

SOME ADDITIVE PROBLEMS ON PRIMES, PRIME SQUARES AND CHEN PRIMES

By
KUMMARI MALLESHAM
MATH08201104004

Harish-Chandra Research Institute, Allahabad

*A thesis submitted to the
Board of Studies in Mathematical Sciences*

*In partial fulfilment of requirements
for the Degree of*

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



May, 2018

Homi Bhabha National Institute¹

Recommendations of the Viva Voce Committee

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Shri Kummari Mallesham entitled "Some Additive Problems on Primes, Prime Squares and Chen Primes" and recommend that it may be accepted as fulfilling the thesis requirement for the award of Degree of Doctor of Philosophy.

Chairman – Prof. B. Ramakrishnan



Date: 06/09/2018

Guide / Convener – Prof. D. Surya Ramana



Date: 06/09/2018

Co-guide – Prof. Gyan Prakash



Date: 06/09/2018

Examiner - Prof. Stephan Baier



Date: 06/09/2018

Member 1- Prof. Kalyan Chakraborty



Date: 06/09/2018

Member 2 - Prof. R. Thangadurai



Date: 06/09/2018

Member 3 - Prof. P. K. Ratnakumar



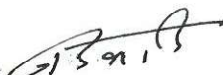
Date: 06/09/2018

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to HBNI.


I/We hereby certify that I/we have read this thesis prepared under my/our direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 06/09/2018

Place: Allahabad



Prof. Gyan Prakash
Co-guide



Prof. D. Surya Ramana
Guide

¹ This page is to be included only for final submission after successful completion of viva voce.

I would like to thank Jahnvi madam, Aarti madam, Prachi and Mannu for treating me as a family member. I am indebted to my parents, Eshwaramma and Ramulu, my brother, Prabhu and my sister, Mangamma, for their invaluable love and affection.

Kummari Mallesham

I express my thanks to Prof. Shigeru Kanemitsu for insightful discussions and for his exciting lectures on Kuznetsov trace formula and to Prof. Yuri Valentinovich Nesterenko for his course on transcendental number theory. My special thanks to Sary Drappeau for his time and discussions which changed my view of mathematics.

I thank all the staff of the Administration of HRI for making my stay comfortable. I am especially thankful to the staff of the library who were helpful at all times. I am thankful to HRI mess, pantry and guest house staff. I thank my school teachers and college lecturers for their support and encouragement, especially N. Ratnakar and Ravinder Reddy who taught me mathematics at school and college level for their help in various ways.

I am thanking my friend Rahul Kumar Singh for his encouragement, moral support and wonderful discussions on mathematics. I am thankful to my friends Asutosh, Balesh, Bibek, Pallab, Debika, Pramod, Bhuwanesh, Manikandan, Aravind, Mithun, Ritika, Pradeep, Sumana, Nabin, Manish, Anoop, Anup, Veekesh, Soumyarup, Tushar, Kasi, Akhilesh, Jay, Manna, Divyang, Prem, Sheshadri, Murali, Umamaheshwaran, Ramdin, Rameez, Parul, Souvik, Sneh, Samrat, Arjit, Ajanta, Krishna mohan tripati, Uttam, Udit kannu, Joshi, Juyal, Eshita, Senthil, Sudhir, Manish gupta, Rajesh, Subbu, Anupam Singh, Varsha, Disha, Aparajita and Anupam Gumber who made my stay at HRI memorable. Also I would like to thank my MSc and BSc friends, especially Mitropam, Arghya, Yellappa, Praveen, Ramana, Jogesh and Srinivas.

Acknowledgments

First and foremost, I express my humble gratitude to my advisor Prof. D. S. Ramana and co-advisor Gyan Prakash for their support throughout my Ph.D. I would like to thank Gyan Prakash for the nice lectures on Additive combinatorics. I would like to thank Surya Ramana for discussing good mathematical ideas with me over walks on the nice road of HRI. He also clarified many of my questions and doubts and suggested nice references to look at which helped me a lot. I thank him for his suggestions and guidance on academics as well as on personal life. I appreciate his patience and composure. Finally, I heartfully thank Gyan Prakash and Surya Ramana for bearing with me for several years.

I thank the members of my doctoral committee, Prof. B. Ramakrishnan, Prof. R. Thangadurai, Prof. P. K. Ratnakumar and Prof. Kalyan Chakraborty, for their helpful comments and encouragement throughout my work on this thesis. I would also like to thank all faculty members of HRI for the courses they gave during my years of study here. My special thanks are to Prof. Rukmini Dey, one of the nicest persons I have ever met, for the inspiration she gave to me, Prof. N. Raghavendra for his inspiring personality and Prof. S. D. Adhikari for sharing many motivational stories, philosophy, many interesting anecdotes about mathematicians and of course about mathematics.

I would like to thank Prof. Joseph Oesterlé for the course he gave at HRI on introduction to Stark's conjectures and for clarifying all my doubts with a lot of patience.

To

My Parents
and My Teachers

What you get by achieving your goals is not as important as what you become by achieving your goals. —Henry David Thoreau

List of Publications arising from the thesis

Journal

1. “Prime in sumsets”, Kummari Mallesham, *Arch. Math.*, **2018**, Vol. 110, 131-143.

Preprints

1. “An improved bound for the additive energy of dense sets of prime numbers”, Kummari Mallesham, Gyan Prakash and D.S. Ramana.
2. “On monochromatic representations of sums of squares of primes”, Kummari Mallesham, Gyan Prakash and D.S. Ramana.
3. “On linear patterns of complexity one in sifted sets”, Kummari Mallesham, Gyan Prakash.

Conferences

1. International Conference on Number Theory (9th to 13th January, 2017, KSOM, Kerala (Topic: Primes in Sumsets)
2. International Conference on Class Groups of Number Fields and Related Topics (ICCGNFRT), HRI, Allahabad, 4th -7th September, 2017. (Topic: Primes in Sumsets)

Kummari Mallesham

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

Kummari Mallesham

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Kummari Mallesham

5.1	Introduction	65
5.2	The finite problem	69
5.3	An application of the Circle method	71
5.3.1	The minor arc contribution	74
5.3.2	The function ψ on a major arc	77
5.3.3	The major arc contribution	84
5.3.4	Proof of Theorem 5.1.2 completed	87
5.4	Monochromatic representation	88
5.5	Optimality	89
6	Linear patterns of complexity one in	
	Chen primes	91
6.1	Introduction	91
6.2	Preliminaries	97
6.3	Inverse theorem for local Gowers U^2 -norm	104
6.4	Translation invariant equations in integers	108
6.5	Translation invariant equations in sifted sets	121
6.5.1	Reduction to a W -tricked set	122
6.5.2	Controlling Ψ -patterns of complexity one	124
6.5.3	Construction of a pseudorandom majorant	130
6.5.4	Transference principle	145
6.5.5	End of the proof of Theorem 6.5.1	157
	Bibliography	159

2.3.2	Proof of Theorem 2.3.1	19
2.4	An optimization principle	21
3	Primes in Sumsets	25
3.1	Introduction	25
3.2	Proof of the Theorem 3.1.2	28
3.3	Preliminaries	29
3.3.1	The large sieve inequality	29
3.3.2	The Brun-Titchmarsh inequality	30
3.3.3	A decomposition of the Von Mangoldt function Λ	30
3.3.4	An application of Davenport's bound	31
3.3.5	An arithmetical function	32
3.4	Reduction to well distributed subsets	33
3.5	Proof of the Theorem 3.1.3	38
3.6	Optimality	45
4	Additive energy of dense subsets of the primes	47
4.1	Introduction	47
4.2	The finite problem	53
4.3	Proof of the main theorem	54
4.4	Optimality	62
4.5	Lower Density of Sumset of Sets of Primes	64
5	Monochromatic sums of squares of primes	65

Contents

List of Key Notations	v
Synopsis	vii
1 Primes in sumsets	viii
2 Additive energy of dense subsets of the primes	x
3 Monochromatic sums of squares of primes	xii
4 Linear patterns of complexity one in Chen primes	xiii
1 Introduction	1
2 The Local Problem and an Optimization Principle	5
2.1 Introduction	5
2.2 Local Problem for invertible elements	8
2.3 Local Problem for squares	16
2.3.1 A Sum in $\mathbf{Z}/p\mathbf{Z}$	17

Theorem 4.1. *Let V be a translation invariant matrix of order $r \times m$ with entries in \mathbf{Z} , of rank r and complexity one. Then there exists a positive constant $C > 0$ depending at most on r, m and V such that for any $\mathcal{A} \subseteq \mathcal{P} \cap [1, N]$ there exists non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{A}^m$ if*

$$|\mathcal{A}| \geq C(\log \log N)^{-\frac{1}{25m}} \pi(N). \tag{7}$$

ments asserting the existence of a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$, where A is a given subset of \mathbf{N} . For instance, let $m \geq 3$ be an integer and let $V = (a_{i,j})$ be the matrix of order $(m-2) \times m$ be defined by $a_{i,j} = 0$ if $j \notin \{i, i+1, i+2\}$, $a_{i,j} = 1$ if $j \in \{i, i+2\}$ and $a_{i,i+1} = -2$. Then V is translation invariant and for any x in \mathbf{Z}^m with $x = (x_1, x_2, \dots, x_m)$, the relation $Vx^t = 0$ is the equivalent to the assertion that x_1, x_2, \dots, x_m are in arithmetical progression. Thus Szemerédi's theorem [35] is equivalent to the statement that for any $A \subseteq \mathbf{N}$ of positive upper density, there is a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$ and the theorem of Roth [29] is the special case when $m = 3$. Similarly, the celebrated result of Green and Tao [11] is equivalent to the statement that for any subset \mathcal{A} of the set of primes \mathcal{P} with positive relative upper density in \mathcal{P} , there is a non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{A}^m$.

Let V of order $r \times m$ with entries in \mathbf{Z} be translation invariant matrix such that the kernel of the associated linear map $x \mapsto Vx$ from \mathbf{Q}^m to \mathbf{Q}^r has dimension at least 2. Then it can in turn be deduced from Szemerédi's theorem that for any subset A of \mathbf{N} of positive upper density there is a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$. One may naturally ask for quantitative versions of this result, but these are rather difficult to obtain in general and are tied up with an appropriate notion of “complexity” of V . The matrix V considered in the preceding paragraph has complexity one precisely when $m = 3$, that is, in the setting of Roth's theorem on three term arithmetical progressions. For arbitrary V of complexity one, Kevin Henriot [15] has obtained the following theorem.

Ramaré [24] showed that $r_{\mathcal{D}}(K) \ll K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$ when \mathcal{D} is the set of squares. By a lower bound in [13], the optimal bound in this case is expected to have a $\log 2$ in place of the $3 \log 2$ in the exponential factor.

A classical theorem of L.K. Hua asserts that the set of squares of primes is an asymptotic basis of finite order. In light of this theorem, F. Hennecart asked if one may extend Sárközy's problems to the case when \mathcal{D} is the set of squares of the prime numbers. Independently of Hennecart, this question was considered by Guohua Chen [4], who showed that $r_{\mathcal{D}}(K) \ll_{\epsilon} K^{2+\epsilon}$. The main result of the third part of this thesis improves on this conclusion of Chen. More precisely, and in analogy with the result of [24] for the squares, we prove the following theorem.

Theorem 3.1. *For any integer $K \geq 2$ we have $r_{\mathcal{D}}(K) \leq K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$ when \mathcal{D} is the set of squares of the prime numbers.*

We prove Theorem 3.1 by an adaption of the method of Gyan Prakash, Ramana and Ramaré [24] for the case of the squares.

4 Linear patterns of complexity one in Chen primes

A matrix V of order $r \times m$ with entries in \mathbf{Z} is said to be translation invariant if $Ve^t = 0$, where e is the vector $(1, 1, \dots, 1)$ in \mathbf{Q}^m and e^t its transpose. By a trivial solution of the system of linear equations $Vx^t = 0$ for such V we mean $x = \lambda e$, for some $\lambda \in \mathbf{Q}$.

A number of celebrated results in additive combinatorics can be recast as a state-

3 Monochromatic sums of squares of primes

A subset \mathcal{D} of the set of natural numbers is said to be an asymptotic basis of finite order if there exists a positive integer m such that every sufficiently large integer can be written as a sum of at most m elements of \mathcal{D} . The smallest m for which the above property holds is called the order of the asymptotic basis \mathcal{D} . There are two classical examples of asymptotic bases of finite order. The first of these is the set of squares of the natural numbers, which by Lagrange's four squares theorem, is certainly an asymptotic basis of order four. The second example is the set of the prime numbers, which is seen to be an asymptotic basis of order at most 4 by the classical theorem of Vinogradov which asserts that every sufficiently large odd integer can be written as a sum of three prime numbers.

Given an asymptotic basis \mathcal{D} of finite order and an integer $K \geq 1$ one may ask the following question. What is the smallest integer $r_{\mathcal{D}}(K)$, if it exists, such that given any colouring, or partition, of \mathcal{D} in K colours, every sufficiently large integer is expressible as a sum of at most $r_{\mathcal{D}}(K)$ elements of \mathcal{D} , all of the same colour ?

When \mathcal{D} is either the set of squares or the set of primes the aforementioned question was posed as Problems 39 and 40 by A. Sárközy on page 26 of [31]. It is easily seen that indeed $r_{\mathcal{D}}(K)$ is finite for all $K \geq 1$ in these cases. In [13] N. Hegyváry and F. Hennecart showed that $r_{\mathcal{D}}(K) \ll (K \log K)^5$ when \mathcal{D} is the set of squares and $r_{\mathcal{D}}(K) \ll K^3$ when \mathcal{D} is the set of primes. Ramana and Ramaré [25] then obtained $r_{\mathcal{D}}(K) \ll K \log \log 2K$ when \mathcal{D} is the set of primes, which is the best possible bound up to the implied constant. Recently, Gyan Prakash, Ramana and

for $N \geq N(\alpha)$. The main result of the second part of this thesis is the following theorem which generalizes and improves on (4).

Theorem 2.1. *Let $\alpha \in (0, 1]$. Then there is an integer $N(\alpha)$ such that for all $N \geq N(\alpha)$ and subsets A of the prime numbers in $[1, N]$ satisfying $|A| \geq \alpha\pi(N)$ we have*

$$E(A, B) \leq (1 + o(\alpha)) e^\gamma \frac{|A|^2 |B|}{\log m(B)} \log \log \frac{4}{\alpha} \quad (5)$$

for any non-empty subset B of the prime numbers in $[1, N]$, where $m(B) = \min_{b \in B} b$.

We prove this theorem by a refinement of the method of [25], which in turn stems from the method of O. Ramaré and I. Ruzsa [26]. Our theorem is optimal in the sense that $1 + o(\alpha)$ cannot be replaced with $c + o(\alpha)$ for any $c < 1$. As a corollary of our theorem we obtain the “main theorem” of K. Matomäki [20].

Corollary 2.1. *Let \mathcal{A} and \mathcal{B} be subsets of the set of prime numbers \mathcal{P} with relative lower densities α and β respectively in \mathcal{P} . Then*

$$\liminf_{N \rightarrow \infty} \frac{(\mathcal{A} + \mathcal{B})(N)}{N} \geq (1 - o(\alpha + \beta)) \frac{\beta}{e^\gamma \log \log(\frac{1}{\alpha})}. \quad (6)$$

Matomäki proves this result by using the methods of B.J. Green, and B.J. Green and T. Tao to directly obtain a lower bound for the left hand side of (6), without recourse to additive energy. Our method via the bound (5) appears to be simpler.

Ramana and O. Ramaré [25]. This method was originally used to obtain an upper bound for additive energy of “dense” subsets of primes, the subject of the second part of the thesis, which we now outline.

2 Additive energy of dense subsets of the primes

The additive energy of subsets A, B of the integers, denoted by $E(A, B)$, is defined by

$$E(A, B) = |\{(x_1, x_2, y_1, y_2) \in A \times A \times B \times B : x_1 + y_1 = x_2 + y_2\}|. \quad (2)$$

The additive energy $E(A, B)$ of A and B is an important quantity in additive combinatorics. It is related to $|A + B|$, the cardinality of the sumset $A + B$, by

$$|A + B| E(A, B) \geq |A|^2 |B|^2. \quad (3)$$

This follows from an application of the Cauchy-Schwarz inequality. Thus, upper bounds on the additive energy $E(A, B)$ translate to lower bounds on $|A + B|$. The converse, however, is not true.

In a paper [25] on a question of A. Sárközy, which is described in Section 3 below, Ramana and Ramaré showed that for any $\alpha \in (0, 1)$ and $A \subset (N, 2N] \cap \mathcal{P}$ with $|A| \geq \alpha |\mathcal{P} \cap (N, 2N]|$, we have

$$E(A, A) \leq (2 + o(\alpha)) e^\gamma \frac{|A|^3}{\log N} \log \log \frac{4}{\alpha} \quad (4)$$

Using the linear sieve they showed that

$$P_{N;A,B} \ll \frac{|A||B|}{\log N} R \tag{1}$$

where $R = \frac{1000N}{|A|^{1/2}|B|^{1/2}}$. Our contribution to the aforementioned question of Sárközy et al., detailed in Chapter 3 of the thesis, begins with observation that the bound (1) can be obtained by a simple counting argument and the Chebyshev upper bound. Note that (1) is implied by the trivial estimate $P_{N;A,B} \leq |A||B|$ unless $|A||B| \gg \frac{N^2}{(\log N)^2}$. We then improve (1) under this assumption on $|A||B|$ by proving the following theorem.

Theorem 1.1. *Let A, B be subsets of $\{1, \dots, N\}$, where $N \geq 1$ is an integer and suppose further that $|A||B| \gg \frac{N^2}{(\log N)^2}$. Then we have that $P_{N;A,B} \ll \frac{|A||B|}{\log N} \log \log R$.*

This upper bound for $P_{N;A,B}$ is optimal, up to the implied constant, in general. In fact, we have the following proposition, which corrects the conclusion of Example 2, page 36 of [3].

Proposition 1.1. *Let $N \geq 1$ and $k \ll \log \log N$ be integers and let $m_k = \prod_{p \leq k} p$. Then if $A = \{1 \leq a \leq N : a \equiv 0 \pmod{m_k}\}$ and $B = \{1 \leq b \leq N : b \equiv 1 \pmod{m_k}\}$ we have that*

$$P_{N;A,B} \gg \frac{|A||B|}{\log N} \log \log R.$$

As we show in Chapter 3 of the thesis, this proposition can be easily deduced from the Siegel-Walfisz theorem. For the proof of Theorem 1.1, we first reduce to the case when A and B are “well-distributed” subsets and then apply the method of D.S.

1 Primes in sumsets

Let S be an infinite subset of the natural numbers \mathbf{N} . An asymptotic additive decomposition of S is a pair of subsets (A, B) of \mathbf{N} with $|A|, |B| \geq 2$ such that all large enough elements of S can be written as $a + b$ for some $(a, b) \in A \times B$. When such an asymptotic decomposition of S exists, S is said to be asymptotically additively decomposable. A famous conjecture of Ostmann (see page 13 of [23]), often called the inverse Goldbach conjecture, asserts that the set of prime numbers \mathcal{P} is not asymptotically additively decomposable. The following result of C. Elsholtz (see page 1 of [7]) gives a necessary condition.

Theorem 1.1. *Suppose that (A, B) is an asymptotic additive decomposition of \mathcal{P} then we have that*

$$\frac{x^{1/2}}{(\log x)^5} \ll A(x), B(x) \ll x^{1/2}(\log x)^4,$$

where $A(x), B(x)$ are the counting functions of the sets A, B respectively.

This result reduces Ostmann's conjecture to the following :

Conjecture 1.1. *For any $\delta > 0$ and all subsets A, B of $\{1, 2, \dots, N\}$ with $|A|, |B| \geq N^\delta$, the sumset $A + B$ contains a composite number when N is a sufficiently large integer.*

In [3, Section 5], A. Sárközy, A. Balog and J. Rivat pose a question which goes in the opposite direction to the above conjecture 1.1. More precisely, given subsets A, B of $\{1, \dots, N\}$, where $N \geq 1$ is an integer, they ask for optimal upper bounds for the number $P_{N;A,B}$ of pairs (a, b) in $A \times B$ such that $a + b$ is a prime number.

SYNOPSIS

This thesis contains four principal chapters. The first of these, Chapter 3, deals with a question of A. Sárközy, A. Balog and J. Rivat which asks for optimal upper bounds for the number of pairs (a, b) of integers in $A \times B$ whose sum is a prime number, where A and B are subsets of $\{1, 2, \dots, N\}$ and N is a large enough integer. Our contribution to this question is described in Section 1 below. Chapter 4 gives an asymptotically sharp bound for the additive energy of a pair of “dense” sets of prime numbers in $[1, N]$. This bound allows us to recover a theorem of K. Matomäki that gives an optimal lower bound for the density of the sumset of a pair of subsets of the primes with given relative lower densities in the set of primes. Matomäki’s uses the methods of B. Green, B. Green and T. Tao, whereas our method is essentially simpler and stems from a method of D.S. Ramana and O. Ramaré, which is a variant of the original method of I. Ruzsa and O. Ramaré [26]. The contents of this part of the thesis are discussed in Section 2 of this synopsis. In Chapter 5 of this thesis we give a chromatic version of the well-known theorem of L.K. Hua on the representation of integers as a sum of squares of prime numbers. A precise statement of our result is in Section 3 of this synopsis. In the sixth and final chapter of this thesis we obtain a slightly improved version of a result of K. Henriot on the existence of non-trivial solutions in the prime numbers to a translation invariant system of linear equations over \mathbf{Z} of “complexity one”. Further, we generalise this result to cover Chen primes in particular. Here by a Chen prime is meant a prime number p such that $p + 2$ has at most two prime factors and each of these factors is at least $p^{\frac{3}{11}}$. The contents of this chapter are discussed in Section 4 below.

List of Key Notations

\mathbf{R}	The set of real numbers.
\mathbf{Q}	The set of rational numbers
\mathbf{Z}	The set of integers.
\mathbf{N}	The set of natural numbers
\mathcal{P}	The set of prime numbers
$e(\theta)$	$e^{2\pi i\theta}$
$\mathbf{Z}/M\mathbf{Z}$	The additive group of integers modulo M
$(\mathbf{Z}/M\mathbf{Z})^*$	The multiplicative group of integers modulo M
$\pi^*(N)$	The number of primes in $(N/2, N]$
$\nu(n)$	The number of prime divisors of an integer n
$\phi(n)$	The Euler's totient function
$\Lambda(n)$	The Von Mangoldt function
$\mu(n)$	The Möbius function
$\tau(n)$	The number of divisors of an integer n
$M_{r \times t}(\mathbf{Z})$	The set of $r \times t$ matrices with integer coefficients
$\hat{f}(t)$	$\int_{\mathbf{R}} f(x)e^{-2\pi ixt} dx$
A^r	r -times Cartesian product of A
$\mathbf{E}_{n \in X} f(n)$	$\frac{1}{ X } \sum_{n \in X} f(n)$

method appears to be substantially simpler.

Chapter 5 of this thesis is also based on joint work with Gyan Prakash and D.S. Ramana. The main result of this chapter is Theorem 5.1.1. This theorem gives a close to optimal upper bound for the smallest integer $r(K)$ such that given any colouring (or partition) of the set of squares of primes \mathcal{D} in K colours, every sufficiently large integer is expressible as a sum of at most $r(K)$ elements of \mathcal{D} , all of the same colour. Indeed, a pair of problems proposed by A. Sárközy in [31] ask for analogous results when \mathcal{D} is replaced by the set of integral squares and then by the set of prime numbers. These problems were given nearly optimal solutions by Gyan Prakash, Ramana and Ramaré [24] and by Ramana and Ramaré [25] respectively. The problem for the squares of prime numbers was orally posed to us by F. Hennecart and was independently considered by G. Chen in [4]. Our Theorem 5.1.1 improves on Chen's result, which is $r(K) \ll_{\epsilon} K^{2+\epsilon}$ and was hithertofore the best known bound on $r(K)$.

Chapter 6 is based on the joint work with Gyan Prakash. In [10] Green and Tao showed that there are infinitely many non-trivial arithmetic progression of length three in the set of Chen prime numbers. Such arithmetic progressions are a particular cases of a linear patterns of complexity one. In Chapter 6 we study more generally linear patterns of “complexity” one in sifted sequences. These sequences include the Chen prime numbers and many other arithmetically interesting sequences. Our main result here, Theorem 6.1.2, generalizes the theorem of Green and Tao [10, Theorem 1.2] and improves on the result of Henriot [15, Theorem 2].

This thesis separates into two parts. In studying the problems of the first part, carried out in Chapters 3 through 5, we rely a common strategy that goes back to a work of I. Ruzsa and O. Ramaré [26]. The principle here is to reduce to the problems considered to what we have called *local problems*, treated in Chapter 2. The final Chapter 6 forms the second part of the thesis. Here we expand on the methods of K. Henriot [15] and Shao [33], stemming from the work of B. J. Green and T. Tao [10].

Chapter 3 of this thesis is based on our paper [19], in which we have partially answered the following question of A. Sárközy, A. Balog and J. Rivat [3]. Given subsets A, B of $\{1, \dots, N\}$, where $N \geq 1$ is an integer, determine optimal upper bounds for the number $P_{N;A,B}$ of pairs (a, b) in $A \times B$ such that $a + b$ is a prime number. Theorem 3.1.3, which is the main result of Chapter 3, gives an essentially optimal answer to this question under the hypothesis $|A||B| \gg \frac{N^2}{(\log N)^2}$. This result considerably improves on the upper bound for $P_{N;A,B}$ originally obtained in [3], which is non-trivial only under the aforementioned hypothesis.

Chapter 4 of this thesis is based on joint work with Gyan Prakash and D.S. Ramana. Here we obtain an essentially optimal upper bound for the additive energy of dense subsets of primes which is stated as Theorem 4.1.1. Using the classical relation between the additive energy and the cardinality of the sumsets we then recover from our theorem the essentially optimal lower bound for the asymptotic lower density of the sumset of a given pair of subsets of the set primes numbers, originally obtained by K. Matomäki in Theorem 2.1 of [20], using the methods of Green and Tao. Our

CHAPTER 1

Introduction

This thesis is centered around certain additive problems on the set of prime numbers, the set of squares of primes numbers and the set of Chen prime numbers. Here by a Chen prime number we shall always mean a prime number p such that $p + 2$ has at most two prime factors, each of which is at least $p^{3/11}$. Each of these sets is defined by multiplicative conditions and, as has been the experience so far, additive problems on sets of integers that are easily described from the multiplicative point of view tend to be difficult to resolve or are, at any rate, interesting.

The purpose of the present chapter is to give a brief description of the problems considered in Chapters 3 through 6, which are the principal chapters of this thesis. A more detailed introduction their contents is given in the leading section of each of these chapters. The reader may also wish to refer to the Synopsis, on pages vii through xvi, which is essentially a collage of the introductory sections of these chapters of the thesis.

also an extreme point of \mathcal{K}_2 , such that $f(x^*, y^*) = f(x^*, v) = f(u, v)$. In particular, we have $f(x, y) \leq f(u, v) = f(x^*, y^*)$ for all $(x, y) \in \mathcal{K}_1 \times \mathcal{K}_2$.

(ii) Let us suppose that $x^* = (x_1^*, \dots, x_m^*)$ be an extreme point of \mathcal{K}_1 with two of its co-ordinates x_i^* and x_j^* with respect to the canonical basis $\{e_1, \dots, e_m\}$ of \mathbb{R}^m lying in the interval $(0, D_1)$. For a small enough $\delta > 0$, we consider two points $y = x^* - \delta(e_i - e_j)$ and $z = x^* + \delta(e_i - e_j)$ of \mathbb{R}^m . Then we have $y, z \in \mathcal{K}_1$ and $x^* = (y + z)/2$. Thus we get a contradiction to the fact that x^* is an extreme point of \mathcal{K}_1 . Thus, we conclude that if x^* is an extreme point of \mathcal{K}_1 , then, excepting at most one, all co-ordinates of x^* are equal to either 0 or D_1 . Moreover, if l is the number of co-ordinates of x^* that are distinct from 0 then we have, from the condition in the definition of \mathcal{K}_1 , that l must satisfy the inequality $lD_1 \geq P_1 \geq (l-1)D_1$. \square

Chapters 3 through 5 to the local problems described here.

Suppose that $n, m \geq 1$ are integers and let P_1, P_2, D_1 and D_2 be real numbers > 0 .

Further let

$$\mathcal{K}_1 = \left\{ (x_1, \dots, x_m) \in \mathbf{R}^m : \sum_{i=1}^m x_i = P_1, 0 \leq x_i \leq D_1 \text{ for all } i \right\},$$

and

$$\mathcal{K}_2 = \left\{ (x_1, \dots, x_n) \in \mathbf{R}^n : \sum_{i=1}^n x_i = P_2, 0 \leq x_i \leq D_2 \text{ for all } i \right\}.$$

Let us also assume that \mathcal{K}_1 and \mathcal{K}_2 are non-empty sets. Then $\mathcal{K}_1, \mathcal{K}_2$ are compact and convex subsets of $\mathbf{R}^m, \mathbf{R}^n$ respectively. Then we have :

Lemma 2.4.1. *If $f : \mathbf{R}^m \times \mathbf{R}^n \mapsto \mathbf{R}$ a bilinear form with real coefficients α_{ij} defined by $f(x, y) = \sum_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} \alpha_{ij} x_i y_j$ then*

(i) *there are extreme points x^* and y^* of \mathcal{K}_1 and \mathcal{K}_2 respectively so that $f(x, y) \leq f(x^*, y^*)$ for all $x \in \mathcal{K}_1, y \in \mathcal{K}_2$.*

(ii) *If $x^* = (x_1^*, x_2^*, \dots, x_m^*)$ is an extreme point of \mathcal{K}_1 then, excepting at most one i , we have either $x_i^* = 0$ or $x_i^* = D_1$ for each i . Also, if l is the number of i such $x_i^* \neq 0$ then $lD_1 \geq P_1 > (l-1)D_1$. A similar result holds for the extreme points of \mathcal{K}_2 .*

Proof. (i) suppose that $f(x, y)$ attains its maximum on the compact set $\mathcal{K}_1 \times \mathcal{K}_2$ at (u, v) . Then the map $x \mapsto f(x, v)$ is linear and thus attains its maximum on the compact convex set \mathcal{K}_1 at an extreme point of \mathcal{K}_1 , say x^* . We must necessarily have $f(x^*, v) = f(u, v)$. Arguing similarly with the linear map $y \mapsto f(x^*, y)$ we obtain y^* ,

$$(\mathcal{E}(k, t))^{\frac{2}{i^2}} = \left(\prod_{p|U} \mathcal{E}_p(k, t) \right)^{\frac{2}{i^2}} \leq \left(\frac{U}{\phi(U)} \right)^{4/t} \exp \left(8 t^3 2^t \sum_{p \leq A^{25}} \frac{1}{p} \right). \quad (2.52)$$

From (3.20) on page 70 of [28] we deduce that $\sum_{p \leq A^\ell} \frac{1}{p} \leq (\log 50) \log \log A$, since $A \geq 4$. On combining this remark with (2.52), (2.38) and (2.50) we then conclude that for any even integer $t \geq 2$ we have

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \left(\frac{U}{\phi(U)} \right)^2 \frac{|\mathcal{X}| |\mathcal{Y}|}{\tau(U)} \exp \left(\frac{3 \log A}{t} + 8 (\log 50) t^3 2^t \log \log A \right). \quad (2.53)$$

Let us now set $v \log 2 = \log \left(\frac{\log A}{(\log \log A)^6} \right)$ and suppose that $A_0 \geq e^e$ is such that we have $\frac{\log \log A}{\log \log \log A} \geq 12$ and $v \geq 4$ for all $A > A_0$. For such A we take t in (2.53) to be an even integer satisfying $v \leq t \leq v + 2$. Also, with $w = \frac{6 \log \log \log A}{\log \log A}$ we have $w \leq \frac{1}{2}$ and $v = \frac{(1-w) \log \log A}{\log 2}$. Thus $\frac{1}{t} \leq \frac{1}{v} \leq \frac{(\log 2)(1+2w)}{\log \log A}$ and $t^3 2^t \leq 32 v^3 2^v \leq \frac{32 \log A}{(\log 2)^3 (\log \log A)^3}$. Substituting these inequalities in (2.53) we obtain (2.39) for $A > A_0$. To obtain (2.39) for $e^{e^2} \leq A \leq A_0$ it suffices to take $t = 2$ in (2.53). □

2.4 An optimization principle

In this section we state and prove an optimization principle, which is a minor variant on a similar principle from [25], whose proof of this principle we follow. As stated in Section 2.1 this principle plays a key role in reducing the problems considered in

this we see that

$$\sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} \prod_{p|U} \epsilon_p(x, y) \right)^t \leq \sum_{y \in G} \sum_{(x_1, x_2, \dots, x_t) \in \mathcal{X}^t} \prod_{1 \leq i \leq t} \prod_{p|U} \epsilon_p(x_i, y). \quad (2.48)$$

Interchanging the summations over G and \mathcal{X}^t on the right hand side of the above relation and applying Hölder's inequality again, this time to exponent $\frac{t}{2}$, we obtain that the right hand side of (2.48) does not exceed

$$|\mathcal{X}|^{t-2} \left(\sum_{(x_1, x_2, \dots, x_t) \in \mathcal{X}^t} \left(\sum_{y \in G} \prod_{1 \leq i \leq t} \prod_{p|U} \epsilon_p(x_i, y) \right)^{\frac{t}{2}} \right)^{\frac{2}{t}}. \quad (2.49)$$

Finally, on expanding the summand in the sum over \mathcal{X}^t in (2.49) and extending the summation to all of G^t we conclude using (2.48) and (2.47) and a rearrangement of terms that

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \frac{|\mathcal{X}||\mathcal{Y}|}{\tau(U)} \left(\frac{\phi(U)^3}{|\mathcal{X}|^2|\mathcal{Y}|} \right)^{\frac{1}{t}} \mathcal{E} \left(\frac{t}{2}, t \right)^{\frac{2}{t^2}}, \quad (2.50)$$

where for any integer k with $1 \leq k \leq t$ we have set

$$\mathcal{E}(k, t) = \frac{1}{\phi(U)^{2t}} \sum_{(y_1, y_2, \dots, y_t) \in G^t} \sum_{(x_1, x_2, \dots, x_t) \in G^t} \prod_{p|U} \prod_{\substack{1 \leq i \leq t, \\ 1 \leq j \leq k}} \epsilon_p(x_i, y_j). \quad (2.51)$$

The Chinese Remainder Theorem gives $G = \prod_{p|U} (G_p \setminus \{0\})$. Moreover, for all $p|U$ and (x, y) in $(\mathbf{Z}/U\mathbf{Z})^2$ we have $\epsilon_p(x, y) = \epsilon_p(x_p, y_p)$. It follows that $\mathcal{E}(k, t) = \prod_{p|U} \mathcal{E}_p(k, t)$, where $\mathcal{E}_p(k, t)$ is as defined by (2.41). Using (2.42) with $k = \frac{t}{2}$, valid on account of Corollary 2.3.3, and recalling that $U = \prod_{p \leq A^{25}} p$ we then obtain

2.3.2 Proof of Theorem 2.3.1

We shall write G for the set $(\mathbf{Z}/U\mathbf{Z})^*$ and continue to use G_p for $\mathbf{Z}/p\mathbf{Z}$. Also, for any x in $\mathbf{Z}/U\mathbf{Z}$ and $p|U$ we denote the canonical image of x in $\mathbf{Z}/p\mathbf{Z}$ by x_p and, to be consistent with the notation of preceding subsection, write $\lambda_p(x)$ for the Legendre symbol $\left(\frac{x_p}{p}\right)$. Then we have that

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \prod_{p|U} \left(\frac{1 + \lambda_p(x^2 + y^2 + c)}{2} \right), \quad (2.45)$$

since $0 \leq 1 + \lambda_p(x^2 + y^2 + c) \leq 2$ for any pair (x, y) in $\mathcal{X} \times \mathcal{Y}$, with equality in the upper bound for every prime $p|U$ when $x^2 + y^2 + c$ is an invertible square in $\mathbf{Z}/U\mathbf{Z}$. On extending the definitions of δ_p and ϵ_p from Subsection 2.3.1 by setting $\delta_p(x, y) = \lambda_p(x^2 + y^2 + c)$ and $\epsilon_p(x, y) = 1 + \delta_p(x, y)$ for any (x, y) in $(\mathbf{Z}/U\mathbf{Z})^2$ and $p|U$, we may rewrite (2.45) as

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \frac{1}{\tau(U)} \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \prod_{p|U} \epsilon_p(x, y). \quad (2.46)$$

Let $t \geq 2$ be an even integer. Then an interchange of summations followed by an application of Hölder's inequality to exponent t to the right hand side of (2.46) gives

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \frac{|\mathcal{Y}|^{1-\frac{1}{t}}}{\tau(U)} \left(\sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} \prod_{p|U} \epsilon_p(x, y) \right)^t \right)^{\frac{1}{t}}. \quad (2.47)$$

To bound the sum over $y \in \mathcal{Y}$ on the right hand side of the inequality above, we first expand the summand in this sum and extend the summation to all $y \in G$. By

$$\mathcal{E}_p(k, t) = \mathbb{E}_{y_1, y_2, \dots, y_t} \mathbb{E}_{x_1, x_2, \dots, x_t} \prod_{\substack{1 \leq i \leq t, \\ 1 \leq j \leq k.}} \epsilon_p(x_i, y_j). \quad (2.41)$$

Using this notation we state our corollary as follows.

Corollary 2.3.3. *For any even integer $t \geq 2$ we have*

$$\mathcal{E}_p(t/2, t) \leq \left(\frac{p}{p-1} \right)^{2t} \exp\left(\frac{4t^5 2^t}{p} \right). \quad (2.42)$$

Proof. Since $t \geq 2$ is an even integer. By taking $k = t/2$ in (2.7) we get that

$$\mathcal{E}_p(t/2, t) = \mathbb{E}_{y_1, y_2, \dots, y_t} \mathbb{E}_{x_1, x_2, \dots, x_t} \prod_{\substack{1 \leq i \leq t, \\ 1 \leq j \leq t/2.}} \epsilon_p(x_i, y_j). \quad (2.43)$$

Observing that the summands in the sum are independent of the variables y_i for $t/2 < i \leq t$, and allowing the sum over full G_p we get

$$\mathcal{E}_p(t/2, t) \leq \frac{p^{t/2}}{(p-1)^{2t}} \sum_{(y_1, y_2, \dots, y_{t/2}) \in G_p^{t/2}} \sum_{(x_1, x_2, \dots, x_t) \in G_p^t} \prod_{\substack{1 \leq i \leq t, \\ 1 \leq j \leq t/2.}} (1 + \lambda_p(x_i^2 + y_j^2 + c)). \quad (2.44)$$

Substituting the upper bound on the sum in the right of (2.44) by Lemma 2.3.3, we conclude the inequality (2.42) holds. □

As stated in Section 2.1, this theorem is only slightly different from Theorem 2.1 of [24]. For the convenience of the reader we give a full proof, reproducing that in [24] with minor modifications. The preliminaries required for the proof are covered in Subsection 2.3.1 below, while the proof itself is presented in Subsection 2.3.1.

2.3.1 A Sum in $\mathbf{Z}/p\mathbf{Z}$

We write G_p for the ring $\mathbf{Z}/p\mathbf{Z}$ when p is a prime number. Also, $\lambda_p(x)$ shall denote the Legendre symbol $\left(\frac{x}{p}\right)$, for any x in G_p .

Lemma 2.3.2. *Let p be a prime number and c an element of G_p . Then for any an even integer $t \geq 2$ we have*

$$\sum_{(y_1, y_2, \dots, y_{t/2}) \in G_p^{t/2}} \sum_{(x_1, x_2, \dots, x_t) \in G_p^t} \prod_{\substack{1 \leq i \leq t, \\ 1 \leq j \leq t/2}} (1 + \lambda_p(x_i^2 + y_j^2 + c)) \leq p^{3t/2} \exp\left(\frac{4t^5 2^t}{p}\right). \quad (2.40)$$

PROOF.— See the proof of the Proposition 2.2 of [24].

