DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS

by

Mustafa Kutay KUTLU

July, 2022
İZMİR

ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS

A Thesis Submitted to the
Graduate School of Natural And Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics

by
Mustafa Kutay KUTLU

July, 2022
İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled "**ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS**" completed by **MUSTAFA KUTAY KUTLU** under supervision of **ASSOC. PROF. DR. ENGİN MERMUT** and **ASST. PROF. DR. HAYDAR GÖRAL** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

....

Assoc. Prof. Dr. Engin MERMUT

Supervisor

....

Asst. Prof. Dr. Haydar GÖRAL

Supervisor

....

Prof. Dr. Engin BÜYÜKAŞIK

Jury Member

....

Prof. Dr. Selçuk DEMİR

Jury Member

....

Assoc. Prof. Dr. Aslı GÜÇLÜKAN İLHAN

Jury Member

Prof. Dr. Okan FISTIKOĞLU
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

First, I would like to offer my most sincere thankfulness to my supervisor Asst. Prof. Dr. Haydar GÖRAL for his encouragement for number theory, endless patience from the moment I started studying with him, and precious guidance. Furthermore, I am deeply grateful to Assoc. Prof. Dr. Engin MERMUT for his help and kind cooperation in every manner during my studies with him.

I am thankful to Prof. Dr. Engin BÜYÜKAŞIK, Prof. Dr. Selçuk DEMİR and Assoc. Prof. Dr. Aslı GÜÇLÜKAN İLHAN for their welcome to take part in the thesis defense jury.

I would like to thank my friends Hikmet Burak ÖZCAN and Sedef ÖZCAN, who always gave me motivation when I was very overwhelmed. When I had problems, they were ready to help.

I would like to thank to Asst. Prof. Dr. Şermin Çam ÇELİK for her time, advice, and support. I sincerely thank Asst. Prof. Dr. Doğa Can SERTBAŞ, Sadık EYİDOĞAN, and Çağatay ALTUNTAŞ for reading my thesis line by line and making critical corrections and valuable comments.

I gratefully acknowledge the support ”2210-A National Scholarship Programme for Master Students” that I received from The Scientific and Technological Research Council of Turkey (TÜBİTAK).

Finally, I must emphasize that I am very grateful to my mother Ayşe KUTLU, and my sister Nurdilay KUTLU for supporting me over the years. They were always with me in any difficulties.

Mustafa Kutay KUTLU

ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS

ABSTRACT

This thesis aims to give two different proofs of Roth's theorem on arithmetic progressions, which is one of the most important results in additive combinatorics. This theorem guarantees the existence of 3-term arithmetic progressions in sufficiently large subsets of the set of positive integers. Aforementioned result was first proved by Klaus Roth in 1953. Roth's original proof is based on the Fourier analysis techniques. We will first give the graph-theoretical proof and then the finite Fourier analytic proof of Roth's theorem on arithmetic progressions.

Keywords: Roth's theorem, arithmetic progression, Fourier analysis, graph theory.

ARİTMETİK DİZİLERDE ROTH TEOREMİ

ÖZ

Bu tezdeki amacımız, toplamsal kombinatorikteki en önemli sonuçlardan biri olan Roth'un aritmetik diziler üzerindeki teoreminin iki farklı kanıtını vermektir. Bu teorem yeteri kadar büyük pozitif tam sayıların alt kümelerindeki 3 terimli aritmetik dizilerin varlığını garanti eder. Bahsi geçen sonuç ilk olarak 1953 yılında Klaus Roth tarafından kanıtlanmıştır. Roth'un orijinal kanıtı Fourier analiz tekniklerine dayanmaktadır. Biz de çalışmamızda Roth teoremine ilk olarak çizge teorisi ile verilen bir kanıtı daha sonra da sonlu Fourier analizi yöntemi ile yapılmış olan kanıtı sunacağız.

Anahtar kelimeler: Roth teoremi, aritmetik dizi, Fourier analizi, çizge teorisi

CONTENTS

LIST OF FIGURES

Page

LIST OF SYMBOLS

n : a natural number

N : a natural number

ε : a real number

δ : a real number

V : vertex set

E : edge set

d : density

ABBREVIATIONS

k–AP	: an arithmetic progression of length k
$i.e.$	: in other words

CHAPTER ONE
INTRODUCTION

Additive combinatorics is a branch of number theory dealing with problems in certain subsets of integers and their behavior under addition. Understanding additive patterns that appear in a sufficiently large set A has always been important in additive combinatorics. For instance, whether A contains non-trivial arithmetic progressions of a given length k has attracted the attention of many mathematicians. One of the important reasons for seeking arithmetic progressions is that they are quite indestructible structures. Indeed, they are preserved under affine transformations.

This branch may be said to begin with an excellent result of van der Waerden in 1927. Van der Waerden proved that any finite coloring of the positive integers contains arbitrarily long non-trivial monochromatic arithmetic progressions. We refer the reader to van der Waerden (1927) for his original proof. After that, in 1936, Erdős and Turán asserted a conjecture that any subset of positive integers with positive upper density (see Definition 2.1.2) contains a non-trivial arithmetic progression of length k, for any $k \geq 3$. This conjecture is a stronger form of van der Waerden's theorem. To see this, we need to have that at least one of the color classes has positive upper density. However, one may immediately see this because if the upper density of all color classes were 0, then the upper density of the positive integers would be 0.

In 1953, Klaus Roth proved Erdős-Turán's conjecture for the case $k = 3$. After this, the case $k = 3$ of Erdős-Turán's conjecture was named Roth's theorem. Roth's proof is based on the Fourier analysis. This proof contains two parts: either the set A looks random, in which case we may easily count the number of three-term progressions, or the set has some structure that can be exploited to find a subset with increased density. It is pretty hard to generalize Roth's method to the case $k \geq 4$. For more details about Roth's Theorem, we direct the reader to Roth (1953).

Szemerédi made the next big breakthrough in 1969. Szemerédi established the case $k = 4$ of Erdős-Turán's conjecture. After that, Szemerédi (1975) obtained the general case of Erdős-Turán's conjecture. Then, the general case of Erdős-Turán's conjecture was named Szemerédi's theorem. His proof is extremely ingenious and contains complicated combinatorial arguments. Because of that reason, his proof is called *tour-de-force*.

An alternative new proof of Szemerédi's theorem was proposed by Furstenberg in 1977 using ergodic theory. Furstenberg's result is important for other results, such as multidimensional and polynomial versions of Szemerédi's theorem. Polynomial Szemerédi's theorem states that if A is a subset of positive integers with positive upper density and $p_1(n), p_2(n), \dots, p_k(n)$ are integer-valued polynomials with $p_i(0) = 0$ for all i, then there are infinitely many $u, n \in \mathbb{Z}$ such that $u + p_i(n) \in A$ for all $1 \leq i \leq k$. Let d be a positive integer, $v_1, v_2, \dots, v_k \in \mathbb{Z}^d$ and $\delta > 0$. Multidimensional Szemerédi's theorem explains that if N is sufficiently large, then every subset A of $\{1, 2, \dots, N\}^d$ with $|A| \geq \delta N^d$ contains a set of the form $a + tv_1, \dots, a + tv_k$, where $a \in \mathbb{Z}^d$ and t is a positive integer. For the results that we mentioned above and more about these results, we refer the reader to Furstenberg (1977), Bergelson & Leibman (1996), and Leibman (1998).

In 1998-2001, Gowers made a significant breakthrough by extending Roth's Fourier analysis techniques to prove Szemerédi's theorem. Gowers proved that there exists a positive constant c_k such that any subset A of $\{1, 2, \dots, N\}$ with $|A| \gg \frac{N}{(\log\log N)^{c_k}}$ contains a non-trivial k-term arithmetic progression. We refer the reader to Gowers (2001) for more details about the result of Gowers.

Although proofs of Szemerédi's theorem of Furstenberg, Szemerédi, and Gowers use different methods, they have much in common. In each of these approaches, the proof has an iterative procedure, and the crucial idea of the proof is the dichotomy between the structure and randomness.

After the work of Gowers, Green and Tao proved that the set of prime numbers contains arbitrarily long non-trivial arithmetic progressions. One may consult Green & Tao (2008). This study is quite important due to the fact that the upper density of primes is zero. Furthermore, Tao and Ziegler established the existence of infinitely many polynomial progressions in the set of primes in Tao & Ziegler (2008).

Erdős asserted another conjecture that for any subset A of positive integers with

$$\sum_{n\in A}\frac{1}{n}=\infty, \tag{1.1}$$

A contains arbitrarily long non-trivial arithmetic progressions. Gowers' result is a significant step toward the proof of Erdős' conjecture. This conjecture was proved by Thomas Bloom and Olof Sisask in 2021 for the case $k = 3$. However, the general case of Erdős' conjecture is still open. For more detailed information about the result of Thomas Bloom and Olof Sisask, we direct the reader to Bloom & Sisask (2021).

CHAPTER TWO
PRELIMINARIES

2.1 Equivalence of Roth's Theorems

From now on, let n and N be natural numbers and ε and δ be real numbers. Also, for brevity, when we call Roth's theorem, we mean Roth's theorem on arithmetic progressions.

In this section, we will show two equivalent versions of Roth's theorem on arithmetic progressions. Before giving identical forms of Roth's theorem, we will introduce some basic definitions from additive combinatorics.

Definition 2.1.1. *A non-trivial arithmetic progression of length k (k-AP) is a sequence of k numbers such that each difference between two consecutive terms is the same positive constant.*
For example, $\{3, 7, 11, 15, 19\}$ *is a non-trivial 5-AP in* $\mathbb{Z}$ *with common difference 4.*

Definition 2.1.2. *The upper density of a subset* $A \subset \mathbb{Z}^+$ *is defined as*

$$\bar{d}(A) = \limsup_{N\to\infty} \frac{|A \cap \{1, 2, \ldots, N\}|}{N}. \tag{2.1}$$

Definition 2.1.3. *Characteristic function* $A(x)$ *of a set A is defined as*

$$A(x) = \begin{cases} 1, & if\ x \in A, \\ 0, & if\ x \notin A. \end{cases} \tag{2.2}$$

Now, we are ready to give the following two equivalent forms of Roth's theorem.

Theorem 2.1.1 (Infinitary form of Roth's theorem)**.** *Let A be a subset of* $\mathbb{Z}^+$ *with positive upper density. Then, A contains a non-trivial three-term arithmetic progression.*

Theorem 2.1.2 (Finitary form of Roth's theorem)**.** *For any $\delta > 0$, there exists a positive integer $N_0 = N_0(\delta)$ such that for every $N \geq N_0$ and every $A \subseteq \{1, 2, \dots, N\}$ with $|A| \geq \delta N$, A contains a non-trivial three-term arithmetic progression.*

Now, we will prove the equivalence of these two forms in the following proof.

Proposition 1. *Finitary version of Roth's theorem is equivalent to infinitary version.*

Proof. First, suppose that the finitary form of Roth's theorem holds. Let $A \subseteq \mathbb{Z}^+$ be a subset with positive upper-density δ. Then, we have that

$$\limsup_{N \to \infty} \frac{|A \cap \{1, 2, \dots, N\}|}{N} = \delta. \tag{2.3}$$

This means that there are infinitely many N such that

$$\frac{|A \cap \{1, 2, \dots, N\}|}{N} > \frac{\delta}{2}. \tag{2.4}$$

Let $B_N = A \cap \{1, 2, \dots, N\}$ with $|B_N| \geq \frac{\delta}{2} N$. By Theorem (2.1.2), there exists a positive integer $N_0 = N_0\left(\frac{\delta}{2}\right)$ such that for every $N \geq N_0$, B_N contains a non-trivial three-term arithmetic progression. As $B_N \subseteq A$, then A also contains a non-trivial three-term arithmetic progression.

