

On a generalisation of Roth's theorem for arithmetic progressions and applications to sum-free subsets

Jehanne Dousse

April 5, 2013

Abstract

We prove a generalisation of Roth's theorem for arithmetic progressions to d -configurations, which are sets of the form $\{n_i + n_j + a\}_{1 \leq i \leq j \leq d}$ with $a, n_1, \dots, n_d \in \mathbb{N}$, using Roth's original density increment strategy and Gowers uniformity norms. Then we use this generalisation to improve a result of Sudakov, Szemerédi and Vu about sum-free subsets [10] and prove that any set of n integers contains a sum-free subset of size at least $\log n \left(\log^{(3)} n\right)^{1/32772 - o(1)}$.

1 Introduction

In 1953 Roth [8] proved his famous theorem about arithmetic progressions of length 3.

Theorem 1.1. *Let $0 < \alpha < 1$. Any subset of $\{1, \dots, N\}$ of size αN with $N \geq N_0(\alpha)$ contains a non-trivial arithmetic progression of length 3.*

A d -configuration is a set of the form $\{n_i + n_j + a\}_{1 \leq i \leq j \leq d}$ with $a, n_1, \dots, n_d \in \mathbb{N}$. The d -configuration is *non-trivial* if for all $i \neq j$, $n_i \neq n_j$.

Our aim is to prove the following generalisation of Roth's theorem:

Theorem 1.2. *Let $0 < \alpha < 1$ and $d \geq 1$. Any subset of $\{1, \dots, N\}$ of size αN with $N \geq N_0(\alpha, d)$ contains a non-trivial d -configuration.*

If $d = 2$, this is equivalent to Roth's theorem because 2-configurations are exactly arithmetic progressions of length 3.

Theorem 1.2 can be proved easily by using Szemerédi's theorem [11], presented here in a quantitative form due to Gowers [4]. We write $p \uparrow q$ for p^q , with the convention that $p \uparrow q \uparrow r$ stands for $p \uparrow (q \uparrow r)$.

Theorem 1.3. *Let $0 < \alpha < 1$, and k be a positive integer. Let $N \geq 2 \uparrow 2 \uparrow (\alpha)^{-1} \uparrow 2 \uparrow 2 \uparrow (k + 9)$ and let A be a subset of $[N]$ with cardinality αN . Then A contains a non-trivial arithmetic progression of length k .*

Now to prove Theorem 1.2 we can locate in A a progression P of length $2d - 1$, set the $2n_i + a$'s to be the elements with odd indices of P , and notice that P is a d -configuration. By Theorem 1.3, this is possible if $N \geq 2 \uparrow 2 \uparrow (\alpha)^{-1} \uparrow 2 \uparrow 2 \uparrow (2d + 8)$, but this bound is very unsatisfactory. Therefore we give a proof that does not involve such a deep theorem as Szemerédi's, and show that Theorem 1.2 is true for $N \geq \exp\left(\exp\left(\left(\frac{C}{\alpha}\right)^{d(d+1)-1}\right)\right)$, where C is some absolute constant.

In the second part of this paper we use Theorem 1.2 to prove a result about sum-free subsets. For two finite sets of real numbers A and B , one says that B is *sum-free* with respect to A if the set $\{b + b' \mid b, b' \in B, b \neq b'\}$ is disjoint from A . Let $\phi(n)$ denote the largest integer such that any set A of size n contains a subset of cardinality $\phi(n)$ which is sum-free with respect to A . An interesting question is to

find a lower bound for $\phi(n)$. Erdős first mentioned in [3] that $\phi(n) \geq c \log n$ for some constant c , and the first published proof of this result was done by Choi [2] who proved that $\phi(n) \geq \log_2 n$. Then Ruzsa [9] improved this result slightly by showing that $\phi(n) \geq 2 \log_3 n - 1$. Recently Sudakov, Szemerédi and Vu [10] gave the first superlogarithmic bound by showing that $\phi(n) \geq \log n \left(\log^{(5)} n \right)^{1-o(1)}$, where $\log^{(i)} n$ denotes the iterated logarithm ($\log^{(0)} x = x$, $\log^{(i+1)} x = \log(\log^{(i)} x)$). By modifying a small part of their proof using Theorem 1.2, we prove that $\phi(n) \geq \log n \left(\log^{(3)} n \right)^{1/32772-o(1)}$.

2 Proof of Theorem 1.2

As in several proofs of Roth's theorem [8, 6, 7, 12], we use the density increment strategy to prove Theorem 1.2. It consists of showing that either A contains a non-trivial d -configuration or it has increased density on some arithmetic progression.

In the following, we will use the notation $[N] := \{1, \dots, N\}$ and $e(\theta) := e^{2i\pi\theta}$.

