Canad. J. Math. Vol. 72 (3), 2020 pp. 805–833
http://dx.doi.org/10.4153/S0008414X19000026

Zeroes of Polynomials With Prime Inputs and Schmidt's h-invariant

Stanley Yao Xiao and Shuntaro Yamagishi

Abstract. In this paper we show that a polynomial equation admits infinitely many prime-tuple solutions, assuming only that the equation satisfies suitable local conditions and the polynomial is sufficiently non-degenerate algebraically. Our notion of algebraic non-degeneracy is related to the h-invariant introduced by W. M. Schmidt. Our results prove a conjecture by B. Cook and Á. Magyar for hypersurfaces of degree 3.

1 Introduction

Solving systems of integral polynomial equations in integers is among the oldest, persistently interesting problems in number theory. It is understood, especially in the context of the Hardy–Littlewood circle method, that systems tend to become easier to solve when the number of variables involved increases. For instance, it is not known whether the equation $x^2 + 1 = p$, where x varies in the integers and p varies among the primes, has infinitely many solutions, but the corresponding 3-variable equation $x^2 + y^2 = p$ was solved by Fermat using elementary means over three centuries ago. One can then ask whether it is possible to interpolate between these situations. That is, given a system of polynomial equations that is solvable in the integers, one can ask whether the system remains solvable when some of the variables are restricted to a thin subset of integers. One particular natural subset is the set of prime numbers. Indeed, many interesting problems involving prime numbers may be phrased in such a manner. For example, the existence of infinitely many solutions to the equation $x - y = 2$ with x, y restricted to primes is precisely the twin prime conjecture.

B. Cook and Á. Magyar broke new ground by applying the Hardy–Littlewood circle method to show, in great generality, that systems of polynomial equations in many variables can be solved when all of the inputs are prime numbers [2]. The key hypothesis they require is that the so-called Birch singular locus must be sufficiently small. For $\mathbf{f} = \{f_1, \dots, f_{r_d}\} \subseteq \mathbb{Q}[x_1, \dots, x_n]$ a system of forms (homogeneous polynomials) of degree d, we define the *Birch singular locus* $V_{\mathbf{f}}^*$ to be the affine variety in $\mathbb{A}_{\mathbb{C}}^n$ given by

$$V_{\mathbf{f}}^* = \left\{ \mathbf{x} \in \mathbb{C}^n : \operatorname{rank} \left(\frac{\partial f_r(\mathbf{x})}{\partial x_j} \right)_{\substack{1 \le r \le r_d \\ 1 \le j \le n}} < r_d \right\},$$

and let the *Birch rank* be $\mathcal{B}(\mathbf{f}) = n - \dim V_{\mathbf{f}}^*$.

Received by the editors June 7, 2018; revised January 1, 2019.
Published online on Cambridge Core February 7, 2019.
AMS subject classification: **11D72**, 11P32.
Keywords: Circle method, h-invariant, Hardy–Littlewood, prime numbers.

The Birch rank is an important invariant that arose in [1]. W. M. Schmidt introduced a different invariant, now called Schmidt's h-invariant, for systems of polynomials [4]. Cook and Magyar conjectured [2, p. 736] that their main theorem ought to hold assuming the largeness of the h-invariant instead of the Birch rank (see (2.1)).

In this paper, we give a partial solution to the conjecture of Cook and Magyar. We establish the conjecture for hypersurfaces with an additional assumption. However, our assumption is redundant for cubic polynomials; therefore, we establish the conjecture unconditionally in this case. Given a form $f \in \mathbb{Q}[x_1, \dots, x_n]$ of degree at least 2, we define the *h-invariant* $h(f)$ of f to be the least positive integer h such that f can be written identically as

$$f = U_1 V_1 + \cdots + U_h V_h, \tag{1.1}$$

where U_i and V_i are forms in $\mathbb{Q}[x_1, \dots, x_n]$ of degree at least 1 $(1 \le i \le h)$. We then define the quantity

$$h^\star(f) = \max(|\{U_i : \deg U_i = 1\}|),$$

where the maximum is over all representations of the shape (1.1). In other words, $h^\star(f)$ is the maximum number of linear forms involved in the representation of f as a sum of $h = h(f)$ products of rational forms. Clearly, we have $h^\star(f) \le h(f)$. For a degree d polynomial $b(\mathbf{x}) \in \mathbb{Q}[x_1, \dots, x_n]$, we define $h(b) = h(f)$ and $h^\star(b) = h^\star(f)$, where $f(\mathbf{x})$ is the degree d portion of $b(\mathbf{x})$. We note that any polynomial $b(\mathbf{x})$ of degree 2 or degree 3 satisfies

$$h(b) = h^\star(b).$$

We define the quantity $\mathcal{M}_b(N) = \sum_{\mathbf{x} \in [0,N]^n \cap \mathbb{Z}^n} \delta_b(\mathbf{x})$, where

$$\delta_b(\mathbf{x}) = \begin{cases} \prod_{1 \le i \le n} \log p_i & \text{if } x_i = p_i^{t_i},\ p_i \text{ is prime},\ t_i \in \mathbb{N}\ (1 \le i \le n),\ b(\mathbf{x}) = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let Λ be the von Mangoldt function, where $\Lambda(x)$ is $\log p$ if x is a power of a prime p and 0 otherwise. We use the notation $e(x)$ to denote $e^{2\pi i x}$. We define

$$T(b; \alpha) = \sum_{\mathbf{x} \in [0,N]^n \cap \mathbb{Z}^n} \Lambda(\mathbf{x})\, e(\alpha \cdot b(\mathbf{x})), \tag{1.2}$$

where $\Lambda(\mathbf{x}) = \Lambda(x_1) \cdots \Lambda(x_n)$ for $\mathbf{x} = (x_1, \dots, x_n) \in (\mathbb{Z}_{\ge 0})^n$. By the orthogonality relation, we have

$$\mathcal{M}_b(N) = \sum_{\mathbf{x} \in [0,N]^n \cap \mathbb{Z}^n} \delta_b(\mathbf{x}) = \int_0^1 T(b; \alpha)\, d\alpha. \tag{1.3}$$

We obtain the following theorem by estimating the integral in (1.3).

Theorem 1.1 *Let $b(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$ be a polynomial of degree d. Then there exists a positive number A_d dependent only on d such that if $h^\star(b) > A_d$, then there exist $c > 0$ and C_b such that*

$$\mathcal{M}_b(N) = C_b N^{n-d} + O\Big(\frac{N^{n-d}}{(\log N)^c}\Big).$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

In fact, we prove that $C_b > 0$, provided that the equation $b(\mathbf{x}) = 0$ has a non-singular solution in $\mathbb{Z}_p^\times$ (the units of p-adic integers) for every prime p and the equation $f(\mathbf{x}) = 0$, where $f(\mathbf{x})$ is the degree d portion of $b(\mathbf{x})$, has a non-singular real zero in the interior of $\mathfrak{B}_0 = [0,1]^n$.

The following result is an immediate consequence of Theorem 1.1, which replaces the assumption of large Birch rank in [2, Theorem 1] with large h-invariant for cubic polynomials. Note it does not require much work to achieve this when $\deg f = 2$, because we know $\mathcal{B}(f) \ll h(f) \ll \mathcal{B}(f)$.

Corollary 1.2 *Let $b(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$ be a cubic polynomial. Then there exists a positive number A_3 such that if $h(b) > A_3$, then there exist $c > 0$ and C_b such that*

$$\mathcal{M}_b(N) = C_b N^{n-3} + O\Big(\frac{N^{n-3}}{(\log N)^c}\Big).$$

We establish Theorem 1.1 in a similar manner to [2], but we shall make use of the fact that the representation (1.1) has sufficient linear terms. We also modify the method in [2] to better suit our purposes, so that it is in terms of the h-invariant instead of the Birch rank.

Despite Theorem 1.1 and Corollary 1.2 being our primary goals in this paper, it is necessary for us to work over a system of polynomials at times. Indeed, our strategy is to decompose a polynomial into a sum of elements in a suitable system of polynomials, and then use methods that apply to systems to deduce results of a single polynomial.

The organization of the rest of the paper is as follows. In Section 2, we prove some basic properties of the h-invariant. A sufficiently large $h^\star(b)$ allows us to massage our polynomial $b(\mathbf{x})$ into something amenable to the circle method through a process called *regularization*. We collect results related to the regularization process in Section 3. In Section 4, we obtain results from [4] based on Weyl differencing in terms of polynomials instead of forms, as in [4]. We chose to present the details in Section 4 to make certain dependencies of the constants explicit, because it plays an important role in our estimates. We then obtain the minor arc estimates in Section 5, and the major arc estimates in Section 6.

2 Properties of the h-invariant

Let $\mathbf{f} = \{f_1, \dots, f_{r_d}\} \subseteq \mathbb{Q}[x_1, \dots, x_n]$ be a system of forms of degree $d > 1$. We generalize the definition of h-invariant for a single form, and define the h-invariant of $\mathbf{f}$ by

$$h(\mathbf{f}) = \min_{\boldsymbol{\mu} \in \mathbb{Q}^{r_d} \setminus \{\mathbf{0}\}} h(\mu_1 f_1 + \cdots + \mu_{r_d} f_{r_d}).$$

Given an invertible linear transformation $T \in \mathrm{GL}_n(\mathbb{Q})$, let $\mathbf{f} \circ T = \{f_1 \circ T, \dots, f_{r_d} \circ T\}$. It follows from the definition of the h-invariant that $h(\mathbf{f}) = h(\mathbf{f} \circ T)$. Let $\mathbf{b} = (b_1, \dots, b_{r_d}) \subseteq \mathbb{Q}[x_1, \dots, x_n]$ be a system of degree d polynomials. We let f_r be the degree d portion of b_r $(1 \le r \le r_d)$, and define

$$h(\mathbf{b}) = h(\{f_r : 1 \le r \le r_d\}).$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

It is known that a large Birch rank implies a large h-invariant, since we have

$$h(\mathbf{f}) \geq 2^{1-d}\mathcal{B}(\mathbf{f}) \tag{2.1}$$

by [4, Lemma 16.1, (10.3), (17.1)]; however, at the present time the authors do not know whether there exists an infinite family of varieties with bounded Birch rank but unbounded h-invariant.

We prove two basic lemmas regarding the properties of the h-invariant in this section. Let $f \in \mathbb{Q}[x_1, \ldots, x_n]$ be a form of degree d. For $1 \leq i \leq n$, let $f|_{x_i=0} = f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]$, which is either identically 0 or a form of degree d. Let $h(f) = 0$ if f is identically 0. We prove the following simple lemma.

Lemma 2.1 *Let $f \in \mathbb{Q}[x_1, \ldots, x_n]$ be a form of degree $d > 1$. Then for any $1 \leq i \leq n$, we have $h(f) - 1 \leq h(f|_{x_i=0}) \leq h(f)$.*

Proof Without loss of generality, we consider the case $i = 1$. Let us write

$$f(x_1, \ldots, x_n) = x_1 g(x_1, \ldots, x_n) + f(0, x_2, \ldots, x_n). \tag{2.2}$$

Clearly, $g(\mathbf{x})$ is either identically 0 or a form of degree $d - 1$. Let $h = h(f)$ and $h' = h(f|_{x_1=0})$. By the definition of h-invariant, we can find rational forms $U_{j'}$, $V_{j'}$ $(1 \leq j' \leq h')$ of positive degree that satisfy

$$f(0, x_2, \ldots, x_n) = U_1 V_1 + \cdots + U_{h'} V_{h'}.$$

Note if $h' = 0$, we assume the right-hand side to be identically 0. By substituting the above equation into (2.2), we obtain

$$f = x_1 g + U_1 V_1 + \cdots + U_{h'} V_{h'}.$$

Because $g(\mathbf{x})$ is either identically 0 or a form of degree $d-1$, it follows that $h \leq 1 + h'$.

For the other inequality, let u_j, v_j $(1 \leq j \leq h)$ be rational forms of positive degree that satisfy

$$f = u_1 v_1 + \cdots + u_h v_h. \tag{2.3}$$

By substituting $x_1 = 0$ into each form on both sides of the equation, it is clear that we obtain $h' \leq h$. This completes the proof of the lemma. We add a remark that in the special case when f satisfies $f = x_1 v_1 + u_2 v_2 + \cdots + u_h v_h$, in other words when we have $u_1 = x_1$ in (2.3), we easily obtain $h' = h - 1$. ∎

The following is an immediate consequence of Lemma 2.1.

Lemma 2.2 *Let $\mathbf{f} = \{f_1, \ldots, f_{r_d}\} \subseteq \mathbb{Q}[x_1, \ldots, x_n]$ be a system of forms of degree $d > 1$. Suppose $h(\mathbf{f}) > 1$. Then for any $1 \leq i \leq n$, we have*

$$h(\mathbf{f}) - 1 \leq h(\mathbf{f}|_{x_i=0}) \leq h(\mathbf{f}),$$

where $\mathbf{f}|_{x_i=0} = \{f_1|_{x_i=0}, \ldots, f_{r_d}|_{x_i=0}\}$.

Let $f(\mathbf{x}) \in \mathbb{Q}[x_1, \ldots, x_n]$ be a form, and let $h = h(f)$ and $0 < M \leq h$. Suppose we have

$$f = u_1 V_1 + \cdots + u_M V_M + U_{M+1} V_{M+1} + \cdots + U_h V_h,$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

where each u_i is a linear rational form $(1 \le i \le M)$, and each $U_{i'}$ and V_j are rational forms of positive degree $(M+1 \le i' \le h, 1 \le j \le h)$. It can be easily verified that the linear forms $u_1, \dots, u_M$ are linearly independent over $\mathbb{Q}$. Then by considering the reduced row echelon form of the matrix formed by the coefficients of $u_1, \dots, u_M$ and relabeling the variables if necessary, we can suppose without loss of generality that

$$f = (x_1 + \ell_1)v_1 + \cdots + (x_M + \ell_M)v_M + u_{M+1}v_{M+1} + \cdots + u_h v_h, \tag{2.4}$$

where each ℓ_i is a linear form in $\mathbb{Q}[x_{M+1}, \dots, x_n]$ $(1 \le i \le M)$, and each $u_{i'}$ and v_j are rational forms of positive degree $(M+1 \le i' \le h, 1 \le j \le h)$. We then define $g_M \in \mathbb{Q}[x_1, \dots, x_n]$ in the following manner,

$$f(x_1, x_2, \dots, x_n) = g_M(x_1, \dots, x_n) + f(-\ell_1, \dots, -\ell_M, x_{M+1}, \dots, x_n). \tag{2.5}$$

We note that there is no ambiguity for defining the polynomial

$$f(-\ell_1, \dots, -\ell_M, x_{M+1}, \dots, x_n) \in \mathbb{Q}[x_{M+1}, \dots, x_n]$$

obtained by substitution, because each $\ell_i \in \mathbb{Q}[x_{M+1}, \dots, x_n]$ $(1 \le i \le M)$. It is also clear that $g(-\ell_1, \dots, -\ell_M, x_{M+1}, \dots, x_n) = 0$.

Lemma 2.3 *Let $1 \le M \le h$. Suppose a degree d form $f(\mathbf{x}) \in \mathbb{Q}[x_1, \dots, x_n]$ satisfies (2.4). Define $g_M(\mathbf{x})$ as in (2.5). Then we have*

$$h(g_M) \ge M \quad \text{and} \quad h(f(-\ell_1, -\ell_2, \dots, -\ell_M, x_{M+1}, \dots, x_n)) = h - M.$$

Proof Since the linear forms $(x_1 - \ell_1), \dots, (x_M - \ell_M)$ are linearly independent over $\mathbb{Q}$, we can find $A \in \mathrm{GL}_n(\mathbb{Q})$ such that

$$\begin{pmatrix} x_1 - \ell_1 \\ \vdots \\ x_M - \ell_M \\ x_{M+1} \\ \vdots \\ x_n \end{pmatrix} = A \circ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Let $\widetilde{f}(\mathbf{x}) = f(A \circ \mathbf{x})$. We then have $\widetilde{f}(A^{-1} \circ \mathbf{x}) = f(\mathbf{x})$, and also that $h(\widetilde{f}) = h(\widetilde{f} \circ A^{-1}) = h(f) = h$. Because $f(\mathbf{x})$ satisfies (2.4), it follows that $\widetilde{f}(\mathbf{x})$ satisfies $\widetilde{f} = x_1 V_1 + \cdots + x_M V_M + U_{M+1} V_{M+1} + \cdots + U_h V_h$, where each U_i and V_j are rational forms of positive degree $(M+1 \le i \le h, 1 \le j \le h)$.

