BACHELOR'S THESIS

Degree in Mathematics

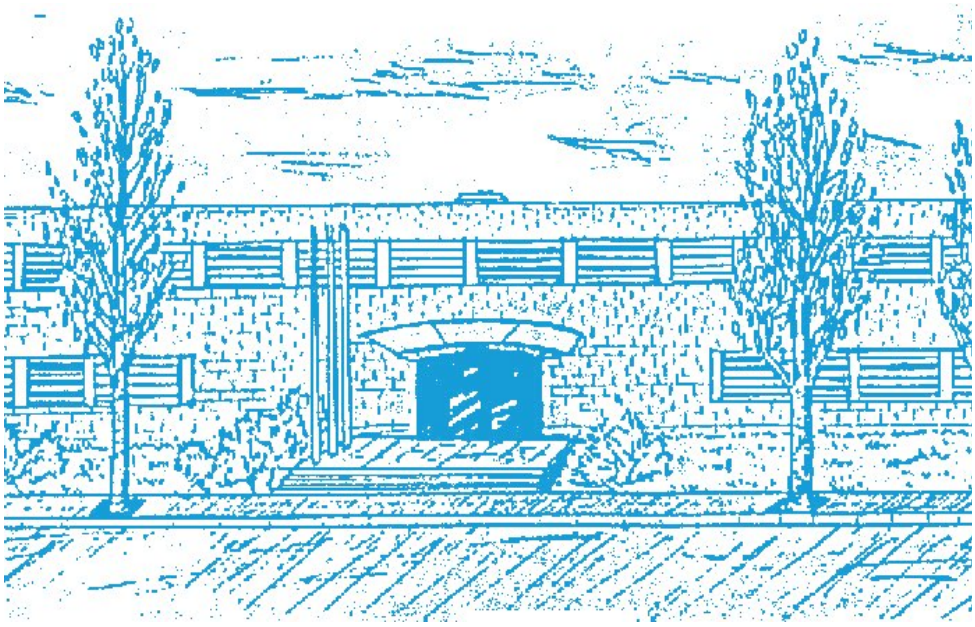
Title: Roth's Theorem: graph theoretical, analytic and combinatorial proofs

Author: García Hernández, Sílvia

Advisor: Rué Perna, Juanjo

Department: Department of Mathematics

Academic year: 2021-2022



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

Roth's Theorem: graph theoretical, analytic and combinatorial proofs

Author: Sílvia García Hernández

Advisor: Juanjo Rué Perna

Facultat de Matemàtiques i Estadística
Universitat Politècnica de Catalunya
Q1 2021-2022

I would like to thank my tutor Juanjo for guiding me through this thesis, for helping me to understand all the difficult concepts and for always being so patient.

Also, I would like to thank Edgar for always being there, supporting and helping me.

Abstract. In 1936, Erdős–Turán conjectured that any set of integers with positive upper density contains arbitrarily long arithmetic progressions. In 1953, Klaus Roth resolved this conjecture for progressions of length three. This theorem, known as Roth's Theorem, is the main topic of this thesis.

In this dissertation we will understand, rewrite and collect some of the proofs of Roth's Theorem that have appeared over the years, while developing some of the problems that arise in each area. This includes the original Fourier analytic proof due to Roth (in a more modern language), the combinatorial proof due to Szemerédi, and finally, the graph theoretical proof based on Szemerédi's Regularity Lemma. We will also explore recent progress around this theorem, as the finite field analogue and the recent breakthrough concerning upper bounds for the cap set problem.

Contents

Introduction

Ramsey theory is a branch of mathematics that focuses around the idea that *any large enough structure will necessarily contain an orderly substructure*. Problems found in this area ask questions such as "how many elements an structure must contain in order to guarantee a particular property?". Founded in 1930, Ramsey theory had already emerge with *Hilbert's cube lemma* [15] in 1892, but it is not until years later that people got influenced. One of the first important results of Ramsey theory appears in 1927, when Van der Waerden proved the *Van der Waerden's Theorem* [27], which entails with it the study of arithmetic progressions: the sequences of numbers such that the difference between the consecutive terms is constant.

For the purpose of proving this theorem, let us introduce some notation that we will use all along this thesis:

- We will write the interval of natural numbers $\{1, \ldots, n\}$ as $[n]$.
- We will call an arithmetic progression of length k a k-AP

0.1 Van der Waerden's Theorem

In this section we will develop the proof of Van der Waerden's Theorem. Let us begin by announcing it:

Theorem 0.1 (Van der Waerden's Theorem, 1927)**.** *Let r and k be positive integers. Then, there exists a number $W(r,k)$ such that if $N \geq W(r,k)$, then any r-coloring of $[N]$ contains a monochromatic k-AP.*

Let us start proving some intermediate results before proving Van der Waerden's Theorem, that for simplicity, we will call VdW Theorem:

The numbers $W(r,k)$ are called *Van der Waerden's numbers*. Proving the theorem consists on showing that for every choice of r and k, $W(r,k) < \infty$. Trivially, $W(r,1) = 1$. By the *Pigeonhole Principle*, once having $r+1$ elements, and r colors to paint with, two of them must have the same color, so $W(r,2) = r+1$.

Henceforth, we will assume a certain coloring of $[N]$ for N large enough. The following definition generalizes the notion of a monochromatic AP, which must be understood taking the object defined inside a bigger colored interval:

Definition 0.2 (Sunflower)**.** A sunflower with m petals of size $k-1$ is a set of integers of the form $\{a\} \cup A_1 \cup \ldots \cup A_m$, where $A_i = \{a + d_i, a + 2d_i, \ldots, a + (k-1)d_i\}$ for $1 \leq i \leq m$, satisfying the following properties:

(*i*) $A_i \cap A_j = \emptyset$ for $i \neq j$ (disjointness of the petals).

(*ii*) All elements in A_i are colored with the same color (monochromaticity of the petals).

(*iii*) If $i \neq j$ the color used to color elements of A_i is different from the one used to color the elements of A_j (different colors for different petals).

In such a situation, the sets $A_1, \ldots, A_m$ are called the *petals* of the sun flower and a the *center* of the sunflower.

Note that in a sunflower, we have a lot of control on the color and the structure of each petal (we have a $(k-1)$-AP), but we do not control the color of the center. In particular, if the color of the center is equal to the color of one of the petals, then the sunflower defines a monochromatic k-AP.

Let us prove a proposition that will be useful to prove VdW Theorem:

Proposition 0.3. *Let us consider an r-coloring of $[N]$ and let $A, A+d, \ldots, A+(k-1)d \subseteq [N]$ with the induced coloring. Assume that:*

(*i*) *$A + d, \ldots, A + (k-1)d$ are colored in the exactly same way.*

(*ii*) *$A + d$ is a sunflower with m petals of size $k-1$ (and in particular all its dilates $A + 2d, \ldots, A + (k-1)d$).*

Then, $A \cup (A + d) \cup \ldots \cup (A + (k-1)d)$ contains either a k-AP or a sunflower with $m+1$ petals of size $k-1$.

Proof. Let us pick the sunflower $A + d$, which has m petals of size $k-1$. Notice that if the center of our sunflower has the same color as one of the petals, we have a k-AP, so we are done.

Otherwise, let us construct a sunflower with $m+1$ petals of size $k-1$ contained in $A \cup (A+d) \cup \cdots \cup (A+(k-1)d)$ supposing that the color of the center of $A+d$ is different to any color of the petals in $A+d$. As we have used before, we will write $A+d = \{a+d\} \cup A_1 \cup \cdots \cup A_n$, where $A_i = \{a + d + d_i, a + d + 2d_i, \ldots, a + d + (k-1)d_i\}$ and A_i is a petal of $A + d$.

Let us consider the set $B = \{a\} \cup B_1 \cup B_2 \cup \cdots \cup B_m \cup B_{m+1}$, where

$$B_i = \{a + (d + d_i), a + 2(d + d_i), \ldots, a + (k-1)(d + d_i)\} \text{ for } 1 \leq i \leq m,$$

and

$$B_{m+1} = \{a + d, \ldots, a + (k-1)d\}.$$

Notice that B_{m+1} is created by taking the center of the sunflowers $A + d, \ldots, A + (k-1)d$, and by construction, $B \subseteq A \cup (A + d) \cup \cdots \cup (A + (k-1)d)$. Let us prove now that B is in fact a sunflower with $m + 1$ petals of size $k - 1$. First, let us remark some properties:

- The color of $a + r(d + d_i)$, for $1 \leq r \leq k - 1$ and $1 \leq i \leq m$ is the same as the color of $a + d + rd_i$, since dilates by d do not change color.
- The color of $a + d + rd_i$ is the same as the color of $a + d + d_i$, since the elements that differ in a multiple of d are in the same petal.

This proves that all elements in B_i for $1 \leq i \leq m$ have the same color.

- The color of B_i, which is the same used to color the element $a + d_i$, is different to the one used in B_j if $i \neq j$.

Now, notice that B_{m+1} is also a monochromatic petal: the color of $a+d, \ldots, a+(k-1)d$ is the same and it is different to the ones used to paint $B_1, \ldots, B_m$ because by hypothesis, the center of $A+d$, which is $a+d$, had a different color to the one used to paint the element $a+d+d_i$ for $1 \leq i \leq m$. So we have that $B = \{a\} \cup B_1 \cup B_2 \cup \cdots \cup B_m \cup B_{m+1}$ is a sunflower with $m + 1$ petals of size $k - 1$. □

Before proving VdW Theorem, we need a lemma that will give us the main idea of its proof: applying an induction argument on the length of the AP and combine it with the existence of a certain sunflower:

Lemma 0.4. *Let k be a positive integer. Assume that $W(r, k-1)$ exists for every choice of r. Then, for every choice of r and m there exists a positive integer $W(r, m, k-1)$ such that if $N \geq W(r, m, k-1)$, then any r-coloring of $[N]$ contains either a monochromatic k-AP or a sunflower with m petals of size $k - 1$.*

Proof. By applying induction on m: for $m = 1$, the statement holds trivially: a sunflower with one petal of size $k - 1$, by definition, has at least a monochromatic $(k - 1)$-AP (if the center a has the same color as the elements of the petal, then we have a k-AP). As by hypothesis we now that $W(r, k-1)$ exists, and in particular, $W(r, 1, k-1) = W(r, k-1) < \infty$.

Now, let us assume that $W(r, m - 1, k - 1)$ exists and prove that $W(r, m, k - 1)$ also exists. Let us write $N_1 = W(r, m - 1, k - 1)$, which exists by induction hypothesis and $N_2 = 2W(r^{N_1}, k - 1)$, that exists since we are assuming the existence for every choice of r.

In order to prove the existence of $W(r, m, k-1)$, it is enough to prove that $W(r, m, k-1) \leq N_1 N_2$: let us consider the interval $[N_1 N_2]$ and color it using r colors. We will look at it as the concatenation of $W(r^{N_1}, k-1)$ blocks of size N_1 each, which will define the *first part* of the partition, followed by extra $W(r^{N_1}, k-1)$ blocks of size N_1 each, that we will call the *second part* of the partition. In other words, both parts of size N_1 has $N_2/2$ blocks.

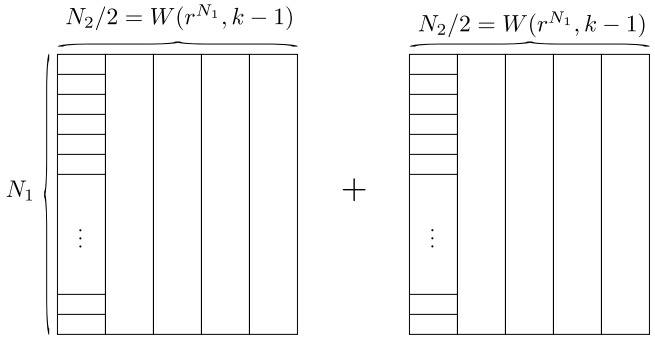


Figure 1: Diagram of the interval $[N_1 N_2]$

We will now focus on the second part of the partition. If we have a block containing a monochromatic k-AP, we are done. So let us assume the contrary. Note that since the size of a block is $N_1 = W(r, m-1, k-1)$ and we can use r colors to paint it, a block can be colored in r^{N_1} different ways. This means that in the second part of the partition we will find $k-1$ blocks identically colored, forming a $(k-1)$-AP, say $B+d, \ldots, B_{k-1}d$ (here we are using the hypothesis which says $W(r, k-1)$ exists applied on the blocks of $[N_1, N_2]$). Notice that $|B+d| = N_1 = W(n, m-1, k-1)$, so since we are supposing that $B+d$ does not contain a monochromatic k-AP, it must contain a sunflower $A+d$ with $m+1$ petals of size $k-1$.

We are now left with a final step: let us consider $B = (B+d) - d$ and $A = (A+d) - d$, and notice that $A \subseteq B$. Observe that we set $B+d, \ldots, B+(k-1)d$ to be in the *second part* of the partition, but B may belongs to the first part (this is the reason why we take 2 in the definition of N_2). We do not now anything about B's coloration, but by Proposition 0.3, $A \cup (A+d) \cup \cdots \cup (A+(k-1)d)$ must contain either a monochromatic k-AP or a sunflower with m petals of size $k-1$, as we wanted to prove. □

At this point we have all we needed to prove VdW. For this purpose, we will apply induction as follows:

Proof of Theorem 0.1 (Van der Waerden). We apply induction on k. We have shown that $W(r, 1)$ and $W(r, 2)$ exists for any r. Assume that $W(r, k-1)$ exists for every choice of r and let us prove that $W(r, k)$ exists for every choice of r.

Lemma 0.4 tells us that $W(r, m, k-1)$ exists for every r and m given. In particular, $W(r, r, k-1)$ exists for every choice of r. Let us prove that $W(r, k) = W(r, r, k-1)$. First,

$W(r, r, k-1)$ tells us that for every $N \geq W(r, r, k-1)$, any r-coloring of $[N]$ has either a monochromatic k-AP or a sunflower with r petals of size $k-1$. In the second case, since we have r petals and r colors to paint with, the color of the center must be equal to the color of one of the petals. Therefore, in any of the cases, we have a monochromatic k-AP as we wanted to show. □

After Van der Waerden's Theorem, in 1936, Erdős and Turán conjectured a big result on arithmetic combinatorics [8]: *every set $A \subset \mathbb{N}$ with positive natural density, this is*

$$\limsup_{n \longrightarrow \infty} \frac{|A \cap \{1, 2, \dots n\}|}{n} > 0,$$

contains a k arithmetic progression for every k. This infinite version has a finite analogue: *for every k and every $\varepsilon > 0$, there exists n_0 such that for $n \geq n_0$, all sets of non-negative integers A contained in $[n]$ with $|A| \geq \varepsilon n$ must contain a k-AP.* Note that this conjecture is a stronger version of Van der Waerden's Theorem, as it states that the most popular colour in any colouring is the right candidate to contain k-APs.

It was not until 1953 that Klaus Roth partially proved this conjecture for the 3-AP case using Fourier analytic methods [18]:

Theorem 0.5 (Roth, 1953)**.** *For every $\alpha > 0$ there exists n such that every subset $A \subseteq [n]$ of size at least αn contains a 3-AP.*

Due to this result, and joint with his investigations in analytic number theory, Roth received the Fields Medal in 1958.

The main focus of this work is to understand, rewrite and collect some of the proofs of Roth's Theorem that have appeared over the years. In this dissertation we will present three proofs while developing some of the problems that arise in each area. The structure of this thesis is as follows:

- In the first chapter, we begin with a **combinatorial proof** given by Szemerédi, which appeared 10 years after Roth's initial proof. In this proof we will be able to point the main ideas Szemerédi used, in a masterpiece of combinatorial reasoning, to settle the general conjecture affirmatively in 1975 [22].
- In the second chapter, we will approach the original proof of Roth's Theorem by a **Fourier analytic proof**, which will give us a much more specific result. We will also give a precise discussion of the main idea of the proof in the finite field model, where we can exploit the particularity of working with subspaces.
- In the third chapter, we will present the **graph theoretical proof**, and with it, the Szemerédi's Regularity Lemma. The use of this technique allows us to translate the result to non-abelian groups.

– In the last chapter, we will finish the work giving results about the upper and lower bound of the largest subset of $[n]$ which contains no 3-AP, which is still an open research problem.

0.2 Preliminaries and notation

Let us finish the introduction explaining the basic concepts and introducing some notation required in order to develop this thesis, and that from now on we will assume that the reader is familiar with:

– As we have already mentioned, we will write the interval of natural numbers $\{1, \ldots, n\}$ as $[n]$.

Definition 0.6 (Arithmetic progression)**.** An arithmetic progression is a sequence of numbers such that the difference between the consecutive terms is constant. In particular, we will call an arithmetic progression of length k a k-AP.

Since our work will be focused on arithmetic progressions of length 3, let us remark those:

Definition 0.7 (Arithmetic progression of length 3)**.** We will call an arithmetic progression of length 3 a 3-AP, which in particular, it is triple of the form

$$(a, a + d, a + 2d).$$

Thus, if we have a triple (a_1, a_2, a_3) satisfying this definition, it must also satisfy the following equality:

$$a_1 + a_2 = 2a_3.$$

In order to give some results about the cardinality of the sets containing 3-AP, we will use the following asymptotic notation. For any given functions f and g,

– We say that $f(n) = O(g(n))$, if there exists some constants c and n_0 such that

$$f(n) \leq c \cdot g(n) \qquad \text{for all} \quad n \geq n_0.$$

– We say that $f(n) = o(g(n))$ if for all $c > 0$ exists n_0 such that

$$g(n) < c \cdot f(n) \qquad \text{for all} \quad n \geq n_0.$$

Chapter 1

Combinatorial proof

In this chapter we present the combinatorial proof of Roth's Theorem, due to Szemerédi 10 years later Roth's Fourier proof. We will follow the theory developed by L. Graham et al. in the book Ramsey Theory, building on the main ideas that Szemerédi used for his proof of Erdős-Turán conjecture [22].

