Roth's theorem on 3-arithmetic progressions in the integers

Anne De Roton

Roth's theorem on 3-arithmetic progressions in the integers

CIMPA Research school Shillong 2013

Anne de Roton
anne.de-roton@univ-lorraine.fr

Institut Elie Cartan
Université de Lorraine
France

Introduction.

0.1. **Statement of the result.** In 1953, K. Roth [13] proved that a set of positive integers with positive upper density must contain a non trivial arithmetic progression of length 3. He more precisely obtained the following quantitative result.

Theorem (Roth (1953))**.** *There exist a positive integer N_0 and a positive constant c such that for any $N \geq N_0$, any set A of positive integers less than N with density larger than $c/\log\log N$ does contain a non trivial 3-arithmetic progression, i.e. there exist $n, r \in \mathbb{N}$ (thus $r \neq 0$) such that the 3 integers n, $n+r$ and $n+2r$ are all elements of A.*

Remark 0.1. Of course any non empty set does contain a trivial 3-arithmetic progression (meaning that $r = 0$). Thus we have to exclude these trivial progressions to get a meaningful result.

The first part of this notes will be devoted to the proof of this theorem whereas a second part will give some short survey on the analogue of Roth's theorem in some infinite subsets of integers of zero density such as the subset of prime numbers.

We tried to give the reader all the details needed in the first part so that a master student can read Roth's theorem proof easily. In the second part, some proofs are only sketched and we rather tried to give an idea of the issues specific to zero density subsets than to explain precisely how all the arguments work.

These notes are based on some previous notes by K. Soundararajan available at http://math.stanford.edu/ ksound/Notes.pdf and by T. Gowers [4] and on Tao and Vu's book [19]. The author is also indebted to the students and colleagues who attended her course in Shillong and especially to Jan-Christoph Schlage-Puchta and Olöf Sisask whose comments and questions were very useful to clarify some of the arguments.
This course was given in the CIMPA research school in Shillong organised in november 2013 by Gautami Bhowmik and Himadri Mukerjee. The author warmly thanks the organisers for this opportunity.
The author is partly supported by the ANR grant ANR "Caesar" 12-BS01-0011.

0.2. **Tools and notations.** Given a positive integer N, we write $\mathbb{Z}_N$ to denote $\mathbb{Z}/N\mathbb{Z}$ and for simplicity, $[1, N]$ for the set of positive integers less than N. Given a subset of positive integers A, we write A_N for the set $A \cap [1, N]$ and if $A \subset \mathbb{Z}$ is finite, we write $|A|$ for the number of elements in A and $\mathbf{1}_A$ for the indicator function of A. Subsets of integers less than N will often be identified with the corresponding subsets of $\mathbb{Z}_N$.
If $f : \mathbb{Z}_N \to \mathbb{C}$ or $f : [1, N] \to \mathbb{C}$, we define the mean value

$$\mathbb{E}f = \frac{1}{N}\sum_{n\in[1,N]} f(n) \quad \text{or} \quad \mathbb{E}f = \frac{1}{N}\sum_{n\in\mathbb{Z}_N} f(n).$$

We shall see that in the proof of Roth's theorem (in integers and in subsets of integers) that Fourier analysis plays a great role. We recall here some definition and introduce some notation related to this field.

Definition 0.1. Given $f, g : \mathbb{Z}_N \to \mathbb{C}$ two functions, we define

- the Fourier transform $\hat{f} : \mathbb{Z}_N \to \mathbb{C}$ of f by
$$\hat{f}(k) = \frac{1}{N} \sum_{n \in \mathbb{Z}_N} f(n) e\left(-\frac{kn}{N}\right) \quad \text{where} \quad e(\theta) = e^{2i\pi\theta};$$
- the convolution of f and g by
$$(f * g)(n) = \mathbb{E}_{m \in \mathbb{Z}_N} f(n-m) g(m) = \mathbb{E}_{m \in \mathbb{Z}_N} f(m) g(n-m).$$

We recall here for $f, g : \mathbb{Z}_N \to \mathbb{C}$,

- the Fourier inversion formula:
$$f(n) = \sum_{k \in \mathbb{Z}_N} \widehat{f}(k) e\left(\frac{kn}{N}\right).$$
- the following property of the convolution: $\widehat{f * g} = \hat{f}\hat{g}$.

Definition 0.2. We also define, for $p \geq 1$ and $f : \mathbb{Z}_N \to \mathbb{C}$, the L^p norm of f by
$$\|f\|_{L^p} = \left(\frac{1}{N} \sum_{n \in \mathbb{Z}_N} |f(n)|^p\right)^{1/p}.$$

For $\hat{f}$ we shall rather use the ℓ_p-norm
$$\|\hat{f}\|_{\ell_p} = \left(\sum_{n \in \mathbb{Z}_N} |f(n)|^p\right)^{1/p}.$$

We shall heavily use Parseval's identities:
$$\sum_{k \in \mathbb{Z}_N} \hat{f}(k)\overline{\hat{g}(k)} = \mathbb{E}_{n \in \mathbb{Z}_N} f(n)\overline{g(n)} \quad \text{and} \quad \sum_{r \in \mathbb{Z}_N} \left|\widehat{f}(k)\right|^2 = \mathbb{E}_{n \in \mathbb{Z}_N} |f(n)|^2,$$
and for $p \geq 1$ and p' such that $\frac{1}{p} + \frac{1}{p'} = 1$, Hölder's inequality:
$$\mathbb{E}_{n \in \mathbb{Z}_N} f(n)\overline{g(n)} \leq \|f\|_{L^p} \|g\|_{L^{p'}} \quad \text{and} \quad \sum_{n \in \mathbb{Z}_N} \hat{f}(n)\overline{\hat{g}(n)} \leq \|\hat{f}\|_{\ell_p} \|\hat{g}\|_{\ell_{p'}}$$
and Cauchy's inequality
$$\mathbb{E}_{n \in \mathbb{Z}_N} f(n)\overline{g(n)} \leq \|f\|_{L^2} \|g\|_{L^2} \quad \text{and} \quad \sum_{n \in \mathbb{Z}_N} \hat{f}(n)\overline{\hat{g}(n)} \leq \|\hat{f}\|_{\ell_2} \|\hat{g}\|_{\ell_2}.$$

We shall also need a function counting weighted arithmetic progressions.
Given $f_1, f_2, f_3 : \mathbb{Z}_N \to \mathbb{C}$, we define
$$\Lambda_3(f_1, f_2, f_3) = \frac{1}{N^2} \sum_{n,r \in \mathbb{Z}_N} f_1(n) f_2(n+r) f_3(n+2r).$$

Remark 0.2. If $\mathbf{1}_A$ is the indicator function of a set $A \subset \mathbb{Z}_N$, then $N^2 \Lambda_3(1_A, 1_A, 1_A)$ counts the number of 3-arithmetic progressions in A (as a subset of $\mathbb{Z}_N$).

One key formula in Roth's argument is the following lemma.

Lemma 0.3. *Given $f_1, f_2, f_3 : \mathbb{Z}_N \to \mathbb{C}$, we have*

$$\Lambda_3(f_1, f_2, f_3) = \sum_{k \in \mathbb{Z}_N} \widehat{f_1}(k)\widehat{f_2}(-2k)\widehat{f_3}(k). \tag{0.1}$$

Proof. We use the definition of the Fourier transform and get

$$\begin{aligned}
\sum_{k \in \mathbb{Z}_N} \widehat{f_1}(k)\widehat{f_2}(-2k)\widehat{f_3}(k) &= \left(\frac{1}{N}\right)^3 \sum_{k \in \mathbb{Z}_N} \sum_{n_1,n_2,n_3 \in \mathbb{Z}_N} f_1(n_1)f_2(n_2)f_3(n_3) e\left(-\frac{k(n_1 - 2n_2 + n_3)}{N}\right) \\
&= \left(\frac{1}{N}\right)^2 \sum_{n_1,n_2,n_3 \in \mathbb{Z}_N} f_1(n_1)f_2(n_2)f_3(n_3) \frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(-\frac{k(n_1 - 2n_2 + n_3)}{N}\right) \\
&= \left(\frac{1}{N}\right)^2 \sum_{\substack{n_1,n_2,n_3 \in \mathbb{Z}_N \\ n_1+n_3=2n_2}} f_1(n_1)f_2(n_2)f_3(n_3)
\end{aligned}$$

where in the last equality we used

$$\frac{1}{N} \sum_{k \in \mathbb{Z}_N} e\left(\frac{nk}{N}\right) = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, a change of variables gives

$$\left(\frac{1}{N}\right)^2 \sum_{\substack{n_1,n_2,n_3 \in \mathbb{Z}_N \\ n_1+n_3=2n_2}} f_1(n_1)f_2(n_2)f_3(n_3) = \Lambda_3(f_1, f_2, f_3).$$

□

Notation 0.4. Given a real number x, we shall write $\|x\|_{\mathbb{R}/\mathbb{Z}}$ for the distance between x and the nearest integer.

