Arithmetic Structures in Random Sets

by

Mariah Hamel

B.A., Colby College, 2002
M.Sc., The University of British Columbia, 2004

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

The Faculty of Graduate Studies

(Mathematics)

The University of British Columbia

June 2008

Abstract

We prove various results in additive combinatorics for subsets of random sets. In particular we extend Sárközy's theorem and a theorem of Green on long arithmetic progressions in sumsets to dense subsets of random sets with asymptotic density 0. Our proofs require a transference argument due to Green and Green-Tao which enables us to apply known results for sets of positive upper density to subsets of random sets which have positive relative density. We also prove a density result which states that if a subset of a random set has positive relative density, then the sumset of the subset must have positive upper density in the integers.

Table of Contents

Acknowledgements

I want to thank my advisor Izabella Łaba for her continued support and guidance throughout my graduate work. I am especially thankful for our many meetings filled with mathematical insight, encouragement and thoughtful advice.

I would like to thank Jozsef Solymosi, David Boyd and Malabika Pramanik for many useful discussions and for offering such an interesting array of graduate courses. I am grateful to David Brydges for always being so kind and for always inquiring about my progress as a graduate student. I would like to thank Greg Martin and Stephen Kwok-Kwong Choi for being members of my examining committee and for thoughtful suggestions for this thesis. I have enjoyed discussions with many of my fellow graduate students including Karsten Chipeniuk, Chris Coulter, Chris Duarte, Zhai Kelan and Terry Soo.

I am deeply grateful to have begun studying mathematics under the supervision of Ben Mathes, Leo Livshits and Fernando Gouvêa at Colby College. I will forever remember their encouragement and their love of mathematics.

I am indepted to my parents, Ed and Mary Hamel, for having supported me throughout my education.

Finally, I would like to thank Mathew Rogers for his support, encouragement and love.

Dedication

For Carina

Statement of Co-Authorship

The manuscript which is contained in this thesis is joint work with Izabella Łaba. It was suggested to us that Theorem 2.1.2 was feasible Ben Green. The statement of Theorem 2.1.4 was first considered by Łaba. The manuscript was researched and jointly written by Mariah Hamel and Izabella Łaba.

Chapter 1

Introduction

1.1 History

Additive combinatorics can be described as a study of the structural properties of sets of integers, or more generally, additive groups. While this area has long been of interest to mathematicians, over the past few years there has been an abundance of collaboration between those who study number theory, harmonic analysis, combinatorics and ergodic theory. In this introduction we will provide motivation to the work included below through a discussion of the history of additive combinatorics as well as some (difficult) conjectures which remain open.

One of the first results in the subject which is today called additive combinatorics is a coloring result due to van der Waerden proved in 1927. He was able to show that given integers r and k, if the set of integers is colored by r colors, then there must be k integers in arithmetic progression which are monochromatic. A related result, due to Schur states that if the integers are r-colored, then there must be x, y and z, monochromatic, such that $x + y = z$. Both of these problems fall into the category of Ramsey theory which is concerned with extremal results. We can rephrase them as follows: Suppose $\{1, ..., n\}$ is colored with r colors. How large must n be (as a function of r) to guarantee the desired monochromatic structure?

In 1936 Erdös and Turán formulated the following conjecture:

Conjecture 1.1.1. *Suppose $A \subset \mathbb{N}$ such that*

$$\sum_{a \in A} \frac{1}{a} = \infty.$$

Then A must contain arithmetic progressions of arbitrary length.

We remark that this can be seen as an extension of the theorem of van der Waerden since if we color the integers with finitely many colors, then certainly one of the color classes will satisfy the hypothesis of Conjecture 1.1.1. However, no such generalization of Schur's theorem exists since, for example, the set of odd integers also satisfies the hypothesis of the Conjecture but there are no three odd numbers which satisfy $x + y = z$.

While Conjecture 1.1.1 remains open, there have been many interesting results in that direction. We say that a subset A of the integers has *positive upper density* if

$$\limsup_{n\to\infty} \frac{|A \cap \{1, ..., n\}|}{n} > 0.$$

In 1953, Roth [13] proved that any subset of the integers with positive upper density must contain arithmetic progressions of length three. A few years later, building on Roth's result, Varnavides [17] showed that, in fact, such a subset must contain many three-term arithmetic progressions. While we will not elaborate on the details of proof here, we note that the reader can compare the proof of Roth's theorem to that of Sárközy's theorem on square differences included in Section 1.4. Similarly the extension of Sárközy's theorem contained in Chapter 2 follows Varnavides' original argument.

In 1975, Szemerédi extended Roth's theorem for arithmetic progressions of arbitrary length. Specifically, he proved the following:

Theorem 1.1.2. Szemerédi's Theorem. *Let k be a positive integer and let $\delta \geq 0$. Then there exists $N = N(k, \delta)$, such that every subset A of $\{1, ..., N\}$ such that $|A| \geq \delta N$ must contain an arithmetic progression of length k.*

Motivated by Szemerédi's work, Furstenberg [4] provided an ergodic proof of the same theorem. More recently, in 2001, Gowers [5] approached the problem with a proof that can be seen as a generalization of Roth's original argument for three term arithmetic progressions. Many of the ideas developed by Gowers were used by Green and Tao [8] in proving that the primes contain arithmetic progressions of arbitrary length.

Inverse problems form another important topic in additive number theory. Suppose that $A \subset \{1, ..., N\}$. Freiman's theorem [3] states that if the size of the sumset $A + A := \{a_1 + a_2 : a_1, a_2 \in A\}$ is small (relative to the size of A) then A itself must be very structured. To state Freiman's theorem precisely, we require a the following definition:

Definition 1.1.3. *Suppose that $x_0, x_1, ..., x_d$ are integers and $m_1, m_2, ..., m_d$ are positive integers. A generalized arithmetic progression of dimension d is defined*

$$P := \{x_0 + \sum_{i=1}^{d} \lambda_j x_j : 0 \leq \lambda_j \leq m_j - 1\}.$$

If the size of the progression is equal to the product of the m_is, we say that P is proper.

An illustrative example of a generalized arithmetic progression is defined by the set $P := 6 + \{0, 3, 6, 9\} + \{0, 100, 200, 300\}$. Freiman's theorem then states:

Theorem 1.1.4. *(Freiman-Ruzsa-Chang) Suppose $A \subset \mathbb{Z}$ and $|A| = N$ and $|A + A| \leq C|A|$. Then there exists a proper d-dimensional arithmetic progression P such that $A \subset P$ and $|P| \leq C_1 N$ where $d \leq [C - 1]$ and $\log(|P|/|A|) \ll C^2(\log C)^3$.*

Despite many works on the above subjects there remain many open problems. Even the case of three term arithmetic progressions has not been satisfactorily solved. In terms of an upper bound, the best known result is due to Bourgain [2] in which he shows a set of density $\delta \gg \frac{(\log \log N)^2}{(\log N)^{2/3}}$ (assuming N is sufficiently large) must contain three term arithmetic progressions. On the other hand, the best lower bound due to Behrend [1] shows that there exists a subset of$\{1, ..., N\}$ of size $N^{1-\frac{\sqrt{2\log 2}}{\sqrt{\log N}}+\frac{\epsilon}{\sqrt{\log N}}}$ which contains no three term arithmetic progressions.

In the case of Roth's theorem, there have been some interesting results for sets which do not satisfy Bourgain's bounds. In 1996, Kohayakawa, Luczak and Rödl [12] proved a version of Roth's theorem for sets of asymptotic density zero which satisfy the additional hypothesis of positive relative density in a random set which has density zero in the integers. In 2002, Green [7] proved a version of Roth's theorem for sets which have positive upper density in the primes.

1.1.1 Statement of main results

The purpose of this section is to state the main results contained in this thesis. Theorems 1.1.5 and 1.1.6 are contained in Chapter 2 and can be found as Theorem 2.1.2 and Theorem 2.1.4 respectively. The proof of Theorem 1.1.5 can be found in Section 2.4. The proof of Theorem 2.1.4 can be found in Sections 2.6, 2.7 and 2.8. The purpose of the remainder of this chapter is to provide an introduction to the ideas used in Chapter 2 and to fill in certain details that were omitted from [11].

In 1978 Sárközy [15] proved a variant on Roth's theorem showing that subsets of positive upper density must contain two elements whose difference is a perfect square. The same result was proved independently by Furstenberg ([4]) using ergodic theory around the same time. In 1988, Pintz, Steiger and Szemerédi, using a combination of Fourier analytic and combinatorial arguments, showed that any subset of $\{1, ..., N\}$ of size at least $(\log N)^{-c \log\log\log\log N} N$ must contain a square difference. On the other hand, a construction of Ruzsa shows that there exists a set of size $N^{1-0.267}$ which contains no square difference. It is conjectured that for any $\epsilon > 0$, and N sufficiently large (depending on ϵ), there exists a set of size $N^{1-\epsilon}$ which contains no square difference.

In a joint paper with Izabella Łaba we prove the following variant of Sárközy's theorem.

Theorem 1.1.5. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (cN^{-\theta}, 1]$ where $0 < \theta < 1/110$. Let $\alpha > 0$. Then the statement*

> *for every set $A \subset W$ with $|A| \geq \alpha W$, there are $x, y \in A$ such that $x - y$ is a non-zero perfect square*

is true with probability $1 - o_\alpha(1)$ as $N \to \infty$.

If A is a subset of the integers, it is known that the sumset $A + A = \{a_1 + a_2 : a_1, a_2 \in A\}$ contains 'more' additive structure than the original set. The case when A is a subset of positive upper density has been studied by Bourgain, Ruzsa, Green and Sanders. For a more detailed discussion of the history of this problem we direct the reader to the introduction of [11] included below. We prove the following analogue for subsets of random sets.

Theorem 1.1.6. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (CN^{-\theta}, 1]$, where $0 < \theta < 1/140$. Assume that α and k obey*

$$\alpha \geq \frac{C_1 \log\log N}{\sqrt{\log N}}, \tag{1.1.1}$$

$$k \leq \exp\left(\frac{\alpha^2 \log\log N}{C_2 \log\frac{1}{\alpha}(\log\log\log N + \log\frac{1}{\alpha})}\right), \tag{1.1.2}$$

where C_1, C_2 are sufficiently large constants. Then the statement

> *for every set $A \subset W$ with $|A| \geq \alpha|W|$, the sumset $A+A$ contains a k-term arithmetic progression*

is true with probability $1 - o_{k,\alpha}(1)$ as $N \to \infty$.

1.2 Preliminaries and Notation

Throughout this paper, we will be concerned with subsets of the integers or subsets of the additive group $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$. We define $e(\theta) := \exp(2\pi i\theta)$. We often use the notation $A(x)$ to denote the characteristic function of the set A. Suppose that f and g are real-valued functions with $f(x) \geq 0$ for all x. We write $g(x) = O(f(x))$ if $|g(x)| \leq Cf(x)$ for some constant $C > 0$. If $f(x) > 0$ we write $g(x) = o(f(x))$ if $\lim_{x\to\infty} g(x)/f(x) = 0$. Finally, if there exists C_1 and C_2 such that $C_1 f(x) \leq g(x) \leq C_2 f(x)$ we write $g(x) = \Theta(f(x))$.

In this setting, we begin with some preliminary definitions and lemmas which will be used throughout.

Definition 1.2.1. *Suppose that $f : \mathbb{Z}_N \to \mathbb{C}$. Then we define the discrete Fourier transform*

$$\widehat{f}(\xi) := \frac{1}{N} \sum_{x\in\mathbb{Z}_N} f(x)e(-x\xi/N).$$

Definition 1.2.2. *If f and g are functions such that $f, g : \mathbb{Z}_N \to \mathbb{C}$, then we define their convolution*

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

Lemma 1.2.3. *Suppose that f and g are functions such that $f, g : \mathbb{Z}_N \to \mathbb{C}$. Then the following identities hold:*

(i) Parseval's identity:

$$\sum_{\xi\in\mathbb{Z}_N} \widehat{f}(\xi)\overline{\widehat{g}(\xi)} = \frac{1}{N} \sum_{x\in\mathbb{Z}_N} f(x)\overline{g(x)},$$

(ii) Fourier inversion:

$$f(x) = \sum_{\xi\in\mathbb{Z}_N} \widehat{f}(\xi)e(-x\xi/N),$$

(iii) Convolution identity:

$$N\widehat{f}(\xi)\widehat{g}(\xi) = \widehat{f * g}(\xi).$$

We will denote the *L^p-norm* of f by

$$\|f\|_{L^p(X)} := (\sum_{x\in X} |f(x)|^p)^{1/p}$$

and if $X = \mathbb{Z}_N$ we will write $\|f\|_p$. We will say that the *probability* of a set A is

$$\mathbb{P}(A) := \frac{|A|}{N}.$$

We define the related *expectation* for a function $f : \mathbb{Z}_N \to \mathbb{C}$ to be

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

It will also be useful to sometimes use the *conditional probability* and *conditional expectation*, defined respectively,

$$\mathbb{P}(A|X) := \frac{|A \cap X|}{|X|}, \mathbb{E}(f|X) := \frac{1}{|X|} \sum_{x\in X} f(x).$$

1.3 Bohr Sets

In this section, we include some background relating to Bohr sets. In particular, we provide a detailed exposition of the results required in [11], including a localized version of a theorem of Chang which is developed in [14].

One of the first uses of Bohr sets in additive number theory came in a paper of Bourgain [2] in which he improved the upper bound on $r_3(N)$ (the function which denotes the size of the largest subset of $\{1, ..., N\}$ which contains no three term arithmetic progression). His proof relies on a density increment argument, however instead of iterating on arithmetic progressions, he increases the relative density of a subset in consecutive Bohr sets. As we will see, one can show that Bohr sets contain long arithmetic progressions.

Definition 1.3.1. *A Bohr set is a set of the form $B = b + B(\Lambda, \delta)$ where $b \in \mathbb{Z}_N$, $\Lambda \subset \mathbb{Z}_N$, $\delta \in (0, 2)$ and*

$$B(\Lambda, \delta) := \{x \in \mathbb{Z}_N : |e(x\xi/N) - 1| \leq \delta \textit{ for all } \xi \in \Lambda\}.$$

We will say that the Bohr set B is of rank $d := |\Lambda|$ and of radius δ.

In the following lemma, we state two basic facts about Bohr sets that we will use throughout.

Lemma 1.3.2. *Suppose $\Lambda \subset \mathbb{Z}_N$ is a set of frequencies, and $\delta, \delta_1, \delta_2 \in (0, 2)$. Then the following are true,*

$$B(\Lambda, \delta_1) + B(\Lambda, \delta_2) \subset B(\Lambda, \delta_1 + \delta_2) \tag{1.3.3}$$

$$|B(\Lambda, \delta)| \leq C^{|\Lambda|} |B(\Lambda, \delta/2)| \tag{1.3.4}$$

for some absolute constant $C > 1$.

