Equations in the Primes

by

Brian Michael Cook

B.Sc., Southern Polytechnic University, 2005
M.Sc., Georgia State University, 2007

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

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2012

Abstract

We provide results related to the study of prime points on level sets of homogeneous integral forms which are linear or quadratic. In the linear case we present an extension of the Green-Tao Theorem, which finds affine copies of finite intervals in relatively dense subsets of the primes, to a higher dimensional setting in which one finds affine copies of suitably generic point configurations in relatively dense subsets of a Cartesian product of the primes.

For general integral quadratic forms we present a result which is a Birch-Goldbach type theorem for a single quadratic form with sufficient rank. This guarantees solubility among the primes on the level set of a quadratic form subject to local conditions. This is an extension of a well known result of Hua.

Preface

The material from Chapter 2 is taken from the following manuscript, which is comprised of joint work A. Magyar.

> Constellations in $\mathbb{P}^d$ (with A. Magyar), Int Math Res Notices (2011) doi: 10.1093/imrn/rnr127

Table of Contents

Abstract ii

Preface iii

Table of Contents iv

Acknowledgements vi

Dedication vii

1 Introduction 1

1.1 Introduction 1

1.1.1 Linear Equations in the Primes 2

1.1.2 Diagonal Forms in the Primes 6

1.2 Overview 8

1.2.1 A Multidimensional Green-Tao Theorem 8

1.2.2 General Quadratic Forms in the Primes 11

2 A Higher Dimensional Green-Tao Theorem 14

2.1 Introduction 14

2.2 Norms, Transference, and Pseudo-random Measures 14

2.3 The Generalized von Neumann Inequality. 18

2.4 The Dual Function Estimate. 22

2.5 Proof of the Main Results. 25

2.5.1 Proof of Theorem 1.2.2. 25

2.5.2 Proof of Theorem 1.2.1. 27

2.6 The Correlation Condition. 30

3 Quadratic Equations 36
3.1 Introduction. 36
3.2 The Minor Arcs 36
3.2.1 Sufficiently Off Diagonal Forms 36
3.2.2 Insufficiently Off Diagonal Forms 39
3.3 The Major Arcs and an Asymptotic Formula 43
3.4 The Singular Series 47
3.5 Conclusions 51

4 Conclusion 53
4.1 Future Projects 53
4.1.1 A Conjecture 53
4.1.2 A Reasonable Approach 55
4.2 Final Remarks 55

Bibliography 56

Acknowledgements

First and foremost, I wish to acknowledge my advisor Prof. A. Magyar. I can think of no other person with whom I would have rather worked with. Other names of note are Prof. G. Chen, Prof. N. Lyall, Prof. J. Ziegler, and Prof. J. Whitenton. The faculty and staff at the following universities also require recognition: University of British Columbia; University of Georgia; Georgia State University; and Southern Polytechnic State University.

Dedication

To Tigger.

Chapter 1

Introduction

1.1 Introduction

The main purpose of this work is to provide some background on the study of solving equations in the primes, as well as to contribute a few results in this area. The main types of equations we are interested in are homogenous polynomials with integer coefficients. We are especially interested in affine varieties defined by the level sets of such equations. This problem is of particular importance due to the large number of related problems in additive number theory and additive combinatorics.

It is not a matter of necessity, but one of convenience that we do not measure directly the density of prime points on surfaces, but a weighted version. All weights are given essentially by the von Mangoldt function, denoted by Λ, which takes the the value $\log(p)$ for a power of a prime p, and zero elsewhere. In every result appearing in this paper one may translate directly to the primes by dividing through by the appropriate power of a logarithm, the appropriate power being the number of variables. This is standard. We note that P is used to denote the set of primes.

We now proceed to overview the main results below, after a brief interlude to overview some notations as well as some of the known results that have previously been obtained in this area.

As usual, we use $\mathbb{Z}$ for the integers, $\mathbb{R}$ for the reals, and $\mathbb{C}$ for set of complex numbers, We use the shorthand notation $\mathbb{Z}_m$ to denote the group of residue classes $\mathbb{Z}/m\mathbb{Z}$. We use $||f||_{L^p(X)}$ to denote the standard L^p norms of a function f on a given measure space X, and unless the situation necessitates we shall omit X and simple write $||f||_{L^p}$. We employ the notation $\mathbb{E}_{x\in X} = |X|^{-1}\sum_{x\in X}$.The Bachmann-Landau notation O and o notation is used frequently. The notation $f \lesssim g$ is also used as an alternative to $f = O(g)$. We use $f \approx g$ to mean $f \lesssim g$ and $g \lesssim f$. Further notational conventions are introduced as they appear.

1.1.1 Linear Equations in the Primes

The study of linear systems of equations in prime unknowns has a long history. However, the recent work of Green and Tao encompasses the majority of what is known in the subject, and this is where we shall focus. We need some preliminary ideas before the main conjecture and the results that partially resolve it may be stated. These definitions are taken directly from [12].

Definition 1.1.1. (Affine-linear forms) Let R, n be integers. An affine-linear form on $\mathbb{Z}^R$ is a function $\psi : \mathbb{Z}^R \to \mathbb{Z}$ which is the sum $\psi = \psi' + \psi(0)$ of a linear form ψ' on $\mathbb{Z}^R$ and $\psi(0)$ is an integer. A system of affine-linear forms is then a collection $\Psi = (\psi_1, ..., \psi_n)$ where each ψ_i is an affine-linear form. The image of the $\mathbb{Z}^R$ under Ψ is referred to as an affine-sublattice of $\mathbb{Z}^n$. The size of such a system relative to the scale N is given by

$$||\Psi||_N = \sum_{i=1}^{n} \sum_{j=1}^{R} |\psi_i'(e_j)| + \sum_{i=1}^{n} |\psi_i(0)N^{-1}|, \tag{1.1}$$

where the e_j are the standard basis elements for $\mathbb{Z}^R$. To avoid trivialities, it is assumed that in a given system of affine-linear forms we have no constant forms and no two affine-linear forms are rational multiples of each other.

With a given system of affine-linear forms, $\Psi = (\psi_1, ..., \psi_n)$, the main problem is to evaluate the sum

$$\sum_{x \in K \cap [-N,N]^R} \prod_{i=1}^{n} \Lambda(\psi_i(x)) \tag{1.2}$$

for a given convex body K. This sum counts prime points represented (with multiplicity) by the system of affine-linear forms where each prime point, more precisely each prime power point, is weighted with the von Mangoldt function. A heuristic argument, which dates back to Hardy and Littlewood at least, provides an expected value for the sum as a singular series, which is actually a product of terms taking into account solubility at each prime place. The Archimedean factor is

$$\beta_\infty = vol(K \cap [-N,N]^R \cap \Psi^{-1}(\mathbb{R}^+)^n). \tag{1.3}$$

Here *vol* denotes the volume and Ψ is extended to $\mathbb{R}$ in the natural way. The order of this term is N^R in general.

The local factors are defined in terms of localized versions of the von Mangoldt function, which for each prime p is given by $\Lambda_p(y) = p/(p-1)$ for each integer y not divisible by p, whence we simply get zero. The local factors are then

$$\beta_p = \mathbb{E}_{x\in\mathbb{Z}_p^R} \prod_{i=1}^{n} \Lambda_p(\psi_i(x)). \tag{1.4}$$

The heuristic argument lends the following conjecture.

Conjecture 1.1.1. *(Generalized Hardy-Littlewood conjecture). Let $N, R, n,$ and L be positive integers. Also, let $\Psi = (\psi_1, ..., \psi_n)$ be a system of affine-linear forms with size $||\Psi||_N \leq L$, and $K \subset [-N,N]^R$ be a convex body. Then we have*

$$\sum_{x\in K\cap[-N,N]^R} \prod_{i=1}^{n} \Lambda(\psi_i(x)) = \beta_\infty \prod_p \beta_p + o_{t,d,L}(N^R). \tag{1.5}$$

The original formulation of Hardy and Littlewood only deals with the systems composed of the affine-linear forms $\psi_i(x) = x + b_i$ on $\mathbb{Z}$. This generalization originally appears in [13]. In this same paper they prove a significant portion of this conjecture, albeit conditionally. The assumptions required in their proof are known as the Möbius and nilsequences conjecture and the Gowers-norm conjecture. Subsequently these results have been shown. The former by Green and Tao in [1], and the latter Green, Tao, and Ziegler in [2]. To state the main result of Green and Tao we need one more definition.

Definition 1.1.2. (Complexity). Let ($\Psi = \psi_1, ..., \psi_n$) be a system of affine-linear forms. If $1 \leq i \leq n$ and $s \geq 0$, we say that Ψ has i-complexity at most s if one can cover the $n-1$ forms $\{\psi_i : i = 1, ..., n\, ; \, i \neq j\}$ by $s+1$ classes such that ϕ_i does not lie in the affine-linear span of any of these classes. The complexity of the system is then defined as the minimal s such that the system has i-complexity at most s for each i. If no such s exists, then the complexity is infinite.

With this we have the following.

Theorem 1.1.1. *(Green and Tao) The generalized Hardy-Littlewood conjecture is true for all systems of finite complexity.*

From the definition of complexity it is easily seen that the only excluded cases are those systems which have two affine-linear forms which are affinely related, meaning that the homogeneous part of two forms are rational multiples of one another. So while this result provides numerous examples, a few of which we point out below, it does nothing for the problems like the Goldbach conjecture.

Thus far the phrasing of the problems and results are presented in the form of simultaneously representing primes by affine-linear forms. In the linear setting, it turns out that this is the same as finding prime points on level sets of a system of linear forms in many situations. Green and Tao apply a little algebra to Theorem 1.1.1 and obtain the following.

Theorem 1.1.2. *(Green and Tao: Linear equations in the primes). Let A be an $R \times n$ integral matrix with $R \leq n$. Assume that A has full rank R, and that the only element of the row-space of A over $\mathbb{Q}$ with two or fewer non-zero entries is the zero vector. Let $N > 1$ and $b \in \mathbb{Z}^R$ be a vector in $A\mathbb{Z}^n$. Then we have*

$$\sum_{x \in [N]^n,\, Ax = b} \prod_{i=1}^{n} \Lambda(x_i) = \mu_\infty \prod_p \mu_p + o(N^{n-R}). \tag{1.6}$$

The local densities μ_p are given by

$$\mu_p = \lim_{M \to \infty} \mathbb{E}_{x \in [-M,M]^n,\, Ax=b} \prod_{i=1}^{n} \Lambda_p(x_i), \tag{1.7}$$

and the global factor μ_∞ is given by

$$\mu_\infty = \#\{x \in \mathbb{Z}^n : x_i \in [0,N]^n,\ Ax = b\}. \tag{1.8}$$

The implied constant in the little o term depends on A, R, and n only.

Let us highlight a couple of examples.

Example 1. The weighted number of solutions to the equation $x_1+x_2+x_3 = N$ with each $x_i \leq N$ prime obeys the asymptotic

$$\mathfrak{S}(N)N^2 + o(N^2), \tag{1.9}$$

with

$$\mathfrak{S}(N) = \prod_{(p,\,N)\neq 1} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{(p,\,N)=1} \left(1 + \frac{1}{(p-1)^3}\right). \tag{1.10}$$

This is Vinogradov's Three Primes Theorem, see e.g. [16]. Relatively recently it has been shown, conditionally on the Generalized Riemann Hypothesis, that there exists at least one prime point solution for every odd integer $N \geq 7$, this is done in [7].

Example 2. The system $L_i(x) = x_i - 2x_{i+1} + x_{i+2} = 0$ for $i = 1, ..., k$ in $k+2$ variables counts k-term arithmetic progressions. The weighted number of solutions with $x_1 < ... < x_{k+2}$ is given by

$$\left(\mathfrak{S}(N) + o(1)\right) N^2 \tag{1.11}$$

where

$$\mathfrak{S}(N) = \frac{1}{2(k-1)} \prod_{p} \beta_p \tag{1.12}$$

with

$$\beta_p = \begin{cases} \frac{1}{p}\left(\frac{p}{p-1}\right)^{k-1} & \text{if } p \leq k \\ \left(1 - \frac{k-1}{p}\right)\left(\frac{p}{p-1}\right)^{k-1} & \text{if } p > k. \end{cases} \tag{1.13}$$

The ease in adding the conditions $x_i < x_{i+1}$ is a nice feature of the addition of a general convex body..

We return to this example shortly. More examples are given in [13], and are in general not overly difficult to arrive at.

1.1.2 Diagonal Forms in the Primes

The study of finding prime points on level sets of homogenous forms of higher degree has received much less attention than the linear case. The major exception is for forms which are diagonal, meaning all terms are of the form $a_i x_i^d$. Forms of this type obey some similar properties to linear ones, and it is this approach Hua takes in showing the results of this section.

The main result Hua obtains is an asymptotic similar in nature to the one presented in Theorem 1.1.2. For convenience we look only at $F(x) = x_1^d + ... + x_n^d$ where $d \geq 1$. This is the most important instance of a diagonal form, the problem of finding prime solutions to $F = v$ for various values of v is known as the Goldbach-Waring problem. This one, and more general diagonal forms, are covered by the work of Chapter 3, albeit not as effectively. Set

$$\mathcal{M}(v, N) = \int_0^1 \sum_{x_1,...,x_n=1}^{N} \Lambda(x_1)...\Lambda(x_n)e((F(x) - v)r)dr, \tag{1.14}$$

that is the weighted number of prime points on the level set $F = v$. Note that this is only relevant when $v \approx N^d$, so we shall just assume the equality and set $\mathcal{M}(N^d, N) = \mathcal{M}(N)$.

Modifying the methods Vinogradov employs, Hua arrives at an asymptotic with the following singular series.

$$
\begin{aligned}
W_{a,q} &= \sum_{s \in \mathbb{Z}_q^*} e(as^d/q) \\
B(v,q) &= \sum_{(a,q)=1} \phi(q)^{-n} W_{a,q}^n e(-va/q) \\
\mathfrak{S}(v) &= \sum_{q=1}^{\infty} B(v,q).
\end{aligned}
\tag{1.15}
$$

Here ϕ is the totient function of Euler, and (a,q) represents the greatest common divisor. Also, $\mathbb{Z}_q^*$ is the multiplicative of the group $\mathbb{Z}_q$.

Theorem 1.1.3. *We have*

$$
\mathcal{M}(N) = \frac{\Gamma^n(1+d^{-1})}{\Gamma(nd^{-1})} \mathfrak{S}(N) N^{n/d-1} + o(N^{s/d-1}) \tag{1.16}
$$

provided that

$$
n \geq \begin{cases} 2^d + 1 & \textit{if } d \leq 11 \\ 2d^2\left(2\log\, d + \log\,\log\, d + 2.5\right) & \textit{if } d > 11. \end{cases} \tag{1.17}
$$

Here Γ denotes the standard gamma function.