1. Proof of Roth's theorem

In this section, we give a proof of Roth's theorem that we recall here.

Theorem 1.1 (Roth (1953))**.** *There exist a positive integer N_0 and a positive constant c such that for any $N \geq N_0$, any set A of positive integers less than N with density larger than $c/\log\log N$ does contain a non trivial 3-arithmetic progression, i.e. there exist $n, r \in \mathbb{N}$ (thus $r \neq 0$) such that the 3 integers n, $n + r$ and $n + 2r$ are all elements of A.*

1.1. Outline of the proof.

First we define arithmetic progressions.

Definition 1.1. A subset P of integers less than N of the form

$$P = \{a + nb, n \in \{0, 1, \cdots, N_1 - 1\}\} \text{ with } a, b \in \mathbb{N}$$

is called an arithmetic progression of length N_1.

We first explain roughly the main ideas of the proof.

(1) One would expect, for a "random" set $A \subset [1, N]$ of cardinality δN, around $\delta^3 N^2$ 3-arithmetic progressions, whereas we only have δN trivial 3-arithmetic progressions in A (much less than $\delta^3 N^2$ if N is much larger than $1/\delta^2$).
(2) A non "random" set has to be structured in some sense. More precisely, Roth proves that a set with much less 3-arithmetic progressions than expected has to concentrate on some large arithmetic progression P in A. To prove this part, Roth uses the discrete Fourier transform and formula (0.1).
(3) If P is an arithmetic progression in A of length N_1 on which A concentrates, the subset $A \cap P$ is isomorphic to some set $A_1 \subset [1, N_1]$ of cardinality $\delta_1 N_1$ with $\delta_1 > \delta$. Furthermore, if A is progression free (with no non trivial 3-arithmetic progression), then so does A_1.
(4) An iteration of the argument finishes the proof. If A is a progression free subset of $[1, N]$ of cardinality δN and if δ and N are large enough, then Roth constructs a sequence of subsets $A_k \subset [1, N_k]$ of integers with cardinality $\delta_k N_k$ such that the sequence $(N_k)_k$ does not decrease too fast and the sequence $(\delta_k)_k$ increases much enough so that $\delta_k > 1$ and we still have $N_k > 1/\delta_k^2$. We thus get an absurdity.

Remark 1.2. The third point is essential in the argument. Actually, all the argument would work for any equation of the form $ax + by = cz$ with $a + b = c$. These equations are called invariant equations (invariant by translation and multiplication) according to Rusza's terminology [14] and [15].

The main steps in the proof will be the following:
Let N be a sufficient large prime integer and A be a subset of $[1, N]$.
(1) If A is a progression free subset of $[1, N]$ of density $\delta > 0$, then there exists $k \neq 0$ such that $\left|\widehat{\mathbf{1}_A}(k)\right|$ is big.
(2) If $\left|\widehat{\mathbf{1}_A}(k)\right|$ is big for some $k \neq 0$ then A concentrates on a large arithmetic progression.
(3) By iteration, if A is large enough, we will eventually reach an arithmetic progression P with relative density of A in P strictly larger than 1, a contradiction.

1.2. **Begining of the proof.** Before entering the core of the proof, we need to take the right setting so that everything works.
Let A be a set of positive integers of upper density strictly larger than some positive δ.
It will be very convenient to consider A_N as a subset of $\mathbb{Z}_N$ (N will be chosen so that $|A_N| \geq \delta N$). This procedure may not be harmless since $N^2 \Lambda_3(\mathbf{1}_{A_N}, \mathbf{1}_{A_N}, \mathbf{1}_{A_N})$ will count the number of 3-arithmetic progressions in A_N modulo N. This way, we may add some arithmetic progressions. For instance, $(3, 0, 4)$ is a 3-arithmetic progression in $\mathbb{Z}_7$ but not in $\mathbb{N}$. To prevent us from counting 3-arithmetic progressions in $\mathbb{Z}_N$ which are not genuine 3-arithmetic progressions in $\mathbb{N}$, we have two choices:
(a) either we choose $f_1 = 1_{A_N}$, $f_2 = f_3 = 1_{B_N}$ with $B_N = \{x \in A : N/3 < x < 2N/3\}$, which is close to the original argument of Roth in [13];
(b) or we choose $f_2 = 1_{A_N}$ and $f_1 = f_3 = 1_{B_N}$ with B_N the set of even or odd numbers in A_N whichever is larger, which is the argument used in Gowers' proof of Roth's theorem.

In these two cases, any 3-arithmetic progression in $A_N \times B_N \times B_N$ in the first case and in $B_N \times A_N \times B_N$ in the second one as subsets of $\mathbb{Z}_N$ is a genuine 3-arithmetic progression in $[1, N]$ and $|B_N|$ is not too small since:

(a) in the first case, if $x \in A_N$ and $y, z \in B_N$, then $2y - N < N/3 < x + z$ and $2y + N > 7N/3 > 5N/3 > x + z$. Furthermore, in this case, we have either that $|B_N| \geq |A_N|/4$ or one of the sets $A_N \cap [1, N/3]$ and $A_N \cap [2N/3, N]$ has cardinality at least $3|A|/8$. This last case implies that the relative density of A on this set is at least $\frac{9}{8}|A_N|/N$, thus that we found an arithmetic progression of length $N_1 = N/3$ on which the relative density of A is $9/8$ times the density of A in $[1, N]$.

(b) in the second case by parity, $y+z = 2x+N$ and $y+z = 2x-N$ are forbidden provided N is odd. In this case we have $|B_N| \geq |A_N|/2$.

We choose to work with the second choice.

In the proof, we shall work modulo N and it will be very convenient to take $N > 2$ prime (thus odd). The following lemma proves that we can take N prime without loosing any factor in the density

Lemma 1.3. *Let A be a subset of $\mathbb{N}$ of upper density α. Then for all $\delta < \alpha$ we have $|A_p| \geq \delta p$ for infinitely many prime numbers p.*

Remark 1.4. One may think of using Bertrand's postulate which states that there is a prime number between n and $2n$ for any integer $n \geq 2$ but this would lead to a loss of $1/2$ in the density of A. If this step is only used once this is harmless but in the proof of Roth's theorem, we have to use this step at each iteration and we cannot afford such a loss.

Proof. Let ε be some positive real number such that $\varepsilon < \min(\alpha, 1-\alpha)$. By definition of the upper density, there exists an infinite increasing sequence $(N_k)_k$ of integers (thus tending to infinity) such that

$$|A_{N_k}| \geq (\alpha - \varepsilon/2)N_k \quad \text{for every } k \in \mathbb{N}.$$

The sequence $(|A_n|)_n$ is increasing and satisfies $|A_n| \leq |A_{n+k}| \leq |A_n| + k$ thus for any $n \in \left[\frac{1-\alpha+\varepsilon/2}{1-\alpha+\varepsilon}N_k, \frac{\alpha-\varepsilon/2}{\alpha-\varepsilon}N_k\right]$, we have $|A_n| \geq (\alpha - \varepsilon)n$.
Now if $\pi(x)$ is the number of the prime numbers less than x, the prime number theorem gives:

$$\begin{aligned}\left|\left\{p \in \left[\frac{1-\alpha+\varepsilon/2}{1-\alpha+\varepsilon}N_k, \frac{\alpha-\varepsilon/2}{\alpha-\varepsilon}N_k\right]\right\}\right| &= \pi\left(\frac{\alpha-\varepsilon/2}{\alpha-\varepsilon}N_k\right) - \pi\left(\frac{1-\alpha+\varepsilon/2}{1-\alpha+\varepsilon}N_k\right) \\ &\sim \left(\frac{\alpha-\varepsilon/2}{\alpha-\varepsilon} - \frac{1-\alpha+\varepsilon/2}{1-\alpha+\varepsilon}\right)\frac{N_k}{\log N_k} \\ &\sim \frac{\varepsilon/2}{(\alpha-\varepsilon)(1-\alpha+\varepsilon)}\frac{N_k}{\log N_k}.\end{aligned}$$