Recall that each ℓ_i is a linear form in $\mathbb{Q}[x_{M+1}, \dots, x_n]$ $(1 \le i \le M)$. Clearly, we have

$$\begin{aligned} \widetilde{f}(0, \dots, 0, x_{M+1}, \dots, x_n) &= f(A \circ (0, \dots, 0, x_{M+1}, \dots, x_n)) \\ &= f(-\ell_1, -\ell_2, \dots, -\ell_M, x_{M+1}, \dots, x_n). \end{aligned}$$

Then we can deduce from Lemma 2.1 (see the remark at the end of the proof of Lemma 2.1) that

$$\begin{aligned} h(f(-\ell_1, -\ell_2, \dots, -\ell_M, x_{M+1}, \dots, x_n)) &= h(\widetilde{f}(0, \dots, 0, x_{M+1}, \dots, x_n)) \\ &= h - M. \end{aligned}$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

It then follows easily from the fact that $h(f) = h$, the definition of h-invariant, and (2.5) that $h(g_M) \geq M$, for otherwise we obtain a contradiction. ∎

3 Regularization Lemmas

In this section, we collect results from [2, 4] related to regular systems (see Definition 3.1) and the regularization process (Proposition 3.5), which played an important role in [2] to obtain the minor arc estimate. Throughout this section we use the following notation. Let $d, n > 1$, and let $\mathbf{f}$ be a system of forms in $\mathbb{Q}[x_1, \dots, x_n]$ of degree less than or equal to d. We let $\mathbf{f} = (\mathbf{f}^{(d)}, \dots, \mathbf{f}^{(1)})$, where $\mathbf{f}^{(i)}$ is the subsystem of all forms of degree i in $\mathbf{f}$ $(1 \leq i \leq d)$. We label the elements of $\mathbf{f}^{(i)}$ by $\mathbf{f}^{(i)} = \{f_1^{(i)}, \dots, f_{r_i}^{(i)}\}$, where $r_i = |\mathbf{f}^{(i)}|$, the number of elements in $\mathbf{f}^{(i)}$.

We shall call a system of polynomials regular if it has at most the expected number of integer solutions; we define this formally below.

Definition 3.1 Let $d > 1$. Let $\boldsymbol{\psi} = (\boldsymbol{\psi}^{(d)}, \dots, \boldsymbol{\psi}^{(1)})$ be a system of polynomials in $\mathbb{Q}[x_1, \dots, x_n]$, where $\boldsymbol{\psi}^{(i)}$ is the subsystem of all polynomials of degree i in $\boldsymbol{\psi}$ $(1 \leq i \leq d)$. We define $V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{Z})$ to be the set of solutions in $\mathbb{Z}^n$ of the equations $\psi_j^{(i)}(\mathbf{x}) = 0$ $(1 \leq i \leq d, 1 \leq j \leq |\boldsymbol{\psi}^{(i)}|)$ that we denote by $\boldsymbol{\psi}(\mathbf{x}) = \mathbf{0}$. Let $r_i = |\boldsymbol{\psi}^{(i)}|$ $(1 \leq i \leq d)$, and let $D_{\boldsymbol{\psi}} = \sum_{i=1}^{d} i r_i$. We say the system $\boldsymbol{\psi}$ is *regular* if $|V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{Z}) \cap [-N, N]^n| \ll N^{n - D_{\boldsymbol{\psi}}}$.

Similarly as above we also define $V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{R})$ to be the set of solutions in $\mathbb{R}^n$ of the equations $\boldsymbol{\psi}(\mathbf{x}) = \mathbf{0}$.

The following is one of the main results of [4] that provides a sufficient condition for a system of polynomials to be regular.

Theorem 3.2 (Schmidt [4]) *Let $d > 1$. Let $\boldsymbol{\psi} = (\boldsymbol{\psi}^{(d)}, \dots, \boldsymbol{\psi}^{(2)})$ be a system of rational polynomials with notation as in Definition 3.1, and also let $\mathbf{f}^{(i)}$ be the system of degree i portion of the polynomials $\boldsymbol{\psi}^{(i)}$ $(2 \leq i \leq d)$. We let $r_i = |\boldsymbol{\psi}^{(i)}| = |\mathbf{f}^{(i)}|$ $(2 \leq i \leq d)$, and $R_{\boldsymbol{\psi}} = \sum_{i=2}^{d} r_i$. If we have*

$$h(\mathbf{f}^{(i)}) \geq d\, 2^{4i} (i!) r_i R_{\boldsymbol{\psi}} \quad (2 \leq i \leq d),$$

then the system $\boldsymbol{\psi}$ is regular.

Let

$$\rho_{d,i}(t) = d\, 2^{4i} (i!) t^2 \quad (2 \leq i \leq d), \tag{3.1}$$

so that for each $2 \leq i \leq d$, we have that $\rho_{d,i}(t)$ is an increasing function, and $\rho_{d,i}(R_{\boldsymbol{\psi}}) \geq d\, 2^{4i} (i!) r_i R_{\boldsymbol{\psi}}$.

Note Theorem 3.2 is regarding a system of polynomials that does not contain any linear polynomials. We prove Corollary 3.3 for systems that contain linear forms as well.

Corollary 3.3 ([2, Corollary 3]) *Let $d > 1$. Let $\boldsymbol{\psi} = (\boldsymbol{\psi}^{(d)}, \dots, \boldsymbol{\psi}^{(1)})$ be a system of rational polynomials with notation as in Definition 3.1. Suppose $\boldsymbol{\psi}^{(1)}$ only contains linear forms and that they are linearly independent over $\mathbb{Q}$. We also let $\mathbf{f}^{(i)}$ be the system*

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

of degree i portion of the polynomials $\boldsymbol{\psi}^{(i)}$ $(1 \le i \le d)$. We let $r_i = |\boldsymbol{\psi}^{(i)}| = |\mathbf{f}^{(i)}|$ $(1 \le i \le d)$, and $R_{\boldsymbol{\psi}} = \sum_{i=1}^{d} r_i$. For each $2 \le i \le d$, let $\rho_{d,i}(\cdot)$ be as in (3.1). If we have

$$h(\mathbf{f}^{(i)}) \ge \rho_{d,i}(R_{\boldsymbol{\psi}} - r_1) + r_1 \quad (2 \le i \le d),$$

then the system $\boldsymbol{\psi}$ is regular.

Proof We have $\boldsymbol{\psi}^{(1)} = \mathbf{f}^{(1)} = \{f_1^{(1)}, \dots, f_{r_1}^{(1)}\}$. Let

$$f_i^{(1)} = a_{i1}x_1 + \cdots + a_{in}x_n \quad (1 \le i \le r_1),$$

and denote the coefficient matrix of these linear forms by

$$A = [a_{ij}]_{\substack{1 \le i \le r_1 \\ 1 \le j \le n}}.$$

Let $\mathbf{e}_j$ be the j-th standard basis of $\mathbb{R}^n$ $(1 \le j \le n)$. Since the linear forms $f_1^{(1)}, \dots, f_{r_1}^{(1)}$ are linearly independent over $\mathbb{Q}$, we can find an invertible linear transformation $T \in \mathrm{GL}_n(\mathbb{Q})$, where every entry of the matrix is in $\mathbb{Z}$, such that $(f_i^{(1)} \circ T^{-1})(\mathbf{x}) = m_{n-i+1}x_{n-i+1}$, where $m_{n-i+1} \in \mathbb{Q}\backslash\{0\}$ $(1 \le i \le r_1)$. For simplicity, let us denote $\mathbf{x}' = (x_{n-r_1+1}, \dots, x_n)$. Let

$$Y = V_{\mathbf{f}^{(1)},\mathbf{0}}(\mathbb{R}) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{f}^{(1)}(\mathbf{x}) = \mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^n : A \circ \mathbf{x} = \mathbf{0}\} = \mathrm{Ker}(A),$$

which is a subspace of codimension r_1. Since $T(Y) = \mathrm{Ker}(A \circ T^{-1})$, it follows from our choice of $T \in \mathrm{GL}_n(\mathbb{Q})$ that $T(Y) = \mathbb{R}\mathbf{e}_1 + \cdots + \mathbb{R}\mathbf{e}_{n-r_1}$. We also know there exist $c', C' > 0$ such that

$$[-c'N, c'N]^n \subseteq T([-N, N]^n) \subseteq [-C'N, C'N]^n.$$

Define $\boldsymbol{\psi}' = (\boldsymbol{\psi}'^{(d)}, \dots, \boldsymbol{\psi}'^{(1)}) = \boldsymbol{\psi} \circ T^{-1}$, and let $\mathbf{f}'^{(i)}$ be the system of degree i portion of the polynomials $\boldsymbol{\psi}'^{(i)}$ $(1 \le i \le d)$. We then have $\mathbf{f}'^{(i)} = \mathbf{f}^{(i)} \circ T^{-1}$. We can also verify that $V_{\boldsymbol{\psi}',\mathbf{0}}(\mathbb{R}) = T(V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{R}))$. Therefore, we obtain

$$T(V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{R}) \cap [-N, N]^n) \subseteq V_{\boldsymbol{\psi}',\mathbf{0}}(\mathbb{R}) \cap [-C'N, C'N]^n,$$

and since every entry of the matrix $T \in \mathrm{GL}_n(\mathbb{Q})$ is in $\mathbb{Z}$, it follows that

$$\left| V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{Z}) \cap [-N, N]^n \right| \le \left| V_{\boldsymbol{\psi}',\mathbf{0}}(\mathbb{Z}) \cap [-C'N, C'N]^n \right|. \tag{3.2}$$

Let $\boldsymbol{\psi}'' = (\boldsymbol{\psi}'^{(d)}|_{\mathbf{x}'=\mathbf{0}}, \dots, \boldsymbol{\psi}'^{(2)}|_{\mathbf{x}'=\mathbf{0}})$. Since $\boldsymbol{\psi}'^{(1)} = \mathbf{0}$ is equivalent to $\mathbf{x}' = \mathbf{0}$, we have

$$\left| V_{\boldsymbol{\psi}',\mathbf{0}}(\mathbb{Z}) \cap [-C'N, C'N]^n \right| = \left| V_{\boldsymbol{\psi}'',\mathbf{0}}(\mathbb{Z}) \cap [-C'N, C'N]^{n-r_1} \right|. \tag{3.3}$$

Since the degree i portion of $\boldsymbol{\psi}'^{(i)}|_{\mathbf{x}'=\mathbf{0}}$ is $\mathbf{f}'^{(i)}|_{\mathbf{x}'=\mathbf{0}}$ for each $2 \le i \le d$, we have by Lemma 2.2 that

$$h(\mathbf{f}'^{(i)}|_{\mathbf{x}'=\mathbf{0}}) \ge h(\mathbf{f}'^{(i)}) - r_1 = h(\mathbf{f}^{(i)}) - r_1 \ge \rho_{d,i}(R_{\boldsymbol{\psi}} - r_1).$$

Thus, it follows by Theorem 3.2 that

$$\left| V_{\boldsymbol{\psi}'',\mathbf{0}}(\mathbb{Z}) \cap [-C'N, C'N]^{n-r_1} \right| \ll N^{(n-r_1)-\sum_{i=2}^{d} ir_i}. \tag{3.4}$$

Therefore, we obtain from (3.2), (3.3), and (3.4) that

$$\left| V_{\boldsymbol{\psi},\mathbf{0}}(\mathbb{Z}) \cap [-N, N]^n \right| \ll N^{n-\sum_{i=1}^{d} ir_i}. \qquad \blacksquare$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Given $\mathbf{g} = \{g_1, \dots, g_{r_d}\} \subseteq \mathbb{Q}[x_1, \dots, x_n]$, a system of forms of degree d, and a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, we let $\overline{\mathbf{g}}$ be the system obtained by removing all the forms of $\mathbf{g}$ that depend only on the $\mathbf{z}$ variables. Clearly, if we have the trivial partition $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, where $\mathbf{z} = \varnothing$, then $\overline{\mathbf{g}} = \mathbf{g}$. For a form $g(\mathbf{x})$ over $\mathbb{Q}$, we define $h(g; \mathbf{z})$ to be the smallest number h_0 such that $g(\mathbf{x})$ can be expressed as $g(\mathbf{x}) = g(\mathbf{y}, \mathbf{z}) = \sum_{i=1}^{h_0} u_i v_i + w_0(\mathbf{z})$, where u_i, v_i are rational forms of positive degree $(1 \le i \le h_0)$, and $w_0(\mathbf{z})$ is a rational form only in the $\mathbf{z}$ variables. We also define $h(\mathbf{g}; \mathbf{z})$ to be

$$h(\mathbf{g}; \mathbf{z}) = \min_{\boldsymbol{\lambda} \in \mathbb{Q}^{r_d} \setminus \{\mathbf{0}\}} h(\lambda_1 g_1 + \cdots + \lambda_{r_d} g_{r_d}; \mathbf{z}).$$

If we have the trivial partition, then clearly we have $h(\mathbf{g}; \varnothing) = h(\mathbf{g})$. We have the following lemma.

Lemma 3.4 (Lemma 2, [2]) *Let $\mathbf{g} = \{g_1, \dots, g_{r_d}\} \subseteq \mathbb{Q}[x_1, \dots, x_n]$ be a system of forms of degree d, and suppose we have a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$. Let $\mathbf{y}'$ be a distinct set of variables with the same number of variables as $\mathbf{y}$. Then we have $h(\mathbf{g}(\mathbf{y}, \mathbf{z}), \mathbf{g}(\mathbf{y}', \mathbf{z}); \mathbf{z}) = h(\mathbf{g}; \mathbf{z})$.*

Given a system of forms, which may not be regular, we want to obtain a regular system in a controlled manner. The process in the following proposition is referred to as the *regularization* of systems in [2], and it is a crucial component of their method. Given a system of rational forms $\mathbf{f}$, via the regularization process we obtain another system $\mathcal{R}(\mathbf{f})$ that is regular, the number of forms it contains is controlled, and its level sets partition the level sets of $\mathbf{f}$. We remark that condition (iii) of Proposition 3.5, with a suitable choice of $\mathcal{F}$, together with Corollary 3.3 implies that the resulting system is regular.

Proposition 3.5 (Propositions 1 and 1′ [2]) *Let $d > 1$, and let $\mathcal{F}$ be any collection of non-decreasing functions $\mathcal{F}_i \colon \mathbb{Z}_{\ge 0} \to \mathbb{Z}_{\ge 0}$ $(2 \le i \le d)$. For a collection of non-negative integers $r_1, \dots, r_d$, there exist constants*

$$C_1(r_1, \dots, r_d, \mathcal{F}), \dots, C_d(r_1, \dots, r_d, \mathcal{F})$$

such that the following holds.

Given a system of integral forms $\mathbf{f} = (\mathbf{f}^{(d)}, \dots, \mathbf{f}^{(1)}) \subseteq \mathbb{Z}[x_1, \dots, x_n]$, where each $\mathbf{f}^{(i)}$ is a system of r_i forms of degree i $(1 \le i \le d)$, and a partition of variables $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, there exists a system of forms $\mathcal{R}(\mathbf{f}) = (\mathbf{a}^{(d)}, \dots, \mathbf{a}^{(1)})$ satisfying the following. Let $r_i' = |\mathbf{a}^{(i)}|$ $(1 \le i \le d)$, and $R' = r_1' + \cdots + r_d'$.