2.1 Complexity and Gowers uniformity norms

Gowers' uniformity norms play an important role in his proof of Szemerédi's theorem [4], and therefore in the particular case of Roth's theorem, as explained in Green's course notes [6].

We will use these norms in our proof too, but we first need the notion of complexity, introduced by Green and Tao in [5].

Definition. Let $\Psi = (\psi_1, \dots, \psi_t)$ be a system of affine-linear forms. If $1 \leq i \leq t$ and $s \geq 0$, we say that Ψ has i -complexity at most s if one can cover the $t - 1$ forms $\{\psi_j : j \in [t] \setminus \{i\}\}$ by $s + 1$ classes, such that ψ_i does not lie in the affine-linear span of any of these classes. The *complexity* of Ψ is defined to be the least s for which the system has i -complexity at most s for all $1 \leq i \leq t$, or ∞ if no such s exists.

We will now define the Gowers uniformity norms.

Let f be a function from $\mathbb{Z}/N\mathbb{Z}$ to \mathbb{C} . The *expectation* of f is defined to be the quantity

$$\mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f(x) := \frac{1}{N} \sum_{x \in \mathbb{Z}/N\mathbb{Z}} f(x).$$

Definition. Suppose that $f : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{C}$ is a function. Let $k \geq 2$ be an integer. The *Gowers U^k -norm* is defined by

$$\|f\|_{U^k} := \left(\mathbb{E}_{x, h_1, \dots, h_k \in \mathbb{Z}/N\mathbb{Z}} \prod_{(\omega_1, \dots, \omega_k) \in \{0, 1\}^k} C^{|\omega_1| + \dots + |\omega_k|} f(x + \omega_1 h_1 + \dots + \omega_k h_k) \right)^{1/2^k},$$

where we use the notation $Cf := \overline{f}$.

In particular the Gowers U^2 -norm which we will use later is defined by

$$\|f\|_{U^2} := \left(\mathbb{E}_{x, h_1, h_2 \in \mathbb{Z}/N\mathbb{Z}} f(x) \overline{f(x + h_1)} \overline{f(x + h_2)} f(x + h_1 + h_2) \right)^{1/4}.$$

To establish a link between complexity and Gowers uniformity norms, we will use the following theorem [5], which allows to control systems of complexity s by the Gowers U^{s+1} -norm.

Theorem 2.1. *Let $f_1, \dots, f_t : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{R}$ be functions such that $|f_i(x)| \leq 1$ for all $i \in [t]$ and all $x \in \mathbb{Z}/N\mathbb{Z}$. Suppose that $\Psi = (\psi_1, \dots, \psi_t)$ is a system of affine-linear forms of complexity s consisting of t forms in d variables. Then*

$$\left| \mathbb{E}_{x_1, \dots, x_d \in \mathbb{Z}/N\mathbb{Z}} \prod_{i=1}^t f_i(\psi_i(x_1, \dots, x_d)) \right| \leq \min_{1 \leq i \leq t} \|f_i\|_{U^{s+1}}.$$

We will now show that d -configurations have complexity 1, in order to be able to control averages using the Gowers U^2 -norm.

Lemma 2.2. *Let $(n_i + n_j + a)_{1 \leq i \leq j \leq d}$ be a d -configuration. Then it has complexity 1.*

Proof: Let $i_0, j_0 \in [d]$ such that $i_0 < j_0$. We take the first class to consist of all forms involving n_{i_0} , and the second one to consist of all other forms. Then the form $n_{i_0} + n_{j_0} + a$ is in the linear span of neither of those classes, because the first one does not involve n_{j_0} at all, and the second one does not involve n_{i_0} .

Let us now consider the form $2n_{i_0} + a$. As in the previous case, we take the first class to consist of all forms involving n_{i_0} , and the second one to consist of all other forms. Clearly $2n_{i_0} + a$ is not in the linear span of the second class because it does not involve n_{i_0} at all. And it is also not in the linear span of the first class, because in each of its forms, n_{i_0} appears with a different n_j , and we cannot cancel them to obtain $2n_{i_0} + a$.

Therefore the system has (i_0, j_0) -complexity 1 for all $1 \leq i_0 \leq j_0 \leq d$, so the d -configuration has complexity at most 1. It does not have complexity 0 because if we consider the form $n_{i_0} + n_{j_0} + a$ and put all other forms in the same class, it is in its linear span. Indeed for example $2(n_{i_0} + n_{j_0} + a) = (2n_{i_0} + a) + (2n_{j_0} + a)$. So the d -configuration has complexity 1. \square

As a particular case of Theorem 2.1, using the fact that d -configurations have complexity 1, we have the following theorem:

Theorem 2.3. *Let $f_{ij} : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{R}$ be functions such that $|f_{ij}(x)| \leq 1$ for all $1 \leq i \leq j \leq d$, and a be an integer. Then*

$$|\mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} f_{ij}(n_i + n_j + a)| \leq \min_{1 \leq i \leq j \leq d} \|f_{ij}\|_{U^2}.$$

2.2 Obtaining a large Gowers U^2 -norm

In this subsection, we prove that either A contains non-trivial d -configurations or a particular function f_A has a large Gowers U^2 -norm.

