Three Problems in Additive Combinatorics



Ruoyi Wang
Balliol College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term 2021

To my mother and father

Acknowledgements

Firstly, I would like to thank my supervisor Prof. Tom Sanders for suggesting interesting problems to work on, for many insightful discussions, for a careful reading of my first paper, and for his continuous encouragement during the pandemic.

Thanks to Prof. James Maynard for pointing out a mistake in an early version of Chapter 1 and many interesting mathematical discussions; to Prof. Ben Green for supervising a mini-project about Kakeya sets which introduced me to additive combinatorics; to Dr. Joni Teräväinen for discussing and suggesting literature relevant to Chapter 2 and Chapter 4; to Dr. Aled Walker, Dr. Sofia Lindqvist, Oliver McGrath and Jared Lichtman for discussions; to everyone in the reading group for presenting interesting papers; to Candida Bowtell for being a wonderful officemate; and to an anonymous referee who provided comments on the work that became part of the thesis.

Thanks to Prof. Coralia Cartis for her generous help and guidance during some of the more difficult times; to the OxPDE group and Dr. Avi Mayorcas for supporting my change of research area; to Dr. Toni Scharle, Alessandro Lodi, Serte Donderwinkel and fellow members of the Clarendon Council for sharing some of the happiest moments; and to the University of Oxford for the fantastic academic resources and extracurricular opportunities, which benefitted me on my way towards completing the thesis.

Thanks to the faculty members of the University of Edinburgh for a stimulating undergraduate education, without which I would not have considered pursuing a further degree in mathematics; and in particular to Prof. Jim Wright and Prof. Ivan Cheltsov for supervising undergraduate projects and continuous encouragement during the postgraduate degree.

My DPhil study has been generously supported by a Clarendon Scholarship. I would like to thank the Clarendon Fund, the Mathematical Institute and Balliol College, for the funding and their continuous support throughout the course.

Finally, and most importantly, this D. Phil degree could not have been completed without the love and support of my family and friends. In particular, I would like to thank my mother for regularly visiting me from abroad throughout the degree, including in recent years when travel became very difficult.

Abstract

The problems considered in this thesis involve several fundamental objects in additive combinatorics. In this project, we mainly focus on the structure of subsets of integers. As integers are the most fundamental objects in number theory, one of our main interests in this project is to understand additive combinatorial problems related to sets which arise naturally in number theory, such as prime numbers and smooth numbers.

In Chapter 1, we study the size of subsets of smooth numbers which are free of geometric progressions. By taking a Freiman isomorphism, we shall use additive results to study the multiplicative structure in smooth numbers. Our result is proved in Section 1.2 and Section 1.3. For completeness, we illustrate a typical counterexample in Section 1.4, namely the set of k-free smooth numbers.

In the second problem, we focus on difference sets which do not contain shifted primes. Let $\mathbb{P}$ denote the set of the prime numbers. Suppose $A \subseteq [N]$ is nonempty and it satisfies $(A-A) \cap (\mathbb{P}-1) = \emptyset$. We are primarily interested in estimating the cardinality of the set A. In Chapter 2, we prove a new upper bound on the cardinality of A. In Chapter 3, we investigate sets A such that $nA - nA$ avoids $\mathbb{P}-1$. We first prove an upper bound on $|A|$ which depends on n, then we generalise an example of Ruzsa.

In the third problem, we are interested in the quantitative aspect of partition regularity of the equation $x+y=z$ when x, y, z are restricted to shifted primes $\mathbb{P}-1$. In Chapter 4, we provide an exposition of the work of Li and Pan, which is the first result concerning partition regularity over shifted primes. We focus on the analytical ingredients of their transference principle and the quantitative bound that follows from it. In Chapter 5, we improve the bound produced by the method of Li and Pan using a different approach. The main ingredient of the proof is an iteration lemma from Chapter 2.

Contents

0 Preliminaries **1**

0.1 Notation 1

0.2 Additive combinatorics and Fourier analysis 3

0.2.1 The basics 3

0.2.2 The large spectrum and Bohr sets 5

0.2.3 Fourier concentration and energy increment 7

0.3 Analytic number theory 9

0.3.1 Zeros of Dirichlet L-functions 9

0.3.2 A minor arc estimate 11

1 Geometric progressions in smooth numbers **13**

1.1 Introduciton 13

1.2 An argument using an approximate structure 16

1.3 Proof of the main result 19

1.4 A counterexample: k-free integers 23

2 Difference sets without shifted primes **27**

2.1 Introduction 27

2.2 Zeros of Dirichlet L-functions 29

2.3 Major arc estimates 30

2.4 Ruzsa–Sanders iteration argument with modifications 41

2.4.1 Modified main iteration lemma 42

2.4.2 Proof of the main result 47

3 Iterated sumset $nA - nA$ without shifted primes **52**

3.1 Introduction 52

3.2 Proof of the upper bound 53

3.3 A generalisation of Ruzsa's example 56

4 Partition regularity over shifted primes, I **59**
4.1 Introduction 59
4.2 A quantitative variant of Schur's theorem 61
4.3 A Fourier-analytic transference principle 64
4.3.1 The L^∞ approximation 66
4.3.2 Proof of Proposition 4.3.3 68
4.4 The Fourier restriction estimate for the primes 69
4.5 Proof of the Li–Pan bound 70

5 Partition regularity over shifted primes, II **75**
5.1 Introduction 75
5.2 The bootstrapping lemma 76
5.3 The main result 82
5.4 Transference principle revisited 86
5.5 Partition regularity of $x + y = z$ over $\mathbb{P} - 1$ under the generalised Riemann hypothesis 89

Bibliography **94**

Chapter 0

Preliminaries

0.1 Notation

We use both the Landau notation and the Vinogradov notation to describe inequalities. For example, both $f(n) = O(g(n))$ and $f(n) \ll g(n)$ mean that there exists an absolute constant $c > 0$ such that $f(n) \leq cg(n)$ for all n; both $f(n) = \Omega(g(n))$ and $f(n) \gg g(n)$ mean that there exists an absolute constant $c > 0$ such that $f(n) \geq cg(n)$ for all n. The Landau notation $f(n) = o_{n\to\infty}(g(n))$ means that $\lim_{n\to\infty} f(n)/g(n) = 0$, and $f(n) = \omega_{n\to\infty}(g(n))$ means that $g(n) = o_{n\to\infty}(f(n))$. We shall omit the variable n when it is clear from the context. We shall use subscripts to indicate that the constant in the notation above implicitly depends on other parameters, for instance, the expression $f(n) \ll_{\epsilon,\delta} g(n)$ says that $f(n) \leq c_{\epsilon,\delta} g(n)$ for all n, where $c_{\epsilon,\delta}$ is a positive constant whose choice depends on parameters ϵ and δ.

We shall use the following notation to denote some commonly used sets.

$\mathbb{R}$: the set of real numbers.

$\mathbb{C}$: the set of complex numbers.

$\mathbb{N}$: the set of positive integers.

$\mathbb{P}$: the set of prime numbers.

$\mathbb{T} := \mathbb{R}/\mathbb{Z}$.

$[N]$: the set $\{1, ..., N\}$.

$S(N, y)$: the set of y-smooth numbers no larger than N, that is, the collection of positive integers $n \leq N$ whose prime divisors are no larger than y.

We use subscripts to define subsets of some of these sets, and the meaning should be clear from the context. For instance, $\mathbb{R}_{\geq 0}$ denotes the set of all nonnegative real numbers.

The notion of sumset and difference set is used throughout the dissertation. Let G be an abelian group and $A, B \subseteq G$. Let m, n be positive integers. We define

$$mA + nB := \{a_1 + \cdots + a_m + b_1 + \cdots + b_n : a_i \in A, b_i \in B\},$$

and similarly

$$mA - nB := \{a_1 + \cdots + a_m - b_1 - \cdots - b_n : a_i \in A, b_i \in B\}.$$

We also define

$$A \setminus B := \{x : x \in A \text{ and } x \notin B\}.$$

We need to use the following arithmetic functions in this thesis.

Λ: the von Mangoldt function, which is defined as $\Lambda(n) = \log p$ if $n = p^k$ for some positive integer k, and 0 otherwise.

ϕ: the Euler totient function. The function $\phi(n)$ counts the number of positive integers up to n that are relatively prime to n.

ζ: the (Riemann) zeta function. For $\Re(s) > 1$, it is given by $\zeta(s) = \sum_{n\geq 1} n^{-s}$. For $\Re(s) \leq 1$, it is defined by analytic continuation.

π: the prime counting function. The function $\pi(n)$ is the number of prime numbers up to n.

μ: the Möbius function. For any positive integer n, $\mu(n) = 1$ if n is a square-free positive integer with an even number of prime factors, $\mu(n) = -1$ if n is a square-free positive integer with an odd number of prime factors, and $\mu(n) = 0$ if n has a squared prime factor.

We use $|\cdot|$ to denote both the cardinality of a set or the measure of a set. The meaning of the notation should be clear from the relevant context.

We use $\|f\|_{L^p(G)}$ to denote the L^p-norm of a function f defined on the group G (equipped with an appropriate measure), which is defined as

$$\|f\|_{L^p(G)} := \left(\int_G |f|^p\right)^{1/p}.$$

When the group is clear from the context, we may omit the group and use the notation $\|f\|_p$. We use $\mathrm{supp} f$ to denote the support of the function f.

We use the Hardy–Littlewood circle method which decomposes the circle into the "major arcs" and the "minor arcs". The setup needed to describe the decomposition

is as follows. Let $\eta > 0$ and let a, q be positive integers. We define

$$\mathfrak{M}_{a,q,\eta} = \{\theta \in \mathbb{T} : |\theta - a/q| \leq \eta\} \text{ and } \mathfrak{M}^*_{q,\eta} = \bigcup_{\substack{1 \leq a \leq q \\ (a,q)=1}} \mathfrak{M}_{a,q,\eta}. \tag{0.1}$$

When $\eta = 1/(qQ)$, we omit the subscript η in the expressions above. For instance, we write $\mathfrak{M}_{a,q,1/(qQ)}$ as $\mathfrak{M}_{a,q}$. We shall identify the torus $\mathbb{T}$ with any interval of length 1 (such as $[1/\eta, 1 + 1/\eta]$ or $[0, 1]$) when necessary.

Throughout the thesis, for any $\xi \in \mathbb{R}$, $e(\xi) := e^{2\pi i \xi}$.

For any positive integers x, q and $0 \leq a < q$, we define

$$\psi(x; q, a) := \sum_{\substack{n \equiv a \pmod q \\ n \leq x}} \Lambda(n).$$

0.2 Additive combinatorics and Fourier analysis

0.2.1 The basics

We list some standard definitions from additive combinatorics and Fourier analysis which are used in the thesis. Further details can be found in Tao and Vu [55] and Rudin [44, Chapter 1].

Let N be a positive integer and $A \subseteq [N]$. The *density of A in $[N]$* is the nonnegative number α such that $|A| = \alpha N$. We also say the density of A when the meaning is clear from the context. It follows from the definition that the density of a set is always between 0 and 1.

A *generalised arithmetic progression of dimension d* in $\mathbb{Z}$ is a set of the form

$$P = \{a + n_1 v_1 + \cdots + n_d v_d : n_i \in [N_i]\}$$

where $a, v_i \in \mathbb{Z}$ and $N_i \in \mathbb{N}$ for all $1 \leq i \leq d$. We say that the progression P is *proper* if $|P| = \prod_{1 \leq i \leq d} N_i$. We say that the progression P is *symmetric* if $P = -P$.

Let G be a locally compact abelian group and let $\widehat{G}$ denote its *dual group*, which is defined to be the collection of continuous complex homomorphisms defined on G whose images are in the unit circle in the complex plane with the group law of pointwise multiplication of functions.

Let μ be a Haar measure[1] on G. For $f : G \to \mathbb{C}$, we define the Fourier transform $\widehat{f} : \widehat{G} \to \mathbb{C}$ of f as follows. Let $\chi \in \widehat{G}$ be a character, we define

$$\widehat{f}(\chi) = \int_G f(x)\chi(x)d\mu(x).$$

The Fourier inversion formula says that there exists a unique measure λ on $\widehat{G}$ such that for any $f \in C_c^\infty(G)$ and $x \in G$, one has

$$f(x) = \int_{\widehat{G}} \widehat{f}(\chi)\overline{\chi(x)}d\lambda(\chi).$$

Throughout the thesis, we shall identify $\widehat{G}$ with any (well-known) group that is isomorphic to $\widehat{G}$. We shall bear in mind the following examples.

- When $G = \mathbb{R}^n$ and μ is the Lebesgue measure, $\widehat{G} \cong \mathbb{R}^n$ and λ is the Lebesgue measure on $\mathbb{R}^n$.
- When G is a finite abelian group equipped with the uniform probability measure, we have $\widehat{G} \cong G$ and $\widehat{G}$ is equipped with the counting measure.
- When $G = \mathbb{Z}$ and μ is the counting measure, the dual group $\widehat{G}$ is isomorphic to $\mathbb{T}$. We may also identify $\mathbb{T}$ with any interval $I \subseteq \mathbb{R}$ of length 1, so that for any $\xi \in I$,
$$\widehat{f}(\xi) = \sum_{n\in\mathbb{Z}} f(n)e^{-2\pi in\xi}.$$
The Fourier inversion formula in this case becomes
$$f(n) = \int_I \widehat{f}(\xi)e^{2\pi i\xi n}d\lambda(\xi)$$
for any $f \in L^1(\mathbb{Z}) \cap L^\infty(\mathbb{Z})$ and $x \in \mathbb{Z}$, where λ is the Lebesgue measure.

By the Pontryagin duality theorem, we have $\widehat{\widehat{G}} \cong G$.

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

$$f * g(x) = \int_G f(y^{-1}x)g(y)d\mu(y).$$

It is easy to see that

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

The *Parseval formula* states that for $f, g \in L^1(G) \cap L^2(G)$, one has

$$\int_G f(x)\overline{g(x)}dx = \int_{\widehat{G}} \widehat{f}(\gamma)\overline{\widehat{g}(\gamma)}d\gamma.$$

[1] A *Haar measure* on G is defined to be a nonnegative regular measure which is translation-invariant and not identically 0.

0.2.2 The large spectrum and Bohr sets

In this subsection, we derive all size bounds needed to prove relevant results in Chapter 4. The proofs are based on Prendiville [37, Appendix A].

Definition 0.2.1 (The large spectrum)**.** Let $f : \mathbb{Z} \to \mathbb{C}$ be a function with finite support[2] and let $\alpha > 0$. We define the *α-large spectrum of f* to be the set

$$\mathrm{Spec}(f, \alpha) := \{\xi \in \mathbb{T} : |\widehat{f}(\xi)| \geq \alpha\}.$$

Since for any $p > 1$, one has

$$\alpha^p |\mathrm{Spec}(f, \alpha)| \leq \int_{\mathrm{Spec}(f,\alpha)} |\widehat{f}(\xi)|^p d\xi \leq \int_{\mathbb{T}} |\widehat{f}(\xi)|^p d\xi.$$

One can deduce the following lemma immediately which will be used in the upcoming section.

Lemma 0.2.2. *Let $p > 1$. For any function $f : \mathbb{Z} \to \mathbb{C}$ and $\alpha > 0$, we have*

$$|\mathrm{Spec}(f, \alpha)| \leq \alpha^{-p} \|\widehat{f}\|_p^p.$$

Definition 0.2.3 (Bohr set)**.** Let N be a positive integer. Let $S \subseteq \mathbb{T}$ be a set of frequencies. Let $\rho, \eta > 0$. We define the *Bohr set* $B(N, S, \rho, \eta)$ as

$$B(N, S, \rho, \eta) := \{x \in [-\eta N, \eta N] : \|x\alpha\|_{\mathbb{T}} \leq \rho \text{ for all } \alpha \in S\}.$$

The following lemma is a generalisation of Prendiville [37, Lemma A.2]. Due to the setup of the transference principle for the colouring problem, we need to handle a more general Bohr set than the one from Prendiville's article.

Lemma 0.2.4. *Let L be a positive integer. Let f be a nonnegative measurable function supported on $[N]$ which has a positive L^1 norm. Let α be a positive real number satisfying $4\pi L^{-1}\|f\|_1 < \alpha < 1/2$, and $S := \mathrm{Spec}(f, \alpha)$. Then, for any $2\eta N L^{-1} < \rho < 1/2$, one has*

$$|B(N, S, \rho, \eta)| \geq \rho^{O_p(L\alpha^{-p}\|\widehat{f}\|_p^p)} \eta N.$$

[2] By finite, we mean $\mu(\mathrm{Supp}(f)) < \infty$ where μ is the counting measure on $\mathbb{Z}$, and $\mathrm{Supp}(f)$ denotes the support of f.

Proof. Partition $\mathbb{T}$ into L intervals $\{I_i\}_{i=1}^{L}$ of length L^{-1}. Assume that $|I_i \cap S| > 0$ for $i = 1, ..., r$, and $|I_i \cap S| = 0$ otherwise. For each $1 \leq i \leq r$, fix $\xi_i \in I_i \cap S$.

For any $\xi \in I_i$, it follows from the triangle inequality that

$$|\widehat{f}(\xi)| \geq |\widehat{f}(\xi_i)| - |\widehat{f}(\xi) - \widehat{f}(\xi_i)|.$$

Since f is supported on $[N]$ and $\|\xi - \xi_i\|_{\mathbb{T}} \leq 1/L$, we have

$$\begin{aligned} |\widehat{f}(\xi) - \widehat{f}(\xi_i)| =& |\sum_{x=0}^{N} f(x)(e(-n\xi) - e(-n\xi_i))| \\ \leq& \sum_{x=0}^{N} 2\pi \|\xi - \xi_i\|_{\mathbb{T}} |f(x)| \\ \leq& 2\pi L^{-1} \|f\|_1, \end{aligned}$$

and so

$$|\widehat{f}(\xi)| \geq \alpha - 2\pi L^{-1}\|f\|_1 \geq \alpha/2.$$

Therefore, it follows from Lemma 0.2.2 that for any $p \geq 1$, one has

$$rL^{-1} \leq \left| \bigcup_{i=1}^{r} I_i \right| \leq |\mathrm{Spec}(f, \alpha/2)| \ll_p \alpha^{-p} \|\widehat{f}\|_p^p,$$

and so

$$r \ll_p L\alpha^{-p} \|\widehat{f}\|_p^p.$$

Notice that for any $x \in B(N, \{\xi_1, ..., \xi_r\}, \rho/2, \eta)$ and $\xi \in S$, there exists $1 \leq i \leq r$ such that

$$\begin{aligned} \|x\xi\|_{\mathbb{T}} \leq& \|x\xi_i\|_{\mathbb{T}} + \|x(\xi - \xi_i)\|_{\mathbb{T}} \\ \leq& \rho/2 + \eta N L^{-1} \leq \rho, \end{aligned}$$

and so

$$B(N, \{\xi_1, ..., \xi_r\}, \rho/2, \eta) \subseteq B(N, S, \rho, \eta).$$

Thus, one may obtain a lower bound for $|B(N, S, \rho, \eta)|$ by bounding the size of $B(N, \{\xi_1, ..., \xi_r\}, \rho/2, \eta)$ from below.

Let L' be the smallest integer which is bigger than $4/\rho$ and partition $\mathbb{T}^r$ into L'^r cubes $\{C_i\}_{i=1}^{L'^r}$ of side length L'^{-1}. Since

$$\sum_{i=1}^{L'^r} \sum_{-\eta N/2 \leq n \leq \eta N/2} 1_{(\|n\xi_1\|_{\mathbb{T}}, ..., \|n\xi_r\|_{\mathbb{T}}) \in C_i} = 2\lfloor \eta N/2 \rfloor + 1,$$

it follows from the pigeonhole principle that there exists a cube, say C_1, such that $(\|n\xi_1\|_{\mathbb{T}}, ..., \|n\xi_r\|_{\mathbb{T}}) \in C_1$ for at least $L'^{-r}\eta N/2$ integers between $-\eta N/2$ and $\eta N/2$. On the other hand, for any $n_1, n_2 \in C_1$, we have

$$\|(n_1 - n_2)\xi_i\|_{\mathbb{T}} \leq \rho/2,$$

and so

$$\{n_1 - n_2 : n_1, n_2 \in C_1\} \subseteq B(N, \{\xi_1, ..., \xi_r\}, \rho/2, \eta).$$

Thus, counting the cardinality yields

$$|B(N, S, \rho, \eta)| \geq |B(N, \{\xi_1, ..., \xi_r\}, \rho/2, \eta)| \geq \rho^{O_p(L\alpha^{-p}\|\widehat{f}\|_p^p)}\eta N,$$

which completes the proof. □

0.2.3 Fourier concentration and energy increment

To connect Fourier concentration with a density increment, we shall use the following proposition[3].

Proposition 0.2.5. *Let N be a positive integer and $A \subseteq [N]$ have density $\alpha > 0$. Let $Q \geq Q_1 \geq 1$. Suppose*

$$\alpha^{-1}|A|^{-1}\sum_{q=1}^{Q_1}\frac{1}{\phi(q)}\int_{\mathfrak{M}_q^*}|(1_A - \alpha 1_{[N]})\widehat{\ }(\theta)|^2 d\theta \geq c$$

for some $c > 0$, then there exists an arithmetic progression P with common difference $q \leq Q_1$ and $|P| \gg Q_1^{-1}\min\{Q, c\alpha N\}$ such that $|A \cap P| \geq \alpha(1 + 2^{-5}c)|P|$.

Here, $\mathfrak{M}_{a,q,\eta}$ denotes the major arc centred at a/q with length 2η and $\mathfrak{M}_q^* = \bigcup_{(a,q)=1}\mathfrak{M}_{a,q,1/(qQ)}$, see (0.1).

To prove the proposition, we need the following L^2 energy increment lemma, see Ruzsa and Sanders [47, Proposition 7.2].

Lemma 0.2.6. *Let $A \subseteq [N]$ have density α. Let $\eta > 0$ and let N, q be positive integers. Suppose*

$$\mathfrak{M}_{a,q,\eta} \cap \mathfrak{M}_{a',q,\eta} = \emptyset \text{ whenever } a \neq a',$$

[3]The statement of Corollary 7.3 of Ruzsa and Sanders [47, Corollary 7.3] is the same as the statement obtained by choosing $Q = N/Q_1$ in Proposition 0.2.5. We need the more flexible choice of parameters in order to obtain the density increment in the next subsection. In fact, they also applied the result stated in Proposition 0.2.5 in Section 8 of their paper.

and

$$\alpha^{-1}|A|^{-1}\sum_{a=1}^{q}\int_{\mathfrak{M}_{a,q,\eta}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta>c$$

for some $c>0$. Then there exists an arithmetic progression P with common difference q and $|P|\gg q^{-1}\min\{\eta^{-1},c|A|\}$ such that $|A\cap P|\geq\alpha(1+c/4)|P|$.

Proof of Proposition 0.2.5. This is the same proof as Ruzsa and Sanders [47, Proof of Corollary 7.3]. Notice first that

$$\begin{aligned}\sum_{q=1}^{Q_1}\frac{q}{\phi(q)}\sum_{a=1}^{q}\int_{\mathfrak{M}_{a,q,Q^{-1}}}&|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta\\&=\sum_{q=1}^{Q_1}\frac{q}{\phi(q)}\sum_{q'h=q}\sum_{\substack{r'=1\\(r',q')=1}}^{q'}\int_{\mathfrak{M}_{r'h,q'h,Q^{-1}}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta\\&=\sum_{q'=1}^{Q_1}\sum_{\substack{a'=1\\(a',q')=1}}^{q'}\int_{\mathfrak{M}_{a',q',Q^{-1}}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta\sum_{h=1}^{\lfloor Q_1/q'\rfloor}\frac{q'h}{\phi(q'h)}.\end{aligned}$$

Since for each $1\leq q'\leq Q_1$, we have $\lfloor Q_1/q'\rfloor\geq Q_1/(2q')$. It follows that

$$\sum_{h=1}^{\lfloor Q_1/q'\rfloor}\frac{q'h}{\phi(q'h)}\geq\frac{q'}{\phi(q')}\sum_{h=1}^{\lfloor Q_1/q'\rfloor}1\geq\frac{Q_1}{2\phi(q')},$$

and so

$$\begin{aligned}\sum_{q=1}^{Q_1}\frac{q}{\phi(q)}\sum_{a=1}^{q}\int_{\mathfrak{M}_{a,q,Q^{-1}}}&|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta\\&\geq Q_1\sum_{q'=1}^{Q_1}\frac{1}{2\phi(q')}\sum_{\substack{a'=1\\(a',q')=1}}^{q'}\int_{\mathfrak{M}_{a',q',Q^{-1}}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta.\end{aligned}$$

Since for each pair (a',q'), the arc $\mathfrak{M}_{a',q',Q^{-1}}$ contains the arc $\mathfrak{M}_{a',q'}=\mathfrak{M}_{a',q',(q'Q)^{-1}}$, it follows from the hypothesis of the proposition that

$$\alpha^{-1}|A|^{-1}\sum_{q=1}^{Q_1}\frac{q}{\phi(q)}\sum_{a=1}^{q}\int_{\mathfrak{M}_{a,q,Q^{-1}}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2d\theta\geq\frac{Q_1c}{2}.$$

On the other hand, one has (see Montgomery and Vaughan [34, Theorem 2.1])

$$\sum_{q=1}^{Q_1}\frac{q}{\phi(q)}\leq 4Q_1,$$

so by the pigeonhole principle, there exists some $1 \leq q \leq Q_1$ such that

$$\alpha^{-1}|A|^{-1}\sum_{a=1}^{q}\int_{\mathfrak{M}_{a,q,Q^{-1}}}|(1_A-\alpha 1_{[N]})\widehat{\ }(\theta)|^2 d\theta \geq \frac{c}{8}.$$

The conclusion of the proposition follows from an application of Lemma 0.2.6. □

0.3 Analytic number theory

We introduce the number-theoretic background we need in the thesis.

0.3.1 Zeros of Dirichlet L-functions

The results in this subsection are well-known due to their role in the proof of Linnik's theorem, see Iwaniec and Kowalski [27].

For any $q \in \mathbb{N}$, we say that χ *is a Dirichlet character of modulus* q if χ is a multiplicative function such that

- $\chi(n) \neq 0$ if and only if $(n,q)=1$;
- $\chi(n+q)=\chi(n)$ for every positive integer n;
- $\chi(mn)=\chi(m)\chi(n)$ for all $m,n \in \mathbb{N}$.

It follows from $n^{\phi(q)} \equiv 1 \pmod{q}$ and the properties above that $|\chi(n)|=1$ for all $n \in \mathbb{N}$ such that $(n,q)=1$. The *principal character of modulus* q is the character whose image set is $\{0,1\}$.

Let χ be a character of modulus q and χ' be the principal character of modulus q. We say that χ is *primitive* if it cannot be written as $\chi=\chi'\chi_1$ for some χ_1 of modulus smaller than q. Otherwise, we say that χ is an *induced character*.

Let χ be a Dirichlet character of modulus q. We define

$$\psi(x,\chi) := \sum_{n\leq x}\chi(n)\Lambda(n),$$

The classical way of studying $\psi(x,\chi)$ is to understand the zeros of the corresponding L-function $L(s,\chi)$, which is defined as the analytic continuation of the function

$$\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}, \quad \Re(s)>1.$$

In fact, to study the location of the zeros of L-functions in the critical strip $0 < \Re(s) < 1$, it suffices to study the zeros of $L(s,\chi)$ for all primitive characters. This

can be seen by using the Euler product expansion. Let χ be a character of modulus q which is induced by a primitive character χ_1 and let χ' be the principal character of modulus q. It follows from the definition of induced characters that $\chi = \chi_1\chi'$, and so

$$L(s,\chi) = \prod_{p|q}(1-\chi_1(p)p^{-s})L(s,\chi_1) \text{ for } \Re(s) > 1. \tag{0.2}$$

Therefore, by analytic continuation, $L(s,\chi) = 0$ if and only if $L(s,\chi_1) = 0$ in the region $\Re(s) > 0$. For the principal character χ', one may use the same method to show that $L(s,\chi')$ has the same zeros as $\zeta(s)$ in the critical strip.

The following theorem can be found in Iwaniec and Kowalski [27, Section 18.1, Principle 1].

Theorem 0.3.1. *There is an absolute constant $c_1 > 0$ such that for any positive integer q and $T \geq 1$, there exists at most one character χ such that χ has modulus q and $L(s,\chi)$ has a zero β in the region*

$$\Re(s) \geq 1 - \frac{c_1}{\log(qT)},\ |\Im(s)| \leq T. \tag{0.3}$$

Moreover, this zero may only occur if χ is a real non-principal character, and the zero has to be real and simple.

When this zero occurs, we shall need an upper bound on its size. This can be obtained by taking $\epsilon = 1/2$ in Theorem 5.28[4] of Iwaniec and Kowalski [27], which yields

$$1 - \beta \gg q^{-1/2}. \tag{0.4}$$

For any $1/2 \leq \alpha \leq 1$ and $T \geq 1$, let $N(\alpha, T, \chi)$ denote the number of zeros of $L(s,\chi)$ counted with multiplicity in the region

$$\alpha < \Re(\rho) \leq 1, |\Im(\rho)| \leq T. \tag{0.5}$$

It follows that

$$N_q(\alpha, T) := \sum_{\chi \pmod q} N(\alpha, T, \chi)$$

is the total number of zeros of all $L(s,\chi)$ (of modulus q; counted with multiplicity) in the region (0.5). We shall need the following log-free zero density estimate [27, Principle 2, Section 18.1] to bound the quantities involved in the next section.

[4]It is pointed out in the theorem that the implicit constant $c(\epsilon)$ is no longer computable when one takes $\epsilon < 1/2$.

Theorem 0.3.2. *There exists an absolute constant $C_1 > 0$ such that for any $1/2 \leq \alpha \leq 1$ and $T \geq 1$, we have*

$$N_q(\alpha, T) \leq C_1(qT)^{C_1(1-\alpha)}. \tag{0.6}$$

The presence of a zero in the region (0.3) has a strong impact on other zeros of L-functions, which will play a crucial role in our quantitative result. The statement below is a quantitative version of exceptional zero repulsion [27, Principle 3, Section 18.1].