Conversely, let us assume that the finitary form does not hold. So, there exists $\delta > 0$, for all $N \in \mathbb{N}$, there exists a subset $A_N \subseteq \{1, 2, \dots, N\}$ with $|A_N| \geq \delta N$, but A_N does not contain a three term arithmetic progression. Now, define $f_N : \mathbb{N} \to \{0, 1\}$ as the characteristic function of A_N and consider a metric d on $2^{\mathbb{N}}$, namely, all functions from $\mathbb{N}$ to $\{0, 1\}$, as

$$d(g, h) = \inf \left\{ \frac{1}{k+1} : g(i) = h(i), for\ all\ i < k \right\}, \tag{2.5}$$

where $g, h \in 2^{\mathbb{N}}$. This gives us a product topology on $2^{\mathbb{N}}$. As $2^{\mathbb{N}}$ is a compact set in this topology, there exists a convergent subsequence $\{f_{N_k}\}$ of $\{f_N\}$. Assume that $f_{N_k} \to f$, where f is the characteristic function of $A \subseteq \mathbb{N}$.

With the abuse of notation, we may write f_N instead of f_{N_k}. Then, for any $\varepsilon > 0$, there exists N_ε such that $d(f_N, f) < \varepsilon$, for all $N > N_\varepsilon$. In other words,

$$\inf\left\{\frac{1}{k+1} : f_N(i) = f(i), \forall i < k\right\} < \varepsilon, \tag{2.6}$$

for all $N > N_\varepsilon$. As f_N converges to f, the number of elements of $A_N \cap \{1, 2, \ldots, N\}$ converges to that of $A \cap \{1, 2, \ldots, N\}$. By assumption

$$\limsup_{N\to\infty} \frac{|A_N \cap \{1, 2, \ldots, N\}|}{N} \geq \delta \tag{2.7}$$

then

$$\limsup_{N\to\infty} \frac{|A \cap \{1, 2, \ldots, N\}|}{N} \geq \delta \tag{2.8}$$

is satisfied.

Thus, the upper density of A is positive. This gives us that A must contain a three-term arithmetic progression, a contradiction. Therefore, the finitary form holds. □

2.2 Basic Concepts from Graph Theory

In this section, we will give a gentle introduction to most of the terminology used in Chapter 3. We may refer the reader to Diestel (2018) for more details about graph theory.

Definition 2.2.1. *A graph is a structure $G = (V, E)$, where V is a set whose elements are vertices and $E \subseteq V \times V$ is a set of paired vertices, whose elements are called edges.*

Throughout this chapter, we will be interested only in simple graphs, which are unweighted, undirected and contain no loops or multiple edges. Let us see a simple graph in the following example.

Example 1. *Let* $V = \{1, 2, 3, 4, 5, 6\}$ *be a vertex set and* $E = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{4, 5\}, \{5, 6\}\}$ *be an edge set. Then,*

$G:$

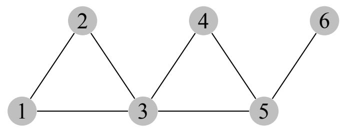

Figure 2.1 Simple graph

is a simple graph.

The number of vertices of a graph G is called its order and denoted by $|G|$. We denote the number of edges of G by $||G||$ and the number of all edges between X and Y in E by $e(X, Y)$.

Definition 2.2.2. *Let* $G = (V, E)$ *be a graph and* $X, Y \subseteq V$. *Then, the density of the pair* (X, Y)*, denoted by* $d(X, Y)$*, is defined as*

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}. \tag{2.9}$$

This definition describes the notion of density in the graph-theoretical sense. Let us illustrate this definition with an example.

Example 2. *Let* $X = (V_1, E_1)$*, where* $V_1 = \{1, 2, 3\}$ *and* $E_1 = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$ *and* $Y = (V_2, E_2)$*, where* $V_2 = \{4, 5\}$ *and* $E_2 = \{\{4, 5\}\}$.

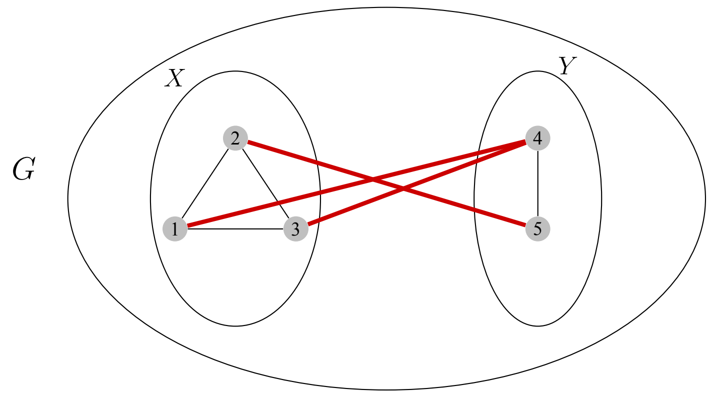

Figure 2.2 Density between two graphs

Then, the density of the pair (X,Y) is

$$d(X,Y) = \frac{e(X,Y)}{|X||Y|} = \frac{3}{6} = \frac{1}{2}. \tag{2.10}$$

Definition 2.2.3. *Let $G = (V,E)$ be a graph and $\varepsilon > 0$ be given. We call a pair (A,B), where A and B are disjoint subsets of V is ε–$regular$, if all $X \subseteq A$ and all $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ satisfy $|d(X,Y) - d(A,B)| \leq \varepsilon$.*

Example 3. *Let $G = (V,E)$ be a graph with $V = \{1,2,3,4,5\}$ and $X = (V_1, E_1)$, where $V_1 = \{1,2,3\}$ and $E_1 = \emptyset$ and $Y = (V_2, E_2)$, where $V_2 = \{4,5\}$ and $E_2 = \emptyset$.*

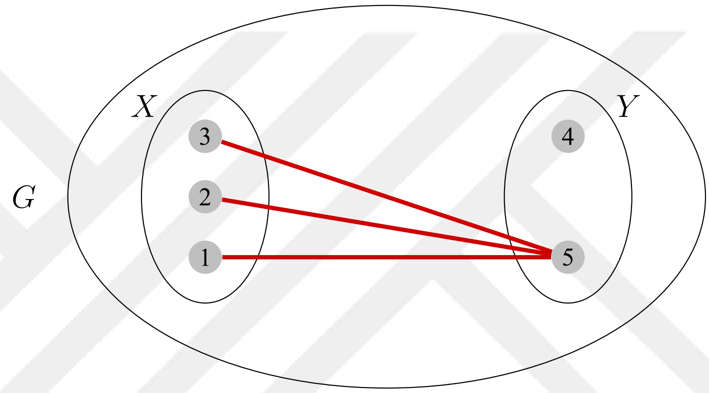

Figure 2.3 Example of non-regularity

The pair (X,Y) is not $\frac{1}{3}$–$regular$. Indeed, if we take $X' = X$ and $Y' = \{5\}$, then $|X'| \geq \frac{1}{3}|X|$ and $|Y'| \geq \frac{1}{3}|Y|$. Thus, we have that $d(X,Y) = \frac{3}{6} = \frac{1}{2}$ and $d(X',Y') = \frac{3}{3} = 1$, which gives

$$|d(X,Y) - d(X',Y')| = \frac{1}{2} > \frac{1}{3}. \tag{2.11}$$

The previous example shows that ε–$regularity$ is meaningful only when the density is large enough.

Remark 1. *Let* X *and* Y *be two disjoint vertex sets with* $d(X,Y) \leq \varepsilon^3$*. Then, the pair* (X,Y) *is* ε*–*$regular$*.*

Proof. If $|X^{'}| \geq \varepsilon|X|$ and $|Y^{'}| \geq \varepsilon|Y|$ then

$$d(X^{'},Y^{'}) = \frac{e(X^{'},Y^{'})}{|X'||Y'|} \leq \frac{e(X,Y)}{\varepsilon|X|\varepsilon|Y|} = \frac{1}{\varepsilon^2}d(X,Y) \leq \frac{\varepsilon^3}{\varepsilon^2} = \varepsilon. \tag{2.12}$$

Hence, $|d(X,Y) - d(X^{'},Y^{'})| \leq \varepsilon$. □

Definition 2.2.4. *A partition* $P = \{V_1, V_2, \ldots, V_k\}$ *of the vertex set* V *is the set of non-empty subsets in* V *such that*

$$\bigcup_{i=1}^{k} V_i = V, \tag{2.13}$$

with $V_i \cap V_j = \emptyset$ *for all* $i \neq j$*.*

Definition 2.2.5. *Any partition* $P^{'}$ *of* V *is called a refinement of* P*, if every element of* $P^{'}$ *is a subset of some element of* P*.*

A partition $P = \{V_1, \ldots, V_k\}$ is called an equipartition, if $|V_1| = |V_2| = \cdots = |V_k|$.

Definition 2.2.6. *Let* $P = \{V_0, V_1, \ldots, V_k\}$ *be a partition of* V *in which one set* V_0 *has been singled out as an exceptional set. Then,* P *is called an* ε*– regular partition of* V*, if it satisfies the following three conditions:*

1. $|V_0| \leq \varepsilon|V|$,
2. $|V_1| = |V_2| = \cdots = |V_k|$,
3. *all but at most* εk^2 *of the pairs* (V_i, V_j) *with* $1 \leq i < j \leq k$ *are* ε*– regular.*

Definition 2.2.7. *The mean square density of a partition* $P = \{V_1, V_2, \ldots, V_k\}$ *of a graph* G *of order* n *is defined by*

$$d_2(P) = \sum_{1 \leq i < j \leq k} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2. \tag{2.14}$$

2.3 Basic Properties of the Fourier Transform

In this section, we will set the stage by introducing the main objects in Fourier analysis that we will use in the proof of Roth's theorem.

Definition 2.3.1. *For any natural number N, the function $\chi : \mathbb{Z}_N \longrightarrow \mathbb{C}^*$ defined as $\chi(t) = \exp\left(\frac{2\pi it}{N}\right) = e^{2\pi it/N}$ is called an additive character on $\mathbb{Z}_N$.*

Note that an additive character χ satisfies that for any $a, b \in \mathbb{Z}_N$,

$$\chi(a+b) = \chi(a)\chi(b), \tag{2.15}$$

i.e. χ respects additive and multiplicative operations of $\mathbb{Z}_N$ and $\mathbb{C}^*$, respectively.

Lemma 2.3.1.

$$\sum_{r\in\mathbb{Z}_\mathbb{N}} \chi(ru) = \begin{cases} 0, & u \neq 0 \\ N, & u = 0. \end{cases} \tag{2.16}$$

Proof. Let

$$S = \sum_{r\in\mathbb{Z}_\mathbb{N}} \chi(ru) = 1 + \chi(u) + \chi(2u) + \ldots + \chi((N-1)u). \tag{2.17}$$

Then,

$$\begin{aligned} \chi(u)S &= \chi(u) + \chi(u)\chi(u) + \ldots + \chi((N-2)u)\chi(u) + \chi((N-1)u)\chi(u) && (2.18) \\ &= \chi(u) + \chi(2u) + \ldots + \chi((N-1)u) + 1 && (2.19) \\ &= S && (2.20) \end{aligned}$$

which implies that either $S = 0$ or $\chi(u) = 1$. Here, $\chi(u) = 1$ if and only if $u = 0$. Clearly, the sum $S = N$ if $u = 0$. Otherwise, if $u \neq 0$ then $S = 0$. $\square$

Now, let us define one of our main tools: Fourier transform.

Definition 2.3.2. *Given a function $f : \mathbb{Z}_N \to \mathbb{C}$, the Fourier transform $\widehat{f} : \mathbb{Z}_N \to \mathbb{C}$ of f is defined as*

$$\widehat{f}(r) = N^{-1} \sum_{x \in \mathbb{Z}_N} \chi(xr) f(x). \tag{2.21}$$

Fourier transform has many useful properties. The following theorem summarizes a number of them, which will be very handy in the proof of Roth's theorem.