One of the main aspects of Hua's work is with the singular series. He provides an an infinite arithmetic progression $\mathcal{Z}$, dependent on d, such that $\mathfrak{S}(N)$ is bounded below on $\mathcal{Z}$. Some specific instances are as follows.

Corollary 1.1.1. *Every sufficiently large odd integer is the sum of three primes.*

Corollary 1.1.2. *Every sufficiently large integer congruent to 5 modulo 24 is the sum of five squares of primes.*

Corollary 1.1.3. *Every sufficiently odd large integer is the sum of nine cubes of primes.*

Corollary 1.1.4. *Every sufficiently large integer congruent to 17 modulo 240 is the sum of 17 fourth powers of primes.*

While these results are for a specific arithmetic progression, it is possible to obtain results on all large numbers for these particular forms. For example, as one can easily check, the 4-fold sum set of squares of the reduced residue class modulo 24 covers all residue classes modulo 24. With this one can obtain from Corollary 1.1.2.

Theorem 1.1.1. *Every sufficiently large integer is the sum of at most nine squares of primes.*

1.2 Overview

We now come to a description of the results presented in this document. These are split into two main parts. The first of which overviews a result that we provide which generalizes the result of Green and Tao that the primes contain arbitrarily long arithmetic progression. The subsequent section introduces an extension of the result of Hua for general quadratic forms.

1.2.1 A Multidimensional Green-Tao Theorem

One way of phrasing the celebrated theorem of Green and Tao [12] is the statement that subsets of positive relative upper density of the primes contain an affine copy of any finite set of the integers, and in particular contain arbitrary long arithmetic progressions. It is natural to ask if similar results hold in the multi-dimensional settings, especially in light of the multi-dimensional extensions of the closely related theorem of Szemerédi [19] on arithmetic progressions in dense subsets of the integers. Indeed such a result was obtained by Tao [20], showing that the Gaussian primes contain arbitrary constellations. In the same paper the problem of finding constellations in dense subsets of P^d was raised and briefly discussed.

The difficulty in this setting comes from two facts. First, the natural majorant of the d-tuples of primes is not pseudo-random with respect to the box norms, which replace the Gowers' uniformity norms in the multi-dimensional case. This may be circumvented by assuming the set e is in *general position* as described below, as is already suggested in [20]. However even under the this non-degeneracy assumption, the so-called *correlation conditions* in [12] do not seem to be sufficient, and a key observation of this note is to use more general correlation conditions to obtain the dual function estimates in the multi-dimensional case. Also, we need a more abstract form of the *transference principle* of Green and Tao [12]. The formulation we use is due to Gowers [10], however essentially equivalent results have been obtained by Tao and Ziegler [21], as well as by Reingold et al. [17].

Finally let us note that we expect the main result of this paper remains true for sets which are not in general position. For example in the simplest case, when $e = \{(x, y), (x + d, y), (x, y + d)\}$, it is easy to see that both subsets of the form $A = B \times C$ and random subsets $A \subseteq P^2$ of positive relative density, contain many affine copies of e. However to prove such a result, our approach needs to be modified in an essential way, as the box norms do not seem to control such constellations in the relative setting.

Let $e = \{e_1, \ldots, e_l\} \in (\mathbb{Z}^d)^l$ be a set of vectors; a constellation defined by e is then a set $e' = \{x, x + te_1, \ldots, x + te_l\}$ where $t \neq 0$ is a scalar, that is an affine image of the set $e \cup \{0\}$.

Definition 1.2.1. We say that a set of l vectors $e \in (\mathbb{Z}^d)^l$ is in general position, if $|\pi_i(e \cup \{0\})| = l + 1$ for each i, where π_i is the orthogonal projection to the i^{th} coordinate axis.

Let us also recall that a subset A of the d-tuples of primes P^d is of positive upper relative density if

$$\limsup_{N\to\infty} \frac{|A \cap [1, N]^d|}{\pi(N)^d} > 0$$

Our main result is then the following

Theorem 1.2.1. *Given any set $A \subseteq P^d$ of positive relative upper density, we have that A contains infinitely many constellations defined by a set of vectors $e \in (\mathbb{Z}^d)^l$ in general position.*

Remarks: We note that for $d = 1$ this translates back to the above described theorem of Green and Tao [12], as any finite subset of $\mathbb{Z}$ is in general position.

Also, one may assume that $l = d$ and the set $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}^d$ forms a basis in $\mathbb{R}^d$ besides being in general position, by passing to higher dimensions. Indeed, if $e \in (\mathbb{Z}^d)^l$ then let $\{f_1, \ldots, f_l\} \subseteq \mathbb{Z}^l$ be linearly independent vectors, and define a basis

$$e' = \{e'_1 = (e_1, f_1), \ldots, e'_l = (e_l, f_l), e'_{l+1}, \ldots, e'_{l+d}\} \subseteq \mathbb{Z}^{d+l}$$

by extending the linearly independent set of vectors $e'_i = (e_i, f_i),\ (1 \leq i \leq l)$. Here we have used to (e_i, f_i) to denote the concatenation of the vectors e_i and f_i. If e was in general position then it is easy to make the construction so that e' is also in general position, and if the set $A' := A \times P^l$ contains a constellation $x' + te'$, then A contains $x + te$. Thus from now on we will always assume that e is also a basis of $\mathbb{R}^d$.

Theorem 1.2.1 may be viewed as a relative version of the so-called Multidimensional Szemerédi Theorem [8], stating that any subset of $\mathbb{Z}^d$ of positive upper density contains infinitely many constellations defined by any finite set of vectors $e \subseteq \mathbb{Z}^d$. As is customary, we will work in the finitary settings, when the underlying space is the group $\mathbb{Z}_N^d = (\mathbb{Z}/N\mathbb{Z})^d$, N being a large prime. In this settings we need the following, more quantitative version:

Theorem 1.2.1. *(Furstenberg-Katznelson [8]). Let $\alpha > 0$, $d \in \mathbb{N}$ and let $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}_N^d$ be a fixed set of vectors. If $f : \mathbb{Z}_N^d \to [0,1]$ is a given function such that $\mathbb{E}(f(x) : x \in \mathbb{Z}_N^d) \geq \alpha$, then one has*

$$\mathbb{E}(f(x)f(x+te_1)\ldots f(x+te_d) :\ x \in \mathbb{Z}_N^d,\ t \in \mathbb{Z}_N) \geq c(\alpha, e) \qquad (1.18)$$

where $c(\alpha, e) > 0$ is a constant depending only on α and the set e.

Here we have used the expectation notation

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

In the relative setting, when $A \subseteq P^d$, the condition: $\mathbb{E}(f(x) : x \in \mathbb{Z}_N^d) \geq$

α (after identifying $[1, N]^d$ with $\mathbb{Z}_N^d$) does not hold for the indicator function $f = \mathbf{1}_A$, however it holds for $f = \mathbf{1}_A \Lambda^d$ where Λ^d is the d-fold tensor product of the von Mangoldt function Λ. The price one pays is that the function f is no longer bounded uniformly in N. Following the strategy of [12] we will show that the d-fold tensor product $\otimes^d \nu$ of the pseudo-random measure ν used in [12] is sufficiently random in our settings in order to apply the transference principle of [10]; we will refer to such measures ν as *d-pseudo-random* measures. We postpone the definition of d-pseudo-random measures until later, but state our main result in the finitary settings below:

Theorem 1.2.2. *Let $\alpha > 0$ be given, and d be fixed. There exists a constant $c(\alpha, e) > 0$ such that the following holds. If $0 \leq f \leq \mu$ is a given function on $\mathbb{Z}_N^d$ such that $\mu = \otimes^d \nu$ where ν is d-pseudo-random, and $\mathbb{E}(f(x) : x \in \mathbb{Z}_N^d) \geq \alpha$, then for any basis $e = \{e_1, ..., e_d\}$ in general position, we have that*

$$\mathbb{E}(f(x)f(x + te_1)...f(x + te_d) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) \geq c(\alpha, e) \tag{1.19}$$

1.2.2 General Quadratic Forms in the Primes

Here we introduce an analogue of Theorem 1.1.3 for general quadratic forms in n variables, i.e. forms which may not be diagonal. Let $Q(x) = \langle x, Ax \rangle$ be an integral quadratic form on the integers in n variables, so that A is a symmetric $n \times n$ matrix with integer coefficients. Also, let $S_v = \{x \in \mathbb{Z}^n : Q(x) = v\}$ be its level surface. With P_N being the set of primes which are at most N, we wish to study $|S_v \cap P_N^n|$, that is the number of solutions of the equation $Q(x_1, \ldots, x_n) = v$ among the prime numbers. Our work is building on that of Hua and the method follows a similar outline, while also taking into account the methods of Birch [3] and Davenport [6] which treated integer solutions of general forms.

The difficulty in treating general quadratic forms is that in the work of Hua, and in fact most subsequent works addressing the number of prime solutions of diophantine equations, has exploited the additive structure of diagonal equations. For general quadratic forms the additive structure is not available. To overcome this we have first considered forms $Q(x) = \langle x, Ax \rangle$, where the underlying matrix A has an off-diagonal block of sufficiently large rank. We start out, as usual, by writing the number of weighted prime solutions via the expression

$$\begin{aligned}\mathcal{M}(v,N) &= \sum_{x_1,...,x_n=1}^{N} \Lambda(x_1)...\Lambda(x_n)S_v(x_1,...,x_n) \\ &= \int_0^1 \sum_{x_1,...,x_n=1}^{N} \Lambda(x_1)...\Lambda(x_n)e((Q(x)-v)r)dr, \qquad (1.20)\end{aligned}$$

where Λ again denotes the von Mangoldt function.

Our approach is to apply the Hardy-Littlewood circle method to obtain an asymptotic for 1.20. We (eventually) define the major arcs, the same way as in [14]; the major arcs shall be the collection of r's such that $|r - a/q| \leq (\log N)^c/N^2$ for some reduced fraction a/q with denominator $q \leq (\log N)^c$, c being a constant depending only on the underlying dimension n. In the case of an off-diagonal block of large enough rank, we first eliminate the von Mangoldt function by two applications of the Cauchy-Schwarz inequality picking up a logarithmic type loss. However using the Birch-Davenport method we can get strong enough bounds on the minor arcs to compensate.

In the opposite case we treat the matrix A as a block diagonal matrix consisting of a small and two large blocks, exploiting that the remaining off-diagonal blocks have small rank. Here the minor arcs estimates are similar to those of Hua [14] and Vinogradov [22], using uniform estimates and rewriting L^2 bounds as solutions of systems of equations.

The treatment of the major arcs is fairly standard reducing the integral over them to a product of local factors and a singular integral by making acceptable errors. This process culminates in an asymptotic formula of the form

$$\mathcal{M}(v,N) = N^{n-2}\mathfrak{S}(v)J(\mu) + O((\log N)^{-\delta}N^{n-2}), \qquad (1.21)$$

where $\delta > 0$, and $\mathfrak{S}$ and J are the singular series and integral, respectively.

The asymptotic may be used to deduce several results. The following is the analogue of Theorem 1 in [3], which in this case is essentially the

Hasse-Minkowski Theorem, see e.g. [4]. In the statement, $\mathcal{B}$ is the unit cube $[0,1]^n$ in $\mathbb{R}^n$.

Theorem 1.2.3. *Let Q be a homogenous quadratic polynomial in n variables. If we have the rank of Q is at least 34, then*

$$\mathcal{M}(v,N) = N^{n-2}\mathfrak{S}(v)J(\mu) + O(N^{n-2}(\log N)^{-\delta}),$$

where $\delta > 0$, the O-term is uniform in v and N, and $\mu = N^{-2}v$. Here $\mathfrak{S}(v)$ is positive so long as Q has a non-singular point in the reduced residue class modulo every sufficiently large prime power; and $J(Q(\bar{x}))$ exceeds a positive lower bound if $\bar{x}$ runs through a closed subset of the interior of $\mathcal{B} - V^*$. Here $\mathcal{B} = [0,1]^n \subset \mathbb{R}^n$, and V^* is the null space of the matrix A over $\mathbb{R}^n$.

Another result of interest is the following.

Theorem 1.2.4. *If Q is a positive definite integral quadratic form with rank at least 34 in n variables, then there exists an arithmetic progression, $\mathcal{Z}$, such that, when restricted to P^n, Q represents all sufficiently large elements in $\mathcal{Z}$.*

Theorem 1.2.4 may be viewed as a generalization of Theorem 1.1.3. As we have previously seen, Hua's main result (for quadratics) is that every sufficiently large integer congruent to 5 modulo 24 can be written as a sum 5 squared prime numbers.

Chapter 2

2.1 Introduction

The aim of this chapter is to prove Theorem 1.2.1. The method of proof follows the lines of the proof given in [12]. In Sections 3-4 we prove two key propositions, the so-called *generalized von Neumann inequality* and the *dual function estimate*. The first roughly says that the number of constellations defined by a set e is controlled by the appropriate box norm. The second is the essential step in showing that the box norms are QAP norms.

In Section 5, we prove our main results assuming that the measure exhibited in [12] is also d-pseudo-random in the sense defined above. First we show Theorem 1.2.2, which follows then easily from the Transference Principle. Next, we prove Theorem 1.2.1 by a standard argument passing from $\mathbb{Z}_N$ to $\mathbb{Z}$.

Finally, in the last section, we provide the additions of the results given in [12], which in turn proves d-pseudo-randomness of the measure ν that is used by Green and Tao. This is done by slightly modifying their arguments of Sec.10 in [12] based on earlier work of Goldston and Yıldırım [9] [5].

2.2 Norms, Transference, and Pseudo-random Measures

First we introduce the d-dimensional box norms. We actually introduce one norm for each linearly independent set of vectors
$\{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$.

For a function $f : \mathbb{Z}_N^d \to \mathbb{C}$ this norm with respect to a basis e is given by

$$||f||_{\square(e)^d}^{2^d} = \mathbb{E}(\prod_{\omega \in \{0,1\}^d} f(x + \omega t e) : x \in \mathbb{Z}_N^d,\ t \in \mathbb{Z}_N^d)$$

with the notation $\omega te = \omega_1 t_1 e_1 + ... + \omega_d t_d e_d$.