Theorem 0.3.3. *There is an absolute constant $c_2 > 0$ such that the following holds for all positive integers q. If there exists a character χ_0 of modulus q such that $L(s, \chi_0)$ has a zero in the region (0.3), then for all χ of modulus q, the functions $L(s, \chi)$ do not have any other zero in the region*

$$\Re(s) \geq 1 - \frac{c_2|\log((1-\beta)\log(qT))|}{\log(qT)}, \quad |\Im(s)| \leq T.$$

On the other hand, it is known that such zeros rarely occur, if they occur at all. More precisely, a result of Landau [27, Theorem 5.28] states that there exists an absolute constant $c_3 > 0$ such that for any distinct real primitive Dirichlet characters χ_1 and χ_2 of moduli q_1 and q_2, and real zeros β_1 and β_2, respectively, we have

$$\min\{\beta_1, \beta_2\} \leq 1 - \frac{c_3}{\log(q_1 q_2)}. \tag{0.7}$$

One of the most famous conjectures in analytic number theory is the generalised Riemann hypothesis.

Conjecture 0.3.4 (The Generalised Riemann Hypothesis)**.** *For any Dirichlet character χ and any s such that $L(s, \chi) = 0$, we have either $\Re(s) = 1/2$ or s is a negative real number.*

0.3.2 A minor arc estimate

We shall need the following minor arc estimate which can be found in Ruzsa and Sanders [47, Corollary 6.2]. This result can be deduced from the well-known minor arc estimate of Vinogradov (see, for instance, Iwaniec and Kowalski [27, Theorem 13.6]).

Theorem 0.3.5. *Let N, Q, d be positive integers such that $d \leq N$. Let $F_{N,d}(n) = \Lambda(dn+1)1_{[N]}(n)$, where Λ is the von Mangoldt function.*

For any positive integers q, a such that $q \leq Q$ and $(a,q) = 1$ and any $|\theta - a/q| \leq 1/(qQ)$, we have

$$|\widehat{F_{N,d}}(\theta)| \ll d(\log N)^4 \left(\frac{N}{\sqrt{q}} + N^{4/5} + \sqrt{NQ} \right).$$

Chapter 1

Geometric progressions in smooth numbers

1.1 Introduciton

The density of subsets of $[N]$ containing no nontrivial arithmetic progression has been extensively studied. The first nontrivial result is due to Roth [42, 43], which states that any subset of positive integers of positive upper density contains a nontrivial 3-term arithmetic progression[1]. The density threshold in Roth's original papers has been improved by several people, and the current best result is due to Bloom and Sisask [4], which states that any subset of $[N]$ free of nontrivial 3-term arithmetic progressions must have cardinality $O\left(N(\log N)^{-1-c}\right)$ for some absolute constant $c > 0$. The case when $k = 4$ was first shown by Szemerédi [53], and the density threshold has been subsequently improved, see Green and Tao [21].

In 1975, Szemerédi [54] proved the celebrated theorem which deals with arbitrarily long progressions. Many proofs of Szemerédi's theorem have been developed more recently, where several different disciplines of mathematics are used. Since the known quantitative bounds on the density threshold is relatively poor when $k > 4$, here we record only Szemerédi's theorem without specifying any quantitative information.

Theorem 1.1.1 (Szemerédi's theorem)**.** *For every positive integer k and every $\delta > 0$, there exists an integer $N = N(k, \delta)$ such that every subset of $[N]$ of density at least δ contains an arithmetic progression of length k.*

Rankin [38] studied subsets of the integers without nontrivial geometric progressions in his work on arithmetic progressions. A basic example of such sets is the

[1]Nontrivial arithmetic progressions or geometric progressions mean progressions whose terms are not equal to each other.

set of squarefree integers, which has asymptotic density $6/\pi^2$. For every $k \geq 3$, he proved that there exist sets free of k-term geometric progressions with positive asymptotic density. Subsequently, estimates of both the asymptotic density and the upper asymptotic density of subsets free of k-term geometric progressions have been obtained by several authors [41, 5, 2, 35]. The most recent result is due to Nathanson and O'Bryant [35], which shows that the upper asymptotic density of sequences of positive integers that contain no k-term geometric progression is bounded above by

$$1 - \frac{1}{2^k - 1} - \frac{2}{5}\left(\frac{1}{5^{k-1}} - \frac{1}{6^{k-1}}\right) - \frac{4}{15}\left(\frac{1}{7^{k-1}} - \frac{1}{10^{k-1}}\right).$$

In the special case when $k = 3$, McNew [33] obtained a better bound using a different method.

Motivated by these investigations, we are interested in the density of subsets of smooth numbers which contain no nontrivial geometric progressions. Let $S(N, y)$ denote the set of y-smooth numbers no larger than N, that is,

$$S(N, y) = \{x \leq N : \text{ all prime factors of } x \text{ are no larger than } y\}.$$

The case when $y = N$ corresponds to the integer case discussed above. In fact, it has been known that for any $y = (\log N)^{k/(k-1)+\epsilon}$, where $\epsilon > 0$ and N is sufficiently large depending on ϵ and k, the set of k-free[2] integers in $S(N, y)$ has positive density [16]. More specifically, let $\alpha(N, y) > 0$ be the unique solution of the equation

$$\log N = \sum_{p \leq y} \frac{\log p}{p^{\alpha(N,y)} - 1}.$$

Hildebrand and Tenenbaum [24] obtained the estimate

$$\alpha(N, y) = \frac{\log(1 + y/\log N)}{\log y}\left(1 + O\left(\frac{\log\log(1+y)}{\log y}\right)\right) \tag{1.1}$$

for all $N \geq y \geq 2$. Given $\epsilon > 0$, by modifying the method of Ivić and Tenenbaum [25], it can be shown that uniformly in the range $N \geq y \geq (\log N)^{k/(k-1)+\epsilon}$ where N is sufficiently large, the number of k-free y-smooth numbers less than N is

$$\frac{|S(N, y)|}{\zeta(k\alpha(N, y))}\left(1 + o(1)\right).$$

Granville [16] also demonstrated that it is extremely difficult to estimate the number of k-free integers in $S(N, y)$ when y is a multiple of $\log N$.

[2] A *k-free* integer is an integer which cannot be divided by any kth power.

The situation when y is small and N is large turns out to be very different. For $y = 2$, by applying Szemerédi's theorem directly on the exponents, one can easily find a geometric progression in any subset of $S(N, 2)$ with positive density.

In this chapter, we shall prove the following result.

Theorem 1.1.2. *Let $\alpha > 0$ and $\epsilon > 0$. Let $k \geq 3$ be an integer. There exists $N_0 = N_0(\alpha, \epsilon, k)$ such that for every $N > N_0$ and $y \leq \log N/(\log\log\log N)^{2+\epsilon}$, any subset A of $S(N, y)$ with density α contains a nontrivial k-term geometric progression.*

Thus, for each k, it is unknown whether any subset of $S(N, y)$ with density α contains a nontrivial k term geometric progression for y between $\log N/(\log\log\log N)^{2+\epsilon}$ and $y = (\log N)^{1+1/(k-1)+\epsilon}$, which is a relatively small gap when k is very large. Since our method works for any fixed k, it is likely that we would need much more understanding of the structure of smooth numbers if we hope to understand the 3-term geometric progression problem.

Let Y denote the number of primes no bigger than y and let $2 = p_1 < \cdots < p_Y \leq y$ be the prime numbers. An integer $p_1^{e_1}...p_Y^{e_Y}$ is contained in $S(N, y)$ if and only if $e_1 \log p_1 + ... + e_Y \log p_Y \leq \log N$. Thus, by mapping $p_1^{e_1}...p_Y^{e_Y}$ to $(e_1, ..., e_Y)$, we can establish a bijection between $S(N, y)$ and the set $B(\log N, y)$, where the *Y-dimensional weighted l^1 ball* is defined by

$$B(R, y) := \left\{ (e_1, ..., e_Y) \in \mathbb{Z}_{\geq 0}^Y : \sum_{i=1}^{Y} e_i \log p_i \leq R \right\}.$$

In particular, this map implies that $(S(N, y), \mathbb{Q}^\times)$ and $(B(\log N, y), \mathbb{Z}^Y)$ are Freiman isomorphic (see Tao and Vu [55, Chapter 5]). Thus, to prove Theorem 1.1.2, it suffices to prove the following result, which actually shows the existence of a geometric progression with common ratio a power of 2.

Theorem 1.1.3. *Let $\alpha > 0$ and $\epsilon > 0$. There exists $N_0 = N_0(\alpha, \epsilon, k)$ such that for every $N > N_0$ and $y \leq \log N/(\log\log\log N)^{2+\epsilon}$, any subset A of $B(\log N, y)$ with density α contains a k-tuple $((e_1, e_2, ..., e_Y), (e_1 + d, e_2, ..., e_Y), ..., (e_1 + (k-1)d, e_2, ..., e_Y))$ where $d > 0$.*

In Section 1.2, we obtain a partial result which holds for $y \leq (\log N)^{1/2}$. There are two main ideas in this section: a suitable approximate structure and an averaging argument. We use this result together with more careful observations on the structure of $B(\log N, y)$ to prove Theorem 1.1.3 in Section 1.3.

It turns out that one cannot always find an arbitrarily long geometric progression in any subset of $S(N,y)$ of positive density when $y \geq (\log N)^{1+\epsilon}$. In Section 1.4, we employ the method of Ivić and Tenenbaum [25], whose paper proves that the square-free integers in $S(N,y)$ has positive density when $y \geq (\log N)^{2+\epsilon}$, to prove the following result.

Proposition 1.1.4. *For fixed positive integer $k \geq 2$ and positive number ϵ, there exist $N_0 = N_0(k,\epsilon)$ and $c = c(k,\epsilon)$ such that the set of k-free integers in $S(N,y)$ has density at least c whenever $y \geq (\log N)^{1+1/(k-1)+\epsilon}$ and $N \geq N_0$.*

The proposition above is sharp in the sense that as $\epsilon \to 0$, the density of k-free numbers in $S(N,y)$ goes to 0 when $y = (\log N)^{1+1/(k-1)+\epsilon}$ and N is sufficiently large.

1.2 An argument using an approximate structure

The proof of our first partial result consists of two main ingredients. Firstly, we count the smooth numbers using elementary geometric ideas. Then we apply a basic averaging argument.

The following lemma will be used in both this section and the upcoming one. It is shown by a "volume packing" argument where we associate the counting of lattice points with the number of unit cubes contained in polytopes.

Lemma 1.2.1. *Let k be a positive integer and $a_1, ..., a_k$ be positive real numbers. For any $C > 0$, let*

$$D_C = \left\{ (n_1, ..., n_k) \in \mathbb{Z}_{\geq 0}^k : \sum_{i=1}^k a_i n_i \leq C \right\}.$$

Then, for any $C_1, C_2 > 0$, one has

$$\left(\frac{C_1}{C_2 + \sum_{i=1}^k a_i} \right)^k \leq \frac{|D_{C_1}|}{|D_{C_2}|} \leq \left(\frac{C_1 + \sum_{i=1}^k a_i}{C_2} \right)^k.$$

Proof. For each $C > 0$, let

$$\mathbf{D}_C = \left\{ (x_1, ..., x_k) \in \mathbb{R}_{\geq 0}^k : \sum_{i=1}^k a_i x_i \leq C \right\}.$$

For every $\bar{n} = (n_1, ..., n_k) \in \mathbb{Z}_{\geq 0}^k$, we associate $\bar{n}$ with the cube

$$\mathbf{X}_{\bar{n}} = \{ (x_1, ..., x_k) \in \mathbb{R}_{\geq 0}^k : n_i \leq x_i < n_i + 1 \text{ for each } 1 \leq i \leq k \}$$

whose volume is 1.

Notice that for every $\bar{n} \in \mathbb{Z}^k_{\geq 0}$, if $\bar{n} \in \mathbf{D}_C$, then for every $(x_1, ..., x_k) \in \mathbf{X}_{\bar{n}}$ one has

$$\sum_{i=1}^{k} a_i x_i \leq C + \sum_{i=1}^{k} a_i.$$

Therefore,

$$|D_C| = \left| \bigcup_{\bar{n} \in D_C} \mathbf{X}_{\bar{n}} \right| \leq |\mathbf{D}_{C+\sum_{i=1}^{k} a_i}| \tag{1.2}$$

where the left hand side is the cardinality of D_C, and the right hand side is the volume of the polytope. On the other hand, for every $(x_1, ..., x_k) \in \mathbf{D}_C$, we have $(x_1, ..., x_k) \in \mathbf{X}_{\bar{n}}$ where $\bar{n} = (\lfloor x_1 \rfloor, ..., \lfloor x_k \rfloor)$. Since $a_1\lfloor x_1 \rfloor + \cdots + a_k \lfloor x_k \rfloor \leq a_1 x_1 + \cdots + a_k x_k \leq C$, it follows that $\mathbf{D}_C \subseteq \bigcup_{\bar{n} \in D_C} \mathbf{X}_{\bar{n}}$. Thus,

$$|\mathbf{D}_C| \leq \left| \bigcup_{\bar{n} \in D_C} \mathbf{X}_{\bar{n}} \right| = |D_C|. \tag{1.3}$$

Combining (1.2) and (1.3) yields

$$\left(\frac{C_1}{C_2 + \sum_{i=1}^{k} a_i} \right)^k \leq \frac{|\mathbf{D}_{C_1}|}{|\mathbf{D}_{C_2+\sum_{i=1}^{k} a_i}|} \leq \frac{|D_{C_1}|}{|D_{C_2}|} \leq \frac{|\mathbf{D}_{C_1+\sum_{i=1}^{k} a_i}|}{|\mathbf{D}_{C_2}|} \leq \left(\frac{C_1 + \sum_{i=1}^{k} a_i}{C_2} \right)^k,$$

which completes the proof. □

Proposition 1.2.2. *Let $\alpha > 0$ and $k \geq 3$. There exists $R_0 = R_0(\alpha, k)$ such that for every $R > R_0$ and $y \leq R^{1/2}$, any subset A of $B(R, y)$ with density α contains a k-tuple $((e_1, e_2, ..., e_Y), (e_1 + d, e_2, ..., e_Y), ..., (e_1 + (k-1)d, e_2, ..., e_Y))$ where $d > 0$.*

Proof. Throughout the proof, we shall use B to denote $B(R, y)$. Let

$$B' := \{(e_1, ..., e_Y) \in \mathbb{Z}^Y_{\geq 0} : \sum_{i=1}^{Y} e_i \log p_i \leq \delta R\},$$

where $\delta > 0$ is a positive parameter such that $\delta R \to \infty$ as $R \to \infty$. The reason for this condition is that we shall later apply Szemerédi's theorem to an arithmetic progression of length δR.

By Lemma 1.2.1, we have

$$\frac{|B + B'|}{|B|} \leq \left(\frac{R + \delta R + \sum_{i=1}^{Y} \log p_i}{R} \right)^Y. \tag{1.4}$$

Since $R \geq y^2$, it follows from the prime number theorem that $Y(\sum_{i=1}^{Y} \log p_i)/R \ll 1$. Let us choose $\delta = \log(y)/y$, so we have $\delta Y \ll 1$. Combining these observations yields

$$\left(\frac{R + \delta R + \sum_{i=1}^{Y} \log p_i}{R} \right)^Y = O(1). \tag{1.5}$$

Let $P = \{(x, 0, ..., 0) \in \mathbb{Z}_{\geq 0}^Y : 0 \leq x < \lfloor \delta R / \log 2 \rfloor\}$, so P is a subset of B'. It follows from (1.4) that

$$\frac{|B + P|}{|B|} \leq \left(\frac{R + \delta R + \sum_{i=1}^{Y} \log p_i}{R} \right)^Y. \tag{1.6}$$

We shall employ a basic averaging argument[3] to locate a progression of length $\lfloor \delta R / \log 2 \rfloor$ to which we can apply Szemerédi's theorem. More specifically, we claim that there exists some $t \in \mathbb{Z}^Y$ such that $|A \cap (t + P)| \geq \alpha' |P|$, where $\alpha' \geq c\alpha$ for some absolute constant c.

Let

$$\Gamma = \{(e_1, ..., e_Y) : e_1 \equiv 0 \pmod{\lfloor \delta R / \log 2 \rfloor}, e_i \in \mathbb{Z} \text{ for all } 2 \leq i \leq Y\}.$$

Let $T = \Gamma \cap B$. Then we have

(i) $\bigcup_{t \in T} |A \cap (t + P)| = A$,

(ii) $\bigcup_{t \in T} (t + P) \subseteq B + P$,

(iii) and $(t + P) \cap (t' + P) = \emptyset$ for any $t, t' \in T$, $t \neq t'$.

By (ii) and (iii) above, we can deduce that

$$|T| \leq \frac{|B + P|}{|P|}.$$

Combining this inequality with (i), (1.5) and (1.6) shows that there must exist $t \in T$ such that

$$|A \cap (t + P)| \geq \frac{|A|}{|B + P|} |P| \gg \alpha |P|,$$

which completes the proof of our claim. Fix this choice of t.

Since the set $A \cap (t + P)$ is contained in a one-dimensional subspace and $|P| \gg \delta R \geq R^{1/2}$, we can apply Szemerédi's theorem to deduce the existence of a nontrivial k-term arithmetic progression, which proves the desired proposition. □

[3]We remark that this particular way of averaging is sensitive to the fact that the collection of translates of P partitions the set we are interested in. A more general averaging argument can be found in Varnavides [56].

Notice that the approximate structure shown in (1.4) does not hold anymore if y is much larger than $R^{1/2}$. The reason is that as y gets large, the approximation of $|B + B'|$ by the volume of corresponding polytope becomes worse. In fact, the argument above can be used to show that the conclusion of Proposition 1.2.2 holds when $y \ll (R \log R)^{1/2}$: the bounds (1.4) and (1.5) still hold since $Yy/R \ll ((R \log R)/\log R)/R = 1$. We did not state this range of y as we shall improve it in the upcoming section.

By taking $R = \log N$, we can deduce the following corollary, where the range of y is much smaller than the range stated in Theorem 1.1.3.

Corollary 1.2.3. *Let $\alpha > 0$ and $k \geq 3$. There exists $N_0 = N_0(\alpha, k)$ such that for every $N > N_0$ and $y \leq (\log N)^{1/2}$, any subset A of $B(\log N, y)$ with density α contains a k-tuple $((e_1, e_2, ..., e_Y), (e_1 + d, e_2, ..., e_Y), ..., (e_1 + (k-1)d, e_2, ..., e_Y))$ where $d > 0$.*

1.3 Proof of the main result

The goal of this section is to prove Theorem 1.1.3, which extends the range of y such that for sufficiently large N, any subset of y-smooth numbers less than N with given positive density α contains a nontrivial k-term geometric progression.

Heuristically speaking, the observation used here is that for this range of y, the portion of y-smooth numbers less than N with 'few' small prime divisors cannot be positive (we shall first consider the prime divisors less than $2(\log\log N)(\log\log\log N)$ to prove the result for $y \leq \log N/((\log\log N)(\log\log\log N)^{2+\epsilon})$). Next, we find a suitable subset by averaging appropriately, and applying results we have already obtained.

Let R, y, r, c be positive numbers. Let Y and C denote the number of primes no larger than y and c, respectively. We define

$$V(R, y, r, c) := \left\{ (e_1, ..., e_Y) \in B(R, y) : \sum_{i=C+1}^{Y} e_i \log p_i \geq R - r \right\}$$

to be a collection of elements of $B(R, y)$ whose weight on the first C coordinates is less than r. The next lemma tells us that, under certain conditions on the size of R, y, r and c, the density of $V(R, y, r, c)$ in $B(R, y)$ is negligible if N is sufficiently large.

Lemma 1.3.1. *There exists an absolute constant $c' > 0$ such that for any $R > y > r > c$ satisfying*

$$10Yr < R - r, c \leq c'r \text{ and } \log Y \leq o(c),$$

one has

$$\frac{|V(R,y,r,c)|}{|B(R,y)|} \to 0 \text{ as } c \to \infty.$$

Proof. Let $v = (e_1, ..., e_Y)$ be an element of $V(R,y,r,c)$. Let j be the first coordinate for which

$$e_j \geq \frac{10r}{\log p_j} \text{ and } j > C.$$

Notice that such index exists due to the hypotheses. We define $\phi : V(R,y,r,c) \to \mathbb{Z}_{\geq 0}^{Y-C}$ by

$$\phi(v) := (e_{C+1}, ..., e_{j-1}, e_j - \lceil 5r/\log p_j \rceil, e_{j+1}, ..., e_Y),$$

and

$$W = \left\{ w \in \mathbb{Z}_{\geq 0}^{Y-C} : w = \phi(v) \text{ for some } v \in V(R,y,r,c) \right\}.$$

Since for each $w \in W$, there are at most $Y|B(r,c)|$ different choices of $v \in V(R,y,r,c)$ such that $\phi(v) = w$ (the factor Y follows from looking at the last $Y - C$ coordinates; the factor $|B(r,c)|$ follows from considering the first C coordinates using the definition of the set $V(R,y,r,c)$), we can deduce that

$$|V(R,y,r,c)| \leq Y|B(r,c)||W|.$$

On the other hand, for each $(e_{C+1}, ..., e_Y) \in W$, it follows from the definition of W that $\sum_{i=C+1}^{Y} e_i \log p_i \leq R - 5r$. Hence, for each $w \in W$, the set

$$\{(e_1, ..., e_Y) \in B(R,y) : (e_{C+1}, ..., e_Y) = w\}$$

has cardinality at least $|B(5r,c)|$. Therefore, we have

$$|B(R,y)| \geq |B(5r,c)||W|.$$

Thus, we can conclude that

$$\frac{|V(R,y,r,c)|}{|B(R,y)|} \leq Y \frac{|B(r,c)|}{|B(5r,c)|}.$$

Now it follows from Lemma 1.2.1 that

$$\frac{|V(R,y,r,c)|}{|B(R,y)|} \leq Y \left(\frac{r + \sum_{i=1}^{C} \log p_i}{5r} \right)^C.$$

By the prime number theorem, there exists $C' > 0$ such that $\sum_{i=1}^{n} \log p_i \leq C'n$ for all $n \in \mathbb{N}$. Hence, by taking $c' = 1/C'$, it follows from our hypotheses that

$$\frac{|V(R,y,r,c)|}{|B(R,y)|} \leq Y \left(\frac{r+r}{5r} \right)^C = o_c(1),$$

which completes the proof. □

Let us make some comments on the proof of this lemma. The relation between the quantities y and r is chosen to guarantee that the map ϕ is well-defined on $V(R, y, r, c)$. This map makes the comparison technically easy. The lower bound on r in terms of c is used to ensure that the geometric comparison used to count smooth number could work efficiently. A large c yields a large exponent in the final inequality used in the proof, yet it requires a large l^1-radius r (of the C-dimensional ball) if we intend to apply Proposition 1.2.2 to find a geometric progression. As a result, it reduces the allowable size of y due to the first condition. On the other hand, if we choose a small c, then we need a large denominator in the final inequality when y is large. To get a large denominator, which requires large r, we still need to reduce the size of y.

Now we use the previous lemma to extend the range of y proved in Proposition 1.2.2.

Lemma 1.3.2. *Let $\alpha > 0$, $\epsilon > 0$ and $k \geq 3$. There exists $R_0 = R_0(\alpha, \epsilon, k)$ such that for every $R > R_0$ and $y \leq R/((\log R)(\log\log R)^{2+\epsilon})$, any subset A of $B(R, y)$ with density α contains a k-tuple $((e_1, e_2, ..., e_Y), (e_1 + d, e_2, ..., e_Y), ..., (e_1 + (k-1)d, e_2, ..., e_Y))$ where $d > 0$.*

Proof. We may assume that $y > 4(\log R)^2(\log\log R)^2$, since otherwise the result follows from Proposition 1.2.2. Let

$$y_1 = y$$

$$r_1 = 4(\log R)^2(\log\log R)^2$$

and

$$c_1 = 2\log R\log\log R.$$

It follows that for sufficiently large R, one has

$$|A \cap B(R, y_1) \setminus V(R, y_1, r_1, c_1)| \geq 0.9\alpha|B(R, y_1)|.$$

Therefore, by averaging, we can find $a_1 \in \mathbb{Z}_{\geq 0}^{Y_1 - C_1}$ such that the set

$$A_1 = \{(e_1, ..., e_{Y_1}) \in A \cap B(R, y_1) \setminus V(R, y_1, r_1, c_1) : (e_{C_1+1}, ..., e_{Y_1}) = a_1\}$$

and the C_1-dimensional affine subspace

$$U_1 = \{(e_1, ..., e_{Y_1}) \in B(R, y_1) \setminus V(R, y_1, r_1, c_1) : (e_{C_1+1}, ..., e_{Y_1}) = a_1\}$$

satisfy

$$|A_1| \geq \alpha|U_1|/2.$$

Since $U_1 \cap V(R, y_1, r_1, c_1) = \emptyset$, it follows from the definition of $V(R, y_1, r_1, c_1)$ that there exists $r' \geq r$ such that

$$B(r', c_1) = \{(e_1, ..., e_{C_1}) : (e_1, ..., e_{C_1}, ..., e_{Y_1}) \in U_1\}.$$

Moreover, the set

$$A' = \{(e_1, ..., e_{C_1}) : (e_1, ..., e_{C_1}, ..., e_{Y_1}) \in A_1\}$$

has

$$|A'| \geq \alpha |B(r', c_1)|/2.$$

Therefore, by our choice of r_1 and c_1, we can apply Proposition 1.2.2 to A' and $B(r', c_1)$, and deduce the existence of the claimed k-tuple in A_1. □

In the proof of the proposition above, we chose $r_1 = c_1^2$ due to the condition of Proposition 1.2.2. Since we have improved the range of y which guarantees the existence of a nontrivial progression, we can apply Lemma 1.3.2 to further improve the range. The argument is identical to the one used to prove Lemma 1.3.2; the only difference is that we shall apply Lemma 1.3.1 with different parameters.

Proof of Theorem 1.1.3. Let $R = \log N$ and $y_2 = y$. Again, by Proposition 1.2.2, we may assume that $y \geq R^{1/2}$. Let

$$r_2 = (\log\log N)(\log\log\log N)^{2+\epsilon/2} \text{ and } c_2 = 2(\log\log N)(\log\log\log N).$$

It follows from Lemma 1.3.1 that for sufficiently large N,

$$|A \cap (B(\log N, y_2) \setminus V(\log N, y_2, r_2, c_2)| \geq 0.9\alpha |B(\log N, y_2)|.$$

By averaging, we can find $a_2 \in \mathbb{Z}_{\geq 0}^{Y_2 - C_2}$ such that the sets

$$A_2 = \{(e_1, ..., e_{Y_2}) \in A \cap B(R, y_2) \setminus V(R, y_2, r_2, c_2) : (e_{C_2+1}, ..., e_{Y_2}) = a_2\}$$

and

$$U_2 = \{(e_1, ..., e_{Y_2}) \in B(R, y_2) \setminus V(R, y_2, r_2, c_2) : (e_{C_2+1}, ..., e_{Y_2}) = a_2\}$$

satisfy

$$|A_2| \geq \alpha |U_2|/2.$$

Therefore, there exists $r' \geq r_2$ such that

$$A' = \{(e_1, ..., e_{C_2}) : (e_1, ..., e_{C_2}, ..., e_{Y_2}) \in A_2\}$$

satisfies $|A'| \geq \alpha |B(r', c_2)|/2$. One can conclude the existence of the geometric progression in the same way as Lemma 1.3.2. □

1.4 A counterexample: k-free integers

Throughout this section, let $\epsilon > 0$ and $k \geq 2$. Our aim is to prove Proposition 1.1.4 using the method of Ivić and Tenenbaum [25], which provides an example of a k-term-geometric-progression-free set.

Proposition 1.4.1. *For sufficiently large N (which depends on ϵ), and y which satisfies $y \geq \log^{1+1/(k-1)+\epsilon} N$, the density of the set of k-free numbers in $S(N, y)$ is*

$$\frac{1}{\zeta(k\alpha(N,y))}\left(1 + o(1)\right).$$

Here $\alpha = \alpha(N, y)$ is defined to be the unique solution of

$$\log N = \sum_{p \leq y} \frac{\log p}{p^{\alpha(N,y)} - 1}.$$

The case when $k = 2$ follows directly from the main result of the paper of Ivić and Tenenbaum [25, Theorem 1]. It is noticed that their method can be adapted to prove an analogue result for k-free numbers, see Granville [16, Section 5]. In this section, we provide a complete written proof of the result by modifying the Ivić–Tenenbaum argument.

To describe the argument precisely, one needs a few more parameters. For $2 \leq y \leq N$, we define

$$u = \frac{\log N}{\log y}.$$

Let $\xi = \xi(u)$ be the unique positive solution of the equation $e^{\xi} = \xi u + 1$ if $u > 1$, and $\xi(1) = 0$. By taking logarithms, $\xi = \log(\xi u + 1) = \log u + \log(\xi + 1/u)$, and so

$$\xi(u) = \log u + \log\log u + O(1). \tag{1.7}$$

Let

$$\beta = \beta(N, y) = 1 - \frac{\xi(u)}{\log y}.$$

It follows from (1.7) that there exists N_ϵ such that in the range

$$\log^{1+1/(k-1)+\epsilon} N < y \leq N, \quad N \geq N_\epsilon,$$

one has

$$\beta \geq 1 - \frac{1}{1 + 1/(k-1) + \epsilon} \geq \frac{1}{k} + \frac{\epsilon}{4}.$$

On the other hand, by Hildebrand and Tenenbaum [24], uniformly in the range

$$N \geq 2, \log^{1+\epsilon} N \leq y \leq N,$$

one has

$$\alpha(N,y) = \beta(N,y) + O_\epsilon\left(\exp\left(-2\sqrt{\log y}\right) + u^{-1}\log^{-2} y\right). \tag{1.8}$$

Thus, to prove the proposition stated at the beginning of the section, it suffices to prove that the cardinality of the set of k-free numbers in $S(N,y)$ is

$$\frac{|S(N,y)|}{\zeta(k\beta)}\left(1 + O_\epsilon\left(\frac{\log\log y}{\log y}\right)\right).$$

To prove the estimate above, we shall need the following two technical lemmas which concern the behaviour of

$$\frac{|S(N/d,y)|}{|S(N,y)|},$$

where N, y are within the range

$$N \geq 2, \log^{1+\epsilon} N \leq y \leq N. \tag{1.9}$$

These lemmas are proven in Ivić and Tenenbaum [25, Section 5]. They allow us to apply different estimates to $|S(N/d,y)|/|S(N,y)|$ according to the range of d.