Before stating a corollary of the above lemma, let us introduce some additional notation. Let p be a fixed prime number and let c be a given element of G_p . For any (x, y) in G_p^2 we set $\delta_p(x, y) = \lambda_p(x^2 + y^2 + c)$ and $\epsilon_p(x, y) = 1 + \delta_p(x, y)$. We endow $G_p \setminus \{0\}$, and likewise $(G_p \setminus \{0\})^t$ for any integer $t \geq 1$, with their uniform probability measures and write \mathbb{E}_x and $\mathbb{E}_{x_1, x_2, \dots, x_t}$ respectively in place of $\frac{1}{p-1} \sum_{x \in G_p \setminus \{0\}}$ and $\frac{1}{(p-1)^t} \sum_{x_1, x_2, \dots, x_t \in G_p \setminus \{0\}}$. Finally, we define $\mathcal{E}_p(k, t)$ for any integer k with $1 \leq k \leq t$ by

We bound the right hand side of (2.35) by using (2.28) as

$$|\mathcal{X}||\mathcal{Y}| \exp\left(-\sum_{p \in J} \frac{1}{p^2}\right) \exp\left(2\left(\mathcal{L}(X, Y) \sum_{p \in J} \frac{1}{p^2}\right)^{\frac{1}{2}} + \sum_{p \in J} \frac{1}{p^2}\right), \quad (2.36)$$

where $\mathcal{L}(X, Y) = \log\left(\frac{U^2}{|\mathcal{X}||\mathcal{Y}|}\right) \leq 9 \log R$. We obtain (2.33) by noticing that

$$\exp\left(-\sum_{p \in J} \frac{1}{p}\right) \leq \frac{\phi(P)}{P} \exp\left(\frac{4}{Q^2}\right) \quad \text{and} \quad \sum_{p \in J} \frac{1}{p^2} \leq \frac{2}{Q^2}. \quad (2.37)$$

□

2.3 Local Problem for squares

Let $A \geq e^{e^2}$ be real number and $U = \prod_{p \leq w} p$, where $w = A^{25}$. Suppose further that \mathcal{X} and \mathcal{Y} are subsets of $(\mathbf{Z}/U\mathbf{Z})^*$ of density at least $\frac{1}{A}$. That is,

$$|\mathcal{X}| \text{ and } |\mathcal{Y}| \geq \frac{\phi(U)}{A}. \quad (2.38)$$

For a given element c of $\mathbf{Z}/U\mathbf{Z}$, let $T_c(\mathcal{X}, \mathcal{Y})$ denote the set of pairs $(x, y) \in \mathcal{X} \times \mathcal{Y}$ such that $x^2 + y^2 + c$ is an invertible square in $\mathbf{Z}/U\mathbf{Z}$.

Theorem 2.3.1. *For all $A, U, \mathcal{X}, \mathcal{Y}$ and c as above, we have*

$$|T_c(\mathcal{X}, \mathcal{Y})| \leq \left(\frac{U}{\phi(U)}\right)^2 \frac{|\mathcal{X}||\mathcal{Y}|}{\tau(U)} \exp\left(\frac{\left(3 \log 2 + O\left(\frac{\log \log \log A}{\log \log A}\right)\right) \log A}{\log \log A}\right). \quad (2.39)$$

since $\sum_{p \in J} \frac{1}{p^2} \leq \sum_{n \geq 1} \frac{1}{n^2} \leq 2 \leq \log T^\ell$.

□

Remark 2.2.2. Note that the upper bound given by (2.28) for $\mathcal{R}_U(\mathcal{U}, \mathcal{V}, s)$ is uniform in s .

Before stating next corollary we fix some notation. Let $R \geq 1000$ be a real number and let $U = \prod_{p \leq R} p$, $M_R = (R \log R / \log \log R)^2$ and $Q = \log R \log \log R$. We write $T(\mathcal{X}, \mathcal{Y})$ for the number of pairs (x, y) in $\mathcal{X} \times \mathcal{Y}$ such that $x + y$ is an invertible element modulo U for any subsets $\mathcal{X}, \mathcal{Y} \subset \mathbf{Z}/U\mathbf{Z}$.

Corollary 2.2.3. *Let $\mathcal{X}, \mathcal{Y} \subset \mathbf{Z}/U\mathbf{Z}$ with $|\mathcal{X}|, |\mathcal{Y}| \geq U/M_R$. Then we have*

$$T(\mathcal{X}, \mathcal{Y}) \leq \frac{\phi(P)}{P} |\mathcal{X}| |\mathcal{Y}| \exp\left(\frac{36}{\log \log R}\right), \quad (2.33)$$

where $P = \prod_{Q^2 < p \leq R} p$.

Proof. Let $\mathcal{Z} = \mathbf{Z}/U\mathbf{Z}$, I be the set of primes dividing U and $\mathcal{Z}_p = \mathbf{Z}/p\mathbf{Z}$ for any $p \in I$. Then the Chinese remainder theorem gives

$$\mathcal{Z} = \prod_{p \in I} \mathcal{Z}_p. \quad (2.34)$$

Let $X = \mathcal{X}, Y = -\mathcal{Y}$ and J be the subset of I consisting of primes p such that $Q^2 < p \leq R$. Then we have

$$T(\mathcal{X}, \mathcal{Y}) \leq |\{(x, y) \in X \times Y \mid x_p \neq y_p \text{ for all } p \in J\}|. \quad (2.35)$$

$w = T^\ell$ and let U be a divisor of W . Let $\mathcal{U}, \mathcal{V} \subset \mathbf{Z}/W\mathbf{Z}$ with $|\mathcal{U}||\mathcal{V}| \geq (W/TD)^2$ where D is positive real number satisfying $TD \leq T^\ell$. For any $s \in \mathbf{Z}/W\mathbf{Z}$ we then have

$$\mathcal{R}_U(\mathcal{U}, \mathcal{V}, s) \leq \frac{\phi(U)}{U} |\mathcal{U}||\mathcal{V}| \exp\left(8\left(\log w \sum_{p|U} \frac{1}{p^2}\right)^{1/2}\right), \quad (2.28)$$

where $\mathcal{R}_U(\mathcal{U}, \mathcal{V}, s) = |\{(a, b) \in \mathcal{U} \times \mathcal{V} : (a - b + s, U) = 1\}|$.

Proof. Let $\mathcal{Z} = \mathbf{Z}/W\mathbf{Z}$, I be the set of primes dividing W and $\mathcal{Z}_p = \mathbf{Z}/p\mathbf{Z}$ for any $p \in I$. Then the Chinese remainder theorem gives

$$\mathcal{Z} = \prod_{p \in I} \mathcal{Z}_p. \quad (2.29)$$

Let $X = \mathcal{U}$, $Y = \mathcal{V} - s$ and $J \subseteq I$ be the set of primes dividing U . Then we have that

$$\mathcal{R}_U(\mathcal{U}, \mathcal{V}, s) = |\{(x, y) \in X \times Y \mid x_p \neq y_p \text{ for all } p \in J\}|. \quad (2.30)$$

Also, $\mathcal{L}(X, Y) = \log\left(\frac{|Z|^2}{|X||Y|}\right) = \log\left(\frac{W^2}{|U||V|}\right) \leq 2\log(TD) \leq 2\log T^{2\ell}$, since $TD \leq 8T^\ell \leq T^{2\ell}$. We now obtain (2.28) using (2.5) to bound the right hand side of (2.30) on recalling that $w = T^\ell$ and taking note of the following remarks. We have

$$\exp\left(-\sum_{p \in J} \frac{1}{p}\right) \leq \exp\left(\sum_{p \in J} \frac{2}{p^2}\right) \prod_{p \in J} \left(1 - \frac{1}{p}\right) = \frac{\phi(U)}{U} \exp\left(\sum_{p \in J} \frac{2}{p^2}\right), \quad (2.31)$$

by the inequality $-\log(1 - u) \leq u + 2u^2$ valid for $0 \leq u \leq 1/2$. Finally,

$$\sum_{p \in J} \frac{1}{p^2} \leq \left(\log T^\ell \sum_{p \in J} \frac{1}{p^2}\right)^{\frac{1}{2}} \quad (2.32)$$

Then on recalling the definition of $\mathcal{L}(X, Y)$ from the statement of Theorem 2.2.1 we may rewrite (2.23) as

$$\mathcal{T}_J(X, Y) \leq |X||Y| \exp\left(\frac{\mathcal{L}(X, Y)}{t}\right) D(Z, t)^{\frac{1}{t^2}}. \quad (2.25)$$

We give Z the uniform probability measure. The random variables $\{\epsilon\}_{i \in I}$ on $Z \times Z$ form an independent family, since $\epsilon_i(x, y)$ depends only on the i th co-ordinates of x and y . Thus, if for each $i \in I$ we write $F(Z_i, t)$ for the left hand side of (2.6) with $A = Z_i$ we then see using (2.6) that

$$D(Z, t) = \prod_{i \in J} F(Z_i, t) \leq \exp\left(-t^2 \sum_{i \in J} \frac{1}{Z_i} + t^3 \sum_{i \in J} \frac{1}{Z_i^2}\right). \quad (2.26)$$

Since (2.25) is true for any integer $t \geq 1$, we conclude using (2.26) that

$$\mathcal{T}_J(X, Y) \leq |X||Y| \exp\left(-\sum_{i \in J} \frac{1}{Z_i}\right) \inf_{\substack{t \geq 1, \\ t \in \mathbf{Z}}} \exp\left(\frac{\mathcal{L}(X, Y)}{t} + t \sum_{i \in J} \frac{1}{Z_i^2}\right). \quad (2.27)$$

We obtain (2.5) from (2.27) on remarking that when $a, b \geq 0$, $\inf_{\substack{t \geq 1, \\ t \in \mathbf{Z}}} \left(\frac{a}{t} + bt\right) \leq 2(ab)^{\frac{1}{2}} + b$. Indeed, denoting this infimum by $m(a, b)$ we see that if $a \leq b$ then $m(a, b) \leq a + b \leq 2(ab)^{\frac{1}{2}} + b$ and if $a > b$ we take $t_0 \in \mathbf{Z}$ such that $\left(\frac{a}{b}\right)^{\frac{1}{2}} \leq t_0 \leq \left(\frac{a}{b}\right)^{\frac{1}{2}} + 1$ to get $m(a, b) \leq \frac{a}{t_0} + bt_0 \leq 2(ab)^{\frac{1}{2}} + b$, as required.

We now derive following two corollaries of this theorem, which will be applied in the following chapters.

Corollary 2.2.1. *Let $T \geq 4$ and $l \geq 2$ be real numbers and $W = \prod_{p \leq w} p$, with*

$$\mathcal{T}_J(X, Y) \leq |X|^{1-\frac{1}{t}} \left(\sum_{x \in X} \left(\sum_{y \in Y} \prod_{i \in J} \epsilon_i(x, y) \right)^t \right)^{\frac{1}{t}}. \quad (2.20)$$

We extend the summation over $x \in X$ on the right hand side of (2.20) to $x \in Z$, then expand the summand and finally interchange summations to get

$$\mathcal{T}_J(X, Y) \leq |X|^{1-\frac{1}{t}} \left(\sum_{(y_1, y_2, \dots, y_t) \in Y^t} \sum_{x \in Z} \prod_{1 \leq k \leq t} \prod_{i \in J} \epsilon_i(x, y_k) \right)^{\frac{1}{t}}. \quad (2.21)$$

A second application of Hölder's inequality to exponent t now shows that the term in the brackets on the right hand side of (2.21) does not exceed

$$|Y|^{t-1} \sum_{(y_1, y_2, \dots, y_t) \in Y^t} \left(\sum_{x \in Z} \prod_{1 \leq k \leq t} \prod_{i \in J} \epsilon_i(x, y_k) \right)^t. \quad (2.22)$$

We extend the sum over Y^t in (2.22) to Z^t , expand the summand and substitute the resulting expression into (2.21) to obtain

$$\mathcal{T}_J(X, Y) \leq |X|^{1-\frac{1}{t}} |Y|^{1-\frac{1}{t}} \left(\sum_{(y_1, y_2, \dots, y_t) \in Z^t} \sum_{(x_1, x_2, \dots, x_t) \in Z^t} \prod_{1 \leq l \leq t} \prod_{1 \leq k \leq t} \prod_{i \in J} \epsilon_i(x_l, y_k) \right)^{\frac{1}{t^2}}. \quad (2.23)$$

Note that the expression in the brackets on the right hand side of (2.23) does not depend on the subsets X and Y . Let us set

$$D(Z, t) = \frac{1}{|Z|^2} \sum_{(y_1, y_2, \dots, y_t) \in Z^t} \sum_{(x_1, x_2, \dots, x_t) \in Z^t} \prod_{1 \leq l \leq t} \prod_{1 \leq k \leq t} \prod_{i \in J} \epsilon_i(x_l, y_k). \quad (2.24)$$

We now substitute (2.14) into (2.13) and use (2.15) to get

$$F(t, t) \leq \left(1 - \frac{(2t-1)}{|A|} + \frac{3\binom{t}{2} + \binom{t-1}{2}}{|A|^2} \right) F(t-1, t-1). \quad (2.16)$$

for any integer t with $1 < t < |A|$. Using $1 + u \leq \exp(u)$, valid for all real u , and recurrence on t together with $F(1, 1) \leq \exp(-\frac{1}{|A|})$ we get

$$F(t, t) \leq \exp \left(-\frac{\sum_{1 \leq r \leq t} (2r-1)}{|A|} + \frac{3\binom{t}{2} + 4\sum_{1 \leq r \leq t-1} \binom{r}{2}}{|A|^2} \right) \quad (2.17)$$

from (2.16) when $1 < t < |A|$. We obtain (2.6) for such t from (2.17) since $\sum_{1 \leq r \leq t} (2r-1) = t^2$ and since for any integer $t > 1$ we have

$$3\binom{t}{2} + 4\sum_{1 \leq r \leq t-1} \binom{r}{2} \leq \frac{3t^2}{2} + 2\int_0^t u(u-1)du \leq t^3. \quad (2.18)$$

□

PROOF OF THEOREM 2.2.1.— For each $i \in J$ and $(x, y) \in X \times Y$, let $\epsilon_i(x, y) = 1$ if $x_i \neq y_i$ and be 0 otherwise. Then we have that

$$\mathcal{T}_J(X, Y) = \sum_{x \in X} \sum_{y \in Y} \prod_{i \in J} \epsilon_i(x, y). \quad (2.19)$$

Let $t \geq 1$ be an integer. An application of Hölder's inequality to exponent t to the right hand side of (2.19) gives

Since for any y, y' in A we have

$$\mathbb{E}_x(\delta(x, y)) = \frac{1}{|A|} \quad \text{and} \quad \mathbb{E}_x(\delta(x, y)\delta(x, y')) = \frac{\delta(y, y')}{|A|}. \quad (2.10)$$

we then deduce that

$$1 - \frac{m}{|A|} \leq \mathbb{E}_{x_n} \prod_{1 \leq j \leq m} \epsilon(x_n, y_j) \leq 1 - \frac{m}{|A|} + \frac{1}{|A|} \sum_{1 \leq j < l \leq m} \delta(y_j, y_l). \quad (2.11)$$

Substituting this into (2.8) we obtain

$$\begin{aligned} \left(1 - \frac{m}{|A|}\right) F(m, n-1) &\leq F(m, n) \\ &\leq \left(1 - \frac{m}{|A|}\right) F(m, n-1) + \frac{\binom{m}{2}}{|A|^2} F(m-1, n-1). \end{aligned} \quad (2.12)$$

Putting successively $m = n = t$ and $m = t-1, n = t$ for $1 < t$ in the upper inequality in (2.12) and using $F(m, n) = F(n, m)$ in the latter case gives

$$F(t, t) \leq \left(1 - \frac{t}{|A|}\right) F(t, t-1) + \frac{\binom{t}{2}}{|A|^2} F(t-1, t-1), \quad (2.13)$$

$$F(t, t-1) \leq \left(1 - \frac{t-1}{|A|}\right) F(t-1, t-1) + \frac{\binom{t-1}{2}}{|A|^2} F(t-1, t-2). \quad (2.14)$$

Also, the lower inequality in (2.12) with $m = n = t-1$ implies

$$\left(1 - \frac{t}{|A|}\right) F(t-1, t-2) \leq F(t-1, t-1). \quad (2.15)$$

A , and likewise A^n , for an integer $n \geq 1$, their uniform probability measures and write \mathbb{E}_x and $\mathbb{E}_{x_1, x_2, \dots, x_n}$ in place of $\frac{1}{|A|} \sum_{x \in |A|}$ and $\frac{1}{|A|^n} \sum_{(x_1, x_2, \dots, x_n) \in A^n}$ respectively. Also, we will use these notations in the same sense with other letters in place of x . For any integers $n, m \geq 1$ we set

$$F(m, n) = \mathbb{E}_{y_1, y_2, \dots, y_m, x_1, x_2, \dots, x_n} \prod_{1 \leq j \leq m} \prod_{1 \leq k \leq n} \epsilon(x_k, y_j). \quad (2.7)$$

and define $F(m, 0) = F(0, n) = F(0, 0) = 0$. Thus $F(m, n) = F(n, m)$ for all integers $m, n \geq 0$.

For an integer $t \geq 1$, the left hand side of (2.6) is the same as $F(t, t)$, which we shall bound by recurrence on t . Since (2.6) is trivial when $|A| \leq t$ and since $F(1, 1) = 1 - \frac{1}{|A|} \leq \exp(-\frac{1}{|A|})$, we shall assume that $1 < t < |A|$. Now let m, n be integers with $|A| \geq m, n \geq 1$. Then we have

$$F(m, n) = \mathbb{E}_{y_1, y_2, \dots, y_m, x_1, x_2, \dots, x_{n-1}} \prod_{\substack{1 \leq j \leq m, \\ 1 \leq k \leq n-1}} \epsilon(x_k, y_j) \mathbb{E}_{x_n} \prod_{1 \leq j \leq m} \epsilon(x_n, y_j) \quad (2.8)$$

To bound the expectation over x_n in (2.8) we let $\delta(x, y) = 1 - \epsilon(x, y)$ and apply the truncation inequalities

$$\begin{aligned} 1 - \sum_{1 \leq j \leq m} \delta(x_n, y_j) &\leq \prod_{1 \leq j \leq m} \epsilon(x_n, y_j) \\ &\leq 1 - \sum_{1 \leq j \leq m} \delta(x_n, y_j) + \sum_{1 \leq j < l \leq m} \delta(x_n, y_j) \delta(x_n, y_l). \end{aligned} \quad (2.9)$$

method of [26] and [25] relies only on one application of Hölder's inequality rather than the two applications that we use following [24].

2.2 Local Problem for invertible elements

Theorem 2.2.1. *Let Z be the product of a finite family of finite sets $\{Z_i\}_{i \in I}$ and X and Y be non-empty subsets of Z . Given a subset J of I , let*

$$\mathcal{T}_J(X, Y) = |\{(x, y) \in X \times Y \mid x_i \neq y_i \text{ for all } i \in J\}|, \quad (2.4)$$

where x_i and y_i are the i -th of co-ordinates of x and y respectively for each $i \in I$.

Then we have that

$$\mathcal{T}_J(X, Y) \leq |X||Y| \exp\left(-\sum_{i \in J} \frac{1}{|Z_i|}\right) \exp\left(2\left(\mathcal{L}(X, Y) \sum_{i \in J} \frac{1}{|Z_i|^2}\right)^{\frac{1}{2}} + \sum_{i \in J} \frac{1}{|Z_i|^2}\right), \quad (2.5)$$

where $\mathcal{L}(X, Y) = \log\left(\frac{|Z|^2}{|X||Y|}\right)$.

The proof of the theorem depends on the following lemma, which we take up first.

Lemma 2.2.1. *Let A be a finite set and for any $(x, y) \in A^2$ let us set $\epsilon(x, y)$ to be 1 if $x \neq y$ and to be 0 otherwise. Then for any integer $t \geq 1$ we have*

$$\frac{1}{|A|^{2t}} \sum_{(y_1, y_2, \dots, y_t) \in A^t} \sum_{(x_1, x_2, \dots, x_t) \in A^t} \prod_{1 \leq l \leq t} \prod_{1 \leq k \leq t} \epsilon(x_l, y_k) \leq \exp\left(-\frac{t^2}{|A|} + \frac{t^3}{|A|^2}\right). \quad (2.6)$$

Proof. It will be convenient to use probabilistic terminology. Accordingly, we give

$$\mathcal{T}_{W,\Omega}(\mathcal{X}, \mathcal{Y}) \leq |\mathcal{X}|^{1-\frac{1}{t}} |\mathcal{Y}|^{1-\frac{1}{s}} \left(\sum_{(y_1, y_2, \dots, y_t) \in \mathcal{Z}^t} \sum_{(x_1, x_2, \dots, x_s) \in \mathcal{Z}^t} \prod_{1 \leq l \leq s} \prod_{1 \leq k \leq t} \epsilon_{W,\Omega}(x_l, y_k) \right)^{\frac{1}{ts}}, \quad (2.3)$$

where $\mathcal{Z} = \mathbf{Z}/W\mathbf{Z} = \prod_{p|W} \mathbf{Z}/p\mathbf{Z}$, by the Chinese remainder theorem. Note that the quantity in the brackets in (2.3) is independent of \mathcal{X} and \mathcal{Y} . We estimate this quantity by exploiting the product structure of \mathcal{Z} and using induction on integers t and s , which is then chosen to optimise the resulting estimates.

The technical details of the steps in the above outline are different in the two cases that interest us, namely, invertible elements and squares. In particular, in the former case, the details are essentially of a combinatorial nature. That is, one may ignore the ring structure on $\mathbf{Z}/W\mathbf{Z}$, retaining only the fact that it splits into a finite product of finite sets. This is the reason for the combinatorial formulation of the local problem for invertible elements expressed by the statement of Theorem 2.2.1 below, from which our results on this local problem are deduced as corollaries.

An inequality essentially equivalent to (2.5) of Theorem 2.2.1 is stated in Theorem 3 of [26], but the proof of this theorem given in [26] is incorrect, as pointed out in Remark 2.4 on page 965 of [25]. When corrected this proof yields Proposition 2.3 of [25]. The method used to obtain these results gives $\mathcal{L}(X, Y)^{\frac{2}{3}} \left(\sum_{i \in J} \frac{1}{|Z_i|^2} \right)^{\frac{1}{3}}$ in place of $\left(\mathcal{L}(X, Y) \sum_{i \in J} \frac{1}{|Z_i|^2} \right)^{\frac{1}{2}}$ in the second exponential factor on the right hand side of (2.5), as shown in [25]. While this difference is unimportant to the problem considered in Chapter 3, it is crucial to the problem in taken up in Chapter 4. The

trivial bound for $\mathcal{T}_{W,\Omega}(\mathcal{X}, \mathcal{Y})$. We shall, however, be interested only in the following two cases :

(i) $\phi(x, y) = x - y + c$ and Ω is the set of invertible elements of $\mathbf{Z}/W\mathbf{Z}$.

(ii) $\phi(x, y) = x^2 + y^2 + c$ and Ω is the set of invertible squares in $\mathbf{Z}/W\mathbf{Z}$.

Here c is a given element of $\mathbf{Z}/W\mathbf{Z}$.

We call (i) the local problem for invertible elements and (ii) the local problem for squares. Our results in the first case are given by Corollaries 2.2.1 and 2.2.3. These corollaries will be applied in Chapters 4 and 3 respectively. The local problem for the squares was treated by Gyan Prakash, Ramana and Ramaré in [24]. We state their result with minor modifications as Theorem 2.3.1 and apply in it Chapter 5. An essential role in the reduction of the problems considered in the following chapters to the local problems studied here is played by the simple optimization principle described in final section 2.4 of this chapter.

We obtain our results on the local problem for invertible elements by extending the method of [24] for the squares to this case. Let us summarise this method in general terms. To bound $\mathcal{T}_{W,\Omega}(\mathcal{X}, \mathcal{Y})$ we begin by noting that

$$\mathcal{T}_{W,\Omega}(\mathcal{X}, \mathcal{Y}) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \epsilon_{W,\Omega}(x, y). \quad (2.2)$$

where $\epsilon_{W,\Omega}(x, y) = 1$ if $\phi(x, y) \notin \Omega$ and is 0 otherwise. Two applications of Hölder's inequality to integer exponents $t \geq 1$ and $s \geq 1$ yield

CHAPTER 2

The Local Problem and an Optimization Principle

2.1 Introduction

We begin by describing the problem considered in this chapter in fairly abstract terms. Suppose that W is a square free integer and that \mathcal{X}, \mathcal{Y} and Ω are subsets of $\mathbf{Z}/W\mathbf{Z}$, endowed with its natural ring structure. Supposing further that ϕ is a function from $\mathbf{Z}/W\mathbf{Z} \times \mathbf{Z}/W\mathbf{Z}$ to $\mathbf{Z}/W\mathbf{Z}$, our problem is to bound

$$\mathcal{T}_{W,\Omega}(\mathcal{X}, \mathcal{Y}) = |\{(x, y) \in \mathcal{X} \times \mathcal{Y} \mid \phi(x, y) \in \Omega\}|. \quad (2.1)$$

This is the local problem of the title of this chapter. Typically, the square-free integer W is the product of all primes numbers not exceeding a certain bound and \mathcal{X}, \mathcal{Y} are large subsets of $\mathbf{Z}/W\mathbf{Z}$. Of course, in the generality stated, one has only the

$$P_{N;A,B} \gg \frac{N^2}{\phi(m_k)m_k \log N} \gg \frac{|A||B|}{\log N} \log \log R,$$

which proves the proposition.

□

3.6 Optimality

We prove Proposition 3.1.4 here. Indeed, if $r_{A,B}(n)$ is, as before, the number of pairs $(a, b) \in A \times B$ such that $a + b = n$ then

$$P_{N;A,B} \geq \sum_{\substack{\frac{N}{2} \leq p \leq N \\ p \equiv 1 \pmod{m_k}}} r_{A,B}(p). \quad (3.50)$$

We observe that for any integer $n \equiv 1 \pmod{m_k}$ and $n \geq 2m_k$, we have

$$r_{A,B}(n) \geq \left\lfloor \frac{n}{m_k} \right\rfloor \geq \frac{n}{2m_k}.$$

Using this lower bound for $r_{A,B}(p)$ in (3.50) when $N \geq 4m_k$ we get

$$P_{N;A,B} \geq \frac{1}{2m_k} \sum_{\substack{\frac{N}{2} \leq p \leq N, \\ p \equiv 1 \pmod{m_k}}} p \geq \frac{N}{4m_k} \sum_{\substack{\frac{N}{2} \leq p \leq N, \\ p \equiv 1 \pmod{m_k}}} 1. \quad (3.51)$$

By the Chebyshev bound $\log m_k = \sum_{p \leq k} \log p \ll \log \log N$. Thus on using the Siegel-Walfisz theorem (see [17, page 419]), we have

$$\sum_{\substack{\frac{N}{2} \leq p \leq N, \\ p \equiv 1 \pmod{m_k}}} 1 \gg \frac{N}{\phi(m_k) \log N}. \quad (3.52)$$

Merten's formula gives the the upper bound $\phi(m_k) \ll \frac{m_k}{\log \log m_k}$. Also, $|A| \sim \frac{N}{m_k}$, $|B| \sim \frac{N}{m_k}$ and therefore $R \sim m_k$, from the definition of R given in the statement of Theorem 3.1.2. From (3.51) and (3.52) we then get

and

$$D_2 \leq \frac{B_1}{|\mathcal{Y}| - 1} \leq \frac{2|B_1|}{|\mathcal{Y}|},$$

where we use $|\mathcal{X}| \geq \frac{|A_3|}{D_1} \geq 2$ and $|\mathcal{Y}| \geq \frac{|B_1|}{D_2} \geq 2$, valid by condition (i) given at the end of Section 3.4. These bounds on D_1, D_2 together with (3.45) and (3.46) give

$$T(U) \ll \frac{U}{\phi(U)} \frac{|A_3||B_1|}{|\mathcal{X}||\mathcal{Y}|} \sum_{(a,b) \in \mathcal{X} \times \mathcal{Y}} c(a,b), \quad (3.47)$$

Note that \mathcal{X}, \mathcal{Y} are subsets of $\mathbb{Z}/U\mathbb{Z}$ with $|\mathcal{X}| \geq \frac{|A_3|}{D_1} \geq \frac{U}{M_R}$ and $|\mathcal{Y}| \geq \frac{|B_1|}{D_2} \geq \frac{U}{M_R}$ and that the sum on the right of above relation is nothing but $T(\mathcal{X}, \mathcal{Y})$ of the Corollary 2.2.3, which gives

$$T(U) \ll \frac{U}{\phi(U)} \frac{\phi(P)}{P} |A_3||B_1| \exp\left(\frac{36}{\log \log R}\right), \quad (3.48)$$

where $P = \prod_{Q^2 < p \leq R} p$, $U = \prod_{p \leq R} p$ and $Q = \log R \log \log R$. By Mertens's formula we then get

$$T(U) \ll |A||B| \log \log R \exp\left(\frac{C}{\log \log R}\right) \ll |A||B| \log \log R. \quad (3.49)$$

From (3.49), (3.43), (3.36), (3.32) and (3.27) we then conclude that

$$P_{N;A_3,B_1} \ll \frac{|A||B|}{\log N} \log \log R,$$

which together with (3.23) yields Theorem 3.1.3.

$$\sum_{a \in \tilde{A}_3} m_{A_3}(a) = |A_3| \text{ with } 0 \leq m_{A_3}(a) \leq D_1,$$

and

$$\sum_{b \in \tilde{B}_1} m_{B_1}(b) = |B_1| \text{ with } 0 \leq m_{B_1}(b) \leq D_2.$$

Let us set $c(a, b)$ to be 1 when $a + b$ is invertible modulo in $\mathbf{Z}/U\mathbf{Z}$ and to be 0 otherwise. Then from (3.44) we get that

$$T(U) = \frac{U}{\phi(U)} \sum_{(a,b) \in \tilde{A}_3 \times \tilde{B}_1} c(a, b) m_{A_3}(a) m_{B_1}(b). \quad (3.45)$$

We estimate above sum with the help of the optimization principle given in Section 2.4 of Chapter 2. From the Lemma 2.4.1, we then have that

$$\sum_{(a,b) \in \tilde{A}_3 \times \tilde{B}_1} c(a, b) m_{A_3}(a) m_{B_1}(b) \leq \sum_{(a,b) \in \tilde{A}_3 \times \tilde{B}_1} c(a, b) x_a^* y_b^*, \quad (3.46)$$

for some x_a^* and y_b^* with a varying over \tilde{A}_3 and b varying over \tilde{B}_1 , satisfying the following conditions. All the x_a^* are either 0 or D_1 , excepting at most one, which must lie in $(0, D_1)$ and similarly, all y_b^* are either 0 or D_2 , excepting at most one, which must lie in $(0, D_2)$. Moreover, if \mathcal{X} and \mathcal{Y} denote, respectively, the subsets of \tilde{A}_3 and \tilde{B}_1 for which $x_a^* \neq 0$ and $y_b^* \neq 0$, then $|\mathcal{X}| D_1 \geq |A_3| \geq (|\mathcal{X}| - 1) D_1$ and $|\mathcal{Y}| D_2 \geq |B_1| \geq (|\mathcal{Y}| - 1) D_2$. Thus, from this we have the bounds

$$D_1 \leq \frac{A_3}{|\mathcal{X}| - 1} \leq \frac{2|A_3|}{|\mathcal{X}|}$$

$q > R$, the triangle inequality applied to (3.40) shows that

$$|T - T(U)| \ll \sum_{R \leq q \leq N^{1/8}} \frac{1}{\phi(q)} \sum_{a \bmod^* q} |\widehat{A}_3(a/q)| |\widehat{B}_1(a/q)| + O(|A||B|). \quad (3.41)$$

We estimate the sum over q in (3.41) by using $\frac{q}{\log \log q} \ll \phi(q)$ and the large sieve inequality (3.8). Since $\frac{\log \log q}{q}$ is decreases with q for $q \geq 10$, we get that

$$|T - T(U)| \ll \frac{\log \log R}{R} N |A|^{1/2} |B|^{1/2} \quad (3.42)$$

$$\ll |A| |B| \log \log R, \quad (3.43)$$

since $R = \frac{1000N}{|A|^{1/2}|B|^{1/2}}$.

Now we estimate $T(U)$. A simple argument using standard properties of Ramanujan sums, given below (3.23) on page 969 of [25] shows that

$$T(U) = \frac{U}{\phi(U)} |\{(a, b) \in A_3 \times B_1 : (a + b, U) = 1\}|. \quad (3.44)$$

As before, we use \tilde{A}_3 to denote the image of A_3 under the natural projection from the set of all integers \mathbf{Z} to $\mathbf{Z}/U\mathbf{Z}$ and similarly denote by \tilde{B}_1 the image of B_1 . Further, for any residue class a modulo U , let $m_{A_3}(a)$ be the number of elements of the set A_3 that belongs to this residue class. Similarly, we define $m_{B_1}(b)$ for any residue class b modulo U . Let $D_1 = \frac{|A_2|}{U} M_R$ and $D_2 = \frac{|B|}{U} M_R$. We then have using condition (ii) given at the end of the preceding section that

(3.36) follows from (3.37).

We now consider the contribution to the sum on the right-hand side of (3.35) from q in the range $1 \leq q \leq N^{1/8}$. We set

$$T = \sum_{1 \leq q \leq N^{1/8}} \omega(q, L) \sum_{a \bmod^* q} \widehat{A}_3(a/q) \widehat{B}_1(a/q). \quad (3.38)$$

and use asymptotic formula for $\omega(q, L)$ given by (3.14) with $\kappa = 100$. The contribution of error term of this asymptotic formula for $\omega(q, L)$ to T is

$$\ll \frac{1}{(\log N)^{100}} \sum_{1 \leq q \leq N^{1/8}} \frac{2^{\nu(q)} \log 2q}{q} \sum_{a \bmod^* q} |\widehat{A}_3(a/q)| |\widehat{B}_1(a/q)|, \quad (3.39)$$

since $L = N^{1/2}$. By the trivial bound $2^{\nu(q)} \log 2q \ll q$ we see that (3.39) is

$$\begin{aligned} &\ll \frac{1}{(\log N)^{100}} \sum_{1 \leq q \leq N^{1/8}} \sum_{a \bmod^* q} |\widehat{A}_3(a/q)| |\widehat{B}_1(a/q)| \\ &\ll \frac{N}{(\log N)^{100}} |A|^{1/2} |B|^{1/2} \ll |A||B|, \end{aligned}$$

by the large sieve inequality (3.8) and $|A||B| \gg \frac{N^2}{(\log N)^2}$. Thus we have

$$T = \sum_{1 \leq q \leq N^{1/8}} \frac{\mu(q)}{\phi(q)} \sum_{a \bmod^* q} \widehat{A}_3(a/q) \widehat{B}_1(a/q) + O(|A||B|). \quad (3.40)$$

We recall that $U = \prod_{p \leq R} p$. Then since $R = \frac{1000N}{|A|^{1/2}|B|^{1/2}}$ and $|A||B| \gg \frac{N^2}{(\log N)^2}$, we see that $U \leq N^{1/8}$ for all large N . We set $T(U)$ to be the sum over q on the right-hand side of (3.40) restricted to $q|U$. Since for all other q we have either $\mu(q) = 0$ or

after an interchange of summations. We note that

$$\sum_{n \equiv 0 \pmod{d}} r(n) = \frac{1}{d} \sum_{a \pmod{d}} \sum_n r(n) e(an/d) = \frac{1}{d} \sum_{q|d} \sum_{a \pmod{q}} \sum_n r(n) e(an/q), \quad (3.34)$$

by orthogonality of characters on the group $\mathbf{Z}/d\mathbf{Z}$. On combining (3.34) with (3.33), interchanging summations and recalling definition of $\omega(q, L)$ from (3.13) together with (3.30), we deduce that

$$\sum_n r(n) \Lambda^\sharp(n) = \sum_{1 \leq q \leq L} \omega(q, L) \sum_{a \pmod{q}} \widehat{A}_3(a/q) \widehat{B}_1(a/q). \quad (3.35)$$

We estimate the contribution to the sum on the right-hand side of (3.35) from q satisfying $N^{1/8} < q \leq L$ by showing that

$$\begin{aligned} \sum_{N^{1/8} < q \leq L} \omega(q, L) \sum_{a \pmod{q}} \widehat{A}_3(a/q) \widehat{B}_1(a/q) &\ll N^{7/8} (\log N)^2 |A|^{1/2} |B|^{1/2} \\ &\ll |A| |B|. \end{aligned} \quad (3.36)$$

Indeed, by (3.15) we have that the absolute value of the left side of (3.36) does not exceed

$$\frac{(\log 2L)^2}{N^{1/8}} \sum_{1 \leq q \leq L} \sum_{a \pmod{q}} |\widehat{A}_3(a/q)| |\widehat{B}_1(a/q)| \leq \frac{(\log 2L)^2 (2N |A_3|^{1/2} |B_1|^{1/2})}{N^{1/8}}, \quad (3.37)$$

where we have applied the large sieve inequality in the form (3.7) to left-hand side of above relation. We see using $|A_3| \leq |A|$, $|B_1| \leq |B|$ and $|A||B| \gg \frac{N^2}{(\log N)^2}$ that

From definition of $r(n)$ we easily see that

$$\sum_n r(n) e(-nt) = \widehat{A}_3(-t) \widehat{B}_1(-t). \quad (3.30)$$

Thus (3.28) and (3.29) give

$$\sum_n r(n) \Lambda^b(n) \ll \frac{N}{(\log N)^{100}} \int_0^1 |\widehat{A}_3(-t)| |\widehat{B}_1(-t)| dt. \quad (3.31)$$

Applying Cauchy-Schwarz inequality to the above integral and using the Parseval relation together with the fact that $|A_3| \leq |A|$ and $|B_1| \leq |B|$ we get that

$$\sum_n r(n) \Lambda^b(n) \ll \frac{N}{(\log N)^{100}} |A|^{1/2} |B|^{1/2}.$$

On recalling our hypothesis that $|A| |B| \gg \frac{N^2}{(\log N)^2}$ we then obtain

$$\sum_n r(n) \Lambda^b(n) \ll \frac{|A| |B|}{(\log N)^{99}} \ll |A| |B|. \quad (3.32)$$

Now we estimate first term on the right of (3.26). From the definition of $\Lambda^\sharp(n)$ in (3.11) we obtain

$$\sum_n r(n) \Lambda^\sharp(n) = - \sum_{1 \leq d \leq L} \mu(d) \log d \sum_{n \equiv 0 \pmod d} r(n), \quad (3.33)$$

3.5 Proof of the Theorem 3.1.3

We shall prove (3.24) here by the method [25] with appropriate alterations. We assume throughout that N is a sufficiently large integer. We begin by noting that if $(a, b) \in A_3 \times B_1$ is such that $a+b$ is a prime number, then by (iii) above $N^{1/8} \leq a+b$ and consequently $\frac{1}{8} \log N \leq \Lambda(a+b)$, where Λ is the Von Mangoldt function. It then follows that

$$P_{N;A_3,B_1} \log N \leq 8 \sum_{a \in A_3, b \in B_1} \Lambda(a+b), \quad (3.26)$$

We set $L = N^{\frac{1}{2}}$ and substitute the decomposition (3.12) into the right hand side of (3.26) to get

$$P_{N;A_3,B_1} \log N \ll \sum_n r(n) \Lambda^\sharp(n) + \sum_n r(n) \Lambda^\flat(n), \quad (3.27)$$

where $r(n)$ is the number of pairs $(a, b) \in A_3 \times B_1$ such that $n = a+b$.