That this is actually a norm is not immediate, but for the standard basis it can be shown by repeated applications of the Cauchy-Schwarz inequality, similarly as for the Gowers norms (see for example [11]). For a different basis, note that we have $||f||_{\square(e)^d} = ||f \circ T||_{\square^d}$ for an appropriate linear transformation T, where $||f||_{\square^d}$ is the norm with respect to the standard basis. The same way one shows [11] that the analogue of the so-called Gowers-Cauchy-Schwarz inequality holds

Proposition 2.2.1. *($\square^d(e)$-Cauchy-Schwarz inequality) Given 2^d functions, indexed by elements of $\{0,1\}^d$, we have*

$$\langle f_\omega : \omega \in \{0,1\}^d \rangle = \mathbb{E}(\prod_{\omega \in \{0,1\}^d} f_\omega(x + \omega te) : x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N^d)$$

$$\leq \prod_{\omega \in \{0,1\}^d} ||f_\omega||_{\square(e)^d}$$

Gowers presents an alternative approach to the Green-Tao Transference Theorem from a more functional analytic point of view, making use of the Hahn-Banach Theorem. The specific version he provides will be presented below after we recall some definitions. First we note that $|| \cdot ||^*$ is defined to be the dual norm of $|| \cdot ||$.

Definition 2.2.1. Let $|| \cdot ||$ be a norm on $\mathcal{H} = L^2(\mathbb{Z}_n)$ such that $||f||_{L^\infty} \leq ||f||^*$, and let $X \subseteq \mathcal{H}$ be bounded. Then $||\cdot||$ is a *quasi algebra predual (QAP) norm* with respect to X if there exists an operator $\mathcal{D} : \mathcal{H} \to \mathcal{H}$, a positive function c on $\mathbb{R}$ and an increasing positive function C on $\mathbb{R}$ satisfying:

(i) $\langle f, \mathcal{D}f \rangle \leq 1$ *for all* $f \in X$,

(ii) $\langle f, \mathcal{D}f \rangle \geq c(\epsilon)$ *for every* $f \in X$ *with* $||f|| \geq \epsilon$, and

(iii) $||\mathcal{D}f_1 ... \mathcal{D}f_K||^* \leq C(K)$ *for any* $f_1, ..., f_K \in X$.

This definition in enough to state the transference principle.

Theorem 2.2.1. *(Gowers [10]) Let μ and ω be non-negative functions on Y, Y finite, with $||\mu||_{L^1}, ||\omega||_{L^1} \leq 1$, and $\eta, \delta > 0$ be given parameters. Also* *let $||\cdot||$ be a QAP norm with respect to X, the set of all functions bounded above by* $\max\{\mu,\omega\}$ *in absolute value. There exists an $\epsilon > 0$ such that the following holds: If we have that $||\mu-\omega|| < \epsilon$, then for every function with $0 \le f \le \mu$ there exists a function g with $0 \le g \le \omega/(1-\delta)$ and $||f-g|| \le \eta$.*

Remarks: By a simple re-scaling of the norms the constants 1 in Definition 2.2.1 and Theorem B can be replaced by any other fixed constants. The actual form given by Gowers is more explicit, in fact giving a specific choice of ϵ. However, for our purposes, we only need such an ϵ that is independent of the size of Y. Also, for our purpose one may choose $\omega \equiv 1$ and $\delta = 1/2$.

The definition of a pseudo-random measure here is slightly stronger than that of Green and Tao, adapted to the higher dimensional settings. Let us begin with the one dimensional case. Following [12], we define a *measure* to be a non-negative function $\nu : \mathbb{Z}_N \to \mathbb{R}$ such that

$$\mathbb{E}(\nu(x) : x \in \mathbb{Z}_N) = 1 + o(1).$$

where the $o(1)$ notation means a quantity which tends to 0 as $N \to \infty$. A measure will be deemed pseudo-random if it satisfies two properties at a specific level. The first of these is known as the linear forms condition, as we will use only forms with integer coefficients we need a slightly simplified version.

Definition 2.2.2. (Green-Tao [12]) Let ν be a measure, and $m_0, t_0 \in \mathbb{N}$ be small parameters. Then ν satisfies the (m_0,t_0)-*linear forms condition* if the following holds. For $m \le m_0$ and $t \le t_0$ arbitrary, suppose that $\{L_{i,j}\}_{1\le i\le m, 1\le j\le t}$ are integers, and that b_i are arbitrary elements of $\mathbb{Z}_N$. Given m linear forms $\phi_i : \mathbb{Z}_N^t \to \mathbb{Z}_N$ with

$$\phi_i(x) = \sum_{j=1}^{t} L_{i,j} x_j + b_i,$$

$x = (x_1, ..., x_t)$ and $b = (b_1, ..., b_t)$, if we have that each ϕ_i is nonzero and that they are pairwise linearly independent, then

$$\mathbb{E}\left(\prod_{i=1}^{m}\nu(\phi_i(x)) : x \in \mathbb{Z}_N^t\right) = 1 + o(1), \tag{2.1}$$

where the $o(1)$ term is independent of the choice of the b_i's.

The next condition is referred to as the correlation condition.

Definition 2.2.3. Let ν be a measure. Then ν satisfies the (m_0, m_1) correlation condition if for every $1 \leq m \leq m_0$ there exists a function $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}_+$ such that for all $k \in \mathbb{N}$

$$\mathbb{E}(\tau^k(x) :\, x \in \mathbb{Z}_N) = O_{m,k}(1)$$

and also

$$\mathbb{E}\left(\prod_{i=1}^{m_1}\prod_{j=1}^{m_0}\nu(\phi_i(y) + h_{i,j}) : y \in \mathbb{Z}_N^r\right) \leq \prod_{i=1}^{m_0}\left(\sum_{1\leq j<j'\leq m_0}\tau(h_{i,j} - h_{i,j'})\right)$$

where the functions $\phi_i : \mathbb{Z}_N^r \to \mathbb{Z}_N$ are pairwise independent linear forms.

Remarks:

This is a stronger condition that what is used in [12], in fact they used the special case when $m_1 = 1$, and ϕ is the identity. We define below a d-pseudo-random measure to be a measure satisfying these conditions at specific levels.

Definition 2.2.4. We call a measure ν a *d-pseudo-random* if ν satisfies the $((d^2 + 2d)2^{d-1},\, 2d^2 + d)$-linear forms condition and the $(d, 2^d)$-correlation condition

We will deal with d-fold tensor product of measures, $\mu = \otimes_{i=1}^{d}\nu$ and call them d-measures. We will call such a d-measure μ to be pseudo-random if the corresponding measure ν is d-pseudo-random. Finally, note that for a

d-measure

$$\mathbb{E}(\mu(x) : x \in \mathbb{Z}_N^d) = \prod_{i=1}^{d} \mathbb{E}(\nu(x_i) : x_i \in \mathbb{Z}_N) = 1 + o(1).$$

2.3 The Generalized von Neumann Inequality.

Let $e = \{e_1, \ldots, e_d\} \subseteq \mathbb{Z}_N^d$ be a base of $\mathbb{Z}_N^d$ which is also in general position, which in this settings means that $|\pi_i(e \cup \{0\})| = d + 1$ for each i where $\pi_i : \mathbb{Z}_N^d \to \mathbb{Z}_N$ is the orthogonal projection to the i-th coordinate axis.

Proposition 2.3.1. *(Generalized von Neumann Inequality) Let $w = \otimes^d \nu$ be a pseudo-random d-measure. Given a function $0 \leq f \leq w$, we have that*

$$\Lambda f := \mathbb{E}\left(f(x)f(x+te_1)...f(x+te_d) : \, x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N\right) = O(||f||_{\square(e')^d}) \quad (2.2)$$

where $e' = \{e_d, e_d - e_1, ..., e_d - e_{d-1}\}$.

Proof. We shall apply the Cauchy-Schwarz inequality several times. Begin by writing

$$\Lambda f \equiv \Lambda = \mathbb{E}(f(x) \prod_{i=1}^{d} f(x + t_1 e_i) : x \in \mathbb{Z}_N^d, t_1 \in \mathbb{Z}_N).$$

Push through the summation on t_1 and split the f to write this as

$$\mathbb{E}(\sqrt{f(x)}\, \mathbb{E}(\sqrt{f(x)} \prod_{i=1}^{d} f(x + t_1 e_i) : t_1 \in \mathbb{Z}_N) : x \in \mathbb{Z}_N^d).$$

Applying Cauchy-Schwarz to get

$$\Lambda^2 \leq \mathbb{E}(w(x) \prod_{i=1}^{d} f(x + t_1 e_i) \prod_{j=1}^{d} f(x + t_1 e_j + t_2 e_j) : t_1, t_2 \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have made the substitution $t_2 \mapsto t_1 + t_2$ for the new variable. Note that there should be a $\mathbb{E}(w(x)) = 1 + o(1)$ multiplier, following from the fact that $f \leq w$ and from the linear forms condition, but for convenience we suppress it and will continue to do so (this is a big O result, so this is not of any consequence). We make one further substitution, $x \mapsto x - t_1 e_1$, yielding

$$\Lambda^2 \leq \mathbb{E}(w(x - t_1 e_1) \prod_{i=2}^{d} \prod_{\omega \in \{0,1\}} f(x + t_1 e_i^{(1)} + \omega t^{(1)} e_i)$$

$$\times \prod_{\omega' \in \{0,1\}} f(x + \omega' t^{(1)} e_1) : t_1, t_2 \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have introduced the notations $e_i^{(j)} = e_i - e_j$, and $t^{(i)} = \{t_{1+j}\}_{j=1}^{i}$. Note that the final product of this expression is independent of t_1.

We now repeat this procedure exactly, pushing through the t_1 sum and splitting the terms independent of t_1, followed by a change of variables. After l applications of the Cauchy-Schwarz inequality, we claim to have

$$\Lambda^{2^l} \leq \mathbb{E}(W_l(x, t_1, ..., t_{l+1}) \prod_{i=l+1}^{d} \prod_{\omega \in \{0,1\}^l} f(x + t_1 e_i^{(l)} + \omega t^{(l)} e_{i;l}))$$

$$\times \prod_{\omega' \in \{0,1\}^l} f(x + \omega' t^{(l)} e_{l;l-1}) : t_1, ..., t_{l+1} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d), \qquad (2.3)$$

for an appropriate weight function W_l which is a product of w's, evaluated on linear forms which are pairwise linearly independent.

The notations introduced here are $e_{i;l} = \{e_i, e_i^{(1)}, ..., e_i^{(l-1)}\}$ (note that $l > 1$), and $\omega t^{(l)} e_{i;l} = \omega_1 t_2 e_i + \omega_2 t_3 e_i^{(1)} + ... + \omega_l t_{l+1} e_i^{(l-1)}$.

To check this form, using induction, apply the Cauchy-Schwarz inequality one more time with the new variable $t_1 + t_{l+2}$ to get

$$\Lambda^{2^{l+1}} \leq \mathbb{E}(W_l(x, t_1, ..., t_{l+1}) W_l(x, t_1 + t_{l+2}, ..., t_{l+1})$$

$$\times \prod_{i=l+1}^{d} \prod_{\omega \in \{0,1\}^l} f(x + t_1 e_i^{(l)} + \omega t^{(l)} e_{i;l}) f(x + t_1 e_i^{(l)} + t_{l+2} e_i^{(l)} + \omega t^{(l)} e_{i;l})$$

$$\times \prod_{\omega' \in \{0,1\}^l} w(x + \omega' t^{(l)} e_{l;l-1}) : t_1, ..., t_{l+2} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

Write

$$W'_{l+1}(x,t_1,...,t_{l+2}) = W_l(x,t_1,...,t_{l+1})W_l(x,t_1+t_{l+2},...,t_{l+1})$$

$$\times \prod_{\omega' \in \{0,1\}^l} w(x+\omega' t^{(l)} e_{l;l-1}).$$

We now apply the substitution $x \mapsto x - t_1 e^{(l)}_{l+1}$, note that $e^{(l)}_i - e^{(l)}_{l+1} = e^{(l+1)}_i$, and set

$$W_{l+1}(x,t_1,...,t_{l+2}) = W'_{l+1}(x - t_1 e^{(l)}_{l+1}, t_1,...,t_{l+2}), \tag{2.4}$$

This gives

$$\Lambda^{2^{l+1}} \le \mathbb{E}(W_{l+1}(x,t_1,...,t_{l+2}) \times \prod_{i=l+2}^{d} \prod_{\omega \in \{0,1\}^{l+1}} f(x+t_1 e^{(l+1)}_i + \omega t^{(l+1)} e_{i;l+1})$$

$$\times \prod_{\omega' \in \{0,1\}^{l+1}} f(x+\omega' t^{(l+1)} e_{l+1;l}) : t_1,...,t_{l+2} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

and this is the form we wanted to obtain.

After $d-1$ iterations, one arrives at the form

$$\Lambda^{2^{d-1}} \le \mathbb{E}(W_{d-1}(x,t_1,...,t_d)$$

$$\times \prod_{\omega' \in \{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_1,,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

This may be written as

$$\Lambda^{2^{d-1}} \le \mathbb{E}(\prod_{\omega' \in \{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_2,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d) + E,$$

where

$$E = \mathbb{E}((W_{d-1}(x,t_1,...,t_d) - 1)$$

$$\times \prod_{\omega' \in \{0,1\}^d} f(x+\omega' t^{(d-1)} e_{d;d-1}) : t_1,...,t_d \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d).$$

To see that the main term is in fact an appropriate box norm, notice that

$$e_{d;d-1} = \{e_d, e_d - e_1, ..., e_d - e_{d-1}\}$$

is also in general position.

To deal with the error term E, we apply the Cauchy-Schwarz inequality one more time to get

$$E \leq \mathbb{E}((W(x, t_2, ..., t_d) - 1)^2$$

$$\times \prod_{\omega' \in \{0,1\}^d} w(x + \omega' t^{(d)} e_{d;d-1}) : t_2, , ..., t_{d+1} \in \mathbb{Z}_N, x \in \mathbb{Z}_N^d),$$

where we have set

$$W(x, t_2, ..., t_d) = \mathbb{E}(W_{d-1}(x, t_1, t_2, ..., t_d) : t_1 \in \mathbb{Z}_N)$$

and again used the fact that $f \leq w$. Now to show that $E = o(1)$, it is enough to show that the linear forms defining W are pairwise independent, after of course expanding $(W - 1)^2$ and applying the linear forms condition. By following the construction of W, this amounts to showing that at each step W_l satisfies pairwise independence, which itself reduces to showing that the coefficient of x is 1 in each form and each form has a nonzero coefficient in t_1 (in each coordinate).

To be more precise, the case $l = 1$ is immediate. Assuming this is so for l fixed, then

$$W'_{l+1}(x, t_1, ..., t_{l+2}) = W_l(x, t_1, ..., t_{l+1}) W_l(x, t_1 + t_{l+2}, ..., t_{l+1})$$

$$\times \prod_{\omega' \in \{0,1\}^l} w(x + \omega' t^{(l)} e_{l;l-1}).$$

certainly satisfies this, as the forms in $W_l(x, t_1, ..., t_{l+1})$ and $W_l(x, t_1 + t_{l+2}, ..., t_{l+1})$ are pairwise independent because the t_1 coefficient is non-zero, and $\prod_{\omega' \in \{0,1\}^l} w(x + \omega' t^{(l)} e_{l;l-1})$ is independent of t_1. The statement about the coefficient of x is obvious. Also, it not hard to see that the vector multiple of t_1 is either e_{l+1} or $e_{l+1}^{(i)}$ (for forms appearing after i applications of Cauchy-Schwarz). Thus the statement is true for $l+1$.