(i) *Each form of the system $\mathbf{f}$ can be written as a rational polynomial expression in the forms of the system $\mathcal{R}(\mathbf{f})$. In particular, the level sets of $\mathcal{R}(\mathbf{f})$ partition those of $\mathbf{f}$.*

(ii) *For each $1 \le i \le d$, r_i' is at most $C_i(r_1, \dots, r_d, \mathcal{F})$.*

(iii) *The subsystem $(\mathbf{a}^{(d)}, \dots, \mathbf{a}^{(2)})$ satisfies $h(\mathbf{a}^{(i)}) \ge \mathcal{F}_i(R')$ for each $2 \le i \le d$. Moreover, the linear forms of subsystem $\mathbf{a}^{(1)}$ are linearly independent over $\mathbb{Q}$.*

(iv) *Let $\overline{\mathbf{a}}^{(i)}$ be the system obtained by removing from $\mathbf{a}^{(i)}$ all forms that depend only on the $\mathbf{z}$ variables $(2 \le i \le d)$. Then the subsystem $(\overline{\mathbf{a}}^{(d)}, \dots, \overline{\mathbf{a}}^{(2)})$ satisfies $h(\overline{\mathbf{a}}^{(i)}; \mathbf{z}) \ge \mathcal{F}_i(R')$ for each $2 \le i \le d$.*

We will be utilizing this proposition in Section 5 to obtain the minor arc estimate.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

4 Technical Estimates

In this section, we provide results from [4] related to Weyl differencing that are necessary in obtaining estimates for the singular series in Section 6.1. The work here is similar to that of [4], which is in terms of forms instead of polynomials as in this section. It is stated in [4] with some explanation that similar results for polynomials also follow, but the details are not shown. We chose to present the necessary details in order to make explicit certain dependencies of the constants that are crucial in our estimates. Let us denote $\mathfrak{B}_1 = [-1,1]^n$. We shall refer to $\mathfrak{B} \subseteq \mathbb{R}^n$ as a box, if $\mathfrak{B}$ is of the form

$$\mathfrak{B} = I_1 \times \cdots \times I_n,$$

where each I_j is a closed or open or half open/closed interval $(1 \le j \le n)$. Given a function $G(\mathbf{x})$, we define

$$\Gamma_{d,G}(\mathbf{x}_1, \ldots, \mathbf{x}_d) = \sum_{t_1=0}^{1} \cdots \sum_{t_d=0}^{1} (-1)^{t_1+\cdots+t_d}\, G(t_1\mathbf{x}_1 + \cdots + t_d\mathbf{x}_d).$$

Then it follows [4, §11] that $\Gamma_{d,G}$ is symmetric in its d arguments, and that

$$\Gamma_{d,G}(\mathbf{x}_1, \ldots, \mathbf{x}_{d-1}, \mathbf{0}) = 0.$$

It is clear from the definition that $\Gamma_{d,G} + \Gamma_{d,G'} = \Gamma_{d,G+G'}$. We also have that if G is a form of degree j, where $d > j > 0$, then $\Gamma_{d,G} = 0$ [4, Lemma 11.2].

For $\alpha \in \mathbb{R}$, let $\|\alpha\|$ denote the distance from α to the closest integer. Given $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, we let $\|\boldsymbol{\alpha}\| = \max_{1\le i\le n} \|\alpha_i\|$.

Lemma 4.1 ([4, Lemma 13.1]) *Suppose*

$$G(\mathbf{x}) = G^{(0)} + G^{(1)}(\mathbf{x}) + \cdots + G^{(d)}(\mathbf{x}),$$

where $G^{(j)}$ is a form of degree j with real coefficients $(1 \le j \le d)$, and $G^{(0)} \in \mathbb{R}$. Let $\mathfrak{B}$ be a box with sides ≤ 1, let $P > 1$, and put

$$S' = S'(G, P, \mathfrak{B}) = \sum_{\mathbf{x} \in P\mathfrak{B} \cap \mathbb{Z}^n} e(G(\mathbf{x})).$$

Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis vectors of $\mathbb{R}^n$. Then for any $\varepsilon > 0$, we have

$$|S'|^{2^{d-1}} \ll P^{(2^{d-1}-d)n+\varepsilon} \sum \Big(\prod_{i=1}^{n} \min\big(P, \|\Gamma_{d,G^{(d)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{d-1}, \mathbf{e}_i)\|^{-1}\big) \Big),$$

where the sum is over $(d-1)$-tuples of integer points $\mathbf{x}_1, \ldots, \mathbf{x}_{d-1}$ in $P\mathfrak{B}_1$, and the implicit constant in $\ll$ depends only on n, d, and ε.

Lemma 4.2 ([4, Lemma 14.2]) *Make all the assumptions of Lemma 4.1. Suppose further that $|S'| \ge P^{n-Q}$, where $Q > 0$. Let $0 < \eta \le 1$. Then the number $N(\eta)$ of integral $(d-1)$-tuples $\mathbf{x}_1, \ldots, \mathbf{x}_{d-1} \in P^{\eta}\mathfrak{B}_1$ with*

$$\|\Gamma_{d,G^{(d)}}(\mathbf{x}_1, \ldots, \mathbf{x}_{d-1}, \mathbf{e}_i)\| < P^{-d+(d-1)\eta} \quad (i = 1, \ldots, n)$$

satisfies $N(\eta) \gg P^{n(d-1)\eta - 2^{d-1}Q - \varepsilon}$, where the implicit constant in $\gg$ depends only on n, d, η, and ε.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Let $\boldsymbol{\psi} = \{\psi_1, \dots, \psi_{r_d}\}$ be a system of rational polynomials of degree d. Let $\mathbf{f} = \{f_1, \dots, f_{r_d}\}$ be the system of forms, where f_i is the degree d portion of ψ_i $(1 \le i \le r_d)$. For the rest of this section, we assume $\mathbf{f}$ to be a system of integral forms. We define the following exponential sum associated with $\boldsymbol{\psi}$ and $\mathfrak{B}$,

$$S(\boldsymbol{\alpha}) = S(\boldsymbol{\psi}, \mathfrak{B}; \boldsymbol{\alpha}) = \sum_{\mathbf{x} \in P\mathfrak{B} \cap \mathbb{Z}^n} e(\boldsymbol{\alpha} \cdot \boldsymbol{\psi}(\mathbf{x})). \tag{4.1}$$

Let $\mathbf{e}_1, \dots, \mathbf{e}_n$ be the standard basis vectors of $\mathbb{C}^n$. We define $\mathfrak{M}_d = \mathfrak{M}_d(\mathbf{f})$ to be the set of $(d-1)$-tuples $(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}) \in (\mathbb{C}^n)^{d-1}$ for which the matrix

$$[m_{ij}] = [\Gamma_{d,f_j}(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}, \mathbf{e}_i)] \quad (1 \le j \le r_d, 1 \le i \le n)$$

has rank strictly less than r_d. For $R > 0$, we let $z_R(\mathfrak{M}_d)$ be the number of integer points $(\mathbf{x}_1, \dots, \mathbf{x}_{d-1})$ on $\mathfrak{M}_d$ such that

$$\max_{1 \le i \le d-1} \max_{1 \le j \le n} |x_{ij}| \le R,$$

where $\mathbf{x}_i = (x_{i1}, \dots, x_{in})$ $(1 \le i \le d-1)$.

Let $P > 1$, $Q > 0$, and $\varepsilon > 0$ be given, and suppose that $d > 1$. We then have the following.

Lemma 4.3 ([4, Lemma 15.1]) *Given a box $\mathfrak{B}$ with sides ≤ 1, define the sum $S(\boldsymbol{\alpha})$ associated with $\boldsymbol{\psi}$ and $\mathfrak{B}$ as in (4.1). Given $0 < \eta \le 1$, one of the following three alternatives must hold.*

(i) $|S(\boldsymbol{\alpha})| \le P^{n-Q}$.

(ii) *There exists $n_0 \in \mathbb{N}$ such that $n_0 \ll P^{r_d(d-1)\eta}$ and $\|n_0\boldsymbol{\alpha}\| \ll P^{-d+r_d(d-1)\eta}$.*

(iii) $z_R(\mathfrak{M}_d) \gg R^{(d-1)n-2^{d-1}(Q/\eta)-\varepsilon}$ *holds with $R = P^\eta$.*

All implicit constants depend at most on $n, d, r_d, \eta, \varepsilon$, and $\mathbf{f}$.

Proof Take $\boldsymbol{\alpha} \in \mathbb{R}^{r_d}$. Let $\boldsymbol{\alpha} \cdot \boldsymbol{\psi}(\mathbf{x}) = G^{(0)} + G^{(1)}(\mathbf{x}) + \cdots + G^{(d)}(\mathbf{x})$, where $G^{(j)}$ is a form of degree j $(1 \le j \le d)$, and $G^{(0)} \in \mathbb{R}$. Suppose (i) fails. Then we may apply Lemma 4.2. The number $N(\eta)$ of integral $(d-1)$-tuples $\mathbf{x}_1, \dots, \mathbf{x}_{d-1}$ in $P^\eta\mathfrak{B}_1$ with

$$\|\Gamma_{d,G^{(d)}}(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}, \mathbf{e}_i)\| < P^{-d+(d-1)\eta} \quad (1 \le i \le n) \tag{4.2}$$

satisfies $N(\eta) \gg R^{n(d-1)-2^{d-1}(Q/\eta)-\varepsilon}$, where $R = P^\eta$ and the implicit constant in $\gg$ depends only on n, d, η, and ε.

Recall that $\boldsymbol{\psi} = \{\psi_1, \dots, \psi_{r_d}\}$. Given $\mathbf{x}_1, \dots, \mathbf{x}_{d-1}$ as above, we form the matrix

$$[m_{ij}]_{\mathbf{x}_1,\dots,\mathbf{x}_{d-1}} = [\Gamma_{d,\psi_j}(\mathbf{x}_1, \dots, \mathbf{x}_{d-1}, \mathbf{e}_i)] \quad (1 \le i \le n, 1 \le j \le r_d).$$

Recall that f_j is the degree d portion of ψ_j $(1 \le j \le r_d)$ and $\mathbf{f} = \{f_1, \dots, f_{r_d}\}$. Since each ψ_j is of degree d, it follows that $\Gamma_{d,\psi_j} = \Gamma_{d,f_j}$ $(1 \le j \le r_d)$. It is also clear that $G^{(d)}(\mathbf{x}) = \boldsymbol{\alpha} \cdot \mathbf{f}(\mathbf{x})$. Now if this matrix $[m_{ij}]_{\mathbf{x}_1,\dots,\mathbf{x}_{d-1}}$ has rank less than r_d for each of the $(d-1)$-tuples counted by $N(\eta)$, then by the definition of $z_R(\mathfrak{M}_d)$ we have that

$$z_R(\mathfrak{M}_d) \ge N(\eta) \gg R^{n(d-1)-2^{d-1}(Q/\eta)-\varepsilon},$$

where again the implicit constant in $\gg$ depends only on n, d, η, and ε. Thus we have (iii) in this case. Hence, we may suppose that at least one of these matrices, which we

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

denote by $[m_{ij}]$, has rank r_d. Without loss of generality, suppose the submatrix M_0, formed by taking the first r_d columns of $[m_{ij}]$, has rank r_d. Let $n_0 = \det(M_0)$.

It follows from the definition of Γ_{d,f_j} that every monomial occurring in $\Gamma_{d,f_j}(\mathbf{x}_1, \dots, \mathbf{x}_d)$ has some component of $\mathbf{x}_i$ as a factor for each $1 \le i \le d$ [4, Proof of Lemma 11.2]. Also, the maximum absolute value of all coefficients of Γ_{d,f_j} is bounded by a constant dependent only on d and the coefficients of f_j [4, Lemma 11.3]. Therefore, by the construction of $[m_{ij}]$ we have $m_{ij} \ll R^{d-1}$, and hence $n_0 \ll R^{r_d(d-1)} = P^{r_d(d-1)\eta}$, where the implicit constants in $\ll$ depend only on r_d and $\mathbf{f}$.

We have $\Gamma_{d,G^{(d)}} = \sum_{j=1}^{r_d} \Gamma_{d,\alpha_j f_j} = \sum_{j=1}^{r_d} \alpha_j \Gamma_{d,f_j}$. Hence, from (4.2) we can write $\sum_{j=1}^{r_d} \alpha_j m_{ij} = c_i + \beta_i$ $(1 \le i \le n)$, where the c_i are integers and the β_i are real numbers bounded by the right-hand side of (4.2). Let $u_1, \dots, u_{r_d}$ be the solution of the system of linear equations

$$\sum_{j=1}^{r_d} u_j m_{ij} = n_0 c_i \quad (1 \le i \le r_d). \tag{4.3}$$

Then

$$\sum_{j=1}^{r_d} (n_0 \alpha_j - u_j) m_{ij} = n_0 \beta_i \quad (1 \le i \le r_d). \tag{4.4}$$

By applying Cramer's rule to (4.3), it follows that the u_j are integers. Also, by applying Cramer's rule to (4.4), we obtain that

$$\|n_0 \alpha_j\| \le |n_0 \alpha_j - u_j| \ll R^{(d-1)(r_d-1)} P^{-d+(d-1)\eta} = P^{-d+(d-1)r_d\eta},$$

where the implicit constant in $\ll$ depends only on r_d and $\mathbf{f}$. This completes the proof of Lemma 4.3 ∎

We define $g_d(\mathbf{f})$ to be the largest real number such that

$$z_P(\mathfrak{M}_d) \ll P^{n(d-1)-g_d(\mathbf{f})+\varepsilon} \tag{4.5}$$

holds for each $\varepsilon > 0$. It was proved [4, Corollary, p. 280] that

$$h(\mathbf{f}) < \frac{d!}{(\log 2)^d} \big(g_d(\mathbf{f}) + (d-1) r_d (r_d - 1) \big). \tag{4.6}$$

Let $\gamma_d = \frac{2^{d-1}(d-1)r_d}{g_d(\mathbf{f})}$ when $g_d(\mathbf{f}) > 0$. We let $\gamma_d = +\infty$ if $g_d(\mathbf{f}) = 0$. We also define

$$\gamma_d' = \frac{2^{d-1}}{g_d(\mathbf{f})} = \frac{\gamma_d}{(d-1)r_d}.$$

Corollary 4.4 ([4, p. 276, Corollary]) *Given a box $\mathfrak{B}$ with sides ≤ 1, we define the sum $S(\boldsymbol{\alpha})$ associated with ψ and $\mathfrak{B}$ as in (4.1). Suppose $\varepsilon' > 0$ is sufficiently small and $Q > 0$ satisfies $Q\gamma_d' < 1$. Then one of the following alternatives must hold.*

(i) $|S(\boldsymbol{\alpha})| \le P^{n-Q}$.

(ii) *There exists $n_0 \in \mathbb{N}$ such that $n_0 \ll P^{Q\gamma_d+\varepsilon'}$ and $\|n_0\boldsymbol{\alpha}\| \ll P^{-d+Q\gamma_d+\varepsilon'}$, where the implicit constants in $\ll$ depend only on $n, d, r_d, \varepsilon', Q$, and $\mathbf{f}$.*

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Note that the fact that the implicit constant depends on $\mathbf{f}$, but not on other lower order terms of $\boldsymbol{\psi}$, is an important feature that we make use of in Section 6.1.