Let us first estimate the second sum on the right of the above inequality. Since $r(n) = 0$ for n not in the interval $[1, 2N]$, we have that

$$\sum_n r(n) \Lambda^\flat(n) = \int_0^1 \left(\sum_n r(n) e(-nt) \right) \left(\sum_{1 \leq n \leq 2N} \Lambda^\flat(n) e(nt) \right) dt, \quad (3.28)$$

by orthogonality of the functions $t \mapsto e(nt)$ on $[0, 1]$. By Lemma 3.3.1, we have that

$$\sum_{1 \leq n \leq 2N} \Lambda^\flat(n) e(nt) \ll \frac{N}{(\log N)^{100}}. \quad (3.29)$$

Therefore, to complete the proof of Theorem 3.1.3 we need to show that

$$P_{N;A_3,B_1} \ll \frac{|A||B|}{\log N} \log \log R. \quad (3.24)$$

To do this, we may assume that

$$(i) |A_3| \geq \frac{2|A_2|}{U} M_R \text{ and } |B_1| \geq \frac{2|B|}{U} M_R.$$

Indeed, if (i) does not hold, say, $|A_3| < \frac{2|A_2|}{U} M_R$, then using (3.5) and recalling the value of M_R we have the stronger conclusion that

$$P_{N;A_3,B_1} \ll \frac{N}{\log N} |A_3|^{\frac{1}{2}} |B_1|^{\frac{1}{2}} \ll \frac{N}{\log N} \frac{|A_2|^{\frac{1}{2}} |B_1|^{\frac{1}{2}} M_R^{1/2}}{U^{1/2}} \ll \frac{|A||B|}{\log N}, \quad (3.25)$$

since $U \geq e^{\frac{R}{2}}$, by the Chebyshev bound. A similar argument disposes the case $|B_1| < \frac{2|B|}{U} M_R$ as well. By the definitions of A_3 and B_1 given at the beginning of this section, we also have

$$(ii) |\{x \in A_3 : x \equiv a \pmod{U}\}| \leq \frac{|A_2|}{U} M_R \text{ and } |\{y \in B_1 : y \equiv b \pmod{U}\}| \leq \frac{|B|}{U} M_R \text{ for any } a \in \tilde{A}_3, b \in \tilde{B}_1,$$

$$(iii) \text{ each element of } A_3 \text{ is larger than } N^{1/8}.$$

These remarks bring us to our final section, where we prove (3.24) taking account of the conditions (i), (ii) and (iii) above, thereby completing the proof of Theorem 3.1.3.

from which and (3.20) we get

$$P_{N;A_3,B_2} \ll \frac{N}{\phi(U) \log N} |\mathcal{C}_1|^{1/2} |\mathcal{D}_2|^{1/2} |A_3|^{1/2} |B_2|^{1/2}.$$

Now using $|\mathcal{C}_1| \leq U$, $|\mathcal{D}_2| \leq \frac{U}{M_R}$ and $\phi(U) \gg \frac{U}{\log R}$, which follows from Mertens's formula, we get that

$$P_{N;A_3,B_2} \ll \frac{N}{\log N} \frac{\log R}{\sqrt{M_R}} |A|^{1/2} |B|^{1/2},$$

since $A_3 \subset A$, $B_2 \subset B$. Now recalling the definitions R and M_R we get that

$$P_{N;A_3,B_2} \ll \frac{|A||B|}{\log N} \log \log R.$$

Similarly, we get the bounds

$$P_{N;A_4,B_1}, P_{N;A_4,B_2} \ll \frac{|A||B|}{\log N} \log \log R.$$

Using these bounds in (3.17), we then see that

$$P_{N;A_2,B} \ll P_{N;A_3,B_1} + \frac{|A||B|}{\log N} \log \log R. \quad (3.22)$$

Thus, from (3.22) and (3.16), we conclude that

$$P_{N;A,B} \ll P_{N;A_3,B_1} + \frac{|A||B|}{\log N} \log \log R. \quad (3.23)$$

Clearly, we then have

$$P_{N;A_3,B_2} = \sum_{a \in \mathcal{C}_1, b \in \mathcal{D}_2} P_{N;A_{3,a},B_{2,b}}. \quad (3.18)$$

The summand on the right of (3.18) can be estimated as

$$P_{N;A_{3,a},B_{2,b}} \ll \frac{N}{\phi(U) \log N} |A_{3,a}|^{1/2} |B_{2,b}|^{1/2}, \quad (3.19)$$

Indeed, if a pair $(x, y) \in A_{3,a} \times B_{2,b}$ is such that $x + y$ is a prime $p_{x,y}$, then $p_{x,y} \equiv a + b \pmod{U}$. Since under the condition $|A||B| \geq \frac{5000N^2}{(\log N)^2}$ we have $R \leq \frac{2}{5} \log N$ and thus $U \leq N^{1/2}$ from the value of U and the Chebyshev bounds, the Brun-Titchmarsh inequality (3.3.2) shows that there are at most $\frac{4N}{\phi(U) \log N}$ such primes $p_{x,y}$. Further, each such prime can be written in at most $\min(|A_{3,a}|, |B_{2,b}|) \leq |A_{3,a}|^{1/2} |B_{2,b}|^{1/2}$ many ways as a sum $x + y$, with $x \in A_{3,a}, y \in B_{2,b}$. These remarks yield (3.19).

Using (3.19) in (3.18) we then get

$$P_{N;A_3,B_2} \ll \frac{N}{\phi(U) \log N} \sum_{a \in \mathcal{C}_1, b \in \mathcal{D}_2} |A_{3,a}|^{1/2} |B_{2,b}|^{1/2}. \quad (3.20)$$

By the Cauchy-Schwarz inequality applied to the sum on the right hand side of (3.20), we have

$$\sum_{a \in \mathcal{C}_1, b \in \mathcal{D}_2} |A_{3,a}|^{1/2} |B_{2,b}|^{1/2} \leq |\mathcal{C}_1|^{1/2} |\mathcal{D}_2|^{1/2} \left(\sum_{a \in \tilde{A}_3, b \in \tilde{B}_2} |A_{3,a}| |B_{2,b}| \right)^{1/2} \quad (3.21)$$

Let us now define

$$A_3 = \{x \in A_2 : x \equiv a \pmod{U} \text{ for some } a \in \mathcal{C}_1\},$$

$$B_1 = \{y \in B : y \equiv b \pmod{U} \text{ for some } b \in \mathcal{D}_1\},$$

$$A_4 = \{x \in A_2 : x \equiv a \pmod{U} \text{ for some } a \in \mathcal{C}_2\},$$

$$B_2 = \{y \in B : y \equiv b \pmod{U} \text{ for some } b \in \mathcal{D}_2\}.$$

Then we have

$$P_{N;A_2,B} = P_{N;A_3,B_1} + P_{N;A_3,B_2} + P_{N;A_4,B_1} + P_{N;A_4,B_2}. \quad (3.17)$$

We first estimate $P_{N;A_3,B_2}$. To do this, for any $a \in \tilde{A}_3 = \mathcal{C}_1$ we define $A_{3,a}$ by

$$A_{3,a} = \{x \in A_3 : x \equiv a \pmod{U}\},$$

and similarly for any $b \in \tilde{B}_2 = \mathcal{D}_2$ we define $B_{2,b}$ by

$$B_{2,b} = \{y \in B_2 : y \equiv b \pmod{U}\}.$$

Then we have a partition of A_3 and B_2 as follows:

$$A_3 = \cup_{a \in \mathcal{C}_1} A_{3,a} \text{ and } B_2 = \cup_{b \in \mathcal{D}_2} B_{2,b}.$$

3.4 Reduction to well distributed subsets

Let N , A and B be as in Theorem 3.1.3. In what follows we take $R = \frac{1000N}{|A|^{1/2}|B|^{1/2}}$, as in Section 3.1, and for this R , we define $U = \prod_{p \leq R} p$, $M_R = (R \log R / \log \log R)^2$ and $Q = \log R \log \log R$. Also, for any subset Z of \mathbf{Z} , we denote by \tilde{Z} the image of Z in $\mathbf{Z}/U\mathbf{Z}$ under the natural projection map from \mathbf{Z} .

Let $A_1 = A \cap [1, N^{1/8}]$, $A_2 = A \cap [N^{1/8}, N]$. Then we have

$$P_{N;A,B} = P_{N;A_1,B} + P_{N;A_2,B} \leq |A_1||B| + P_{N;A_2,B} \ll \frac{|A||B|}{\log N} + P_{N;A_2,B}, \quad (3.16)$$

when N is large enough, since $|A_1| \leq N^{1/8}$, $|B| \leq N$ and $\frac{N^2}{(\log N)^2} \ll |A||B|$.

We now estimate $P_{N;A_2,B}$. To this end, for any a in $\mathbf{Z}/U\mathbf{Z}$ we define $m(a)$ and $n(a)$ to be, respectively, $|\{x \in A_2 : x \equiv a \pmod{U}\}|$ and $|\{y \in B : y \equiv a \pmod{U}\}|$ and then set

$$\mathcal{C}_1 = \{a \in \tilde{A}_2 : m(a) \leq \frac{|A_2|}{U} M_R\},$$

$$\mathcal{D}_1 = \{b \in \tilde{B} : n(b) \leq \frac{|B|}{U} M_R\},$$

$$\mathcal{C}_2 = \{a \in \tilde{A}_2 : m(a) > \frac{|A_2|}{U} M_R\},$$

$$\mathcal{D}_2 = \{b \in \tilde{B} : n(b) > \frac{|B|}{U} M_R\}.$$

Since $\sum_{a \in \tilde{A}} m(a) = |A|$ and $\sum_{b \in \tilde{B}} n(b) = |B|$, it follows that $|\mathcal{C}_2| \leq \frac{U}{M_R}$ and $|\mathcal{D}_2| \leq \frac{U}{M_R}$.

PROOF.— The lemma follows from Davenport’s classical bound for $\sum_{1 \leq n \leq x} \mu(n)e(nt)$, given by Theorem 13.10 on page 348 of [17], by an integration by parts. See [25, Section 3] for the details.

3.3.5 An arithmetical function

For any integer $q \geq 1$ and a positive real number $L \geq 1$, let us set

$$\omega(q, L) = - \sum_{\substack{1 \leq l \leq L, \\ l \equiv 0 \pmod{q}}} \frac{\mu(l) \log l}{l}, \quad (3.13)$$

where μ is the Möbius function. We then have the following estimates for $\omega(q, L)$, proved in [25, Section 2.1]. Here $\nu(q)$ denotes the number of prime divisors of q .

Lemma 3.3.2. (i) For $1 \leq q \leq L^{1/2}$, we have the asymptotic formula

$$\omega(q, L) = \frac{\mu(q)}{\phi(q)} + O_\kappa \left(\frac{2^{\nu(q)} \log 2q}{q (\log L)^\kappa} \right), \quad (3.14)$$

for any $\kappa \geq 1$ and

(ii) for any $q, L \geq 1$, we have

$$|\omega(q, L)| \leq \frac{(\log 2L)^2}{q}. \quad (3.15)$$

$$\Lambda(n) = - \sum_{d|n} \mu(d) \log d \quad \text{for each integer } n \geq 1.$$

For a given $L \geq 1$ we set

$$\Lambda^\sharp(n) = - \sum_{\substack{d|n, \\ d \leq L}} \mu(d) \log d, \quad \Lambda^\flat(n) = - \sum_{\substack{d|n, \\ d > L}} \mu(d) \log d. \quad (3.11)$$

for any integer $n \geq 1$. Naturally, both functions defined depend on L ; we have not made this dependence explicit in the notation for brevity. We then evidently have the following decomposition of Λ , which is a sieve identity originally due to H. Iwaniec, to the extent we are aware.

$$\Lambda(n) = \Lambda^\sharp(n) + \Lambda^\flat(n) \quad \text{for each integer } n \geq 1. \quad (3.12)$$

3.3.4 An application of Davenport's bound

Let $N \geq 1$ be an integer and set $L = N^{1/2}$. Then for $\Lambda^\flat(n)$ as in (3.11), the following lemma gives a uniform bound for the Fourier transform of the restriction of $n \mapsto \Lambda^\flat(n)$ to the interval $[1, 2N]$.

Lemma 3.3.1. *We have*

$$\sum_{1 \leq n \leq 2N} \Lambda^\flat(n) e(nt) \ll \frac{N}{(\log N)^{100}},$$

for all $t \in [0, 1]$.

in $[1, N]$ and $\widehat{A}(t) = \sum_{n \in A} e(nt)$ then

$$\sum_{P \leq q \leq Q} \sum_{a \bmod^* q} |\widehat{A}\left(\frac{a}{q}\right)|^2 \leq 2N|A| \quad (3.7)$$

for any $1 \leq P \leq Q$ and $Q \leq N^{\frac{1}{2}}$. Further, if B is also a subset of $[1, N]$ then

$$\sum_{P \leq q \leq Q} \sum_{a \bmod^* q} |\widehat{A}\left(\frac{a}{q}\right) \widehat{B}\left(\frac{a}{q}\right)| \leq 2N\sqrt{|A||B|}, \quad (3.8)$$

as follows by an application of Cauchy-Schwarz inequality together with (3.7).

3.3.2 The Brun-Titchmarsh inequality

If q, a are positive integers with $(a, q) = 1$, then for all $q \leq x$, we have

$$\pi(x; q, a) \leq \frac{2x}{\phi(q) \log(x/q)}, \quad (3.9)$$

where $\pi(x; q, a)$ denotes the number of primes not exceeding x and congruent to a modulo q . For a proof see [21, page 121]. In particular, we have

$$\pi(x; q, a) \leq \frac{4x}{\phi(q) \log x}, \quad (3.10)$$

when $q \leq x^{\frac{1}{2}}$.

3.3.3 A decomposition of the Von Mangoldt function Λ

The fundamental theorem of arithmetic can be written in the form

$$P_{N;A,B} \ll \frac{N}{\log N} \min(|A|, |B|) \ll \frac{N}{\log N} |A|^{\frac{1}{2}} |B|^{\frac{1}{2}} \ll \frac{|A||B|}{\log N} R, \quad (3.5)$$

as required.

3.3 Preliminaries

In this section we collect various facts, some standard and the others from [25], in preparation for the proof Theorem 3.1.3.

3.3.1 The large sieve inequality

The inequality (3.6) is the classical large sieve inequality, proved on [38, page 68], for example.

Let $N \geq 1$ be a integer and $Q \geq 1$ be a real number. Then for any sequence of complex numbers $\{a_n\}_{n=1}^N$ and real number t if we set

$$S(t) = \sum_{1 \leq n \leq N} a_n e(nt),$$

we have

$$\sum_{1 \leq q \leq Q} \sum_{a \bmod^* q} |S(a/q)|^2 \leq (N + Q^2) \sum_{1 \leq n \leq N} |a_n|^2. \quad (3.6)$$

We will often use this result in the following forms : if A is a subset of the integers

3.1.3 is finally completed in Section 3.5.

The method Ramana and Ramaré [25] is a variant of the original method of Ruzsa and Ramaré [26] and was devised to obtain an upper bound for additive energy of “dense” subsets of primes, the subject of Chapter 4 of this thesis, where we will describe an essentially optimal form of this method for that purpose.

3.2 Proof of the Theorem 3.1.2

We give here a simple proof of Theorem 3.1.2. To this end, for any $A, B \subset \mathbf{N}$ and an integer n we set

$$r_{A,B}(n) = |\{(a, b) \in A \times B : a + b = n\}|, \quad (3.2)$$

which for brevity we denote by $r(n)$. Clearly, we have that

$$r(n) \leq \min(|A|, |B|), \quad (3.3)$$

and that

$$P_{N;A,B} = \sum_{\substack{1 \leq p \leq 2N, \\ p \text{ prime}}} r(p). \quad (3.4)$$

Using the bound (3.3) for $r(n)$ and the Chebyshev bound for the number of primes not exceeding x , we then get from (3.4) that

The upper bound for $P_{N;A,B}$ given by this theorem is optimal, up to the implied constant, in general. In fact, we have the following proposition, which corrects the conclusion of Example 2, page 36 of [3].

Proposition 3.1.4. *Let N be a positive integer, $k \ll \log \log N$ be an integer and $m_k = \prod_{p \leq k} p$. Then if $A = \{1 \leq a \leq N : a \equiv 0 \pmod{m_k}\}$ and $B = \{1 \leq b \leq N : b \equiv 1 \pmod{m_k}\}$, we have that*

$$P_{N;A,B} \gg \frac{|A||B|}{\log N} \log \log R.$$

This proposition, which is an easy consequence of the Siegel-Walfisz theorem, will be proved in Section 3.6.

Let us now turn to the proof of Theorem 3.1.3, which is readily seen to amount to obtaining an upper bound for

$$\sum_{a \in A, b \in B} \Lambda(a + b). \tag{3.1}$$

We bound this sum using a method that was suggested by the method of Ramana and Ramaré [25]. To use this method, however, we will first need to reduce to the case when A and B are themselves “large sets” that are “well-distributed” with respect to certain moduli and contain “relatively large elements”. This reduction is carried out in Section 3.4, at the end of which the reader will find precise formulations of the conditions we have just stated on A and B . In addition, we will require a number of preliminary facts which we record in Section 3.3. The proof of Theorem

where $A(x), B(x)$ are the counting functions of the sets A, B respectively.

This result shows that Ostmann's conjecture will easily follow if the conjecture below is shown to hold.

Conjecture 3.1.1. *For any $\delta > 0$ and all subsets A, B of $\{1, 2, \dots, N\}$ with $|A|, |B| \geq N^\delta$, the sumset $A + B$ contains a composite number when N is a sufficiently large integer.*

In [3, Section 5], A. Sárközy, A. Balog and J. Rivat pose a question which goes in the opposite direction to the conjecture stated above. More precisely, given subsets A, B of $\{1, \dots, N\}$, where $N \geq 1$ is an integer, they ask for optimal upper bounds for the number $P_{N;A,B}$ of pairs (a, b) in $A \times B$ such that $a + b$ is a prime number. Using the linear sieve it is shown in [3] that

Theorem 3.1.2. *We have $P_{N;A,B} \ll \frac{|A||B|}{\log N} R$ where $R = \frac{1000N}{|A|^{1/2}|B|^{1/2}}$.*

Our contribution to the aforementioned question of Sárközy et al., described in this chapter, begins with observation that the bound (1) can be obtained by a simple counting argument and the Chebyshev upper bound. We show this in Subsection 3.2 below. Also, note that (1) is implied by the trivial estimate $P_{N;A,B} \leq |A||B|$ unless $|A||B| \gg \frac{N^2}{(\log N)^2}$. We then improve (1) under this assumption on $|A||B|$ by proving the following theorem, which is the main result of this Chapter.

Theorem 3.1.3. *Let A, B be subsets of $\{1, \dots, N\}$, where $N \geq 1$ is an integer and suppose further that $|A||B| \gg \frac{N^2}{(\log N)^2}$. Then we have that $P_{N;A,B} \ll \frac{|A||B|}{\log N} \log \log R$.*

CHAPTER 3

Primes in Sumsets

3.1 Introduction

Let S be an infinite subset of the natural numbers \mathbf{N} . An asymptotic additive decomposition of S is a pair of subsets (A, B) of \mathbf{N} with $|A|, |B| \geq 2$ such that $S \cap [x_0, \infty) = A + B \cap [x_0, \infty)$ for some $x_0 \in \mathbf{N}$, where $A + B$ is the sumset of A and B . When such an asymptotic decomposition of S exists, S is said to be asymptotically additively decomposable. A famous conjecture of Ostmann (see page 13 of [23]), often called the inverse Goldbach conjecture, asserts that the set \mathcal{P} of prime numbers is not asymptotically additively decomposable. The following result of C. Elsholtz (see page 1 of [7]) gives a necessary condition.

Theorem 3.1.1. *Suppose that (A, B) is an asymptotic additive decomposition of \mathcal{P} then we have that*

$$\frac{x^{1/2}}{(\log x)^5} \ll A(x), B(x) \ll x^{1/2}(\log x)^4,$$

Since δ in (4.50) can be arbitrarily small to start with, we see that the conclusion of the Theorem 4.1.1 is optimal in general.

4.5 Lower Density of Sumset of Sets of Primes

Here we show that the bound (2.49) from Matomäki [20] can be easily recovered from Theorem 4.1.1 using the inequality

$$|A + B| E(A, B) \geq |A|^2 |B|^2. \quad (4.51)$$

Indeed, let ϵ, θ be real numbers in $(0, 1)$ and for any integer $N \geq 1$ set $A = \mathcal{A} \cap [1, \epsilon N]$ and $B = \mathcal{B} \cap (N^{1-\theta}, (1-\epsilon)N]$. Then $|B| \geq \mathcal{B}((1-\epsilon)N) - N^{1-\theta}$, $m(B) \geq N^{1-\theta}$ and $|A| \geq \frac{\epsilon\alpha}{2}\pi(N)$, when N is large enough. Since $A + B \subseteq [1, N]$ we have

$$(\mathcal{A} + \mathcal{B})(N) \geq |A + B| \geq \frac{(\mathcal{B}((1-\epsilon)N) - N^{1-\theta}) \log N^{1-\theta}}{(1 + o(\alpha))e^\gamma \log \log \left(\frac{8}{\epsilon\alpha}\right)}, \quad (4.52)$$

where the second inequality follows from (4.4) and (4.51). We then conclude that

$$\liminf_{N \rightarrow +\infty} \frac{(\mathcal{A} + \mathcal{B})(N)}{N} \geq \frac{(1-\theta)(1-\epsilon)\beta}{(1 + o(\alpha))e^\gamma \log \log \left(\frac{8}{\epsilon\alpha}\right)}, \quad (4.53)$$

from which (4.5) follows on setting $\epsilon = \alpha$ and letting $\theta \rightarrow 0$.

□

invertible residue class modulo M . We set $A = \mathcal{S} \cap [1, \delta R]$ and $B = \mathcal{S} \cap [R, 2R]$. By appropriately modifying the application of the circle method described in [2] we then see that

$$E(A, B) \sim \frac{c_\infty M}{\phi(M)^4} \prod_{p>L} \left(1 + \frac{1}{(p-1)^3}\right) \frac{R^3}{(\log R)^4} \quad (4.47)$$

as $R \rightarrow +\infty$, where

$$c_\infty = 1_{[-\delta, 0]} * 1_{[0, \delta]} * 1_{[-2, -1]} * 1_{[1, 2]}(0) = \int_{|u| \leq \delta} (1 - |u|)(\delta - |u|) du = \left(1 - \frac{\delta}{3}\right) \delta^2. \quad (4.48)$$

We have $\log M \sim L$ and $\frac{M}{\phi(M)} \sim e^\gamma \log \log M$ as $L \rightarrow +\infty$ by the Prime Number Theorem and Mertens's formula respectively and therefore $\frac{M}{\phi(M)} \geq e^\gamma \log \log \left(\frac{16\phi(M)}{\delta}\right)$ for all large enough $L \geq L(\delta)$, where $L(\delta)$ depends only on δ . For a fixed $L \geq L(\delta)$, the Prime Number Theorem for arithmetical progressions gives $|A| \sim \frac{\delta R}{\phi(M) \log R}$ and $|B| \sim \frac{R}{\phi(M) \log R}$ as $R \rightarrow +\infty$. We then conclude from (4.47) that there is an $R(\delta)$, depending only on δ , such that

$$E(A, B) \geq (1 - \delta) \prod_{p>L} \left(1 + \frac{1}{(p-1)^3}\right) \frac{e^\gamma |A|^2 |B|}{\log R} \log \log \left(\frac{16\phi(M)}{\delta}\right) \quad (4.49)$$

whenever $R \geq R(\delta)$. We put $\alpha = \frac{1}{2\phi(M)}$ and $N = 2R$. Then $R \geq R(\delta)$ is large enough, A and B are sets of prime numbers in $[1, N]$ with $|A| \geq \alpha \pi(N)$ and $m(B) \geq R$. Moreover, it follows from (4.49) that

$$E(A, B) \geq (1 - \delta) e^\gamma \frac{|A|^2 |B|}{\log m(B)} \log \log \left(\frac{4}{\delta \alpha}\right), \quad (4.50)$$

for any divisor U of W . We then conclude from (4.8) and (4.39) that

$$\mathcal{S}(W) \leq \frac{\frac{W}{U}}{\phi\left(\frac{W}{U}\right)} |A|^2 |B| \exp\left(8 \left(\log w \sum_{p|U} \frac{1}{p^2}\right)^{\frac{1}{2}} + \frac{32w}{W}\right). \quad (4.45)$$

for any divisor U of W . We set $u = \log w \log \log w$. Then $2 < u < w$ since $w \geq 16$ and we apply (4.45) with $U = \prod_{u < p \leq w} p$. We have $\sum_{p|U} \frac{1}{p^2} \leq \sum_{n > u} \frac{1}{n^2} \leq \frac{2}{u}$. Using (3.16) on page 70 of [28], we easily see that $\exp\left(\frac{w}{2}\right) \leq W$ and that $16w^2 \log \log w \leq w^4 \leq \exp(w)$ since $w \geq 16$. It follows that $\frac{32w}{W} \leq \frac{8}{(\log \log w)^{\frac{1}{2}}}$. By (3.30) of [28] and with $V = \frac{W}{U}$ we have $\frac{V}{\phi(V)} \leq e^\gamma (\log u) \exp\left(\frac{1}{(\log \log w)^2}\right)$. From these remarks and recalling that $w = \left(\frac{4}{\alpha}\right)^2$ we easily conclude that

$$\mathcal{S}(W) \leq e^\gamma |A|^2 |B| \log \log \left(\frac{4}{\alpha}\right) \exp\left(\frac{27 \log \log \log w}{(\log \log w)^{\frac{1}{2}}}\right). \quad (4.46)$$

Substituting this into (4.11) and combining with (4.9) we obtain (4.4) for all integers N such that $N \geq W^2$ and $N \geq C \exp(w)$. Since we certainly have $\exp(4w) \geq W^2$ by (3.15) on page 70 of [28], these conditions on N are met when $N \geq C \exp(4w)$, where $w = \left(\frac{4}{\alpha}\right)^2$ as before. This completes the proof of Theorem 4.1.1. \square

4.4 Optimality

We show here that conclusion of Theorem 4.1.1 is essentially the best possible in general. Thus let $R \geq 1$ and δ in $(0, 1)$ be real numbers. Also, let $M = \prod_{p \leq L} p$ for an integer $L \geq 1$ and let \mathcal{S} be the set of all prime numbers lying in a fixed

Substituting (4.41) into (4.40) and combining the resulting inequality with (4.37), (4.36), (4.33), (4.32), (4.31) and finally (4.27) we conclude that

$$S(A, B) \leq \mathcal{S}(W) + \frac{C_4(\log 2w)^2 N |A| |B|}{w \log N} + \frac{C_5 N |A| |B|}{(\log N)^2}, \quad (4.42)$$

when $N \geq C_2$ and $(\log N)^4 \geq w = \left(\frac{4}{\alpha}\right)^2$, for some real numbers $C_4, C_5 > 0$ and $C_2 \geq 1$. Thus if $N \geq C_2 \exp(w)$, then $\log N \geq w$ and it follows from (4.42) that we have (4.11) with $o(\alpha) = \frac{(C_4+C_5)(\log 2w)^2}{\alpha w}$, since $|A| \geq \alpha\pi(N) \geq \frac{\alpha N}{\log N}$ by (3.5) on page 69 of [28].

We now estimate $\mathcal{S}(W)$ using Theorem 4.1.2. We suppose that $N \geq W^2$ and begin by noting that

$$|A| \geq \frac{4N}{T \log N} \quad \text{and} \quad |\{x \in A \mid x \equiv a \pmod{W}\}| \leq \frac{4N}{\phi(W) \log N}, \quad (4.43)$$

where the second inequality is trivial when $(a, W) \neq 1$ since A is a set of prime numbers and follows from the Brun-Titchmarsh Theorem (see paragraph 3.3.2) when $(a, W) = 1$ since $\log N \geq 2 \log W$. Thus on setting $\mathcal{X} = A$, $M = \frac{4N}{\log N} \geq 1$ and $D = \frac{W}{\phi(W)} > 0$ we see from (4.43) that the condition (4.6) is satisfied. Also, (3.30) on page 70 of [28] gives $D \leq 2e^\gamma \log w$ from which we see that $8T^\ell \geq TD$ holds. We set $S = B$ and $c(b) = b$ for all $b \in B$. From (4.7) we evidently have

$$|\{(x_1, x_2, y_1) \in A^2 \times B \mid (x_1 - x_2 + y_1, W) = 1\}| \leq \mathcal{R}_U(A, B), \quad (4.44)$$

when N is sufficiently large. We are now left with the sum

$$\mathcal{S} = \sum_{1 \leq q \leq (\log N)^4} \frac{\mu(q)}{\phi(q)} \sum_{a \bmod^* q} \left| \widehat{A} \left(\frac{a}{q} \right) \right|^2 \widehat{B} \left(\frac{a}{q} \right), \quad (4.38)$$

which is the crux of the matter. We set $T = \frac{4}{\alpha} \geq 4$ and $\ell = 2$. Then with $w = T^\ell$ and $W = \prod_{p \leq w} p$ we suppose that $(\log N)^4 \geq w$ and write $\mathcal{S}(W)$ for the sub sum of \mathcal{S} over all $q|W$. Reasoning as in the justification of (3.23) on page 969 of [25] using standard properties of Ramanujan sums that

$$\mathcal{S}(W) = \frac{W}{\phi(W)} \left| \{(x_1, x_2, y_1) \in A^2 \times B \mid (x_1 - x_2 + y_1, W) = 1\} \right|. \quad (4.39)$$

Since for any integer $q \geq 1$ that does not divide W we have either $\mu(q) = 0$ or $q > w$, we conclude from (4.38) and the triangle inequality that

$$\mathcal{S} \leq \mathcal{S}(W) + 2|B| \sum_{w < q \leq (\log N)^4} \frac{1}{\phi(q)} \sum_{a \bmod^* q} \left| \widehat{A} \left(\frac{a}{q} \right) \right|^2. \quad (4.40)$$

We set $E(u) = \sum_{1 \leq q \leq u} \sum_{a \bmod^* q} \left| \widehat{A} \left(\frac{a}{q} \right) \right|^2$ for any $u \geq 1$. Then Proposition 4.3.1 tells us that $E(u) \ll \frac{N|A|}{\log N} \log 2u$ when $u \leq 2(\log N)^4$, since $2(\log N)^4 \leq N^{\frac{1}{2}}$ when N is sufficiently large. By partial integration, as in the justification of (3.20) on page 969 of [25], we deduce that

$$\sum_{w < q \leq (\log N)^4} \frac{1}{\phi(q)} \sum_{a \bmod^* q} \left| \widehat{A} \left(\frac{a}{q} \right) \right|^2 \ll \frac{(\log 2w)^2 N|A|}{w \log N}. \quad (4.41)$$

This is easily seen, as in the previous chapter, by inserting

$$\begin{aligned} \sum_{n \equiv 0 \pmod d} r(n) &= \frac{1}{d} \sum_{a \pmod d} \sum_n r(n) e(an/d) \\ &= \frac{1}{d} \sum_{q|d} \sum_{a \pmod^* q} \sum_n r(n) e(an/q), \end{aligned} \tag{4.35}$$

which is a consequence of orthogonality of characters on the group $\mathbf{Z}/d\mathbf{Z}$, into the left hand side of (4.34), interchanging summations, using (3.13) and (4.30).

By (3.15) we have $\omega(q, L) \leq \frac{(\log 2L)^2}{q} \leq \frac{1}{(\log N)^2}$ when $(\log N)^4 < q$. This together with the trivial bound $|\widehat{B}(a/q)| \leq |B|$ and the large sieve inequality (3.7) gives

$$\left| \sum_{(\log N)^4 < q \leq L} \omega(q, L) \sum_{a \pmod^* q} \left| \widehat{A}\left(\frac{a}{q}\right) \right|^2 \widehat{B}\left(\frac{a}{q}\right) \right| \leq \frac{2N|A||B|}{(\log N)^2}. \tag{4.36}$$

When $1 \leq q \leq L^{\frac{1}{2}}$, (3.14) gives the asymptotic formula $\omega(q, L) = \frac{\mu(q)}{\phi(q)} + E(q, L)$ with $E(q, L) \ll_{\kappa} \frac{2^{\nu(q)} \log q}{q(\log L)^{\kappa}}$, for any $\kappa \geq 1$. This applies when $1 \leq q \leq (\log N)^4$ if N is large enough. Taking $\kappa = 2$ and using the trivial inequality $2^{\nu(q)} \log q \leq 4q$ for each $q \geq 1$ we get $E(q, L) \ll \frac{1}{(\log L)^2}$. As above, the large sieve inequality and the trivial bound $|\widehat{B}(a/q)| \leq |B|$ give

$$\left| \sum_{1 \leq q \leq (\log N)^4} E(q, L) \sum_{a \pmod^* q} \left| \widehat{A}\left(\frac{a}{q}\right) \right|^2 \widehat{B}\left(\frac{a}{q}\right) \right| \ll \frac{N|A||B|}{(\log N)^2}, \tag{4.37}$$

for all t in $(0, 1]$. The trivial bound $|\widehat{B}(t)| \leq |B|$ together with the Parseval relation gives $\int_0^1 |\widehat{A}(t)|^2 \widehat{B}(t) dt \leq |A||B|$. Combining these remarks with (4.28) and (4.29) we get

$$\sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0.}} r(n) \Lambda^{\flat}(n) \ll \frac{N|A||B|}{(\log N)^2}, \quad (4.31)$$

We now estimate the second term on the right hand side of (4.27). Using the definition of Λ^{\sharp} in (4.26), interchanging summations and removing the condition $n \neq 0$ in the sum over n by rearranging terms, we get

$$\begin{aligned} \sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0.}} r(n) \Lambda^{\sharp}(n) &= - \sum_{1 \leq d \leq L} \mu(d) \log d \sum_{\substack{n \equiv 0 \pmod{d}, \\ n \neq 0.}} r(n) \\ &= - \sum_{1 \leq d \leq L} \mu(d) \log d \sum_{n \equiv 0 \pmod{d}} r(n) + r(0) \sum_{1 \leq d \leq L} \mu(d) \log d. \end{aligned} \quad (4.32)$$

It follows from the definition of $r(n)$ that $r(n) \leq |A||B|$ for all integers n . Using this bound for $r(0)$ together with a trivial estimate we get

$$|r(0)| \left| \sum_{1 \leq d \leq L} \mu(d) \log d \right| \leq |A||B|L \log L \leq \frac{N|A||B|}{(\log N)^2}. \quad (4.33)$$

when N is large enough, since $L = N^{\frac{1}{2}}$. With $\omega(q, L)$ as defined in (3.13) we then note that

$$- \sum_{1 \leq d \leq L} \mu(d) \log d \sum_{n \equiv 0 \pmod{d}} r(n) = \sum_{1 \leq q \leq L} \omega(q, L) \sum_{a \pmod{*} q} \left| \widehat{A}\left(\frac{a}{q}\right) \right|^2 \widehat{B}\left(\frac{a}{q}\right). \quad (4.34)$$

$$\Lambda^\sharp(n) = - \sum_{\substack{d|n, \\ 1 \leq d \leq L}} \mu(d) \log d \quad \text{and} \quad \Lambda^\flat(n) = - \sum_{\substack{d|n, \\ d > L}} \mu(d) \log d, \quad (4.26)$$

for any integer $n \neq 0$, with $L = N^{\frac{1}{2}}$. Then by the definition of $S(A, B)$ as the sum on the right hand side of (4.10) and using (4.12) we have

$$S(A, B) = \sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0}} r(n) \Lambda^\sharp(n) + \sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0}} r(n) \Lambda^\flat(n), \quad (4.27)$$

where, we recall, $r(n)$ for any integer n is the number of triples $(x_1, x_2, y_1) \in A^2 \times B$ such that $n = x_1 - x_2 + y_1$.

We first estimate the last term in (4.27). Since $r(n) = 0$ when $|n| > 2N$, the Parseval relation gives

$$\sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0}} r(n) \Lambda^\flat(n) = \int_0^1 \left(\sum_{n \in \mathbf{Z}} r(n) e(-nt) \right) \left(\sum_{1 \leq |n| \leq 2N} \Lambda^\flat(n) e(nt) \right) dt. \quad (4.28)$$

Note that $\Lambda^\flat(n) = \Lambda^\flat(-n)$ for any $n \neq 0$. Thus on using Lemma 3.3.1 we obtain

$$\sum_{1 \leq |n| \leq 2N} \Lambda^\flat(n) e(nt) \ll \frac{N}{(\log N)^{100}}, \quad (4.29)$$

for all t in $(0, 1]$. Further, from the definition of $r(n)$ we have

$$\sum_{n \in \mathbf{Z}} r(n) e(nt) = |\widehat{A}(t)|^2 \widehat{B}(t) \quad (4.30)$$

Indeed, otherwise we have distinct (x, y) and (x', y') with $x + y = x' + y'$ so that $z = x - x' = y - y'$ is a non-zero element of \mathbf{T} whose order divides the co-prime integers $\text{ord}(x)\text{ord}(x')$ and $\text{ord}(y)\text{ord}(y')$, which is absurd. Since the order of any element of $X + Y$ does not exceed $N^{\frac{1}{2}}$, $X + Y$ is a δ -spaced set of points in \mathbf{T} with $\delta = \frac{1}{N}$ and the large sieve inequality gives

$$\sum_{x \in X} \sum_{y \in Y} |\widehat{A}_1(x + y)|^2 = \sum_{t \in X + Y} |\widehat{A}_1(t)|^2 \leq 2N|A_1|. \quad (4.23)$$

On the other hand, since A_1 consists of integers co-prime to $\prod_{p \leq N^{\frac{1}{2}}} p$, the arithmetical form of the large sieve inequality (see Section 10.4 on page 105 and Exercise 1 on page 108 of [32] as also [37]) applied to the translate $t \mapsto \widehat{A}_1(x + t)$ of \widehat{A}_1 together with (1.5), page 211 of [39] and $\frac{N^{\frac{1}{2}}}{Q} \geq N^{\frac{1}{4}}$ gives

$$|\widehat{A}_1(x)|^2 \leq \frac{\sum_{y \in Y} |\widehat{A}_1(x + y)|^2}{\sum_{\substack{1 \leq d \leq \frac{N^{\frac{1}{2}}}{Q}, \\ (d, R) = 1.}} \frac{\mu(d)^2}{\phi(d)} \leq \frac{4R}{\phi(R) \log N} \sum_{y \in Y} |\widehat{A}_1(x + y)|^2 \quad (4.24)$$

for each $x \in X$. Summing over all $x \in X$ in (4.24) and combining the result with (4.23) and $\frac{R}{\phi(R)} \leq 4e^\gamma \log Q$ when $Q \geq 2$, obtained from (3.30) on page 70 of [28], we conclude that

$$\sum_{x \in X} \left| \widehat{A}_1(x) \right|^2 \leq 32e^\gamma \frac{\log(2Q)}{\log N} N|A_1|. \quad (4.25)$$

Since $|A| = |A_0| + |A_1|$, we obtain (4.20) from (4.25), (4.22) and (4.21). \square

PROOF OF THEOREM 4.1.1.— We shall first prove (4.11). We begin by extending the definitions introduced in (3.11) by setting

Proposition 4.3.1. *Let $N \geq 2$ be an integer and A be a subset of the primes in $[1, N]$. Then if $1 \leq Q \leq N^{\frac{1}{2}}$ we have that*

$$\sum_{1 \leq q \leq Q} \sum_{a \bmod^* q} \left| \widehat{A} \left(\frac{a}{q} \right) \right|^2 \leq 96e^\gamma \frac{\log(2Q)}{\log N} N|A|. \quad (4.20)$$

Proof. For completeness we give a short proof, tailored from Tao [37], where a more general inequality is proved. First note that by the usual large sieve inequality (see Theorem 7.7 on page 175 of [17]) the left hand side does not exceed $2N|A|$ since $Q \leq N^{\frac{1}{2}}$. It thus suffices to prove (4.20) assuming, as we will, that $1 \leq Q \leq N^{\frac{1}{4}}$.

With $A_0 = A \cap [1, N^{\frac{1}{2}}]$ and $A_1 = (N^{\frac{1}{2}}, N]$ we have $|\widehat{A}(t)|^2 \leq 2(|\widehat{A}_0(t)|^2 + |\widehat{A}_1(t)|^2)$. For any t in the additive group $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, let $\text{ord}(t)$ denote its order. Let X be the set of points x in \mathbf{T} with $\text{ord}(x) \leq Q$. Then the left hand side of (4.20) is $\sum_{x \in X} |\widehat{A}(x)|^2$ and we have

$$\sum_{x \in X} |\widehat{A}(x)|^2 \leq 2 \sum_{x \in X} |\widehat{A}_0(x)|^2 + 2 \sum_{x \in X} |\widehat{A}_1(x)|^2 \quad (4.21)$$

Since $A_0 \subseteq [1, N^{\frac{1}{2}}]$ and $Q \leq N^{\frac{1}{4}}$, the large sieve inequality shows that the first term on the right hand side of (4.21) is at most $4N^{\frac{1}{2}}|A_0|$. Since $2N \geq (\log N)^2$ for $N \geq 1$ we certainly have

$$\sum_{x \in X} |\widehat{A}_0(x)|^2 \leq 16 \frac{\log(2Q)}{\log N} N|A_0|. \quad (4.22)$$

Now let $R = \prod_{p \leq Q} p$ and Y be the set of points in \mathbf{T} with $\text{ord}(y) \leq \frac{N^{\frac{1}{2}}}{Q}$ and $(\text{ord}(y), R) = 1$. Then the map $(x, y) \mapsto x + y$ is bijection from $X \times Y$ onto $X + Y$.