The fact that $E = o(1)$ then follows directly from the $(d(d+2)2^{d-1}, d(2d+1))$ linear forms condition. □

2.4 The Dual Function Estimate.

As before we assume that a basis $e = \{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$ is given which is in general position. We will use the notation $\omega y e = \omega_1 y_1 e_1 + ... + \omega_d y_d e_d$, for $\omega \in \{0,1\}^d$ and $y \in \mathbb{Z}_N^d$. First we define the dual of a function $f : \mathbb{Z}_N^d \to \mathbb{R}$ with respect to the $\| \|_{\square(e)^d}$ norm.

Definition 2.4.1. . Let $f : \mathbb{Z}_N^d \to \mathbb{R}$ be a given function and let $e = \{e_1, ..., e_d\} \subseteq \mathbb{Z}_N^d$ be a basis of $\mathbb{Z}_N^d$. The dual of the the function f is the function

$$\mathcal{D}f(x) = \mathbb{E}\,(\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} f(x + \omega t e) : t \in \mathbb{Z}_N^d) \tag{2.5}$$

Proposition 2.4.1. *With X and $\mathcal{D}$ as above, and e in general position, we have*

$$||\mathcal{D}f_1 ... \mathcal{D}f_K||^*_{\square(e)^d} \leq C(K)$$

for any $f_1, ..., f_K \in X$.

Proof. We must show that

$$\langle f, \mathcal{D}f_1 ... \mathcal{D}f_K \rangle \leq C_K ||f||_{\square(e)^d}$$

by the definition of the dual norm. By applying the definition of $\mathcal{D}f$, the LHS gives

$$\langle f, \mathcal{D}f_1 ... \mathcal{D}f_K \rangle = \mathbb{E}(f(x) \prod_{i=1}^{K} \mathbb{E}(\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} f_i(x + \omega t^i e) : t^i \in \mathbb{Z}_N^d) : x \in \mathbb{Z}_N^d).$$

Expanding out the products then gives the right hand side as

$$\mathbb{E}(\mathbb{E}(f(x) \prod_{\omega\in\{0,1\}^d,\,\omega\neq 0} \times$$

$$\times \prod_{i=1}^{K} f_i(x+\omega t^i e + \omega t e) : x, t \in \mathbb{Z}_N^d) : T = (t^1, ..., t^K) \in (\mathbb{Z}_N^d)^K)$$

after a substitution $t^i \mapsto t + t^i$ for each i for some fixed t, and adding a redundant summation in t. Now we call $F_{(\omega,T)}(x) = \prod_{i=1}^K f_i(x+\omega t^i e)$ for non-zero ω, and $F_{(0^d,T)}(x) = f(x)$. The last expression then becomes

$$\mathbb{E}(\langle F_{(\omega,T)} : \omega \in \{0,1\}^d\rangle : T \in \mathbb{Z}_N^d).$$

By applying the $\Box(e)$-Cauchy-Schwarz inequality, we have arrived at

$$||\mathcal{D}f_1 ... \mathcal{D}f_K||^*_{\Box(e)^d} \leq \mathbb{E}(\prod_{\omega\in\{0,1\}^d,\,\omega\neq 0^d} ||F_{(\omega,T)}||_{\Box(e)^d} : T \in \mathbb{Z}_N^d).$$

An application of the Holder inequality gives that the right hand side is bounded above by

$$\prod_{\omega\in\{0,1\}^d,\,\omega\neq 0^d} \mathbb{E}(||F_{(\omega,T)}||^{2^d}_{\Box(e)^d} : T \in (\mathbb{Z}_N^d)^K),$$

where we added one factor of the constant 1 function, which has L^q-norm one for each q. Thus, we now just need to show that for a fixed $\omega \neq 0^d$ we have

$$\mathbb{E}(||F_{(\omega,T)}||^{2^d}_{\Box(e)^d} : T \in (\mathbb{Z}_N^d)^K) = O(K)$$

for $T = (t^1, ..., t^K)$.

We continue by expanding the last expression for a fixed $\omega \neq 0^d$,

$$||F_{(\omega,T)}||^{2^d}_{\Box(e)^d} : T \in (\mathbb{Z}_N^d)^K) = O(K)$$

$$= \mathbb{E}(\prod_{\omega' \in \{0,1\}^d} \prod_{i=1}^{K} f_i(x + \omega t^i e + \omega' te) : x, t, t^1, ..., t^K \in \mathbb{Z}_N^d).$$

The right hand side factorizes as

$$\mathbb{E}(\prod_{i=1}^{K} \mathbb{E}(\prod_{\omega' \in \{0,1\}^d} f_i(x + \omega ye + \omega' te) : y \in \mathbb{Z}_N^d) : x, t \in \mathbb{Z}_N^d).$$

Applying the bound $f \leq \nu$ gives

$$\mathbb{E}(\mathbb{E}^K(\prod_{\omega' \in \{0,1\}^d} \nu(x + \omega ye + \omega' te) : y \in \mathbb{Z}_N^d) : x, t \in \mathbb{Z}_N^d).$$

The inner sum is now split component wise

$$\mathbb{E}(\prod_{j=1}^{d} \prod_{\omega' \in \{0,1\}^d} \mu((\omega ye)_j + (\omega' te + x)_j) : y \in \mathbb{Z}_N^d),$$

where the notation $(x)_j$ denotes the j^{th} coordinate. The terms $(\omega ye)_j$ represent the linear forms $\sum_{s=1}^{d} \omega_s y_s (e_s)_j$, which satisfy the hypothesis in the $(d, 2^d)$ correlation condition by the assumptions on e. Hence we have

$$\mathbb{E}(\prod_{j=1}^{d} \prod_{\omega' \in \{0,1\}^d} \mu((\omega ye)_j + (\omega' te + x)_j) : y \in \mathbb{Z}_N^d) \leq \prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \tau(((\omega' - \omega'')te)_j),$$

as the $(x)_j$ terms drop out in the subtraction.

Plugging this bound back in gives

$$\mathbb{E}((\prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \tau(((\omega' - \omega'')te)_j))^K : t \in \mathbb{Z}_N^d).$$

Making use of the triangle inequality in $\mathcal{L}^{dK}$, after another application of

Holder, reduces our task to bounding

$$\prod_{j=1}^{d} \sum_{\omega' \neq \omega''} \mathbb{E}(\tau^{dK}(((\omega' - \omega'')te)_j) : t \in \mathbb{Z}_N^d).$$

By the assumptions on e and the fact that $\omega' - \omega'' \neq 0^d$, $((\omega' - \omega'')te)_j$ provides a uniform cover of $\mathbb{Z}_N$, and we may reduce this to

$$\mathbb{E}(\tau^{dK}(t) : t \in \mathbb{Z}_N).$$

This expression is $O_K(1)$. □

2.5 Proof of the Main Results.

In this section we prove our main results under the assumption that the measure exhibited in [12] is d-pseudo-random, i.e. it satisfies Definition 2.2.4.

2.5.1 Proof of Theorem 1.2.2.

Let $e = \{e_1, \dots, e_d\} \subseteq \mathbb{Z}_N^d$ be a basis which is in general position. For a function $f : \mathbb{Z}_N^d \to \mathbb{R}$ we define its dual by

$$\mathcal{D}f(x) = \mathbb{E}(\prod_{\omega \in \{0,1\}^d, \omega \neq 0} f(x + \omega te) : t \in \mathbb{Z}_N^d). \tag{2.6}$$

Then clearly

$$\langle f, Df \rangle = \|f\|_{\Box(e)^d}^{2^d} \tag{2.7}$$

Let $\mu = \otimes^d \nu$ be a pseudo-random d-measure, and let X be the set of functions f on $\mathbb{Z}_N^d$ such that $|f| \leq \mu$ pointwise.

Lemma 2.5.1. *The norm $\| \|_{\Box(e)^d}$ is a quasi algebra predual (QAP) norm, with respect to the set X and the operator D.*

Proof. We have already shown part (iii) of Definition 2.2.1, which was the content of Proposition 2.4.1. If $\|f\|^d_{\Box(e)^d} \leq \varepsilon$ then

$$\langle f, Df\rangle = \|f\|^{2^d}_{\Box(e)^d} \leq \varepsilon^{2^d}$$

and part (ii) follows. Finally, since $|f| \leq \mu$ it follows that

$$\langle f, Df\rangle \leq \|\mu\|^{2^d}_{\Box(e)^d} = 1 + o(1)$$

as the linear forms $(x + \omega te)_j$ are pairwise linearly independent (for each j) and ν satisfies the linear forms condition. □

We are in the position to apply the transference principle to decompose a function $0 \leq f \leq \mu$ into the sum of a bounded function g and a function h which has small contribution to the expression in (1.2).

Proof of Theorem 1.2.2. Let $\alpha > 0$ and let $0 \leq f \leq \mu$ be function such that $\mathbb{E}f \geq \alpha$, where μ is a pseudo-random d-measure on $\mathbb{Z}^d_N$. We apply Theorem 2.2.1, with $Y = \mathbb{Z}^d_N$, $\delta = 1/2$ and $\eta > 0$. Note that since μ is a measure one has that $\|\mu\|_{L^1} = \mathbb{E}\mu = 1 + o(1)$. Since $\| \ \|_{\Box(e)^d}$ is a *QAP* norm with respect to the set $X = \{f : Y \to \mathbb{R},\ |f| \leq \mu\}$, it follows that there is an $\varepsilon > 0$ such that if

$$\|\mu - 1\|_{\Box(e)^d} < \varepsilon \tag{2.8}$$

then there is a decomposition $f = g + h$ such that

$$0 \leq g \leq 2 \qquad \text{and} \qquad \|h\|_{\Box(e)^d} < \eta. \tag{2.9}$$

Since μ is pseudo-random $\|\mu - 1\|_{\Box(e)^d} = o(1)$ thus (2.8) holds for large enough N. Using this decomposition together with Theorem 1.2.1 and Proposition 2.3.1 one may write

$$\mathbb{E}(f(x)f(x+te_1)...f(x+te_d) : x \in \mathbb{Z}^d_N, t \in \mathbb{Z}_N) =$$

$$= \ \mathbb{E}(g(x)g(x+te_1)...g(x+te_d) : x \in \mathbb{Z}^d_N, t \in \mathbb{Z}_N) \ + \ O(\|h\|_{\Box(e)^d})$$

$$\geq c'(\alpha, e) - C_d\eta \geq c'(\alpha, e)/2$$

by choosing η sufficiently small with respect to α and e. This proves Theorem 1.2.2. □

2.5.2 Proof of Theorem 1.2.1.

Let us identify $[1,N]^d$ with $\mathbb{Z}_N^d$. First we show that constellations in $\mathbb{Z}_N^d$ defined by e which are contained in a box $B \subseteq [1,N]^d$ of size εN, are in fact genuine constellations contained in B. We say that $e = \{e_1, \dots, e_d\} \in \mathbb{Z}^{d^2}$ is *primitive* if the segment $[0,e]$ does not contain any other lattice points other than its endpoints in $\mathbb{Z}^{d^2}$ considered as a lattice point in $\mathbb{Z}^{d^2}$. Let us also define the positive quantity $\tau(e)$ by

$$\tau(e) = \inf_{m \notin \{0,e\},\, x \in [0,e]} ||m - x||_{L^\infty} \qquad \text{where} \qquad |x|_\infty = \max_{1 \le j \le d^2} |x_j|$$

m is running through the lattice points $\mathbb{Z}^{d^2}$ other than 0 and e.

Lemma 2.5.2. *Let $0 < \varepsilon < \tau(e)$. Let N be sufficiently large, and let $B = I^d$ be a box of size εN contained in $[1,N]^d \simeq \mathbb{Z}_N^d$. If there exist $x \in \mathbb{Z}_N^d$ and $t \in \mathbb{Z}_N \backslash \{0\}$ such that $x \in B$ and $x + te \subseteq B$ as a subset on $\mathbb{Z}_N^d$, then there exists a scalar $t' \neq 0$ such that $x + t'e \subseteq B$ also as a subset of $\mathbb{Z}^d$. Moreover if e is primitive (and $1 \le t < N$) then one may take $t' = t$ or $t' = t - N$.*

Proof. First, note that one can assume e is primitive as $x + te = x + tse'$ for a fixed primitive e' and $s \in \mathbb{N}$. By our assumption, there is an $x \in [1,N]^d$ and $t \in [1, N-1]$ such that $x \in B$ and $x + te_j \in B + (N\mathbb{Z})^d$ for all $1 \le j \le d$. Thus for each j there exits $m_j \in \mathbb{Z}^d$ such that $||te_j - Nm_j||_{L^\infty} \le \varepsilon N$ and hence $|\lambda e - m|_\infty \le \varepsilon$, where $m = \{m_1, \dots, m_d\} \in \mathbb{Z}^{d^2}$ and $\lambda = t/N$. Since $0 < \lambda < 1$ and $\varepsilon < \tau(e)$ this implies that $m = 0$ or $m = e$. If $m = 0$ then $|te|_\infty \le \varepsilon N$ and since $x \in B$ it follows that $x + te \subseteq B \subseteq \mathbb{Z}^d$. If $m = e$ then $||(t-N)e_j||_{L^\infty} \le \varepsilon N$ thus $x + (t-N)e \subseteq B \subseteq \mathbb{Z}^d$, so $x + t'e \subseteq B$ as a subset of $\mathbb{Z}^d$. This proves the lemma. □

Let us briefly recall the pseudo-random measure ν defined in Sec.9 [12]. Let $w = w(N)$ be a sufficiently slowly growing function (choosing $w(N) = O(\log\log N)$ is sufficient as in [12]) and let $W = \prod_{p \le w} p$ be the product of primes up to w. For given b relative prime to W define the modified von Mangoldt function $\bar{\Lambda}_b$ by

$$\bar{\Lambda}_b(n) = \begin{cases} \frac{\phi(W)}{W} \log(Wn+b) & \text{if } Wn+b \text{ is a prime;} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

where ϕ is the Euler function. Note that by Dirichlet's theorem on the distribution of primes in residue classes one has that $\sum_{n\leq N} \bar{\Lambda}_b(n) = N(1+o(1))$. Also, if $A \subseteq P^d$ is of positive relative density α and if $\bar{\Lambda}_b^d := \otimes^d \bar{\Lambda}_b$ is the d-fold tensor product of $\bar{\Lambda}_b$ the it is easy to see that there exists a b such that

$$\limsup_{N\to\infty} N^{-d} \sum_{x\in[1,N]^d} \mathbf{1}_A(x)\bar{\Lambda}_b^d(x) \; > \; \alpha/2 \tag{2.11}$$

We will fix such b and choose N sufficiently large N for which the expression in (2.11) is at least $\alpha/2$. Let $R = N^{d^{-1}2^{-d-5}}$ and recall the Goldston-Yildirim divisor sum [12], [9]

$$\Lambda_R(n) = \sum_{d|n, d\leq R} \mu(d) \, \log(R/d)$$

μ being the Möbius function. For given small parameters $0 < \varepsilon_1 < \varepsilon_2 < 1$ (whose exact values will be specified later) recall the Green-Tao measure

$$\nu(n) = \begin{cases} \frac{\phi(W)}{W} \frac{\Lambda_R(Wn+b)^2}{\log R} & \text{if } \varepsilon_1 N \leq n \leq \varepsilon_2 N; \\ 1 & \text{otherwise.} \end{cases} \tag{2.12}$$

Note that $\nu(n) \geq 0$ for all n, and also it is easy to see that for N sufficiently large, one has that

$$\nu(n) \geq d^{-1} 2^{-d-6} \, \bar{\Lambda}_b(n) \tag{2.13}$$

for all $\varepsilon_1 N \leq n \leq \varepsilon_2 N$. Indeed, this is trivial unless $Wn+b$ is a prime. In that case, since $\varepsilon_1 N > R$, $\Lambda_R(Wn+b) = \log R = d^{-1}2^{-d-5} \log N$ and (2.13) follows.