Definition 1.3.3. *Let $c_0 \in (0, 1/2)$ be a small parameter. We will say that a Bohr set $B(\Lambda, \delta)$ is c_0-regular if*

$$\mathbb{P}(B(\Lambda, (1 + c_0^2)\delta) \setminus B(\Lambda, (1 - c_0^2)\delta)) \leq c_0 \mathbb{P}(B(\Lambda, \delta)).$$

We will also say that $B = b + B(\Lambda, \delta)$ is regular if $B(\Lambda, \delta)$ is regular.

If B is a c_0-regular Bohr set, then we have the following size bound.

Lemma 1.3.4. *Let $B(\Lambda, \delta)$ be a c_0-regular Bohr set. Then*

$$|B(\Lambda, \delta)| \geq (c^{-1} c_0^2 \delta)^{|\Lambda|}.$$

Proof. We first notice that $B(\Lambda, \delta) \subset B(\Lambda, (1 + c_0^2)\delta)$. Hence, by regularity, we have

$$\begin{aligned} |B(\Lambda, \delta) \backslash B(\Lambda, (1 - c_0^2)\delta)| &\leq |B(\Lambda, (1 + c_0^2)\delta) \backslash B(\Lambda, (1 - c_0^2)\delta)| \\ &< c_0 |B(\Lambda, \delta)|. \end{aligned}$$

Hence we must have that $B(\Lambda, (1 - c_0^2)\delta)$ is nonempty. Say $b \in B(\Lambda, (1 - c_0^2)\delta)$. Now by Property 1.3.3 we have

$$B(\Lambda, c_0^2\delta) + B(\Lambda, (1 - c_0^2)\delta) \subset B(\Lambda, \delta)$$

and hence

$$B(\Lambda, c_0^2\delta) + b \subset B(\Lambda, \delta).$$

Finally using Property 1.3.4 we have $|B(\Lambda, c_0^2\delta)| \geq (c^{-1} c_0^2 \delta)^{|\Lambda|}$ as desired. □

Lemma 1.3.5. *Assume c_0 is small enough. Then for any $\Lambda \subset \mathbb{Z}_N$ with $|\Lambda| \leq \sqrt{c_0} N$ and for any $\delta_0 > 0$ there exists $\delta \in (\delta_0/2, \delta_0)$ so that $B(\Lambda, \delta)$ is c_0-regular.*

Proof. We pick radii $\delta_0 > \delta_1 > ... > \delta_{k+1} := \delta_0/2$ where $k = \Theta(N)$ such that $\delta_{i+1} \leq (1-c_0^2)\delta_i \leq (1+c_0^2)\delta_i \leq \delta_{i-1}$. By Property 1.3.4 we have

$$|B(\Lambda, \delta_0)| \leq C^{|\Lambda|}|B(\Lambda, \delta_0/2)|.$$

Hence, by the pigeonhole principle there exists $1 \leq i \leq k$ such that

$$|B(\Lambda, \delta_{i-1})| \leq C^{2|\Lambda|/k}|B(\Lambda, \delta_{i+1})|. \tag{1.3.5}$$

Using this, and the fact that $B(\Lambda, \delta_{i+1}) \subset B(\Lambda, \delta_{i-1})$, we have

$$\begin{aligned}|B(\Lambda, (1+c_0^2)\delta_i)| &\leq |B(\Lambda, \delta_{i-1})| \\ &\leq C^{2|\Lambda|/k}|B(\Lambda, \delta_{i+1})| \\ &\leq C^{2|\Lambda|/k}|B(\Lambda, (1-c_0^2)\delta_i)|.\end{aligned}$$

Since $B(\Lambda, (1-c_0^2)\delta_i) \subset B(\Lambda, (1+c_0^2)\delta_i)$, we must have

$$\begin{aligned}|B(\Lambda, (1+c_0^2)\delta_i)\backslash B(\Lambda, (1-c_0^2)\delta_i)| &\leq |B(\Lambda, (1+c_0^2)\delta_i)| - \frac{1}{C^{2|\Lambda|/k}}|B(\Lambda, (1+c_0^2)\delta_i)| \\ &\leq \frac{(C^{2|\Lambda|/k}-1)}{C^{2|\Lambda|/k}}|B(\Lambda, (1+c_0^2)\delta_i)|.\end{aligned}$$

Using the assumptions on $|\Lambda|$ and k we have $C^{2|\Lambda|/k} = 1 + O(\sqrt{c_0})$ which gives the desired result. □

1.3.1 A localized version of Chang's theorem

In this section we would like to explain a localized version of Chang's structure theorem due to Sanders. We begin with dissociated sets and the statement of Chang's theorem. We also include a description of Sanders' result, its relation to the work of Chang and other ingredients of its proof.

Definition 1.3.6. *We say that the set $\Lambda := \{\lambda_1, ..., \lambda_k\}$ is dissociated if*

$$\sum_{i=1}^{k} c_i\lambda_i = 0,$$

where $c_i \in \{-1, 0, 1\}$, implies $c_i = 0$ for every $1 \leq i \leq k$.

Definition 1.3.7. *If $\Lambda := \{\lambda_1, ..., \lambda_k\}$ we say that the cube spanned by Λ is the set of the form*

$$\overline{\Lambda} := \{\sum_{i=1}^{j} c_i\lambda_i : \lambda_i \in \Lambda, c_i \in \{-1, 0, 1\}\}.$$

The following theorem shows that the set of large Fourier coefficients of a given subset of the integers must be structured.

Theorem 1.3.8. (Chang's theorem) *Let $\rho, \alpha \in (0,1]$, let $A \subset \mathbb{Z}_N$ such that $|A| = \alpha N$ and define $R \subset \mathbb{Z}_N$ by*

$$R := \{\xi \in \mathbb{Z}_N : |\widehat{A}(\xi)| \geq \rho\alpha\}.$$

Then there exists a maximal dissociated subset $\Lambda \subset R$ so that $R \subset \overline{\Lambda}$ and $|\overline{\Lambda}| \leq 2\rho^{-2}\log(1/\alpha)$.

In the remainder of this section, we will explain the following proposition of Sanders which we require to prove Proposition 2.7.1.

Proposition 1.3.9. *Suppose $B := B(\Gamma, \delta)$ is a regular Bohr set and that ϵ, $\eta \in (0,1]$. Assume $f : \mathbb{Z}_N \to \mathbb{R}$ such that supp(f) $\subset B$. Then there exists $\Lambda \subset \mathbb{Z}_N$ and $\delta' \in (0,1]$ such that*

$$|\Lambda| \ll \epsilon^2 |B| \log\left(\|f\|_{L^1(B)}^{-2} \|f\|_{L^2(B)}^2\right),$$

$$\delta' \gg \frac{\delta\eta\epsilon^2}{d^2|B|\log\left(\|f\|_{L^1(B)}^{-2}\|f\|_{L^2(B)}^2\right)}$$

and

$$\{\xi \in \mathbb{Z}_N : |\widehat{f}(\xi)| \geq \frac{\epsilon|B|}{N}\|f\|_{L^1(B)}\}$$
$$\subset \{\xi \in \mathbb{Z}_N : |1 - e(-x\xi/N)| \leq \eta \; \forall \; x \in B(\Gamma \cup \Lambda, \delta')\}.$$

Sanders proves his theorem with two lemmas which can be compared with the two conclusions of Theorem 1.3.8. The first step shows that the set of large Fourier coefficients of f must satisfy a structural inclusion. The second step shows that this inclusion is 'nice' in a quantitative sense. Before we can state the two required lemmas, we must generalize the definition of dissociated sets in the context of Bohr sets.

Definition 1.3.10. *Suppose $\Lambda := \{\lambda_1, ..., \lambda_d\}$ and $m := (m_i)_{i=1}^d$ where $m_i \in \mathbb{Z}$. Define*

$$m\Lambda := \sum_{i=1}^{d} m_i\lambda_i$$

and

$$|m| := \sum_{i=1}^{d} m_i.$$

Let S be a non-empty symmetric neighborhood of 0. Then we say that Λ is S-dissociated if $m\Lambda \in S$ implies $m_i = 0$ for every $1 \leq i \leq d$.

With this definition we use the following lemma to show that the set of large Fourier coefficients of a function f is contained in a highly structured set.

Lemma 1.3.11. *Suppose $B := B(\Gamma, \delta)$ is a regular Bohr set and let ϵ, $\eta \in (0, 1]$. Suppose $f : \mathbb{Z}_N \to \mathbb{R}$ such that supp(f) $\subset B$ and define $S := \{\xi \in \mathbb{Z}_N : |\widehat{f}(\xi)| \geq \frac{\epsilon|B|}{N}\|f\|_{L^1(B)}\}$. If Λ is a maximal $\{\xi \in \mathbb{Z}_N : |\widehat{B}(\xi)| \geq \frac{N}{3|B|}\}$-dissociated subset of S then there exists*

$$\delta' \gg \min\{\frac{\eta}{|\Lambda|}, \frac{\eta\delta}{3d}\}$$

such that

$$S \subset \{\xi \in \mathbb{Z}_N : |1 - e(-x\xi/N)| \leq \eta \ \forall \ x \in B(\Gamma \cup \Lambda, \delta')\}.$$

The proof of this lemma is straightforward and follows from basic facts about Bohr sets. Since the proof requires several steps, we recommend the reader consult [14] for a detailed proof.

Lemma 1.3.12. *Suppose B, f, ϵ, η and S are as in the previous lemma. Further, assume that $f \not\equiv 0$. Then there exists $\delta'' \gg \frac{\delta}{\mathrm{d}|\Lambda|}$ such that $B(\Gamma, \delta'')$ is regular and such that if Λ is a $\{\xi \in \mathbb{Z}_N : |\widehat{B}(\xi)| \geq \frac{N}{3|B|}\}$-dissociated subset of S then*

$$|\Lambda| \ll \epsilon^{-2}|B| \log(\|f\|_{L^1(B)}^{-2}\|f\|_{L^2(B)}^2).$$

While the proof of this lemma requires a localized version of Rudin's inequality, several steps are similar to the proof of Chang's theorem. Again we direct the reader to [14] for a complete proof.

Finally, we would like to mention that Proposition 1.3.9 can also be seen as a quantitative improvement of the dual version of a local Bessel inequality of Green and Tao [10]. Using the ideas from Chang's structure theorem, Sanders greatly improves the size on $|\Lambda|$ with a cost on the size of δ'.

1.4 A proof of Sárközy's Theorem

In this section we include a proof of Sárközy's theorem following Green [6] which we require in order to prove Theorem 2.3.1. We also remark that as in [6] we do not require the best known bounds, and so minimize the technicalities which results in a less than optimal quantitative bound.

Theorem 1.4.1. (Sárközy's theorem) *Let $\delta > 0$ and suppose $A \subset \{1, ..., N\}$ so that $|A| \geq \delta N$. Assume N is sufficiently large. Then there exist elements x and y in A ($x \neq y$) so that $x - y$ is a perfect square.*

We begin with some preliminary lemmas and propositions. For the remainder of this section, we assume S to be the set of non-zero squares less than or equal to $N/2$.

Proposition 1.4.2. (Squares in uniform sets) *Let $\delta > 0$. Suppose $A \subset \{1, ..., N\}$ such that $|A| = \delta N$. Assume that $|\widehat{A}(\xi)| \leq \alpha$ for every $\xi \neq 0$ where $\alpha := 2^{-30}\delta^{13/2}$. Then there exist x, $y \in A$ such that $x - y$ is a perfect square.*

Proof. We first note that A must have density δ on at least one of the intervals $[0, N/2]$ or $[N/2, N]$. Without loss of generality, assume $B := A \cap [0, N/2]$ is such that $|B| \geq \delta N/2$. If we then consider A and B as subsets of $\mathbb{Z}_N$ we can bound the number of integer square differences from below using the number of square differences modulo N. The number of square differences modulo N is larger than

$$\begin{aligned}\sum_{x,y\in\mathbb{Z}_N} S(x)B(y)A(y+x) &= N^2 \sum_{\xi\in\mathbb{Z}_N} \widehat{S}(\xi)\widehat{B}(\xi)\widehat{A}(-\xi) \\ &= N^2\widehat{S}(0)\widehat{B}(0)\widehat{A}(0) + N^2\sum_{\xi\neq0}\widehat{(S)}(\xi)\widehat{B}(\xi)\widehat{A}(-\xi) \\ &\geq N^{3/2}\delta^2/4 \\ &\quad - N^2\max_{\xi\neq0}|\widehat{A}(\xi)|^{1/6}\sum_{\xi\in\mathbb{Z}_N}|\widehat{S}(\xi)||\widehat{B}(\xi)||\widehat{A}(\xi)|^{5/6} \\ &\geq N^{3/2}\delta^2/4 - N^2\max_{\xi\neq0}|\widehat{A}(\xi)|^{1/6}\Big(\sum_{\xi\in\mathbb{Z}_N}|\widehat{S}(\xi)|^{12}\Big)^{1/12}\cdot \\ &\quad \Big(\sum_{\xi\in\mathbb{Z}_N}|\widehat{B}(\xi)|^2\Big)^{1/2}\Big(\sum_{\xi\in\mathbb{Z}_N}|\widehat{A}(\xi)|^2\Big)^{5/12}\end{aligned}$$

where we have used Parseval's identity and Hölder's inequality. Using Parseval again, we have the bounds $\sum_\xi |\widehat{B}(\xi)|^2 \leq \delta$ and $\sum_\xi |\widehat{A}(\xi)|^2 \leq \delta$. From [6] we have $\|\widehat{S}\|_{12} \leq 2^{19/12}N^{-1/2}$. Putting this all together, with the assumption of α-uniformity, we have

$$\sum_{x,y\in\mathbb{Z}_N} S(x)B(y)A(y+x) \geq \frac{N^{3/2}\delta^2}{4} - \alpha^{1/6}\cdot 2^{19/12}N^{3/2}\delta^{11/12}.$$

This quantity is positive with our choice of α as desired. □

Proposition 1.4.3. (Density increment) *Suppose there exists $\xi \in \mathbb{Z}_N$ such that $|\widehat{A}(\xi)| \geq 2^{-30}\delta^{13/2}|A|/N$. Then there exists a progression P with square difference such that*
(i) $\frac{|A \cap P|}{|P|} \geq \delta + 2^{-63}\delta^{12}$
and
(ii) $|P| \geq \frac{1}{20}N^{1/64}$

We require the following lemma cited in [6] from [5]. The proof uses Weyl's inequality.