$(0, H)$. Let \mathcal{U} and \mathcal{V} be, respectively, the subsets of $\tilde{\mathcal{X}}$ for which $x_a^* \neq 0$ and $y_b^* \neq 0$. Then (ii) of Lemma 2.4.1 gives $|\mathcal{U}|H \geq |\mathcal{X}| > (|\mathcal{U}| - 1)H$. Combining this with (4.6) we get $|\mathcal{U}| \geq \frac{|\mathcal{X}|}{H} \geq \frac{W}{TD} > 2$, on noting that $W \geq 30T^\ell$ since $T^\ell \geq 16$ and $8T^\ell \geq TD$. This gives $H \leq \frac{|\mathcal{X}|}{|\mathcal{U}| - 1} \leq \frac{|\mathcal{X}|}{|\mathcal{U}|} \exp\left(\frac{2}{|\mathcal{U}|}\right)$ since $\frac{1}{1-u} \leq \exp(2u)$ when $0 \leq u \leq \frac{1}{2}$. Since the same inequalities hold with $|\mathcal{U}|$ replaced by $|\mathcal{V}|$, we obtain

$$H^2 \leq \frac{|\mathcal{X}|^2}{|\mathcal{U}||\mathcal{V}|} \exp\left(\frac{2}{|\mathcal{U}|} + \frac{2}{|\mathcal{V}|}\right) \leq \frac{|\mathcal{X}|^2}{|\mathcal{U}||\mathcal{V}|} \exp\left(\frac{4TD}{W}\right). \quad (4.17)$$

Since $\alpha_i(a, b) \geq 0$ and $0 \leq x_a^*, y_b^* \leq H$ for all (a, b) and $TD \leq 8T^\ell$, we then deduce that

$$\sum_{(a,b) \in \tilde{\mathcal{X}}^2} \alpha_s(a, b) x_a^* y_b^* \leq \frac{|\mathcal{X}|^2}{|\mathcal{U}||\mathcal{V}|} \left(\sum_{(a,b) \in \mathcal{U} \times \mathcal{V}} \alpha_s(a, b) \right) \exp\left(\frac{32T^\ell}{W}\right). \quad (4.18)$$

Using Corollary 2.2.3, we get an upper bound for right hand side of (4.18) as

$$\frac{\phi(U)}{U} |\mathcal{X}|^2 \exp\left(8(\log w \sum_{p|U} \frac{1}{p^2})^{1/2} + \frac{32w}{W}\right). \quad (4.19)$$

Then the conclusion (4.8) of the Theorem 4.1.2 follows from (4.19), (4.18), (4.16) and (4.14).

4.3 Proof of the main theorem

We first record the following proposition which is a variant of Theorem 5.3 of [27].

of (4.11). For the close relation between Ramaré's large sieve inequality and the restriction theorem for primes in the sense of Green, Green-Tao, we refer to Tao [37].

4.2 The finite problem

We prove Theorem 4.1.2 here. With notation as in the statement of this theorem, let U be a given divisor of W . For any a, b in $\mathbf{Z}/W\mathbf{Z}$ and s in S we set $\alpha_s(a, b) = 1$ if $a - b + c(s)$ is invertible modulo U and 0 otherwise. Further, we let $m(a)$ be the number of x in \mathcal{X} such that $x \equiv a \pmod{W}$. Then on writing $\tilde{\mathcal{X}}$ for image of \mathcal{X} in $\mathbf{Z}/W\mathbf{Z}$, we see that

$$\mathcal{R}_U(\mathcal{X}, \mathbf{c}) = \sum_{s \in S} \sum_{(a, b) \in \tilde{\mathcal{X}}^2} \alpha_s(a, b) m(a) m(b). \quad (4.14)$$

Also, we have

$$\sum_{a \in \tilde{\mathcal{X}}} m(a) = |\mathcal{X}| \quad \text{and} \quad 0 \leq m(a) \leq H, \quad (4.15)$$

with $H = \frac{DM}{W}$, from the second condition in (4.6). For a given s in S we now bound the inner sum on the right hand side of (4.14). Using Lemma 2.4.1 of Chapter 2 and (4.15) we obtain

$$\sum_{(a, b) \in \tilde{\mathcal{Z}}^2} \alpha_s(a, b) m(a) m(b) \leq \sum_{(a, b) \in \tilde{\mathcal{Z}}^2} \alpha_s(a, b) x_a^* y_b^*, \quad (4.16)$$

for some x_a^* and y_b^* , with a and b varying over $\tilde{\mathcal{X}}$, such that the x_a^* , and similarly all the y_b^* , are either equal to 0 or to H excepting at most one, which must lie in

we show that

$$S(A, B) \leq \frac{W}{\phi(W)} |\{(x_1, x_2, y_1) \in A^2 \times B \mid (x_1 - x_2 + y_1, W) = 1\}| + o(\alpha)|A|^2|B| \quad (4.11)$$

for all sufficiently large N , where $W = \prod_{p \leq w} p$ with $w = \left(\frac{A}{\alpha}\right)^2$. The proof of (4.4) is completed from (4.11) on using Theorem 4.1.2 to estimate the first term on the right hand side of (4.11) and combining the result with (4.10). To prove (4.11) we begin by analogy with (3.12) and introduce the decomposition

$$\Lambda(|n|) = - \sum_{\substack{d|n, \\ 1 \leq d}} \mu(d) \log d = - \sum_{\substack{d|n, \\ 1 \leq d \leq L}} \mu(d) \log d - \sum_{\substack{d|n, \\ d > L}} \mu(d) \log d, \quad (4.12)$$

valid for any $L \geq 1$ and all integers $n \neq 0$, not necessarily positive. We then set $L = N^{1/2}$, for the given N which is eventually taken to be suitably large, and insert (4.12) into the sum on the right hand side of (4.10). By various steps that run essentially parallel to those following (3.27) in Section 3.5, we reduce to the estimation of the sum

$$\sum_{1 \leq q \leq (\log N)^4} \frac{\mu(q)}{\phi(q)} \sum_{a \bmod^* q} \sum_{(x_1, x_2, y_1) \in A^2 \times B} e\left(\frac{a(x_1 - x_2 + y_1)}{q}\right). \quad (4.13)$$

The contribution to the above sum from $q|W$ is easily seen to be equal to the first term on the right hand side of (4.11). The contribution to (4.13) from the remaining q is estimated using a variant, given here as Proposition 4.3.1, of an improved large sieve inequality for the primes due to Ramaré [27], thereby completing the proof

Theorem 2.2.1 bears fruit. More precisely, the treatment of this local problem in [25] only yields a version of the above theorem with the $(\log w \sum_{p|U} \frac{1}{p^2})^{\frac{1}{2}}$ replaced by $((\log w)^2 \sum_{p|U} \frac{1}{p^2})^{\frac{1}{3}}$ in the exponential factor on the right hand side of (4.8). This in turn is responsible for the factor $2 + o(\alpha)$ in (4) of [25], in place of the $1 + o(\alpha)$ that given by our Theorem 4.1.1.

With Theorem 4.1.2 in hand, it is once again a matter of putting into effect the method of [25], with some modifications, to arrive at Theorem 4.1.1. We conclude this introduction with an outline of the method, deferring the details to Section 4.3. Thus with notation as in the statement of Theorem 4.1.1, let us set $r(n)$ for any integer n to be the number of triples $(x_1, x_2, y_1) \in A^2 \times B$ such that $n = x_1 - x_2 + y_1$. Then we have from (4.1) that

$$E(A, B) = \sum_{n \in B} r(n) . \quad (4.9)$$

This implies the following inequality, which is our point of departure :

$$E(A, B) \log m(B) \leq \sum_{\substack{n \in \mathbf{Z}, \\ n \neq 0.}} r(n) \Lambda(|n|) , \quad (4.10)$$

where $m(B)$, we recall, is $\min_{b \in B} b$ and Λ is the Von Mangoldt function. Indeed, all terms on the right hand side of (4.10) are ≥ 0 and for $n \in B$ we have $\log m(B) \leq \Lambda(n) = \Lambda(|n|)$, since B is a set of prime numbers. The sum on the right hand side of (4.10) is finite, since $r(n) = 0$ when $|n| > 2N$. Let $S(A, B)$ denote this sum. Then

the advantage of being technically simpler as well.

The proof of the Theorem 4.1.1 goes via the following theorem, which may be viewed as a sieve bound.

Theorem 4.1.2. *Let $T \geq 4$ and $\ell \geq 2$ be real numbers and let $W = \prod_{p \leq w} p$, with $w = T^\ell$. Also, let $\mathbf{c} = \{c(s)\}_{s \in S}$ be a finite sequence of integers and \mathcal{X} be a subset of the integers for which there exist real numbers $M \geq 1$ and $D > 0$ with $8T^\ell \geq TD$ such that*

$$|\mathcal{X}| \geq \frac{M}{T} \quad \text{and} \quad |\{x \in \mathcal{X} \mid x \equiv a \pmod{W}\}| \leq \frac{DM}{W} \quad (4.6)$$

for all $a \in \mathbf{Z}/W\mathbf{Z}$. Finally, for any divisor U of W let

$$\mathcal{R}_U(\mathcal{X}, \mathbf{c}) = |\{(x_1, x_2, s) \in \mathcal{X} \times \mathcal{X} \times S \mid (x_1 - x_2 + c(s), U) = 1\}|. \quad (4.7)$$

Then for each divisor U of W we have

$$\mathcal{R}_U(\mathcal{X}, \mathbf{c}) \leq \frac{\phi(U)}{U} |\mathcal{X}|^2 |S| \exp\left(8(\log w \sum_{p|U} \frac{1}{p^2})^{1/2} + \frac{32w}{W}\right). \quad (4.8)$$

Note that $\frac{\phi(U)}{U} |\mathcal{X}|^2 |S|$ is the “expected ” upper bound for $\mathcal{R}_U(\mathcal{X}, \mathbf{c})$ from a naive probabilistic point of view.

We prove Theorem 4.1.2 in Section 4.2 by combining the Corollary 2.2.1 with the optimization principle of Lemma 2.4.1 in the manner we put together Corollary 2.2.3 and Lemma 2.4.1 at the end of Section 3.5. It is in this theorem that the full strength of our finer result on the local problem for invertible elements given by

stated except in the dependencies of $o(\alpha)$ and $N(\alpha)$ on α . In particular, the $1 + o(\alpha)$ on the right hand side of (4.4) cannot be replaced with $c + o(\alpha)$ for any $c < 1$. We shall verify this in Section 4.4.

To place Theorem 4.1.1 in context, we recall that Matomäki [20] gives an optimal lower bound for the relative lower density in the integers of the sum $\mathcal{A} + \mathcal{B}$ of subsets \mathcal{A} and \mathcal{B} of the prime numbers \mathcal{P} in terms of their relative lower densities in \mathcal{P} . More precisely, suppose that $\liminf_{N \rightarrow +\infty} \frac{\mathcal{A}(N)}{\pi(N)} \geq \alpha$ and $\liminf_{N \rightarrow +\infty} \frac{\mathcal{B}(N)}{\pi(N)} \geq \beta$, where $N \mapsto \mathcal{A}(N)$ and $N \mapsto \mathcal{B}(N)$ are the respective counting functions. Then Theorem 1.1 of [20], which is the main theorem of that work, tells us that

$$\liminf_{N \rightarrow +\infty} \frac{(\mathcal{A} + \mathcal{B})(N)}{N} \geq (1 - o_{\alpha+\beta \downarrow 0}(1)) \frac{\beta}{e^\gamma \log \log \left(\frac{1}{\alpha} \right)}, \quad (4.5)$$

where $o_{\alpha+\beta \downarrow 0}(1)$ means that the quantity which tends to 0 as $\alpha + \beta$ tends to 0, and that (4.5) is the best possible in general. This result improves upon the theorem of Chipeniuk and Hamel [5]. As noted in [20], a general result of Ramaré and Ruzsa (Theorem 1 of [26]) when applied to the primes yields a similar conclusion but with $(1 - o(\alpha + \beta))$ replaced with $(c - o(\alpha + \beta))$, for an unspecified c . We shall show in Section 4.5 that (2.49) may be easily deduced from Theorem 4.1.1 and (4.2).

Matomäki obtains (2.49) via a lower bound for $|A + B|$ given by Theorem 2.1 of [20], without recourse to additive energy. This theorem is proved in [20] by the methods of Green and Green-Tao. We shall, however, obtain Theorem 4.1.1 by a refinement of the method of Ramana and Ramaré [25]. Our method appears to have

This follows from an application of the Cauchy-Schwarz inequality. Thus, upper bounds on the additive energy $E(A, B)$ translate to lower bounds on $|A + B|$. The converse, however, is not true (see Sec. 2.3 of [36], for example).

In [25], Ramana and Ramaré showed that for any $\alpha \in (0, 1)$ and $A \subset (N, 2N] \cap \mathcal{P}$ with $|A| \geq \alpha |\mathcal{P} \cap (N, 2N]|$, we have

$$E(A, A) \leq (2 + o(\alpha)) e^\gamma \frac{|A|^3}{\log N} \log \log \frac{4}{\alpha} \quad (4.3)$$

for $N \geq N(\alpha)$. Here, as before, \mathcal{P} denotes the set of prime numbers. The main result of the present chapter of this thesis is the following theorem which generalizes and improves on (4.3).

Theorem 4.1.1. *Let α be in $(0, 1]$. Then there is an $N(\alpha)$ depending only on α such that for all $N \geq N(\alpha)$ and $A \subset [1, N] \cap \mathcal{P}$ satisfying $|A| \geq \alpha \pi(N)$ we have*

$$E(A, B) \leq (1 + o(\alpha)) e^\gamma \frac{|A|^2 |B|}{\log m(B)} \log \log \left(\frac{4}{\alpha} \right) \quad (4.4)$$

for any non-empty $B \subset [1, N] \cap \mathcal{P}$, where $m(B) = \min_{b \in B} b$.

In particular, we replace the $2 + o(\alpha)$ in (4.3) with $1 + o(\alpha)$. In (4.4), $o(\alpha)$ denotes a function of α that tends to 0 with α , $N \mapsto \pi(N)$ is the counting function of the primes and e, γ are the usual numerical constants.

The trivial bound $E(A, B) \leq \min(|A|^2 |B|, |B|^2 |A|)$ shows that (4.4) is non-trivial only if most elements of B are large enough. On the other hand, natural examples show that the conclusion of the above theorem cannot be improved in the generality

CHAPTER 4

Additive energy of dense subsets of the primes

4.1 Introduction

When A and B are subsets the integers, the additive energy of A and B is the quantity $E(A, B)$ defined by

$$E(A, B) = |\{(x_1, x_2, y_1, y_2) \in A \times A \times B \times B \mid x_1 + y_1 = x_2 + y_2\}|. \quad (4.1)$$

The additive energy $E(A, B)$ of A and B is an important quantity in additive combinatorics and additive number theory. It is related to $|A + B|$, the cardinality of the sumset $A + B$, by the classical inequality

$$|A + B| E(A, B) \geq |A|^2 |B|^2. \quad (4.2)$$

primes Ω

$$\Omega = \{p_1^2\} \cup \dots \cup \{p_r^2\} \cup_{l \in R} \{p^2 : p^2 \equiv l \pmod{M_r}\}$$

note that this is a K_r -partition of Ω , where

$$K_r = r + (p_1 - 1) \cdot \frac{p_2 - 1}{2} \cdot \dots \cdot \frac{p_r - 1}{2}.$$

Now we claim that $r(K_r) \geq M_r$. To prove this, Let us take n to be a large square free number which is a multiple of M_r . Since n is square free number, it can not be a multiple of p^2 for any prime p . So n can not be of the form $h p_i^2$ for $1 \leq i \leq r$. And suppose, we express n as a sum of h elements of the set $\{p^2 : p^2 \equiv l \pmod{M_r}\}$ for some invertible square l modulo M_r . This implies, n is of the form $n = h l + (q_1 + \dots + q_h) M_r$. Since n is a multiple of M_r , we have $M_r | h$ which implies that $h \geq M_r$. Thus our claim follows.

Note that the sequence of integers $(K_r)_{r=1}^{\infty}$ is strictly increasing. Therefore, any $K \geq 2$ is in the interval $K_r \leq K < K_{r+1}$ for some r . And noting that $r(K)$ is a non decreasing sequence of K , we get $r(K) \geq r(K_r) \geq M_r$, where the second inequality follows from the above claim. Writing M_r in terms of K , and using known estimates on primes we get that

$$r(K) \gg K \exp\left(\frac{(\log 2 + o(1)) \log K}{\log \log K}\right).$$

such that $\mathfrak{Q}_i \cap (N, 4N]$ contains at least $\frac{\sqrt{N}}{K \log N}$ elements of \mathfrak{Q} . For such i we set $S = \mathfrak{Q}_i \cap (N, 4N]$. Then S is set of square of primes in $(N, 4N]$ with $|S| \geq \frac{\sqrt{N}}{K \log N}$ and no integer in S is divisible by a prime $p \leq K^{25}$. It now follows from (5.70) that (5.71) holds with $D \ll K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$. Since elements of S does not divisible by any prime $p \leq [6D]$ when N is large enough, we may apply Lemma 5.4.1 to S to deduce that every integer $n \geq (288D + 72)N$ is a sum of no more than $\frac{n}{N}$ elements of S . In particular, there is a $C_1 > 0$ such that every integer in $I(N) = ((288D + 72)N, (288D + 73)N]$ is a sum of at most $C_1 D$ squares of primes all belonging to S and therefore to \mathfrak{Q}_i . Thus for all large enough N , every integer in the interval $I(N)$ can be expressed as a sum of no more than $C_1 D$ squares of primes all of the same colour. On remarking that the interval $I(N)$ meets $I(N + 1)$ for all large enough N , we obtain that $r(K) \leq C_1 D$. This yields the conclusion of Theorem 5.1.1 since $C_1 D \ll K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$.

5.5 Optimality

The bound on $r(K) \ll K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$ given by Theorem 5.1.1 is the best possible up to a constant. Indeed, we expect that $3 \log 2$ can be replaced by $\log 2$ in the exponent of the exponential because of the following example of a partition of squares of primes.

Let $r \geq 1$ be an integer. Let $M_r = p_1 \dots p_r$ and let R be the invertible squares modulo M_r . We now consider the following partition of the set of all squares of

5.4 Monochromatic representation

Here we deduce Theorem 5.1.1 from Theorem 5.1.2. Before this deduction we give a lemma with the following notation. For any subset S of the integers, we write $e_6(S)$ for the number of tuples $(x_1, x_2, \dots, x_{12})$ in S^{12} satisfying $x_1 + \dots + x_6 = x_7 + \dots + x_{12}$. We observe that if $S \subset [N, 4N]$ satisfying the hypothesis of Theorem 5.1.2, then we can conclude from (5.2) that

$$e_6(S) \ll \frac{|S|^{11}}{N^{\frac{1}{2}} \log N} \exp\left(\frac{(3 \log 2 + o(1)) \log A}{\log \log A}\right). \quad (5.70)$$

Now we state the lemma as follows.

Lemma 5.4.1. *Let N be positive integer and let $D \geq 1$ be a real number satisfying the condition $N \geq 72D + 12$. If S be a subset of the interval $(N, 4N]$ such that*

$$e_6(S) \leq \frac{|S|^{12} D}{3N} \quad (5.71)$$

and if S contains an integer that is not divisible by any prime $p \leq \lceil 6D \rceil$ then every integer $n \geq 30N(2\lceil 6D \rceil + 1)$ is a sum of no more than $\frac{n}{N}$ elements of S .

PROOF.— See [24, Lemma 1.2].

We now give the proof of Theorem 5.1.1. Since $r(K)$ is increasing with K , it suffices to prove Theorem 5.1.1 for all K sufficiently large. For such a K , let $\cup_{1 \leq i \leq K} \mathfrak{Q}_i$ be a partition of the set of square of primes \mathfrak{Q} into K disjoint subsets.

We set $N_0 = 2K^{50}$. Let N be an integer $\geq N_0$. There is an $i, 1 \leq i \leq K$,

integral in (5.10), we get

$$E_6(S) \ll \frac{2W\tau(U)(\log N)^{11}}{\phi(W)} |\{x \in S^{11} \mid f(x) \text{ an invertible square mod } 2W\}| + O\left(\frac{|S|^{11}(\log N)^{11}}{A}\right) \quad (5.68)$$

5.3.4 Proof of Theorem 5.1.2 completed

It remains only to bound the cardinality of the set

$$\{x \in S^{11} \mid f(x) \text{ an invertible square mod } 2W\}.$$

We find an upper bound the cardinality of this set using Theorem 5.2.1. Let \mathcal{Z} be the set of integers $n > 0$ such that $n^2 \in S$. The set \mathcal{Z} is contained in $[\sqrt{N}, 2\sqrt{N}]$ and satisfies $|\mathcal{Z}| \geq \frac{\sqrt{N}}{A \log N}$ and $|\{z \in \mathcal{Z} \mid z \equiv a \pmod{U}\}| \leq \frac{3\sqrt{N}}{\phi(U) \log N}$, when N is sufficiently large depending on A . Finally, let $I = S^9$ and for any $x = (x_1, x_2, \dots, x_9) \in S^9$ we set $c(x) = x_1 + \dots + x_4 - x_5 - \dots - x_9$. Then with $R_U(\mathcal{Z}, \mathbf{c})$ as in Theorem 5.2.1 we have that

$$|\{x \in S^{11} \mid f(x) \text{ an invertible square modulo } 2W\}| \leq |R_U(\mathcal{Z}, \mathbf{c})|, \quad (5.69)$$

since $U|2W$. On combining the bound for $|R_U(\mathcal{Z}, \mathbf{c})|$ given by Theorem 5.2.1 with (5.69) and (5.68) and after noticing that $\frac{W}{\phi(W)}\left(\frac{U}{\phi(U)}\right)^2 \ll (\log A)^3$ we finally obtain (5.2), as required.

that The value of $T(W)$ is

$$\frac{1}{2\phi(W)} \sum_{\substack{0 \leq r < 2W, \\ (r, 2W)=1.}} \sum_{q|2W} \sum_{\substack{0 \leq a < q, \\ (a, q)=1.}} \int_{-\frac{1}{M}}^{\frac{1}{M}} \widehat{\beta}(t) \sum_{x \in S^{11}} \prod_{i=1}^{11} \log x_i e(tf(x)) e\left(\frac{a(f(x) - r^2)}{q}\right) dt. \quad (5.64)$$

Finally, on interchanging summations and remarking that

$$\frac{1}{2W} \sum_{q|2W} \sum_{\substack{0 \leq a < q, \\ (a, q)=1.}} e\left(\frac{a(f(x) - r^2)}{q}\right) = \frac{1}{2W} \sum_{0 \leq a < 2W} e\left(\frac{a(f(x) - r^2)}{2W}\right) \quad (5.65)$$

we conclude that the right hand side of (5.64) is the same as the left hand side of

$$\begin{aligned} \frac{W}{\phi(W)} \sum_{\substack{0 \leq r < 2W, \\ (r, 2W)=1.}} \sum_{\substack{x \in S^{11}, \\ f(x) \equiv r^2 \pmod{2W}}} \log x_1 \log x_2 \dots \log x_{11} \int_{-\frac{1}{M}}^{\frac{1}{M}} \widehat{\beta}(t) e(tf(x)) dt \\ \leq \frac{W(\log N)^{11}}{\phi(W)} \sum_{\substack{0 \leq r < 2W, \\ (r, 2W)=1.}} \sum_{\substack{x \in S^{11}, \\ f(x) \equiv r^2 \pmod{2W}}} 1, \end{aligned} \quad (5.66)$$

where we have used $|\int_{-\frac{1}{M}}^{\frac{1}{M}} \widehat{\beta}(t) e(tf(x)) dt| \leq \int_{\mathbf{R}} \widehat{\alpha}(t) dt = 1$, since $|\widehat{\beta}(t)| = \widehat{\alpha}(t)$ for all $t \in \mathbf{R}$. For each invertible square b in $\mathbf{Z}/2W\mathbf{Z}$, the number of r in $[0, 2W)$ co-prime to $2W$ and such that $r^2 \equiv b$ modulo $2W$ is $2\tau(U)$. Then it follows from (5.66) and (5.64) that

$$T(W) \leq \frac{2W\tau(U)(\log N)^{11}}{\phi(W)} |\{x \in S^{11} \mid f(x) \text{ an invertible square mod } 2W\}|. \quad (5.67)$$

On combining (5.67) with (5.62), (5.33) and recalling that (5.18) is the same as the

then by (ii) of Lemma 5.3.3 combined with the triangle inequality and (5.22) we get

$$T_1 - T(W) \ll \frac{\|\widehat{\beta}\|_\infty |S|^9 (\log N)^9 A^3}{w^{97/200}} \ll \frac{A^3 |S|^9 (\log N)^9 N}{w^{97/200}}, \quad (5.61)$$

since $\|\widehat{\beta}\|_\infty = \sup_{t \in \mathbf{R}} |\widehat{\beta}(t)| \leq \frac{5N}{2}$. From (5.61), (5.59) and on recalling that $|S| \geq \frac{\sqrt{N}}{A \log N}$ and $w = A^{25}$ we conclude that

$$T = T(W) + O\left(\frac{|S|^{11}}{A}\right), \quad (5.62)$$

when N is sufficiently large, depending only on A . Let us now estimate $T(W)$. When $q|2W$ we have $\phi(qW) = q\phi(W)$ and $(r+mW)^2 \equiv r^2$ modulo q for all integers m and the condition $(r+mW) = 1$ holds always. Therefore we have $V_q(a, r) = qe\left(-\frac{ar^2}{q}\right)$ when $q|2W$, for all $0 \leq a < q$. Furthermore, since $r \mapsto r+W$ is a bijection from the integers co-prime to $2W$ in $[0, W)$ to those in $(W, 2W]$ co-prime to $2W$, we obtain

$$\frac{1}{\phi(qW)} \sum_{\substack{0 \leq r < W, \\ (r, W) = 1}} V_q(-a, r) = \frac{1}{2\phi(W)} \sum_{\substack{0 \leq r < 2W, \\ (r, 2W) = 1}} e\left(-\frac{ar^2}{q}\right) \quad (5.63)$$

for any $q|2W$ and all $0 \leq a < q$. Also, we have

$$\widehat{S}(t)^6 \widehat{S}(-t)^5 = \sum_{x \in S^{11}} \log x_1 \log x_2 \dots \log x_{11} e(f(x)t)$$

where $f(x)$ denotes $x_1 + x_2 + \dots + x_6 - x_7 - \dots - x_{11}$ for any $x = (x_1, \dots, x_{11}) \in S^{11}$.

By means of the change of variable $t - \frac{a}{q} \mapsto t$ in the integrals in (5.60) we then see

$$\frac{1}{\phi(qW)} |V_q(a, r)| \ll_{\epsilon} \frac{\log \log V}{\phi(W) V^{\frac{1}{2}-\epsilon}}. \quad (5.57)$$

Taking $\epsilon = 1/200$ and using the bound $\log \log V \leq V^{1/100}$ in (5.57) we conclude (ii).

5.3.3 The major arc contribution

In this subsection we reduce the problem of bounding $E_6(S)$ to a finite problem. Let us first dispose of the first term in (5.18), which we denote here by T . Then on writing T_1 for

$$\sum_{\substack{0 \leq r < W, \\ (r, W) = 1.}} \sum_{1 \leq q \leq Q} \frac{1}{\phi(qW)} \sum_{\substack{0 \leq a < q, \\ (a, q) = 1.}} V_q(-a, r) \int_{\mathfrak{M}(\frac{a}{q})} \widehat{\beta} \left(t - \frac{a}{q} \right) \widehat{S}(t)^6 \widehat{S}(-t)^5 dt \quad (5.58)$$

we deduce by substituting the complex conjugate of right hand side of (5.34) for $\psi(-t) = \overline{\psi}(t)$ in T and using the triangle inequality together with (5.22) that

$$\begin{aligned} T - T_1 &\ll \phi(W) N \exp(-c \sqrt{\log N}) \int_0^1 |\widehat{S}(t)|^{11} dt \\ &\ll \phi(W) N \exp(-c \sqrt{\log N}) |S|^9 (\log N)^9 A^3. \end{aligned} \quad (5.59)$$

If we now set

$$T(W) = \sum_{\substack{0 \leq r < W, \\ (r, W) = 1.}} \sum_{q|2W} \frac{1}{\phi(qW)} \sum_{\substack{0 \leq a < q, \\ (a, q) = 1.}} V_q(-a, r) \int_{\mathfrak{M}(\frac{a}{q})} \widehat{\beta} \left(t - \frac{a}{q} \right) \widehat{S}(t)^6 \widehat{S}(-t)^5 dt. \quad (5.60)$$

$$e\left(\frac{-ar^2}{q}\right) V_q(a, r) = U \sum_{\substack{0 \leq m_2 < V, \\ (r+m_2W, V)=1}} e\left(\frac{a_2W^2m_2^2 + 2a_2Wrm_2}{V}\right). \quad (5.53)$$

Multiplying by a function $e\left(\frac{-a_2r^2}{V}\right)$ both side of above equation and change the variable $r + m_2W \mapsto x$ in the summation on right of (5.53) and using the fact that $(V, W) = 1$, we get

$$e\left((-r^2(a_2/V + a/q))\right) V_q(a, r) = U \sum_{\substack{0 \leq x < V, \\ (x, V)=1}} e\left(\frac{a_2x^2}{V}\right). \quad (5.54)$$

We have a following bound on the summation on right of (5.54)

$$\sum_{\substack{0 \leq x < V, \\ (x, V)=1}} e\left(\frac{a_2x^2}{V}\right) \ll_{\epsilon} V^{\frac{1}{2}+\epsilon}, \quad (5.55)$$

see [16, Lemma 8.5], for example.

From (5.55) and (5.54), it follows that

$$\frac{1}{\phi(qW)} |V_q(a, r)| \ll_{\epsilon} \frac{UV^{\frac{1}{2}+\epsilon}}{\phi(qW)} = \frac{UV^{\frac{1}{2}+\epsilon}}{\phi(V)\phi(UW)}, \quad (5.56)$$

here we use the fact that $(V, W) = 1$ in the equality on the right of the above equation. Since $U|2W$, we have $\phi(UW) = U\phi(W)$ and we have the lower bound on $\phi(V)$, namely $\phi(V) \gg V/\log \log V$. From this and (5.56) we get

applying the Lemma 5.3.4, we get that

$$\sum_{0 \leq m < U} e\left(\frac{a_1 W^2 m^2 + 2a_1 W r m}{U}\right) = \left(\sum_{0 \leq m_1 < U_1} e\left(\frac{a_1 W^2 m_1^2 + 2a_1 W r m_1}{U}\right)\right) \times \left(\sum_{0 \leq m_2 < U_2} e\left(\frac{2a_1 W r m_2}{U_2}\right)\right). \quad (5.51)$$

We can conclude from (5.51), (5.50) and (5.49) that $V_q(a, r) = 0$ unless $U_2 | 2a_1 W r$. That is, unless $(U, W^2) | 2a_1 W r$ we have $V_q(a, r) = 0$.

Since $(a_1, U) = 1$ and $(r, W) = 1$, it follows that $V_q(a, r) = 0$ unless $(U, W^2) | 2W$. We note that $(U, W^2) = (q, W^2)$ and that $(q, W^2) | 2W$ is equivalent to $\inf(v_p(q), 2v_p(W)) \leq v_p(2W)$ for all primes $p | 2W$. From the definition of W we have $2v_p(W) > v_p(2W)$ for all primes $p | 2W$. Consequently, $V_q(a, r) = 0$ unless $v_p(q) \leq v_p(2W)$ for all primes $p | 2W$, which is the same as (i).

To prove (ii), we may assume that $(q, W^2) | 2W$ and $(q, W^2) | 2W$ and $V > 1$. We can conclude from $(q, W^2) = (U, W^2)$ and $(q, W^2) | 2W$ that $U | 2W$, in particular $U | W^2$. Thus, we have $(U, W^2) = U$. It follows that

$$\sum_{\substack{0 \leq m_1 < U, \\ (r + m_1 W, U) = 1}} e\left(\frac{a_1 W^2 m_1^2 + 2a_1 W r m_1}{U}\right) = U, \quad (5.52)$$

again using the same fact that $(r + m_1 W, U) = 1$ is always holds, as U is a w -smooth number. On combining (5.49) and (5.52) we get that

PROOF.— See [24, page 26], for example.

We now give a proof of the above Proposition with the aid of this lemma. Since $(r, W) = 1$, the condition $(r + mW, qW) = 1$ in the definition of $V_q(a, r)$ can be replaced by the condition $(r + mW, q) = 1$, thus we have

$$V_q(a, r) = \sum_{\substack{0 \leq m < q, \\ (r+mW, q)=1}} e\left(\frac{a(r+mW)^2}{q}\right). \quad (5.48)$$

Let $q = UV$, where U is w -smooth and $(V, W) = 1$ and let $a = a_2U + a_1V$, $(a_1, U) = 1$ and $(a_2, V) = 1$. Then from (5.48) follows that

$$e\left(\frac{-ar^2}{q}\right) V_q(a, r) = \left(\sum_{\substack{0 \leq m_2 < V, \\ (r+m_2W, V)=1}} e\left(\frac{a_2W^2m_2^2 + 2a_2Wrm_2}{V}\right) \right) \times \\ \left(\sum_{\substack{0 \leq m_1 < U, \\ (r+m_1W, U)=1}} e\left(\frac{a_1W^2m_1^2 + 2a_1Wrm_1}{U}\right) \right). \quad (5.49)$$

Now we analyze the second term in the product of right of the above equation. Since U is w -smooth and $(r, W) = 1$, the condition $(r + m_1W, U) = 1$ is always holds, thus we get

$$\sum_{\substack{0 \leq m < U, \\ (r+mW, U)=1}} e\left(\frac{a_1W^2m^2 + 2a_1Wrm}{U}\right) = \sum_{0 \leq m < U} e\left(\frac{a_1W^2m^2 + 2a_1Wrm}{U}\right). \quad (5.50)$$

We write $U = U_1U_2$, where $U_1 = \frac{U}{(U, W^2)}$, $U_2 = (U, W^2)$. Note that $U_2 | a_1W^2$, thus

and so

$$\sum_{n=3}^X \int_{n-1}^n \frac{f(n)}{\log x} dx = \int_2^X 2x \beta(x^2) e(\theta x^2) dx + O\left(\sqrt{N} (\log N)^{2B+1}\right). \quad (5.45)$$

Note that the integral on right of above is nothing but $\overline{\beta}(\theta)$. Thus, we have an asymptotic formula for the inner sum on right of (5.35) as follows

$$\sum_{n \equiv r \pmod{W}} 1_P(n) n 2n \log n \beta(n^2) e(n^2 t) = \frac{V_q(a, r)}{\phi(qW)} \overline{\beta}(\theta) + O\left(N \exp(-c\sqrt{\log N})\right), \quad (5.46)$$

for large enough N depends on only on A . Substituting this into (5.35) we get an asymptotic formula for $\psi(t)$ as in (5.34).

We need the following proposition, which provides information about $V_q(a, r)$.

Proposition 5.3.3. *Let a and q be integers satisfying (5.16) and r any integer co-prime to W . Then we have*

- (i) $V_q(a, r) = 0$ unless $q|2W$ or there is a prime $p > w$ such that $p|q$.
- (ii) $\frac{1}{\phi(qW)} |V_q(a, r)| \ll \frac{1}{\phi(W)w^{97/200}}$ when q does not divide $2W$.

We prove this Proposition with the help of following lemma.

Lemma 5.3.4. *Let $P(z) = c_0 z^2 + c_1 z + c_2$ be a polynomial with integer coefficients and let d be a positive integer with $d = d_1 d_2$ and d_2 divides c_0 . Then*

$$\sum_{0 \leq m < d} e\left(\frac{P(m)}{d}\right) = \sum_{0 \leq m_1 < d_1} e\left(\frac{P(m_1)}{d}\right) \sum_{0 \leq m_2 < d_2} e\left(\frac{c_1 m_2}{d_2}\right). \quad (5.47)$$

Using the functions $f(u)$ and S_n , we have

$$\sum_{n \equiv r \pmod{W}} 1_P(n) n 2n \log n \beta(n^2) e(n^2 t) = \sum_{n=2}^X (S_n - S_{n-1}) f(n) \quad (5.39)$$

$$= S_X f(X+1) + \sum_{n=2}^X S_n (f(n) - f(n+1)) . \quad (5.40)$$

As $|\theta| \leq \frac{(\log N)^{2B}}{N}$, the Mean-Value theorem implies that

$$f(n) - f(n+1) \ll (\log N)^{2B+1} . \quad (5.41)$$

Hence the sum on left of (5.39) becomes

$$\begin{aligned} & \frac{V_q(a, r)}{\phi(qW)} \left[Li(X) f(X+1) + \sum_{n=2}^X Li(n) (f(n) - f(n+1)) \right] \\ & + O \left(N \exp(-c\sqrt{\log N}) \right) . \end{aligned} \quad (5.42)$$

As $Li(2) = 0$, we then rewrite (5.42) as

$$\frac{V_q(a, r)}{\phi(qW)} \sum_{n=3}^X \int_{n-1}^n \frac{f(n)}{\log x} dx + O \left(N \exp(-c\sqrt{\log N}) \right) . \quad (5.43)$$

When $n-1 < x < n$, the Mean-value theorem reveals that

$$f(n) = f(x) + O \left((\log N)^{2B+1} \right) , \quad (5.44)$$

on recalling expression of $\psi(t)$, noticing the fact that all primes more than U are co-prime to W , and trivially estimating contribution to the sum over the interval $[0, U]$ using an upper bound 3^{A^l} on U .

We now find an asymptotic formula for the inner sum on right of (5.35). To this end, we let $X = \lceil \sqrt{5N} \rceil$, and for $n \in [1, X]$ let

$$S_n = \sum_{\substack{m \leq n, \\ m \equiv r \pmod{W}}} 1_P(n) e\left(\frac{am^2}{q}\right). \quad (5.36)$$

Using the fact that $q \leq Q$, we get

$$S_n = \sum_{\substack{0 \leq m < q, \\ (r+mW, qW)=1}} e\left(\frac{a(r+mW)^2}{q}\right) \sum_{\substack{d, \\ (r+mW)+dqW \leq 1, \\ (r+mW)+dqW \text{ is a prime}}} 1 + O(Q). \quad (5.37)$$

As $n \leq \sqrt{5N}$ and $qW \leq (\log N)^{B+1}$ for large values of N , by appealing to the SiegelWalfisz theorem, we get the asymptotic expression

$$S_n = \frac{Li(n)}{\phi(qW)} \sum_{\substack{0 \leq m < q, \\ (r+mW, qW)=1}} e\left(\frac{a(r+mW)^2}{q}\right) + O\left(\sqrt{N} \exp(-c\sqrt{\log N})\right), \quad (5.38)$$

where $Li(n) = \int_2^n \frac{dt}{\log t}$, and c is a positive absolute constant.

on A , we have

$$\begin{aligned} \int_{\mathfrak{m}} |\widehat{S}(t)|^{11} |\psi(t)| dt &\ll \frac{N}{A^6} \int_0^1 |\widehat{S}(t)|^{11} dt \\ &\ll \frac{N|S|^9(\log N)^9}{A^3} \ll \frac{|S|^{11}(\log N)^{11}}{A}, \end{aligned} \quad (5.32)$$

since $|S| \geq N^{\frac{1}{2}}/A \log N$. An application of the triangle inequality now allows us to conclude that

$$\int_{\mathfrak{m}} \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt \ll \frac{|S|^{11}(\log N)^{11}}{A}. \quad (5.33)$$

□

5.3.2 The function ψ on a major arc

Let us set $W = 2U$, where U is as defined starting of Section 5.3. For any integers a , q and r , with $q > 0$, we set $V_q(a, r) = \sum_{\substack{0 \leq m < q, \\ (r+mW, qW)=1}} e\left(\frac{a(r+mW)^2}{q}\right)$.

