TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 361, Number 6, June 2009, Pages 2929–2957
S 0002-9947(09)04821-1
Article electronically published on January 27, 2009

ASYMPTOTIC ESTIMATES FOR RATIONAL LINEAR SPACES ON HYPERSURFACES

SCOTT T. PARSELL

Abstract. We develop a repeated efficient differencing procedure for estimating mean values of certain multidimensional exponential sums over smooth numbers. As a consequence, we obtain asymptotic lower bounds for the number of linear spaces of fixed dimension up to a given height lying on the hypersurface defined by an additive equation.

1. Introduction

The problem of counting integral points lying on the hypersurface defined by an additive equation has occupied a prominent place in number theory over the past century. Specifically, one often asks how large s must be in terms of k in order to ensure that the hypersurface

$$c_1 z_1^k + \cdots + c_s z_s^k = 0 \tag{1.1}$$

contains a non-trivial integral point for all integers $c_1, \dots, c_s$. Frequently, one also wishes to establish asymptotic estimates for the number of integral points lying within a box as the box size tends to infinity. Subject to a local solubility hypothesis, the ground-breaking work of Wooley [12] on Waring's problem can be used to show that the number of integral solutions of (1.1) in the box $[-P, P]^s$ has the expected order of magnitude of P^{s-k} whenever $s \geq (1+o(1))k \log k$. Moreover, an asymptotic formula for the number of solutions can be established when $s \geq (1+o(1))k^2 \log k$, and in this case no local solubility hypothesis is needed (except for indefiniteness) since a classical result of Davenport and Lewis [5] shows that k^2+1 variables suffice to satisfy the congruence conditions.

Because of the homogeneity of (1.1), if the hypersurface in question contains one non-trivial integral point, then it contains all scalar multiples of that point as well. One may choose to express this by saying that the hypersurface contains a rational linear space of projective dimension zero, and it is therefore natural to ask about linear spaces of higher dimension. While results concerning the existence of such spaces date to work of Brauer [4] and Birch [3], asymptotic estimates for the number of such spaces up to a given height seem to have been considered only in recent work of the author (see [6], [7], and [9]). If $\mathbf{x}_1, \dots, \mathbf{x}_d$ are linearly independent vectors in $\mathbb{Z}^s$, then we are interested in determining whether the linear space of projective dimension $d-1$ spanned by these vectors is contained in (1.1). By collecting the

Received by the editors January 8, 2007.
2000 *Mathematics Subject Classification.* Primary 11D45, 11D72; Secondary 11L07, 11P55.
The author was supported in part by a National Science Foundation Postdoctoral Fellowship (DMS-0102068) and by a grant from the Holcomb Research Institute.

©2009 American Mathematical Society

coefficients of $t_1^{i_1}\cdots t_d^{i_d}$ for each d-tuple $(i_1,\dots,i_d)$ satisfying $i_1+\cdots+i_d=k$, one sees that this occurs if and only if $\mathbf{x}_1,\dots,\mathbf{x}_d$ satisfy the system of equations

$$c_1x_{11}^{i_1}\cdots x_{1d}^{i_d}+\cdots+c_sx_{s1}^{i_1}\cdots x_{sd}^{i_d}=0 \qquad (i_1+\cdots+i_d=k). \tag{1.2}$$

We note that the number of equations in the above system is given by

$$\ell=\binom{k+d-1}{k}. \tag{1.3}$$

We shall frequently abbreviate a monomial of the shape $x_1^{i_1}\cdots x_d^{i_d}$ by $\mathbf{x}^{\mathbf{i}}$, and we shall also write $|\mathbf{i}|=i_1+\cdots+i_d$. Our strategy for counting solutions of (1.2) is to focus on solutions in which most of the variables are free of large prime factors. Thus our main tool will be the exponential sum

$$f(\boldsymbol{\alpha})=f(\boldsymbol{\alpha};P,R)=\sum_{x_1,\dots,x_d\in\mathcal{A}(P,R)}e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right),$$

where $e(y)=e^{2\pi iy}$, and where

$$\mathcal{A}(P,R)=\{n\in[1,P]\cap\mathbb{Z}:p|n,\ p\text{ prime }\Rightarrow p\le R\}$$

denotes the set of R-smooth numbers up to P. In order to account for negative solutions to (1.2), we let $f^*(\boldsymbol{\alpha})$ denote the analogue of $f(\boldsymbol{\alpha})$ in which the variables $x_1,\dots,x_d$ range over $\pm\mathcal{A}(P,R)\cup\{0\}$. By orthogonality, the number of solutions $N_{s,k,d}(P)$ of the system (1.2) with $x_{ij}\in[-P,P]\cap\mathbb{Z}$ satisfies

$$N_{s,k,d}(P)\ge\int_{\mathbb{T}^\ell}\prod_{j=1}^{s}f^*(c_j\boldsymbol{\alpha})\,d\boldsymbol{\alpha},$$

where $\mathbb{T}^\ell$ denotes the ℓ-dimensional unit hypercube. Our aim is to show that $N_{s,k,d}(P)\gg P^{sd-k\ell}$ whenever s is sufficiently large in terms of k and d. This then leads to a similar estimate for the number of linear spaces of height at most P lying on (1.1), except that each space is counted with a weight equal to the number of different bases. We return to the issue of counting distinct spaces later in this section.

In order to count solutions of the system (1.2) via the Hardy-Littlewood method, one needs upper bounds for the number of solutions of an auxiliary symmetric system. We find it convenient to do the bulk of our analysis with the exponential sum $f(\boldsymbol{\alpha})$, which restricts us for the moment to positive solutions. We let $S_{s,k,d}(P,R)$ denote the number of solutions of the system

$$\mathbf{x}_1^{\mathbf{i}}+\cdots+\mathbf{x}_s^{\mathbf{i}}=\mathbf{y}_1^{\mathbf{i}}+\cdots+\mathbf{y}_s^{\mathbf{i}} \qquad (|\mathbf{i}|=k)$$

with $x_{ij},y_{ij}\in\mathcal{A}(P,R)$, and we observe that

$$S_{s,k,d}(P,R)=\int_{\mathbb{T}^\ell}|f(\boldsymbol{\alpha};P,R)|^{2s}d\boldsymbol{\alpha}.$$

Before considering upper bounds for $S_{s,k,d}(P,R)$, it is useful to derive an elementary lower bound. Let $S_{s,k,d}(P,R;\mathbf{h})$ denote the number of solutions of the system

$$\sum_{m=1}^{s}(\mathbf{x}_m^{\mathbf{i}}-\mathbf{y}_m^{\mathbf{i}})=h_{\mathbf{i}} \qquad (|\mathbf{i}|=k)$$

with $\mathbf{x}_m, \mathbf{y}_m \in \mathcal{A}(P,R)^d$, and observe that

$$S_{s,k,d}(P,R;\mathbf{h}) = \int_{\mathbb{T}^\ell} |f(\boldsymbol{\alpha};P,R)|^{2s} e(-\boldsymbol{\alpha}\cdot\mathbf{h})\, d\boldsymbol{\alpha} \le S_{s,k,d}(P,R).$$

Thus, by summing over all values of $\mathbf{h}$ for which $S_{s,k,d}(P,R;\mathbf{h})$ is non-zero, we find that

$$|\mathcal{A}(P,R)|^{2sd} \ll P^{k\ell} S_{s,k,d}(P,R).$$

If R is at least a positive power of P, then it is well known (see for example [11], section 12.1) that $\mathcal{A}(P,R) \gg P$, so in this case we have

$$S_{s,k,d}(P,R) \gg P^{2sd-k\ell}. \tag{1.4}$$

By considering diagonal solutions, one also obtains the lower bound $S_{s,k,d}(P,R) \gg P^{sd}$, but the expression in (1.4) dominates whenever $s > k\ell/d$. Moreover, a heuristic argument suggests that $P^{2sd-k\ell}$ represents the true order of magnitude, since there are $O(P^{2sd})$ choices for the variables and a random choice should satisfy each of the ℓ equations (independently) with probability $O(P^{-k})$.

Thus we aim to establish estimates of the shape

$$S_{s,k,d}(P,R) \ll P^{2sd-k\ell+\Delta_s+\varepsilon}, \tag{1.5}$$

where $\Delta_s = \Delta_{s,k,d}$ is small whenever s is sufficiently large in terms of k and d. If the estimate (1.5) holds with $R = P^\eta$ whenever η is sufficiently small, we say that Δ_s is an *admissible* exponent for (s,k,d). Our main theorem in this direction is the following.

Theorem 1.1. *Suppose that k is sufficiently large in terms of d, and write*

$$s_0 = dk\ell\left(\tfrac{1}{2}\log(dk) - \log\log k\right). \tag{1.6}$$

Then the estimate (1.5) *holds with*

$$\Delta_s = \begin{cases} dk\ell\, e^{3-2s/(dk\ell)} & \text{if } 1 \le s \le s_0, \\ \ell\,(\log k)^2\, e^{3-3(s-s_0)/(2dk\ell)} & \text{if } s > s_0. \end{cases}$$

We now state our estimate for $N_{s,k,d}(P)$ that follows from Theorem 1.1 through an application of the Hardy-Littlewood method.

Theorem 1.2. *Suppose that k is sufficiently large in terms of d and that $s \ge s_1$, where*

$$s_1 = dk\ell(\tfrac{4}{3}\log(k\ell) + \log(dk) + 2\log\log k + 8). \tag{1.7}$$

Further suppose that the system (1.2) *has a non-singular real solution and a non-singular p-adic solution for every prime p. Then there are positive constants $\mathcal{C} = \mathcal{C}(s,k,d;\mathbf{c})$ and $P_0 = P_0(s,k,d;\mathbf{c})$ such that whenever $P \ge P_0$ one has*

$$N_{s,k,d}(P) \ge \mathcal{C} P^{sd-k\ell}.$$

A simple counting argument shows that the number of choices for $\mathbf{x}_1, \dots, \mathbf{x}_d \in [-P,P]^s \cap \mathbb{Z}^s$ that are linearly dependent over $\mathbb{Q}$ is $O(P^{sd-s+d})$. Whenever $s > k\ell+d$, as is the case in (1.7), this is of smaller order of magnitude than $P^{sd-k\ell}$, so the conclusion of Theorem 1.2 holds with $N_{s,k,d}(P)$ replaced by $N^*_{s,k,d}(P)$, the number of solutions of (1.2) for which the vectors $\mathbf{x}_1, \dots, \mathbf{x}_d$ are linearly independent.

Heights for subspaces. In view of the preceding remark, Theorem 1.2 may be interpreted as providing lower bounds for the number of linear spaces of bounded

"height" lying on the hypersurface (1.1). We now aim to make this assertion somewhat more precise. First of all, we define the height of a vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ to be

$$H(\mathbf{x}) = \frac{\max_{1 \le i \le n} |x_i|}{\gcd(x_1, \dots, x_n)}.$$

Notice that this height is also well defined on points viewed as elements of projective space. Now for a subspace $\mathcal{L} \subset \mathbb{Z}^s$ with basis vectors $\mathbf{x}_1, \dots, \mathbf{x}_d$, we write

$$H(\mathcal{L}) = H(\mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_d),$$

where we set $n = \binom{s}{d}$ and identify the wedge product with its natural embedding in $\mathbb{Z}^n$. If $\mathbf{y}_1, \dots, \mathbf{y}_d$ is another basis of $\mathcal{L}$, then we have $Y = XB$, where X and Y denote the $s \times d$ matrices corresponding to each basis and where B is an invertible $d \times d$ change-of-basis matrix. Since

$$\mathbf{y}_1 \wedge \cdots \wedge \mathbf{y}_d = (\det B)\ \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_d,$$

we see that our definition of height does not depend on the basis.

Let $\mathcal{N}_{s,k,d}(P)$ denote the number of distinct linear spaces $\mathcal{L}$, with projective dimension $d-1$ and height at most P, lying on the hypersurface (1.1). If $\beta_Q(\mathcal{L})$ denotes the number of integral bases for $\mathcal{L}$ with all components bounded by $Q = (P/d!)^{1/d}$, then we have

$$N^*_{s,k,d}(Q) \ll \sum_{H(\mathcal{L}) \le P} \beta_Q(\mathcal{L}) \le \Big(\max_{\mathcal{L}} \beta_Q(\mathcal{L}) \Big) \mathcal{N}_{s,k,d}(P).$$

The number of possibilities for the change-of-basis matrix can be estimated by choosing a prime p with $Q < p \le 2Q$ and viewing the computations over $\mathbb{F}_p$. It is then clear that $\beta_Q(\mathcal{L}) \ll p^{d^2} \ll Q^{d^2}$, so we have

$$\mathcal{N}_{s,k,d}(P) \gg Q^{sd - k\ell - d^2} \gg P^{s - k\ell/d - d}$$

whenever the conditions of Theorem 1.2 hold. This provides an estimate of the type advertised in the title. It seems that a more sophisticated approach would be required to obtain asymptotic formulas for $\mathcal{N}_{s,k,d}(P)$ from the results of [9].

The author expresses his sincere thanks to Bob Vaughan and Trevor Wooley for constant encouragement and for many helpful discussions concerning this problem and related ideas. The author also acknowledges the useful comments of the referee.

2. Preliminary estimates

Fundamental to our iterative method is an estimate for the number of non-singular solutions to an associated system of congruences. In order to retain adequate control over the singular solutions, however, we are forced to work with systems somewhat smaller than (1.2). We find it convenient to place the indices $\mathbf{i}$ in lexicographic order, so that one writes $\mathbf{i} \prec \mathbf{j}$ if and only if there exists l with $0 \le l < d$ such that $i_1 = j_1, \dots, i_l = j_l$ and $i_{l+1} < j_{l+1}$. We temporarily think of j as being fixed and write $\mathbf{j}_1$ for the multi-index $(j, 0, \dots, 0)$. Further, let

$$\ell_j = \binom{k - j + d - 1}{d - 1} \tag{2.1}$$

denote the number of equations in (1.2) with $\mathbf{i} \succ \mathbf{j}_1$. It turns out that these equations form a maximal set to which we can apply the argument in §3 for counting singular solutions with j efficient differences.

Suppose that $f_{\mathbf{i}}(\mathbf{x})$ is a polynomial in t variables with $t \ge \ell_j$, and let $\mathbf{u} \in \mathbb{Z}^{\ell_j}$. When $\mathbf{q} \in \mathbb{N}^d$, we again adopt the notation $\mathbf{q}^{\mathbf{i}} = q_1^{i_1} \cdots q_d^{i_d}$ and further write $q = q_1 \cdots q_d$. We now define $\mathcal{B}_{\mathbf{q}}(\mathbf{f}; \mathbf{u})$ to be the number of non-singular solutions $\mathbf{x} \in (\mathbb{Z}/q^k\mathbb{Z})^t$ of the system of congruences

$$f_{\mathbf{i}}(\mathbf{x}) \equiv u_{\mathbf{i}} \pmod{\mathbf{q}^{\mathbf{i}}} \qquad (\mathbf{i} \succ \mathbf{j}_1,\ |\mathbf{i}| = k).$$

By a non-singular solution, we mean a solution for which the Jacobian matrix $(\partial f_{\mathbf{i}}/\partial x_l)$ of the left-hand side has at least one $\ell_j \times \ell_j$ sub-matrix whose determinant is relatively prime to q.