Theorem 2.3.2. *Let $f : \mathbb{Z}_N \to \mathbb{C}$. The Fourier transform $\widehat{f}$ has the following properties:*

(a) Inversion Formula *For any $x \in \mathbb{Z}_N$,*

$$f(x) = \sum_{r \in \mathbb{Z}_N} \chi(xr) \widehat{f}(r). \tag{2.22}$$

(b) Pointwise Estimate *For any $r \in \mathbb{Z}_N$ we have*

$$|\widehat{f}(r)| \leq N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|. \tag{2.23}$$

(c) Plancherel's Identity

$$\sum_{r \in \mathbb{Z}_N} |\widehat{f}(r)|^2 = N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|^2. \tag{2.24}$$

(d) Convolution Identity *For $f, g : \mathbb{Z}_N \to \mathbb{C}$, we define the convolution of f and g as follows*

$$(f * g)(x) = \sum_{y \in \mathbb{Z}_N} f(y) g(x - y). \tag{2.25}$$

Then for any r

$$\widehat{(f * g)}(r) = N \widehat{f}(r) \widehat{g}(r). \tag{2.26}$$

Proof.

(a) As

$$\widehat{f}(r) = N^{-1} \sum_{y \in \mathbb{Z}_N} \chi(-yr) f(y) \tag{2.27}$$

$$\sum_{r \in \mathbb{Z}_N} \chi(xr) \widehat{f}(r) = N^{-1} \sum_{r \in \mathbb{Z}_N} \sum_{y \in \mathbb{Z}_N} \chi(xr) \chi(-yr) f(y). \tag{2.28}$$

This implies that

$$\sum_{r \in \mathbb{Z}_N} \chi(xr) \widehat{f}(r) = N^{-1} \sum_{y \in \mathbb{Z}_N} f(y) \sum_{r \in \mathbb{Z}_N} \chi(r(x-y)). \tag{2.29}$$

Then, by Proposition (2.3.1), which proves (a).

(b) As

$$\widehat{f}(r) = N^{-1} \sum_{x \in \mathbb{Z}_N} \chi(-xr) f(x), \tag{2.30}$$

$$|\widehat{f}(r)| = \left| N^{-1} \sum_{x \in \mathbb{Z}_N} \chi(-xr) f(x) \right| \leq N^{-1} \sum_{x \in \mathbb{Z}_N} |\chi(-xr)||f(x)|. \tag{2.31}$$

As $|\chi(-xr)| = |\exp(\frac{2\pi i(-xr)}{N})| = 1$, we get

$$|\widehat{f}(r)| \leq N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|. \tag{2.32}$$

(c) We know that for any $a \in \mathbb{C}$, $|a|^2 = a\bar{a}$. So,

$$|\widehat{f}(r)|^2 = \widehat{f}(r)\overline{\widehat{f}(r)} = N^{-1} \sum_{x \in \mathbb{Z}_N} \chi(-xr) f(x) \overline{N^{-1} \Bigg(\sum_{y \in \mathbb{Z}_N} \chi(-yr) f(y) \Bigg)} \tag{2.33}$$

$$= N^{-2} \sum_{x \in \mathbb{Z}_N} \sum_{y \in \mathbb{Z}_N} f(x)\overline{f(y)} \chi(-xr) \overline{\chi(-yr)}. \tag{2.34}$$

Since $\overline{\chi(-yr)} = \chi(yr)$, we obtain

$$|\widehat{f}(r)|^2 = N^{-2} \sum_{x,y \in \mathbb{Z}_N} f(x)\overline{f(y)}\chi(-xr)\chi(yr). \tag{2.35}$$

This gives us

$$\sum_{r \in \mathbb{Z}_N} |\widehat{f}(r)|^2 = N^{-2} \sum_{r \in \mathbb{Z}_N} \sum_{x,y \in \mathbb{Z}_N} f(x)\overline{f(y)}\chi(-xr)\chi(yr) \tag{2.36}$$

$$= N^{-2} \sum_{x,y \in \mathbb{Z}_N} f(x)\overline{f(y)} \sum_{r \in \mathbb{Z}_N} \chi(r(y-x)). \tag{2.37}$$

As we know $\sum_r \chi(r(y-x)) = N$ for $y = x$ and 0 otherwise, we have

$$\sum_{r \in \mathbb{Z}_N} |\widehat{f}(r)|^2 = N^{-1} \sum_{x \in \mathbb{Z}_N} |f(x)|^2. \tag{2.38}$$

(d) The convolution of f and g is

$$(f * g)(x) = \sum_y f(y)g(x-y). \tag{2.39}$$

Applying the Fourier transform we get

$$\widehat{(f * g)}(r) = N^{-1} \sum_x \chi(-xr)(f * g)(x) \tag{2.40}$$

$$= N^{-1} \sum_x \chi(-xr) \sum_y f(y)g(x-y). \tag{2.41}$$

Since $\chi(-yr)\chi(yr) = 1$,

$$\widehat{(f * g)}(r) = N^{-1} \sum_{x,y} \chi(-xr)f(y)g(x-y)\chi(-yr)\chi(yr) \tag{2.42}$$

$$= N^{-1} \sum_y f(y)\chi(-yr) \sum_x g(x-y)\chi(-r(x-y)). \tag{2.43}$$

As

$$\sum_{y} f(y)\chi(-yr) = N\widehat{f}(r) \quad and \quad \sum_{x} g(x-y)\chi(-r(x-y)) = N\hat{g}(r) \tag{2.44}$$

we have,

$$\widehat{(f*g)}(r) = N\widehat{f}(r)\widehat{g}(r). \tag{2.45}$$

□

CHAPTER THREE
PROOF OF ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS VIA GRAPH THEORY

In this chapter, we will prove Roth's theorem using some techniques from graph theory. To do this, we first give the proof of Szemerédi's regularity lemma and the triangle removal lemma. Thanks to these two results, we will obtain the proof of Roth's theorem.

Before giving the proof of Szemerédi's regularity lemma, it is necessary to establish the following auxiliary lemmas.

Lemma 3.0.1. *(Diestel (2018), [Lemma 7.2.2]) Let G be a graph with $|G| = n$. Then,*

1. *For all partitions P of V, we have $0 \leq d_2(P) \leq \frac{1}{2}$.*
2. *If P' refines P, then $d_2(P') \geq d_2(P)$.*

Proof. For (1), let $P = \{V_1, \dots, V_k\}$. Then,

$$d_2(P) = \sum_{1 \leq i < j \leq k} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 \leq \sum_{1 \leq i < j \leq k} \frac{|V_i||V_j|}{n^2}. \tag{3.1}$$

As

$$\frac{(|V_1| + \cdots + |V_k|)^2}{n^2} = 1, \tag{3.2}$$

holds, we have

$$1 = \frac{(|V_1| + \cdots + |V_k|)^2}{n^2} = \sum_{i=1}^{k} \frac{|V_i|}{n^2} + 2 \sum_{1 \leq i < j \leq k} \frac{|V_i||V_j|}{n^2}. \tag{3.3}$$

Since we have that

$$\sum_{i=1}^{k} \frac{|V_i|}{n^2} > 0, \tag{3.4}$$

$$2 \sum_{1 \le i < j \le k} \frac{|V_i||V_j|}{n^2} < 1 \tag{3.5}$$

is satisfied.

This implies that

$$\sum_{1 \le i < j \le k} \frac{|V_i||V_j|}{n^2} < \frac{1}{2}. \tag{3.6}$$

Thus, we get $0 \le d_2(P) \le \frac{1}{2}$.

For (2), let $P = \{X, Y\}$ and $P^{'} = \{X_1, X_2, Y_1, Y_2\}$, where $X = X_1 \cup X_2$ with $X_1 \cap X_2 = \emptyset$ and $Y = Y_1 \cup Y_2$ with $Y_1 \cap Y_2 = \emptyset$. Then,

$$e(X, Y) = \sum_{i,j=1}^{2} e(X_i, Y_j) \tag{3.7}$$

and so

$$|X||Y|d(X, Y) = \sum_{i,j=1}^{2} |X_i||Y_j|d(X_i, Y_j). \tag{3.8}$$

Hence,

$$d(X, Y) = \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) \tag{3.9}$$

holds.

Then, we obtain that

$$d(X, Y)^2 = \left(\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j) \right)^2 \tag{3.10}$$

$$= \left(\sum_{i,j=1}^{2} \sqrt{\frac{|X_i||Y_j|}{|X||Y|}} \sqrt{\frac{|X_i||Y_j|}{|X||Y|}} d(X_i, Y_j) \right)^2. \tag{3.11}$$

Using the Cauchy-Schwarz inequality, we get that

$$d(X,Y)^2 \leq \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} \cdot \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i,Y_j)^2. \tag{3.12}$$

Since we have

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} = 1, \tag{3.13}$$

we conclude that

$$\frac{|X||Y|}{n^2} d(X,Y)^2 \leq \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{n^2} d(X_i,Y_j)^2. \tag{3.14}$$

Hence, $d_2(P) \leq d_2(P')$ is satisfied. □

Lemma 3.0.2. *(Diestel (2018), [Lemma 7.2.3]) Suppose that the pair (X,Y) is not ε–regular. Suppose further that $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ with $|X_1| \geq \varepsilon|X|$, $|Y_1| \geq \varepsilon|Y|$ and*

$$|d(X,Y) - d(X_1,Y_1)| > \varepsilon. \tag{3.15}$$

Then, we have

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i,Y_j)^2 \geq d(X,Y)^2 + \varepsilon^4. \tag{3.16}$$

Proof. Note that

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} (d(X_i,Y_j) - d(X,Y))^2 \tag{3.17}$$

$$= \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i,Y_j)^2 - 2d(X,Y) \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i,Y_j) + d(X,Y)^2 \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|}. \tag{3.18}$$

Since

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i,Y_j) = d(X,Y) \quad \text{and} \quad \sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} = 1, \tag{3.19}$$

we get that

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 - d(X,Y)^2 \geq \frac{|X_1||Y_1|}{|X||Y|}(d(X_1,Y_1) - d(X,Y))^2 \tag{3.20}$$

$$\geq \frac{\varepsilon|X|\varepsilon|Y|}{|X||Y|}\varepsilon^2 = \varepsilon^4. \tag{3.21}$$

Thus, we conclude that

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{|X||Y|} d(X_i, Y_j)^2 \geq \varepsilon^4 + d(X,Y)^2. \tag{3.22}$$

□

One may immediately deduce from the previous lemma that

$$\sum_{i,j=1}^{2} \frac{|X_i||Y_j|}{n^2} d(X_i, Y_j)^2 \geq \frac{|X||Y|}{n^2} d(X,Y)^2 + \frac{|X||Y|}{n^2}\varepsilon^4 \tag{3.23}$$

holds.

Lemma 3.0.3. *(Diestel (2018), [Lemma 7.2.4]) Let G be a graph with n vertices and $0 < \varepsilon \leq \frac{1}{4}$. Let $P = \{V_0, V_1, \ldots, V_k\}$ be a partition of G such that $|V_0| \leq \varepsilon n$ and $|V_1| = \cdots = |V_k|$. If P is not ε–regular, then there exists a refinement $P' = \{V_0', V_1', \ldots, V_\ell'\}$ of P such that the following are satisfied:*

1. $|V_0'| \leq |V_0| + \frac{1}{2}\varepsilon^6 n$,
2. $|V_1'| = |V_2'| = \cdots = |V_\ell'|$,
3. $k \leq \ell \leq k\varepsilon^{-6}2^{k+1}$,
4. $d_2(P') \geq d_2(P) + \frac{\varepsilon^5}{2}$.