Proposition 5.3.2. *Let a and q be any integers satisfying (5.16). Then for all t in the major arc $\mathfrak{M}(\frac{a}{q})$ we have*

$$\psi(t) = \frac{1}{\phi(qW)} \sum_{\substack{0 \leq r < W, \\ (r, W)=1}} V_q(a, r) \overline{\beta}\left(t - \frac{a}{q}\right) + O\left(\phi(W)N \exp(-c\sqrt{\log N})\right). \quad (5.34)$$

PROOF.— Let $\theta = t - \frac{a}{q}$ and $f(u) = 2u \log u \beta(u^2) e(u^2 \theta)$ for any real u . we have

$$\psi(t) = \sum_{\substack{0 \leq r < W, \\ (r, W)=1}} \sum_{n \equiv r \pmod{W}} 1_P(n) n 2n \log n \beta(n^2) e(n^2 t) + O\left(9^{A^t}\right), \quad (5.35)$$

Again using the fact that the right of (5.27) is increasing function and recalling the values of Q, M ; we get that

$$\max_{0 \leq u \leq \sqrt{5N}} |T(u)| \ll \sqrt{N} (\log N)^{c+1} \left(\frac{1}{(\log N)^B A^{48}} + \frac{1}{N^{1/4}} + \frac{1}{(\log N)^{2B}} \right)^{\frac{1}{8}}. \quad (5.28)$$

For large values of N depends on A , large absolute value of B depends on c , we then get

$$\max_{0 \leq u \leq \sqrt{5N}} |T(u)| \ll \frac{\sqrt{N}}{A^6}, \quad (5.29)$$

this proves the lemma. □

Now we return to bound $\psi(t)$ on minor the arcs. By the Properties of Riemann-Stieltjes integral we have

$$\psi(t) = \int_0^{\sqrt{5N}} 2u \beta(u^2) dT(u), \quad (5.30)$$

where $T(u)$ defined as in Lemma 5.3.1. Thus, on integrating by parts and using the inequality (5.25), we have

$$\psi(t) \ll \sqrt{N} \max_{0 \leq u \leq \sqrt{5N}} |T(u)| \ll \frac{N}{A^6}, \quad (5.31)$$

on remarking that $2u\beta(u^2)$ is piecewise monotonic in the interval $[0, \sqrt{5N}]$.

From (5.23) and (5.19) it now follows that for all N large enough, depending only

\mathfrak{M} , we must then have $Q < q$ on account (5.16). We then conclude using $q^2 \leq qM$ that for each t in \mathfrak{m} there are integers a and $q \neq 0$ with $(a, q) = 1$ satisfying

$$\left|t - \frac{a}{q}\right| \leq \frac{1}{q^2} \quad \text{and} \quad Q < q \leq M. \quad (5.24)$$

To get a bound on $\psi(t)$ when $t \in \mathfrak{m}$, we appeal to the following lemma.

Lemma 5.3.1. *Let α be a real number such that*

$$\alpha = \frac{a}{b} + \lambda, \quad (a, q) = 1, \quad |\lambda| \leq \frac{1}{q^2}, \quad Q < q \leq M.$$

and let $T(u) = \sum_{0 \leq n \leq u} 1_P(n) \log n e(n^2 \alpha)$. Then we have

$$\max_{0 \leq u \leq \sqrt{5N}} |T(u)| \ll \frac{\sqrt{N}}{A^6}. \quad (5.25)$$

Proof. On the assumption on α , we have

$$\sum_{x < n \leq 2x} 1_P(n) \log n e(n^2 \alpha) \ll x (\log x)^c \left(\frac{1}{Q} + \frac{1}{x^{1/2}} + \frac{M}{x^2} \right)^{\frac{1}{8}}, \quad (5.26)$$

for some absolute constant $c > 0$; see [4, Lemma 2.1], for example. From this it follows that

$$T(x) = \sum_{0 \leq n \leq x} 1_P(n) \log n e(n^2 \alpha) \ll x (\log x)^{c+1} \left(\frac{1}{Q} + \frac{1}{x^{1/2}} + \frac{M}{x^2} \right)^{\frac{1}{8}}, \quad (5.27)$$

by dividing the interval $[0, x]$ into dyadic intervals $(\frac{x}{2^{j+1}}, \frac{x}{2^j}]$; $j = 0, 1, \dots, \log x$ and using the fact that the right of (5.26) is increasing function of x .

from $|S| \geq N^{\frac{1}{2}}/A(\log N)$ and

$$E_5(S) = \sum_{\substack{p_1^2+p_2^2+\dots+p_5^2=p_7^2+p_8^2+\dots+p_{10}^2, \\ p_i^2 \in S.}} \log p_1 \log p_2 \dots \log p_{10} \quad (5.20)$$

$$\leq (\log N)^{10} \sum_{1 \leq n \leq 20N} R_5^2(n) \ll N^{\frac{3}{2}} (\log N)^5 |S|^5, \quad (5.21)$$

where $R_5(n)$ denotes the number of representations of an integer n as a sum of five elements of S . To verify (5.21) we note that $R_5(n) = 0$ when $n > 20N$ and $R_5(n) \leq r_5(n)$, the number of representations of n as a sum of five squares of prime numbers, and we have that $r_5(n) \ll n^{\frac{3}{2}}/(\log n)^5$ [16, Theorem 11], by an application of the circle method. As a consequence of (5.19) we have

$$\sum_{1 \leq q \leq Q} \sum_{\substack{0 \leq a < q, \\ (a,q)=1.}} \int_{\mathfrak{M}(\frac{a}{q})} |\widehat{S}(t)|^{11} dt \leq \int_{-\frac{1}{M}}^{1-\frac{1}{M}} |\widehat{S}(t)|^{11} dt \ll |S|^9 (\log N)^9 A^3. \quad (5.22)$$

5.3.1 The minor arc contribution

Here we bound the second term in (5.18). Let us first verify that for all $t \in \mathfrak{m}$ we have

$$|\psi(t)| \ll \frac{N}{A^6}, \quad (5.23)$$

when N is large enough, depending only on A . Indeed, for any real t Dirichlet's approximation theorem gives a rational number $\frac{a}{q}$ satisfying $|t - \frac{a}{q}| \leq \frac{1}{qM}$ together with $1 \leq q \leq M$ and $(a, q) = 1$. When t is in \mathfrak{m} we see that $\frac{a}{q}$ is in $[0, 1]$ since $\mathfrak{m} \subseteq [\frac{1}{M}, 1 - \frac{1}{M}]$. Consequently, we also have $0 \leq a \leq q$. Since, however, t is not in

we call the interval $[\frac{a}{q} - \frac{1}{M}, \frac{a}{q} + \frac{1}{M})$ the major arc $\mathfrak{M}(\frac{a}{q})$. It is easily checked that distinct major arcs are in fact disjoint when $M > 2Q^2$, which holds when N is sufficiently large depending only on A . We denote by \mathfrak{M} the union of the family of major arcs $\mathfrak{M}(\frac{a}{q})$. Each interval in the complement of \mathfrak{M} in $[0, 1)$ is called a minor arc. We denote the union of the minor arcs by \mathfrak{m} .

We have

$$\int_0^1 \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt = \int_{-\frac{1}{M}}^{1-\frac{1}{M}} \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt, \quad (5.17)$$

by the periodicity of the integrand. From the definitions given above it is easily seen that the interval $[-\frac{1}{M}, 1 - \frac{1}{M})$ is the union of \mathfrak{m} and $\mathfrak{M} \setminus [1 - \frac{1}{M}, 1 + \frac{1}{M})$. Since distinct major arcs are disjoint, it then follows that the right hand side of (5.17) is the same as

$$\sum_{1 \leq q \leq Q} \sum_{\substack{0 \leq a < q, \\ (a, q) = 1}} \int_{\mathfrak{M}(\frac{a}{q})} \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt + \int_{\mathfrak{m}} \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt. \quad (5.18)$$

We shall presently estimate each of the two terms in (5.18). We begin by observing that

$$\int_0^1 |\widehat{S}(t)|^{11} dt \ll |S|^9 (\log N)^9 A^3. \quad (5.19)$$

In effect, the integral in (5.19) does not exceed $|S| \log N E_5(S)$. Thus (5.19) follows

Indeed,

$$E_6(S) = \sum_{\substack{p_1^2+p_2^2+\dots+p_6^2=p_7^2+p_8^2+\dots+p_{12}^2, \\ p_i^2 \in S.}} \log p_1 \log p_2 \dots \log p_{12} \quad (5.11)$$

$$\leq \sum_{\substack{p_1^2+p_2^2+\dots+p_6^2=p_7^2+p_8^2+\dots+p_{11}^2+q^2, \\ p_i^2 \in S; q^2 \text{ is a prime square in } (N, 4N].}} \log p_1 \log p_2 \dots \log p_{11} \log q \quad (5.12)$$

$$\leq 1/2\sqrt{N} \sum_{\substack{p_1^2+p_2^2+\dots+p_6^2=p_7^2+p_8^2+\dots+p_{11}^2+q^2, \\ p_i^2 \in S; q^2 \text{ is a prime square in } (N, 4N].}} \log p_1 \log p_2 \dots \log p_{11} 2q \log q \quad (5.13)$$

$$\leq 5/4\sqrt{N} \sum_{\substack{p_1^2+p_2^2+\dots+p_6^2=p_7^2+p_8^2+\dots+p_{11}^2+q^2, \\ p_i^2 \in S; q^2 \text{ is a prime square.}}} \log p_1 \log p_2 \dots \log p_{11} 2q \log q \beta(q^2) \quad (5.14)$$

$$\leq 5/4\sqrt{N} \int_0^1 \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt, \quad (5.15)$$

in above inequalities: from (5.13) to (5.14) we use the lower bound $2/5$ on $\beta(t^2)$ in the interval $(N, 4N]$ and from (5.14) to (5.15) we use the orthogonality of the functions $t \mapsto e(nt)$ on $[0, 1]$.

We apply the circle method to estimate the integral on the right hand side of (5.10). To this end, we set $Q = (\log N)^B A^{48}$, $M = \frac{N}{(\log N)^{2B}}$, where B is a large absolute constant and, for any integers a and q satisfying

$$0 \leq a \leq q \leq Q \text{ and } (a, q) = 1, \quad (5.16)$$

therefore that the right hand side of (5.8) does not exceed $\frac{4|T_{c(i)}(\mathcal{X}, \mathcal{Y})||Z|^2}{|\mathcal{X}||\mathcal{Y}|}$. Using this in (5.6) together with the bound supplied by (2.39) for $|T_{c(i)}(\mathcal{X}, \mathcal{Y})|$, applicable since $3A \geq e^{e^2}$, we then conclude that (5.5) holds. □

5.3 An application of the Circle method

We prove Theorem 5.1.2 in this section. As stated in Section 5.1, we will first reduce the problem of bounding $E_6(S)$ to Theorem 5.2.1. This is carried out in Subsections 5.3.1 through 5.3.3 starting with the preliminaries given below. We then complete the proof of Theorem 5.1.2 in Subsection 5.3.4 by applying Theorem 5.2.1.

We suppose that $A \geq e^{e^2}$ are real number and assume that N is a sufficiently large integer depending only on A , its actual size varying to suit our requirements at various stages of the argument. We set $\alpha(t) = 1 - \left| \frac{2t}{5N} \right|$ when $|t| \leq \frac{5N}{2}$ and 0 for all other $t \in \mathbf{R}$ and set $\beta(t) = \alpha(t - \frac{5N}{2})$. Thus $\beta(t) \geq 0$ for all t in \mathbf{R} and $\beta(t) \geq \frac{2}{5}$ when $t \in [N, 4N]$. Finally, we set

$$\psi(t) = \sum_n 1_{\mathbb{P}}(n) 2n \log n \beta(n^2) e(n^2 t) \tag{5.9}$$

and write $\widehat{S}(t) = \sum_{p^2 \in S} \log p e(p^2 t)$ for any $t \in \mathbf{R}$ for a given subset S of the squares in $(N, 4N]$ satisfying the hypotheses of Theorem 5.1.2. We observe that

$$\frac{4}{5} \sqrt{N} E_6(S) \leq \int_0^1 \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt. \tag{5.10}$$

$a^2 + b^2 + c(i)$ is an invertible square in $\mathbf{Z}/U\mathbf{Z}$ and 0 otherwise. Further, we write $m(a)$ for the number of z in \mathcal{Z} such that $z \equiv a \pmod{U}$. Then if $\tilde{\mathcal{Z}}$ denotes the image of \mathcal{Z} in $\mathbf{Z}/U\mathbf{Z}$ we have

$$|R_U(\mathcal{Z}, \mathbf{c})| = \sum_{i \in I} \sum_{(a,b) \in \tilde{\mathcal{Z}}^2} \alpha_i(a,b) m(a)m(b). \quad (5.6)$$

Moreover, on account of the second assumption in (5.4) we have that

$$\sum_{a \in \tilde{\mathcal{Z}}} m(a) = |\mathcal{Z}| \quad \text{and} \quad 0 \leq m(a) \leq D, \quad (5.7)$$

where $D = \frac{3\sqrt{N}}{\phi(U)\log N}$. For large values of N , depending on A ; $\tilde{\mathcal{Z}}$ is contained in $(\mathbf{Z}/U\mathbf{Z})^*$. Let us bound the inner sum on the right hand side of (5.6) for a fixed i in I . By means of Lemma 2.4.1 and (5.7) we obtain

$$\sum_{(a,b) \in \tilde{\mathcal{Z}}^2} \alpha_i(a,b) m(a)m(b) \leq \sum_{(a,b) \in \tilde{\mathcal{Z}}^2} \alpha_i(a,b) x_a^* y_b^*, \quad (5.8)$$

for some x_a^* and y_b^* , with a and b varying over $\tilde{\mathcal{Z}}$, satisfying the following conditions. All the x_a^* , and similarly all the y_b^* , are either 0 or D excepting at most one, which must lie in $(0, D)$. Moreover, if \mathcal{X} and \mathcal{Y} are, respectively, the subsets of $\tilde{\mathcal{Z}}$ for which $x_a^* \neq 0$ and $y_b^* \neq 0$ then $|\mathcal{X}|D \geq |\mathcal{Z}| > (|\mathcal{X}| - 1)D$. From the first condition in (5.4) we then get $|\mathcal{X}| \geq \frac{|\mathcal{Z}|}{D} \geq \frac{\phi(U)}{3A} \geq 2$. Consequently, we also have $D \leq \frac{|\mathcal{Z}|}{|\mathcal{X}| - 1} \leq \frac{2|\mathcal{Z}|}{|\mathcal{X}|}$. The same inequalities hold with $|\mathcal{X}|$ replaced by $|\mathcal{Y}|$. Then with $T_{c(i)}(\mathcal{X}, \mathcal{Y})$ as in the Section 2.3 of the Chapter 2 we have that $\sum_{(a,b) \in \mathcal{X} \times \mathcal{Y}} \alpha_i(a,b) = |T_{c(i)}(\mathcal{X}, \mathcal{Y})|$ and

in the form of the finite addition theorem of Sárközy [30]. We give details of this deduction, whose principle goes back to [13], in Section 5.4.

5.2 The finite problem

The main result of this section is Theorem 5.2.1, which is a slightly modification of Theorem 2.1 of [24]. For the sake of completeness, we will provide a full proof of Theorem 5.2.1, which as the reader may expect, runs in parallel with that of Theorem 4.1.2. We begin with some notation. We shall suppose that $A \geq e^{e^2}$ and let $U = \prod_{p \leq w} p$, where $w = A^{25}$. In addition, we let \mathcal{Z} be a set of primes in the interval $(\sqrt{N}, 2\sqrt{N}]$ with

$$|\mathcal{Z}| \geq \frac{\sqrt{N}}{A \log N} \quad \text{and} \quad |\{z \in \mathcal{Z} | z \equiv a \pmod{U}\}| \leq \frac{3\sqrt{N}}{\phi(U) \log N}, \quad (5.4)$$

for all classes a in $\mathbf{Z}/U\mathbf{Z}$. Also, we denote by $\mathbf{c} = \{c(i)\}_{i \in I}$ a given finite sequence of integers and let $R_U(\mathcal{Z}, \mathbf{c})$ denote the set of triples (x, y, i) in $\mathcal{Z} \times \mathcal{Z} \times I$ such that $x^2 + y^2 + c(i)$ is an invertible square modulo U . Finally, let $\tau(U) = 2^{\pi(w)}$ be the number of divisors of U . Now we can state the theorem as follows.

Theorem 5.2.1. *We have*

$$|R_U(\mathcal{Z}, \mathbf{c})| \leq \left(\frac{U}{\phi(U)} \right)^2 \frac{|\mathcal{Z}|^2 |I|}{\tau(U)} \exp \left(\frac{(3 \log 2 + o(1)) \log 3A}{\log \log 3A} \right), \quad (5.5)$$

where $o \ll \frac{(\log \log \log 3A)}{\log \log 3A}$.

Proof. Let a, b be any elements of $\mathbf{Z}/U\mathbf{Z}$. For any i in I we set $\alpha_i(a, b) = 1$ if

where

$$\psi(t) = \sum_n 1_{\mathbb{P}}(n) 2n \log n \beta(n^2) e(n^2 t)$$

and $\widehat{S}(t) = \sum_{p^2 \in S} \log p e(p^2 t)$ for any $t \in \mathbf{R}$, which is given by (5.10) in the Section 5.3. We then split the integral into minor and major arcs whose definitions are given in the Section 5.3. First we note the inequality

$$\int_0^1 |\widehat{S}(t)|^{11} dt \ll |S|^9 (\log N)^9 A^3, \quad (5.3)$$

which follows from the bound $r_5(n) \ll n^{3/2}/\log^5 n$, where $r_5(n)$ is the number of ways n can be written as sum of 5 prime squares. For the minor arcs contribution we use the estimates on the exponential sums over prime squares and (5.3), details are given in the subsection 5.3.1. And major arcs analysis lead us to estimate of certain cardinalities, details are given in the Subsections 5.3.2 and 5.3.3. Finally on combining minor and major arcs estimates we get

$$E_6(S) \ll \frac{2W\tau(U)(\log N)^{11}}{\phi(W)} |\{x \in S^{11} \mid f(x) \text{ an invertible square mod } 2W\}| \\ + O\left(\frac{|S|^{11}(\log N)^{11}}{A}\right),$$

where $W = 2 \prod_{p \leq A^{25}} p$ and $f(x) = x_1 + x_2 + \dots + x_6 - x_7 - \dots - x_{11}$. We then complete the proof of Theorem 5.1.2 in the Subsection 5.3.4, by giving estimation on the above cardinality.

Finally, Theorem 5.1.1 is deduced from Theorem 5.1.2 by means of a classical application of the Cauchy-Schwarz inequality and the use of additive combinatorics

Theorem 5.1.1. *For any integer $K \geq 2$ we have $r_{\mathcal{D}}(K) \leq K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$ when \mathcal{D} is the set of squares of the prime numbers.*

We will prove Theorem 5.1.1 via the theorem below, which we state with the help of following notation. For any subset S of the squares of primes in the interval $(N, 4N]$, we shall write

$$E_6(S) = \sum_{\substack{p_1^2 + p_2^2 + \dots + p_6^2 = p_7^2 + p_8^2 + \dots + p_{12}^2, \\ p_i^2 \in S, 1 \leq i \leq 12}} \log p_1 \log p_2 \dots \log p_{12}. \quad (5.1)$$

Theorem 5.1.2. *Let $A \geq e^{e^2}$ be real number. Then for all sufficiently large integers N , depending only on A , and any subset S of the squares of primes in the interval $(N, 4N]$ with $|S| \geq \frac{N^{\frac{1}{2}}}{A \log N}$ we have*

$$E_6(S) \ll \frac{|S|^{11} (\log N)^{11}}{N^{\frac{1}{2}}} \exp\left(\frac{(3 \log 2 + o(1)) \log A}{\log \log A}\right), \quad (5.2)$$

where $o(1) \ll \frac{\log \log \log A}{\log \log A}$.

Our proofs of Theorems 5.1.1 and 5.1.2 are an adaptation of the method of Gyan Prakash, Ramana and Ramaré [24] to the case of the squares of the prime numbers. Now we sketch the proof of Theorem 5.1.2. To prove this theorem we apply the circle method, suggested by [6]. Our starting point is the following inequality

$$\frac{4}{5} \sqrt{N} E_6(S) \leq \int_0^1 \widehat{S}(t)^6 \widehat{S}(-t)^5 \psi(-t) dt$$

the following question. What is the smallest integer $r_{\mathcal{D}}(K)$, if it exists, such that given any colouring, or partition, of \mathcal{D} in K colours, every sufficiently large integer is expressible as a sum of at most $r_{\mathcal{D}}(K)$ elements of \mathcal{D} , all of the same colour ?

When \mathcal{D} is either the set of squares or the set of primes the aforementioned question was posed as Problems 39 and 40 by A. Sárközy on page 26 of [31]. It is easily seen that indeed $r_{\mathcal{D}}(K)$ is finite for all $K \geq 1$ in these cases. In [13] N. Hegyváry and F. Hennecart showed that $r_{\mathcal{D}}(K) \ll (K \log K)^5$ when \mathcal{D} is the set of squares and $r_{\mathcal{D}}(K) \ll K^3$ when \mathcal{D} is the set of primes. Ramana and Ramaré [25] then obtained $r_{\mathcal{D}}(K) \ll K \log \log 2K$ when \mathcal{D} is the set of primes, which is the best possible bound up to the implied constant. Recently, Gyan Prakash, Ramana and Ramaré [24] showed that $r_{\mathcal{D}}(K) \ll K \exp\left(\frac{(3 \log 2 + o(1)) \log K}{\log \log K}\right)$ when \mathcal{D} is the set of squares. By a lower bound in [13], the optimal bound in this case is expected to have a $\log 2$ in place of the $3 \log 2$ in the exponential factor.

A classical theorem of L.K. Hua implies that the set of squares of primes is an asymptotic basis of finite order. Indeed, Hua's theorem tells that every sufficiently large integer n which is congruent to 5 modulo 24 can be written as sum of at most 5 prime squares and hence any large integer is a sum of at most 9 prime squares. F. Hennecart asked (orally) if one may extend Sárközy's problems to the case when \mathcal{D} is the set of squares of the prime numbers. Independently of Hennecart, this question was considered by Guohua Chen [4], who showed that $r_{\mathcal{D}}(K) \ll_{\epsilon} K^{2+\epsilon}$. The main result of this chapter improves on this conclusion of Chen. More precisely, and in analogy with the result of [24] for the squares, we prove the following theorem.

Monochromatic sums of squares of primes

5.1 Introduction

A subset \mathcal{D} of the set of natural numbers is said to be an asymptotic basis of finite order if there exists a positive integer m such that every sufficiently large integer can be written as a sum of at most m elements of \mathcal{D} . The smallest m for which the above property holds is called the order of the asymptotic basis \mathcal{D} . There are two classical examples of asymptotic bases of finite order. The first of these is the set of squares of the natural numbers, which by Lagrange's four squares theorem, is certainly an asymptotic basis of order four. The second example is the set of the prime numbers, which is seen to be an asymptotic basis of order at most 4 by the classical theorem of Vinogradov which asserts that every sufficiently large odd integer can be written as a sum of three prime numbers.

Given an asymptotic basis \mathcal{D} of finite order and an integer $K \geq 1$ one may ask

Then we get a lower bound

$$\Lambda(\lambda_A, \dots, \lambda_A) \geq \exp\left(-c(V, \Psi, F, m, \kappa) \alpha^{-8m-\kappa}\right)$$

By Lemma 6.4.3 and since $\lambda_A \leq \left(\frac{\log N}{\log w}\right)^k 1_A$, we have

$$|\{y \in A^m : Vy^t = 0\}| \geq \exp\left(-c(V, \Psi, F, m, \kappa) \alpha^{-8m-\kappa}\right) N^{m-r} \left(\frac{\log N}{\log w}\right)^{-mk}.$$

Note that the number of $y \in [N]^m$ with two identical coordinates and such that $Vy = 0$ is $\ll N^{m-r-1}$. Let us take $\kappa = m$, then we see that the above cardinality is $\geq C N^{m-r-1}$ for any large constant C . Thus we see that there exists at least one $y \in A^m$ with distinct coordinates such that $Vy^t = 0$.

a D -pseudorandom weight $\nu : \mathbf{Z}_M \rightarrow \mathbf{R}^+$ of level $(\log N)^{1-o(1)}$ such that

$$0 \leq \lambda_A \ll \lambda_{b,W} \ll \nu.$$

Let $\nu' = \frac{1}{2}(\nu + \nu * \mu_B)$, so that $|\lambda'_A| \ll \nu'$ and $|\lambda_A - \lambda'_A| \ll \nu'$. By Proposition 6.5.2, ν' is also D -pseudorandom of level $(\log N)^{1-o(1)}$.

We recall now that Ψ is in exact 1-normal form at i_0 . Applying Proposition 6.5.3 to the functions f_1, \dots, f_m , and inserting the estimate on $\|\lambda_A - \lambda'_A\|_{U^2}$ from Proposition 6.5.6, we obtain the desired result. \square

6.5.5 End of the proof of Theorem 6.5.1

By recalling the expression of $\Lambda(f_1, \dots, f_m)$ from (6.5.1), we see that

$$\Lambda(\lambda_A, \dots, \lambda_A) = \Lambda(\lambda'_A, \dots, \lambda'_A) + \sum \Lambda(f_1, \dots, f_m),$$

where the above sum varies over $f_i \in \{\lambda'_A, \lambda_A - \lambda'_A\}$ and at least one of which is $\lambda_A - \lambda'_A$. By Lemma 6.5.10 and Proposition 6.5.8 we see that for any $\kappa > 0$

$$\Lambda(\lambda_A, \dots, \lambda_A) \geq \exp\left(-c(V, \Psi, F, m, \kappa) \alpha^{-8m-\kappa}\right) - O(\epsilon^{1/4} + \delta^{1/4} + (\log N)^{\frac{-1}{4}+o(1)}),$$

whenever $\delta^{-4} \log \epsilon^{-1} \leq c \log N$. We choose $\epsilon = \delta = \exp(-c'(V, \Psi, F, m, \kappa) \alpha^{-8m-\kappa})$ for a some large constant $c'(V, \Psi, F, m, \kappa)$, and assume that

$$\alpha \geq c(V, \Psi, F, m, \kappa) (\log \log N)^{\frac{-1}{8m+\kappa}}.$$

therefore in $[-2N, 2N]$. Since $\lambda'_A \geq \frac{\alpha}{2} 1_{A'}$, we have

$$\Lambda(\lambda'_A, \dots, \lambda'_A) \geq \left(\frac{\alpha}{2}\right)^m \Lambda(1_{A'}, \dots, 1_{A''}). \quad (6.128)$$

By the above Proposition 6.5.7, we know that A' has density $\gg_{V, \Psi, F, \kappa} \alpha^{1+\kappa}$ in $[-2N, 2N]$ for any $\kappa > 0$. By Lemma 6.4.3 and Theorem 6.1.3 we obtain

$$\begin{aligned} \Lambda(1_{A'}, \dots, 1_{A'}) &= M^{m-r} \left| \left\{ y \in (A')^m : Vy = 0 \right\} \right| \\ &\gg \exp\left(-c \alpha^{-(1+\kappa)8m} \log \frac{1}{\alpha}\right) \\ &\gg \exp\left(-c \alpha^{-8m-\kappa} \log \frac{1}{\alpha}\right), \end{aligned}$$

where c depends at most on V, Ψ, F, m and κ . Thus writing this lower bound in (6.128), we get the assertion of the proposition. □

Now we compare $\Lambda(\lambda_A, \dots, \lambda_A)$ with $\Lambda(\lambda'_A, \dots, \lambda'_A)$.

Lemma 6.5.10. *Suppose that f_1, \dots, f_m are functions all equals to λ'_A or $\lambda_A - \lambda'_A$ with at least one of them equals to $\lambda_A - \lambda'_A$. Then we have*

$$\Lambda(f_1, \dots, f_m) \ll \epsilon^{1/4} + \delta^{1/4} + (\log N)^{\frac{-1}{4}+o(1)}. \quad (6.129)$$

Proof. Let us consider $i_0 \in [m]$ such that $f_{i_0} = \lambda_A - \lambda'_A$. Let $Q = \|\theta\|$ and let $D = D_{d, m, Q}$ be a constant from Proposition 6.5.3. By Proposition 6.5.5 there exists

Moreover, λ'_A is “large” on the integers. More precisely, we have

Proposition 6.5.7. *For any $\kappa > 0$, the level set $A' = \{\lambda'_A \geq \alpha/2\}$ has density $\gg_\kappa \alpha^{1+\kappa}$ in \mathbf{Z}_M .*

Proof. Having bound on the L^{2l} -norm of λ'_A , we can get a lower bound on the density of A' in \mathbf{Z}_M . In fact we have

$$\alpha = \mathbf{E}_{x \in \mathbf{Z}_M} \lambda'_A(x) + \mathbf{E}_{x \in \mathbf{Z}_M} \lambda'_A(x) 1_{A'}(x).$$

By an application of Hölder’s inequality we get

$$\alpha/2 \leq \|1_{A'}\|_r \|\lambda'_A\|_s$$

where $\frac{1}{r} + \frac{1}{s} = 1$. Note that if s approaches infinity then r approaches 1 from the right.

Thus the proof of the proposition follows by noting the bound $\|\lambda'_A\|_{2l} \ll_k l^k$. \square

We have a lower bound on the average of λ'_A over the Ψ -configuration.

Proposition 6.5.8. *Suppose $\delta^{-4} \log \epsilon^{-1} \leq c \log N$, then we have*

$$\Lambda(\lambda'_A, \dots, \lambda'_A) \geq \exp\left(-c \alpha^{-8m-\kappa} \log \frac{1}{\alpha}\right), \quad (6.127)$$

where the constant $c = c(V, \Psi, F, m, \kappa)$ depends at most on V, Ψ, F, m, κ .

Proof. Consider the level set $A' = \{\lambda'_A \geq \alpha/2\}$ contained in the support of λ'_A and

and get

$$\begin{aligned}
\sum_{m \in \mathbf{Z}_M} \left| \hat{\lambda}_A(m) \right|^q &= \sum_{m \in \mathbf{Z}_M} \left| \frac{1}{M} \sum_{n \in \mathbf{Z}_M} \lambda_A(n) e\left(\frac{-nm}{M}\right) \right|^{q/2} \\
&\ll_q \left(\frac{1}{M} \sum_{n \in \mathbf{Z}_M} \frac{\lambda_A^2(n)}{\beta(n)} \right)^{q/2} \\
&\ll_q \left(\frac{1}{M} \sum_{n \in \mathbf{Z}_M} \lambda_A(n) \right)^{q/2} \\
&\ll_{k,q} 1
\end{aligned}$$

where we use the lower bound on $\beta(n)$ whenever $n \in A$ and the fact that $\mathbf{E}_{x \in \mathbf{Z}_M} \lambda_A(n) = \alpha \leq 1$. \square

We now see that λ'_A approximates λ_A in a Fourier l^4 sense. In fact we have that $\|\lambda_A - \lambda'_A\|_{U^2} \ll \epsilon^{1/4} + \delta^{1/4}$. To see this, by the above Lemma 6.5.9, we have $\sum_r |\hat{\lambda}_A(r)|^q \ll_{k,q} 1$ for any $q > 2$. Therefore,

$$\begin{aligned}
\|\lambda_A - \lambda'_A\|_{U^2}^4 &= \sum_r |\hat{\lambda}_A(r)|^4 |1 - \hat{\mu}_B(r)|^4 \\
&\ll \epsilon \sum_{r: |\hat{\lambda}_A(r)| \geq \delta} |\hat{\lambda}_A(r)|^4 + \delta \sum_{r: |\hat{\lambda}_A(r)| \leq \delta} |\hat{\lambda}_A(r)|^3 \\
&\ll \epsilon + \delta
\end{aligned}$$

where we used the fact that $|1 - \hat{\mu}_B(r)| = |\mathbf{E}_{x \in \mathbf{Z}_M} 1 - e(\frac{-rx}{M})| \ll \epsilon$ for all $r \in \Gamma$. \square

for any complex sequence (b_n) , where $\beta(n)$ is an enveloping sieve function with $R = M^{1+c_0/80k+20}$. This means that $\beta : \mathbf{Z}^+ \rightarrow \mathbf{R}$ is a non-negative function satisfying the majorant property

$$\beta(n) \gg G_F^{-1} \log^k R 1_{X_{R!}}(n) \quad (6.126)$$

with $G_F = \prod_p \frac{\gamma(p)}{(1-1/p)^k}$, where

$$\gamma(p) = \frac{1}{p} |\{n \in \mathbf{Z}_p : (p, F(n)) = 1\}| = \begin{cases} 1 - k/p & \text{if } p > w \\ 1 & \text{if } p \leq w \end{cases}$$

and $X_{R!} = \{n \in \mathbf{Z} : (F(n), d) = 1, \forall d \leq R\}$. In particular, for any integer $n \in A$, we have $n \in X_{R!}$ and

$$\beta(n) \gg \log^k R \prod_p \frac{(1-1/p)^k}{\gamma(p)} \gg \left(\frac{\log R}{\log w}\right)^k.$$

We apply (6.125) to the sequence (b_n) defined by

$$b_n = \begin{cases} \frac{\lambda_A(n)}{\beta(n)} & \text{if } n \in A \\ 0 & \text{otherwise} \end{cases}$$

which implies that

$$\|\lambda'_A\|_{2l} \ll c(k)^k ((2kl)!)^{1/2l} + \frac{1}{|B|^{1/2l}} \left(\frac{\log N}{\log w} \right)^{(1-\frac{1}{2l})k}.$$

From this, (6.110) follows by recalling the trivial bound $n! \leq n^n$ on the factorial of n .

Note that $|B| \geq N^{1/2}$ by the condition on ϵ and δ . Thus we get the bound $\|\lambda'_A\|_{2l} \ll_k l^k$ for large integers l .

STEP 3: We see that λ'_A is close to λ_A in a Fourier l^4 sense. To see this we need the following restriction estimate of Green and Tao.

Lemma 6.5.9. *We have*

$$\sum_r |\hat{\lambda}_A(r)|^q \ll_{q,k} 1 \tag{6.124}$$

for any $q > 2$.

Proof. We apply the Proposition 4.2 of [10] with

$$F(n) = \prod_{i=1}^k (a_i W n + a_i b + b_i) \quad \text{and} \quad R = M^{\frac{1+c_0}{80k+20}} \leq M^{1/20}.$$

Then we get, for any $q > 2$,

$$\sum_{m \in \mathbf{Z}_M} \left| \frac{1}{M} \sum_{n=1}^M b_n \beta(n) e\left(\frac{-mn}{M}\right) \right|^q \ll_q \left(\frac{1}{M} \sum_{n=1}^M |b_n|^2 \beta(n) \right)^{q/2} \tag{6.125}$$

Now we estimate the product in (6.121). We have

$$\prod_p \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-2kl} \ll (\log w)^{2kl} \prod_{p>w} \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-2kl}.$$

Since $\prod_{p>w} \left(1 - \frac{2kl}{p}\right) \left(1 - \frac{1}{p}\right)^{-2kl} \asymp 1$, the contribution to the above product comes from those $p > w$ which divide $y_i - y_j$ for some $i \neq j$.

Since $|y_i - y_j| \leq M$, and any integer $y \leq M$ can have at most $\log M / \log w$ prime factors greater than w , we see that there are at most $4l^2 \frac{\log M}{\log w}$ primes greater than w which could divide some difference $y_i - y_j$. For each p that divides $y_i - y_j$ for some $i \neq j$, we take the smallest possible value k for $\rho(p)$. Thus, by noticing the bound $(1 - 1/p)^{-1} \leq \frac{w}{w-1}$ for $p > w$, we have

$$\prod_{\substack{p>w \\ p|y_i-y_j}} \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-2kl} \leq \left(\frac{w}{w-1}\right)^{4l^2(2kl-1)\frac{\log M}{\log w}}. \quad (6.122)$$

By our assumption of on the bound $l \leq \frac{1}{20} \log^{1/3} w$, we see that the right hand side of (6.122) is at most $5e^{c(k)}$. Thus, we get an upper bound

$$|\mathcal{A}(y_1, \dots, y_{2l})| \ll c(k)^{2kl} (2kl)! \frac{M}{(\log N)^{2kl}}. \quad (6.123)$$

From (6.123), (6.120), (6.117), (6.119) and (6.118), we see that

$$\|\lambda'_A\|_{2l}^{2l} \ll c(k)^{2kl} (2kl)! + \frac{l}{|B|} \left(\frac{\log N}{\log w}\right)^{(2l-1)k}$$

we see that the term, given by (6.117), is

$$\ll_k \left(\frac{\log N}{\log w} \right)^{2lk-k} \sum_{r=1}^{2l-1} \frac{1}{M^{2l}} \sum_{(y_1, \dots, y_{2l}) \in I_r} \mu_B(y_1) \cdots \mu_B(y_{2l}) \quad (6.119)$$

Each term in the sum over r may be bounded by $\frac{1}{|B|^{2l-r}}$. We can assume that $|B| \geq 2$. Thus the quantity given in (6.119) is

$$\ll_k \left(\frac{\log N}{\log w} \right)^{2lk-k} \frac{l}{|B|}.$$

Now we bound the term given in (6.117). We observe that this quantity is at most

$$\frac{1}{M^{2l+1}} \left(\frac{\log N}{\log w} \right)^{2lk} \sum_{(y_1, \dots, y_{2l}) \in J_{2l}} \mu_B(y_1) \cdots \mu_B(y_{2l}) |\mathcal{A}(y_1, \dots, y_{2l})|. \quad (6.120)$$

So to bound this, we first need to estimate the cardinality $|\mathcal{A}(y_1, \dots, y_{2l})|$, when $(y_1, \dots, y_{2l}) \in J_{2l}$. This can be done by applying Klimov's lemma. Note that the hypotheses of the Klimov's lemma are satisfied by our parameters. Thus we get

$$|\mathcal{A}(y_1, \dots, y_{2l})| \ll c(k)^{2kl} (2kl)! \frac{M}{(\log N)^{2kl}} \prod_p \left(1 - \frac{\rho(p)}{p} \right) \left(1 - \frac{1}{p} \right)^{-2kl} \quad (6.121)$$

The function ρ which appears on the right hand side of (6.121) takes the values as follows: $\rho(p) = 0$ if $p \leq w$, $\rho(p) = 2kl$ if $p > w$ and $p \nmid y_i - y_j$ for all $i \neq j$.

Applying Klimov's lemma, we get

$$|\mathcal{A}(y_1, \dots, y_{2l})| \ll M \left(\frac{\log w}{\log v} \right)^k \quad (6.115)$$

$$\ll c(k) M \left(\frac{\log w}{\log N} \right)^k. \quad (6.116)$$

Actually, we can apply Klimov's lemma in the case of $k \geq 2$. However (6.116) is true even if $k = 1$, by the Brun-Titchmarsh inequality.