We begin proving and announcing some results in order to develop the proof, which is based on a combinatorial basic structure, named after David Hilbert: the *Hilbert Cube*. Hilbert was one of the pioneers of Ramsey theory; in 1892, he published a paper [15] regarding this structure, 30 years before its foundation.

Definition 1.1. Let $[N]$ be the set of non-negative integers from 1 to N. We call $M \subset [N]$ a k-cube if there exists $a > 0$ and $d_1, \ldots, d_k > 0$ such that

$$M = M(a : d_1, \ldots, d_k) = \left\{ a + \sum_{i=1}^{k} \varepsilon_i d_i \text{ such that } \varepsilon_i = 0, 1 \right\}.$$

The following lemma shows that we can assure the existence of Hilbert cubes as soon as our set has linear size:

Lemma 1.2. *Let n, α, k be such that the sequence $\alpha = \alpha_0, \alpha_1, \ldots, \alpha_k$ satisfying*

$$\alpha_{i+1} = \left\{ \binom{\alpha_i}{2} / (n-1) \right\}$$

has $\alpha_k \geq 1$. If $A \subseteq [n]$ with $|A| = \alpha$, there exists a k-cube $M \subset A$.

Notice that, in particular, if $|A| = cn$, for c fixed, there exists a k-cube $M \subset A$ with $k = \log \log n + O(1)$.

Proof. There exists $\binom{\alpha}{2}$ positive differences $a' - a$ with $a, a' \in A$. By the *Pigeonhole Principle*, since all the differences can only take $n-1$ different values, at least

$$\binom{\alpha}{2} / (n-1)$$

of the differences must be equal. Setting d_1 equal to the most frequently occurring difference, and $A_1 = \{a \in A : a + d_1 \in A\}$, we have

$$A_1 \subseteq A, \qquad d_1 + A_1 \subseteq A, \qquad |A_1| \geq \alpha_1.$$

Now, applying this argument to A_1, we get

$$A_2 \subseteq A_1, \qquad d_2 + A_2 \subseteq A_1, \qquad |A_2| \geq \frac{\binom{|A_1|}{2}}{n-1} \geq \alpha_2.$$

Now, by induction,

$$A_i \subset A_{i-1}, \qquad d_1 + A_i \subseteq A_{i-1}, \qquad |A_i| \geq \alpha_i.$$

Since $\alpha_k \geq 1$, there exists $a \in A_k$. Now, $M(a : d_i, \ldots, d_k) \subseteq A_{i-1}$ by a simple backward induction on i so that

$$M = M(a : d_1, \ldots, d_k) \subseteq A.$$

□

The analytic result is indicated by

$$\alpha_{i+1} = \binom{\alpha_i}{2}/n - 1 = \frac{\alpha_i(\alpha_i - 1)}{2(n-1)} \sim \frac{\alpha_i^2}{2n} \tag{1.1}$$

which implies $\log\log(n/\alpha_i) \sim i + O(1)$. Let us show this result using induction:

Claim.

$$\log\log(n/\alpha_i) \sim i + O(1).$$

Proof. Using relation (1.1),

$$\frac{n}{\alpha_{i+1}} \sim \frac{2n^2}{\alpha_i^2}.$$

Now, writing $\beta_i = n/\alpha_i$, we have

$$\beta_{i+1} \sim 2\beta_i^2.$$

Let $\log\log(\beta_{i-1}) \sim i - 1 + O(1)$ be true and let us prove $\log\log(\beta_i) \sim i + O(1)$ by induction:

$$\begin{aligned} \log\log(\beta_i) \sim \log\log\left(2\beta_{i-1}^2\right) &= \log(\log 2 + 2\log\beta_{i-1}) \\ &\sim \log(2\log\beta_{i-1}) \\ &= \log_2 2 + \log\log\beta_{i-1} \\ &\sim 1 + i - 1 + O(1) = i + O(1). \end{aligned}$$

So $\log\log(\beta_i) \sim i + O(1)$ as we wanted. □

At that point, let us introduce a function that will play a significant role in the proof: for every $l \in \mathbb{N}^+$, let $S(l)$ denote the largest number of elements of $[1, l]$ that can be chosen so that no 3-AP appears. Trivially, this function verifies the following property:

$$S(l_1 + l_2) \leq S(l_1) + S(l_2). \tag{1.2}$$

The function $S(l)$ is what we call a *subadditive* function. The next result, also known as *Fekete's lemma*, gives limiting answer for the limit of suabadditive sequences:

Lemma 1.3 (Subadditivity lemma)**.** *If $S : \mathbb{N} \longrightarrow \mathbb{R}^+$ is subadditive, then*

$$\alpha = \lim \frac{S(n)}{n} \text{ exists} \quad \text{and} \quad \frac{S(n)}{n} \geq \alpha \ \forall n \in \mathbb{N}$$

Proof. Let us first set $\alpha = \limsup_{n\to\infty} S(n)/n$. Let $n \in \mathbb{N}$. We can write any $x \in \mathbb{N}$ as $x = qn + r$ with $0 \leq r < n$. So in particular, S verifies

$$S(x) = S(qn + r) \leq S((q + 1)n) \overset{(1.2)}{\leq} (q + 1)S(n).$$

Thus,

$$\frac{S(x)}{x} \leq \frac{(q + 1)S(n)}{qn}.$$

Recall that if $x \longrightarrow \infty$, we also have $q \longrightarrow \infty$. Now, taking lim sup in both sides of the last inequality,

$$\alpha \leq S(n)/n.$$

In particular $\alpha \leq \liminf_{n\to\infty} S(n)/n$. So $\alpha = \lim_{n\to\infty} S(n)/n$ as we wanted to show. □

At that point we are ready to prove Roth's Theorem:

Theorem 1.4 (Szemerédi)**.** *If A is a set of positive integer with positive upper density, then A contains a* 3-AP*.*

Proof. Let $A \subset [l]$, $|A| \geq cl$ and suppose that A do not contain any 3-AP. By hypothesis, A is a set with positive upper density. Recall that $S(l)$ satisfies $S(l_1 + l_2) = S(l_1) + S(l_2)$, so there exists $c > 0$ such that $c = \lim S(l)/l$ and $S(l) > cl$ for all l. Let $\varepsilon > 0$ be very small and l_0 be such that

$$c \leq \frac{S(l)}{l} < c + \varepsilon \quad \forall l \geq l_0. \tag{1.3}$$

Let l be large so the inequality $0.01c^2 \log\log l > l_0$ holds, which will be clear at the end of the proof.

The main objective will be to show the existence of a large-dimensional cube $M \subseteq A$ whose elements will be located far from the edges of $[l]$. To do so, let us partition the interval $[1, l]$ into three sets as shown in Figure 1.1.

Since A has no 3-AP, $A \cap [1, 0.49l]$ and $A \cap [0.5l, l]$ must be 3-AP free. Thus, we can apply (1.3) to these intervals:

$$|A \cap [1, 0.49l]| < 0.49l(c + \varepsilon), \qquad |A \cap [0.5l, l]| < 0.5l(c + \varepsilon).$$

So on $[1, 0.49l] \cup [0.5l, l]$ A has at most $0.99l(c + \varepsilon)$ elements. Now, since $|A| \geq cl$ and ε is small, the density of A on $(0.49l, 0, 5l)$ is

$$\frac{|A \cap (0.49l, 0.5l)|}{0.01l} \geq \frac{cl - 0.99l(c + \varepsilon)}{0.01l} = \frac{0.01cl - 0.99l\varepsilon}{0.01l} = c - 99\varepsilon > \frac{c}{2}.$$

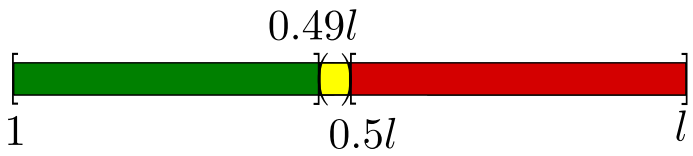


Figure 1.1: Partition of the interval $[1, l]$

At that point, as l is very large, we can split the interval $(0.49l, 0.5l)$ into disjoints sub-intervals of size $l^{1/2} + O(1)$. Notice that on one of these sub-intervals, the density of A must be at least $c/2$, otherwise A would not have density greater than $c/2$ on $(0.49l, 0, 5l)$. In that very same interval, by Lemma 1.2 it must exists a k-cube M so that satisfying:

(i) $M = M(a : d_1, \ldots, d_k) \subseteq (0, 49l, 0.5) \subseteq A$,

(ii) $k = \log\log l^{1/2} + O(1) = \log\log l + O(1)$,

(iii) $d_i \leq 2l^{1/2}$, $1 \leq i \leq k$.

Note that (iii) comes from the fact that M is contained in an interval of size $l^{1/2} + O(1)$, so if a is the smaller element of our k-cube M, then the element $a + d_i$ must also be contained, and in particular $d_i < 2l^{1/2}$.

Let $M_1 = \{a\}$ and $M_i = M(a : d_1, \ldots, d_{i-1})$ for $2 \leq i \leq k$. We define

$$N_i = \{2m - x \text{ such that } x \in A,\ x < A,\ m \in M_i\}$$

as the set of the third term of progressions $\{x, m, y\}$ with $x, m \in A$. By hypothesis, A is 3-AP free, so $A \cap N_i = \emptyset$. Let us now show that A has density greater $c/2$ on $[1, 0, 49l]$ by *reductio ad absurdum*:

Lemma 1.5. *The following estimate holds*

$$\frac{|A|}{|[1, 0.49l]|} > c/2. \tag{1.4}$$

Proof. Suppose that A has density lower or equal to $c/2$. Let us define the sets

$$S_1 = |A \cap [1, 0.49l]|, \qquad S_2 = |A \cup (0, 49l, 0.5l)|, \qquad S_3 = |A \cap [0.5l, l]|.$$

So we have

$$\begin{aligned} \frac{|S_1|}{|[1, 0.49l]|} < \frac{c}{2} \qquad &\Longrightarrow \qquad |S_1| < \frac{0.49cl}{2} = 0.245cl, \\ \frac{|S_2|}{|(0.49l, 0.5l)|} \overset{(1.3)}{<} c + \varepsilon \qquad &\Longrightarrow \qquad |S_2| < 0.01l(c + \varepsilon), \\ \frac{|S_3|}{|[0.5l, l]|} \overset{(1.3)}{<} c + \varepsilon \qquad &\Longrightarrow \qquad |S_3| < 0.5l(c + \varepsilon). \end{aligned}$$

Now, we have that $|A| = |S_1| + |S_2| + |S_3| < l(0.755c + 0.51\varepsilon) < cl$, which contradicts the fact $|A| \geq cl$. So the density of A on $[1, 0.49l]$ must be greater than $c/2$. □

This last result and the fact that $a \in (0.49l, 0.5l)$ give us a bound for the N_i sets we have defined above:

$$|N_i| \geq |N_1| = |A \cap [1, a)| > 0.245cl.$$

Notice we can write $M_{i+1} = M_i \cup (M_i + d_i)$, $N_{i+1} = N_i \cup (N_i + 2d_i)$, where the N_i forms an ascending sequence with $|N_k| < l$. So in particular, by the *Pigeonhole Principle* there must exists i so that

$$|N_{i+1} - N_i| < \frac{l}{k}.$$

From now on, let us focus our attention on this pair, whose cardinality is similar. Let us call an AP with difference $2d_i$ a *block*. Given a maximal block $\{x, x + 2d_i, \ldots, x + s(2d_i)\}$ of N_i, by definition of N_{i+1}, we can find the following element of the progression $x + (s+1)(2d_i)$ in $N_{i+1} - N_i$. Thus, the number of blocks is at most the cardinal of the set $N_{i+1} - N_i$, which is smaller than l/k.

Let us now split $[l]$ into residue classes modulo $2d_i$. Note that the elements of a block will belong to the same class after the partition. In particular, if t_j is the number of blocks into which N_i is partitioned in the class of congruence j, then $[l] - N_i$ will be partitioned into at most $t_j + 1$ blocks (the gaps plus the ends). So, using that l/k is the maximum number of blocks of N_i, we have

$$\sum_{j=0}^{2d_i - 1} t_j < \frac{l}{k},$$

and using $d_i < 2l^{1/2}$, on $[l] - N_i$

$$\sum_{j=0}^{2d_i - 1} t_j + 1 = \sum_{j=0}^{2d_i - 1} t_j + 2d_i < \frac{l}{k} + 2d_i = \frac{l}{\log\log l}(1 + O(1)).$$

Now we may begin the final argument. We call a block of $[l] - N_i$ *small* if its cardinal is less than $0.01c^2 \log\log l$, and *large* otherwise. Recall that since the number of blocks have to be less than $\frac{l}{\log\log l}(1 + O(1))$, all the small blocks together must have cardinality smaller than

$$\frac{l}{\log\log l}(1 + O(1)) \cdot 0.01c^2 \log\log l = 0.01c^2 l + O(l).$$

We have defined l so that A has density lower than $c + \varepsilon$ on every large block, and hence on its union. Now, recall that $A \cap N_i = \emptyset$, so each element of A must be either in a large block or in a small one:

$$\begin{aligned} |A| &= |A \cup ([l] - N_i)| \\ &< (c + \varepsilon)(l - |N_i| + 0.01c^2 l + O(l) \\ &\overset{(1.4)}{<} cl - c(0.245cl) + \varepsilon l + 0.01c^2 l + O(l) \\ &< cl, \end{aligned}$$

contradicting the assumption $|A| \geq cl$. □

Notice that this result is not quantitative. In the next chapter, we will develop techniques that allow us to get a much more specific result.

1.1 Discussion

As we have already said, in this proof we can see the main ideas that Szemerédi used in order to prove the general case of Roth's Theorem. For this purpose, in 1972 he proved the case for $k = 4$ [21], and 3 years later, in 1975, Szemerédi presented the proof of the Erdös-Turán conjecture,

Theorem 1.6 (Szemerédi, 1975)**.** *A subset of the natural numbers with positive upper density contains infinitely many arithmetic progressions of length k for all positive integers k.*

The original proof by Szemerédi, while extremely difficult and with a complex structure, was purely combinatorial and almost entirely self-contained, except for an invocation of Van der Waerden's Theorem.

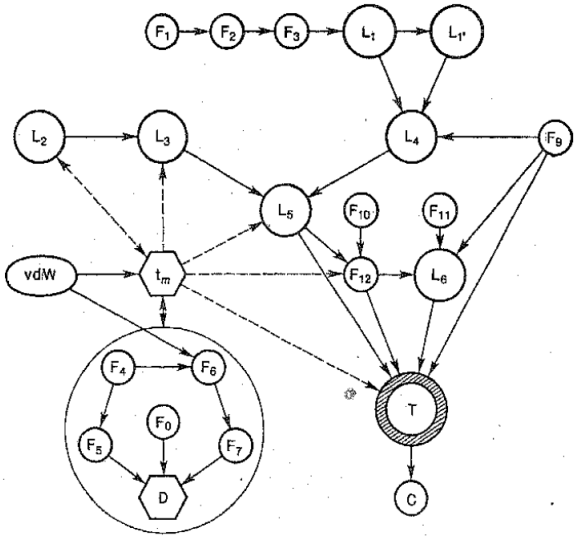


Figure 1.2: Diagram representing an approximate flow chart of Szemerédi's Theorem. Available at [11]

The last picture depicts the structure of the paper by Szemerédi proving fully Erdős-Turán conjecture. This has been a masterpiece in combinatorics and one of the main reasons to propose Szemerédi for the Abel prize in 2012.

In order to prove Szemerédi's Theorem, it was required the development of what nowadays is known as a powerful tool in extremal graphs methods: the *Szemerédi Regularity Lemma*. In Szemerédi's Theorem proof, he introduced a weaker version of this lemma, restricted to bipartite graphs. It was not until years later that a extension proof for hypergraphs appeared. This can be found in the paper *The hypergraph regularity method and its applications* [17]. In chapter 3 we will develop Szemerédi's Regularity Lemma for hypergraphs in order to obtain a very simple proof of Roth's Theorem.

Chapter 2

Fourier analytic proof

In this chapter we will approach Roth's Theorem analytic proof, the first of the multiple proofs that we can find in relation to this theorem. Roth's Theorem was proved in 1953 [18], when Roth partially resolved Erdős-Turán conjecture [8] for $k = 3$. As we already mention, this result, in addition to other important work, earned him the Fields Medal in 1958.

2.1 Introduction to discrete Fourier analysis

The original proof of Roth's Theorem used Fourier analytic methods. In order to be able to develop the proof, let us begin this chapter by presenting some basic results of the discrete Fourier analysis.

From now on let us consider n to be a positive integer and let $\mathbb{Z}_n$ be the field of integers mod n. Let $\omega = e^{2\pi i/n}$ be a primitive n-th root, so that $\omega^n = 1$.

Definition 2.1 (Discrete Fourier transform)**.** Given a function $f : \mathbb{Z}_n \longrightarrow \mathbb{C}$, we define its discrete Fourier transform by

$$\hat{f}(r) = \frac{1}{n} \sum_{k=0}^{n-1} f(k)\omega^{-rk}.$$

We will use the following ergodic-theory notation for the Fourier transform:

$$\hat{f}(r) = \mathbb{E}_x f(x)\omega^{-rx},$$

which is interpreted as a *mean*, as we have announced.