Proof of Theorem 1.2.1. Set $\mu = \otimes^d \nu$, and let

$$g(x) := c_d\, \bar{\Lambda}_b^d(x)\, \mathbf{1}_A(x)\, \mathbf{1}_{[\varepsilon_1 N, \varepsilon_2 N]^d}(x) \qquad (c_d = d^{-d} 2^{-d^2-6d}) \tag{2.14}$$

Then by (2.13) one has that $g(x) \leq \mu(x)$ for all $x \in \mathbb{Z}_+^d$. By (2.12) one may choose a sufficiently large number N' for which

$$(N')^{-d} \sum_{x \in [1,N']^d} \mathbf{1}_A(x) \bar{\Lambda}_b^d(x) \;>\; \alpha/2 \tag{2.15}$$

and a prime N such that

$$(1 - \frac{\alpha}{100d})N' \leq \varepsilon_2 N \leq N'$$

If ε_1 is such that $\varepsilon_1/\varepsilon_2 \leq \alpha/100d$, then by the Prime Number Theorem in arithmetic progressions

$$(N')^{-d} \sum_{x \in [1,N']^d \setminus [\varepsilon_1 N, \varepsilon_2 N]^d} \bar{\Lambda}_b^d(x) \;\leq\; \alpha/10 \tag{2.16}$$

It follows from (2.15) and (2.16)

$$N^{-d} \sum_{x \in [1,N']^d} g(x) \;\geq\; c_d\, N^{-d} \sum_{x \in [\varepsilon_1 N, \varepsilon_2 N]^d} \mathbf{1}_A(x) \bar{\Lambda}_b^d(x) \;\geq\; c_d \varepsilon_2^d \alpha/4 \tag{2.17}$$

Using the identification $[1, N]^d \simeq \mathbb{Z}_N^d$, one has that $\mathbb{E}(g(x) : x \in \mathbb{Z}_N^d) \geq \alpha'$ (with $\alpha' = c_d^d \varepsilon_2^d \alpha/4$), and $0 \leq g(x) \leq \mu(x)$ for all x. Thus, save for proving that the measure ν is d-pseudo-random, Theorem 1.2.2 implies that

$$\mathbb{E}(g(x)g(x + te_1) \ldots g(x + te_d) : \; x \in \mathbb{Z}_N^d, t \in \mathbb{Z}_N) \;\geq\; c'(\alpha, e) > 0.$$

Note that the contribution of trivial constellations, corresponding to $t = 0$, is at most $O(N^{-1} \log^d N)$, as $|\bar{\Lambda}_b^d| \leq \log^d N$ uniformly on $[1, N]^d$. Since the support of g is contained in $A \cap [\varepsilon_1 N, \varepsilon_2 N]^d$, Lemma 2.5.2 implies that $A \cap [\varepsilon_1 N, \varepsilon_2 N]^d$ must contain genuine constellations of the form $\{x, x + te_1, \ldots, x + te_d\}$ as a subset of $\mathbb{Z}^d$. Choosing an infinite sequence of N's it follows that A contains infinitely many constellations defined by e. □

2.6 The Correlation Condition.

To complete the proof of Theorem 1.2.1, one needs to show that the measure ν defined in (2.12) satisfies both the linear forms conditions and the $(d, 2^d)$ correlation conditions given above. Since the measure ν is the same (apart from the slight change in the interval where $\nu \equiv 1$) is the one given in [12] (see Definition 9.3, there), the linear forms condition is already established in Prop. 9.8 in [12]. It turns out that the arguments given in [12] (see Prop. 9.6, Lemma 9.9 and Prop.9.10) generalize in a straightforward manner to obtain the more general (m_0, m_1) correlation condition for any given specific values of m_0 and m_1.

Proposition 2.6.1. For a fixed m_0, m_1, there exists a function τ such that

$$\mathbb{E}\tau^k = O_k(1)$$

and also

$$\mathbb{E}(\prod_{i=1}^{m_1}\prod_{j=1}^{m_0} \nu(\phi_i(y) + h_{i,j}) : y \in \mathbb{Z}_N^r) \leq \prod_{i=1}^{m_0}(\sum_{1\leq j<j'\leq m_0} \tau(h_{i,j} - h_{i,j'})) \quad (2.18)$$

where the $\phi_i : \mathbb{Z}_N^r \to \mathbb{Z}_N$ are pairwise linearly independent linear forms.

Let us first note that the arguments of Lemma 9.9 and Prop. 9.10 of [12] applies to our case and it is enough to establish the following inequality (see Prop. 9.6 [12])

$$\mathbb{E}\,(\prod_{i=1}^{m_1}\prod_{j=1}^{m_0} \Lambda_R^2(W(\phi_i(y) + h_{i,j}) + b) :\ y \in B)$$

$$\leq\ C_M \left(\frac{W \log R}{\phi(W)}\right)^M \prod_{i=1}^{m_1}\prod_{p|\triangle_i} (1 + O_M(p^{-1/2})) \quad (2.19)$$

where $M = m_1 m_0$ and B is a box of size at most R^{10M}. Moreover one can assume that $h_{i,j} \neq h_{i,j'}$ for all i, $j \neq j'$.

The next step is, following [12], to write the the expression

$$\mathbb{E}(\prod_{i=1}^{M} \Lambda_R^2(\theta_i(y)) : y \in B),$$

where $\theta_i = W(\phi_{\lfloor i/m_1 \rfloor}(y) + h_{\lfloor i/m_1 \rfloor, (i\,(p))}) + b$ ($\lfloor x \rfloor$ is the floor function, $i\,(m_1)$ is i modulo m_1), as a contour integral of the the following form plus a small error

$$(2\pi i)^{-M} \int_{\Gamma_1} ... \int_{\Gamma_1} F(z, z') \prod_{j=1}^{M} \frac{R^{z_j + z'{}_j}}{z_j^2 z'{}_j^2} dz_j dz'_j, \tag{2.20}$$

where $z = (z_1, ..., z_M)$, $z' = (z'_1, ..., z'_M)$, and function $F(z, z')$ is taking form of an Euler product

$$F(z, z') = \prod_p E_p(z, , z'),$$

where

$$E_p(z, z') = \sum_{X, X' \subseteq [M]} \frac{(-1)^{|X| + |X'|} \omega_{X \bigcup X'}(p)}{p^{\sum_{j \in X} z_j + \sum_{j \in X'} z'_j}}.$$

The function ω relates this expression to the particular forms. Specifically

$$\omega_X(p) = \mathbb{E}(\prod_{i \in X} \mathbf{1}_{\theta_i \equiv 0\,(p)} : x \in \mathbb{Z}_N^r).$$

Lemma 2.6.1. *(Local factor estimate). Set the intervals $I_i = [(i-1)m_1 + 1, im_1]$ as a partition of $[M]$. For $\alpha \in I_i$, the homogeneous part of θ_α is $W\phi_i$. Also, set $\Delta_i = \prod_{j<j;\, j,j' \in I_i} |h_{i,j} - h_{i,j'}|$.The following estimates hold: $\omega_X(p)$:*

1. *If $p \leq w(N)$, then $\omega_X(p) = 1$ if $|X| = 0$, and is 0 otherwise.*
2. *If $p > w(N)$ and $|X| = 0$, then $w_X(p) = 1$.*
3. *If $p > w(N)$ and $X \subseteq I_i$ is nonempty, we have $w_X(p) = p^{-1}$ when $|X| = 1$, and $w_X(p) \leq p^{-1}$ when $|X| > 1$. In the latter case, if $p \nmid \Delta_\alpha$, we have that $\omega_X(p) = 0$.*

4. *If $p > w(N)$ and $X \cap I_i \neq \emptyset$ and $X \cap I_{i'} \neq \emptyset$ for some $i \neq i'$, we have $\omega_X(p) \leq p^{-2}$.*

Proof. When $p \leq w(N)$, then $W\phi_i + b \equiv b\,(p)$, giving the first result. The second statement is trivial.

For the third statement, let us start with $X \subseteq I_i$ with $|X| = 1$. Then we have

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)} : y \in \mathbb{Z}_N^r) = p^{-1}$$

for any fixed j, proving the first part. The second part requires an estimate of

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)}\mathbf{1}_{W(\phi_i(y)+h_{i,\,j'})+b\equiv 0\,(p)} : y \in \mathbb{Z}_N^r),$$

with $j \neq j'$. If $p|\,|h_{\alpha,\,j} - h_{\alpha,\,j'}|$, then the we are left with simply a single equation $(p \nmid W)$, and may refer to the first part. When $p \nmid \Delta_\alpha$, $\omega_X(p) = 0$ as $h_{i,\,j}$ is not congruent to $h_{i,\,j'}$, modulo p.

For the last statement, we have the upper bound

$$\mathbb{E}(\mathbf{1}_{W(\phi_i(y)+h_{i,\,j})+b\equiv 0\,(p)}\mathbf{1}_{W(\phi_i'(y)+h_{i',\,j'})+b\equiv 0\,(p)} : y \in \mathbb{Z}_N^r)$$

for some $i \neq i'$ and j, j'. The forms ϕ_i and $\phi_{i'}$ are linearly independent modulo p (see the proof of Lemma 10.1 in [12]), hence we have the intersection of two distinct linear algebraic sets, which has size at most p^{r-2}. □

The terms E_p in the Euler product can be separated as

$$E_p(z, z') = 1 - \mathbf{1}_{p>w(N)} \sum_{j=1}^{M} (p^{-1-z_j} + p^{-1-z_j'} - p^{-1-z_j-z_j'})$$

$$+ \sum_{i=1}^{m_1} \mathbf{1}_{p>w(N);\, p|\Delta_i} \lambda_p^{(i)}(z, z') + \sum_{X \bigcup X' \not\subseteq I_\alpha,\, \alpha\in[m_1];\, |X \bigcup X'|>1} \frac{O_M(p^{-2})}{p^{\sum_X z_j + \sum_{X'} z_j'}},$$

where

$$\lambda_p^{(i)}(z, z') = \sum_{X \bigcup X' \subset I_i;\, |X \bigcup X'|>1} \frac{O_M(p^{-1})}{p^{\sum_X z_j + \sum_{X'} z_j'}}.$$

We define the terms

$$E_p^{(0)} = 1 + \sum_{i=1}^{m_1} \mathbf{1}_{p>w(N);\, p|\Delta_i} \lambda_p^{(i)}(z,z'),$$

and factorize $E_p = E_p^{(0)} E_p^{(1)} E_p^{(2)} E_p^{(3)}$ as follows:

$$E_p^{(1)} = (E_p^{(0)})^{-1} \times$$

$$\times (\frac{E_p}{\prod_{j=1}^M (1 - \mathbf{1}_{p>w(N)} p^{-1-z_j})(1 - \mathbf{1}_{p>w(N)} p^{-1-z'_j})(1 - \mathbf{1}_{p>w(N)} p^{-1-z_j-z'_j})^{-1}},$$

and

$$E_p^{(2)} =$$

$$\prod_{j=1}^M (1 - \mathbf{1}_{p\le w(N)} p^{-1-z_j})^{-1} (1 - \mathbf{1}_{p\le w(N)} p^{-1-z'_j})^{-1} (1 - \mathbf{1}_{p\le w(N)} p^{-1-z_j-z'_j}),$$

and

$$E_p^{(3)} = \prod_{j=1}^M (1 - p^{-1-z_j})(1 - p^{-1-z'_j})(1 - p^{-1-z_j-z'_j})^{-1},$$

Also set $G_i = \prod_p E_p^{(i)}$, noting that

$$G_3 = \prod_{j=1}^M \frac{\zeta(1+z_j+z'_j)}{\zeta(1+z_j)\zeta(1+z'_j)}.$$

The the following is the analogue of lemma 10.6 in [12]. To state it, Let us recall the domain $\mathcal{D}_\sigma^M$ to be the set

$$\{z_j, z'_j : \Re z_j, \Re z'_j \in (-\sigma, 100)\,,\, 1 \le j \le M\}.$$

We also have the norms on for f analytic on $\mathcal{D}_\sigma^M$, denoted $||f||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)}$, given by

$$||f||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)} = \sup ||(\frac{\partial}{\partial z_1})^{\alpha_1}...(\frac{\partial}{\partial z_M})^{\alpha_1}(\frac{\partial}{\partial z_1'})^{\alpha_1}...(\frac{\partial}{\partial z_M'})^{\alpha_1} f||_{\mathcal{L}^\infty(\mathcal{D}_\sigma^M)},$$

where the supremum is taken over all $\alpha_1, ..., \alpha_M, \alpha_1', ..., \alpha_M'$ whose sum is at most k.

Lemma 2.6.2. *Let $0 < \sigma = 1/(6M)$. Then the Euler products G_i are absolutely convergent for $i = 0, 1, 2$ in the domain $\mathcal{D}_\sigma^M$, and hence represent analytic functions on this domain. We also have the estimates*

$$||G_0||_{\mathcal{C}^r(\mathcal{D}_\sigma^M)} = O_M(\log(R)/\log\log(R))^r \prod_{p|\prod_{i=1}^{m_1}\Delta_i} (1 + O_M(p^{2M\sigma-1}))$$

$$||G_0||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq \exp(O_M(\log^{1/3}(R)))$$

$$||G_1||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq O_M(1)$$

$$||G_2||_{\mathcal{C}^M(\mathcal{D}_{1/6M}^M)} \leq O_{M,w(N)}(1)$$

$$G_0(0,0) = \prod_{i=1}^{m_1}\prod_{p|\Delta_i}(1 + O_M(p^{-1/2}))$$

$$G_1(0,0) = 1 + o_M(1)$$

$$G_2(0,0) = (W/\phi(W))^M,$$

where the first bound is for all $0 \leq r \leq M$.