Proof. Let $I = \{(i,j) : 1 \leq i < j \leq k$ and (V_i, V_j) is ε–$irregular\}$. As the set I contains ε–$irregular$ pairs and P is not ε–$regular$ with $|V_0| \leq \varepsilon n$ and $|V_1| = |V_2| = \cdots = |V_k|$, we have $|I| \geq \varepsilon k^2$. Consider an ε–$irregular$ pair (V_i, V_j). Then, there are subsets $V_{ij}^{(1)} \subseteq V_i$ and $V_{ji}^{(1)} \subseteq V_j$ such that $|V_{ij}^{(1)}| \geq \varepsilon |V_i|$ and $|V_{ji}^{(1)}| \geq \varepsilon |V_j|$ and

$$|d(V_{ij}^{(1)}, V_{ji}^{(1)}) - d(V_i, V_j)| > \varepsilon \tag{3.24}$$

holds.

Let $V_i = V_{ij}^{(1)} \cup V_{ij}^{(2)}$ and $V_j = V_{ji}^{(1)} \cup V_{ji}^{(2)}$ with $V_{ij}^{(1)} \cap V_{ij}^{(2)} = \emptyset$ and $V_{ji}^{(1)} \cap V_{ji}^{(2)} = \emptyset$. By Lemma 3.0.2, we see that

$$\sum_{k,l=1}^{2} \frac{|V_{ij}^{(k)}||V_{ji}^{(l)}|}{n^2} d(V_{ij}^{(k)}, V_{ji}^{(l)})^2 \geq \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 + \varepsilon^4 \frac{|V_i||V_j|}{n^2}. \tag{3.25}$$

Then, for any $i \in \{1, 2, \ldots, k\}$, we have two possibilities:

1. If $(i,j) \notin I$ for any $j \in \{1, 2, \ldots, k\}$, then set $P_i = \{V_i\}$.
2. If $(i,j) \in I$ for some $j > i$, then set $P_{ij} = \{V_{ij}^{(1)}, V_{ij}^{(2)}\}$ and P_i to be the common refinement of all P_{ij}'s.

For instance, let us take $P = \{V_0, V_1, V_2, V_3\}$ and $i = 1$. If for all $(i,j) \notin I$, then we set $P_1 = \{V_1\}$. For $i = 2$, if $(2,1), (2,3) \in I$, then we set $P_{23} = \{V_{23}^{(1)}, V_{23}^{(2)}\}$ and $P_{21} = \{V_{21}^{(1)}, V_{21}^{(2)}\}$. Hence

$$P_2 = \{V_{23}^{(1)} \cap V_{21}^{(1)}, V_{23}^{(1)} \cap V_{21}^{(2)}, V_{23}^{(2)} \cap V_{21}^{(1)}, V_{23}^{(2)} \cap V_{21}^{(2)}\}. \tag{3.26}$$

As we split V_i's into two parts every time when we create P_{ij}'s, we have that $|P_i| \geq 2^k$ for all $i \in \{1, 2, \ldots, k\}$. Now, let $P_i = \{V_{i1}, V_{i2}, \ldots, V_{ik_i}\}$ for each i and

$$Q = \{V_0\} \cup P_1 \cup P_2 \cup \cdots \cup P_k = \{V_0, V_{11}, V_{12}, \ldots, V_{1k_1}, \ldots, V_{k1}, \ldots, V_{kk_k}\}. \tag{3.27}$$

So, $|Q| \geq k2^k$ is satisfied. Now, we want to show that

$$d_2(Q) \geq d_2(P) + \frac{\varepsilon^5}{2}. \tag{3.28}$$

As $P_i = \{V_{i1}, \ldots, V_{ik_i}\}$ is a refinement of $P_{ij} = \{V_{ij}^{(1)}, V_{ij}^{(2)}\}$, by Lemma (3.0.1), we deduce that

$$\sum_{k,\ell=1}^{2} \frac{|V_{ij}^{(k)}||V_{ji}^{(\ell)}|}{n^2} d(V_{ij}^{(k)}, V_{ji}^{(\ell)})^2 \leq \sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2. \tag{3.29}$$

By (3.25), we have that

$$\sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2 \geq \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 + \varepsilon^4 \frac{|V_i||V_j|}{n^2}. \tag{3.30}$$

As $Q = \{V_0 \cup P_1 \cup, \ldots, \cup P_k\}$, where $P_i = \{V_{i1}, \ldots, V_{ik_i}\}$, we see that

$$d_2(Q) = \sum_{i<j} \sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2 \tag{3.31}$$

$$= \sum_{1\leq i<j\leq k} \sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2 + \sum_{i=1}^{k} \sum_{a=1}^{k_i} \frac{|V_0||V_{ia}|}{n^2} d(V_0, V_{ia})^2. \tag{3.32}$$

Let

$$R = \sum_{i=1}^{k} \sum_{a=1}^{k_i} \frac{|V_0||V_{ia}|}{n^2} d(V_0, V_{ia})^2. \tag{3.33}$$

Then, we get that

$$d_2(Q) = \sum_{(i,j)\notin I} \sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2 + \sum_{(i,j)\in I} \sum_{a=1}^{k_i} \sum_{b=1}^{k_j} \frac{|V_{ia}||V_{jb}|}{n^2} d(V_{ia}, V_{jb})^2 + R \tag{3.34}$$

$$\geq \sum_{(i,j)\notin I} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 + \sum_{(i,j)\in I} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 \sum_{(i,j)\in I} \frac{|V_i||V_j|}{n^2} \varepsilon^4 \tag{3.35}$$

$$+ \sum_{i=1}^{k} \frac{|V_0||V_i|}{n^2} d(V_0, V_i)^2. \tag{3.36}$$

As $|V_i| \geq \frac{n-\varepsilon n}{k}$ and

$$d_2(P) = \sum_{(i,j)\notin I} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 + \sum_{(i,j)\in I} \frac{|V_i||V_j|}{n^2} d(V_i, V_j)^2 + \sum_{i=1}^{k} \frac{|V_0||V_i|}{n^2} d(V_0, V_i)^2, \tag{3.37}$$

we obtain that

$$d_2(Q) \geq d_2(P) + \varepsilon^4 \sum_{(i,j)\in I} \frac{|V_i||V_j|}{n^2} \tag{3.38}$$

$$\geq d_2(P) + \varepsilon^4 \varepsilon k^2 \frac{n^2(1-\varepsilon)^2}{k^2 n^2} \tag{3.39}$$

$$= d_2(P) + \varepsilon^5 (1-\varepsilon)^2 \tag{3.40}$$

$$= d_2(P) + \frac{\varepsilon^5}{2}. \tag{3.41}$$

Thus, $d_2(Q) \geq d_2(P) + \frac{\varepsilon^5}{2}$.

Now, let $A = k2^k$. Then, to make Q an equipartition and to show the third hypothesis of Lemma 3.0.3, we must divide Q by partitioning each part into parts of sizes exactly $\lfloor \frac{\varepsilon^6 n}{2A} \rfloor$ and at most one part of size less than that. Then, we get $\frac{2A}{\varepsilon^6} = \frac{2k2^k}{\varepsilon^6}$ parts. In addition, if we move all small parts into V_0, then we have $V_0^{'}$. In this way, we can obtain an equipartition $P^{'} = \{V_0^{'}, V_1^{'}, \ldots, V_\ell^{'}\}$ with

$$|V_0^{'}| \leq |V_0| + A\frac{\varepsilon^6 n}{2A} = |V_0| + \frac{\varepsilon^6 n}{2} \tag{3.42}$$

and $d_2(P^{'}) \geq d_2(Q)$.

Therefore,

$$d_2(P^{'}) \geq d_2(P) + \frac{\varepsilon^5}{2}, \tag{3.43}$$

which completes the proof. □

3.1 Szemerédi's Regularity Lemma

Now, we are ready to prove Szemerédi's regularity lemma.

Szemerédi's Regularity Lemma. (*Lima (2011)*) *For all $\varepsilon > 0$ and $t \in \mathbb{N}$, there exist $N = N(\varepsilon, t)$ and $T = T(\varepsilon, t)$ such that every n– vertex graph G, where $n > N$, admits an ε– regular partition P with $t \leq |P| \leq T$.*

Proof. Let $0 < \varepsilon \leq \frac{1}{4}$ and $t \in \mathbb{N}$ be given. Let $N = \lceil \frac{2t}{\varepsilon} \rceil$ and $n > N$. Consider an n– vertex graph G and let $P_0 = \{V_0, V_1, \ldots, V_t\}$ be a partition of G, where $|V_0| \leq t - 1$ and $|V_1| = |V_2| = \cdots = |V_t|$. As $n > N \geq \frac{2t}{\varepsilon}$, $\frac{\varepsilon n}{2} > t$ and also $\frac{\varepsilon n}{2} \geq t - 1$ hold. This implies that $|V_0| \leq \frac{\varepsilon n}{2}$. Now, we have two cases. If P_0 satisfies the third property of Definition (2.2.6), then P_0 is ε– regular and we are done.

Otherwise, by Lemma (3.0.3), there exists a refinement P_1 of P_0 which is an equipartition such that the size of the exceptional set increases by at most $\frac{1}{2}\varepsilon^6 n$ and $d_2(P_1) \geq d_2(P_0) + \frac{\varepsilon^5}{2}$. We have again two cases for the new partition P_1. If P_1 satisfies the third property of Definition (2.2.6), then P_1 is ε– regular and we are done. Otherwise, by Lemma (3.0.3), there exists a refinement P_2 of P_1 which is an equipartition such that the size of the exceptional set increases by at most $\frac{1}{2}\varepsilon^6 n$ and $d_2(P_2) \geq d_2(P_1) + \frac{\varepsilon^5}{2}$.

We can repeat this process and find the k^{th} partition P_k with

$$d_2(P_k) \geq d_2(P_{k-1}) + \frac{\varepsilon^5}{2}. \tag{3.44}$$

Also, we get from the previous steps that $d_2(P_k) \geq d_2(P_0) + k\frac{\varepsilon^5}{2}$. But, by Lemma (3.0.1), we know that $d_2(P_k) \leq \frac{1}{2}$ implying that $\frac{k\epsilon^5}{2} \leq \frac{1}{2}$. Then, we have that $k \leq \frac{1}{\varepsilon^5}$. Thus, our process must end in at most ε^{-5}-steps. So, in at most ε^{-5}-steps, we get an equipartition P such that P satisfies the third property of Definition (2.2.6). Moreover, the exceptional set of P has size at most $\frac{\varepsilon n}{2} + \varepsilon^{-5}\frac{1}{2}\varepsilon^6 n = \varepsilon n$. Therefore, P is an ε-regular partition of the graph G with $t \leq |P| \leq T$, for some T depending on t. $\square$

3.2 Triangle Removal Lemma

To state the triangle removal lemma, we need to give the following two definitions.

Definition 3.2.1. *In graph theory, a triangle is a simple graph with three vertices and three edges.*

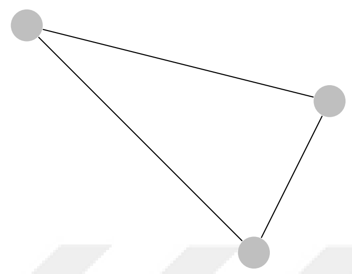

Figure 3.1 A triangle

Definition 3.2.2. *A graph G is said to be $k-$ far from being triangle free, if one has to delete at least k-edges of G to delete all triangles in it.*

Example 4. *Let G be a $7-$vertex graph with vertex set $V = \{1, 2, 3, 4, 5, 6, 7\}$ and edge set*

$$E = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{3, 6\}, \{3, 7\}, \{5, 6\}, \{5, 7\}\}. \tag{3.45}$$

That is,

G :

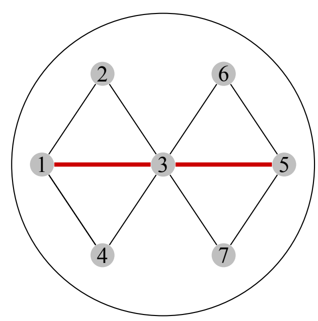

Figure 3.2 Example of 2-far from being triangle free graph

It is $2-$ far from being triangle free.