There is one important fact about the functions ω^{rx} that appear in the definition of the discrete fourier transform:

Proposition 2.2. *The functions ω^{rx} form an orthonormal basis with respect to his inner product.*

Proof. Notice that the inner product of the functions ω^{rx} and ω^{sr} is $\mathbb{E}_x\omega^{(r-s)x}$. Now, if $r = s$, then $\mathbb{E}_x\omega^{(r-s)x} = 1$. Otherwise, if $r \neq s$, we have a geometric progression. Using its formula

and the fact that $\omega^n = 1$, we have

$$\mathbb{E}_x \omega^{(r-s)x} = \frac{1 - \omega^{(r-s)n}}{n(1 - \omega^{r-s})} = 0.$$

□

There exists three notions about Fourier transform that we will use repeatly. The first one is *the inversion formula:*

Lemma 2.3 (The inversion formula)**.** *Let $f : \mathbb{Z}_n \longrightarrow \mathbb{C}$ and $\hat{f}$ its Fourier transform. Then for every $x \in \mathbb{Z}_n$ we have*

$$f(x) = \sum_r \hat{f}(r)\omega^{rx}.$$

Proof. By expanding out $\hat{f(r)}$, we can observe:

$$\begin{aligned}\sum_r \hat{f}(r)\omega^{rx} &= \sum_r \mathbb{E}_y f(y)\omega^{-ry}\omega^{rx} \\ &= \mathbb{E}_y f(y) \sum_r \omega^{r(x-y)} = \begin{cases} 1 & \text{if } y = x, \\ 0 & \text{if } y \neq x. \end{cases}\end{aligned}$$

Note that in the first case we have used that the probability of $x = y$ if y is randomly chosen is $1/n$. □

Notice that in the following theorem it will appear for the first time the inner product between two functions. All along this chapter we will make an abuse of notation, since we will be using two of the following inner products at the same time, where (ii) is (i) normalized: for $f, g : X \longrightarrow \mathbb{C}$, we can define:

(i) $\mathbb{E}_x f(x)\overline{g(x)}$

(ii) $\sum_x f(x)\overline{g(x)}$

The context will make clear which one is intended. Let us now pass to prove the next result, which is attributed to Plancherel and sometimes to Parseval, but we will call it *Parseval identity*:

Theorem 2.4 (Parseval identity)**.** *Let $f, g : \mathbb{Z}_n \longrightarrow \mathbb{C}$ and let $\hat{f}, \hat{g}$ be their respective Fourier transforms. Then,*

$$\langle \hat{f}, \hat{g} \rangle = \langle f, g \rangle. \tag{2.1}$$

Proof. This result follows immediately from two facts: the functions ω^{rs} form an orthonormal basis and the Fourier coefficients of $\hat{f}(r)$ are the coefficients of f in the ω^{rs} basis. So now:

$$\begin{aligned}\langle \hat{f}, \hat{g} \rangle = \sum_r \hat{f}(r)\overline{\hat{g}(r)} &= \sum_r \mathbb{E}_{x,y} f(x)\overline{f(y)}\omega^{-r(x-y)} \\ &= \mathbb{E}_{x,y} f(x)\overline{f(y)} \sum_r \omega^{-r(x-y)} \\ &= \mathbb{E}_x f(x)\overline{g(y)} = \langle f, g \rangle.\end{aligned}$$

where in the last inequality we have used that if $x = y$ then the sum over r is m, while if $x \neq y$ it is 0, and the probability of $x = y$ if y is randomly chosen is $1/n$. □

Let us now present the next result, which is the *convolution identity*. This result sets the Fourier transform apart from any other unitary map. First let us define the *convolution* of two functions:

Definition 2.5 (Convolution)**.** Given two functions $f, g : \mathbb{Z}_n \longrightarrow \mathbb{C}$, we define their convolution $f * g$ by

$$f * g(r) = \mathbb{E}_{y+z=r} f(y)g(z).$$

If we want to use the counting measure of $\mathbb{Z}_n$, then the convolution is defined using sums:

$$\hat{f} * \hat{g}(r) = \sum_{s+t=r} \hat{f}(s)\hat{g}(t).$$

The following lemma gives us a strong property of the Fourier transform:

Lemma 2.6. *let $f, g : \mathbb{Z}_n \longrightarrow \mathbb{C}$ Then, for every $r \in \mathbb{Z}_n$ we have:*

$$\widehat{f * g}(r) = \hat{f}(r)\hat{g}(r).$$

Proof. By expanding out the definition of convolution:

$$\begin{aligned}\widehat{f * g}(r) &= \mathbb{E}_x f * g(x)\omega^{-rx} \\ &= \mathbb{E}_x(\mathbb{E}_{y+z=x} f(y)g(z)\omega^{-rx}) \\ &= \mathbb{E}_x \mathbb{E}_{y+z=x}(f(y)\omega^{-ry})(g(z)\omega^{-rz}) \\ &= (\mathbb{E}_y f(y)\omega^{-ry})(\mathbb{E}_z g(z)\omega^{-rz}) = \hat{f}(r)\hat{g}(r).\end{aligned}$$

as we wanted to show. □

After proving these basics of Fourier transform, we can introduce some interesting identities that can be useful. First, let us begin with a definition:

Definition 2.7. Given a function $f : \mathbb{Z}_n \longrightarrow \mathbb{C}$, we define

$$||f||_4 = \left(\sum_r |f(r)|^4 \right)^{\frac{1}{4}}.$$

Proposition 2.8. *Let $f : \mathbb{Z}_n \longrightarrow \mathbb{C}$. Then,*

$$||\hat{f}||_4^4 = \mathbb{E}_{x,a,b} f(x)\overline{f(x+a)f(x+b)}f(x+a+b),$$

where $\mathbb{E}_{x,a,b}$ can be written as $\frac{1}{n^3}\sum_{r,a,b}$.

Proof. First notice that there exists a correspondence between quadruples of the form $(x, x+a, x+b, x+a+b)$ and the quadruples (x, y, z, w), such that $x + w = y + z$. Hence, the right hand side of the equation can be rewritten as

$$\mathbb{E}_{x+w=y+z} f(x)f(w)\overline{f(y)f(z)} = \mathbb{E}_u \mathbb{E}_{x+w=y+z=u} f(x)f(w)\overline{f(y)f(z)} = \langle f * f, f * f\rangle.$$

Using now Parseval and the convolution identity, we can write

$$\langle f * f, f * f\rangle = \langle \hat{f}^2, \hat{f}^2\rangle = \sum_r |\hat{f}(r)|^4 = ||\hat{f}||_4^4,$$

as we wanted to prove. □

With the definition of the characteristic function of a set we can find a couple more Fourier-analytic properties that would be interesting to use. As we will see, applying the Fourier transform to constant functions is a nice way to codify the configuration of a set:

Definition 2.9 (Characteristic function of a set)**.** Let $A \subset \mathbb{Z}_n$ be a set. We define its characteristic function by

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

Lemma 2.10. *Let $A \subset \mathbb{Z}_n$ be a set of density α. Then*

1. $\hat{A}(0) = \alpha$.
2. $\sum_r |\hat{A}(r)|^2 = \alpha$.
3. $\hat{A}(-r) = \overline{\hat{A}(r)}$ *for every* r.
4. $||\hat{A}||_4^4 = \sum_r |\hat{A}(r)|^4 \leq \alpha^4$.

Proof. For the first property, we can write:

$$\hat{A}(r) = \mathbb{E}_x A(x)\omega^{-rx} = \frac{1}{n}\sum_{k=0}^{n-1} A(k)\omega^{-rk} \iff \hat{A}(0) = \frac{1}{n}\sum_{k=0}^{n-1} A(k) = \frac{|A|}{n} = \alpha.$$

Let us now prove the second property announced:

$$\sum_r |\hat{A}(r)|^2 = \sum_r \langle \hat{A}(r), \hat{A}(r)\rangle = \frac{1}{n}\sum_{x=0}^{n-1}\langle A(x), A(x)\rangle = \frac{1}{n}\sum_{x=0}^{n-1}|A(x)| = \frac{|A|}{n} = \alpha.$$

For the third one:

$$\hat{A}(-r) = \mathbb{E}_x A(x)\omega^{-rx} = \mathbb{E}_x \overline{A(x)\omega^{rx}} = \overline{\hat{A}(r)}.$$

Let us prove the last one using Proposition 2.8:

$$\begin{aligned}\alpha^4 &= (\mathbb{E}_x A(x))^4 \\ &= ((\mathbb{E}_x A(x))^2 (\mathbb{E}_y A(y))^2 \\ &= (\mathbb{E}_x \mathbb{E}_y A(x)A(y))^2 \overset{y=x+h}{=} (\mathbb{E}_h \mathbb{E}_x (A(x)A(x+h))^2.\end{aligned}$$

Using now Cauchy-Schwarz (C-S) inequality

$$\begin{aligned}(\mathbb{E}_h \mathbb{E}_x (A(x)A(x+h))^2 &\leq \mathbb{E}_h (\mathbb{E}_x A(x)A(x+h)^2)^2 \\ &= \mathbb{E}_h (\mathbb{E}_x (A(x)A(x+h))(\mathbb{E}z(A(z)(A(z+h)) \\ &= \mathbb{E}_{x,h,z} A(x)A(x+h)A(z+h).\end{aligned}$$

Since $z = x + k$, we can write

$$\mathbb{E}_{x,h,z} A(x)A(x+h)A(z+h) \geq \mathbb{E}_{x,h,k} A(x)\overline{A(x+h)A(x+k)}A(x+h+k) = ||\hat{A}||_4^4,$$

as we wanted. □

Let us call quadruples of the form $(x, x+a, x+b, x+a+b)$ *additive quadruples*. As we have already stated, there exists a bijection between these quadruples and (x, y, z, w) such that $x + y = z + w$. The densities of these quadruples and of the 3-APs, which are the probabilities that a random additive quadruple or a random 3-AP live entirely in A, are given by the following definition:

Definition 2.11 (Additive quadruple and 3-AP densities)**.** We define the additive quadruple density of a set A as

$$\mathbb{E}_{x,a,b} A(x)A(x+a)A(x+b)A(x+a+b),$$

and the density of a 3-AP by

$$P_3(A) = \mathbb{E}_{x,d} A(x)A(x+d)A(x+2d).$$

The following lemma presents one of the first important results about 3-APs on a set. In a general sense, it says that if the spectrum of A is concentrated, then we can have control of the number of 3-APs.

Lemma 2.12. *Let n be odd and let A be a subset of $\mathbb{Z}_n$ of density α. Suppose that $\max_{r \neq 0} |\hat{A}(r)| \leq c\alpha^2$. Then $|P_3(A) - \alpha^3| \leq c\alpha^3$.*

Proof. In order to give the 3-AP density using Fourier, let us define the set

$$A_2 = \{u \in A \text{ such that } u/2 \in A\}.$$

So we can write

$$\hat{A}_2(r) = \mathbb{E}_x A_2(x)\omega^{-rx} = \mathbb{E}_x A(x/2)\omega^{-rx} = \mathbb{E}_x A(x)\omega^{-2rx} = \hat{A}(2r).$$

Hence,

$$\begin{aligned} P_3(A) &= \mathbb{E}_{x+y=2z} A(x)A(y)A(z) = \mathbb{E}_{x+y=2z} A(x)A(y)A(z/2) \\ &= \langle A * A, A_2\rangle \overset{(2.1)}{=} \langle \hat{A}^2, \hat{A}_2\rangle = \sum_r \hat{A}(r)^2\hat{A}(-2r). \end{aligned}$$

Since $\hat{A}(0) = \alpha$, we have

$$|P_3(A) - \alpha^3| = |\sum_{r\neq 0} \hat{A}(r)^2\hat{A}(-2r)| \leq \max_{r\neq 0}|\hat{A}(r)| \sum_{r\neq 0} |\hat{A}(r)||\hat{A}(-2r)|.$$

Now, by Cauchy-Schwartz, we know that the last sum can be at most

$$\left(\sum_{r\neq 0} |\hat{A}(r)|^2\right)^{1/2} \left(\sum_{r\neq 0} |\hat{A}(-2r)|^2\right)^{1/2}.$$

Notice that the second of these factors is equal to the first, so we can write:

$$\begin{aligned} \max_{r\neq 0}|\hat{A}(r)| \sum_{r\neq 0} |\hat{A}(r)||\hat{A}(-2r)| &\leq \max_{r\neq 0}|\hat{A}(r)| \left(\sum_{r\neq 0} |\hat{A}(r)|^2\right)^{1/2} \left(\sum_{r\neq 0} |\hat{A}(-2r)|^2\right)^{1/2} \\ &= \max_{r\neq 0}|\hat{A}(r)| \sum_{r\neq 0} |\hat{A}(r)^2| \\ &\overset{(2.10)}{=} \max_{r\neq 0}|\hat{A}(r)| \sum_{r\neq 0} (\alpha - \alpha^2) \\ &\leq \alpha \cdot \max_{r\neq 0}|\hat{A}(r)|. \end{aligned}$$

Therefore, taking $c = \max_{r\neq 0}|\hat{A}(r)|$, we have $|P_3(A) - \alpha^3| \leq c\alpha$. □

2.2 Finite field analogue proof

Before proving Roth's Theorem in $\mathbb{Z}_n$, it is interesting to develop the proof in $\mathbb{F}_3^n$ in order to get the main idea. First, notice that all the theory we have been working on $\mathbb{Z}_n$ can be similarly developed in $\mathbb{F}_3^n$. The corresponding question in $\mathbb{F}_3^n$, also known as the *cap set problem*, was studied by Meshulam in 1995 [16], using an argument that remarks the advantages of the finite model and the potential of the Fourier transform. Note that in a finite field

we can work with subspaces, where the idea of *increasing density* (which we will see soon) seems more natural and it gives us an intuition of the proof in $\mathbb{Z}_n$. This problem is very relevant in a lot of areas, such as code theory. Our approach is based on the exposition in [29].

Let us start this chapter announcing the analogue result of Roth's Theorem in the finite field, which states that any sufficient dense set must contain a non-trivial 3-AP:

Theorem 2.13 (Meshulam)**.** *Let $A \subset G = \mathbb{F}_3^n$. Suppose that A does not contain any non-trivial 3-AP. Then*

$$|A| = \frac{|G|}{\log |G|}(1 + o(1)).$$

With our first lemma below we will be ready to introduce the dichotomy, a concept that will be present in all the following Roth's Theorem proofs. It states that the number of 3-AP in a set with small non-trivial (different from the 0 coefficient) Fourier coefficients is equal to the number of 3-AP that can be found in a set whose elements are chosen from G with probability $\alpha = |A|/|G|$:

Lemma 2.14. *Let $A \subset G = \mathbb{F}_3^n$ be a subset of density α satisfying $\sup_{t\neq 0} |\widehat{1}_A(t)| = o(1)$. Then, A contains $(\alpha^3 + o(1))|G|^2$ 3-term arithmetic progressions.*

Proof. Let $2 \cdot A = \{2a : a \in A\}$. We will write the normalised number of 3-AP in A as $T_3(1_A, 1_A, 1_A)$, which is defined by

$$T_3(1_A, 1_A, 1_A) := \mathbb{E}_{x,d} 1_A(x)1_A(x + d)1_A(x + 2d) = \langle 1_A * 1_A, 1_{2\cdot A}\rangle,$$

where the last equality comes from the use of the definition of convolution and a change of variable, $x + d = y$. But using Parseval's identity, we can also write

$$T_3(1_A, 1_A, 1_A) \overset{P.}{=} \langle \widehat{1}_A^2, \widehat{1}_{2\cdot A}\rangle = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)\widehat{1}_A(-2t) = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3,$$

where we have use that $-2 \equiv 1 \mod 3$. Now, bounding the last sum:

$$\begin{aligned}\sum_{t=0} \widehat{1}_A(t)^3 &\leq \sup_{t\neq 0} |\widehat{1}_A(t)| \sum_t |\widehat{1}_A(t)|^2 \\ &= \sup_{t\neq 0} |\widehat{1}_A(t)| \cdot \mathbb{E}_x |\widehat{1}_A(x)|^2 \\ &= \alpha \cdot \sup_{t\neq 0} |\widehat{1}_A(t)|.\end{aligned}$$

Hence,

$$T_3(1_A, 1_A, 1_A) = \alpha^3 + o(1).$$

Since it was normalized, by multiplying by $|G|^2$ we get the desired result. □

In this case, the dichotomy is as follows: either the non-trivial Fourier coefficients of A are small (in this case we will say that A is uniform), and thus by Lemma 2.14 A contains lots of 3-APs; or A it is not uniform. In that last case we will be able to see that A contains non-trivial progressions.

Let us now continue proving a lemma that states that the characteristic function of a 3-AP free set must possess a large Fourier coefficient:

Lemma 2.15. *Let $A \subset G$ be a subset of density α, with $|G| \geq 2\alpha^{-1}$. Since A does not contain any non-trivial 3-AP, there exits $t \neq 0$ such that*

$$|\widehat{1}_A(t)| \geq \frac{1}{2}\alpha^2.$$

Proof. As we did in Lemma 2.14, we use the definition of the normalised number of 3-APs in a set, that can be written as $T_3(1_A, 1_A, 1_A) = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3$. By hypothesis, we now that A does not contain any non-trivial 3-AP. Therefore, we have

$$T_3(1_A, 1_A, 1_A) = |A| = \frac{|A| \cdot |G|}{|G|} = \frac{\alpha}{|G|}.$$

Recall that we can also write the number of 3-APs as

$$T_3(1_A, 1_A, 1_A) = \alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3.$$

So we have that

$$\alpha^3 + \sum_{t\neq 0} \widehat{1}_A(t)^3 = \frac{\alpha}{|G|} \iff \sum_{t\neq 0} \widehat{1}_A(t)^3 = \frac{\alpha}{|G|} - \alpha^3.$$

Taking now absolute values in both sides,

$$\begin{aligned}\left|\sum_{t\neq 0} \widehat{1}_A(t)^3\right| &= \left|\frac{\alpha}{|G|} - \alpha^3\right| \leq \left|\frac{1}{2}\alpha^3 - \alpha^3\right| \\ &= \left|\frac{1}{2}\alpha^3\right| \leq \left|\sum_{t\neq 0} \widehat{1}_A(t)^3\right| \leq \sup_{t\neq 0} |\widehat{1}_A(t)| \sum_t |\widehat{1}_A(t)|^2 \leq \alpha \cdot \sup_{t\neq 0} |\widehat{1}_A(t)|\end{aligned}$$

In particular, we have

$$\left|\frac{1}{2}\alpha^3\right| \leq \alpha \cdot \sup_{t\neq 0} |\widehat{1}_A(t)| \iff \text{ there exists } t \neq 0 \text{ such that } |\widehat{1}_A(t)| \geq \frac{1}{2}\alpha^2,$$

as we wanted. □

Let us show next that if there exists a large non-trivial Fourier coefficient, then A has a bias towards an affine subspace of $\mathbb{F}_3^n$ of co-dimension 1 (dimension $n-1$). With this lemma, we present another element that will play an elementary role: the *density increment*.

