CHAPTER 4

Roth’s Theorem on Arithmetic Progressions

This chapter is available free to all individuals,
on the understanding that it is not to be used for financial gain,
and may be downloaded and/or photocopied,
with or without permission from the author.
However, this document may not be kept on any information storage and retrieval system
without permission from the author,
unless such system is not accessible to any individuals other than its owners.

4.1. Introduction

A famous theorem of van der Waerden states that given any natural numbers \( \ell \) and \( r \), there exists \( N_0(\ell, r) \) such that for every natural number \( n > N_0(\ell, r) \), every partition of the set \( \{1, 2, \ldots, n\} \) into \( r \) subsets will yield a subset which contains \( \ell \) terms in arithmetic progression.

This result leads naturally to the following question. Suppose that \( A \) is a set of natural numbers. For every natural number \( n \in \mathbb{N} \), let

\[
A(n) = A(n, \mathcal{A}) = \sum_{a \in \mathcal{A}} 1
\]

and

\[
D(n) = D(n, \mathcal{A}) = \frac{A(n)}{n};
\]

in other words, \( A(n) \) and \( D(n) \) denote respectively the number and proportion of elements of the set \( \{1, 2, \ldots, n\} \) that are also in \( A \). Define the upper asymptotic density of the set \( \mathcal{A} \) by

\[
\overline{d} = \overline{d}(\mathcal{A}) = \limsup_{n \to \infty} D(n).
\]

Erdős and Turán conjectured that every set \( \mathcal{A} \) of natural numbers with positive upper asymptotic density contains arbitrarily long arithmetic progressions. This is equivalent to the statement that if there is a natural number \( \ell \) such that the set \( \mathcal{A} \) contains no arithmetic progression of \( \ell \) terms, then \( \overline{d}(\mathcal{A}) = 0 \).

The Hardy–Littlewood method can be adapted to establish the case \( \ell = 3 \) of this conjecture, as demonstrated by Roth in the 1950’s. The novelty of this approach is that the Hardy–Littlewood method is applied to study a sequence that is not explicitly given, such as \( k \)-powers of natural numbers or primes.

For every \( n \in \mathbb{N} \), let

\[
M(n) = \max\{|S| : S \subseteq \{1, 2, \ldots, n\}, S \text{ does not contain } 3 \text{ terms in arithmetic progression}\},
\]

where \( |S| \) denotes the number of elements of the set \( S \). In other words, \( M(n) \) denotes the largest number of elements which can be taken from the set \( \{1, 2, \ldots, n\} \) with no 3 of them in arithmetic progression. Also, for every \( n \in \mathbb{N} \), let

\[
\delta(n) = \frac{M(n)}{n}.
\]

**Theorem 4.1.** Suppose that \( n \in \mathbb{N} \) and \( n \geq 3 \). Then \( \delta(n) \ll (\log \log n)^{-1} \).

The Erdős–Turán conjecture is now known to be true for every positive integer \( \ell \), and is now universally known as Szemerédi’s theorem. Szemerédi’s proof is a tour de force in combinatorics, and does not use the Hardy–Littlewood technique.
Roth’s technique involves working with a set $\mathcal{M} \subseteq \{1, 2, \ldots, n\}$ that satisfies $|\mathcal{M}| = M(n)$ and does not contain 3 terms in arithmetic progression. We keep this set $\mathcal{M}$ fixed throughout our discussion, and apply the Hardy–Littlewood technique on this set. More precisely, consider the generating function

$$f(\alpha) = \sum_{x \in \mathcal{M}} e(\alpha x).$$

Then

$$\int_0^1 f^2(\alpha)f(-2\alpha) \, d\alpha = \int_0^1 \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \sum_{x_3 \in \mathcal{M}} e(\alpha(x_1 + x_2 - 2x_3)) \, d\alpha$$

$$= \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \sum_{x_3 \in \mathcal{M}} \int_0^1 e(\alpha(x_1 + x_2 - 2x_3)) \, d\alpha$$

$$= \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \sum_{x_3 \in \mathcal{M}} 1 = M(n),$$

since the only possible solutions of the equation

$$x_1 + x_2 = 2x_3, \quad x_1, x_2, x_3 \in \mathcal{M},$$

are the trivial solutions $x_1 = x_2 = x_3$.