Let P be an arithmetic progression of length N , and $A \subseteq P$ a set of size αN . By linear rescaling, we may assume that $P = [N]$, because it does not change either the density α or the number of d -configurations in A . Indeed let us assume that the forms $(n_i + n_j + a)_{1 \leq i \leq j \leq d}$ are located in the progression P . Adding some constant to each element of P , and changing a accordingly, we may assume that P is equal to $\{k, 2k, \dots, Nk\}$ for some integer k and that for every $1 \leq i \leq j \leq d$, $n_i + n_j + a = y_{i,j}k$. By linearly rescaling P to $[N]$, our d -configuration becomes $(y_{i,j})_{1 \leq i \leq j \leq d}$. Let us show that this is still a d -configuration. We have for every $1 \leq i \leq d-1$, $n_{i+1} - n_i = (y_{i,i+1} - y_{i,i})k$, which means that the n_i 's are also inside an arithmetic progression with common difference k . Let us write $n_i = \beta_i k + b$ where $\beta_i \in \mathbb{N}$ for all $1 \leq i \leq d$. Then we have $n_i + n_j + a = \beta_i k + \beta_j k + 2b + a = y_{i,j}k$, so if we set $a' := \frac{2b+a}{k}$, we have $\beta_i + \beta_j + a' = y_{i,j}$ for all $1 \leq i \leq j \leq d$, which means that we still have a d -configuration.

Conversely if $(m_i + m_j + a)_{1 \leq i \leq j \leq d}$ is a d -configuration in $[N]$, then the linear system $(k(m_i + m_j + a) + b)_{1 \leq i \leq j \leq d} = (km_i + km_j + (ak + b))_{1 \leq i \leq j \leq d}$ is a d -configuration located in the progression $\{k + b, \dots, Nk + b\}$.

Let us set $N' := 2N + 1$ and let \tilde{A} denote A considered as a subset of $\mathbb{Z}/N'\mathbb{Z}$. The number of d -configurations in \tilde{A} is the same as that in A , so we will identify \tilde{A} and A . Let us also note that to count once and only once each d -configuration $(n_i + n_j + a)_{1 \leq i \leq j \leq d}$ we shall assume that a is equal either to 0 or 1.

Given functions $(f_{ij})_{1 \leq i \leq j \leq d}$, set

$$\Pi_d((f_{ij})_{1 \leq i \leq j \leq d}) := \mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} f_{ij}(n_i + n_j) + \mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} f_{ij}(n_i + n_j + 1).$$

The quantity $\Pi_d(1_A, \dots, 1_A)$, where 1_A is the characteristic function of A ($1_A(x) = 1$ if $x \in A$, 0 otherwise), is equal to $1/N'^d$ times the number of d -configurations in A , including the trivial ones. We shall compare this with $\Pi_d(\alpha 1_{[N]}, \dots, \alpha 1_{[N]})$, where $\alpha 1_{[N]}(x)$ is defined to be α if $x \in [N]$ and 0 if $x \in \mathbb{Z}/N'\mathbb{Z} \setminus [N]$.

To compute the difference between the two we introduce the *balanced function* of A , defined by

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

Let us note that the expectation of f_A is equal to 0. This property will be useful later in the proof.

Since Π_d is multilinear, we may expand $\Pi_d(1_A, \dots, 1_A)$ as a main term $\Pi_d(\alpha 1_{[N]}, \dots, \alpha 1_{[N]}) = \alpha^{d(d+1)/2} \Pi_d(1_{[N]}, \dots, 1_{[N]})$ plus $2^{d(d+1)/2} - 1$ other terms $\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})$ where at least one the g_{ij} 's is equal to f_A .

Lemma 2.4. *Suppose that $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$ and that A contains fewer than $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ non-trivial d -configurations. Then there are 1-bounded functions $(g_{ij})_{1 \leq i \leq j \leq d}$, at least one of which being equal to f_A , such that*

$$|\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})| \geq \left(\frac{\alpha}{C}\right)^{\frac{d(d+1)}{2}}$$

for some absolute constant $C > 0$, where $\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})$ is one of the $2^{d(d+1)/2} - 1$ other terms.