Lemma 2.16. *Let $A \subset G$ be a subset of density α. Suppose $t \neq 0$ is such that $|\widehat{1}_A(t)| \geq \alpha^2/2$. Then, there exists a subspace $V \leq \mathbb{F}_3^n$ of co-dimension 1 on some translate of which A has density at least $\alpha(1+\alpha/4)$.*

Proof. Let us write $V := \langle t \rangle^\perp$ the ortogonal complement of vector t, and let $v_j + V$ for $1 \leq j \leq 3$ be the complet set of cosets of V. Let us call $f_A := 1_A - \alpha$ the *balanced function of* A. Then,

$$\begin{aligned}\widehat{1}_A(t) = \widehat{f}_A(t) &= \frac{1}{|G|}\sum_{j=1}^{3}\sum_{x\in v_j+V}(1_A(x)-\alpha)\omega^{tx} \\ &= \sum_{j=1}^{3}\frac{1}{|G|}\sum_{x\in v_j+V}(1_A(x)-\alpha)\omega^{j} = \sum_{j=1}^{3}a_j\omega^j,\end{aligned}$$

where $a_j = (1/|G|)\sum_{x\in v_j+V}(1_A(x)-\alpha) = (|A\cap(v_j+V)| - \alpha\,|V|)/\,|G|$. First, notice that since all elements of A are in $v_j + V$ for some j, then $\sum_j a_j = 0$. Taking absolute values and applying the triangle inequality we find

$$\sum_j |a_j| \geq \alpha^2/2,$$

so $\sum_j |a_j| + a_j \geq \alpha^2/2$. Therefore, by the *Pigeonhole principle* there exists at least one value of j for which $|a_j| + a_j \geq \alpha^2/6$. So $a_j \geq \alpha^2/12$ and, as $|G| = 3\,|V|$, we can write

$$(|A\cap(v_j+V)| - \alpha\,|V|)/\,|G| = a_j \geq \alpha^2/12,$$

$$|A\cap(v_j+V)| \geq \alpha\,|V| + \frac{\alpha^2\cdot|G|}{12} = \alpha\,|V| + \frac{\alpha^2\cdot 3\,|V|}{12} = \alpha(1+\frac{\alpha}{4})\,|V|\,.$$

Finally,

$$|A\cap(v_j+V)| \geq \alpha(1+\frac{\alpha}{4})\,|V|\,,$$

as we wanted to prove. □

Now we are ready to complete Meshulam's Theorem by iterating Lemmas 2.15 and 2.16:

Proof of Theorem 2.13 (Meshulam). Let us suppose that A has density α in G and contains no 3-APs. Then, by Lemma 2.15, there exists a non-trivial coefficient of A with size at least $\alpha^2/2$. Therefore, by Lemma 2.16, there must exists a subspace V of dimension n-1 and a vector v_j such that A has density at least $\alpha + \alpha^2/4$ on V.

Let us write $A_1 = (A\cap(v_j+V))$. Since 3-AP are translation invariant, then if A has no 3-AP neither do A_1. Let us now focus on the set $A_1 \subset V \cong \mathbb{F}_3^{n-1}$ of density $\alpha + \alpha^2/4$ and iterate the explained procedure taking $G = \mathbb{F}_3^{n-1}$.

First, notice that since the density can never exceed 1, our procedure will come to an end if the density of the set A_k surpass this value. Given that in every step we increase the density in a $\alpha^2/4$ factor, we have that it grows up from α to 2α in $\alpha/(\alpha^2/4) = 4\alpha^{-1}$ steps, from 2α to 4α in $2\alpha/((2\alpha)^2/4) = (4\alpha^{-1})/2$ steps, so inductively the density will reach 1 in at most

$$(1 + 1/2 + 1/4 + \ldots)4\alpha^{-1} = 8\alpha^{-1}$$

iterates. Thus, the size of G after $8\alpha^{-1}$ iterates is at least $3^{n-8\alpha^{-1}}$. By Lemma 2.15, G must verify $|G| < 2\alpha^{-2}$ when we finish the procedure, otherwise we could apply again this very same lemma. So $3^{n-8\alpha^{-1}} \leq 2\alpha^{-2}$, thus:

$$3^n \leq 2\alpha^{-2} \cdot 3^{8\alpha^{-1}} \iff n \leq 8\alpha^{-1} + \log_3(2\alpha^{-2}).$$

Since n grows with α, then $1/\alpha$ gets smaller, so we can omit the second term and write:

$$n \leq \frac{8}{\alpha} \leq \frac{1}{\alpha} \iff \alpha \leq \frac{1}{n} \iff |A| \leq \frac{|G|}{\log_3(|G|)},$$

where we have used $|G| = 3^n$ and $\alpha = |A| / |G|$. □

2.3 Analytic proof of Roth's Theorem

After working on the analogue proof for finite fields, we will understand much more easily the idea of Roth's Theorem 0.5 using the Fourier transform. If we wanted to prove the same result in $\mathbb{Z}$, we will reduce it to $\mathbb{Z}_n$, but there exists a problem: in $\mathbb{Z}_n$ we do not have an algebraic structure such a vectorial subspace, so the density increment we have seen before has to be done differently. Nevertheless, the same idea persists: if we have a 3-AP free set, then there exists a large non-trivial Fourier coefficient, which implies a density increment on a sub-AP.

The structure of the proof is as follows: let us think of A as a subset of $\mathbb{Z}_n$. Supposing n odd, we can apply Lemma 2.12 to state that either A contains a 3-AP or there exits $r \neq 0$ such that $|\hat{A}(r)| \geq \delta^2/2$. In the second case we can relate A with one of the functions ω_r with $r \neq 0$. This will allow us to deduce that the intersection of A with some $\sqrt{n}$-AP has density at least $\delta + \frac{1}{4}\delta^2$.

Let us begin with a generalisation of Lemma 2.12 to manage the fact that a 3-AP inside $\mathbb{Z}_n$ does not have to correspond to a 3-AP inside $[n]$:

Lemma 2.17. *Let n be odd and let A, B and C be subsets of $\mathbb{Z}_n$ with densities α, β and γ, respectively. Let $\max_{r\neq 0} |\hat{A}(r)| \leq \theta$. Then,*

$$|\mathbb{E}_{x,d} A(x)B(x+d)C(x+2d) - \alpha\beta\gamma| \leq \theta(\beta\gamma)^{1/2}.$$

Proof. This proof is analogous to the one we have already done in Lemma 2.12. Changing the functions to A, B and C, we get

$$\mathbb{E}_{x,d}A(x)B(x+d)C(c+2d) = \sum_r \hat{A}(r)\hat{B}(-2r)\hat{C}(r).$$

Now,

$$\begin{aligned}|\sum_{r\neq 0}\hat{A}(r)\hat{B}(-2r)\hat{C}(r)| &\leq \max_{r\neq 0}|\hat{A}(r)|\sum_{r\neq 0}\hat{B}(-2r)\hat{C}(r)\\ &\overset{C\text{-}S}{\leq} \max_{r\neq 0}|\hat{A}(r)|\sum_{r\neq 0}(\hat{B}(-2r)^2)^{1/2}(\sum_{r\neq 0}\hat{C}(r)^2)^{1/2}\\ &= \max_{r\neq 0}|\hat{A}(r)|(\beta\gamma)^{1/2} \leq \theta(\beta\gamma)^{1/2},\end{aligned}$$

as we wanted to show. □

Let us now approach a corollary that will be used in order to prove Roth's Theorem. It states that either there A contains a 3-AP or there exists a large non-trivial Fourier coefficient of $\hat{A}(r)$: the dichotomy argument.

Corollary 2.18. *Let n be odd and let A be a subset of $[n]$ of density δ. Suppose that $A\cap[n/3,2n/3]$ has cardinality at least $\delta n/5$. Then either A contains a 3-AP or when A is considered as a subset of $\mathbb{Z}_n$, there exists $r\neq 0$ such that $|\hat{A}(r)|\geq\delta^2/10$.*

Proof. Let $B=C=A\cap[n/3,2n/3]$ and consider A, B and C as subsets of $\mathbb{Z}_n$. Notice that any 3-AP of the form $(x,x+d,x+2d)\in A\times B\times C$ has $d\in(-n/3,n/3)$ (recall that B and C are the intersection of A with the interval $[n/3,2n/3]$), from which we can conclude that $(x,x+d,x+2d)$ is a 3-AP even if it is considered as a subset of $[n]$. Thus, if A does not contain any non-trivial 3-AP, by Lemma 2.12 we obtain, as a big bound,

$$\theta(\beta\gamma)^{1/2}\geq\alpha\beta\gamma/2,$$

so $\theta\geq\alpha(\beta\gamma)^{1/2}/2$. Using the hypothesis, the density of β and γ is at least $\delta/5$ and $\alpha=\delta$, so

$$\alpha(\beta\gamma)^{1/2}/2\geq\delta^2/10,$$

as we wanted to show. □

The next step is to prove that if there exits a large non-trivial Fourier coefficient of $\hat{A}(r)$, then A has a slightly greater intersection with a progression of length $\sqrt{n}$:

Lemma 2.19. *Let n be a positive integer, let $r\in\mathbb{Z}_n$ and let $\delta>0$. Then there is a partition of $[n]$ into arithmetic progressions P_i of length at least $\delta\sqrt{n}/16$ such that for every i and every $x,y\in P_i$ we have $|\omega^{rs}-\omega^{ry}|<\delta$.*

Proof. Let $m = \lfloor \sqrt{n} \rfloor$. If we have m elements distributed into the circumference, the sum of its distances must be equal to 2π. So by *Pigeonhole principle*, there must exist $u, v \in [m]$ such that $|\omega^{ru} - \omega^{rv}| \leq 2\pi m^{-1}$. Let us set $d = |u - v|$, so we can write $|\omega^{r(x+d)} - \omega^{rx}| \leq 2\pi m^{-1}$. Since $|\omega^a| = 1 \ \forall a$, it follows

$$|\omega^{r(x+d)} - \omega^{rx}| \leq 2\pi m^{-1} \iff |\omega^{rd}(\omega^{r(x+d)} - \omega^{rx})| = |\omega^{r(x+2d)} - \omega^{r(x+d)}| \leq 2\pi m^{-1}.$$

Multiplying two times by ω^{rd}, we have

$$|\omega^{r(x+3d)} - \omega^{r(x+2d)}| \leq 2\pi m^{-1}.$$

So taking $x = y + td$,

$$\begin{aligned} |\omega^{r(x+t)} - \omega^{rx}| = |\omega^{r(y+td)} - \omega^{ry}| &= |\omega^{r(y+td)} - \omega^{r(y+(t-1)d)} + \omega^{r(y+(t-1)d)} - \omega^{ry}| \\ &\leq |\underbrace{\omega^{r(y+td)} - \omega^{r+(t-1)d}}_{\leq 2\pi m^{-1}}| + |\omega^{r(y+(t-1)d)} - \omega^{ry}| \leq \ldots \leq \frac{2\pi t}{m} \end{aligned}$$

Now, let us partition $[n]$ into residue classes $\mod d$:

1	$1+d$	$1+2d$	$\ldots$
2	$2+d$	$2+2d$	$\ldots$
$\vdots$	$\vdots$	$\vdots$	
$d-1$	$2d-1$	$3d-1$	$\ldots$
d	$2d$	$3d$	$\ldots$

Notice that every class will have length greater $\lfloor \sqrt{n} \rfloor = m$. Now, we want that for all elements x, y in every class, $|\omega^{rx} - \omega^{ry}| < \delta$. Since we know that for all $u, v \in [m]$, $|\omega^{ru} - \omega^{rv}| \leq 2\pi t m^{-1}$, we have $t = \delta m / 2\pi$. Thus, partitioning the residue classes into arithmetic progressions of length between $\delta m/4\pi$ and $\delta m/2\pi$, we get the diameter $\omega_r(P) \leq \delta$, as we wanted.

Recall that we want to work with progressions from at least $\delta\sqrt{n}/16$. Suppose that we partition a class into progressions of length $\delta m/2\pi$ and the length of the last one is less than this number. We can always unify the last class with the penultimate and divide by 2. This procedure will give us two classes of length $\delta m/4\pi$, which is greater than $\delta\sqrt{n}/16$, as we wanted. □

The next corollary introduces us the idea of density increment for the Fourier analytic proof:

Corollary 2.20. *Let n be a positive integer and let $A \subset \mathbb{Z}_n$ be a subset of density α. Suppose that there exists $r \neq 0$ such that $|\hat{A}(r)| \geq \theta$. Then, there exits an arithmetic progression $P \subset [n]$ of cardinality at least $\theta\sqrt{n}/32$ such that if we consider P as a subset of $\mathbb{Z}_n$, then*

$$\frac{|A \cap P|}{|P|} \geq \alpha + \frac{\theta}{4}.$$

Proof. First, let us consider the following function $f : \mathbb{Z}_n \longrightarrow \mathbb{Z}_n$ such that

$$f(x) = \begin{cases} 1-\alpha & x \in A, \\ -\alpha & x \notin A. \end{cases}$$

where the average is 0. Notice $\mathbb{E}_x f(x) = \mathbb{E}_x A(x) - \alpha = \alpha - \alpha = 0$. Also, $\hat{f}(r) = \hat{A}(r)$ for every $r \neq 0$. Thus, by hypothesis, there exists $r \neq 0$ such that $|\hat{f}(r)| \geq \theta$.

Applying Lemma 2.19, we can partition $[n]$ into arithmetic progressions $P_1, \ldots, P_r$. Taking $\delta = \theta/2\pi$ we have that the length of every P_i must be at least $\theta\sqrt{n}/32$. Since $|\mathbb{E}_x f(x)\omega^{-rx}| \geq \theta$, it follows that

$$\sum_i |P_i||\mathbb{E}_{x\in P_i} f(x) w^{-rx} \geq \theta \sum_i |P_i| = \theta n.$$

Let $x_i \in P_i$. Recall that the diameter $\omega_r(P_i) \leq \theta/2$ and $|f(x)| \leq 1$ for each i, so

$$\begin{aligned} |\mathbb{E}_{x\in P_i f(x)}\omega^{-rx}| &\leq |\mathbb{E}_{x\in P_i} f(x)\omega^{-rx_i}| + \mathbb{E}_{x\in P_i}|f(x)||\omega^{-rx} - \omega^{-rx_i}| \\ &\leq \mathbb{E}_{x\in P_i}|f(x)||\omega^{-rx_i}| + \frac{\theta}{2} \leq \mathbb{E}_{x\in P_i}|f(x)| + \frac{\theta}{2}. \end{aligned}$$

Thus,

$$\sum_i |P_i||\mathbb{E}_{x\in P_i} f(x)| \geq \frac{\theta}{2}\sum_i |P_i|.$$

Since $\sum_i |P_i|\mathbb{E}_{x\in P_i} f(x) = 0$ and using the inequality we got, there must exists i such that

$$|P_i|(|\mathbb{E}_{x\in P_i} f(x)| + \mathbb{E}_{x\in P_i} f(x)) \geq \frac{\theta|P_i|}{2}.$$

Now, we know that $\mathbb{E}_{x\in P_i} f(x)$ must be positive so we do not have to work with the absolute value anymore:

$$2|P_i|\mathbb{E}_{x\in P_i} f(x) \geq \frac{\theta|P_i|}{2} \quad \Longrightarrow \quad \mathbb{E}_{x\in P_i} f(X) \geq \frac{\theta}{4}.$$

By definition,

$$\mathbb{E}_{x\in P_i} f(x) = \mathbb{E}_{x\in P_i} A(x) - \alpha \geq \frac{\theta}{4} \iff \mathbb{E}_{x\in P_i} A(x) = \frac{|A\cap P_i|}{|P_i|} \geq \alpha + \frac{\theta}{4},$$

as we wanted. □

In the Fourier analytic proof, we also have an iteration argument, which is described by the following lemma:

Lemma 2.21. *Let $A \subseteq [n]$ be a set of density α and suppose that A does not contain any non-trivial 3-AP. Then, there exists a subprogression $P \subseteq [n]$ of cardinality at least $\alpha^2\sqrt{n}/500$ such that*

$$\frac{|A\cap P|}{|P|} \geq \alpha + \alpha^2/40.$$

Proof. Let us choose a set in order to apply Corollary 2.18. If n is odd, we set $t = n$. If n is even, we have two cases:

- If $n/2$ is odd, either $[n/2]$ or $\{n/2+1, \dots, n\}$ must have density α, so let us take that subset and set $t = n/2$.
- If $n/2$ is even, either $[n/2-1]$ or $\{n/2, \dots, n\}$ must have density α, so we take that one and set $t = n/2 - 1$ or $n/2 + 1$.