This last quantity is clearly larger than 1 if k, thus N_k, is large enough thus there are infinitely many prime numbers p such that $|A_p| \geq (\alpha - \varepsilon)p$. Taking $\varepsilon = \alpha - \delta$ yields the Lemma. □

We are now ready to begin the proof of Roth's Theorem.
Let A be a progression free subset of $\mathbb{N}$ of density strictly larger than $\delta > 0$. For infinitely many prime numbers N we have $|A_N| \geq \delta N$. For such a prime number N, we choose for B_N the set of odd or even integers in A_N whichever is larger and identify the sets with their corresponding sets in $\mathbb{Z}_N$. For simplicity, we write $f_1 = f_3 = 1_{B_N}$ and $f_2 = 1_{A_N}$. We have

$$\mathbb{E}_{n\in\mathbb{Z}_N} f_1(n) \geq \delta/2, \quad \mathbb{E}_{n\in\mathbb{Z}_N} f_2(n) \geq \delta \quad \text{and} \quad \#\{3AP \text{ in } A_N \subset [1,N]\} \geq N^2\Lambda_3(f_1,f_2,f_1).$$

Furthermore, A being progression-free, we have

$$\frac{|B_N|}{N^2} = \Lambda_3(f_1,f_2,f_1) = \sum_{k\in\mathbb{Z}_N} \widehat{f_1}(k)^2\widehat{f_2}(-2k) = \widehat{f_1}(0)^2\widehat{f_2}(0) + \sum_{k\neq 0} \widehat{f_1}(k)^2\widehat{f_2}(-2k).$$

Since $\widehat{f_1}(0)^2\widehat{f_2}(0) = |B_N|^2|A_N|/N^3$ is much greater than $|B_N|/N^2$ when $N \geq 4/\delta^2$, it means that the function f_2 must have large non-zero Fourier coefficients. The purpose of the next section is to make this statement more precise.

Remark 1.5. Note that $N^2\widehat{f_1}(0)^2\widehat{f_2}(0)$ measures the expected number of 3-arithmetic progressions in $B_N\times A_N\times B_N \subset \mathbb{Z}_N$ since we have N^2 triples (x,y,z) in 3-arithmetic progression modulo N and the probability that $(x,y,z) \in B_N \times A_N \times B_N$ is $|A_N||B_N|^2/N^3$.

1.3. **Few 3APs implies big Fourier coefficient.** The following proposition states that a lack of 3-arithmetic progressions in A leads to the existence of a large Fourier coefficient of f_2.

Proposition 1.6. *Let N be a large prime number, α be a positive real number and $f_1, f_2 : \mathbb{Z}_N \to [0,1]$ be some functions satisfying $\mathbb{E}_{n\in\mathbb{Z}_N} f_2(n) = \alpha$, $\mathbb{E}_{n\in\mathbb{Z}_N} f_1(n) \geq \alpha/2$ and $\|f_1\|_2 \leq \alpha$. Then either $\Lambda_3(f_1,f_2,f_1) > \alpha^3/8$ or there exist $k \in \mathbb{Z}_N \setminus \{0\}$ such that $|\hat{f}_2(k)| \geq \alpha^2/8$.*

Proof. Suppose f_1 and f_2 satisfy the hypotheses and $\Lambda_3(f_1,f_2,f_1) \leq \alpha^3/8$. Then

$$\frac{\alpha^3}{8} \geq \Lambda_3(f_1,f_2,f_1) = \frac{|A_N||B_N|^2}{N^3} + \sum_{k\neq 0} \widehat{f_1}(k)^2\widehat{f_2}(-2k).$$

Since $|A_N||B_N|^2 \geq \alpha^3/4N^3$, this gives

$$\begin{aligned}\frac{\alpha^3}{8} &\leq \sup_{k\neq 0} |\widehat{f_2}(-2k)| \sum_{k\in\mathbb{Z}_N} |\widehat{f_1}(k)|^2 \\ &= \sup_{k\neq 0} |\widehat{f_2}(k)| \mathbb{E}_{n\in\mathbb{Z}_N} |f_1(n)|^2 \\ &\leq \alpha \sup_{k\neq 0} |\widehat{f_2}(k)|.\end{aligned}$$

This proves the announced result. □

1.4. **Big Fourier coefficient implies density increment.**

Lemma 1.7. *Let x be a real number. Then $|e(x) - 1| \leq 2\pi\|x\|_{\mathbb{R}/\mathbb{Z}}$ where $\|x\|_{\mathbb{R}/\mathbb{Z}}$ denotes the distance between x and the nearest integer.*

Proof. The function $(x \mapsto e(x))$ is 1-periodic so this is enough to prove the result for $x \in [-1/2, 1/2]$. We use

$$|e(x) - 1| = \left|e^{2i\pi x} - e^{2i\pi 0}\right| = 2|\sin(\pi x)| \leq 2\pi|x| = 2\pi\|x\|_{\mathbb{R}/\mathbb{Z}}.$$

□

Proposition 1.8. *Let σ be a positive real number, $N \geq 16^2\pi^2/\sigma^2$ be a large prime number and $g : \mathbb{Z}_N \to [-1, 1]$ be a function satisfying $\sum_{n\in\mathbb{Z}_N} g(n) = 0$. Assume that there exists $k \in \mathbb{Z}_N$ such that $|\widehat{g}(k)| \geq \sigma$. Then there exists an arithmetic progression P of length at least $\frac{\sigma\sqrt{N}}{8\pi}$, such that g has mean value at least $\sigma/8$ on P.*

Proof. Assume that for some $k \in \mathbb{Z}_N$,

$$|\widehat{g}(k)| := \left|\frac{1}{N}\sum_{n=1}^{N} g(n)e\left(\frac{nk}{N}\right)\right| \geq \sigma > 0.$$

Note that $\widehat{g}(0) = 0$, thus $k \neq 0$. By Dirichlet's theorem there exist coprime integers b and h such that $1 \leq b \leq h \leq \sqrt{N}$ and that

$$\left|\frac{k}{N} - \frac{b}{h}\right| \leq \frac{1}{h\sqrt{N}} \quad \text{thus} \quad \left\|\frac{kh}{N}\right\|_{\mathbb{R}/\mathbb{Z}} \leq \frac{1}{\sqrt{N}}.$$

We divide $[1, N]$ in h congruence classes modulo h

$$\mathcal{C}(a) = \{a, a + h, a + 2h, \cdots\} \cap [1, N] = \{a, a + h, a + 2h, \cdots, a + (N_a - 1)h\}.$$

We then choose a positive integer M and divide $[0, N_a - 1] \subset \mathbb{R}$ in M intervals I_m of same length and define

$$J_m(a) = \{a + jh,\ j \in I_m\} = \{a + j_m h, a + (j_m + 1)h, \cdots\}.$$

We therefore have

$$\{1, \cdots, N\} = \cup_{a\in\{1,\cdots,h\}}\mathcal{C}(a) = \cup_{a=1}^{h}\cup_{m=1}^{M} J_m(a).$$

each class $\mathcal{C}(a)$ has size at least $\frac{N}{h} - 1$ and each $J_m(a)$ has size $L_m(a) = \lfloor\frac{N_a}{M}\rfloor \in \left[\frac{N}{hM} - 2, \frac{N}{hM}\right]$. Furthermore, we have

$$\sigma \leq |\widehat{g}(k)| = \left|\frac{1}{N}\sum_{a=1}^{h}\sum_{m=1}^{M}\sum_{n\in J_m(a)} g(n)e\left(\frac{nk}{N}\right)\right|$$

Now, we know that $\frac{k}{N} = \frac{b}{h} + \theta$ with $|\theta| \leq \frac{1}{h\sqrt{N}}$ and for $n \in J_m(a)$, we have $n = a + h(j_m + l)$ with $0 \leq l \leq L_m(a) - 1$ thus by 1-periodicity of $(x \mapsto e(x))$, we have

$$\begin{aligned} e\left(\frac{nk}{N}\right) &= e\left(\frac{nb}{h}\right) e(n\theta) = e\left(\frac{ab}{h}\right) e((a + h(j_m + l))\theta) \\ &= e\left(\frac{ab}{h}\right) \left[e((a + hj_m)\theta) + e((a + hj_m)\theta)(e(hl\theta) - 1)\right]. \end{aligned}$$

Using Lemma 1.7, we get for $l \leq L_m(a) - 1$ and $|\theta| \leq \frac{1}{h\sqrt{N}}$,

$$\begin{aligned} |e(hl\theta) - 1| &\leq 2\pi \|hl\theta\|_{\mathbb{R}/\mathbb{Z}} \\ &\leq \frac{2\pi L_m(a)}{\sqrt{N}}. \end{aligned}$$