Lemma 1.4.4. *Let $a \in \mathbb{Z}_N$ and suppose $2^{2^{128}} \leq t \leq N$. Then there exists some $p \leq t$ such that $|p^2 a| \leq t^{-1/16}N$.*

Proof. (Of proposition 1.4.3) By Lemma 1.4.4, we take $t = N^{1/4}$, and hence there exists $p \leq N^{1/4}$ such that $|p^2\xi| \leq N^{63/64}$. We then define the square difference progression $P := \{p^2, 2p^2, ..., Lp^2\}$ where $L = \frac{1}{20}N^{1/64}$. Then, we have the estimate

$$\begin{aligned}
\widehat{P}(\xi) &= N^{-1} \sum_{x \in \mathbb{Z}_N} P(x)e(-x\xi/N) \\
&= N^{-1} \sum_{j=1}^{L} e(-(jp^2)\xi/N) \\
&= N^{-1} \sum_{j=1}^{L} \big(1 - (1 - e(-(jp^2)\xi/N))\big) \\
&= L/N - N^{-1} \sum_{j=1}^{L} \big(1 - e(-(jp^2)\xi/N)\big).
\end{aligned}$$

Hence, using the triangle inequality, we have

$$\begin{aligned}
|\widehat{P}(\xi)| &\geq \frac{L}{N} - \frac{1}{N}\sum |1 - e(-(jp^2)\xi/N)| \\
&\geq \frac{L}{N}(1 - 2\pi\frac{L|p^2\xi|}{N}) \\
&\geq \frac{L}{2N}.
\end{aligned}$$

Now we would like to show that A has increased density on some translate of P. We begin by summing the relative densities over all translates. In the following calculation we use the definition of convolution as well as Parseval's identity and the convolution identity given in Lemma 1.2.3.

$$\begin{aligned}\sum_x |A \cap (x-P)|^2 &= \sum_x \Big|\sum_y A(y)(x-P)(y)\Big|^2 \\ &= \sum_x |A * P(x)|^2 \\ &= N \sum_\xi |\widehat{(A*P)}(\xi)|^2 \\ &= N^3 \sum_\xi |\widehat{A}(\xi)|^2 |\widehat{P}(\xi)|^2 \\ &\geq N\delta^2 L^2 (1 + 2^{-62}\delta^{11}).\end{aligned}$$

This bound shows that we have the desired density increment for a given progression modulo N, however we require true $\mathbb{Z}$ progressions. Hence, we need to check that our bound holds for some square progression that doesn't split. We say that a progression P is 'good' if it is a genuine $\mathbb{Z}$ progression and we denote the set of such progressions by G. Our choice of L gives $Lp^2 \leq \frac{1}{20}N^{1/64}N^{1/2} < N^{2/3}$ and hence we know that the set of 'good' progressions must have size at least $N^{1/3}$. Therefore, the contribution to our estimate above from 'bad' progressions is given by

$$\sum_{x \notin G} |A \cap (x-P)|^2 \leq \sum_{x \notin G} L^2 \leq L^2 N^{2/3}.$$

Hence, we have

$$\begin{aligned}\sum_{x \in G} |A \cap (x-P)|^2 &\geq N\delta^2 L^2(1 + 2^{-62}\delta^{11}) - L^2 N^{2/3} \\ &\geq N\delta^2 L^2 (1 + 2^{-63}\delta^{11})\end{aligned}$$

for large enough N. On the other hand, we have

$$\begin{aligned}\sum_{x \in G} |A \cap (x-P)|^2 &\leq \max_{x \in G} |A \cap (x-P)| \sum_{x \in G} |A \cap (x-P)| \\ &\leq |A| L \max_{x \in G} |A \cap (x-P)|.\end{aligned}$$

Hence, there exists $x_0 \in G$ so that

$$|A \cap (x_0 - P)| \geq \delta(1 + 2^{-63}\delta^{11})|P|$$

as desired. □

Proof of Sárközy's theorem: We use the following iterative argument:

Step 1: Set $P_0 := \{1, ..., N_0\}$ where $N_0 := N$, $A_0 := A$ and $\delta_0 := |A_0|/|P_0|$. If A_0 is such that $\widehat{A_0}(\xi)| \leq \alpha$ for each $\xi \neq 0$, we terminate. Otherwise, there must exist ξ_0 such that $|\widehat{A}(\xi_0)| \geq \delta^{13/2} \cdot 2^{-30}$. Hence by Proposition 1.4.3 there exists a square difference progression P_1 such that

$$\frac{|A_0 \cap P_1|}{|P_1|} \geq \delta_0 + 2^{-63}\delta_0^{12}$$

and

$$|P_1| \geq \frac{1}{20}N_0^{1/64}.$$

Step 2: Map the set $A_0 \cap P_1$ to $A_1 \subset \{1, ..., \lfloor 1/20N_0^{1/64} \rfloor\} =: P_1$. Any square difference found in A_1 will correspond to a square difference in A_0. Now apply the same argument as in Step 1.

Iterating this argument, we will reach a density of $2\delta_0$ after at most $2^{63}\delta^{-11}$ steps and a density of 1 after less than $2^{64}\delta^{-11}$ steps. Finally, we check that after our iterations we still have a set P_k which has at least two elements. We can check that this happens as long as $c_1 N^{(1/(64))^{c_2\delta^{11}}} \geq 2$ which we can guarantee as long as $N \geq \exp(\exp(c\delta^{-11}))$.

1.5 The transference principle and pseudorandom sets

The idea of transference enables us deduce results for sets which obey certain random properties from results which are known for sets of positive upper density. The first such result of this type is a version of Roth's theorem in random sets due to Kohayakawa, Łuczak and Rödl [12] which we briefly described in Section 1.1 and is stated precisely in Section 2.1. A Fourier analytic proof of the same result is given by Tao and Vu in [16]. The first such application developed in the Fourier analytic setting was by Green [7] in order to prove his version of Roth's theorem in the primes.

Theorem 1.5.1. (Green) *Suppose that $A \subset P$, where P is defined to be the set of primes, such that*

$$\limsup_{n \to \infty} \frac{|A \cap P_n|}{|P_n|} = \alpha > 0$$

where P_n is the set of primes less than or equal to n. Then there exists x, y, $z \in A$ such that $x + z = 2y$.

Green's theorem states that if one considers a subset of the primes with positive relative density, then Roth's theorem still holds. In his proof he exploits the fact that the primes obey certain random properties.

Later, Green and Tao [9] reformulate the transference principle in a form we use below. The application in their paper allows them to deduce a version of Roth's theorem for subsets of Chen primes (those primes p such that $p+2$ is the product of at most 2 primes) with positive relative density. This form of the transference principle says that if a set is majorized by a 'pseudorandom' measure then the set can be decomposed into a large bounded component plus a small uniform component. On the bounded component we can apply known results for sets of positive upper density, while the uniform component only contributes a small error term.

Lemma 1.5.2. *Assume that $f : \mathbb{Z}_N \to [0, \infty)$ satisfies $\mathbb{E}(f) \geq \delta > 0$ and*

$$\|\widehat{f}\|_q \leq M \tag{1.5.6}$$

for some $2 < q < 3$. Assume also that $f \leq \nu$, where $\nu : \mathbb{Z}_N \to [0, \infty)$ obeys the pseudorandom condition

$$\|\hat{\nu}(\xi) - \mathbf{1}_{\xi=0}\|_\infty \leq \eta \tag{1.5.7}$$

for some $0 < \eta \leq 1$. Let

$$f_1(x) = \mathbb{E}(f(x + y_1 - y_2) : \ y_1, y_2 \in B_0),$$

where

$$B_0 = \{x : \ |e^{-2\pi i \xi x/N} - 1| \leq \epsilon_0, \ \xi \in \Lambda_0\}, \ \Lambda_0 = \{\xi : \ |\widehat{f}(\xi)| \geq \epsilon_0\}$$

for some ϵ_0 to be fixed later. Let also $f_2(x) = f(x) - f_1(x)$. Then

(i) $0 \leq f_1 \leq 1 + (1 + \mathbb{P}(B_0)^{-1})\eta$,

(ii) $\mathbb{E}f_1 = \mathbb{E}f \geq \delta$,

(iii) $\|\widehat{f_2}(\xi)\|_\infty \leq 3(1+\eta)\epsilon_0$,

(iv) $|\widehat{f_i}(\xi)| \leq |\widehat{f}(\xi)|$ for all $\xi \in \mathbb{Z}_N$ and $i = 1, 2$. In particular, (1.5.6) holds with f replaced by f_2.

Proof. (i) To verify that $|f_1| \leq 1 + \eta N/|B_0|$ we use the definition of f_1, the fact that f is bounded by the pseudorandom function ν, and the Fourier inversion formula 1.2.3 (ii). We have,

$$\begin{aligned}
|f_1| &= \left|\mathbb{E}_{y_1,y_2\in B_0}\big(f(x+y_1-y_2)\big)\right| \\
&\le \left|\mathbb{E}_{y_1,y_2\in B_0}\big(\nu(x+y_1-y_2)\big)\right| \\
&= \left|\mathbb{E}_{y_1,y_2\in B_0}\sum_{\xi\in\mathbb{Z}_N}\widehat{\nu}(\xi)e(-(x+y_1-y_2)\xi/N)\right| \\
&= \left|\mathbb{E}_{y_1,y_2\in B_0}\sum_{\xi\in\mathbb{Z}_N}\widehat{\nu}(\xi)e(-x\xi/N)|\mathbb{E}_{y\in B_0}e(-y\xi/N)|^2\right| \\
&\le \sum_{\xi\in\mathbb{Z}_N}|\widehat{\nu}(\xi)||\mathbb{E}_{y\in B_0}e(-y\xi/N)|^2 \\
&\le 1+\eta N/|B_0|
\end{aligned}$$

which completes the proof of (i).

(ii) The fact that $\mathbb{E}f_1=\mathbb{E}f$ follows directly from the definition of f_1.

(iii) and (iv) We begin with the observation that $|\widehat{f_1}(\xi)|\le|\widehat{f}(\xi)|$ since

$$\begin{aligned}
\widehat{f_1}(\xi) &:= N^{-1}\sum_{x\in\mathbb{Z}_N}f_1(x)e(-x\xi/N) \\
&= N^{-1}\sum_{x\in\mathbb{Z}_N}\mathbb{E}_{y_1,y_2\in B_0}f(x+y_1-y_2)e(-x\xi/N) \\
&= \widehat{f}(\xi)|\mathbb{E}_{y\in B_0}e(y\xi/N)|^2
\end{aligned}$$

which verifies (iv) in the case that $i=1$. Using this equality, we have

$$\begin{aligned}
\widehat{f_2}(\xi) &= \widehat{f}(\xi)-\widehat{f_1}(\xi) \\
&= \widehat{f}(\xi)\big(1-|\mathbb{E}_{y\in B_0}(e(y\xi/N))|^2\big)
\end{aligned}$$

which shows (iv) with $i=2$. Now, if $\xi\in\Lambda_0$ then for each $y\in B_0$ we must have $|e(y\xi/N)-1|\le\epsilon_0$ by the definition of B_0. Combining this with the identity for $\widehat{f_2}$ above, we must have

$$|\widehat{f_2}|\le 3\epsilon_0\mathbb{E}(\nu)\le 3(1+\eta)\epsilon_0.$$

On the other hand, if $\xi\notin\Lambda_0$ then by the definition of Λ_0 we have $|\widehat{f}(\xi)|<\epsilon_0$ and hence, combining this with our identity for $\widehat{f_2}$ we have

$$|\widehat{f_2}|\le\epsilon_0$$

establishing (iii). □

The following lemma, from [16], is needed to deduce Lemma 2.2.6 which allows us to exploit the random properties of ν in order to deduce an l^q estimate from an l^2 one.

Lemma 1.5.3. *Suppose that* $f : \mathbb{Z}_N \to [0, \infty)$ *satisfies* $\mathbb{E}f \geq \delta > 0$ *and assume that* $f \leq \nu$, *where* ν *satisfies the pseudorandom condition*

$$||\widehat{\nu}(\xi) - \mathbf{1}_{\xi=0}||_\infty \leq \eta$$

for some $0 < \eta \leq 1$. *Define* $\Lambda := \{\xi \in \mathbb{Z}_N : |\widehat{f}(\xi)| \geq \alpha\}$ *for any* $\alpha > 0$. *Then*

$$|\Lambda| \leq 4/\alpha^2$$

for all $\alpha \geq 2\eta^{1/2}$.

Proof of Lemma 2.2.6: By assumption, we have $||\widehat{f}||_2^2 \leq c\eta^{-\epsilon/2}$. Hence, we have

$$\begin{aligned}
||\widehat{f}||_{2+\epsilon}^{2+\epsilon} &:= \sum_{\xi \in \mathbb{Z}_N} |\widehat{f}(\xi)|^{2+\epsilon} \\
&= \sum_{\xi : |\widehat{f}(\xi)| \leq 4\eta} |\widehat{f}(\xi)|^{2+\epsilon} . + \sum_{\xi : |\widehat{f}(\xi)| > 4\eta} |\widehat{f}(\xi)|^{2+\epsilon} \\
&\leq 4\eta^\epsilon \sum_{\xi \in \mathbb{Z}_N} |\widehat{f}(\xi)|^2 + \sum_{k \geq \lfloor \frac{\log \eta}{\log 2} \rfloor} \sum_{\xi : \widehat{f}(\xi) \in [2^k, 2^{k+1}]} |\widehat{f}(\xi)|^{2+\epsilon} \\
&\leq c\eta^{\epsilon/2} + \sum_k 2^{k(2+\epsilon)} |\{\xi : |\widehat{f}(\xi)| \approx 2^k\}| \\
&\leq c\eta^{\epsilon/2} + \sum_k 2^{k(2+\epsilon)} \cdot 2^{-2k+1} \\
&\leq M
\end{aligned}$$

where we have used Lemma 1.5.3 for the second to last step.

Bibliography

[1] F. Behrend, *On the sets of integers which contain no three terms in arithmetic progression.* Proc. Nat. Acad. Sci., 23 (1946), 331-332.

[2] J. Bourgain, *On triples in arithmetic progressions*, Geom. Funct. Anal. 9 (1999), 968–984.

[3] G. A. Freiman, *Foundation of a structural theory of set addition (translated from Russian)*, Translations of Mathematical Monographs, 37, 1973.

[4] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. d'Analyse Math, 71, 1977, 204-256.

[5] T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal., 12 (2001), 465-588.

[6] B.J. Green, *On arithmetic structures in dense sets of integers*, Duke Math. Jour., 114 (2002), 215-238.

[7] B.J. Green, *Roth's Theorem in the primes*, Annals of Math. 161 (2005), 1609-1636.

[8] B.J. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Annals of Math., to appear.

[9] B.J. Green, T. Tao, *Restriction theory of the Selberg Sieve, with applications*, Journal de Théorie des Nombers de Bordeaux, **18** (2006), 137–172.