Lemma 2.1. *Suppose that $f_{\mathbf{i}} \in \mathbb{Z}[x_1, \dots, x_t]$, where $t \ge \ell_j$. Further let $\mathbf{u} \in \mathbb{Z}^{\ell_j}$ and $\mathbf{q} \in \mathbb{N}^d$, and write $q = q_1 \cdots q_d$. Then one has*

$$\operatorname{card} \mathcal{B}_{\mathbf{q}}(\mathbf{f}; \mathbf{u}) \ll q^{kt - \frac{\ell_j}{d}(k-j)+\varepsilon} q_1^{-j\ell_j}.$$

Proof. We start by choosing integers $a_{\mathbf{i}} \equiv u_{\mathbf{i}} \pmod{\mathbf{q}^{\mathbf{i}}}$ with $1 \le a_{\mathbf{i}} \le q^k$ for each $\mathbf{i}$ with $\mathbf{i} \succ \mathbf{j}_1$ and $|\mathbf{i}| = k$. Since the number of prime divisors of q is $O(\log q/\log\log q)$, it follows from the main theorem of Wooley [15] and the Chinese Remainder Theorem that the number of non-singular solutions of the system of congruences

$$f_{\mathbf{i}}(\mathbf{x}) \equiv a_{\mathbf{i}} \pmod{q^k} \qquad (\mathbf{i} \succ \mathbf{j}_1,\ |\mathbf{i}| = k)$$

is $O(q^{k(t-\ell_j)+\varepsilon})$ for each choice of $\mathbf{a}$. Now the number of choices for $\mathbf{a}$ is $q_1^{\omega_1} \cdots q_d^{\omega_d}$, where

$$\omega_m = \sum_{\mathbf{i} \succ \mathbf{j}_1, |\mathbf{i}| = k} (k - i_m).$$

Furthermore, the number of indices $\mathbf{i}$ with $\mathbf{i} \succ \mathbf{j}_1$ and $i_m = r$ is

$$\binom{k-r-j+d-2}{d-2}$$

whenever $2 \le m \le d$,

$$\binom{k-r+d-2}{d-2}$$

if $m = 1$ and $r \ge j$, and zero if $m = 1$ and $r < j$. Thus we obtain the formulas

$$\omega_m = \sum_{r=0}^{k-j} (k-r) \binom{k-r-j+d-2}{d-2} = \sum_{l=0}^{k-j} (l+j) \binom{l+d-2}{d-2}$$

for $2 \le m \le d$ and

$$\omega_1 = \sum_{r=j}^{k} (k-r) \binom{k-r+d-2}{d-2} = \sum_{l=0}^{k-j} l \binom{l+d-2}{d-2}.$$

We now observe that

$$\omega_1 = (d-1) \sum_{l=1}^{k-j} \binom{l+d-2}{l-1} = (d-1) \sum_{l=1}^{k-j} \left[\binom{l+d-1}{l-1} - \binom{l+d-2}{l-2} \right],$$

with the convention that $\binom{m}{n} = 0$ when $n < 0$. The latter sum telescopes to give

$$\omega_1 = (d-1) \binom{k-j+d-1}{k-j-1} = \tfrac{1}{d}(d-1)(k-j)\ell_j,$$

and it is easy to see that $\omega_m = \omega_1 + j\ell_j$ for $m \geq 2$. Thus for $m \geq 2$ one has

$$k(t-\ell_j) + \omega_m = kt - k\ell_j + k\ell_j(1-\tfrac{1}{d}) - j\ell_j(1-\tfrac{1}{d}) + j\ell_j = kt - \frac{\ell_j}{d}(k-j),$$

and the lemma follows. □

Definition 2.2. When $0 \leq j < k$, we say that $(\boldsymbol{\Psi})$ is of type (j, P, A) if

(1) The system consists of polynomials $\Psi_{\mathbf{i}} \in \mathbb{Z}[z_1, \dots, z_d]$, each of total degree $k-j$, indexed by the vectors $\mathbf{i}$ satisfying $|\mathbf{i}| = k$.

(2) The coefficient of each term of degree $k-j$ in $\Psi_{\mathbf{i}}$ is bounded by AP^j.

(3) For each $\mathbf{i}$ with $\mathbf{i} \succ (j, 0, \dots, 0)$, the polynomial $\Psi_{\mathbf{i}}$ contains a term of degree $k-j$ that does not appear explicitly in any of the $\Psi_{\mathbf{i}'}$ with $\mathbf{i}' \succ \mathbf{i}$.

Write $s_0(n)$ for the square-free kernel of the integer n, defined to be the product of all primes dividing n. We conclude this section by recalling an estimate for the number of integers in an interval with a given square-free kernel.

Lemma 2.3. *Suppose that X is a positive real number and n is a positive integer such that $\log n \ll \log X$. Then, for every $\varepsilon > 0$, one has*

$$\operatorname{card}\{y \leq X : s_0(y) = s_0(n)\} \ll X^{\varepsilon}.$$

Proof. This is Lemma 2.1 of Wooley [12]. □

3. Efficient differencing

Our goal in this section is to develop an iterative method for bounding $S_{s,k,d}(P, R)$ as s increases, and it is convenient to increase s to $s + \ell$, where ℓ is as in (1.3), at each iteration. Moreover, within each iteration, we aim to employ a repeated differencing process that injects new congruence information at each stage.

We suppose throughout that k is sufficiently large in terms of d and that P is sufficiently large in terms of k. We let $\boldsymbol{\Psi}$ be a system of type (j, P, A) for some constant A, where $0 \leq j < k$. We further let $C_1, \dots, C_u$ be constants and write $\widetilde{C} = C_1 C_2 \cdots C_u$. For each $\mathbf{i}$ with $|\mathbf{i}| = k$, we let $D_{\mathbf{i}} \in \mathbb{Z}[x_1, \dots, x_u]$ have the property that $D_{\mathbf{i}}(\mathbf{c}) \neq 0$ whenever $1 \leq |c_l| \leq C_l$. Generally, the variables ε and η denote small positive numbers whose values may change from statement to statement. Typically, η will be chosen sufficiently small in terms of ε, and the implicit constants in our analysis may depend at most on ε, η, s, k, and d. Since our methods involve only a finite number of steps, all implicit constants that arise remain under control, and the values assumed by η remain uniformly bounded away from zero.

We let $S_{s,r}(P, Q, R; \boldsymbol{\Psi})$ denote the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n(\Psi_{\mathbf{i}}(\mathbf{z}_n, \mathbf{c}) - \Psi_{\mathbf{i}}(\mathbf{w}_n, \mathbf{c})) = D_{\mathbf{i}}(\mathbf{c}) \sum_{m=1}^{s} (\mathbf{x}_m^{\mathbf{i}} - \mathbf{y}_m^{\mathbf{i}}) \qquad (|\mathbf{i}| = k) \tag{3.1}$$

with $1 \leq z_{nl}, w_{nl} \leq P$, with $x_{ml}, y_{ml} \in \mathcal{A}(Q, R)$, with $\eta_n \in \{\pm 1\}$, and with $1 \leq c_l \leq C_l$. Note that we have suppressed the dependence on k, d, $\mathbf{C}$, and $\mathbf{D}$ for simplicity; likewise, we shall often abbreviate $S_{s,k,d}(P, R)$ by $S_s(P, R)$.

We further write $\operatorname{Jac}(\boldsymbol{\Psi}; \mathbf{z}, \mathbf{w}, \mathbf{c})$ for the $\ell_j \times 2rd$ Jacobian matrix formed with the polynomials on the left-hand side for which $\mathbf{i} \succ \mathbf{j}_1$, and we write $\operatorname{Jac}(\boldsymbol{\Psi}; \mathbf{z}, \mathbf{c})$

and $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{w},\mathbf{c})$ for the corresponding $\ell_j \times rd$ Jacobians. Here we think of $\mathbf{z}$ as $(\mathbf{z}_1,\dots,\mathbf{z}_r)$.

When $\mathbf{u}\in\mathbb{Z}^m$ and $\mathbf{v}\in\mathbb{Z}^n$ with $m\le n$, we write $\mathbf{u}\hookrightarrow\mathbf{v}$ if there exists a strictly increasing function $\sigma:\{1,\dots,m\}\to\{1,\dots,n\}$ with the property that $u_i=v_{\sigma(i)}$ for each i with $1\le i\le m$. If $\mathbf{z}^*\in\mathbb{Z}^{\ell_j}$ and $\mathbf{z}^*\hookrightarrow\mathbf{z}$, then we also write $J(\boldsymbol{\Psi};\mathbf{z}^*,\mathbf{c})$ for the determinant of the $\ell_j\times\ell_j$ matrix $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{z}^*,\mathbf{c})$. Now let $\widetilde{S}_{s,r}(P,Q,R;\boldsymbol{\Psi})$ denote the number of solutions of (3.1) with the variables as above and additionally

$$J(\boldsymbol{\Psi};\mathbf{z}^*,\mathbf{c})\ne 0 \quad\text{and}\quad J(\boldsymbol{\Psi};\mathbf{w}^*,\mathbf{c})\ne 0 \tag{3.2}$$

for some $\mathbf{z}^*,\mathbf{w}^*\in\mathbb{Z}^{\ell_j}$ with $\mathbf{z}^*\hookrightarrow\mathbf{z}$ and $\mathbf{w}^*\hookrightarrow\mathbf{w}$.

Finally, we let $T_{s,r}(P,Q,R,\theta;\boldsymbol{\Psi})$ denote the number of solutions of the system

$$\sum_{n=1}^{r}\eta_n(\Psi_{\mathbf{i}}(\mathbf{z}_n,\mathbf{c})-\Psi_{\mathbf{i}}(\mathbf{w}_n,\mathbf{c}))=D_{\mathbf{i}}(\mathbf{c})\mathbf{q}^{\mathbf{i}}\sum_{m=1}^{s}(\mathbf{u}_m^{\mathbf{i}}-\mathbf{v}_m^{\mathbf{i}})\qquad(|\mathbf{i}|=k) \tag{3.3}$$

with $\mathbf{z}$, $\mathbf{w}$, $\mathbf{c}$, and $\boldsymbol{\eta}$ as above, with $u_{ml},v_{ml}\in\mathcal{A}(QP^{-\theta},R)$, with $\mathbf{q}\in[P^\theta,P^\theta R]^d$, and with

$$z_{nl}\equiv w_{nl}\pmod{(q_1\cdots q_d)^k}. \tag{3.4}$$

We are now ready to state the fundamental lemma that provides the basis for our efficient differencing procedure.

Lemma 3.1. *Suppose that $(\boldsymbol{\Psi})$ is a system of type (j,P,A), where $\ell_j\le rd$, and let θ be a parameter at our disposal. For each $\varepsilon>0$, there exists $\eta=\eta(\varepsilon,s,k,d)>0$ such that whenever $R\le P^\eta$ one has*

$$\begin{aligned}S_{s,r}(P,Q,R;\boldsymbol{\Psi})\ll{}&\widetilde{C}P^{2r(d-1)+\ell_j-1}S_s(Q,R)+P^{\theta+\varepsilon}Q^{2d-1}\widetilde{S}_{s-1,r}(P,Q,R;\boldsymbol{\Psi})\\&+(P^\theta R)^{(2s-1)d+k(rd^2-\ell_j)+\varepsilon}T_{s,r}(P,Q,R,\theta;\boldsymbol{\Psi}).\end{aligned}$$

Proof. Let S_1 denote the number of solutions counted by $S_{s,r}(P,Q,R,\boldsymbol{\Psi})$ for which the rank of $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{z},\mathbf{w},\mathbf{c})$ is less than ℓ_j, and let S_2 denote the number of solutions for which $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{z},\mathbf{w},\mathbf{c})$ has rank ℓ_j. We sometimes find it convenient to write $\Psi_{\mathbf{i}}(\mathbf{Z},\mathbf{c})$ as a polynomial in the variable $\mathbf{Z}=(Z_1,\dots,Z_d)$, which may then be evaluated at any $\mathbf{Z}\in\{\mathbf{z}_1,\dots,\mathbf{z}_r,\mathbf{w}_1,\dots,\mathbf{w}_r\}$.

First of all, suppose that $S_1\ge S_2$, and consider a choice of $\mathbf{z}$ and $\mathbf{w}$ counted by S_1. Then the rows of the corresponding Jacobian matrix are linearly dependent, so there exist $a_{\mathbf{i}}\in\mathbb{Z}$, not all zero, such that

$$\sum_{\mathbf{i}\succ\mathbf{j}_1}a_{\mathbf{i}}\frac{\partial\Psi_{\mathbf{i}}}{\partial Z_l}\bigg|_{\mathbf{Z}=\mathbf{Z}_0}=0 \tag{3.5}$$

whenever $\mathbf{Z}_0\in\{\mathbf{z}_1,\dots,\mathbf{z}_r,\mathbf{w}_1,\dots,\mathbf{w}_r\}$ and $1\le l\le d$. We now choose a prime $p\in[P,2P]$ that does not divide any coefficient of a term of maximal degree in any of the polynomials $\partial\Psi_{\mathbf{i}}/\partial Z_l$. The number of choices for the coefficients $a_{\mathbf{i}}$ modulo p is $O(p^{\ell_j-1})$, since one of them may be normalized to 1 in $\mathbb{F}_p$. Now let $\mathbf{i}$ denote the smallest index (in the lexicographic ordering defined above) for which $a_{\mathbf{i}}$ is non-zero modulo p. By condition (3) of Definition 2.2, there is an l with $1\le l\le d$ such that $\partial\Psi_{\mathbf{i}}/\partial Z_l$ contains a term of degree $k-j-1$ that is not present in any $\partial\Psi_{\mathbf{j}}/\partial Z_l$ with $\mathbf{j}\succ\mathbf{i}$. Thus, by considering terms of degree $k-j-1$, it follows that the polynomial $\sum_{\mathbf{i}\succ\mathbf{j}_1}a_{\mathbf{i}}\,\partial\Psi_{\mathbf{i}}/\partial Z_l$ is not identically zero in $\mathbb{F}_p[Z_1,\dots,Z_d]$. Hence (3.5) shows that each $\mathbf{z}_n$ and $\mathbf{w}_n$ satisfy a non-trivial polynomial equation in d variables over the field $\mathbb{F}_p$, so the argument of the proof of Lemma 2 of Wooley

[14] shows that the number of choices for $\mathbf{z}$ and $\mathbf{w}$ modulo p is $O(p^{2r(d-1)})$ for each fixed choice of the $a_{\mathbf{i}}$. Thus the total number of possibilities for $\mathbf{z}$ and $\mathbf{w}$ modulo p is $\ll p^{2r(d-1)} \cdot p^{\ell_j - 1} \ll P^{2r(d-1)+\ell_j-1}$. Since $p \ge P$, it follows that there are $O(P^{2r(d-1)+\ell_j-1})$ choices for $\mathbf{z}$ and $\mathbf{w}$ over $\mathbb{Z}$ as well. Trivially, there are $O(\widetilde{C})$ choices for $\mathbf{c}$ and $\boldsymbol{\eta}$. Now write

$$f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R) = \sum_{\mathbf{x} \in \mathcal{A}(Q,R)^d} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} D_{\mathbf{i}}(\mathbf{c}) \mathbf{x}^{\mathbf{i}} \right).$$

Then for any fixed choice of $\mathbf{z}$, $\mathbf{w}$, $\mathbf{c}$, and $\boldsymbol{\eta}$, there is an integral vector $\mathbf{n}$ such that the number of choices for $\mathbf{x}$ and $\mathbf{y}$ satisfying (3.1) is given by

$$\int_{\mathbb{T}^\ell} |f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)|^{2s} e(\boldsymbol{\alpha} \cdot \mathbf{n}) d\boldsymbol{\alpha} \le S_s(Q, R),$$

where this last inequality follows on considering the underlying Diophantine equations. We therefore have

$$S_{s,r}(P, Q, R; \boldsymbol{\Psi}) \le 2S_1 \ll \widetilde{C} P^{2r(d-1)+\ell_j-1} S_s(Q, R). \tag{3.6}$$

Next, suppose that $S_2 \ge S_1$, and consider a solution counted by S_2. After relabeling variables and making appropriate sign changes, we may suppose that $\mathrm{Jac}(\boldsymbol{\Psi}; \mathbf{z}, \mathbf{c})$ has rank ℓ_j. When $\mathbf{z} \in \mathbb{Z}^{rd}$ and $\mathcal{I} \subseteq \{1, \dots, rd\}$ with $|\mathcal{I}| = \ell_j$, we define the vector $\mathbf{z}_{\mathcal{I}} \in \mathbb{Z}^{\ell_j}$ to have ith component $\mathbf{z}_{\sigma(i)}$, where $\sigma(i)$ is the ith element (in increasing order) of $\mathcal{I}$. In particular, we have $\mathbf{z}_{\mathcal{I}} \hookrightarrow \mathbf{z}$. Here again we typically think of the components of $\mathbf{z}$ to be ordered as $(\mathbf{z}_1, \dots, \mathbf{z}_r)$.