Proof. The estimates proceed exactly as in Lemma 10.3 and Lemma 10.6 in [12] with $\Delta = \prod_{i=1}^{m_1}\Delta_i$, barring the statement about $G_0(0,0)$. To see this, we have

$$G_0(0,0) = \prod_{p|\Delta} E_p^{(0)} = \prod_{p|\Delta}(1 + \sum_{i=1}^{m_1}\lambda_p^{(i)}(0,0)) \leq \prod_{i=1}^{m_1}\prod_{p|\Delta_i}(1 + |\lambda_p^{(i)}(0,0)|)$$

and we crudely have $|\lambda_p^{(i)}(0,0)| = 1 + O_M(p^{-1/2})$. □

The expression in (5.3) takes the form

$$(2\pi i)^{-M}\int_{\Gamma_1}...\int_{\Gamma_1}G(z,z')\prod_{j=1}^{M}\frac{\zeta(1+z_j+z_j')R^{z_j+z_j'}}{\zeta(1+z_j)\zeta(1+z_j')z_j^2z_j'^2}dz_jdz_j'$$

with $G=G_0G_1G_2$. To estimate it let us recall the following general result on contour integration from [12], see Lemma 10.4 there.

Lemma 2.6.3. *(Goldston-Yıldırım [12][5]) Let R be a positive number. If $G(z,z')$ is analytic in the $2M$ variables on $\mathcal{D}_\sigma^M$ for some $\sigma>0$, and we have the estimate*

$$||G||_{\mathcal{C}^k(\mathcal{D}_\sigma^M)}=\exp(O_{M,\sigma}(\log^{1/3}(R))),$$

then

$$(2\pi i)^{-M}\int_{\Gamma_1}...\int_{\Gamma_1}G(z,z')\prod_{j=1}^{M}\frac{\zeta(1+z_j+z_j')R^{z_j+z_j'}}{\zeta(1+z_j)\zeta(1+z_j')z_j^2z_j'^2}dz_jdz_j'$$

$$=G(0,...,0)\log^M(R)+\sum_{j=1}^{M}O_{M,\sigma}(||G||_{\mathcal{C}^j(\mathcal{D}_\sigma^M)})\log^{M-j}(R)$$

$$+O_{M,\sigma}(\exp(-\delta\sqrt{\log(R)}))$$

for some $\delta>0$.

Estimate (2.19) follows easily applying Lemma 5 (with $\sigma=1/6M$) to $G=G_0G_1G_2$ using Lemma 4, which in turn implies Proposition 2.6.1, where the function τ is defined precisely as in [12]. This finishes the proof of Theorem 1.2.1.

Chapter 3

3.1 Introduction.

The main goal for this portion of our work is to provide the analogue of Theorem 1.1.3 for a general integral quadratic form. As previously noted, we are applying the circle method of Hardy and Littlewood. The minor arcs are dealt with in section 2, which is done is two separate cases. The methods for the major arcs are standard, and worked out in section 3. Section 4 is dedicated to the singular series. The implications stated in the Chapter 1 are dealt with in the final section.

3.2 The Minor Arcs

3.2.1 Sufficiently Off Diagonal Forms

For this section, we make the stronger assumption that A has an m_1 by m_2 off-diagonal block say B, of rank at least R, which we shall determine later. The ability to handle this scenario is first noticed by Liu [15].

We set

$$T(r) = \sum_{x_1,\ldots,x_n=1}^{N} \Lambda(x_1)\ldots\Lambda(x_n)e(Q(x)r). \tag{3.1}$$

One may write in the form

$$T(r) = \sum_{y=(x_1,\ldots,x_{m_1})} \sum_{z=(x_{m_1+1},\ldots,x_n)} F(y)G(z)e(Q(y,z)r).$$

We use F and G simply as shorthand for the corresponding products of the von Mangoldt function. With two applications of the Cauchy-Schwartz inequality, and the fact that

$$\sum_{x=1}^{N} \Lambda(x)^2 \lesssim N \log N,$$

we have the Weyl-type inequality

$$|T(r)|^4 \lesssim N^{3n}(\log N)^{2n} \\ \times \sum_{h\in[-N,N]^{m_1}, l\in[-N,N]^{m_2}} e(2\,\langle l, Bh\rangle r) = N^{3n}(\log N)^{2n}\, V(r). \tag{3.2}$$

Writing $w = (h, l)$, we have that $w\,\langle l, Bh\rangle = \langle w, A'w\rangle$, where A' is obtained from A by making all entries a_{ij} zero when both $i \le m_1$ and $j \le m_1$, or both $i > m_1$ and $j > m_1$. In other words A_1 consists of the off-diagonal block B and its transpose B^T, hence it has rank $2R$.

Let us define the set of major arcs according to a parameter $0 < \theta < 1$ as

$$M(\theta) = \bigcup_{1\le q\le N^\theta} \bigcup_{(a,q)=1} M_{a,q}(\theta)$$

where

$$M_{a,q}(\theta) = \{r :\; 2|qr - a| \le N^{-2+\theta}\},$$

and the minor arcs are simply the complement of the major arcs. Then it is fairly standard (see e.g. Lemma 3.3 [3] and Lemma 3.2 in [6]), that one has the estimate for $r \notin M(\theta)$

$$|V(r)| \,\le C_n\,(\log N)^n\, N^{n-R\theta}.$$

Thus we have shown

Lemma 3.2.1. *Suppose A is a symmetric $n\times n$ matrix, which has an $m_1 \times m_2$ off-diagonal block of rank R. Then for $r \notin M(\theta)$, we have*

$$|T(r)| \,\le\, C_n\,(\log N)^n\, N^{n-\frac{R\theta}{4}}. \tag{3.3}$$

A much more precise formulation of this result is given by Liu [15] in his treatment of this case.

We will assume from now on that $R \geq 9$, and fix $\theta = \theta_1$ such that $R\theta_1 > 8$. Then in particular $T(r) = O(N^{n-2-\delta})$ for some fixed $\delta > 0$ for $r \notin M(\theta_1)$. We will use now a "sliding scale" argument due to Birch (see Lemma 4.4 [3]) to reduce the major arcs corresponding to value θ such that $N^{\theta} \approx (\log N)^C$ while keeping the error terms of size $O(N^{n-2}(\log N)^{-c})$ for some fixed constant c which depends on n. To do that we'll use the fact

$$|\mathsf{M}(\theta)| \leq N^{-2+2\theta}$$

which is immediate from the definition. We set up a sequence $\theta_1, ..., \theta_t$, such that $9\theta_1 > 8$, and $\theta_{i+1} = \frac{17}{18}\theta_i$, which will ensure that $2\theta_i - \frac{R}{4}\theta_{i+1} \leq -\frac{\theta_i}{8}$ for $R \geq 9$, thus by (3.3)

$$\begin{aligned} \int_{\mathsf{M}(\theta_i)-\mathsf{M}(\theta_{i+1})} |T(r)|\, dr &\lesssim (\log N)^n \, N^{n-\frac{R}{4}\theta_{i+1}} \, |\mathsf{M}(\theta_i)| \\ &\leq (\log N)^n \, N^{n-2-\frac{\theta_i}{8}}. \end{aligned} \tag{3.4}$$

Now if we fix θ_t such that $N^{\theta_t} \approx (\log N)^c$ for some fixed constant $c > 0$, then $t \approx c\frac{\log N}{\log\log N}$, thus we have shown

Lemma 3.2.2. *Assume that the matrix A has an off-diagonal block of rank $R \geq 9$. Let $c > 0$ be fixed, and let $0 < \theta < 1$ be such that $N^{\theta_t} \approx (\log N)^c$. Then one has the minor arcs estimate*

$$\int_{\mathsf{m}(\theta_t)} |T(r)|\, dr = O(N^{n-2}\frac{\log\,\log\,N}{(\log\,N)^C}), \tag{3.5}$$

with $C = \frac{c}{8} - n$, assuming N is sufficiently large.

3.2.2 Insufficiently Off Diagonal Forms

We decompose the form matrix A which defines Q into the following form

$$A = \begin{bmatrix} a & l_1 & l_2 \\ l_1 & A_1 & B \\ l_2 & B^t & A_2 \end{bmatrix}.$$

Here l_1, l_2 are vectors and A_1, A_2, B are matrices, and of course a comes from the pure quadratic term (which we assume is for x_1). Then we write

$$Q(x) = ax_1^2 + 2x_1L_1(y) + 2x_1L_2(z) + Q_1(y) + Q_2(z) + 2By \cdot z,$$

where we have decomposed $\mathbb{Z}^n = \mathbb{Z} \times \mathbb{Z}^{m_1} \times \mathbb{Z}^{m_2}$ $(x = (x_1, y, z))$.

This first thing we need to discuss is the decomposition of A, which is accomplished once we pick A_1. If we had any such decomposition giving B rank larger than 8 then we may use the previous section, so we assume $\text{rank}(B) \leq 8$. If we assume that A has overall rank of $R \geq 34$, then we can choose n_1 such that the matrix $\begin{bmatrix} A_1 & B \end{bmatrix}$ from the above form has rank precisely 20. Then the rank of A_1 is at least 20-8=12. It follows that the rank of $\begin{bmatrix} B^t & A_2 \end{bmatrix}$ is at least $R - 20 - 2 \geq 12$. So we have that the rank of the matrix A_2 is at least $R - 22 - 8 \geq 4$. So assuming $R \geq 34$ gives the ability to select A_1 with rank $R_{A_1} \geq 12$, and A_2 with rank $R_{A_2} \geq 4$.

For now let us fix a generic minor arc m, and look at the integral

$$\begin{aligned} I_{\mathsf{m}} := \int_{\mathsf{m}} & \sum_{(x_1,y,z)\in[N]\times[N]^{m_1}\times[N]^{m_2}} \Lambda(x_1)F(y)G(z) \\ & \times e((ax_1^2 + x_1L_1(y) + x_1L_2(z) + Q_1(y) + Q_2(z) + By \cdot z)r)dr. \end{aligned} \tag{3.6}$$

We partition the sum in the integral along the level sets of the linear forms $L_1(y)$, $L_2(z)$, and By. Then we have

$$\begin{aligned} & \sum_{t_1,t_2,t_3} \int_{\mathsf{m}} \sum_{x_1\in[N],\, y\in L_1^{-1}(t_1)\cap B^{-1}(t_3),\, z\in L_2^{-1}(t_2)} \Lambda(x_1)F(y)G(z) \\ & \times e((ax_1^2 + t_1x_1 + t_2x_1 + Q_1(y) + Q_2(z) + t_3 \cdot z)r)dr, \end{aligned} \tag{3.7}$$

where $L_1^{-1}(t_1) = \{y \in [N]^{m_1} : L_1(y) = t_1\}$, $L_2^{-1}(t_2) = \{z \in [N]^{m_2} : L_1(y) = t_2\}$ and $B^{-1}(t_3) = \{y \in [N]^{m_1} : B(y) = t_3\}$. Note that since that rank of B is $R_B \leq 8$, t_3 runs through $\Gamma_B \cap [-CN, CN]^{m_2}$, where $\Gamma_B = B(\mathbb{Z}^{m_1})$ is a sublattice of rank R_B.

First lets assume, the generic case, when the linear form $L_1(y)$ is linearly independent of the forms defining By. Otherwise, the value t_3 would uniquely determine t_1 so we would not need to restrict to the level set of the form $L_1(y)$, a case we will get back to later. Similarly, we assume first that $L_2(z)$ is not identically zero.

The innermost sums now split into a product. Call the x_1 sum S_0, the y sum S_1, and the z sum S_2, and we have the form

$$\sum_{t_1,t_2,t_3} \int_{\mathsf{m}} S_0(r, t_1, t_2) S_1(r, t_1, t_3) S_2(r, t_2, t_3) dr := \sum_{t_1,t_2,t_3} U(t_1, t_2, t_3).$$

We have the simple bound

$$U(t_1, t_2, t_3) \leq ||S_0(\cdot, t_1, t_2)||_{L^\infty(\mathsf{m})} ||S_1(\cdot, t_1, t_3)||_{L^2(\mathbb{T})} ||S_2(\cdot, t_2, t_3)||_{L^2(\mathbb{T})},$$

where $\mathbb{T}$ denotes $\mathbb{R}/\mathbb{Z}$. If $t_1 + t_2 \neq 0$, then we may apply Hua's bound on S_0 (see e.g. Lemma 10.8 [14]). If we have $t_1 + t_2 = 0$, then the following argument may be rerun to give a power gain. Let us assume then that we have $t_1 + t_2 \neq 0$. Then we may choose the parameter c defining the minor arcs such that

$$||S_0(\cdot, t_1, t_2)||_{L^\infty(\mathsf{m})} \lesssim N \, (\log \, N)^{-C}, \tag{3.8}$$

on m for any given constant C uniformly in t_1 and t_2. It now follows from the Cauchy-Schwarz inequality and the fact that the parameters (t_1, t_2, t_3) can take $O(N^{R_B+2})$ values, that

$$|U_{\mathsf{m}}|^2 \lesssim N^{R_B+4} (\log \, N)^{-C} \sum_{t_1,t_2,t_3} ||S_1(\cdot, t_1, t_3)||^2_{L^2(\mathbb{T})} ||S_2(\cdot, t_2, t_3)||^2_{L^2(\mathbb{T})} \tag{3.9}$$

For fixed t_1, t_2, t_3, the L^2 estimates are the weighted number of solutions in the primes to the systems of equations

$$Q_1(y) = Q_1(y')$$

$$L_1(y) = L_1(y') = t_1$$

$$By = By' = t_3,$$

and

$$Q_2(z) + t_3 \cdot z = Q_2(z') + t_3 \cdot z'$$

$$L_2(z) = L_2(z') = t_2.$$

If we sum these over t_1 and t_2 then the systems become

$$Q_1(y) = Q_1(y')$$

$$L_1(y) = L_1(y')$$

$$By = By' = t_3,$$

and

$$Q_2(z) + t_3 \cdot z = Q_2(z') + t_3 \cdot z'$$

$$L_2(z) = L_2(z').$$

Let $u(t_3)$ and $v(t_3)$ denote the number of solutions to these systems over the in the natural numbers of size at most N. Then $u(t_3)v(t_3)$ is the number of solutions to the system of equations

$$Q_1(y) = Q_1(y')$$

$$L_1(y) = L_1(y')$$

$$By = By' = t_3$$

$$Q_2(z) + By \cdot z = Q_2(z') + By' \cdot z'$$

$$L_2(z) = L_2(z').$$

The sum over t_3 reduces this to the number of solutions of

$$\begin{aligned} Q_1(y) &= Q_1(y') \qquad (3.10) \\ L_1(y) &= L_1(y') \\ By &= By' \\ Q_2(z) + By \cdot z &= Q_2(z') + By' \cdot z' \\ L_2(z) &= L_2(z'), \end{aligned}$$

which we denote by W. Since the weights are at most log N, the integral over the minor arcs is then bounded above by

$$|I_{\mathsf{m}}|^2 \lesssim N^{R_B+4}(\log\, N)^{-C+n}\, W.$$

The following and a few additional remarks finish the argument for Lemma 3.2.4 below.