[10] B.J. Green, T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc., to appear.

[11] M. Hamel and I. Laba, *Arithmetic structures in random sets*, INTEGERS:Electronic Journal of Combinatorial Number Theory, 8, 2008.

[12] Y. Kohayakawa, T. Łuczak, V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), 133–163.

[13] K. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.

[14] T. Sanders, *Additive structures in sumsets*, to appear in Math. Proc. Cambridge Philos. Soc.

[15] A. Sárkozy, *On difference sets of integers i*, Acta Math. Acad. Sci. Hungar, 31:125-149, 1978.

[16] T. Tao, V. Vu, *Additive combinatorics*, Cambridge Univ. Press, 2006.

[17] P. Varnavides, *On certain sets of positive density*, Journal London Math. Soc., **34** (1959), 358–360

Department of Mathematics, University of British Columbia,
Vancouver, B.C. V6T 1Z2, Canada
mehamel@math.ubc.ca

Chapter 2

Arithmetic structures in random sets

Arithmetic structures in random sets[1]

Mariah Hamel and Izabella Łaba

Abstract

We extend two well-known results in additive number theory, Sárközy's theorem on square differences in dense sets and a theorem of Green on long arithmetic progressions in sumsets, to subsets of random sets of asymptotic density 0. Our proofs rely on a restriction-type Fourier analytic argument of Green and Green-Tao.

2.1 Introduction

The purpose of this paper is to extend several basic results in additive number theory, known for sets of positive density in $\mathbb{Z}_N$, to the setting of random sets of asymptotic density 0. This line of work originated in the paper of Kohayakawa-Luczak-Rödl [29], who proved a random-set analogue of Roth's theorem on 3-term arithmetic progressions. Roth's theorem [32] asserts that for any fixed $\delta > 0$ there is a large integer N_0 such that if $N > N_0$ and if A is a subset of $\{1, \ldots, N\}$ with $|A| \geq \delta N$, then A contains a non-trivial 3-term arithmetic progression $a, a+r, a+2r$ with $r \neq 0$. The article [29] raises the following question: are there any sets W, sparse in $\{1, \ldots, N\}$, with the property that any set A containing a positive proportion of the elements of W must contain a 3-term arithmetic progression? The authors proceed to answer it in the affirmative for random sets:

[1] A version of this paper is published. M. Hamel and I. Łaba, Additive structures in random sets, INTEGERS: Electronic Journal of Combinatorial Number Theory, 8 2008.

Theorem 2.1.1. *[29] Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (CN^{-1/2}, 1]$. Fix $\alpha > 0$. Then the statement*

> *every set $A \subset W$ with $|A| \geq \alpha|W|$ contains a 3-term arithmetic progression*

is true with probability $1 - o_\alpha(1)$ as $N \to \infty$.

The current interest in questions of this type is motivated by the work of Green [25] and Green-Tao [26], [27] on arithmetic progressions in the primes, where the "pseudorandomness" of the almost-primes plays a key role. For example, Tao-Vu [39, Section 10.2] give an alternative (and simpler) proof of Theorem 2.1.1 under the stronger assumption that $p \geq CN^{-\theta}$ with θ small enough. While the argument in [29] is combinatorial and uses Szemerédi's regularity lemma, the proof in [39] is Fourier-analytic and relies in particular on a restriction-type estimate from [25], [27].

It is natural to ask which other results from additive number theory can be extended to the random set setting. While the methods of [29] do not seem to extend to other questions, the decomposition technique in [27] turns out to be more robust. We are able to use it to prove random set analogues of two well-known results: Sárközy's theorem on square differences, and a theorem of Green on long arithmetic progressions in sumsets.

We note that the concept of pseudorandomness has played a major role in many of the basic extremal results in additive number theory, such as Szemerédi's theorem on arithmetic progressions. Specifically, in order to find a certain type of an arithmetic structure (such as an arithmetic progression) in sets of positive density, one often begins by showing that such structures are common in appropriately defined pseudorandom sets. It is not clear whether our results will have applications of this type, as the corresponding extremal results for sets of positive density are already known. On the other hand, we expect that the methods developed here will be useful in proving similar results in settings where the background set W is a given set of density zero with sufficiently good pseudorandom properties (e.g. the primes, the Chen primes). For example, one could inquire about the arithmetic properties of sets of the form $A + B$, where A and B are subsets of the primes with relative positive density.

We now give the precise statement of our results. Throughout the paper, W is a random subset of $\mathbb{Z}_N$, with each $x \in \mathbb{Z}_N$ belonging to W independently with probability $p \in (0, 1]$. We will assume that $p \geq N^{-\theta}$, where θ is a sufficiently small positive number. In particular, we allow p to go to 0 as $N \to \infty$. We also fix $\delta > 0$ and let $A \subset W$, $|A| = \delta|W|$.

Sárközy's theorem (proved also independently by Furstenberg) states that for any fixed positive number δ there is a large integer N_0 such that if $N > N_0$ and if A is a subset of $\{1, \ldots, N\}$ with $|A| \geq \delta N$, then A contains two distinct elements x, y such that $x - y$ is a perfect square. The best known quantitative bound, due to Pintz, Steiger and Szemerédi [31], is that one may take $N_0 = (\log N)^{-c \log\log\log\log N}$. In the converse direction, Ruzsa [33] constructed a set of size $N^{1-0.267}$ which contains no square difference.

We are able to prove the following.

Theorem 2.1.2. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (cN^{-\theta}, 1]$ where $0 < \theta < 1/110$. Let $\alpha > 0$. Then the statement*

> *for every set $A \subset W$ with $|A| \geq \alpha W$, there are $x, y \in A$ such that $x - y$ is a non-zero perfect square*

is true with probability $1 - o_\alpha(1)$ as $N \to \infty$.

We also have an analogous result for higher power differences, see Section 2.5.

If A, B are two sets of integers, we will write $A + B = \{a + b : a \in A, b \in B\}$. Let W be a random set as described above, but with $\theta \in (0, 1/2]$. One can show using a probabilistic argument that it holds with probability $1 - o(1)$ that the sumset $A + A$ of every subset $A \subset W$ with $|A| > \alpha|W|$ has density at least α^2 in $\mathbb{Z}_N$ [2]. If θ is close enough to 0, then we can prove the following stronger result using Fourier-analytic methods.

Proposition 2.1.3. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (CN^{-\theta}, 1]$, where $0 < \theta < 1/140$. Then for every $\beta < \alpha$, the statement*

> *for every set $A \subset W$ with $|A| \geq \alpha|W|$, we have $|A + A| \geq \beta N$*

is true with probability $1 - o_{\alpha,\beta}(1)$ as $N \to \infty$.

It is easy to see that one can have $|A + A| \approx \alpha N$ in the setting of the proposition: let $A_x = W \cap (P + x)$, where P is an arithmetic progression in $\mathbb{Z}_N$ of step about α^{-1} and length about αN. An averaging argument shows that $|A_x| \gg \alpha|W|$ for some x, while $|A_x + A_x| \leq 2|P| \approx \alpha N$.

[2] We are grateful to Mihalis Kolountzakis for pointing this out to us and communicating a short proof.

Our second main result concerns the existence of long arithmetic progressions in sumsets. Bourgain [18] proved that if A, B are sumsets of $\{1, \dots, N\}$ with $|A| > \alpha N$, $|B| > \beta N$, then $A + B$ contains a k-term arithmetic progression with

$$k > \exp(c(\alpha\beta \log N)^{1/3} - \log\log N). \tag{2.1.1}$$

The point here is that a sumset has much more arithmetic structure, and therefore contains much longer arithmetic progressions, than would be normally expected in a set of a similar size (based on Szemerédi's theorem, for example). Bourgain's bound was improved by Green [23] to

$$k > \exp(c(\alpha\beta \log N)^{1/2} - \log\log N), \tag{2.1.2}$$

which is the best known result in this direction so far. An alternative proof of essentially the same bound was given more recently by Sanders [35]. On the other hand, Ruzsa [34] gave a construction showing that the exponent $1/2$ in (2.1.2) cannot be improved beyond $2/3$. Note that if $A = B$, the estimate (2.1.2) gives a non-trivial result only when $\alpha > (\log N)^{-1/2}$, and in particular sets with density $N^{-\epsilon}$ are not allowed.

The case of sparse sets was considered more recently by Croot-Ruzsa-Schoen [21]. The authors proved that if $A, B \subset \mathbb{Z}_N$ obey $|A||B| \geq (6N)^{2-\frac{2}{k-1}}$, then $A + B$ contains a k-term arithmetic progression. They also gave a construction of sets $A \subset \mathbb{Z}_N$ with $|A| \geq N^{1-\theta}$, where θ is small enough depending on $\epsilon > 0$, such that $A + A$ does not contain an arithmetic progression longer than $\exp(c\theta^{-\frac{2}{3}-\epsilon})$.

Our result is the following.

Theorem 2.1.4. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (CN^{-\theta}, 1]$, where $0 < \theta < 1/140$. Assume that α and k obey*

$$\alpha \geq \frac{C_1 \log\log N}{\sqrt{\log N}}, \tag{2.1.3}$$

$$k \leq \exp\left(\frac{\alpha^2 \log\log N}{C_2 \log\frac{1}{\alpha}(\log\log\log N + \log\frac{1}{\alpha})}\right), \tag{2.1.4}$$

where C_1, C_2 are sufficiently large constants. Then the statement

> *for every set $A \subset W$ with $|A| \geq \alpha|W|$, the sumset $A + A$ contains a k-term arithmetic progression*

is true with probability $1 - o_{k,\alpha}(1)$ as $N \to \infty$.

A non-quantitative version of the result, namely that the displayed statement in the theorem is true with probability $1 - o(1)$ as $N \to \infty$ if α and k are fixed, can be obtained by applying Szemerédi's theorem to the positive density set $A+A$. Our point, as in [18] or [23], is that the arithmetic progressions indicated by Theorem 2.1.4 are much longer than those in Szemerédi's theorem, and that they can be found using a much easier argument. For comparison, the current best bounds in Szemerédi's theorem [22] imply that a set of relative density α in $\mathbb{Z}_N$ should contain k-term arithmetic progressions with

$$k \le \log\log\left(\frac{\log\log N}{\log\frac{1}{\alpha}}\right),$$

which is much weaker than (2.1.4).

The bounds on θ in Theorems 2.1.2 and 2.1.4 are due to our choices of exponents in the proofs and are probably not optimal. The natural threshold would be 1/2, as in [29]. However, it does not seem possible to extend our results to all $\theta < 1/2$ using the same type of arguments as in this paper.

The article is organized as follows. In the next section we explain the notation and summarize the known results that will be used repeatedly. Theorem 2.1.2 is proved in Sections 2.3 and 2.4. Its analogue for higher power differences, Theorem 2.5.1, is stated and proved in Section 2.5. The proof of Theorem 2.1.4 is given in Section 2.6, with the proofs of the main estimates postponed to Sections 2.7 and 2.8. The proof of Proposition 2.1.3, which involves a simplified version of the argument in the proof of Theorem 2.1.4, concludes the paper.

2.2 Preliminaries

We first explain the notation. We use $|A|$ to denote the cardinality of a set $A \subset \mathbb{Z}_N$. The *probability* of a set A is $\mathbb{P}(A) = N^{-1}|A|$, and the *expectation* of a function $f : \mathbb{Z}_N \to \mathbb{C}$ is defined as

$$\mathbb{E}f = \mathbb{E}_x f = N^{-1} \sum_{x\in\mathbb{Z}_N} f(x).$$

We will also sometimes use conditional probability and expectation

$$\mathbb{P}(A|X) = \frac{|A \cap X|}{|X|},\ \mathbb{E}(f|X) = \mathbb{E}_{x\in X} f(x) = \frac{1}{|X|}\sum_{x\in X} f(x).$$

Whenever the range of a variable (in a sum, expectation, etc.) is not indicated, it is assumed to be all of $\mathbb{Z}_N$. We will also use the notation

$\|f\|_p = (\sum_x |f(x)|^p)^{1/p}$ and $\|f\|_{L^p(X)} = (\sum_{x\in X} |f(x)|^p)^{1/p}$. All constants throughout the paper will be independent of N, α, and k.

The discrete Fourier transform of f is defined by

$$\widehat{f}(\xi) = \mathbb{E}_x f(x) e^{-2\pi i x\xi/N}.$$

We have the usual Plancherel identity $\sum \widehat{f}\widehat{\bar{g}} = N^{-1}\sum f\bar{g}$ and the inversion formula $f(x) = \sum_{\xi\in\mathbb{Z}_N} \widehat{f}(\xi)e^{-2\pi i x\xi/N}$.

We define the convolution of two functions $f, g : \mathbb{Z}_N \to \mathbb{C}$ by the formula

$$(f * g)(x) = \sum_y f(y)g(x-y) = \sum_{t,s:t+s=x} f(t)g(s).$$

We have the identity $N\widehat{f}\widehat{g} = \widehat{f * g}$.

We recall a few basic results about Bohr sets, all of which are standard in the literature and can be found e.g. in [28], [39], or in [19] where regular Bohr sets were first introduced.

Definition 2.2.1. *A* Bohr set *is a set of the form $B = b + B(\Lambda, \delta)$, where $b \in \mathbb{Z}_N$, $\Lambda \subset \mathbb{Z}_N$, $\delta \in (0, 2)$, and*

$$B(\Lambda, \delta) = \{x \in \mathbb{Z}_N : \ |e^{2\pi i x\xi/N} - 1| \leq \delta \text{ for all } \xi \in \Lambda\}.$$

We will often refer to $|\Lambda|$ and δ as the rank and radius of B, respectively.

Definition 2.2.2. *Let c_0 be a small positive constant which will remain fixed throughout the paper. We will say that a Bohr set $B(\Lambda, \delta)$ is* regular *if*

$$\mathbb{P}(B(\Lambda, (1 + c_0^2)\delta) \setminus B(\Lambda, (1 - c_0^2)\delta)) \leq c_0\mathbb{P}(B(\Lambda, \delta)).$$

We will also say that $B = b + B(\Lambda, \delta)$ is regular if $B(\Lambda, \delta)$ is regular.

Lemma 2.2.3. *If $B = B(\Lambda, \delta)$ is a regular Bohr set, then $\mathbb{P}(B) \geq (cc_0^2\delta)^{|\Lambda|}$.*

Lemma 2.2.4. *Assume that c_0 is small enough. Then for any $\Lambda \subset \mathbb{Z}_N$ with $|\Lambda| \leq \sqrt{c_0}N$ and any $\delta_0 > 0$ there is a $\delta \in (\frac{\delta_0}{2}, \delta_0)$ such that $B(\Lambda, \delta)$ is regular.*

We will need a Fourier-analytic argument which first appeared in [25] in a slightly different formulation and in [27] as stated, and was adapted in [39] to a random set setting. Specifically, [25] and [27] introduced the decomposition $f = f_1 + f_2$ defined below, where f_1 is the "structured" bounded part, and f_2 is unbounded but random. We will need several results concerning the properties of f_1 and f_2, which we collect in the next two lemmas. The first one is contained in the proofs of [27, Proposition 5.1] or [39, Theorem 10.20].