We therefore have

$$S_{s,r}(P, Q, R; \boldsymbol{\Psi}) \ll \sum_{\mathbf{c}, \boldsymbol{\eta}, \boldsymbol{\omega}, \mathcal{I}} \int_{\mathbb{T}^\ell} \mathcal{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi}) \mathcal{H}^*_{\mathbf{c},\boldsymbol{\omega},\mathcal{I}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi}) |f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)|^{2s} d\boldsymbol{\alpha},$$

where

$$\mathcal{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi}) = \sum_{\mathbf{z}} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} \sum_{n=1}^{r} \eta_n \Psi_{\mathbf{i}}(\mathbf{z}_n, \mathbf{c}) \right)$$

and

$$\mathcal{H}^*_{\mathbf{c},\boldsymbol{\omega},\mathcal{I}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi}) = \sum_{\substack{\mathbf{z} \\ J(\boldsymbol{\Psi}; \mathbf{z}_{\mathcal{I}}, \mathbf{c}) \ne 0}} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} \sum_{n=1}^{r} \omega_n \Psi_{\mathbf{i}}(\mathbf{z}_n, \mathbf{c}) \right).$$

By the Cauchy-Schwarz inequality, we have

$$S_{s,r} \ll \left(\sum_{\mathbf{c},\boldsymbol{\eta}} \int_{\mathbb{T}^\ell} |\mathcal{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s}| \, d\boldsymbol{\alpha} \right)^{1/2} \left(\sum_{\mathbf{c},\boldsymbol{\omega},\mathcal{I}} \int_{\mathbb{T}^\ell} |\mathcal{H}^*_{\mathbf{c},\boldsymbol{\omega},\mathcal{I}}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s}| \, d\boldsymbol{\alpha} \right)^{1/2},$$

where we have abbreviated $S_{s,r}(P, Q, R; \boldsymbol{\Psi})$ and the exponential sums $f_{\mathbf{c}}(\boldsymbol{\alpha}; Q, R)$, $\mathcal{H}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi})$, and $\mathcal{H}^*_{\mathbf{c},\boldsymbol{\omega},\mathcal{I}}(\boldsymbol{\alpha}; P; \boldsymbol{\Psi})$ in the obvious ways. It follows on considering the underlying Diophantine equations that the first factor on the right-hand side is bounded above by $S_{s,r}(P, Q, R; \boldsymbol{\Psi})^{1/2}$ and hence that

$$S_{s,r}(P, Q, R; \boldsymbol{\Psi}) \ll \sum_{\mathcal{I} \subseteq \{1, \dots, rd\}} S_3(\mathcal{I}) \ll \max_{\mathcal{I}} S_3(\mathcal{I}), \tag{3.7}$$

where $S_3(\mathcal{I})$ is the number of solutions counted by $S_{s,r}(P,Q,R;\mathbf{\Psi})$ for which

$$J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})\neq 0 \quad \text{and} \quad J(\mathbf{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c})\neq 0.$$

We now fix $\mathcal{I}$ for which $S_3(\mathcal{I})$ is maximal and further classify the solutions counted by $S_3(\mathcal{I})$. Write $x\,\mathcal{D}(L)\,y$ if there exists $d|x$ with $d\le L$ such that x/d has all its prime factors amongst those of y. We let S_4 denote the number of solutions counted by $S_3(\mathcal{I})$ for which

$$x_{ml}\ \mathcal{D}(P^\theta)\ J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c}) \quad \text{or} \quad y_{ml}\ \mathcal{D}(P^\theta)\ J(\mathbf{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c}) \tag{3.8}$$

for some m and l with $1\le m\le s$ and $1\le l\le d$, and let S_5 denote the number of solutions such that (3.8) fails for all m and l.

Suppose that $S_4\ge S_5$, and write

$$\mathcal{S}(\mathbf{\Psi};\mathbf{z},\mathbf{c})=\{x\in\mathcal{A}(Q,R): x\ \mathcal{D}(P^\theta)\ J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})\}$$

and

$$\widetilde{\mathcal{H}}_{\mathbf{c},\boldsymbol{\eta},l}(\boldsymbol{\alpha})=\sum_{\substack{\mathbf{z}\\ J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})\neq 0}}\ \sum_{\substack{\mathbf{x}\in\mathcal{A}(Q,R)^d\\ x_l\in\mathcal{S}(\mathbf{\Psi};\mathbf{z},\mathbf{c})}} e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}\left(\sum_{n=1}^{r}\eta_n\Psi_{\mathbf{i}}(\mathbf{z}_n,\mathbf{c})-D_{\mathbf{i}}(\mathbf{c})\mathbf{x}^{\mathbf{i}}\right)\right).$$

Then by the Cauchy-Schwarz inequality we have

$$\begin{aligned}S_4&\ll\sum_{\mathbf{c},\boldsymbol{\eta},l}\int_{\mathbb{T}^\ell}\widetilde{\mathcal{H}}_{\mathbf{c},\boldsymbol{\eta},l}(\boldsymbol{\alpha})\mathcal{H}^*_{\mathbf{c},\boldsymbol{\eta}}(-\boldsymbol{\alpha})f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s-1}d\boldsymbol{\alpha}\\ &\ll\left(\sum_{\mathbf{c},\boldsymbol{\eta},l}\int_{\mathbb{T}^\ell}|\widetilde{\mathcal{H}}_{\mathbf{c},\boldsymbol{\eta},l}(\boldsymbol{\alpha})^2f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s-2}|d\boldsymbol{\alpha}\right)^{1/2}\left(\sum_{\mathbf{c},\boldsymbol{\eta},l}\int_{\mathbb{T}^\ell}|\mathcal{H}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha})^2f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s}|d\boldsymbol{\alpha}\right)^{1/2},\end{aligned}$$

so on considering the underlying Diophantine equations and recalling (3.7) we find that

$$S_{s,r}(P,Q,R;\mathbf{\Psi})\ll\sum_{g,h,l,\mathbf{c}}V(g,h,l,\mathbf{c}), \tag{3.9}$$

where $V(g,h,l,\mathbf{c})$ denotes the number of solutions of the system

$$\sum_{n=1}^{r}\eta_n(\Psi_{\mathbf{i}}(\mathbf{z}_n,\mathbf{c})-\Psi_{\mathbf{i}}(\mathbf{w}_n,\mathbf{c}))+D_{\mathbf{i}}(\mathbf{c})\sum_{m=1}^{s-1}(\mathbf{x}_m^{\mathbf{i}}-\mathbf{y}_m^{\mathbf{i}})=D_{\mathbf{i}}(\mathbf{c})(a^{i_l}\mathbf{u}^{\mathbf{i}}-b^{i_l}\mathbf{v}^{\mathbf{i}})$$

with $g|J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})\neq 0$ and $h|J(\mathbf{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c})\neq 0$, with $s_0(u_l)=g$ and $s_0(v_l)=h$, with $1\le a,b\le P^\theta$, with $\mathbf{u},\mathbf{v}\in[1,Q]^d$, with $u_l\le Q/a$ and $v_l\le Q/b$, and with the remaining variables as indicated in the discussion surrounding (3.1). Now write

$$G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha})=\sum_{\substack{\mathbf{z}\\ g|J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})\neq 0}} e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}(\eta_1\Psi_{\mathbf{i}}(\mathbf{z}_1,\mathbf{c})+\cdots+\eta_r\Psi_{\mathbf{i}}(\mathbf{z}_r,\mathbf{c}))\right)$$

and

$$\mathcal{G}_{\mathbf{c},\boldsymbol{\eta},l}(\boldsymbol{\alpha})=\sum_{g\le Q}G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha})\sum_{a\le P^\theta}\ \sum_{\substack{\mathbf{u}\in[1,Q]^d\\ u_l\le Q/a\\ s_0(u_l)=g}} e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}D_{\mathbf{i}}(\mathbf{c})a^{i_l}\mathbf{u}^{\mathbf{i}}\right).$$

From (3.9), we find that

$$S_{s,r}(P,Q,R;\mathbf{\Psi}) \ll \sum_{\mathbf{c},\boldsymbol{\eta},l} \int_{\mathbb{T}^\ell} |\mathcal{G}_{\mathbf{c},\boldsymbol{\eta},l}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\boldsymbol{\alpha})^{2s-2}| \, d\boldsymbol{\alpha}, \tag{3.10}$$

and on applying Cauchy's inequality and interchanging the order of summation we obtain

$$|\mathcal{G}_{\mathbf{c},\mathbf{n},l}(\boldsymbol{\alpha})|^2 \le \mathcal{E}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}) \mathcal{F}_{\mathbf{c},l}(\boldsymbol{\alpha}), \tag{3.11}$$

where

$$\mathcal{E}_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha}) = \sum_{g \le Q} |G_{\mathbf{c},\boldsymbol{\eta},g}(\boldsymbol{\alpha})|^2$$

and

$$\mathcal{F}_{\mathbf{c},l}(\boldsymbol{\alpha}) = \sum_{g \le Q} \left| \sum_{\substack{\mathbf{u} \in [1,Q]^d \\ s_0(u_l)=g}} \sum_{a \le P^\theta, \, a \le Q/u_l} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} D_{\mathbf{i}}(\mathbf{c}) a^{i_l} \mathbf{u}^{\mathbf{i}} \right) \right|^2 .$$

After another application of Cauchy's inequality, we deduce from Lemma 2.3 that

$$\mathcal{F}_{\mathbf{c},l}(\boldsymbol{\alpha}) \ll \sum_{g \le Q} Q^{d-1+\varepsilon} \sum_{\substack{\mathbf{u} \in [1,Q]^d \\ s_0(u_l)=g}} \frac{QP^\theta}{u_l} \ll Q^{2d-1+\varepsilon} P^\theta. \tag{3.12}$$

Thus, on substituting (3.11) and (3.12) into (3.10) and recalling a standard estimate for the divisor function, we conclude that

$$S_{s,r}(P,Q,R;\mathbf{\Psi}) \ll Q^{2d-1} P^{\theta+\varepsilon} \widetilde{S}_{s-1,r}(P,Q,R;\mathbf{\Psi}). \tag{3.13}$$

Finally, suppose that $S_5 \ge S_4$, and consider a solution to (3.1) counted by S_5. Write $\tilde{q}_{ml}$ for the largest divisor of x_{ml} that is coprime to $J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})$, and write $\tilde{p}_{ml}$ for the largest divisor of y_{ml} that is coprime to $J(\mathbf{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c})$. Since for each m and l the condition (3.8) fails to hold, we have $\tilde{q}_{ml} > P^\theta$ and $\tilde{p}_{ml} > P^\theta$. Moreover, since these integers are R-smooth, we may divide out a suitable product of prime factors to obtain integers q_{ml} dividing x_{ml} and p_{ml} dividing y_{ml} with the property that

$$P^\theta < q_{ml}, p_{ml} \le P^\theta R \tag{3.14}$$

and

$$(q_{ml}, J(\mathbf{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})) = (p_{ml}, J(\mathbf{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c})) = 1 \tag{3.15}$$

for each m and l with $1 \le m \le s$ and $1 \le l \le d$. Thus we have $S_5 \ll V_1$, where V_1 denotes the number of solutions of the system

$$\sum_{n=1}^{r} \eta_n(\Psi_{\mathbf{i}}(\mathbf{z}_n,\mathbf{c}) - \Psi_{\mathbf{i}}(\mathbf{w}_n,\mathbf{c})) = D_{\mathbf{i}}(\mathbf{c}) \sum_{m=1}^{s} (\mathbf{q}_m^{\mathbf{i}} \mathbf{u}_m^{\mathbf{i}} - \mathbf{p}_m^{\mathbf{i}} \mathbf{v}_m^{\mathbf{i}}) \qquad (|\mathbf{i}| = k)$$

with $\mathbf{z},\mathbf{w},\mathbf{c},\boldsymbol{\eta}$ as in the discussion surrounding (3.1), with $\mathbf{u},\mathbf{v} \in \mathcal{A}(QP^{-\theta},R)^{sd}$, and with $\mathbf{q}$ and $\mathbf{p}$ satisfying (3.14) and (3.15). We now write

$$q = \prod_{\substack{1 \le m \le s \\ 1 \le l \le d}} q_{ml} \quad \text{and} \quad p = \prod_{\substack{1 \le m \le s \\ 1 \le l \le d}} p_{ml}$$

and introduce the exponential sum

$$F_{\mathbf{c},\boldsymbol{\eta},q}(\boldsymbol{\alpha}) = \sum_{\substack{\mathbf{z}\in[1,P]^{rd}\\(q,J(\boldsymbol{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c}))=1}} e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}\left(\eta_1\Psi_{\mathbf{i}}(\mathbf{z}_1,\mathbf{c})+\cdots+\eta_r\Psi_{\mathbf{i}}(\mathbf{z}_r,\mathbf{c})\right)\right).$$

Then we have

$$V_1 \le \sum_{\mathbf{c},\boldsymbol{\eta},\mathbf{q},\mathbf{p}}\int_{\mathbb{T}^\ell} F_{\mathbf{c},\boldsymbol{\eta},q}(\boldsymbol{\alpha})F_{\mathbf{c},\boldsymbol{\eta},p}(-\boldsymbol{\alpha})\prod_{m=1}^{s}\mathcal{F}_{\mathbf{c},m}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}, \tag{3.16}$$

where

$$\mathcal{F}_{\mathbf{c},m}(\boldsymbol{\alpha}) = f_{\mathbf{c}}(\mathbf{q}_m\boldsymbol{\alpha};QP^{-\theta},R)f_{\mathbf{c}}(-\mathbf{p}_m\boldsymbol{\alpha};QP^{-\theta},R),$$

and where we have written $\mathbf{q}\boldsymbol{\alpha}$ for the ℓ-dimensional vector whose component indexed by $\mathbf{i}$ is given by $\mathbf{q}^{\mathbf{i}}\alpha_{\mathbf{i}}$. We now let

$$X_{\mathbf{c},\boldsymbol{\eta},m}(\boldsymbol{\alpha}) = |F_{\mathbf{c},\boldsymbol{\eta},q}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\mathbf{q}_m\boldsymbol{\alpha};QP^{-\theta},R)^{2s}|$$

and

$$Y_{\mathbf{c},\boldsymbol{\eta},m}(\boldsymbol{\alpha}) = |F_{\mathbf{c},\boldsymbol{\eta},p}(\boldsymbol{\alpha})^2 f_{\mathbf{c}}(\mathbf{p}_m\boldsymbol{\alpha};QP^{-\theta},R)^{2s}|.$$

Then by interchanging the order of summation and applying Hölder's inequality twice in (3.16), we find that

$$\begin{aligned} V_1 &\le \sum_{\mathbf{q},\mathbf{p}}\prod_{m=1}^{s}\left(\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^\ell} X_{\mathbf{c},\boldsymbol{\eta},m}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}\right)^{1/2s}\left(\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^\ell} Y_{\mathbf{c},\boldsymbol{\eta},m}(\boldsymbol{\alpha})\,d\boldsymbol{\alpha}\right)^{1/2s}\\ &= \sum_{\mathbf{q},\mathbf{p}}\prod_{m=1}^{s} W(P,Q,R,\mathbf{q}_m)^{1/2s}W(P,Q,R,\mathbf{p}_m)^{1/2s}, \end{aligned}$$

where $W(P,Q,R,\mathbf{q})$ denotes the number of solutions of (3.3) with $\mathbf{z},\mathbf{w}\in[1,P]^{rd}$, with $\mathbf{u},\mathbf{v}\in\mathcal{A}(QP^{-\theta},R)^{sd}$, and with

$$(q_1\cdots q_d, J(\boldsymbol{\Psi};\mathbf{z}_{\mathcal{I}},\mathbf{c})) = (q_1\cdots q_d, J(\boldsymbol{\Psi};\mathbf{w}_{\mathcal{I}},\mathbf{c})) = 1. \tag{3.17}$$

It therefore follows from Hölder's inequality that

$$V_1 \ll (P^\theta R)^{(2s-1)d}\left(\prod_{m=1}^{s}\sum_{\mathbf{q}_m}W(P,Q,R,\mathbf{q}_m)\right)^{1/2s}\left(\prod_{m=1}^{s}\sum_{\mathbf{p}_m}W(P,Q,R,\mathbf{p}_m)\right)^{1/2s}$$

and hence

$$V_1 \ll (P^\theta R)^{(2s-1)d}U_{s,r}(P,Q,R,\theta;\boldsymbol{\Psi}), \tag{3.18}$$

where $U_{s,r}(P,Q,R,\theta;\boldsymbol{\Psi})$ denotes the number of solutions of (3.3) with $\mathbf{c},\boldsymbol{\eta},\mathbf{z},\mathbf{w},\mathbf{u},\mathbf{v}$ as above, with $\mathbf{q}\in(P^\theta,P^\theta R]^d$, and with (3.17).

