorem. *Journal of Combinatorial Theory, Series A*, 32(1):20–34, 1982.

[BG] Stephanie Bell and Will Grodzicki. Using szemerédi's regularity lemma to prove roth's theorem. Available at `http://math.uga.edu/~lyall/REU2010/stephaniewill.pdf`.

[BK12] Michael Bateman and Nets Hawk Katz. New bounds on cap sets. *Journal of the American Mathematical Society*, 25(2):585–613, 2012.

[CF13] David Conlon and Jacob Fox. Graph removal lemmas. *Surveys in Combinatorics*, page 1–50, 2013.

[CLP17] Ernie Croot, Vsevolod Lev, and Péter Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Annals of Mathematics*, 185(1):331–337, 2017.

[Ede04] Yves Edel. Extensions of generalized product caps. *Designs, Codes and Cryptography*, 31(1):5–14, 2004.

[EG17] Jordan Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Annals of Mathematics*, 185(1):339–343, 2017.

[Gre] Ben Green. Additive and combinatorial number theory lecture note. Available at `https://courses.maths.ox.ac.uk/node/view_material/41193`.

[KS96] János Komlós and Miklós Simonovits. Szemerédi's regularity lemma and its applications in graph theory. 1996.

[Lee] Choongbum Lee. Topics in combinatorics : Extremal combinatorics lecture note. Available at `https://math.mit.edu/~cb_lee/18.318/lecture3.pdf`.

[Mes95] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *Journal of Combinatorial Theory, Series A*, 71(1):168–172, 1995.

[Rot53] Klaus Roth. On certain sets of integers. *Journal of the London Mathematical Society*, s1-28(1):104–109, 1953.

[Sze75] Endre Szemerédi. Regular partitions of graphs. Technical report, STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE, 1975.