Now, we can state the triangle removal lemma.

Triangle Removal Lemma. *For all $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon)$ and $N = N(\varepsilon)$ such that for all $n \geq N$, whenever $n-$ vertex graph $G = (V, E)$ is εn^2- far from being triangle free, it contains at least δn^3 triangles.*

Proof. Let $\varepsilon > 0$ and $t = \lceil \frac{4}{\varepsilon} \rceil$. Then by Szemerédi's regularity lemma, there exist N and T such that every $n-$ vertex graph G with $n \geq N$ has an $\frac{\varepsilon}{4}-$ regular partition whose length is between t and T. Also, by our assumption, G is εn^2- far from being triangle free. Now, let $P = \{V_0, V_1, \ldots, V_k\}$ be an $\frac{\varepsilon}{4}-$ regular partition of G, where $t \leq k \leq T$ and $|V_1| = |V_2| = \cdots = |V_k| = c$, where c is a constant. Then, we can obtain a subgraph G' of G by deleting the following edges:

1. All edges incident in V_0. In this case, there exist at most $|V_0|n \leq \frac{\varepsilon}{4}n^2$ edges are deleted.

2. All edges inside V_i, for all $i \in \{1, 2, \ldots, k\}$. In this case,

$$|V_i|^2 k = c^2 k \leq \left(\frac{n}{k}\right)^2 k = \frac{n^2}{k} \leq \frac{n^2}{t}. \tag{3.46}$$

   Since $t = \lceil \frac{4}{\varepsilon} \rceil$, there exist at most $\frac{n^2}{t} \leq \frac{\varepsilon n^2}{4}$ edges are deleted.

3. All edges between irregular pairs. In this case, there exist at most

$$\frac{\varepsilon}{4}k^2|V_i|^2 = \frac{\varepsilon}{4}k^2c^2 \leq \frac{\varepsilon}{4}k^2\left(\frac{n}{k}\right)^2 \leq \frac{\varepsilon n^2}{4} \tag{3.47}$$

   edges are deleted.

4. All edges between regular pairs in which the density is smaller than $\frac{\varepsilon}{2}$. In this case,

$$e(V_i, V_j) = |V_i||V_j|d(V_i, V_j) < \frac{\varepsilon}{2}\left(\frac{n}{k}\right)^2 \tag{3.48}$$

   holds.

   So, there exists at most $\frac{k}{2}\frac{\varepsilon}{2}\frac{n^2}{k^2} < \frac{\varepsilon n^2}{4}$ edges are deleted.

Hence, the number of deleted edges is smaller than εn^2. This implies that G' still contains a triangle. Now, we will show that G' has at least δn^3 triangles for some $\delta = \delta(\varepsilon)$. As there is no edge between the irregular pairs and also there is no edge inside V_i, for all i then, three vertices of a triangle in G' must belong to the three distinct parts.

For instance, say these parts V_1, V_2, V_3 of P, then we get a triangle as follows;

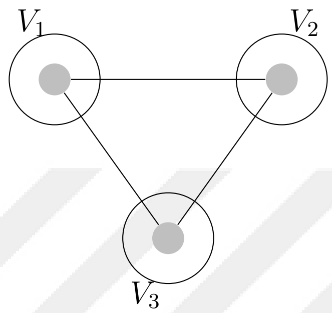

Figure 3.3 How a triangle occur between three distinct parts

Now, we will show that V_1, V_2, V_3 also support many triangles. We know that $(V_1, V_2), (V_1, V_2), (V_2, V_3)$ are ε–regular and $d(V_1, V_2), d(V_1, V_3), d(V_2, V_3)$ are all greater than or equal to $\frac{\varepsilon}{2}$ from the above deleting part. Before finishing the proof, we need the following definition.

Definition 3.2.3. *A vertex $v_1 \in V_1$ is called typical, if it has at least $\frac{\varepsilon}{4}c$ adjacent vertices in V_2 and at least $\frac{\varepsilon}{4}c$ adjacent vertices in V_3.*

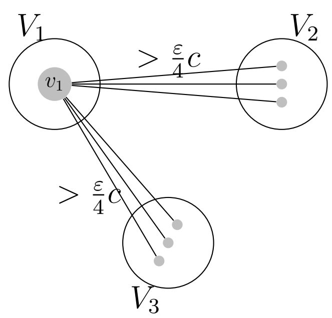

Figure 3.4 A typical vertex

As (V_i, V_j) is $\frac{\varepsilon}{4}$– regular with $d(V_i, V_j) \geq \frac{\varepsilon}{2}$ for $1 \leq i < j \leq 3$, for any $V_i^{'} \subseteq V_i$ and $V_j^{'} \subseteq V_j$ we have cardinality at least $\frac{\varepsilon}{4}c$. Then

$$|d(V_i, V_j) - d(V_i^{'}, V_j^{'})| \leq \frac{\varepsilon}{4} \tag{3.49}$$

which in turn implies that

$$\frac{\varepsilon}{4} \leq d(V_i^{'}, V_j^{'}) \leq \frac{3\varepsilon}{4} \tag{3.50}$$

holds.

Remark 2. *There are more than $\frac{c}{2}$ typical vertices in V_1.*

Proof of Remark 2. Let V_{12} be the set of all vertices in V_1 with less than $\frac{\varepsilon c}{4}$ adjacent vertices in V_2. Similarly, let V_{13} be the set of all vertices in V_1 with less than $\frac{\varepsilon c}{4}$ adjacent vertices in V_3. So, the set of non-typical elements of V_1 is a subset of $V_{12} \cup V_{13}$. This yields that

$$d(V_{12}, V_2) = \frac{e(V_{12}, V_2)}{|V_{12}||V_2|} < \frac{|V_{12}|\frac{\varepsilon c}{4}}{|V_{12}|c} = \frac{\varepsilon}{4}. \tag{3.51}$$

So, $|V_{12}| < \frac{\varepsilon c}{4}$ if not $d(V_{12}, V_2)$ must be greater than or equal to $\frac{\varepsilon}{4}$ by inequality (3.51). Similarly, $|V_{13}| < \frac{\varepsilon c}{4}$ is satisfied. Thus, the number of non-typical elements in V_1 is less than $\frac{\varepsilon c}{2}$. This implies that the number of typical elements greater than or equal to

$$c - \frac{\varepsilon c}{2} > \frac{c(2-\varepsilon)}{2} > \frac{c}{2}, \tag{3.52}$$

because $\varepsilon < 1$.

Therefore, there are more than $\frac{c}{2}$ typical vertices in V_1. This proves the remark.

Now, we will continue the proof of the triangle removal lemma. Let $v_1 \in V_1$ be a typical element and consider $V_2^{'} \subset V_2$ and $V_3^{'} \subset V_3$ where $V_2^{'}$ and $V_3^{'}$ are the set of vertices that are adjacent to v_1 and

$$|V_2^{'}| \geq \frac{\varepsilon c}{4} \text{ and } |V_3^{'}| \geq \frac{\varepsilon c}{4} \tag{3.53}$$

holds.

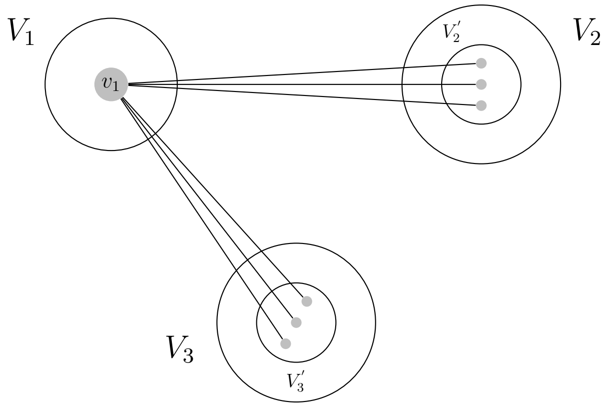

Figure 3.5 Generating triangles that contains typical vertices

Now, every edge between $V_2^{'}$ and $V_3^{'}$ will generate a triangle. To investigate these triangles, we must find the number of edges between $V_2^{'}$ and $V_3^{'}$, that is

$$e(V_2^{'}, V_3^{'}) = d(V_2^{'}, V_3^{'})|V_2^{'}||V_3^{'}| \geq \frac{\varepsilon}{4} \cdot \frac{\varepsilon c}{4} \cdot \frac{\varepsilon c}{4} = \frac{\varepsilon^3 c^2}{4^3}. \tag{3.54}$$

So, the number of triangles in $G^{'} \subset G$ is at least

$$\frac{c}{2} \cdot \frac{\varepsilon^3 c^2}{4^3} = \frac{\varepsilon^3 c^3}{2^7}. \tag{3.55}$$

We know that $P = \{V_0, V_1, \ldots, V_k\}$ with $|V_0| \leq \frac{\varepsilon}{4} n$ and so that for any i,

$$|V_i| \geq \left(n - \frac{\varepsilon n}{4}\right)\frac{1}{k} \tag{3.56}$$

is satisfied from the definition of P. This gives us

$$|V_i| = c \geq \frac{n}{k}\left(1 - \frac{\varepsilon}{4}\right) \geq \left(1 - \frac{\varepsilon}{4}\right)\frac{n}{T} \quad \text{where} \quad t \leq k \leq T. \tag{3.57}$$

So, the number of triangles in $G^{'} \subset G$ is at least

$$\frac{\varepsilon^3 c^3}{2^7} \geq \frac{\varepsilon^3 (1 - \frac{\varepsilon}{4})^3 n^3}{2^7 T^3} = \delta n^3 \tag{3.58}$$

where $\delta = \frac{\varepsilon^3 (1 - \frac{\varepsilon}{4})^3}{2^7 T^3}$. Thus, G contains at least δn^3 triangles. $\square$

3.3 Roth's Theorem

In this section, we will give the proof of Roth's theorem.

Roth's Theorem. (*Lima (2011)*) *For all $\varepsilon > 0$, there exists $N = N(\varepsilon)$ such that for all $n \geq N$, if $A \subseteq \{1, 2, \ldots, n\}$ with $|A| \geq \varepsilon n$, then A contains a non-trivial three term arithmetic progression.*

Proof. Let $\varepsilon > 0$ and $A \subseteq \{1, 2, \ldots, n\}$ with $|A| \geq \varepsilon n$. Now, we will construct a 3– partite graph $G = (V, E)$ and apply the triangle removal lemma. Let G be a 3– partite graph with vertex set $V = V_1 \cup V_2 \cup V_3$ where $V_1 = \{1, 2, \ldots, n\}$, $V_2 = \{1, 2, \ldots, 2n\}$ and $V_3 = \{1, 2, \ldots, 3n\}$ with V_1, V_2, V_3 are pairwise disjoint. This implies that $|V| = 6n$. Then, define edges in G as follows:

1. There is an edge from $i \in V_1$ to $j \in V_2$ if and only if $j - i \in A$.
2. There is an edge from $j \in V_2$ to $k \in V_3$ if and only if $k - j \in A$.
3. There is an edge from $i \in V_1$ to $k \in V_3$ if and only if $\frac{k-i}{2} \in A$.