Corollary 2.18 tells us that either:

- A contains a 3-AP or
- $|A \cap [t/3, 2t/3)| < \frac{\alpha t}{5}$ or
- there exists $r \neq 0$ such that $|\hat{A}(r)| \geq \alpha^2/10$ when A is considered as a subset of $\mathbb{Z}_t$.

In the first case, we are done. In the second case, we now that one of the sets $A \cap [1, t/3)$ or $A \cap [2t/3, t)$ has cardinality at least $2\alpha t/5$ and therefore, density at least $11\alpha/10$ (the natural bound will be $(2\alpha t/5)/(t/3) = 6\alpha/5$, but notice that if t is a multiple of 3, then the set $A \cap [2t/3, t)$ has cardinality $t/3 + 1$ rather than $t/3$, so $(2\alpha t/5)/(t/3 + 1) \geq 11\alpha/10$), so we are done.

In the third case, we can use Corollary 2.18, which gives us $\theta = \alpha^2/10$, the hypothesis we need in order to apply Corollary 2.20. Thus, we have a progression P given by this last corollary satisfying the conclusion of this lemma, since $\theta\sqrt{t}/32 \geq \alpha^2\sqrt{n}/500$. □

Now we have all the ingredients to prove Theorem 0.5:

Proof of Theorem 0.5 (Roth). Let $n \geq (500/\alpha^2)^6$, so $\alpha^2\sqrt{n}/500 \geq n^{1/3}$. Thus, by Lemma 2.21, either A contains a 3-AP or there exist a progression P of size $n^{1/3}$ inside of which A has density at least $\alpha + \alpha^2/40$. If we are in the first case of Lemma 2.21, we are done.

In the second case, we just need to repeat the argument and recall that the critical value of the density is 1, so we need to be sure that after reaching this value we find a 3-AP. Note that iterating it $40/\alpha$ times, the density will reach 2α. After a further $40/2\alpha$ iterations, it will reach 4α, and so on. So the maximum number of iteration required is $80/\alpha$, since

$$(40/\alpha)(1 + 1/2 + 1/4 + \dots) = 80/\alpha.$$

Therefore, we need to choose n such that $n^{3^{-80/\alpha}} \geq (500/\alpha^2)^6$ in order to ensure that after every iteration, the cardinality of our progression is at least $\alpha^2\sqrt{n}/500$. Taking logs

$$3^{-80/\alpha} \log(n) \geq 6(\log(500) + 2\log(1/\alpha)),$$

and taking logs again:

$$\log\log(n) \geq 80/\alpha \log(3) + \log(6) + \log(\log(500) + 2\log(1/\alpha)) \geq c/\alpha$$

for some absolute constant c. So we have that for a set of density $\alpha \geq c/\log\log(n)$, so with size

$$c \cdot \left(\frac{n}{\log\log(n)} \right),$$

there must exist a 3-AP. This proves Roth's Theorem. □

Note that this proof gives us a stronger result compared to the combinatorial one: in this case, we have an extra $\log\log(n)$ dividing.

2.4 Problems on the generalisation to 4-AP

Let us now return to the finite field. Recall that one of the basic principles in order to prove Meshulam's Theorem is the use of the Fourier identity that we have seen in Lemma 2.14:

$$T_3(1_A, 1_A, 1_A) := \mathbb{E}_{x,d} 1_A(x) 1_A(x+d) 1_A(x+2d) = \sum_t \hat{1}_A(t)^3,$$

which establishes a relation between the sum of Fourier coefficients and the estimation of the number of 3-AP progressions in A. One of the questions that may arise is if this identity can be construct for 4-AP. Since the right-hand side sum will require two different parameters, it will be impossible to give an estimation. However, the analogue proof of Lemma 2.14, which is a necessary premise of the Meshulam's Theorem, does not hold for 4-term progressions.

The following theorem reveals an example of a set which is uniform and contains a more than expected number of 4-progressions: the dichotomy, which was a very relevant element of the proof, is not present for this case:

Theorem 2.22. *Let $p > 4$ be a prime. There exits $\varepsilon > 0$ such that for every $\delta > 0$ there exists n and a set $A \subseteq G = \mathbb{F}_3^n$ of density α with the following properties:*

(i) The set A is uniform in the sens that $sup_{t \neq 0} |\widehat{1}_A(t)| \leq \delta$.

(ii) The set A contains significantly more than the expected number of 4-AP, namely at least a proportion of $\alpha^4 + \varepsilon$.

Proof. Let us prove the set

$$A := \{x \in \mathbb{F}_p^n : x \cdot x = 0\}$$

verifies the properties. First, the characteristic function of this set can be given by

$$1_A(x) = \mathbb{E}_u \omega^{u(x \cdot x)},$$

where the expectation of u is taken over $\mathbb{F}_p$. Let us bound the Fourier coefficient 1_A at $t \neq 0$ in absolute value. Writing $q_u(x) = u(x \cdot x) + x \cdot t$, we have

$$|\widehat{1}_A(t)| \overset{\widehat{1}_A(t)=\widehat{1}_A(-t)}{=} |\mathbb{E}_{x\in\mathbb{F}_p^n}(\mathbb{E}_{u\in\mathbb{F}_p}\omega^{u(x\cdot x)})\omega^{-x\cdot t}| = |\mathbb{E}_{x\in\mathbb{F}_p^n,u\in\mathbb{F}_p}\omega^{u(x\cdot x)+x\cdot t}| \leq |\mathbb{E}_{u\neq 0}|\mathbb{E}_{x\in\mathbb{F}_p^n}\omega^{q_u(x)}|.$$

Now, using a Gauss sum to estimate the quadratic exponential sum,

$$\mathbb{E}_{u\neq 0}|\mathbb{E}_{x\in\mathbb{F}_p^n}\omega^{q_u(x)}| \leq p^{-n/2}.$$

So we have proved for any δ given n large enough, (i) is verified. Using the same procedure, we find that $|\mathbb{E}_x 1_A(x) - p^{-1}| \leq p^{-n/2}$, since $u = 0$ with probability p^{-1} and the rest of the elements are small. Therefore, the density of A can be approximately p^{-1} for n large enough.

In order to prove (ii), we need to count the number of 4-AP in A. To do so, let us define

$$L_i(x, d) = x + (i - 1)d,$$

for $(x, d) \in (\mathbb{F}_p^n)^2$. We say that we have a 4-AP in A if $L_i \cdot L_i = 0$ for $1 \leq i \leq 4$. But notice that $L_1 \cdot L_1 - 3L_2 \cdot L_2 - 3L_3 \cdot L_3 - L_4 \cdot L_4$ is equal to

$$x \cdot x - 3(x + d) \cdot (c + d) + 3(x + 2d) \cdot (x + 2d) - (x + 3d) \cdot (x + 3d) = 0 \quad \text{for } x, d \in G.$$

Therefore, since $L_1 \cdot L_1 - 3L_2 \cdot L_2 - 3L_3 \cdot L_3 - L_4 \cdot L_4 = 0$, we can assure that $L_4 \cdot L_4$ is zero whenever $L_1 \cdot L_1, L_1 \cdot L_1, L_3 \cdot L_3$ are simultaneously zero. Indeed, counting the number of 3-AP in A is enough. We have seen that the non-trivial coefficients of A are small, so we can use Lemma 2.14 to assure that A contains the number of 3-AP expected in a random case.

Thus, the number of 4-AP in A is equal to the number of 3-AP, which $\sim \alpha^3 >> \alpha^4 + \varepsilon$, for ε small enough. □

Even if we cannot use the same procedure, Gowers generalized the Fourier analysis to what he called *quadratic Fourier analysis* [12]. This generalization earned him the Fields Medal in 1998.

Chapter 3

Graph theoretical proof

In this chapter we will present the generalized graph theory developed by Endre Szemerédi in order to prove Erdős–Turán conjecture. He first created a technique for bipartite graphs [24] for his original proof, and once proved the conjecture, it was generalized. We will use it with the purpose of presenting a different proof for the particular case of $k = 3$, Roth's Theorem.

3.1 Szemerédi's Regularity Lemma

In this section we prove a result that Szemerédi developed in one of his courses; a graph-theoretical tool that has dominated in extremal graphs methods ever since: the *Regularity Lemma*. We will be following Diestel's book *Graph Theory* [6].

We begin with some definitions:

Definition 3.1 (Density of (X, Y))**.** Let $G = (V, E)$ be a graph with V vertices and E edges. Let $X, Y \subseteq V$ be disjoint subsets of vertices. Let us denote by $||X, Y||$ the number of edges between X and Y. Then we call the *density* of the pair (X, Y)

$$d(X, Y) := \frac{||X, Y||}{|X||Y|}.$$

Notice that $0 \leq d(X, Y) \leq 1$.

Definition 3.2 (ε-regular pair)**.** Given $\varepsilon > 0$, we call a pair (A, B) of disjoints sets $A, B \subseteq V$ ε-regular if all $X \subseteq A$ and $Y \subseteq B$ with

$$|X| \geq \varepsilon|A| \quad \text{and} \quad |Y| \geq \varepsilon|B|$$

satisfies

$$|d(X, Y) - d(A, B)| \leq \varepsilon.$$

So the smaller we take ε, the more uniformly distributed our edges will be in an ε-regular pair.

Definition 3.3 (ε-regular partition)**.** Consider a partition $\{V_0, V_1, \ldots, V_k\}$ in which V_0 has been singled out as an *exceptional set*. We call such a partition an ε-regular partition of G if it satisfies the following three conditions:

(i) $|V_0| \leq \varepsilon |V|$,

(ii) $|V_1| = |V_2| = \cdots = |V_k|$,

(iii) all but at most εk^2 of the pairs (V_i, V_j) with $1 \leq i < j \leq k$ are ε-regular.

The set V_0 can be seen as a *residual* set; it makes possible to demand to all the other sets to have the same size, so its vertices are ignored when the regularity of the partition is achieved.

After seeing the definitions above, let us formulate the *Regularity Lemma*:

Theorem 3.4 (Szemerédi's Regularity Lemma)**.** *For every $\varepsilon > 0$ and every integer $m \geq 1$, there exists an integer M such that every graph of order at least m admits and ε-regular partition $\{V_0, V_1, \ldots, V_k\}$ with $m \leq k \leq M$.*

In other words, the Regularity Lemma assures that given any $\varepsilon > 0$, every graph has an ε-regular partition into a bounded number of sets. The upper bound M on the number of sets guarantees that for large graphs the partition is large too. The lemma also allows us to specify a lower bound m for the number of partition sets; this can be used to increase the proportion of sets running between different partition sets over edges inside partition sets. Note that the Regularity Lemma as it has been announced is used with dense graphs. For sparse graphs it becomes trivial, because all densities of pairs tend to zero. In order to prove the Regularity Lemma, let us begin proving results:

Lemma 3.5. *For $\eta_1, \ldots, \eta_k > 0$ and $e_1, \ldots, e_k \geq 0$:*

$$\sum \frac{e_i^2}{\eta_i} \geq \frac{(\sum e_i)^2}{\sum \eta_i}.$$

Proof. Let $a_i := \sqrt{\eta_i}$ and $b_i := e_i/\sqrt{\eta_i}$. Using Cauchy-Schwarz

$$\sum a_i^2 \sum b_i^2 \geq (\sum a_i b_i)^2, \tag{3.1}$$

we have

$$\sum b_i^2 \geq \frac{(\sum a_i b_i)^2}{\sum b_i^2} \iff \sum \frac{e_i^2}{\eta_i} \geq \frac{(\sum e_i)^2}{\sum \eta_i}.$$

as we wanted. □

Now, let us define what we will call the *energy function*:

Definition 3.6. Let $G = (V, E)$ be a graph and $n := |V|$. For disjoint sets $A, B \subseteq V$ we define

$$q(A, B) := \frac{|A||B|}{n^2} d^2(A, B) = \frac{||A, B||^2}{|A||B|n^2}. \tag{3.2}$$

For partitions $\mathcal{A}$ of A and $\mathcal{B}$ of B we set

$$q(\mathcal{A}, \mathcal{B}) := \sum_{A' \in \mathcal{A};\ B' \in \mathcal{B}} q(A', B'),$$

and for a partition $\mathcal{P} = \{C_1, \ldots, C_k\}$ of V we let

$$q(\mathcal{P}) := \sum_{i<j} q(C_i, C_j).$$

If $\mathcal{P} = \{C_0, C_1, \ldots, C_k\}$ is a partition of V with exceptional set C_0, we treat C_0 as a set of singletons and define

$$q(\mathcal{P}) := q(\tilde{\mathcal{P}}),$$

with $\tilde{\mathcal{P}} := \{C_1, \ldots, C_k\} \cup \{\{v\} : v \in C_0\}$.

This function, that looks like an *energy function*, will play a principal role in the proof of the Regularity Lemma. First, it measures the regularity of the partition $\mathcal{P}$: if $\mathcal{P}$ has too many irregular pairs (A, B), we may take the pairs (X, Y) of subsets which do not obey the regularity of the pairs (A, B), and make those sets X and Y into partition sets of their own. As we will see, it refines $\mathcal{P}$ into a partition where this *energy function* q is greater. Let us prove that we would not be able to increase forever the value of q:

Proposition 3.7. *Let $\varepsilon > 0$ and let $\mathcal{P}$ be a partition of G. Then,*

$$q(\mathcal{P}) \leq 1.$$

Proof.

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i<j} q(C_i, C_j) \\ &= \sum_{i<j} \frac{|C_i||C_j|}{n^2} d^2(C_i, C_j) \\ &\leq \frac{1}{n^2} \sum_{i<j} |C_i||C_j| \\ &\leq 1. \end{aligned}$$

□

Notice that if the function $q(\mathcal{P})$ is bounded, the number of times that it can be increased is also bounded. This means that we can always construct an ε-regular partition for any

given graph after a bounded number of refinements.

In order to complete the proof, all we have to do is to note how many sets that last partition (which will be ε-regular) can have if we start with a partition into m sets and chose this number as our bound M. We begin showing that when we refine a partition, the value of q will not decrease:

Lemma 3.8.

(*i*) *Let $C, D \subseteq V$ be disjoint. If $\mathcal{C}$ is a partition of C and $\mathcal{D}$ is a partition of D, then $q(\mathcal{C}, \mathcal{D}) \geq q(C, D)$.*

(*ii*) *If $\mathcal{P}, \mathcal{P}'$ are partitions of V and $\mathcal{P}'$ refines $\mathcal{P}$, then $q(\mathcal{P}') \geq q(\mathcal{P})$.*

Proof.

(*i*) Let $\mathcal{C} =: \{C_1, \ldots, C_k)$ and $\mathcal{D} =: \{D_1, \ldots, D_l\}$. Then,

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \sum_{i,j} q(C_i, D_j) \\ &= \frac{1}{n^2} \sum_{i,j} \frac{||C_i, D_j||}{|C_i||D_j|}. \end{aligned}$$

Now, using Cauchy-Schwartz inequality (3.1) and the fact $\sum_{i,j} ||C_i, D_j|| =$ total of the edges,

$$\begin{aligned} \frac{1}{n^2} \sum_{i,j} \frac{||C_i, D_j||}{|C_i||D_j|} &\geq \frac{1}{n^2} \frac{(\sum_{i,j} ||C_i, D_j||)^2}{\sum_{i,j} |C_i||D_j|} \\ &= \frac{1}{n^2} \frac{||C, D||^2}{(\sum_i |C_i|)(\sum_j |D_j|)} \\ &= q(C, D). \end{aligned}$$

So $q(\mathcal{C}, \mathcal{D}) \geq q(C, D)$ as we wanted.