It remains to impose the condition (3.4). Write

$$\Upsilon_{\mathbf{i}}(\mathbf{z},\mathbf{c},\boldsymbol{\eta}) = \sum_{n=1}^{r}\eta_n\Psi_{\mathbf{i}}(\mathbf{z}_n,\mathbf{c}),$$

and let $\mathcal{B}_{\mathbf{q}}(\mathbf{u},\mathbf{c},\boldsymbol{\eta})$ denote the set of solutions modulo $(q_1\dots q_d)^k$ of the system of congruences

$$\Upsilon_{\mathbf{i}}(\mathbf{z},\mathbf{c},\boldsymbol{\eta}) \equiv u_{\mathbf{i}} \pmod{\mathbf{q}^{\mathbf{i}}} \qquad (|\mathbf{i}|=k)$$

satisfying $(q_1 \cdots q_d, J(\boldsymbol{\Psi}; \mathbf{z}_{\mathcal{I}}, \mathbf{c})) = 1$. For each solution that is counted by $U_{s,r}(P, Q, R, \theta; \boldsymbol{\Psi})$, we have

$$\Upsilon_{\mathbf{i}}(\mathbf{z}, \mathbf{c}, \boldsymbol{\eta}) \equiv \Upsilon_{\mathbf{i}}(\mathbf{w}, \mathbf{c}, \boldsymbol{\eta}) \pmod{\mathbf{q}^{\mathbf{i}}} \qquad (|\mathbf{i}| = k),$$

so we can classify the solutions according to this common residue class. Let

$$H_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{z}; \mathbf{c}, \boldsymbol{\eta}) = \sum_{\substack{\mathbf{x} \in [1,P]^{rd} \\ \mathbf{x} \equiv \mathbf{z} \pmod{(q_1 \cdots q_d)^k}}} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} \Upsilon_{\mathbf{i}}(\mathbf{x}, \mathbf{c}, \boldsymbol{\eta}) \right)$$

and

$$\widetilde{H}_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{c}, \boldsymbol{\eta}) = \sum_{\substack{\mathbf{u} \\ 1 \le u_{\mathbf{i}} \le \mathbf{q}^{\mathbf{i}}}} \left| \sum_{\mathbf{z} \in \mathcal{B}_{\mathbf{q}}(\mathbf{u}, \mathbf{c}, \boldsymbol{\eta})} H_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{z}; \mathbf{c}, \boldsymbol{\eta}) \right|^2.$$

Then we have

$$U_{s,r}(P, Q, R, \theta; \boldsymbol{\Psi}) \ll \sum_{\mathbf{q}, \mathbf{c}, \boldsymbol{\eta}} \int_{\mathbb{T}^\ell} \widetilde{H}_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{c}, \boldsymbol{\eta}) |f_{\mathbf{c}}(\mathbf{q}\boldsymbol{\alpha}; QP^{-\theta}, R)|^{2s} d\boldsymbol{\alpha}.$$

By Cauchy's inequality and Lemma 2.1, we have

$$\widetilde{H}_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{c}, \boldsymbol{\eta}) \ll \sum_{\substack{\mathbf{u} \\ 1 \le u_{\mathbf{i}} \le \mathbf{q}^{\mathbf{i}}}} (P^\theta R)^{\Omega + \varepsilon} \sum_{\mathbf{z} \in \mathcal{B}_{\mathbf{q}}(\mathbf{u}, \mathbf{c}, \boldsymbol{\eta})} |H_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{z}; \mathbf{c}, \boldsymbol{\eta})|^2,$$

where

$$\Omega = d\left(krd - \frac{\ell_j}{d}(k - j) \right) - j\ell_j = k(rd^2 - \ell_j).$$

After inserting this upper bound for $\widetilde{H}_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{c}, \boldsymbol{\eta})$, considering the underlying Diophantine equations, and recalling (3.3) and (3.4), we deduce that

$$U_{s,r}(P, Q, R, \theta; \boldsymbol{\Psi}) \ll (P^\theta R)^{\Omega + \varepsilon} T_{s,r}(P, Q, R, \theta; \boldsymbol{\Psi}).$$

The lemma now follows on assembling (3.6), (3.13), and (3.18). □

When s is sufficiently large, it turns out that the second term in the estimate of Lemma 3.1 can be eliminated. Thus we obtain the following simplification, which will be useful in our iterative processes.

Lemma 3.2. *Suppose that $(\boldsymbol{\Psi})$ is a system of type (j, P, A), that $0 \le \theta \le 1/(dk)$, that $s \ge (k+2)\ell$, and that $r \le \ell_j \le rd$. For each $\varepsilon > 0$, there exists $\eta_0 = \eta_0(\varepsilon, s, k, d) > 0$ such that whenever $R = P^\eta$ and $\eta < \eta_0$ one has*

$$S_{s,r}(P, Q, R; \boldsymbol{\Psi}) \ll \widetilde{C} P^{2r(d-1)+\ell_j - 1} S_s(Q, R) + P^{\gamma + \varepsilon} T_{s,r}(P, Q, R, \theta; \boldsymbol{\Psi}),$$

where $Q = P^{1-\theta}$ and

$$\gamma = \theta[(2s-1)d + k(rd^2 - \ell_j)].$$

Proof. We may clearly suppose that the second term in the estimate of Lemma 3.1 dominates, for the above estimate is certainly true otherwise. That is, we may assume that

$$S_{s,r}(P, Q, R; \boldsymbol{\Psi}) \ll P^{\theta + \varepsilon} Q^{2d-1} \widetilde{S}_{s-1,r}(P, Q, R; \boldsymbol{\Psi}).$$

Write

$$\mathcal{H}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P;\boldsymbol{\Psi})=\sum_{\mathbf{z}} e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}\left(\eta_1\Psi_{\mathbf{i}}(\mathbf{z}_1,\mathbf{c})+\cdots+\eta_r\Psi_{\mathbf{i}}(\mathbf{z}_r,\mathbf{c})\right)\right),$$

where the summation is over all $\mathbf{z}\in[1,P]^{rd}$ for which $\mathrm{Jac}(\boldsymbol{\Psi};\mathbf{z},\mathbf{c})$ has rank ℓ_j. Then

$$\widetilde{S}_{s-1,r}(P,Q,R;\boldsymbol{\Psi})=\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^\ell}|\mathcal{H}^*(\boldsymbol{\alpha};P;\boldsymbol{\Psi})^2 f_{\mathbf{c}}(\boldsymbol{\alpha};Q,R)^{2s-2}|\,d\boldsymbol{\alpha},$$

and following two applications of Hölder's inequality we deduce that

$$\begin{aligned}S_{s,r}&\ll P^{\theta+\varepsilon}Q^{2d-1}\left(\sum_{\mathbf{c},\boldsymbol{\eta}}\int_{\mathbb{T}^\ell}|\mathcal{H}^*_{\mathbf{c},\boldsymbol{\eta}}(\boldsymbol{\alpha};P;\boldsymbol{\Psi})|^2 d\boldsymbol{\alpha}\right)^{1/s}\left(S_{s,r}\right)^{1-1/s}\\&\ll \widetilde{C}P^{2rd-\ell_j+s\theta+\varepsilon}Q^{(2d-1)s},\end{aligned}$$

where we have abbreviated $S_{s,r}(P,Q,R;\boldsymbol{\Psi})$ by $S_{s,r}$. We now claim that this bound is smaller than the first term in the lemma whenever $s\ge(k+2)\ell$ and $\theta\le 1/(dk)$. By (1.4), we have $S_s(Q,R)\gg Q^{2sd-k\ell}$, so it suffices to show that

$$P^{2rd-\ell_j+s\theta}Q^{(2d-1)s}\ll P^{2r(d-1)+\ell_j-1}Q^{2sd-k\ell},$$

and this is equivalent to

$$s(2\theta-1)-\ell_j\le -2r+\ell_j-1-k\ell(1-\theta),$$

or

$$\theta(2s-k\ell)\le s+2\ell_j-2r-1-k\ell.$$

Since $\theta\le 1/(dk)$, it now suffices to show that

$$2s-k\ell\le dk(s+2\ell_j-2r-1-k\ell),$$

and a simple calculation reveals that this indeed holds whenever the conditions $s\ge(k+2)\ell$ and $r\le\ell_j$ are satisfied. □

We now describe the polynomials $\Psi_{\mathbf{i}}$ to which we want to apply Lemma 3.1 and verify that they satisfy the hypotheses of the lemma. To this end, we first define the difference operator Δ_j recursively by

$$\Delta_1(f(\mathbf{z});\mathbf{h}_1)=f(\mathbf{z}+\mathbf{h}_1)-f(\mathbf{z})$$

and

$$\Delta_{j+1}(f(\mathbf{z});\mathbf{h}_1,\dots,\mathbf{h}_{j+1})=\Delta_1(\Delta_j(f(\mathbf{z});\mathbf{h}_1,\dots,\mathbf{h}_j);\mathbf{h}_{j+1}),$$

and we adopt the convention that $\Delta_0(f(\mathbf{z}))=f(\mathbf{z})$. Next we define $\Psi_{\mathbf{i},j}$ recursively by taking $\Psi_{\mathbf{i},0}(\mathbf{z})=\mathbf{z}^{\mathbf{i}}$ and setting

$$\Psi_{\mathbf{i},j}(\mathbf{z};\mathbf{h};\mathbf{m})=\Delta_j(\mathbf{z}^{\mathbf{i}};\mathbf{h}_1(m_1\cdots m_d)^k,\dots,\mathbf{h}_j(m_1\cdots m_d)^k).$$

We typically think of $\mathbf{h}$ and $\mathbf{m}$ as fixed and regard $\Psi_{\mathbf{i},j}$ as a polynomial in $\mathbf{z}$. When $\mathbf{h}=(\mathbf{h}_1,\dots,\mathbf{h}_j)$ is a j-tuple of d-dimensional vectors, we find it useful to let $\mathbf{h}^*$ denote the corresponding d-tuple of j-dimensional vectors formed by taking the transpose of the underlying matrix, so that $\mathbf{h}_l^*=(h_{1l},\dots,h_{jl})$. We start by relating our vector difference operator to the more familiar scalar one. When

$\mathcal{A} = \{i_1, \dots, i_m\}$ and $\mathcal{B} = \{j_1, \dots, j_t\}$ with $\mathcal{A} \cap \mathcal{B} = \emptyset$, we write

$$D_t(f(z); \mathbf{h}; \mathcal{A}; \mathcal{B}) = \Delta_t(f(z + h_{i_1} + \cdots + h_{i_m}); h_{j_1}, \dots, h_{j_t}),$$

where Δ_t is the one-dimensional version of the difference operator defined above.

Lemma 3.3. *One has*

$$\Delta_j(\mathbf{z}^{\mathbf{i}}; \mathbf{h}_1, \dots, \mathbf{h}_j) = \sum_{\mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_d = \{1, \dots, j\}} \prod_{l=1}^{d} D_{|\mathcal{A}_l|}(z_l^{i_l}; \mathbf{h}_l^*; \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{l-1}; \mathcal{A}_l).$$

Proof. We proceed by induction on j. First of all, we have

$$\Delta_0(\mathbf{z}^{\mathbf{i}}) = z_1^{i_1} \cdots z_d^{i_d} = \prod_{l=1}^{d} D_0(z_l^{i_l}; \emptyset; \emptyset).$$

Now suppose that the result holds with j replaced by $j-1$. Then by the induction hypothesis and the linearity of Δ_1, we have

$$\begin{aligned}\Delta_j(\mathbf{z}^{\mathbf{i}}; \mathbf{h}_1, \dots, \mathbf{h}_j) &= \Delta_1(\Delta_{j-1}(\mathbf{z}^{\mathbf{i}}; \mathbf{h}_1, \dots, \mathbf{h}_{j-1}); \mathbf{h}_j) \\ &= \sum_{\mathcal{A}_1 \sqcup \cdots \sqcup \mathcal{A}_d = \{1, \dots, j-1\}} \left(\prod_{l=1}^{d} f_l(z_l + h_{jl}) - \prod_{l=1}^{d} f_l(z_l) \right),\end{aligned}$$

where

$$f_l(z) = D_{|\mathcal{A}_l|}(z^{i_l}; \mathbf{h}_l^*; \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{l-1}; \mathcal{A}_l).$$

Note that, for any complex numbers a_l and b_l, one has

$$\prod_{l=1}^{d} a_l - \prod_{l=1}^{d} b_l = \sum_{l=1}^{d} (a_l - b_l) \prod_{m>l} a_m \prod_{m<l} b_m.$$

We therefore find that

$$\prod_{l=1}^{d} f_l(z_l + h_{jl}) - \prod_{l=1}^{d} f_l(z_l) = \sum_{l=1}^{d} D_{|\mathcal{A}_l|+1}(z_l^{i_l}; \mathbf{h}_l^*; \mathcal{C}_{l-1}; \mathcal{A}_l \cup \{j\})\, Y_l(\mathbf{z}; \mathbf{h}),$$

where we have written $\mathcal{C}_{l-1}$ for $\mathcal{A}_1 \cup \cdots \cup \mathcal{A}_{l-1}$, and where

$$Y_l(\mathbf{z}; \mathbf{h}) = \prod_{m>l} D_{|\mathcal{A}_m|}(z_m^{i_m}; \mathbf{h}_m^*; \mathcal{C}_{m-1} \cup \{j\}; \mathcal{A}_m) \prod_{m<l} D_{|\mathcal{A}_m|}(z_m^{i_m}; \mathbf{h}_m^*; \mathcal{C}_{m-1}; \mathcal{A}_m).$$

On writing $\mathcal{B}_l = \mathcal{A}_l \cup \{j\}$ and $\mathcal{B}_m = \mathcal{A}_m$ for $m \neq l$, we see that

$$\Delta_j(\mathbf{z}^{\mathbf{i}}; \mathbf{h}_1, \dots, \mathbf{h}_j) = \sum_{l=1}^{d} \sum_{\substack{\mathcal{B}_1 \sqcup \cdots \sqcup \mathcal{B}_d = \{1, \dots, j\} \\ j \in \mathcal{B}_l}} \prod_{m=1}^{d} D_{|\mathcal{B}_m|}(z_m^{i_m}; \mathbf{h}_m^*; \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_{m-1}; \mathcal{B}_m),$$

and the result follows on summing over l. □

We are now in a position to analyze the polynomials $\Psi_{\mathbf{i},j}$ defined above.