Proof Since $Q\gamma_d' < 1$, we can choose $\varepsilon_1 > 0$ sufficiently small so that $\eta = Q\gamma_d' + \varepsilon_1$ satisfies $0 < \eta \le 1$. Also, with this choice of η, we have

$$\frac{2^{d-1}Q}{\eta} = \frac{2^{d-1}Q}{Q\gamma_d' + \varepsilon_1} = \frac{g_d(\mathbf{f})}{1 + \varepsilon_1 g_d(\mathbf{f})/(2^{d-1}Q)} < g_d(\mathbf{f}).$$

Then choose $\varepsilon_0 > 0$ such that $2^{d-1}Q/\eta + \varepsilon_0 < g_d(\mathbf{f})$. By the definition of $g_d(\mathbf{f})$ we have $z_R(\mathfrak{M}_d) \ll R^{n(d-1)-2^{d-1}Q/\eta-\varepsilon_0}$. Thus, in this case we see that statement (iii) in Lemma 4.3 cannot occur with $0 < \varepsilon < \varepsilon_0$. Also the equation $\eta = Q\gamma_d' + \varepsilon_1$ implies $r_d(d-1)\eta = Q\gamma_d + r_d(d-1)\varepsilon_1$. Therefore, from Lemma 4.3 (applying it with $0 < \varepsilon < \varepsilon_0$) we obtain our result with $\varepsilon' = r_d(d-1)\varepsilon_1$. ∎

For the rest of this section, we assume $\boldsymbol{\psi}$ to be a system of integral polynomials of degree d. When the polynomials $\boldsymbol{\psi}$ in question are over $\mathbb{Z}$, we consider the following.

Hypothesis $(\star)$ Let $\mathfrak{B}$ be a box in $\mathbb{R}^n$. For any $\Delta > 0$, there exists $P_1 = P_1(\mathbf{f}, \Omega, \Delta, \mathfrak{B})$ such that for $P > P_1$, each $\boldsymbol{\alpha} \in \mathbb{T}^{r_d}$ satisfies at least one of the following two alternatives.

(i) $|S(\boldsymbol{\alpha})| \le P^{n-\Delta\Omega}$.

(ii) There exists $q = q(\boldsymbol{\alpha}) \in \mathbb{N}$ such that $q \le P^{\Delta}$ and $\|q\boldsymbol{\alpha}\| \le P^{-d+\Delta}$.

We will say that the restricted Hypothesis $(\star)$ holds if the above condition holds for each Δ in $0 < \Delta \le 1$.

The important thing to note here is that the lower bound for P in Hypothesis $(\star)$ only depends on $\mathbf{f}$, and not on $\boldsymbol{\psi}$. In other words, only the highest degree portion of the polynomials $\boldsymbol{\psi}$ play a role in this estimate.

Proposition 4.5 ([4, Proposition II_0]) *Given a box $\mathfrak{B}$ with sides ≤ 1, Hypothesis $(\star)$ is true for any Ω in*

$$0 < \Omega < \frac{g_d(\mathbf{f})}{2^{d-1}(d-1)r_d}. \tag{4.7}$$

Proof It follows from (4.7) that $\Omega\gamma_d < 1$. We set $Q = \Delta\Omega$, and let $\varepsilon > 0$ be sufficiently small so that $Q\gamma_d + \varepsilon < \Delta$. First, we suppose $\Delta \le (d-1)r_d$. In this case, it follows that $Q\gamma_d' < 1$. Thus it follows from Corollary 4.4 that there exists $P_0 = P_0(\mathbf{f}, \Omega, \Delta)$ such that whenever $P > P_0$, one of the following must hold.

(i) $|S(\boldsymbol{\alpha})| \le P^{n-\Delta\Omega}$.

(ii) There exists $q \in \mathbb{N}$ such that $q \le P^{\Delta}$ and $\|q\boldsymbol{\alpha}\| \le P^{-d+\Delta}$.

On the other hand, if $\Delta > (d-1)r_d$, then case (ii) is always true by Dirichlet's theorem on Diophantine approximation. ∎

For each $q \in \mathbb{N}$, we denote $\mathbb{U}_q$ as the group of units in $\mathbb{Z}/q\mathbb{Z}$. Given $\mathbf{m} \in \mathbb{U}_q^{r_d}$, we define

$$E(q^{-1}\mathbf{m}) = E(\boldsymbol{\psi}, q; q^{-1}\mathbf{m}) = q^{-n} \sum_{\mathbf{x} \pmod q} e(q^{-1}\,\mathbf{m} \cdot \boldsymbol{\psi}(\mathbf{x})).$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Lemma 4.6 ([4, Lemma 7.1]) *Suppose Ω satisfies (4.7). Then for $0 < Q < \Omega$, we have*

$$|E(q^{-1}\mathbf{m})| \ll q^{-Q}, \tag{4.8}$$

where the implicit constant in $\ll$ depends only on $\mathbf{f}$, Q, and Ω.

Again the fact that the implicit constant depends on $\mathbf{f}$, but not on other lower order terms of $\boldsymbol{\psi}$, becomes crucial when we apply this lemma in Section 6.1.

Proof Since $E(q^{-1}\mathbf{m}) = q^{-n}S(\boldsymbol{\alpha})$ with $\boldsymbol{\alpha} = q^{-1}\mathbf{m}$, $P = q$, and $\mathfrak{B} = [0,1)^{r_d}$, and with our choice of Ω we know that Hypothesis $(\star)$ is satisfied by Proposition 4.5. Thus we apply it with $\Delta = Q/\Omega < 1$. Let q be sufficiently large, and suppose we are in case (ii) of Hypothesis $(\star)$. Then we know there exists $q_0 \leq q^{\Delta} < q$ (when $q \neq 1$) with $\|q_0 q^{-1}\mathbf{m}\| \leq q^{-d+\Delta} < q^{-1}$. Since $(\mathbf{m}, q) = 1$, this is not possible. Therefore, we must have case (i) of Hypothesis $(\star)$, which is precisely the inequality (4.8). ∎

5 Hardy–Littlewood Circle Method: Minor Arcs

For each $q \in \mathbb{N}$, recall that we let $\mathbb{U}_q$ be the group of units in $\mathbb{Z}/q\mathbb{Z}$. When $q = 1$, we let $\mathbb{U}_1 = \{0\}$. Let us denote $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. For a given value of $C > 0$ and an integer $1 \leq q \leq (\log N)^C$, we define the *major arc*

$$\mathfrak{M}_{m,q}(C) = \left\{ \alpha \in \mathbb{T} : \|\alpha - m/q\| \leq N^{-d}(\log N)^C \right\}$$

for each $m \in \mathbb{U}_q$. Recall that $\|\beta\|$ is the distance from $\beta \in \mathbb{R}$ to the nearest integer, which induces a metric on $\mathbb{T}$ via $d(\alpha, \beta) = \|\alpha - \beta\|$. These arcs are disjoint for N sufficiently large, and we define

$$\mathfrak{M}(C) = \bigcup_{q \leq (\log N)^C} \bigcup_{m \in \mathbb{U}_q} \mathfrak{M}_{m,q}(C).$$

We then define the *minor arcs* to be $\mathfrak{m}(C) = \mathbb{T} \backslash \mathfrak{M}(C)$.

We obtain the following bound on the minor arcs in this section.

Proposition 5.1 *Let $b(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$ be a polynomial of degree d. Let $T(b;\alpha)$ be defined as in (1.2). Then there exists a positive number A_d dependent only on d such that the following holds. Suppose $b(\mathbf{x})$ satisfies $h^\star(f_b) > A_d$. Then, given any $c > 0$, there exists $C > 0$ such that*

$$\int_{\mathfrak{m}(C)} T(b;\alpha)\, d\alpha \ll \frac{N^{n-d}}{(\log N)^c}.$$

The proposition is achieved by splitting the exponential sum $T(b;\alpha)$ over certain level sets based on a decomposition of the polynomial $b(\mathbf{x})$. Thus before we get into the proof of Proposition 5.1, we first establish this decomposition in six steps, where the resulting decomposition is given in (5.12). For simplicity, we let $f(\mathbf{x})$ be the degree d portion of $b(\mathbf{x})$ for the remainder of the paper. We let $h = h(f)$, and let $0 < M < h^\star(f) \leq h$ be chosen later.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Step 1: Decomposition of the variables As explained in the paragraph before (2.4), by relabeling the variables if necessary, we have

$$f = (x_1 + \ell_1)v'_1 + \cdots + (x_M + \ell_M)v'_M + u'_{M+1}v'_{M+1} + \cdots + u'_h v'_h,$$

where each ℓ_i is a linear form in $\mathbb{Q}[x_{M+1}, \ldots, x_n]$ $(1 \le i \le M)$, and each $u'_{i'}$ and v'_j are rational forms of positive degree $(M+1 \le i' \le h, 1 \le j \le h)$. We can then find a monomial $x_{i_1}x_{i_2}\cdots x_{i_d}$, where $M < i_1 \le i_2 \le \cdots \le i_d$, of f with a non-zero coefficient. This is the case, for otherwise it means that every monomial of f is divisible by one of $x_1, \ldots, x_M$, and consequently that $h = h(f) \le M$, which is a contradiction. We denote the distinct variables of $\{x_{i_1}, x_{i_2}, \ldots, x_{i_d}\} \subseteq \{x_{M+1}, \ldots, x_n\}$ by $\{w_1, \ldots, w_K\}$, and let $\mathbf{w} = (w_1, \ldots, w_K)$. Clearly, we have $K \le d$. We selected these K variables for the purpose of applying Weyl differencing later. We also label $\mathbf{y} = (x_1, \ldots, x_M) = (y_1, \ldots, y_M)$ for notational convenience, let $\mathbf{z} = \{x_{M+1}, \ldots, x_n\} \backslash \mathbf{w}$, and let $\mathbf{z} = (z_1, \ldots, z_{n-M-K})$. We note that each ℓ_i is a rational linear form only in the $\mathbf{w}$ and the $\mathbf{z}$ variables $(1 \le i \le h)$.

Step 2: Decomposition of $f(\mathbf{x})$ We define g_M with respect to f as in (2.5). By Lemma 2.3, we have

$$f(\mathbf{x}) = f(\mathbf{w}, \mathbf{y}, \mathbf{z}) = g_M(\mathbf{w}, \mathbf{y}, \mathbf{z}) + f(\mathbf{w}, (-\ell_1, \ldots, -\ell_M), \mathbf{z}),$$

where

$$h(g_M(\mathbf{w}, \mathbf{y}, \mathbf{z})) \ge M \quad \text{and} \quad h(f(\mathbf{w}, (-\ell_1, \ldots, -\ell_M), \mathbf{z})) = h - M. \tag{5.1}$$

We then have

$$f(\mathbf{0}, \mathbf{y}, \mathbf{z}) = g_M(\mathbf{0}, \mathbf{y}, \mathbf{z}) + f(\mathbf{0}, (-\ell_1|_{\mathbf{w}=\mathbf{0}}, \ldots, -\ell_M|_{\mathbf{w}=\mathbf{0}}), \mathbf{z}). \tag{5.2}$$

Let $f_M(\mathbf{z}) = f(\mathbf{0}, (-\ell_1|_{\mathbf{w}=\mathbf{0}}, \ldots, -\ell_M|_{\mathbf{w}=\mathbf{0}}), \mathbf{z})$. Consequently, we obtain from Lemma 2.1 and (5.1) that

$$h(g_M(\mathbf{0}, \mathbf{y}, \mathbf{z})) \ge M - K \ge M - d, \tag{5.3}$$

$$h(f_M(\mathbf{z})) \ge h - M - K \ge h - M - d. \tag{5.4}$$

Step 3: Decomposition of $b(\mathbf{x})$ with respect to $\mathbf{w}$, $\mathbf{y}$, and $\mathbf{z}$ Let $b_M(\mathbf{z}) = b(\mathbf{0}, (-\ell_1|_{\mathbf{w}=\mathbf{0}}, \ldots, -\ell_M|_{\mathbf{w}=\mathbf{0}}), \mathbf{z})$. It is clear that the degree d portion of the polynomial $b(\mathbf{0}, \mathbf{y}, \mathbf{0})$ is $g_M(\mathbf{0}, \mathbf{y}, \mathbf{0})$. Let

$$\begin{aligned} b(\mathbf{0}, \mathbf{y}, \mathbf{z}) - b_M(\mathbf{z}) = \sum_{j=1}^{d-1} \sum_{1 \le t_1 \le \cdots \le t_j \le M} \Big(\sum_{k=0}^{d-j} \Psi^{(k)}_{t_1, \ldots, t_j}(\mathbf{z}) \Big) y_{t_1} \cdots y_{t_j} \\ + \Big(\sum_{k=1}^{d} \Psi^{(k)}_{\varnothing}(\mathbf{z}) \Big) + g_M(\mathbf{0}, \mathbf{y}, \mathbf{0}), \end{aligned} \tag{5.5}$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

where $\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})$ and $\Psi^{(k)}_{\varnothing}(\mathbf{z})$ are forms of degree k. With these notations, we have the following decomposition,

$$\begin{aligned}(5.6)\qquad b(\mathbf{w},\mathbf{y},\mathbf{z}) &= b(\mathbf{w},\mathbf{0},\mathbf{0})\\ &\quad+\sum_{j=1}^{d-1}\sum_{1\le i_1\le\cdots\le i_j\le K}\Big(\sum_{k=1}^{d-j}\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y},\mathbf{z})\Big)w_{i_1}\cdots w_{i_j}\\ &\quad+\sum_{j=1}^{d-1}\sum_{1\le t_1\le\cdots\le t_j\le M}\Big(\sum_{k=0}^{d-j}\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})\Big)y_{t_1}\cdots y_{t_j}\\ &\quad+\Big(\sum_{k=1}^{d}\Psi^{(k)}_{\varnothing}(\mathbf{z})\Big)+g_M(\mathbf{0},\mathbf{y},\mathbf{0})+b_M(\mathbf{z})-b(\mathbf{0},\mathbf{0},\mathbf{0}),\end{aligned}$$

which we describe below. We note that $\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y},\mathbf{z})$ are forms of degree k. The above decomposition establishes the following. The term

$$b(\mathbf{w},\mathbf{0},\mathbf{0})+\sum_{j=1}^{d-1}\sum_{1\le i_1\le\cdots\le i_j\le K}\Big(\sum_{k=1}^{d-j}\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y},\mathbf{z})\Big)w_{i_1}\cdots w_{i_j}$$

consists of all the monomials of $b(\mathbf{x})$ that involve any variables of $\mathbf{w}$. Consequently, we have

$$\begin{aligned}b(\mathbf{0},\mathbf{y},\mathbf{z}) &= \sum_{j=1}^{d-1}\sum_{1\le t_1\le\cdots\le t_j\le M}\Big(\sum_{k=0}^{d-j}\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})\Big)y_{t_1}\cdots y_{t_j}\\ &\quad+\Big(\sum_{k=1}^{d}\Psi^{(k)}_{\varnothing}(\mathbf{z})\Big)+g_M(\mathbf{0},\mathbf{y},\mathbf{0})+b_M(\mathbf{z}),\end{aligned}$$

and the degree d portion of $b(\mathbf{0},\mathbf{y},\mathbf{z})$ is $f(\mathbf{0},\mathbf{y},\mathbf{z})$. Clearly, the degree d portion of $b_M(\mathbf{z})$ is $f_M(\mathbf{z})=f(\mathbf{0},(-\ell_1|_{\mathbf{w}=\mathbf{0}},\dots,-\ell_M|_{\mathbf{w}=\mathbf{0}}),\mathbf{z})$. It then follows from (5.2) and (5.5) that the degree d portion of

$$\sum_{j=1}^{d-1}\sum_{1\le t_1\le\cdots\le t_j\le M}\Big(\sum_{k=0}^{d-j}\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})\Big)y_{t_1}\cdots y_{t_j}+\Big(\sum_{k=1}^{d}\Psi^{(k)}_{\varnothing}(\mathbf{z})\Big)+g_M(\mathbf{0},\mathbf{y},\mathbf{0})$$

is

$$g_M(\mathbf{0},\mathbf{y},\mathbf{z})=\sum_{j=1}^{d-1}\sum_{1\le t_1\le\cdots\le t_j\le M}\Psi^{(d-j)}_{t_1,\dots,t_j}(\mathbf{z})\,y_{t_1}\cdots y_{t_j}+\Psi^{(d)}_{\varnothing}(\mathbf{z})+g_M(\mathbf{0},\mathbf{y},\mathbf{0}).$$