Lemma 1.4.2. *Uniformly for all N, y in the range (1.9) and $1 \leq d \leq y$, one has*

$$\left|S\left(\frac{N}{d}, y\right)\right| = |S(N,y)|\, d^{-\beta}\left(1 + O_\epsilon\left(\frac{\log d}{\log N} + \frac{\log\log y}{\log y}\right)\right).$$

Lemma 1.4.3. *Uniformly for all N, y in the range (1.9) and $1 \leq d \leq N$, one has*

$$\left|S\left(\frac{N}{d}, y\right)\right| \ll |S(N,y)|\, d^{-\beta + c/\log y},$$

where $c > 0$ is an absolute constant.

Proof of Proposition 1.4.1. We proceed by simplifying the more general proof of Ivić and Tenenbaum [25, Section 6]. Let 1_{k-free} denote the characteristic function of k-free numbers. Let $\lambda^{(k)}(n)$ be the multiplicative function defined by $\lambda^{(k)}(1) = 1$ and

$$\lambda^{(k)}(p^a) := \begin{cases} -1 & \text{if } a = k, \\ 0 & \text{otherwise} \end{cases}$$

for every prime number p. It follows from direct computation that

$$\sum_{d|n} \lambda^{(k)}(d) = 1_{k-free}(n).$$

Substitute the identity above and change the order of summation yields

$$\begin{aligned}\sum_{n\in S(N,y)} 1_{k-free}(n) &= \sum_{n\in S(N,y)}\sum_{d|n}\lambda^{(k)}(d)\\ &= \sum_{d\in S(N,y)}\lambda^{(k)}(d)\left|S\left(\frac{N}{d},y\right)\right|\\ &= \sum_{d\in S(N,y),d\le Y}\lambda^{(k)}(d)\left|S\left(\frac{N}{d},y\right)\right| + \sum_{d\in S(N,y),d>Y}\lambda^{(k)}(d)\left|S\left(\frac{N}{d},y\right)\right|\\ &=: I+II,\end{aligned}$$

where Y is chosen to be $\log^{6/\epsilon} y$. We shall see that the main contribution comes from the first sum, whereas the second sum is considerably smaller.

We begin with the first sum. Firstly, notice that for sufficiently large N (which implies that y is sufficiently large), we have $y > Y$, and so

$$\sum_{d\in S(N,y),d\le Y}\lambda^{(k)}(d)\left|S\left(\frac{N}{d},y\right)\right| = \sum_{d\le Y}\lambda^{(k)}(d)\left|S\left(\frac{N}{d},y\right)\right|.$$

In order to obtain the desired main term, we extend the sum over all integers d, and show that this leads to an error term of size $O_\epsilon(|S(N,y)|\log\log y/\log y)$. By Lemma 1.4.2, in this range of d we have

$$\left|S\left(\frac{N}{d},y\right)\right| = |S(N,y)|\,d^{-\beta}\left(1+O_\epsilon\left(\frac{\log\log y}{\log y}\right)\right).$$

Since $k\beta > 1$, we may use the usual series expansion of zeta function to see that $\sum_{d=1}^{\infty}\lambda^{(k)}(d)d^{-\beta} = \prod_{p\text{ prime}}(1-p^{-k\beta}) = \zeta^{-1}(k\beta)$. It follows that

$$I = |S(N,y)|\left(\zeta^{-1}(2\beta) + O_\epsilon\left(\sum_{d>Y}|\lambda^{(k)}(d)|d^{-\beta} + \frac{\log\log y}{\log y}\sum_{d\le Y}|\lambda^{(k)}(d)|d^{-\beta}\right)\right).$$

Thus, it suffices to bound the error terms in the equation above. Since $\lambda^{(k)}$ is multiplicative, it is supported on kth powers, and $|\lambda^{(k)}(n)| \le 1$ for all $n \in \mathbb{N}$. It follows that

$$\sum_{d>Y}|\lambda^{(k)}(d)|d^{-\beta} \le \sum_{m>Y^{1/k}} 1_{k-free}(m)m^{-k\beta} \ll_\epsilon Y^{1/k-\beta} \le \frac{1}{\log y}.$$

The other term can be bounded by the same method; indeed, one has

$$\sum_{d\le Y}|\lambda^{(k)}(d)|d^{-\beta} \le \sum_{m=1}^{\infty} 1_{k-free}(m)m^{-k\beta} \ll_\epsilon 1.$$

To bound the second sum, we use Lemma 1.4.3 to obtain

$$II \ll |S(N,y)| \sum_{d>Y} |\lambda^{(k)}(d)| d^{-\beta+c/\log y} = |S(N,y)| \sum_{m \geq Y^{1/k}} m^{-k(\beta+c/\log y)}.$$

For sufficiently large y, we can bound the second sum as before, getting

$$II \ll_\epsilon |S(N,y)| Y^{-\beta+1/k+c/\log y} \leq \frac{|S(N,y)|}{\log y}.$$

Combining the bounds of sums I and II above, we can conclude that

$$\sum_{n \in S(N,y)} 1_{k-free}(n) = \frac{|S(N,y)|}{\zeta(k\beta)} \left(1 + O_\epsilon\left(\frac{\log\log y}{\log y}\right)\right),$$

which implies Proposition 1.4.1 by (1.8). □

Proof of Proposition 1.1.4. Since the first order derivative of the function $f(y) = \log(1 + y/\log N)/\log y$ is positive for all $y \geq (\log N)^{1+1/(k-1)+\epsilon}$, it follows from (1.1) that

$$\begin{aligned}
\alpha(N,y) =& \frac{\log(1+y/\log N)}{\log y}\left(1 + O\left(\frac{\log\log(1+y)}{\log y}\right)\right) \\
\geq& \left(1 - \frac{1}{1+1/(k-1)+\epsilon}\right)(1+o(1)) \\
\geq& \frac{1}{k} + \frac{\epsilon}{4}.
\end{aligned}$$

In particular, $k\alpha(N,y)$ is bounded away from 1 and so

$$\frac{1}{\zeta(k\alpha(N,y))} \ll_{k,\epsilon} 1.$$

The result then follows from Proposition 1.4.1. □

Chapter 2

Difference sets without shifted primes

2.1 Introduction

A set X consisting of positive integers is called *difference intersective*, if $X \cap (Y - Y) \neq \emptyset$ whenever Y has positive upper density. The notion of intersectivity is due to Ruzsa and partly to Tijdeman [45]. In 1978, Sárközy [49, 50, 51] published a series of papers studying difference sets of sequences of the integers. He showed that both the set $\mathbb{P} - 1$ and the set of perfect squares are difference intersective. More recently, people have also considered other related problems, which involve both the integers and the polynomials, see Rice [39, 40], see also Lê [29].

In this chapter, we shall prove the following result [57].

Theorem 2.1.1. *Let N be a positive integer. Suppose that the difference of any two elements of a set $A \subseteq [N]$ is never one less than a prime number, then $|A| = O(N \exp(-c(\log N)^{1/3}))$ where $c > 0$ is absolute.*

Sárközy [51] established the first explicit bound of $|A|$ using the Hardy–Littlewood method and an iteration argument. He showed that the cardinality of the set is bounded by $|A| = O(N \exp(-(2 + o(1)) \log\log\log N))$. Subsequently, the bound $|A| = O(N \exp(-\omega(N) \log\log\log N))$, where $\omega(N)$ tends to infinity as $N \to \infty$,[1] was obtained by Lucier [31] who improved Sárközy's argument using the methods of Pintz, Steiger and Szemerédi [36][2]. By exploiting a dichotomy depending on whether an exceptional zero of Dirichlet L-functions occurs or not, Ruzsa and Sanders [47] proved that $|A| = O(N \exp(-c(\log N)^{1/4}))$, where $c > 0$ is an absolute constant.

[1] More precisely, $\omega(N) \sim c \log\log\log\log\log N$ for some $c > 0$.

[2] The Pintz–Steiger–Szemerédi bound was improved by Bloom and Maynard [3] in 2020. It is not yet known whether the Bloom–Maynard approach can be used to improve Theorem 2.1.1.

Our key ingredient is a new major arc estimate. The underlying motivation of building the estimate is to reduce the size of the error terms, which plays a crucial role in the method of Ruzsa and Sanders. To achieve this, we need to control the contribution of the zeros of L-functions and avoid additional errors caused by the length of the major arcs.

The main tool for bounding the error caused by zeros of L-functions is the exceptional zero repulsion [27, Principle 3, Section 18.1]. Roughly speaking, it compensates for the reduced size of the main term when an exceptional zero occurs. In fact, on the way of proving Linnik's theorem, one obtains a better error term in the prime number theorem in arithmetic progressions (see Iwaniec and Kowalski [27, Equations (18.87) and (18.89)]) than the one used by Ruzsa and Sanders [47, Proposition 4.7].

The error terms would be too large for the combinatorial arguments should we integrate the error term of the prime number theorem directly. To deal with this issue, we shall exploit all terms involving zeros with large real parts and small imaginary parts using the classical explicit formula to improve the major arc estimate.

In fact, the sets $\mathbb{P} \pm 1$ are the only shifts of the primes which are difference intersective. To see this, let $|t| \geq 2$ and consider the shift[3] $\mathbb{P} - t$. Let $b > 3$ be a positive integer. Consider the congruence class

$$\{bt, 2bt, ..., kbt, ...\}$$

which has a positive upper density. Since any (positive) difference between two elements of this congruence class can be written as ibt for some integer $i \geq 1$, it follows that $ibt \pm t$ is composite. For $t = 0$, we may use any arithmetic progression whose common difference is composite to show that $\mathbb{P}$ is not intersective.

There is a large gap between the upper bound from Theorem 2.1.1 and the cardinality of any known example of a subset of an initial segment of the integers whose difference set avoids any number which is one less than a prime. Erdős and Sárközy [10] proved the first nontrivial result, which was later improved by the following theorem due to Ruzsa [46, Section 5][4].

Theorem 2.1.2. *Let N be a positive integer. There exists $A \subseteq [N]$ with $|A| \geq \exp(((\log 2)/2 + o(1)) \log N / \log \log N)$ such that the difference set $A - A$ does not contain any number which is one less or one more than a prime.*

[3]Strictly speaking, the set we are interested in is $(\mathbb{P} - t) \cap \mathbb{N}$, since our definition of intersectivity concerns with subsets of positive integers.

[4]Ruzsa also gives a slightly different construction to find a slightly bigger set whose difference set avoids $\mathbb{P} - 1$.

Even assuming the generalised Riemann hypothesis, which would simplify our work in Section 2.3 (see also Ruzsa and Sanders [47]), we are still unable to produce a much better bound. We record the following question asked by Ruzsa [45].

Question 2.1.3. Under the generalised Riemann hypothesis, can one obtain a bound $|A| = O(N^{1-c+o(1)})$ for some positive absolute constant c in Theorem 2.1.1?

The chapter is organised as follows. We provide a detailed proof of the major arc estimate in Section 2.3. The details of the set-up used in the proof is provided in Section 2.2. In Section 2.4, we use our improved major arc estimates to modify the iteration scheme of Ruzsa and Sanders and get our result.

2.2 Zeros of Dirichlet L-functions

Let N, q be positive integers. For any Dirichlet character χ of modulus q and $T \geq 1$, we define

$$Z(\chi; T) := \{\rho : L(\rho, \chi) = 0, \Re(\rho) \geq 1/2, |\Im(\rho)| \leq T\}.$$

and

$$Z(q; T) := \bigcup_{\chi \pmod q} Z(\chi; T).$$

We treat a zero with multiplicity m as m elements in the zero sets above, and we use $|Z(\chi; T)|$ to denote the cardinality of this multiset.

The following lemma is a consequence of various results concerning zeros of Dirichlet L-functions listed in Chapter 0. It shows that Definition 2.3.1 of the next section is exhaustive.[5]

Lemma 2.2.1. *There are positive absolute constants $C_2 \geq 10$ and c_4 such that for any $D \geq 2$ and $T \geq 1$, the following assertions hold.*

Suppose there exists a primitive character χ_D such that χ_D has modulus $d_D \leq D$ and $L(s, \chi_D)$ has a zero β_D in the region

$$\Re(s) \geq 1 - \frac{c_4}{C_2 \log(DT)}, \quad |\Im(s)| \leq T. \tag{2.1}$$

Then,

(i) the zero β_D is real and simple, and it is the only zero of $L(s, \chi_D)$ in the region (2.1),

[5] Due to the use of constants from the relevant section of Chapter 0, the subscripts of the constants in this chapter are chosen accordingly.

(ii) there does not exist any other primitive character χ of modulus $d \leq D^{C_2}$ such that $L(s,\chi)$ has a zero in the region (2.1),

(iii) (exceptional zero repulsion) for any $d_D \mid d$, all other zeros in $Z(dq;T)$ are in the region

$$\Re(s) \leq 1 - \frac{c_2|\log((1-\beta_D)\log(dqT))|}{\log(dqT)},\ |\Im(s)| \leq T.$$

Here, c_2 is the constant from Chapter 0.

Proof. Let $c_4 = \min\{c_1/2, c_3/2\}$. Since

$$\Re(\beta_D) \geq 1 - \frac{c_4}{C_2\log(DT)} > 1 - \frac{c_1}{\log(dT)},$$

it follows from Theorem 0.3.1 that assertion (i) is true. Suppose assertion (ii) fails, then there exists χ'_D of modulus $d'_D \leq D^{C_2}$ and a zero β'_D of $L(s,\chi'_D)$ such that β'_D is in the region (2.1). Again it follows from Theorem 0.3.1 that β'_D is real and simple. Thus, by (0.7), the zeros β_D, β'_D satisfy

$$\begin{aligned} 1 - \frac{c_4}{C_2\log(DT)} &\leq \min\{\beta_D, \beta'_D\} \leq 1 - \frac{c_3}{\log(DD^{C_2})} \\ &\leq 1 - \frac{2c_4}{(C_2+1)\log D} < 1 - \frac{c_4}{C_2\log(DT)}, \end{aligned}$$

which is a contradiction. Therefore, our assumption is false and the claim (ii) follows.

Assertion (iii) follows from our choice of the constant c_4 and Theorem 0.3.3. □

2.3 Major arc estimates

For any positive integers N, d, let

$$F_{N,d}(n) := \Lambda(dn+1)1_{[N]}(n),$$

where Λ is the von Mangoldt function and $1_{[N]}$ is the characteristic function of the set $[N]$.

We shall use such functions to detect primes and prime powers in arithmetic progressions, which will in turn provide the desired structure in the iteration scheme. The iteration is done by an energy increment argument, and we shall need appropriate estimates of the Fourier transform of $F_{N,d}$ to effect this.

To establish such estimates on the major arcs, we shall split into two different cases depending on whether a possible exceptional zero, as defined in Definition 2.3.1

below, exists or not. More precisely, our set-up involves two parameters, one of which controls the modulus of the exceptional primitive character and another clarifies the height of the rectangle which contains the zeros we need to consider, and they jointly quantify our notion of being exceptional.

Definition 2.3.1. Let C_2 and c_4 be the constants from Lemma 2.2.1. Let $D \geq 2$ and $T \geq 1$. We say that (D, T) is *exceptional* if there exists a unique primitive character χ_D such that χ_D has modulus $d_D \leq D$, and $L(s, \chi_D)$ has a zero β_D which is real and simple and satisfies $\beta_D \geq 1 - c_4/(C_2 \log(DT))$. We call χ_D the *exceptional character* and β_D the *exceptional zero*. Otherwise, we say that (D, T) is *unexceptional*.

The goal of this section is to prove the following result.

Proposition 2.3.2 (Major arc estimates)**.** *There exist positive absolute constants C_3 and c_5 such that the following holds.*

Let $T, D \geq 2$ and let N be a positive integer such that $N > (DT)^{C_3}$.

(1) Suppose that (D, T) is unexceptional. Then for any $\delta \in [-1/2, 1/2]$ and any positive integers a, d, q satisfying $(a, q) = 1$ and $dq \leq D$, we have

$$\begin{aligned}\left|\widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right)\right| &\leq \frac{2|\widehat{F_{N,d}}(0)|}{\phi(q)} + O\left(\frac{dNq}{\phi(d)\phi(q)}\exp\left(-c_5\frac{\log N}{\log(DT)}\right)\right) \\ &\quad + O\left((1+N|\delta|)\frac{dqN\log^2 N}{T}\right).\end{aligned}$$

We also have

$$\left|\widehat{F_{N,d}}(0)\right| \geq \frac{dN}{2\phi(d)} - O\left(\frac{dN\log^2 N}{T}\right).$$

(2) Suppose that (D, T) is exceptional. Then for any $\delta \in [-1/2, 1/2]$ and any positive integers a, d, q satisfying $(a, q) = 1$, $dq \leq D^{C_2}$ and $d_D \mid d$, we have

$$\begin{aligned}\left|\widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right)\right| &\leq \frac{2|\widehat{F_{N,d}}(0)|}{\phi(q)} \\ &\quad + O\left(\frac{dNq}{\phi(d)\phi(q)}(1-\beta_D)\log(dqT)\exp\left(-c_5\frac{\log N}{\log(dqT)}\right)\right) \\ &\quad + O\left((1+N|\delta|)\frac{dqN\log^2 N}{T}\right).\end{aligned}$$

We also have

$$\left|\widehat{F_{N,d}}(0)\right| \geq \frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{4c_4} - O\left(\frac{dN\log^2 N}{T}\right).$$

Here C_2 and c_4 are the constants in Definition 2.3.1.

The purpose of our first lemma is to write the Fourier transform of the function $F_{N,d}$ in a way which allows us to use known techniques related to exponential sums and $\psi(x,\chi)$.

Lemma 2.3.3. *Let N, a, d, q be positive integers and let $-1/2 \leq \kappa \leq 1/2$. One has*

$$\widehat{F_{N,d}}\left(\frac{a}{q}+\kappa\right) = \frac{1}{\phi(dq)}\sum_{\chi \pmod{dq}} e\left(\frac{\kappa}{d}\right) S_{dN+1}\left(\frac{\kappa}{d},\chi\right) G_{a,q,d,\chi} + O\left((\log(dN))(\log q)\right),$$

where

$$S_x(\delta,\chi) := \sum_{n\leq x}\Lambda(n)\chi(n)e\left(-n\delta\right)$$

and

$$G_{a,q,d,\chi} := \sum_{m=0}^{q-1} e\left(-\frac{am}{q}\right)\overline{\chi}(dm+1). \tag{2.2}$$

Proof. Firstly, notice that

$$\begin{aligned}\widehat{F_{N,d}}\left(\frac{a}{q}+\kappa\right) &= \sum_{n\leq N} F_{N,d}(n)e\left(-\left(\frac{a}{q}+\kappa\right)n\right)\\ &= \sum_{n\leq N}\Lambda(dn+1)e\left(-\left(\frac{a}{q}+\kappa\right)n\right).\end{aligned}$$

By applying a change of variables and rearranging the sum, it follows that

$$\widehat{F_{N,d}}\left(\frac{a}{q}+\kappa\right) = \sum_{m=0}^{q-1} e\left(-\frac{am}{q}\right)\sum_{\substack{l\leq dN+1\\ l\equiv 1 \pmod d\\ \frac{l-1}{d}\equiv m \pmod q}}\Lambda(l)e\left(-\frac{\kappa(l-1)}{d}\right).$$

Since the inner sum above runs over an arithmetic progression, it follows from the orthogonality of Dirichlet characters (see Iwaniec and Kowalski [27, p.45 and p.121]) that

$$\widehat{F_{N,d}}\left(\frac{a}{q}+\kappa\right) = \sum_{m=0}^{q-1} e\left(-\frac{am}{q}\right)\sum_{\chi \pmod{dq}}\frac{1}{\phi(dq)}\overline{\chi}(dm+1)e\left(\frac{\kappa}{d}\right)S\left(\frac{\kappa}{d},\chi\right).$$

Since the sum above is finite and there are only two terms which contain m, we can change the order of summations and obtain the result. □

The expansion using Dirichlet characters helps us to reduce the task of obtaining major arc estimates to estimating sums of form $S_x(\delta, \chi)$ where δ is relatively close to 0. This turns out to provide technical convenience in partial summation. Moreover, by partial summation and classical complex-analytic number theory, we can express the sum $S_x(\delta, \chi)$ in terms of certain zeros of $L(s, \chi)$ up to a small error term. Consequently, we can perform integration by parts when we estimate the impact of these zeros, instead of simply integrating the absolute value of the error term caused by them. This allows us to improve the major arc estimates at points which are not very close to rational numbers with small denominators.

By the truncated explicit formula, one has the following estimates, see Iwaniec and Kowalski [27, Proposition 5.25].[6]

Proposition 2.3.4. *Let d, q be positive integers and $x > 0$. For any non-principal character χ of modulus dq and any $1 \leq T \leq x^{1/4}$, one has*

$$\sum_{n \leq x} \Lambda(n)\chi(n) = -\sum_{\rho \in Z(\chi;T)} \frac{x^\rho}{\rho} + O\left(\frac{x \log^2(dqx)}{T}\right).$$

For the principal character χ' of modulus dq and any $1 \leq T \leq x^{1/4}$, one has

$$\sum_{n \leq x} \Lambda(n)\chi'(n) = x - \sum_{\rho \in Z(\chi';T)} \frac{x^\rho}{\rho} + O\left(\frac{x \log^2(dqx)}{T}\right).$$

The aim of the next lemma is to connect the sum $S(\delta, \chi)$ with zeros of $L(s, \chi)$, which is done by partial summation.

Lemma 2.3.5. *Let N, q, d be positive integers. For any Dirichlet character χ of modulus dq, $-1/2 \leq \delta \leq 1/2$ and[7] $1 \leq T, dq \leq N^{1/32}$, we have*

$$\begin{aligned} S_{dN+1}(\delta, \chi) = &\int_{N^{1/8}}^{dN+1} \left(1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} t^{\rho-1}\right) e^{-2\pi i \delta t} dt \\ &+ O\left((1 + dN|\delta|)\frac{dN \log^2(dqN)}{T}\right). \end{aligned}$$

Here $1_{\chi'}(\chi) = 1$ if χ is the principal character, and $1_{\chi'}(\chi) = 0$ otherwise.

[6] Here we have an extra restriction $\Re(\rho) \geq 1/2$ compared to the explicit formula given in Iwaniec and Kowalski [27, Proposition 5.25]. The reason is that we have absorbed the error caused by the zeros whose real parts are smaller than $1/2$ into the error term. This can be done by using the classical estimate of the number of zeros whose imaginary parts are less than T, see Iwaniec and Kowalski [27, Theorem 5.24].

[7] The exponent $1/32$ is introduced purely for technical reasons. It follows from the restriction of T in Proposition 2.3.4 and our truncation of integral in the proof of Lemma 2.3.5.

Proof. For $\delta = 0$, we have $S_{dN+1}(0,\chi) = \sum_{n \le dN+1} \Lambda(n)\chi(n)$, and so an application of Proposition 2.3.4 yields

$$\begin{aligned} S_{dN+1}(0,\chi) &= (dN+1)1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{(dN+1)^\rho}{\rho} + O\left(\frac{dN\log^2(dqN)}{T}\right) \\ &= \int_{N^{1/8}}^{dN+1} \left(1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} t^{\rho-1}\right) dt + O\left(\frac{dN\log^2(dqN)}{T}\right). \end{aligned}$$

From now on, we assume $\delta \neq 0$. By Abel's summation formula, we have

$$S_{dN+1}(\delta,\chi) = e^{-2\pi i\delta(dN+1)} \sum_{n \le dN+1} \Lambda(n)\chi(n) + 2\pi i\delta \int_1^{dN+1} \left(\sum_{1<n\le t} \Lambda(n)\chi(n)\right) e^{-2\pi it\delta} dt.$$

It follows that

$$\begin{aligned} S_{dN+1}(\delta,\chi) = e^{-2\pi i\delta(dN+1)} &\sum_{n \le dN+1} \Lambda(n)\chi(n) \\ &+ 2\pi i\delta \int_{N^{1/8}}^{dN+1} \left(\sum_{1<n\le t} \Lambda(n)\chi(n)\right) e^{-2\pi it\delta} dt + O\left(|\delta| N^{1/4}\log N\right), \end{aligned}$$

where the error term is obtained by bounding the integral over $[1, N^{1/8}]$ by

$$\left|2\pi i\delta \int_1^{N^{1/8}} \left(\sum_{1<n\le t} \Lambda(n)\chi(n)\right) e^{-2\pi it\delta} dt\right| \le 2\pi|\delta| N^{1/8} N^{1/8}\log N = O\left(|\delta| N^{1/4}\log N\right).$$

By substituting the expressions given in Proposition 2.3.4, we have

$$\begin{aligned} S_{dN+1}(\delta,\chi) = e^{-2\pi i\delta(dN+1)} &\left((dN+1)1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{(dN+1)^\rho}{\rho}\right) \\ + 2\pi i\delta \int_{N^{1/8}}^{dN+1} &\left(t1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{t^\rho}{\rho}\right) e^{-2\pi i\delta t} dt + O\left((1+dN|\delta|)\frac{dN\log^2(dqN)}{T}\right). \end{aligned}$$

The error term above follows from integrating the error term in the explicit formula.

Since $\delta \neq 0$, by integration by parts, we have

$$\begin{aligned}
&2\pi i\delta \int_{N^{1/8}}^{dN+1} \left(t1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{t^\rho}{\rho} \right) e^{-2\pi i\delta t} dt \\
&= - \left[e^{-2\pi i\delta t} \left(t1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{t^\rho}{\rho} \right) \right]_{N^{1/8}}^{dN+1} \\
&\quad + \int_{N^{1/8}}^{dN+1} \left(1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} t^{\rho-1} \right) e^{-2\pi i\delta t} dt \\
&= - e^{-2\pi i\delta(dN+1)} \left((dN+1)1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} \frac{(dN+1)^\rho}{\rho} \right) \\
&\quad + \int_{N^{1/8}}^{dN+1} \left(1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} t^{\rho-1} \right) e^{-2\pi i\delta t} dt + O\left(N^{1/8} dqT \log(dqT) \right),
\end{aligned}$$

where the error term above comes from bounding the term $e^{-2\pi i\delta N^{1/8}}(N^{1/8}1_{\chi'}(\chi) - \sum_{\rho \in Z(\chi;T)} N^{\rho/8}/\rho)$ using the known upper bound[8] for $|Z(\chi;T)|$ (see Iwaniec and Kowalski [27, Theorem 5.24]). The lemma follows from the equations above.

□

Following the lemma above, we need to bound the contribution of the zeros in the region $Z(\chi;T)$. The proposition below is introduced for this purpose, and it can be shown by using arguments from Iwaniec and Kowalski [27, Section 18.4]. It turns out that in the unexceptional case, the zero density estimate recorded in Theorem 0.3.2 will be strong enough to produce the desired bound. In the exceptional situation, this is no longer the case, since the size of the main term could be reduced due to the term which contains the exceptional zero. To compensate for this, we shall use the exceptional repulsion given in Theorem 0.3.3.

Proposition 2.3.6. *There exist absolute constants*[9] C_4 *and* c_6 *such that for all* $x, D, T \geq 2$ *satisfying* $x > (DT)^{C_4}$*, we have the following.*

[8]It follows from Iwaniec–Kowalski [27, Theorem 5.24] that for every primitive χ modulo q one has

$$|Z(\chi;T)| \leq \frac{T}{\pi} \log \frac{qT}{2\pi e} + O(q(T+3))$$

where the implicit constant is absolute. Hence, for every χ modulo dq, one has

$$|Z(\chi;T)| \ll dqT \log(dqT).$$

[9]The constants C_4 and c_6 depend on c_2, c_4, C_1 and the implicit constant in the Landau notation.

(i) If (D,T) is unexceptional, then for any $q,d\geq 1$ satisfying $dq\leq D$, we have

$$\sum_{\rho\in Z(dq;T)}|x^{\rho-1}|=O\left(\exp\left(-c_6\frac{\log x}{\log(DT)}\right)\right).$$

(ii) If (D,T) is exceptional, then for any $q,d\geq 1$ satisfying $dq\leq D^{C_2}$ and $d_D\mid d$, we have

$$\sum_{\substack{\rho\in Z(dq;T)\\ \rho\neq\beta_D}}|x^{\rho-1}|=O\left((1-\beta_D)\log(dqT)\exp\left(-c_6\frac{\log x}{\log(dqT)}\right)\right).$$

Here, C_2 is the constant from Definition 2.3.1.