Lemma 3.4. *Fix j with $0 \le j < k$, and suppose that $\mathbf{h}_1, \dots, \mathbf{h}_j \in \mathbb{Z}^d$ and $\mathbf{m}_1, \dots, \mathbf{m}_j \in \mathbb{Z}^d$ have the property that $0 < |h_{nl} m_{nl}^k| \le cP$ whenever $1 \le n \le j$ and $1 \le l \le d$. Then the polynomials $\Psi_{\mathbf{i},j}$ form a system of type (j, P, A), where $A = c^j (k!)^{d+1}$.*

Proof. It is easy to show (see, for example, Vaughan [11, Exercise 2.1]) that the leading term of $D_t(z^i; \mathbf{h}; \mathcal{A}; \mathcal{B})$ is

$$g(z) = \frac{i!}{(i-t)!}\left(\prod_{n\in\mathcal{B}} h_n\right) z^{i-t},$$

and it therefore follows from Lemma 3.3 that the terms of highest degree in the polynomial $\Psi_{\mathbf{i},j}(\mathbf{z}; \mathbf{h}; \mathbf{m})$ are given by

$$G_{\mathbf{i},j}(\mathbf{z}) = \sum_{\mathcal{A}_1\sqcup\cdots\sqcup\mathcal{A}_d=\{1,\dots,j\}} \left(\prod_{l=1}^{d} \frac{i_l!}{(i_l-|\mathcal{A}_l|)!} \prod_{n\in\mathcal{A}_l} h_{nl} m_{nl}^k\right) z_1^{i_1-|\mathcal{A}_1|}\cdots z_d^{i_d-|\mathcal{A}_d|}.$$

Conditions (1) and (2) of Definition 2.2 follow immediately. To check condition (3), we fix $\mathbf{i}$ with $\mathbf{i} \succ \mathbf{j}_1$ (so in particular $i_1 \geq j$) and consider the term $z_1^{i_1-j} z_2^{i_2}\cdots z_d^{i_d}$ arising from the choice $\mathcal{A}_1 = \{1,\dots,j\}$ in the expression for $G_{\mathbf{i},j}(\mathbf{z})$ above. Suppose now that there is some $\mathbf{i}'$ such that $\Psi_{\mathbf{i}',j}(\mathbf{z})$ (and hence $G_{\mathbf{i}',j}(\mathbf{z})$) contains the term $z_1^{i_1-j} z_2^{i_2}\cdots z_d^{i_d}$. If $i_1' = i_1$, then this term must again arise from the choice $\mathcal{A}_1 = \{1,\dots,j\}$, and it follows that $\mathbf{i}' = \mathbf{i}$. Otherwise, we must have $i_1' < i_1$, which implies that $\mathbf{i}' \prec \mathbf{i}$. □

We now consider the effect of substituting $\Psi_{\mathbf{i},j}(\mathbf{z}; \mathbf{h}; \mathbf{m})$ for $\Psi_{\mathbf{i}}(\mathbf{z}, \mathbf{c})$ in the Fundamental Lemma. Suppose that $0 \leq \phi_j \leq 1/(dk)$, and write

$$M_j = P^{\phi_j}, \quad H_j = PM_j^{-dk}, \quad \text{and} \quad Q_j = P(M_1\cdots M_j)^{-1},$$

with the convention that $Q_0 = P$. We also set $\widetilde{M}_j = M_1\cdots M_j$ and $\widetilde{H}_j = H_1\cdots H_j$, and we replace the conditions on $\mathbf{c}$ by

$$1 \leq |h_{il}| \leq H_i \quad \text{and} \quad M_i < m_{il} \leq M_i R \tag{3.19}$$

for each i and l with $1 \leq i \leq j$ and $1 \leq l \leq d$. Finally, we take $D_{\mathbf{i}}(\mathbf{m}) = \mathbf{m}_1^{\mathbf{i}}\cdots\mathbf{m}_j^{\mathbf{i}}$.

The following lemma allows us to relate $T_{s,r}(P, Q_j, R, \phi_{j+1}; \mathbf{\Psi}_j)$ to $S_s(Q_{j+1}, R)$ and $S_{s,w}(P, Q_{j+1}, R; \mathbf{\Psi}_{j+1})$ and hence to repeat the differencing process.

Lemma 3.5. *Suppose that $r \leq 2w$. For every $\varepsilon > 0$ there exists $\eta_0 = \eta_0(\varepsilon, s, k, d) > 0$ such that whenever $R = P^\eta$ and $\eta < \eta_0$ one has*

$$\begin{aligned} T_{s,r}(P, Q_j, R, \phi_{j+1}; \mathbf{\Psi}_j) &\ll P^{(2d-1-d(d-1)k\phi_{j+1})r+\varepsilon} \widetilde{H}_j^d \widetilde{M}_{j+1}^d S_s(Q_{j+1}, R) \\ &+ P^\varepsilon H_{j+1}^{d(r-1)} \left(\widetilde{H}_{j+1}^d \widetilde{M}_{j+1}^d S_s(Q_{j+1}, R)\right)^{1-r/2w} \left(S_{s,w}(P, Q_{j+1}, R; \mathbf{\Psi}_{j+1})\right)^{r/2w}. \end{aligned}$$

Proof. We introduce the exponential sums

$$\mathcal{L}_{\mathbf{a},q}(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = \sum_{\substack{\mathbf{z}\in[1,P]^d \\ \mathbf{z}\equiv\mathbf{a} \pmod q}} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} \Psi_{\mathbf{i},j}(\mathbf{z}; \mathbf{h}; \mathbf{m})\right),$$

$$\mathcal{K}_q(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m}) = \sum_{\mathbf{a}\in[1,q]^d} |\mathcal{L}_{\mathbf{a},q}(\boldsymbol{\alpha}; \mathbf{h}; \mathbf{m})|^2,$$

and

$$g_{\mathbf{q}}(\boldsymbol{\alpha}; \mathbf{m}) = \sum_{\mathbf{x}\in\mathcal{A}(Q_{j+1},R)^d} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} D_{\mathbf{i}}(\mathbf{m}) \mathbf{q}^{\mathbf{i}} \mathbf{x}^{\mathbf{i}}\right).$$

Then on considering the underlying Diophantine equations we find that

$$T_{s,r} \asymp \sum_{\mathbf{h},\mathbf{m}} \sum_{\mathbf{q}\in[M_{j+1},M_{j+1}R]^d} \int_{\mathbb{T}^\ell} \mathcal{K}_{(q_1\cdots q_d)^k}(\boldsymbol{\alpha};\mathbf{h};\mathbf{m})^r |g_{\mathbf{q}}(\boldsymbol{\alpha};\mathbf{m})|^{2s} d\boldsymbol{\alpha},$$

where the summation ranges for $\mathbf{h}$ and $\mathbf{m}$ are given by (3.19) and where we have abbreviated $T_{s,r}(P,Q,R,\phi_{j+1};\boldsymbol{\Psi}_j)$ by $T_{s,r}$. We let U_0 denote the number of solutions counted by $T_{s,r}(P,Q,R,\phi_{j+1};\boldsymbol{\Psi}_j)$ for which $z_{nl}=w_{nl}$ for some n and l with $1\le n\le r$ and $1\le l\le d$, and let U_1 denote the number of solutions with $z_{nl}\ne w_{nl}$ for all n and l.

First suppose that $U_0\ge U_1$. In view of the condition (3.4), we have

$$U_0 \ll P^{2d-1-d(d-1)k\phi_{j+1}} \sum_{\mathbf{h},\mathbf{m},\mathbf{q}} \int_{\mathbb{T}^\ell} \mathcal{K}_{(q_1\cdots q_d)^k}(\boldsymbol{\alpha};\mathbf{h};\mathbf{m})^{r-1} |g_{\mathbf{q}}(\boldsymbol{\alpha};\mathbf{m})|^{2s} d\boldsymbol{\alpha},$$

so after two applications of Hölder's inequality we find that

$$T_{s,r}(P,Q,R,\phi_{j+1};\boldsymbol{\Psi}_j) \ll P^{(2d-1-d(d-1)k\phi_{j+1})r+\varepsilon} \widetilde{H}_j^d \widetilde{M}_{j+1}^d S_s(Q_{j+1},R). \tag{3.20}$$

Next suppose that $U_1\ge U_0$, and consider a solution counted by U_1. For each n and l, (3.4) allows us to write

$$w_{nl} = z_{nl} + g_{nl}(q_1\cdots q_d)^k,$$

where the g_{nl} and q_l are integers with

$$1\le |g_{nl}|\le H_{j+1} \quad \text{and} \quad M_{j+1}<q_l\le M_{j+1}R. \tag{3.21}$$

Thus we see that U_1 is bounded above by the number of solutions of the system

$$\sum_{n=1}^r \Psi_{\mathbf{i},j+1}(\mathbf{z}_n;\mathbf{h},\mathbf{g}_n;\mathbf{m},\mathbf{q}) = D_{\mathbf{i}}(\mathbf{m})\mathbf{q}^{\mathbf{i}} \sum_{m=1}^s (\mathbf{u}_m^{\mathbf{i}} - \mathbf{v}_m^{\mathbf{i}}) \qquad (|\mathbf{i}|=k),$$

with all the variables in the ranges described above. Now write

$$G(\boldsymbol{\alpha};\mathbf{g};\mathbf{q}) = \sum_{\mathbf{z}\in[1,P]^d} e\left(\sum_{|\mathbf{i}|=k} \alpha_{\mathbf{i}} \Psi_{\mathbf{i},j+1}(\mathbf{z};\mathbf{h},\mathbf{g};\mathbf{m},\mathbf{q}) \right).$$

Then we find after an application of Hölder's inequality that

$$U_1 \ll H_{j+1}^{d(r-1)} \sum_{\mathbf{h},\mathbf{g},\mathbf{m},\mathbf{q}} \int_{\mathbb{T}^\ell} |G(\boldsymbol{\alpha};\mathbf{g};\mathbf{q})^r g_{\mathbf{q}}(\boldsymbol{\alpha};\mathbf{m})^{2s}| d\boldsymbol{\alpha},$$

where the summation conditions are given by (3.19) and (3.21). Applying Hölder's inequality twice more gives

$$U_1 \ll H_{j+1}^{d(r-1)} S_{s,w}(P,Q_{j+1},R;\boldsymbol{\Psi}_{j+1})^{r/2w} \left(P^\varepsilon \widetilde{H}_{j+1}^d \widetilde{M}_{j+1}^d S_s(Q_{j+1},R) \right)^{1-r/2w},$$

and this, together with (3.20), completes the proof of the lemma. □

4. Mean value theorems

We begin by deriving a simple result using only first differences. When $0 \le \theta \le 1/(dk)$, we employ the notation

$$M = P^{\theta}, \qquad H = PM^{-dk}, \qquad \text{and} \qquad Q = PM^{-1}.$$

Theorem 4.1. *Suppose that $s \ge (k+2)\ell$ and that $\Delta_s \le k\ell$ is an admissible exponent. Then the exponent $\Delta_{s+\ell} = \Delta_s(1 - \frac{1}{dk})$ is also admissible.*

Proof. We take $\theta = 1/(dk)$ in the above notation. Then by Lemma 3.2 we have

$$S_{s,\ell}(P,P,R;\mathbf{\Psi}_0) \ll P^{2\ell d-\ell-1}S_s(P,R) + M^{(2s-1)d+k\ell(d^2-1)+\varepsilon}T_{s,\ell}(P,P,R,\theta;\mathbf{\Psi}_0).$$

By employing the argument of the proof of Lemma 3.5, we find that

$$T_{s,\ell}(P,P,R,\theta;\mathbf{\Psi}_0) \ll P^{d\ell+\varepsilon}M^d S_s(Q,R)$$

after making a trivial estimate and noting that $H = 1$. It follows that

$$S_{s,\ell}(P,P,R;\mathbf{\Psi}_0) \ll P^{2\ell d-\ell-1}S_s(P,R) + M^{2sd+k\ell(d^2-1)+d}P^{d\ell+\varepsilon}S_s(Q,R).$$

Moreover, since the exponent $\lambda_s = 2sd - k\ell + \Delta_s$ is permissible, we find after a little calculation that

$$S_{s+\ell}(P,R) \ll S_{s,\ell}(P,P,R;\mathbf{\Psi}_0) \ll P^{\Lambda_1+\varepsilon} + P^{\Lambda_2+\varepsilon},$$

where

$$\Lambda_1 = 2(s+\ell)d - k\ell - (\ell+1) + \Delta_s$$

and

$$\Lambda_2 = 2(s+\ell)d - k\ell + \Delta_s(1-\theta).$$

Since $\Delta_s \le k\ell$, one sees easily that $\theta\Delta_s \le \ell + 1$ and hence that $\Lambda_1 \le \Lambda_2$. We therefore deduce that the exponent $\Delta_{s+\ell} = \Delta_s(1-\theta)$ is admissible, as required. □

We note that the above theorem yields admissible exponents that decay roughly like $k\ell e^{-s/(dk\ell)}$. Good results therefore begin to appear when s is a bit larger than $dk\ell\log(k\ell)$, and we can improve this somewhat by employing repeated efficient differencing. However, we are hampered by the fact that, after j differences, the singularity issues considered in sections 2 and 3 force us to restrict attention to ℓ_j of the ℓ available equations. We find it convenient to introduce the notation

$$\Omega_j = k(\ell - \ell_j), \tag{4.1}$$

which may be thought of as a measure of the resulting loss of potential congruence information. Our results arising from repeated differencing are summarized in the following theorem.

Theorem 4.2. *Suppose that $k \ge 2d$, and let u be a positive integer with $u \ge (k+2)\ell$. Further suppose that $\Delta_u \le k\ell$ is an admissible exponent, and let j be an integer with $1 \le j \le k/2$. For each positive integer m, we write $s = u + m\ell$ and define the numbers $\phi(j,s,J)$, θ_s, and Δ_s recursively as follows. Given a value of $\Delta_{s-\ell}$, we set $\phi(j,s,j) = 1/(dk)$ and evaluate $\phi(j,s,J-1)$ successively for $J = j,\dots,2$ by setting*

$$\phi^*(j,s,J-1) = \frac{1}{2dk} + \left(\frac{1}{2} + \frac{\Omega_{J-1} - \Delta_{s-\ell}}{2dk\ell_{J-1}}\right)\phi(j,s,J) \tag{4.2}$$

and

$$\phi(j,s,J-1) = \min\{1/dk, \phi^*(j,s,J-1)\}.$$

Finally, we set

$$\theta_s = \min_{1 \le j \le k/2} \phi(j, s, 1)$$

and

$$\Delta_s = \Delta_{s-\ell}(1 - \theta_s) + \ell(dk\theta_s - 1). \tag{4.3}$$

Then Δ_s is an admissible exponent for $s = u + m\ell$ for all positive integers m.

Proof. Take j to be the least integer with $1 \le j \le k/2$ for which $\phi(j, s+\ell, 1) = \theta_{s+\ell}$, and write $\phi_J = \phi(j, s+\ell, J)$ for $J = j, \dots, 1$. By following the argument of the proof of Theorem 6.1 in [16], we find that the minimality of j ensures that $\phi_J < 1/(dk)$ whenever $J < j$. We recall the notation established in section 3 and set

$$M_J = P^{\phi_J}, \qquad H_J = PM_J^{-dk}, \qquad \text{and} \qquad Q_J = P(M_1 \cdots M_J)^{-1}$$

for $1 \le J \le j$, with the usual convention that $Q_0 = P$.