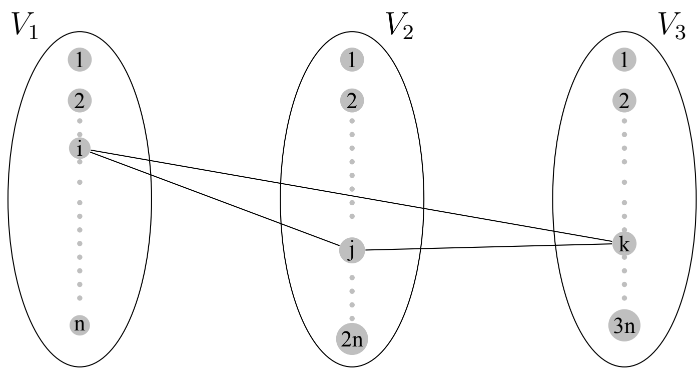

Figure 3.6 Relation between triangles and arithmetic progressions

This implies that $(i,j,k) \in V_1 \times V_2 \times V_3$ defines a triangle in G if and only if $j-i \in A$, $k-j \in A$ and $\frac{k-i}{2} \in A$. In other words, $(i,j,k) \in V_1 \times V_2 \times V_3$ defines a triangle in G if and only if $\{j-i, \frac{k-i}{2}, k-j\}$ is a three term arithmetic progression in A. A triangle in G is formed by $(i,j,k) \in V_1 \times V_2 \times V_3$ yields to a trivial arithmetic progression if and only if $j-i = \frac{k-i}{2} = k-j = a$, for some $a \in A$ if and only if $(i,j,k) = (i, i+a, i+2a)$, for some $i \in \{1,2,\ldots,n\}$ and $a \in A$.

Now, call a triangle in G as *trivial triangle* if it yields to a trivial arithmetic progression in A. Each trivial triangle corresponds to a pair $(i,a) \in \{1,2,\ldots,n\} \times A$. Hence, the number of trivial triangles is equal to $n|A| \geq \varepsilon n^2$.

Note that there are no two trivial triangles sharing an edge. So we can not delete two trivial triangles by deleting an edge.

This implies that G is at least $\varepsilon n^2 = \frac{\varepsilon}{36}(6n)^2$- far from being triangle free. Then, by the Triangle Removal lemma for $\frac{\varepsilon}{36}$, there exists $\delta_0 = \delta_0(\varepsilon)$ and $N_0 = N_0(\varepsilon)$ such that whenever $|G| > N_0$ and G is $\frac{\varepsilon}{36}|G|^2$– far from being triangle free, then G contains at least $\delta_0|G|^3$-many triangles. So, we get $|G| = 6n > N_0$ where $n > N_0/6$. Also, G contains at least

$$\delta_0(6n)^3 = 216\delta_0 n^3 = \delta n^3 \tag{3.59}$$

many triangles where $216\delta_0 = \delta$. Then, we have at least $\delta n^3 - n^2$ many triangles which are non-trivial. Thus, if $\delta n^3 - n^2 > 1$ holds, then we can find a non-trivial triangle:

$$n^2(\delta n - 1) > 1 \Leftarrow \delta n - 1 > 1 \Leftarrow n > \frac{2}{\delta} \tag{3.60}$$

So, there exists a non-trivial triangle, if $n > \frac{2}{\delta}$. Equivalently, A contains a non-trivial three term arithmetic progression, when $N >$ max$\{\frac{N_0}{6}, \frac{2}{\delta}\}$. $\square$

CHAPTER FOUR
ROTH'S THEOREM ON ARITHMETIC PROGRESSIONS VIA FOURIER ANALYSIS

In this chapter, we will prove Roth's theorem via techniques of Fourier analysis.

Let $X \subseteq \{1, 2, \ldots, N\}$ and $|X| = dN$. For some technical reasons, we will want 2 to be invertible in $\mathbb{Z}_N$, so we assume that N is odd. This does not cause any problem because if N is even, we can take $N + 1$ instead of N. Indeed, the density of X in $\{1, 2, \ldots, N, N + 1\}$ is very close to the density of X in $\{1, 2, \ldots, N\}$. Also, we will consider X as a subset of $\mathbb{Z}_N$. Hence, we can use the tools of Fourier analysis that we introduced in Chapter 1.

Our proof has two main parts, which depend on the magnitude of Fourier coefficients. The first one is that if all Fourier coefficients $|\widehat{X}(m)|$ of the characteristic function $X(x)$ are sufficiently small for all $m \neq 0$, then we can find a non-trivial three term arithmetic progression directly in the set X. On the other hand, if there exists $m \neq 0$ such that $|\widehat{X}(m)|$ is large, then we can find an arithmetic progression set P in $\mathbb{Z}_{\mathbb{N}}$, and using it and some density arguments, we obtain a real three term arithmetic progression in $X \subseteq \mathbb{Z}$.

Remark 3. *Note that* $\{a, b, c\}$ *is a* 3*-AP in* $\mathbb{Z}_N$*, if* $a + c \equiv 2b \pmod{N}$*. The point that we have to be careful about is that arithmetic progressions in* $\mathbb{Z}_N$ *are not necessarily progressions in* $\mathbb{Z}$ *(they might 'wrap around'). For instance in* $\mathbb{Z}_{102}$*,* $\{65, 100, 33\}$ *is a* 3*-term arithmetic progression but not in* $\mathbb{Z}$*.*

So, we need to be sure that we count 3-APs in $\mathbb{Z}$. To guarantee this, we define a set

$$Y = X \cap \left[\frac{N}{3}, \frac{2N}{3}\right]. \tag{4.1}$$

Note that, if $a, b \in Y$ and $c = 2b - a$ then, $\{a, b, c\}$ is a 3AP in $\mathbb{Z}$.

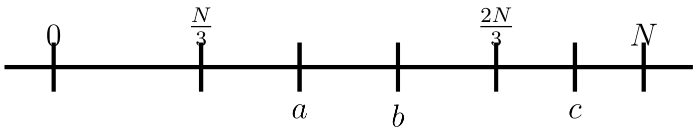

Figure 4.1 How to choose $\{a, b, c\} \in [0, N]$

Since $a, b \in Y$, $|a-b| \leq \frac{N}{3}$ and $|a-b| = |b-c|$. So, $|b-c| \leq \frac{N}{3}$. Hence

$$|a-c| \leq |a-b| + |b-c| \leq \frac{2N}{3} \leq N. \tag{4.2}$$

This implies that $\{a, b, c\}$ is a 3-AP in $\mathbb{Z}$. In other words, there is no wrapping around. If we choose a and b in such a way that

$$|a-b| > \frac{N}{3} \quad \text{and} \quad |b-c| > \frac{N}{3}, \tag{4.3}$$

then, c will wrap around. That is $\{a, b, c\}$ does not represent a 3AP in $\mathbb{Z}$.

Now, let us denote the number of 3APs contained in X by ϑ. Then, we can count ϑ as follows

$$\vartheta \geq \#\Big\{(a, b, c) \in Y \times Y \times X : a + c \equiv 2b \pmod N\Big\} \tag{4.4}$$

$$= \sum_{a,b,c,\, a+c\equiv 2b} Y(a)Y(b)X(c). \tag{4.5}$$

By Proposition (2.3.1), we know that $\sum_{m} \chi(xm) = N$ if and only if $x = 0$. Then we obtain that

$$\vartheta \geq N^{-1} \sum_{a,b,c} Y(a)Y(b)X(c) \sum_{r \in \mathbb{Z}_N} \chi(-r(a + c - 2b)) \tag{4.6}$$

$$= N^{-1} \sum_{r} \sum_{a} Y(a)\chi(-ra) \sum_{b} Y(b)\chi(2rb) \sum_{c} X(c)\chi(-rc). \tag{4.7}$$

Using the Fourier transform, we have that $\sum_x f(x)\chi(-rx) = N\widehat{f}(r)$. Then

$$\vartheta \geq N^2 \sum_r \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r) \tag{4.8}$$

$$= N^2\widehat{Y}(0)\widehat{Y}(0)\widehat{X}(0) + N^2 \sum_{r\neq 0} \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r). \tag{4.9}$$

As

$$\widehat{Y}(r) = N^{-1} \sum_{x\in\mathbb{Z}_\mathbb{N}} \chi(-xr)Y(x) \tag{4.10}$$

then for $r = 0$

$$\widehat{Y}(0) = N^{-1} \sum_{x\in\mathbb{Z}_\mathbb{N}} Y(x) = N^{-1}|Y|. \tag{4.11}$$

This implies that

$$\vartheta \geq N^{-1}|Y|^2|X| + N^2 \sum_{r\neq 0} \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r). \tag{4.12}$$

Now, let us denote $R = \sum_{r\neq 0} \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r)$. So,

$$\vartheta \geq N^{-1}|Y|^2|X| + R. \tag{4.13}$$

At this point, we may assume that $|Y| \geq \frac{|X|}{5}$. If not, then either $X \cap \left[0, \frac{N}{3}\right]$ or $X \cap \left[\frac{2N}{3}, N\right]$ must have size at least $\frac{2|X|}{5}$.

Let us assume that

$$\left|X \cap \left[0, \frac{N}{3}\right]\right| \geq \frac{2|X|}{5}, \tag{4.14}$$

then

$$\frac{\left|X \cap [0, \frac{N}{3}]\right|}{N/3} \geq \frac{2/5|X|}{N/3} = \frac{6}{5}\frac{|X|}{N} = \frac{6d}{5} > d. \tag{4.15}$$

Hence, the relative density is at least $\frac{6d}{5} > d$. Thus, we may replace X by $X \cap \left[0, \frac{N}{3}\right]$ and $\{1, 2, \ldots, N\}$ by $[1, \frac{N}{3}]$ and repeat the same argument for these. Clearly if we find a 3-AP in the set X that has density d, then we can find a 3-AP in the set $X \cap \left[0, \frac{N}{3}\right]$ that has density $\frac{6d}{5}$.

4.1 Small Fourier Coefficients

In this section, we will give the first part of the proof of Roth's theorem. When all the non-zero Fourier coefficients of X are small, we will use some estimations on R to show the existence of a non-trivial 3-AP in the set X.

Theorem 4.1.1. *(Lott (2017)) If* $|\widehat{X}(r)| < \frac{d^2}{10}$, *for all* $r \neq 0$, *then* A *contains a non-trivial* 3-*AP.*

Proof. We prove this theorem by showing that the remainder term R in $\vartheta \geq N^{-1}|Y|^2|X| + R$ is small. We know that

$$R = N^2 \sum_{r \neq 0} \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r). \tag{4.16}$$

Then we have that

$$\begin{aligned} |R| &= N^2 \Big| \sum_{r \neq 0} \widehat{Y}(r)\widehat{Y}(-2r)\widehat{X}(r) \Big| &(4.17)\\ &\leq N^2 \max_{r \neq 0} |\widehat{X}(r)| \cdot \Big| \sum_{r \neq 0} \widehat{Y}(r)\widehat{Y}(-2r) \Big|. &(4.18) \end{aligned}$$

By the Cauchy-Schwarz inequality,

$$|R| \leq N^2 \max_{r \neq 0} |\widehat{X}(r)| \left(\sum_r |\widehat{Y}(r)|^2 \right)^{1/2} \left(\sum_r |\widehat{Y}(-2r)|^2 \right)^{1/2}. \tag{4.19}$$

Then, by Plancherel's identity (2.24),

$$|R| \leq N^2 \max_{r \neq 0} |\widehat{X}(r)| \, N^{-1/2} \left(\sum_x |Y(x)|^2 \right)^{1/2} N^{-1/2} \left(\sum_x |Y(-2x)|^2 \right)^{1/2}. \tag{4.20}$$

Since $\sum_x |Y(x)|^2 = |Y|$, then we have that

$$|R| \leq N \max_{r \neq 0} |\widehat{X}(r)||Y|. \tag{4.21}$$

Since $|\widehat{X}(r)| < \frac{d^2}{10}$, then

$$|R| \leq N\frac{d^2}{10}|Y|. \tag{4.22}$$

Substituting $d = \frac{|X|}{N}$ into the last inequality, one sees that

$$\begin{aligned} |R| &\leq N\frac{|X|^2}{N^2}\frac{|Y|}{10} && (4.23)\\ &= N^{-1}\frac{|X|}{2}\frac{|X|}{5}|Y|. && (4.24) \end{aligned}$$