Combining this and using that $|g(n)| \leq 1$ and $|e(x)| = 1$, we have that

$$\begin{aligned} \sigma &\leq \left| \frac{1}{N} \sum_{a=1}^{h} e\left(\frac{ab}{h}\right) \sum_{m=1}^{M} e((a + hj_m)\theta) \sum_{n \in J_m(a)} g(n) \right| + \frac{1}{N} \sum_{n=1}^{N} |g(n)| \frac{2\pi L_m(a)}{\sqrt{N}} \\ &\leq \frac{1}{N} \sum_{a=1}^{h} \sum_{m=1}^{M} \left| \sum_{n \in J_m(a)} g(n) \right| + \frac{2\pi L_m(a)}{\sqrt{N}}. \end{aligned}$$

Now we choose M large enough so that $\frac{2\pi L_m(a)}{\sqrt{N}} \leq \frac{\sigma}{2}$ and get

$$\frac{1}{N} \sum_{a=1}^{h} \sum_{m=1}^{M} \left| \sum_{n \in J_m(a)} g(n) \right| \geq \frac{\sigma}{2}.$$

Since $\sum_{n=1}^{N} g(n) = 0$, we also have

$$\frac{1}{N} \sum_{a=1}^{h} \sum_{m=1}^{M} \max\left(0, \sum_{n \in J_m(a)} g(n)\right) \geq \frac{\sigma}{4}$$

Thus there exist some arithmetic progression $J_m(a)$ of length at least $\frac{\sigma\sqrt{N}}{8\pi}$ (provided $N \geq (16\pi/\sigma)^2$ for some positive constant c) such that

$$\frac{1}{|J_m(a)|} \sum_{n \in J_m(a)} g(n) \geq \frac{\sigma}{4}.$$

□

1.5. **Iteration.** Let A be a subset of $\mathbb{N}$ of upper density strictly larger than δ. For infinitely many prime numbers N, we can define $f_1 = 1_{B_N}$ and $f_2 = 1_{A_N}$ as before and we have

$$\frac{1}{N}\sum_{n\in\mathbb{Z}_N} f_1(n) \geq \delta/2 \quad \text{and} \quad \frac{1}{N}\sum_{n\in\mathbb{Z}_N} f_2(n) \geq \delta.$$

Furthermore, the number of 3-arithmetic progressions in A_N as a subset of $[1, N]$ is larger than $N^2\Lambda_3(f_1, f_2, f_1)$.
Given such a prime N larger than $(16\pi)^2/\alpha^2$, we define α by $\alpha = \mathbb{E}_{n\in\mathbb{Z}_N} f_2(n)$. Then, according to Proposition 1.6, either we have $\Lambda_3(f_1, f_2, f_1) \geq \alpha^3/8$ or there exist $k \in \mathbb{Z}_N\backslash\{0\}$ such that $\hat{f}_2(k) \geq \alpha^2/8$.
If A does not contain any non trivial 3-arithmetic progression, then we must be in the second case. We write $g = f_2 - \alpha 1_{[1,N]}$ and apply Proposition 1.8 with $\sigma = \alpha^2/8$. Thus we find an arithmetic progression $P = \{x+lh,\, l \leq N_1\}$ in $[1, N]$ on which the mean value of g is larger than $\sigma/4$. We write $N_1 = |P| \geq \sigma\sqrt{N}/(8\pi)$ and $A_1 = \{l \in [1, N_1] : n+lh \in P\cap A\}$. We get a subset A_1 of $[1, N_1]$ with

$$|A_1| \geq \left(\alpha + \frac{\sigma}{4}\right) N_1 \geq \left(\delta + \frac{\delta^2}{4\times 8}\right) N_1 =: (\delta + c_2\delta^2)N_1 =: \delta_1 N_1$$

and

$$N_1 \geq \frac{\delta^2\sqrt{N}}{64\pi} =: c_1\delta^2\sqrt{N} \text{ provided } N \geq c/\delta^2.$$

We iterate the argument on A_1. We get that if $N_1 \geq c/\delta_1^2$ (which is implied by $N_1 \geq c/\delta^2$), then there exists a subset A_2 of $\mathbb{Z}_{N_2}$ such that

$$N_2 \geq c_1(\delta + c_2\delta^2)^2\sqrt{N_1} \geq (c_1\delta^2)^{1+1/2}N^{1/4} \quad \text{and} \quad \frac{|A_2|}{N_2} \geq \delta + 2c_2\delta^2.$$

After k iteration, we get that there exists a subset A_k of $\mathbb{Z}_{N_k}$ such that

$$\frac{|A_k|}{N_k} \geq \delta + kc_2\delta^2 \quad \text{and} \quad N_k \geq \left(c_1\delta^2\right)^{\left(1+\frac{1}{2}+\cdots+\frac{1}{2^{k-1}}\right)} N^{\frac{1}{2^k}} = \left(c_1\delta^2\right)^{2(1-2^{-k})} N^{\frac{1}{2^k}}.$$

After $k_1 = \lceil \delta^{-1}/c_2 \rceil$ iterations, we get a relative density larger than 2δ on a progression of length larger than $N_{k_1} \geq \min(c_1\delta^2, 1)N^{1/2^{k_1}}$ which is larger than $c/\delta_{k_1}^2$ if N is large enough (depending on δ, c_1 and c_2). After k_2 iterations, we reach a density larger than 4δ on a progression of length larger than $N_{k_2} \geq \min(c_1(2\delta)^2, 1)N_{k_1}^{1/2^{k_2}}$. We finally reach a density larger than 1 and get a contradiction. Some precise study of admissible values for the constants give the quantitative version of Roth's theorem.

1.6. **Quantitative improvements of Roth's theorem.** Since 1953, significative quantitative improvements of Roth's theorem have been made. If $r_3(N)$ denotes the maximum size of a subset of positive integers less than N with no nontrivial 3-arithmetic progression, Roth proved that

$$r_3(N) \ll \frac{N}{\log\log N}.$$

Much later, Heath-Brown [8] (1987) and Szemerédi [18] (1990) improved (independantly) this result by showing that

$$r_3(N) \leq CN(\log N)^{-c}$$

for some small positive c and some large constant C. The main idea of Heath-Brown and Szemerédi's works is to consider diophantine approximations of a large bench of frequencies rather than one frequency at a time in the application of Dirichlet's Theorem. They prove that a lack of 3-arithmetic progression in A implies that the ℓ_2 norm on a subset S of $\mathbb{Z}_N$ (rather than the ℓ_∞ norm in Roth's work) of the Fourier transform of the balanced function $\mathbf{1}_A - \delta \mathbf{1}_{[1,N]}$ has to be large. Then they prove that if this ℓ_2 norm on S is large then A concentrates on some large arithmetic progression. To do so, they simultaneously approximate the elements ξ/N with $\xi \in S$.

By considering Bohr sets where previous arguments had used arithmetic progressions, Bourgain obtained

$$r_3(N) \leq CN(\log\log N)^2(\log N)^{-2/3}$$

in [2,3] (1999, 2008). Bourgain proves that a lack of 3-arithmetic progression in A implies a density increment in regular Bohr sets.
Part of these improvements are summarized in [19].
Very recently, Sanders [16] and [17] obtained the best known result so far by proving that

$$r_3(N) \leq CN\frac{(\log\log N)^5}{\log N}.$$

These quantitative results are related to some very strong conjecture.

Conjecture 1.9 (Erdös-Turán)**.** *Let A be a subset of the positive integers. If the series $\sum_{n\in A}\frac{1}{n}$ is divergent then A contains arbitrarily long arithmetic progressions.*

This is easy (cf tutorial) to prove that this conjecture is implied by the bounds

$$r_k(N) \ll_k \frac{N}{\log N(\log\log N)^{1+\varepsilon}}$$

with some positive ε (which may depend on k). Thus, Sanders' result is very close to proving that any subset A such that $\sum_{n\in A}\frac{1}{n}$ is divergent contains 3-arithmetic progressions.

2. Roth's theorem in sets of density 0.

2.1. **The general strategy.** Suppose that we have an infinite subset $\mathcal{A}$ of the integers which contains infinitely many 3-arithmetic progressions but which has zero density (think of the prime numbers for example). When can we say that Roth's theorem holds in this set ? That is, must a subset of positive relative density in $\mathcal{A}$ contain a 3-arithmetic progression?