Proof: We have

$$\alpha^{\frac{d(d+1)}{2}} \Pi_d(1_{[N]}, \dots, 1_{[N]}) \geq \alpha^{\frac{d(d+1)}{2}} \times \frac{1}{N'^d} \times 2 \left(\frac{N}{2}\right)^d,$$

because if we choose $1 \leq n_i \leq N/2$ for all $1 \leq i \leq d$, then for all $1 \leq i \leq j \leq d$, $1 \leq n_i + n_j \leq N$, giving at least $\left(\frac{N}{2}\right)^d$ d -configurations of the form $(n_i + n_j)_{1 \leq i \leq j \leq d}$, and if we choose $0 \leq n_i \leq N/2 - 1$ for all $1 \leq i \leq d$, then for all $1 \leq i \leq j \leq d$, $1 \leq n_i + n_j + 1 \leq N$, giving at least $\left(\frac{N}{2}\right)^d$ d -configurations of the form $(n_i + n_j + 1)_{1 \leq i \leq j \leq d}$. Therefore we have at least $2 \left(\frac{N}{2}\right)^d$ d -configurations in total (including trivial ones).

Recall that a d -configuration $(n_i + n_j)_{1 \leq i \leq j \leq d}$ is trivial if at least two of the n_i 's are equal. Let us find an upper bound for the number of trivial d -configurations of the form $(n_i + n_j)_{1 \leq i \leq j \leq d}$ in A . First, we have $\binom{d}{2}$ ways to choose i_0 and j_0 for which we set $n_{i_0} = n_{j_0}$. We know that $2n_{i_0}$ must be in A for all i_0 , so there are at most $|A| = \alpha N$ choices for n_{i_0} (we have an equality here if A only contains even integers). Then n_{j_0} is forced to be equal to n_{i_0} , and we have at most $|A| = \alpha N$ choices for each one of the $d-2$ other n_i 's too, giving at most $\binom{d}{2} (\alpha N)^{d-1}$ trivial d -configurations of the form $(n_i + n_j)_{1 \leq i \leq j \leq d}$ in A . Note that we may have counted some d -configurations that are not fully contained in A , but we only want an upper bound so this is not a problem here. The d -configurations of the form $(n_i + n_j + 1)_{1 \leq i \leq j \leq d}$ work exactly in the same way, and in total we have at most $2 \binom{d}{2} (\alpha N)^{d-1} = d(d-1) (\alpha N)^{d-1}$ trivial d -configurations in A . Therefore, using the fact that A contains fewer than $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ non-trivial d -configurations, we obtain

$$\Pi_d(1_A, \dots, 1_A) \leq \frac{\alpha^{\frac{d(d+1)}{2}}}{2^{d-1} + 1} \left(\frac{N}{N'}\right)^d + d(d-1) \frac{(\alpha N)^{d-1}}{N'^d}.$$

After some calculation we find that if N satisfies the condition $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$ then the second term is negligible and we obtain

$$\Pi_d(1_A, \dots, 1_A) \leq \frac{2}{2^d + 1} \alpha^{\frac{d(d+1)}{2}} \left(\frac{N}{N'}\right)^d.$$

Note that the bound we chose for N is not optimal, but it has the advantage of not being too complicated and does not change the final bound in Theorem 1.2.

Therefore the sum of the $2^{d(d+1)/2} - 1$ other terms involving f_A must have magnitude at least

$$\left(\frac{1}{2^{d-1}} - \frac{2}{2^d + 1} \right) \alpha^{\frac{d(d+1)}{2}} \left(\frac{N}{N'} \right)^d.$$

Since $N' \leq 3N$, one of those terms must have magnitude larger than

$$\frac{1}{3^d \times 2^{\frac{d(d+1)}{2} + 2d}} \alpha^{\frac{d(d+1)}{2}}.$$

To avoid heavy expressions, we will say that $|\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})| \geq \left(\frac{\alpha}{C}\right)^{\frac{d(d+1)}{2}}$ for some absolute constant $C > 0$. \square

Now we will use Lemma 2.4 and Theorem 2.3 to establish the following corollary.

Corollary 2.5. *Let α , $0 < \alpha < 1$, be a real number. Suppose that $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$ and that A is a subset of $[N]$ with $|A| = \alpha N$ containing fewer than $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ non-trivial d -configurations. Let $f_A : \mathbb{Z}/N'\mathbb{Z} \rightarrow \mathbb{R}$ be the balanced function of A . Then*

$$\|f_A\|_{U^2} \geq \left(\frac{\alpha}{C}\right)^{\frac{d(d+1)}{2}}$$

for some constant $C > 0$.

Proof: Let $\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})$ be the same term as in Lemma 2.4.