Lemma 2.2.5. *Assume that* $f : \mathbb{Z}_N \to [0,\infty)$ *satisfies* $\mathbb{E}(f) \geq \delta > 0$ *and*

$$\|\widehat{f}\|_q \leq M \tag{2.2.1}$$

for some $2 < q < 3$. *Assume also that* $f \leq \nu$, *where* $\nu : \mathbb{Z}_N \to [0,\infty)$ *obeys the pseudorandom condition*

$$\|\widehat{\nu}(\xi) - \mathbf{1}_{\xi=0}\|_\infty \leq \eta \tag{2.2.2}$$

for some $0 < \eta \leq 1$. *Let*

$$f_1(x) = \mathbb{E}(f(x+y_1-y_2) : \ y_1, y_2 \in B_0),$$

where

$$B_0 = \{x : \ |e^{-2\pi i\xi x/N} - 1| \leq \epsilon_0, \ \xi \in \Lambda_0\}, \ \Lambda_0 = \{\xi : \ |\widehat{f}(\xi)| \geq \epsilon_0\}$$

for some ϵ_0 *to be fixed later. Let also* $f_2(x) = f(x) - f_1(x)$. *Then*

(i) $0 \leq f_1 \leq 1 + (1 + \mathbb{P}(B_0)^{-1})\eta$,

(ii) $\mathbb{E}f_1 = \mathbb{E}f$,

(iii) $\|\widehat{f_2}\|_\infty \leq 3(1+\eta)\epsilon_0$,

(iv) $|\widehat{f_i}(\xi)| \leq |\widehat{f}(\xi)|$ *for all* $\xi \in \mathbb{Z}_N$ *and* $i = 1, 2$. *In particular, (2.2.1) holds with* f *replaced by* f_2.

In order to be able to apply Lemma 2.2.5, we need to have the estimate (2.2.1) for some $2 < q < 3$. To this end we have the following result, based on the Stein-Tomas argument as used in [25], [27], and contained in the form we need in [39, Lemma 10.22 and proof of Theorem 10.18].

Lemma 2.2.6. *Let* f *and* ν *be as in Lemma 2.2.5, except that instead of (2.2.1) we assume that*

$$\|\widehat{f}\|_2 \leq C\eta^{-\epsilon/4}$$

for some $\epsilon > 0$. *Then (2.2.1) holds with* $q = 2 + \epsilon$.

We adapt this argument to the random setting as in [39, Section 10.2]. Suppose that W is a random subset of $\mathbb{Z}_N$ such that each $x \in \mathbb{Z}_N$ belongs to W independently with probability $p \in (0,1)$. We will assume that $p \geq N^{-\theta}$, where $0 < \theta < 1/100$. We also fix $\delta > 0$ and let $A \subset W$, $|A| = \delta|W|$. We let

$$\nu(x) = p^{-1}W(x), \ f(x) = p^{-1}A(x).$$

Lemma 2.2.7. *Let* ν *and* f *be the random variables defined above. Then*

(i) $\|\widehat{\nu(\xi)} - \mathbf{1}_{\xi=0}\|_\infty = O(N^{-1/5})$ *with probability* $1 - o(1)$,

(ii) $\|\widehat{f}\|_2^2 = N^{-1}\|f\|_2^2 = O(p^{-1}) \leq N^\theta$ *with probability* $1 - o(1)$.

Part (i) of the lemma follows from well-known probabilistic arguments. It can be found e.g. in [39, Corollary 1.9 and Lemma 4.15], or extracted from the proof of Lemma 14 in [23]. Observe in particular that (i) with $\xi = 0$ says that $\mathbb{P}(W) = p(1 + O(N^{-1/5}))$ with probability $1 - o(1)$. Part (ii) follows from this and the Plancherel identity.

2.3 A Varnavides-type theorem for square differences

The purpose of this section is to prove the following theorem.

Theorem 2.3.1. *Let* $0 < \delta \leq 1$ *and* $N \geq 1$ *be a prime integer. Let* $f : \mathbb{Z}_N \to [0, 1]$ *be a bounded function such that*

$$\mathbb{E}f \geq \delta.$$

Then we have

$$\mathbb{E}(f(n)f(n + r^2)|n, r \in \mathbb{Z}_N,\ 1 \leq r \leq \lfloor\sqrt{N/3}\rfloor) \geq c(\delta) - o_\delta(1).$$

Theorem 2.3.1 strengthens Sárközy's theorem (stated in the introduction) in the same way in which a theorem of Varnavides [40] strengthens Roth's theorem on 3-term arithmetic progressions. It guarantees the existence of "many" square differences in a set of positive density, instead of just one.

Proof. The proof combines Sárközy's theorem with a modification of Varnavides's combinatorial argument [40]. We first note that it suffices to prove the result for characteristic functions. To see this, let f be as in the theorem, and define $A := \{n \in \mathbb{Z}_N : f(n) \geq \delta/2\}$. Then $|A| \geq \delta N/2$ and $f \geq \frac{\delta}{2}$ on A. Hence, assuming the result for characteristic functions, we have

$$\mathbb{E}(f(n)f(n + r^2)) \geq \frac{\delta^2}{4}\mathbb{E}(A(n)A(n + r^2)) \geq \frac{\delta^2}{4}c(\delta/2).$$

We now turn to the proof of the result for characteristic functions. Let $A \subset \mathbb{Z}_N$ such that $|A| \geq \delta N$ and N is sufficiently large. We will consider arithmetic progressions

$$P_{x,r} = \{x, x + r^2, \dots, x + (k-1)r^2\},\ 1 \leq x \leq x + (k-1)r^2 \leq N \quad (2.3.1)$$

where $x, r \in \mathbb{Z}_N$, $r \leq \sqrt{3N}$, and where $k \in \mathbb{N}$ is chosen so that the conclusion of Sárközy's theorem holds for subsets of $\{1, ..., k\}$ which have size at least $\frac{1}{2}\delta k$.

Suppose that

$$r^2 < \frac{\delta N}{k^2}. \tag{2.3.2}$$

We say that a progression $P_{x,r}$ as in (2.3.1) is *good* if

$$|P_{x,r} \cap A| \geq \frac{1}{2}\delta k. \tag{2.3.3}$$

Let $G_r(N)$ denote the set of good progressions $P_{x,r}(N)$ for a fixed r. We claim that

$$|G_r(N)| > \frac{1}{4}\delta N. \tag{2.3.4}$$

Indeed, we have

$$|A \cap (kr^2, N - kr^2)| \geq |A| - 2kr^2 \geq \delta N - 2kr^2 \geq \delta(1 - \frac{2}{k})N,$$

where at the last step we use (2.3.2). Each $a \in A \cap (kr^2, N - kr^2)$ is contained in exactly k progressions $P_{x,r}$. Hence

$$\sum_{x:1\leq x<x+(k-1)r^2\leq N} |A \cap P_{x,r}| \geq k\delta(1 - \frac{2}{k})N > \frac{3}{4}\delta kN \quad (k > 8).$$

On the other hand, the number of progressions $P_{x,r}$ for a fixed r is clearly bounded by N, hence we have an upper bound

$$\sum_{x:1\leq x<x+(k-1)r^2\leq N} |A \cap P_{x,r}| < N \cdot \frac{1}{2}\delta k + G_r(N)k.$$

Combining these bounds yields (2.3.4) as claimed.

Let $G(N) := \sum_{r:r^2<\frac{\delta N}{k^2}} G_r(N)$. Then

$$G(N) \geq \frac{\sqrt{\delta N}}{k}\frac{\delta N}{4} = c_1(\delta)N^{3/2}, \tag{2.3.5}$$

since k depends only on δ.

By Sárközy's theorem, each good progression $P_{x,r}$ contains a square difference. We now count the number of good progressions which may contain a fixed square difference pair $x, x + r^2$. Clearly, $x, x + r^2$ can be contained in at most $k - 1$ progressions with step size r^2 and at most $\frac{1}{2}k(k - 1)$ progressions with step size r^2/t for integers $t > 1$. Since k depends only on δ, the total number of progressions containing $x, x+r^2$ is bounded by $c_2(\delta)$. Thus the total number of square differences in A must be at least

$$\frac{c_1(\delta)}{c_2(\delta)} N^{3/2} = c(\delta) N^{3/2}.$$

Subtracting off the trivial progressions (with $r^2 = 0$) gives the desired result. □

2.4 Proof of Theorem 2.1.2

Let W, A be as in Theorem 2.1.2. At least one of the sets $A_1 = A \cap [0, N/3)$, $A_2 = A \cap [N/3, 2N/3)$, $A_3 = A \cap [2N/3, N)$, say A_1 (the other two cases are identical), has size at least $|A|/3$. Define ν, f as in Lemma 2.2.7, but with A replaced by A_1. By Lemma 2.2.7, the assumptions of Lemma 2.2.6 with $\eta = N^{-1/5}$ and $\epsilon = 1/11$ are satisfied with probability $1 - o(1)$, thus (2.2.1) holds with $q = 23/11$. We will henceforth condition on these events. Let $f = f_1 + f_2$ as in Lemma 2.2.5, with $\epsilon_0 = \epsilon_0(\alpha)$ small enough to be fixed later. We would like to ensure that

$$\|f_1\|_\infty \leq 2. \tag{2.4.1}$$

By Lemma 2.2.5, this will follow if

$$N^{-1/5}(1 + \mathbb{P}(B_0)^{-1}) < 1. \tag{2.4.2}$$

By Lemma 2.2.3, we can estimate $\mathbb{P}(B_0) \gg (c\epsilon_0)^{|\Lambda_0|}$, while by (2.2.1) and Chebyshev's inequality we have $|\Lambda_0| \leq (M/\epsilon_0)^{23/11}$. Now a short calculation shows that if

$$\log \frac{1}{\epsilon_0} < c_1 \log\log N \tag{2.4.3}$$

with c_1 small enough, which we will assume henceforth, then (2.4.2) and (2.4.1) hold.

It suffices to prove that

$$\mathbb{E}(f(x)f(x+r^2) | x, r \in \mathbb{Z}_N, 1 \leq r \leq \sqrt{N/3}) \geq c(\delta) - o_\delta(1). \tag{2.4.4}$$

Indeed, since $A_1 \subset [0, N/3)$, any square difference $a - a' = r^2$ with $a, a' \in A_1$ and $1 \leq r^2 \leq N/3$ must be an actual square difference in $\mathbb{Z}$, not just a square difference mod N.

We write $f(x)f(x+r^2) = \sum_{i,j=1}^{2} f_i(x)f_j(x+r^2)$, and estimate the expectation of each term. Applying Theorem 2.3.1 to f_1, we get a lower bound on the main term

$$\mathbb{E}(f_1(x)f_1(x+r^2)|x,r \in \mathbb{Z}_N, 1 \le r \le \sqrt{N/3}) \ge c_1(\delta) - o_\delta(1), \qquad (2.4.5)$$

if N is large enough so that (2.4.3) holds. We now turn to the error estimates. We write

$$\begin{aligned}&\mathbb{E}(f_2(x)f_2(x+r^2)|x,r \in \mathbb{Z}_N, 1 \le r \le \sqrt{N/3})\\ &= \sqrt{3N}\,\mathbb{E}(f_2(x)f_2(x+t)S(t)|x,t \in \mathbb{Z}_N),\end{aligned} \qquad (2.4.6)$$

where $S(\cdot)$ denotes the characteristic function of the squares less than $N/3$. From Green [24] we have the estimate

$$\|\hat{S}\|_{12} \le 2^{19/12}N^{-1/2},$$

based on a number theoretic bound on the number of representations of an integer as the sum of six squares. Using also Parseval's identity and Hölder's inequality, we have

$$\begin{aligned}&\mathbb{E}(f_2(x)f_2(x+t)S(t)|x,t \in \mathbb{Z}_N)\\ &= \sum_{\xi\in\mathbb{Z}_N} |\hat{f}_2(\xi)|^2|\hat{S}(\xi)|\\ &\le (\sum_{\xi\in\mathbb{Z}_N} |\hat{S}(\xi)|^{12})^{1/12}(\sum_{\xi\in\mathbb{Z}_N} |\hat{f}_2(\xi)|^{24/11})^{11/12}\\ &\le 2^{19/12}N^{-1/2}\|\hat{f}_2\|_{23/11}^{23/12}\|\hat{f}_2\|_\infty^{1/12}\\ &\le CN^{-1/2}\epsilon_0^{1/12}.\end{aligned}$$

Plugging this into (2.4.6), we see that

$$\mathbb{E}(f_2(x)f_2(x+r^2)|x,r \in \mathbb{Z}_N, 1 \le r \le \sqrt{N/3}) \le c_1(\delta)/4$$

if ϵ_0 was chosen sufficiently small depending on δ. The "mixed" error terms are estimated similarly. Combining the error estimates with (2.4.5) yields (2.4.4) as desired.

2.5 Power differences

In this section we show that a modification of the proof of Theorem 2.1.2 yields an analogous result for higher power differences.

Theorem 2.5.1. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (cN^{-\theta}, 1]$ with $0 < \theta < \theta_k$, where θ_k is small enough depending on $k \in \mathbb{N}$. Let $\alpha > 0$. Then the statement*

> *for every set $A \subset W$ with $|A| \geq \alpha W$, $\exists\ x, y \in A$ such that $x - y = n^k$ for some $n \in \mathbb{N}$*

is true with probability $o_{k,\alpha}(1)$ as $N \to \infty$.

Since the proof is very similar to that of Theorem 2.1.2, we only sketch the main steps. Instead of Theorem 2.3.1, we will need a similar result for higher powers, which can be proved by exactly the same argument.