2.1.1. *Varnavides Theorem.* Roth's argument not only prove that a subset of integers less than N with not too small upper density does contain some non trivial 3-arithmetic progressions. It also leads to the fact that such a set must contain many non trivial 3-arithmetic progressions. We shall explain how Varnavides [20] proved such a result and how this result is used to prove Roth's theorem in sets of integers of zero density.

Theorem 2.1 (Varnavides, 1959)**.** *Let $A \subset \mathbb{Z}_N$, where N is a prime. Assume $|A| \geq \eta N$ with $\eta > 0$. The number of 3-term arithmetic progressions in A is then at least $N^2 h(\eta)$ where c_0 and c_1 are absolute constants and*

$$h(\eta) := \frac{\eta}{c_0 \exp\left(c_1/\eta(\log(1/\eta))^5\right)}.$$

Proof. Recall Sanders' result [16]: for given L and $\eta \gg (\log\log L)^5/\log L$, every subset of $\{1, 2, \ldots, L\}$ with at least ηL elements contains at least one non-trivial three-term arithmetic progression. This result can be rephrased as follows: there are constants c_0 and c_1 such that, if $L \geq c_0 \exp\left(c_1/\eta(\log(1/\eta))^5\right)$, then any subset of $\{1, \cdots, L\}$ of density at least $\eta/2$ contains a non-trivial three-term arithmetic progression. It follows that, given an arithmetic progression $S_{a,d} = \{a + d, a + 2d, a + 3d, \ldots, a + Ld\}$ in $\mathbb{Z}_N$ ($a, d \in \mathbb{Z}_N$, $d \neq 0$, $L \leq N$) whose intersection with A has at least $(\eta/2)L$ elements, there is at least one non-trivial three-term arithmetic progression in $A \cap S \subset \mathbb{Z}_N$, provided L is large enough.

If we consider all arithmetic progressions of length L and given modulus $d \neq 0$ in $\mathbb{Z}_N$, we see that each element of A is contained in exactly L of them. Hence, $\sum_a |S_{a,d} \cap A| = L|A| \geq \eta N L$, and so (for $d \neq 0$ fixed) $|S_{a,d} \cap A| \geq (\eta/2)L$ for at least $(\eta/2)N$ values of a. Varying d, we get that $|S_{a,d} \cap A| \geq (\eta/2)L$ for at least $(\eta/2)N(N-1)$ arithmetic progressions $S_{a,d}$. By the above, each such intersection $S_{a,d} \cap A$ contains at least one non-trivial three-term arithmetic progression.

Each non-trivial three-term arithmetic progression $\{a_1, a_2, a_3\}$ in $\mathbb{Z}_N$ can be contained in at most $L(L-1)$ arithmetic progressions $\{a + d, a + 2d, \ldots, a + Ld\}$ of length L (the indices of a_1 and a_2 in the progression of length L determine the progression). Hence, when we count the three-term arithmetic progressions coming from the intersections $S_{a,d} \cap A$, we are counting each such progression at most $L(L-1)$ times. Thus we have shown that A contains at least

$$\frac{\eta}{2}\frac{N(N-1)}{L(L-1)} \geq \frac{\eta}{2}\frac{N^2}{L^2}$$

distinct non-trivial three-term arithmetic progressions for

$$L = \left\lceil c_0 \exp\left(c_1/\eta(\log(1/\eta))^5\right)\right\rceil,$$

provided that $L \leq N$. If $L > N$, the bound in the statement of the lemma is trivially true (as there is always at least one trivial three-term arithmetic progression in A.) □

Using this result on characteristic function of sets, we can prove the general following result.

Lemma 2.2. *Let N be a large prime number. For a real number $\alpha \in (0, 1)$ and a positive real number M, there exists some constant $c(\alpha, M)$ such that for any function $f : [1, N] \to [0, M]$ satisfying $\mathbb{E}_{n \in \mathbb{Z}_N} f(n) \geq \alpha$, we have the lower bound $\Lambda_3(f, f, f) > c(\alpha, M)$. Furthermore, we can take $c(\alpha, M) = \left(\frac{\alpha}{2}\right)^3 h\left(\frac{\alpha}{2M}\right)$.*

Proof. Let A be the set $A := \{n \in \mathbb{Z}_N : f(n) \geq \alpha/2\}$. Alors

$$\begin{aligned}\alpha \leq \mathbb{E}_{n\in\mathbb{Z}_N} f(n) &= \frac{1}{N}\sum_{n\in A} f(n) + \frac{1}{N}\sum_{n\notin A} f(n) \\ &\leq \frac{|A|}{N}M + \frac{N-|A|}{N}\frac{\alpha}{2} = \frac{\alpha}{2} + \left(M - \frac{\alpha}{2}\right)\frac{|A|}{N},\end{aligned}$$

thus $|A| \geq \frac{\alpha}{2M-\alpha}N \geq \frac{\alpha}{2M}N$. Now we apply Theorem 2.1 and get that

$$\Lambda_3(\mathbf{1}_A, \mathbf{1}_A, \mathbf{1}_A) \geq h\left(\frac{\alpha}{2M}\right),$$

therefore

$$\Lambda_3(f, f, f) \geq \left(\frac{\alpha}{2}\right)^3 h\left(\frac{\alpha}{2M}\right).$$

□

One could weaken the hypothesis of the theorem. Actually, this is not mandatory to have a L^∞ bound for f. A L^2 bound of f is enough as stated in the following Lemma.

Lemma 2.3. *Let α, c be positive real numbers and $f : \mathbb{Z}_N \to \mathbb{C}$ be a function satisfying $\mathbb{E}_{\mathbb{Z}_N} f \geq \alpha$ and $\|f\|_{L^2} \leq c$. Then*

$$\Lambda_3(f, f, f) \geq \left(\frac{\alpha}{2}\right)^3 h\left(\left(\frac{\alpha}{2c}\right)^2\right).$$

Proof. Let f be a function satisfying the hypotheses. Define $A = \{n \in \mathbb{Z}_N :\ f(n) \geq \alpha/2\}$. Then we have

$$\alpha \leq \mathbb{E}_{n\in\mathbb{Z}_N} f(n) \leq \frac{1}{N}\sum_{n\in A} f(n) + \frac{1}{N}\frac{\alpha}{2}(N - |A|).$$

But

$$\frac{1}{N}\sum_{n\in A} f(n) = \mathbb{E}_{n\in\mathbb{Z}_N} f(n)\mathbf{1}_A(n) \leq \|f\|_2 \|1_A\|_2 \leq c\sqrt{\frac{|A|}{|N|}},$$

thus

$$\alpha \leq c\sqrt{\frac{|A|}{|N|}} + \frac{1}{N}\frac{\alpha}{2}(N - |A|)$$

and $g(\sqrt{|A|}) \leq 0$ with $g(x) = x^2 - x\frac{2c}{\alpha}\sqrt{N} + N$. This leads to

$$|A| \geq N\left(\frac{c}{\alpha} - \sqrt{\left(\frac{c}{\alpha}\right)^2 - 1}\right)^2 \geq \left(\frac{\alpha}{2c}\right)^2 N.$$

Now, we can apply Varnavides Theorem to 1_A and we get

$$\Lambda_3(1_A, 1_A, 1_A) \geq h\left(\left(\frac{\alpha}{2c}\right)^2\right).$$

Since $f(n) \geq \alpha/2$ when $n \in A$, we have

$$\Lambda_3(f,f,f) \geq \left(\frac{\alpha}{2}\right)^3 \Lambda_3(1_A, 1_A, 1_A),$$

thus the result. □

Unfortunately, these theorems are not enough to directly deal with the case of subsets of integers of density 0. Indeed, taking a subset A of $\mathcal{A}$ of relative density $\delta > 0$ in $\mathcal{A}$, we get

$$\mathbb{E}_{[1,N]} 1_{A_N} \sim \delta \frac{|\mathcal{A}_N|}{N} \to_{N\to+\infty} 0$$

thus these theorems do not apply. We may think of normalizing the function so that the density becomes positive but then the normalized characteristic function $f = \frac{N}{|\mathcal{A}_N|}\mathbf{1}_{A_N}$ is not bounded anymore, neither in L^∞-norm nor in L^2-norm.

2.1.2. *Tranference principles.* In order to deal with the previous issue, we will approximate the normalized function f by some function f_1 satisfying the hypothesis of Lemma 2.2 (Green and Tao transference principle) or of Lemma 2.3 (new transference principle). Here "approximate" means that we will choose f_1 so that $\Lambda_3(f,f,f)$ is close to $\Lambda_3(f_1,f_1,f_1)$.