We prove by induction on m that the exponents Δ_s defined above are admissible for $s = u + m\ell$, where m is a non-negative integer. The $m = 0$ case is trivial, since the admissibility of Δ_u is a hypothesis of the theorem. Now suppose that $s = u + m\ell$, where $m \ge 0$, and that Δ_s is admissible. Then when R is a sufficiently small power of P we have $S_s(P, R) \ll P^{\lambda_s + \varepsilon}$, where $\lambda_s = 2sd - k\ell + \Delta_s$. We need to establish that $\Delta_{s+\ell}$ is admissible as well. In order to do this, we first show inductively that

$$T_{s,\ell_J}(P, Q_J, R, \phi_{J+1}; \mathbf{\Psi}_J) \ll P^{(2d-1-d(d-1)k\phi_{J+1})\ell_J + \varepsilon} \widetilde{H}_J^d \widetilde{M}_{J+1}^d Q_{J+1}^{\lambda_s} \tag{4.4}$$

for $J = j-1, j-2, \dots, 1, 0$. To establish (4.4) for $J = j-1$, we apply Lemma 3.5 with j replaced by $j-1$ and with $r = \ell_{j-1}$ and $w = \ell_j$. It is easy to verify that $\ell_{j-1} \le 2\ell_j$ whenever $k \ge 2d$ and $j \le k/2$. On making the trivial estimate

$$S_{s,\ell_j}(P, Q_j, R; \mathbf{\Psi}_j) \ll P^{2d\ell_j} \widetilde{H}_j^d \widetilde{M}_j^d Q_j^{\lambda_s + \varepsilon}$$

and noting that $\phi_j = 1/(dk)$, and hence $H_j = 1$, it follows easily that

$$T_{s,\ell_{j-1}}(P, Q_{j-1}, R, \phi_j; \mathbf{\Psi}_{j-1}) \ll P^{d\ell_{j-1} + \varepsilon} \widetilde{H}_{j-1}^d \widetilde{M}_j^d Q_j^{\lambda_s},$$

as required. Now suppose that (4.4) holds for J, where $1 \le J \le j-1$. Then Lemma 3.2 gives

$$S_{s,\ell_J}(P, Q_J, R; \mathbf{\Psi}_J) \ll \widetilde{H}_J^d \widetilde{M}_J^d P^{(2d-1)\ell_J - 1} Q_J^{\lambda_s + \varepsilon} + M_{J+1}^{\gamma} T_{s,\ell_J}(P, Q_J, R, \phi_{J+1}; \mathbf{\Psi}_J),$$

where $\gamma = (2s-1)d + k\ell_J(d^2-1)$. Substituting (4.4) and noting that $Q_J = Q_{J+1}M_J$ then yields

$$S_{s,\ell_J}(P, Q_J, R; \mathbf{\Psi}_J) \ll \widetilde{H}_J^d \widetilde{M}_J^d P^{(2d-1)\ell_J - 1} Q_J^{\lambda_s + \varepsilon} \left(1 + M_{J+1}^{\gamma + d - \lambda_s} P^{1 - d(d-1)k\phi_{J+1}\ell_J}\right),$$

and a simple calculation shows that the second term in parentheses can be expressed as PM_{J+1}^{β}, where

$$\beta = k\ell_J(d-1) + k\ell - \Delta_s \ge 0.$$

Hence the second term dominates, and we get

$$S_{s,\ell_J}(P, Q_J, R; \mathbf{\Psi}_J) \ll \widetilde{H}_J^d \widetilde{M}_J^d M_{J+1}^{\beta} Q_J^{\lambda_s} P^{(2d-1)\ell_J + \varepsilon}.$$

Thus an application of Lemma 3.5 gives

$$T_{s,\ell_{J-1}}(P, Q_{J-1}, R, \phi_J; \mathbf{\Psi}_{J-1}) \ll P^{(2d-1-d(d-1)k\phi_J)\ell_{J-1}} \widetilde{H}_{J-1}^d \widetilde{M}_J^d Q_J^{\lambda_s + \varepsilon} + P^{\varepsilon} W_J,$$

where

$$W_J = H_J^{d(\ell_{J-1}-1)} \Big(\widetilde{H}_J^d \widetilde{M}_J^d Q_J^{\lambda_s} \Big)^{1-\frac{\ell_{J-1}}{2\ell_J}} \Big(\widetilde{H}_J^d \widetilde{M}_J^d M_{J+1}^{\beta} Q_J^{\lambda_s} P^{(2d-1)\ell_J} \Big)^{\frac{\ell_{J-1}}{2\ell_J}}.$$

After simplifying and putting $\sigma = \frac{\ell_{J-1}}{2\ell_J}$, we find that

$$W_J = \widetilde{H}_{J-1}^d \widetilde{M}_J^d Q_J^{\lambda_s} H_J^{d\ell_{J-1}} M_{J+1}^{\sigma\beta} P^{(d-\frac{1}{2})\ell_{J-1}},$$

and hence

$$T_{s,\ell_{J-1}}(P, Q_{J-1}, R, \phi_J; \mathbf{\Psi}_{J-1}) \ll \widetilde{H}_{J-1}^d \widetilde{M}_J^d Q_J^{\lambda_s+\varepsilon} \Big(P^{\Lambda_1} + H_J^{d\ell_{J-1}} M_{J+1}^{\sigma\beta} P^{(d-\frac{1}{2})\ell_{J-1}} \Big),$$

where $\Lambda_1 = (2d-1-d(d-1)k\phi_J)\ell_{J-1}$. Moreover, the second term in parentheses can be expressed as P^{Λ_2}, where

$$\Lambda_2 = (1-dk\phi_J)d\ell_{J-1} + \frac{\phi_{J+1}\ell_{J-1}}{2\ell_J}(dk\ell_J + \Omega_J - \Delta_s) + (d-\tfrac{1}{2})\ell_{J-1}.$$

Finally, from (4.2) and the observation that $\phi_J < 1/(dk)$ for $J < j$, we obtain the relation

$$(2dk\phi_J - 1)\ell_J = (dk\ell_J + \Omega_J - \Delta_s)\phi_{J+1},$$

which yields

$$\Lambda_2 = (1-dk\phi_J)d\ell_{J-1} + (dk\phi_J - \tfrac{1}{2})\ell_{J-1} + (d-\tfrac{1}{2})\ell_{J-1} = \Lambda_1.$$

Thus we find that (4.4) holds with J replaced by $J-1$, as required. Applying this bound with $J=0$, we get

$$T_{s,\ell}(P, P, R, \phi_1; \mathbf{\Psi}_0) \ll P^{(2d-1-d(d-1)k\phi_1)\ell + d\phi_1 + \lambda_s(1-\phi_1)+\varepsilon}.$$

Lemma 3.2 therefore yields

$$S_{s+\ell}(P,R) \ll S_{s,\ell}(P,P,R;\mathbf{\Psi}_0) \ll P^{\Lambda_3+\varepsilon} + P^{\Lambda_4+\varepsilon},$$

where

$$\Lambda_3 = (2d-l)\ell - 1 + \lambda_s \quad \text{and} \quad \Lambda_4 = (2d-1)\ell + \lambda_s + \phi_1(dk\ell - \Delta_s),$$

and it is obvious that $\Lambda_3 \le \Lambda_4$. Thus we find that the exponent

$$\lambda_{s+\ell} = 2d(s+\ell) - k\ell + \Delta_s(1-\phi_1) + \ell(dk\phi_1 - 1)$$

is permissible, and this completes the proof. □

We now need to gain some understanding of the size of the admissible exponents provided by Theorem 4.2, and this is achieved by a fairly standard argument (see, for example, [7], [9], [12], [13], and [16] for similar analyses). The following lemma provides the starting point by relating these exponents to the roots of a transcendental equation.

Lemma 4.3. *Suppose that $s \ge (k+3)\ell$ and that $\Delta_{s-\ell}$ is an admissible exponent satisfying $\ell(\log k)^2 < \Delta_{s-\ell} \le k\ell$. Write $\delta_{s-\ell} = \Delta_{s-\ell}/(dk\ell)$, and define δ_s to be the unique (positive) solution of the equation*

$$\delta_s + \log\delta_s = \delta_{s-\ell} + \log\delta_{s-\ell} - \frac{2}{dk} + \frac{2}{dk(\log k)^{3/2}}. \tag{4.5}$$

Then the exponent $\Delta_s = dk\ell\delta_s$ is admissible.

Proof. We apply Theorem 4.2 with $j = [(\log k)^{1/3}]$. Then on writing $\theta_s = \phi(j, s, 1)$, we find that the exponent

$$\Delta_s^* = \Delta_{s-\ell}(1-\theta_s) + \ell(dk\theta_s - 1) = dk\ell\delta_{s-\ell} - \ell + dk\ell\theta_s(1-\delta_{s-\ell}) \tag{4.6}$$

is admissible. On recalling (2.1) and (4.1), a simple calculation shows that

$$\Omega_J \le \ell(\log k)^{1/2}$$

for $0 \le J < j$, provided that k is sufficiently large in terms of d. Thus on writing ϕ_J for $\phi(j, s, J)$ and noting that $\ell_{J-1} \le \ell$, we deduce from (4.2) that

$$\phi_{J-1} \le \frac{1}{2dk} + \frac{1}{2}(1-\delta')\phi_J \qquad (2 \le J \le j), \tag{4.7}$$

where

$$\delta' = \frac{\Delta_{s-\ell} - \ell(\log k)^{1/2}}{dk\ell} > \delta_{s-\ell}(1 - (\log k)^{-3/2}), \tag{4.8}$$

the last inequality following from the hypothesis that $\Delta_{s-\ell} > \ell(\log k)^2$. Using a downward induction via (4.7), one easily verifies that

$$\phi_J \le \frac{1}{dk(1+\delta')}\left(1 + \delta'\left(\frac{1-\delta'}{2}\right)^{j-J}\right) \qquad (1 \le J \le j),$$

so in particular we have

$$\theta_s = \phi_1 \le \frac{1 + \delta' 2^{1-j}}{dk(1+\delta')}, \tag{4.9}$$

since $0 < \delta' < 1$. Let us temporarily introduce the notation $L = (\log k)^{-3/2}$. Since $(1+\alpha x)/(1+x)$ is a decreasing function of x whenever $\alpha < 1$, we deduce from (4.8) and (4.9) that

$$\theta_s \le \frac{1 + \delta_{s-\ell}(1-L)2^{1-j}}{dk(1+\delta_{s-\ell}(1-L))} \le \frac{1 + \delta_{s-\ell}(2^{1-j}+L)}{dk(1+\delta_{s-\ell})} \le \frac{1 + 2\delta_{s-\ell}L}{dk(1+\delta_{s-\ell})},$$

provided that k is large enough so that $j \ge 1 + \log_2(\log k)^{3/2}$. It now follows with a little computation from (4.6) that

$$\frac{\Delta_s^*}{dk\ell} \le \delta_{s-\ell}\left(1 - \frac{2-w}{dk\ell(1+\delta_{s-\ell})}\right),$$

where $w = 2(1-\delta_{s-\ell})L$. Since $\log(1-x) \le -x$ for $0 < x < 1$, we therefore obtain

$$\begin{aligned}\frac{\Delta_s^*}{dk\ell} + \log\frac{\Delta_s^*}{dk\ell} &\le \delta_{s-\ell}\left(1 - \frac{2-w}{dk(1+\delta_{s-\ell})}\right) + \log\delta_{s-\ell} - \frac{2-w}{dk(1+\delta_{s-\ell})} \\ &\le \delta_{s-\ell} + \log\delta_{s-\ell} - \frac{2}{dk} + \frac{2}{dk(\log k)^{3/2}}\end{aligned}$$

on inserting the bound $w \le 2L$. Now $\delta + \log\delta$ is an increasing function of δ, so if δ_s is defined by (4.5), then it must be the case that $\Delta_s^*/(dk\ell) \le \delta_s$, and it follows that $dk\ell\delta_s$ is an admissible exponent. □

We are now in a position to state the stronger mean value estimates arising from repeated differencing in a form convenient for applications.

Theorem 4.4. *Suppose that $d \geq 2$ and that k is sufficiently large in terms of d, define s_0 and s_1 as in* (1.6) *and* (1.7)*, and write $L = (\log k)^2$. Then the exponents Δ_s defined by*

$$\Delta_s = \begin{cases} dk\ell e^{3-\frac{2s}{dk\ell}} & \text{if } 1 \leq s \leq s_0, \\ \ell L e^{8/3}\left(1 - \frac{3}{2dk}(1-\frac{d}{2L})\right)^{(s-s_0)/\ell} & \text{if } s > s_0 \end{cases}$$

are admissible.

Proof. Write $r = (k+2)\ell$. We start by observing that the theorem is trivially true when $s \leq r$. Next, we define δ_s to be the unique positive solution of the equation

$$\delta_s + \log \delta_s = 1 - \frac{2(s-r)}{dk\ell} + \frac{2(s-r)}{dk\ell(\log k)^{3/2}}, \tag{4.10}$$

and we show inductively that the exponent $\Delta_s = dk\ell\delta_s$ is admissible whenever $r < s \leq s_0$. First of all, suppose that $r < s \leq r + \ell$. The exponent $\Delta_s^* = k\ell$ is trivially admissible, and furthermore

$$\frac{\Delta_s^*}{dk\ell} + \log \frac{\Delta_s^*}{dk\ell} = \frac{1}{d} + \log \frac{1}{d} < 1 - \frac{2}{dk} < \delta_s + \log \delta_s,$$

since $0 < s - r \leq \ell$. It follows that $\Delta_s^*/(dk\ell) < \delta_s$ and hence that $\Delta_s = dk\ell\delta_s$ is admissible. Now suppose that $r+\ell < s \leq s_0$ and that the exponent $\Delta_{s-\ell} = dk\ell\delta_{s-\ell}$ is admissible. Then we have

$$\delta_{s-\ell} + \log \delta_{s-\ell} > 1 - \frac{2(s_0 - r)}{dk\ell} > 1 - \log(dk) + 2\log\log k.$$

Since $\delta_{s-\ell} < 1$, we deduce that $dk\delta_{s-\ell} > (\log k)^2$, and thus the exponent

$$\Delta'_{s-\ell} = \min\{k\ell,\ \Delta_{s-\ell}\}$$

satisfies the hypotheses of Lemma 4.3. We therefore conclude that the exponent $\Delta'_s = dk\ell\gamma_s$ is admissible, where γ_s is the positive root of the equation

$$\gamma_s + \log \gamma_s = \delta'_{s-\ell} + \log \delta'_{s-\ell} - \frac{2}{dk} + \frac{2}{dk(\log k)^{3/2}},$$

and where $\delta'_{s-\ell} = \Delta'_{s-\ell}/(dk\ell) \leq \delta_{s-\ell}$. On applying (4.10) with s replaced by $s-\ell$, we find that $\gamma_s + \log\gamma_s \leq \delta_s + \log\delta_s$, and hence $\gamma_s \leq \delta_s$. It follows that $\Delta_s = dk\ell\delta_s$ is admissible. Moreover, when $s \leq s_0$, we see from (4.10) that

$$\log \delta_s \leq \frac{7}{3} - \frac{2s}{dk\ell},$$

provided that k is sufficiently large, and the desired bound for Δ_s follows. Finally, if $s > s_0$, we take t to be the integer with $s_0 - \ell < t \leq s_0$ and $t \equiv s \pmod{\ell}$. Then we know that $\Delta_t = dk\ell e^{7/3 - 2t/(dk\ell)}$ is an admissible exponent, and we have

$$e^{7/3}\ell(\log k)^2 \leq \Delta_t < e^{8/3}\ell(\log k)^2. \tag{4.11}$$

We now apply Theorem 4.2 with $j = 2$ and s replaced by $t + \ell$. In the notation of that theorem, we have $\phi(2, t+\ell, 2) = 1/(dk)$, and thus

$$\phi^*(2, t+\ell, 1) = \frac{1}{2dk} + \left(\frac{1}{2} + \frac{\Omega_1 - \Delta_t}{2dk\ell_1}\right)\frac{1}{dk} = \frac{1}{dk} + \frac{\Omega_1 - \Delta_t}{2d^2k^2\ell_1}.$$

It therefore follows from (4.3) that the exponent

$$\Delta_{t+\ell} = \Delta_t\left(1 - \frac{1}{dk}\left(1 + \frac{\ell}{2\ell_1}\right) + \frac{\Delta_t - \Omega_1}{2d^2k^2\ell_1}\right) + \frac{\Omega_1\ell}{2dk\ell_1} \tag{4.12}$$

is admissible. A simple calculation reveals that $\Omega_1 \le d\ell$ for k sufficiently large, and thus (4.11) gives $\Omega_1 \le d(\log k)^{-2}\Delta_t = dL^{-1}\Delta_t$. Hence on iterating (4.12) and noting that $\ell/\ell_1 = 1 + O(1/k)$, we find that the exponent

$$\Delta_s = \Delta_t \left(1 - \frac{3}{2dk}\left(1 - \frac{d}{2L}\right)\right)^{(s-t)/\ell}$$

is admissible for k sufficiently large, and the theorem follows on substituting the upper bound in (4.11) and recalling that $t \le s_0$. □

To deduce Theorem 1.1, we first note that

$$1 - \frac{3}{2dk}\left(1 - \frac{d}{2L}\right) \le \left(1 - \frac{3}{2dk}\right)\left(1 + \frac{1}{kL}\right)$$

for $dk \ge 6$. Thus on using the inequality $(1 + b/x)^x \le e^b$ we find that

$$\left(1 - \frac{3}{2dk}\left(1 - \frac{1}{\log k}\right)\right)^{dk} \le e^{-3/2} \cdot e^{d/L}.$$

Theorem 1.1 now follows immediately from Theorem 4.4 when $s \le s_1$. Finally, the argument given in section 6 below to prove Theorem 1.2 may be modified to show that $\Delta_s = 0$ for $s > s_1$, so Theorem 1.1 holds in that case as well.