Theorem 2.5.2. *Let $0 < \delta \leq 1$, and let $N \geq 1$ be a prime integer. Let $f : \mathbb{Z}_N \to [0, 1]$ be a bounded function such that $\mathbb{E}f \geq \delta$. Then we have*

$$\mathbb{E}(f(n)f(n + r^k)|n, r \in \mathbb{Z}_N,\ 1 \leq r \leq \lfloor \sqrt[k]{N/3} \rfloor) \geq c(\delta) - o_\delta(1).$$

We now follow the argument in Section 2.4. Define ν, f, f_1, f_2 as in the proof of Theorem 2.1.2. Applying Theorem 2.5.2 to f_1, we see that

$$\mathbb{E}(f_1(x)f_1(x + r^k)|x, r \in \mathbb{Z}_N, 1 \leq r \leq \sqrt[k]{N/3}) \geq c(\delta) - o_{\delta,\epsilon_0,M}(\eta).$$

To estimate the error terms, we invoke the asymptotic formula for Waring's problem (see e.g. [30]), which implies that

$$R_{k,3k}(x) := |\{(a_1, ..., a_{3k}) \in \mathbb{Z}_N | a_1^k + ... + a_{3k}^k \equiv x(\mod N)\}| \leq cN^2.$$

By convolution and Parseval identities, this translates to

$$||\widehat{P_k}||_{6k} \leq c_1 N^{1/k-1},$$

where P_k denotes the characteristic function of the set of k-th powers smaller than $N/3$, and c, c_1 are constants depending on k. Now we are able to estimate the error terms as in Section 2.4, for example we have

$$\begin{aligned}\mathbb{E}(f_2(x)f_2(x + r)P_k(r)) &\leq ||\hat{P}_k||_{6k} ||\hat{f}_2||_{(12k-1)/(6k-1)}^{(12k-2)/(12k)} ||\hat{f}_2||_\infty^{1/6k} \\ &\leq c_1 C N^{1/k-1} \epsilon_0^{1/6k}.\end{aligned}$$

At the last step we used that (2.2.1) holds with $q = \frac{12k-1}{6k-1}$ if θ_k is small enough. The proof is finished as in Section 2.4.

2.6 Long arithmetic progressions in sumsets

We now turn to Theorem 2.1.4. In this section we prove the theorem, modulo the two main estimates (2.6.1), (2.6.7) which will be proved in the next two sections.

Our proof will combine the arguments of Sanders [35] with those of Green-Tao [27]. Let W, A be as in Theorem 2.1.4, and define ν, f as in Lemma 2.2.7. We will show that, with high probability, there is a reasonably large Bohr set B on which we have $f * f(x) > 0$ for all but a few values of x. But $f * f$ is supported on $A + A$, hence all but a small fraction of B is contained in $A + A$. The proof is concluded by invoking a pigeonholing argument from [35], which says that the portion of B contained in $A + A$ contains a long arithmetic progression.

The details are as follows. Fix k (the length of the progression), and let $\sigma = (16k)^{-1}$. We will also assume that $k > k_0$ and $\alpha < \alpha_0$, where $k_0 \in \mathbb{N}$ is a sufficiently large absolute constant and $\alpha_0 > 0$ is a sufficiently small absolute constant.

By Lemma 2.2.7, the assumptions of Lemma 2.2.6 with $\eta = N^{-1/5}$ and $\epsilon = 1/9$ are satisfied with probability $1 - o(1)$, thus (2.2.1) holds with $q = 19/9$. Let $f = f_1 + f_2$ as in Lemma 2.2.5, with $\epsilon_0 = \epsilon_0(\alpha, \sigma)$ small enough to be fixed later. We will assume that (2.4.3) holds with c_1 sufficiently small; as in Section 2.4, it follows that $\|f_1\|_\infty \leq 2$.

We need an extension of a result of Sanders [35]: there are regular Bohr sets $B := b + B(\Gamma, \delta)$ and $B' := b + B(\Gamma, \delta')$ such that

$$\left|\{x \in B' : (f_1 * f_1)(x) \geq \frac{\alpha^2}{2}|B|\}\right| > (1 - \sigma)|B'|, \tag{2.6.1}$$

and

$$\delta' \gg \frac{\alpha^2 \delta}{|\Gamma|}, \tag{2.6.2}$$

$$\delta \gg (\frac{\alpha}{\log(\sigma^{-1})})^{C \log(\alpha^{-1})}, \tag{2.6.3}$$

$$|\Gamma| \ll \alpha^{-2} \log(\sigma^{-1}). \tag{2.6.4}$$

We establish this in Proposition 2.7.2. We then verify in Section 2.8, via a restriction-type argument, that if

$$\log \frac{1}{\epsilon_0} \gg \alpha^{-2}(\log \frac{1}{\alpha})(\log k)(\log \log k + \log \frac{1}{\alpha}), \tag{2.6.5}$$

with a large enough implicit constant, then

$$\left|\{x \in B' : |f_2 * f_i(x)| \geq \frac{\alpha^2}{10}|B|\}\right| < \sigma|B'|, \ i = 1, 2. \tag{2.6.6}$$

It follows that

$$\left|\{x \in B' : (f * f)(x) \geq \frac{\alpha^2}{10}|B|\}\right| > (1 - 4\sigma)|B'|, \tag{2.6.7}$$

provided that both (2.4.3) and (2.6.5) hold. A somewhat cumbersome calculation shows that ϵ_0 can be chosen so as to satisfy both (2.4.3) and (2.6.5), provided that

$$\log k \ll \frac{\alpha^2 \log\log N}{\log \frac{1}{\alpha}(\log\log\log N + \log \frac{1}{\alpha})}, \tag{2.6.8}$$

which is equivalent to (2.1.4).

We now invoke Lemma 6.5 in [35], which says that if

$$(4\sigma)^{-1} \ll |\Gamma|^{-1}\delta' N^{1/|\Gamma|}, \tag{2.6.9}$$

then the set on the left side of (2.6.7) contains an arithmetic progression of length $(16\sigma)^{-1} = k$. Plugging in (2.6.2)-(2.6.4) and solving for N, we see that (2.6.9) holds if

$$\log N \gg \alpha^{-2}(\log^2 k + \log^2(\frac{1}{\alpha}) + \log\frac{1}{\alpha}\log\log k). \tag{2.6.10}$$

Another cumbersome calculation shows that if we assume (2.6.8), then the additional condition (2.1.3) suffices to guarantee that (2.6.10) holds. Thus, assuming both (2.1.3) and (2.1.4), the set on the left side of (2.6.7) contains a k-term arithmetic progression. Since that set is contained in $A + A$, the conclusion of the theorem follows.

In the next two sections we complete the proof by verifying the inequalities (2.6.1), (2.6.6).

2.7 The main term estimate

Proposition 2.7.1. *Let $B = b + B(\Gamma, \delta)$ be a regular Bohr set. Let $f : \mathbb{Z}_N \to \mathbb{R}$ be a function such that supp$(f) \subset B$, $0 \leq f \leq 1$ and $\mathbb{E}_B f = \alpha > 0$. Fix $\sigma \in (0, 1]$ and let $d = |\Gamma|$. Then one of the following must be true:*

(i) There is a $\delta' \gg \frac{\alpha^2\delta}{d}$ such that $B' = b + B(\Gamma, \delta')$ is regular and

$$\left|\{x \in B' : (f * f)(x) \geq \frac{\alpha^2}{2}|B|\}\right| \geq (1 - \sigma)|B'|, \tag{2.7.1}$$

or

(ii) There is a regular Bohr set $B'' = b'' + B(\Gamma \cup \Lambda, \delta'')$ such that

$$\mathbb{E}(f|B'') \geq \alpha(1 + 2^{-5}), \tag{2.7.2}$$

where $|\Lambda| \ll \alpha^{-2} \log \sigma^{-1}$ and $\delta'' \gg \frac{\alpha^4 \delta}{d^3 \log \sigma^{-1}}$.

Proof: We essentially follow the argument of Sanders [35]; however, some care must be taken to get the right quantitative version. Replacing f by $f(\cdot + b)$ if necessary, we may assume that $b = 0$. Let c_0 be a small enough constant which will be fixed later. By [28], Lemma 8.2, we can find δ' such that

$$\delta' \in (c_0\alpha^2\delta d^{-1}, 2c_0\alpha^2\delta d^{-1}) \tag{2.7.3}$$

and that the set B' defined in (i) is regular. Suppose that (2.7.1) fails for this choice of δ'; we have to prove that this implies (ii).

The failure of (2.7.1) means that we can find a set $S \subset B' \cap \{x : (f * f)(x) < \frac{\alpha^2}{2}|B|\}$ such that $|S| = \sigma|B'|$. Let $g = f - \alpha B$ be the "balanced function" of f. We first claim that

$$\frac{1}{|B||B'|}\sum_{x \in S} g * g(x) \leq -\frac{\alpha^2\sigma}{2} + O(d\delta'\delta^{-1}\sigma). \tag{2.7.4}$$

To prove this, we write

$$\begin{aligned}&\frac{1}{|B||B'|}\sum_{x \in S}(g * g)(x)\\ &\quad = \frac{1}{|B||B'|}\Big(\sum_{x \in S}(f * f)(x) - 2\alpha\sum_{x \in S}(B * f)(x) + \alpha^2\sum_{x \in S}(B * B)(x)\Big).\end{aligned}$$

The first term obeys

$$\frac{1}{|B||B'|}\sum_{x \in S}(f * f)(x) \leq \frac{\alpha^2|B|}{2|B||B'|}|S| = \frac{\alpha^2\sigma}{2}, \tag{2.7.5}$$

by the choice of S. The second term is estimated as in [35]. By [35], Corollary 3.4, we have for $x \in B'$

$$|f * \frac{B}{|B|}(x) - f * \frac{B}{|B|}(0)| \ll d\delta'\delta^{-1}.$$

But $f * \frac{B}{|B|}(0) = \alpha$, so that $f * \frac{B}{|B|}(x) = \alpha + O(d\delta'\delta^{-1})$ for $x \in B'$. Hence

$$\frac{1}{|B'|}\sum_{x\in S}\frac{B}{|B|} * f(x) = \frac{|S|}{|B'|}(\alpha + O(d\delta'\delta^{-1})) = \alpha\sigma + O(d\delta'\delta^{-1}\sigma). \qquad (2.7.6)$$

Finally, we trivially have $B * B(x) \leq |B|$ for all x, hence

$$\frac{1}{|B||B'|}\sum_{x\in S} B * B(x) \leq \sigma + O(d\delta'\delta^{-1}\sigma). \qquad (2.7.7)$$

Combining (2.7.5), (2.7.6), (2.7.7), we get (2.7.4).

We now convert this to a Fourier analytic statement. We have

$$\begin{aligned}\sum_{x\in S} g * g(x) &= \sum_{x\in\mathbb{Z}_N} g * g(x)S(x)\\ &= N\sum_{\xi\in\mathbb{Z}_N}\widehat{g * g}(\xi)\widehat{S}(\xi)\\ &= N^2\sum_{\xi\in\mathbb{Z}_N}|\widehat{g}(\xi)|^2\widehat{S}(\xi).\end{aligned}$$

Hence, by the triangle inequality, (2.7.4) implies that

$$\frac{N^2}{|B||B'|}\sum_{\xi}|\widehat{g}(\xi)|^2|\widehat{S}(\xi)| \geq \frac{\alpha^2\sigma}{2} + O(d\delta'\delta^{-1}\sigma). \qquad (2.7.8)$$

Define

$$\mathcal{L} := \{\xi \in \mathbb{Z}_N : |\widehat{S}(\xi)| \geq \frac{\alpha\sigma|B'|}{4N}\}.$$

We claim that the main contribution to the sum in (2.7.8) comes from $\mathcal{L}$.

In fact

$$\begin{aligned}
\frac{N^2}{|B||B'|}\sum_{\xi\notin\mathcal{L}}|\widehat{g}(\xi)|^2|\widehat{S}(\xi)| &\leq \frac{\alpha\sigma N}{4|B|}\sum_{\xi\notin\mathcal{L}}|\widehat{g}(\xi)|^2\\
&\leq \frac{\alpha\sigma N}{4|B|}\sum_{\xi\in\mathbb{Z}_N}|\widehat{g}(\xi)|^2\\
&= \frac{\alpha\sigma}{4|B|}\sum_{x\in\mathbb{Z}_N}|g(x)|^2\\
&= \frac{\alpha\sigma}{4|B|}\sum_{x\in\mathbb{Z}_N}|f(x)-\alpha B(x)|^2\\
&= \frac{\alpha\sigma}{4|B|}\sum_{x\in\mathbb{Z}_N}f(x)^2 - 2\frac{\alpha^2\sigma}{4|B|}\sum_{x\in\mathbb{Z}_N}f(x)B(x)\\
&\quad + \frac{\alpha^3\sigma}{4|B|}\sum_{x\in\mathbb{Z}_N}B(x)^2\\
&\leq \frac{\alpha^2\sigma}{4} - \frac{2\alpha^3\sigma}{4} + \frac{\alpha^3\sigma}{4}\\
&= \frac{\alpha\sigma}{4}(\alpha-\alpha^2)\\
&\leq \frac{\alpha^2\sigma}{4}.
\end{aligned}$$

Hence

$$\frac{N^2}{|B||B'|}\sum_{\xi\in\mathcal{L}}|\widehat{g}(\xi)|^2|\widehat{S}(\xi)| \geq \frac{\alpha^2\sigma}{4} + O(d\delta'\delta^{-1}).$$

Since $\frac{N}{|B'|}|\widehat{S}(\xi)|$ is trivially bounded by σ, we have

$$\frac{N}{|B|}\sum_{\xi\in\mathcal{L}}|\widehat{g}(\xi)|^2 \geq \frac{\alpha^2}{4} + O(d\delta'\delta^{-1}\sigma). \tag{2.7.9}$$

We now apply the localized version of Chang's theorem proved in [35] (Proposition 4.2) to $S \subset B'$, with $\epsilon = \alpha/4$ and $\eta = 1/2$. We conclude that there is a set $\Lambda \subset \mathbb{Z}_N$ and a $\delta_0'' > 0$ such that

$$|\Lambda| \ll \frac{2^4}{\alpha^2}\log\sigma^{-1},$$

$$\delta_0'' \gg \frac{\delta'\alpha^2 4}{d^2\log\sigma^{-1}},$$

and

$$\mathcal{L} \subset \{\xi \in \mathbb{Z}_N : |1 - e^{-2\pi i x\xi/N}| \leq 1/2 \; \forall \; x \in B(\Gamma \cup \Lambda, \delta_0'')\}.$$

Choose $\delta'' \in (\delta_0'', 2\delta_0'')$ such that $B'' := B(\Gamma \cup \Lambda, \delta'')$ is regular. Note that this together with (2.7.3) implies that δ'' obeys the condition in (ii). We may also assume that $\delta'' < \delta'$. Our goal is to get the L^2 density increment as in (2.7.2) on a translate of B''.