Let

$$I_s = \{(y_1, \dots, y_{2l}) : y_i \leq M/2 \text{ with at most } s \text{ distinct coordinates } y_i\}$$

and

$$J_s = \{(y_1, \dots, y_{2l}) : y_i \leq M/2 \text{ with exactly } s \text{ distinct coordinates } y_i\}$$

for any $1 \leq s \leq 2l$. Thus, (6.113) can be at most

$$\frac{1}{M^{2l+1}} \sum_{r=1}^{2l-1} \sum_{(y_1, \dots, y_{2l}) \in I_r} \mu_B(y_1) \cdots \mu_B(y_{2l}) \sum_x \lambda_A(x - y_1) \cdots \lambda_A(x - y_{2l}) \quad (6.117)$$

$$+ \frac{1}{M^{2l+1}} \sum_{(y_1, \dots, y_{2l}) \in J_{2l}} \mu_B(y_1) \cdots \mu_B(y_{2l}) \sum_x \lambda_A(x - y_1) \cdots \lambda_A(x - y_{2l}). \quad (6.118)$$

We now estimate the term given in (6.117). Using the trivial bound on

$$\sum_x \lambda_A(x - y_1) \cdots \lambda_A(x - y_{2l}) \ll_k \left(\frac{\log N}{\log w} \right)^{2lk-k}$$

$n \in \{1, \dots, p\}$ such that

$$\prod_{i=1}^r (q_i n + l_i) \equiv 0 \pmod{p}.$$

We first prove that for any $l \leq \frac{1}{20} \log^{1/3} w$, we have

$$\|\lambda'_A\|_{2l} \ll_k l^k + \left(\frac{\log N}{\log w}\right)^{(1-\frac{1}{2l})k} |B|^{\frac{-1}{2l}}, \quad (6.110)$$

where $B = \text{Bohr}(\Gamma, \delta)$. For this let us start with the expression

$$\|\lambda'_A\|_{2l}^{2l} = \mathbf{E}_x |(\lambda_A * \mu_B(x))|^{2l} \quad (6.111)$$

$$= \mathbf{E}_x |\mathbf{E}_y \lambda_A(x+y) \mu_B(y)|^{2l} \quad (6.112)$$

$$\leq \mathbf{E}_{y_1, \dots, y_{2l}} \mu_B(y_1) \cdots \mu_B(y_{2l}) \mathbf{E}_x \lambda_A(x-y_1) \cdots \lambda_A(x-y_{2l}). \quad (6.113)$$

For each $2l$ -tuple $(y_1, \dots, y_{2l}) \in \mathbf{Z}_M^{2l}$, the inner sum of (6.113) can be estimated as

$$\mathbf{E}_x \lambda_A(x-y_1) \cdots \lambda_A(x-y_{2l}) \leq \frac{1}{M} \left(\frac{\log N}{\log w}\right)^{2kl} |\mathcal{A}(y_1, \dots, y_{2l})|, \quad (6.114)$$

where

$$\mathcal{A}(y_1, \dots, y_{2l}) = \left\{ 1 \leq x \leq M : \left(\prod_{i=1}^k (a_i W x + a_i b + b_i - a_i y_j), \prod_{p \leq (NW)^{1/4k+1}} p \right) = 1 \text{ for each } 1 \leq j \leq 2l \right\}.$$

$$\|\lambda_A - \lambda'_A\|_{U^2}.$$

STEP 1: Let us now take $\epsilon \in (0, c_1]$ and $\delta \in (0, 1]$ for some small enough $c_1 > 0$. Let us fix the Bohr set $B(\Gamma, \epsilon)$ of \mathbf{Z}_M with $\Gamma = \{r \in \mathbf{Z}_M : |\lambda_A(\hat{r})| \geq \delta\} \cup \{1\}$. We note that $B = B(\Gamma, \epsilon) \subset [-\epsilon M, \epsilon M]$. We define λ'_A as the convolution function over \mathbf{Z} given by $\lambda'_A := \lambda_A * \lambda_B$ where $\lambda_B = \frac{1_B}{|B|}$. If ϵ is small enough, then we see that the support of λ'_A is contained in the interval $[-2N, 2N]$. Since $M > 2N$, we may also consider λ'_A as a function on \mathbf{Z}_M . Since

$$\lambda'_A(n) = \lambda_A * \mu_B(n) := \mathbf{E}_{x \in \mathbf{Z}_M} \lambda_A(n - x) \mu_B(x),$$

where $\mu_B = \left(\frac{|B|}{M}\right)^{-1} 1_B$. Thus λ'_A can also be seen as a convolution over \mathbf{Z}_M .

STEP 2: We follow E. Naslund [22, Section 2.1]. We need the following lemma which is due to Klimov [18, Theorem 3].

Lemma 6.5.8. (*Klimov*) *Let $1 \leq i \leq r, 1 \leq n \leq X, v_0 \leq v \leq \frac{\sqrt{X}}{\log^{2r} X}$, for a fixed v_0 , and define $X_v(q_i, l_i)$ to be the number of integers n for which $p \nmid q_i n + l_i$, for each $p \leq v$, and each $1 \leq i \leq r$. Then if $u_0 = O(\exp(\log^B v))$ for a fixed constant $B > 0$ we have, for $r \geq 2$*

$$X_v(q_i, l_i) \leq \frac{X}{\log^r v} r! \prod_p \left(1 - \frac{\rho(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-r} \left(1 + \left(\frac{\log \log u_0}{\log v}\right)\right),$$

where $u_0 = \max(q_i, u_{i,j})$, $u_{i,j} = |l_i q_j - l_j q_i|$ ($1 \leq i, j \leq r$), and $\rho(p)$ is the number of

Accordingly, we define the normalized indicator function of A by

$$\lambda_A = L \left(\frac{\log N}{\log w} \right)^k 1_A,$$

where $L = \frac{M}{N} \asymp_{\Psi, V, d, F} 1$. With this normalization, we see that $\mathbf{E}_{[M]} \lambda_A = \alpha$ and $0 \leq \lambda_A \ll \lambda_{b, W}$, recalling the definition of $\lambda_{b, W}$ from Section 6.5.3.

Given a function $f : \mathbf{Z} \rightarrow \mathbf{C}$ with support in $[-2N, 2N]$, we define an M -periodic function \tilde{f} by $\tilde{f} = f(n + lM)$ where l is the unique integer such that $n + lM \in [-M/2, M/2]$, thus we may consider \tilde{f} as a function on \mathbf{Z}_M .

We define the convolution $f * g$ of functions $f, g : \mathbf{Z} \rightarrow \mathbf{C}$ with finite support by $f * g(n) := \sum_{x \in \mathbf{Z}} f(n - x)g(x)$.

We can treat λ_A as a function on \mathbf{Z}_M (by abuse of notation we take $\tilde{\lambda}_A = \lambda_A$). In this section our aim is to find a decomposition $\lambda_A = F_1 + F_2$ of λ_A such that $\|F_1\|_{L^r(\mathbf{Z}_M)}$ is “small” and $\|F_2\|_{U^2}$ is “small”. In fact we have the following proposition.

Proposition 6.5.6. *Let λ_A as above. Let $\epsilon \in (0, c_1], \delta \in (0, 1]$ be such that $\delta^{-4} \log \epsilon^{-1} \leq c_2 \log N$ for small positive absolute constants c_1, c_2 . Then there exists $\lambda'_A : \mathbf{Z}_M \rightarrow \mathbf{R}$ such that $\|\lambda'_A\|_{L^{2l}(\mathbf{Z}_M)} \ll_k l^k$ for any $l \leq 1/20 \log^{1/3} w$ and $\|\lambda_A - \lambda'_A\|_{U^2} \ll \epsilon^{1/4} + \delta^{1/4}$.*

Proof. Proof of this proposition has three steps. In the first step we give a function λ'_A , in second the step we estimate $\|\lambda'_A\|_{L^{2l}(\mathbf{Z}_M)}$ and in the third step we estimate

□

6.5.4 Transference principle

We begin the proof of Theorem 6.5.1 in this subsection and complete it in the next subsection. Let $V \in M_{r \times m}(\mathbf{Z})$ be a translation invariant matrix of complexity one with $m \geq 3$. We can choose a linear parametrization $\Psi : \mathbf{Z}^d \rightarrow \mathbf{Z}^m \cap \ker_{\mathbf{Q}}(V)$ in an exact 1-normal form over \mathbf{Z} at every $i \in [m]$. We assume that N is large enough with respect to d, m, Ψ, V and F . Let us take a prime M such that $2(\|V\|^m + \|\Psi\|^m) \leq M \leq 4V(\|V\|^m + \|\Psi\|^m)N$.

We need to analyse functions supported on $[-2N, 2N]$. For this, we embed the interval $[-2N, 2N]$ in a cyclic group \mathbf{Z}_M . The linear map Ψ reduces modulo M to a linear map $\theta : \mathbf{Z}_M^d \rightarrow \ker_{\mathbf{Z}_M}(V)$ in exact 1-normal form over \mathbf{Z}_M at every $i \in [m]$ and such that $\|\theta\| = \|\Psi\|$.

Let $N \geq 1$ be an integer and let us recall the values

$$W = \prod_{p \leq w} p, \quad w = c_0 \log N, \quad (b, W) = 1 \quad \text{with} \quad c_0 \in \left[\frac{1}{8k+2}, \frac{1}{4k+1} \right].$$

We consider a subset $A \subset \{1, \dots, N\}$ such that $|A| \geq \alpha \left(\frac{\log w}{\log N} \right)^k N$ and $b + AW \subset S(NW, (NW)^{1/4k+1})$.

where \tilde{x} is equals to $y \in [1, M]$ so that $x \equiv y \pmod{M}$ for any $x \in x_M$. Thus we have

$$\nu_1(\phi_i(n+x)) = \nu(\tilde{\phi}_i(n) + \widetilde{\phi_i(x)}).$$

Hence using (6.67), for any good x , we have

$$\begin{aligned} \mathbf{E}_{n \in [P]^d} \prod_{i=1}^t \nu_1(\phi_i(x+n)) &= \mathbf{E}_{n \in [P]^d} \prod_{i=1}^t \nu(\tilde{\phi}_i(n) + \widetilde{\phi_i(x)}) \\ &= 1 + O_D\left(\frac{1}{(\log N)^{1-o(1)}}\right). \end{aligned}$$

since ϕ_i is a non-zero affine linear map, for any i , the number of bad x is at most $O_{D,t,m}(PM^{d-1})$. Since $\nu(n) \ll_\epsilon M^\epsilon$, we obtain

$$\begin{aligned} \mathbf{E}_{\substack{x \in \mathbf{Z}_M^d \\ n \in [P]^d}} I\{x \text{ good}\} \prod_{i=1}^t \nu_1(\phi_i(x+n)) \\ = \left(1 - O_{D,t,m}\left(\frac{P}{M}\right)\right) \left(1 + O_{D,t,m}\left(\frac{1}{(\log N)^{1-o(1)}}\right)\right) \end{aligned} \quad (6.108)$$

and

$$\mathbf{E}_{\substack{x \in \mathbf{Z}_M^d \\ n \in [P]^d}} I\{x \text{ not good}\} \prod_{i=1}^t \nu_1(\phi_i(x+n)) = O_{D,t,m,\epsilon}\left(\frac{PM^\epsilon}{M}\right) \quad (6.109)$$

Using (6.109), (6.108) and noting that $N \asymp_{D,t,m} M$, we obtain the pseudorandom asymptotic property

$$\mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{i=1}^t \nu_1(\phi_i(x)) = 1 + O_{D,t,m}\left(\frac{1}{(\log N)^{1-o(1)}}\right).$$

of finite complexity such that the map $\tilde{\Phi} : \mathbf{Z}^d \rightarrow \mathbf{Z}^t$ induces the map $\tilde{\Phi} : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ and $\|\dot{\Phi}\|_M = \|\tilde{\Phi}\|$.

Now let us consider

$$\mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{i=1}^t \nu_1(\phi_1(x)).$$

Let $P = M^{1/2}$ and through the embedding $[P] \hookrightarrow \mathbf{Z}_M$, we identify $[P]^d$ as a subset of \mathbf{Z}_M^d . Note that for any $n \in \mathbf{Z}_M^d$ we have

$$\mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{i=1}^t \nu_1(\phi_1(x)) = \mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{i=1}^t \nu_1(\phi_1(x+n)).$$

Hence we have

$$\mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{i=1}^t \nu_1(\phi_1(x)) = \mathbf{E}_{x \in \mathbf{Z}_M^d} \mathbf{E}_{n \in [P]^d} \prod_{i=1}^t \nu_1(\phi_1(x+n)).$$

We say that $x \in \mathbf{Z}_M^d$ is good if for any i with $1 \leq i \leq t$, we have

$$\phi_i(x) \notin \left[-\frac{M}{2}, -\frac{M}{2} + DP \right]_M \cup \left[\frac{M}{2}, \frac{M}{2} - DP \right]_M$$

where $[a, b]_M$ is a interval in the integers with respect to the metric

$$\|x - y\|_M = \min_{n \in \mathbf{Z}} |x - y + Mn|.$$

If x is good, then for every $n \in [P]^d$, we have

$$\tilde{\phi}_i(x+n) = \tilde{\phi}_i + \tilde{\phi}_i(n)$$

of [11] to give an explicit error term, which involves the level of the pseudorandom majorant. The following proposition, due to Henriot [15, Proposition 11], gives a pseudorandom majorant on $\tilde{\lambda}_{b,W}$.

Proposition 6.5.5. *Let $D \geq 1$. There exists a constant C_D such that if $N \geq C_D$ and $c_1N \leq M \leq c_2N$ is a prime, then there exists a D -pseudorandom weight $\nu_1 : \mathbf{Z}_M \rightarrow \mathbf{R}^+$ of level $(\log N)^{1-o(1)}$ such that*

$$0 \leq \tilde{\lambda}_{b,W} \ll_D \nu_1. \quad (6.105)$$

Proof. Let ν be a pseudorandom majorant as in Proposition 6.5.4. Now we define ν_1 as follows.

$$\nu_1(n) := \nu(n) \quad \text{for } n \in \{1, \dots, M\}$$

and we extend ν_1 to whole the set of integers \mathbf{Z} as periodically with period M . Thus ν_1 defines a function on \mathbf{Z}_M . Clearly we have

$$0 \leq \tilde{\lambda}_{b,W} \ll_D \nu_1.$$

Now we verify the pseudorandom property. For this let us consider an affine linear map

$$\Phi = (\phi_1, \dots, \phi_t) : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t \quad (6.106)$$

of finite complexity with $d, t, \|\dot{\Phi}\| \leq D$. Then there exists an affine linear map

$$\tilde{\Phi} = (\tilde{\phi}_1, \dots, \tilde{\phi}_t) : \mathbf{Z}^d \rightarrow \mathbf{Z}^t \quad (6.107)$$

$$c_{\chi,k,2}^t + O\left(e^{-c|L|^{1/2}}(\log R)^{O_{k,t}(1)} + \frac{1}{w} + \frac{L \log w}{\log R} + \frac{R^{5t}}{P}\right) \quad (6.103)$$

provided that $L \leq c \log R / \log w$. Assume that $P \geq N^{c_1}$, for some absolute constant c_1 which depends at most on t and k . Choose $L = C(\log \log N)^2$ and $R = N^\eta$ with $5\eta t \leq c_0/2$, so that

$$\mathbf{E}_{n \in [P]^d} \prod_{i \in [t]} \Lambda_{\chi,R,W}[\psi_i(n)] = c_{\chi,k,2}^t + O((\log N)^{-1+o(1)}). \quad (6.104)$$

By Lemma 6.5.6, we have $c_{\chi,k,2} > 0$ and therefore we may define a normalized weight function $\nu := c_{\chi,k,2}^{-1} \Lambda_{\chi,R,W}$, which satisfies the desired pseudorandomness asymptotic by (6.104), and which majorizes a constant multiple of $\lambda_{b,W}$.

□

Let us take N to be a large positive integer and consider an embedding $[N] \hookrightarrow \mathbf{Z}_M$, where M is a prime larger than N . We are interested in finding a pseudorandom majorant over \mathbf{Z}_M for the function $\lambda_{b,W}$, which we can think of as a function on \mathbf{Z}_M . More precisely, given a function $f : \mathbf{Z} \rightarrow \mathbf{C}$ with support in $[N]$, we define an M -periodic function $\tilde{f}(n) = f(n + lM)$, where l is the unique integer such that $n + lM \in [N]$. Note that \tilde{f} may in turn be viewed as a function on \mathbf{Z}_M .

It is relatively easy to construct a pseudorandom majorant on \mathbf{Z}_M from the pseudorandom majorant on the integers \mathbf{Z} , given in Proposition 6.5.4, by cutting \mathbf{Z}_M^d into small boxes as explained in [11, page 527]. But Henriot refined the arguments

Thus, we have

$$c_{\chi,k,2} = \int_0^\infty \cdots \int_0^\infty \left(\int_{\mathbf{R}} \varphi(\xi)(1+i\xi)^k e^{-(1+i\xi)(x_1+\cdots+x_k)} d\xi \right)^2 dx_1 \cdots dx_k.$$

We observe that

$$\chi^{(k)}(x_1 + \cdots + x_k) = (-1)^k \int_{\mathbf{R}} \varphi(\xi)(1+i\xi)^k e^{-(1+i\xi)(x_1+\cdots+x_k)} d\xi.$$

Hence,

$$c_{\chi,k,2} = \int_0^\infty \cdots \int_0^\infty \left(\chi^{(k)}(x_1 + \cdots + x_k) \right)^2 dx_1 \cdots dx_k, \quad (6.101)$$

which shows that $c_{\chi,k,2} > 0$.

□

Lemma 6.5.7. *Let $1 \leq L \leq \frac{c \log R}{\log w}$. Then we have*

$$\begin{aligned} h_{R,W}^t & \int \cdots \int_{[-L,L]^\Omega} \prod_p E_{p,\xi} \prod_{(i,j) \in \Omega} \varphi(\xi_{ij}) d\xi_{ij} \\ & = c_{\chi,k,2}^t + O\left(e^{-c|L|^{1/2}} + \frac{1}{w} + \frac{L \log w}{\log R} \right). \end{aligned} \quad (6.102)$$

Proof. The proof of this lemma follows by using the Euler product $\prod_p E_{p,\xi}$ value from lemma 6.5.5 and recalling the growth of φ . □

Now we are in a position to give a proof of Proposition 6.5.4

PROOF OF PROPOSITION 6.5.4.— Let $P \geq 1$. Using lemmas 6.5.2, 6.5.4 and 6.5.7,

we see that the average $\mathbf{E}_{n \in [P]^d} \prod_{i \in [t]} \Lambda_{\chi,R,W}[\psi_i(n)]$ is equal to

By Lemma 6.5.3, we have

$$\begin{aligned} \prod_p E_{p,\xi} &= \prod_{p>w} \left(1 + k \sum_{B \text{ vertical}} (-1)^{|B|} p^{-1-\sum_B z_{ij}} + O_{t,k}(p^{-2}) \right) \\ &= (1 + O_{t,k}(w^{-1})) \prod_{p>w} \prod_{B \text{ vertical}} (1 - p^{-1-\sum_B z_{ij}})^{(-1)^{|B|}k}. \end{aligned}$$

Since $p^{-z} = 1 + O(L \log p / \log R)$ for $p \leq w$ and $|z| \leq L / \log R$, we have

$$\prod_p E_{p,\xi} = (1 + O_{t,k}(\frac{1}{w} + \frac{L \log w}{\log R})) (\log w)^{kt} \prod_{B \text{ vertical}} \zeta(1 + \sum_B z_{ij})^{(-1)^{|B|}k}. \quad (6.99)$$

Using the fact that $\zeta(s) = \frac{1}{s-1}(1 + O(|s-1|))$ for $\Re(s) > 1$, it follows that

$$\prod_p E_{p,\xi} = (1 + O_{t,k}(\frac{1}{w} + \frac{L \log w}{\log R})) (\log w)^{kt} \prod_{B \text{ vertical}} (\sum_B z_{ij})^{(-1)^{|B|}k}. \quad (6.100)$$

If we substitute the value of z_{ij} in the above equation (6.100), we get the conclusion of the lemma. \square

Definition 6.5.6. (Sieve factor) $c_{\chi,k,2} = \iint_{\mathbf{R}^2} \left(\frac{(1+i\xi_1)(1+i\xi_2)}{2+i(\xi_1+\xi_2)} \right)^k \varphi(\xi_1)\varphi(\xi_2)d\xi_1d\xi_2$

Lemma 6.5.6. For any $k \geq 0$, we have $c_{\chi,k,2} > 0$

Proof. Note that

$$\frac{1}{(2+i(\xi_1+\xi_2))^k} = \int_0^\infty \cdots \int_0^\infty \prod_{j=1}^k e^{-(2+i(\xi_1+\xi_2))} dx_1 \cdots dx_k.$$

We now estimate the error term in (6.89). By the multiplicativity of $\alpha(m_1, \dots, m_t)$ we can write the error term in (6.89) as

$$e^{-cL^{1/2}} \prod_p \sum_{(r_{ij}) \in \{0,1\}^\Omega} \alpha(p^{r_1}, \dots, p^{r_t}) p^{-\frac{\sum_{i,j} r_{ij}}{\log R}} \quad (6.92)$$

$$= e^{-cL^{1/2}} \prod_p \sum_{B \subset \Omega} \alpha(p, B) p^{-\frac{|B|}{\log R}} \quad (6.93)$$

Note that if $B \neq \emptyset$, then we have $\alpha(p, B) \leq K^2/p$. Keeping this in mind we obtain an upper bound for the error term as

$$\ll_k e^{-cL^{1/2}} \prod_p \left(1 + \frac{k^2 |\Omega|}{p^{1+1/\log R}}\right) \quad (6.94)$$

$$\ll_k e^{-cL^{1/2}} \prod_p \left(1 - \frac{1}{p^{1+1/\log R}}\right)^{-|\Omega|k^2} \quad (6.95)$$

$$\ll_k e^{-cL^{1/2}} \zeta \left(1 + \frac{1}{\log R}\right)^{|\Omega|k^2} \quad (6.96)$$

$$\ll_k e^{-cL^{1/2}} (\log R)^{|\Omega|k^2}. \quad (6.97)$$

Thus we get the conclusion of the lemma by (6.97), (6.91), (6.88) and (6.89). \square

With the estimates on $\alpha(p, B)$ in our hand we can evaluate the Euler product.

Lemma 6.5.5. *Let $1 \leq L \leq \frac{c \log R}{\log w}$ be a parameter. For every $\xi \in [-L, L]^\Omega$, we have*

$$\prod_p E_{p,\xi} = \left(1 + O_{k,t} \left(\frac{1}{w} + \frac{L \log w}{\log R}\right)\right) h_{R,W}^{-t} \prod_{B \text{ vertical}} \left(\sum_{(i,j) \in B} (1 + i\xi_{ij})\right)^{-(-1)^{|B|}k} \quad (6.98)$$

Proof. Note the identity $\sum_{B \text{ vertical}} (-1)^{|B|} = -t$, and write $z_{ij} = (1 + i\xi_{ij})/\log R$.

where we use the fact that $\varphi(\xi) \ll e^{-c|\xi|^{1/2}}$. Therefore we have

$$\prod_{(i,j) \in \Omega} \chi\left(\frac{\log m_{ij}}{\log R}\right) = \int \cdots \int_{[-L,L]^\Omega} \prod_{(i,j) \in \Omega} m_{ij}^{-(1+i\xi_{ij})/\log R} \varphi(\xi_{ij}) d\xi_{ij} \quad (6.86)$$

$$+ O\left(e^{-cL^{1/2}} \prod_{(i,j) \in \Omega} m_{ij}^{-\frac{1}{\log R}}\right). \quad (6.87)$$

We write the value of the above product (6.86) in (6.82), then we see that the value given in (6.82) becomes

$$\int \cdots \int_{[-L,L]^\Omega} \sum'_{(m_{i,j}) \in \mathbf{N}^\Omega} \alpha(m_1, \dots, m_t) \prod_{(i,j) \in \Omega} \mu(m_{ij}) m_{ij}^{-(1+i\xi_{ij})/\log R} \varphi(\xi_{ij}) d\xi_{ij} \quad (6.88)$$

$$+ O\left(e^{-cL^{1/2}} \sum'_{(m_{i,j}) \in \mathbf{N}^\Omega} \alpha(m_1, \dots, m_t) \prod_{(i,j) \in \Omega} m_{ij}^{-\frac{1}{\log R}}\right). \quad (6.89)$$

Using the multiplicativity of $\alpha(m_1, \dots, m_t)$, we can write the main term in (6.88) as

$$\int \cdots \int_{[-L,L]^\Omega} \prod_p \sum_{(r_{ij}) \in \{0,1\}^\Omega} (-1)^{\sum_{(i,j) \in \Omega} r_{ij}} \alpha(p^{r_1}, \dots, p^{r_t}) \times \quad (6.90)$$

$$p^{-\sum_{(i,j) \in \Omega} r_{ij} z_{ij}} \prod_{(i,j) \in \Omega} \varphi(\xi_{ij}) d\xi_{ij},$$

where $z_{ij} = (1+i\xi_{ij})/\log R$ and $r_i = \max(r_{i,1}, r_{i,2})$. We can write the above quantity, by using the definition of Euler factor, in a closed form as

$$\int \cdots \int_{[-L,L]^\Omega} \prod_p E_{p,\xi} \prod_{(i,j) \in \Omega} \varphi(\xi_{ij}) d\xi_{ij}. \quad (6.91)$$

Therefore, in this case, we get $\alpha(p, B) \ll_k 1/p^2$. \square

We define the following Euler factor.

Definition 6.5.5. (*Euler factor*) Let $\xi \in \mathbf{R}^\Omega$ and $z_{ij} = (1 + i\xi_{ij})/\log R$. We define

$$E_{p,\xi} = \sum_{B \subset \Omega} (-1)^{|B|} \alpha(p, B) p^{-\sum_{(i,j) \in \Omega} z_{ij}}. \quad (6.81)$$

The estimates on $\alpha(p, B)$, by Lemma 6.5.3, and the fact that $\Re(z_{ij}) > 0$ guarantees the absolute convergence of the product $\prod_p E_{p,\xi}$.

Lemma 6.5.4. For any $L \geq 1$, we have

$$\sum'_{(m_{i,j}) \in \mathbf{N}^\Omega} \alpha(m_1, \dots, m_t) \prod_{(i,j) \in \Omega} \mu(m_{i,j}) \chi\left(\frac{\log m_{i,j}}{\log R}\right) \quad (6.82)$$

$$= \int \cdots \int_{[-L,L]^\Omega} \prod_p E_{p,\xi} \prod_{(i,j) \in \Omega} \varphi(\xi_{ij}) d\xi_{ij} + O(e^{-cL^{1/2}} (\log R)^{|\Omega|k^2}). \quad (6.83)$$

Proof. We have that

$$\chi(x) = \int_{-\infty}^{\infty} \varphi(\xi) e^{-(1+i\xi)x} d\xi. \quad (6.84)$$

Let $L \geq 1$ be a real number. In (6.84) taking the value $x = \log m_{ij}/\log R$ and truncating the integral at $-L$ to L we get that

$$\chi\left(\frac{\log m_{ij}}{\log R}\right) = \int_{-L}^L m_{ij}^{-(1+i\xi_{ij})/\log R} \varphi(\xi_{ij}) d\xi_{ij} + O(e^{-cL^{1/2}} m_{ij}^{-\frac{1}{\log R}}), \quad (6.85)$$

the variables $m_{i,j}$, keeping in mind that $m_i = [m_{i,1}, m_{i,2}]$. Writing $m_{i,j} = p^{r_{ij}}$, $r_i = \max(r_{i1}, r_{i2})$, and $B = \{(i, j) \in \Omega : r_{ij} = 1\}$, we have $r_i = 1$ if and only if $r_{ij} = 1$ for some $j \in [2]$, since our $m_{i,j}$'s are square-free, that is, if and only if the slice B_i of B at i is non-empty. Therefore

$$\alpha(p^{r_1}, \dots, p^{r_t}) = \mathbf{E}_{n \in \mathbf{Z}_p^d} (p | F(b + W\psi_i(n)) \forall i : B_i \neq \emptyset) =: \alpha(p, B) \quad (6.79)$$

A non-empty set $B \subset \Omega$ is said to be vertical if $B \subset \{i\} \times [2]$ for some $i \in [t]$. We now estimate the size of the local factors $\alpha(p, B)$ in the following lemma.

Lemma 6.5.3. *For $B \neq \emptyset$, we have*

$$\alpha(p, B) = \begin{cases} 0 & \text{if } p \leq w \\ \frac{k}{p} & \text{if } p > w \text{ and } B \text{ vertical} \\ O_k(\frac{1}{p^2}) & \text{if } p > w \text{ and } B \text{ is not vertical.} \end{cases}$$

Proof. (i) Let $p \leq w$, then $p|W$. Recall that our choice of b satisfies $(F(b), W) = 1$. Therefore p does not divide $F(b + \psi_l(n)W)$ for any l and n . Thus we have $\alpha(p, B) = 0$. (ii) Let $p > w$. Let us recall the polynomial $F(n) = \prod_{i=1}^k (a_i n + b_i)$. Since $p > w$, we can choose w to be a sufficiently large enough so that a_i 's are invertible modulo p . When B is vertical, there is only one i such that that B_i is non-empty and therefore $\alpha(p, B) = k/p$. (iii) Let $p > w$ and B is not vertical, then there are at least two indices i, j such that B_i and B_j are non-empty. Note that

$$|\{n \in \mathbf{Z}_p^d : p | F(b + \psi_i(n)W) \text{ and } p | F(b + \psi_j(n)W)\}| \ll_k p^{d-2}. \quad (6.80)$$

Proof. Note that the value in (6.72) is equal to

$$\sum_{n \in [P]^d} \prod_{i=1}^t \left(\sum_{m|F(b+\psi_i(n)W)} \mu(m) \chi \left(\frac{\log m}{\log R} \right) \right)^2. \quad (6.74)$$

After opening the square in (6.74) and interchanging the summation we see that the quantity in (6.74) equals

$$\sum'_{(m_{i,j}) \in \mathbf{N}^\Omega} \left(\prod_{(i,j) \in \Omega} \mu(m_{i,j}) \chi \left(\frac{\log m_{i,j}}{\log R} \right) \right) \sum_{\substack{n \in [P]^d \\ m_{i,j}|F(b+\psi_i(n)W) \forall (i,j) \in \Omega}} 1. \quad (6.75)$$

The counting inner sum in (6.75) can be rewritten as

$$\sum_{n \in [P]^d} \prod_{(i,j) \in \Omega} 1_{m_{i,j}|F(b+\psi_i(n)W)} \quad (6.76)$$

and which is equal to

$$\sum_{n \in [P]^d} \prod_{i=1}^t 1_{m_i|F(b+\psi_i(n)W)}, \quad (6.77)$$

where $m_i = [m_{i,1}, m_{i,2}]$. Letting $m = [m_1, \dots, m_t]$ we see that (6.77) equals to

$$P^d \mathbf{E}_{n \in \mathbf{Z}_m^d} \prod_{i=1}^t 1_{m_i|F(b+\psi_i(n)W)} + O_t(P^{d-1}). \quad (6.78)$$

By using the facts that $\max_x |\chi(x)| \leq R$ and $R = N^\eta \gg 1$, we get the conclusion of the lemma from (6.78), (6.75) and (6.72). \square

Observe that by the Chinese remainder theorem, $\alpha(m_1, \dots, m_t)$ is multiplicative in

as $\varphi(\xi) \ll e^{-c|\xi|^{1/2}}$. The advantage of this choice is that, by truncation at any parameter $L \geq 1$, we have

$$\rho(m) = \int_{-L}^L m^{-(1+i\xi)/\log R} \varphi(\xi) d\xi + O(e^{-cL^{1/2}}) \text{ for } m \leq R. \quad (6.70)$$

We now make preparation for the proof of Proposition 6.5.4. We fix $D \geq 1$ and $w = c_0 \log N$, so that we may assume w is larger than any fixed constant depending on D and our polynomial F . We then consider a system of affine-linear forms $\Psi : \mathbf{Z}^d \rightarrow \mathbf{Z}^t$ of finite complexity such that $d, t, \|\dot{\Psi}\| \leq D$.

We first expand the divisor sums inside the correlation of divisor sums, and it is useful to introduce a notation $\Omega = [t] \times [2]$. Note also that the prime in \sum' means that the summation is restricted to square-free numbers.

Lemma 6.5.2. *Let $(m_{i,j}) \in \mathbf{N}^\Omega$. We write m_i for the lcm $[m_{i,1}, m_{i,2}]$ of $m_{i,1}$ and $m_{i,2}$. Let*

$$\alpha(m_{1,1}, m_{1,2}, \dots, m_{t,1}, m_{t,2}) := \alpha(m_1, \dots, m_t) = \mathbf{E}_{n \in \mathbf{Z}_m^d} (m_i | F(b + W\psi_i(n)) \forall i \in [t]). \quad (6.71)$$

Let also $P \geq 1$. Then

$$h_{R,W}^{-t} \sum_{n \in [P]^t} \Lambda_{\chi,R,W}[\psi_1(n)] \dots \Lambda_{\chi,R,W}[\psi_t(n)] \quad (6.72)$$

$$= P^d \sum'_{(m_{i,j}) \in \mathbf{N}^\Omega} \alpha(m_1, \dots, m_t) \prod_{(i,j) \in \Omega} \mu(m_{i,j}) \chi \left(\frac{\log m_{i,j}}{\log R} \right) + O(R^{2|\Omega|} P^{d-1}). \quad (6.73)$$

η is a positive constant to be chosen later. Let us take a sequence of real numbers $\rho : \mathbf{N} \rightarrow \mathbf{R}$ such that $\rho(1) = 1$ and with support on $[R]$ which we will choose later.

Definition 6.5.4. We let $h_{R,W} = \left(\frac{\log R}{\log w}\right)^k$ and

$$\Lambda_{\rho,R,W}(n) = h_{R,W} \left(\sum_{m|F(b+Wn)} \mu(m)\rho(m) \right)^2. \quad (6.68)$$

Our pseudorandom majorant turns out to be a constant multiple of the above function. Indeed the constant is $c_{\chi,k,2}^{-1}$ where $c_{\chi,k,2}$ is as in Lemma 6.5.6.

Lemma 6.5.1. Let $R = N^\eta$ with $0 < \eta \leq c_0/2 \leq 1/8k + 2$, where c_0 is as in 6.64.

We have

$$0 \leq \lambda_{b,W} \ll_{\eta,k} \Lambda_{\rho,R,W} \ll_{\epsilon} N^{\epsilon} \quad (6.69)$$

for every $\epsilon > 0$.

Proof. The second inequality follows from the divisor bound. For the first inequality: note that if $\lambda_{b,W}(n)$ is non zero, then all the divisors of $F(Wn+b)$ exceed $N^{1/4k+1}(> R)$ except 1. \square

Now we specify the weights $\rho(m)$. We let

$$\rho(m) = \chi \left(\frac{\log m}{\log R} \right) \quad \text{where} \quad \chi(x) = 1_{[-1,1]}(x) e^{x+1} e^{-1/(1-x^2)}.$$

By Fourier inversion formula, we have $\chi(x) = \int_{-\infty}^{\infty} \varphi(\xi) e^{-(1+i\xi)x} d\xi$ for every $x \in [-1, 1]$, where φ is the Fourier transform of $1_{[-1,1]}(x) e^{1-1/(1-x^2)}$, and thus decays

long arithmetic progressions and linear patterns in primes. Our majorant takes the form as in (6.68).

Let us fix a large integer N and recall the values

$$W = \prod_{p \leq w} p, \quad w = c_0 \log N, \quad (b, W) = 1 \quad \text{with} \quad c_0 \in \left[\frac{1}{8k+2}, \frac{1}{4k+1} \right]. \quad (6.64)$$

Let us define the measure associated to W -tricked sifted sets as

$$\lambda_{b,W}(n) := \left(\frac{\log N}{\log w} \right)^k \mathbf{1}(n \in [N] : b + Wn \in S(NW, (NW)^{1/4k+1})). \quad (6.65)$$

We will construct a weight function over \mathbf{Z} which majorizes $\lambda_{b,W}$ and satisfies pseudorandomness asymptotics. Indeed we have the following proposition.

Proposition 6.5.4. *Let $D \geq 1$ be a parameter. There exist a constant C_D such that the following holds. For $N \geq C_D$ and $w = c_0 \log N$, there exists $\nu : \mathbf{Z} \rightarrow \mathbf{R}^+$ such that, for every $\epsilon > 0$,*

$$0 \leq \lambda_{b,W} \ll_D \nu \ll_\epsilon N^\epsilon \quad (6.66)$$

and, for any $P \geq N^{c_1}$, absolute constant c_1 , and any affine system $\Psi : \mathbf{Z}^d \rightarrow \mathbf{Z}^t$ of finite complexity and such that $d, t, \|\dot{\Psi}\| \leq D$,

$$\mathbf{E}_{n \in [P]^d} \nu[\psi_1(n)] \dots \nu[\psi_t(n)] = 1 + O_D \left(\frac{1}{(\log N)^{1-o(1)}} \right). \quad (6.67)$$

We closely follow [15, Section 5] to prove this proposition. We let $R = N^\eta$, where

that the functions f_i are dominated by a pseudorandom measure. More precisely we have

Proposition 6.5.3. (*Generalized Von Neumann theorem*)

Let $d, t, Q, H \geq 1$ be parameters. There exists a constant D depending on d, t, Q such that the following holds. Suppose $M > D$ is a prime and $\Psi : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ be an affine linear system of complexity one and it is in exact 1-normal form at each $i \in \{1, \dots, t\}$, and such that $\|\dot{\Psi}\| \leq Q$. Suppose also that $\nu : \mathbf{Z}_M \rightarrow \mathbf{R}^+$ is D -pseudorandom of level H , and $f_1, \dots, f_t : \mathbf{Z}_M \rightarrow \mathbf{R}$ with $|f_i| \leq \nu$ for each i . Then we have

$$|\Lambda(f_1, \dots, f_t)| \leq \min_i (\|f_i\|_{U^2(\mathbf{Z}_M)}) + O_D(H^{-1/4}). \quad (6.63)$$

6.5.3 Construction of a pseudorandom majorant

In this section we will construct pseudorandom majorant of the normalized indicator function of W -tricked sifted sets. The idea of constructing this majorant is primarily based on the Selberg Λ^2 -sieve. The essence of Selberg sieve is to consider sums of the form $(\sum_{d|n} \rho(d))^2$ with weights $\rho(d) \in \mathbf{R}$ such that $\rho(1) = 1$ and $\rho(d) = 0$ for $d > R$ for some sieve level R . Here $\rho(d)$'s are called Selberg weights. Nowadays it has become a powerful tool to deal with the set of primes. For instance Goldston, Pintz and Yildirim [9, 8] considered a particular type of Selberg weights while studying small gaps between primes. Green and Tao [11, 12] studied correlations of $(\sum_{d|n} \rho(d))^2$ for smooth weights $\rho(d)$ while establishing the presence of arbitrarily

complexity with $d, t, \|\dot{\Psi}\| \leq D$ (where $\dot{\Psi} = \Psi - \Psi(0)$ is linear part of Ψ), we have

$$\mathbf{E}_{n \in \mathbf{Z}_M^d} \nu[\psi_1(n)] \cdots \nu[\psi_t(n)] = 1 + O_D\left(\frac{1}{H}\right). \quad (6.61)$$

This pseudorandomness property turns out to be invariant under averaging. In fact we have the following proposition.

Proposition 6.5.2. *Let $D, H \geq 1$ be parameters. Suppose that $\nu : \mathbf{Z}_M \rightarrow \mathbf{R}^+$ is D -pseudorandom of level H , B is a symmetric subset of \mathbf{Z}_M and $\mu_B = (|B|/M)^{-1}1_B$. Then $\nu' = (\nu + \nu * \mu_B)/2$ is also D -pseudorandom of level H .*

Proof. Let $\nu_0 = \nu$ and $\nu_1 = \nu * \mu_B$. Thus we have

$$\nu_\epsilon(x) = \mathbf{E}_{y \in B} \nu(x + \epsilon y) \quad \text{for each } \epsilon \in \{0, 1\}. \quad (6.62)$$

Let $T = \mathbf{E}_{n \in \mathbf{Z}_M^d} \nu'[\psi_1(n)] \cdots \nu'[\psi_t(n)]$. We have

$$\begin{aligned} T &= \mathbf{E}_{\epsilon \in \{0,1\}^t} \mathbf{E}_{n \in \mathbf{Z}_M^d} \nu_{\epsilon_1}[\psi_1(n)] \cdots \nu_{\epsilon_t}[\psi_t(n)] \\ &= \mathbf{E}_{\epsilon \in \{0,1\}^t} \mathbf{E}_{y \in B^t} \mathbf{E}_{n \in \mathbf{Z}_M^d} \nu[\psi_1(n) + \epsilon_1 y_1] \cdots \nu[\psi_t(n) + \epsilon_t y_t]. \end{aligned}$$

Note that for every $\epsilon \in \{0, 1\}^t$ and $y \in B^t$, the affine linear system $(\psi_i + \epsilon_i y_i)_{1 \leq i \leq t}$ has the same linear part as $(\psi_i)_{1 \leq i \leq t}$. Since ν is D -pseudorandom of level H , we have $T = 1 + O_D(H^{-1})$. Thus ν' is also D -pseudorandom of level H . \square

We now have come to the main part of this subsection. The following proposition, which is Theorem 3 in [15], gives us estimates on the count $\Lambda(f_1, \dots, f_t)$ provided

the quantity in (6.59) becomes

$$\frac{\mathbf{E}_{x_1^{(0)}, x_2^{(0)}} \mathbf{E}_{x_1^{(1)}, x_2^{(1)}, x} f_1(x) \overline{f_1(x + x_2^{(1)} - x_2^{(0)})} \times}{f_1(x + x_1^{(1)} - x_1^{(0)}) f_1(x + x_1^{(1)} - x_1^{(0)} + x_2^{(1)} - x_2^{(0)})} \quad (6.60)$$

Now let us change the variables in (6.60). Let $x_2^{(1)} - x_2^{(0)} = h_1$ and $x_1^{(1)} - x_1^{(0)} = h_2$, then we see that the quantity in (6.60) becomes

$$\mathbf{E}_{x, h_1, h_2 \in \mathbf{Z}_M} \prod_{w \in \{0,1\}^2} C^{|w|} f_1(x + \sum_{j=1}^2 w_j h_j)$$

which is exactly equals to $\|f_1\|_{U^2}^4$.