By Theorem 2.3, we have

$$\begin{aligned} |\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})| &= |\mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} g_{ij}(n_i + n_j) + \mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} g_{ij}(n_i + n_j + 1)| \\ &\leq |\mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} g_{ij}(n_i + n_j)| + |\mathbb{E}_{n_1, \dots, n_d \in \mathbb{Z}/N'\mathbb{Z}} \prod_{1 \leq i \leq j \leq d} g_{ij}(n_i + n_j + 1)| \\ &\leq 2 \min_{1 \leq i \leq j \leq d} \|g_{ij}\|_{U^2} \leq 2 \|f_A\|_{U^2}. \end{aligned}$$

And by Lemma 2.4, we have

$$|\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})| \geq \left(\frac{\alpha}{C}\right)^{\frac{d(d+1)}{2}}$$

for some constant $C > 0$.

Therefore

$$2 \|f_A\|_{U^2} \geq |\Pi_d((g_{ij})_{1 \leq i \leq j \leq d})| \geq \left(\frac{\alpha}{C}\right)^{\frac{d(d+1)}{2}}.$$

\square

2.3 Inverse results for the Gowers U^2 -norm

In this subsection, we use the fact that f_A has a large Gowers U^2 -norm to show that it also has a large Fourier coefficient.

Let f be a function from $\mathbb{Z}/N\mathbb{Z}$ to \mathbb{C} . The *Fourier transform* \hat{f} of f is defined, for all $\xi \in \mathbb{Z}/N\mathbb{Z}$, by the formula:

$$\hat{f}(\xi) := \mathbb{E}_{x \in \mathbb{Z}/N\mathbb{Z}} f(x) e(-\xi x/N).$$

Let us recall the following theorem proved in [6].

Theorem 2.6. *Suppose that $f : \mathbb{Z}/N'\mathbb{Z} \rightarrow \mathbb{C}$ is a 1-bounded function with $\|f\|_{U^2} \geq \delta$. Then there is some $r \in \mathbb{Z}/N'\mathbb{Z}$ such that*

$$|\hat{f}(r)| \geq \delta^2.$$

By Corollary 2.5 and Theorem 2.6, we have the following result.

Proposition 2.7. *Let α , $0 < \alpha < 1$, be a real number. Suppose that $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$ and that A is a subset of $[N]$ with $|A| = \alpha N$ containing fewer than $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ non-trivial d -configurations. Let $f_A : \mathbb{Z}/N'\mathbb{Z} \rightarrow \mathbb{R}$ be the balanced function of A . Then there is some $r \in \mathbb{Z}/N'\mathbb{Z}$ and some constant $C > 0$ such that*

$$|\hat{f}_A(r)| \geq \left(\frac{\alpha}{C}\right)^{d(d+1)}.$$

2.4 Obtaining a density increment

In this subsection, we will use the fact that f_A has a large Fourier coefficient to find a density increment on some arithmetic progression and thus complete the proof of Proposition 2.10.

Since f is supported on $[N]$ and $N' = 2N + 1$, Proposition 2.7 immediately implies the next proposition.

Proposition 2.8. *Let α , $0 < \alpha < 1$, be a real number. Suppose that $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$ and that A is a subset of $[N]$ with $|A| = \alpha N$ containing fewer than $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ non-trivial d -configurations. Let f_A be the balanced function of A , considered now as a function on $[N]$. Then there is some $\theta \in [0, 1]$ and some constant $C > 0$ such that*

$$\left| \sum_{x \in [N]} f_A(x) e(\theta x) \right| \geq \left(\frac{\alpha}{C}\right)^{d(d+1)} N. \quad (2.1)$$

Let us recall a lemma proved in [13], adapted from Roth's original argument [8].

Lemma 2.9. *Let $f : \mathbb{Z} \rightarrow \mathbb{R}$ be a function supported on $[N]$ such that $|f(n)| \leq 1$ for all n , $\sum_n f(n) = 0$ and*

$$|\mathbb{E}_{x \in [N]} f(x) e(\theta x)| \geq \sigma$$

for some $\theta \in [0, 1]$ and $\sigma > 0$. Then there exists a non-trivial arithmetic progression $P \subseteq [N]$ with $|P| \geq c\sigma^2 \sqrt{N}$ and

$$\mathbb{E}_{x \in P} f(x) \geq \frac{\sigma}{4}.$$

We have $-\alpha \leq f_A(n) \leq 1 - \alpha$ for all n , so $|f_A(n)| \leq 1$ for all n . We also have $\sum_n f_A(n) = 0$. Therefore, using the conclusion of Proposition 2.8, we can apply Lemma 2.9 with $f = f_A$ and $\sigma = \left(\frac{\alpha}{C}\right)^{d(d+1)}$. We obtain an arithmetic progression $P \subseteq [N]$ of length $|P| \geq \left(\frac{\alpha}{C}\right)^{2d(d+1)} \sqrt{N}$ and

$$\mathbb{E}_{x \in P} f_A(x) \geq \frac{1}{4} \left(\frac{\alpha}{C}\right)^{d(d+1)} \geq \left(\frac{\alpha}{C'}\right)^{d(d+1)},$$

which means that

$$\frac{|A \cap P|}{|P|} \geq \alpha + \left(\frac{\alpha}{C'}\right)^{d(d+1)}.$$

We obtain the following proposition.