The main idea of the proof of Theorem 4.1 is that if $M(n)$ were close to $n$, then the integral

$$\int_0^1 f^2(\alpha)f(-2\alpha) \, d\alpha$$

would be close to $M^2(n)$, thus contradicting (4.2).

### 4.2. A Major Arc Type Argument

The first step of the argument is to approximate the generating function (4.1). This can be achieved with a relatively small error if we make use of the disorderly arithmetical structure of the set $\mathcal{M}$. Sums of the form

$$\sum_{x=1}^n e(\alpha x)$$

$$\sum_{x \in \mathcal{A}}$$

tend to have large modulus near rational points $a/q$ if the elements of $\mathcal{A}$ are well distributed in residue classes modulo $q$.

More precisely, suppose that the natural number $m < n$. Write

$$v(\alpha) = \delta(m) \sum_{x=1}^n e(\alpha x) \quad \text{and} \quad E(\alpha) = v(\alpha) - f(\alpha).$$

If we let $\chi_\mathcal{M}$ denote the characteristic function of the set $\mathcal{M}$, then

$$f(\alpha) = \sum_x \chi_\mathcal{M}(x)e(\alpha x).$$

Hence

$$E(\alpha) = \sum_{x=1}^n c(x)e(\alpha x),$$

where

$$c(x) = \delta(m) - \chi_\mathcal{M}(x).$$

**Theorem 4.2.** Suppose that

$$g(\alpha) = \sum_{z=0}^{m-1} e(\alpha z).$$

Suppose further that the natural number $q < n/m$. Then

$$g(\alpha q)E(\alpha) = \sum_{h=1}^{n-mq} \sigma(h)e(\alpha(h + mq - q)) + R(\alpha),$$

where, for every $h = 1, \ldots, n - mq$,

$$\sigma(h) = \sum_{z=0}^{m-1} c(h + xq) \geq 0,$$

and where

$$|R(\alpha)| < 2m^2q.$$

**Proof.** It is easy to see that

$$g(\alpha q)E(\alpha) = \sum_{z=0}^{m-1} \sum_{x=1}^{n} c(x)e(\alpha(x + qz)).$$

Note that $x + qz \in [1, n + mq - q]$. Writing $x + qz = h + mq - q$, we have

$$g(\alpha q)E(\alpha) = \sum_{h=1}^{n} e(\alpha(h + mq - q)) \sum_{z=0}^{m-1} c(h + mq - q - qz)$$

$$= \sum_{h=1}^{n-mq} e(\alpha(h + mq - q)) \sum_{z=0}^{m-1} c(h + q(m - 1 - z)) + R(\alpha),$$

where

$$R(\alpha) = \sum_{h=1+q-mq}^{0} e(\alpha(h + mq - q)) \sum_{z=0}^{m-1} c(h + q(m - 1 - z))$$

$$+ \sum_{h=n-mq+1}^{n} e(\alpha(h + mq - q)) \sum_{z=0}^{m-1} c(h + q(m - 1 - z)).$$

Now the inner sums in (4.8) clearly do not exceed $m$ in absolute value, so

$$|R(\alpha)| \leq m(mq - q + mq) < 2m^2q.$$

On the other hand, if $1 \leq h \leq n - mq$, then for every integer $z$ in the range $0 \leq z \leq m - 1$, the inequality $1 \leq h + q(m - 1 - z) \leq n$ is always satisfied. It follows from (4.7) that

$$g(\alpha q)E(\alpha) = \sum_{h=1}^{n-mq} \left( \sum_{z=0}^{m-1} c(h + q(m - 1 - z)) \right) e(\alpha(h + mq - q)) + R(\alpha),$$

where

$$\sum_{z=0}^{m-1} c(h + q(m - 1 - z)) = \sum_{z=0}^{m-1} c(h + xq) = \sigma(h).$$

The inequalities (4.5) and (4.6) now follow from (4.9)–(4.11). Note next that

$$\sigma(h) = \sum_{x=0}^{m-1} (\delta(m) - \chi_M(h + xq)) = M(m) - \sum_{x=0}^{m-1} \chi_M(h + xq).$$

The sum

$$r = \sum_{x=0}^{m-1} \chi_M(h + xq)$$

is the number of elements of $M$ in the arithmetic progression $h, h + q, \ldots, h + (m - 1)q$. 