As $|Y| \geq \frac{|X|}{5}$, then we get that

$$|R| \leq N^{-1}\frac{|X|}{2}|Y|^2. \tag{4.25}$$

Since $\vartheta \geq N^{-1}|Y|^2|X| + R$, then

$$\begin{aligned} \vartheta &\geq |N^{-1}|Y|^2|X| - |R|| && (4.26)\\ &\geq N^{-1}|Y|^2|X| - \left|\frac{1}{2}N^{-1}|Y|^2|X|\right| && (4.27)\\ &\geq \frac{1}{2}N^{-1}|Y|^2|X|. && (4.28) \end{aligned}$$

As $|Y| \geq \frac{|X|}{5}$ and $|X| = dN$, then we get that

$$\begin{aligned} \vartheta &\geq \frac{1}{2}N^{-1}\frac{|X|^3}{5^2} && (4.29)\\ &= \frac{d^3N^2}{50}. && (4.30) \end{aligned}$$

Finally, we have found a lower bound for the ϑ. This bound contains the number of trivial progressions, but we search for only non-trivial arithmetic progressions. So we must subtract those off, but there are only $|X| = dN$ of them. So the number of non-trivial 3-APs contained in X is at least $\frac{d^3N^2}{50} - dN$ which is positive if $N > \frac{50}{d^2}$. Thus, for sufficiently large N, the set X contains a non-trivial 3-AP. □

4.2 Large Fourier Coefficients

In this section, we will give the second part of the proof of Roth's theorem. If X has a large Fourier coefficient, then the estimate on R is not applicable. So we must use a different method to count arithmetic progressions in the set X. To do so, we will work with the function $g(x) = X(x) - d$. Let us describe some properties of $g(x)$.

Proposition 2. *The function $g(x)$ has the following properties:*

(a)

$$\sum_{x \in \mathbb{Z}_N} g(x) = 0. \tag{4.31}$$

(b) $\widehat{g}(r) = \widehat{X}(r)$ *for all* $r \neq 0$ *and* $\widehat{g}(0) = 0$.

Proof.

(a) From the definition of the function $g(x)$, we have that

$$\sum_{x \in \mathbb{Z}_N} g(x) = \sum_{x \in \mathbb{Z}_N} (X(x) - d) = \sum_{x \in \mathbb{Z}_N} X(x) - \sum_{x \in \mathbb{Z}_N} d. \tag{4.32}$$

Clearly $\sum_x X(x) = |X|$ and $\sum_x d = dN$, then we get that

$$\sum_{x \in \mathbb{Z}_N} g(x) = |X| - dN. \tag{4.33}$$

Initially we set $|X| = dN$. This gives us

$$\sum_{x \in \mathbb{Z}_N} g(x) = 0. \tag{4.34}$$

(b) Using the Fourier transformation, we get that

$$\widehat{g}(r) = N^{-1}\sum_{x\in\mathbb{Z}_N}\chi(-xr)g(x) = N^{-1}\sum_{x}\chi(-xr)(X(x)-d) \tag{4.35}$$

$$= N^{-1}\sum_{x}\chi(-xr)X(x) - dN^{-1}\sum_{x}\chi(-xr). \tag{4.36}$$

Now, we have two cases. If $r \neq 0$, then

$$\widehat{g}(r) = N^{-1}\sum_{x\in\mathbb{Z}_N}\chi(-xr)X(x) = \widehat{X}(r) \tag{4.37}$$

because $\sum_{x}\chi(-xr) = 0$ for $r \neq 0$.
Otherwise if $r = 0$ then,

$$\widehat{g}(0) = N^{-1}\sum_{x\in\mathbb{Z}_N} g(x)\chi(0) = N^{-1}\sum_{x\in\mathbb{Z}_N} g(x) = 0, \tag{4.38}$$

using part (a) in the last equality above. □

Now, we will prove an essential theorem for the large coefficient part.

Theorem 4.2.1. *(Lott (2017)) Suppose that $|\widehat{X}(m)| > \varepsilon > 0$ for some $m \neq 0$. Then, there exists an arithmetic progression set Q satisfying that*

$$|Q| \geq \left(\frac{\varepsilon}{64}\right)\sqrt{N} \quad \textit{and} \quad |X\cap Q| \geq \left(d+\frac{\varepsilon}{8}\right)|Q|. \tag{4.39}$$

Proof. We will prove this in three stages. First, we construct a long arithmetic progression Q such that $\widehat{Q}(m)$ is also large. Second, we show that the set X has a large intersection with some translate of Q, say it $Q^{'}$. Third, we lift the $\mathbb{Z}_N$ progression $Q^{'}$ to a $\mathbb{Z}$ progression $Q^{''}$ without losing too much of its length or intersection with X. Suppose that $|\widehat{X}(m)| > \varepsilon > 0$ for some $m \neq 0$. Consider the set of points

$$\{(0,0),(1,m),(2,2m),\ldots,(N-1,(N-1)m)\} \subseteq \{1,\ldots,N-1\}\times\{1,\ldots,N-1\}. \tag{4.40}$$

We divide $\{1, \ldots, N-1\} \times \{1, \ldots, N-1\}$ into a $\lfloor\sqrt{N-1}\rfloor \times \lfloor\sqrt{N-1}\rfloor$ grid. Then, there are less than N total boxes and exactly N points, so by the pigeonhole principle there are two points that are in the same box. i.e., there exist α, β such that both $\alpha - \beta \leq \sqrt{N}$ and $m(\alpha - \beta) \leq \sqrt{N}$ when considered modulo N. That is, $(\alpha, m\alpha)$ and $(\beta, m\beta)$ are in the same box.

For instance, take $N = 6$, $m = 2$, then $\lfloor\sqrt{N-1}\rfloor = 2$, and we get the following figure:

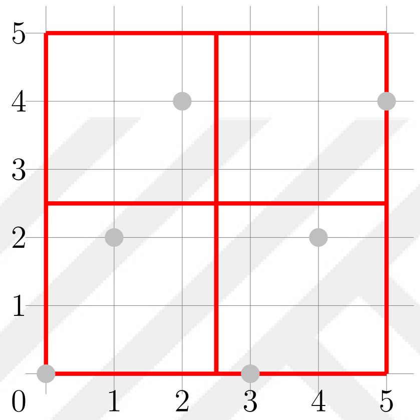

Figure 4.2 This figure shows that "$(0,0)$ and $(1,2)$" and "$(3,0)$ and $(4,2)$" are in the same box.

Now let $k = \alpha - \beta$ and Q be the progression $\{\ldots, -2k, -k, 0, k, 2k, \ldots\}$ of length $\lfloor\frac{\sqrt{N}}{8}\rfloor$ where the end points of Q are chosen to be as symmetric around 0 as possible. Now we want $|\widehat{Q}(m)|$ to be large. Consider

$$|N\widehat{Q}(m) - |Q|| = \left|NN^{-1}\sum_{x\in\mathbb{Z}_N} Q(x)\chi(-xm) - |Q|\right| \tag{4.41}$$

$$= \left|\sum_{x\in\mathbb{Z}_N} Q(x)\chi(-xm) - Q(x)\right| \tag{4.42}$$

$$= \left|\sum_{x\in\mathbb{Z}_N} Q(x)(\chi(-xm) - 1)\right|. \tag{4.43}$$

As $Q(x)$ is a characteristic function of the set Q, then we get that

$$|N\widehat{Q}(m) - |Q|| = \left| \sum_{x \in Q} \chi(-xm) - 1 \right| \tag{4.44}$$

$$\leq \sum_{x \in Q} |\chi(-xm) - 1|. \tag{4.45}$$

We know that elements of Q are of the form kl, then we have

$$|N\widehat{Q}(m) - |Q|| \leq \sum_{kl \in Q} |\chi(-mkl) - 1|. \tag{4.46}$$

$$\tag{4.47}$$

Since Q is as symmetric as possible and the elements of Q depend on only the value of l, we deduce that

$$|N\widehat{Q}(m) - |Q|| \leq \sum_{|l| \leq \frac{|Q|}{2}} |\chi(mkl) - 1| \tag{4.48}$$

$$\leq |Q| \max_{|l| \leq \frac{|Q|}{2}} |\chi(mkl) - 1|. \tag{4.49}$$

As $|Q| \leq \frac{\sqrt{N}}{8}$, $|l| \leq \frac{|Q|}{2}$ and $mk \leq \sqrt{N}$, we have that

$$mkl \leq \sqrt{N}\left(\frac{|Q|}{2}\right) \leq \frac{N}{16}. \tag{4.50}$$

As we know $\chi(mkl)$ is an N^{th} root of unity, we get that

$$\chi(mkl) = e^{\frac{2\pi imkl}{N}} \leq e^{\frac{2\pi i}{16}}. \tag{4.51}$$

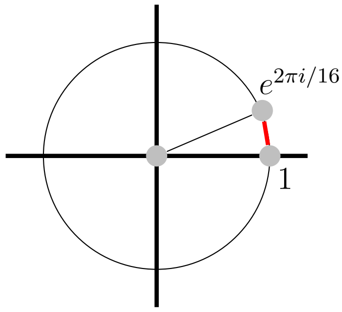

Figure 4.3 Arc length of unit circle

Then the difference between $\chi(mkl)$ and 1 is at most the length of the arc between $e^{\frac{2\pi i}{16}}$ and 1.

Hence, for any l

$$|\chi(mkl) - 1| \leq \frac{2\pi}{16} < \frac{1}{2}. \tag{4.52}$$

We conclude that

$$|N\widehat{Q}(m) - |Q|| < \frac{|Q|}{2}. \tag{4.53}$$

So

$$-\frac{|Q|}{2} < N\widehat{Q}(m) - |Q| < \frac{|Q|}{2}. \tag{4.54}$$

This implies that

$$\widehat{Q}(m) > \frac{1}{2}N^{-1}|Q|. \tag{4.55}$$

In the next step, we show that the set X has a large intersection with some translate of Q. Let us define a function $F : \mathbb{Z}_N \to \mathbb{C}$ such that

$$F(x) = (g * Q)(x) = \sum_y g(y)Q(x-y) \tag{4.56}$$

where $g(x) = X(x) - d$ and $Q(x)$ is the characteristic function of the set Q.

At this point of the proof, we need the following fact:

Fact 1.

$$\sum_x F(x) = 0 \tag{4.57}$$

Proof of Fact 1. From the definition of the function F and Proposition (2), we get that

$$\sum_x F(x) = \sum_x \sum_y g(y)Q(x-y) = \sum_y g(y) \sum_x Q(x-y) = |Q| \sum_y g(y) = 0. \tag{4.58}$$

This proves the Fact 1.

Then, by pointwise estimate and convolution identity in (2.3.2), we get that

$$N^{-1} \sum_x |F(x)| \geq |\widehat{F}(m)| = N|\widehat{g}(m)||\widehat{Q}(m). \tag{4.59}$$

We know that $|\widehat{g}(m)| = |\widehat{X}(m)|$ and $|\widehat{Q}(m)| \geq |\frac{1}{2}N^{-1}|Q||$ from the first part. Then,

$$N^{-1} \sum_x |F(x)| \geq N|\widehat{X}(m)|\frac{1}{2}N^{-1}|Q| \geq \varepsilon\frac{|Q|}{2}. \tag{4.60}$$

Hence

$$\sum_x |F(x)| \geq N\varepsilon\frac{|Q|}{2}. \tag{4.61}$$

Then, by the above fact, we have

$$\sum_{x\in\mathbb{Z}_N} |F(x)| + F(x) \geq \frac{1}{2}N\varepsilon\frac{|Q|}{2}. \tag{4.62}$$

So, for some $x \in \mathbb{Z}_N$

$$|F(x)| + F(x) \geq \frac{1}{2}\varepsilon\frac{|Q|}{2}, \tag{4.63}$$

which implies that

$$F(x) \geq \frac{1}{4}\varepsilon|Q|. \tag{4.64}$$

Then, by the definition of $F(x)$, we see that

$$\frac{1}{4}\varepsilon|Q| \leq \sum_y g(y)Q(x-y) = \sum_y X(y)Q(x-y) - d\sum_y Q(x-y) \tag{4.65}$$

$$\leq |X \cap (x-Q)| - d|(x-Q)|. \tag{4.66}$$

As $|Q| = |(x - Q)|$,

$$|X \cap (x - Q)| \geq \big(d + \frac{1}{4}\varepsilon\big)|(x - Q)|, \tag{4.67}$$

which proves the second step.