Proof. Assertion (i) follows from an integration by parts argument and the log-free zero density estimate (0.6). Indeed, by Definition 2.3.1 and Lemma 2.2.1, every zero in the set $Z(dq;T)$ has real part less than $1-c_4/(C_2\log(DT))$, and so

$$\sum_{\rho\in Z(dq;T)}|x^{\rho-1}|=x^{-1/2}N_{dq}(1/2,T)+\log x\int_{1/2}^{1-c_4/(C_2\log(DT))}N_{dq}(\alpha,T)x^{\alpha-1}d\alpha.$$

By the log-free zero density estimate (0.6), we have

$$\sum_{\rho\in Z(dq;T)}|x^{\rho-1}|\leq C_1x^{-1/2}(dqT)^{C_1/2}+C_1\log x\int_{1/2}^{1-c_4/(C_2\log(DT))}((dqT)^{C_1}/x)^{1-\alpha}d\alpha.$$

Notice that for any $C\geq 2C_1$ and $x\geq(DT)^C$, we have $(dqT)^{C_1}/x\leq(DT)^{C_1}/x\leq x^{-1/2}$, and so for any $c\leq c_4/(2C_2)$ we have

$$\sum_{\rho\in Z(dq;T)}|x^{\rho-1}|\ll\left(\frac{1}{x}\right)^{c_4/(2C_2\log(DT))}\ll\exp\left(-c\frac{\log x}{\log(DT)}\right).\tag{2.3}$$

In the situation considered in assertion (ii), it follows from Definition 2.3.1, $d_D\mid d$ and (0.2) that there exists a character χ of modulus dq such that $L(s,\chi)$ has a zero $\beta_D\geq 1-c_4/(C_2\log(DT))$. Since $d_D\mid d$, it follows from Lemma 2.2.1 that any other zero in the set $Z(dq;T)$ has real part less than

$$a:=1-\frac{c_2|\log((1-\beta_D)\log(dqT))|}{\log(dqT)}.$$

Hence, applying a similar integration by parts argument and the log-free zero density estimate (0.6) yields

$$\begin{aligned}\sum_{\substack{\rho\in Z(dq;T)\\ \rho\neq\beta_D}}|x^{\rho-1}|=&x^{-1/2}N_{dq}(1/2,T)+\log x\int_{1/2}^{1-a}N_{dq}(\alpha,T)x^{\alpha-1}d\alpha\\ \leq&C_1x^{-1/2}(dqT)^{C_1/2}+C_1\log x\int_{1/2}^{1-a}((dqT)^{C_1}/x)^{1-\alpha}d\alpha.\end{aligned}$$

Thus, for any $C \geq 2C_1$ and $x \geq (DT)^C$, we have

$$\sum_{\substack{\rho \in Z(dq;T) \\ \rho \neq \beta_D}} |x^{\rho-1}| \ll x^{-1/4} + \exp\left(-\frac{c_2}{2} \frac{|\log((1-\beta_D)\log(dqT))|}{\log(dqT)} \log x\right). \tag{2.4}$$

On the other hand, since $(1-\beta_D)\log(dqT) \leq c_4 < 1$ and $x > (DT)^C$, it follows that for all d, q satisfy $dq \leq D^{C_2}$ and $C > 4C_2/c_2$, we have $c_2 \log x / \log(dqT) > 4$, and so we have

$$\begin{aligned} &\exp\left(-\frac{c_2}{2} \frac{|\log((1-\beta_D)\log(dqT))|}{\log(dqT)} \log x\right) \\ =&((1-\beta_D)\log(dqT))^{c_2 \log x/(2\log(dqT))} \\ \leq&(1-\beta_D)\log(dqT) c_4^{c_2 \log x/(4\log(dqT))}. \end{aligned}$$

Since $c_4 < 1$, we can conclude that for any $c \leq (c_2 \log c_4^{-1})/4$ we have

$$\sum_{\substack{\rho \in Z(dq;T) \\ \rho \neq \beta_D}} |x^{\rho-1}| \leq O\left((1-\beta_D)\log(dqT) \exp\left(-c\frac{\log x}{\log(dqT)}\right)\right). \tag{2.5}$$

Combining the conditions stated before (2.3) (2.4) and (2.5), we can deduce that our assertions hold. □

We shall also need the following bound. Recall that for any given $a, d, q \geq 1$ such that $(a, q) = 1$ and any character χ of modulus dq, the quantity $G_{a,q,d,\chi}$ is defined in (2.2) as

$$G_{a,q,d,\chi} = \sum_{m=0}^{q-1} e\left(-\frac{am}{q}\right) \overline{\chi}(dm+1).$$

Lemma 2.3.7. *Let d, q be positive integers. Let χ' be the principal character of modulus dq. Let a be an integer which is coprime to q. We have*

$$|G_{a,q,d,\chi'}| \leq 1.$$

Proof. First, suppose $(d, q) = 1$. Then let d^{-1} be the inverse of d mod q. By substituting $t = dm + 1$ we have

$$G_{a,q,d,\chi'} = \sum_{\substack{m=0 \\ (dm+1,q)=1}}^{q-1} e\left(-\frac{am}{q}\right) = e\left(\frac{d^{-1}a}{q}\right) \sum_{\substack{t=0 \\ (t,q)=1}}^{q-1} e\left(-\frac{ad^{-1}t}{q}\right).$$

Since $(ad^{-1}, q) = 1$, we can conclude that the claimed bound holds by taking $s = -ad^{-1}$ in the Ramanujan sum formula [27, Section 3.2, Equation (3.3)]

$$\sum_{t=0,(t,q)=1}^{q-1} e\left(\frac{st}{q}\right) = \mu\left(\frac{q}{(s,q)}\right)\frac{\phi(q)}{\phi(q/(s,q))}.$$

If $(d,q) = g > 1$, then we can write $q = gq_1$ for some $q_1 < q$. It follows that each $0 \leq m \leq q-1$ can be written uniquely as $m = kq_1 + l$ for some $0 \leq k \leq g-1$ and $0 \leq l \leq q_1 - 1$. Since

$$(dl+1, dq) = (dl+1, q) = (d(l+kq_1)+1, q) = (d(l+kq_1)+1, dq)$$

for all integers k, it follows that

$$\begin{aligned} G_{a,q,d,\chi'} &= \sum_{\substack{l=0 \\ (dl+1,dq)=1}}^{q_1-1} \sum_{k=0}^{g-1} e\left(-\frac{a(kq_1+l)}{q}\right) \\ &= \sum_{\substack{l=0 \\ (dl+1,dq)=1}}^{q_1-1} e\left(-\frac{al}{q}\right) \sum_{k=0}^{g-1} e\left(-\frac{ak}{g}\right). \end{aligned}$$

On summing the geometric progressions (using $(a,g) = 1$), we see that the inner sum is 0 giving the result. □

Proof of Proposition 2.3.2. We shall first prove the easier assertion where (D,T) is unexceptional.

By Lemma 2.3.3, Lemma 2.3.5 and the trivial bound $|G_{a,q,d,\chi}| \leq q$, we have

$$\begin{aligned} \widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right) = \frac{e^{2\pi i\delta/d}}{\phi(dq)} \int_{N^{1/8}}^{dN} e^{-2\pi i\delta t/d} G_{a,q,d,\chi'} dt \\ - \frac{e^{2\pi i\delta/d}}{\phi(dq)} \sum_{\chi \pmod{dq}} \int_{N^{1/8}}^{dN} \left(\sum_{\rho \in Z(\chi;T)} t^{\rho-1}\right) e^{-2\pi i\delta t/d} G_{a,q,d,\chi} dt \\ + O\left((1+N|\delta|)\frac{dqN\log^2 N}{T}\right), \end{aligned}$$

where χ' is the principal character of modulus dq. By Lemma 2.3.7, for each $N^{1/8} \leq t \leq dN$ we have

$$\left|\frac{e^{2\pi i\delta/d}}{\phi(dq)} \int_{N^{1/8}}^{dN} e^{-2\pi i\delta t/d} G_{a,q,d,\chi'} dt\right| \leq \frac{1}{\phi(d)\phi(q)} \int_{N^{1/8}}^{dN} dt = \frac{dN}{\phi(d)\phi(q)} + O(N^{1/4}\log N).$$

By using[10] $\phi(dq) \geq \phi(d)\phi(q)$ and taking the trivial bound $|G_{a,q,d,\chi}| \leq q$, we can bound the second term by

$$\begin{aligned}&\left|\frac{e^{2\pi i\delta/d}}{\phi(dq)}\sum_{\chi \pmod{dq}}\int_{N^{1/8}}^{dN}\left(\sum_{\rho\in Z(\chi;T)} t^{\rho-1}\right)e^{-2\pi i\delta t/d}G_{a,q,d,\chi}dt\right| \\ \leq&\frac{q}{\phi(d)\phi(q)}\int_{N^{1/8}}^{dN}\left(\sum_{\chi \pmod{dq}}\sum_{\rho\in Z(\chi;T)}|t^{\rho-1}|\right)dt.\end{aligned}$$

Thus, by Proposition 2.3.6, we have

$$\begin{aligned}&\left|\frac{e^{2\pi i\delta/d}}{\phi(dq)}\sum_{\chi \pmod{dq}}\int_{N^{1/8}}^{dN}\left(\sum_{\rho\in Z(\chi;T)} t^{\rho-1}\right)e^{-2\pi i\delta t/d}dtG_{a,q,d,\chi}\right| \\ &\qquad = O\left(\frac{dqN}{\phi(d)\phi(q)}\exp\left(-\frac{c_6}{8}\left(\frac{\log N}{\log(DT)}\right)\right)\right). \quad (2.6)\end{aligned}$$

To obtain the second estimate, we take $q = 1$ and $\delta = 0$, so the contribution of the main term in the asymptotic can be computed explicitly; it is equal to

$$\frac{1}{\phi(d)}\int_{N^{1/8}}^{dN+1} dt = \frac{dN}{\phi(d)} + O(N^{1/4}\log N).$$

Since $N \geq (DT)^{C_3}$, by taking C_3 large enough (depending on the implicit constant and c_4), we can make the error term in (2.6) at most $dN/(4\phi(d))$ and the lower bound for $|\widehat{F_{N,d}}(0)|$ follows. This completes the proof of the unexceptional case.

The exceptional case considered in (ii) is technically more complicated, since we need to evaluate the impact of the exceptional zero and compare it with the amount of the repulsion.

By substituting Lemma 2.3.5 into Lemma 2.3.3, we have

$$\begin{aligned}\widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right) =& \frac{e^{2\pi i\delta/d}}{\phi(dq)}\int_{N^{1/8}}^{dN} e^{-2\pi it\delta/d}\left(G_{a,q,d,\chi'} - t^{\beta_D-1}G_{a,q,d,\chi'\chi_D}\right)dt \\ &- \frac{e^{2\pi i\delta/d}}{\phi(dq)}\sum_{\chi \pmod{dq}}\int_{N^{1/8}}^{dN}\left(\sum_{\substack{\rho\in Z(\chi;T)\\ \rho\neq\beta_D}} t^{\rho-1}\right)e^{-2\pi it\delta/d}G_{a,q,d,\chi}dt \\ &+ O\left((1+N|\delta|)\frac{dqN\log^2 N}{T}\right),\end{aligned}$$

[10]This inequality can be shown by using the formula $\phi(p^k) = p^k(1-1/p)$ where p is prime and the well-known fact that ϕ is multiplicative.

where χ' is the principal character of modulus dq.

Since the modulus of χ_D is a divisor of d, we have $\chi_D(dm+1)=1$ and it follows that

$$\chi'(dm+1)=\chi'\chi_D(dm+1)$$

for all $0\le m\le q-1$. Therefore, the first integral appearing in the expression for $\widehat{F_{N,d}}(a/q+\delta)$ above is equal to

$$\frac{e^{2\pi i\delta/d}G_{a,q,d,\chi'}}{\phi(dq)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)e^{-2\pi i\delta t/d}dt.$$

On the other hand, by the second assertion of Proposition 2.3.6, for sufficiently large C_3 (which is at least $8C_2$) and $t>N^{1/8}$ we have

$$\begin{aligned}
&\frac{1}{\phi(dq)}\sum_{\chi\ (\mathrm{mod}\ dq)}\left|-e^{2\pi i(\delta/d-t\delta/d)}\sum_{\substack{\rho\in Z(\chi;T)\\ \rho\neq\beta_D}}t^{\rho-1}G_{a,q,d,\chi}\right|\\
\le&\frac{|G_{a,q,d,\chi}|}{\phi(d)\phi(q)}\left(\sum_{\chi\ (\mathrm{mod}\ dq)}\sum_{\substack{\rho\in Z(\chi;T)\\ \rho\neq\beta_D}}|t^{\rho-1}|\right)\\
\ll&\frac{q}{\phi(d)\phi(q)}(1-\beta_D)\log(dqT)\exp\left(-\frac{c_6}{8}\frac{\log N}{\log(dqT)}\right),
\end{aligned}$$

where we take the trivial bound $|G_{a,q,d,\chi}|\le q$ and use the inequality $\phi(dq)\ge\phi(d)\phi(q)$. It follows that

$$\begin{aligned}
&\left|\widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right)-\frac{e^{2\pi i\delta/d}G_{a,q,d,\chi'}}{\phi(d)\phi(q)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)e^{-2\pi i\delta t/d}dt\right|\\
&\qquad\le O\left(\frac{dqN}{\phi(d)\phi(q)}(1-\beta_D)\log(dqT)\exp\left(-\frac{c_6}{8}\frac{\log N}{\log(dqT)}\right)\right)\\
&\qquad\qquad+O\left((1+N|\delta|)\frac{dqN\log^2N}{T}\right).
\end{aligned}$$

By taking $q=1$ and $\delta=0$ in the estimates above, we can deduce that

$$\left|\widehat{F_{N,d}}(0)-\frac{1}{\phi(d)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)dt\right|\le\frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{4c_4}+O\left(\frac{dN\log^2N}{T}\right),\tag{2.7}$$

where we used the assumption $N\ge(DT)^{C_3}$ to obtain the first term on the right hand side; recall that c_4 is the constant involved in Definition 2.3.1. Thus, by the triangle

inequality and Lemma 2.3.7, we can conclude that

$$\left|\widehat{F_{N,d}}\left(\frac{a}{q}+\delta\right)\right| \le \frac{|\widehat{F_{N,d}}(0)|}{\phi(q)} + \frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{4c_4} + O\left(\frac{dqN}{\phi(d)\phi(q)}(1-\beta_D)\log(dqT)\exp\left(-\frac{c_6}{8}\frac{\log x}{\log(dqT)}\right)\right) + O\left((1+N|\delta|)\frac{dqN\log^2 N}{T}\right).$$

Therefore, we can deduce the first inequality we manage to prove the second one.

In order to prove the lower bound on $|\widehat{F_{N,d}}(0)|$, we need to bound the integral

$$\frac{1}{\phi(d)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)dt$$

from below. To do so we use the inequality $1-e^{-x} \ge x/(x+1)$ which holds[11] for all $x>0$. For $t \ge N^{1/8}$, one has $t \ge (DT)^{C_3/8} \ge (dqT)^{C_3/(8C_2)}$, and so

$$1-t^{\beta_D-1} \ge 1-(dqT)^{-C_3(1-\beta_D)/(8C_2)} \ge \frac{C_3(1-\beta_D)\log(dqT)}{8C_2+C_3(1-\beta_D)\log(dqT)}.$$

Since $1-\beta_D \le c_4/(C_2\log(DT)) \le c_4/\log(dqT)$ and C_3 is sufficiently large, we have

$$1-t^{\beta_D-1} \ge \frac{C_3(1-\beta_D)\log(dqT)}{8C_2+C_3c_4} \ge \frac{1-\beta_D}{2c_4}\log(dqT).$$

Thus,

$$\frac{1}{\phi(d)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)dt \ge \frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{2c_4} - O(N^{1/8}).$$

Therefore, by the triangle inequality in (2.7),

$$\left|\widehat{F_{N,d}}(0)\right| \ge \frac{1}{\phi(d)}\int_{N^{1/8}}^{dN}\left(1-t^{\beta_D-1}\right)dt - \frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{4c_4} - O\left(\frac{dN\log^2 N}{T}\right) \ge \frac{dN}{\phi(d)}\frac{(1-\beta_D)\log(dT)}{4c_4} - O\left(\frac{dN\log^2 N}{T}\right),$$

as claimed. □

2.4 Ruzsa–Sanders iteration argument with modifications

We prove Theorem 2.1.1 in this section. We shall follow the iteration scheme used by Ruzsa and Sanders and apply our estimate from the previous section.

[11] This can be verified by taking second order derivatives of e^x and $x+1$.

2.4.1 Modified main iteration lemma

In this subsection, we derive an analogue of the main iteration lemma given in Ruzsa and Sanders [47, Section 8]. Under certain restrictions on several input parameters, the iteration lemma allows one to find a denser subset, located on a sub-progression, given a set whose difference set does not contain shifted primes. This will be done by obtaining L^2-concentration and applying Proposition 0.2.5.

In order to show our modification in an obvious way, we choose to write the iteration lemma in a similar way to Ruzsa and Sanders [47, Lemma 8.1]. More specifically, we replace their notion of "N is small" (outcome (iii)), which is defined as $\log N \ll (\log D)^2$ for them, by $\log N \ll (\log \alpha^{-1} + \log\log\log D + 1)(\log D + \log\log N + 1)$. This strengthening leads to our improvement.

The overall strategy of our proof is similar as the proof of Ruzsa and Sanders [47, Lemma 8.1]. We follow the same steps as them to detect structure and to obtain minor arc bounds, and replace their major arc estimates by ours to improve the result.

Lemma 2.4.1. *There exist positive absolute constants*[12] $C_5, C_6, c_7, c_8, c_9, c_{10}$ *such that we can obtain the following result.*[13]

Let D be large enough so that $\log\log\log D$ is well-defined. Let N be a positive integer such that $N \leq \exp(D^{1/10})$ and $N > D^{C_5}$. Let $T = D^{C_2^2}$. Suppose that $A \subseteq [N]$ has density $\alpha > 0$. Assume one of the following:

(a) (D, T) is unexceptional and $1 \leq d \leq D$;

(b) (D, T) is exceptional and $d \leq D^{C_2}$ is a positive integer which is a multiple of d_D.

Then one of the following assertions holds:

- *(i) (Density increment) There exists a positive integer d' with $d' \leq c_7\alpha^{-3}$ and a progression P' with common difference d' and length $\geq (c_8\alpha/d\log N)^8 N$ such that $|A \cap P'| \geq \alpha(1 + c_9)|P'|$.*
- *(ii) (Structure) We have $\langle 1_A * 1_{-A}, F_{N',d}\rangle \geq \alpha^2 N|\widehat{F_{N',d}}(0)|/2$ where $N' \gg \alpha N$. In particular, there exist $a, a' \in A$ such that $(a - a')d = p - 1$ for some prime number p.*

[12] The constant C_5 depends on the constants c_4, c_5 and C_2, C_3 involved in Definition 2.3.1 and the major arc estimates in Section 4.

[13] We need to introduce an upper bound on N due to the factor $(\log N)^4$ in the minor arc estimate.

(iii) *(N is small) We have* $\log N \le C_6(\log \alpha^{-1} + \log\log\log D + 1)(\log D + \log\log N + 1)$.

(iv) *($d\alpha^{-1}$ is large) either (D,T) is unexceptional and $d\alpha^{-1} \ge c_{10}D^{c_{10}}$; or (D,T) is exceptional and $d\alpha^{-1} \ge c_{10}D^{1+c_{10}}$.*

Proof. Let $c > 0$ be an absolute constant to be optimised later. Then, either $N \ll c^{-1}\alpha^{-1}$ or $N' := \lfloor c\alpha N \rfloor \ge 1$. In the first case, we are in outcome (iii). From now on, we consider the situation when $N' \ge 1$. By Proposition 2.3.2 and the upper bound of β_D given in (0.4), we have

$$|\widehat{F_{N',d}}(0)| \gg \frac{N'}{d}. \tag{2.8}$$

Suppose first that

$$\langle 1_A * 1_{-A}, F_{N',d}\rangle \ge \frac{\alpha^2 N|\widehat{F_{N',d}}(0)|}{2}. \tag{2.9}$$

Notice that if $dn+1 = p^k$ for some integers $k \ge 2$ and $n \le N'$, then one must have $p \le \sqrt{dN'+1}$. Thus,

$$\begin{aligned}
&\big|\{n \le N' : dn+1 = p^k \text{ where } k \ge 2, k \in \mathbb{N}\}\big| \\
\ll &\log(dN') \sum_{p \le (dN'+1)^{1/2}} \frac{1}{\log p} \\
\le &(dN')^{1/4}\log(dN') + 4 \sum_{(dN')^{1/4} \le p \le (dN'+1)^{1/2}} 1 \\
\ll &(dN')^{1/2}.
\end{aligned}$$

Combine the bound above with the trivial upper bounds $1_A * 1_{-A}(n) \le \alpha N$ and $F_{N',d}(n) \ll \log(dN')$, we have

$$\sum_{\substack{n:dn+1=p^k \\ p \text{ prime}, k\ge 2}} 1_A * 1_{-A}(n)F_{N',d}(n) \ll \alpha N\sqrt{dN'}\log(dN').$$

By (2.8), we have $|\widehat{F_{N',d}}(0)|^{2/3} \gg \sqrt{dN'}\log(dN')$ unless we are in outcome (iii), and so

$$\sum_{\substack{n:dn+1=p^k \\ p \text{ prime}, k\ge 2}} 1_A * 1_{-A}(n)F_{N',d}(n) \ll \alpha^2 N|\widehat{F_{N',d}}(0)|^{2/3}.$$

By (2.8) and the definition of N', if $|\widehat{F_{N',d}}(0)| = O(1)$ then we are in outcome (iii). Thus, subtracting the inequality above from (2.9) yields

$$\sum_{n:dn+1\in\mathbb{P}} 1_A * 1_{-A}(n)F_{N',d}(n) \gg \alpha^2 N|\widehat{F_{N',d}}(0)|.$$

Since $1_A * 1_{-A}$ is supported on the difference set $A - A$, it follows that there exist $a, a' \in A$ such that $a - a' = (p-1)/d$ for some prime number p, and so we are in outcome (ii).

Otherwise, we have

$$\langle 1_A * 1_{-A}, F_{N',d} \rangle \leq \frac{\alpha^2 N |\widehat{F_{N',d}}(0)|}{2}. \tag{2.10}$$

Let I denote the interval $[N]$ and consider the inner product

$$\langle (1_A - \alpha 1_I) * (1_{-A} - \alpha 1_{-I}), F_{N',d} \rangle.$$

By direct computations, we have

$$\langle 1_A * 1_{-I}, F_{N',d} \rangle = \alpha N \widehat{F_{N',d}}(0) + O\left(|\widehat{F_{N',d}}(0)| N' \right),$$

$$\langle 1_I * 1_{-A}, F_{N',d} \rangle = \alpha N \widehat{F_{N',d}}(0) + O\left(|\widehat{F_{N',d}}(0)| N' \right),$$

and

$$\langle 1_I * 1_{-I}, F_{N',d} \rangle = N \widehat{F_{N',d}}(0) + O\left(|\widehat{F_{N',d}}(0)| N' \right).$$

Combining the bounds above with (2.10) yields

$$\langle (1_A - \alpha 1_I) * (1_{-A} - \alpha 1_{-I}), F_{N',d} \rangle \leq \alpha^2 N |\widehat{F_{N',d}}(0)| (-1/2 + O(c)).$$

Therefore, we can choose $c \gg 1$ which guarantees

$$|\langle (1_A - \alpha 1_I) * (1_{-A} - \alpha 1_{-I}), F_{N',d} \rangle| \gg \alpha^2 N |\widehat{F_{N',d}}(0)|.$$

It follows from Plancherel's theorem that

$$\int_{\mathbb{T}} \left| \left(\widehat{1_A} - \alpha \widehat{1_I} \right)(\theta) \right|^2 |\widehat{F_{N',d}}(\theta)| d\theta \gg \alpha^2 N |\widehat{F_{N',d}}(0)|. \tag{2.11}$$

Let c', c'' be two positive absolute constants to be chosen later. Let

$$Q' := \frac{d^4 \log^8 N'}{c'^2 \alpha^2},\ Q := \frac{N}{Q'} \text{ and } Q'' := c'' \alpha^{-3}.$$

If $Q' \geq Q''$ then let

$$\mathfrak{M}' := \bigcup_{Q' < q \leq Q} \mathfrak{M}_q^*,\ \mathfrak{M}'' := \bigcup_{Q'' < q \leq Q'} \mathfrak{M}_q^* \text{ and } \mathfrak{M}''' := \bigcup_{q \leq Q''} \mathfrak{M}_q^*,$$

where $\mathfrak{M}_q^*$ and $\mathfrak{M}_{a,q}$ (involved in the definition of $\mathfrak{M}_q^*$) are defined with respect to Q as in (0.1). If $Q' < Q''$ we define $\mathfrak{M}'' = \emptyset$ and $\mathfrak{M}', \mathfrak{M}'''$ as above. Notice that it follows from Dirichlet's pigeonhole principle that

$$\mathbb{T} = \mathfrak{M}' \cup \mathfrak{M}'' \cup \mathfrak{M}''',$$

and so

$$\int_{\theta\in\mathbb{T}} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)|d\theta \leq \int_{\theta\in\mathfrak{M}'} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)|d\theta$$
$$+ \int_{\theta\in\mathfrak{M}''} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)|d\theta + \int_{\theta\in\mathfrak{M}'''} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)|d\theta. \quad (2.12)$$

Unless $N' \leq d$ or $dN'^{-1/10} \geq c'\alpha$ (either of which implies outcome (iii)), we can apply the minor arc estimate given in Theorem 0.3.5 with our choices[14] of N', Q' and Q to see that

$$|\widehat{F_{N',d}}(\theta)| \ll \frac{N'}{d}\left(c'\alpha + dN'^{-1/10}\right) \ll \frac{c'\alpha N'}{d} \ll c'\alpha|\widehat{F_{N',d}}(0)| \quad (2.13)$$

for all $\theta \in \mathfrak{M}'$. Since

$$\int_{\mathbb{T}} \left|(\widehat{1_A} - \alpha\widehat{1_I})(\theta)\right|^2 d\theta \leq \alpha N$$

by Plancherel's theorem, it follows from Hölder's inequality and (2.13) that

$$\int_{\theta\in\mathfrak{M}'} \left|(\widehat{1_A} - \alpha\widehat{1_I})(\theta)\right|^2 |\widehat{F_{N',d}}(\theta)|d\theta \ll c'\alpha^2 N|\widehat{F_{N',d}}(0)|. \quad (2.14)$$

We employ our major arc estimates on $\mathfrak{M}''$ and $\mathfrak{M}'''$. By the relevant hypotheses, we can apply the first assertion of Proposition 2.3.2 unless $N' \leq (DT)^{C_3}$ or $dQ' > D$ (which would imply outcome (iii) or (iv), respectively) when (D, T) is unexceptional; or we can apply the second assertion of Proposition 2.3.2 unless $N' \leq (DT)^{C_3}$ or $dQ' > D^{C_2}$ (which would imply outcome (iii) or (iv), respectively) when (D, T) is exceptional. Thus, either we are in outcome (iii) or (iv), or it follows from the relevant major arc estimate that for $a/q + \delta \in \mathfrak{M}_{a,q}$ we have

$$\left|\widehat{F_{N',d}}\left(\frac{a}{q} + \delta\right)\right|$$
$$\leq \frac{2|\widehat{F_{N',d}}(0)|}{\phi(q)} + O\left(\frac{|\widehat{F_{N',d}}(0)|q\log(dqT)}{\phi(q)\log(dT)} \exp\left(-\frac{c_5 \log N'}{\log(D^{C_2}T)}\right)\right) + O\left(\frac{N'}{D^{2C_2}}\right), \quad (2.15)$$

where we use $\log N \leq D^{1/10}$ and our choice of T (in terms of D, and $C_2 \geq 10$) to bound the last error term in each inequality from Proposition 2.3.2.

By (2.8) and $dq \leq D^{C_2}$, the third term in (2.15) is always bounded by the first term. To deal with the second term, notice that

$$\frac{|\widehat{F_{N',d}}(0)|q\log(dqT)}{\phi(q)\log(dT)} \exp\left(-c_5\frac{\log N'}{\log(D^{C_2}T)}\right) \ll c''\alpha^3|\widehat{F_{N',d}}(0)| \quad (2.16)$$

[14]We also used $N \leq \exp(D^{1/10})$ in the inequality below.

(where the implicit constant comes from the big O notation) unless[15]

$$\begin{aligned} c_5 \frac{\log N'}{\log(D^{C_2}T)} &\leq \log 3\alpha^{-1} + \log(q/\phi(q)) + \log c''' \\ &\leq \log \alpha^{-1} + \log\log\log D + \log c''' \end{aligned}$$

which implies (iii) or (iv). Since

$$\alpha^3 \ll \min\left\{\alpha, \min_{q\leq Q''} \frac{1}{\phi(q)}\right\}, \tag{2.17}$$

it follows from (2.15), (2.16), (2.17) and our choice of Q'' that either we are in outcome (iii) or (iv), or we have

$$\sup_{\theta\in\mathfrak{M}''} |\widehat{F_{N',d}}(\theta)| \ll c''\alpha |\widehat{F_{N',d}}(0)| \tag{2.18}$$

and

$$\sup_{\theta\in\mathfrak{M}'''} |\widehat{F_{N',d}}(\theta)| \ll \frac{|\widehat{F_{N',d}}(0)|}{\phi(q)}. \tag{2.19}$$

By substituting (2.18) and applying Plancherel's theorem, we have

$$\int_{\theta\in\mathfrak{M}''} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)| d\theta \ll c''\alpha^2 N |\widehat{F_{N',d}}(0)|. \tag{2.20}$$

Therefore, combining the lower bound obtained in (2.11) and the upper bounds (2.14) and (2.20), and the inequality (2.12), we can choose some $c' \gg 1$ and $c'' \gg 1$ so that

$$\int_{\theta\in\mathfrak{M}'''} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)| d\theta \gg \alpha^2 N |\widehat{F_{N',d}}(0)|.$$

By the definition of $\mathfrak{M}'''$ and the triangle inequality, we have

$$\sum_{q\leq Q''} \int_{\theta\in\mathfrak{M}_q^*} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 |\widehat{F_{N',d}}(\theta)| d\theta \gg \alpha^2 N |\widehat{F_{N',d}}(0)|.$$

Combine this with (2.19), we can deduce that

$$\sum_{q\leq Q''} \frac{|\widehat{F_{N',d}}(0)|}{\phi(q)} \int_{\theta\in\mathfrak{M}_q^*} |(\widehat{1_A} - \alpha\widehat{1_I})(\theta)|^2 d\theta \gg \alpha^2 N |\widehat{F_{N',d}}(0)|.$$

Since $|\widehat{F_{N',d}}(0)| > 0$, we obtain the density increment outcome (i) by applying Proposition 0.2.5. □

[15]Here we used the well-known bound $q/\phi(q) \ll \log\log q$.

2.4.2 Proof of the main result

Using the lemma above, we can perform the iteration and deduce the main result. For technical reasons, we use the following lemma given in Ruzsa and Sanders [47, Lemma 9.1] to obtain a modified initial set when an exceptional zero presents.

Lemma 2.4.2. *Let $A \subseteq [N]$ have density $\alpha > 0$ and let $d \geq 1$. There exists an arithmetic progression P with common difference d and length $\geq c_{11}\alpha N/d$ such that $|A \cap P| \geq \alpha|P|/2$. Here c_{11} is a positive absolute constant.*

Proof of Theorem 2.1.1. Suppose N is sufficiently large. Let $C' \geq C_5$ be a large constant which depends on the previous constants $c_4, ..., c_{11}$ and $C_2, ..., C_5$ (but does not depend on N and α) and set

$$D := \exp\left(\frac{\log N}{C'(\log \alpha^{-1} + \log\log N + 1)}\right).$$

Let $T = D^{C_2^2}$ as in Lemma 2.4.1.

Firstly, let us choose the initial set where we begin the iteration. Since the initial hypotheses of Lemma 2.4.1 differ depending on whether (D, T) is exceptional or not, our choices here split into two cases accordingly.

(a) (D, T) is unexceptional. In this case we choose $A_1 = A$ and we begin the iteration with the given set A. We have

$$N_1 = N,\ \alpha_1 = \alpha \text{ and } d_1 = 1.$$

(b) (D, T) is exceptional. In this case, we apply Lemma 2.4.2 above with

$$d_1 = d_D.$$

Let P_0 be the arithmetic progression obtained by applying the lemma and let $A_0' = A \cap P_0$. Notice that the length of the progression P_0 satisfies

$$N_1 := |P_0| \geq c_{11}\alpha N/d_D.$$

Let A_1 be the affine transformation of A_0' so that it lies in the interval $[N_1]$. It follows that A_1 is dilated by a factor d_1^{-1} times a shifted subset of A, and $|A_1| = \alpha_1 N_1$ with $\alpha_1 \geq \alpha/2$.