4.2. A MAJOR ARC TYPE ARGUMENT
Let these elements be \( h + x_1q, \ldots, h + x_rq \). Now no three of these are in arithmetic progression. Hence no three of \( x_1, \ldots, x_r \) are in arithmetic progression, whence no three of \( 1 + x_1, \ldots, 1 + x_r \) are in arithmetic progression. Also \( 1 + x_j \leq m \) for every \( j = 1, \ldots, r \). It follows that we must have \( r \leq M(m) \), whence \( \sigma(h) \geq 0 \), in view of (4.12). \( \square \)

**Theorem 4.3.** Suppose that \( 2m^2 < n \). Then for every real number \( \alpha \), we have

\[
|E(\alpha)| < 2n(\delta(m) - \delta(n)) + 16m^2.
\]

**Proof.** By Dirichlet’s theorem, there exist integers \( a \) and \( q \) satisfying \( (a, q) = 1 \) and \( 1 \leq q \leq 2m \) such that

\[
\left| \alpha - \frac{a}{q} \right| \leq \frac{1}{2mq}.
\]

Then

\[
g(\alpha q) = g(\alpha q - a) = g(\beta),
\]

where

\[
|\beta| = |\alpha q - a| \leq \frac{1}{2m}.
\]

It follows from (4.4) and (4.13) that

\[
|g(\alpha q)| = |g(\beta)| = \left| \frac{\sin \pi n \beta}{\sin \pi \beta} \right| \geq \frac{2m}{\pi}.
\]

Note next that \( q \leq 2m < n/m \). In view of Theorem 4.2, we have

\[
\frac{m}{2} |E(\alpha)| \leq \frac{2m}{\pi} |E(\alpha)| \leq |g(\alpha q)E(\alpha)| = \sum_{h=1}^{n-mq} \sigma(h) + 2m^2q
\]

\[
= g(0)E(0) - R(0) + 2m^2q < mE(0) + 4m^2q \leq mE(0) + 8m^3.
\]

On the other hand,

\[
E(0) = \sum_{x=1}^{n} (\delta(m) - \chi_\mathcal{M}(x)) = n\delta(m) - M(n) = n(\delta(m) - \delta(n)).
\]

The result follows on combining (4.14) and (4.15). \( \square \)

### 4.3. Completion of the Proof

Write

\[
I = \int_{0}^{1} f^2(\alpha) \nu(-2\alpha) \, d\alpha.
\]

In view of (4.1) and (4.3), we have

\[
I = \int_{0}^{1} \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} e(\alpha(x_1 + x_2 - 2y)) \, d\alpha
\]

\[
= \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \sum_{y=1}^{n} \delta(m) \int_{0}^{1} e(\alpha(x_1 + x_2 - 2y)) \, d\alpha
\]

\[
= \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \sum_{y=1}^{n} \delta(m) + \sum_{x_1 \in \mathcal{M}} \sum_{x_2 \in \mathcal{M}} \delta(m).
\]

Let \( M_1 \) and \( M_2 \), where \( M_1 + M_2 = M(n) \), denote respectively the number of odd and even elements of \( \mathcal{M} \). Then

\[
I = \delta(m)(M_1^2 + M_2^2) \geq \frac{1}{2} \delta(m)(M_1 + M_2)^2 = \frac{1}{2} \delta(m)M^2(n).
\]
On the other hand, it follows from (4.2), (4.3) and (4.16) that

\[ |M(n) - I| = \left| \int_0^1 f^2(\alpha) (f(-2\alpha) - v(-2\alpha)) \, d\alpha \right| \leq \left( \max_{\alpha} |E(\alpha)| \right) \int_0^1 |f(\alpha)|^2 \, d\alpha. \]

Clearly

\[ \int_0^1 |f(\alpha)|^2 \, d\alpha = \int_0^1 f(\alpha) f(-\alpha) \, d\alpha = M(n). \]

It follows from Theorem 4.3 that if \(2m^2 < n\), then

(4.18) \[ |M(n) - I| \leq (2n(\delta(m) - \delta(n)) + 16m^2)M(n). \]