Proposition 2.10. *Suppose that $0 < \alpha < 1$ and $N > 16^d \alpha^{-\frac{d(d+1)}{2}}$. Suppose that $P \subseteq \mathbb{Z}$ is an arithmetic progression of length N and that $A \subseteq P$ is a set with cardinality αN . Then one of the following two alternatives holds:*

- A contains at least $\frac{\alpha^{d(d+1)/2} N^d}{2^{d-1} + 1}$ d -configurations;
- There is an arithmetic progression P' of length $\geq \left(\frac{\alpha}{C}\right)^{2d(d+1)} N^{1/2}$ such that, writing $A' := A \cap P'$ and $\alpha' := |A'|/|P'|$, we have $\alpha' > \alpha + \left(\frac{\alpha}{C}\right)^{d(d+1)}$ for some absolute constant $C > 0$.

2.5 The final bound

Finally, using Proposition 2.10, we obtain the following quantitative version of Theorem 1.2:

Theorem 2.11. *There is an absolute constant C such that any subset $A \subseteq [N]$ with cardinality at least $\frac{CN}{(\log \log N)^{\frac{1}{d(d+1)-1}}}$ contains a non-trivial d -configuration.*

Proof: Set $P_0 := [N]$ and assume that we have a set $A \subseteq P_0$ with $|A| = \alpha N$ containing no non-trivial d -configuration. Then we attempt to use Proposition 2.10 repeatedly to obtain a sequence P_0, P_1, \dots of progressions and sets $A_i := A \cap P_i$. The densities $\alpha_i := |A_i|/|P_i|$ will satisfy $\alpha_{i+1} > \alpha_i + \left(\frac{\alpha_i}{C}\right)^{d(d+1)}$ and the length of P_i will be

$$N_i \geq \left(\frac{\alpha_{i-1}}{C}\right)^{2d(d+1)} N_{i-1}^{1/2} \geq \left(\frac{\alpha}{C}\right)^{2d(d+1) \sum_{k=0}^{i-1} \left(\frac{1}{2}\right)^k} N^{(1/2)^i} \geq \left(\frac{\alpha}{C}\right)^{4d(d+1)} N^{(1/2)^i},$$

using the fact that $\alpha_i \geq \alpha$ and $\sum_{k=0}^{i-1} \left(\frac{1}{2}\right)^k \leq 2$ for all $i \geq 1$.

But this iteration cannot last too long, otherwise we would obtain a set A_i with density more than 1 over the arithmetic progression P_i for some i , which is impossible. In particular there cannot be more than $\left(\frac{C'}{\alpha}\right)^{d(d+1)-1}$ steps in total. We conclude that the applications of Proposition 2.10 must have been invalid, which means that the condition $N_i > 16^d \alpha_i^{-\frac{d(d+1)}{2}}$ was violated. Since

$$N_i \geq \left(\frac{\alpha}{C}\right)^{4d(d+1)} N^{(1/2)^i} \left(\frac{C'}{\alpha}\right)^{d(d+1)-1}$$

and $\alpha_i \geq \alpha$, we infer the bound

$$16^d \alpha^{-\frac{d(d+1)}{2}} \geq \left(\frac{\alpha}{C}\right)^{4d(d+1)} N^{(1/2)^i} \left(\frac{C'}{\alpha}\right)^{d(d+1)-1}.$$

Rearranging leads to

$$\begin{aligned} \log \log N &\leq \log \left(\log \left(16^d \alpha^{-\frac{d(d+1)}{2}} \right) + 4d(d+1) \log \left(\frac{C}{\alpha} \right) \right) + \log 2 \left(\frac{C'}{\alpha} \right)^{d(d+1)-1} \\ &\leq \left(\frac{C''}{\alpha} \right)^{d(d+1)-1}, \end{aligned}$$

for some constant C'' .

Therefore if $\alpha \geq \frac{C''}{(\log \log N)^{\frac{1}{d(d+1)-1}}}$ (ie. $N \geq \exp \left(\exp \left(\left(\frac{C''}{\alpha} \right)^{d(d+1)-1} \right) \right)$), A contains a non-trivial d -configuration. Theorem 2.11 is proved. \square

The bound obtained with this proof is better than the one using Szemerédi's theorem because we only have 3 exponentials instead of 5.