Lemma 3.2.3. *We have the bound*

$$W \lesssim N^{2n-R_B-8}.$$

Proof. We will use the well-known fact (see e.g. [18]), that if $Q'(x)$ is an integral quadratic form of rank at least 5 in n variables and if $v \in \mathbb{Z}^n$, then the number of solutions of the equation $Q'(x) + v \cdot x = 0$ in $[-N, N]^n$ is of $O(N^{n-2})$.

Now for the system (3.10), we have that $Q_1(y) - Q_1(y') = 0$, that is $Q'(y, y') = 0$ with the quadratic form Q' of rank twice the rank of A_1, so is at least 14 by our construction. Now restricting $Q'(y, y')$ to the subspace M defined by the linear equations: $L_1(y) - L_1(y') = 0$, $By - By' = 0$, which is by our assumptions is a subspace of codimension $R_B + 1 \leq 9$ in $\mathbb{R}^{2m_1}$, its rank is still at least $2R_{A_1} - 18 \geq 6$. Thus the number of solutions in $(y, y') \in M \cap [N]^{2m_1}$ is of $O(N^{2m_1-R_B-3})$, where the implicit constant depends only on the coefficients on the matrix A.

Next, fix a solution (y, y') and consider the equations the number of pairs

$(z,z') \in [-N,N]^{2m_2}$ for which $Q_2(z) - Q_2(z') + By \cdot z - By' \cdot z' = 0$ and $L_2(z) - L_2(z') = 0$. Since the rank of the form $Q_2(z) - Q_2(z')$ is $2R_{A_2} \geq 8$ it follows that its restriction to the hyperplane $\{L_2(z) - L_2(z') = 0\}$ has rank at least 6. Thus the number of solutions (z,z') is of magnitude $O(N^{2m_2-3})$, where the implicit constant depend only on the matrix A. This yields

$$W \lesssim N^{2m_1+2m_2-R_B-6} = N^{2n-R_B-8}.$$

□

For the case when $L_1(y)$ is linearly dependent of By, that is: $L_1(y) = By{\cdot}\gamma$ for some fixed rational vector γ, we only need to restrict the summation along the level sets of By and $L_2(z)$. Thus one has

$$\begin{aligned} |I_{\mathsf{m}}| \leq \sum_{t_2,t_3} \int_{\mathsf{m}} \sum_{x_1 \in [N],\, y \in B^{-1}(t_3),\, z \in L_2^{-1}(t_2)} & \Lambda(x_1) F(y) G(z) \\ \times e((ax_1^2 + t_3 \cdot \gamma\, x_1 + t_2 x_1 + Q_1(y) + Q_2(z) + t_3 \cdot z)r) dr, & \end{aligned} \tag{3.11}$$

and the rest of the analysis goes along the same lines. Similarly if $L_2(z)$ is identically 0, then there is of course no need for the parameter t_2.

We now have achieved

Lemma 3.2.4. *Assume that the matrix A has rank $R \geq 34$. Let $C > 0$ be a fixed constant. If $c > 0$ is a fixed constant, sufficiently large with respect to C and N, and if $0 < \theta < 1$ is such that $N^\theta = (\log N)^c$, then one has the minor arcs estimate*

$$\int_{\mathsf{m}(\theta)} |T(r)| \lesssim N^{n-2} (\log N)^{-C}. \tag{3.12}$$

3.3 The Major Arcs and an Asymptotic Formula

The major arcs are now a union of intervals of the form $\mathsf{M}_{a,q}((\log N)^c)$ $(q \leq (\log N)^c)$, where c is given by Lemma 3.2.4, and is fixed throughout this section. For a fixed a, q we look at the exponential sum T, and as the major arcs are small, we may use any approximation that has a logarithmic gain in the error.

To start we fix a $q \leq (\log\, N)^c$ and some $a \in \mathbb{Z}_q^*$. We follow the standard arguments, albeit with a slightly different look. Write

$$\begin{aligned} T(r) &= \sum_{x \in [N]^n} F(x) e(Q(x)r) \\ &= \sum_{s \in \mathbb{Z}_q^n} \sum_{x \in [N]^n} \mathbf{1}_{x \equiv s\,(q)} F(x) e(aQ(s)/q) e(Q(x)\tau) \\ &= \sum_{s \in \mathbb{Z}_q^n} e(Q(s)a/q) \int_{z \in N\mathcal{B}} e(Q(z)\tau) d\psi_s(z), \end{aligned} \tag{3.13}$$

where we have set $\tau = r - a/q$, and $\psi_s(z) = \psi_{s_1}(z_1)...\psi_{s_n}(z_n)$ for $\psi_l(y) = \sum_{t \equiv l(q),\, t \leq y} \Lambda(t)$, and $\mathcal{B}$ is $[0,1]^n \subset \mathbb{R}^n$.

Lemma 3.3.1. *On each major arc $\mathsf{M}_{a,q}((\log\, N)^c)$, the following holds: Fix a constant $C > 2c$. For each $s \in \mathbb{Z}_q^n$ we have*

$$\begin{aligned} \int_{z \in N\mathcal{B}} e(Q(z)\tau) d\psi_s(z) = \mathbf{1}_{s \in (\mathbb{Z}_q^*)^n} \phi(q)^{-n} \int_{z \in N\mathcal{B}} e(Q(z)\tau) dz \\ + O(N^n (\log\, N)^{c - C/2}). \end{aligned} \tag{3.14}$$

Proof. Define for a fixed l the one dimensional signed measure $d\nu_l = d\psi_l - d\omega_l$, where $d\omega_l$ is the Lebesgue measure divided by the reciprocal of the totient of q if $l \in \mathbb{Z}_q^*$, and zero otherwise. For a continuous function f one then has

$$\int_{[0,N]} f(z) d\nu_l(z) = \sum_{x \in [N],\, x \equiv l\,(q)} f(x) - \phi(q)^{-1} \int_0^N f(z) dz.$$

Also set $d|\nu_l| = d\omega_l + d\psi_l$.

We have

$$\begin{aligned} \int_{z \in N\mathcal{B}} e(Q(z)\tau) d\psi_s(z) = \int_{z \in N\mathcal{B}} e(Q(z)\tau)(d\nu_{s_1}(z_1) + d\omega_{s_1}(z_1)) \\ \times ... \times (d\nu_{s_n}(z_n) + d\omega_{s_n}(z_n)). \end{aligned} \tag{3.15}$$

Expanding out the products in the last integral gives the form

$$\int_{z\in N\mathcal{B}} e(Q(z)\tau)d\omega_s(z) + \sum_{i=1}^{2^n-1} \int_{z\in N\mathcal{B}} e(Q(z)\tau)d\mu_{i,s}(z), \qquad (3.16)$$

where $d\mu_{i,s}$ runs over all the corresponding products, barring the $d\omega_s(z)$ term.

Consider

$$\int_{z\in N\mathcal{B}} e(Q(z)\tau)d\mu_{i,s}(z)$$

for some fixed i. Assume without loss of generality that $d\mu_{i,s}$ is of the form $d\nu_{s_1}(z_1)d\sigma_s(z_2, ..., z_n)$, where $d\sigma_s$ may be signed in some variables (and is of course independent of s_1). Now for the first component we shall split the continuous interval $[0, N]$ into smaller disjoint intervals of size $N\,(\log\, N)^{-C}$. Here C' is simply chosen to be between $C/2$ and C such that $(\log\, N)^C$ is an integer, say J. The equality $[0, N] = \bigcup_{j=1}^{J} I_j$ follows. Also let us set $\mathcal{B}_j = I_j \times [0, N]^{n-1}$, which absorbs the factor of N.

Now for a fixed interval I_j select some $y \in I_j$ and we have

$$\begin{aligned}\int_{z\in\mathcal{B}_|} e(Q(z)\tau)d\mu_{i,s} &= \int_{z\in\mathcal{B}_|} e(Q(y, z_2, ..., z_n)\tau)d\nu_{s_1}(z_1)d\sigma_s(z_2, ..., z_n)\\ &\quad + \int_{z\in\mathcal{B}_|} (e(Q(z_1, ..., z_n)\tau) - e(Q(y, z_2, ..., z_n)\tau))\\ &\quad \times d\nu_{s_1}(z_1)d\sigma_s(z_2, ..., z_n).\\ &:= E_1 + E_2\end{aligned}$$

We have

$$|E_1| \le \int_{z_2,...,z_n\in[0,N]} |\int_{I_j} d\nu_{s_1}(z_1)|\ d|\sigma_s|(z_2, ..., z_n) = O(N^n e^{-c_0\sqrt{\log\, N}})$$

by the Siegel-Walfisz Theorem, as $q \le (\log\, N)^c$. To bound E_2 we note that on I_j the integrand is $O(N^n\,(\log\,N)^{c-C'})$. In turn,

$$|E_2| \lesssim N^{-c\theta_N} \int_{z\in\mathcal{B}_|} d|\nu_{s_1}|(z_1)d|\sigma_s|(z_2,...,z_n) \lesssim N^n\,(\log\,N)^{c-2C'}\,.$$

Summing over the intervals gives the result for each error term. There are $2^n - 1$ error terms and the proof is complete. □

The integral appearing in the last result, namely

$$\int_{N\mathcal{B}} e(Q(z)\tau)dz = N^n \int_{\zeta\in\mathcal{B}} e(Q(\zeta)N^2\tau)d\zeta,$$

is denoted by $N^n\mathcal{I}(\mathcal{B}, N^2\tau)$ in [3]. This function is independent of a and q. Thus the integral over any major arc yields the common integral

$$\int_{\tau=(\log\,N)^c} \mathcal{I}(\mathcal{B}, N^2\tau)e(-\tau v)d\tau.$$

With $\mu = N^{-2}v$, set

$$J(\mu;\Phi) = \int_{|\tau|\leq\Phi} \mathcal{I}(\mathcal{B},\tau)e(-\tau v)d\tau,$$

and

$$J(\mu) = \lim_{\Phi\to\infty} J(\mu).$$

The following is Lemma 5.3 in [3].

Lemma 3.3.2. *$J(\mu)$ is continuous and uniformly bounded in μ. Moreover,*

$$|J(\mu) - J(\mu,\Phi)| \lesssim \Phi^{-\frac{1}{2}}$$

holds uniformly in μ.

If we define

$$W_{a,q} = \sum_{s\in(\mathbb{Z}_q^*)^n} e(Q(s)a/q),$$

then we now have

Lemma 3.3.3. *For a fixed major arc $\mathsf{M}_{a,q}((\log N)^c)$ and fixed constant $C > 3c$ we have*

$$\int_{\mathsf{M}_{a,q}} T(r)e(-vr)dr = N^{n-2}\phi(q)^{-n}W_{a,q}e(-va/q)J(\mu)+O(N^n (\log N)^{3c/2-C/2}),$$

where $\mu = N^{-2}v$.

Recall that the measure of the major arcs is at most $N^{-2+2c\theta_N}$, and define

$$\begin{aligned} B(v,q) &= \sum_{(a,q)=1} \phi(q)^{-n}W_{a,q}e(-va/q) \\ \mathfrak{S}(v,N) &= \sum_{q\leq(\log N)^c} B(v,q). \end{aligned}$$

It follows that

Lemma 3.3.4. *By choosing $C = 4c$ in the above arguments and setting $\delta = c/2 > 0$ we have*

$$\mathcal{M}(N,v) = \mathfrak{S}(v,N)J(\mu)N^{n-2} + O(N^{n-2} (\log N)^{-\delta}).$$

3.4 The Singular Series

Here we analyze the singular series $\mathfrak{S}(v,N)$ following the outline of [14].

Lemma 3.4.1. *For a given prime p, let R_p denote the rank of A over $\mathbb{Z}_p^n$. For all $a \in \mathbb{Z}_q^*$ the estimate*

$$|W_{a,p}| \lesssim p^{n-R_p/2}$$

holds, and the implied constant is dependent only on n. In turn

$$|B(v,p)| \lesssim p^{1-R_p/2}$$

holds uniformly in v.

Proof. Define the sets $Y_i = \{s \in \mathbb{Z}_p^* : s_i \equiv 0\,(p)\}$, $i = 1, ..., n$, and $Y = \bigcup_i Y_i$. Using the principle of inclusion-exclusion gives

$$\sum_{s\in\mathbb{Z}_p^n} 1_{s\in(\mathbb{Z}_p^*)^n} e(Q(s)a/p) = \sum_{s\in\mathbb{Z}_p^n} e(Q(s)a/p) -$$

$$-\sum_{k=1}^{n}(-1)^{k-1} \sum_{L\subseteq[n],\,|L|=k}\ \sum_{s\in\mathbb{Z}_p^n} 1_{s\in Y_L} e(Q(s)a/p),$$

where $Y_L = \bigcup_{i\in L} Y_i$.

The first term on the right hand side above is a Gaussian sum, and has the the upper bound $p^{n-R_p/2}$. For a fixed k above we have $\binom{n}{k}$ choices for L. For each choice we again have a Gaussian sum in $n-k$ variables which has rank at least $\alpha_k = \max\{0, R_p - 2k\}$. Hence for k fixed we again have a bound of $\binom{n}{k}p^{n-k-\alpha_k/2} \le \binom{n}{k}p^{n-R_p/2}$. The result then follows. □

Define, for a given prime p, β_p to the largest power of p to divide all the coefficients of A. Then set $\gamma_p = \beta_p + 1$ for $p > 2$ and $\gamma_2 = \beta_2 + 2$ for $p = 2$.

Lemma 3.4.2. *Fix a prime p, let $t \ge 2\gamma_p$, and for $\alpha > 0$ define R_t to be the rank of the map*

$$A : \mathbb{Z}^n_{p^{t-\alpha}} \to \mathbb{Z}^n_{p^\alpha}.$$

It follows that

$$|W_{a,p^t}| \le p^{tn - R_t(t-\lfloor t/2\rfloor)}.$$

Hence for large enough t, dependent only on A, we have

$$|W_{a,p^t}| \le p^{tn-R(t-\lfloor t/2\rfloor)},$$

where R is the rank of A over $\mathbb{R}^n$. Moreover, if A is nonsingular modulo p, then $W_{a,t} = 0$ for all $t > \gamma_p$.