(*ii*) Let $\mathcal{P} =: \{C_1, \ldots, C_k\}$, and for $i = 1 \ldots, k$ let C_i be the partition of C_i induced by $\mathcal{P}'$. Notice that $q(\mathcal{P}') = \sum_i q(\mathcal{C}_i) + \sum_{i<j} q(C_i, C_j)$, so

$$\begin{aligned} q(\mathcal{P}) &= \sum_{i<j} q(C_i, C_j) \\ &\overset{(i)}{\leq} \sum_{i<j} q(\mathcal{C}_i, \mathcal{C}_j) \\ &\leq q(\mathcal{P}'). \end{aligned}$$

as we wanted. □

The next lemma proves that refining a partition by subpartitioning a not ε-regular pair of partition sets increases the value of our *energy function* q. Here, we find for the first time one of the most important arguments that appeared in the Fourier analytic proof, the *increasing density* version for graphs; *the energy increment.*

Lemma 3.9. *Let $\varepsilon > 0$ and let $C, D \subseteq V$ be disjoint. If (C, D) is not ε-regular, then there exist partitions $\mathcal{C} = \{C_1, C_2\}$ of C and $\mathcal{D} = \{D_1, D_2\}$ of D such that*

$$q(\mathcal{C}, \mathcal{D}) \geq q(C, D) + \varepsilon^4 \frac{|C||D|}{n^2}.$$

Proof. Suppose (C, D) is not ε-regular. Then there exist sets $C_1 \subseteq C$ with $|C_1| \geq \varepsilon|C|$ and $D_1 \subseteq D$ with $|D_1| \geq \varepsilon|D|$ such that

$$|\eta| := |d(C_1, D_1) - d(C, D)| > \varepsilon. \tag{3.3}$$

Let $\mathcal{C} := \{C_1, C_2\}$ and $\mathcal{D} := \{D_1, D_2\}$, where $C_2 := C \setminus C_1$ and $D_2 := D \setminus D_1$ so the sets in $\mathcal{C}$ and $\mathcal{D}$ are disjoint. Let us introduce a notation which will be easier to work with: let us write $c_i := |C_i|$, $d_i := |D_i|$, $e_{ij} := ||C_i, D_j||$, $c := |C|$, $d := |D|$, and $e := ||C, D||$. Now, let us show that $\mathcal{C}$ and $\mathcal{D}$ satisfy the conclusion of the lemma. Remember that $||A, B||$ denotes the number of edges between the sets A and B and $|A|$ denotes the cardinality of the set A. Firstly, by Lemma 3.8, we can write

$$\begin{aligned} q(\mathcal{C}, \mathcal{D}) &= \frac{1}{n^2} \sum_{i,j} \frac{||C_i, D_j||^2}{|C_i||D_j|} \\ &= \frac{1}{n^2} \left(\frac{||C_1, D_1||^2}{|C_1||D_1|} + \sum_{i+j>2} \frac{||C_i, D_j||^2}{|C_i||D_j|} \right) \\ &\geq \frac{1}{n^2} \left(\frac{||C_1, D_1||^2}{|C_1||D_1|} + \frac{(||C, D|| - ||C_1, D_1||)^2}{|C||D| - |C_1||D_1|} \right). \end{aligned}$$

where we have used Cauchy Schwartz inequality. Now, writing it in terms of our notation we have

$$q(\mathcal{C}, \mathcal{D}) \geq \frac{1}{n^2} \left(\frac{e_{11}^2}{c_1 d_1} + \frac{(e - e_{11})^2}{cd - c_1 d_1} \right).$$

Using definition 3.2 and working a little bit with it we can write:

$$e_{11} = \frac{c_1 d_1 e}{cd} + \eta c_1 d_1,$$

so now,

$$\begin{aligned} n^2 q(\mathcal{C},\mathcal{D}) &\geq \frac{1}{c_1 d_1}\left(\frac{c_1 d_1 e}{cd}+\eta c_1 d_1\right)^2 + \frac{1}{cd-c_1 d_1}\left(\frac{cd-c_1 d_1}{cd}e\eta c_1 d_1\right)^2 \\ &= \frac{c_1 d_1 e^2}{c^2 d^2} + \frac{2e\eta c_1 d_1}{cd} + \eta c_1 d_1 + \frac{cd-c_1 d_1}{c^2 d^2}e^2 - \frac{2e\eta c_1 d_1}{cd} + \frac{\eta^2 c_1^2 d_1^2}{cd-c_1 d_1} \\ &= \frac{e^2}{cd} + \eta^2 c_1 d_1 + \frac{\eta^2 c_1^2 d_1^2}{cd-c_1 d_1} \\ &\geq \frac{e^2}{cd} + \eta^2 c_1 d_1 \\ &\overset{(3.3)}{\geq} \frac{e^2}{cd} + \varepsilon^4 cd, \end{aligned}$$

where we have used $c_1 \geq \varepsilon c$ and $d_1 \geq \varepsilon d$ by the choice of C_1 and D_1. □

The final step left is showing that if a partition has enough irregular pairs of partition sets that do not verify the definition of an ε-regular partition, then subpartitioning all them at once triggers an increase of the *energy function* q by a constant. In this lemma we can remark that the dichotomy we had in the Fourier analytic proof is also present:

Lemma 3.10. *Let $0 < \varepsilon < 1/4$ and let $\mathcal{P} = \{C_0, C_1, \ldots, C_k\}$ be a partition of V, with exceptional set C_0 of size $|C_0| \leq \varepsilon n$ and $|C_1| = \ldots = |C_k| =: c$. If $\mathcal{P}$ is not ε-regular, then there exists a partition $\mathcal{P}' = \{C_0', C_1', \ldots, C_l'\}$ of V with exceptional set C_0', where $k \leq l \leq k4^{k+1}$ such that $|C_0'| \leq |C_0| + n/2^k$, all other sets C_i' have equal size and either $\mathcal{P}'$ is ε-regular or*

$$q(\mathcal{P}') \leq q(\mathcal{P}) + \frac{\varepsilon^5}{2}.$$

Proof. For all $1 \leq i < j \leq k$, let us define a partition $\mathcal{C}_{ij}$ of C_i and a partition $\mathcal{C}_{ji}$ of C_j as follows: if the pair (C_i, C_j) is ε-regular, we let $C_{ij} := C_i$ and $C_{ji} := \{C_j\}$. If not, by Lemma 3.9 there exists partitions $\mathcal{C}_{ij}$ of C_i and $\mathcal{C}_{ji}$ of C_j with $|\mathcal{C}_{ij}| = |\mathcal{C}_{ji}| = 2$ and

$$q(\mathcal{C}_{ij}, \mathcal{C}_{ji}) \geq q(C_i, C_j) + \varepsilon \frac{|C_i||C_j|}{n^2} = q(C_i, C_j) + \frac{\varepsilon^4 c^2}{n^2}. \tag{3.4}$$

So we take the not ε-regular couples (C_i, C_j) and create partitions $\mathcal{C}_{ij}$ and $\mathcal{C}_{ji}$ respecting the conditions of Lemma 3.9. Given all the partitions of C_i, let $\mathcal{C}_i$ for $i = 1, \ldots, k$ be the unique minimal partition of C_i that refines every partition $\mathcal{C}_{ij}$ with $j \neq i$: if an element does not coincide with another one in all of the partitions $\mathcal{C}_{ij}$ with $j \neq i$ in which it belongs to, then both of elements will not lie at the same $\mathcal{C}_i$ ($\mathcal{C}_i$ is the set of equivalence classes). Thus, $|\mathcal{C}_i| \leq 2^{k-1}$. Now, let us consider the partition

$$\mathcal{C} := \{C_0\} \cup \bigcup_{i=1}^{k} \mathcal{C}_i$$

of V, with C_0 as the exceptional set. Since $\mathcal{C}$ refines $\mathcal{P}$ and $|\mathcal{C} \setminus \{C_0\}| \leq k2^{k-1}$, we have

$$k \leq |\mathcal{C}| \leq k2^{k-1} + \varepsilon n \leq k2^k.$$

We define $\mathcal{C}_0 := \{\{v\} \text{ such that } v \in C_0\}$ as the exceptional set of the new partition. If $\mathcal{P}$ is not ε-regular, it means that for more than εk^2 of the pairs (C_i, C_j) with $1 \leq j \leq k$ the partition $\mathcal{C}_{ij}$ is non-trivial. Therefore, by Lemma 3.8 (i) and using the definition of the *energy function* q when existing an exceptional set, we can write

$$\begin{aligned} q(\mathcal{C}) &= \sum_{1\leq<j} q(\mathcal{C}_i, \mathcal{C}_j) + \sum_{1\leq i} q(\mathcal{C}_0, \mathcal{C}_i) + \sum_{0\leq i} q(\mathcal{C}_i) \\ &\geq \sum_{i\leq i<j} q(C_{ij}, C_{ji}) + \sum_{1\leq i} q(\mathcal{C}_0, \{C_i\}) + q(\mathcal{C}_0) \\ &\overset{(3.4)}{\geq} q(C_i, C_j) + \varepsilon k^2 \frac{\varepsilon^4 c^2}{n^2} + \sum_{i\leq i} q(\mathcal{C}_0, \{C_i\}) + q(\mathcal{C}_0) \\ &= q(\mathcal{P} + \varepsilon^5 \left(\frac{kc}{n}\right)^2 \\ &> q(\mathcal{P}) + \varepsilon^5/2. \end{aligned}$$

where for the last inequality, we have used that $|C_0| \leq \varepsilon n \leq n/4$, so $kc/n \geq 3/4 > 1/2$.

The last step that remains is turning $\mathcal{C}$ into the correct partition $\mathcal{P}'$. In order to do so, we will simply cut the sets of $\mathcal{C}$ into another ones with the same size and small and send the rest to the exceptional set. Since the set C_0 cannot grow to much, these new sets will be large enough.

If $c < 4^k$, we are done. Notice that in this case, we can set $l = kc = k4^k < 4k^{k+1}$, so we can set $C_0' := C_0$ and the singletons $\{v\}$ with $v \in V \setminus C_0'$ as desired, because we can make more partitions than elements we have, and for this case, the definition of ε-regularity becomes trivial.

Let us assume that $c \geq 4^k$. We will take the disjoints subsets $C_1', \ldots, C_l'$ of $\mathcal{C}$ with size $\lfloor c/4^k \rfloor \geq 1$ such that every C_i' is a subset of some $C \in \mathcal{C} \setminus C_0$, and we will define the new exceptional set as $C_0' = V \setminus \cup C_i'$. So our new partition $\mathcal{P}' = \{C_0, C_1' \ldots, C_l'\}$ is a partition of V which refines $\mathcal{C}$. In particular, by Lemma 3.8 (ii),

$$q(\mathcal{P}') \geq q(\mathcal{C}) \geq q(\mathcal{P}) + \frac{\varepsilon^5}{2}.$$

Now all the sets of the new partition $\mathcal{P}'$ have the same size except for the exceptional set C_0'. Recall that each set $C_i' \neq C_0'$ is included in one of the sets $C_1, \ldots, C_k$, but since we had set the size of our sets to $\lfloor c/4^k \rfloor$, no more than 4^k sets can lie in the same C_i for $1 \leq i \leq k$,

our partition verifies $k \leq l \leq k4^{k+1}$.

At this point all that remains is to look after the set C_0'. Using the fact that $C_1', \ldots, C_l'$ have at most $\lfloor c/4^k \rfloor$ vertices of each set, the vertices added to C_0' have to be less than $\lfloor c/4^k \rfloor$ for every C of $\mathcal{C} \setminus \{C_0\}$:

$$\begin{aligned} |C_0'| &\leq |C_0| + \lfloor c/4^k \rfloor |\mathcal{C}| \overset{3.10}{\leq} |C_0| + \frac{c}{4^k}(k2^k) \\ &= |C_0| + \frac{ck}{2^k} \leq |C_0| + \frac{n}{2^k}. \end{aligned}$$

as we wanted. □

At that point we are almost finished. The only part left is to find the value of M and thus the bounds for k and n. To do so, we will iterate Lemma 3.10:

Proof of Theorem 3.4 (Szemerédi's Regularity Lemma). Without lost of generality, let $0 < \varepsilon \leq 1/4$ and $m \geq 1$ be given. Recall that by Lemma 3.10, the maximum number of iterations required in order to obtain an ε-regular partition is $i := 2/\varepsilon^5$, due to the fact that our energy function verifies $q(\mathcal{P}) \leq 1$ for any partition $\mathcal{P}$.

In order to apply Lemma 3.10, our partition $\{C_0, C_1, \ldots, C_k\}$ has to satisfy:

– $|C_1| = \ldots = |C_k|$,

– $|C_0| \leq \varepsilon n$.

We have shown that with every iteration of the lemma, the size of the exceptional set grows by at most $n/2^{\tilde{k}}$ for a partition of $\tilde{k}+1$ elements, and we want to be sure that after every iteration, $|C_0| \leq \varepsilon n$. In order to do so, notice that with every iteration, the number of the sets of the partition rest equal or increase, so we can bound every iteration by $n/2^{\tilde{k}}$. Consequently, we have to choose n and k such that after i iterations, we never exceed εn. The first move will be to set the initial size of C_0: we will choose it to be $|C_0| \leq \frac{1}{2}\varepsilon$. Notice that to be able to achieve $|C_1| = |C_2| = \ldots = |C_k| = \lfloor n/k \rfloor$, we need to let $|C_0| < k$ (otherwise we can increase by one the cardinal of the non-exceptional sets). So in particular, we need to choose n such that $k < \frac{1}{2}\varepsilon n$, so $n > 2k/\varepsilon$.

Recall that by Lemma 3.10, the maximum number of iterations before getting an ε-regular partition must be $i = 2/\varepsilon^5$. In order to find a bound for k, we need to bound the total increase once we have the ε-regular partition. Therefore we set $k \geq m$ to be large enough so $in/2^k \leq \frac{1}{2}\varepsilon n$. Thus

$$k + \frac{i}{2^k} n \leq \varepsilon n \tag{3.5}$$

since $k < \frac{1}{2}\varepsilon n$. This way after $i = 2/\varepsilon^5$ iterations, the exceptional set will have grown up to at most to εn.

Let us now set M. Note that by Lemma 3.10, if we have a partition of r elements, in the next iteration the partition will grow to at most $r4^{r+1}$. So let us define $f : x \longrightarrow x4^{x+1}$. If we set $M := \max\{f^i(k), 2k/\varepsilon\}$, then M will be greater than the number of partitions after i iterations. In particular $n \geq M$ will be large enough to satisfy (3.5).

The last step that remains is to show that every graph $G = (V, E)$ of order at least m has an ε-regular partition $\{C_0, C_1, \ldots, C'_k\}$ with $m \leq k' \leq M$. To do so let us work with two cases: if $n \leq M$ or $n > M$, where $n := |G|$:

- If $n \leq M$, we partition G into $k' = n$ singletons verifying $|V_1| = \ldots = |V'_k| = 1$ and $|V_0| = \emptyset$, which is trivially an ε-regular partition, so we are done.
- If $n > M$, we let C_0 to be a set with the minimal cardinal such that $k' := k$ (where this k is the one verifying (3.5)) divides $|V \setminus C_0|$, and $\{C_1, \ldots, C_k\}$ be a partition of $|V \setminus C_0|$ with sets of the same size. Since the chose k verifies $k \leq 2k/\varepsilon < n$, then $|C_0| < k \leq \varepsilon n$ by (3.5). Beginning with the described partition, we iterate Lemma 3.10 until we obtain an ε-regular partition of G. Recall that the partition will be obtained after at most i iterations.

□

As we had already mentioned, the Regularity Lemma is one of the most important pieces in extremal graph theory. Unfortunately, the dependence of the parameters in the theorem is very bad: the algorithm which refines a partition requires $\frac{2}{\varepsilon^5}$ iterations in order to create a ε-regular partition. In particular, if we start with a partition of k sets, Lemma 3.10 in one iteration, will transform it into a partition of at most $k2^k \leq 2^{2^k}$ parts. Thus, in the end, our set will have

$$2^{2^{\cdot^{\cdot^{\cdot^{2}}}}} \Big\} \leq \frac{2}{\varepsilon^5}$$

parts.

3.2 Triangle counting Lemma and Triangle removal Lemma

One of the multiple applications of the Regularity Lemma is the *Triangle removal Lemma*. Informally, this lemma states that when a graph contains few copies of a triangle, then all of the copies can be eliminated by removing a small number of edges. In this section and in the next one we will present the notes that David Conlon followed in his course of extremal graph theory [5].

In order to prove this powerful lemma, we need to prove an other lemma first, the *Triangle counting Lemma*:

Lemma 3.11 (Triangle counting Lemma)**.** *Let $G = (V, E)$ be a graph and $X \cup Y \cup Z$ a partition of V. Suppose that $d(X, Y) = \alpha$, $d(X, Z) = \beta$ and $d(Y, Z) = \gamma$. Let $\varepsilon > 0$ such that $\min\{\alpha, \beta, \gamma\} \geq 2\varepsilon$. Suppose that all pairs $\{X, Y\}, \{Y, Z\}$ and $\{X, Y\}$ are ε-regular. If we write $\triangle xyz$ as the number of triangles with $x \in Y$, $y \in Y$ and $z \in Z$, then*

$$\triangle xyz \geq (1 - 2\varepsilon)|X|(\alpha - \varepsilon)|Y|(\beta - \varepsilon)|Z|(\gamma - \varepsilon).$$

Proof. For a vertex $v \in V$, we set $d_X(v)$, $d_Y(v)$, and $d_Z(v)$ the number of neighborhoods of v in X, Y and Z, respectively. Let us begin proving that we can control the number of vertices in X with small degree. We need to prove this result in order to apply the ε-regularity condition. Let us show

$$|\{x \in d_Y(x) < (\alpha - \varepsilon)|Y|\}| < \varepsilon|X|.$$

We will write $|X'| = |\{x \in d_Y(x) < (\alpha - \varepsilon)|Y|\}|$.