3 The improvement in Sudakov, Szemerédi and Vu's theorem

Now that we proved Theorem 2.11, we will use it to improve Sudakov, Szemerédi and Vu's theorem about sum-free subsets [10].

Let $\phi(A)$ denote the maximum cardinality of a subset of A which is sum-free with respect to A . Let $\phi(n)$ be the minimum of $\phi(A)$ over all sets A of n integers. Sudakov, Szemerédi and Vu obtained the first superlogarithmic lower bound for $\phi(n)$ by proving the following theorem [10].

Theorem 3.1. *There is a function $g(n)$ tending to infinity with n such that the following holds. Any set A of n integers contains a subset B with cardinality $g(n) \log n$ such that B is sum-free with respect to A .*

Their proof shows that we can take $g(n)$ to be of the order $(\log^{(5)} n)^{1-o(1)}$. Thus they proved that

$$\phi(n) \geq \log n \left(\log^{(5)} n \right)^{1-o(1)}.$$

It is easier to describe $g(n)$ as the inverse of an iterative exponential function. They set $g(n)$ in Theorem 3.1 to be $c(m/\log m)$, where c is a sufficiently small positive constant, $m = F^{-1}(n^{1/2})$ and

$$F(h) = \exp \left(h^{182} \times \left(2 \uparrow \left(e^{h^{32770}} \right) \uparrow 2 \uparrow 2 \uparrow (2h + 9) \right) \right).$$

We will show that, using Theorem 2.11, we can set $g(n)$ to be $c(m/\log m)$, where $m = G^{-1}(n^{1/2})$ and

$$G(h) = \exp \left(h^{182} \times \left(e \uparrow \left(ce^{h^{32770}} \right) \uparrow (h(h + 1) - 1) \right) \right).$$

In their proof, Sudakov, Szemerédi and Vu deduce Theorem 3.1 from the following theorem.

Theorem 3.2. *Let X, Y be two finite sets of positive integers with $\frac{1}{h^{29}}|Y| \geq |X| \geq F(h)$, where F is the function described above and h is a sufficiently large integer. Then Y contains a subset Z of size h which is disjoint from X and is sum-free with respect to $X \cup Y$.*

We will show that a modification of the proof of Theorem 3.2 allows us to replace F by G and obtain the following theorem.

Theorem 3.3. *Let X, Y be two finite sets of positive integers with $\frac{1}{h^{29}}|Y| \geq |X| \geq G(h)$, where G is the function described above and h is a sufficiently large integer. Then Y contains a subset Z of size h which is disjoint from X and is sum-free with respect to $X \cup Y$.*

In particular, the proof of Theorem 3.2 uses the following corollary, proved using Szemerédi's theorem.

Corollary 3.4. *If $A \subseteq \{1, \dots, N\}$ is a set of size αN , and if $N > 2 \uparrow 2 \uparrow (\alpha)^{-1} \uparrow 2 \uparrow 2 \uparrow (2k + 9)$, then there is a subset $A' \subseteq A$ of k elements such that, for any two elements $x, y \in A'$, there is an element z of A satisfying $x + y = 2z$.*

But by using Theorem 2.11, we can replace Corollary 3.4 by the following corollary, which comes with a better bound.

Corollary 3.5. *If $A \subseteq \{1, \dots, N\}$ is a set of size αN , and if $N > e \uparrow e \uparrow \left(\frac{C}{\alpha}\right) \uparrow (k(k + 1) - 1)$, then there is a subset $A' \subseteq A$ of k elements such that, for any two elements $x, y \in A'$, there is an element z of A satisfying $x + y = 2z$.*

Proof: By Theorem 2.11, A contains a non-trivial k -configuration $(n_i + n_j + a)_{1 \leq i \leq j \leq k}$. Now we can take $A' = \{2n_1 + a, \dots, 2n_k + a\}$, and so for any two elements $2n_i + a$ and $2n_j + a \in A'$, $(2n_i + a) + (2n_j + a) = 2(n_i + n_j + a)$ with $n_i + n_j + a \in A$. \square

To prove Theorem 3.3 we leave all Sudakov, Szemerédi and Vu’s proof of Theorem 3.2 unchanged, except at the end, where we replace their bound by ours. We advise the reader to read the proof in [10], because we won’t copy the beginning of the proof here. Thus when they need $m_1 \geq 2 \uparrow 2 \uparrow \left(e^{h^{32770}} \right) \uparrow 2 \uparrow 2 \uparrow (2h + 9)$, we only need $m_1 \geq e \uparrow e \uparrow \left(C e^{h^{32770}} \right) \uparrow (h(h + 1) - 1)$. As in their proof, we have $\log m_1 \geq \frac{\log |Y|}{h^{182}}$. Then we only need to verify that $\log |Y| \geq h^{182} \times \left(e \uparrow \left(C e^{h^{32770}} \right) \uparrow (h(h + 1) - 1) \right)$. Since the right-hand side is equal to $\log G(h)$, this inequality follows from the assumption of Theorem 3.3 that $|Y| \geq |X| \geq G(h)$. This completes the proof.