Notice that both choices of N_1 above satisfy the size hypotheses $D^{C_5} \leq N_1 \leq \exp(D^{1/10})$ of Lemma 2.4.1 unless $\log \alpha^{-1} \gg \log N / \log\log N$, which itself suffices to show the main theorem. Thus, regardless of whether (D, T) is exceptional or not, we can apply Lemma 2.4.1 with A_1, N_1, α_1 and d_1 chosen above.

Let

$$k_0 := \min\left\{ \frac{2\log \alpha^{-1}}{\log(1+c_9)}, \frac{c' \log N}{\log^2 \alpha^{-1}}, c'\sqrt{\frac{\log N}{\log \alpha^{-1}}} \right\},$$

where $c' \gg 1$ is a positive absolute constant to be specified later.

Suppose for some $k < k_0$ we have obtained A_k, N_k, α_k and d_k such that $|A_k| = \alpha_k N_k$, A_k is dilated by a factor d_k^{-1} times a shifted subset of A, the inputs A_k, D, α_k, d_k satisfy the hypotheses of Lemma 2.4.1, and N_k, α_k, d_k satisfy the bounds

$$\alpha_k \geq (1+c_9)^{k-1}\alpha_1, \tag{2.21}$$

$$d_k \leq c_7^{k-1}\alpha_1^{-3(k-1)} d_1, \tag{2.22}$$

and

$$N_k \geq \frac{c_8^{8(k-1)}\alpha_1^{32k^2} N_1}{c_7^{8k^2} d_1^{8(k-1)} (\log N)^{8(k-1)}}. \tag{2.23}$$

We claim that for a suitably chosen c', we shall either obtain a contradiction from outcome (ii) or be in outcome (i) when we apply Lemma 2.4.1 with A_k, D, α_k, d_k. Moreover, if outcome (ii) does not occur at the kth step, we can obtain A_{k+1}, N_{k+1}, α_{k+1} and d_{k+1} such that $|A_{k+1}| = \alpha_{k+1}N_{k+1}$, A_{k+1} is dilated by a factor d_{k+1}^{-1} times a shifted subset of A, the inputs A_{k+1}, D, d_{k+1} satisfy the relevant hypotheses of Lemma 2.4.1 and the bounds

$$\alpha_{k+1} \geq (1+c_9)^k \alpha_1,\ d_{k+1} \leq c_7^k \alpha_1^{-3k} d_1 \text{ and } N_{k+1} \geq \frac{c_8^{8k}\alpha_1^{32(k+1)^2} N_1}{c_7^{8(k+1)^2} d_1^{8k} (\log N)^{8k}}. \tag{2.24}$$

By our induction hypothesis, we apply the iteration lemma to A_k, N_k, α_k, D and d_k, so that we will end up in one of the following outcomes. Here, the constants $c_5, ..., c_9$ are the constants appearing in Lemma 2.4.1.

(i) There exists a positive integer d_k' with $d_k' \leq c_7 \alpha_k^{-3}$ and a progression P_k with common difference d_k' and length $\geq (c_8\alpha_k/d_k \log N_k)^8 N_k$, such that the set $A_k' := A_k \cap P_k$ satisfies $|A_k'| \geq \alpha_k(1+c_9)|P_k|$.

(ii) There exist $a, a' \in A_k$ such that $(a-a')d_k = p-1$ for some prime p. Since A_k is dilated by a factor d_k^{-1} times a shifted subset of the original set A, it follows that $A - A$ contains a number which is one less than a prime.

(iii) $\log N_k \le C_6(\log D + \log\log N_k + 1)(\log \alpha_k^{-1} + \log\log\log D + 1)$.

(iv) (D,T) is unexceptional and $d_k\alpha_k^{-1} \ge c_{10}D^{c_{10}}$; or (D,T) is exceptional and $d_k\alpha_k^{-1} \ge c_{10}D^{1+c_{10}}$.

If outcome (ii) happens, we obtain a contradiction to the hypothesis of Theorem 2.1.1. Thus, we turn to other possible outcomes.

By taking logarithms of the bounds (2.21), (2.22) and (2.23) given in the induction hypothesis, we have

$$\log N_k \ge 32k^2 \log \alpha_1 + \log N_1 - 8k\log d_1 - 8k\log\log N - ck^2$$

where c is a positive absolute constant depending on c_7 and c_8, and

$$\log d_k \le 3(k-1)(\log \alpha_1^{-1} - \log 2) + \log d_1 + (k-1)\log c_7.$$

Suppose we are in outcome (iii), then we have

$$32k^2 \log \alpha_1 + \log N - 8k\log d_1 - 8k\log\log N - ck^2 \le \left(\frac{\log N}{C'\log \alpha^{-1}} + \log\log N\right)\log \alpha^{-1},$$

which implies either $\log \alpha^{-1} \gg \log N/\log\log N$ hence yielding the theorem, or $k \gg \sqrt{\log N/\log \alpha^{-1}}$ contradicting to the choice of k_0.

Suppose we are in outcome (iv). In this case, if (D,T) is unexceptional then

$$\log d_k + \log \alpha_1^{-1} - k\log(1+c_9) \ge c_{10}\frac{\log N}{C'\log \alpha^{-1}} + \log c_{10}.$$

Therefore, either $\log \alpha^{-1} \gg \log N/\log \alpha^{-1}$ which implies our theorem, or

$$3k\log \alpha^{-1} \gg \frac{\log N}{\log \alpha^{-1}},$$

which implies $k \gg \log N/(\log \alpha^{-1})^2$, contradicting to our choice of k_0.

If (D,T) is exceptional, then outcome (iv) means that

$$\log d_k + \log \alpha_1^{-1} - k\log(1+c_9) \ge (1+c_{10})\log D + \log c_{10}.$$

Since $d_1 \le D$, the above inequality holds only if

$$k\log \alpha^{-1} \gg \frac{\log N}{\log \alpha^{-1}},$$

which implies $k \gg \log N/(\log \alpha^{-1})^2$, contradicting to our choice of k_0.

Therefore, we cannot enter outcome (iii) or (iv) if we choose $c' \gg 1$ appearing in the definition of k_0 sufficiently small. By excluding all other possible outcomes, we conclude that we are in outcome (i). Let

$$N_{k+1} := |P_k| \text{ and } d_{k+1} := d_k d'_k. \tag{2.25}$$

We define A_{k+1} to be the affine transformation of A'_k so that it lies in the interval $[N_{k+1}]$. Let α_{k+1} be the density of the set A_{k+1} in $[N_{k+1}]$. Since A_k is dilated by a factor d_k^{-1} times a shifted subset of A, it follows that A_{k+1} is dilated by a factor d_{k+1}^{-1} times a shifted subset of A. The claimed bounds (2.24) follow from (2.25), the bounds given by iteration outcome (i) and the induction hypothesis (2.21), (2.22) and (2.23). To check that $A_{k+1}, D, \alpha_{k+1}, d_{k+1}$ satisfy the hypotheses of Lemma 2.4.1, notice first that if (D,T) is exceptional, then $d_D \mid d_{k+1}$ since $d_k \mid d_{k+1}$. Thus, we only need to check the size bounds

$$N_{k+1} \geq D^{C_5}$$

and

$$d_{k+1} \leq D \text{ if } (D,T) \text{ is unexceptional; } d_{k+1} \leq D^{C_2} \text{ if } (D,T) \text{ is exceptional.}$$

hold for all $k \leq k_0$.

Firstly, unless $k \gg \log N/(\log \alpha^{-1})^2$ or $k \gg \sqrt{\log N/\log \alpha^{-1}}$, it follows from (2.24) that

$$\begin{aligned}\log N_{k+1} \geq 32(k+1)^2 \log \alpha_1 + \log N_1 - 8(k+1)\log d_1 - 8(k+1)\log\log N - c(k+1)^2 \\ \geq \frac{C_5 \log N}{C'(\log \alpha^{-1} + \log\log N + 1)}\end{aligned}$$

where c is a positive absolute constant depending only on c_7 and c_8. On the other hand, unless $k \gg \sqrt{\log N/\log \alpha^{-1}}$, it follows from (2.24) that

$$c_7^k \alpha_1^{-3k} \leq D.$$

The second condition follows from a combination of the bound above and the definition of d_1.

Since every set has density at most 1, we will no longer be in outcome (i) after at most $2\log \alpha^{-1}/\log(1+c_9)$ steps, if

$$c' \min\left\{\frac{\log N}{\log^2 \alpha^{-1}}, \sqrt{\frac{\log N}{\log \alpha^{-1}}}\right\} \geq \frac{2\log \alpha^{-1}}{\log(1+c_9)},$$

we arrive at a contradiction. It follows that

$$\log \alpha^{-1} \geq c(\log N)^{1/3}$$

for some absolute constant $c > 0$. This completes the proof of Theorem 2.1.1.

□

Chapter 3

Iterated sumset $nA - nA$ without shifted primes

3.1 Introduction

In the previous chapter, we proved an upper bound on the cardinality of $A \subseteq [N]$ whose difference set $A - A$ does not contain any number which is one less than a prime. The goal of this chapter is to study $A \subseteq [N]$ whose iterated sumset $nA - nA$ avoids $\mathbb{P} - 1$. There are two aspects which we are primarily interested in — the upper bound on the size of such sets, and the complementary question asking for such examples.

It is well-known that $2A-2A$ has more additive structure compared to the original set A, see Sanders [48]. Heuristically, one expects the set $nA - nA$ to be denser as n becomes larger. In the second section of this chapter, we shall give a short proof of the following result.

Theorem 3.1.1. *Let l and k be positive integers. There exist constants $c > 0$ and N_0 such that the following holds for all $N \geq N_0$. Let A be a subset of $[N]$ such that $2^k A - 2^k A$ does not contain any number which can be written as $(p-1)^l$ for some prime number p. If $k = o(\log N)$, then $|A| = O(N \exp(-ck))$.*

The theorem above provides a stronger upper bound than Theorem 2.1.1 when one looks at $nA - nA$ for some $n = \exp(\Omega((\log N)^{1/3}))$. As n becomes larger, the bound above becomes closer to the one proposed in Question 2.1.3 due to Ruzsa [45]. Unfortunately, we still cannot show a power saving unless $n = \exp(\Omega(\log N))$, which itself appears to reflect only a rather basic property of sumsets with large doubling constants.

The proof is essentially an easy application of the Bogolyubov–Ruzsa lemma due to Sanders [48], and a result due to Croot, Lyall and Rice [8] concerning generalised arithmetic progressions which avoid $h(\mathbb{P})$ where h is an integer polynomial satisfying certain conditions. To illustrate the difference between the approach in the previous chapter and the one using the theorems of Sanders and Croot–Lyall–Rice, we show that the latter can be used to prove a nontrivial upper bound of $|A|$ when $n = 2$. At the end of this section, we also discuss how a conjectured bound concerning the Croot–Lyall–Rice result can be used to deduce an improved upper bound[1] of $|A|$ such that $A \subseteq [N]$ and $2A - 2A$ avoids lth powers of shifted primes.

The final section of this chapter provides an example of $A \subseteq [N]$ such that $nA - nA$ does indeed avoid $\mathbb{P} - 1$. This is a generalisation of Ruzsa's example stated in Theorem 2.1.2, and in the same manner as Ruzsa, the constructed set $nA - nA$ also avoids $\mathbb{P} + 1$. More specifically, we shall derive the following.

Theorem 3.1.2. *Let $n \in \mathbb{N}$ be fixed. For sufficiently large N, there exists a subset $A \subseteq [N]$ whose cardinality is at least $\exp(((\log 2)/(2n) + o(1)) \log N / \log\log N)$, such that the set $nA - nA$ avoids $\mathbb{P} - 1$.*

The proof follows the same ideas as Ruzsa. In fact, Ruzsa's example is exactly what one gets by taking $n = 1$ in Section 3.3. We shall define a series of congruence equations and use their solutions to define the set A. Using the congruence relationships, we conclude that every nonzero element of $nA - nA + 1$ must be a composite number. The reason why the cardinality of A becomes smaller as n becomes larger is that our construction requires more congruence equations for larger n.

3.2 Proof of the upper bound

We need the following two results in our proof. The first one is a special case of the Bogolyubov–Ruzsa lemma due to Sanders [48, Theorem 1.1] when the underlying group is $\mathbb{Z}$.

Theorem 3.2.1. *Suppose $A \subseteq \mathbb{Z}$ is finite. Suppose also that $|A + A| \leq K|A|$. Then $2A-2A$ contains a proper symmetric d-dimensional generalised arithmetic progression M of size at least $\exp(-s(K))|A|$ with $d = d(K) \leq C \log^C 2K$ and $s = s(K) \leq C \log^C 2K$. Here C is an absolute constant.*

[1]The bound that follows from the Croot–Lyall–Rice conjecture for $A \subseteq [N]$ such that $2A - 2A$ avoids lth powers ($l \geq 2$) of shifted primes is stronger than what we know [1, 3] about the cardinality of $A \subseteq [N]$ such that $A - A$ avoids lth powers of the integers.

The second required result, due to Croot, Lyall and Rice [8, Theorem 3], concerns the size of a proper symmetric generalised arithmetic progression which intersects the values of a $\mathbb{P}$-intersective polynomial at prime numbers.

Theorem 3.2.2. *Let $h \in \mathbb{Z}[x]$ be a polynomial of degree l such that for every $q \in \mathbb{N}$ there exists some $r \in \mathbb{Z}$ such that $(r, q) = 1$ and $q \mid h(r)$. Suppose $M \subseteq [-N, N]$ is a proper symmetric generalised arithmetic progression of dimension $d \leq c(\log\log N)/l$ and $M \cap h(\mathbb{P}) \subseteq \{0\}$, then*

$$|M| \ll_h N^{1-c^d 5^{-ld}}.$$

Here, $c > 0$ is a small absolute constant.

Proof of Theorem 3.1.1. Throughout the proof, we fix $K = 4$. By Theorem 3.2.1, for each $1 \leq i \leq k$, we have

(1) either $2^i A - 2^i A$ contains a proper symmetric generalised arithmetic progression M_i which has size at least $\exp(-C\log^C 4K)|2^{i-1}A|$ and dimension at most $C\log^C 4K$, and so it follows[2] from Theorem 3.2.2 with $h(p) = (p-1)^l$ that

$$|M_i| \ll (2^i N)^{1-(c/5^l)^{C\log^C 4K}};$$

(2) or $|2^i A| \geq 2K|2^{i-1}A|$.

Let k_0 be the smallest integer such that outcome (1) is true, if such k_0 exists and satisfies $k_0 \leq k$. Then, we have

$$|M_{k_0}| \ll 2^{k_0} N^{1-(c/5^l)^{C\log^C 4K}}.$$

Since the second outcome holds for all $i < k_0$, it follows that

$$\begin{aligned} |A| \ll (2K)^{-k_0}|2^{k_0-1}A| &\ll (2K)^{-k_0}\exp(C\log^C 4K)|M_{k_0}| \\ &\ll K^{-k_0}\exp(C\log^C 4K)N^{1-(c/5^l)^{C\log^C 4K}}. \end{aligned} \quad (3.1)$$

Otherwise, the first outcome never occurs and so

$$|A| \leq (2K)^{-k}|2^k A| \leq K^{-k}N. \quad (3.2)$$

We conclude our proof of Theorem 3.1.1 by combining (3.1) and (3.2). □

[2]We assume N is sufficiently large.

We remark that, using the lemmas cited in this section in a rather straightforward manner without the assistance of Theorem 2.1.1, one can also obtain a nontrivial upper bound on the cardinality of set $A \subseteq [N]$ such that $2A - 2A$ avoids lth powers of $\mathbb{P} - 1$.

Proposition 3.2.3. *Let l be a fixed positive integer. There exists $\delta > 0$ such that the following holds for every positive integer N that is sufficiently large. For any $A \subseteq [N]$ such that $2A - 2A$ avoids numbers of the form $(p-1)^l$ for some prime p, we have*

(i) ($|2A - 2A|$ is small) either $|2A - 2A| \ll N \exp(-(\log N)^{1-\delta})$,

(ii) (A has a large doubling constant) or $|A + A| \geq \exp(\delta(\log\log N)^{1/C})|A|/2$.

Here, C is the constant from Theorem 3.2.1. The implicit constant in (i) follows from Theorem 3.2.2, which does not depend on δ, N or the set A.

In particular, if $2A - 2A$ avoids numbers of the form $(p-1)^l$ where p is some prime number, then $|A| = O(N \exp(-\delta(\log\log N)^{1/C}))$.

Proof. Throughout the proof, the constant C is the same one as what appears in Theorem 3.2.1, and c is the constant from Theorem 3.2.2. Let δ be a constant such that $\delta \leq (c/lC)^{1/C}$ and $\delta \leq (2C\log(5^l/c))^{-1/(C-1)}$.

Suppose the second outcome does not occur, which means that the doubling constant is at most $\exp(\delta(\log\log N)^{1/C})/2$. Then, by Theorem 3.2.1, $2A - 2A$ contains a proper symmetric generalised arithmetic progression M of dimension at most $C\delta^C \log\log N$ and size at least $\exp(-C\delta^C \log\log N)|A|$.

Let $h(p) = (p-1)^l$. It is easy to check that h satisfies all the relevant hypotheses concerning h listed in Theorem 3.2.2. Since $C\delta^C \log\log N \leq c(\log\log N)/l$, one can apply Theorem 3.2.2 to obtain

$$|M| \ll N^{1-c^{C\delta^C \log\log N} 5^{-lC\delta^C \log\log N}}.$$

Since $\delta \leq (2C\log(5^l/c))^{-1/(C-1)}$, we have

$$\begin{aligned}\log\log\left(N^{c^{C\delta^C \log\log N} 5^{-lC\delta^C \log\log N}}\right) =& C\delta^C \log(c/5^l)\log\log N + \log\log N \\ \geq& (1-\delta/2)\log\log N,\end{aligned}$$

and so

$$|M| \ll N^{1-c^{C\delta^C \log\log N} 5^{-lC\delta^C \log\log N}} \ll N\exp(-(\log N)^{1-\delta/2}).$$

Therefore,

$$|A| \leq \exp(C\delta^C \log\log N)|M| \leq N\exp(-(\log N)^{1-\delta}).$$

□

Since Theorem 3.2.2 is applicable only when the dimension of the generalised arithmetic progression $M \subseteq [N]$ is of size $O(\log\log N)$, one cannot replace the doubling constant used in Proposition 3.2.3 by anything substantially larger (notice that a generalised arithmetic progression of dimension k roughly has doubling constant 2^k, so the dimension bound from the Bogolyubov–Ruzsa theorem cannot be better than $O(\log K)$). On the other hand, it is conjectured [8, Section 5] that the expected bound for Theorem 3.2.2 should be of the form

$$|M| \ll_h N^{1-c^l/d}. \tag{3.3}$$

It is also observed that such a bound holds for dimensions up to about $d = \sqrt{\log N}$ in specific cases. We remark that, if one can indeed prove the bound (3.3) for the extended dimension range as mentioned above, then one can immediately improve the bound of $|A|$ from Proposition 3.2.3 to $|A| = O(N\exp(-(\log N)^\delta))$ where $\delta > 0$ is a small absolute constant. Although this is still weaker than the bound from Theorem 2.1.1 when $l = 1$, it is stronger than the existing upper bounds [3, 1] on the cardinality of $A \subseteq [N]$ whose difference set avoids lth powers when $l \geq 2$.

3.3 A generalisation of Ruzsa's example

In this section, we prove Theorem 3.1.2 stated at the end of the first section, which claims that for given positive integer n and sufficiently large N, there exists a subset $A \subseteq [N]$ whose cardinality is at least $\exp(((\log 2)/2^n + o(1))\log N/\log\log N)$, such that the set $nA - nA$ avoids $\mathbb{P} - 1$. We remark that when $n = 1$, this is exactly the example of Ruzsa [46].

Proof of Theorem 3.1.2. Let $r_1 = 1$ and let p_i denote the ith smallest prime number.

For each $j \geq 2$, we define r_j to be the smallest integer satisfying the following conditions:

$$r_j > 2nr_{j-1} + 1, \;\; r_j > p_{2nj+2n} + 1, \tag{3.4}$$

$$r_j \equiv 0 \pmod{p_1 p_2 \cdots p_{2nj-2n}}, \tag{3.5}$$

and

$$r_j \equiv -\frac{1}{a} \pmod{p_{2nj-2n+(2a-1)}}, \;\; r_j \equiv \frac{1}{a} \pmod{p_{2nj-2n+2a}} \tag{3.6}$$

for each $1 \leq a \leq n$. Since the congruence equations are defined with respect to distinct primes, by the Chinese remainder theorem, the size of r_j is bounded by

$$p_1 \cdots p_{2nj-2n} \leq r_j \leq 2nr_{j-1} + p_{2nj+2n} + p_1 \cdots p_{2nj}.$$

Thus, by the prime number theorem, for all large j one has

$$r_j = \exp\left((2n + o(1))j \log j\right).$$

Let k be the number of elements of the sequence $\{r_j\}$ which are smaller than $N/2$. By our estimate of r_j, we have $k = (1/(2n) + o(1))(\log N)/\log\log N$. Let

$$A := \left\{ \sum_{i=1}^{k} c_i r_i : c_i = 0 \text{ or } 1 \right\}.$$

It follows that

$$|A| = 2^k = \exp\left(\left(\frac{\log 2}{2n} + o(1)\right)\frac{\log N}{\log\log N}\right).$$

Let $a \in nA - nA$. We claim that either $a = 0$, or $a + 1$ is composite.

Let

$$x_i = \sum_{j=1}^{k} c_j^i r_j, i = 1, ..., 2n.$$

It suffices to show that $x_1 + \cdots + x_n - x_{n+1} - \cdots - x_{2n} + 1$ is not a prime, where without loss of generality we assume that $x_1 \geq \cdots \geq x_n$ and $x_{n+1} \geq \cdots \geq x_{2n}$.

Let $j \leq k$ be an integer. We say that the coefficients $c_j^1, ..., c_j^{2n}$ are *identical* if

$$\sum_{i=1}^{n} c_j^i = \sum_{i=n+1}^{2n} c_j^i.$$

Notice that if the coefficients are identical for all $j \leq k$, then $x_1 + \cdots x_n - x_{n+1} - \cdots - x_{2n} = 0$ and our claim is clearly true. Therefore, in the rest of the proof, we assume that there exists $j \leq k$ such that the coefficients are not identical.

Let m be the biggest index such that $c_m^1, ..., c_m^{2n}$ are not identical. We claim that

$$\left| \sum_{i=1}^{n} x_i - \sum_{i=n+1}^{2n} x_i \right| > p_{2nm}. \tag{3.7}$$

This essentially follows from (3.4). More specifically, by the triangle inequality

$$\begin{aligned}\left| \sum_{i=1}^{n} x_i - \sum_{i=n+1}^{2n} x_i \right| \geq & r_m - \sum_{i=1}^{m-1} n r_i \\ \geq & 2n r_{m-1} - n(1 + 1/(2n) + 1/(2n)^2 + ...) r_{m-1} > (n-1) r_{m-1} \\ > & p_{2nm}.\end{aligned}$$

Let l be the smallest index such that $c_l^1, ..., c_l^{2n}$ are not identical. Since $c_l^i = 0$ or 1 for each $1 \leq i \leq 2n$, it follows that

$$\sum_{i=1}^{n} c_l^i - \sum_{i=n+1}^{2n} c_l^i \in \{-n, ..., -1, 1, ..., n\}.$$

Therefore, by (3.6), there exists some $0 \leq a \leq 2n-1$ such that

$$\left(\sum_{i=1}^{n} c_l^i - \sum_{i=n+1}^{2n} c_l^i \right) r_l + 1 \equiv 0 \pmod{p_{2nl-a}}. \tag{3.8}$$

On the other hand, by (3.5), we have

$$\sum_{i=1}^{n} x_i - \sum_{i=n+1}^{2n} x_i - \left(\sum_{i=1}^{n} c_l^i - \sum_{i=n+1}^{2n} c_l^i \right) r_l \equiv 0 \pmod{p_i} \tag{3.9}$$

for all $i \leq 2nl$. Hence, it follows from (3.8) and (3.9) that

$$x_1 + \cdots + x_n - x_{n+1} - \cdots - x_{2n} + 1 \equiv 0 \pmod{p_{2nl-a}}.$$

By (3.7), $|x_1 + \cdots + x_n - x_{n+1} - \cdots - x_{2n}| > p_{2nm} \geq p_{2nl-a}$, and so it cannot be p_{2nl-a} itself. This completes our proof. □

Chapter 4

Partition regularity over shifted primes, I

4.1 Introduction

Schur [52] proved that the equation $x + y = z$ is partition regular[1] over the set of the natural numbers in his celebrated paper published over a hundred years ago. More specifically, let $r(k)$ be the smallest positive integer such that every k-colouring of the set $[r(k)]$ contains a monochromatic solution to $x + y = z$ where $x, y, z \in [r(k)]$. It follows from Schur's paper that the following upper bound of $r(k)$ holds for all k.

Theorem 4.1.1. *One has* $r(k) \leq \exp(O(k \log k))$.

The exact bound from Schur's paper has been improved slightly (see Irving [26]), nevertheless there is not any known upper bound of size $\exp(o(k \log k))$. It is closely related to the Ramsey number $R_k(3)$, which is defined to be the smallest integer such that any k-colouring of the edges of the complete graph $K_{R_k(3)}$ contains a monochromatic triangle.

Schur's paper also provides the following lower bound on $r(k)$.

Theorem 4.1.2. *One has* $r(k) \geq \exp(\Omega(k))$.

The exact value of $r(k)$ (for small k) has been extensively investigated by both mathematicians and computer scientists, and each of these improved bound leads to an improvement of the implicit constant in Theorem 4.1.2, see, for instance, Exoo [11], Heule [23] and Fredricksen–Sweet [12].

[1]Let S be a subset of the integers. We say that an equation is *partition regular* over S if for any finite colouring of S, there exists a non-trivial solution to the equation where all the variables are of the same colour. A solution $(x_1, ..., x_n)$ is called *non-trivial* if $x_i \neq x_j$ for any $i \neq j$.

Li and Pan [30] showed that the equation $x + y = z$ is also partition regular over $\mathbb{P} - 1$, which is a sparse subset of the integers. The goals of this chapter are to introduce their work concerning the following quantitative version of this problem, and to study the key ingredients of the Li–Pan proof. Let k, N be positive integers. Let $r_p(k)$ be the smallest integer such that for any $N \geq r_p(k)$, any k-colouring of $\mathbb{P} \cap [N]$ contains a monochromatic solution of

$$p_1 - p_2 = p_3 - 1.$$

Although it is not explicitly stated in the paper of Li and Pan, one can deduce the following quantitative bound on $r_p(k)$ by tracing the quantitative relationships involved in their work.

Theorem 4.1.3. *One has* $r_p(k) = O(\exp(\exp(\exp(\exp(k^{O(1)})))))$.

Li and Pan used a somewhat similar strategy as Green's [18] proof of Roth's theorem in the primes[2] which can be summarised as follows. Firstly, one needs the counting result in the setting of the integers, which are essentially quantitative versions of Schur's theorem and Roth's theorem, respectively. Secondly, one aims to prove the corresponding counting result in the primes. This is done by the so called transference principle, which typically involves two parts — an analytic framework which transfers the result in the setting of the integers to suitably weighted functions, and finding an appropriate way to weight the integers in a way that resembles the primes[3].

We remark that Lê [28] proved a more general partition regularity result for linear equations over shifted primes, showing that any system of linear equations which is partition regular over $\mathbb{N}$ is also partition regular over $\mathbb{P} - 1$. The author uses ideas from the paper of Green and Tao [20] concerning more general linear equations in the primes instead of Green's work on 3-term arithmetic progressions, together with Deuber's work [9] on partition regularity.

In the next chapter, we shall prove an improved bound on $r_p(k)$ using the work conducted in Chapter 2. We remark that, due to the nature of the difference set

[2]We remark that a stronger result for Roth's theorem in the primes follows directly from Bloom and Sisask's bound on Roth's theorem [4]. Nevertheless, the transference principle developed in Green's work has been applied to many problems in additive combinatorics concerning sparse subsets of the integers.

[3]Precisely speaking, the weight used in the work of Green and Li–Pan resembles a subset of primes which is contained in an arithmetic progression. One needs to choose the common difference of the progression carefully to avoid the Fourier bias caused by small primes.

problem and the Fourier analytic method there which no longer works for longer patterns, this approach cannot be used to study the general partition regularity problem of Lê [28].

A well-known open problem of Erdős and Graham [14] asks whether the equation $x + y = z$ is partition regular over the squares, another sparse subset of the integers which has attracted a lot of interest. Unfortunately, due to the density of the squares in the integers, one cannot use the transference principle to prove an analogue result in this setting.

There has been a continuous interest in establishing certain combinatorial results in the random setting (see, for instance, the survey of Conlon [6]). By the work of Graham, Rödl and Ruciński [15], for almost all subset $A \subseteq [N]$ with $|A| \geq \omega(N^{1/2})$, every 2-colouring of A contains a monochromatic solution to $x+y = z$. This result is generalised to k-colouring for general k by Conlon and Gowers [7], and by Friedgut, Rödl and Schacht [13]. It is interesting to ask whether one can develop a random version of the result of Li and Pan.

This chapter is organised as follows. In Section 4.2, we develop a quantitative variant of Schur's theorem, which is a generalisation of Schur's theorem concerning functions that look like a partition. This is the result that will be further generalised using the transference principle in the rest of chapter, which ultimately leads to the proof of Theorem 4.1.3. In Section 4.3, we focus on the analytic aspect — we shall prove an L^∞ transference principle which does not involve any properties specific to the primes. Section 4.4 is a brief summary of the key estimates we shall need for the primes, which follows from an application of sieve theory. In the final section, we apply the results from the previous sections to prove the bound stated in Theorem 4.1.3.

4.2 A quantitative variant of Schur's theorem

For any functions f_1, f_2, f_3 defined on the integers, let

$$T(f_1, f_2, f_3) := \sum_{x,y\in\mathbb{Z}} f_1(x)f_2(y)f_3(x+y).$$

We say $\{\mathcal{P}_i\}_{i=1}^k$ is *a k-partition of the set S* if $\cup_{i=1}^k \mathcal{P}_i = S$ and $\mathcal{P}_i \cap \mathcal{P}_j = \emptyset$ whenever $i \neq j$. For any partition $\{\mathcal{P}_i\}_{i=1}^k$ of $\mathbb{Z}$, we define $f|_{\mathcal{P}_i}(x) := f(x)1_{\mathcal{P}_i}(x)$.