Combining (4.17) and (4.18), we have

\[ \frac{1}{2} nM(n(\delta(m) - \delta(n)) = \frac{1}{2} \delta(m)M^2(n) \leq I \leq M(n) + (2n(\delta(m) - \delta(n)) + 16m^2)M(n), \]

so that

(4.19) \[ \delta(m)\delta(n) \leq 2n^{-1} + 4(\delta(m) - \delta(n)) + 32m^2n^{-1} \leq 4(\delta(m) - \delta(n)) + 34m^2n^{-1}, \]

so long as \(2m^2 < n\).

**Theorem 4.4.** The limit

(4.20) \[ \tau = \lim_{n \to \infty} \delta(n) \]

exists. Furthermore, \(\delta(n_2) \leq 2\delta(n_1)\) for all natural numbers \(n_1 \leq n_2\).

**Proof.** It is trivial that \(M(m + n) \leq M(m) + M(n)\). Hence for \(n_2 \geq n_1\),

(4.21) \[ M(n_2) = M \left( n_1 \left[ \frac{n_2}{n_1} \right] + \left[ n_2 - n_1 \left[ \frac{n_2}{n_1} \right] \right] \right) \leq \left[ \frac{n_2}{n_1} \right] M(n_1) + M \left( n_2 - n_1 \left[ \frac{n_2}{n_1} \right] \right). \]

Clearly

\[ M(n_2) \leq \frac{n_2}{n_1} M(n_1) + n_1, \]

so that

\[ \delta(n_2) \leq \delta(n_1) + \frac{n_1}{n_2}. \]

Hence

\[ \lim_{n_2 \to \infty} \delta(n_2) \leq \delta(n_1) \quad \text{and} \quad \lim_{n_2 \to \infty} \delta(n_2) \leq \lim_{n_1 \to \infty} \delta(n_1), \]

so the limit (4.20) exists. Also, it follows from (4.21) that

\[ M(n_2) \leq \frac{n_2}{n_1} M(n_1) + M(n_1) \leq 2 \frac{n_2}{n_1} M(n_1). \]

The second assertion follows immediately. \(\square\)

**Remark.** Letting \(n \to \infty\), the inequality (4.19) becomes

\[ \delta(m)\tau \leq 4(\delta(m) - \tau). \]

Letting \(m \to \infty\), we conclude that \(\tau^2 \leq 0\), so that \(\tau = 0\). This is a weaker form of Theorem 4.1.

To complete the proof of Theorem 4.1, we write

\[ \lambda(x) = \delta(2^x). \]

In view of Theorem 4.4, it suffices to prove that \(\lambda(x) \ll x^{-1}\). By (4.19), we have

\[ \lambda(y)\lambda(y + 1) \leq 4(\lambda(y) - \lambda(y + 1)) + 34 \cdot 2^{-3y}, \]

so that

\[ 1 \leq \frac{4(\lambda(y) - \lambda(y + 1)) + 34 \cdot 2^{-3y}}{\lambda(y)\lambda(y + 1)}. \]
Summing this over $y = x, x + 1, \ldots, 2x - 1$, we have

$$x \leq \sum_{y=x}^{2x-1} \frac{4(\lambda(y) - \lambda(y + 1))}{\lambda(y)\lambda(y + 1)} + \sum_{y=x}^{2x-1} \frac{34 \cdot 2^{-3y}}{\lambda(y)\lambda(y + 1)}$$

$$= 4 \sum_{y=x}^{2x-1} \left( \frac{1}{\lambda(y + 1)} - \frac{1}{\lambda(y)} \right) + \sum_{y=x}^{2x-1} \frac{34 \cdot 2^{-3y}}{\lambda(y)\lambda(y + 1)}$$

$$\leq \frac{4}{\lambda(2x)} + \frac{200x2^{-3x}}{\lambda^2(2x)},$$

in view of Theorem 4.4. When $\lambda(2x) > 1/x$, then

$$\frac{200x2^{-3x}}{\lambda^2(2x)} < \frac{x}{2}$$

for all sufficiently large $x$, so that $\lambda(2x) < 8/x$. 