□

We remark here that the validity of the inequality (6.57) requires $|f_i| \leq 1$ for each i . But while working with sifted sets we encounter functions which are not bounded by 1 and in general such an inequality does not hold for those functions. However, a weak form of the inequality (6.57) holds for the class of functions which are dominated by a “pseudorandom measure”.

We now define, following [15, Section 6], functions called pseudorandom majorants and discuss their basic properties.

Definition 6.5.3. *Let $D, H \geq 1$ be parameters. We say that $\nu : \mathbf{Z}_M \rightarrow \mathbf{R}^+$ is D -pseudorandom of level H if, for every affine system $\Psi : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ of finite*

which is same as

$$\mathbf{E}_{x_1^{(0)}, x_1^{(1)}, x_3, \dots, x_d \in \mathbf{Z}_M} \mathbf{E}_{x_2} \frac{\prod_{j \in Y_1} f_j(a_{j1}x_1^{(0)} + a_{j2}x_2 + \sum_{i=3}^d a_{ji}x_i + b_j) \times}{\prod_{j \in Y_1} f_j(a_{j1}x_1^{(1)} + a_{j2}x_2 + \sum_{i=3}^d a_{ji}x_i + b_j)},$$

and which is at most

$$\mathbf{E}_{x_1^{(0)}, x_1^{(1)}, x_3, \dots, x_d \in \mathbf{Z}_M} \left| \frac{\mathbf{E}_{x_2} f_1(a_{11}x_1^{(0)} + a_{12}x_2 + \sum_{i=3}^d a_{1i}x_i + b_1) \times}{f_1(a_{11}x_1^{(1)} + a_{12}x_2 + \sum_{i=3}^d a_{1i}x_i + b_1)} \right|.$$

Let us change the variables. Given $x_3, \dots, x_d \in \mathbf{Z}_M$, we set $b_1 + \sum_{j=3}^d a_{1j}x_j = y$.

Then we see that

$$|\Lambda(f_1, \dots, f_t)|^2 \leq \mathbf{E}_{x_1^{(0)}, x_1^{(1)}, y} \left| \mathbf{E}_{x_2} f_1(y + a_{11}x_1^{(0)} + a_{12}x_2) \overline{f_1(y + a_{11}x_1^{(1)} + a_{12}x_2)} \right|. \quad (6.58)$$

Applying Cauchy-Schwarz to the right hand side of (6.58), we see that the square of this is at most

$$\mathbf{E}_{x_1^{(0)}, x_2^{(0)}} \mathbf{E}_{x_1^{(1)}, x_2^{(1)}, y} \frac{f_1(y + x_1^{(0)} + x_2^{(0)}) \overline{f_1(y + x_1^{(0)} + x_2^{(1)})} \times}{f_1(y + x_1^{(1)} + x_2^{(0)}) f_1(y + x_1^{(1)} + x_2^{(1)})}. \quad (6.59)$$

We again change the variables in (6.59). Let $x = y + x_1^{(0)} + x_2^{(0)}$, then we see that

x_1, x_2 for $j \neq 1$. Let $Y_1 = \{j \in [d] : \psi_j \text{ depends at most on } x_1\}$ and similarly let $Y_2 = \{j \in [d] : \psi_j \text{ depends at most on } x_2\}$.

To prove our proposition, it is enough to prove that

$$|\Lambda(f_1, \dots, f_t)| \leq \|f_1\|_{U^2}.$$

By the definition 6.5.1 of $\Lambda(f_1, \dots, f_t)$, we have

$$\begin{aligned} |\Lambda(f_1, \dots, f_t)| &= \left| \mathbf{E}_{x \in \mathbf{Z}_M^d} f_1[\psi_1(x)] \cdots f_t[\psi_t(x)] \right| \\ &= \left| \mathbf{E}_{x \in \mathbf{Z}_M^d} \prod_{j \in Y_1} f_j[\psi_j(x)] \prod_{j \in Y_2} f_j[\psi_j(x)] \right| \\ &= \left| \mathbf{E}_{x_2, \dots, x_d \in \mathbf{Z}_M} \prod_{j \in Y_2} f_j[\psi_j(x)] \mathbf{E}_{x_1 \in \mathbf{Z}_M} \prod_{j \in Y_1} f_j[\psi_j(x)] \right| \\ &\leq \mathbf{E}_{x_2, \dots, x_d \in \mathbf{Z}_M} \left| \mathbf{E}_{x_1 \in \mathbf{Z}_M} \prod_{j \in Y_1} f_j[\psi_j(x)] \right|. \end{aligned}$$

As a consequence of Cauchy-Schwarz inequality and noting the fact that $|f_i| \leq 1$, we see that the square of $|\Lambda(f_1, \dots, f_t)|$ is at most

$$\begin{aligned} \mathbf{E}_{x_2, \dots, x_d \in \mathbf{Z}_M} \mathbf{E}_{x_1^{(0)}, x_1^{(1)}} \prod_{j \in Y_1} f_j(a_{j1}x_1^{(0)} + a_{j2}x_2 + \sum_{i=3}^d a_{ji}x_i + b_j) \times \\ \prod_{j \in Y_1} f_j(a_{j1}x_1^{(1)} + a_{j2}x_2 + \sum_{i=3}^d a_{ji}x_i + b_j), \end{aligned}$$

Note that this operator Λ counts Ψ -patterns with weight functions f_1, \dots, f_t . This count can be controlled by Gowers U^{s+1} -norm.

We now first define Gowers uniformity norms and see how $\Lambda(f_1, \dots, f_t)$ can be controlled by $\|f_i\|_{U^2}$ in the case of complexity one systems provided that $|f_i| \leq 1$ for each $i = 1, \dots, t$. See [12, Appendix B and C] for more details about this norm.

Definition 6.5.2. Let $f : \mathbf{Z}_M \rightarrow \mathbf{C}$. We define the Gowers uniformity norm $\|f\|_{U^{s+1}(\mathbf{Z}_M)}$ by the formula

$$\|f\|_{U^{s+1}(\mathbf{Z}_M)}^{2^{s+1}} := \mathbf{E}_{x \in \mathbf{Z}_M, h \in \mathbf{Z}_M^{s+1}} \prod_{w \in \{0,1\}^{s+1}} C^{|w|} f(x + \sum_{j=1}^{s+1} w_j h_j), \quad (6.56)$$

where C is the complex conjugate operator on the space of complex valued functions on \mathbf{Z}_M .

Remark 6.5.1. For any function $f : \mathbf{Z}_M \rightarrow \mathbf{C}$, we have $\|f\|_{U^2}^4 = \sum_{r \in \mathbf{Z}_M} |\hat{f}(r)|^4$.

Proposition 6.5.1. Let $\Psi : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ be an affine linear system of complexity one and it is in exact 1-normal form at each $i \in \{1, \dots, t\}$. Let $f_1, \dots, f_t : \mathbf{Z}_M \rightarrow \mathbf{C}$ with $|f_i| \leq 1$ for each i . Then we have

$$|\Lambda(f_1, \dots, f_t)| \leq \min_i (\|f_i\|_{U^2}). \quad (6.57)$$

Proof. Let $\psi = (\psi_1, \dots, \psi_t) : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ be an affine linear system of complexity one and it is in exact 1-normal form at each $i \in [t]$. Let $\psi_i = \sum_{j=1}^d a_{ij} x_j + b_i$ for $1 \leq j \leq t$. Since ψ_1 is in exact 1-normal form, there exists two indices $j_1, j_2 \in [d]$, say $j_1 = 1, j_2 = 2$, such that $a_{11} a_{12} \neq 0$ and ψ_j does not depend on both the variables

Hence, in order to prove Theorem 6.1.2 it is enough to prove the following result.

Theorem 6.5.1. (*W-tricked version*)

Let V, F and $S(N, N^{1/4k+1}) := S_F(N, N^{1/4k+1})$ be as in Theorem 6.1.2. Let $A \subset \{1, \dots, N\}$ be such that $b + AW \subset S(NW, (NW)^{1/4k+1})$, where

$$W = \prod_{p \leq w} p, \quad w = c_0 \log N, \quad (b, W) = 1 \quad \text{with} \quad c_0 \in \left[\frac{1}{8k+2}, \frac{1}{4k+1} \right].$$

Let us also assume that, for any $\epsilon > 0$,

$$|A| \geq \alpha \left(\frac{\log w}{\log N} \right)^k N \quad \text{with} \quad \alpha \geq c (\log \log N)^{-\frac{1}{8m} + \epsilon}. \quad (6.54)$$

Then there exists $y \in A^t$ with distinct coordinates such that $Vy^t = 0$.

In the following three subsections we prepare for the proof of Theorem 6.5.1, and we conclude the section with the proof of it.

6.5.2 Controlling Ψ -patterns of complexity one

Let M be a prime number and let $\mathbf{Z}_M := \mathbf{Z}/M\mathbf{Z}$ be the group of residue classes modulo M . Let $\Psi : \mathbf{Z}_M^d \rightarrow \mathbf{Z}_M^t$ be an affine system of finite complexity s and it is in exact s -normal form at each $i \in \{1, \dots, t\}$.

Definition 6.5.1. The operator Λ on functions $f_1, \dots, f_t : \mathbf{Z}_M \rightarrow \mathbf{R}$ defined by

$$\Lambda(f_1, \dots, f_t) := \mathbf{E}_{n \in \mathbf{Z}_M^d} f_1[\psi_1(n)] \cdots f_t[\psi_t(n)]. \quad (6.55)$$

Let

$$X_W = \left\{ 1 \leq b \leq W : \left(\prod_{i=1}^k a_i b + b_i, W \right) = 1 \right\}. \quad (6.47)$$

By Brun sieve, we have

$$|X_W| \leq c_F \frac{W}{(\log w)^k}, \quad (6.48)$$

where c_F depends at most on F . Since $W \leq N^{1/4k+1}$, we have

$$\sum_{b \in X_W} |A \cap P(b)| = |A| \geq c_F \alpha \frac{N}{(\log N)^k}, \quad (6.49)$$

where $P(b) = \{n : n \equiv b \pmod{W}\}$. By Pigeonhole principle, there exists $b_0 \in X_W$ such that

$$|A \cap P(b_0)| \geq c \alpha \frac{N}{|X_W| (\log N)^k} \geq c_F \alpha \left(\frac{\log w}{\log N} \right)^k \frac{N}{W}, \quad (6.50)$$

where we used an upper for $|X_W|$ in the second inequality.

Let

$$N' := \left\lfloor \frac{N}{W} \right\rfloor = N^{\frac{8k+1}{8k+2} + o(1)}, \quad (6.51)$$

where we use the value of $W = \prod_{p \leq w} p$, from (6.46), with $c_0 = 1/8k + 2$.

Put

$$A' := \left\{ 1 \leq n \leq \frac{N - b_0}{W} : b_0 + nW \in A \subset S(N, N^{1/4k+1}) \right\}. \quad (6.52)$$

Then we have, from (6.50), that

$$|A'| \geq c_F \alpha \left(\frac{\log w}{\log N'} \right)^k N'. \quad (6.53)$$

Here λ'_A approximates λ_A means that the sum of Ψ -patterns of complexity one with weight function $\lambda_A - \lambda'_A$ is “small”. In fact, this Ψ -pattern count can be controlled by Gowers U^2 -norm $\|\lambda_A - \lambda'_A\|_{U^2}$. To prove Theorem 6.1.2, it is enough to find a function λ'_A so that $\|\lambda_A - \lambda'_A\|_{U^2}$ is small and $\sum_{n \in \mathbf{Z}/p\mathbf{Z}} \lambda'_A(n) \gg p$, which is the theme of the Transference principle. By Theorem 6.1.3, the count of Ψ -patterns of complexity one with weight function λ'_A is “large” as λ'_A is a dense subset of the integers. Therefore the count of Ψ -patterns with weight function λ_A is large. We discuss this in detail in the following subsections.

6.5.1 Reduction to a W -tricked set

Let us assume that N is a sufficiently large integer. Let $S(N, N^{1/4k+1})$ be a set as in Theorem 6.1.2 and let

$$A \subset S(N, N^{1/4k+1}) \quad \text{with } |A| = \alpha |S(N, N^{1/4k+1})|. \quad (6.44)$$

Using the estimate (6.3) on the cardinality of $S(N, N^{1/4k+1})$, we have

$$|A| \geq c_F \alpha \frac{N}{(\log N)^k}, \quad (6.45)$$

where c_F depends at most only upon F .

Let

$$W = \prod_{p \leq w} p, \quad \text{where } w = c_0 \log N, \quad \text{with } c_0 \in \left[\frac{1}{8k+2}, \frac{1}{4k+1} \right]. \quad (6.46)$$

□

We end this section with the proof of Theorem 6.4.1.

Proof of Theorem 6.4.1: We have $\mathbf{Z}/p\mathbf{Z} = B(\{1\}, \frac{1}{2}) := B_0$ and $A_p \subset \mathbf{Z}/p\mathbf{Z}$ with $|A_p| = \alpha|B_0|$. We apply Proposition 6.4.1 repeatedly, starting with $A_p^0 = A_p$ and $B_0 = \mathbf{Z}/p\mathbf{Z}$, to obtain a sequence of regular Bohr sets B_0, B_1, \dots, B_k and $A_p^0 \subset B_0, \dots, A_p^k \subset B_k$ of densities $\alpha_0, \dots, \alpha_k$. The iteration stops if A_p^k satisfies the case (i) of Proposition 6.4.1. We have $\alpha_i \geq \alpha_{i-1} + (c_1\alpha_{i-1})^{6m} \geq \alpha + i(c_1\alpha)^{6m}$ and hence $k \leq (c_1\alpha)^{-6m}$. Moreover we have $d(B_i) \leq d(B_{i-1}) + 1$ and if $d(B_i) = d(B_{i-1}) + 1$, then $\alpha_i \geq \alpha_{i-1} + (c_1\alpha_{i-1})^{2m}$. Therefore $d(B_k) \leq (c_1\alpha)^{-2m} + 1$. Moreover $\delta(B_i) \geq \frac{(c_1\alpha)^{6mq}}{4^q(1+\|\Phi_p\|_p)^{2q^2m}d_{i-1}^{2q}}\delta(B_{i-1})$ and hence $\delta(B_k) \geq \left(\frac{(c_1\alpha)^{10mq}}{4^q(1+\|\Phi_p\|_p)^{2q^2m}}\right)^k$. Since A_p^k satisfies the case (i) of Proposition 6.4.1, the result follows by using the bounds of $\delta(B_k)$, $d(B_k)$ and k , as well as the following inequality

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} \prod_{i=1}^m A_p(\phi^i(n)) \geq \mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} \prod_{i=1}^m A_p^k(\phi^i(n)).$$

6.5 Translation invariant equations in sifted sets

In this section we prove Theorem 6.1.2. We closely follow Henriot [15, Sections 6 and 7]. We first reduce our problem to the case of a W -tricked set which we also call A (by abuse of notation) as shown in subsection 6.5.1. We then consider this set as a subset of $\mathbf{Z}/p\mathbf{Z}$ for some suitable prime p . We appropriately normalize the indicator function of this set, i.e., λ_A , so that we can “approximate” this normalized indicator function λ_A by a function λ'_A which has “positive density” in the integers.

which implies that

$$\begin{aligned} & \left| \mathbf{E}_{n_l \in B_l} \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) e_p((n_0 + n_1 n_l + n_1 n'_l)y) \right| \\ & \geq |\mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l) e_p(n_1 n_l y)| - 200c'd \geq (c_1 \alpha)^{2m} / 4. \end{aligned}$$

On the other hand, for $n'_l \in B'_l$, we have $|1 - e_p(n_1 y n'_l)| \leq 8cc'\delta$. Hence,

$$\begin{aligned} & \left| \mathbf{E}_{n_l \in B_l} \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) e_p((n_0 + n_1 n_l + n_1 n'_l)y) \right| \\ & \leq \mathbf{E}_{n_l \in B_l} \left| \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) e_p(n_1 n'_l) \right| \\ & \leq \mathbf{E}_{n_l \in B_l} \left| \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) \right| + 8c'cd. \end{aligned}$$

So, from the two bounds above, it follows that

$$\mathbf{E}_{n_l \in B_l} \left| \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) \right| \geq \frac{(c_1 \alpha)^{2m}}{8}. \quad (6.42)$$

Note that

$$\mathbf{E}_{n_l \in B_l} \mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) \geq \delta(n_0) - 200c'd \geq -\frac{(c_1 \alpha)^{2m}}{16},$$

since $n_0 \notin E$. Hence, there exists $n_l \in B_l$ such that

$$\mathbf{E}_{n'_l \in B'_l} f(n_0 + n_1 n_l + n_1 n'_l) \geq \frac{(c_1 \alpha)^{2m}}{16}, \quad (6.43)$$

which proves the conclusion (3) of Proposition 6.4.1.

Note that

$$\begin{aligned}\mathbf{E}_{n_0 \in B_0} \delta(n_0) &= \mathbf{E}_{n_0 \in B_0} \mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l) \\ &\geq -200(1 + \|\phi\|_p)^{qm} \rho^q d \geq -\frac{(c_1 \alpha)^{6m}}{128},\end{aligned}$$

since $\mathbf{E}_{n_0 \in B_0} f(n_0) = 0$ and by our choice of ρ . This gives an upper bound

$$|E| \leq 3(c_1 \alpha)^{4m} |B_0|/4.$$

It follows that there exists $n_0 \in (1 - c)B_0 \setminus E$ and $y \in \mathbf{Z}/p\mathbf{Z}$ such that

$$|\mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l) e_p(n_1 n_l y)| \geq \frac{(c_1 \alpha)^{2m}}{2}. \quad (6.41)$$

Fix such an n_0 and y . We define B'_l by

$$B'_l := \text{Bohr}(S \cup \{y n_1\}, c' c \delta)$$

with $c' \leq (c_1 \alpha)^{2m} d^{-1}/2^{13}$. Then for any $n'_l \in B'_l$,

$$\left| \mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l + n_1 n'_l) - \mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l) e_p(n_1 n_l y) \right| \leq 200c' d,$$

Now let us assume that the conclusion (1) of Proposition 6.4.1 is false. Then by using the inequality (6.33) and the value of ρ , given in (6.36), we get that

$$\mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m A_p(\phi^i(n)) \leq \frac{\alpha^m}{4}. \quad (6.37)$$

Hence by Corollary 6.4.1, we have

$$\mathbf{E}_{n_0 \in B_0} \max_{y \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{E}_{n_l \in B_l} f(n_0 + a_{il} b_l n_l) e_p(y a_{il} b_l n_l)|^2 \geq (c_1 \alpha)^{4m}, \quad (6.38)$$

for some $1 \leq l \leq q$ and $1 \leq i \leq m$ with $a_{il} \neq 0$ and b_l defined as in (6.32). We also have $B_l = B_0|_c = \text{Bohr}(S, c\delta)$ with $(\rho/2)^q \leq c \leq \rho^q$.

Let us now also assume that the conclusion of Proposition 6.4.1 is false. Then we have

$$\mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l) < \frac{(c_1 \alpha)^{6m}}{128}, \quad (6.39)$$

for all $n_0 \in \mathbf{Z}/p\mathbf{Z}$ with $n_0 + n_1 B_l \subset B_0$, where $n_1 = a_{il} b_l$.

For $n_0 \in B_0$, we write $\delta(n_0) = \mathbf{E}_{n_l \in B_l} f(n_0 + n_1 n_l)$. Thus we have

$$\delta(n_0) \leq \frac{(c_1 \alpha)^{6m}}{128}, \quad (6.40)$$

for each $n_0 \in (1-c)B_0$. Let E be the set consisting of those integers n_0 in $(1-c)B_0$ with $\delta(n_0) \leq -(c_1 \alpha)^{2m}/32$.

where $c_2 > 0$ is an absolute constant. Let $b'_1 = a_{i_0j}b_j$ and

$$b'_2 = a_{i_0l} \prod_{\substack{1 \leq i \leq m \\ j \leq k \leq l-1 \\ a_{il} \neq 0 \\ (i,k) \neq (i_0,j)}} a_{ik}.$$

Then $a_{i_0l}b_l = b'_1b'_2$ and we may rewrite (6.35) as

$$\|f_{i_0}\|_{U^2(B_0, B_j, B_l, b'_1, b'_1 b'_2)}^4 \geq (c_2 \alpha)^{4m}.$$

Since we have $(|b'_1|_p + |b'_2|)^2 \leq (1 + \|\Phi\|_p)^{2qm}$, using Theorem 6.3.1 we see that (ii) of the claim holds. Hence the result follows. \square

Proof of Proposition 6.4.1:

We can assume that q is a large fixed integer. Let us assume that

$$A_p \subset B_0 := \text{Bohr}(S, \delta)$$

with $|A_p| = \alpha|B_0|$, and let set $f = A_p - \alpha|B_0|$.

Let ρ be a real number such that

$$(1 + \|\phi\|_p)^{2qm} \rho d = (c_1 \alpha)^{4m}, \tag{6.36}$$

where $c_1 > 0$ is a small absolute constant (at most depends on q, ϕ, m). Let B_1, \dots, B_q be regular Bohr sets with $B_1 = B_0|_{\rho_1}$ and $B_j = B_{j-1}|_{\rho_j}$ where $\rho_j \in [\rho/2, \rho)$ for all $1 \leq j \leq q$.

$$1. \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m A_p(\phi^i(n)) \leq \frac{\alpha^m}{4}.$$

2. There exist i, l with $1 \leq i \leq m$ and $1 \leq l \leq q$ for which $a_{il} \neq 0$. For such an a_{il} we have

$$\mathbf{E}_{n_0 \in B_0} \max_{y \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{E}_{n_l \in B_l} f(n_0 + a_{il} b_l n_l) e_p(y a_{il} b_l n_l)|^2 \geq (c_1 \alpha)^{4m},$$

where $f = A_p - \alpha B_0$.

Proof. We set $f_0 = \alpha B_0$ and $f_1 = A_p - \alpha B_0$. Then we have

$$\begin{aligned} \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m A_p(\phi^i(n)) &= \alpha^m \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m B_0(\phi^i(n)) \\ &\quad + \sum_{\substack{w \in \{0,1\}^m \\ w \neq 0}} \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m f_{w(i)}(\phi^i(n)). \end{aligned}$$

Using Lemma 6.2.7 (ii) and the fact that $b_0 = 1$, we see that the first term in the right hand side of the above equality is equal to $\alpha^m(1 + O((1 + \|\Phi\|_p)^{qm} \rho d))$ and hence at least $\frac{\alpha^m}{2}$. Therefore if the case (i) of the claim does not hold, then for some $w \in \{0, 1\}^m$ with $w \neq 0$, we have

$$-\mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m f_{w(i)}(\phi^i(n)) \geq \frac{\alpha^m}{2^{m+2}},$$

where $f_{w(i_0)} = f_1 = A_p - \alpha B_0$ for some i_0 . Hence using Lemma 6.4.4 there exist $1 \leq j < l \leq q$ with $a_{i_0 j} a_{i_0 l} \neq 0$, such that

$$\|f_{i_0}\|_{U^2(B_0, B_j, B_l, a_{i_0 j} b_j, a_{i_0 l} b_l)}^4 \geq (c_2 \alpha)^{4m}, \quad (6.35)$$

summations, we see that the right hand side of (6.34) is

$$\mathbf{E}_{n_j, n'_j \in b_j B_j} \mathbf{E}_{n \in \prod_{\substack{0 \leq i \leq q \\ i \neq j}} b_i B_i} \prod_{s \in X_j} f_s(n_0 + a_{sj} n_j + \overline{\sum_{\substack{k=1 \\ k \neq j}}^q a_{sk} n_k}) f_s(n_0 + a_{sj} n'_j + \overline{\sum_{\substack{k=1 \\ k \neq j}}^q a_{sk} n_k}).$$

By one more application of Cauchy-Schwarz inequality, we see that $|\Lambda|^4$ is at most

$$\mathbf{E}_{n_j, n'_j \in b_j B_j} \mathbf{E}_{n \in \prod_{\substack{0 \leq i \leq q \\ i \neq j, l}} b_i B_i} \left| \mathbf{E}_{n_l \in b_l B_l} f_1(n_0 + a_{1j} n_j + a_{1l} n_l + \sum_{\substack{k=1 \\ k \neq j, l}}^q a_{1k} n_k) \times \right. \\ \left. \overline{f_1(n_0 + a_{1j} n'_j + a_{1l} n_l + \sum_{\substack{k=1 \\ k \neq j, l}}^q a_{1k} n_k)} \right|^2$$

which equals , by Lemma 6.2.7, to

$$\mathbf{E}_{n_0 \in B_0} \mathbf{E}_{n_j, n'_j \in b_j B_j} \left| \mathbf{E}_{n_l \in b_l B_l} f_1(n_0 + a_{1j} n_j + a_{1l} n_l) \overline{f_1(n_0 + a_{1j} n'_j + a_{1l} n_l)} \right|^2 \\ + O((1 + \|\phi\|_p)^{qm} \rho d).$$

After opening the square in the above sum, we see that the sum is same as

$$\|f_1\|_{U^2(B_0, B_j, B_l, a_{1j} b_j, a_{1l} b_l)}^4.$$

Thus, the assertion of the lemma holds. □

Corollary 6.4.1. *If $(1 + \|\Phi\|_p)^{2qm} \rho d \leq (c_1 \alpha)^{4m}$ for a sufficiently small absolute constant $c_1 > 0$, then one of the following holds.*

Proof. Let us assume that $i_0 = 1$. By the assumption on $\Phi = (\phi^1, \dots, \phi^m) : (\mathbf{Z}/p\mathbf{Z})^{q+1} \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$, we have

$$\phi^1(n_0, \dots, n_q) = n_0 + \sum_{i=1}^q a_{1i} n_i$$

and there exists $1 \leq j < l \leq q$ such that $a_{1j} a_{1l} \neq 0$, and for $i \neq 1$, ϕ^i may depend on at most one of n_j and n_l . For $k = j, l$, let $X_k = \{1 \leq r \leq q : \phi^r \text{ depends at most on } n_k\}$.

Let

$$\Lambda := \mathbf{E}_{n \in \prod_{i=0}^q b_i B_i} \prod_{r=1}^m f_r(\phi^r(n))$$

Thus, we have

$$\begin{aligned} \Lambda &= \mathbf{E}_{n \in \prod_{0 \leq i \leq q} b_i B_i} \prod_{s \in X_l \setminus \{1\}} f_s(\phi^s(n)) \prod_{s \in X_j} f_s(\phi^s(n)) \\ &= \mathbf{E}_{n \in \prod_{\substack{0 \leq i \leq q \\ i \neq j}} b_i B_i} \prod_{s \in X_l \setminus \{1\}} f_s(\phi^s(n)) \mathbf{E}_{n_j \in b_j B_j} \prod_{s \in X_j} f_s(\phi^s(n)) \end{aligned}$$

By applying Cauchy-Schwarz inequality, we see that

$$|\Lambda|^2 \leq \mathbf{E}_{n \in \prod_{\substack{0 \leq i \leq q \\ i \neq j}} b_i B_i} \left| \mathbf{E}_{n_j \in b_j B_j} \prod_{s \in X_j} f_s(\phi^s(n)) \right|^2. \quad (6.34)$$

After opening the square on the right hand side of (6.34) and interchanging the

such that with $A_p^1 = \left(\frac{A_p - n_0}{n_1}\right)$, we have

$$|A_p^1 \cap B_1| \geq (\alpha + (c_1 \alpha)^{2m}) |B_1|,$$

where $c_1 > 0$ is an absolute constant.

Let ρ be a real number and B_1, \dots, B_q be regular Bohr sets with $B_1 = B_{0, \rho_1}$ and $B_j = B_{j-1, \rho_j}$ where $\rho_j \in (\frac{\rho}{2}, \rho]$ for all $1 \leq j \leq q$. Let $\Phi_p = (\phi^1, \dots, \phi^m) : (\mathbf{Z}/p\mathbf{Z})^{q+1} \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ be the linear map as in Theorem 6.4.1. Then for any i , with $1 \leq i \leq m$, we have $\phi^i(n_0, \dots, n_q) = n_0 + \sum_{j=1}^q a_{ij} n_j$ for some $a_{ij} \in \mathbf{Z}/p\mathbf{Z}$. We set $b_0 = 1$ and for any j with $1 \leq j \leq q$, we set

$$b_j = \prod_{\substack{1 \leq i \leq m \\ 1 \leq l \leq j-1 \\ a_{il} \neq 0}} a_{il}. \quad (6.32)$$

Using Lemma 6.2.5, we have

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} \prod_{i=1}^m A_p(\phi^i(n)) \geq \left(\frac{\rho \delta}{2}\right)^{2q^2 d} \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m A_p(\phi^i(n)). \quad (6.33)$$

The following lemma is the local version of generalized Von Neumann theorem.

Lemma 6.4.4. *Let $f_1, \dots, f_m : \mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$ be a function with $\|f_i\|_\infty \leq 1$. Then for any i_0 , there exist $1 \leq j < l \leq q$, with $a_{i_0 j} a_{i_0 l} \neq 0$ such that*

$$\left| \mathbf{E}_{n \in \prod_{j=0}^q b_j B_j} \prod_{i=1}^m f_i(\phi^i(n)) \right|^4 \leq \|f_{i_0}\|_{U^2(B_0, B_j, B_l, a_{i_0 j} b_j, a_{i_0 l} b_l)}^4 + O((1 + \|\Phi\|_p)^{qm} \rho d).$$

To prove Theorem 6.4.1, we need the following density increment argument which is noted as a following proposition.

Proposition 6.4.1. *Let $\alpha > 0$ be a real number, p be a prime and $B_0 \subset \mathbf{Z}/p\mathbf{Z}$ be a regular Bohr set of dimension d and radius δ . Let $A_p \subset B_0$ with $|A_p| \geq \alpha|B_0|$. Let $\Phi = (\phi^1, \dots, \phi^m) : (\mathbf{Z}/p\mathbf{Z})^{q+1} \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ be a linear map as in Theorem 6.4.1. Then one of the following holds.*

1. *We have*

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} \prod_{i=1}^m A_p(\phi^i(n)) \geq \left(\frac{c_1 \delta \alpha}{(1 + \|\Phi\|_p) d} \right)^{10q^3 m d},$$

where $c_1 > 0$ is an absolute constant.

2. *There exists a regular Bohr set B_1 , $n_0 \in \mathbf{Z}/p\mathbf{Z}$, $n_1 \in (\mathbf{Z}/p\mathbf{Z})^*$ with*

$$\delta(B_1) \geq \frac{(c_1 \alpha)^{6mq}}{4^q (1 + \|\Phi_p\|_p)^{2q^2 m} d^{2q}} \delta(B_0) \text{ and } d(B_1) = d(B_0)$$

such that with $A_p^1 = \left(\frac{A_p - n_0}{n_1} \right)$, we have

$$|A_p^1 \cap B_1| \geq (\alpha + (c_1 \alpha)^{6m}) |B_1|,$$

where $c_1 > 0$ is an absolute constant.

3. *There exists a regular Bohr set B_1 , $n_0 \in \mathbf{Z}/p\mathbf{Z}$, $n_1 \in (\mathbf{Z}/p\mathbf{Z})^*$ with*

$$\delta(B_1) \geq \frac{(c_1 \alpha)^{6mq}}{4^q (1 + \|\Phi_p\|_p)^{2q^2 m} d^{2q}} \delta(B_0) \text{ and } d(B_1) \leq d(B_0) + 1$$

follows that

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} f_p(\psi_p^1(n)) \cdots f_p(\psi_p^m(n)) = p^{-(m-r)} \sum_{y=(y_1, \dots, y_m) \in \ker(V_p)} f_p(y_1) \cdots f_p(y_m). \quad (6.29)$$

Since we have $\text{supp}(f) \subset [-N, N]$ and $p > 2\|V\|N$, it follows that

$$\begin{aligned} \sum_{y \in \ker(V_p)} f_p(y_1) \cdots f_p(y_m) &= \sum_{y \in \ker(V_p) \cap [-N, N]^m} f_p(y_1) \cdots f_p(y_m) \\ &= \sum_{y \in \ker(V)} f(y_1) \cdots f(y_m). \end{aligned} \quad (6.30)$$

Hence the result follows. □

PROOF OF PROPOSITION 6.1.3.—

Choose a prime p with $2(\|V\|^m + \|\psi\|^m)N < p \leq 4(\|V\|^m + \|\psi\|^m)N$. Then p satisfies the assumptions of Lemma 6.4.2 and Lemma 6.4.3. Let A_p denotes the image of A under the natural projection map $\pi : \mathbf{Z} \rightarrow \mathbf{Z}/p\mathbf{Z}$. Then we have $|A_p| = |A| \geq \frac{\alpha}{4(\|V\|^m + \|\psi\|^m)} p = c\alpha p$, where $c > 0$ is a constant depending only upon V and m . Then using Theorem 6.4.1 with $\Phi = \Psi_p$, we obtain

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} A_p(\psi_p^1(n)) \cdots A_p(\psi_p^m(n)) \geq c_1 \exp(-c_2 \alpha^{-8m} \log \frac{1}{\alpha}), \quad (6.31)$$

where $c_1, c_2 > 0$ are constants depending only upon V and m . Now using Lemma 6.4.3 with f equal to the indicator function of A , we obtain the result. □

For any prime p , let $\Psi_p = (\psi_p^1, \dots, \psi_p^m) : (\mathbf{Z}/p\mathbf{Z})^{q+1} \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ be the linear map induced by ψ and $V_p : (\mathbf{Z}/p\mathbf{Z})^m \rightarrow (\mathbf{Z}/p\mathbf{Z})^r$ be the linear map induced by V . We have the following lemma which is a simple consequence of Lemma 6.2.4.

Lemma 6.4.2. *For any prime p with $p > \max(\|\psi\|^m, \|V\|^m)$, we have*

$$\text{Im}(\Psi_p) = \ker(V_p) \text{ and } \dim(\ker(\Psi_p)) = q + 1 - m + r. \quad (6.27)$$

Moreover Ψ_p is in exact 1-normal form with $\|\psi\| = \|\Psi_p\|_p$.

Proof. Since $\text{Im}(\psi) \subset \ker(V)$, it follows that for any p , we have $\text{Im}(\Psi_p) \subset \ker(V_p)$. Using Lemma 6.2.4, we have $\dim(\text{Im}(\Psi_p)) = \dim(\text{Im}(\Psi))$ and $\dim(\ker(V_p)) = \dim(\ker(V))$. Since we have $\dim(\text{Im}(\Psi)) = \dim(\ker(V))$, it follows that $\dim(\text{Im}(\Psi_p)) = \dim(\ker(V_p))$. Hence the first claim follows. The second claim is easy to verify. \square

The following lemma is due to Henriot [15, Lemma 5].

Lemma 6.4.3. *Let ψ and Ψ_p be as above. Let $f : \mathbf{Z} \rightarrow \mathbf{C}$ be a function with $\text{supp}(f) \subset [-N, N]$ and let $p > 2(\|V\|^m + \|\psi\|^m)N$. Then f also induces a map $f_p : \mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$ and we have*

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} f_p(\psi_p^1(n)) \dots f_p(\psi_p^m(n)) = p^{-(m-r)} \sum_{y=(y_1, \dots, y_t) \in \text{Ker}(V)} f(y_1) \dots f(y_t). \quad (6.28)$$

Proof. From Lemma 6.4.2, we have $\dim(\ker(\Psi_p)) = q + 1 - (m - r)$. Therefore it

Proof. Since $\dim(\ker(V)) = m - r$, we have that $\ker(V) \cap \mathbf{Z}^m$ is a free \mathbf{Z} -module of rank $m - r$. Thus, there exist \mathbf{Z} -linear vectors $X_1 = (1, \dots, 1)$, $X_2 = (a_{12}, \dots, a_{m2})$, \dots and $X_{m-r} = (a_{1m-r}, \dots, a_{mm-r})$ such that

$$\ker(V) \cap \mathbf{Z}^m = \mathbf{Z}X_1 + \dots + \mathbf{Z}X_{m-r}.$$

Thus, we have a \mathbf{Z} -linear isomorphism

$$\psi = (\psi^1, \dots, \psi^m) : \mathbf{Z}^{m-r} \rightarrow \ker(V) \cap \mathbf{Z}^m$$

where $\psi^i(x_1, \dots, x_{m-r}) = x_1 + \sum_{j=2}^{m-r} a_{ij}x_j$ for $i = 1, \dots, m$.

We also have the complexity of ψ is 1, as the complexity of V is 1. In particular, ψ has complexity at most 1 at $i \in [m]$. Thus, there exists a partition $Y_1 \cup Y_2$ of $[m] \setminus \{i\}$ such that $\psi^i \notin \langle \psi^j : j \in Y_l \rangle$ for $l = 1, 2$. Thus, we can find vectors $f_1, f_2 \in \mathbf{Z}^{m-r}$ such that $\psi^i(f_1) \neq 0$, $\psi^i(f_2) \neq 0$ and $\psi^j(f_1) = 0$ for $j \in Y_1$ and $\psi^l(f_2) = 0$ for $l \in Y_2$. Let $d' = m - r + 2$. We now define a linear system $\psi' : \mathbf{Z}^{d'} \rightarrow \mathbf{Z}^m$ as follows

$$\psi'(x, y_1, y_2) = \psi(x + y_1f_1 + y_2f_2).$$

Note that in this newly defined linear system ψ^i depends on the variables y_1, y_2 but ψ^j may depend at most on one of y_1 and y_2 . Thus, the system is in exact 1-normal form at $i \in [m]$. If we continue this process, we can make a linear system which is in exact 1-normal form at each $i \in [m]$. By our choice of ψ itself, second condition of the lemma holds. \square

□

6.4 Translation invariant equations in integers

In this section we give a proof of Theorem 6.1.3 by deducing it from the following theorem.

Theorem 6.4.1. *Let p be a prime and $A_p \subset \mathbf{Z}/p\mathbf{Z}$ with $|A_p| \geq \alpha p$. Let $\Phi = (\phi^1, \dots, \phi^m) : (\mathbf{Z}/p\mathbf{Z})^{q+1} \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ be a linear map with $\phi^i(n_0, \dots, n_q) = n_0 + \phi^i(0, n_1, \dots, n_q)$ and Φ is in 1-normal form. Then we have*

$$\mathbf{E}_{n \in (\mathbf{Z}/p\mathbf{Z})^{q+1}} A_p(\phi^1(n)) \dots A_p(\phi^m(n)) \geq c_1 \exp(-c_2 \alpha^{-8m} \log \frac{1}{\alpha}), \quad (6.26)$$

where $c_1, c_2 > 0$ are constants depending at most on $m, \|\Phi\|_p$ and q .

Let $V \in M_{r \times m}(\mathbf{Z})$ be a translation invariant matrix of complexity one and rank r . We think of V as a linear map from \mathbf{Q}^m to \mathbf{Q}^r . Then the dimension of $\ker(V)$ is $m - r$. In the following lemma we give a linear parametrization of $\ker(V)$.