We need the following quantitative bound on the number of monochromatic solutions, which follows from counting the number of monochromatic triangles of any k-edge-colouring of the complete graph K_N. See Li and Pan [30, Lemma 2.1 and Lemma 2.2] for a proof.

Proposition 4.2.1. *Let M, k be positive integers. There exist positive absolute constants C_1, C_2 such that for any $M \geq \exp(k^{C_1})$ and any k-partition $\mathcal{P}_1, ..., \mathcal{P}_k$ of the integers, we have*

$$\sum_{i=1}^{k} T(1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i}) \geq M^2 \exp(-k^{C_2}).$$

The proposition below is a variant of Corollary 2.1 of Li and Pan.

Proposition 4.2.2. *Let M, k be positive integers and C_1, C_2 be the constants from Proposition 4.2.1. There exists a constant $\beta_k = \Omega(\exp(-k^{O(1)}))$ such that for any $c > 0$ and any k-partition $\mathcal{P}_1, ..., \mathcal{P}_k$ of the integers, the following result holds for all $M \geq \exp(k^{C_1})$. For any $f : \mathbb{Z} \to \mathbb{R}$ supported on $[1, M]$ such that $\sum_{x=1}^{M} f(x) \geq cM(1 - \beta_k)$ and $0 \leq f \leq c(1 + \beta_k)$, one has*

$$\sum_{i=1}^{k} T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}) \gg \beta_k c^3 M^2.$$

Here, the implicit constant does not depend on the constant c, the parameters M, k, or the function f.

Proof. We write $\sum_{i=1}^{k} T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i})$ as the sum of $\sum_{i=1}^{k} T(1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i})$ and $8k$ other terms. Notice that

$$\begin{aligned}
&\left|\sum_{i=1}^{k} T(f|_{\mathcal{P}_i} - c1_{[M]}|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i})\right| \\
\leq &\left|\sum_{i=1}^{k} T((f - c\beta_k - c1_{[M]})|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i})\right| + \left|\sum_{i=1}^{k} T(c\beta_k|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i}, c1_{[M]}|_{\mathcal{P}_i})\right| \\
\leq &c^2 \sum_{i=1}^{k} \sum_{x=1}^{M} (c1_{[M]} - f + c\beta_k)|_{\mathcal{P}_i}(x) + c^3 M\beta_k \\
\leq &3c^3 M\beta_k
\end{aligned}$$

since $f - c\beta_k \leq c1_{[M]}$ pointwisely. One may bound the other terms in a similar manner. Therefore, by Proposition 4.2.1 and the triangle inequality,

$$\begin{aligned}
\sum_{i=1}^{k} T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}) &\geq \sum_{i=1}^{k} T(1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i}, 1_{[M]}|_{\mathcal{P}_i}) - 24c^3 M\beta_k \\
&\geq c^3 M^2 \exp(-k^{C_2}) - 24c^3 M\beta_k.
\end{aligned}$$

Thus one may choose $\beta_k = \exp(-k^{C_2})/30$ to prove our assertion. □

The next lemma is essentially Li and Pan [30, pp. 124-126], which says that a quantitative version of Schur's theorem holds for any "dense" collection of functions. The pointwise upper bound hypothesis in the lemma is purposely designed for the L^∞ transference principle to be introduced in the next section.

Proposition 4.2.3. *Let M, k be positive integers and let C_1 be the constant from Proposition 4.2.1. Suppose $M \geq \exp(k^{C_1})$. Then there exist constants*[4] *$\beta'_k, \beta''_k = \Omega(\exp(-k^{O(1)}))$ such that for any $c > 0$ and any nonnegative functions $f_1, ..., f_k : \mathbb{Z} \to \mathbb{R}$ supported on $[-\beta'_k M, M + \beta'_k M]$ satisfying*

1. *$\sum_{i=1}^k \sum_{x \in \mathbb{Z}} f_i(x) \geq cM(1 - \beta'_k)$,*
2. *and for any $x \in \mathbb{Z}$, $\sum_{i=1}^k f_i(x) \leq c(1 + \beta'_k)$,*

we have

$$\sum_{i=1}^k T(f_i, f_i, f_i) \gg \beta''_k c^3 M^2.$$

Here, the implicit constant does not depend on the constant c, the parameters M, k, or the function f.

Proof. Let

$$X := \{x \in \mathbb{Z} : \sum_{i=1}^k f_i(x) \geq c\beta'_k\}.$$

By our assumptions and the definition of X,

$$c(1 + \beta'_k)|X| + c\beta'_k(M + 2\beta'_k M - |X|) \geq \sum_{x \in \mathbb{Z}} \sum_{i=1}^k f_i(x) \geq cM(1 - \beta'_k),$$

and so

$$|X| \geq M(1 - 2\beta'_k) - 2M\beta'^2_k \geq M(1 - 3\beta'_k).$$

Let $X' := X \cap [M]$. It follows that

$$|X'| \geq M(1 - 5\beta'_k).$$

Let $\mathcal{P}_1, ..., \mathcal{P}_k$ be a k-partition such that for each $n \in X'$, $n \in \mathcal{P}_i$ if and only if i is the smallest integer such that $f_i(n) = \max_{1 \leq j \leq k} f_j(n)$. For each $n \in X'$, we have

$$\max_{1 \leq j \leq k} f_j(n) \geq \frac{c\beta'_k}{k}.$$

[4] We assume $\beta'_k, \beta''_k < 1/4$, as this is the situation relevant to us.

Combine this with $f_i \geq 0$ yields

$$\sum_{i=1}^{k} T(f_i, f_i, f_i) \geq \sum_{i=1}^{k} T\left(f'|_{\mathcal{P}_i}, f'|_{\mathcal{P}_i}, f'|_{\mathcal{P}_i}\right),$$

where

$$f'(n) := \frac{c\beta'_k}{k} 1_{[X']}|_{\mathcal{P}_i}(n).$$

Notice that f' is supported on $[M]$, $\sum_{n\in\mathbb{Z}} f'(n) \geq |X'|c\beta'_k/k \geq c\beta'_k M(1-5\beta'_k)/k$, and $0 \leq f' \leq c\beta'_k/k$. Therefore, by taking $\beta'_k = \beta_k/5$, one can apply Proposition 4.2.2 to conclude that

$$\sum_{i=1}^{k} T(f_i, f_i, f_i) \geq \sum_{i=1}^{k} T\left(f'_{\mathcal{P}_i}, f'_{\mathcal{P}_i}, f'_{\mathcal{P}_i}\right) \gg c^3 k^{-3} \beta'^3_k \beta_k M^2.$$

Choose $\beta''_k = \beta'^3_k k^{-3} \beta_k$, which is of size $\Omega(\exp(-k^{O(1)}))$, completes the proof.

□

4.3 A Fourier-analytic transference principle

The ultimate goal of the transference principle, originally introduced by Green [17], is to transfer certain additive combinatorics results from (dense subsets of) integers to suitable pseudorandom sets (weighting the integers with measures satisfying suitable conditions). In this section, we present a general framework of applying such techniques to count monochromatic solutions of $x + y = z$ in sparse sets. This approach is used in the proof of Li and Pan [30], where the authors focus on the specific case of shifted primes. We study the general Fourier analytical properties needed to achieve this without specifically looking at shifted primes, which will be used in the upcoming section.

Definition 4.3.1 (majorant)**.** Let M be a positive integer and $S \subseteq [M]$. We say $\nu : \mathbb{Z} \to \mathbb{C}$ is *a majorant on* S if ν is a nonnegative function supported on S such that $\sum_{x\in\mathbb{Z}} \nu(x) = M(1 + o(1))$.

Definition 4.3.2. Let $\nu : \mathbb{Z} \to \mathbb{R}$ be a nonnegative function supported on $[M]$ such that $\sum_{x\in\mathbb{Z}} \nu(x) = M(1 + o(1))$. We say ν is *(Schur) pseudorandom* if for any positive integer k, there exists a constant $c_k > 0$ such that the inequality

$$\sum_{i=1}^{k} T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}) \geq c_k M^2 + o_{M\to\infty}(1)$$

holds for any k-partition $\mathcal{P}_1, ..., \mathcal{P}_k$ of the integers and any nonnegative function f satisfying the following conditions:

1. $f(n) \leq \nu(n)$ for all n,

2. $\sum_{x\in\mathbb{Z}} f(x) \geq (1-\beta_k')M$.

Here, β_k' is the constant from Proposition 4.2.3.

To obtain a lower bound on $\sum_{i=1}^k T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i})$, we would like to approximate $\{f|_{\mathcal{P}_i}\}_{i=1}^k$ by functions $\{g_i\}_{i=1}^k$, and show that $\sum_{i=1}^k T(g_i, g_i, g_i)$ is large using Proposition 4.2.3.

By Parseval's theorem, for any $f_1, f_2, f_3 \in \ell^1(\mathbb{Z}) \cap \ell^2(\mathbb{Z})$, we have

$$T(f_1, f_2, f_3) = \int_{\mathbb{T}} \widehat{f_1}(\xi)\widehat{f_2}(\xi)\widehat{f_3}(-\xi)d\xi.$$

This suggests that we can establish criteria for the pseudorandomness if ν has nice properties on the Fourier transform side.

We shall take an L^∞-approach in this section. This is first introduced by Green [18] in order to prove Roth's theorem in the primes.

Proposition 4.3.3. *Let $\nu : \mathbb{Z} \to \mathbb{R}$ be a non-negative function supported on $[M]$ such that $\sum_{x\in\mathbb{Z}} \nu(x) = M(1+o(1))$. Suppose ν satisfies the following conditions:*

(1) (Fourier decay) there exists some decay function $d(M) = \exp(-\exp(k^C))$, where C is a sufficiently large constant, such that

$$\|(\widehat{\nu} - \widehat{1_{[M]}})(\xi)\|_{L^\infty(\mathbb{T})} \leq d(M)M;$$

(2) (Restriction estimate) there exists some $2 < p < 3$ such that for any complex-valued function $|f| \leq \nu$, one has

$$\|\widehat{f}\|_{L^p(\mathbb{T})} \ll_p M^{1-1/p}.$$

Then ν is pseudorandom.

Before proceeding to details, let us briefly sketch the proof. Let $f \in L^1(\mathbb{Z})$ be a function such that $|f| \leq \nu$ and $\sum_{x\in\mathbb{Z}} f(x) \geq (1-\beta_k)M$. We shall use our tools from Fourier analysis to find functions $\{g_i\}_{i=1}^k$ which satisfy all the hypotheses of Proposition 4.2.3 (with $c = 1$) such that

$$\max_{1\leq i\leq k} \left\| \widehat{f|_{\mathcal{P}_i}} - \widehat{g_i} \right\|_{L^\infty(\mathbb{T})} \leq d'(M)M \tag{4.1}$$

for some decay function $d'(M) = O(\exp(-k^C))$.

On the other hand, by Hölder's inequality,

$$|T(f_1,f_2,f_3)| \leq \|\widehat{f_1}\|_{L^\infty(\mathbb{T})}^{1/(2p)} \|\widehat{f_1}\|_{L^{(2p-1)/(2(p-2))}(\mathbb{T})}^{1-1/(2p)} \|\widehat{f_2}\|_{L^p(\mathbb{T})} \|\widehat{f_3}\|_{L^p(\mathbb{T})}. \tag{4.2}$$

By expanding the right hand side of

$$\sum_{i=1}^k T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}) = \sum_{i=1}^k T(g_i + (f|_{\mathcal{P}_i} - g_i), g_i + (f|_{\mathcal{P}_i} - g_i), g_i + (f|_{\mathcal{P}_i} - g_i)),$$

it suffices to show that all terms except $\sum_{i=1}^k T(g_i,g_i,g_i)$ contribute only a small error term. To do this, we use (4.1) to bound the L^∞-norm in (4.2). To bound the L^p terms, we notice that since $0 \leq g_i \leq 2$ and g is supported on $[M]$, for any $1 \leq i \leq k$ we have

$$\|\widehat{g_i}\|_{L^\infty(\mathbb{T})} \leq \|g_i\|_{L^1(\mathbb{Z})} \ll M$$

and

$$\|\widehat{g_i}\|_{L^2(\mathbb{T})} = \|g_i\|_{L^2(\mathbb{Z})} \ll M^{1/2},$$

where the second inequality is a consequence of Plancherel's theorem. Thus, by Hölder's inequality, one has

$$\|\widehat{g_i}\|_{L^p(\mathbb{T})} \ll M^{1-1/p}$$

for all $2 \leq p \leq \infty$. Since for any $1 \leq i \leq k$, the function $f|_{\mathcal{P}_i}$ still satisfies the pointwise bound $|f|_{\mathcal{P}_i}| \leq \nu$, it follows from the restriction estimate that

$$\|\widehat{f|_{\mathcal{P}_i}}\|_{L^p(\mathbb{T})} \ll_p M^{1-1/p}.$$

Therefore, by the triangle inequality, for any $1 \leq i \leq k$ one has

$$\|\widehat{f|_{\mathcal{P}_i}} - \widehat{g_i}\|_{L^p(\mathbb{T})} \ll_p M^{1-1/p}.$$

Combine this with (4.1) and (4.2) yields

$$\sum_{i=1}^k T(f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}, f|_{\mathcal{P}_i}) = \sum_{i=1}^k T(g_i,g_i,g_i) + O\left(kd'(M)^{1/(2p)}M^2\right). \tag{4.3}$$

4.3.1 The L^∞ approximation

Let α, ρ, η be positive parameters to be chosen later. Let

$$R := \bigcup_{1 \leq i \leq k} \mathrm{Spec}(f|_{\mathcal{P}_i}, \alpha)$$

denote the large spectrum. Let

$$B := B(M, R, \rho, \eta).$$

Let

$$\sigma(x) := \frac{1_B(x)}{|B|}$$

be the normalised characteristic function of B.

For each $1 \leq i \leq k$, we define

$$g_i := f|_{\mathcal{P}_i} * \sigma * \sigma \text{ and } g := \sum_{i=1}^{k} g_i,$$

so that

$$g = f * \sigma * \sigma.$$

The reason of taking the two-fold convolution is to obtain pointwise upper bounds which are good enough for our purpose, as taking convolutions make functions look "smoother". Notice that

$$\|g\|_{L^1(\mathbb{T})} = \|f\|_{L^1(\mathbb{T})}.$$

We shall first prove a pointwise upper bound on g. By choosing the parameters appropriately in the next subsection, this would ensure that g satisfies the boundedness condition needed in order to apply Proposition 4.2.3.

Lemma 4.3.4. *For every $n \in \mathbb{Z}$, we have*

$$g(n) \leq 1 + \frac{d(M)M}{|B|}.$$

Proof. Since $f(n) \leq \nu(n)$ for all $n \in \mathbb{Z}$, we have

$$g(n) = f * \sigma * \sigma(n) \leq \nu * \sigma * \sigma(n).$$

Therefore, it follows from the Fourier inversion formula, the triangle inequality and Plancherel's theorem that

$$\begin{aligned}
g(n) \leq & \int_{\mathbb{T}} \widehat{\sigma}^2(\xi) \left(\widehat{1_{[M]}}(\xi) + \widehat{\nu}(\xi) - \widehat{1_{[M]}}(\xi) \right) e(n\xi) d\xi \\
\leq & \left| \int_{\mathbb{T}} \widehat{\sigma}^2(\xi) \widehat{1_{[M]}}(\xi) e(n\xi) d\xi \right| + \int_{\mathbb{T}} \widehat{\sigma}^2(\xi) |\widehat{\nu}(\xi) - \widehat{1_{[M]}}(\xi)| d\xi \\
\leq & \frac{1}{|B|^2} \sum_{x+y+z=n} 1_{[M]}(x) 1_B(y) 1_B(z) + \|\widehat{\nu} - \widehat{1_{[M]}}\|_\infty \int_{\mathbb{T}} \widehat{\sigma}^2(\xi) d\xi \\
\leq & 1 + \frac{d(M)M}{|B|}.
\end{aligned}$$

□

Lemma 4.3.5. *For any $1 \leq i \leq k$ and $\xi \in \mathbb{T}$, we have*

$$|\widehat{f|_{\mathcal{P}_i}}(\xi) - \widehat{g}_i(\xi)| \ll \|\nu\|_1 \rho + \alpha.$$

Proof. Notice that for each i,

$$\widehat{f|_{\mathcal{P}_i}}(\xi) - \widehat{g}_i(\xi) = \widehat{f|_{\mathcal{P}_i}}(\xi)(1 - \widehat{\sigma}^2(\xi)).$$

If $\xi \in R$, then it follows from the definition of the Bohr set that

$$|1 - \widehat{\sigma}(\xi)| = \left| \frac{1}{|B|} \sum_{x \in B} (1 - e(-\xi x)) \right| \ll \rho.$$

Combining this with $|\widehat{f|_{\mathcal{P}_i}}(\xi)| \leq \|\nu\|_1$ yields

$$\begin{aligned} |\widehat{f|_{\mathcal{P}_i}}(\xi) - \widehat{g}_i(\xi)| \leq & |\widehat{f|_{\mathcal{P}_i}}(\xi)||1 - \widehat{\sigma}(\xi)||1 + \widehat{\sigma}(\xi)| \\ \ll & \|\nu\|_1 \rho. \end{aligned}$$

If $\xi \notin R$, then $|\widehat{f|_{\mathcal{P}_i}}(\xi)| \leq \alpha$, and so

$$|\widehat{f|_{\mathcal{P}_i}}(\xi) - \widehat{g}_i(\xi)| \ll \alpha.$$

□

4.3.2 Proof of Proposition 4.3.3

In this subsection, we prove the L^∞ transference principle stated in Proposition 4.3.3. In particular, we keep a clear record of the necessary conditions on the parameters.

In order to apply Proposition 4.2.3 to count $\sum_{i=1}^k T(g_i, g_i, g_i)$, we need to choose parameters involved in the definition of the Bohr set B and the large spectrum R, so that the following conditions hold:

1. $g = f * \sigma * \sigma$ is supported on $[-\beta'_k M, M + \beta'_k M]$,
2. $\sum_{i=1}^k \sum_{x \in \mathbb{Z}} g_i(x) \geq (1 - \beta'_k)M$, and
3. $\sum_{i=1}^k g_i(x) \leq 1 + \beta'_k$.

The first condition requires the "width" parameter η to be bounded by

$$\eta \leq \beta'_k/2,$$

since $f * \sigma * \sigma$ is supported on $[-2\eta M, M + 2\eta M]$ — this follows directly from the definition of convolution.

The second condition follows directly from our choice of g, since

$$\sum_{i=1}^{k}\sum_{x\in\mathbb{Z}} g_i(x) = \sum_{x\in\mathbb{Z}} g(x) = \|f*\sigma*\sigma\|_{\ell^1(\mathbb{Z})} = \|f\|_{\ell^1(\mathbb{Z})} \geq (1-\beta_k')M.$$

To meet the third condition, by Lemma 4.3.4, we need $|B|$ to be large enough so that $d(M)M/|B|$ is of size at most β_k'. Choose L to be the smallest integer which is bigger than $\max\{4\pi\alpha^{-1}M, 2\eta M\rho^{-1}\}$, then it follows from Lemma 0.2.4 that we need

$$(\alpha^{-1}M + \eta M\rho^{-1})\alpha^{-p}M^{p-1}\log\rho \gg \log d(M) - \log\beta_k'. \tag{4.4}$$

To ensure that $\sum_{i=1}^{k} T(g_i,g_i,g_i) \gg \beta_k''M^2$ is the main term, the parameter α from the definition of the large spectrum R needs to be sufficiently small. More specifically, it follows from (4.3) that one needs

$$d'(M) \ll (k^{-1}\beta_k)^{2p}.$$

By Lemma 4.3.5 and the previous inequality, we have the bound above when we choose

$$\rho = \exp(-k^C) \text{ and } \alpha = M\exp(-k^C)$$

where the constant C is large. Substitute this to (4.4), we can deduce that the decay function should satisfy the bound

$$d(M) \ll \exp(-\exp(k^C)),$$

which completes the proof of Proposition 4.3.3.

4.4 The Fourier restriction estimate for the primes

To prove the restriction estimate needed for the transference principle, we shall use an enveloping sieve[5] considered by Green and Tao [19]. Everything in this section follows from taking

$$F(n) = an + b$$

and using results from their paper, where the function F there is a product of linear forms which reflects the desired properties of the primes we would like to find.

[5]In the paper of Green [17] and the paper of Li and Pan [30], the authors actually proved the restriction estimate for the measure ν defined in the next section. They used Brun's pure sieve instead of the Selberg sieve.

The key properties possessed by the sieve function β_R are the majorant property and the Fourier extension estimate. The former [19, Proposition 3.1 (i)] states that for any $|a|, |b| \leq M$ and large integer $R \leq M$, one has

$$\beta_R(n) \gg 1_{X_{R!}}(n)\mathfrak{S}_F^{-1} \log R \tag{4.5}$$

where

$$\mathfrak{S}_F = \prod_p \frac{\gamma(p)}{1-1/p},$$

$$\gamma(p) = \frac{1}{p}|\{x \in \mathbb{Z}/p\mathbb{Z} : (p, an+b) = 1\}|,$$

and

$$X_{R!} = \{x \in \mathbb{Z} : (an+b, d) = 1 \text{ for all } d \leq R\}.$$

The L^p Fourier extension estimate for β_R stated below is a consequence of various Fourier properties [19, Proposition 3.1 (iii)(iv)] of β_R, together with a crude upper bound [19, Proposition 3.1 (ii)] on β_R. The proof can be found in Green and Tao [19, Proposition 4.2].

Proposition 4.4.1. *Let M be a positive integer and $1 \leq R \leq M^{1/10}$, $F(n) = an + b$ where $|a|, |b| \leq M$, and β_R be an enveloping sieve. Given any $p > 2$, for any complex-valued sequence $\{a_i\}_{i \leq M}$ we have*

$$\left\| \sum_{1 \leq x \leq M} a_x \beta_R(x) e(x\theta) \right\|_{L^p(\mathbb{T})} \ll_p M^{1/2-1/p} \left(\sum_{1 \leq x \leq M} |a_x|^2 \beta_R(x) \right)^{1/2}.$$

4.5 Proof of the Li–Pan bound

In order to control the Fourier transform and apply the transference principle, we fix

$$W := \log\log N/4$$

and

$$d := \prod_{1 \leq p \leq W} p.$$

Let M be the largest integer such that $dM + 1 \leq N$ and let

$$\nu(n) := \begin{cases} \phi(d)\log(dn+1)/d & \text{if } dn+1 \text{ is a prime and } dn+1 \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

The choice of d above will ultimately lead to good control on the Fourier bias caused by small prime numbers. This is first used in Green's paper [17, Section 6], and it is later called the W-trick in relevant literature due to the notation used in his work.

By the Siegel–Walfisz theorem, for $q \leq (\log x)^A$ one has

$$\psi(x; q, a) = \frac{x}{\phi(q)} \left(1 + O\left(x \exp\left(-C_A \sqrt{\log x}\right)\right)\right),$$

It follows that

$$\sum_{n \in \mathbb{Z}} \nu(n) = M \left(1 + O\left(\exp\left(-C\sqrt{\log N}\right)\right)\right),$$

for some $C > 0$, and so ν is a majorant on the set $\{n : dn + 1 \leq N \text{ and } dn + 1 \text{ is a prime}\}$, which is a subset of $[M]$.

We shall prove the Li–Pan bound by showing that ν is pseudorandom, which requires us to verify the Fourier decay condition and the restriction estimate for some $p > 2$. Similar to the previous chapter, we shall use the circle method which allows us to deal with the major arcs and the minor arcs using different techniques. The choice of major and minor arcs is different from the previous chapter since we need an asymptotic instead of an upper bound on the major arcs. This is also the reason of using the W-trick and choosing a smaller common difference d.

Following the choice of Green [17], let $A = 4/(p-2)$ where p is the exponent in the restriction estimate, and $B = 2A + 20$. Fix $Q = M(\log M)^{-B}$. The major arcs are defined to be the collection of $\mathfrak{M}_{a,q}$ (see Chapter 0 for the notation) where $q \leq (\log M)^B$, and the minor arcs $\mathfrak{m}$ are the rest of the circle. We remark that the choice of major arcs involved in proving the restriction estimate is highly non-uniform in p.

The proof of the Fourier decay condition below essentially follows from the one of Green [17, Section 4 and Lemma 6.2]; the only difference is that we are working with the integer group here instead of the finite field.

Proposition 4.5.1. *There exists an absolute constant $c > 0$ such that*

$$\|(\widehat{\nu} - \widehat{1_{[M]}})(\xi)\|_{L^\infty(\mathbb{T})} \leq M(\log\log M)^{-c}.$$

Proof. By Green [17, Lemma 4.8], on the major arcs, for any $q \leq (\log M)^B$, $(a, q) = 1$ and $\theta \in \mathfrak{M}_{a,q}$, we have

$$\widehat{\nu}(\theta) = \begin{cases} \dfrac{\mu(q)}{\phi(q)} e\left(\dfrac{a\overline{d}}{q}\right) \sum_{x \leq M} e\left(-\left(\theta - \dfrac{a}{q}\right) x\right) + O\left(M\left(\log M\right)^{-A}\right) & \text{if } (d, q) = 1, \\ O\left(M\left(\log M\right)^{-A}\right) & \text{otherwise.} \end{cases}$$

On the minor arcs, it follows from the minor arc estimate (see Theorem 0.3.5) which essentially follows from a result of Vinogradov that

$$\widehat{\nu}(\theta) \ll M(\log M)^{10}(q^{-1/2} + M^{-1/5} + M^{-1/2}Q^{1/2}).$$

On the other hand, a direct computation yields

$$\begin{aligned}\widehat{1_{[M]}}(\theta) &= \sum_{1\le x\le M} e(-x\theta) \\ &= \sum_{1\le y\le q} e\left(-\frac{ay}{q}\right) \sum_{\substack{1\le z\le M \\ x\equiv y \pmod q}} e\left(-z\left(\theta-\frac{a}{q}\right)\right) + O(q).\end{aligned}$$

By Green [17, Lemma 4.3], for every $q \le (\log M)^B$ and $|\theta - a/q| \le 1/qQ$ we have

$$\sum_{\substack{1\le z\le M \\ x\equiv y \pmod q}} e\left(-z\left(\theta-\frac{a}{q}\right)\right) = q^{-1}\sum_{x\le M} e\left(-x\left(\theta-\frac{a}{q}\right)\right) + O(q^{-1}M(\log M)^{-A}).$$

Using the asymptotics above, we can conclude the following.

1. The arc $\mathfrak{M}_{0,1}$: let $\theta \in \mathfrak{M}_{0,1}$. We have $\widehat{\nu}(\theta) = \sum_{x\le M} e(-\theta x) + O(M(\log M)^{-A})$, and $\widehat{1_{[M]}}(\theta) = \sum_{x\le M} e(-x\theta)$, and so
$$\|\widehat{\nu}(\theta) - \widehat{1_{[M]}}\|_{L^\infty(\mathbb{T})} \ll M(\log M)^{-A}. \tag{4.6}$$

2. The rest of major arcs: let $\theta \in \mathfrak{M}_{a,q}$ for some $2 \le q \le Q$. For $q \le W$, we have $(d,q) = q > 1$, and so $|\widehat{\nu}(\theta)| \ll M(\log M)^{-A}$. For $W \le q \le Q$, $|\widehat{\nu}(\theta)| \le M/\phi(W) + O(M(\log M)^{-A}) \le M\log\log W/W \ll M/(\log\log M)^{1/2}$. Meanwhile, using the cancellation of complete exponential sums,
$$\begin{aligned}\widehat{1_{[M]}}(\theta) = q^{-1}\sum_{x\le M} e\left(-x\left(\theta-\frac{a}{q}\right)\right)\sum_{1\le y\le q} e\left(-\frac{ay}{q}\right) + O(q^{-1}M(\log M)^{-A}) \\ = O(q^{-1}M(\log M)^{-A}).\end{aligned} \tag{4.7}$$
Therefore
$$\|\widehat{\nu}(\theta) - \widehat{1_{[M]}}\|_{L^\infty(\mathbb{T})} \ll M(\log\log M)^{-1/2}. \tag{4.8}$$

3. The minor arcs: it follows from the minor arc estimate and (4.7) that
$$\|\widehat{\nu}(\theta) - \widehat{1_{[M]}}\|_{L^\infty(\mathbb{T})} \ll M(\log M)^{-A} \le N(\log N)^{-A+1/8}. \tag{4.9}$$

Combining (4.6), (4.8) and (4.9) yields the assertion.

□

The Fourier restriction estimate follows from a rather straightforward application of Proposition 4.4.1.

Proposition 4.5.2. *Let f be any complex-valued function such that $|f| \leq \nu$. For every $p > 2$, one has*

$$\|\widehat{f}\|_{L^p(\mathbb{T})} \ll_p M^{1-1/p}.$$

Proof. Fix $R := M^{1/10}$. Let

$$f'(n) = f(n)1_{X_{R!}}(n).$$

Since all but at most $M^{1/10}$ elements of $\mathrm{supp} f$ is contained in the set $X_{R!}$ and $|f| \leq \nu$, it follows that for every $p > 2$,

$$\left| \|\widehat{f'}\|_{L^p(\mathbb{T})} - \|\widehat{f}\|_{L^p(\mathbb{T})} \right| \ll M^{1/10} \log M \leq M^{1/2-1/p}. \tag{4.10}$$

Let β_R be the sieve function from Section 4.4, and so by (4.5) we have $\beta_R(n) \gg (\log M) \prod_{p \leq \log\log N/4}(1-1/p)$ for all $n \in X_{R!}$. Take $a_x = f'(x)/\beta_R(x)$ and apply the conclusion of Proposition 4.4.1, we can conclude that

$$\|\widehat{f'}\|_{L^p(\mathbb{T})} \leq M^{1/2-1/p} \left(\sum_{1 \leq x \leq M} \frac{f'(x)^2}{\beta_R(x)} \right)^{1/2}.$$

Since $\phi(d) = d \prod_{1 \leq p \leq \log\log N/4}(1-1/p)$, we have

$$|f'(x)| \leq \nu(x) \leq \frac{\phi(d)}{d} \log M \leq \log M \prod_{1 \leq p \leq \log\log N/4} (1-1/p).$$

On the other hand, by our choice of d, we have $\gamma(p) = 1$ for $p \leq \log\log N/4$ and $\gamma(p) = 1 - 1/p$ otherwise. It follows that $\mathfrak{S}_F = \prod_{p \leq \log\log N/4} 1/(1-1/p)$ and

$$\beta_R^{-1}(x) \ll 1_{X_{R!}}(x) \frac{\prod_{p \leq \log\log N/4} 1/(1-1/p)}{\log R} \ll \frac{1_{X_{R!}}}{\log M \prod_{p \leq \log\log N/4}(1-1/p)}.$$

Therefore, we have

$$\begin{aligned}
\|\widehat{f'}\|_{L^p(\mathbb{T})} &\leq M^{1/2-1/p} \left(\sum_{1 \leq x \leq M} \frac{f'(x)}{\beta_R(x)} f'(x) \right)^{1/2} \\
&\ll M^{1/2-1/p} \left(\sum_{\substack{1 \leq x \leq M \\ x \in \mathrm{supp} f}} f'(x) \right)^{1/2} \\
&\ll M^{1-1/p}.
\end{aligned}$$

The claim follows from combining the inequality above with (4.10). □

Proof of Theorem 4.1.3. By Lemma 4.5.1 and Lemma 4.5.2, one can apply Proposition 4.3.3 if

$$(\log\log M)^{-c} \leq \exp(-\exp(k^C)).$$

Since $M \gg N/d \gg N/(\log N)^{1/4}$, the above condition holds when

$$N \gg \exp(\exp(\exp(\exp(k^{C'})))),$$

where C' is an absolute constant. □

Chapter 5

Partition regularity over shifted primes, II

5.1 Introduction

In this chapter, we use the iteration lemma from Chapter 2 to improve the bound from the previous chapter.