Proof. Fix a prime p, and for simplicity set $\gamma_p = \gamma$. Let $t \ge 2\gamma$ and set

$\alpha = \lfloor t/2 \rfloor$. Then we apply the substitution $s = s_1 + p^{t-\alpha}s_2$ to get

$$\begin{aligned} W_{a,p^t} &= \sum_{s \in (\mathbb{Z}^*_{p^t})^n} e(Q(s)a/p^t) \\ &= \sum_{s_1 \in (\mathbb{Z}^*_{p^{t-\alpha}})^n} \sum_{s_2 \in \mathbb{Z}^n_{p^\alpha}} e(Q(s_1 + p^{t-\alpha}s_2)a/p^t) \\ &= \sum_{s_1 \in (\mathbb{Z}^*_{p^{t-\alpha}})^n} \sum_{s_2 \in \mathbb{Z}^n_{p^\alpha}} e(Q(s_1)a/p^t)e((2As_1 \cdot s_2)a/p^\alpha), \end{aligned}$$

as $\alpha \le t/2$. The inner sum is zero if $2As_1$ has a nonzero coordinate. Thus

$$|W_{a,p^t}| \le p^{\alpha n} |\{s_1 \in (\mathbb{Z}_{p^{t-\alpha}})^*)^n : 2As_1 \equiv 0 \,(p^\alpha)\}|.$$

This gives the upper bound

$$|W_{a,p^t}| \le p^{(t-\alpha)(n-R_t)} p^{\alpha n} \le p^{tn - R_t(t-\alpha)}.$$

Also, if t is sufficiently large with respect to A then we have that $R_t = R$.

Finally, since $s_1 \in (\mathbb{Z}^*_{p^t})^n$, it follows that $2As_1 \equiv 0$ has no solutions if A is nonsingular modulo p. Applying the above argument with $\alpha = 1$ completes the proof. □

We note that Lemma 3.4.1 covers $W_{a,p}$ for all sufficiently large primes. Also, for any prime p, Lemma 3.4.2 provides bounds for W_{a,p^t} for all sufficiently large t.

Lemma 3.4.3. *If* $(q_1, q_2) = 1$*, then*

$$W_{a,q_1q_2} = W_{aq_2,q_1} W_{aq_1,q_2}$$

and

$$B(v, q_1q_2) = B(v, q_1)B(v, q_2).$$

See [14] for the proof (Lemma 8.1). We are now in a position to provide a bound for $B(v, q)$

Lemma 3.4.4. *Given $\epsilon > 0$, we have*

$$B(v,q) \lesssim q^{1-R/2+\epsilon}$$

uniformly in v.

Proof. Applying Lemma 3.4.3 gives

$$B(v,q) = \phi(q)^{-n} \sum_{(a,q)=1} W_{a,q} e(-va/q) =$$

$$= \sum_{(a,q)=1} \phi(p_1^{t_1})^{-n} W_{a,p_1^{t_1}} ... \phi(p_l^{t_l})^{-n} W_{a,p_l^{t_l}}.$$

Now we apply Lemma 3.4.1 and Lemma 3.4.2 to get

$$|B(v,q)| \lesssim q \prod_{i=1}^{l} (1-1/p_i)^{-n} p_i^{-R(t_i - \lfloor t_i/2 \rfloor)},$$

where the implied constant absorbs the finite number of pairs (p_i, t_i) for which the rank is insufficient. We have

$$\prod_{i=1}^{l} (1-1/p_i)^{-n} \leq \prod_{p \leq q} (1-1/p)^{-n} \lesssim (\log\, q)^n.$$

Thus

$$|B(v,q)| \lesssim q^{1+\epsilon} \prod_{i=1}^{l} p_i^{-R(t_i - \lfloor t_i/2 \rfloor)} \leq q^{1+\epsilon-R/2}$$

as claimed. □

It easily follows now that the singular series is absolutely convergent when the the rank of A is at least 5. The infinite product representation follows as usual: with

$$\chi_p(v) = 1 + \sum_{t=1}^{\infty} B(v, p^t),$$

we have

$$\mathfrak{S}(v) = \lim_{N\to\infty} \mathfrak{S}(N, v) = \prod_p \chi_p(v).$$

Define $M(p^t, v)$ to be the number of solutions of $Q(x) \equiv v\,(p^t)$ where $x \in (\mathbb{Z}^*_{p^t})^n$. We have the analogue of Lemma 8.6 in [14].

Lemma 3.4.5. *We have $M(p^t, v) = \phi(p^t)^n p^{-t}(1 + \sum_{j=1}^t B(v, p^j))$.*

We conclude this section with one final result.

Lemma 3.4.6. *If A has rank at least 5, then there exists integers λ and K, and a positive number δ such that*

$$\mathfrak{S}(v) \geq \delta$$

whenever $v \equiv \lambda\,(K)$.

Proof. With the above estimates for $|B(v, q)|$, there exists a p_0 such that

$$\prod_{p>p_0} \chi_p(v) \geq \delta' > 0$$

holds for some positive δ' for all v. Set $\chi_p(v, t)$ to be the tth partial sum of the series defining χ_p. The estimates for $|B(v, p^t)|$ provide a t_0 such that

$$|\chi_p(v, t) - \chi_p(v)| < \frac{1}{2^{p_0+1}}$$

holds for all v. By simple averaging, Lemma 3.4.5 provides a v_p in $\mathbb{Z}_{p^t}$ such that $\chi_p(v_p, t) \geq 1$.

We now set $\lambda = \prod_{p\leq p_0} v_p$ and $K = \prod_{p\leq p_0} p$ and the result follows from the Chinese Remainder Theorem. □

3.5 Conclusions

Here we simply collect the pieces to prove Theorem 1.2.3 and Theorem 1.2.4 which are stated in the opening chapter.

Proof of Theorem 1.2.3. The asymptotic formula has already been showed to hold under this hypothesis. The statement about the positivity of $\mathfrak{S}$ is a consequence of Hensel's Lemma, see e.g. [4]. The statement regarding the function J is precisely the same as the one given in [3]. This completes the proof. $\square$

Proof of Theorem 1.2.4. We have seen in Lemma 3.4.6 that there is a infinite arithmetic progression $\mathcal{Z}$ such that $\mathfrak{S}(v) \geq \delta > 0$ for all sufficiently large elements $v \in \mathcal{Z}$. Also, over $\mathbb{R}$ it is easily seen that $Q = v$ has a non-singular solution (as Q is canonically quivalent to $x_1^2 + ... + x_n^2$). Thus the function J can be bounded below by a positive constant for these $v \approx N^2$, and the proof is complete. $\square$

Chapter 4

We take some time to conclude with some a discussion future projects.

4.1 Future Projects

4.1.1 A Conjecture

The most natural continuation of this work is to extend the results of Chapter 3 to equations of higher degree. We put forth a general conjecture, which is an analogue of Birch's Theorem for prime points. Let us overview Birch's results given in [3].

Let $f = (f_1, ..., f_R)$ be a system of homogeneous integral forms of common degree d in n variables. For a fixed $v \in \mathbb{Z}^R$ set $V(v)$ to be the complex affine variety defined by $f = v$. Set V^* to be the collection of points where the rank of Jac_f is strictly less than R. Set $K = 2^{1-d} codim(V*)$, where *codim* denotes the codimension. One should note that in the case that f is represented by a single quadratic form $Q = \langle x, Ax \rangle$, we have that the $codim(V^*)$ is simply the rank of the matrix A.

The main result of Birch states that if $K > R(R+1)(d-1)$, then the number of integer points in the box $x \in [N]^n$ satisfying $f(x) = v$, call this $\mathcal{N}(v, N)$, obeys

$$\mathcal{N}(v, N) = \mathfrak{S}(v) J(N^{-Rd} v) N^{n-Rd} + O(P^{n-Rd-\delta}) \tag{4.1}$$

for some $\delta > 0$, where $\mathfrak{S}$ is given here by the product of p-adic densities for the equation $f = v$, and J is precisely the same as in the previous chapter.

Here we conjecture the following. Define the singular series as

$$\begin{aligned} W_{a,q} &= \sum_{s\in(\mathbb{Z}_q^*)^n} e(Q(s)a/q) \\ B(v,q) &= \sum_{(a\in(\mathbb{Z}^*)^R} \phi(q)^{-n} W_{a,q} e((-v\cdot a)/q) \\ \mathfrak{S}(v,N) &= \sum_{q=1}^{\infty} B(v,q). \end{aligned}$$

Conjecture 4.1.1. *There exists a constant $K_0 = K(R,d)$ such that if the singular variety associated to the set $V = \{f = v\}$, f as above, has codimension $K \geq K_0$, then we have the weighted number of prime points on the $V \cap [N]^n$ satisfies*

$$\mathcal{M}(v,N) = \mathfrak{S}(v,N) J(N^{-Rd}v) N^{n-Rd} + o(N^{n-Rd}), \tag{4.2}$$

where $\mathfrak{S}$ is given in 4.2, J is the same as in 4.1, and the implied constant depends in the little o depends only on $n, R,$ and d.

It is worth noting that the constant K_0 gives a lower bound on the number of variables n. The case $d = 1$ is rendered moot by the results of Green and Tao discussed in Chapter 1. The results of Chapter 3 resolve this case when $d = 2$, $R = 1$. From the point of view of the transference principle of Green and Tao in [12], which is essentially the one presented in Chapter 2, it seems reasonable that one should be able to take $K_0 = R(R+1)(d-1)2^{d-1}$. This says nothing of the positivity of the singular series however.

It is also worth noting that this is essentially a minor arc question. The treatise of the major arcs given in Chapter 2 is easily modified to above situation, and provides precisely the main term as stated above.

One more final note, in relation to the work on linear equations, this difficulty of this conjecture is on par with linear systems of complexity one. Essentially this boils down to the fact that we allow n to be taken large compared to R and d. In comparison, systems of R linear forms in $n > 2R+1$ variables have complexity one.

4.1.2 A Reasonable Approach

Here we shall discuss an approach for the case $R = 1$, $d > 2$ of Conjecture 4.1.1. First we shall point out a few partial results that are obtainable from the methods we have used. Take F to be a homogenous integral form of degree d. Associated to F is a d-ary symmetric linear form $F(x^{(1)}, ..., x^{(d)})$ over $(\mathbb{C}^n)^d$. Thus we recover our original form when we restrict ourself to the diagonal, which is of course a copy of $\mathbb{C}^n$. If there exists a splitting of the variables $x^{(1)} = (y, 0)$ and $x^{(2)} = (0, z)$ such that $(y, z) \in \mathbb{C}^n$ with $codim(\mathcal{L}$ large, where $\mathcal{L} = \{((y, 0), (0, z), ..., x^{(d)}) : F(((y, 0), (0, z), ..., x^{(d)}) = 0)\}$ (dependent only on d), then the methods in section 3.2.1 can provide an appropriate asymptotic.

The method applied in section 3.2.2 is not directly generalizable to higher degree polynomials, as the notion rank loses meaning. However it does provide a framework in which to approach such a generalization. The work of Schmidt [18] may prove to be quite useful here. His variant of Birch's method provides a more thorough treatment of systems of forms which are highly singular. His is approach is to decompose a form as

$$Q = R_1 S_1 + ... R_m S_m, \tag{4.3}$$

where R_i and S_i are forms of positive degree. The minimal value of m provides a natural generalization of the rank condition above. Moreover, over $\mathbb{C}$ this notion is essentially equivalent to condition of Birch.

The goal is then to modify the decomposition which appears in section 3.2.2 to the extent that when the off diagonal analogue fails, then one may apply a similar mean value type estimate for the 'good' parts of the form (those with a large Schmidt condition).

4.2 Final Remarks

This section brings our presentation to a close, and the author would like to take this time to thank the reader.

Bibliography

[1] T. Tao B. Green. The Möbius function is strongly orthogonal to nilsequences. *Annals of Math.*, to appear.

[2] T. Ziegler B. Green, T. Tao. An inverse theorem for the Gowers' $u^{s+1}[n]$-norm. *Annals of Math.*, to appear.

[3] B.J. Birch. Forms in many variables. *Proc. Roy. Soc. Ser. A*, 265:245–263, 1962.

[4] J. W. S. Cassels. *Rational Quadratic Forms.* Dover, 2008.

[5] C. Yıldırım D. Goldston. Higher correlations of divisor sums related to primes iii: small gaps between primes. *Proc. London Math. Soc.*, 95:653–686, 2003.

[6] H. Davenport. Cubic forms in 32 variables. *Phil. Trans. A*, 251:193–232, 1975.

[7] Te Riele Deshouillers, Effinger and Zinoviev. A complete Vinogradov 3-primes theorem under the Riemann hypothesis. *Electronic Research Announcements of the American Mathematical Society*, 3 (15):99–104, 2008.

[8] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi's theorem for commuting transformations. *J. d'Analyse Math.*, 34:275–291, 1978.

[9] D. Goldstond and C. Yıldırım. Higher correlations of divisor sums related to primes i: triple correlations. *Integers: Electronic Journal of Combinatorial Number theory*, 3:1–66, 2003.

[10] T. Gowers. Decompositions, approximate structure, transference, and the Hahn-Banach theorem. *Bull. Lon. Math. Soc.*, 42 (4):573–606, 2010.

[11] W.T. Gowers. Hypergraph regularity and the multidimensional szemerédi theorem. *Annals of Math.*, 166/3:897–946, 2007.

[12] B.J. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of Math.*, 167:481–547, 2008.

[13] B.J. Green and T. Tao. Linear equations in the primes. *Annals of Math.*, 171-3:1753–1850, 2010.

[14] L.K. Hua. *Additive Theory of Prime Numbers.* Translations of Mathematical Monographs, 13, American Mathematical Society, Providence, R.I., 1965.

[15] J. Liu. Integral points on quadrics with prime coordinates. *Monatsh Math*, 164:439–465, 2011.

[16] M. B. Nathonson. *Additive Number Theory: The Classical Bases.* Springer, 1996.

[17] M. Tulsiani O. Reingold, L. Trevisan and S. Vadham. Dense subsets of pseudorandom sets. *Electronic Colloquium of Computational Complexity*, pages Report TR08–045, 2008.

[18] W. Schmidt. The density of integer points on homogeneous varieties. *Acta Math.*, 154:243–296, 1985.

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

[20] T. Tao. The Gaussian primes contain arbitrarily shaped constellations. *Journal d'Analyse Mathmatique*, 99 (1):109–176, 2006.

[21] T. Tao and T. Ziegler. The primes contain arbitrarily long polynomial progressions. *Acta Math.*, 201:213–305, 2008.

[22] I.M. Vinogradov. *The method of trigonometrical sums in the theory of numbers*. London: Interscience Publishers, Translated, revised, and annotated by K. F. Roth and A. Davenport.