By *reductio ad absurdum*, suppose $|X'| \geq \varepsilon|X|$. Using the definition of ε-regular pair, $|d(X', Y) - d(X, Y)| < \varepsilon$. Since $d(X, Y) = \alpha$ and

$$d(X', Y) = \frac{||X', Y||}{|X'||Y|} = \frac{\sum_{x \in X'} d_Y(x)}{|X'||Y|} < \frac{\sum_{x \in X'} (\alpha - \varepsilon)|Y|}{|X'||Y|} = \alpha - \varepsilon,$$

we have

$$d(X', Y) - d(X, Y) < \alpha - \varepsilon - \alpha < -\varepsilon,$$

which contradicts the definition of ε-regularity. The same result is valid when studying the vertices with small degree in Y and Z. Thus, we can conclude

$$|\{x \in X : d_Y(x) \geq (\alpha - \epsilon) \text{ and } d_Z(x) \geq (\alpha - \varepsilon)|Z|\}| \geq (1 - 2\varepsilon)|X|.$$

In order to study in how many triangles an element is involved, let us take $x \in X \setminus W$, where W is the set verifying $|W| < 2\varepsilon|X|$ and which contains the elements with the *wrong* kind of degrees. If $N(x)$ denote the neighborhood of x, using that $\min\{\alpha, \beta, \gamma\} \geq 2\varepsilon$, we have

$$|N(x) \cap Y| = d_Y(x) \geq (\alpha - \varepsilon)|Y| \geq \varepsilon|Y| \qquad \text{and} \qquad |N(x) \cap Z| = d_Z(x) \geq (\beta - \varepsilon)|Z| \geq \varepsilon|Z|.$$

So we can apply the definition of ε-regularity to the pair $\{Y, Z\}$ to deduce that there exists *lots* of edges between this two sets. The number of edges will be:

$$||N(x) \cap Y, N(x) \cap Z|| \leq \underbrace{(\alpha - \varepsilon)|Y| \cdot (\beta - \varepsilon)|Z|}_{\text{product of the sizes}}, \qquad \cdot \qquad \underbrace{(\gamma - \varepsilon)}_{\text{edge density between them}}.$$

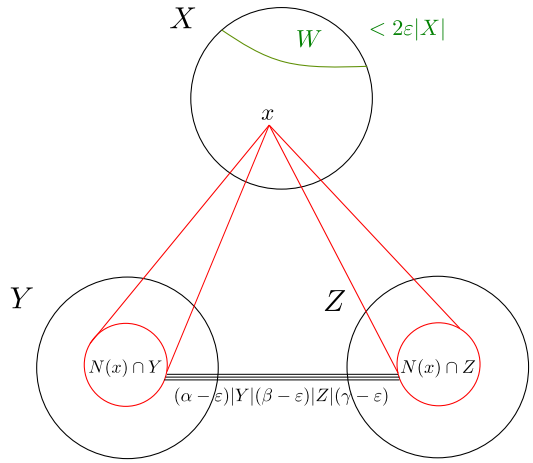


Figure 3.1: Diagram of the Triangle counting Lemma

Putting everything together, we have that

$$\triangle xyz \geq (1 - 2\varepsilon)|X|(\alpha - \varepsilon)|Y|(\beta - \varepsilon)|Z|(\gamma - \varepsilon),$$

as we wanted to show. □

Let us now prove the *Triangle removal Lemma* also known as the $(6, 3)$-problem, developed by Szemerédi and Ruszsa [19]:

Theorem 3.12 (Triangle removal Lemma)**.** *For every $\varepsilon > 0$, there exists $\delta := \delta(\varepsilon) > 0$ such that $\delta \longrightarrow 0$ when $\varepsilon \longrightarrow 0$, and $n_0 := n_0(\varepsilon)$ such that every graph G with $n \geq n_0$ vertices and at most δn^3 triangles, can be transformed into a triangle-free graph by removing at most εn^2 edges.*

Proof. Let us show the contrapositive proof: if we need to remove at least εn^2 edges in order to get a triangle-free graph, then the initial graph must have more than δn^3 triangles.

Let us take $\varepsilon > 0$ and set the minimum order of the graph G at $m = \lfloor \frac{4}{\varepsilon} \rfloor$. The first step of the proof is to consider a $\frac{\varepsilon}{4}$-regular partition of $G = V_0 \cup V_1 \cup \cdots V_k$, which exists by Szemerédi's Regularity Lemma. Let $c = |V_1| = |V_2| = \cdots = |V_k|$. Recall that $\frac{\varepsilon}{4} < k$ and that the number of vertices n verifies

$$n = |V| = \sum_{i\geq} |V_i| = |V_0| + \sum_{i\geq 1} c > kc,$$

so $kc < n$.

Let us begin with the second step of the proof: cleaning our partition. We will remove the edges of G from the following sets:

– Edges that are incident in V_0: we have at most $|V_0|n = \frac{\varepsilon}{4}n^2$ of these edges.

– Inner edges of $V_1, \dots, V_k$: we have at most $k\binom{c}{2} < ck^2 < \frac{n^2}{k} < \frac{\varepsilon}{4}n^2$ of these edges.

– Edges between not ε-regular pairs. Recall that by definition of an $\frac{\varepsilon}{4}$-regular partition, we have at most $\frac{\varepsilon}{4}k^2$ of these pairs, so the number of edges in this situation is less than $\frac{\varepsilon}{4}k^2c^2 < \frac{\varepsilon}{4}c^2$.

– All edges between low density pairs, in other words, edges between $\frac{\varepsilon}{4}$-regular pairs verifying $d(V_i, V_j) \leq \varepsilon n^2$. We have at most $\binom{k}{2}$ of these pairs, so the number of edges is bounded by $\sum_{i \neq j} d(V_i, V_j)|V_i||V_j| < \frac{k^2}{2}\frac{\varepsilon}{2}c^2 < \frac{\varepsilon}{4}n^2$.

Summarizing, adding all the contributions above, we will have removed at most εn^2 edges, so now we can begin with the last step of the proof. If at this point the resultant graph is triangle-free, we are done. So let us suppose that some triangle remains and find a contradiction. In this case, in order to obtain a triangle-free graph, we need to remove more edges. Recall that the edges that have survived must been defined between $\frac{\varepsilon}{4}$-regular pairs whose density is bigger than $\frac{\varepsilon}{2}$. This triangle must be sitting between three different sets, that we will call V_i, V_j, V_k. Now let us see that the hypothesis of Lemma 3.11 with $\varepsilon = \frac{\varepsilon}{4}$ are verified by these sets:

– $d(V_i, V_j) = \alpha$, $d(V_i, V_k) = \beta$, $d(V_j, V_k) = \gamma$, with $\min\{\alpha, \beta, \gamma\} \geq \frac{\varepsilon}{2}$.

– V_i, V_j and V_k form $\frac{\varepsilon}{4}$-regular pairs by hypothesis.

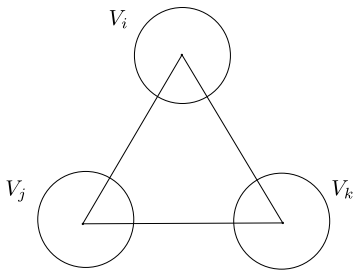


Figure 3.2: Triangle formed by the sets V_i, V_j and V_k

So by Lemma 3.11, these three sets define at least

$$(1 - \frac{\varepsilon}{2})(\frac{\varepsilon}{4})^2c^3$$

triangles. The last step that remains is writing this result in terms of n. Recall that $n = |V_0| + kc$, $|V_0| \leq \frac{\varepsilon}{4}n$ and $k \leq M(m, \frac{\varepsilon}{4}) := M(\varepsilon)$, so we have

$$n = |V_0| + ck \Longrightarrow c > \frac{1}{k}(1 - \frac{\varepsilon}{4})n > \underbrace{\frac{1}{M(\varepsilon)^3}(1 - \frac{\varepsilon}{4})^3n^3}_{\delta}.$$

Hence, we have proved the theorem: if we had not removed all of the triangles when removing δn^2 edges, our graph would had have more than δn^3 triangles, contradicting the hypothesis. □

3.3 Graph theoretical proof of Roth's Theorem

Now we have all the tools to provide a third proof of Roth's Theorem by only graph theoretical means. This version is equivalent to the announced before in Theorem 0.5.

Theorem 3.13 (Roth's Theorem)**.** *Let $A \subseteq [n]$. If A does not contain a 3-AP, then $|A| = o(n)$.*

Proof. Let us prove that for all $\varepsilon > 0$ and $A \subseteq [n]$ verifying $|A| > \varepsilon n$ for n large enough, then A must contain a 3-AP. For this purpose, we build an appropriate graph G in order to apply the Triangle removal Lemma 3.12. So for $A \subseteq [n]$, let us define the graph $H(A) = (V, E)$, whose set of vertices V can be written as the union of the three following disjoint sets

$$V = \{(i,1) : i \in [n]\} \cup \{(j,1) : j \in [2n]\} \cup \{(k,1) : k \in [3n]\}$$

(therefore $|V| = 6n$), and its set of edges is defined by:

- $(i,1)$ and $(j,2)$ are adjacent if and only if $j - i \in A$.
- $(j,2)$ and $(k,3)$ are adjacent if and only if $k - j \in A$.
- $(i,1)$ and $(k,3)$ are adjacent if and only if $k - i \in 2 \cdot A = \{2a : a \in A\}$.

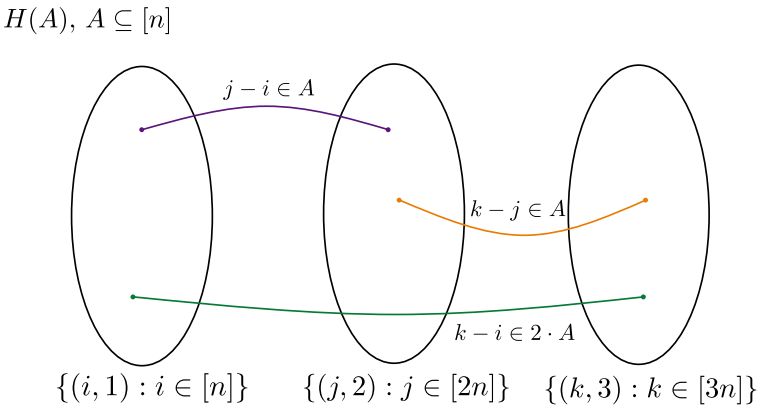


Figure 3.3: Diagram of the edges of $H(A)$

Observe that if there exists $(i,1)$, $(j,2)$ and $(k,3)$ defining a triangle in $H(S)$, then we can write

$$j - i = a_1, \qquad k - i = a_2, \qquad k - i = 2a_3,$$

with $a_i \in A$, so that $\{a_1, a_2, a_3\}$ defines a 3-AP. We can also take the elements $(i, 1)$, $(i+a, 2)$ and $(i + 2a, 3)$ with $a \in A$, which define the triangles associates to the *trivial* 3-APs: $a, a + 0, a + 2 \cdot 0$. Since there exists $|A| \cdot n$ triangles of that form, it will be required to remove at least $|A| \cdot n$ edges in order to get a free-triangle graph.

Let us show that if $|A| > \varepsilon n$ and trivially $|A| \leq n$, then there must exists a triangle, and thus a 3-AP. By Lemma 3.12:

- The number of edges that have to be removed in order to get a free-triangle graph, is at least $\varepsilon n^2 = \frac{\varepsilon}{36}(6n)^2 = \frac{\varepsilon}{36}|V|^2$.
- There exists $\delta := \delta(\varepsilon)$ such that $H(A)$ has at least $\delta|V|^3 = \delta 6^3 n^3$.

Thus, the number of non-trivial triangles is at least $\delta 6^3 n^3 - n^2$. Therefore, taking n such that $0 < \delta 6^3 n^3 - n^2$, so

$$n > \frac{1}{6^3 \delta},$$

we can assure the existence of a non-trivial triangle and therefore, A must contain a 3-AP. □

Notice that the property of supersaturation gives as something much more stronger that what we have proved: if the condition given for n is verified, then the number of triangles is cubic on n.

3.4 Roth's Theorem for non-abelian groups

One of the results we can prove using the Lemma 3.12 is the analogous version of Roth's Theorem for non-abelian groups. We present here a modified proof of the one that can be found in Lluís Vena's master thesis [28].

Theorem 3.14 (Roth's Theorem)**.** *Let G be a finite group of odd order N and let A be a subset of its elements. If there not exists $x, y, z \in A$ such that $xy = z^2$, then the size of A is $o(N)$.*

Proof. Let us prove that for all $\varepsilon > 0$ and $A \subseteq G$ verifying $|A| > \varepsilon N$ for N large enough, then A must contain a 3-AP. Let us begin defining a tripartite graph, G_1, G_2, G_3, where each G_i is a copy of the group G, so $A \subseteq G_i$ for $i = 1, 2, 3$. Suppose that for any $x, y, z \in A$, $xy \neq z^2$. To fix some notation, let us call a_i an element of G_i. We define an edge between G_i and G_j with $i \neq j$ whenever:

- $\exists a_1, a_2$ such that $g_1 = a_1 \cdot a_2^{-1} \in A$
- $\exists a_2, a_3$ such that $g_2 = a_2 \cdot a_3^{-1} \in A$
- $\exists a_1, a_3$ such that $g_3 = a_1 \cdot a_3^{-1} \in A$

This way $g1g2 = g3$.

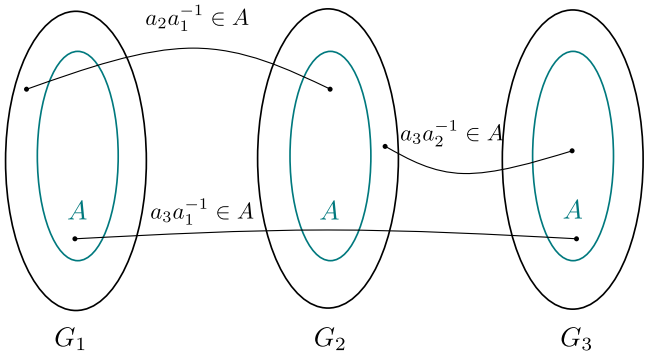


Figure 3.4: Diagram of the edges of the tripartite graph

Note that the equation we want g_1, g_2 and g_3 to verify is $xy = z^2$. Let us prove that if order of the group is odd, the function

$$\begin{aligned} f : G &\longrightarrow G \\ x &\longrightarrow x^2 \end{aligned}$$

describes a bijection. Note that it is enough to prove that the function is injective, since the order of the image is finite and the same as the order of the domain. Suppose that $x^2 = y^2$. Then,

$$(x^2)^{\frac{N+1}{2}} = (y^2)^{\frac{N+1}{2}} \iff x = y.$$

Therefore, we can change the third condition to define an edge to $g_3^2 = a_1a_3^{-1}$, with $g_3 \in A$. At this point, we can apply the Theorem 3.12 as we did before in 3.13. □

Remark that one of the advantages of using this graph techniques is that we can extend the result to non abelian groups, an impossible thing to do using other techniques.

Chapter 4

Bounds

Improving bounds of Roth's Theorem is a problem with a long history and work done. On Roth's Theorem, we can be interested to bound the largest subset of $[N]$ which contains no 3-APs, that we will call $r_3([N])$. Let us begin this section presenting some improvements on its upper bound.

4.1 Upper bound improvements

As we have already seen in chapter 2, the original proof given by Roth using Fourier methods showed that

$$r_3([N]) = c \cdot \frac{N}{\log \log N}$$

for some constant c. Over the years, the upper bound has been improved

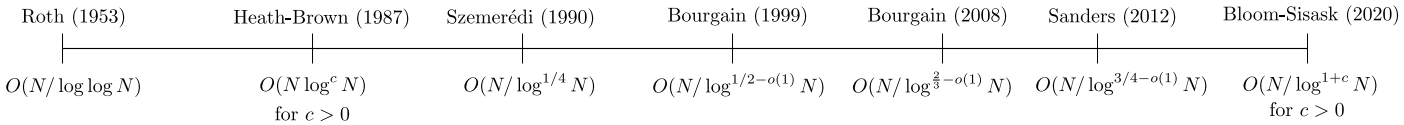


Figure 4.1: Timeline of upper bound improvements, Heath-Brown [14], Szemerédi [23], Bourgain[3][4], Sanders[20], Bloom-Sisask[26].

and as shown, in 2020, Thomas Bloom and Olof Siszac broke the logarithmic barrier in Roth's Theorem. The proof consist on developing an integer analogue of the result of Bateman and Katz [1] for the model setting of vector spaces over a finite field, also known as the cap set problem, which we have already seen the first approach given by Meshulam [16] in section 2.2 of chapter 2. Let us now explore the most recent improvement on the finite field model.

4.1.1 The polynomial method proof in the finite field model

In 2017, Ellenberg and Giswijt [7] found the best known bound for Roth's Theorem in $\mathbb{F}_3^n$,

$$r_3(\mathbb{F}_n^3) = O(2.76^n),$$

which improves upon the $O(3^n/n^{1+\epsilon})$ bound proved by Bateman and Katz [1]. The proof consists of a four-page paper that and on it can be seen the power of algebraic methods. To develop this proof, we will follow Yufei Zhao's notes on graph theory and additive combinatorics [25]. Let us begin with one of the main identities on the cap set problem:

Definition 4.1. Let $A \subseteq \mathbb{F}_3^n$ be a 3-AP-free subset. Then, we have the following identity,

$$\delta_0(x) := \sum_{a \in A} \delta_a(x)\delta_a(y)\delta_a(z). \tag{4.1}$$

for $x, y, z \in A$, where δ_a is the Dirac delta function, defined as:

$$\delta_a(x) := \begin{cases} 1 & \text{if } x = a, \\ 0 & \text{if } x \neq a. \end{cases}$$

Notice that the identity 4.1 holds because

$$x + y + z = 0 \iff z - y = y - x$$

in $\mathbb{F}_3^n$. It follows that if x, y, z form an arithmetic progression, then for some $a \in \mathbb{F}_3^n$ we have $x = y = z = a$. Now we will show that the right-hand side is "high-rank" and the left hand-side is "low-rank", using the following definition:

Definition 4.2. Let $F : A \times A \longrightarrow \mathbb{F}$, for a field $\mathbb{F}$. We say F is rank 1 if it is non-zero and can be written in the form $F(x, y) = f(x)g(y)$ for some functions $f, g : A \longrightarrow \mathbb{F}$.

Definition 4.3 (Slice-rank of F)**.** Let F : $A \times A \times A \longrightarrow \mathbb{F}$. We define the *slice-rank* of F as the minimum number of slice-rank 1 functions required to write F as a linear combination.