Theorem 5.1.1. *Let $r_p(k)$ denote the smallest integer such that for any $N \geq r_p(k)$ we can find a monochromatic solutions of $p_1 - p_2 = p_3 - 1$ from any k-colouring of $\mathbb{P} \cap [N]$. Then*

$$r_p(k) \leq \exp(\exp(\exp(k^{O(1)}))).$$

There are three major steps involved in the proof of Theorem 5.1.1. Firstly, we derive an iteration lemma from Lemma 2.4.1, which helps us to find many shifted primes in the difference set of any set A with large density. By the pigeonhole principle, there exists a large monochromatic set amongst those shifted primes unless the theorem is true, and it turns out that a translate of this monochromatic set has large intersection with the original set A. These considerations ultimately yield a colouring iteration lemma. The final part of the proof is to bootstrap the colouring iteration lemma. We eventually conclude that there is always a monochromatic solution to $p_1 - p_2 = p_3 - 1$, where $p_1, p_2, p_3 \in [N] \cap \mathbb{P}$, unless N is small in terms of k, which completes the proof. The details of the aforementioned work will be given in Section 5.2 and Section 5.3.

Following Green's work on Roth's theorem in the primes, Helfgott and de Roton [22] and Naslund [32] subsequently obtained stronger bounds using variants of Green's transference principle. In Section 5.4, we list some problems which prevent us from improving the Li–Pan bound using the transference principle of Helfgott–de

Roton. These difficulties demonstrate the different nature of the density problem and the colouring problem, and somehow reflect the difficulty of adapting the Fourier analytic method to the partition regularity problems. We remark that the current best bound for Roth's theorem in the primes follows directly from the recent work of Bloom and Sisask [4], who improved the bound for Roth's theorem (in the integers) and broke the so called "logarithmic barrier". Their method is completely different from the transference principle approaches in the previous work, and there does not seem to be any obvious way to apply it to our problem.

At the end of this section, we review our bootstrapping approach under the assumption that the generalised Riemann hypothesis is true. A highlight of this section is that we are able to deduce "a nearly optimal bound" compared to Schur's theorem over the integers, although the implicit constant [1] in Theorem 5.1.2 from our method cannot be improved to $1+\epsilon$ for any $\epsilon > 0$. This strongly suggests that the "correct" bound for this problem should be way stronger than what we can prove unconditionally in Theorem 5.1.1.

Theorem 5.1.2. *Assuming the generalised Riemann hypothesis, we have*

$$r_p(k) = \exp(k^{O(1)}).$$

Since the major arc estimates used to prove Lemma 2.4.1 hold only for $\mathbb{P} \pm 1$, the method presented in this chapter cannot be used to study the partition regularity of equations $p_1 - p_2 = p_3 - t$ when $t \neq \pm 1$. The easiest example compare the problems is when $t = p$ where p is a prime number. In this case, the equation $p_1 - p_2 = p_3 - p$ always has the trivial monochromatic solution $p_1 = p_2 = p_3 = p$. Nevertheless, the argument from Chapter 2 fails, as there always exists an arithmetic progression whose difference set avoids $\mathbb{P} - p$.

5.2 The bootstrapping lemma

The aim of this section is to establish the bootstrapping lemma which is the key tool used in the proof of Theorem 5.1.1.

Firstly, we use the iteration lemma from Chapter 2 and run through a similar argument as is at the end of that chapter. Instead of looking for an upper bound on the density of the set, we shall assume a large density and locate plenty of shifted

[1] Our primary interest is to reduce the number of exponents in the bound on $r_p(k)$. By adjusting Lemma 2.4.1 using the generalised Riemann hypothesis, it is possible to improve the implicit constant that follows from our proof of Theorem 5.1.2 in Section 5.5.

primes. The following modified iteration lemma is shown by using only the results from Chapter 2. It provides us quantitative bounds that are needed to improve the quantitative aspect of Theorem 5.1.1.

The setup is slightly different from Chapter 2. Here, the exceptional modulus is rather explicitly presented in the statement of the lemma as a factor of the common difference of the arithmetic progression. The main reason of doing so is that we shall need different definitions of being exceptional in the upcoming section, and we believe an explicit expression is easier for the purpose of tracing the quantitative bounds.

Lemma 5.2.1. *Let N and d be positive integers. Let $D \geq 2$ and $\alpha > 0$. Suppose we have $D^{C_5} < N < \exp(D^{1/10})$ and $\alpha = O((\log N)^{-1})$. Suppose also that*

$$\log N \geq C_7((\log \alpha^{-1})^3 + \log\log D)\log D \text{ and } \log d + (\log(\alpha^{-1}))^2 \leq C_7^{-1}\log D \quad (5.1)$$

for some sufficiently large constant C_7. Fix $T = D^{C_2^2}$. Let $\bar{d} = d_D$ be the modulus of the exceptional zero if (D,T) is exceptional in the sense of Chapter 2, and $\bar{d} = 1$ if (D,T) is unexceptional.

Suppose that A is contained in an arithmetic progression of length N with common difference $d\bar{d}$, and has $|A| = \alpha N$. Then, there exist $A' \subseteq A - A$ which is contained in an arithmetic progression of length $N' \geq \alpha^{C_8(\log\alpha^{-1})^2}(d\bar{d})^{-C_8\log\alpha^{-1}}(\log N)^{-C_8\log\alpha^{-1}}N$ with common difference $d'\bar{d}$, where $d' \leq \alpha^{-C_8\log\alpha^{-1}}d$ and $d \mid d'$, such that

$$|A' \cap (\mathbb{P} - 1)| \geq \frac{\alpha^{C_8\log\alpha^{-1}}}{d\bar{d}\log N}N'.$$

Here, C_8 is an absolute constant which does not depend on $N, \alpha, d, \bar{d}$ or D. All small constants c with subscripts in the proof and large constants C_2, C_5 are the same ones as in Lemma 2.4.1.

Proof. Let A_0 be the affine transformation of A so that $A_0 \subseteq [N]$. We choose

$$A_1 = A_0, N_1 = N, \alpha_1 = \alpha \text{ and } d_1 = d\bar{d}.$$

Similar as what we did in Chapter 2, suppose for some positive integer k, we have obtained A_k, N_k, α_k and d_k such that A_k is a subset of $[N_k]$, $|A_k| = \alpha_k N_k$, $d\bar{d} \mid d_k$, A_k is dilated by a factor d_k^{-1} times a shifted subset of A, the inputs A_k, D, d_k satisfy the hypotheses of Lemma 2.4.1, and N_k, α_k, d_k satisfy the bounds

$$\alpha_k \geq (1 + c_9)^{k-1}\alpha, \quad (5.2)$$

$$d_k \leq c_7^{k-1}\alpha^{-3(k-1)}d\bar{d}, \quad (5.3)$$

and

$$N_k \geq \frac{c_8^{8(k-1)}\alpha^{32k^2}N}{c_7^{8k^2}(d\bar{d})^{8(k-1)}(\log N)^{8(k-1)}}. \tag{5.4}$$

Let us apply Lemma 2.4.1 to A_k, N_k, α_k and d_k. By appropriate choice of C_7 in our assumption (5.1), outcomes (iii) or (iv) of Lemma 2.4.1 cannot occur unless $k > 2(\log \alpha^{-1})/\log(1+c_9)$, where c_9 is the constant from Lemma 2.4.1. Thus, for any $k \leq \lfloor 2(\log \alpha^{-1})/\log(1+c_9)\rfloor$, we are either in outcome (i) or outcome (ii).

Since outcome (i) at every step of the iteration would produce a set with density bigger than 1 when $k > (\log \alpha^{-1})/\log(1+c_9)$, it follows that we must be in outcome (ii) for some

$$k_0 \leq \frac{\log \alpha^{-1}}{\log(1+c_9)}. \tag{5.5}$$

Therefore, using the size bounds (5.2), (5.3) and (5.4) and the quantitative bound given in outcome (ii), we have

$$\langle 1_{A_{k_0}} * 1_{-A_{k_0}}, F_{N'_{k_0},d_{k_0}}\rangle \geq \alpha_{k_0}^2 N_{k_0} |\widehat{F_{N'_{k_0},d_{k_0}}}(0)|/2$$

where

$$N'_{k_0} \gg \alpha_{k_0}N_{k_0} \gg \frac{c_8^{8(k_0-1)}\alpha^{32k_0^2+2}N}{(d\bar{d})^{8k_0}(\log N)^{8k_0}}, \tag{5.6}$$

and

$$|\widehat{F_{N'_{k_0},d_{k_0}}}(0)| \gg \frac{N'_{k_0}}{d_{k_0}}.$$

Since $1_{A_{k_0}} * 1_{-A_{k_0}}(n) \leq |A_{k_0}| = \alpha_{k_0}N_{k_0}$ and $|F_{N'_{k_0},d_{k_0}}(n)| \leq \log(N'_{k_0}d_{k_0}+1)$ for all $n \in \mathbb{Z}$, it follows that[2]

$$\begin{aligned}|\mathrm{supp}(1_{A_{k_0}} * 1_{-A_{k_0}}F_{N'_{k_0},d_{k_0}})| &\gg \frac{1}{\alpha_{k_0}N_{k_0}\log(N'_{k_0}d_{k_0})}\alpha_{k_0}^2N_{k_0}\frac{N'_{k_0}}{d_{k_0}}\\ &\geq \frac{\alpha_{k_0}N'_{k_0}}{d_{k_0}\log(N'_{k_0}d_{k_0})}\\ &\gg \frac{\alpha^{6k_0}}{d\bar{d}\log N}N'_{k_0}.\end{aligned}$$

Let

$$A'_{k_0} := \{n \in A_{k_0} - A_{k_0} : n \leq N'_{k_0}, d_{k_0}n+1 \text{ is a prime}\}.$$

Since $F_{N'_{k_0},d_{k_0}}$ is supported on those n such that $d_{k_0}n+1 = p^k$ for some prime p and $k \geq 1$, and there are at most $O(\sqrt{N'_{k_0}d_{k_0}})$ composite prime powers which

[2]In the last inequality below, we use the assumption that α is significantly smaller than c_i for any i. This is the situation we will need in the upcoming sections.

are contained in the support above, it follows from (5.1) and the lower bound of $|\mathrm{supp}(1_{A_{k_0}} * 1_{-A_{k_0}} F_{N'_{k_0},d_{k_0}})|$ that

$$|A'_{k_0}| \gg \frac{\alpha^{6k_0}}{d\bar{d}\log N} N'_{k_0}. \tag{5.7}$$

Let A' be the pre-image of A'_{k_0} under the affine transformations carried out in the proof above. Since A_{k_0} is dilated by a factor $d_{k_0}^{-1}$ times a shifted subset of A, it follows that A' is contained in an arithmetic progression of length N'_{k_0} with common difference d_{k_0}, and A' consists of numbers which are one less than a prime. Therefore, one can conclude that the assertion of the lemma holds for sufficiently large C_8 by (5.3), (5.5), (5.6) and (5.7).

□

The second lemma we need for the bootstrapping follows from an averaging argument.

Lemma 5.2.2. *Let $X \subseteq \mathbb{Z}$ be a set contained in an arithmetic progression of length N with common difference d. Suppose $Y \subseteq \mathbb{Z}$ is contained in an arithmetic progression of length N' with common difference dd' where d' is a positive integer, then there exists $n \in \mathbb{Z}$ such that*

$$|(n+Y) \cap X| \geq \frac{|X||Y|}{N + d'N'}.$$

Proof. We assume $d = 1$ in the proof below, since the general result follows from applying the $d = 1$ result to a dilation of the sets X and Y.

Notice that the set

$$\{x \in \mathbb{Z} : (x+Y) \cap X \neq \emptyset\}$$

is contained in an interval of length at most $N + d'N'$. Also, by a direct counting argument, we have

$$\sum_{x \in \mathbb{Z}} |(x+Y) \cap X| = |X||Y|.$$

Therefore, it follows from the pigeonhole principle that there exists some $n \in \mathbb{Z}$ such that

$$|(n+Y) \cap X| \geq \frac{|X||Y|}{N + d'N'},$$

which completes the proof. □

We combine the iteration lemma and the averaging lemma above to deduce the bootstrapping lemma, which will be the main lemma for the argument in the upcoming section. To make the statement less complicated, we shall also simplify the constants involved in the iteration lemma.

There are two steps involved in the proof below. The first one is to apply the iteration lemma to an initial set and the pigeonhole principle to find either a desired monochromatic solution, or a large monochromatic subset of shifted primes in its difference set. In the latter case, we then use the averaging lemma above to locate one translate whose intersection with the initial set is large. Since difference sets are translation invariant, the difference set of this particular translate lies in the difference set of the initial set, and so it does not contain the colour of the initial set.

Lemma 5.2.3. *Let N_0 and k be positive integers. Let $N \leq N_0$ and d be positive integers. Let D be a positive number such that $D^{C_5} < N < \exp(D^{1/10})$ and $\alpha = O((\log N)^{-1})$. Fix $T = D^{C_2^2}$. Let $\bar{d} = d_D$ be the modulus of the exceptional zero if (D, T) is exceptional in the sense of Chapter 2, and $\bar{d} = 1$ if (D, T) is unexceptional.*

Suppose N, d, α, D satisfy (5.1), then for any k-colouring of $(\mathbb{P} - 1) \cap [N_0]$ and $j \leq k$, we have the following.

Suppose $A \subseteq [N_0]$ is a monochromatic subset of $\mathbb{P} - 1$ contained in an arithmetic progression of length N with common difference $d\bar{d}$, and has $|A| = \alpha N$. Suppose also $A \cup ((A - A) \cap (\mathbb{P} - 1))$ is j-coloured, then

1. *either there exists a monochromatic solution of $x - y = z$ where $x, y, z \in (\mathbb{P} - 1) \cap [N_0]$;*

2. *or there exists a monochromatic set $A'' \subseteq \mathbb{P} - 1$ satisfying the following:*

   (a) *A'' is contained in an arithmetic progression of length N'' and common difference $d''\bar{d}$, where $N'' \geq \alpha^{C(\log \alpha^{-1})^2}(dD)^{-C \log \alpha^{-1}}(\log N)^{-C \log \alpha^{-1}} N$, $d \mid d''$ and $d'' \leq \alpha^{-C \log \alpha^{-1}} d$;*

   (b) *$|A''| \geq \alpha^{C \log \alpha^{-1}}((j-1)dD \log N)^{-1} N''$;*

   (c) *$A'' \cup ((A'' - A'') \cap (\mathbb{P} - 1))$ is $(j-1)$-coloured.*

Proof. Suppose there exists $z \in (A - A) \cap (\mathbb{P} - 1)$ which shares the same colour as A, then we are in the first outcome. Hence, from now on, we assume that $(A - A) \cap (\mathbb{P} - 1)$ is at most $(j-1)$-coloured, and every element of $(A - A) \cap (\mathbb{P} - 1)$ has a different colour from A.

By applying Lemma 5.2.1 with A, α, d, D, N, we can find a set $A' \subseteq A - A$ such that A' is contained in an arithmetic progression of length

$$N' \geq \frac{\alpha^{C_8(\log \alpha^{-1})^2} N}{(dD)^{C_8 \log \alpha^{-1}} (\log N)^{C_8 \log \alpha^{-1}}}$$

with common difference $d'\bar{d}$ where

$$d' \leq \alpha^{-C_8 \log \alpha^{-1}} d, \quad d \mid d',$$

and

$$|A' \cap (\mathbb{P} - 1)| \geq \frac{\alpha^{C_8 \log \alpha^{-1}}}{dD \log N} N'.$$

Let B be the largest colour class in $A' \cap (\mathbb{P} - 1)$. It follows that B is contained in an arithmetic progression of length N' with common difference $d'\bar{d}$, and

$$|B| \geq \frac{\alpha^{C_8 \log \alpha^{-1}}}{(j-1)dD \log N} N'.$$

By Lemma 5.2.2, there exists some n such that

$$|(n + B) \cap A| \geq \frac{|A||B|}{N + (d'/d)N'}.$$

Since $A' \subseteq A - A$, we have $d'N' \leq 2dN$, and so

$$\begin{aligned} |(n + B) \cap A| &\geq \frac{|A||B|}{3N} \\ &\geq \frac{\alpha N \alpha^{C_8 \log \alpha^{-1}}}{3N(j-1)dD \log N} N' \\ &\geq \frac{\alpha^{(C_8+1) \log \alpha^{-1}}}{(j-1)dD \log N} N'. \end{aligned}$$

Let

$$A'' := B \cap (A - n).$$

Since $A'' \subseteq B$, it follows that A'' is monochromatic and satisfies condition (a), with $N'' = N'$ and $d'' = d'$. We have

$$|A''| = |(n + B) \cap A| \geq \frac{\alpha^{(C_8+1) \log \alpha^{-1}}}{(j-1)dD \log N} N''$$

which completes the proof of assertion (b). By our assumption which rules out outcome (i), we know that $A'' \subseteq B$ is monochromatic and has a different colour from A. Since $A'' \subseteq A - n$, we have $A'' - A'' \subseteq A - A$, and so $(A'' - A'') \cap (\mathbb{P} - 1)$ contains at most $j - 1$ colour classes and they are different from the colour of A. By combining these observations, we can conclude that assertion (c) is true.

□

5.3 The main result

By introducing an affine transformation and applying Lemma 2.4.2, one can deduce the following generalisation of Lemma 2.4.2.

Lemma 5.3.1. *Let A be a subset of an arithmetic progression of length N with common difference d and have $|A| = \alpha N$ for some $\alpha > 0$. For every positive integer $\bar{d}$, there exists an arithmetic progression P with common difference $d\bar{d}$ and length at least $c_{11}\alpha N/\bar{d}$ such that $|A \cap P| \geq \alpha|P|/2$. Here c_{11} is the positive absolute constant from Lemma 2.4.2.*

The proof of the main theorem essentially follows from bootstrapping Lemma 5.2.3 from the previous section. The initial monochromatic set follows from the prime number theorem and a basic application of the pigeonhole principle. Afterwards, we shall repeatedly apply Lemma 5.2.3 until we find a monochromatic solution of $x + y = z$ in $\mathbb{P} - 1$, or we shall run out of the colour classes which is impossible. In the proof below, we shall clarify our choices of sets at each step of the bootstrapping, and keep a detailed record of the quantities involved there.

Proof of Theorem 5.1.1. Let $N_0 = N$. By the prime number theorem and the pigeonhole principle, the largest colour class in $(\mathbb{P} - 1) \cap [N_0]$ must have cardinality at least $(1 + o(1))N_0(k \log N_0)^{-1}$. Thus, we can always find a monochromatic subset $A_0 \subseteq [N_0]$ such that $|A_0| = \alpha_0 N_0$ with $\alpha_0 = (2k \log N_0)^{-1}$.

We shall proceed by contradiction. Suppose the theorem is false, then we claim that for each $1 \leq i \leq k$, we can find D_i, A_i, N_i, α_i, $\overline{d_i}$, A_i'', N_i'', d_i'', α_i'' which satisfy the following conditions:

(I) $2^{-k-4+i}(\log(2\alpha_i)^{-1})^3 \leq \log D_i \leq 2^{-k-4+i}(\log \alpha_i^{-1})^3$;

(II) $2^{-k-4+i}(\log(2\alpha_{i-1})^{-1})^3 \leq \log \alpha_i^{-1} \leq 2^{-k-3+i}(\log \alpha_{i-1}^{-1})^3$;

(III) $\log d_i \leq 2^{-k-4+i}(\log \alpha_{i-1}^{-1})^3$;

(IV) $\log N_i \geq \log N_{i-1} - C'(\log \alpha_{i-1}^{-1})^9$ where C' is a large absolute constant;

(V) A_i'' is contained in an arithmetic progression of length N_i'' with common difference $d_i''\overline{d_i}$, where $N_i'' \geq \alpha_i^{C(\log \alpha_i^{-1})^2}(d_i D_i)^{-C \log \alpha_i^{-1}}(\log N_i)^{-C \log \alpha_i^{-1}} N_i$, $d_i \mid d_i''$ and $d_i'' \leq \alpha_i^{-C \log \alpha_i^{-1}} d_i$;

(VI) $|A_i''| = \alpha_i'' N_i''$ where $\alpha_i''/2 \leq \alpha_i^{C \log \alpha_i^{-1}}((k-i)d_i D_i \log N_i)^{-1} \leq \alpha_i''$; and

(VII) $A_i'' \cup ((A_i'' - A_i'') \cap (\mathbb{P} - 1))$ is $(k-i)$-coloured;

In particular, since condition (VI) and condition (VII) are impossible for $i = k$, we come to a contradiction at the kth step and so the theorem must be true.

Let $D_1 = \exp((\log \alpha_0^{-1})^3/2^{k+3})$ and $d_1 = 1$.

- if $(D_1, D_1^{C_2^2})$ is unexceptional, we let $N_1 = N_0$, $\overline{d_1} = 1$, and A_1 be a subset of A_0 such that $|A_1| = \alpha_1 N_1$ where[3] $2^{-k-3}(\log(2\alpha_0)^{-1})^3 \leq \log \alpha_1^{-1} \leq 2^{-k-2}(\log \alpha_0^{-1})^3$;
- if $(D_1, D_1^{C_2^2})$ is exceptional, then let $\overline{d_1}$ be the exceptional modulus. We choose A_1 to be a subset of the set obtained from Lemma 5.3.1, so that A_1 is contained in an arithmetic progression with common difference $\overline{d_1}$ and length $N_1 \gg \alpha_0 N_0/\overline{d_1}$, and A_1 has density α_1 in this arithmetic progression where $2^{-k-3}(\log(2\alpha_0)^{-1})^3 \leq \log \alpha_1^{-1} \leq 2^{-k-2}(\log \alpha_0^{-1})^3$.

By our choice of D_1 and the bound $k \leq (\log\log\log N_0)^c$ for some small constant c, the parameters satisfy the conditions listed in (5.1), and so we can apply Lemma 5.2.3 to conclude that either the theorem is true (when the first outcome of Lemma 5.2.3 occurs) which contradicts our assumption, or we can find a monochromatic set $A_1'' \subseteq \mathbb{P} - 1$ such that

- A_1'' is contained in an arithmetic progression of length N_1'' and common difference $d_1''\bar{d}_1$, where $N_1'' \geq \alpha_1^{C(\log \alpha_1^{-1})^2}(d_1 D_1)^{-C\log \alpha_1^{-1}}(\log N)^{-C\log \alpha_1^{-1}} N_1$, $d \mid d''$ and $d_1'' \leq \alpha_1^{-C\log \alpha_1^{-1}} d_1$;
- $|A_1''| = \alpha_1'' N_1''$ where $\alpha_1''/2 \leq \alpha_1^{C\log \alpha_1^{-1}}((k-1)d_1 D_1 \log N_1)^{-1} \leq \alpha_1''$;
- $A_1'' \cup ((A_1'' - A_1'') \cap (\mathbb{P} - 1))$ is $(k-1)$-coloured.

In particular, it follows from the properties above that conditions (I)-(VII) hold for $i = 1$.

Suppose for some $1 \leq j \leq k-1$, we have defined D_i, A_i, N_i, α_i, $\overline{d_i}$, A_i'', N_i'', d_i'', α_i'' which satisfy conditions (I)-(VII) for all $1 \leq i \leq j$. We choose D_{j+1}, A_{j+1}, N_{j+1}, α_{j+1}, $\overline{d_{j+1}}$, A_{j+1}'', N_{j+1}'', d_{j+1}'', α_{j+1}'' as follows. Fix

$$D_{j+1} = \exp((\log \alpha_j''^{-1})^3/2^{k-j+3}) \text{ and } d_{j+1} = d_j''\overline{d_j}.$$

Depending on whether $(D_{j+1}, D_{j+1}^{C_2^2})$ is exceptional, we define $A_{j+1}, N_{j+1}, \overline{d_{j+1}}, \alpha_{j+1}$ as follows:

[3]This is always possible since α_1 stated here is way smaller than α_0 since $k \leq (\log\log\log N_0)^c$ for some $c < 1$. Similarly, we can make such a choice in the exceptional case.

(A.I) if $(D_{j+1}, D_{j+1}^{C_2^2})$ is unexceptional, we take $A_{j+1} = A_j''$, $N_{j+1} = N_j''$, $\overline{d_{j+1}} = 1$ and $\alpha_{j+1} = \alpha_j''$;

(A.II) if $(D_{j+1}, D_{j+1}^{C_2^2})$ is exceptional, then let $\overline{d_{j+1}}$ be the exceptional modulus. Let A_{j+1} be the set obtained from Lemma 5.3.1, so that A_{j+1} is contained in an arithmetic progression with common difference $d_{j+1}\overline{d_{j+1}}$ and length $N_{j+1} \gg \alpha_j'' N_j''/\overline{d_{j+1}}$, and A_{j+1} has density $\alpha_{j+1} \geq \alpha_j''/2$ in this arithmetic progression.

We shall first show that conditions (I)-(IV) hold for $j+1$.

Condition (I) is a direct consequence of the definitions of D_{j+1} and α_{j+1}.

By (A.I), (A.II) and (VI), we have

$$\begin{aligned}\log \alpha_{j+1} \geq & \log \alpha_j'' - \log 2 \\ \geq & -C(\log \alpha_j^{-1})^2 - \log(k-j) - \log d_j - \log D_j - \log\log N_j - \log 2.\end{aligned}$$

Since $k = (\log\log\log N_0)^c$ for some small constant $c < 1$, and $\alpha_i \leq (k \log N_0)^{-1}$ for all $i \leq j$, we have $C(\log \alpha_j^{-1})^2 + \log(k-j) + \log\log N_j + \log 2 \leq 2^{-k-5+j}(\log \alpha_j^{-1})^3$. By (II), (III) and $\alpha_j \leq (k \log N_0)^{-1}$, we also have

$$\log d_j \leq 2^{-k-4+j}(\log \alpha_{j-1}^{-1})^3 \leq 2^{-k-1+j} \log \alpha_j^{-1} \leq 2^{-k-5+j}(\log \alpha_j^{-1})^3.$$

Combine the bounds above with (I), we conclude that

$$\log \alpha_{j+1}^{-1} \leq 2^{-k-4+j}(\log \alpha_j^{-1})^3 + \log D_j \leq 2^{-k-3+j}(\log \alpha_j^{-1})^3.$$

On the other hand, since $\log \alpha_{j+1}^{-1} \geq \log \alpha_j^{''-1}$, it follows from (I) and (VI) that

$$\log \alpha_{j+1}^{-1} \geq \log D_j \geq 2^{-k-4+j}(\log(2\alpha_j)^{-1})^3.$$

Thus, we conclude that condition (II) holds for $j+1$.

To see that condition (III) is true, notice that combining our definitions of d_{j+1} and $\overline{d_j}$, (I), (III) and (V) yields

$$\begin{aligned}\log d_{j+1} = & \log d_j'' + \log \overline{d_j} \\ \leq & C(\log \alpha_j^{-1})^2 + \log d_j + \log D_j \\ \leq & C(\log \alpha_j^{-1})^2 + 2^{-k-4+j}(\log \alpha_{j-1}^{-1})^3 + 2^{-k-4+j}(\log \alpha_j^{-1})^3.\end{aligned}$$

Since $k \leq (\log\log\log N_0)^c$ and $\alpha_j \leq (k \log N_0)^{-1}$, it follows that $C(\log \alpha_j^{-1})^2 \leq 2^{-k-5+j}(\log \alpha_j^{-1})^3$. On the other hand, by condition (II), $k \leq (\log\log\log N_0)^c$ and

$\alpha_j \le (k \log N_0)^{-1}$, we have $2^{-k-4+j}(\log \alpha_{j-1}^{-1})^3 \le \log 2^{-k-1+j}\alpha_j^{-1} \le 2^{-k-5+j}(\log \alpha_j^{-1})^3$. Therefore,

$$\log d_{j+1} \le 2^{-k-3+j}(\log \alpha_j^{-1})^3.$$

By (A.I), (A.II) and (V), one obtains

$$\begin{aligned}\log N_{j+1} \ge & \log N_j'' - \log D_{i+1} + \log \alpha_j'' + O(1)\\ \ge & - C(\log \alpha_j^{-1})^3 - C(\log d_j + \log D_j)\log \alpha_j^{-1} - C\log \alpha_j^{-1}\log\log N_j\\ & + \log N_j - \log D_{j+1} - C(\log \alpha_j^{-1})^2\\ & - \log(k-j) - \log d_j - \log D_j - \log\log N_j + O(1).\end{aligned}$$

Notice that

$$(\log d_j)(\log \alpha_j^{-1}) + \log \alpha_j^{-1}\log\log N_j \ll (\log \alpha_j^{-1})^2$$

which follows from (II), (III) and the inequality $\alpha_j \le \alpha_0 \le (k \log N)^{-1}$. Hence, by (I), (II), (A.I) and (A.II), one has

$$\begin{aligned}\log N_{j+1} \ge & - C(\log \alpha_j^{-1})^3 - 2C(\log D_j)\log \alpha_j^{-1} + \log N_j - \log D_{j+1} + O((\log \alpha_j^{-1})^2)\\ \ge & \log N_j - C'(\log \alpha_j^{-1})^9,\end{aligned}$$

and so (IV) holds for $j+1$.