5. Major arc asymptotics

In this section, we obtain an asymptotic formula for the exponential sum $f(\boldsymbol{\alpha})$ when $\boldsymbol{\alpha}$ lies within a narrow set of major arcs. We let W be a parameter with $W \le (\log P)^{1/2-\varepsilon}$ and define $\mathfrak{N}(q, \mathbf{a}; W)$ to be the set of all $\boldsymbol{\alpha} \in \mathbb{T}^\ell$ such that

$$|\alpha_{\mathbf{i}} - a_{\mathbf{i}}/q| \le WP^{-k} \qquad (|\mathbf{i}| = k).$$

Further, write $\mathfrak{N}(W)$ for the union of all the $\mathfrak{N}(q, \mathbf{a}; W)$ with $0 \le a_{\mathbf{i}} \le q \le W$ and $(q, \mathbf{a}) = 1$. Here and throughout we adopt the notation $(x, \mathbf{y})$ to represent $\gcd(x, y_1, \dots, y_\ell)$ whenever $x \in \mathbb{Z}$ and $\mathbf{y} \in \mathbb{Z}^\ell$. In what follows, we find it convenient to introduce the notation

$$S_j(q, \mathbf{a}; x_{j+1}, \dots, x_d) = \sum_{1 \le r_1, \dots, r_j \le q} e\left(q^{-1} \sum_{|\mathbf{i}|=k} a_{\mathbf{i}} r_1^{i_1} \cdots r_j^{i_j} x_{j+1}^{i_{j+1}} \cdots x_d^{i_d}\right)$$

and

$$w_j(\boldsymbol{\beta}; x_{j+1}, \dots, x_d) = \int_{[R,P]^j} \widetilde{\rho}_j(\boldsymbol{\gamma}, R)\, e\left(\sum_{|\mathbf{i}|=k} \beta_{\mathbf{i}} \gamma_1^{i_1} \cdots \gamma_j^{i_j} x_{j+1}^{i_{j+1}} \cdots x_d^{i_d}\right) d\gamma_1 \cdots d\gamma_j,$$

where

$$\widetilde{\rho}_j(\boldsymbol{\gamma}, R) = \prod_{i=1}^{j} \rho\left(\frac{\log \gamma_i}{\log R}\right)$$

and where ρ denotes the well-known Dickman function (see, for example, Vaughan [11, section 12.1]). We shall write $S(q, \mathbf{a})$ and $w(\boldsymbol{\beta})$ for $S_d(q, \mathbf{a})$ and $w_d(\boldsymbol{\beta})$, respectively, and we also write

$$S(q; g) = \sum_{r=1}^{q} e\left(\frac{g(r)}{q}\right) \qquad \text{and} \qquad \rho^*(\gamma, R) = \rho\left(\frac{\log \gamma}{\log R}\right).$$

Lemma 5.1. *For any polynomial $g(x)$ with integer coefficients, one has*

$$\sum_{x\in\mathcal{A}(\gamma,R)} e\left(\frac{g(x)}{q}\right) = q^{-1}S(q;g)\rho^*(\gamma,R)\gamma + E(\gamma),$$

where $E(\gamma)$ is a piecewise-differentiable function satisfying $E(\gamma) \ll qP/\log P$.

Proof. First of all, by Lemmas 5.3 and 5.4 of Vaughan [10], one has

$$\sum_{\substack{x\in\mathcal{A}(\gamma,R)\\ x\equiv r \pmod q}} 1 = q^{-1}\rho^*(\gamma,R)\gamma + O\left(\frac{P}{\log P}\right). \tag{5.1}$$

Next, by sorting the sum according to residue classes modulo q and noting that $g(x) \equiv g(r) \pmod q$ whenever $x \equiv r \pmod q$, we obtain

$$\sum_{x\in\mathcal{A}(\gamma,R)} e\left(\frac{g(x)}{q}\right) = \sum_{r=1}^{q} e\left(\frac{g(r)}{q}\right) \sum_{\substack{x\in\mathcal{A}(\gamma,R)\\ x\equiv r \pmod q}} 1,$$

and the lemma now follows easily. □

We can analyze the effect of a small twist on the above sum via partial summation.

Lemma 5.2. *Suppose that $g(x) \in \mathbb{Z}[x]$ and $h(x) \in \mathbb{R}[x]$ and that $h'(x) \ll WP^{-1}$ whenever $|x| \le P$. Then one has*

$$\sum_{x\in\mathcal{A}(P,R)} e\left(\frac{g(x)}{q} + h(x)\right) = q^{-1}S(q;g)\int_R^P \rho^*(\gamma,R)e(h(\gamma))\,d\gamma + O\left(\frac{qWP}{\log P}\right).$$

Proof. For fixed g and q, we write

$$T(\gamma) = \sum_{x\in\mathcal{A}(\gamma,R)} e\left(\frac{g(x)}{q}\right).$$

By Lemma 5.1 and properties of the Riemann-Stieltjes integral, we have

$$\sum_{x\in\mathcal{A}(P,R)} e\left(\frac{g(x)}{q} + h(x)\right) = \int_R^P e(h(\gamma))T'(\gamma)\,d\gamma + O(R) \tag{5.2}$$

and

$$T'(\gamma) = q^{-1}S(q;g)\frac{\partial}{\partial\gamma}\big[\rho^*(\gamma,R)\gamma\big] + E'(\gamma),$$

where $E(\gamma) \ll qP/\log P$. Moreover, we have

$$\frac{\partial}{\partial\gamma}\big[\rho^*(\gamma,R)\gamma\big] = \rho\left(\frac{\log\gamma}{\log R}\right) + \frac{1}{\log R}\,\rho'\left(\frac{\log\gamma}{\log R}\right).$$

Since $\log R \gg \log P$, we deduce that

$$T'(\gamma) = q^{-1}S(q;g)\rho^*(\gamma,R) + E'(\gamma) + O\left(\frac{1}{\log P}\right)$$

for $\gamma \ge R$, since $\rho'(x) \ll 1$ whenever $x \ge 1$. We therefore obtain

$$\int_R^P e(h(\gamma))T'(\gamma)\,d\gamma = q^{-1}S(q;g)\left(\int_R^P \rho^*(\gamma,R)e(h(\gamma))\,d\gamma + \mathcal{E}(q,P)\right),$$

where

$$\mathcal{E}(q,P)=\int_R^P E'(\gamma)e(h(\gamma))\,d\gamma+O\left(\frac{P}{\log P}\right).$$

Integrating by parts and using the assumption that $h'(\gamma)\ll WP^{-1}$, we find that

$$\mathcal{E}(q,P)\ll\frac{qP}{\log P}+\int_R^P|E(\gamma)h'(\gamma)|\,d\gamma\ll\frac{qWP}{\log P}.$$

The lemma now follows immediately on noting the trivial bound $S(q;g)\ll q$ and recalling the formula (5.2). □

Lemma 5.3. *Suppose that $\boldsymbol{\alpha}\in\mathfrak{N}(q,\mathbf{a};W)\subseteq\mathfrak{N}(W)$, and write $\beta_{\mathbf{i}}=\alpha_{\mathbf{i}}-a_{\mathbf{i}}/q$. Then one has*

$$f(\boldsymbol{\alpha})=q^{-d}S(q,\mathbf{a})w(\boldsymbol{\beta})+O\left(\frac{qWP^d}{\log P}\right).$$

Proof. We prove by induction that one has

$$\sum_{x_1,\dots,x_j\in\mathcal{A}(P,R)}e\left(\sum_{|\mathbf{i}|=k}\alpha_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right)=q^{-j}S_j(q,\mathbf{a};\tilde{\mathbf{x}}_{j+1})w_j(\boldsymbol{\beta};\tilde{\mathbf{x}}_{j+1})+O\left(\frac{qWP^j}{\log P}\right)\tag{5.3}$$

for $1\le j\le d$, where we have written $\tilde{\mathbf{x}}_{j+1}$ for the vector $(x_{j+1},\dots,x_d)$. For $j=1$, we fix $x_2,\dots,x_d$ and let

$$g(x_1)=\sum_{|\mathbf{i}|=k}a_{\mathbf{i}}x_1^{i_1}x_2^{i_2}\cdots x_d^{i_d}\qquad\text{and}\qquad h(x_1)=\sum_{|\mathbf{i}|=k}\beta_{\mathbf{i}}x_1^{i_1}x_2^{i_2}\cdots x_d^{i_d}.$$

Then since $\boldsymbol{\alpha}\in\mathfrak{N}(q,\mathbf{a};W)\subseteq\mathfrak{N}(W)$, we have

$$h'(x_1)\ll\sum_{|\mathbf{i}|=k}|\beta_{\mathbf{i}}|P^{k-1}\ll WP^{-1}$$

whenever $|x_1|\le P$, so (5.3) follows immediately from Lemma 5.2 in the case $j=1$. Now suppose that (5.3) holds for some $j<d$, and write $U_j(\mathbf{x})$ for the left-hand side of (5.3). Then one has

$$U_{j+1}(\mathbf{x})=q^{-j}\sum_{x_{j+1}\in\mathcal{A}(P,R)}S_j(q,\mathbf{a};\tilde{\mathbf{x}}_{j+1})w_j(\boldsymbol{\beta};\tilde{\mathbf{x}}_{j+1})+O\left(\frac{qWP^{j+1}}{\log P}\right),\tag{5.4}$$

and the first term on the right-hand side may be rewritten as

$$q^{-j}\sum_{1\le r_1,\dots,r_j\le q}\int_{[R,P]^j}\widetilde{\rho}_j(\boldsymbol{\gamma},R)\sum_{x_{j+1}\in\mathcal{A}(P,R)}e(q^{-1}g(\mathbf{x},\mathbf{r})+h(\mathbf{x},\boldsymbol{\gamma}))\,d\gamma_1\cdots d\gamma_j,$$

where

$$g(\mathbf{x},\mathbf{r})=\sum_{|\mathbf{i}|=k}a_{\mathbf{i}}r_1^{i_1}\cdots r_j^{i_j}x_{j+1}^{i_{j+1}}\cdots x_d^{i_d}\quad\text{and}\quad h(\mathbf{x},\boldsymbol{\gamma})=\sum_{|\mathbf{i}|=k}\beta_{\mathbf{i}}\gamma_1^{i_1}\cdots\gamma_j^{i_j}x_{j+1}^{i_{j+1}}\cdots x_d^{i_d}.$$

Let us also write $V(\mathbf{x},\mathbf{r},\boldsymbol{\gamma})$ for the sum over x_{j+1} on the right-hand side. We may temporarily fix the variables $\mathbf{r}$, $\boldsymbol{\gamma}$, x_{j+2}, …, x_d and view $g(\mathbf{x},\mathbf{r})$ and $h(\mathbf{x},\boldsymbol{\gamma})$ as functions of x_{j+1} alone. Then by applying Lemma 5.2 as above, we find that

$$V(\mathbf{x},\mathbf{r},\boldsymbol{\gamma})=q^{-1}\sum_{r_{j+1}=1}^{q}e(q^{-1}G(\mathbf{x},\mathbf{r}))\int_R^P\rho^*(\gamma_{j+1},R)e(H(\mathbf{x},\boldsymbol{\gamma}))\,d\gamma_{j+1}+O\left(\frac{qWP}{\log P}\right),$$

where

$$G(\mathbf{x},\mathbf{r})=\sum_{|\mathbf{i}|=k} a_{\mathbf{i}} r_1^{i_1}\cdots r_{j+1}^{i_{j+1}} x_{j+2}^{i_{j+2}}\cdots x_d^{i_d}$$

and

$$H(\mathbf{x},\boldsymbol{\gamma})=\sum_{|\mathbf{i}|=k} \beta_{\mathbf{i}} \gamma_1^{i_1}\cdots \gamma_{j+1}^{i_{j+1}} x_{j+2}^{i_{j+2}}\cdots x_d^{i_d}.$$

On substituting this into (5.4), we find that

$$U_{j+1}(\mathbf{x})=q^{-j-1}S_{j+1}(q,\mathbf{a};\tilde{\mathbf{x}}_{j+2})w_{j+1}(\boldsymbol{\beta};\tilde{\mathbf{x}}_{j+2})+O\left(\frac{qWP^{j+1}}{\log P}\right),$$

and thus (5.3) holds with $j+1$ in place of j. The lemma now follows immediately by taking $j=d$ in (5.3). □

Finally, we record the asymptotics for our exponential sums over complete intervals, which are valid on a wider set of major arcs. We let $X\le P^{1-\varepsilon}$ be a parameter, and define $\mathfrak{M}(q,\mathbf{a};X)$ to be the set of $\boldsymbol{\alpha}\in\mathbb{T}^\ell$ such that

$$|q\alpha_{\mathbf{i}}-a_{\mathbf{i}}|\le XP^{-k}\qquad (|\mathbf{i}|=k).$$

We further write $\mathfrak{M}(X)$ for the union of the $\mathfrak{M}(q,\mathbf{a};X)$ with $0\le a_{\mathbf{i}}\le q\le X$ and $(q,\mathbf{a})=1$. Finally, write

$$v(\boldsymbol{\beta})=\int_{[0,P]^d} e\left(\sum_{|\mathbf{i}|=k}\beta_{\mathbf{i}}\boldsymbol{\gamma}^{\mathbf{i}}\right)d\boldsymbol{\gamma}$$

in analogy with the function $w(\boldsymbol{\beta})$ defined above.

Lemma 5.4. *Suppose that $\boldsymbol{\alpha}\in\mathfrak{M}(q,\mathbf{a};X)\subseteq\mathfrak{M}(X)$, and write $\beta_{\mathbf{i}}=\alpha_{\mathbf{i}}-a_{\mathbf{i}}/q$. Then one has*

$$F(\boldsymbol{\alpha})=q^{-d}S(q,\mathbf{a})v(\boldsymbol{\beta})+O(XP^{d-1}).$$

Proof. This follows immediately from Lemma 5.3 of [9]. For an alternative proof, one may follow the argument of Lemmas 5.1–5.3 above with the right-hand side of (5.1) replaced by $\gamma/q+O(1)$ to deduce the result. □

The argument of this section may obviously be applied to exponential sums over more general sets than $[1,P]$ and $\mathcal{A}(P,R)$, the two considered here. All one needs is a formula of the type (5.1), which ensures that the elements of the set are well distributed in residue classes. The rest of the argument is essentially partial summation.