Now we can derive Theorem 3.1 from Theorem 3.3 in the same way as in [10], except that we replace F by G .

After some calculation, we obtain that we can take $g(n)$ to be of the order $(\log^{(3)} n)^{\frac{1}{32772} - o(1)}$. Therefore we obtain the following stronger version of Theorem 3.1.

Theorem 3.6. *There is a function $g(n)$ of the order $(\log^{(3)} n)^{\frac{1}{32772} - o(1)}$ such that any set A of n integers contains a subset B with cardinality $g(n) \log n$ such that B is sum-free with respect to A .*

4 Conclusion

We generalised Roth’s theorem to d -configurations and showed that any set $A \subseteq \{1, \dots, N\}$ with density α such that $N > e \uparrow e \uparrow \left(\frac{C}{\alpha} \right) \uparrow (d(d + 1) - 1)$ contains a non-trivial d -configuration. Then we used this result to improve Sudakov, Szemerédi and Vu’s theorem about sum-free subsets and proved that $\phi(n) \geq \log n \left(\log^{(3)} n \right)^{\frac{1}{32772} - o(1)}$, which is the best lower bound known to date for $\phi(n)$.

Bourgain [1] modified Roth’s original Fourier analytic proof [8] of Roth’s theorem by increasing the density of A on Bohr sets instead of arithmetic progressions. By doing so, he improved Roth’s bound $N \geq \exp \left(\exp \left(\frac{C}{\alpha} \right) \right)$ and showed that $N \geq \left(\frac{C}{\alpha} \right)^{C'/\alpha^2}$ suffices. Therefore it should be possible to do a similar modification to our proof in order to obtain a stronger version of Theorem 1.2 with a bound of the type $N \geq \left(\frac{c(d)}{\alpha} \right)^{\left(\frac{c'(d)}{\alpha} \right)^{d(d-1)}}$. This would hopefully improve our result about sum-free subsets and lead to a bound of the form $\phi(n) \geq \log n (\log^{(2)} n)^{c - o(1)}$. Even if the technical details of such a proof might be considerable, it would constitute an interesting subject for further research. However this lower bound is still far from the best upper bound currently known, $\phi(n) \leq O(e^{\sqrt{\log n}})$, proved by Ruzsa in [9], so we can assume that many interesting results about sum-free subsets are still to be found.

Acknowledgements

The author would like to thank Ben Green for introducing her to this very interesting subject and for the precious advice he gave her during the elaboration of this work. She also thanks Fernando Shao for drawing her attention to a calculation mistake in an early version of this article. She would like to express her gratitude to the referee for a careful reading of this paper and for insightful comments.

References

- [1] J. BOURGAIN, *On triples in arithmetic progression*, Geometric And Functional Analysis, 9 (1999), pp. 968–984.
- [2] S. L. G. CHOI, *On a combinatorial problem in number theory*, Proc. London Math. Soc, 3 (1971), pp. 629–642.

- [3] P. ERDŐS, *Extremal problems in number theory*, in Proceedings of Symposia in Pure Mathematics, vol. 8, American Mathematical Society, 1965, pp. 181–189.
- [4] W. T. GOWERS, *A new proof of Szemerédi's theorem*, Geometric And Functional Analysis, 11 (2001), pp. 465–588.
- [5] B. GREEN AND T. TAO, *Linear equations in primes*, Ann. of Math., 171 (2010), pp. 1753–1850.
- [6] B. J. GREEN, *Roth's theorem on progressions of length 3*. Course notes.
- [7] D. R. HEATH-BROWN, *Integer sets containing no arithmetic progressions*, J. London Math. Soc., 35 (1987), pp. 385–394.
- [8] K. ROTH, *On certain sets of integers*, J. London Math. Soc., 28 (1953), pp. 104–109.
- [9] I. Z. RUZSA, *Sum-avoiding subsets*, The Ramanujan Journal, 9 (2005), pp. 77–82.
- [10] B. SUDAKOV, E. SZEMERÉDI, AND V. VU, *On a question of Erdős and Moser*, Duke Math. J., 129 (2005), pp. 129–155.
- [11] E. SZEMERÉDI, *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica, 27 (1975), pp. 199–245.
- [12] ———, *Integer sets containing no arithmetic progressions*, Acta Math. Hungar., 56 (1990), pp. 155–158.
- [13] T. TAO AND V. H. VU, *Additive Combinatorics*, Cambridge University Press, 2010.