In fact, it follows from the verifications above that the estimates from (5.1) hold for $A_{j+1}, D_{j+1}, \alpha_{j+1}, d_{j+1}$. The first inequality there follows from iterating (II) and (IV) and using (I). More specifically, one has

$$\log \alpha_{j+1}^{-1} \le (\log \alpha_0^{-1})^{3^{j+1}}$$

and

$$\log N_{j+1} \ge \log N_0 - (j+1)C'(\log \alpha_j^{-1})^9.$$

Therefore, the first inequality from (5.1) holds if

$$(\log \alpha_0^{-1})^{3^{j+2}} 2^{-k-3+j}(\log \alpha_0^{-1})^{3^{j+2}} \ll \log N_0 - (j+1)C'(\log \alpha_0^{-1})^{3^{j+3}},$$

for some sufficiently large implicit constant, which is true since $j \le (\log\log\log N_0)^c$ where $c < 1$ and $\alpha_0 \gg (k \log N_0)^{-1}$.

By (III), we have

$$\log d_{j+1} \le 2^{-k-3+j}(\log \alpha_j^{-1})^3.$$

By (VI), we have $\log \alpha_j^{''-1} + \log 2 \ge C(\log \alpha_j^{-1})^2$, and so

$$\log \alpha_j^{-1} \le (2(\log \alpha_j^{''-1})/C)^{1/2}.$$

Thus, it follows that

$$\log d_{j+1} \leq 2^{-k-3+j}(2/C)^{3/2}(\log \alpha_j^{''-1})^{3/2}.$$

Combining this inequality with the definition of D_{j+1}, $k \leq (\log\log\log N_0)^c$ and $\alpha_j \leq (k \log N_0)^{-1}$ yields

$$\log d_{j+1} \leq C_7^{-1}(\log D_{j+1})/2.$$

On the other hand, it follows from condition (I), $k \leq (\log\log\log N_0)^c$ and $\log \alpha_{j+1}^{-1} \gg \log(k \log N_0)$ that

$$(\log \alpha_{j+1}^{-1})^2 \leq C_7^{-1}(\log D_{j+1})/2.$$

Thus, the second inequality of (5.1) also holds.

Therefore, by Lemma 5.2.3, either the theorem holds (when the first outcome of Lemma 5.2.3 occurs) which contradicts our assumption, or we can find a monochromatic set $A''_{j+1} \subseteq \mathbb{P} - 1$ such that

(B.I) A''_{j+1} is contained in an arithmetic progression of length N''_{j+1} and common difference $d''_{j+1}\overline{d_{j+1}}$, where $d_{j+1} \mid d''_{j+1}$, $d''_{j+1} \leq \alpha_{j+1}^{-C\log\alpha_{j+1}^{-1}} d_{j+1}$, and $N''_{j+1} \geq \alpha_{j+1}^{C(\log\alpha_{j+1}^{-1})^2}(d_{j+1}D_{j+1})^{-C\log\alpha_{j+1}^{-1}}(\log N_{j+1})^{-C\log\alpha_{j+1}^{-1}} N_{j+1}$;

(B.II) $|A''_{j+1}| = \alpha''_{j+1}N''_{j+1}$ where

$$\alpha''_{j+1}/2 \leq \alpha_{j+1}^{C\log\alpha_{j+1}^{-1}}((k-j-1)d_{j+1}D_{j+1}\log N_{j+1})^{-1} \leq \alpha''_{j+1};$$

(B.III) $A''_{j+1} \cup ((A''_{j+1} - A''_{j+1}) \cap (\mathbb{P} - 1))$ is $(k-j-1)$-coloured.

In particular, this implies conditions (V)-(VII) hold for $i = j+1$.

□

5.4 Transference principle revisited

The key difference between the L^2 and L^∞ transference principle is the type of estimates for the majorant, which is reflected by how the methods are named. Instead of using the size of

$$\|\widehat{\nu} - \widehat{1_{[M]}}\|_{L^\infty(\mathbb{T})},$$

one looks at the size of

$$\|\widehat{\nu}\|_{L^2(\mathbb{T})} = \|\nu\|_{L^2(\mathbb{Z})},$$

where the equality follows from Plancherel's theorem[4]. To estimate the right hand side, one expands the expression directly and estimates the diagonal and off-diagonal terms separately. Roughly speaking, by choosing the majorant appropriately, one has the following correlation estimate

$$\sum_n \nu(n+m_1)\nu(n+m_2) \ll M.$$

There are two major problems arising from our adaption of the Helfgott–de Roton proof. Firstly, there is no guarantee that the residue class $dn+1$ contains enough primes when d is very large (e.g. $d = N^{1/100}$), and so the error terms are not small enough for any transference principle. Secondly, the loss of absolute constants in the L^2 transference principle would lead to functions that are not dense enough (in the sense of Proposition 4.2.3) for the colouring problem.

To make this more precise, let us follow the steps of Helfgott and de Roton below.

The very first step of their proof is slightly different from Green's proof of Roth's theorem. Specifically, they use the pigeonhole principle to find one residue class such that

$$|\{dn+b : dn+b \leq N \text{ and } dn+b \text{ is a prime}\}| \gg \frac{N}{\phi(d)\log N} \tag{5.8}$$

where d goes beyond the range covered by the Siegel–Walfisz theorem. By the Fourier restriction theorem of Green and Tao [19], one can still obtain the restriction estimate for the Selberg sieve as long as $d \leq N^c$ for some small positive constant c.

Since Roth's theorem in the primes concerns with the equation $p_1+p_2 = 2p_3$, any choice of b would eventually suffice to prove the theorem. However, in our case, we can only work with $b = 1$, since otherwise the equation associated with the argument would be $p_1 + p_2 = p_3 + b$ instead.

Unfortunately, once d becomes larger, the residue class $dn+1$ does not necessarily contain plenty of primes, due to the possible Siegel zero. As a consequence, neither the L^2 estimates from Helfgott-de Roton nor the L^∞ estimates from the previous chapter would be strong enough to handle the error terms arising from the relevant argument.

In the rest of this section, we assume the residue class $dn+1$ does contain enough primes so that (5.8) holds for $b = 1$.

The main problem of applying the L^2 transference principle to our problem now is due to Proposition 4.2.3, which essentially requires the functions to be very "dense" on its support. Heuristically speaking, the functions $f_1, ..., f_k$ there look like sets, and

[4]By definition, any majorant satisfies the analytic conditions needed for Plancherel's theorem.

their sum resembles a set with density $1 - O(\exp(-k^{O(1)}))$. This is very different from Roth's theorem — if Roth's theorem holds for a set, then the same qualitative result[5] holds for any subset with positive density, regardless how small the density is. This is not the case for the colouring problem. An example to see this is the set of all integers contained in $[N/2, N]$, which has density $1/2$ in $[N]$, does not contain any solutions to $x + y = z$.

The lemma used by Helfgott and de Roton to find a dense set is the following [22, Lemma 2.4].

Lemma 5.4.1. *Suppose S is a set with M elements, and $a : \mathbb{Z} \to \mathbb{R}$ is a nonnegative function supported on S. Suppose also that there exist positive constants α and c such that*

$$\|a\|_1 \geq \alpha M \quad \text{and} \quad \|a\|_2^2 \leq cM.$$

Then there exists a subset A' of S such that

(i) $|A'| \geq \alpha^2 M/(4c)$ *and*

(ii) $a(n) \geq \alpha/2$ *for all* $n \in A'$.

Proof. This is a sketch of the proof due to Helfgott and de Roton [22, Lemma 2.4]. Let

$$A'_t := \{n \in \mathbb{Z} : a(n) \geq t\}.$$

By our definition of A' and Cauchy–Schwarz, we have

$$\alpha M \leq \sum_n a(n) \leq t(M - |A'_t|) + \sum_{n \in A'_t} a(n) \leq t(M - |A'_t|) + \sqrt{|A'_t|}\sqrt{cM}.$$

By using properties of the quadratic equation $f(x) = tx^2 - \sqrt{cM}x + \alpha M - tM$, one concludes that

$$\sqrt{|A'_t|} \geq \frac{\sqrt{cM} - \sqrt{cM - 4t(\alpha - t)}}{2t}.$$

Taking $t = \alpha/2$ yields the desired result. □

It is not known whether the proof above produces the best possible constant (1/4 in (i), and it is easy to see that one cannot replace it by anything better than 1 using the method above). Since the implicit constant c, which is a consequence of the pair correlation estimate and bounding certain products of primes, cannot be as

[5]Compared to the quantitative bound on the number of 3-term arithmetic progressions in the original set, the quantitative bound for the subset depends only on the density and the quantitative bound for the original set.

small as desired, we shall end up getting a set with much smaller density than the set X' from the proof of Proposition 4.2.3. On the other hand, the size of the implicit constant c seems a rather fundamental problem for our adaption of the Helfgott–de Roton method. By a careful inspection of Lemma 2.3 of their paper, one may deal with the contribution of the diagonal terms by an appropriate choice of the Bohr set, nevertheless it seems unlikely one can obtain a sufficiently small absolute constant when one tries to bound the off-diagonal terms.

5.5 Partition regularity of $x+y=z$ over $\mathbb{P}-1$ under the generalised Riemann hypothesis

In this section, we look at the same problem assuming the generalised Riemann hypothesis, and investigate the quantitative bounds we can deduce using the method of this chapter. Since the ideas are identical to the previous sections, we shall only give full proofs when the generalised Riemann hypothesis significantly simplifies the technical details.

Throughout this section, we assume that the generalised Riemann hypothesis is true.

Since the generalised Riemann hypothesis asserts that all zeros of the Dirichlet L-functions are on the $1/2$ line, we no longer need to deal with the complication caused by the possible exceptional zero. We can take $D = \exp(\sqrt{\log N})$ and deduce the following from Lemma 5.2.1

Lemma 5.5.1. *Let N, d be positive integers and $\alpha \ll (\log N)^{-1}$. Suppose that*

$$\log d + (\log(\alpha^{-1}))^3 \leq C_9^{-1}\sqrt{\log N}. \tag{5.9}$$

Suppose A is contained in an arithmetic progression of length N with common difference d and has $|A| = \alpha N$. Then, there exist $A' \subseteq A - A$ which is contained in an arithmetic progression of length $N' \geq \alpha^{C_{10}(\log \alpha^{-1})^2} N d^{-C_{10}\log \alpha^{-1}} (\log N)^{-C_{10}\log \alpha^{-1}}$ with common difference d', where $d' \leq \alpha^{-C_{10}\log \alpha^{-1}} d$ and $d \mid d'$, such that

$$|A' \cap (\mathbb{P}-1)| \gg \frac{\alpha}{\log N} N'.$$

Here, C_9, C_{10} are absolute constants.

Sketch of proof. Let $\alpha_1 = \alpha$. Following the same argument as Lemma 5.2.1, one can find some $k_0 \leq \lfloor 2(\log \alpha^{-1})/\log(1+c_9) \rfloor$ (where c_9 is the constant from Lemma 2.4.1) such that

$$\langle 1_{A_{k_0}} * 1_{-A_{k_0}}, F_{N'_{k_0}, d_{k_0}} \rangle \geq \alpha_{k_0}^2 N_{k_0} |\widehat{F_{N'_{k_0}, d_{k_0}}}(0)|/2$$

where N_{k_0}, d_{k_0} are positive integers, $A_{k_0} \subseteq [N_{k_0}]$, $|A_{k_0}| = \alpha_{k_0} N_{k_0}$, A_{k_0} is dilated by a factor $d_{k_0}^{-1}$ times a shifted subset of A, $d \mid d_{k_0}$, and they satisfy the following bounds

$$\alpha_k \geq (1 + c_9)^{k-1}\alpha,$$

$$N'_{k_0} \gg \alpha_{k_0} N_{k_0} \gg \frac{c_8^{8(k_0-1)} \alpha^{32k_0^2+2} N}{d^{8k_0} (\log N)^{8k_0}},$$

$$d_{k_0} \leq c_7^{k_0-1} \alpha^{-3(k_0-1)} d.$$

Here the constants with subscripts are the same ones as Lemma 2.4.1.

Instead of using the major arc estimate $|\widehat{F_{N'_{k_0}, d_{k_0}}}(0)| \gg N'_{k_0}/d_{k_0}$, where the factor $1/d_{k_0}$ reflects the impact of the exceptional zero on the distribution of primes in arithmetic progressions, we apply the unexceptional case estimate from Proposition 2.3.2 to deduce

$$|\widehat{F_{N'_{k_0}, d_{k_0}}}(0)| \gg N'_{k_0}.$$

Using $1_{A_{k_0}} * 1_{-A_{k_0}}(n) \leq |A_{k_0}| = \alpha_{k_0} N_{k_0}$ and $|F_{N'_{k_0}, d_{k_0}}(n)| \leq \log(N'_{k_0} d_{k_0} + 1)$ for all $n \in \mathbb{Z}$ as before, we have

$$|\mathrm{supp}(1_{A_{k_0}} * 1_{-A_{k_0}} F_{N'_{k_0}, d_{k_0}})| \gg \frac{\alpha_{k_0}^2 N_{k_0} N'_{k_0}}{\alpha_{k_0} N_{k_0} \log(N'_{k_0} d_{k_0})} \gg \frac{\alpha_{k_0} N'_{k_0}}{\log N'_{k_0}}.$$

Let

$$A'_{k_0} := \{n \in A_{k_0} - A_{k_0} : n \leq N'_{k_0}, d_{k_0} n + 1 \text{ is a prime}\}.$$

Since there are at most $O(\sqrt{N'_{k_0} d_{k_0}})$ composite prime powers which are contained in the support above and the density increment yields $\alpha_{k_0} \geq \alpha$, it follows from (5.9) and the lower bound on $|\mathrm{supp}(1_{A_{k_0}} * 1_{-A_{k_0}} F_{N'_{k_0}, d_{k_0}})|$ that

$$|A'_{k_0}| \gg \frac{\alpha}{\log N} N'_{k_0}.$$

Let A' be the pre-image of A'_{k_0} under the affine transformations carried out in the proof above. Since A_{k_0} is dilated by a factor $d_{k_0}^{-1}$ times a shifted subset of A, it follows that A' is contained in an arithmetic progression of length N'_{k_0} with common difference d_{k_0}, and A' consists of numbers which are one less than a prime. Therefore, one can conclude that the assertion of the lemma holds for sufficiently large C_{10}.

□

Notice that the density bound from the lemma above is considerably larger than the one from Lemma 5.2.1. In particular, due to the absence of the exceptional zero, the density bound obtained here does not involve the growth of the common difference

caused by the iterations. We will see later that this is the main reason why a stronger result can be proved under the generalised Riemann hypothesis.

Using the improved density bound from the lemma above, one immediately obtains the corresponding bootstrapping lemma following exactly the same proof as Lemma 5.2.3.

Lemma 5.5.2. *Let N_0 and k be positive integers and let N be a positive integer such that $N \leq N_0$. Let $\alpha \ll (\log N)^{-1}$.*

Suppose N, d, α satisfy (5.9), then for any k-colouring of $(\mathbb{P}-1)\cap[N_0]$ and $j \leq k$, we have the following.

Suppose $A \subseteq [N_0]$ is a monochromatic subset of $\mathbb{P}-1$ contained in an arithmetic progression of length N with common difference d, and has $|A| = \alpha N$. Suppose also $A \cup ((A-A)\cap(\mathbb{P}-1))$ is j-coloured, then

1. *either there exists a monochromatic solution of $x - y = z$ where $x, y, z \in (\mathbb{P} - 1) \cap [N_0]$;*

2. *or there exists a monochromatic set $A'' \subseteq \mathbb{P} - 1$ satisfying the following:*

   (a) *A'' is contained in an arithmetic progression of length N'' and common difference d'', where $N'' \geq \alpha^{C(\log\alpha^{-1})^2} d^{-C\log\alpha^{-1}} (\log N)^{-C\log\alpha^{-1}} N$, $d \mid d''$ and $d'' \leq \alpha^{-C\log\alpha^{-1}} d$;*

   (b) *$|A''| \geq \alpha(C(j-1)\log N)^{-1} N''$;*

   (c) *$A'' \cup ((A''-A'')\cap(\mathbb{P}-1))$ is $(j-1)$-coloured.*

The proof of the partition regularity result turns out to be much less technical, as we do not need to adjust "the level of being exceptional" at each bootstrapping step under the generalised Riemann hypothesis, which leads to the improved density bound. Instead, we shall only need to keep applying Lemma 5.5.2 until we find a monochromatic solution.

Proof of Theorem 5.1.2. Let $N_0 = N$. By the prime number theorem, the largest colour class of $[N_0]\cap(\mathbb{P}-1)$ must have cardinality $\geq N_0(k\log N_0)^{-1}$. Thus, there exists a monochromatic subsets $A_0 \subseteq [N_0]$ such that $|A_0| = \alpha_0 N_0$ where $\alpha_0 = (2k\log N_0)^{-1}$. Let $d_0 = 1$.

We shall prove the result by contradiction. We claim that if the theorem fails to be true, then for each $1 \leq i \leq k$ there exist A_i, d_i, N_i and α_i such that

(GRH.I) A_i is contained in an arithmetic progression of length N_i with common difference d_i, where $N_i \geq \alpha_{i-1}^{C(\log \alpha_{i-1}^{-1})^2} d_{i-1}^{-C\log \alpha_{i-1}^{-1}} (\log N_{i-1})^{-C\log \alpha_{i-1}^{-1}} N_{i-1}$, $d_{i-1} \mid d_i$ and $d_i \leq \alpha_{i-1}^{-C\log \alpha_{i-1}^{-1}} d_{i-1}$;

(GRH.II) $|A_i| = \alpha_i N_i > 1$ where $\alpha_{i-1} \geq \alpha_i \geq \alpha_{i-1}(C(k-i)\log N_{i-1})^{-1}$; and

(GRH.III) $A_i \cup ((A_i - A_i) \cap (\mathbb{P} - 1))$ is $(k-i)$-coloured.

Here and in the rest of the proof, the constant C is the one from Lemma 5.5.2. In particular, since (GRH.II) and (GRH.III) are impossible for $i = k$, we come to a contradiction and so the theorem is true.

The parameters A_0, d_0, N_0, α_0 clearly satisfy the bound (5.9), and so we can apply Lemma 5.5.2 to find A_1, d_1, N_1 and α_1 such that

- A_1 is contained in an arithmetic progression of length N_1 with common difference d_1, where $N_1 \geq \alpha_0^{C(\log \alpha_0^{-1})^2} (\log N_0)^{-C\log \alpha_0^{-1}} N_0$ and $d_1 \leq \alpha_0^{-C\log \alpha_0^{-1}}$;
- $|A_1| = \alpha_1 N_1$ where $\alpha_0 \geq \alpha_1 \geq \alpha_0(C(k-1)\log N_0)^{-1}$; and
- $A_1 \cup ((A_1 - A_1) \cap (\mathbb{P} - 1))$ is $(k-1)$-coloured.

In particular, our choices A_1, d_1, N_1 and α_1 satisfy (GRH.I)-(GRH.III).

Suppose for some $1 \leq j \leq k-1$, we have chosen A_i, d_i, N_i, α_i which satisfy (GRH.I)-(GRH.III) for all $1 \leq i \leq j$. We claim that A_j, d_j, N_j and α_j satisfy the following estimates[6][7]

(GRH.i) $\log \alpha_j^{-1} \leq 3j \log(k \log N_0)$;

(GRH.ii) $\log d_j \leq 9Cj^3(\log(k \log N_0))^2$;

(GRH.iii) $\log N_j \geq \log N_0 - C'j^5(\log(k \log N_0))^3$ where C' is a large absolute constant.

By (GRH.I)-(GRH.III), we have the following iterative bounds for all $1 \leq i \leq j$:

$$\log \alpha_i \geq \log \alpha_{i-1} - \log C - \log(k-i) - \log\log N_0; \tag{5.10}$$

$$\log d_i \leq C(\log \alpha_{i-1}^{-1})^2 + \log d_{i-1}; \tag{5.11}$$

[6]The estimate (GRH.iii) is the main reason why we fail to obtain $r_p(k) \leq \exp(k^{1+\epsilon})$ for some sufficiently small $\epsilon > 0$.

[7]Under the generalised Riemann hypothesis, one may simplify Lemma 2.4.1 and refine some of the bounds used in this proof. However, one still cannot significantly improve the size bound on N_i in (GRH.I), which ultimately leads to the power of j in (GRH.iii).

and

$$\log N_i \geq \log N_{i-1} - C(\log \alpha_{i-1}^{-1})(\log^2 \alpha_{i-1}^{-1} + \log d_{i-1} + \log\log N_0). \tag{5.12}$$

Iterating (5.10) yields

$$\log \alpha_j \geq \log \alpha_0 - 2j \log(k \log N_0) \geq -3j \log(k \log N_0),$$

where we used $k \leq (\log N_0)^c$ for some $k < 1$, $\log \alpha_0 = \log(2k \log N_0)^{-1} + o(1)$ and $\alpha_i \leq \alpha_{i-1}$. This completes the proof of (GRH.i). (GRH.ii) follows from iterating (5.11), $d_0 = 1$ and (GRH.i). To prove (GRH.iii), notice that by iterating (5.12) and using (GRH.i), (GRH.ii) and (GRH.II), we have

$$\begin{aligned}
&\log N_j \\
&\geq \log N_0 \\
&\quad - \sum_{i=0}^{j-1} 3Ci \log(k \log N_0)(9i^2 \log^2(k \log N_0) + 9Ci^3(\log(k \log N_0))^2 + \log\log N_0) \\
&\geq \log N_0 - C' j^5 (\log(k \log N_0))^3.
\end{aligned}$$

Since $k \leq (\log N_0)^c$ for some sufficiently small constant c, it follows from (GRH.i) that

$$\log \alpha_j^{-1} \leq 3k \log(k \log N_0) \leq (\log N_0)^{2c} = o((\log N_0)^{1/12}).$$

Similarly, it follows from (GRH.ii) and $k \leq (\log N_0)^c$ that

$$\log d_j \leq 9Ck^3(\log(k \log N_0))^2 = o((\log N_0)^{1/4}).$$

On the other hand, by (GRH.iii) and $k \leq (\log N_0)^c$ we have

$$\log N_j \geq \log N_0 - C'' k^5 (\log(k \log N))^3 \geq 2^{-1} \log N_0.$$

Thus,

$$\log d_j + (\log \alpha_j^{-1})^3 \leq C_9^{-1} \sqrt{\log N_j}$$

where C_9 is the constant from Lemma 5.5.2. Therefore, all hypotheses of Lemma 5.5.2 hold, and so it follows from Lemma 5.5.2 that we can find $A_{j+1}, d_{j+1}, N_{j+1}$ and α_{j+1} which satisfy (GRH.I)-(GRH.III). □

Bibliography

[1] A. Balog, J. Pelikan, J. Pintz, and E. Szemerédi, *Difference sets without kth powers*, Acta. Math. Hungar. **65** (2) (1994), 165–187.

[2] M. Beiglböck, V. Bergelson, N. Hindman and D. Strauss, *Multiplicative structures in additively large sets*, J. Combin. Theory Ser. A **113** (2006), no. 7, 1219–1242.

[3] T. F. Bloom and J. Maynard, *A new upper bound for sets with no square differences*, arXiv preprint available at https://arxiv.org/abs/2011.13266.

[4] T. F. Bloom and O. Sisask, *Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions*, arXiv preprint available at https://arxiv.org/abs/2007.03528.

[5] B. E. Brown and D. M. Gordon, *On sequences without geometric progressions*, Math. Comp. **65** (1996), no. 216, 1749–1754.

[6] D. Conlon, *Combinatorial theorems relative to a random set*, Proceedings of the International Congress of Mathematicians **4** (2014), 303–328.

[7] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Annals of Mathematics **184** (2016), No. 2, 367–454.

[8] E. Croot, N. Lyall and A. Rice, *Polynomials and primes in generalized arithmetic progressions*, International Mathematics Research Notices **2015** (2015), 6021–6043.

[9] W. Deuber, *Partitionen und lineare Gleichungssysteme*, Math. Z. **133** (1973), 109–123.

[10] P. Erdős and A. Sárközy, *On differences and sums of integers II*, Bull. Greek Math. Soc. **18** (1977), 204–223.

[11] G. Exoo, *A lower bound for Schur numbers and multicolor Ramsey numbers of* K_3, Electron. J. Combin. **1** (1994), Research Paper 8.

[12] H. Fredricksen and M. M. Sweet, *Symmetric sum-free partitions and lower bounds for Schur numbers*, Electron. J. Combin. **7** (2000), Research Paper 32.

[13] E. Friedgut, V. Rödl and M. Schacht, *Ramsey properties of random discrete structures*, Random Structures and Algorithms **37** (2010), 407–436.

[14] R. Graham, *Some of my favorite problems in Ramsey theory.* Combinatorial number theory, 229–236, de Gruyter, Berlin, 2007.

[15] R. Graham, V. Rödl, and A. Ruciński, *On Schur properties of random subsets of integers*, J. Number Theory **61** (1996), 388–408.

[16] A. Granville, *On positive integers* $\leq x$ *with prime factors* $\leq t \log x$, Number theory and applications (Banff, AB, 1988), edited by R. A. Mollin, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **265**, Kluwer, Dordrecht, 1989.

[17] B. Green. *Finite field models in additive combinatorics.* In Bridget S Webb, editor, Surveys in combinatorics 2005, pages 1–27. Cambridge Univ. Press, Cambridge, Cambridge, 2005.

[18] B. Green, *Roth's theorem in the primes*, Annals of Mathematics **161** (2005), 1609–1636.

[19] B. Green and T. Tao, *Restriction theory of the Selberg sieve, with applications*, Journal de théorie des nombres de Bordeaux **18** (2006), 147–182.

[20] B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. **171** (2010) 1753–1850.

[21] B. Green and T. Tao, *New bounds for Szemerédi's theorem, III: A polylogarithmic bound for* $r_4(N)$, Mathematika **63** (2017), 944–1040.

[22] H. Helfgott and A. de Roton, *Improving Roth's theorem in the primes.* Int. Math. Res. Not. **2011** (2011), 767–783.

[23] M. Heule, *Schur number five*, Proceedings of the AAAI Conference on Artificial Intelligence **32** (2018), 6598–6606.

[24] A. Hildebrand and G. Tenenbaum, *On integers free of large prime factors*, Trans. Amer. Math. Soc. **296** (1986), 265–290.

[25] A. Ivić and G. Tenenbaum, *Local densities over integers free of large prime factors*, Quart. J. Math. Oxford **37** (1986), 401–417.

[26] R. W. Irving, *An extension of Schur's theorem on sum-free partitions*, Acta Arithmetica **25** (1973), 55–64.

[27] H. Iwaniec and E. Kowalski, *Analytic number theory*, Colloquium Publications, **53**. American Mathematical Society, Providence, RI, 2004.

[28] T. H. Lê, *Partition regularity and the primes*, C. R. Math. Acad. Sci. Paris **350** (2012), 439–441.

[29] T. Lê, *Problems and results on intersective sets*, Combinatorial and Additive Number Theory. Springer Proceedings in Mathematics and Statistics, vol 101. Springer, New York, NY, 2014.

[30] H. Li and H. Pan, *A Schur-type addition theorem for primes*, J. Number Theory **132** (2012), 117–126.

[31] J. Lucier, *Difference sets and shifted primes*, Acta. Math. Hungar. **120** (2008), 79–102.

[32] E. Naslund, *On improving Roth's theorem in the primes*, Mathematika **61** (2015), 49–62.

[33] N. McNew, *On sets of integers which contain no three terms in geometric progression*, Math. Comp. **84** (2015), 2893–2910.

[34] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory I. Classical theory*, Cambridge University Press, 2007.

[35] M. B. Nathanson and K. O'Bryant, *On sequences without geometric progressions*, Integers **13** (2013), Paper No. A73, 5.

[36] J. Pintz, W. L. Steiger and E. Szemerédi, *On sets of natural numbers whose difference set contains no squares.* J. London Math. Soc. (2) **37** (1988), 219–231.

[37] S. Prendiville, *Four variants of the Fourier analytic transference principle*, On line J. Anal. Comb. **12** (2017), 25 pp.

[38] R. A. Rankin, *Sets of integers containing not more than a given number of terms in arithmetical progression*, Proc. Roy. Soc. Edinburgh Sect. A **65** (1961), 332–344.

[39] A. Rice, *Sárközy's theorem for $\mathcal{P}$-intersective polynomials*, Acta. Arith. **157** (2013), 69–89.

[40] A. Rice, *A maximal extension of the best-known bounds for the Furstenberg–Sárközy theorem*, Acta Arith. **187** (2019), 1–41.

[41] J. Riddell, *Sets of integers containing no n terms in geometric progression*, Glasgow Math. J. **10** (1969), 137–146.

[42] K. F. Roth, *Sur quelques ensembles d'entiers*, C. R. Acad. Sci. Paris **234** (1952), 388–390.

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

[44] W. Rudin, *Fourier analysis on groups*. Wiley Classics Library. John Wiley and Sons Inc., New York, 1990. Reprint of the 1962 original, A Wiley–Interscience Publication.

[45] I. Z. Ruzsa, *Uniform distribution, positive trigonometric polynomials and difference sets*, Seminar on Number Theory, 1981/1982, University Bordeaux I, Talence (1982), Exp. No. 18.

[46] I. Z. Ruzsa, *On measures on intersectivity*, Acta Math. Hungar. **43** (1984), 335–340.

[47] I. Z. Ruzsa and T. Sanders, *Difference sets and the primes*, Acta Arith. **131** (2008), 281–301.

[48] T. Sanders, *On the Bogolyubov–Ruzsa lemma*, Analysis and PDE **5** (2012), 627–655.

[49] A. Sárközy, *On difference sets of sequences of integers. I*, Acta Math. Acad. Sci. Hungar. **31** (1978), 125–149.

[50] A. Sárközy, *On difference sets of sequences of integers. II*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **21** (1978), 45–53.

[51] A. Sárközy, *On difference sets of sequences of integers. III*, Acta Math. Acad. Sci. Hungar. **31** (1978), 355–386.

[52] I. Schur, *Über die Kongruenz* $x^m + y^m \equiv z^m \pmod{p}$, Jahresber. Dtsch. Math. **25** (1916), 114–117.

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

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

[55] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge University Press, 2006.

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

[57] R. Wang, *On a theorem of Sárközy for difference sets and shifted primes*, Journal of Number Theory **221** (2020), 220–234.