We also know from (5.2) that $g_M(\mathbf{0},(-\ell_1|_{\mathbf{w}=\mathbf{0}},\dots,-\ell_M|_{\mathbf{w}=\mathbf{0}}),\mathbf{z})=0$, and consequently,

$$\begin{aligned}\Psi^{(d)}_{\varnothing}(\mathbf{z}) &= \Big(-\sum_{j=1}^{d-1}\sum_{1\le t_1\le\cdots\le t_j\le M}\Psi^{(d-j)}_{t_1,\dots,t_j}(\mathbf{z})\,y_{t_1}\cdots y_{t_j}\Big)\Big|_{y_i=-\ell_i|_{\mathbf{w}=\mathbf{0}}\,(1\le i\le M)}\\ &\quad-g_M(\mathbf{0},(-\ell_1|_{\mathbf{w}=\mathbf{0}},\dots,-\ell_M|_{\mathbf{w}=\mathbf{0}}),\mathbf{0}).\end{aligned}$$

In other words, $\Psi^{(d)}_{\varnothing}(\mathbf{z})$ can be expressed as a rational polynomial in the forms $\big\{\Psi^{(d-j)}_{t_1,\dots,t_j}(\mathbf{z}):1\le j\le d-1,1\le t_1\le\cdots\le t_j\le M\big\}\cup\big\{\ell_i|_{\mathbf{w}=\mathbf{0}}:1\le i\le M\big\}$.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Step 4: Regularization of systems Φ and Ψ We let

$$\Phi = \left\{ \Phi^{(k)}_{i_1,\dots,i_j} : 1 \le j \le d-1, 1 \le i_1 \le \cdots \le i_j \le K, 1 \le k \le d-j \right\}.$$

Note every polynomial of Φ has degree strictly less than d, and involves only the $\mathbf{y}$ and the $\mathbf{z}$ variables. Clearly, we have $|\Phi| \le d^2K^d \le d^{d+2}$. We apply Proposition 3.5 to the system Φ with respect to the functions $\mathcal{F} = \{\mathcal{F}_2, \dots, \mathcal{F}_{d-1}\}$, where $\mathcal{F}_i(t) = \rho_{d,i}(2+2t)+2t$ for $2 \le i \le d-1$, and obtain $\mathcal{R}(\Phi) = (\mathbf{a}^{(d-1)}, \dots, \mathbf{a}^{(1)})$. For each form $a_i^{(s)} \in \mathbf{a}^{(s)}$ $(1 \le s \le d-1, 1 \le i \le |\mathbf{a}^{(s)}|)$, we write

$$a_i^{(s)}(\mathbf{y}, \mathbf{z}) = \sum_{k=0}^{s} \sum_{1 \le i_1 \le \cdots \le i_k \le M} \widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z}) y_{i_1} \cdots y_{i_k}, \tag{5.7}$$

where each $\widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z})$ is a form of degree $s-k$. Thus each form $a_i^{(s)}$ introduces at most $(s+1)M^s \le dM^d$ forms in $\mathbf{z}$. Also, for each $1 \le i \le d-1$, we define $\overline{\mathbf{a}}^{(i)}$ to be the system obtained by removing from $\mathbf{a}^{(i)}$ all forms that depend only on the $\mathbf{z}$ variables. Let $\overline{\mathcal{R}}(\Phi) = (\overline{\mathbf{a}}^{(d-1)}, \dots, \overline{\mathbf{a}}^{(1)})$, $R_2 = \sum_{i=1}^{d-1} |\overline{\mathbf{a}}^{(i)}|$, and $D_2 = \sum_{i=1}^{d-1} i\,|\overline{\mathbf{a}}^{(i)}|$. By relabeling if necessary, we denote the elements of $\overline{\mathbf{a}}^{(s)}$ by $\overline{\mathbf{a}}^{(s)} = \{a_i^{(s)} : 1 \le i \le |\overline{\mathbf{a}}^{(s)}|\}$ for each $1 \le s \le d-1$.

Let

$$\begin{aligned}\Psi = {} & \{\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z}) : 1 \le j \le d-1, 1 \le t_1 \le \cdots \le t_j \le M, 0 \le k \le d-j\} \\ & \cup \{\Psi^{(k)}_{\varnothing}(\mathbf{z}) : 1 \le k < d\} \cup \{\ell_i|_{\mathbf{w}=\mathbf{0}} : 1 \le i \le M\} \\ & \cup \{\widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z}) : 1 \le s \le d-1, 1 \le i \le |\mathbf{a}^{(s)}|, \\ & \qquad\qquad 1 \le k \le s, 1 \le i_1 \le \cdots \le i_k \le M\}.\end{aligned}$$

In other words, Ψ is the collection of $\ell_i|_{\mathbf{w}=\mathbf{0}}$, and all $\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})$, $\widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z})$, and $\Psi^{(k)}_{\varnothing}(\mathbf{z})$, except $\Psi^{(d)}_{\varnothing}(\mathbf{z})$. In particular, every polynomial of Ψ has degree strictly less than d. We can see that $|\Psi| \le d^2M^d + d + M + |\mathcal{R}(\Phi)|dM^d$. We let $\mathcal{R}(\Psi)$ be a regularization of Ψ with respect to the functions $\mathcal{F} = \{\mathcal{F}_2, \dots, \mathcal{F}_{d-1}\}$, where again $\mathcal{F}_i(t) = \rho_{d,i}(2+2t)+2t$ for $2 \le i \le d-1$. Let $\mathcal{R}(\Psi) = (\mathbf{v}^{(d-1)}, \dots, \mathbf{v}^{(1)})$, $R_1 = \sum_{i=1}^{d-1} |\mathbf{v}^{(i)}|$, and $D_1 = \sum_{i=1}^{d-1} i\,|\mathbf{v}^{(i)}|$.

Let $\mathcal{R}^{(i)}(\Phi)$, $\Phi^{(i)}$, and $\mathcal{R}^{(i)}(\Psi)$ denote the degree i forms of $\mathcal{R}(\Phi)$, Φ, and $\mathcal{R}(\Psi)$, respectively. From Proposition 3.5, we know that each $|\mathcal{R}^{(i)}(\Phi)| = |\mathbf{a}^{(i)}|$ $(1 \le i \le d-1)$, and consequently R_2, is bounded by some constant dependent only on $\mathcal{F}$ and $|\Phi^{(d-1)}|$, …, $|\Phi^{(1)}|$. Thus we see that R_2 is bounded by a constant dependent only on d. We set $M = \rho_{d,d}(2+2R_2)+2R_2+d$, and note that M is bounded by a constant dependent only on d. By Proposition 3.5 again, we have that each $|\mathcal{R}^{(i)}(\Psi)| = |\mathbf{v}^{(i)}|$ $(1 \le i \le d-1)$, and consequently R_1, is bounded by some constant dependent only on d, $\mathcal{F}$, M, and $|\Phi^{(d-1)}|$, …, $|\Phi^{(1)}|$. Thus R_1 is bounded by a constant dependent only on d as well.

We define

$$A_d = \max\Big\{ 2\rho_{d,d}(2+2R_1) + 4R_1 + 2d, 2\rho_{d,d}(2+2R_2) + 4R_2 + 2d, \frac{5 \cdot 2^{d-1} \cdot (d-1) \cdot d!}{(\log 2)^d} + 5d \Big\}, \tag{5.8}$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

and suppose $h^*(f) \geq A_d$. We note that the third term inside the maximum function above is not required in this section, but this lower bound on A_d becomes necessary in Section 6. With this choice of A_d, we have from (5.3) and (5.4) that

$$h(f_M(\mathbf{z})) \geq h - M - d \geq \rho_{d,d}(2 + 2R_1) + 2R_1, \tag{5.9}$$

$$h(g_M(\mathbf{0}, \mathbf{y}, \mathbf{z})) \geq M - d \geq \rho_{d,d}(2 + 2R_2) + 2R_2. \tag{5.10}$$

Step 5: Definition of the level sets $Z(\mathbf{H})$ and $Y(\mathbf{G};\mathbf{H})$ For each $\mathbf{H} \in \mathbb{Z}^{R_1}$, we define the following set

$$Z(\mathbf{H}) = \left\{\mathbf{z} \in [0, N]^{n-M-K} \cap \mathbb{Z}^{n-M-K} : \mathcal{R}(\Psi)(\mathbf{z}) = \mathbf{H}\right\}.$$

By Proposition 3.5, we know that each of the polynomials $\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z})$ and $\Psi^{(k)}_{\varnothing}(\mathbf{z})$ in (5.6) can be expressed as a rational polynomial in the forms of $\mathcal{R}(\Psi)$. Let $\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z}) = \hat{c}^{(k)}_{t_1,\dots,t_j}(\mathcal{R}(\Psi))$ and $\Psi^{(k)}_{\varnothing}(\mathbf{z}) = \hat{c}^{(k)}_{\varnothing}(\mathcal{R}(\Psi))$, where $\hat{c}^{(k)}_{t_1,\dots,t_j}$ and $\hat{c}^{(k)}_{\varnothing}$ are rational polynomials in R_1 variables. Therefore, for any $\mathbf{z}_0 \in Z(\mathbf{H})$, we have $\Psi^{(k)}_{t_1,\dots,t_j}(\mathbf{z}_0) = \hat{c}^{(k)}_{t_1,\dots,t_j}(\mathbf{H})$ and $\Psi^{(k)}_{\varnothing}(\mathbf{z}_0) = \hat{c}^{(k)}_{\varnothing}(\mathbf{H})$.

Since each of the forms $\widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z})$ in (5.7) can be expressed as a rational polynomial in the forms of $\mathcal{R}(\Psi)$, let

$$\widetilde{\Psi}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{z}) = \widetilde{c}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathcal{R}(\Psi)),$$

where each $\widetilde{c}^{(s-k)}_{s:i:i_1,\dots,i_k}$ is a rational polynomial in R_1 variables. Therefore, for each $a^{(s)}_i \in \mathcal{R}(\Phi) = (\mathbf{a}^{(d-1)}, \dots, \mathbf{a}^{(1)})$, where $1 \leq s \leq d-1$ and $1 \leq i \leq |\mathbf{a}^{(s)}|$, we can write

$$a^{(s)}_i(\mathbf{y}, \mathbf{z}) = \sum_{k=0}^{s} \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq M} \widetilde{c}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathcal{R}(\Psi)) y_{i_1} \cdots y_{i_k}.$$

Consequently, we can define the following polynomial for each $1 \leq s \leq d-1$ and $1 \leq i \leq |\mathbf{a}^{(s)}|$,

$$a^{(s)}_i(\mathbf{y}, Z(\mathbf{H})) = \sum_{k=0}^{s} \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq M} \widetilde{c}^{(s-k)}_{s:i:i_1,\dots,i_k}(\mathbf{H}) y_{i_1} \cdots y_{i_k}, \tag{5.11}$$

so that given any $\mathbf{z}_0 \in Z(H)$, we have $a^{(s)}_i(\mathbf{y}, \mathbf{z}_0) = a^{(s)}_i(\mathbf{y}, Z(\mathbf{H}))$. We also define

$$\overline{R}(\Phi)(\mathbf{y}, Z(\mathbf{H})) = \left\{a^{(s)}_i(\mathbf{y}, Z(\mathbf{H})) : 1 \leq s \leq d-1, 1 \leq i \leq |\overline{\mathbf{a}}^{(s)}|\right\},$$

which consists of R_2 polynomials with possible repetitions. For each $\mathbf{G} \in \mathbb{Z}^{R_2}$, we let $Y(\mathbf{G};\mathbf{H}) = \left\{\mathbf{y} \in [0, N]^M \cap \mathbb{Z}^M : \overline{\mathcal{R}}(\Phi)(\mathbf{y}, Z(\mathbf{H})) = \mathbf{G}\right\}$.

Step 6: Decomposition of $b(\mathbf{w}, \mathbf{y}, \mathbf{z})$ when $(\mathbf{y}, \mathbf{z}) \in Y(\mathbf{G};\mathbf{H}) \times Z(\mathbf{H})$ Recall that Φ is the collection of all $\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y}, \mathbf{z})$ in (5.6), and that each $\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y}, \mathbf{z})$ can be expressed as a rational polynomial in the forms of $\mathcal{R}(\Phi)$. Thus, it follows from this fact and (5.11) that each $\Phi^{(k)}_{i_1,\dots,i_j}(\mathbf{y}, \mathbf{z})$ is constant on $(\mathbf{y}, \mathbf{z}) \in Y(\mathbf{G};\mathbf{H}) \times Z(\mathbf{H})$, and we denote this constant value by $c^{(k)}_{i_1,\dots,i_j}(\mathbf{G}, \mathbf{H})$. Therefore, for any choice of $\mathbf{z} \in Z(\mathbf{H})$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

and $\mathbf{y} \in Y(\mathbf{G};\mathbf{H})$, the polynomial $b(\mathbf{x})$ takes the following shape

$$\begin{aligned} b(\mathbf{w},\mathbf{y},\mathbf{z}) = b(\mathbf{w},\mathbf{0},\mathbf{0}) &+ \sum_{j=1}^{d-1} \sum_{1\le i_1\le\cdots\le i_j\le K} \Big(\sum_{k=1}^{d-j} c^{(k)}_{i_1,\dots,i_j}(\mathbf{G},\mathbf{H}) \Big) w_{i_1}\cdots w_{i_j} \\ &+ \sum_{j=1}^{d-1} \sum_{1\le t_1\le\cdots\le t_j\le M} \Big(\sum_{k=0}^{d-j} \hat{c}^{(k)}_{t_1,\dots,t_j}(\mathbf{H}) \Big) y_{t_1}\cdots y_{t_j} \\ &+ \Big(\sum_{k=1}^{d} \hat{c}^{(k)}_{\varnothing}(\mathbf{H}) \Big) + g_M(\mathbf{0},\mathbf{y},\mathbf{0}) + b_M(\mathbf{z}) - b(\mathbf{0},\mathbf{0},\mathbf{0}). \end{aligned}$$

We label

$$\begin{aligned} \mathfrak{C}_0(\mathbf{w},\mathbf{G},\mathbf{H}) &= b(\mathbf{w},\mathbf{0},\mathbf{0}) + \sum_{j=1}^{d-1} \sum_{1\le i_1\le\cdots\le i_j\le K} \Big(\sum_{k=1}^{d-j} c^{(k)}_{i_1,\dots,i_j}(\mathbf{G},\mathbf{H}) \Big) w_{i_1}\cdots w_{i_j}, \\ \mathfrak{C}_1(\mathbf{y},\mathbf{H}) &= \sum_{j=1}^{d-1} \sum_{1\le t_1\le\cdots\le t_j\le M} \Big(\sum_{k=0}^{d-j} \hat{c}^{(k)}_{t_1,\dots,t_j}(\mathbf{H}) \Big) y_{t_1}\cdots y_{t_j} \\ &\quad + \Big(\sum_{k=1}^{d} \hat{c}^{(k)}_{\varnothing}(\mathbf{H}) \Big) + g_M(\mathbf{0},\mathbf{y},\mathbf{0}), \end{aligned}$$

so that for $\mathbf{z} \in Z(\mathbf{H})$ and $\mathbf{y} \in Y(\mathbf{G};\mathbf{H})$, we have

$$b(\mathbf{w},\mathbf{y},\mathbf{z}) = \mathfrak{C}_0(\mathbf{w},\mathbf{G},\mathbf{H}) + \mathfrak{C}_1(\mathbf{y},\mathbf{H}) + b_M(\mathbf{z}) - b(\mathbf{0},\mathbf{0},\mathbf{0}). \tag{5.12}$$

5.1 Proof of Proposition 5.1

We are now in position to prove Proposition 5.1.