Before stating such results, we introduce here a few tools that will be needed in the proofs.

Definition 2.1. Let $f : \mathbb{Z}_N \to \mathbb{C}$ be some function and ε be some positive real number. We define the ε-spectrum R by

$$R = \{k \in \mathbb{Z}_N : |\hat{f}(k)| \geq \varepsilon\}$$

and the Bohr set B of set of frequencies R and of radius ε by

$$B = \{n \in \mathbb{Z}_N : \forall k \in R, \|nk/N\|_{\mathbb{R}/\mathbb{Z}} \leq \varepsilon\}.$$

We also define the normalized characteristic function of B, $\beta : \mathbb{Z}_N \to \mathbb{C}$ by $\beta = \frac{N}{|B|}\mathbf{1}_B$.

Bohr sets are usual tools in Fourier analysis. We already noticed that this was heavily used by Bourgain in is work on Roth's Theorem. Note that according to Lemma 1.7, we have

$$\forall k \in R, \quad |\hat{\beta}(k) - 1| \ll \varepsilon. \tag{2.1}$$

We shall also need a lower bound for the size of B. A pigeonhole argument leads to the lower bound

$$|B| \gg \varepsilon^r N \quad \text{where } r = |R|. \tag{2.2}$$

We are now ready to state the first transference principle.

Theorem 2.4 (Transference principle, Green-Tao [6])**.** *Let N be a large prime number. Let $f : \mathbb{Z}_N \to \mathbb{R}^+$ be a function satisfying the following conditions:*

(1) $\mathbb{E}_{\mathbb{Z}_N} f \geq \alpha$ for some $\alpha > 0$,

(2) $\|\hat{f}\|_p \leq M$ for some $p \in]2,3[$ and some $M > 0$,

(3) $f \leq \nu$ *with* $\sup_{k\in\mathbb{Z}_N} |\hat{\nu} - \mathbf{1}_0(k)| \leq \eta$ *for some* $\eta \in (0,1)$.

Then $\Lambda_3(f,f,f) \geq \left(\frac{\alpha}{2}\right)^3 h\left(\frac{\alpha}{4}\right) - O_{M,p}\left((\log\log(1/\eta)/\log(1/\eta))^{3/p-1}\right)$ *with* h *defined in Lemma 2.1.*

When the assumption *(3)* is satisfied, we say that ν is η-pseudorandom.

Remark 2.5. According to Lemma 2.3, assumptions *(1)* and *(2)* with $p = 2$ yields the conclusion.

Proof. We write $f = f_1 + f_2$ with $f_1 = f * \sigma$.
We choose σ so that f_1 satisfies the hypothesis of Roth's theorem and that $\Lambda_3(f_1,f_1,f_1)$ and $\Lambda_3(f,f,f)$ are close to each other.

(1) The formula

$$\Delta = |\Lambda_3(f,f,f) - \Lambda_3(f_1,f_1,f_1)| \leq \sum_k |\hat{f}(-2k)\hat{f}(k)^2||1 - \hat{\sigma}(-2k)\hat{\sigma}(k)^2|$$

leads us to choose σ so that $\hat{\sigma}(k)$ is close to 1 when $\hat{f}(k)$ is large. According to (2.1), the function β defined in Definition 2.1 should be a good candidate but it will be easier to work with a function σ such that $\hat{\sigma}$ is positive so we rather work with $\sigma = \beta * \beta$ with some $\varepsilon > 0$. According to (2.1) and the definitions, we have

$$|\hat{f}_2(k)| = |\hat{f}(k)(1-\hat{\sigma}(k))| \ll \begin{cases} \varepsilon \sum_{n\in\mathbb{Z}} |f(n)| \leq \varepsilon|\hat{\nu}(0)| \leq \varepsilon(1+\eta) & \text{if } k \in R, \\ |\hat{f}(k)| \leq \varepsilon & \text{if } k \notin R. \end{cases}$$

We write p' for the real number such that $\frac{1}{p} + \frac{1}{p'} = 1$. Hölder's inequality and hypothesis *(2)* yields

$$\begin{aligned} |\Lambda_3(f_2,f_2,f_2)| &\leq \|\hat{f}_2\|_{\ell_p} \|\hat{f}_2^{\,2}\|_{\ell_{p'}} \leq \|\hat{f}_2\|_{\ell_p} \left(\|\hat{f}_2\|_{\ell_p}^p \|\hat{f}_2\|_\infty^{2p'-p}\right)^{1/p'} \\ &= \|\hat{f}_2\|_{\ell_p}^p \|\hat{f}_2\|_\infty^{3-p} \ll_p M^p \left(\varepsilon(1+\eta)+\varepsilon\right)^{3-p} \ll_p M^p \varepsilon^{3-p}. \end{aligned}$$

The quantities $\Lambda_3(f_i,f_j,f_h)$ with at least two 2 among $\{i,j,h\}$ are similarly bounded: we apply Hölder's inequality with an ℓ_p norm for the function $\hat{f}_1$ and an $\ell_{p'}$ norm for the product of the two $\hat{f}_2$ functions; if needed, we apply Cauchy's inequality for this last product. The quantities $\Lambda_3(f_i,f_j,f_h)$ with only one 2 among $\{i,j,h\}$ are bounded this way: first, we apply Hölder's inequality with an $\ell_{p/2}$ norm for the product of the functions $\hat{f}_1$ and an $\ell_{(p/2)'}$-norm the function $\hat{f}_2$; if needed, we apply Cauchy's inequality for the product of the functions $\hat{f}_1$. The idea remains the same in any case: we know how to control the ℓ_p norm of both $\hat{f}_1$ and $\hat{f}_2$ and we want to be able to pick up some ℓ_∞ norm of $\hat{f}_2$ after Hölder's inequality.

With these estimates, we obtain the upper bound

$$\Delta = |\sum_{(i,j,h)\neq(1,1,1)} \Lambda_3(f_i,f_j,f_h)| \ll_p M^p \varepsilon^{3-p}.$$

(2) We have $\mathbb{E}_{\mathbb{Z}_N} f_1 = \mathbb{E}_{\mathbb{Z}_N} f \geq \alpha$. It remains to prove that f_1 is bounded.

$$\begin{aligned}
|f_1(n)| &= |B|^{-2} \sum_{m_1,m_2\in B} f(n+m_1-m_2) \\
&\leq |B|^{-2} \sum_{m_1,m_2\in B} \nu(n+m_1-m_2) \\
&= \left| |B|^{-2} \sum_{m_1,m_2\in B} \sum_{r\in\mathbb{Z}_N} \hat{\nu}(r) e_N(-r(n+m_1-m_2)) \right| \\
&\leq \sum_{r\in\mathbb{Z}_N} |\hat{\nu}(r)| \left| |B|^{-1} \sum_{m\in B} e_N(-rm) \right|^2 \\
&= \sum_{r\in\mathbb{Z}_N} |\hat{\nu}(r)| |\hat{\beta}(r)|^2.
\end{aligned}$$

Using hypothesis *(3)*, we get

$$|f_1(n)| \leq 1 + \eta \|\beta\|_2^2 = 1 + \eta \frac{N}{|B|}.$$

Since $\sum_m |\hat{f}(m)|^p \leq M$, we have $r \leq (M/\varepsilon)^p$, thus using (2.2), we get $|B| \gg \varepsilon^{(M/\varepsilon)^p} N$ and $|f_1(n)| \leq 1 + \eta(1/\varepsilon)^{(M/\varepsilon)^p}$. We take ε such that $\eta(1/\varepsilon)^{(M/\varepsilon)^p} = 1$ and apply Theorem 2.2 to f_1. We get

$$\Lambda_3(f_1, f_1, f_1) \geq c(\alpha, 2) \quad \text{with } c(\alpha, M) = \left(\frac{\alpha}{2}\right)^3 h\left(\frac{\alpha}{2M}\right).$$

Putting everything together we obtain

$$\Lambda_3(f, f, f) \geq c(\alpha, 2) - O_p\left(M^3(\varepsilon/M)^{3-p}\right) \geq c(\alpha, 2) - O_p\left(M^3 \left(\frac{\log(1/\varepsilon)}{\log(1/\eta)}\right)^{3/p-1}\right).$$

Now we use that with our choice of ε we have $\log(1/\varepsilon) \leq \frac{1}{p} \log\log(1/\eta)$ and get the announced result. □

There are many applications of this principle in the literature: in [5] for the prime numbers, [6] for Chen primes and [19] for random subsets of torsion groups. We shall see later how Green used this principle in [5] to prove Roth's theorem in the primes. We also state an alternative transference principle based on Lemma 2.3, this alternative principle is used in [9] to sharpen Green's quantitative result in the primes.