By the definition of $\mathcal{L}$, we have $\frac{N}{|B''|}|\widehat{B''}(\xi)| \geq 1/2$ for all $\xi \in \mathcal{L}$. Hence

$$\frac{N^3}{|B||B''|^2}\sum_{\xi \in \mathcal{L}} |\widehat{g}(\xi)|^2 |\widehat{B''}(\xi)|^2 \geq \frac{\alpha^2}{16} + O(d\delta'\delta^{-1}).$$

Again using Plancherel's identity and the convolution identity we have

$$\begin{aligned}\alpha^2\Big(\frac{1}{16} + O(\alpha^{-2}d\delta'\delta^{-1})\Big) &\leq \frac{N^3}{|B||B''|^2}\sum_{\xi \in \mathbb{Z}_N} |\widehat{g}(\xi)|^2 |\widehat{B''}(\xi)|^2 \\ &= \frac{N^3}{|B||B''|^2}\sum_{\xi \in \mathbb{Z}_N} |N^{-1}\widehat{g * B''}(\xi)|^2 \\ &= \frac{1}{|B||B''|^2}\sum_{x \in \mathbb{Z}_N} |g * B''(x)|^2.\end{aligned}$$

We now apply Lemma 5.2 from [35] and conclude that

$$\begin{aligned}\frac{1}{|B''|}\sup_{x \in \mathbb{Z}_N} |f * B''(x)| &\geq \alpha\big(1 + 2^{-4} + O(\alpha^{-2}d\delta'\delta^{-1})\big) + O(d\delta''\delta^{-1}) \\ &\geq \alpha\big(1 + 2^{-4}\big) + O(d\alpha^{-1}\delta'\delta^{-1}).\end{aligned}$$

We now let the constant c_0 in (2.7.3) be small enough, so that the error term is bounded by $\alpha 2^{-5}$. The conclusion (ii) follows if we choose b'' to maximize $|f * B''(b'')|$. This proves Proposition 2.7.1.

Proposition 2.7.2. *Let $f : \mathbb{Z}_N \to [0,1]$ be defined such that*

$$\mathbb{E}_{x \in \mathbb{Z}_N} f(x) = \alpha > 0.$$

Let $\sigma \in (0,1]$. Then there exist Bohr sets $B := b + B(\Gamma, \delta)$ and $B' := b + B(\Gamma, \delta')$ such that

$$\Big|\{x \in B' : \; (f * f)(x) \geq \frac{\alpha^2}{2}|B|\}\Big| > (1 - \sigma)|B'|,$$

and

$$\delta' \gg \frac{\alpha^2\delta}{|\Gamma|},$$

$$\delta \gg \big(\frac{\alpha}{\log(\sigma^{-1})}\big)^{C\log(\alpha^{-1})},$$

and

$$|\Gamma| \ll \alpha^{-2}\log(\sigma^{-1}).$$

Proof of Proposition 2.7.2: We construct the Bohr sets B and B' by iterating Proposition 2.7.1. Let $\Gamma_0 := \{0\}$, and pick $\delta_0 \gg 1$ so that $B(\Gamma_0, \delta_0)$ is regular. Define $\alpha_0 := \alpha$. Averaging over translates of $B(\Gamma_0, \delta_0)$, we see that there is a b_0 such that $\mathbb{E}(f|B_0) \geq \alpha_0$ for $B_0 = b_0 + B(\Gamma_0, \delta_0)$. By Proposition 2.7.1, one of the following must hold:

(i) There is a $\delta_0' \gg \frac{\alpha_0^2\delta_0}{|\Gamma_0|}$ such that $B_0' := b_0 + B(\Gamma_0, \delta_0')$ is regular and

$$\Big|\{x \in B_0' : (f * f)(x) \geq \frac{\alpha_0^2}{2}|B_0|\}\Big| \geq (1-\sigma)|B_0'|, \tag{2.7.10}$$

(ii) There is a regular Bohr set $B_1 := b_1 + B(\Gamma_0 \cup \Lambda_0, \delta_1)$ such that

$$\mathbb{E}(f|B_1) \geq \alpha_0(1 + 2^{-5}), \tag{2.7.11}$$

where $|\Gamma_0| \ll \alpha_0^{-2}\log(\sigma^{-1})$ and $\delta_1 \gg \frac{\alpha_0^4\delta_0}{|\Gamma_0|^3\log(\sigma^{-1})}$.

If (i) holds, we let $B' = B_0'$ and we are done. If on the other hand (ii) holds, we repeat the procedure with B_0 replaced by B_1, and continue by induction. If we have not satisfied (i) by the end of the kth step, we have found a regular Bohr set $B_k := b_k + B(\Gamma_k, \delta_k)$ such that

$$\mathbb{E}(f|B_k) = \alpha_k|B_k|,$$

where

$$\alpha_k \geq \alpha_{k-1}(1 + 2^{-5}), \tag{2.7.12}$$

$$\delta_k \gg \frac{\alpha_{k-1}^4\delta_{k-1}}{|\Gamma_{k-1}|^3\log(\sigma^{-1})}, \tag{2.7.13}$$

and

$$|\Gamma_k| - |\Gamma_{k-1}| \ll \alpha_{k-1}\log(\sigma^{-1}). \tag{2.7.14}$$

The iteration must terminate (upon reaching density 1 on a large enough Bohr set) after at most

$$k \ll \log(\alpha^{-1})$$

steps, since from (2.7.12) we have

$$\alpha_k^2 \geq \alpha^2(1+2^{-5})^{k-1}.$$

By (2.7.14) we have

$$\begin{aligned}|\Gamma_k| &\ll \alpha_{k-1}^{-2}\log(\sigma{-}1) + \alpha_{k-2}^{-2}\log(\sigma^{-1}) + \cdots + \alpha_0^{-2}\log(\sigma^{-1}) \\ &\leq \alpha^{-2}\log(\sigma^{-1})\sum_{j=0}^{\infty}(1+2^{-5})^{-j} \ll \alpha^{-2}\log(\sigma^{-1}).\end{aligned}$$

Finally, using our bounds for α_k and $|\Gamma_k|$, we have

$$\delta_k \gg \Big(\frac{\alpha}{\log(\sigma^{-1})}\Big)^{C\log(\alpha^{-1})},$$

for some absolute constant $C > 0$. This proves Proposition 2.7.2.

2.8 The restriction argument

Assume that the hypotheses of Theorem 2.1.4 hold. We need to show that if f_1, f_2 are as in Lemma 2.2.5 and B, B' are the Bohr sets chosen in Proposition 2.7.2, then (2.6.6) holds, i.e.

$$\Big|\{x \in B' : \ |f_2 * f_i(x)| \geq \frac{\alpha^2}{10}|B|\}\Big| \leq \sigma|B'|, \ i = 1, 2. \tag{2.8.15}$$

It suffices to prove that

$$\|f_i * f_2\|_{L^2(B')}^2 \leq \frac{\alpha^4}{200}\sigma|B|^2|B'|. \tag{2.8.16}$$

We have

$$\begin{aligned}
\|f_i * f_2\|_{L^2(B')}^2 &= \sum_{x\in B'} (f_i * f_2)^2(x) \\
&= \sum_{x\in B'} \Big(\sum_y f_i(y) f_2(x-y)\Big)\Big(\sum_z f_i(z) f_2(x-z)\Big) \\
&= \sum_{x,y,z,u,v} B'(x) f_i(y) f_2(z) \frac{1}{N} \sum_\xi e^{-2\pi i(y+z-x)\xi/N} \\
&\quad \cdot f_i(u) f_2(v) \frac{1}{N} \sum_\eta e^{-2\pi i(u+v-x)\eta/N} \\
&= N^3 \sum_{\xi,\eta} \widehat{B'}(-\eta-\xi)\widehat{f_i}(\xi)\widehat{f_2}(\xi)\widehat{f_i}(\eta)\widehat{f_2}(\eta) \\
&= N^3 \sum_\xi (\widehat{B'} * \widehat{f_i}\widehat{f_2})(-\xi)\widehat{f_i}(\xi)\widehat{f_2}(\xi).
\end{aligned}$$

By Hölder's inequality,

$$\|f_i * f_2\|_{L^2(B')}^2 \le N^3 \|\widehat{B'} * \widehat{f_i}\widehat{f_2}\|_{10}\, \|\widehat{f_i}\widehat{f_2}\|_{10/9}. \tag{2.8.17}$$

Applying Young's inequality, we get

$$\|\widehat{B'} * \widehat{f_i}\widehat{f_2}\|_{10} \le \|\widehat{B'}\|_5\, \|\widehat{f_i}\widehat{f_2}\|_{10/9}. \tag{2.8.18}$$

Furthermore,

$$\begin{aligned}
\|\widehat{f_i}\widehat{f_2}\|_{10/9}^{10/9} &\le \|\widehat{f_2}\|_\infty^{1/9} \sum_\xi |\widehat{f_2}(\xi)|\, |\widehat{f_i}(\xi)|^{10/9} \\
&\le \|\widehat{f_2}\|_\infty^{1/9}\, \|\widehat{f_2}\|_{19/9}\, \|\widehat{f_i}(\xi)\|_{19/9}^{10/9},
\end{aligned}$$

where at the last step we used Hölder's inequality again. Plugging this together with (2.8.18) in (2.8.17), we see that

$$\begin{aligned}
\|f_i * f_2\|_{L^2(B')}^2 &\le N^3 \|\widehat{B'}\|_5\, \|\widehat{f_i}\widehat{f_2}\|_{10/9}^2 \\
&\le N^3 \|\widehat{B'}\|_5 \left(\|\widehat{f_2}\|_\infty^{1/9}\, \|\widehat{f_2}\|_{19/9}\, \|\widehat{f_i}\|_{19/9}^{10/9}\right)^{9/5} \\
&\le N^3 \|\widehat{B'}\|_5\, \|\widehat{f_2}\|_\infty^{1/5}\, \|\widehat{f_2}\|_{19/9}^{9/5}\, \|\widehat{f_i}\|_{19/9}^2.
\end{aligned}$$

By Plancherel's theorem and Lemma 2.2.5(iv), we have

$$\|\widehat{f_i}\|_2^2 \le \|\widehat{f}\|_2^2 = N^{-1}\|f\|_2^2 \ll \alpha p^{-1} = \alpha N^\theta.$$

Since $\theta < 1/20$, it follows from Lemma 2.2.6 that

$$\|\widehat{f}\|_{19/9} = O(1) \text{ and } \|\widehat{f}_i\|_{19/9} = O(1), \;\; i = 1, 2.$$

By Lemma 2.2.5(iii), we have

$$\|\widehat{f}_2\|_\infty \leq C\epsilon_0.$$

Finally,

$$\|\widehat{B'}\|_5^5 \leq \|\widehat{B'}\|_\infty^3 \|\widehat{B'}\|_2^2 \leq \frac{|B'|^3}{N^3} \|\widehat{B'}\|_2^2 = \frac{|B'|^4}{N^4}.$$

Combining these estimates, we get

$$\|f_i * f_2\|_{L^2(B')}^2 \ll N^3 \epsilon_0^{1/5} \frac{|B'|^{4/5}}{N^{4/5}}. \tag{2.8.19}$$

We need the right side of this to be smaller than $\frac{\alpha^4}{200}\sigma|B|^2|B'|$, i.e. we need to have

$$\epsilon_0^{1/5} \leq c\alpha^4 \sigma \frac{|B|^2}{N^2} \frac{|B'|^{1/5}}{N^{1/5}} = c\alpha^4 \sigma \mathbb{P}(B)^2 \mathbb{P}(B')^{1/5}. \tag{2.8.20}$$

But by Lemma 2.2.3 and (2.6.2)–(2.6.4), $\mathbb{P}(B)$ and $\mathbb{P}(B')$ are bounded from below by

$$\mathbb{P}(B) \geq \mathbb{P}(B') \gg (c\delta'')^{|\Gamma|} \gg \left(\frac{c\alpha}{\log k}\right)^{c\alpha^{-2} \log \frac{1}{\alpha} \log k},$$

where we plugged in $\sigma = (16k)^{-1}$. Hence (2.8.20) holds if

$$\epsilon_0 \ll \alpha^{28} k^{-9} \left(\frac{c\alpha}{\log k}\right)^{c\alpha^{-2} \log \frac{1}{\alpha} \log k}. \tag{2.8.21}$$

A short calculation shows that (2.6.5) is sufficient to guarantee that (2.8.21) is satisfied.

2.9 Proof of Proposition 2.1.3

Let $0 < \sigma < (\alpha - \beta)/10$. Define ν, f, f_1, f_2 as in Section 2.6, except that instead of (2.4.1) we will require

$$\|f_1\|_\infty \leq 1 + \sigma, \tag{2.9.1}$$

which holds for large enough N (depending on σ and on the ϵ_0 in the definition of f_i) by the same argument as in Section 2.4.

It clearly suffices to prove that

$$\Big|\{x \in \mathbb{Z}_N : \ f * f(x) > 0\}\Big| \geq (\alpha - 10\sigma)N. \tag{2.9.2}$$

Indeed, (2.9.2) shows that the sumset $A + A$ in $\mathbb{Z}_N$ has size at least βN, hence so does the sumset $A + A$ in $\mathbb{Z}$.

We first claim that if N is large enough, then

$$\Big|\{x \in \mathbb{Z}_N : \ f_1 * f_1(x) \geq \sigma\alpha N\}\Big| \geq (\alpha - 3\sigma)N. \tag{2.9.3}$$

To see this, we first note that

$$\|f_1 * f_1\|_1 = \|f_1\|_1^2 = \alpha^2 N^2 (1 + O(N^{-1/5})). \tag{2.9.4}$$

On the other hand, if (2.9.3) failed, we would have

$$\begin{aligned}\|f_1 * f_1\|_1 &\leq \sigma\alpha N \cdot N + \alpha N(1 + \sigma + O(N^{-1/5})) \cdot (\alpha - 3\sigma)N \\ &= \alpha^2 N^2 (1 + O(N^{-1/5})) - \sigma\alpha N^2,\end{aligned}$$

which contradicts (2.9.4). This proves (2.9.3).

The proof of (2.9.2) will be complete if we can show that

$$\Big|\{x \in \mathbb{Z}_N : |\ f_i * f_2(x)| \geq \frac{\sigma\alpha}{10} N\}\Big| \leq \sigma N. \tag{2.9.5}$$

To this end, we repeat the argument in Section 2.8. It suffices to prove that

$$\|f_i * f_2\|_2^2 \leq \frac{\sigma^2\alpha^2}{200} \sigma N^3. \tag{2.9.6}$$

As in Section 2.8 (with $B = B' = \mathbb{Z}_N$), we have

$$\|f_i * f_2\|_2^2 \ll \epsilon_0^{1/5} N^3, \tag{2.9.7}$$

and the right side is smaller than the right side of (2.9.6) if $\epsilon_0 \ll \sigma^{27}\alpha^{18}$, with a small enough implicit constant. Thus (2.9.5) holds for large enough N if ϵ_0 was chosen small enough. This proves Proposition 2.1.3

2.10 Acknowledgements

The authors were supported in part by an NSERC Discovery Grant. We are grateful to Ben Green for suggesting the feasibility of Theorem 2.1.2, and to Ernie Croot and Mihalis Kolountzakis for helpful discussions and suggestions.