Proof Using the notations above we define the following three exponential sums,

$$\begin{aligned} S_0(\alpha,\mathbf{G},\mathbf{H}) &= \sum_{\mathbf{w}\in[0,N]^K\cap\mathbb{Z}^K} \Lambda(\mathbf{w})\, e(\alpha\cdot\mathfrak{C}_0(\mathbf{w},\mathbf{G},\mathbf{H})), \\ S_1(\alpha,\mathbf{G},\mathbf{H}) &= \sum_{\mathbf{y}\in Y(\mathbf{G};\mathbf{H})} \Lambda(\mathbf{y})\, e(\alpha\cdot\mathfrak{C}_1(\mathbf{y},\mathbf{H})), \\ S_2(\alpha,\mathbf{H}) &= \sum_{\mathbf{z}\in Z(\mathbf{H})} \Lambda(\mathbf{z})\, e(\alpha\cdot b_M(\mathbf{z}) - \alpha\cdot b(\mathbf{0},\mathbf{0},\mathbf{0})). \end{aligned}$$

Let $\mathcal{L}_1(N) = \{\mathbf{H} \in \mathbb{Z}^{R_1} : Z(\mathbf{H}) \neq \varnothing\}$, and for each $\mathbf{H} \in \mathcal{L}_1(N)$, let $\mathcal{L}_2(N;\mathbf{H}) = \{\mathbf{G} \in \mathbb{Z}^{R_2} : Y(\mathbf{G},\mathbf{H}) \neq \varnothing\}$. It then follows that

$$|\mathcal{L}_1(N)| \ll N^{D_1} \quad \text{and} \quad |\mathcal{L}_2(N;\mathbf{H})| \ll N^{D_2}, \tag{5.13}$$

where the implicit constant in the second inequality is independent of $\mathbf{H}$. In order to prove the first inequality, let C_0 be the largest absolute value of all coefficients of the polynomials in $\mathcal{R}(\Psi)$. Also let M_0 be the largest number of monomials with non-zero coefficients in any of the polynomials in $\mathcal{R}(\Psi)$. Then we have $|\mathcal{L}_1(N)| \le (2C_0\cdot M_0)^{R_1}\cdot(N+1)^{D_1}$. To see the second inequality, we let C_0' be the largest absolute value of all coefficients of the polynomials $a_i^{(s)}(\mathbf{y},\mathbf{z})$ in $\overline{\mathcal{R}}(\Phi)$, and let M_0' be the largest number of monomials with non-zero coefficients in any of these polynomials.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Then we see that the number of values taken by $a_i^{(s)}(\mathbf{y},\mathbf{z})$ as $(\mathbf{y},\mathbf{z})$ varies in $[0,N]^{n-K}$ is $\le (2C_0'\cdot M_0')\cdot(N+1)^s$. Therefore, we have

$$\begin{aligned}\mathcal{L}_2(N;\mathbf{H}) &= \{\mathbf{G}\in\mathbb{Z}^{R_2} : Y(\mathbf{G},\mathbf{H})\neq\varnothing\}\\ &= \{\mathbf{G}\in\mathbb{Z}^{R_2} : \exists \mathbf{y}\in[0,N]^M\cap\mathbb{Z}^M, \overline{\mathcal{R}}(\Phi)(\mathbf{y},Z(\mathbf{H}))=\mathbf{G}\}\\ &\subseteq \{\mathbf{G}\in\mathbb{Z}^{R_2} : \exists (\mathbf{y},\mathbf{z})\in[0,N]^{n-K}\cap\mathbb{Z}^{n-K}, \overline{\mathcal{R}}(\Phi)(\mathbf{y},\mathbf{z})=\mathbf{G}\},\end{aligned}$$

and the cardinality of the last set is $\le (2C_0'\cdot M_0')^{R_2}\cdot(N+1)^{D_2}$.

By the Cauchy–Schwarz inequality and (5.13), we obtain

(5.14)

$$\begin{aligned}\Big|\int_{\mathfrak{m}(C)} T(b;\alpha)\,d\alpha\Big|^2 &\le \Big|\sum_{\mathbf{H}\in\mathcal{L}_1(N)}\sum_{\mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}\int_{\mathfrak{m}(C)}\sum_{\substack{\mathbf{w}\in[0,N]^K\cap\mathbb{Z}^K\\ \mathbf{z}\in Z(\mathbf{H})\\ \mathbf{y}\in Y(\mathbf{G};\mathbf{H})}}\Lambda(\mathbf{w})\Lambda(\mathbf{y})\Lambda(\mathbf{z})\\ &\quad\times e\Big(\alpha\cdot\big(\mathfrak{C}_0(\mathbf{w},\mathbf{G},\mathbf{H})+\mathfrak{C}_1(\mathbf{y},\mathbf{H})+b_M(\mathbf{z})-b(\mathbf{0},\mathbf{0},\mathbf{0})\big)\Big)\,d\alpha\Big|^2\\ &\ll N^{D_1+D_2}\sum_{\mathbf{H}\in\mathcal{L}_1(N)}\sum_{\mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}\Big|\int_{\mathfrak{m}(C)} S_0(\alpha,\mathbf{G},\mathbf{H})S_1(\alpha,\mathbf{G},\mathbf{H})S_2(\alpha,\mathbf{H})\,d\alpha\Big|^2\\ &\ll N^{D_1+D_2}\Big(\sup_{\substack{\mathbf{H}\in\mathcal{L}_1(N)\\ \mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}}\sup_{\alpha\in\mathfrak{m}(C)}|S_0(\alpha,\mathbf{G},\mathbf{H})|^2\Big)\\ &\quad\times\sum_{\mathbf{H}\in\mathcal{L}_1(N)}\sum_{\mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}\|S_1(\cdot,\mathbf{G},\mathbf{H})\|_2^2\,\|S_2(\cdot,\mathbf{H})\|_2^2,\end{aligned}$$

where $\|\cdot\|_2$ denotes the L^2-norm on $[0,1]$. By the orthogonality relation, it follows that

$$\|S_1(\cdot,\mathbf{G},\mathbf{H})\|_2^2\,\|S_2(\cdot,\mathbf{H})\|_2^2 \le (\log N)^{2n-2K}\mathcal{N}_1(\mathbf{G};\mathbf{H})\mathcal{N}_2(\mathbf{H}),$$

where

$$\begin{aligned}\mathcal{N}_1(\mathbf{G};\mathbf{H}) &= \big|\big\{(\mathbf{y},\mathbf{y}')\in Y(\mathbf{G};\mathbf{H})\times Y(\mathbf{G};\mathbf{H}) : \mathfrak{C}_1(\mathbf{y},\mathbf{H})=\mathfrak{C}_1(\mathbf{y}',\mathbf{H})\big\}\big|,\\ \mathcal{N}_2(\mathbf{H}) &= \big|\big\{(\mathbf{z},\mathbf{z}')\in Z(\mathbf{H})\times Z(\mathbf{H}) : b_M(\mathbf{z})=b_M(\mathbf{z}')\big\}\big|.\end{aligned}$$

With these notations, we can further bound (5.14) as follows

(5.15)

$$\begin{aligned}&\Big|\int_{\mathfrak{m}(C)} T(b;\alpha)\,d\alpha\Big|^2\\ &\qquad\ll (\log N)^{2n-2K}N^{D_1+D_2}\Big(\sup_{\substack{\mathbf{H}\in\mathcal{L}_1(N)\\ \mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}}\sup_{\alpha\in\mathfrak{m}(C)}|S_0(\alpha,\mathbf{G},\mathbf{H})|^2\Big)\,\mathcal{W},\end{aligned}$$

where

$$\mathcal{W}=\sum_{\mathbf{H}\in\mathcal{L}_1(N)}\sum_{\mathbf{G}\in\mathcal{L}_2(N;\mathbf{H})}\mathcal{N}_1(\mathbf{G};\mathbf{H})\mathcal{N}_2(\mathbf{H}).$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

We can express $\mathcal{W}$ as the number of solutions $\mathbf{y}, \mathbf{y}' \in [0,N]^M \cap \mathbb{Z}^M$ and $\mathbf{z}, \mathbf{z}' \in [0,N]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ of the system

$$\begin{aligned} \mathcal{R}(\Psi)(\mathbf{z}) &= \mathcal{R}(\Psi)(\mathbf{z}') = \mathbf{H}, \\ \overline{\mathcal{R}}(\Phi)(\mathbf{y}, Z(\mathbf{H})) &= \overline{\mathcal{R}}(\Phi)(\mathbf{y}', Z(\mathbf{H})) = \mathbf{G}, \\ \mathfrak{C}_1(\mathbf{y}, \mathbf{H}) &= \mathfrak{C}_1(\mathbf{y}', \mathbf{H}), \\ b_M(\mathbf{z}) &= b_M(\mathbf{z}'), \end{aligned} \tag{5.16}$$

for any $\mathbf{H} \in \mathcal{L}_1(N)$ and $\mathbf{G} \in \mathcal{L}_2(N;\mathbf{H})$. We know that the system

$$\overline{\mathcal{R}}(\Phi)(\mathbf{y}, Z(\mathbf{H}))$$

is identical to $\overline{\mathcal{R}}(\Phi)(\mathbf{y}, \mathbf{z}_0)$ for any choice of $\mathbf{z}_0 \in Z(\mathbf{H})$ and any $\mathbf{y} \in [0,N]^M \cap \mathbb{Z}^M$. Similarly, we know that the polynomial $\mathfrak{C}_1(\mathbf{y}, \mathbf{H})$ is identical to

$$b(\mathbf{0}, \mathbf{y}, \mathbf{z}_0) - b_M(\mathbf{z}_0)$$

for any choice of $\mathbf{z}_0 \in Z(\mathbf{H})$. Therefore, since $\mathcal{R}(\Psi)(\mathbf{z}) = \mathbf{H}$ implies that $\mathbf{z} \in Z(\mathbf{H})$, we can rearrange the system (5.16) and deduce that $\mathcal{W}$ is the number of solutions $\mathbf{y}, \mathbf{y}' \in [0,N]^M \cap \mathbb{Z}^M$ and $\mathbf{z}, \mathbf{z}' \in [0,N]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ of the following system:

$$\begin{aligned} \mathcal{R}(\Psi)(\mathbf{z}) &= \mathcal{R}(\Psi)(\mathbf{z}'), \\ \overline{\mathcal{R}}(\Phi)(\mathbf{y}, \mathbf{z}) &= \overline{\mathcal{R}}(\Phi)(\mathbf{y}', \mathbf{z}), \\ b(\mathbf{0}, \mathbf{y}, \mathbf{z}) - b_M(\mathbf{z}) &= b(\mathbf{0}, \mathbf{y}', \mathbf{z}) - b_M(\mathbf{z}), \\ b_M(\mathbf{z}) &= b_M(\mathbf{z}'). \end{aligned} \tag{5.17}$$

Our result follows from the following two claims.

Claim 1 Given any $c > 0$, there exists $C > 0$ such that the following bound holds,

$$\sup_{\substack{\mathbf{H} \in \mathcal{L}_1(N) \\ \mathbf{G} \in \mathcal{L}_2(N;\mathbf{H})}} \sup_{\alpha \in \mathfrak{m}(C)} |S_0(\alpha, \mathbf{G}, \mathbf{H})| \ll \frac{N^K}{(\log N)^c}.$$

Claim 2 We have the following bound on $\mathcal{W}$,

$$\mathcal{W} \ll N^{2n-2K-2d-D_1-D_2}.$$

By substituting the bounds from the above two claims into (5.15), we obtain that for any $c > 0$ there exists $C > 0$ such that

$$\int_{\mathfrak{m}(C)} T(b;\alpha)\, d\alpha \ll \frac{N^{n-d}}{(\log N)^c},$$

and this completes the proof of our proposition. ∎

Therefore, we only need to establish Claims 1 and 2. Claim 1 is obtained via Weyl differencing. Since the set up for our Claim 1 is the same as that of [2], we omit the proof of Claim 1 and refer the reader to [2, p. 725].

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Proof of Claim 2 We now present the proof of Claim 2. From (5.17), we can write

$$\mathcal{W} = \sum_{\mathbf{z}\in[0,N]^{n-M-K}\cap\mathbb{Z}^{n-M-K}} T_1(\mathbf{z})\cdot T_2(\mathbf{z}),$$

where $T_1(\mathbf{z})$ is the number of solutions $\mathbf{y},\mathbf{y}'\in[0,N]^M\cap\mathbb{Z}^M$ to the system

$$\begin{aligned} b(\mathbf{0},\mathbf{y},\mathbf{z}) &= b(\mathbf{0},\mathbf{y}',\mathbf{z}),\\ \overline{\mathcal{R}}(\Phi)(\mathbf{y},\mathbf{z}) &= \overline{\mathcal{R}}(\Phi)(\mathbf{y}',\mathbf{z}),\end{aligned}$$

and $T_2(\mathbf{z})$ is the number of solutions $\mathbf{z}'\in[0,N]^{n-M-K}\cap\mathbb{Z}^{n-M-K}$ to the system

$$\begin{aligned} b_M(\mathbf{z}) &= b_M(\mathbf{z}'),\\ \mathcal{R}(\Psi)(\mathbf{z}) &= \mathcal{R}(\Psi)(\mathbf{z}').\end{aligned}$$

Define $\mathcal{W}_i=\sum_{\mathbf{z}}T_i(\mathbf{z})^2$ $(i=1,2)$ so that we have $\mathcal{W}^2\le\mathcal{W}_1\mathcal{W}_2$ by the Cauchy–Schwarz inequality. We first estimate $\mathcal{W}_1$, which we can deduce to be the number of solutions $\mathbf{y},\mathbf{y}',\mathbf{u},\mathbf{u}'\in[0,N]^M\cap\mathbb{Z}^M$ and $\mathbf{z}\in[0,N]^{n-M-K}\cap\mathbb{Z}^{n-M-K}$ satisfying the equations

$$\begin{aligned} b(\mathbf{0},\mathbf{y},\mathbf{z})-b(\mathbf{0},\mathbf{y}',\mathbf{z}) &= 0,\\ b(\mathbf{0},\mathbf{u},\mathbf{z})-b(\mathbf{0},\mathbf{u}',\mathbf{z}) &= 0,\\ \overline{\mathcal{R}}(\Phi)(\mathbf{y},\mathbf{z})-\overline{\mathcal{R}}(\Phi)(\mathbf{y}',\mathbf{z}) &= \mathbf{0},\\ \overline{\mathcal{R}}(\Phi)(\mathbf{u},\mathbf{z})-\overline{\mathcal{R}}(\Phi)(\mathbf{u}',\mathbf{z}) &= \mathbf{0}.\end{aligned} \tag{5.18}$$

We consider the h-invariant of the system of forms on the left-hand side of (5.18), and show that it is a regular system. The first two equations of (5.18) are the degree d polynomials of the system, and we let h_d be the h-invariant of these two polynomials. Suppose for some $\lambda,\mu\in\mathbb{Q}$, not both 0, we have

$$\lambda\cdot(f(\mathbf{0},\mathbf{y},\mathbf{z})-f(\mathbf{0},\mathbf{y}',\mathbf{z}))+\mu\cdot(f(\mathbf{0},\mathbf{u},\mathbf{z})-f(\mathbf{0},\mathbf{u}',\mathbf{z}))=\sum_{j=1}^{h_d}U_j\cdot V_j,$$

where $U_j=U_j(\mathbf{y},\mathbf{y}',\mathbf{u},\mathbf{u}',\mathbf{z})$ and $V_j=V_j(\mathbf{y},\mathbf{y}',\mathbf{u},\mathbf{u}',\mathbf{z})$ are rational forms of positive degree $(1\le j\le h_d)$. Without loss of generality, suppose $\lambda\neq0$. Let $\boldsymbol{\ell}=(-\ell_1|_{\mathbf{w}=\mathbf{0}},\dots,-\ell_M|_{\mathbf{w}=\mathbf{0}})$. If we set $\mathbf{u}=\mathbf{u}'=\mathbf{y}'=\boldsymbol{\ell}$, then the above equation becomes

$$g_M(\mathbf{0},\mathbf{y},\mathbf{z})=f(\mathbf{0},\mathbf{y},\mathbf{z})-f_M(\mathbf{z})=\frac{1}{\lambda}\sum_{j=1}^{h_d}U_j(\mathbf{y},\boldsymbol{\ell},\boldsymbol{\ell},\boldsymbol{\ell},\mathbf{z})\cdot V_j(\mathbf{y},\boldsymbol{\ell},\boldsymbol{\ell},\boldsymbol{\ell},\mathbf{z}).$$