Theorem 2.6. *Let N be a large prime number. Let $f : Z_N \to \mathbb{R}^+$ be a function satisfying the following conditions:*

(1) $\alpha \leq \mathbb{E} f \leq 1$,

(2) $\|\hat{f}\|_p \leq M$ *for some* $p \in (2, 3)$ *and some* $M > 0$,

(3) $\|f_1\|_2 \leq c$ *with* $f_1 = f * \beta$ *where* $\beta = \frac{N}{|B|} \mathbf{1}_B$ *and* B *is the Bohr set defined as previously.*

Then

$$\Lambda_3(f,f,f) \geq \left(\frac{\alpha}{2}\right)^3 h\left(\left(\frac{\alpha}{2c}\right)^2\right) - O_p\left(M^p\varepsilon^{3-p}\right).$$

Proof. We choose $\varepsilon > 0$ and write $f = f_1 + f_2$ with $f_1 = f * \beta$ where β is defined as in Definition 2.1. We shall mimic the proof of Theorem 2.4.
In the first step of the proof of Theorem 2.4, we only used the third hypothesis to get an upper bound of $\|f\|_1$. Furthermore, replacing σ by β in this step is harmless. Therefore, using β rather than σ and replacing $\|f\|_1 \leq |\hat{\nu}(0)|$ by $\|f\|_1 \leq 1$ in the first step leads to the upper bounds

$$\Delta = |\sum_{(i,j,h)\neq(1,1,1)} \Lambda_3(f_i,f_j,f_h)| \ll_p M^p\varepsilon^{3-p}.$$

It remains to prove that $\Lambda_3(f_1,f_1,f_1)$ is large. We have again $\mathbb{E}f_1 = \mathbb{E}f \geq \alpha$ and now $\|f_1\|_2 \leq c$. We apply Lemma 2.3 and get

$$\Lambda_3(f_1,f_1,f_1) \geq \left(\frac{\alpha}{2}\right)^3 h\left(\left(\frac{\alpha}{2c}\right)^2\right)$$

and therefore

$$\Lambda_3(f,f,f) \geq \Lambda_3(f_1,f_1,f_1) - \Delta \geq \left(\frac{\alpha}{2}\right)^3 h\left(\left(\frac{\alpha}{2c}\right)^2\right) - O_p(M^p\varepsilon^{3-p}).$$

□

2.2. **Roth's Theorem for the primes.** In this section, we will mostly give some sketches of the proves which can be found in the literature. The first proof of Roth's Theorem in the primes was given by Green in [5]. He proved that there exist some constant C such that if a subset A of the primes satisfies

$$\delta_P(N) := \frac{|A_N|}{|\mathcal{P}_N|} \geq C\sqrt{\log\log\log\log\log N}/\sqrt{\log\log\log\log N}$$

for infinitely many integers N, then A must contain some non trivial 3-arithmetic progression. His quantitative result was sharpen by Helfgott and de Roton in [9] who proved that

$$\delta_P(N) \geq C(\log\log\log N)/(\log\log N)^{1/3}$$

is enough. Their method would actually lead, with the use of Sander's result rather than Bourgain's one on Roth's Theorem, to the better bound :

$$\delta_P(N) \geq C(\log\log\log N)^{5/2}/(\log\log N)^{1/2}.$$

The best known result so far is due to Naslund [10] who proved that

$$\delta_P(N) \geq 1/(\log\log N)^{1-\varepsilon} \quad \text{for some positive } \varepsilon$$

is enough.

2.2.1. *Pseudo random measure for the primes.* Following the strategy of transference principle (Theorem 2.4), we would like to majorate the normalized indicator function of the primes by some uniform measure. Unfortunately, prime numbers are ill distributed in arithmetic progressions with small modulus. For example, there are only 1 prime number with residue 0 modulo 3 and half of the primes with residue 1 and residue 2. This fact explains the reason why we need to switch from a set of primes to a set of integers (non necessarily prime) to find a "uniform" measure on our set. The W-trick (terminology in [6] coming from the use of the letter W in [5]) consists in focussing on the intersection of the primes with an arithmetic progression of large modulus, rather than working on all the primes.

Lemma 2.7. *Let α, z be positive real numbers and N' be a large integer. We define $W = \prod_{p\leq z} p$. Let A be a subset of the primes less than N' such that $|A| \geq \alpha N'/\log N'$. Then there exists some arithmetic progression $P(b) = \{b+nW : 1 \leq n \leq N'/W\}$ such that*

$$|P(b) \cap A| \gg \alpha \frac{\log z}{\log N'} \frac{N'}{W} - \log z,$$

where the implied constant is absolute.

Proof. If $(b, W) \neq 1$, the set $\{m \in P(b) : m \text{ prime}\}$ is empty. Since the progressions $P(b)$ with $(b, W) = 1$ are distinct, we have

$$\sum_{b:(b,W)=1} |A \cap P(b)| = |A| - |A \cap [1, W-1]| \geq \alpha \frac{N}{\log N} - W.$$

But $|\{b \leq W : (b, W) = 1\}| \sim W/\log z \sim e^z/\log z$. Therefore there exists some progression $P(b)$ such that

$$|A \cap P(b)| \gg \left(\alpha \frac{N}{\log N} - W\right) \frac{\log z}{W} \gg \alpha \frac{\log z}{\log N} \frac{N}{W} - \log z.$$

□

Now, we fix $z = \frac{1}{3}\log N'$, $W = \prod_{p\leq z} p$, and let N be the least prime larger than $\lceil 2N'/W \rceil$. (The requirement $N > \lceil 2N'/W \rceil$ will ensure that no new three-term arithmetic progressions are created when we apply the reduction map $\pi : \mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ to a set contained in $[1, N'/W]$.) By Bertrand's postulate, $N \leq 2N'/W$. Let A be a subset of the primes less than N' such that $|A| \geq \alpha N'/\log N'$. We assume $\alpha \geq (\log N')N'^{-1/2}$ (say) and obtain from Lemma 2.7 that there is an arithmetic progression $P(b)$ such that $|P(b) \cap A| \gg \alpha(\log z/\log N)N$. We define A_0 to be

$$A_0 = \left\{ n = \frac{m-b}{W} : m \in P(b) \cap A \right\}. \tag{2.3}$$

This is a subset of $X := \{n \in [1, N] : b + nW \text{ is prime}\}$ satisfying

$$|A_0| \gg \alpha \frac{\log z}{\log N} N.$$

Our task is to show that there is a non-trivial three-term arithmetic progression in $A_0 \subset \mathbb{Z}/N\mathbb{Z}$. It will follow immediately that there is a non-trivial three-term arithmetic progression in $A \subset \mathbb{Z}$. According to the transference principles, this is enough to have $\|f\|_{L^1}$ not too small, $\|\hat{f}\|_{\ell_p}$ bounded for some $p \in (2,3)$ and either f majorated by some pseudorandom measure of $\|f * \beta\|_{L^2}$ bounded.

Let f be the normalized characteristic function of A_0, i.e., $f = (\log N/(N \log z))\mathbf{1}_{A_0}$. To begin with, we remark that $\|f\|_1 = (\log N/(N' \log z))|A_0| \gg \alpha$. Furthermore, by the definition of f, we have $0 \leq f(n) \leq \lambda(n)$, where $\lambda : \mathbb{Z} \to \mathbb{R}$ is defined by

$$\lambda = \frac{\log N}{N \log z}\mathbf{1}_X \quad \text{with} \quad X = \{n : 1 \leq n \leq N \text{ and } b + nW \text{ is prime}\}. \tag{2.4}$$

Green uses in [5] that the W-trick (passage to an arithmetic progression $b+nW$ of large modulus) removes all but the largest peaks in the Fourier transform of the primes. He uses the circle method to get an upper bound for $\hat{\lambda}$.
Helfgott and de Roton [9] use in a more direct way the fact that the elements of $\{n : b + nW \text{ prime}\}$ are not forbidden from having small divisors. They use Ramaré's enveloping sieve [12], based partially on work on sieves in [11] (see also [6]). This allows them to take a larger W but this step remains necessary.

We will not explain here how this enveloping use of a sieve works. We just note that Green's work uses some restriction theorems on primes and refer to [5], [1] for these theorems and Helfgott and de Roton use some restriction theorem for an upper-bound sieve (see [6] for the enveloping sieve).