In the final stage, we consider $Q^{'} = (x - Q)$. In this part, we will carry the $\mathbb{Z}_N$ progression $Q^{'}$ to a $\mathbb{Z}$ progression $Q^{''}$ without losing too much length or intersection with X.

As Q is a set of arithmetic progression in $\mathbb{Z}_{\mathbb{N}}$, then also $Q^{'}$ is. Nevertheless, it may not be a set of arithmetic progression in $\mathbb{Z}$. So, we need to find an arithmetic progression in $\mathbb{Z}$ contained in $Q^{'}$ that is still long and has a large intersection with X. We will find this $Q^{''}$ in the following way. As $|Q| = |Q^{'}| \leq \sqrt{N}$, then the common difference of elements in $Q^{'}$ is at most $\sqrt{N}$. So, the last term of $Q^{'}$ does not wrap around on the first term of $Q^{'}$ modulo N. Then, we can divide the set $Q^{'}$ into two sets from the first wrapping point in $Q^{'}$. Hence, $Q^{'} = Q_1 \cup Q_2$ where Q_1 and Q_2 are disjoint arithmetic progressions in $\mathbb{Z}$.

Now, without loss of generality, suppose that $|Q_1| \leq |Q_2|$. If $|Q_1| \leq \frac{1}{8}\varepsilon|Q^{'}|$ then,

$$|X \cap Q_2| \geq |X \cap Q^{'}| - |Q_1| \geq \left(d + \frac{1}{4\varepsilon}\right)|Q^{'}| - \frac{1}{8}\varepsilon|Q^{'}| \tag{4.68}$$

$$\geq \left(d + \frac{1}{8}\varepsilon\right)|Q^{'}|. \tag{4.69}$$

On the other hand, if $|Q_1| \geq \frac{1}{8}\varepsilon|Q^{'}|$, then both Q_1 and Q_2 have length at least $\frac{1}{8}\epsilon|Q^{'}|$. Since $Q^{'} = Q_1 \cup Q_2$, then X must have density at least $d + \frac{1}{4}\varepsilon$ in one of them. If X has density less than $d + \frac{1}{4}\varepsilon$ in each of Q_1 and Q_2 then, we get that

$$|X \cap Q_1| < \left(d + \frac{1}{4}\varepsilon\right)|Q_1| \quad and \quad |X \cap Q_2| < \left(d + \frac{1}{4}\varepsilon\right)|Q_2|. \tag{4.70}$$

Then,

$$|X \cap Q_1| + |X \cap Q_2| < \left(d + \frac{1}{4}\varepsilon\right)(|Q_1| + |Q_2|). \tag{4.71}$$

We know $|Q_1| + |Q_2| = |Q^{'}|$ so,

$$|X \cap Q^{'}| < \left(d + \frac{1}{4}\varepsilon\right)|Q^{'}| \tag{4.72}$$

which contradicts the previous step. Thus, we can choose $Q^{''}$ as one of Q_1 and Q_2 that satisfies the following inequality

$$|X \cap Q_i| \geq \left(d + \frac{1}{8}\varepsilon\right)|Q^{'}|. \tag{4.73}$$

Therefore, we have established the existence of an arithmetic progression $Q^{''}$ in $\mathbb{Z}$ such that

$$|Q^{''}| \geq \frac{1}{8}|Q^{'}| \text{ and } |X \cap Q^{''}| \geq \left(d + \frac{1}{8}\varepsilon\right)|Q^{''}|. \tag{4.74}$$

$\square$

4.3 Completing the Proof of Roth's Theorem via Density Increment

Suppose X is a subset of $\{1, 2, \ldots, N\}$ and X does not contain a 3-AP. Then, by theorem (4.1.1), for some $m \neq 0$, $|\widehat{X}(m)| \geq \frac{d^2}{10}$. Then in Theorem (4.2.1), taking $\varepsilon = \frac{d^2}{10}$, there exists an arithmetic progression Q_1 such that $|Q_1| \geq \frac{d^2}{640}\sqrt{N}$ and $|X \cap Q_1| \geq \left(d + \frac{d^2}{80}\right)|Q_1|$.

Now, let us take $X_1 = X \cap Q_1$. As arithmetic progressions are preserved under affine transformations, then we can identify Q_1 with $\{1, 2, \ldots, |Q_1|\}$ and X_1 with a subset of $\{1, 2, \ldots, |Q_1|\}$ of density $d_1 \geq d + \frac{d^2}{80}$. Clearly, X_1 contains no 3-AP, too.

We may repeat the same argument for X_1. Then, by Theorem (4.1.1), there exists $m_1 \neq 0$ such that $|\widehat{X_1}(m_1)| \geq \frac{d_1^2}{10}$. Then, by Theorem (4.2.1), we can obtain another progression Q_2 such that $|Q_2| \geq \frac{d_1^2}{640}|Q_1|^{1/2}$ and $|X_1 \cap Q_2| \geq (d_1 + \frac{d_1^2}{80})|Q_2|$. Let $X_2 = X_1 \cap Q_2$ where the density of X_2 in Q_2 is $d_2 \geq d_1 + \frac{d_1^2}{80}$, also X_2 contains no 3-AP.

Repeating this process we get a sequence of progressions Q_k and subsets X_k with densities d_k satisfying the following

$$|Q_k| \geq \frac{d_{k-1}^2}{640}\sqrt{|Q_{k-1}|} \quad \text{and} \quad d_k \geq d_{k-1} + \frac{d_{k-1}^2}{80}. \tag{4.75}$$

Using the above recurrence relation of d_k, then we deduce that

$$\begin{aligned}
d_k &\geq d_{k-1} + \frac{d_{k-1}^2}{80} \geq d_{k-2} + \frac{d_{k-2}^2}{80} + \frac{d_{k-2}^2}{80} &(4.76)\\
&\geq d_{k-3} + \frac{d_{k-3}^2}{80} + \frac{d_{k-3}^2}{80} + \frac{d_{k-3}^2}{80} &(4.77)\\
&\vdots &(4.78)\\
&\geq d + k\left(\frac{d^2}{80}\right). &(4.79)
\end{aligned}$$

Thus after $k = \frac{80}{d}$ iterations, we have that $d_k \geq 2d$. Then for $k > \frac{80}{d}$, we have that

$$\begin{aligned}
&d_{k+1} \geq d_k + \frac{d_k^2}{80} \geq 2d + \frac{(2d)^2}{80} &(4.80)\\
&d_{k+2} \geq d_{k+1} + \frac{d_{k+1}^2}{80} \geq d_k + \frac{d_k^2}{80} + \frac{d_k^2}{80} \geq 2d + \frac{(2d)^2}{80} + \frac{(2d)^2}{80} = 2d + 2\frac{(2d)^2}{80}. &(4.81)
\end{aligned}$$

So,

$$d_{k+\delta} \geq 2d + \delta\left(\frac{(2d)^2}{80}\right). \tag{4.82}$$

Then after more than $\frac{k}{2} = \frac{80}{2d}$ iterations, we have that

$$d_{k+\frac{k}{2}} \geq 2d + \frac{k}{2}\left(\frac{(2d)^2}{80}\right) = 2d + \frac{80}{2d}\left(\frac{(2d)^2}{80}\right) = 4d. \tag{4.83}$$

This means that after $\frac{k}{2} = \frac{80}{2d}$ more iterations, our density has increased from $2d$ to $4d$. Then

$$d_{k+\frac{k}{2}+\frac{k}{4}} \geq 4d + \frac{k}{4}\left(\frac{(4d)^2}{80}\right) = 4d + \frac{80}{4d}\frac{(4d)^2}{80} = 8d = 2^3 d. \tag{4.84}$$

Continuing in this way, one may prove:

$$d_{k+\frac{k}{2}+\cdots+\frac{k}{2^a}} \geq 2^{a+1} d. \tag{4.85}$$

In general, the density will have increased to $2^a d$ after more than

$$\frac{80}{d}\left(1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^a}\right) \leq \frac{160}{d} \tag{4.86}$$

iterations. Indeed,

$$\frac{80}{d}\left(1+\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^a}\right) \leq \frac{80}{d}\left(1+\frac{1}{2}+\dots\right) \leq \frac{80}{d}\frac{1}{1-1/2} = \frac{160}{d}. \tag{4.87}$$

Thus, picking a sufficiently large, we can see that the density has increased so that it passed 1 after finitely many iterations, which is a contradiction. This completes the proof of Roth's theorem.

CHAPTER FIVE
CONCLUSION

In this thesis, we aimed to give two different proofs for Roth's theorem on arithmetic progressions. Firstly, in graph theoretic part, our main goal was to prove Szemerédi's Regularity lemma. Besides, we proved the Triangle Removal lemma. Then, these two theorems gave us the proof of Roth's theorem. In the Fourier analytic part, we first stated our main tools, namely the Fourier transform and some of its properties. Then, we divided the proof into two parts. In the small Fourier coefficient part, we used the Fourier transform of the characteristic functions. Then, we could count three term arithmetic progressions with a basic combinatorial method. In the large Fourier coefficient part, we used the density increment argument. Finally, we obtained a contradiction, and this completed the proof of Roth's theorem on arithmetic progressions.

REFERENCES

Bergelson, V., & Leibman, A. (1996). Polynomial extensions of van der Waerden's and Szemerédi's theorems. *Journal of the American Mathematical Society*, *9*(3), 725–753.

Bloom, T. F., & Sisask, O. (2021). Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions. *arXiv.2007.03528*.

Diestel, R. (2018). *Graph Theory*. Graduate Texts in Mathematics. Springer Berlin Heidelberg.

Furstenberg, H. (1977). Ergodic behavior of diagonal measures and a theorem of Szemeredi on arithmetic progressions. *Journal d'Analyse Mathématique*, *31*, 204–256.

Gowers, W. T. (2001). A new proof of Szemerédi's theorem. *Geometric and Functional Analysis*, *11*(3), 465–588.

Green, B., & Tao, T. (2008). The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics (2)*, 481–547.

Leibman, A. (1998). Multiple recurrence theorem for measure preserving actions of a nilpotent group. *Geometric & Functional Analysis GAFA*, *8*(5), 853–931.

Lima, Y. (2011). Szemerédi's regularity lemma. https://matheuscmss.wordpress.com/2011/12/24/szemeredis-regularity-lemma

Lott, A. (2017). Roth's theorem on arithmetic progressions. https://people.math.rochester.edu/faculty/iosevich/lott17.pdf

Roth, K. F. (1953). On certain sets of integers. *Journal of the London Mathematical Society*, *28*, 104–109.

Szemerédi, E. (1975). On sets of integers containing no k elements in arithmetic progression. *Acta Arithmetica*, *27*, 199–245.

Tao, T., & Ziegler, T. (2008). The primes contain arbitrarily long polynomial progressions. *Acta Mathematica*, *201*(2), 213–305.

van der Waerden, B. L. (1927). Beweis einer Baudetschen Vermutung. *Nieuw Archief voor Wiskunde. (2)*, *15*, 212–216.