Therefore, we obtain from (5.10),

$$h_d\ge h(g_M(\mathbf{0},\mathbf{y},\mathbf{z}))\ge\rho_{d,d}(2+2R_2)+2R_2\ge\rho_{d,d}(2+2R_2-2|\overline{\mathbf{a}}^{(1)}|)+2|\overline{\mathbf{a}}^{(1)}|.$$

For each $1\le i\le d-1$, denote by

$$\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z})-\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y}',\mathbf{z})=\left\{a_j^{(i)}(\mathbf{y},\mathbf{z})-a_j^{(i)}(\mathbf{y}',\mathbf{z}):1\le j\le|\overline{\mathbf{a}}^{(i)}|\right\},$$

the system of degree i polynomials of $\overline{\mathcal{R}}(\Phi)(\mathbf{y},\mathbf{z})-\overline{\mathcal{R}}(\Phi)(\mathbf{y}',\mathbf{z})$. We also define

$$\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u},\mathbf{z})-\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u}',\mathbf{z})$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

in a similar manner. We apply Lemma 3.4 to estimate the h-invariant of the degree i forms of the system (5.18) for each $2 \le i \le d-1$,

$$\begin{aligned}
&h\big(\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y}',\mathbf{z}), \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u}',\mathbf{z})\big)\\
&\quad \ge h\big(\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y}',\mathbf{z}), \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{u}',\mathbf{z}); \mathbf{z}\big)\\
&\quad = h\big(\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y}',\mathbf{z}); \mathbf{z}\big)\\
&\quad \ge h\big(\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z}), \overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y}',\mathbf{z}); \mathbf{z}\big)\\
&\quad \ge h\big(\overline{\mathcal{R}}(\Phi)^{(i)}(\mathbf{y},\mathbf{z}); \mathbf{z}\big)\\
&\quad \ge \rho_{d,i}(2+2R_2) + 2R_2\\
&\quad \ge \rho_{d,i}(2+2R_2-2|\overline{\mathbf{a}}^{(1)}|) + 2|\overline{\mathbf{a}}^{(1)}|.
\end{aligned}$$

We also need to show that the linear forms of the system (5.18) are linearly independent over $\mathbb{Q}$. Recall that the linear forms of $\overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{y},\mathbf{z})$ are linearly independent over $\mathbb{Q}$ and do not include any linear forms that depend only on the $\mathbf{z}$ variables, and similarly for $\overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{y}',\mathbf{z})$, $\overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{u},\mathbf{z})$, and $\overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{u}',\mathbf{z})$. We leave it as an exercise for the reader to verify that the linear forms of

$$\overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{y},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{y}',\mathbf{z}) \bigcup \overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{u},\mathbf{z}) - \overline{\mathcal{R}}(\Phi)^{(1)}(\mathbf{u}',\mathbf{z})$$

are linearly independent over $\mathbb{Q}$.

Therefore, it follows from Corollary 3.3 that $\mathcal{W}_1 \ll N^{n+3M-K-(2d+2D_2)}$.

We now estimate $\mathcal{W}_2$, which we can deduce to be the number of solutions $\mathbf{z}$, $\mathbf{z}'$, $\mathbf{z}'' \in [0,N]^{n-M-K} \cap \mathbb{Z}^{n-M-K}$ satisfying the equations

$$\begin{aligned}
b_M(\mathbf{z}) - b_M(\mathbf{z}') &= 0,\\
b_M(\mathbf{z}) - b_M(\mathbf{z}'') &= 0,\\
\mathcal{R}(\Psi)(\mathbf{z}) - \mathcal{R}(\Psi)(\mathbf{z}') &= \mathbf{0},\\
\mathcal{R}(\Psi)(\mathbf{z}) - \mathcal{R}(\Psi)(\mathbf{z}'') &= \mathbf{0}.
\end{aligned} \tag{5.19}$$

We consider the h-invariant of the system of forms on the left-hand side of (5.19), and show that it is a regular system. The first two equations of (5.19) are the degree d polynomials of the system, and we let h_d be the h-invariant of these two polynomials. Suppose for some $\lambda, \mu \in \mathbb{Q}$, not both 0, we have

$$\lambda \cdot (f_M(\mathbf{z}) - f_M(\mathbf{z}')) + \mu \cdot (f_M(\mathbf{z}) - f_M(\mathbf{z}'')) = \sum_{j=1}^{h_d} U_j \cdot V_j,$$

where $U_j = U_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ and $V_j = V_j(\mathbf{z},\mathbf{z}',\mathbf{z}'')$ are rational forms of positive degree $(1 \le j \le h_d)$. We consider two cases, $(\lambda+\mu) \ne 0$ and $(\lambda+\mu) = 0$. Suppose $(\lambda+\mu) \ne 0$. If we set $\mathbf{z}' = \mathbf{z}'' = \mathbf{0}$, then the above equation becomes

$$(\lambda+\mu) \cdot f_M(\mathbf{z}) = \sum_{j=1}^{h_d} U_j(\mathbf{z},\mathbf{0},\mathbf{0}) \cdot V_j(\mathbf{z},\mathbf{0},\mathbf{0}).$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Thus we obtain $h_d \geq h(f_M(\mathbf{z}))$. On the other hand, suppose $(\lambda + \mu) = 0$. Then the above equation (5.1) simplifies to

$$f_M(\mathbf{z}') - f_M(\mathbf{z}'') = \frac{-1}{\lambda} \sum_{j=1}^{h_d} U_j \cdot V_j.$$

From this equation, we substitute $\mathbf{z}'' = \mathbf{0}$ to obtain $h_d \geq h(f_M(\mathbf{z}'))$. Therefore, in either case we obtain from (5.9) that

$$h_d \geq h(f_M(\mathbf{z})) \geq \rho_{d,d}(2 + 2R_1) + 2R_1 \geq \rho_{d,d}(2 + 2R_1 - 2|\mathbf{v}^{(1)}|) + 2|\mathbf{v}^{(1)}|.$$

Recall that we defined $\mathcal{R}(\Psi) = (\mathbf{v}^{(d-1)}, \ldots, \mathbf{v}^{(1)})$, where $\mathbf{v}^{(i)} = \mathcal{R}^{(i)}(\Psi)$ are the degree i forms of $\mathcal{R}(\Psi)$ $(1 \leq i \leq d-1)$. Take $2 \leq i \leq d-1$. Let $m_i = |\mathbf{v}^{(i)}|$, and we label the forms of $\mathbf{v}^{(i)}$ to be $v_1^{(i)}, \ldots, v_{m_i}^{(i)}$. Let h_i be the h-invariant of the degree i forms of the system (5.19). Then for some $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{Q}^{m_i}$, not both $\mathbf{0}$, we have

$$\sum_{j=1}^{m_i} \lambda_j \cdot (v_j^{(i)}(\mathbf{z}) - v_j^{(i)}(\mathbf{z}')) + \sum_{j=1}^{m_i} \mu_j \cdot (v_j^{(i)}(\mathbf{z}) - v_j^{(i)}(\mathbf{z}'')) = \sum_{t=1}^{h_i} U_t \cdot V_t, \tag{5.20}$$

where $U_t = U_t(\mathbf{z}, \mathbf{z}', \mathbf{z}'')$ and $V_t = V_t(\mathbf{z}, \mathbf{z}', \mathbf{z}'')$ are forms of positive degree $(1 \leq t \leq h_i)$. We consider two cases, $(\boldsymbol{\lambda} + \boldsymbol{\mu}) \neq \mathbf{0}$ and $(\boldsymbol{\lambda} + \boldsymbol{\mu}) = \mathbf{0}$.

Suppose $(\boldsymbol{\lambda} + \boldsymbol{\mu}) \neq \mathbf{0}$. In this case, we set $\mathbf{z}' = \mathbf{z}'' = \mathbf{0}$, and equation (5.20) simplifies to

$$\sum_{j=1}^{m_i} (\lambda_j + \mu_j) \cdot v_j^{(i)}(\mathbf{z}) = \sum_{t=1}^{h_i} U_t(\mathbf{z}, \mathbf{0}, \mathbf{0}) \cdot V_t(\mathbf{z}, \mathbf{0}, \mathbf{0}).$$

Therefore, it follows that

$$h_i \geq h(\mathbf{v}^{(i)}) \geq \rho_{d,i}(2 + 2R_1) + 2R_1 \geq \rho_{d,i}(2 + 2R_1 - 2|\mathbf{v}^{(i)}|) + 2|\mathbf{v}^{(1)}|.$$

On the other hand, suppose $(\boldsymbol{\lambda} + \boldsymbol{\mu}) = \mathbf{0}$. Then equation (5.20) simplifies to

$$\sum_{j=1}^{m_i} -\lambda_j \cdot (v_j^{(i)}(\mathbf{z}') - v_j^{(i)}(\mathbf{z}'')) = \sum_{t=1}^{h_i} U_t \cdot V_t.$$

From this equation, we substitute $\mathbf{z}'' = \mathbf{0}$ to obtain

$$h_i \geq h(\mathbf{v}^{(i)}) \geq \rho_{d,i}(2 + 2R_1) + 2R_1 \geq \rho_{d,i}(2 + 2R_1 - 2|\mathbf{v}^{(i)}|) + 2|\mathbf{v}^{(1)}|.$$

We also need to show that the linear forms of the system (5.19),

$$\left\{\mathbf{v}^{(1)}(\mathbf{z}) - \mathbf{v}^{(1)}(\mathbf{z}')\right\} \cup \left\{\mathbf{v}^{(1)}(\mathbf{z}) - \mathbf{v}^{(1)}(\mathbf{z}'')\right\}, \tag{5.21}$$

are linearly independent over $\mathbb{Q}$. Recall that the linear forms of $\mathbf{v}^{(1)}(\mathbf{z})$ are linearly independent over $\mathbb{Q}$. The linear independence over $\mathbb{Q}$ of the system of linear forms (5.21) follows from this fact, and we leave the verification as an exercise for the reader.

Therefore, we obtain by Corollary 3.3 that $\mathcal{W}_2 \ll N^{3(n-M-K)-(2d+2D_1)}$. Combining the bounds for $\mathcal{W}_1$ and $\mathcal{W}_2$ together, we obtain

$$\mathcal{W} \leq \mathcal{W}_1^{1/2} \mathcal{W}_2^{1/2} \ll N^{2n-2K-(2d+D_1+D_2)},$$

which proves Claim 2. ∎

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

6 Hardy–Littlewood Circle Method: Major Arcs

Recall that $f(\mathbf{x})$ is the degree d portion of the degree d polynomial $b(\mathbf{x}) \in \mathbb{Z}[x_1, \dots, x_n]$. In this section we assume that $f(\mathbf{x})$ satisfies $h(f) > A_d$, where A_d is defined in (5.8). We define $g_d(f)$ as in (4.5) with $\mathbf{f} = \{f\}$ and $r_d = 1$. It then follows from (4.6) that $A_d < h(f) \le (\log 2)^{-d} \cdot d! \cdot g_d(f)$. From this bound and our choice of A_d in (5.8), we have

$$\frac{2^{d-1}}{g_d(f)} < \frac{d!2^{d-1}}{(\log 2)^d A_d} < \frac{d!2^{d-1}}{(\log 2)^d (A_d - 5d)} \le \frac{1}{5(d-1)}.$$

We take Ω to be

$$4 < \Omega < 5 \le \frac{(A_d - 5d) \cdot (\log 2)^d}{2^{d-1}(d-1)d!} \le \frac{g_d(f)}{2^{d-1}(d-1)}.$$

Therefore, with this choice of Ω, we have that $b(\mathbf{x})$ satisfies Hypothesis $(\star)$ with $\mathfrak{B}_0$ by Proposition 4.5. We then choose Q to satisfy $0 < Q < \Omega$ and

$$Q \cdot \frac{2^{d-1}}{g_d(f)} < 1. \tag{6.1}$$

In particular, we can choose Q to satisfy $Q > 4$. We fix these values of Ω and Q throughout this section. We note that with these choices of Ω and Q we have

$$0 < \Omega \le \frac{(A_d - dQ) \cdot (\log 2)^d}{2^{d-1}(d-1)d!}. \tag{6.2}$$

The work of this section is based on [2] and is similar to their treatment of the major arcs. However, we needed to tailor their argument to be in terms of the h-invariant instead of the Birch rank.

We define the sums

$$\begin{aligned} \widetilde{S}_{m,q} &= \sum_{\mathbf{k} \in \mathbb{U}_q^n} e(b(\mathbf{k}) \cdot m/q), \\ B(q) &= \sum_{m \in \mathbb{U}_q} \frac{1}{\phi(q)^n} \widetilde{S}_{m,q}, \\ \mathfrak{S}(N) &= \sum_{q \le (\log N)^C} B(q), \end{aligned} \tag{6.3}$$

where ϕ is Euler's totient function. Recall that we have $\mathfrak{B}_0 = [0,1]^n$. We have the following estimate on the major arcs, which is a consequence of [2, (6.1) and Lemma 6], and we leave the details to the reader. We remark that although it is assumed in [2, Lemma 6] that C is sufficiently large, it in fact follows from their proof that assuming $C > 0$ is sufficient.

Lemma 6.1 ([2, Lemma 6]) *Let $c > 0$, $C > 0$, $q \le (\log N)^C$, and $m \in \mathbb{U}_q$. Then we have*

$$\int_{\mathfrak{M}_{m,q}(C)} T(b;\alpha)\, d\alpha = \frac{1}{\phi(q)^n} \widetilde{S}_{m,q}\, J_0 + O\Big(\frac{N^{n-d}}{(\log N)^c} \Big),$$

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

where

$$J_0 = \int_{|\tau| \le N^{-d}(\log N)^C} \int_{\mathbf{u} \in N\mathfrak{B}_0} e(\tau b(\mathbf{u}))\, \mathbf{du}\, d\tau.$$

Note that J_0 is independent of m and q. We now simplify the expression for J_0. Let $\mathcal{I}(\eta) = \int_{\mathfrak{B}_0} e(\eta f(\boldsymbol{\xi}))\, \mathbf{d\xi}$. For any $\varepsilon > 0$, the inner integral of J_0 can be expressed as

$$\begin{aligned}\int_{\mathbf{u} \in N\mathfrak{B}_0} e(\tau b(\mathbf{u}))\, \mathbf{du} &= \int_{\mathbf{u} \in N\mathfrak{B}_0} e(\tau f(\mathbf{u}))\, \mathbf{du} + O(N^{n-1+\varepsilon}) \\ &= N^n \int_{\boldsymbol{\xi} \in \mathfrak{B}_0} e(N^d \tau f(\boldsymbol{\xi}))\, \mathbf{d\xi} + O(N^{n-1+\varepsilon}) \\ &= N^n \cdot \mathcal{I}(N^d \tau) + O(N^{n-1+\varepsilon}),\end{aligned}$$

where we used the change of variable $\mathbf{u} = N\boldsymbol{\xi}$ to obtain the second equality above.

We define $J(L) = \int_{|\eta| \le L} \mathcal{I}(\eta)\, d\eta$. Then we can simplify J_0 as

$$J_0 = N^{n-d} \cdot J((\log N)^C) + O(N^{n-d-1+\varepsilon}(\log N)^C).$$

Since we have $\Omega > 2$ and Hypothesis $(\star)$, and in particular the restricted Hypothesis $(\star)$, it follows by [4, Lemma 8.1] that

$$\mathcal{I}(\eta) \ll \min(1, |\eta|^{-2}). \tag{6.4}$$

As stated in [4, §3], it follows from (6.4) that $\mu(\infty) = \int_{\mathbb{R}} \mathcal{I}(\eta)\, d\eta$ exists. Furthermore, we have $|\mu(\infty) - J(L)| \ll L^{-1}$. We also have $\mu(\infty) > 0$ if the equation $f(\mathbf{x}) = 0$ has a non-singular real solution in the interior of $\mathfrak{B}_0 = [0,1]^n$ [2, p. 704].