The application of Green and Tao's results on the enveloping sieve yields

$$\sum_{m \in \mathbb{Z}_N} |\hat{f}(m)|^q \ll_q 1 \quad \text{for } q > 2. \tag{2.5}$$

To summarize, we now have a function $f : \mathbb{Z}_N \to \mathbb{R}$ satisfying the hypotheses *(1)* and *(2)* of both Theorem 2.4 and Theorem 2.6. Whereas Green worked with the first Theorem, Helfgott and de Roton worked with the second one. We shall now explain how they gained the third hypothesis. Using Theorem 2.6 then lead to a quantitative Roth's Theorem in the primes.

2.2.2. *A bound for the L^2 norm of f_1.* The main idea is the following: if we convolve the function λ defined in (2.4) with β, then the L^2 norm of λ will go down dramatically. This is simply due to the fact that, for any $k \neq 0$, we have a good upper bound (essentially $\ll \frac{k}{\phi(k)} \frac{N}{(\log N)^2}$) on the number of primes $p \leq N$ such that $p + k$ is also a prime. The only fact about β we shall use is that it is not concentrated in too small a set, i.e., the Bohr set B is not too small. The fact that we have restricted ourselves to a congruence class $b+nW$ modulo W will finally come into play, in that it will in effect eliminate the potentially harmful factor $\frac{k}{\phi(k)}$.

Since β is nonnegative, we have $\|f_1\|_2 = \|f * \beta\|_2 \le \|\lambda * \beta\|_2$ with

$$\lambda(n) = \begin{cases} \log N/\log z & \text{if } 1 \le n \le N \text{ and } b + nW \text{ prime,} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, we can assume that β and f are defined on $[-(N-1)/2, (N-1)/2]$ rather that on $[1, N]$ is we take $\varepsilon < 1/4$. Therefore, by symmetry of Bohr sets, we have $\beta(-m) = \beta(m)$ and

$$\|\lambda * \beta\|_2^2 = \mathbb{E}_n \left|\mathbb{E}_m \beta(n)\lambda(n-m)\right|^2 \tag{2.6}$$

$$= \mathbb{E}_{m_1}\mathbb{E}_{m_2}\beta(m_1)\beta(m_2)\mathbb{E}_n\lambda(n+m_1)\lambda(n+m_2). \tag{2.7}$$

Lemma 2.8. *For any integers m_1, m_2, we have*

$$\mathbb{E}_n\lambda(n+m_1)\lambda(n+m_2) \ll \begin{cases} \log N/(\log z) & \text{if} \quad m_1 = m_2, \\ \prod_{p|(m_1-m_2),\, p>z} \frac{p}{p-1} & \text{if} \quad m_1 \neq m_2, \end{cases} \tag{2.8}$$

where the implied constant is absolute.

Proof. The case $m_1 = m_2$ follows from Brun-Titchmarsh:

$$\begin{aligned} \mathbb{E}_n\lambda^2(n+m) &= \frac{1}{N}\left(\frac{\log N}{\log z}\right)^2 \left|\{m \le n \le N' + m : b + (n-m)W \text{ is prime }\}\right| \\ &\ll \left(\frac{\log N}{\log z}\right)^2 \frac{W}{\varphi(W)\log N} \\ &\ll \frac{\log N}{\log^2 z}\prod_{p\le z}(1-1/p)^{-1} \ll \frac{\log N}{\log z}. \end{aligned}$$

To obtain the case $m_1 \neq m_2$, we will use a result based on Selberg's sieve. It is clear that $\mathbb{E}_n\lambda(n+m_1)\lambda(n+m_2)$ equals $\frac{1}{N}\left(\log N/(\log z)\right)^2$ times

$$\left|\{1 \le n \le N' : b + nM \text{ and } b + (n + m_2 - m_1)M \text{ are primes}\}\right|. \tag{2.9}$$

By [7, Thm. 5.7],

$$\begin{aligned} (2.9) &\ll \prod_{p\le z}\left(1-\frac{1}{p}\right)^{-2} \prod_{\substack{p>z \\ p|m_1-m_2}} \left(1-\frac{1}{p}\right)^{-1} \frac{N'}{(\log N')^2} \\ &\ll \prod_{\substack{p>z \\ p|m_1-m_2}} \frac{p}{p-1}\frac{N'(\log z)^2}{(\log N')^2}. \end{aligned}$$

The statement follows. □

Let us now evaluate the last line of (2.7), with Lemma 2.8 in hand. The contribution of the diagonal terms ($m_1 = m_2$) in (2.7) is $\ll \log N/(|B|N\log z)$. The contribution of the

non-diagonal terms ($m_1 \neq m_2$) is

$$\ll \frac{1}{N} \sum_{\substack{m_1 \\ m_2 \neq m_1}} \sum_{m_2} \beta(m_1)\beta(m_2) \prod_{\substack{p>z \\ p|m_1-m_2}} \frac{p}{p-1}. \tag{2.10}$$

Some consideration on the number of large prime divisors of an integer less than N yield to

$$\prod_{\substack{p>z \\ p|m}} \frac{p}{p-1} \ll 1 \quad \text{for any } m \neq 0 \quad \text{with } |m| \leq N'.$$

Thus

$$\sum_n |\beta * \lambda(n)|^2 \ll \frac{1}{N} \left(\frac{\log N}{|B| \log z} + 1 \right).$$

The right side is $\ll 1/N$ as long as $|B| \gg \log N / \log z$. Using (2.2) we get some condition on the parameters so that it happens. With all the constraints, we get the announced result.

To conclude, we briefly explain the main difference between Naslund's work [10] and Helfgott and de Roton's one [9]. In his work, Naslund uses the L^{2q} norm of f_1 rather than de L^2 norm. Since there are very few prime numbers p such that $p, p+k_1, p+k_2, \cdots p+k_l$ are all prime, he gets a very good bound for this norm. He then applies a transference principle, similar to Theorem 2.6 with an L^{2q} norm rather than an L^2-norm to conclude.

References

[1] J. Bourgain. On $\Lambda(p)$-subsets of squares. *Israel J. Math.*, 67(3):291–311, 1989.
[2] J. Bourgain. On triples in arithmetic progression. *Geom. Funct. Anal.*, 9(5):968–984, 1999.
[3] Jean Bourgain. Roth's theorem on progressions revisited. *J. Anal. Math.*, 104:155–192, 2008.
[4] W. T. Gowers. Fourier analysis and Szemerédi's theorem. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, number Extra Vol. I, pages 617–629 (electronic), 1998.
[5] Ben Green. Roth's theorem in the primes. *Ann. of Math. (2)*, 161(3):1609–1636, 2005.
[6] Ben Green and Terence Tao. Restriction theory of the Selberg sieve, with applications. *J. Théor. Nombres Bordeaux*, 18(1):147–182, 2006.
[7] H. Halberstam and H.-E. Richert. *Sieve methods*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974. London Mathematical Society Monographs, No. 4.
[8] D. R. Heath-Brown. Integer sets containing no arithmetic progressions. *J. London Math. Soc.*, 35(2):385–394, 1987.
[9] Harald Andrés Helfgott and Anne de Roton. Improving Roth's theorem in the primes. *Int. Math. Res. Not. IMRN*, (4):767–783, 2011.
[10] Naslund. On improving roth's theorem in the primes. *arXiv*, 1302.2299, 2013.
[11] Olivier Ramaré. On snirel'man's constant. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 22(4):645–706, 1995.
[12] Olivier Ramaré and Imre Z. Ruzsa. Additive properties of dense subsets of sifted sequences. *J. Théor. Nombres Bordeaux*, 13(2):559–581, 2001.
[13] K. F. Roth. On certain sets of integers. *J. London Math. Soc.*, 28:104–109, 1953.
[14] Imre Z. Ruzsa. Solving a linear equation in a set of integers. I. *Acta Arith.*, 65(3):259–282, 1993.
[15] Imre Z. Ruzsa. Solving a linear equation in a set of integers. II. *Acta Arith.*, 72(4):385–397, 1995.
[16] Tom Sanders. On Roth's theorem on progressions. *Ann. of Math. (2)*, 174(1):619–636, 2011.

[17] Tom Sanders. On certain other sets of integers. *J. Anal. Math.*, 116:53–82, 2012.
[18] E. Szemerédi. Integer sets containing no arithmetic progressions. *Acta Math. Hungar.*, 56(1-2):155–158, 1990.
[19] Terence Tao and Van H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Paperback edition [of MR2289012].
[20] P. Varnavides. Note on a theorem of Roth. *J. London Math. Soc.*, 30:325–326, 1955.