6. Counting linear spaces

In order to establish Theorem 1.2, we introduce the exponential sums

$$F_j(\boldsymbol{\alpha})=\sum_{\mathbf{x}\in[-P,P]^d} e\left(\sum_{|\mathbf{i}|=k} c_j\alpha_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right)$$

and

$$f_j(\boldsymbol{\alpha})=\sum_{\mathbf{x}\in\mathcal{A}^*(P,R)^d} e\left(\sum_{|\mathbf{i}|=k} c_j\alpha_{\mathbf{i}}\mathbf{x}^{\mathbf{i}}\right),$$

where $\mathcal{A}^*(P,R) = \pm\mathcal{A}(P,R) \cup \{0\}$. It is fairly easy to argue that the exponential sum estimates established in §4 and in [9] carry over to the above situations, in which the components of $\mathbf{x}$ can take on negative values. On writing $s = t+2u$, one sees by orthogonality that $N_{s,k,d}(P) \geq \mathcal{I}(P;\mathbb{T}^\ell)$, where

$$\mathcal{I}(P;\mathfrak{B}) = \int_{\mathfrak{B}} \left(\prod_{j=1}^{t} F_j(\boldsymbol{\alpha}) \right) \left(\prod_{j=t+1}^{s} f_j(\boldsymbol{\alpha}) \right) d\boldsymbol{\alpha}.$$

We define the set of major arcs by $\mathfrak{M} = \mathfrak{M}(cP^{1/2})$, where we have written $c = \max|c_j|$, and further write $\mathfrak{m} = \mathbb{T}^\ell \setminus \mathfrak{M}$ for the set of minor arcs. We deal with the minor arcs by applying Theorem 1.1 in combination with the following Weyl-type estimate. Here we write $F(\boldsymbol{\alpha})$ for the coefficient-free version of $F_j(\boldsymbol{\alpha})$.

Lemma 6.1. *Suppose that k is sufficiently large in terms of d and that $|F(\boldsymbol{\alpha})| \geq P^{d-\sigma+\varepsilon}$ for some $\varepsilon > 0$, where $\sigma^{-1} \geq 3k^2\ell \log k$. Then there exist integers $a_{\mathbf{i}}$ and q, with $(q,\mathbf{a}) = 1$, satisfying $1 \leq q \leq P^{k\sigma}$ and $|q\alpha_{\mathbf{i}} - a_{\mathbf{i}}| \leq P^{k(\sigma-1)}$ for each $\mathbf{i}$ with $|\mathbf{i}| = k$.*

Proof. This is an immediate consequence of Parsell [9, Theorem 1.2], together with the above remarks. □

By making trivial estimates, one finds that

$$\mathcal{I}(P;\mathfrak{m}) \ll \int_{\mathfrak{m}} |F_i(\boldsymbol{\alpha})^t f_j(\boldsymbol{\alpha})^{2u}| \, d\boldsymbol{\alpha}$$

for some i and j, where $1 \leq i \leq t$ and $t+1 \leq j \leq s$. Moreover, the argument of [7, Lemma 10.3] implies that $c_i\boldsymbol{\alpha} \notin \mathfrak{M}(P^{1/2})$ whenever $\boldsymbol{\alpha} \in \mathfrak{m}$. Therefore, after a change of variable, we obtain from Lemma 6.1 that

$$\mathcal{I}(P;\mathfrak{m}) \ll \left(\sup_{\boldsymbol{\alpha} \in \mathfrak{m}} |F(c_i\boldsymbol{\alpha})| \right)^t S_u(P,R) \ll P^{sd-k\ell-\sigma t+\Delta_u},$$

where $\sigma^{-1} = 3k^2\ell\log k$, and where Δ_u is as in Theorem 1.1. Taking $t = 3k\ell$ and

$$u = \lceil dk\ell \left(\tfrac{2}{3}\log(k\ell) + \tfrac{1}{2}\log(dk) + \log\log k + 3 \right) \rceil$$

gives $\Delta_u < (k\log k)^{-1}$ and hence $\sigma t > \Delta_u$. Thus one has

$$\mathcal{I}(P;\mathfrak{m}) \ll P^{sd-k\ell-\tau} \tag{6.1}$$

for some $\tau > 0$, which completes the analysis of the minor arcs.

Next we prune the major arcs down to a thin set $\mathfrak{N} = \mathfrak{N}(W)$, where $W = (\log P)^\nu$ for some sufficiently small $\nu > 0$. From Lemmas II.2 and II.8 of [1], we have the estimates

$$v(\boldsymbol{\beta}) \ll P^d \left(1 + \sum_{|\mathbf{i}|=k} |\beta_{\mathbf{i}}| P^{|\mathbf{i}|} \right)^{-1/k} \tag{6.2}$$

and

$$S(q,\mathbf{a}) \ll q^{d-1/k+\varepsilon} \tag{6.3}$$

for every $\varepsilon > 0$, provided that $(q, \mathbf{a}) = 1$. In conjunction with these estimates, a routine application of Lemma 5.4 shows that

$$\mathcal{I}(P; \mathfrak{M} \setminus \mathfrak{N}) \ll P^{sd-k\ell} W^{-\delta} \tag{6.4}$$

for some $\delta > 0$, so it now suffices to deal with the set $\mathfrak{N}$.

Write

$$\mathfrak{B} = [-1, 1]^{td} \times ([-1, -R/P] \cup [R/P, 1])^{2ud}$$

and

$$\mathcal{H}(\boldsymbol{\gamma}) = \prod_{j=t+1}^{s} \rho\left(\frac{\log(P\gamma_j)}{\log R}\right).$$

On using Lemma 5.3, together with (6.2), (6.3), and the observation that $\mathrm{meas}(\mathfrak{N}) \ll W^{2\ell+1} P^{-k\ell}$, one finds that

$$\mathcal{I}(P; \mathfrak{N}) = \mathcal{J}\mathfrak{S} P^{sd-k\ell} + O(P^{sd-k\ell} W^{-\delta}) \tag{6.5}$$

for some $\delta > 0$, where

$$\mathcal{J} = \int_{\mathbb{R}^\ell} \int_{\mathfrak{B}} \mathcal{H}(\boldsymbol{\gamma})\, e\left(\sum_{|\mathbf{i}|=k} \beta_{\mathbf{i}} (c_1 \boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + c_s \boldsymbol{\gamma}_s^{\mathbf{i}})\right) d\boldsymbol{\gamma}$$

denotes the singular integral and

$$\mathfrak{S} = \sum_{q=1}^{\infty} \sum_{\substack{\mathbf{a} \in [1,q]^\ell \\ (q,\mathbf{a})=1}} \prod_{j=1}^{s} q^{-d} S(q, c_j \mathbf{a})$$

denotes the singular series. It now suffices to show that $\mathcal{J}$ and $\mathfrak{S}$ are both positive.

To deal with the singular integral, we follow the argument of [8, Lemma 7.4]. We let T be a positive real number and introduce the functions

$$K_T(\beta) = \left(\frac{\sin \pi\beta T^{-1}}{\pi\beta T^{-1}}\right)^2 \quad \text{and} \quad \mathcal{K}_T(\boldsymbol{\beta}) = \prod_{|\mathbf{i}|=k} K_T(\beta_{\mathbf{i}}).$$

It follows from Lemma 14.1 of Baker [2] that

$$\widehat{K}_T(y) = \int_{-\infty}^{\infty} K_T(\beta)\, e(\beta y)\, d\beta = T \max(0, 1 - T|y|) \tag{6.6}$$

for all real numbers y. We introduce the auxiliary singular integral

$$\mathcal{J}_T = \int_{\mathbb{R}^\ell} \mathcal{K}_T(\boldsymbol{\beta}) \int_{\mathfrak{B}} \mathcal{H}(\boldsymbol{\gamma})\, e\left(\sum_{|\mathbf{i}|=k} \beta_{\mathbf{i}} (c_1 \boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + c_s \boldsymbol{\gamma}_s^{\mathbf{i}})\right) d\boldsymbol{\gamma}$$

and note that (6.2) yields

$$\mathcal{J} - \mathcal{J}_T \ll \int_{\mathbb{R}^\ell} (1 - \mathcal{K}_T(\boldsymbol{\beta})) \prod_{|\mathbf{i}|=k} (1 + |\beta_{\mathbf{i}}|)^{-\frac{t}{k\ell}}\, d\boldsymbol{\beta}. \tag{6.7}$$

A simple calculation reveals that

$$1 - \mathcal{K}_T(\boldsymbol{\beta}) \ll \min(1, |\boldsymbol{\beta}|^2 T^{-2}),$$

so by considering the integral in (6.7) over the regions $|\boldsymbol{\beta}| \le T$ and $|\boldsymbol{\beta}| > T$ separately, one easily shows that $\mathcal{J} - \mathcal{J}_T \ll T^{-\delta}$ for some $\delta > 0$. Hence we have

$$\mathcal{J} = \lim_{T \to \infty} \mathcal{J}_T, \tag{6.8}$$

and so it suffices to analyze $\mathcal{J}_T$. We first note that

$$\mathcal{J}_T = \int_{\mathfrak{B}} \mathcal{H}(\boldsymbol{\gamma}) \prod_{|\mathbf{i}|=k} \widehat{K}_T(g_{\mathbf{i}}(\boldsymbol{\gamma}))\, d\boldsymbol{\gamma}, \tag{6.9}$$

where we have written

$$g_{\mathbf{i}}(\boldsymbol{\gamma}) = c_1\boldsymbol{\gamma}_1^{\mathbf{i}} + \cdots + c_s\boldsymbol{\gamma}_s^{\mathbf{i}}.$$

Since we have assumed that the system $g_{\mathbf{i}}(\boldsymbol{\gamma}) = 0$ ($|\mathbf{i}| = k$) possesses a non-singular real solution $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_s)$, the Implicit Function Theorem ensures that locally near $\boldsymbol{\eta}$ there is an $(sd - \ell)$-dimensional space of real solutions, continuously parameterized by $sd - \ell$ of the coordinates. Therefore, by using continuity of the determinant, we may perturb $\boldsymbol{\eta}$ slightly to obtain another non-singular solution in which at most ℓ coordinates are zero. Furthermore, we may suppose after a rearrangement of variables that each j for which some $\eta_{jl} = 0$ satisfies $1 \le j \le t$. It then follows that $\boldsymbol{\eta}$ lies in the interior of $\mathcal{B}$ whenever P is sufficiently large. Now let κ be any bijection from the set of $\mathbf{i}$ with $|\mathbf{i}| = k$ to the set $\{1, \dots, \ell\}$, and consider the map $\varphi : \mathbb{R}^{sd} \to \mathbb{R}^{sd}$ defined by

$$\varphi_{\kappa(\mathbf{i})}(\boldsymbol{\gamma}) = g_{\mathbf{i}}(\boldsymbol{\gamma}) \quad \text{and} \quad \varphi_j(\boldsymbol{\gamma}) = \gamma_j \quad (\ell + 1 \le j \le sd).$$

By the Inverse Function Theorem, there is an open set $U \subseteq \mathfrak{B}$ containing $\boldsymbol{\eta}$, and an open set V containing $(0, \dots, 0, \eta_{\ell+1}, \dots, \eta_{sd})$, such that φ maps U injectively onto V. Since $\mathcal{H}(\boldsymbol{\gamma}) \gg 1$ on $\mathfrak{B}$ and the integrand in (6.9) is non-negative, we have

$$\mathcal{J}_T \gg \int_V \widehat{K}_T(u_1) \cdots \widehat{K}_T(u_\ell)\, du_1 \cdots du_s. \tag{6.10}$$

In particular, the projection of V onto the first ℓ components contains the set $\mathfrak{D} = [-1/2T, 1/2T]^\ell$ whenever T is sufficiently large. Moreover, (6.6) shows that the integrand in (6.10) is bounded below on $\mathfrak{D}$ by $(T/2)^\ell$. Since $\mathrm{meas}(\mathfrak{D}) = T^{-\ell}$, it follows immediately that $\mathcal{J}_T \gg 1$ for T sufficiently large, and we therefore conclude from (6.8) that $\mathcal{J} > 0$.

In order to show that $\mathfrak{S} > 0$, we first note that Lemma II.4 of [1] may be used to deduce that the function

$$S(q) = \sum_{\substack{\mathbf{a} \in [1,q]^\ell \\ (q,\mathbf{a})=1}} \prod_{j=1}^{s} q^{-d} S(q, c_j\mathbf{a})$$

is multiplicative. Moreover, the series

$$T(p) = \sum_{h=0}^{\infty} S(p^h)$$

is absolutely convergent in view of the bound (6.3), so we find that $\mathfrak{S}$ is represented by the absolutely convergent product $\mathfrak{S} = \prod_p T(p)$ and that there exists an integer p_0 such that

$$\frac{1}{2} \le \prod_{p \ge p_0} T(p) \le \frac{3}{2}.$$

It therefore suffices to show that $T(p) > 0$ for all primes $p < p_0$. Let $M_s(q)$ denote the number of solutions of the system of congruences

$$c_1\mathbf{x}_1^{\mathbf{i}} + \cdots + c_s\mathbf{x}_s^{\mathbf{i}} \equiv 0 \pmod{q} \qquad (|\mathbf{i}| = k).$$

By applying the argument of [11, Lemma 2.12], as in [7, Lemma 9.7], we find that

$$\sum_{d|q} S(d) = q^{\ell - sd} M_s(q),$$

and it follows that

$$T(p) = \lim_{h\to\infty} \sum_{d|p^h} S(d) = \lim_{h\to\infty} p^{h(\ell-sd)} M_s(p^h).$$

Since we have assumed that the system (1.2) possesses a non-singular p-adic solution for each prime p, we may apply a Hensel's Lemma argument as in [7, Lemma 9.9], to conclude that there exists an integer $u = u(p) < \infty$ such that for all h one has

$$M_s(p^h) \geq p^{(h-u)(sd-\ell)}.$$

It follows that $T(p) \geq p^{u(\ell-sd)}$ for each $p < p_0$, and we therefore deduce that $\mathfrak{S} > 0$. Theorem 1.2 now follows on recalling (6.1), (6.4), and (6.5), together with the positivity of the singular integral.

References

1. G. I. Arkhipov, A. A. Karatsuba, and V. N. Chubarikov, *Multiple trigonometric sums*, Trudy Mat. Inst. Steklov **151** (1980), 1–126. MR608411 (82i:10045)
2. R. C. Baker, *Diophantine inequalities*, The Clarendon Press, Oxford University Press, New York, 1986. MR865981 (88f:11021)
3. B. J. Birch, *Homogeneous forms of odd degree in a large number of variables*, Mathematika **4** (1957), 102–105. MR0097359 (20:3828)
4. R. Brauer, *A note on systems of homogeneous algebraic equations*, Bull. Amer. Math. Soc. **51** (1945), 749–755. MR0013127 (7:108i)
5. H. Davenport and D. J. Lewis, *Homogeneous additive equations*, Proc. Roy. Soc. Ser. A **274** (1963), 443–460. MR0153655 (27:3617)
6. S. T. Parsell, *The density of rational lines on cubic hypersurfaces*, Trans. Amer. Math. Soc. **352** (2000), 5045–5062. MR1778504 (2001j:11010)
7. ———, *Multiple exponential sums over smooth numbers*, J. Reine Angew. Math. **532** (2001), 47–104. MR1817503 (2001m:11140)
8. ———, *Pairs of additive equations of small degree*, Acta Arith. **104** (2002), 345–402. MR1911162 (2003e:11036)
9. ———, *A generalization of Vinogradov's mean value theorem*, Proc. London Math. Soc. (3) **91** (2005), 1–32. MR2149529 (2006j:11040)
10. R. C. Vaughan, *A new iterative method in Waring's problem*, Acta Math. **162** (1989), 1–71. MR981199 (90c:11072)
11. ———, *The Hardy-Littlewood method*, 2nd ed., Cambridge University Press, Cambridge, 1997. MR1435742 (98a:11133)
12. T. D. Wooley, *Large improvements in Waring's problem*, Ann. of Math. (2) **135** (1992), 131–164. MR1147960 (93b:11129)
13. ———, *On Vinogradov's mean value theorem*, Mathematika **39** (1992), 379–399. MR1203293 (94d:11074)
14. ———, *A note on symmetric diagonal equations*, Number Theory with an Emphasis on the Markoff Spectrum (A. D. Pollington and W. Moran, eds.), Marcel Dekker, 1993, pp. 317–321. MR1219345 (94d:11076)
15. ———, *A note on simultaneous congruences*, J. Number Theory **58** (1996), 288–297. MR1393617 (97h:11037)
16. ———, *On exponential sums over smooth numbers*, J. Reine Angew. Math. **488** (1997), 79–140. MR1465368 (98g:11110)

Department of Mathematics and Actuarial Science, Butler University, 4600 Sunset Avenue, JH 270, Indianapolis, Indiana 46208

E-mail address: sparsell@butler.edu