Lemma 6.4.1. *There exists a linear parametrization $\psi = (\psi_1, \dots, \psi_m) : \mathbf{Z}^{q+1} \rightarrow \ker(V) \cap \mathbf{Z}^m$ such that*

1. ψ is in exact 1-normal form at each $i = 1, \dots, m$ and
2. For each $i = 1, \dots, m$, we have

$$\psi^i(x_0, \dots, x_q) = x_0 + \psi^i(0, x_1, \dots, x_q).$$

and $(b_1 b_2 \widehat{B} * f_{t_2}) = \widehat{b_1 b_2 B} \cdot \widehat{f_{t_2}}$. Thus we have

$$\begin{aligned} h(n_0) &\leq \frac{|\text{supp}(f)| p^2}{|B_1|^2 |B_2|^2} \max_{t_2 \in \mathbf{Z}_p} \sum_{t_1 \in \mathbf{Z}_p} |(b_1 b_2 B * f_{t_2})(t_1)|^2 \\ &\leq \frac{|\text{supp}(f)|}{|B_1|^2 |B_2|^2} \max_{t_2 \in \mathbf{Z}_p} \sum_{t_1 \in \mathbf{Z}_p} \left| \sum_x b_1 b_2 B(t_1 - x) f_{t_2}(x) \right|^2 \end{aligned}$$

Thus we have

$$h(n_0) \leq \frac{|\text{supp}(f)|}{|B_1|^2 |B_2|^2} \sum_{n \in \mathbf{Z}/p\mathbf{Z}} \max_{t \in \mathbf{Z}/p\mathbf{Z}} \left| \sum_{n_2 \in B_2} f(n + b_1 b_2 n_2) e_p(t b_1 b_2 n_2) \right|^2. \quad (6.23)$$

Since $\text{supp}(f) \subset n_0 + b_1 B_1 + b_1 b_2 B_2$, the right hand side of (6.23) is not more than

$$\frac{|\text{supp}(f)|}{|B_1|^2 |B_2|^2} \sum_{n \in n_0 + b_1 B_1 + 2b_1 b_2 B_2} \max_{t \in \mathbf{Z}/p\mathbf{Z}} \left| \sum_{n_2 \in B_2} f(n + b_1 b_2 n_2) e_p(t b_1 b_2 n_2) \right|^2. \quad (6.24)$$

Using the fact that $|n_0 + b_1 B_1 + 2b_1 b_2 B_2| = (1 + O(|b_2|_p \rho'' d)) |B_1|$ and we choosing ρ'' small enough so that $(1 + O(|b_2|_p \rho'' d)) \leq 2$, we see that

$$g(n_0) \leq 4 \mathbf{E}_{n_1 \in b_1 B_1 + 2b_1 b_2 B_2} \max_{t \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{E}_{n_2 \in B_2} f(n_0 + n_1 + b_1 b_2 n_2) e_p(b_1 b_2 n_2 t)|^2. \quad (6.25)$$

By the assumptions $\|f\|_{U^2(X_0, X_1, X_2, b_1, b_1 b_2)}^4 = \mathbf{E}_{n_0 \in B_0} g(n_0) \geq \eta^4$ and

$$(|b_1|_p + |b_2|_p)^2 (\rho' + \rho'') \leq c \eta^4 / d,$$

the assertion of the theorem follows from (6.25) and Lemma 6.2.7.

which shows that $g(n_0) \geq 0$ for all $n_0 \in B_0$. We also have an inequality

$$g(n_0) \leq \frac{1}{|B_1|^2|B_2|^2} \sum_{n_1^0, n_1^1 \in \mathbf{Z}/p\mathbf{Z}} \left| \sum_{n_2 \in B_2} f(n_0 + b_1 n_1^0 + b_1 b_2 n_2) \overline{f(n_0 + b_1 n_1^1 + b_1 b_2 n_2)} \right|^2. \quad (6.21)$$

For each $n_0 \in B_0$, we now estimate $g(n_0)$. For this we can assume, by (6.20), that $\text{supp}(f) \subseteq n_0 + b_1 B_1 + b_1 b_2 B_2$, thus we have $|\text{supp}(f)| \leq (1 + O(|b_2|_p \rho'' d)) |B_1|$. Let us write $h(n_0)$ for the right hand side of the above inequality (6.21), then note that

$$h(n_0) = \frac{1}{|B_1|^2|B_2|^2} \sum_{n_2, n_2' \in B_2} \sum_{\substack{x, y, z, w \in \mathbf{Z}/p\mathbf{Z} \\ x - b_1 b_2 n_2 = z - b_1 b_2 n_2' \\ x - y = z - w}} f(x) \overline{f(y)} \overline{f(z)} f(w). \quad (6.22)$$

The expression for $h(n_0)$ can be rewritten, using Fourier transform, as follows

$$\begin{aligned} h(n_0) &= \frac{p^4}{|B_1|^2|B_2|^2} \sum_{t_1, t_2 \in \mathbf{Z}/p\mathbf{Z}} \left| \widehat{B}_2(b_1 b_2 t_1) \right|^2 \left| \widehat{f}(t_1 + t_2) \right|^2 \left| \widehat{f}(t_2) \right|^2 \\ &\leq \frac{p^4}{|B_1|^2|B_2|^2} \max_{t_2 \in \mathbf{Z}_p} \sum_{t_1 \in \mathbf{Z}_p} \left| \widehat{B}_2(b_1 b_2 t_1) \right|^2 \left| \widehat{f}(t_1 + t_2) \right|^2 \sum_{t_2 \in \mathbf{Z}_p} \left| \widehat{f}(t_2) \right|^2 \end{aligned}$$

By Parseval's identity $\mathbf{E}_x |g(x)|^2 = \sum_{\xi} |\widehat{g}(\xi)|^2$, we see that

$$h(n_0) \leq \frac{|\text{supp}(f)| p^3}{|B_1|^2|B_2|^2} \max_{t_2 \in \mathbf{Z}_p} \sum_{t_1 \in \mathbf{Z}_p} \left| \widehat{B}_2(b_1 b_2 t_1) \right|^2 \left| \widehat{f}(t_1 + t_2) \right|^2.$$

Note that $\widehat{b_1 b_2 B}(t_1) = \widehat{B}(b_1 b_2 t_1)$, $\widehat{f}_{t_2}(t_1) = \widehat{f}(t_1 + t_2)$ where

$$f_{t_2}(x) = f(x) e(-xt_2/p)$$

We call the following theorem as an inverse theorem for the local Gowers U^2 -norm. The proof of this theorem closely follows that of Theorem 3.2 in [33], but here we take advantage of the positivity of each “individual” summand to get an improvement over Theorem 3.2 in [33], whereas Shao worked with “certain” summands only.

Theorem 6.3.1. *Let $f : \mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$ be a function with $|f| \leq 1$. Let $\eta \in (0, 1)$ be a real number and $b_1, b_2 \in (\mathbf{Z}/p\mathbf{Z})^*$. Let B_0, B_1, B_2 be regular Bohr sets of rank d with*

$$B_1 = B_{0|\rho'}, \text{ and } B_2 = B_{1|\rho''}, \quad (6.17)$$

where ρ' and ρ'' are positive real numbers with $(|b_1|_p + |b_2|_p)^2(\rho' + \rho'') \leq \frac{c\eta^4}{d}$ for some sufficiently small absolute constant $c > 0$. If $\|f\|_{U^2(B_0, B_1, B_2, b_1, b_2)} \geq \eta$, then

$$\mathbf{E}_{n_0 \in B_0} \max_{t \in \mathbf{Z}/p\mathbf{Z}} |\mathbf{E}_{n_2 \in B_2} f(n_0 + b_1 b_2 n_2) e_p(b_1 b_2 n_2 t)|^2 \geq \frac{\eta^4}{40}. \quad (6.18)$$

Proof. By the definition of the local Gowers U^2 -norm $\|f\|_{U^2(B_0, B_1, B_2, b_1, b_2)}^4$ of a function f , we have

$$\|f\|_{U^2(X_0, X_1, X_2, b_1, b_2)}^4 = \mathbf{E}_{n_0 \in B_0} g(n_0), \quad (6.19)$$

where $g(n_0) = \mathbf{E}_{n_1^0, n_1^1 \in B_1} \mathbf{E}_{n_2^0, n_2^1 \in B_2} \prod_{\omega \in \{0, 1\}^2} C^{|\omega|} f(n_0 + b_1 n_1^{w_1} + b_1 b_2 n_2^{w_2})$.

Observe that for each $n_0 \in B_0$, we have

$$g(n_0) = \frac{1}{|B_1|^2 |B_2|^2} \sum_{n_1^0, n_1^1 \in B_1} \left| \sum_{n_2 \in B_2} f(n_0 + b_1 n_1^0 + b_1 b_2 n_2) \overline{f(n_0 + b_1 n_1^1 + b_1 b_2 n_2)} \right|^2, \quad (6.20)$$

Lemma 6.2.8. *Let $f, g : \mathbf{Z}_N \rightarrow \mathbf{C}$. Then*

1. *(Plancherel theorem)*

$$\mathbf{E}_{n \in \mathbf{Z}_N} f(n) \overline{g(n)} = \sum_{\xi \in \mathbf{Z}_N} \hat{f}(\xi) \overline{\hat{g}(\xi)}.$$

2. *(Fourier inversion)*

$$f(x) = \sum_{\xi \in \mathbf{N}_N} \hat{f}(\xi) e\left(\frac{\xi x}{N}\right)$$

3. $\widehat{f * g} = \hat{f} \hat{g}$.

6.3 Inverse theorem for local Gowers U^2 -norm

In this section we define a local Gowers U^2 -norm and obtain an inverse theorem for it, which roughly says that if the local Gowers U^2 -norm of a function is large, then this function must possess some structure.

Definition 6.3.1. *(Local Gowers U^2 -norm) Let p be a prime. Given $X_0, X_1, X_2 \subset \mathbf{Z}/p\mathbf{Z}$ and $b_1, b_2 \in (\mathbf{Z}/p\mathbf{Z})^*$. Then the local Gowers U^2 -norm of any function $f : \mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$ is defined as follows*

$$\|f\|_{U^2(X_0, X_1, X_2, b_1, b_2)}^4 := \mathbf{E}_{\substack{n_0 \in X_0 \\ n_1^0, n_1^1 \in X_1 \\ n_2^0, n_2^1 \in X_2}} \prod_{\omega \in \{0,1\}^2} C^{|\omega|} f(n_0 + b_1 n_1^{w_1} + b_2 n_2^{w_2}), \quad (6.16)$$

where $|\omega| = w_1 + w_2$ and C is the involution defined as $Cf = \bar{f}$ on the space of functions from $\mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$.

Lemma 6.2.7. *Let $B \subset \mathbf{Z}/p\mathbf{Z}$ be a regular Bohr set of rank d and $f : \mathbf{Z}/p\mathbf{Z} \rightarrow \mathbf{C}$ be a function with $|f| \leq 1$. Then for any $0 \leq \rho \leq \frac{1}{100d}$ the following holds.*

1. $\mathbf{E}_{n \in B} \mathbf{I}_{n \in B_{|1-\rho}} f(n) = \mathbf{E}_{n \in B} f(n) + O(\rho d)$.

2. *We have $\mathbf{E}_{n_1 \in X} \mathbf{E}_{n \in B} f(n + n_1) = \mathbf{E}_{n \in B} f(n) + O(\rho d)$ for any $X \subset B_{|\rho}$.*

We now recall some basic finite Fourier analysis. For more about this we refer to sections 4.1 and 4.2 of [36].

Let N be an integer and let $\mathbf{Z}_N := \mathbf{Z}/N\mathbf{Z}$ be a cyclic group of order N .

Definition 6.2.7. *Let $f, g : \mathbf{Z}_N \rightarrow \mathbf{C}$. Then*

(i) *The Fourier transform \hat{f} of f defined as follows*

$$\hat{f}(r) = \mathbf{E}_{n \in \mathbf{Z}_N} f(n) e\left(\frac{-nr}{N}\right).$$

(ii) *The convolution $f * g$ of f and g is defined as*

$$f * g(n) = \mathbf{Z}_{m \in \mathbf{Z}_N} f(n - m) g(m).$$

(iii) *For any $0 < p < \infty$, we define the $L^p(\mathbf{Z}_N)$ norm of f to be the quantity*

$$\|f\|_{L^p(\mathbf{Z}_N)} := \|f\|_p = \left(\mathbf{E}_{n \in \mathbf{Z}_N} |f(n)|^p\right)^{1/p}.$$

We have the following properties of Fourier transform.

$\delta > 0$ is

$$\text{Bohr}(S, \delta) := \{x \in \mathbf{Z}/p\mathbf{Z} : |xs|_p \leq \delta p \text{ for all } s \in S\}, \quad (6.13)$$

and its rank d is defined as $d := |S|$.

The following lemma, [36, Lemma 4.20], says that Bohr sets are not small in size.

More precisely we have

Lemma 6.2.5. *Let $S \subset \mathbf{Z}/p\mathbf{Z}$ and let δ be a real number with $0 \leq \delta \leq 1/2$, then we have*

$$|\text{Bohr}(S, \delta)| \geq \delta^{|S|} p. \quad (6.14)$$

We often write B to denote the set $\text{Bohr}(S, \delta)$ and write $\delta(B)$ and $d(B)$ to denote its radius and rank respectively. The ρ -dilate $B|_\rho$ of the Bohr set B is defined by $\text{Bohr}(S, \delta)|_\rho := \text{Bohr}(S, \rho\delta)$.

We say that a Bohr set B is regular if for every $0 \leq \rho \leq \frac{1}{100d}$, we have

$$(1 - 100\rho d) |B| \leq |B|_{1 \pm \rho} \leq (1 + 100\rho d) |B|. \quad (6.15)$$

The following lemma, see [36, Lemma 4.25], says that any Bohr set can be made into a regular Bohr set by a small dilation. More precisely we have

Lemma 6.2.6. *Given any real number ρ with $0 < \rho < 1$ and a Bohr set B , there exists $\rho' \in [\frac{\rho}{2}, \rho]$ such that $B|_{\rho'}$ is a regular Bohr set.*

The following lemma is a crucial one used repeatedly in the sequel and its proof follows straightaway from the definition of a regular Bohr set.

say that V is translation invariant if

$$a_{i1} + \cdots + a_{im} = 0 \quad \forall i = 1, \dots, r. \quad (6.11)$$

The following proposition, due to Henriot [15, Proposition 2], gives a matrix complexity criterion.

Proposition 6.2.1. *Consider a matrix $V \in M_{r \times m}(\mathbf{Z})$ with rows R_1, \dots, R_r and $m \geq 2$, and a system of linear forms $\Psi : \mathbf{Q}^q \rightarrow \ker(V)$. Then Ψ has complexity at most s_0 at i if and only if there exists $0 \leq s \leq s_0$ and a partition $\{1, \dots, m\} \setminus \{i\} = X_1 \sqcup \cdots \sqcup X_{s+1}$ into non-empty sets such that, for every $k = 1, \dots, s+1$,*

$$(e_i + \sum_{j \in X_k} \mathbf{Q}e_j) \cap \langle R_1^t, \dots, R_r^t \rangle = \emptyset, \quad (6.12)$$

where $(e_i)_{1 \leq i \leq t}$ is the canonical basis of \mathbf{Q}^t .

Definition 6.2.5. *(Complexity of a matrix) Let $V = (a_{ij}) \in M_{r \times m}(\mathbf{Z})$. Then V defines a linear map from \mathbf{Q}^m to \mathbf{Q}^r . Let us take any surjective linear map $\Psi : \mathbf{Q}^q \rightarrow \ker(V) \cap \mathbf{Q}^m$. The complexity of V is defined as the complexity of Ψ .*

Note that the definition of complexity of a matrix is well defined by Proposition 6.2.1, that is, which does not depend on the choice of linear parametrisation of the $\ker(V)$.

We now recall the definition of Bohr and regular Bohr sets and their properties.

Definition 6.2.6. *(Bohr set) A Bohr set of a frequency set $S \subset \mathbf{Z}/p\mathbf{Z}$ and radius*

induced by ψ satisfy the following property:

$$\text{rk}(\text{Im}(\psi)) = \dim(\text{Im}(\Psi)) = \dim(\text{Im}(\Psi_p)). \quad (6.10)$$

Proof. By clearing denominators, we may choose a basis $\{\mathbf{y}^1, \dots, \mathbf{y}^{r_1}\}$ of $\text{Im}(\Psi)$ in such a way that for every j , we have $\mathbf{y}^j \in \mathbf{Z}^m$ and there exists $\mathbf{x}^j \in \mathbf{Z}^q$ such that $\psi(\mathbf{x}^j) = \mathbf{y}^j$. Hence $\text{Im}(\psi)$ contains a submodule of rank equal to $\dim(\text{Im}(\Psi))$ which implies that $\text{rk}(\text{Im}(\psi)) \geq \dim(\text{Im}(\Psi))$. Suppose $\{\mathbf{z}^1, \dots, \mathbf{z}^r\}$ is a basis of $\text{Im}(\psi)$, then \mathbf{z}^j 's are also linearly independent over \mathbf{Q} . Hence it follows that $\text{rk}(\text{Im}(\psi)) = \dim(\text{Im}(\Psi))$.

Since $\psi = (\psi^1, \dots, \psi^m) : \mathbf{Z}^q \rightarrow \mathbf{Z}^m$ is a \mathbf{Z} -linear map, we have $\psi^i(x_1, \dots, x_q) = \sum_{j=1}^q a_{ij}x_j$ for some $a_{ij} \in \mathbf{Z}$. The matrix of the linear map Ψ is equal to $(a_{ij}) \in M_{q \times m}(\mathbf{Q})$, whereas the matrix of the linear map Ψ_p is equal to $(\overline{a_{ij}})$. Here $\overline{a_{ij}} \in \mathbf{Z}/p\mathbf{Z}$ is the image of a_{ij} under the natural projection map from \mathbf{Z} to $\mathbf{Z}/p\mathbf{Z}$.

The dimension of the image of Ψ is equal to the maximum integer r such that there exists a subset $I \subset [1, q]$, $J \subset [1, m]$ with $|I| = |J| = r$ and the determinant of the matrix $A = (a_{ij})_{i \in I, j \in J}$ is not equal to zero. Since $\det(A) \leq \|\psi\|^r \leq \|\psi\|^m < p$, it follows that the determinant of A is not equal to zero modulo p . Hence we have $\det(\overline{a_{ij}})_{i \in I, j \in J}$ is not equal to zero and $\dim(\text{Im}(\Psi_p)) \geq \dim(\text{Im}(\Psi))$. Similar arguments show that $\dim(\text{Im}(\Psi_p)) \leq \dim(\text{Im}(\Psi))$. Hence the result follows. \square

Definition 6.2.4. (*Translation invariant matrix*) Let $V = (a_{ij}) \in M_{r \times m}(\mathbf{Z})$. We

The following lemma is due to Green and Tao [12, Lemma 4.4].

Lemma 6.2.3. *(Normal extension) Let $\Psi : R^q \rightarrow R^m$ be an affine linear map of finite complexity s . Then there exists an extension $\Psi' : R^{q'} \rightarrow R^m$ of Ψ which is in exact s -normal form.*

Let $\Psi = (\Psi^1, \dots, \Psi^m) : \mathbf{Z}^q \rightarrow \mathbf{Z}^m$ be an affine linear map. Then $\Psi^i(x_1, \dots, x_q) = \sum_{j=1}^q a_{ij}x_j + b_i$ for some $a_{ij}, b_i \in \mathbf{Z}$ and the norm of Ψ is defined by

$$\|\Psi\| := \sum_{i,j} |a_{ij}| + \sum_{i=1}^m |b_i|. \quad (6.8)$$

For any prime p and $x \in \mathbf{Z}/p\mathbf{Z}$, let $|x|_p = |y|$, where $y \in (-\frac{p}{2}, \frac{p}{2}]$ is the unique integer such that its image under the natural projection map from \mathbf{Z} to $\mathbf{Z}/p\mathbf{Z}$ is equal to x . Given a linear map $\Psi_p = (\Psi^1, \dots, \Psi^m) : (\mathbf{Z}/p\mathbf{Z})^q \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ with $\Psi^i(x_1, \dots, x_q) = \sum_{j=1}^q a_{ij}x_j$ for some $a_{ij} \in \mathbf{Z}/p\mathbf{Z}$, the norm of Ψ_p is defined as

$$\|\Psi_p\|_p := \sum_{i,j} |a_{ij}|_p. \quad (6.9)$$

Given a linear map $\psi : \mathbf{Z}^q \rightarrow \mathbf{Z}^m$, it induces a linear map $\Psi : \mathbf{Q}^q \rightarrow \mathbf{Q}^m$ as well as a linear map $\Psi_p : (\mathbf{Z}/p\mathbf{Z})^q \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$. Note that the image of ψ is a free \mathbf{Z} -module with the rank equal to $\dim(\text{Im}(\Psi))$. In fact we have the following lemma.

Lemma 6.2.4. *Let $\psi : \mathbf{Z}^q \rightarrow \mathbf{Z}^m$ be a linear map. Let p be a prime such that $p > \|\psi\|^m$. Then the linear maps $\Psi : \mathbf{Q}^q \rightarrow \mathbf{Q}^m$ and $\Psi_p : (\mathbf{Z}/p\mathbf{Z})^q \rightarrow (\mathbf{Z}/p\mathbf{Z})^m$ are*

with $|\tau_i| \leq s+1$ and $\tau_i \not\subseteq \sigma_j$ for any $j \neq i$. We say that Ψ is in exact s -normal form at i if there exists such τ_i with $|\tau_i| = s+1$. We say that Ψ is in (exact) s -normal form if it is in (exact) s -normal form at all i .

The following lemma says that every affine linear map which is in s -normal form has a finite complexity. More precisely we have

Lemma 6.2.2. *Let $\Psi = (\Psi^1, \dots, \Psi^m) : k^q \rightarrow k^m$ be an affine linear map which is in s -normal form, then the complexity of Ψ is at most s .*

Proof. Let $i \in \{1, \dots, m\}$ and let $\tau_i = \{i_1, \dots, i_r\}$ be a subset of the support σ_i of Ψ^i with $\tau_i \not\subseteq \sigma_j$ for any $j \neq i$ and $r \leq s+1$. Let us take the partition $\{\Psi^1, \dots, \Psi^m\} \setminus \{\Psi^i\} = X_1 \sqcup \dots \sqcup X_r$, where $X_h = \{\Psi^j : j \neq i, i_h \notin \sigma_j\}$ for each $h = 1, \dots, r$. Since i_h belongs to support of Ψ^i , it follows that Ψ^i does not belong to linear span $\langle X_h \rangle$ of X_h for each $h = 1, \dots, r$. Thus, complexity of Ψ at i is at most s . Since we started with an arbitrary index i , so the complexity of Ψ is itself at most s .

□

The converse of the above lemma is false. However, Green and Tao [12, Lemma 4.4] proved that every affine linear map of complexity s can be “extended” to one that is in s -normal form. Let $\Psi : R^q \rightarrow R^m$ be a R -affine linear map. We can also think of this linear map as a k -affine linear map from k^q to k^m .

Definition 6.2.3. *(Extension of an affine linear map) An extension of Ψ is an affine linear map $\Psi' : R^{q'} \rightarrow R^m$ with $q' \geq q$, $\text{Im}(\Psi) = \text{Im}(\Psi')$ and furthermore if we identify R^q with the subset $R^q \times \{0\}^{q'-q}$ of $R^{q'}$ in a natural way, then Ψ is the restriction of Ψ' .*

6.2 Preliminaries

Let R be an integral domain and k be its field of fractions. We are interested in the case when R is \mathbf{Z} or $\mathbf{Z}/p\mathbf{Z}$, where p is a prime number.

Definition 6.2.1. (*Complexity of a affine linear map*) Let $\Psi = (\Psi^1, \dots, \Psi^m) : k^q \rightarrow k^m$ be a affine linear map. The complexity of Ψ at i , $1 \leq i \leq m$, is the smallest integer $s \geq 0$ for which there is a partition $\{\Psi^1, \dots, \Psi^m\} \setminus \{\Psi^i\} = X_1 \sqcup \dots \sqcup X_{s+1}$ into non-empty sets such that Ψ^i does not belong to the linear span $\langle X_j \rangle$ of X_j for any $1 \leq j \leq s+1$, when such an integer exists. Otherwise we set the complexity of Ψ at i to be ∞ . The complexity of Ψ is the maximum of complexities of ψ at i over all $i = 1, \dots, m$.

Using the following observation, due to Green and Tao [12, Lemma 1.6], one can easily decide, given an affine linear map whether it has a finite complexity or not.

Lemma 6.2.1. *An affine linear map $\Psi = (\Psi^1, \dots, \Psi^m) : k^q \rightarrow k^m$ has a finite complexity if and only if for any $i \neq j$ we have $\Psi^i \neq \alpha\Psi^j$ for all $\alpha \in k$.*

We will recall the notion of normal form. For this, let us first set some notation here. Given a affine linear form Ψ in q variables x_1, \dots, x_q , the support of Ψ is the set of indices j such that Ψ depends on x_j . That is, if $\Psi(x_1, \dots, x_q) = \lambda_1 x_1 + \dots + \lambda_q x_q + \lambda$ then the support of Ψ is the set $\{j : \lambda_j \neq 0\}$.

Definition 6.2.2. (*Normal form*) Let $\Psi = (\Psi^1, \dots, \Psi^m) : k^q \rightarrow k^m$ be an affine linear map. Then each of Ψ^i is an affine linear form in q variables and let the support of Ψ^i be σ_i . We say that Ψ is in s -normal form at i if there exists $\tau_i \subseteq \sigma_i$

and $|S_F(N, N^{1/9})| \asymp N/(\log N)^2$. Then the conclusion of the corollary follows from Theorem 6.1.2.

□

Now we sketch the proof of Theorem 6.1.2. First we reduce to the case “ W -tricked” set which we also call A as shown in the subsection 6.5.1. We consider this set as a subset of \mathbf{Z}_M for some prime M . We then appropriately normalize the indicator function of this set and we denote it by λ_A . We decompose $\lambda_A = F_1 + F_2$ such that the $L^r(\mathbf{Z}_M)$ -norm $\|F_1\|_{L^r(\mathbf{Z}_M)}$ of F_1 is “small” and the Gowers U^2 -norm $\|F_2\|_{U^2}$ of F_2 is “small” as shown in Proposition 6.5.6. Indeed F_1 will be smoothed version of λ_A (convolving λ_A with a indicator function of a Bohr set). Since $\|F_1\|_{L^r(\mathbf{Z}_M)}$ is small F_1 “behaves” like a dense subset of integers as shown in Proposition 6.5.7. Thus, the count of complexity one patterns with weight function F_1 is large by Theorem 6.1.3 as shown in Proposition 6.5.8. On the other-hand λ_A and F_2 will be dominated by “pseudorandom” measure as it is shown in the Section 6.5.3. Therefore we conclude that the count of complexity one patterns with weight function λ_A is “large” as shown in the Section 6.5.5. Thus we conclude the proof of the W -tricked version theorem.

In the following section we give the definition of complexity of a matrix and its properties, mostly taken from [15], and introduce the notion of Bohr and regular Bohr sets and their properties.

where c is a constant depending only upon V .

The proof of this theorem depends on both the inverse theorem for the local Gowers U^2 -norm, which we discuss in Section 6.3, and density increment argument. We give details in Section 6.4.

In 2006, Green and Tao found infinitely many non-trivial three term arithmetic progressions in Chen primes. A prime number p is said to be a Chen prime if $p + 2$ has at most two prime factors, each of which is at least $p^{3/11}$. We denote the set of such prime numbers by \mathcal{Q} . Then we have the following corollary of Theorem 6.1.2, which improves upon and generalizes the result obtained by Green and Tao [10, Theorem 1.2].

Corollary 6.1.4. *Let V be a translation invariant matrix of order $r \times m$ with entries in \mathbf{Z} , of rank r and complexity one. Given any $\epsilon > 0$, there exists a positive constant $C > 0$ depending at most on r, m, V and ϵ such that for any $\mathcal{B} \subseteq \mathcal{Q} \cap [1, N]$ there exists a non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{B}^m$ if*

$$|\mathcal{B}| \geq C (\log \log N)^{-\frac{1}{8m} + \epsilon} |\mathcal{Q}(N)|. \quad (6.6)$$

Proof. Let $F(X) = X(X + 2)$ and let $\mathcal{A} = \mathcal{B} \cap (N^{1/2}, N]$, then we have $\mathcal{A} \subset S_F(N, N^{1/9})$ with

$$|\mathcal{A}| \geq |\mathcal{B}| - N^{1/2} \gg (\log \log N)^{-\frac{1}{8m} + \epsilon} |\mathcal{Q}(N)| \geq c (\log \log N)^{-\frac{1}{8m} + \epsilon} |S_F(N, N^{1/9})|, \quad (6.7)$$

here we use the facts that $|\mathcal{Q}(N)| \gg N/(\log N)^2$, see for example [10, Theorem 6.1],

let

$$S_F(N, z) = \{n \leq N : \gcd(F(n), P(z)) = 1\}. \quad (6.2)$$

Then by a standard application of the Brun's sieve we have the bounds

$$|(S_F(N, N^{1/(4k+1)}))| \asymp_F \frac{N}{\log^k N}, \quad (6.3)$$

provided that F is admissible of degree k and has non-zero discriminant.

Theorem 6.1.2. *Let V be a translation invariant matrix of order $r \times m$ with entries in \mathbf{Z} , of rank r and complexity one and $m \geq 3$. Further, let $F(X) = \prod_{i=1}^k (a_i X + b_i)$, with $a_i, b_i \in \mathbf{Z}$, be an admissible polynomial of degree k with non-zero discriminant. Then given any $\epsilon > 0$ there exists a positive constant $C > 0$ and an $N_0 \geq 1$ depending at most on V, F and ϵ such that for all $N \geq N_0$ and $\mathcal{A} \subset S_F(N, N^{1/(4k+1)})$ there exists a non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{A}^m$ if*

$$|\mathcal{A}| \geq C (\log \log N)^{-\frac{1}{8m} + \epsilon} |S_F(N, N^{1/(4k+1)})|. \quad (6.4)$$

This theorem can be proved using the following theorem combined with the transference principle of Green and Tao, in fact we use a version of transference principle due to Helfgott-De Roton [14].

Theorem 6.1.3. *Let V be a translation invariant matrix of order $r \times m$ with entries in \mathbf{Z} , of rank r and complexity one. Let $\mathcal{A} \subseteq [-N, N]$ be a subset of density α . Then*

$$|\{\mathbf{x} \in \mathcal{A}^m : V\mathbf{x}^t = 0\}| \geq \exp(-c \alpha^{-8m} \log \frac{1}{\alpha}) N^{m-r}, \quad (6.5)$$

the kernel of the associated linear map $x \mapsto Vx$ from \mathbf{Q}^m to \mathbf{Q}^r has dimension at least 2. Then it can in turn be deduced from Szemerédi’s theorem that for any subset A of \mathbf{N} of positive upper density there is a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$. One may naturally ask for quantitative versions of this result, but these are rather difficult to obtain in general and are tied up with an appropriate notion of “complexity” of V . For arbitrary V of complexity one, Kevin Henriot [15] has obtained the following theorem.

Theorem 6.1.1. *Let V be a translation invariant matrix of order $r \times m$ with entries in \mathbf{Z} , of rank r and complexity one and $m \geq 3$. Then there exists a positive constant $C > 0$ depending at most on r, m and V such that for any $\mathcal{A} \subseteq \mathcal{P} \cap [1, N]$ there exists non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{A}^m$ if*

$$|\mathcal{A}| \geq C(\log \log N)^{-\frac{1}{25m}} \pi(N). \tag{6.1}$$

Henriot’s arguments for the above theorem were inspired by those in Shao [33].

We slightly modify the arguments of Henriot and Shao to get an improvement in the exponent of $\log \log N$ in (6.1). Indeed, we improve “inverse theorem for the local Gowers U^2 -norm” which leads us to the improvement in the exponent. Also, we give a generalized version of Henriot’s theorem. More precisely, our main result is the theorem below, which we state with the aid of the following notation

Let $F(X) = \prod_{i=1}^k (a_i X + b_i) \in \mathbf{Z}[X]$ be a polynomial in one variable with coefficients in \mathbf{Z} . We say $F(X)$ is admissible if for each prime p there exists $n \in \mathbb{Z}$ such that p does not divide $F(n)$. Further, for any real number $z > 0$, let $P(z) = \prod_{p \leq z} p$ and

$j \notin \{i, i+1, i+2\}$, $a_{i,j} = 1$ if $j \in \{i, i+2\}$ and $a_{i,i+1} = -2$. Then V is translation invariant and for any x in \mathbf{Z}^m with $x = (x_1, x_2, \dots, x_m)$, the relation $Vx^t = 0$ is equivalent to the assertion that x_1, x_2, \dots, x_m are in arithmetical progression. Thus Szemerédi's theorem [35] is equivalent to the statement that for any $A \subseteq \mathbf{N}$ of positive upper density, there is a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$. Similarly, the celebrated result of Green and Tao [11] is equivalent to the statement that for any subset \mathcal{A} of the set of primes \mathcal{P} with positive relative upper density in \mathcal{P} , there is a non-trivial solution x to $Vx^t = 0$ with $x \in \mathcal{A}^m$.

The above example of a translation invariant matrix V has “complexity” $m - 2$. Indeed we can parametrize the kernel of V by the system of linear forms

$$\begin{aligned} \Psi : \mathbf{Z}^2 &\rightarrow \mathbf{Z}^m \\ (a, d) &\mapsto (a, a + d, \dots, a + (m - 1)d) \end{aligned}$$

of complexity $m - 2$. Thus, for example finding a 3-term arithmetic progression is of complexity one problem. Another example of complexity one system is given by

$$\begin{aligned} \Psi : \mathbf{Z}^{d+1} &\rightarrow \mathbf{Z}^{d(d+1)/2} \\ (x_0, x_1, \dots, x_d) &\mapsto (x_0 + x_i + x_j)_{1 \leq i < j \leq d}. \end{aligned}$$

In fact this is a model for the complexity one systems. X. Shao [33] worked with this system.

Let V of order $r \times m$ with entries in \mathbf{Z} be a translation invariant matrix such that

CHAPTER 6

Linear patterns of complexity one in Chen primes

6.1 Introduction

A matrix V of order $r \times m$ with entries in \mathbf{Z} is said to be translation invariant if $Ve^t = 0$, where e is the vector $(1, 1, \dots, 1)$ in \mathbf{Q}^m and e^t its transpose. For such V by a trivial solution of the system of linear equations $Vx^t = 0$ we mean $x = \lambda e$, for some $\lambda \in \mathbf{Q}$.

A number of celebrated results in additive combinatorics can be recast as a statements asserting the existence of a non-trivial solution x to $Vx^t = 0$ with $x \in A^m$, where A is a given subset of \mathbf{N} . For instance, let $m \geq 3$ be an integer and let $V = (a_{i,j})$ be the matrix of order $(m - 2) \times m$ be defined by $a_{i,j} = 0$ if

- [35] E. Szemerédi, On sets of integer containing no K elements in arithmetic progression, *Acta arithmetica*, 1975.
- [36] T. Tao and V. Vu, Additive combinatorics, *Cambridge University Press*, 2006.
- [37] T. Tao, <https://terrytao.wordpress.com/tag/large-sieve/>.
- [38] G. Tenenbaum, Introduction to Analytic and Probabilistic Number Theory, *Cambridge Studies in Advanced Mathematics*, **46**, Cambridge University Press (1995).
- [39] J. H. Van Lint and H. E. Richert, On Primes in Arithmetical Progressions, *Acta Arithmetica*, Vol. 11, pp. 209-216, 1965.

- [26] O. Ramaré and I. Z. Ruzsa, Additive Properties of Dense Subsets of Sifted Sequences, *Journal de théorie des nombres de Bordeaux*, Vol. 13 (2), pp 559-581, 2001.
- [27] O. Ramaré, Arithmetical aspects of the large sieve inequality with the collaboration of D. S. Ramana. *Harish-Chandra Research Institute Lecture Notes, 1.*, Hindustan Book Agency, New Delhi, 2009.
- [28] B. Rosser and I. Schoenfeld, Approximate Formulas for Some Functions of Prime Numbers, *Illinois Journal of Mathematics*, Vol. 6 (1), pp 64–94, 1962.
- [29] K. F. Roth, On certain sets of integers, *J. London Math. Soc.*, Vol. 28, pp. 104-109, 1953.
- [30] A. Sárközy, Finite Addition Theorem I, *Journal of Number Theory*, Vol. 48, pp. 197-218, 1994.
- [31] A. Sárközy, Unsolved Problems in Number Theory, *Period. Math. Hungar.*, Vol. 42, pp. 17-35, 2001.
- [32] J. P. Serre, Topics in Galois Theory, *Research Notes in Mathematics*, Vol. 1, Jones and Bartlett Publishers, 1992.
- [33] X. Shao, Finding linear patterns of complexity one, *International Mathematics Research Notices*, 2013.
- [34] X. Shao, On inverse ternary Goldbach problem, *American Journal of Mathematics*, Vol. 138 (5), pp. 1167-1191, 2016.

- [17] H. Iwaniec and I. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications 53, A.M.S., 2004.
- [18] N. I. Klimov, Combination of elementary and analytic methods in the theory of numbers, *Uspehi Mat. Nauk (N.S)*, Vol. 4, pp. 145–164, 1958.
- [19] K. Mallesham, Primes in sumsets, *Archiv der Mathematik*, Vol. 110, pp. 131–143, 2018.
- [20] K. Matomaki, Sums of positive density subsets of the primes, *Acta Arith.*, Vol. 159, pp. 201-225, 2013.
- [21] H. L. Montgomery and R. C. Vaughan, The large sieve, *Mathematika*, Vol. 20, pp. 119-134, 1973.
- [22] E. Naslund, On improving Roth’s theorem in the primes, *Mathematika*, Vol. 61 (1), pp. 49–62, 2015.
- [23] H. H. Ostmann, Additive Zahlentheorie. 1.Tel. *Allgemeine Untersuchungen*, springer verlag, Berlin-Heidelberg-Newyork , 1968.
- [24] G. Prakash, D. S. Ramana and O. Ramaré, On monochromatic sums of squares, *Math. Z*, 2017.
- [25] D. S. Ramana and O. Ramaré, Additive Energy of Dense Sets of Primes and Monochromatic Sums, *Israel Journal of Math.*, Vol. 199 (2), pp. 955–974, 2014.

- [7] C. Elsholtz, The inverse Goldbach problem, *Mathematika*, Vol. 48 (1-2), pp. 151-158, 2003.
- [8] A. Goldston, Pintz and Yildirim, Higher correlations of divisor sums related to primes, *Proc. London math.*, 2007.
- [9] A. Goldston, Pintz and Yildirim, Primes in tuples, *Annals of Mathematics*, 2009.
- [10] B. Green and T. Tao, Restriction theory of the Selberg sieve with applications, *journal de theorie des nombres de Bordeaux*, Vol. 18 (1), pp. 147–182, 2006.
- [11] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Annals of mathematics*, Vol. 167 (2), pp. 481–547, 2008.
- [12] B. Green and T. Tao, Linear equations in primes, *Annals of mathematics*, Vol. 18 (1), pp. 147–182, 2008.
- [13] N. Hegyvári and F. Hennecart, On monochromatic sums of squares and primes, *Journal of Number Theory*, Vol. 124, 2007, pp. 314-324.
- [14] H. A. Helfgott and A. de Roton, Improving Roths theorem in the primes, *Int. Math. Res. Not.*, IMRN, Vol. 4, 2011, pp. 767-783.
- [15] K. Henriot, On systems of complexity one in the primes, *Proc. Edinb. Math. Soc.*, 2016.
- [16] L. K. Hua, Additive Theory of Prime Numbers, *Tranal. Math. Monogr.*, Vol. 13, American Mathematical Society, Providence,RI, 1965.

Bibliography

- [1] P. Akhilesh and D. S. Ramana, A chromatic version of Lagrange's four squares theorem, *Monatsh. Math.*, Vol. 176 (1), pp. 17–29, 2015.
- [2] R. Ayoub, On Rademacher's extension of the Goldbach-Vinogradov theorem, *Trans. Amer. Math. Soc.*, Vol. 74, pp. 482-491, 1953.
- [3] A. Balog, J. Rivat and A. Sárközy, On arithmetic properties of sumsets, *Acta Math. Hungar.*, Vol. 144 (1), pp. 18–42, 2014.
- [4] G. Chen, On monochromatic sums of squares of primes, *Journal of Number Theory*, Vol. 162, pp. 180-189, 2016.
- [5] K. Chipeniuk and M. Hamel, On sums of sets of primes with positive relative density. *Jour. Lond. Math. Soc. (2)*, Vol. 83 (3), pp. 673-690, 2011.
- [6] S. Chow, Roth-Waring-Goldbach, *International Mathematics Research Notices*, 2017.