In particular, we say that F has slice-rank 1 if it is non-zero and can be written one of the following forms: $f(x)g(y, z)$, $f(y)g(x, z)$, $f(z)g(x, y)$. Recall from linear algebra that the rank of a diagonal matrix is the number of non-zero entries. We have a similar result for the slice-rank:

Lemma 4.4. *If $F : A \times A \times A \longrightarrow \mathbb{F}$ can be written as*

$$F(x, y, z) = \sum_{a \in A} c_a \delta_a(x)\delta_a(y)\delta_a(z),$$

then slice-rank $F = |\{a \in A : c_a \neq 0\}|$, *where c_a correspond to diagonal entries.*

Proof. Trivially, the slice-rank $F \leq |\{a \in A : c_a \neq 0\}|$, since we can write F as a sum of slice-rank 1 functions using the diagonal form

$$F(x, y, z) = \sum_{\substack{a \in A \\ c_a \neq 0}} c_a \delta_a(x)(\delta_a(y)\delta_a(z)).$$

For the other direction, let us suppose $F < |A|$ and find a contradiction. So we can write:

$$\begin{aligned} F(x,y,z) = & f_1(x)g_1(y,z) + \cdots + f_l(x)g_l(y,z) \\ & + f_{l+1}(y)g_{l+1}(x,z) + \cdots f_m(y)g_m(x,z) \\ & + f_{m+1}(z)g_{m+1}(x,y) + \cdots f_{|A|-1}(z)g_{|A|-1}(x,y). \end{aligned}$$

First, let us prove the following claim:

Claim. *There exists* $h : A \longrightarrow \mathbb{F}_3$ *with* $|\text{supp } h| > m$ *such that*

$$\sum_{z \in A} h(z)f_i(z) = 0$$

for all $i = m+1, \ldots, |A|-1$ *where* supp h *is the set* $\{z \in A : h(z) \neq 0\}$

Proof. Notice that in the vector space of functions $A \longrightarrow \mathbb{F}_3$, the set of h satisfying Lemma 4.4 is a subspace of dimension greater than m (recall that $|\text{supp } h| > m$). Let us prove that every subspace of dimension $m+1$ has a vector whose support (the number of non-zero elements) is at least $m+1$.

For a subspace X of dimension $m+1$, suppose we write $m+1$ vectors forming a basis of X in a $|A| \times (m+1)$ matrix Y. Since this matrix has rank $m+1$, it must exists some non-vanishing minor of order $m+1$. Let us suppose that we can find this minor in the first $m+1$ columns (otherwise, we can rearrange the columns), formed by the vectors $v_1, \ldots, v_{m+1}$, which are a basis of $\mathbb{F}_3^{m+1}$. So, taking the linear combination of those vectors we get, a new vector of support at least $m+1$. □

Let us now pick h from the claim. We find

$$\sum_{z \in A} F(x,y,z)h(z) = \sum_{a \in A}\sum_{z \in A} c_a \delta_a(x)\delta_a(y)\delta_a(z)h(z) = \sum_{a \in A} c_a h(a)\delta_a(x)\delta_a(y),$$

and using the claim:

$$\sum_{z \in A} F(x,y,z)h(z) = f_1(x)\tilde{g_1}(y) + \cdots + f_l(y)\tilde{g_l}(x) + f_{l+1}(y)\tilde{g_{l+1}}(x) + \cdots + f_m(y)\tilde{g_m}(x),$$

where $\tilde{g_i}(y) = \sum_{z \in A} g_i(y,z)h(z)$ for $1 \leq i \leq l$ and $\tilde{g_i}(x) = \sum_{z \in A} g_i(x,z)h(z)$ for $l+1 \leq i \leq m$.

Since this function is the sum of m rank 1 functions, then it will have at most rank m. But we can also write

$$\begin{aligned} \sum_{a \in A} c_a h(a)\delta_a(x)\delta_a(y) = & f_1(x)\tilde{g_1}(y) + \cdots + f_l(y)\tilde{g_l}(x) \\ & + f_{l+1}(y)\tilde{g_{l+1}}(x) + \cdots + f_m(y)\tilde{g_m}(x). \end{aligned}$$

Recall that the support of h is at least $m+1$, we know that this function will have rank greater than m, which is a contradiction given by the hypothesis slice-rank $F < |A|$. □

Now we are ready to prove Ellenberg and Giswijt's result:

Lemma 4.5. *Let us define* $F : A \times A \times A \longrightarrow \mathbb{F}$ *as:*

$$F(x+y+z) := \delta_0(x+y+z).$$

Then, the slice-rank $F \leq 3M$*, where*

$$M := \sum_{\substack{a,b,c\geq 0 \\ a+b+c=n \\ b+2c\leq 2n/3}} \frac{n!}{a!b!c!}.$$

Proof. In $\mathbb{F}_3^n$, we have $\delta_0(x) = 1 - x^2$. Now, imposing that the equality has to hold for each coordinate, we can write

$$\delta_0(x) = \prod_{i=1}^{n}(1-(x_i+y_i+z_i)^2), \tag{4.2}$$

where x_i are the coordinates of $x \in \mathbb{F}_3^n$. Note that on the right-hand side we have a polynomial in $3n$ variables with degree $2n$, formed by the sum of the following monomials

$$x_1^{i_1}\cdots x_n^{i_n} y_1^{j_1}\cdots y_n^{j_n} z_1^{k_1}\cdots z_n^{k_n},$$

where $i_1, i_2, \ldots, i_n, j_1, \ldots, j_n, k_1, \ldots, k_n \in \{1,2,3\}$. For each term, by the *Pigeonhole principle*, at least one of the sums $i_1+\ldots+i_n, j_1+\ldots+j_n, k_1+\ldots+k_n$ is at most $2n/3$ (otherwise the degree of the polynomial would be bigger than $2n$). Now, we can write (4.2) explicitly as

$$\prod_{i=1}^{n}(1-(x_i+y_i+z_i)^2) = \sum_{\substack{i_1,\ldots,i_n \\ j_1,\ldots,j_n \\ k_1,\ldots,k_n}} c_{i_1,i_2,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n} x_1^{i_1}\cdots x_n^{i_n} y_1^{j_1}\cdots y_n^{j_n} z_1^{k_1}\cdots z_n^{k_n},$$

where $c_{i_1,i_2,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n} \in \mathbb{F}_n^3$. Then, we can group the terms of the last equality as sum of slice-rank 1 functions in the following way:

$$\begin{aligned}\prod_{i=1}^{n}(1-(x_i+y_i+z_i)^2) = &\sum_{i_1+\ldots+i_n\leq\frac{2n}{3}} x_1^{i_1}\cdots x_n^{i_n} f_{i_1,\ldots,i_n}(y,z) \\ &+ \sum_{j_1+\ldots+j_n\leq\frac{2n}{3}} y_1^{j_1}\cdots y_n^{j_n} g_{j_1,\ldots,j_n}(x,z) \\ &+ \sum_{k_1+\ldots+k_n\leq\frac{2n}{3}} z_1^{i_1}\cdots z_n^{i_n} h_{k_1,\ldots,k_n}(x,y),\end{aligned}$$

where

$$f_{i_1,\ldots,i_n} = \sum_{\substack{j_1,\ldots,j_n \\ k_1,\ldots,k_n}} c_{i_1,i_2,\ldots,i_n,j_1,\ldots,j_n,k_1,\ldots,k_n} y_1^{j_1}\cdots y_n^{j_n} z_1^{k_1}\cdots z_n^{k_n},$$

and $g_{j_1,\ldots,j_n}$ and $h_{k_1,\ldots,k_n}(x,y)$ are likewise defined, but omitting the terms that already appear in the previous sums.

Since $\sum_k^n i_k \leq 2n/3$ for $i_k \in \{0,1,2\}$, let us define the sets

$$A = \{k \text{ such that } i_k = 0\}, \quad B = \{k \text{ such that } i_k = 1\} \quad \text{and} \quad C = \{k \text{ such that } i_k = 2\}$$

with $|A| = a$, $|B| = b$ and $|C| = c$. Now, using this notation we can write

$$|A| \cdot 0 + |B| \cdot 1 + |C| \cdot 2 \leq 2n/3.$$

Thus, the number of the monomials verifying this property can be counted by the following formula:

$$\sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \binom{n}{a,b,c} = \frac{n!}{a!b!c!} = M.$$

Doing the same procedure for each exponent sum, we have

$$\text{slice-rank} \left\{ \prod_{i=1}^{n} (1 - (x_i + y_i + z_i)^2) \right\} \leq 3M.$$

where the inequality appears to avoid overcounting. □

The last step that remains is to estimate M. First, notice that we can write

$$(1 + x + x^2)^n = \sum_{\substack{a,b,c\geq 0\\ a+b+c=n}} \binom{n}{a,b,c} x^{a\cdot 0} x^b x^{c\cdot 2} = \sum_{\substack{a,b,c\geq 0\\ a+b+c=n}} \binom{n}{a,b,c} x^{b+2c}.$$

Imposing now $b + 2c \leq 2n/3$, we get

$$\begin{aligned} \sum_{\substack{a,b,c\geq 0\\ a+b+c=n}} \binom{n}{a,b,c} x^{b+2c} &\geq \sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \binom{n}{a,b,c} x^{b+2c} \geq \sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \binom{n}{a,b,c} x^{2n/3} \\ &= \sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \frac{n!}{a!b!c!} x^{2n/3} \geq M x^{2n/3}. \end{aligned}$$

Thus, for $0 \leq x \leq 1$ (otherwise $x^{2n/3} > x^{b+2c}$ does not hold),

$$M x^{2n/3} \leq \sum_{\substack{a,b,c\geq 0\\ a+b+c=n\\ b+2c\leq 2n/3}} \frac{n!}{a!b!c!} x^{b+2c} \leq (1 + x + x^2)^n.$$

And now, looking for the minimum value of the following middle expression, which is found at $x = 0,6$, we can state

$$M \leq \inf_{0<x<1} \frac{(1 + x + x^2)^n}{x^{2n/3}} \leq (2.76)^n.$$

4.2 Lower bound improvements: Behrend's construction

We can also be interested in the lower bound: the largest set with no three-term arithmetic progressions. The best construction was given by Behrend in 1946 [2].

Theorem 4.6 (Behrend, 1948)**.** *Let N be large integer. Then there exists a subset $A \in [N]$ with $|A| = N\ e^{-c\sqrt{log(N)}}$ which does not contain any arithmetic progression of length three.*

Proof. Let us base our demonstration in the following geometrical observation: a straight line can intersect a sphere in $\mathbb{Z}^n$ in at most two points. Let n and M be larger integers and let us consider the following set:

$$S(r) = \{\bar{x} \in [M]^n : x_1^2 + x_2^2 + \ldots + x_n^2 = r^2\}.$$

Note that $n \leq r^2 \leq nM^2$. So r^2 can take $M^2n - n - 1 = n(M^2 - 1) - 1$ values.

These sets cover the cube $\{1, \ldots, M\}^n$, which has cardinal M^n. Therefore, by the *Pigeonhole Principle*, there exist a radius $\sqrt{n} \leq r_0 \leq \sqrt{n}M$ such that the sphere $S(r_0)$ has cardinal

$$|S(r_0)| \geq \frac{M^n}{n(M^2 - 1)} \geq \frac{M^{n-2}}{n}.$$

Let us now project these vertices into $\mathbb{N}$ using the following mapping:

$$P(x_1, x_2, \ldots, x_n) = \frac{1}{2M} \sum_{i=1}^{n} x_i (2M)^i,$$

which has the following properties:

(i) P is linear

(ii) P is a one to one mapping

(iii) $x + y = 2z \iff P(x) + P(y) = P(2z)$

(iv) $\max_{x \in S(r_0)} P(x) \leq (2M)^n$

Proof. Note that property (i) can be proven straightforward from the definition of P: if $x, y \in \mathbb{Z}^n$ and $a, b \in \mathbb{Z}$

$$\begin{aligned} P(ax + by) &= \frac{1}{2M} \sum_{i=1}^{n} (ax_i + by_i)(2M)^i \\ &= a\left(\frac{1}{2M} \sum_{i=1}^{m} x_i (2M)^i\right) + b\left(\frac{1}{2M} \sum_{i=1}^{m} y_i (2M)^i\right) \\ &= aP(x) + bP(y). \end{aligned}$$

In order to prove property (ii), note that $P(x) = 0 \iff x = 0$ for $x \in (-2M, 2M)^n$. Suppose that $P(x) = P(y)$ for $x, y \in [M]^n$. Thus,

$$P(x) - P(y) = P(x - y) = 0.$$

Since $x, y \in (-M, M)^n \subset (-2M, 2M)^n$, we have $x - y = 0$, so $x = y$ as we wanted to show.

To see that property (iii) holds, suppose that $P(x) + P(y) = P(2z)$, with $x, y, z \in [M]^n$. We can write

$$P(x) - P(z) = P(z) - P(y) \iff P(x) - 2P(z) + P(y) = P(x - 2z + y) = 0.$$

Note that $x - 2z + y \in (-2M, 2M)^n$, so $x - 2z + y$if and only if $x + y = 2z$, as we wanted.

Property (iv) holds due to the fact that each summand is strictly increasing with each of the coordinates x_i. So for $x \in [M]^n$,

$$\begin{aligned} P(x) \leq P((M, \dots, M)) &= \frac{1}{2M} \sum_{i=1}^{n} M(2M)^i \\ &= M \sum_{i=1}^{n-1} (2M)^i = M \frac{(2M)^n - 1}{2M - 1} \\ &\leq M \frac{(2M)^n}{M} = (2M)^n. \end{aligned}$$

□

Then, taking $(2M)^n \sim N$, we have $M \sim \frac{N^{1/n}}{2}$. Now, if we take $M = \lceil \frac{N^{1/n}}{2} \rceil$, it follows that all points given by $P(S(r_0))$ will belong to $[n] \subset \mathbb{N}$, so $P(r_0) \subset [n]$ and contains no 3-AP. Setting $n = \sqrt{\log N}$, we see that

$$\begin{aligned} |P(S(r_0))| = |S(r_0)| &\geq \frac{N^{1-2/n}}{n2^n} \\ &= Ne^{-\log n - n \log 2 - \frac{2}{n} \log N} \\ &= Ne^{-C\sqrt{\log n}}. \end{aligned}$$

as we wanted. □

Since no big improvements have been done in over 75 years, it is conjectured that Behrand's set is the largest possible set with no three-term progressions.

Conclusions

During this thesis, we have collected three proofs of Roth's Theorem, while developing some important notions around each area such as *Szemerédi's Regularity Lemma*. We have tried to explain the proofs in the best possible way, while correcting some errors, in particular, for the combinatorial proof. We will like to remark how this problem can be approached by very different mathematical fields such as graph theory and Fourier analysis, but keeping the same crucial point: the dichotomy between the existence of a 3-AP or an *energy* or density increment.

What we have shown in this thesis is just an introduction to an active field: as we have explained, improving the upper and lower bounds on $r_3([N])$, the largest subset of $[N]$ which contains no 3-APs, is still an open research problem.

One way to extend this work will be by including a fourth proof: the ergodic theoretical one, which appeared in 1977 when Fustenberg proved Szemerédi's Theorem by these techniques [9], and thus, Roth's Theorem.

Bibliography

[1] Michael Bateman and Nets Hawk Katz. New bounds on cap sets. *J. Amer. Math. Soc.*, 25(2):585–613, 2012.

[2] F. A. Behrend. On sets of integers which contain no three terms in arithmetical progression. *Proc. Nat. Acad. Sci. U.S.A.*, 32:331–332, 1946.

[3] J. Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.

[4] J. Bourgain. Roth's theorem on progressions revisited. *J. Anal. Math.*, 104:155–192, 2008.

[5] David Conlon. Extremal Graph Theory. Available at http://www.its.caltech.edu/~dconlon/Extremal-course.html.

[6] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, Berlin, fifth edition, 2018. Paperback edition of [MR3644391].

[7] Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017.

[8] Paul Erdös and Paul Turán. On Some Sequences of Integers. *J. London Math. Soc.*, 11(4):261–264, 1936.

[9] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.*, 34:275–291 (1979), 1978.

[10] W. T. Gowers. Techniques in combinatorics – lecture notes. Available at https://gowers.files.wordpress.com/2014/11/techniques2014notes.pdf.

[11] W. T. Gowers. The work of Endre Szermerédi. Available at https://abelprisen.no/sites/default/files/2021-04/Abel%20prize%202012%20Endre%20Szemeredi%20W%20T%20Gowers%20The%20work%20of%20Szemeredi%20eng_0.pdf.

[12] W. T. Gowers. A new proof of Szemerédi's theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.

[13] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer. *Ramsey theory*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, 2013. Paperback edition of the second (1990) edition [MR1044995].

[14] D. R. Heath-Brown. Integer sets containing no arithmetic progressions. *J. London Math. Soc. (2)*, 35(3):385–394, 1987.

[15] David Hilbert. Ueber die irreducibilität ganzer rationaler functionen mit ganzzahligen coefficienten. 1892(110):104–129, 1892.

[16] Roy Meshulam. On subsets of finite abelian groups with no 3-term arithmetic progressions. *J. Combin. Theory Ser. A*, 71(1):168–172, 1995.

[17] V. Rödl, B. Nagle, J. Skokan, M. Schacht, and Y. Kohayakawa. The hypergraph regularity method and its applications. *Proc. Natl. Acad. Sci. USA*, 102(23):8109–8113, 2005.

[18] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.

[19] I. Z. Ruzsa and E. Szemerédi. Triple systems with no six points carrying three triangles. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 939–945. North-Holland, Amsterdam-New York, 1978.

[20] Tom Sanders. On certain other sets of integers. *J. Anal. Math.*, 116:53–82, 2012.

[21] E. Szemerédi. On sets of integers containing no four elements in arithmetic progression. *Acta Math. Acad. Sci. Hungar.*, 20:89–104, 1969.

[22] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[23] E. Szemerédi. Integer sets containing no arithmetic progressions. *Acta Math. Hungar.*, 56(1-2):155–158, 1990.

[24] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[25] Yuffei Thao. Graph theory and additive combinatorics. Available at https://yufeizhao.com/gtac/gtac.pdf.

[26] Olof Sisask Thomas F. Bloom. Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions.

[27] B. L. Van der Waerden. Beweis einer Baudetschen Vermutung. 1927.

[28] Lluís Vena Cros. The Removal Lemma: algebraic versions and applications. Available at https://www.tesisenred.net/handle/10803/132098#page=20.

[29] J. Wolf. Finite field models in arithmetic combinatorics—ten years on. *Finite Fields Appl.*, 32:233–274, 2015.