Bibliography

[18] J. Bourgain, *On arithmetic progressions in sums of sets of integers*, in *A tribute to Paul Erdős*, pp. 105-109, Cambridge University Press, 1990.

[19] J. Bourgain, *On triples in arithmetic progressions*, Geom. Funct. Anal. **9** (1999), 968–984.

[20] M.-C. Chang, *A polynomial bound in Freiman's theorem*, Duke Math. J. **113** (2002), 399–419.

[21] E. Croot, I.Ruzsa, T. Schoen, *Long arithmetic progressions in sparse sumsets*, Integers: The Electronic Journal of Combinatorial Number Theory, **7**(2) (2007), #A10.

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

[23] B.J. Green, *Arithmetic progressions in sumsets*, Geom. Funct. Anal. **12** (2002), 584–597.

[24] B.J. Green, *On arithmetic structures in dense sets of integers*, Duke Math. Jour., **114** (2002), 215-238.

[25] B.J. Green, *Roth's Theorem in the primes*, Annals of Math. 161 (2005), 1609-1636.

[26] B.J. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Annals of Math., to appear.

[27] B.J. Green, T. Tao, *Restriction theory of the Selberg Sieve, with applications*, Journal de Théorie des Nombers de Bordeaux, **18** (2006), 137–172.

[28] B.J. Green, T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc., to appear.

[29] Y. Kohayakawa, T. Łuczak, V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), 133–163.

[30] M. B. Nathanson, *Additive Number Theory: The Classical Bases*, Springer, New York, 1996.

[31] J. Pintz, W.L. Steiger, E. Szemerédi, *On sets of natural numbers whose difference set contains no squares*, J. London Math. Soc. **37** (1988), 219-231.

[32] K. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.

[33] I. Ruzsa, *Difference sets without squares*, Period. Math. Hungar. **15** (1984), no. 3, 205-209.

[34] I. Ruzsa, *Arithmetic progressions in sumsets*, Acta Arith. **60** (1991), 191–202.

[35] T. Sanders, *Additive structures in sumsets*, to appear in Math. Proc. Cambridge Philos. Soc.

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

[37] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299–345.

[38] T. Tao, *Arithmetic progressions and the primes*, Collect. Math. (2006), Vol. Extra, 37-88.

[39] T. Tao, V. Vu, *Additive combinatorics*, Cambridge Univ. Press, 2006.

[40] P. Varnavides, *On certain sets of positive density*, Journal London Math. Soc., **34** (1959), 358–360

Department of Mathematics, University of British Columbia,
Vancouver, B.C. V6T 1Z2, Canada
mehamel@math.ubc.ca, ilaba@math.ubc.ca

Chapter 3

Conclusion

3.1 A potential application to primes

One application of the methods used in this paper involves translating results from the random setting into that of the primes. A theorem of Green [47] states that if A is a subset of the primes with positive relative density, then A must contain three terms in arithmetic progression. The strategy developed by Green in [47] is essentially to exploit certain 'random' properties of the primes (in a Fourier analytic sense). One of the first hurdles to overcome is that the primes aren't actually random. For example, the set of primes isn't randomly distributed modulo a small prime. To deal with this, one employs what is today referred to as the $W - trick$ (a nice explanation of this can be found in [49]). This entails defining a number m which is the product of small primes. The next step is to consider the residue classes modulo m. The set of primes contained in a given residue class will then behave in a satisfactorily random manner. The proof then restricts itself to the portion of the set A which falls in one particular residue class (pick the residue class on which A has largest relative density). Working with this chosen portion of A one applies Roth's theorem in a manner similar to the application of Sárközy's theorem in the proof of Theorem 2.1.2. Without providing details, we note that the appropriate choice for ν is a modified version of the von Mangoldt function supported on the chosen residue class.

Combining the results from the manuscript [51] contained in this dissertation, and the framework for the primes of Green [47] or Green and Tao [49] we can prove a version of Sárközy's theorem in the primes (and a similar extension for long arithmetic progressions in sumsets). Specifically, it is possible to show that if A is a subset of the primes with positive relative density, then A must contain a square difference. However, we should mention, that in the case of Sárközy's theorem such a variation holds in the primes for density reasons alone. Pintz, Steiger and Szemerédi [55] show that if a subset $A \subset \{1, ..., N\}$ contains no square difference, the $|A|/N \leq (\log N)^{-c \log \log \log \log N}$. For comparison, the prime number theorem states that the number of primes less than an integer N is asymptotic to $N/\log N$.

In joint work with Karsten Chipeniuk, we are attempting to prove an analogue of Theorem 2.1.3 in the setting of the primes. We begin with a motivating example.

Let P be the set of all primes and define $A := \{p \in P : p \equiv 1 \mod n\}$ and consider $A_N := \{p \in A : p \leq N\}$. Then, the prime number theorem for arithmetic progressions gives us quantitative information on the size of A_N. Namely, we have

$$|A_N| = \frac{1}{\phi(n)} \frac{N}{\log N} + O(N/\log^2 N).$$

Set $\delta := \frac{1}{\phi(n)}$. Then, assuming N is sufficiently large, we have $|A_N| \gg \delta \frac{N}{\log N}$. On the other hand, $A_N + A_N \subset \{m \equiv 2 \mod n\}$ and hence,

$$|A_N + A_N| \leq \frac{2N}{n} \sim \frac{\delta}{\log\log n} N.$$

The previous example leads us to the following question:

Question 3.1.1. *Suppose $A \subset \mathrm{P}$ such that*

$$\limsup_{N \to \infty} \frac{|A \cap \{1, ..., N\}|}{|\mathrm{P}_N|} = \delta$$

where $\mathrm{P}_N := \{p \leq N : p \in \mathrm{P}\}$. Is it true that

$$|A_N + A_N| \gg \frac{\delta}{\log\log(1/\delta)} \cdot N?$$

The question above differs from other applications in the primes since for a density result, we are not able to consider only one residue class modulo m. If p is a prime number, then the residue class in which p lies must be contained in the multiplicative subgroup $\mathbb{Z}_m^* \subset \mathbb{Z}_m$. While we are able to apply a modified version of the convolution lemma from [51] to a given residue class on which a subset of the primes has large enough relative density, we must also ensure that the sumset determined by these residue classes covers enough of $\mathbb{Z}_m$. In particular we must answer the following question:

Question 3.1.2. *Suppose $S \subset \mathbb{Z}_m^*$ such that $|S| \geq \delta\phi(m)$. Is it then true that $|S + S| \geq \frac{\delta}{\log\log(1/\delta)} m$?*

We intend to answer these questions in an upcoming paper.

3.2 An improvement to Theorem 2.1.2

Recall that in the statement of Theorem 2.1.2 we define the random set $W \subset \mathbb{Z}_N$ with each element chosen with probability $p(N) \in (cN^{-\theta}, 1]$ where we require $0 < \theta < 1/110$. As we noted in Section 2.1, while we expect that the range should be extended to all $\theta < 1/2$ the methods of our paper do not seem sufficient to establish such a result. In a work [54] that is currently in preparation, H. Nguyen and V. Vu have proved:

Theorem 3.2.1. *Suppose that W is a random subset of $\mathbb{Z}_N$ such that the events $x \in W$, where x ranges over $\mathbb{Z}_N$, are independent and have probability $p = p(N) \in (cN^{-\theta}, 1]$ where $0 < \theta < 1/2$. Let $\alpha > 0$. Then the statement*

> *for every set $A \subset W$ with $|A| \geq \alpha W$, there are $x, y \in A$ such that $x - y$ is a non-zero perfect square*

is true with probability $1 - o_\alpha(1)$ as $N \to \infty$.

They are also able to prove similar results for k-th powers with the $\theta < 1/2$ replaced by $\theta < 1/k$. Their proof uses methods from graph theory relating to the cited work of Kohayakawa, Luczak, and Rödl and in particular they use Szeméredi's regularity lemma.

3.3 Future directions

A famous result of Bergelson and Leibman [42] is a polynomial version of Szemerédi's theorem.

Theorem 3.3.1. Bergelson-Leibman *Let $\delta > 0$ and let $P_1, ..., P_k$ be polynomials in $\mathbb{Z}[d]$ such that $P_i(0) = 0$ for every $i = 1, ..., k$. Suppose that N is sufficiently large depending on δ and $P_1, ..., P_k$. Then, there exist integers m and d so that $m + P_i(d) \in A$ for all $i = 1, ..., k$.*

Bergelson and Leibman's proof uses ergodic theory and is currently the only known proof of a polynomial version of Szemerédi's theorem. Recently, Tao and Ziegler [59] proved that the primes contain arbitrarily long polynomial progressions, relying on a certain quantitative version of Theorem 3.3.1. Their proof, similarly to the proof of the Green-Tao theorem on primes in arithmetic progression, can be divided into three main steps: a polynomial version of Szemerédi's theorem, a transference principle and a method to treat the primes as 'random'. We should note that despite the similarities in the outline, the polynomial version requires several arguments which are not needed in the case of finding arithmetic progressions in the primes.

There have been certain results using Fourier analysis relating to the polynomial version of Szemerédi's theorem. One reason that this is of interest is that the ergodic proof, relying on the axiom of choice, does not provide any information on when the first occurrence of a given polynomial pattern must occur in a subset of positive upper density. The first quantitative result in this direction is due to Green [46]:

Theorem 3.3.2. *Suppose that $A \subset \{1, ..., N\}$ so that $|A|/N \geq (\log\log N)^{-c}$. Then there exists x, $x+y$, $x+2y \in A$ such that $y = a^2 + b^2$.*

His proof requires quadratic Fourier analysis. We notice, that using these methods, Green is able to deduce quantitative bounds on the required density for a subset of the integers to contain this particular arithmetic progression. Since Green's current bound is not sufficient for handling subsets of the primes, we believe that it would be of interest to consider this problem in the random setting. We expect that the method of Green-Tao for handing four-term arithmetic progressions in the primes should provide the necessary framework for such an application. More specifically, if we assume that W is a random subset of $\mathbb{Z}_N$ and $A \subset W$ has positive relative density then we would like to construct f and ν as in Section 2.2. For such an argument, we expect that it would be necessary to replace the pseudorandom condition $\|\widehat{\nu}(\xi) - \mathbf{1}_{\xi=0}\|_\infty \leq \eta$ with an appropriate quadratic condition.

A result of Lyall and Magyar [53] can be seen as a special case of the polynomial version of Szemerédi's theorem or as a generalization of Sárközy's theorem.

Theorem 3.3.3. *Assume that $P_1, ..., P_l \in \mathbb{Z}[d]$ are linearly independent polynomials so that $P_i(0) = 0$ for each $i = 1, ..., l$ and assume the largest degree of the polynomials P_i is k. Suppose that $A \subset \{1, ..., N\}$ so that $|A|/N \geq \left((\log\log N)^2/\log N\right)^{1/l(k-1)}$. Then there exists an integer d such that $P_i(d) \in A - A$ for every $i = 1, ..., k$.*

Their proof is Fourier analytic and again provides quantitative bounds that the ergodic proof does not. We expect that using similar methods to those in this manuscript we could extend Theorem 3.3.3 to subsets of random sets. We should remark that Lyall and Magyar are optimistic that combining their proof with the methods of Pintz, Steiger and Szemerédi would result in a similar bound to that known for Sárközy's theorem. This method would have the advantage of including all sets of such density, rather than subsets of random sets or subsets of the primes.

Finally, we believe that it would be of interest to find a Fourier analytic proof of the following special case of the Bergelson-Leibman theorem which can be compared with both Theorem 3.3.2 and Theorem 3.3.3:

Theorem 3.3.4. *Suppose that $\delta > 0$ and assume A is a subset of the integers with positive upper density. Then there must exist integers x and y such that x, $x + y^2$, $x + y^3 \in A$.*

We expect that such a result would require quadratic Fourier analysis, which would enable us to determine a bound on how soon we could find three elements of the form x, $x + y^2$ and $x + y^3$ in A. If we are then able to extend this to show that, in fact, A must contain 'many' triples of the desired form, then we would hope to be able to prove a version of Theorem 3.3.4 in the random setting or in the primes. Such a result would provide new quantitative information for a special case of the Theorem of Tao-Ziegler on polynomial progressions in the primes.

Bibliography

[41] F. Behrend, *On the sets of integers which contain no three terms in arithmetic progression.* Proc. Nat. Acad. Sci., 23:331-332, 1946.

[42] V. Bergelson and A. Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. 9 (1996), no. **3**, 725-753.

[43] J. Bourgain, *On triples in arithmetic progressions*, Geom. Funct. Anal. **9** (1999), 968–984.

[44] G. A. Freiman, *Foundation of a structural theory of set addition (translated from Russian)*, Translations of Mathematical Monographs, 37, 1973.

[45] T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal., 12:465-588, 2001.

[46] B.J. Green, *On arithmetic structures in dense sets of integers*, Duke Math. Jour., **114** (2002), 215-238.

[47] B.J. Green, *Roth's Theorem in the primes*, Annals of Math. 161 (2005), 1609-1636.

[48] B.J. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Annals of Math., to appear.

[49] B.J. Green, T. Tao, *Restriction theory of the Selberg Sieve, with applications*, Journal de Théorie des Nombers de Bordeaux, **18** (2006), 137–172.

[50] B.J. Green, T. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*, Proc. Edinburgh Math. Soc., to appear.

[51] M. Hamel and I. Laba, *Arithmetic structures in random sets*, INTEGERS: Electronic Journal of Combinatorial Number Theory, 8, 2008.

[52] Y. Kohayakawa, T. Łuczak, V. Rödl, *Arithmetic progressions of length three in subsets of a random set*, Acta Arith. **75** (1996), 133–163.

[53] N. Lyall and A. Magyar, *Polynomial configurations in difference sets*, preprint.

[54] H. Nguyen and V. Vu, in preparation.

[55] J. Pintz, W.L. Steiger, E. Szemerédi, *On sets of natural numbers whose difference set contains no squares*, J. London Math. Soc. **37** (1988), 219-231.

[56] K. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245–252.

[57] T. Sanders, *Additive structures in sumsets*, to appear in Math. Proc. Cambridge Philos. Soc.

[58] A. Sárkozy, *On difference sets of integers I*, Acta Math. Acad. Sci. Hungar, 31:125-149, 1978.

[59] T. Tao and T. Ziegler, *The primes contain arbitrarily long polynomial progressions*, to appear in Acta. Math.

[60] T. Tao, V. Vu, *Additive combinatorics*, Cambridge Univ. Press, 2006.

[61] P. Varnavides, *On certain sets of positive density*, Journal London Math. Soc., **34** (1959), 358–360

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA,
VANCOUVER, B.C. V6T 1Z2, CANADA
mehamel@math.ubc.ca