Therefore, we obtain the following estimate as a consequence of the definition of the major arcs and Lemma 6.1.

Lemma 6.2 *Suppose $h(f) > A_d$, where we define A_d as in (5.8). Then, given any $c > 0$, there exists $C > 0$ such that we have*

$$\int_{\mathfrak{M}(C)} T(b;\alpha)\, d\alpha = \mathfrak{S}(N)\mu(\infty)N^{n-d} + O\Big(\mathfrak{S}(N)\frac{N^{n-d}}{(\log N)^C} + \frac{N^{n-d}}{(\log N)^c}\Big).$$

6.1 Singular Series

We obtain the following estimate on the exponential sum $\widetilde{S}_{m,q}$ defined in (6.3).

Lemma 6.3 *Suppose $h(f) > A_d$, where we define A_d as in (5.8). Let p be a prime and let $q = p^t$, $t \in \mathbb{N}$. For $m \in \mathbb{U}_q$, we have the following bounds*

$$\widetilde{S}_{m,q} \ll \begin{cases} q^{n-Q} & \text{if } t \le d, \\ p^Q q^{n-Q} & \text{if } t > d, \end{cases}$$

where the implicit constants are independent of p.

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

Proof We consider the two cases $t \le d$ and $t > d$ separately. We apply the inclusion-exclusion principle to bound $\widetilde{S}_{m,q}$ when $q = p^t$ and $t \le d$. Then

$$\begin{aligned}\widetilde{S}_{m,q} &= \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} \prod_{i=1}^{n}\Big(1 - \sum_{u_i\in\mathbb{Z}/p^{t-1}\mathbb{Z}} \mathbf{1}_{k_i=pu_i}\Big) e(b(\mathbf{k})\cdot m/q) \\ &= \sum_{I\subseteq\{1,2,\dots,n\}} (-1)^{|I|} \sum_{\mathbf{u}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}} \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} F_I(\mathbf{k};\mathbf{u}) e(b(\mathbf{k})\cdot m/q),\end{aligned} \tag{6.5}$$

where $\mathbf{1}_{k_i=pu_i}$ denotes a characteristic function and $F_I(\mathbf{k};\mathbf{u}) = \prod_{i\in I}\mathbf{1}_{k_i=pu_i}$ for $\mathbf{u} \in (\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}$. In other words, $F_I(\mathbf{k};\mathbf{u})$ is the characteristic function of the set $H_{I,\mathbf{u}} = \{\mathbf{k} \in (\mathbb{Z}/q\mathbb{Z})^n : k_i = pu_i\ (i \in I)\}$. We now bound the summand in the final expression of (6.5) by further considering two cases, $|I| \ge tQ$ and $|I| < tQ$. In the first case $|I| \ge tQ$, we use the following trivial estimate

$$\begin{aligned}\Big|\sum_{\mathbf{u}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}} \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} F_I(\mathbf{k};\mathbf{u}) e(b(\mathbf{k})\cdot m/q)\Big| &\le p^{(t-1)|I|}(p^t)^{n-|I|} \\ &= q^{n-|I|/t} \le q^{n-Q}.\end{aligned}$$

On the other hand, suppose $|I| < tQ$. Let $\mathfrak{g}_b(\mathbf{x})$ be the polynomial obtained by substituting $x_i = pu_i$ $(i \in I)$ to $b(\mathbf{x})$. Thus $\mathfrak{g}_b(\mathbf{x})$ is a polynomial in $n-|I|$ variables. We can also easily deduce that the degree d portion of $\mathfrak{g}_b(\mathbf{x})$, which we denote by $f_{\mathfrak{g}_b}$, is obtained by substituting $x_i = 0$ $(i \in I)$ to the degree d portion of $b(\mathbf{x})$. Hence, we have $f_{\mathfrak{g}_b} = f|_{x_i=0\,(i\in I)}$. Consequently, by Lemma 2.1 we obtain that $h(f_{\mathfrak{g}_b}) \ge h(f) - |I| > h(f) - dQ > A_d - dQ$. By our choice of Q and Ω, and from (4.6) and (6.2), we have

$$0 < Q < \Omega < \frac{h(f_{\mathfrak{g}_b})\cdot(\log 2)^d}{2^{d-1}(d-1)d!} \le \frac{g_d(f_{\mathfrak{g}_b})}{2^{d-1}(d-1)}.$$

Therefore, with these notations we have by Lemma 4.6 that

$$\begin{aligned}\sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} F_I(\mathbf{k};\mathbf{u}) e(b(\mathbf{k})\cdot m/q) &= \sum_{\mathbf{s}\in(\mathbb{Z}/q\mathbb{Z})^{n-|I|}} e(\mathfrak{g}_b(\mathbf{s})\cdot m/q) \\ &= q^{n-|I|}E(\mathfrak{g}_b, q; m/q) \ll q^{n-|I|-Q}.\end{aligned}$$

Thus, we obtain

$$\sum_{\mathbf{u}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^{|I|}} \sum_{\mathbf{k}\in(\mathbb{Z}/q\mathbb{Z})^n} F_I(\mathbf{k};\mathbf{u}) e(b(\mathbf{k})\cdot m/q) \ll (p^{t-1})^{|I|} q^{n-|I|-Q} \le q^{n-Q}.$$

Consequently, combining the two cases $|I| \ge tQ$ and $|I| < tQ$, we obtain $\widetilde{S}_{m,q} \ll q^{n-Q}$ when $t \le d$.

We now consider the case $q = p^t$ when $t > d$. By the definition of $\widetilde{S}_{m,q}$, we have

$$\begin{aligned}\widetilde{S}_{m,q} &= \sum_{\mathbf{k}'\in\mathbb{U}_p^n} \sum_{\mathbf{s}\in(\mathbb{Z}/p^{t-1}\mathbb{Z})^n} e(b(\mathbf{k}'+p\mathbf{s})\cdot m/q) \\ &= \sum_{\mathbf{k}'\in\mathbb{U}_p^n} \sum_{\mathbf{s}\in[0,p^{t-1})^n} e(b(\mathbf{k}'+p\mathbf{s})\cdot m/q).\end{aligned} \tag{6.6}$$

For each fixed $\mathbf{k}' \in \mathbb{U}_p^n$, we have $b(\mathbf{k}'+p\mathbf{s}) = p^d f(\mathbf{s}) + \chi_{p,\mathbf{k}'}(\mathbf{s})$, where $\chi_{p,\mathbf{k}'}(\mathbf{x})$ is a polynomial of degree at most $d-1$ and its coefficients depend on p and $\mathbf{k}'$. We apply Corollary 4.4 with $r_d = 1$, $\psi(\mathbf{x}) = f(\mathbf{x}) + \frac{1}{p^d}\chi_{p,\mathbf{k}'}(\mathbf{x})$, $\alpha = m/p^{t-d}$, $\mathfrak{B} = [0,1)^n$,

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

and $P = p^{t-1}$. Let $\varepsilon' > 0$ be sufficiently small. Recall from (6.1) that our choice of $Q > 0$ satisfies

$$Q \cdot \frac{2^{d-1}}{g_d(f)} < 1.$$

Let γ_d and γ_d' be as in the paragraph before Corollary 4.4 with $\mathbf{f} = \{f\}$ and $r_d = 1$. Suppose the alternative (ii) of Corollary 4.4 holds. Then we know there exists $n_0 \in \mathbb{N}$ such that $n_0 \ll (p^{t-1}-1)^{Q\gamma_d+\varepsilon'}$ and

$$\|n_0(m/p^{t-d})\| \ll (p^{t-1}-1)^{-d+Q\gamma_d+\varepsilon'} \le \Big(\frac{1}{2}p^{t-1}\Big)^{-d+Q\gamma_d+\varepsilon'}. \tag{6.7}$$

However, this is not possible once p^t is sufficiently large with respect to n, d, ε', Q, and f, for the following reason. First note that n_0 cannot be divisible by p^{t-d} for p^t sufficiently large because $Q\gamma_d + \varepsilon' < Q\gamma_d' < 1$. Since $n_0 \in \mathbb{N}$ is not divisible by p^{t-d} and $(m, p) = 1$, we have

$$\|n_0(m/p^{t-d})\| \ge \frac{1}{p^{t-d}},$$

which contradicts (6.7) for p^t sufficiently large. Thus by Corollary 4.4, we can bound the inner sum of (6.6) by

$$\sum_{\mathbf{s}\in[0,p^{t-1})^n} e\Big(\Big(f(\mathbf{s}) + \frac{1}{p^d}\,\chi_{p,\mathbf{k}'}(\mathbf{s})\Big)\cdot m/p^{t-d}\Big) \ll (p^{t-1})^{n-Q},$$

where the implicit constant depends at most on n, d, ε', Q, and f. Therefore, we can bound (6.6) as follows

$$\begin{aligned}\widetilde{S}_{m,q} &\le \sum_{\mathbf{k}'\in\mathbb{U}_p^n}\Big|\sum_{\mathbf{s}\in[0,p^{t-1})^n} e\Big(\Big(f(\mathbf{s}) + \frac{1}{p^d}\chi_{p,\mathbf{k}'}(\mathbf{s})\Big)\cdot m/p^{t-d}\Big)\Big| \\ &\ll p^n(p^{t-1})^{n-Q} = p^Q q^{n-Q}.\end{aligned}$$

■

For each prime p, we define $\mu(p) = 1 + \sum_{t=1}^{\infty} B(p^t)$, which converges absolutely provided that $h(f) > A_d$, as we see in the following lemma. As stated in [2], by following the outline of L. K. Hua [3, Chapter VII, §2, Lemma 8.1] one can show that $B(q)$ is a multiplicative function of q. Therefore, we consider the following identity

$$\mathfrak{S}(\infty) := \lim_{N\to\infty}\mathfrak{S}(N) = \prod_{p \text{ prime}}\mu(p). \tag{6.8}$$

Lemma 6.4 *There exists $\delta_1 > 0$ such that for each prime p, we have $\mu(p) = 1 + O(p^{-1-\delta_1})$, where the implicit constant is independent of p. Furthermore, we have $|\mathfrak{S}(N) - \mathfrak{S}(\infty)| \ll (\log N)^{-C\delta_2}$ for some $\delta_2 > 0$.*

Therefore, the limit in (6.8) exists, and the product in (6.8) converges. We leave the details that these two quantities are equal to the reader.

Proof Recall that our choice of Q satisfies $Q > 4$. Let $\varepsilon_0 > 0$ be sufficiently small such that $\widetilde{Q} = Q - \varepsilon_0 > 4 \ge 2d/(d-1)$. We substitute $Q = \widetilde{Q} + \varepsilon_0$ into the bounds in Lemma 6.3. It is then clear that we can assume that the implicit constant in Lemma 6.3 is 1 for p sufficiently large with the cost of using $\widetilde{Q}$ in place of Q. For any $t \in \mathbb{N}$, we

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press

know that $\phi(p^t) = p^t(1-1/p) \geq \frac{1}{2}p^t$. Therefore, by considering the two cases as in Lemma 6.3, we obtain

$$\begin{aligned}|\mu(p)-1| &\ll \sum_{1\leq t\leq d} p^t p^{-nt} p^{nt-t\widetilde{Q}} + \sum_{t>d} p^t p^{-nt} p^{\widetilde{Q}+nt-t\widetilde{Q}} \\ &\ll p^{1-\widetilde{Q}} + p^{\widetilde{Q}} p^{-(d+1)(\widetilde{Q}-1)} \ll p^{-1-\delta_1},\end{aligned}$$

for some $\delta_1 > 0$. We note that the implicit constants in $\ll$ are independent of p here.

Let $q = p_1^{t_1}\cdots p_v^{t_v}$ be the prime factorization of $q \in \mathbb{N}$. Without loss of generality, suppose we have $t_j \leq d$ $(1 \leq j \leq \nu_0)$ and $t_j > d$ $(\nu_0 < j \leq \nu)$. By the multiplicativity of $B(q)$, it also follows from Lemma 6.3 that

$$B(q) = B(p_1^{t_1})\cdots B(p_v^{t_v}) \ll q^{1-\widetilde{Q}} \cdot \Big(\prod_{j=\nu_0+1}^{\nu} p_j^{\widetilde{Q}}\Big) \leq q^{1-\widetilde{Q}} \cdot q^{\widetilde{Q}/d} \leq q^{-1-\delta_2},$$

for some $\delta_2 > 0$. We note that the implicit constant in $\ll$ is independent of q here, because the implicit constant in Lemma 6.3 is 1 for p sufficiently large, as mentioned above. Therefore, we obtain

$$|\mathfrak{S}(N) - \mathfrak{S}(\infty)| \leq \sum_{q>(\log N)^C} |B(q)| \ll \sum_{q>(\log N)^C} q^{-1-\delta_2} \ll (\log N)^{-C\delta_2}. \qquad \blacksquare$$

Let $v_t(p)$ denote the number of solutions $\mathbf{x} \in (\mathbb{U}_{p^t})^n$ to the congruence

$$b(\mathbf{x}) \equiv 0 \pmod{p^t}.$$

It can be deduced that

$$1 + \sum_{j=1}^{t} B(p^j) = \frac{1}{\phi(p^t)^n} \sum_{\mathbf{k}\in(\mathbb{U}_{p^t})^n} \sum_{m\in\mathbb{Z}/p^t\mathbb{Z}} e(b(\mathbf{k})\cdot m/p^t) = \frac{p^t}{\phi(p^t)^n} v_t(p).$$

Therefore, provided $h(b) > A_d$, we obtain

$$\mu(p) = \lim_{t\to\infty} \frac{p^t v_t(p)}{\phi(p^t)^n}.$$

At this point we refer the reader to [2, p. 704, 736] to conclude that $\mu(p) > 0$ if the equation $b(\mathbf{x}) = 0$ has a non-singular solution in $\mathbb{Z}_p^\times$, the units of p-adic integers. It then follows from Lemma 6.4 that if the equation $b(\mathbf{x}) = 0$ has a non-singular solution in $\mathbb{Z}_p^\times$ for every prime p, then $\prod_{p \text{ prime}} \mu(p) > 0$. Finally, we let $C_b = \mu(\infty)\prod_{p \text{ prime}} \mu(p)$, and Theorem 1.1 follows as a consequence of Lemmas 6.2 and 6.4 and Proposition 5.1.

Acknowledgments We would like to thank D. Schindler and the anonymous referees for many helpful comments. We would also like to thank the Department of Pure Mathematics at the University of Waterloo for their support as portions of this work were completed while both of the authors were there as graduate students.

References

[1] B. J. Birch, *Forms in many variables*. Proc. Roy. Soc. Ser. A 265(1961/1962), 245–263. https://doi.org/10.1098/rspa.1962.0007
[2] B. Cook and Á. Magyar, *Diophantine equations in the primes*. Invent. Math. 198(2014), 701–737. https://doi.org/10.1007/s00222-014-0508-1
[3] L. K. Hua, *Additive theory of prime numbers*. *(Translations of Mathematical Monographs, 13)*, American Mathematical Society, Providence, RI, 1965.
[4] W. M. Schmidt, *The density of integer points on homogeneous varieties*. Acta Math. 154(1985), 3–4, 243–296. https://doi.org/10.1007/978-1-4939-3201-6_9

Department of Mathematics, University of Toronto, Toronto, ON, Canada
e-mail : syxiao@math.toronto.edu

Department of Mathematics & Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
e-mail : sy46@queensu.ca

https://doi.org/10.4153/S0008414X19000026 Published online by Cambridge University Press